MEASUREMENT OF SINGLE TOP QUARK PRODUCTION IN THE TAU+JETS CHANNEL USING BOOSTED DECISION TREES AT D0

by

Zhiyi Liu

Master of Science, China Institute of Atomic Energy, 2002

THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE DEPARTMENT

OF

PHYSICS

© Zhiyi Liu 2009

SIMON FRASER UNIVERSITY

Fall 2009

APPROVAL

Name: Zhiyi Liu

Degree: Doctor of Philosophy

Title of thesis: Measurement of Single Top Quark Production in the Tau+jets
Channel Using Boosted Decision Trees at D0

Examining Committee: J. Steven Dodge (Chair)

Dugan Clive O'Neil (Senior Supervisor)

Michel Vetterli (Supervisor)

Isabel Trigger (Supervisor)

Levon Pogosian (Internal Examiner)

Reda Tafirout (External Examiner)

Date Approved: December 7, 2009

ii

Abstract

The top quark is the heaviest known matter particle and plays an important role in the Standard Model of particle physics. At hadron colliders, it is possible to produce single top quarks via the weak interaction. This allows a direct measurement of the CKM matrix element V_{tb} and serves as a window to new physics.

The first direct measurement of single top quark production with a tau lepton in the final state (the tau+jets channel) is presented in this thesis. The measurement uses 4.8 fb^{-1} of Tevatron Run II data in $p\bar{p}$ collisions at $\sqrt{s} =$ 1.96 TeV acquired by the D0 experiment. After selecting a data sample and building a background model, the data and background model are in good agreement. A multivariate technique, boosted decision trees, is employed in discriminating the small single top quark signal from a large background. The expected sensitivity of the tau+jets channel in the Standard Model is 1.8 standard deviations. Using a Bayesian statistical approach, an upper limit on the cross section of single top quark production in the tau+jets channel is measured as 7.3 pb at 95% confidence level, and the cross section is measured as $3.4^{+2.0}_{-1.8}$ pb.

The result of the single top quark production in the tau+jets channel is also combined with those in the electron+jets and muon+jets channels. The expected sensitivity of the electron, muon and tau combined analysis is 4.7 standard deviations, to be compared to 4.5 standard deviations in electron and muon alone. The measured cross section in the three combined final states is $\sigma(p\bar{p} \to tb+X, tqb+X) = 3.84^{+0.89}_{-0.83}$ pb. A lower limit on $|V_{tb}|$ is also measured in the three combined final states to be larger than 0.85 at 95% confidence level. These results are consistent with Standard Model expectations.

Keywords: single top quark; electroweak top quark; tau lepton; tau+jets channel; cross section; Tevatron; D0; boosted decision trees

Acknowledgments

First and foremost, I owe my deepest gratitude to my supervisor Prof. Dugan O'Neil. This thesis would not have been possible without his continuous support, encouragement and advising. He is a productive intelligent professor who always ignites my ideas by our discussions. He is a good friend more than a boss.

I would like to thank all of my colleagues working for the D0 experiment. Specially, I would give my thanks to Dr. Ernest Aguilo Chivite for his continuous and close collaboration with me on this project. Although we live in different countries, we work closely like at the same office. I would also thank our single top quark group conveners: Dr. Reinhard Schwienhorst for his online patient help, Drs. Ann Heinson and Cecilia Gerber for their good advice.

I would thank all members of my Ph.D. committee. To Prof. Michel Vetterli and Dr. Isabel Trigger for their nice advice and keeping me on the right track, for Drs. Reda Tafirout and Levon Pogosian for their time on this thesis.

It is also my pleasure to give thanks to all of the graduates and postdocs of the SFU high energy group. Our office is always a place filled with discussions, jokes and laughs. Unforgettable: watching Dag's wavy photon dance and teaching me special words, enjoying tennis games with Erfan and sharing funny stories, talking about taus with Jen and telling her Emacs can also close png files, discussing beer belly and decision tree algorithms with Doug, ... I realize that I have to compose another "thesis" if I want to write down all funny things.

I am deeply indebted to my parents and parents-in-law. Thank them for their consideration, encouragement and support. Last but not least, it is an honor for me to give my special thanks to my lovely wife, Adele, for her love and understanding.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

"What is the universe made of?" is one of the biggest questions facing science over the next quarter-century [1]. Particle physics is the field in which elementary matter and its interactions are studied. Since the civilization of humans, people have never stopped exploring our universe, including the Earth on which we live. With rapid development of science and technology in modern times, we have obtained unprecedented progress in understanding our universe.

All matter has both wave-like and particle-like properties (wave-particle duality). This is a central concept in quantum mechanics. A de Broglie wavelength λ of a particle with momentum p is given by the relation $\lambda = h/p$, where h is Planck's constant. In order to reach a smaller probe scale, an incident particle must be accelerated to a higher energy and guided to smash on a target (in a target-fixed experiment) or collide with another particle to achieve a higher Center-of-Mass (CM) energy. The higher the energy, the richer the species of the output particles [2]. The Tevatron, at the Fermi National Accelerator Laboratory (Fermilab), is a synchrotron that accelerates protons and antiprotons ($p\bar{p}$) up to almost 1 TeV and was the highest energy particle collider in the world[1]. Currently it is the only place in the world to produce and directly study top quarks, since the Tevatron's energy is higher than the top quark production threshold. Datasets on which this thesis is based were taken from Tevatron $p\bar{p}$ collisions.

Theoretically, the Standard Model of particle physics is a non-abelian gauge theory that

---

[1]The Large Hadron Collider (LHC), a proton-proton collider designed with CM energy of 14 TeV, is in operation currently at CM energy of 2.36 TeV.

explains electroweak and strong interactions with sets of fields and the gauge symmetries $SU(3)_C \times SU(2)_L \times U(1)_Y$. The Standard Model is very successful since it explains almost all available experimental data in particle physics. The Higgs mechanism in the Standard Model is used to understand the origin of mass and plays a decisive role in validating the Standard Model [3, 4]. The existence of the Higgs boson is not yet experimentally confirmed, though it may be discovered by experiments at the LHC in the near future.

The top quark is a Standard-Model matter particle and the heaviest among all known elementary particles. Top quark pairs were discovered by the CDF and D0 experiments at the Tevatron in 1995 and first evidence of electroweak top quark production was published by D0 in 2006. Electroweak top quark production is of interest for several reasons. One outstanding reason is that it can be used to measure $|V_{tb}|$, a Cabibbo-Kobayashi-Maskawa matrix (*CKM*) element, indicating the strength of electroweak coupling between a bottom quark and a top quark. If the *CKM* matrix is not unitary, an extra generation of quark flavors may exist. This thesis is reporting a measurement of electroweak top quark production in the tau+jets channel by boosted decision trees at the D0 experiment in 4.8 fb^{-1} of collision data.

Chapter 2 briefly introduces elementary matter and its interactions in the Standard Model. Top quark physics, including the motivation for measuring electroweak top quark production, is also included in the chapter. Chapter 3 introduces the Tevatron accelerator chain and the detector system of the D0 experiment. Data and MC samples including trigger simulations are described in Chapter 4 and object identifications (ID) and event selections in Chapter 5. The technique of boosted decision trees (BDT) is used in both the tau ID optimization and electroweak top signal discrimination. BDT are discussed in Chapter 6. The default D0 tau ID algorithm, which is based on neural networks and $Z \to \tau\tau$ samples, is not optimal for this search. So, a tau ID optimization is performed and is described in Chapter 7. The background model will be presented in detail in Chapter 8. In Chapter 9, a Bayesian statistical approach is applied to study experimental sensitivity and calculate cross sections considering various systematic uncertainties. The chapter also presents cross section combination results between the tau+jets channel and electron/muon+jets channels and a measurement of $|V_{tb}|$. Chapter 10 is a summary of the analysis.

Chapter 2

Standard Model and Top Physics

The Standard Model of particle physics is a successful theoretical framework which describes elementary particles and their fundamental interactions. Section 2.1 gives a brief introduction to elementary matter particles and their interactions in the Standard Model. In Section 2.2, top quark physics, including top quark pair production and single top quark production, is briefly introduced and motivation for the measurement presented in this thesis is provided.

2.1 The Standard Model

2.1.1 Elementary Particles

All elementary particles in the Standard Model can be classified as leptons, quarks and elementary bosons. Leptons and quarks are fermions with spin=1/2 and are the matter particles which build our universe. Leptons and quarks are each organized in three generations or families. Gauge bosons are force carriers and are responsible for interactions between particles. Table 2.1 summarizes some important properties of leptons and quarks.

2.1.2 Elementary Interactions

There are four known and fundamental interactions in nature: electromagnetic (EM), weak, strong and gravitational. The first three are described by the Standard Model.

Table 2.1: Fundamental fermions in the Standard Model [5]

		Generation I		Generation II		Generation III	
Quarks	Flavor	**Up (u)**	**Down (d)**	**Charm (c)**	**Strange (s)**	**Top (t)**	**Bottom (b)**
	Charge	$+\frac{2}{3}$e	$-\frac{1}{3}$e	$+\frac{2}{3}$e	$-\frac{1}{3}$e	$+\frac{2}{3}$e	$-\frac{1}{3}$e
	Mass (MeV)	1.5 to 3.3	3.5 to 6.0	$\sim$1270	$\sim$104	$\sim$171200	$\sim$4200
Leptons	Flavor	**Electron (e)**	**Electron neutrino (ν_e)**	**Muon (μ)**	**Muon neutrino (ν_μ)**	**Tau (τ)**	**Tau neutrino (ν_τ)**
	Charge	-e	0	-e	0	-e	0
	Mass (MeV)	0.511	$< 2.2 \times 10^{-6}$	105.7	$< 1.7 \times 10^{-4}$	1777	< 15.5

EM Interaction The classical theory of the EM interaction was formulated by Maxwell over one hundred years ago. Quantum ElectroDynamics (QED) is a quantum field theory of electromagnetism and the is the most successful of all dynamical theories. Any interaction which is mediated by a photon is an EM interaction.

Weak Interaction The weak interaction was first parameterized by Fermi in the 1930s in his theory explaining β decay in which the weak transition rate is proportional to the strength of the coupling between four contact fermions. All fundamental particles, leptons and quarks, join in the weak interaction. The weak flavor group in which quark and lepton doublets are basic representations is asserted to be $SU(2)$. The weak force carriers are the charged bosons $W^\pm$ and neutral Z^0 boson. The weak and EM interactions are unified by a $SU(2) \times U(1)$ gauge theory in the Standard Model. The theory predicts the existence of four force carriers and explains why three of them are massive and one among them is massless via the Higgs mechanism [3, 6]. The Higgs mechanism implies the existence of a Higgs boson, however, it does not predict the mass of the new particle.

Strong Interaction The theory describing dynamics of the strong interaction is called Quantum Chromodynamics (QCD). In QCD, color plays a similar role to charge in QED. The strong interaction is mediated by massless gluons coupled to the color charge of quarks. In QCD, the interaction terms of the colored quarks with the vector

fields in a Lagrangian requires the existence of vector gauge fields. The strong gauge group is $SU(3)$ where 3 is for three colors. Thus there are 8 generators and hence 8 kinds of gluons.

Gravitational Interaction Since the gravitational interaction is very weak, it is often ignored in the context of other three forces. It is not part of the Standard Model of particle physics and will not be discussed here.

Table 2.2 is a summary table comparing information from the four fundamental interactions in nature.

Table 2.2: Summary table comparing properties of the four known fundamental interactions in nature. Relative strengths are estimated at distances $\sim 10^{-18}$ cm [7].

Interaction	Electromagnetic	Electroweak	Strong	Gravitational
Relative Strength	10^{-2}	10^{-5}	1	10^{-39}
Mediator	Photon	W and Z	Gluon	Graviton
Gauge Symmetry	U(1)	Broken SU(2)$\times$U(1)	SU(3)	–
Range	Infinite	$\sim 10^{-18}$ m	$\sim 10^{-15}$ m	Infinite

2.2 Top Quark

2.2.1 Discovery

In the Standard Model, all left-handed quarks exist in weak isospin doublets. For example, the up quark u_L isospin partner is the down quark d'_L. When the bottom quark was discovered at Fermilab in the 1970s, it was predicted that there should exist a new left-handed isospin partner. After almost two decades of searching, the top quark was directly discovered by the D0 and CDF experiments in $\sqrt{s} = 1.8$ GeV $p\bar{p}$ collisions at the Fermilab Tevatron Collider [8]. This discovery was based on top quark pair production via the strong interaction. Reviews of recent measurements of top quark properties and interactions at the Tevatron can be found in Refs. [9–11].

More than a decade later, the first evidence of the production of top quarks via the weak interaction was published [12] by D0 and observation was published by D0 [13] and

CDF [14] independently. The analysis in this thesis is the first search for electroweak top quark production using tau leptons.

2.2.2 Top Quark Mass

The top quark's mass is comparable to that of a gold atom and it is the heaviest known elementary particle. It is also the only fermion with the mass greater than the mass of the weak force carriers, the *W* and *Z* bosons, and it may play a special role in the mechanism of electroweak symmetry breaking. The top quark mass and the *W* mass are two of the most important parameters constraining the Higgs mass in the Standard Model. The most precise top quark mass is measured in top quark pair samples. The current combined top quark mass is evaluated by the Particle Data Group from Tevatron datasets: 171.3±1.1±1.2 GeV [5].

2.2.3 Top Quark Pair Production
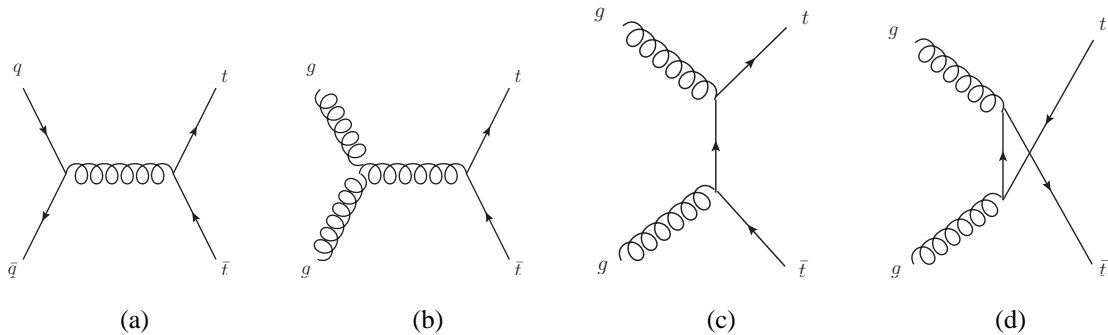
Figure 2.1: Leading order Feynman diagrams of pair production of top quarks via the strong interaction in hadron collisions through quark-antiquark annihilation (a) and gluon fusion (b), (c) and (d).

In $p\bar{p}$ collisions, top quark pairs ($t\bar{t}$) are produced in the strong interactions mainly through quark-antiquark annihilation and gluon fusion as shown in Fig. 2.1. Top pair production has been measured in various channels at the D0 and CDF experiments [5].

2.2.4 Top Quark Decays

In the Standard Model, the top quark decays to a W boson and a quark only by flavor changing weak interactions as shown in Fig. 2.2. It is predicted that $B(t \to bW) > 0.998$ and other decay modes are rare and difficult to extract. The top quark has a very short lifetime of approximately 4×10^{-25} s, although its decay is due to the weak interaction. It decays so quickly that it does not have time to form bound states and depolarise spin states. Unlike light quarks, the lifetime of the top quark is often associated with its intrinsic width. The W boson then decays to leptons with a branching fraction of 11% for each lepton type and to quarks with a branching fraction of 67%. The search performed in this analysis is done in the decay $t \to bW \to b\ \tau\nu_\tau$. The detector signature is an isolated τ, missing transverse energy ($\not{E}_T$) from neutrinos, and at least one b-tagged jet.
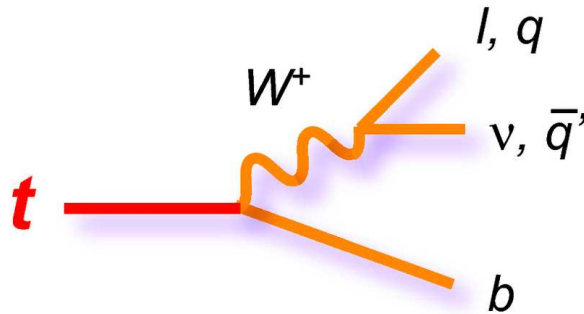
Figure 2.2: Feynman diagram for top quark decay.

2.3 Electroweak Top Quark Production

2.3.1 Single Top Quark Production and Searches

Top quarks can not only be produced in pairs by the strong interaction, but also singly by the electroweak interaction. At the Tevatron, electroweak top quark production is often called single top quark production. Single top quark production is accompanied by a bottom quark in the s-channel mode or by both a bottom quark and a light quark in the t-channel mode as illustrated in Fig. 2.3. The s-channel involves production of an off-shell and time-like W boson which decays into a top and a bottom quark. The t-channel is a W-gluon fusion mode involving the exchange of a space-like W-boson between a light quark and a bottom quark

resulting in a single top quark and a jet. Besides the *s*- and *t*-channels, there are associated *tW* processes at hadron colliders via $bg \to tW$ to generate single top quarks. However, because the processes have a small cross section at the Tevatron, they are often ignored. In this thesis, the goal is to search for the *s*- and *t*-channels at once while neglecting the *tW* process.
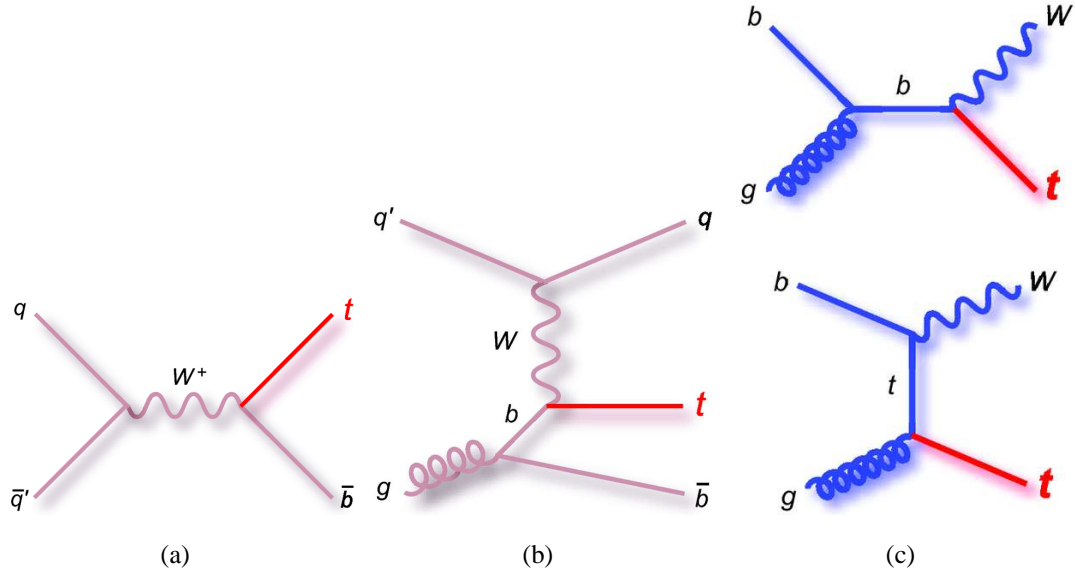
Figure 2.3: Feynman diagrams for single top quark production in (a) *s*-channel, (b) *t*-channel, and (c) *tW* processes. (a) and (b) are the single top quark production modes of interest at the Tevatron.

Cross sections of the *s*- and *t*-channels have been theoretically studied and predicted in many references suchas [15–18]. The theoretical cross sections 1.12 ± 0.04 pb for the *s*-channel and 2.34 ± 0.12 pb for the *t*-channel are used in this analysis as calculated in Ref. [18].

Considering the Standard-Model decay modes of the top quark and *W* boson, single top production results in four channels: electron plus jets, muon plus jets, tau lepton plus jets, and all jets channels. Evidence [12,19,20] and observation [13,14] of single top production with an electron or a muon in the final state (electron+jets or muon+jets channels) and the first direct measurement of $|V_{tb}|$ [12] have been published recently. The study in Ref. [14] also includes the result in the $\not{E}_T$ plus jets channel. However, this is the world's first direct search for single top quarks using tau leptons.

2.3.2 Motivation

There are several motivations for the measurement of single top quarks in the tau+jets channel. These are detailed below.

CKM Matrix and Single Top Quark Production

It is observed that the change of flavors of quarks or leptons is allowed only in the weak interaction. This occurs because there is a mismatch between quantum states of quarks in the weak interaction and in the strong interaction. A unitary matrix called the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix was introduced to characterize the strength of the couplings in the weak decays of quarks and specify the mismatch. This relationship can be formulated as follows:

$$
\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{bmatrix} d \\ s \\ b \end{bmatrix} \tag{2.1}
$$

where $[d', s', b']^T$ and $[d, s, b]^T$ are the eigenstates of quarks d, s and b which participate in the weak and strong interactions respectively. Magnitudes of some CKM matrix elements such as V_{ud} have been well determined experimentally. The current measured values are listed in the following matrix. [5]

$$
\begin{bmatrix} 0.97418 \pm 0.00027 & 0.2255 \pm 0.0019 & (3.93 \pm 0.36) \times 10^{-3} \\ 0.230 \pm 0.011 & 1.04 \pm 0.06 & (41.2 \pm 1.1) \times 10^{-3} \\ (8.1 \pm 0.6) \times 10^{-3} & (38.7 \pm 2.3) \times 10^{-3} & 0.77^{+0.18}_{-0.24} \end{bmatrix} \tag{2.2}
$$

The Wtb vertex through which the single top quark is produced contributes by the vertex factor

$$
-\frac{ig_w}{2\sqrt{2}} V_{tb} \gamma^\mu (1 - \gamma^5) \tag{2.3}
$$

to the matrix element for the production, where g_w is the coupling strength of the weak interaction. Since the cross section for a given process is proportional to the square of matrix element according to Fermi's "Golden Rule", thus the cross section of single top production is proportional to $|V_{tb}|^2$. By measuring the cross section, $|V_{tb}|$ can be extracted with no assumption on the number of flavor generations in the Standard Model. Direct extraction of

$|V_{tb}|$ is one of the primary motivations for measurements of single top production presented in this thesis.

With other well-measured CKM matrix elements $|V_{ub}|$ and $|V_{cb}|$, the direct measurement of $|V_{tb}|$ can also be used to prove the unitarity of the matrix by checking whether or not $|V_{ub}|^2 + |V_{cb}|^2 + |V_{tb}|^2$ is equal to 1 and indicate if any extra quark generation exists beyond the Standard Model.

Sensitivity to New Physics Beyond the Standard Model

Single top quark production can be affected by new physics through either unconventional weak interactions, or new particles or new mechanisms. For example, anomalous Wtb couplings would modify the rate of single top production and the angular correlations of top quark decay products. Recently the D0 experiment combined information from W boson helicity of top quark decay daughters and anomalous coupling searches in the single top final state [21] to present limits on the anomalous Wtb couplings [22]. Different modes of single top quark production (*s*- or *t*-channel production) are sensitive to different new physics beyond the Standard Model. A new heavy vector boson $W'^{\pm}$ could contribute additional processes that would affect the rates and kinematics of the *s*-channel production mode [23]. Recently, D0 published the first analysis to isolate an individual single top quark production channel and measured a cross section of $3.14^{+0.94}_{-0.80}$ pb for the *t*-channel and 1.05 ± 0.81 pb for the *s*-channel (see Fig. 2.4 for a 2D plot of the *t*-channel cross section vs. the *s*-channel cross section) [24]. Another interesting aspect is thepossibility of non-Standard-Model top quark couplings [25]. The flavor changing neutral current terms, for example, via a tZc, or tcg or tug [26] vertex can exhibit large effects on single top *t*-channel production [25].

Enhancement of Signal Acceptance

The tau lepton is a powerful tool often used to study physics in and beyond the Standard Model [27]. A tau lepton has more than 30 decay modes due to its heavy mass of 1.78 GeV. A pie chart in Fig. 2.5 shows tau lepton decay modes and their branching ratios. As shown in the chart, the leptonic channels of the tau decay have about35% branching fraction while the hadronic tau decays have about 65% branching fraction [5].
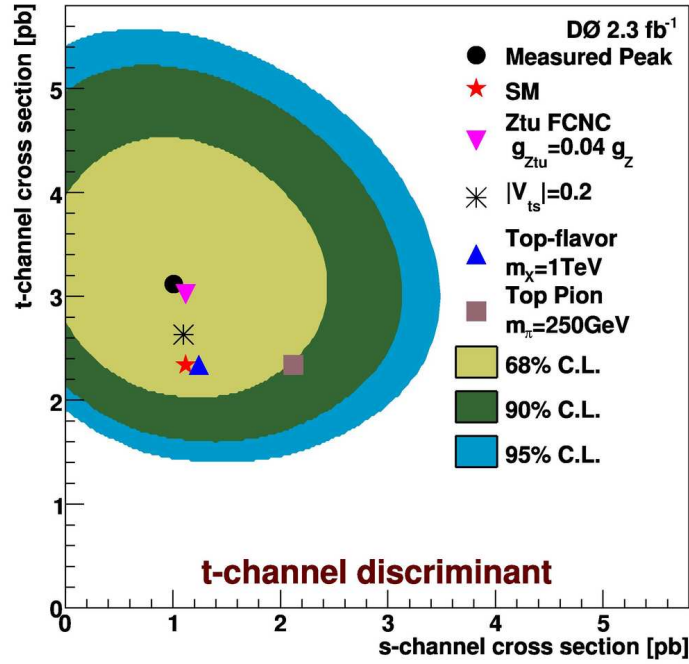
Figure 2.4: 2D cross section plot of *t*-channel vs. *s*-channel single top quark production. Contours are of equal posterior probability density. The points are the measured peak, the Standard Model expectation and several representative new physics models: flavor-changing neutral currents with a *Z* boson coupling to the top and up quark with a strength of 4% of the SM coupling, a four-quark-generation scenario with CKM matrix element $|V_{ts}| = 0.2$, a top-flavor model with new heavy bosons at a scale mx = 1 TeV, and a top-color model with a $t\bar{b}$ bound state (Top Pion) with a mass of $m_\pi = 250$ GeV. (after Fig. 5 of Ref. [24] and references herein.)
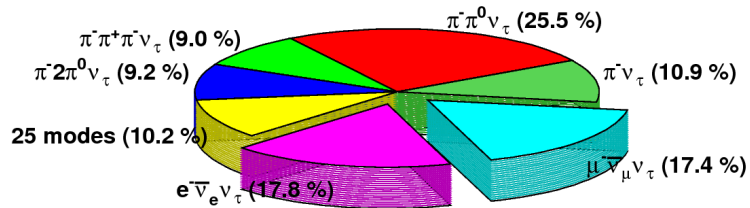
Figure 2.5: Pie chart of τ decay modes (Data from Ref. [5])

Although single top production has been observed in the electron and muon channels, the result is still statistically limited. Considering single top quarks in the tau lepton channel will increase signal acceptance. In this analysis, the hadronically decaying tau lepton will be identified. Adding the hadronic tau channel can increase single top quark acceptance as much as 30%, hence allowing more precise measurement of the single top cross section and $|V_{tb}|$.

Independent Single Top Quark Channel

The electron and muon channelshave many similarities. In particular, they suffer from the same dominant backgrounds (W+jets) and hence have very similar background models. This analysis has a completely different dominant background (QCD-multijets) and different sources of systematics. The tau+jets channel is thus an excellent independent check of the measurements in the other channels.

Higgs Search

As the heaviest lepton, the tau is related to many interesting signatures such as Higgs decays to taus. The best strategy to use in a Higgs search is very dependent on the unknown Higgs mass [28]. Fig. 2.6 shows branching ratios of the Higgs as a function of Higgs mass for the different decay channels predicted by the Standard Model. The tau lepton pair $\tau^+\tau^-$ is one of the most important decay modes for a low mass Higgs boson. This low mass region is favoured by the current electroweak data and is the focus of a lot of effort at the Tevatron. Single top quarks in the tau+jets channel are an important background to this search.
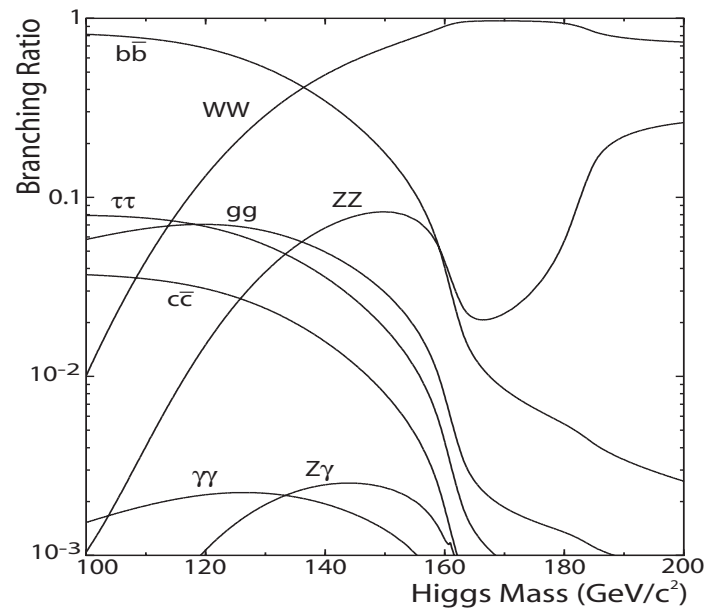
Figure 2.6: Branching ratios of Higgs for the different decay channels in the Standard Model

Chapter 3

Experimental Apparatus

Experimental data used in this analysis were generated by proton anti-proton ($p\bar{p}$) collisions at a centre-of-mass (CM) energy of $\sqrt{s} =$ 1.96 TeV and acquired by the D0 detector system. This chapter has two main sections: Section 3.1 introduces the Tevatron accelerator chain by which proton and anti-proton beams are boosted to high energy step by step; Section 3.3 presents the D0 detector.

3.1 Fermilab Accelerator Chain

The Tevatron [29–32] at Fermilab was the highest energy particle collider in the world. Fermilab is located in Batavia near Chicago, IL, USA. Fig. 3.1 is an aerialphotograph of Fermilab. The Tevatron has two experiments: D0 and Collider Detector Facility (CDF), both of which announced the discovery of top quarks in 1995. The Tevatron collides a beam of protons with a beam of anti-protons, and there are five accelerators with different purposes in the Fermilab accelerator chain: Cockcroft-Walton pre-accelerator, LINear ACcelerator (LINAC), booster, main injector and main ring. All of them except the Cockcroft-Walton are shown in Fig. 3.2. Below is a brief introduction to these components of the Tevatron acceleration chain.
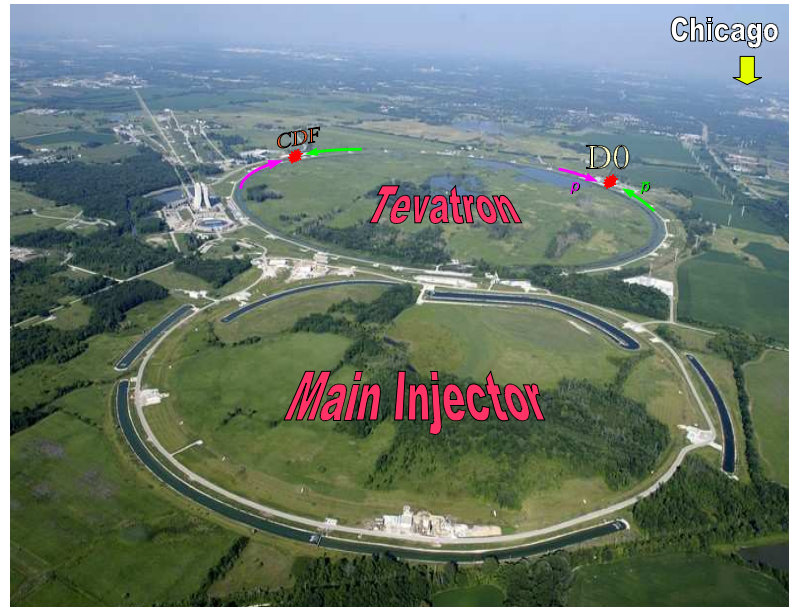
Figure 3.1: Aerial view of Tevatron at Fermi National Accelerator Laboratory located to the west of Chicago, Il, USA [33]

3.1.1 Acceleration of Protons

The proton source is ionized hydrogen gas (H^-) which is pre-accelerated to an energy of 750 keV by a Cockcroft-Walton pre-accelerator [34]. The negatively charged hydrogen ion beam from the Cockcroft-Walton pre-accelerator is led to the LINAC [34]. The LINAC is 130-meters long and is composed of metallic drift tubes isolated by vacuum gaps. It uses a Radio Frequency (RF) technique to accelerate the hydrogen ions to an energy of 400 MeV. Following the LINAC, the booster provides a boost to an energy of 8 GeV. It consists of a series of magnets arranged around a circle with a 75 meter radius [34]. 400 MeV negative hydrogen ions from the LINAC are guided to the booster and have their electrons stripped off by a thin Carbon foil, which leaves only the protons, then it accelerates the protons to 8 GeV.

Upon leaving the booster, the proton beam enters the main injector by a transport enclosure line and is ready to circulate for further acceleration by the main injector [35, 36]. The main injector is also a synchrotron that was built in a Tevatron upgrade (Run II). It can provide two energies of beams: 120 GeV and 150 GeV. During fixed-target operations, the main injector accelerates protons to the desired energy of 120 GeV followed by extraction

FermilabTevatron Accelerator With Main Injector
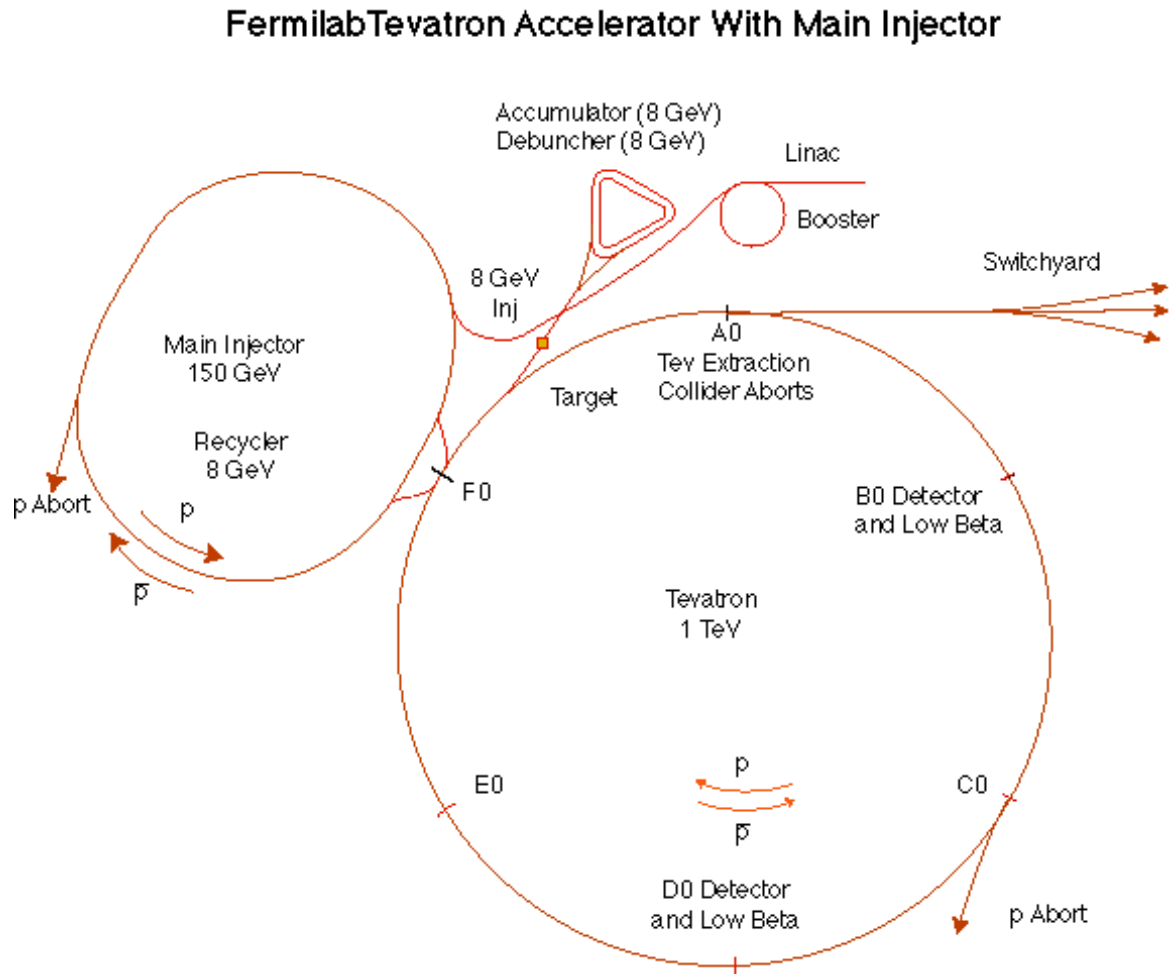
Figure 3.2: Schematic illustration of the Tevatron accelerator chain

to a fixed-target. The 120 GeV protons are used either for anti-proton generation, which will be introduced later, or a separate neutrino experiment called Neutrinos at the Main Injector (NuMI). The main injector also provides injection of 150 GeV protons into the next accelerator: the Tevatron main ring.

The Tevatron main ring is the last and largest accelerator in the chain where protons and anti-protons are accelerated from 150 GeV to 980 GeV each and are collided at two experiments: CDF and D0 [37]. It was the world's first superconducting synchrotron, starting its first operation in 1983. The Tevatron consists of about 1000 superconducting magnets running at $4.3°$ K providing a magnetic field strength of 4.2 Tesla. These magnets bend proton and anti-proton circulating beams traveling in the circle with a $\sim$2 km diameter.

3.1.2 Generation of Anti-protons

Generation of anti-protons is realized by smashing the 120 GeV proton beam from the main injector onto a Nickel target [38]. Anti-protons from reactions between protons and target material have a very large energy spread, so it will be difficult for downstream accelerators to accept. A debuncher after the target station is employed to switch the wide energy spread and the narrow time spread at the RF cavity. Then anti-protons are sent to an accumulator and a feedback stochastic cooling system so that they can be collected successively and stored for further acceleration. Like protons, anti-protons with energy of 8 GeV will be accelerated to 150 GeV in the main injector and to 980 GeV in the Tevatron main ring.

3.1.3 Tevatron Operation

Data have been collected in two run periods: Run I (1992-1996) and Run II (2001-present). During Run I, the Tevatron collided proton and antiproton beams at a center-of-mass energy of $\sqrt{s} =$1.8 TeV. Protons and anti-protons each had 6 bunches per ring and the interval between bunch crossings was 3500 ns. During the Run I period, the typical peak luminosity was 1–2$\times 10^{31}$ cm^{-2}s^{-1} and the D0 experiment recorded approximately 120 pb^{-1} of data. After Run I, the Tevatron was shut down for upgrading. Run II started in March, 2001. After upgrading, including building the new main injector, the CM energy increased to 1.96 TeV, the number of bunches per ring per speciesincreased to 36, and the interval between bunch

crossings is shortened to 396 ns [39]. Fig. 3.3 shows the Run II integrated luminosity delivered to and recorded by D0 as a function of time.
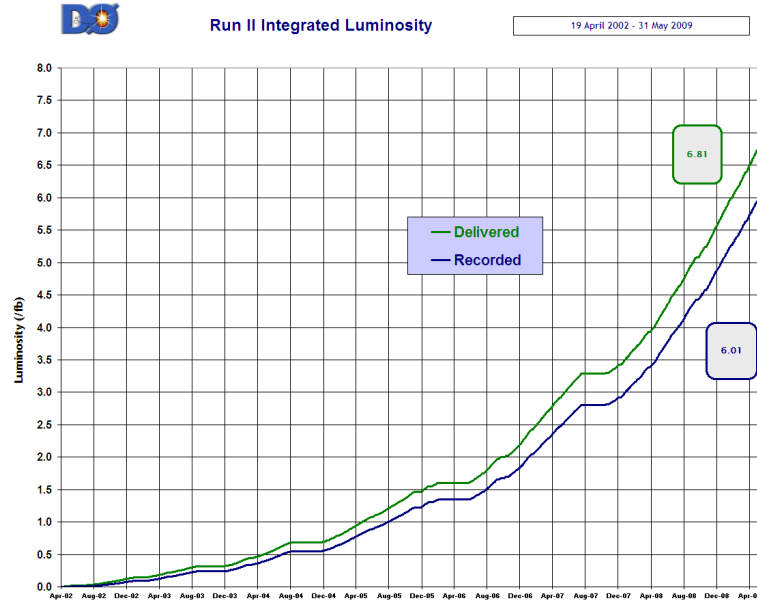
Figure 3.3: Run II integrated luminosity recorded by the D0 experiment.

3.2 High Transverse Momentum Physics

Due to the high CM energy of the incident p and $\bar{p}$, the outgoing particles span a large range of final states and possible kinematic features. Inelastic $p\bar{p}$ collisions, in which one or more particles are scattered at low angles with respect to the beam, dominate the total $p\bar{p}$ cross section. However, the most interesting physics processes take place in hard scattering by which partons in p ($\bar{p}$) are scattered such that large energy is transferred between them [28]. Cross sections of the processes partially rely on the Parton Density Function (PDF), the probability density to observe a parton of a certain momentum fraction within a hadron (p or $\bar{p}$).

Figure 3.4 illustrates two of the most common cases in hard scattering: (a) shows hard scattering that produces outgoing partons with high transverse momentum (p_T), particles from initial and final state radiation, and underlying particles from the remnants of p ($\bar{p}$); (b)
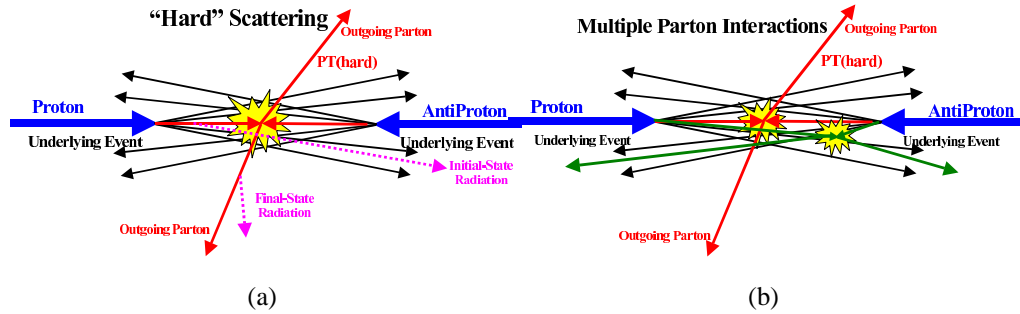
Figure 3.4: Two QCD processes in hard scattering during $p\bar{p}$ collisions. (a) a hard parton scattering with high transverse momentum p_T (b) a multiple parton interaction (after Ref. [40])

shows a similar process in which there are multiple parton interactions. Events belonging to these two cases are referred as QCD-multijet events since the two cases involve the strong interaction and several jets are present at the detector level[1] The QCD-multijet events are the main background in studies of electroweak phenomena. With even higher energy transfer, some heavier particles like *W* bosons and top quarks are created. The topic of this thesis is the measurement of electroweak top quark production in high p_T collisions, and its dominant background is from QCD-multijet events.

3.3 The D0 Detectors

3.3.1 Overview

The detection of particles is made possible by their interactions with matter. The working principle is similar to an eye seeing images. We can see the world around us because light enters our eyes and interacts with the retina to produce the signals that our brain then analyzes. Shortly after collisions, there are many particles generated. Experimentalists need to employ a "retina", i.e., a detector system, to collect information about emitted particles and thus study physics processes of interest. Such information includes momentum, energy and trajectory.

---

[1]A collision which has been recorded without any selection criteria applied (i.e. no high transverse energy requirement [9]) is referred to as a minimum bias event. These real data events are superimposed on MC events in order to simulate multiple interactions and pileup.
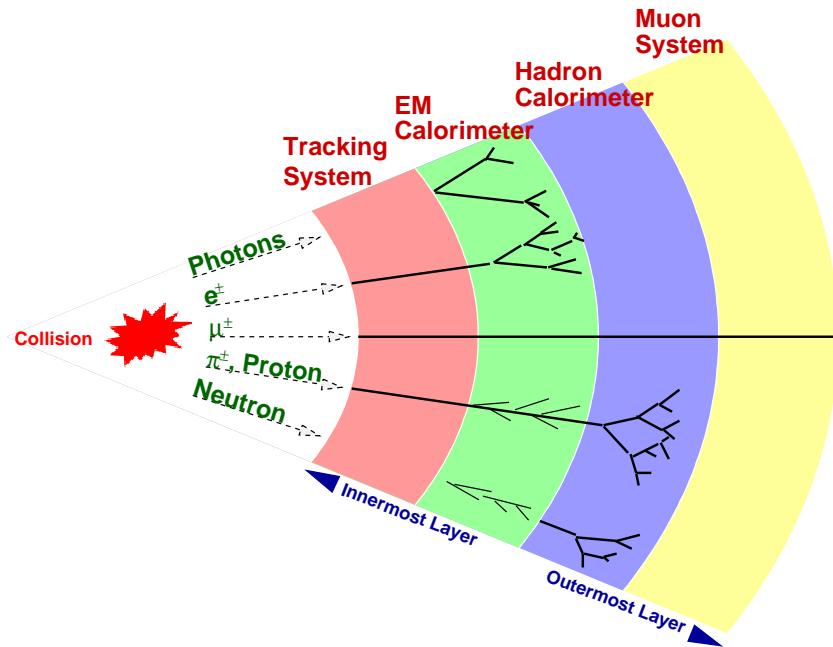
Figure 3.5: Illustration of detecting different particle types after collisions.

Different interactions between incident particles and matter require different types of detectors. Fig. 3.5 illustrates how a typical detector system works for different particles. From innermost to outermost layers, there is a tracking system, an electromagnetic (EM) calorimeter, a hadron calorimeter and a muon system. Some of the most common particles are photons, electrons/positrons, muons, pions, protons and neutrons and are shown in the figure. Photons do not carry electric charge hence their trajectory can not be bent by the magnetic field, however, they do deposit energy in the EM calorimeter. Electrons/positrons ($e^{\pm}$) have charges and their momentum can be measured by the tracking system without too much energy lost before they deposit their energy in the EM calorimeter. High energy EM showers are developed via bremsstrahlung and electron/positron pair production. A muon (μ) has a low interaction rate with matter and at Tevatron energies is expected to behave as a minimum ionizing particle. So, if a fast charged particle passes through a large number of absorbers in the calorimeter with minor energy losses and small angular displacement, then such a particle is usually identified as a muon. Energetic secondary hadrons like protons, neutrons and pions form hadronic showers and eventually lose energy mainly by ionization. However, not all the incident energy due to nuclear interactions,

particle recoil and production of neutrinos *etc.* is recorded [7, 41, 42].

Responding to the improvements in the Tevatron mentioned in Section 3.1.3, the D0 detector system was also significantly upgraded to enhance its detection ability for the physics reach of the experiment. Fig. 3.6 shows a diagram of the upgraded D0 detector system. The D0 detector system consists of 3 major subsystems: the central tracking system, the EM and hadron calorimeters and the muon spectrometer. Section 3.3.3 describes the tracking system. Section 3.3.4 introduces the EM and hadronic calorimeters. The muon spectrometer is presented in Section 3.3.5. Introduction to the luminosity detector is in Section 3.3.6. Section 3.3.7 shows D0's trigger system and data acquisition system. The original D0 detector before upgrading is presented by Ref. [43] and the detector after upgrading can be referred to Ref. [39].
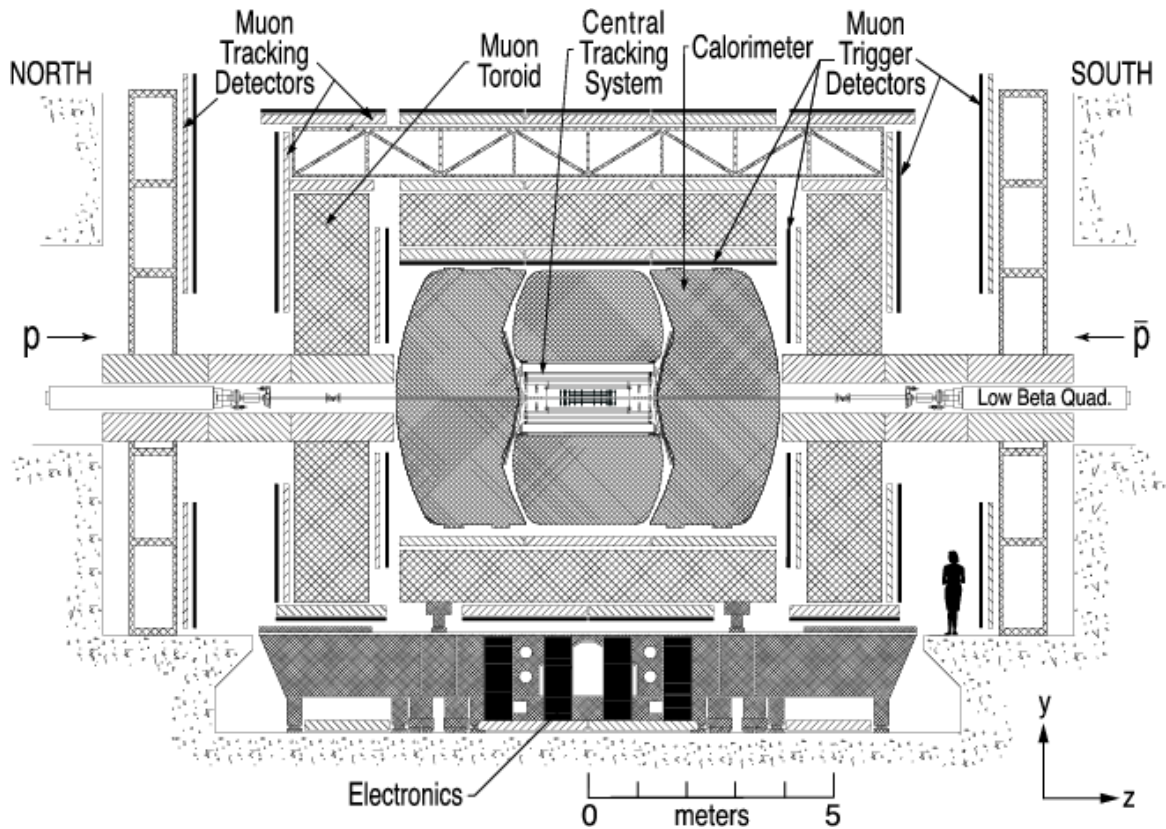
Figure 3.6: Schematic diagram of the upgraded D0 detector viewed inside the Tevatron ring (after Ref. [39]).

3.3.2 Coordinate System of D0 Detector

A right-handed coordinate system as shown in Fig. 3.7 will be used in the detector description and data analysis: the proton beam is along the *z*-axis direction and the *y*-axis is upward in the detector; the angles ϕ and θ are the azimuthal and polar angles respectively. The pseudorapidity is defined as $\eta = -\ln[\tan\theta/2]$, which approximates the rapidity, $y = 1/2\ln[(E + p_z c)/(E - p_z c)]$, as $(mc^2/E) \to 0$.
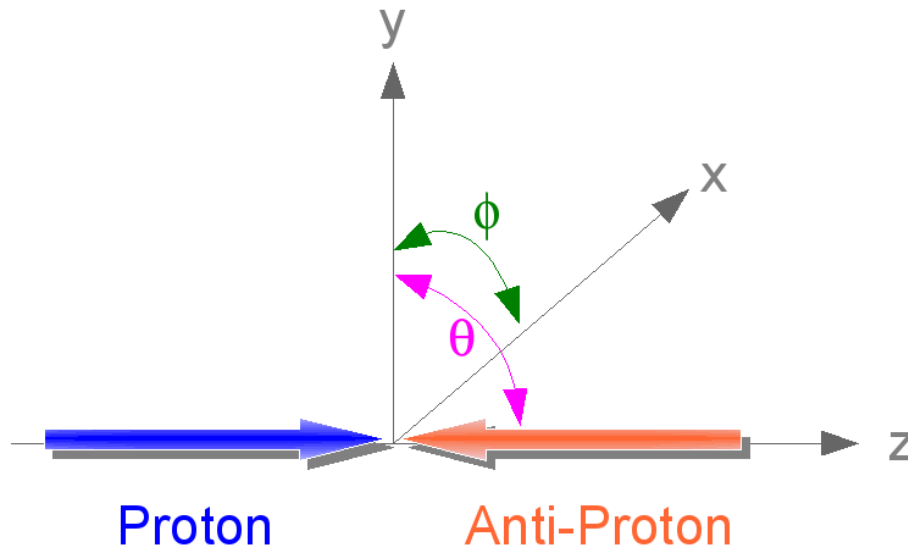
Figure 3.7: Coordinate system of the D0 detector

3.3.3 Central Tracking System

The D0 central tracking system is used to detect the passage of charged particles, to measure their momenta and to locate primary and secondary vertices. It consists of the Silicon Microstrip Tracker and the Central Fiber Tracker both of which are located inside a solenoidal magnet with a field of 2 T [39].

Silicon Microstrip Tracker (SMT)

The D0 SMT is designed to provide high resolution position measurements of charged particles for both tracking and vertexing over almost the full η coverage of the calorimeter and muon spectrometer [39]. It is located immediately outside the Tevatron beryllium beam
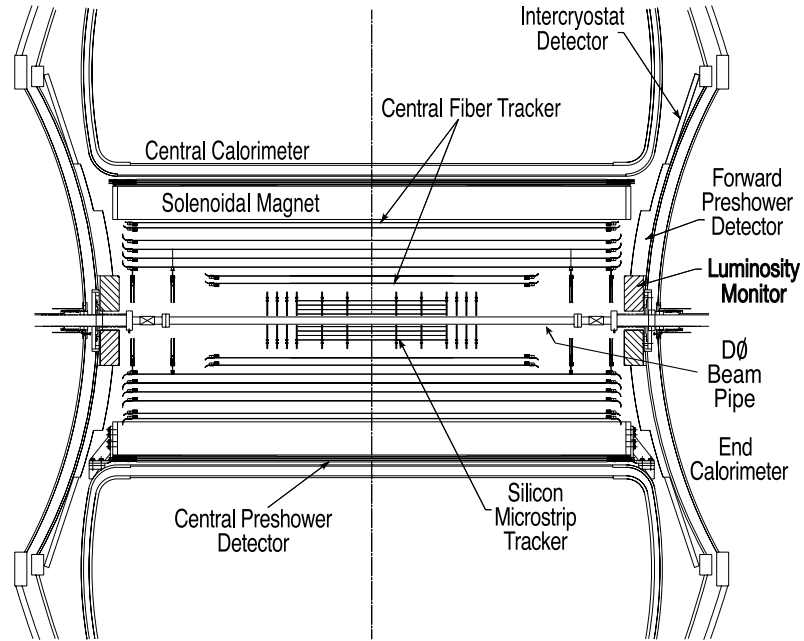
Figure 3.8: Schematic diagram of the D0 central tracking system [39].

pipe. As a charged particle traverses the silicon layers, electron-hole pairs are generated along its track and drift in anelectric field towards active sensors.

The length scale of the SMT is comparable with the scale of the interaction region ($\sim$25 cm). In order to meet the requirements that tracks are generally perpendicular to the detector surface for all η, the SMT uses a design of barrel modules interspersed with disk modules in the center region and in the forward regions. A 3D view of the SMT is shown in Fig. 3.9.

Central Fiber Tracker (CFT)

The CFT, located outside the SMT, occupies a radial space from 20 to 52 cm from the center of the beam pipe [39]. Scintillating fibers are mounted on eight concentric cylinders. The two innermost cylinders are 1.66 meters long in order to provide room for the forward SMT disks while the six outer ones are 2.52 meters long. With this geometry, the η coverage is $\leq$1.7. Two sets of fibers are supported by each cylinder, one set is oriented along the beam direction, another is rotated 3° with respect to the beam line. The base cores of the scintillating fibers are made of polystyrene doped with the organic fluorescent
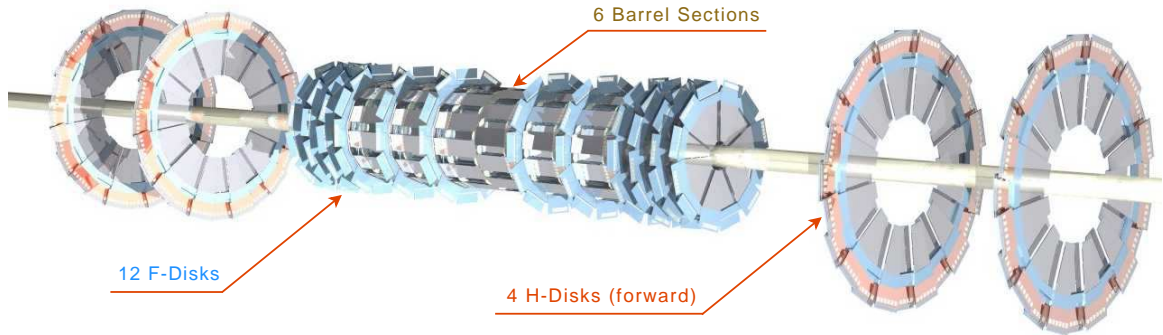
Figure 3.9: Isometric view of silicon microstrip tracker [39].

dye paraterphenyl. The small fiber diameter, 835 μm, provides a resolution of about 100 μm. The scintillation light from the fibers is led via clear fiber waveguides to Visible Light Photon Counters (VLPCs) for read out.

3.3.4 Calorimetry

The D0 calorimeter system consists of three sampling calorimeters: one central calorimeter and two end calorimeters, and an intercryostat detector [39]. Fig. 3.10 shows a schematic 3D view of the system. The Central Calorimeter (CC) covers $|\eta| \lesssim 1$ while two End Calorimeters (EC): South EC and North EC extend coverage $|\eta|$ up to 4.0. The region between CC and EC is filled by an intercryostat detector covering the pseudorapidity region $0.8 < |\eta| < 1.4$. The calorimeters each work within their own cryostat with temperature at $\sim$90 K. The calorimeters' active medium is liquid argon. The absorber plates in the EM sections are made from uranium while those in the fine hadronic sections are made from uranium-niobium alloy and those in the coarse hadronic sections are copper or steel.

A typical readout cell unit of the calorimeters is shown in Fig. 3.11. The metal absorber plates are grounded and the signal boards with resistive surfaces are connected to positive high voltage (2.0 KV). With this structure, the electric field between an absorber plate and a signal board causes electrons to drift in the liquid argon gap. Several such pads having the same η and ϕ are ganged together in depth and compose a typical readout cell unit.

Fig. 3.12, a cross section view of one quarter of the D0 calorimeters, shows a transverse
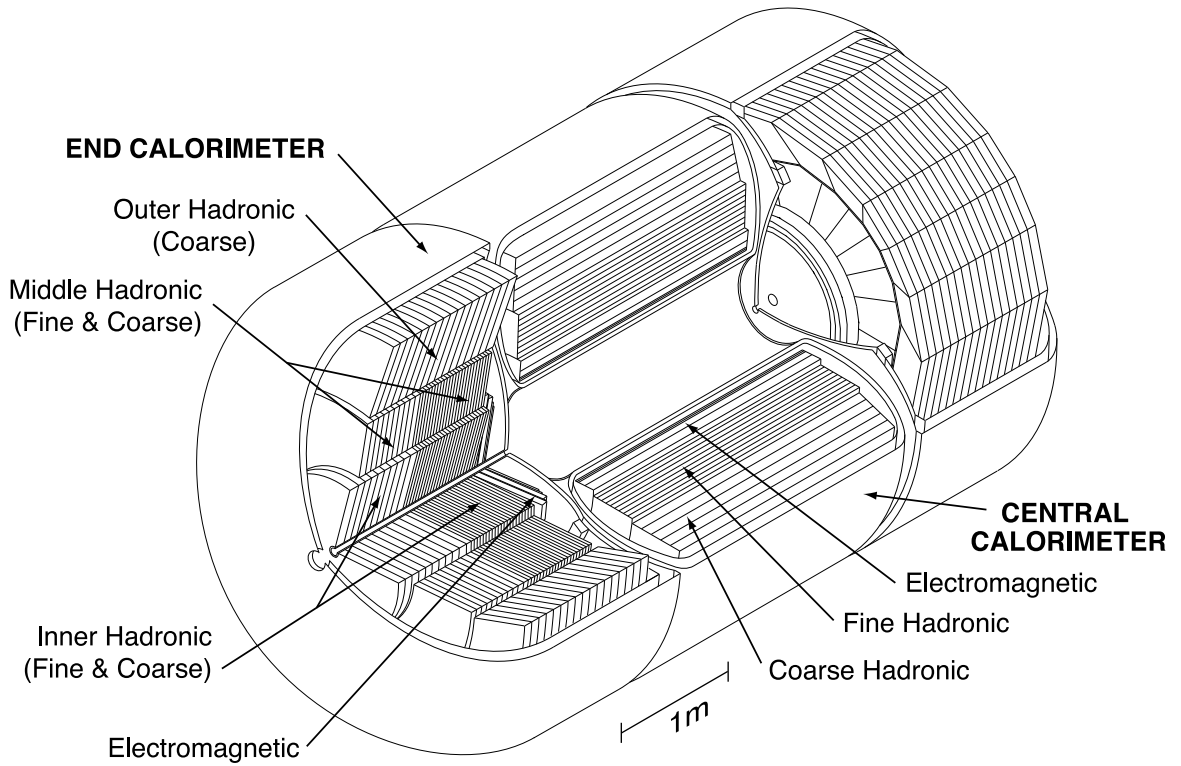
Figure 3.10: Schematic 3D view of the D0 calorimeter system [39].
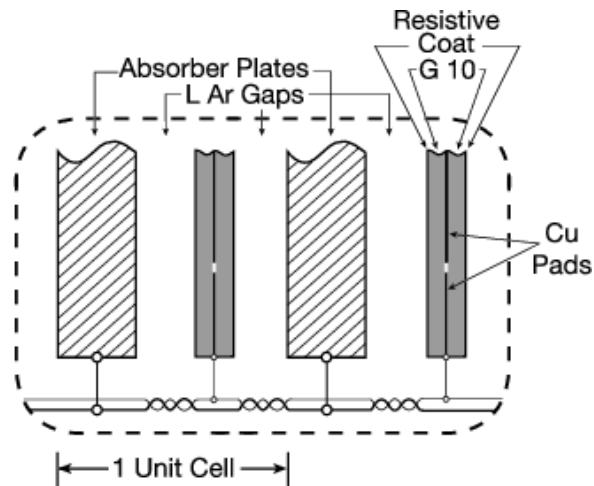
Figure 3.11: A typical cell unit of the liquid argon gap and absorbers [39].
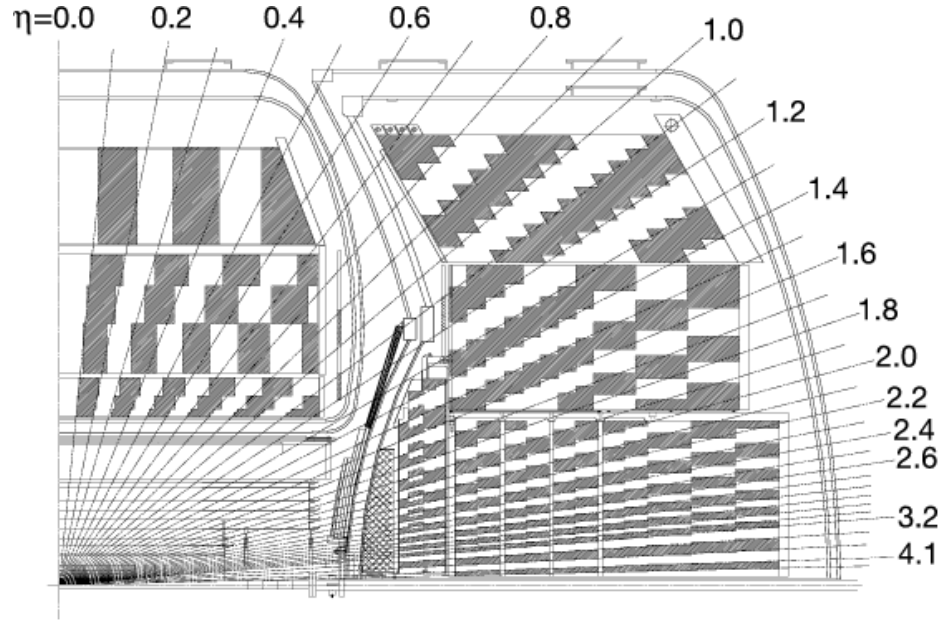
Figure 3.12: Cross section view of one portion of the D0 calorimeters [39].

and longitudinal segmentation pattern. Groups of the cells that are organized together for signal readout are indicated by the shading pattern. The cells within one pattern form pseudo-projective towers because the centers of the cells are on the rays projecting from the center of the interaction region. Subsequent object reconstruction benefits from this feature. Sizes of regular towers in both EM and hadronic sections are $\Delta\eta = 0.1$ and $\Delta\phi = \frac{2\pi}{64} = 0.1$.

The EM modules in the CC and ECs are arrayed in four separate depth layers. The layers in the CC are 1.4, 2.0, 6.0 and 9.8 radiation lengths (X_0) thick and ones in the ECs are 1.6, 2.0, 6.8 and 9.8 X_0 thick. An EM shower reaches maximum at the location around the 3rd layer of the EM section, so in this layer EM modules are split twice as finely in both η and ϕ in order to conduct more precise measurement of EM showers. The fine hadronic modules in the CC are organized in three groups in radius with thicknesses of 1.3, 1.0 and 0.76 absorption lengths (λ_A) while thickness of the single coarse section is about 3.2 λ_A.

3.3.5 Muon Spectrometer

The D0 muon spectrometer system [39] is the outermost layer of the D0 detector system. It is visible in Fig. 3.6, and it consists of Proportional Drift Tubes (PDTs) in the central region, Mini Drift Tubes (MDTs) in the forward regions, toroidal magnets and scintillation counters.

The central muon system covers the pseudo-rapidity range $|\eta| \leq 1.0$ using the PDTs while the forward one extends the muon detection to $|\eta| \sim 2.0$ using the MDTs. One layer (A-layer) among the three layers of the central PDTs is located inside the central magnet while another two are outside the central magnet. Position resolution of the drift chambers is ~ 1 mm.

The muon trigger scintillation counters in the muon system contribute to background rejection by providing triggering. More information can be found in Ref. [39, 44].

3.3.6 Luminosity Monitor

The D0 Luminosity Monitor (LM) system is used to (I) monitor the Tevatron luminosity near the D0 interaction area by detecting inelastic $p\bar{p}$ collisions, (II) estimate beam halo and (III) measure the z coordinate of the interaction vertex [39]. As shown in (a) of Fig. 3.13, two sets of twenty-four plastic scintillation detectors are mounted at $z = \pm 140$ cm with PhotoMultiplier Tube (PMT) readouts whose geometry and location are shown in (b) of Fig. 3.13. The detectors have a pseudorapidity coverage $2.7 < |\eta| < 4.4$ and occupy the area outside the beam pipe.

The luminosity $\mathcal{L}$ is calculated by

$$\mathcal{L} = \frac{f \cdot \overline{N}_{LM}}{\sigma_{LM}} \tag{3.1}$$

where f is the beam bunch crossing frequency, $\overline{N}_{LM}$ is the average number of inelastic collisions per beam crossing collected by the LM, and σ_{LM} is the effective cross section of $p\bar{p}$ which takes into account the acceptance and efficiency of the LM.

The particles from the beam halo are backgrounds when measuring average numbers of $p\bar{p}$ interactions. A technique using precise time-of-flight measurement is used to distinguish the particles in $p\bar{p}$ interactions from those of the beam halo. A particle hits the two LM detectors located at ± 140 cm at the times t_+ and t_-. The difference in the z coordinates of

the interaction vertex z_v by time-of-flight is $z_v = \frac{c}{2}(t_- - t_+)$. If we apply a cut $|z_v| < 100$ cm, then nearly all $p\bar{p}$ collisions are encompassed.
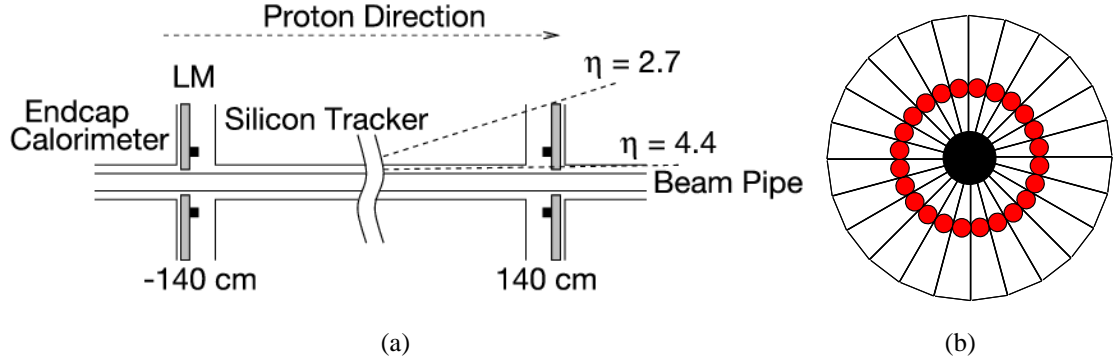
Figure 3.13: Schematic view of (a) the location of the D0 LM detectors; (b) the geometry of the LM and the locations of the PMTs (red solid circles) [39].

3.3.7 Trigger and Data Acquisition Systems

The D0 Run II trigger system was upgraded to meet the requirements from the increased Tevatron luminosity and reduced bunch crossing intervals compared to Run I. The D0 Run II trigger system is a complex trigger system with three distinct levels. As shown in Fig. 3.14, the detector data readout is closely integrated with the trigger system. Level 1 and Level 2 triggers conduct hardware processing while Level 3 employs sophisticated algorithms to reduce the event rate to 50 Hz. Afterwards, the data are stored on tape by online hosts.

L1 reduces an input rate of 1.7 MHz to a L2-accepted rate of about 2 kHz. The trigger framework collects information from each L1 trigger and decides whether or not a specific event is to be accepted for the next level. The L2 trigger system is capable of input rates of up to 10 kHz and maximum output rates of 1 kHz. L2 is implemented with different preprocessors specific for different detectors and a global processor for integration of the data. Like L1, its subsystems include tracking, calorimeter, preshower and muon systems. The final L2 decision is made in the L2Global stage based on physics objects reconstructed in the preprocessors.

In order to enrich the physics samples and further reduce event rates to a level for
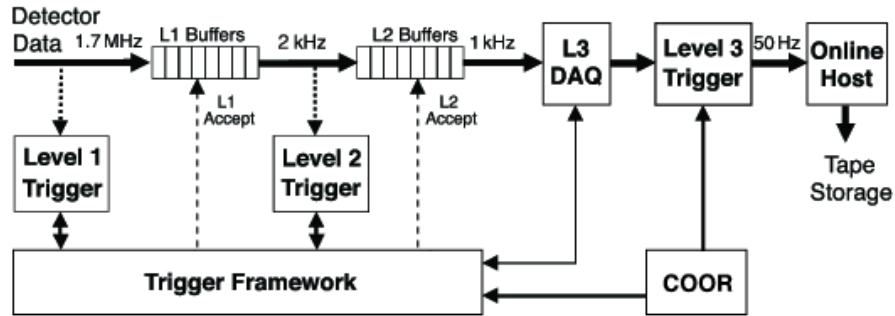
Figure 3.14: Data flowchart of the D0 trigger system and data acquisition system [39].

which the data can be stored to tape, the L3 trigger, as a high level and fully programmable software trigger, is used to do additional rejection with reduction of event rates from 1 kHz to 50 Hz [39]. Physics objects including electrons, muons, taus, jets, vertices and missing E_T are reconstructed by sophisticated algorithms at this level, and the L3 decisions are based on these objects or their relationships such as invariant mass, and spatial difference.

The data acquisition system (L3DAQ) is in charge of data transportation from the VME readout crates of the L2 system to the processing nodes of the L3 farm [39]. Commodity single-board computers coordinate to all nodes in the system.

Chapter 4

Datasets and Monte Carlo Samples

In this chapter, Section 4.1 introduces experimental datasets used in the study, including integrated luminosity. Section 4.2 presents triggers used to skim data and efficiency turn-on curves for simulated samples. Section 4.3 explains Monto Carlo samples used in the background model.

4.1   Data Samples

The data sample was collected between August 2002 and April 2009. In 2005, the Tevatron underwent a luminosity upgrade by a factor of 3. At the same time, the D0 experiment completed significant detector and trigger upgrades [45]. For example, the SMT suffered from radiation damage and a new radiation-hard inner silicon layer installed on the beam pipe at R=1.6 cm [46]. The trigger system also was improved at Level 1 in the calorimeter and tracker, and at Level 2 in the silicon trigger and software trigger [47]. With the upgrades, the D0 detector performs better, meeting the more stringer requirements of higher peak and integrated luminosity. For example, this leads to an increase in b tagging efficiency. So, it is reasonable to split data into Run IIa (until the 2005 upgrade) and Run IIb (since the 2005 upgrade) periods. The Run IIa and Run IIb raw datasets are reconstructed with different D0 software production releases. The D0 common sample group skims the whole dataset and provides analyzers with some skimmed sub-dataset definitions. A skim called the "new phenomena" (NP) skim is used in the analysis mainly because the definition of this skim includes interesting triggers to be introduced below. Integrated luminosity values

corresponding to different trigger versions are shown in Table 4.1. The total integrated luminosity of good quality data is 4.8 fb^{-1} of which Run IIa provides ~ 1.0 fb^{-1} and Run IIb ~ 3.8 fb^{-1}.

Table 4.1: Integrated luminosity values corresponding to different trigger versions used in this analysis. *Shutdown* in the table means the shutdown period during which the D0 detector system underwent maintenance and upgrades in 2007.

Integrated Luminosity [pb^{-1}]		
Channel	Triggers Version	Luminosity
	V11.0 - V12.0	63.1
	V12.0 - V13.0	227.4
Run IIa	V13.0 - V14.0	378.6
	V14.0 - V15.0	334.5
Run IIb (Pre-shutdown)	V15.0 - V15.2	208.7
	V15.2 - V16.2	1,006.1
Run IIb (Post-shutdown)	V15.2 - V16.2	2,553.7
Total Run II Integrated Luminosity		**4,772.1**

4.2 Triggers and Simulation

As of the date when the analysis was finished, there is no specific trigger designed for the tau+jets single top signal. Given this fact and the event topology that is being looked for (one narrow tau jet and two or more other jets), similar triggers to the ones used in the Higgs boson search in the $WH \to \tau_h \nu_\tau bb$ channel [48] are interesting and have been used.

4.2.1 Run IIa Triggers and Parameterization

The triggers used to skim Run IIa events are defined as:

Calorimeter jet trigger (MHT30_3CJT5)

– L1: there must be three calorimeter trigger towers with $E_T >5$ GeV

    – L2: missing transverse energy for jets is greater than 20 GeV

    – L3: vectorial sum of all jets E_T is larger than 30 GeV

Acoplanarity trigger (JT1_ACO_MHT_HT)

    – L1: three calorimeter trigger towers with E_T >5 GeV

    – L2: missing transverse energy for L2 jets is greater than 20 GeV; $\Delta\phi(\text{jet1},\text{jet2})$ <168.75°. Jet1 and jet2 are the two most energetic jets.

    – L3: at least one jet with p_T>9 GeV. Vectorial sum of all jets, p_T must be above 30 GeV and scalar sum of all jets, p_T with p_T>9 GeV is above 50 GeV. $\Delta\phi(\text{jet1},\text{jet2})$ <170°

MC samples do not have trigger level objects, so there is no direct way to apply the requirements above to simulate the triggers. The basic idea of the simulation of Run IIa triggers is to (1) start from simulated offline uncorrected jets and taus, and derive estimated and equivalent trigger level objects and (2) test the trigger requirements to determine if an event passes.

    – In the L1 parameterization, the probability that a certain number of towers are above a certain threshold is calculated as a function of offline jet p_T and in η regions. In the simulation, random numbers are sampled from the probability function in a certain η region and the numbers are counted as the simulated number of towers.

    – In the L2 and L3 parameterizations, equivalent L2 and L3 jet objects are "reconstructed" from offline uncorrected jets with p_T >15 GeV and uncorrected tau objects. The reconstruction of L2 and L3 employs a jet-shift-removal procedure in which effects from trigger-level jet resolution and reconstruction efficiency are combined. The trigger-level jet resolution functions are obtained by binning offline jetsp_T in 20 GeV intervals. The trigger-level jet reconstruction efficiency is also called a turn-on curve and is a function of offline jet p_T. L2 and L3 trigger turn-on curves for jets in the calorimeter regions, CC, ICR and EC, are shown in Figs. 4.1 and 4.2. The fitting functions shown are of the form $\frac{p_0}{|p_1+\exp{(-p_2 \cdot x)}|}$ where p_0, p_1 and p_2 are fitting parameters and x is jet p_T.
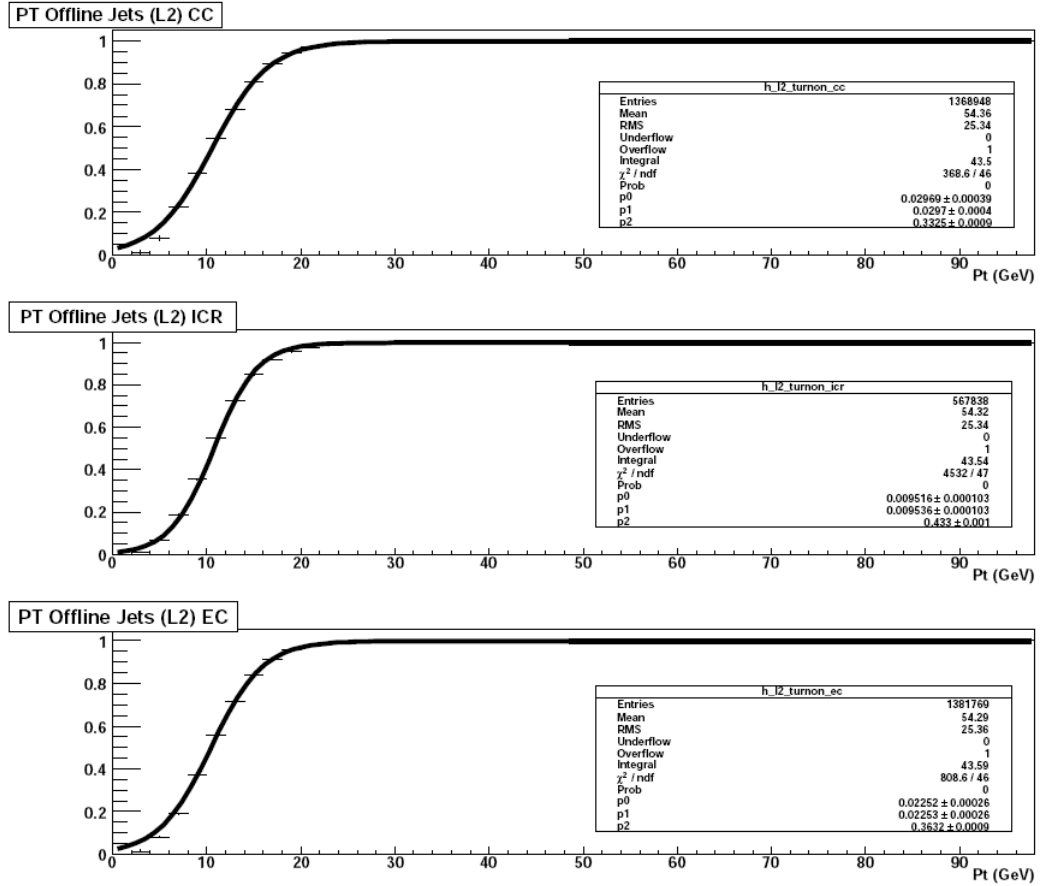
Figure 4.1: L2 jet trigger turn-on as a function of offline jet p_T for Run IIa (after Ref. [49]).

4.2.2 Run IIb Triggers and Parameterization

The triggers used to select Run IIb events in the analysis are defined as:

Acoplanarity trigger (JT1_ACO_MHT_HT) [1] with the same requirements as for Run IIa but with more complex L1 & L2 terms.

Missing H_T triggers (JT1_MHTACO, JT2_MHTACO, MJ_MHTACO) [2] require missing $H_T >35$ GeV where $H_T = \sum_{\text{trigger jets}} \vec{p}_T$, the two leading jets being acoplanar and the missing H_T vector (calculated using all trigger jets with $E_T >9$ GeV) being no closer than 25 degrees to any jet.

---

[1]The trigger version is from v15.00 to v15.19.

[2]The trigger version is from v15.20 to v16.20.
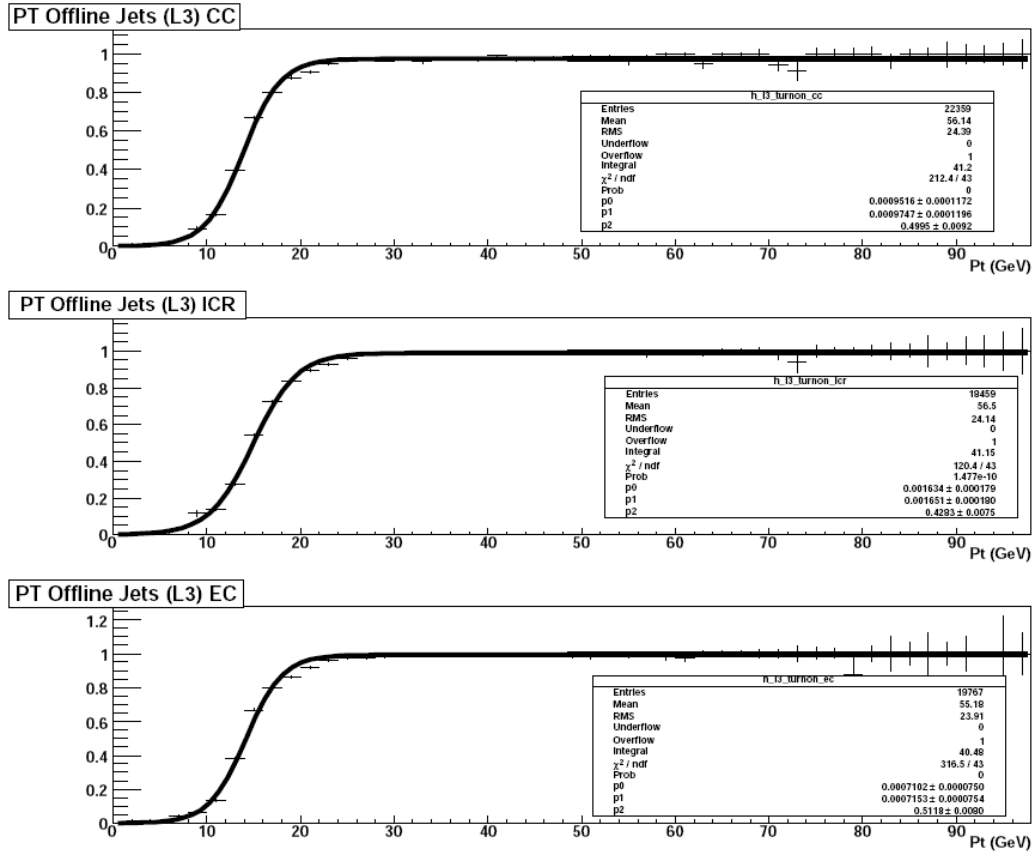
Figure 4.2: L3 jet trigger turn-on as a function of offline jet p_T for Run IIa (after Ref. [49]).

Missing E_T triggers (JT1_MET, JT2_MET, MJ_MET) [3] require missing H_T >25 GeV, missing E_T >25 GeV, the two leading jets being acoplanar and the missing H_T vector being no closer than 25 degrees to any jet.

The trigger parameterization for Run IIb is performed by modified software that was designed to simulate triggers of jet+$\not{E}_T$ and improved from its original form to include proper treatment of tau leptons. Below is a brief explanation of the trigger simulation strategy:

L1: Use trigger efficiency turn-on curves to determine the probability that a reconstructed object fired the trigger. Calculate the probability that the trigger is fired by combining the probability of different reconstructed objects. A random number is generated to compare the combined probability and check if the L1 trigger is fired.

---

[3]The trigger version is from v15.20 to v16.20.

L2: The trigger efficiency in data at this level is 100%

L3: Reconstructed objects are used to create equivalent trigger-level objects (i.e., L3 objects). Two quantities are used: the probability that a reconstructed object is matched to a trigger object and a correction to the transverse momentum of the reconstructed object. These simulated objects are used to reproduce a simulated H_T and missing H_T to check if the trigger is fired.

If the event passes selection of both L1 and L3 simulated triggers, then it is kept. The improvements from the original software include the derivation of new turn-on curves to simulate the L1 & L3 response to jets with a high EM fraction that are usually marked as bad. The L1 turn-on curves for tau leptons need to be derived. The L1 turn-on curves derived from data can be seen in Figs. 4.3, 4.4, 4.5 and 4.6. The figures are plotted in 3 tau decay types (tau type 1 and 2 correspond to 1-prong tau decay modes while tau type 3 corresponds to 3-prong decay modes. Section 7.2 will provide more detailed information about this classification).

(a)   (b)   (c)
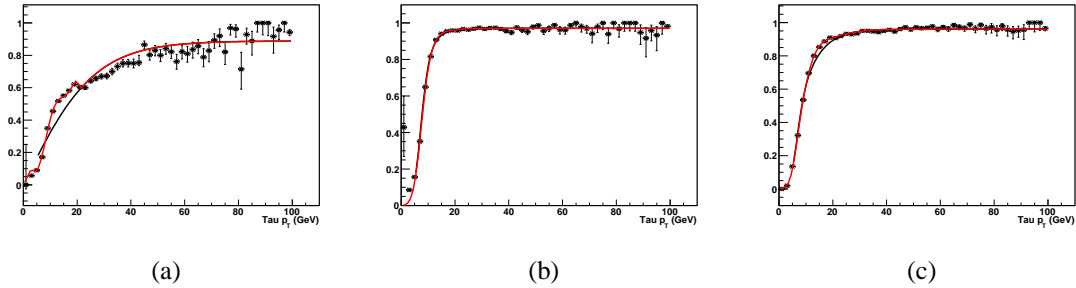
Figure 4.3: L1 turn-on curves for 8 GeV L1 jet objects arising from (a) type 1, (b) type 2, (c) type 3 taus. The red lines are the ones actually used as the turn-on curves, the black lines are a first approximation.

4.3 Monte Carlo Samples

Monte Carlo event generation is a key technique in particle physics to help experimentalists to understand their experimental data. The procedure for generating MC simulated events is as follows:

(a)   (b)   (c)
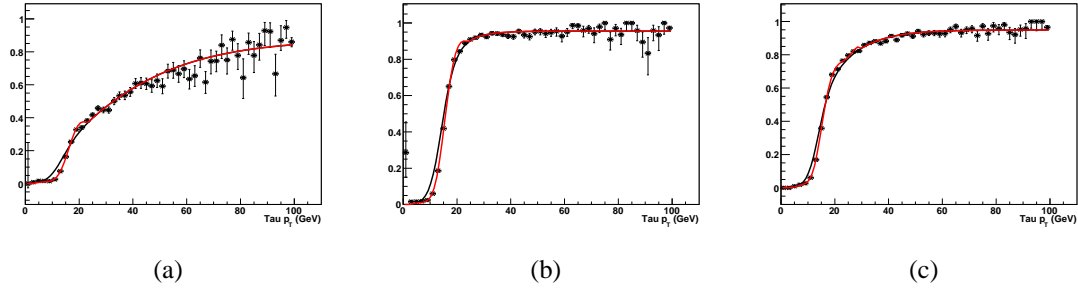
Figure 4.4: L1 turn-on curves for 15 GeV L1 jet objects arising from (a) type 1, (b) type 2, (c) type 3 taus. The red lines are the ones actually used as the turn-on curves, the black lines are a first approximation.

(a)   (b)   (c)
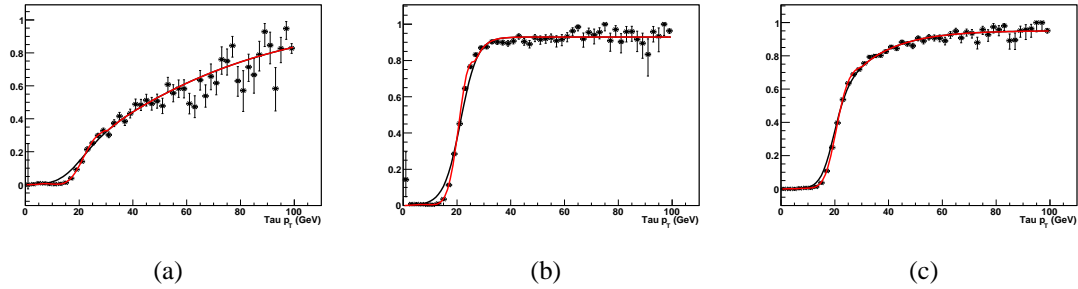
Figure 4.5: L1 turn-on curves for 20 GeV L1 jet objects arising from (a) type 1, (b) type 2, (c) type 3 taus. The red lines are the ones actually used as the turn-on curves, the black lines are a first approximation.

(a)   (b)   (c)
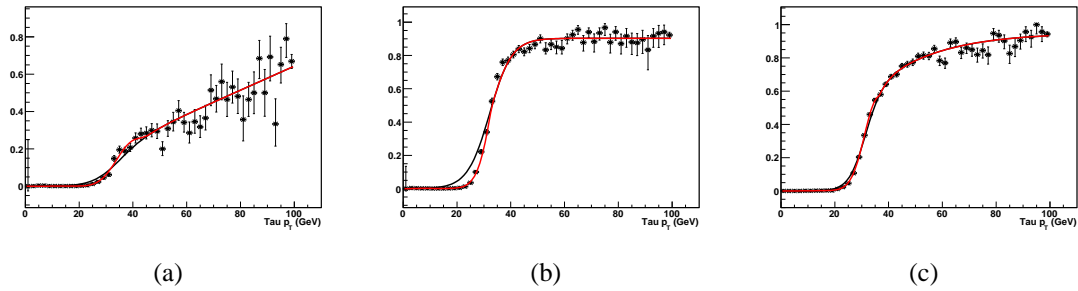
Figure 4.6: L1 turn-on curves for 30 GeV L1 jet objects arising from (a) type 1, (b) type 2, (c) type 3 taus. The red lines are the ones actually used as the turn-on curves, the black lines are a first approximation.

1. employ a MC simulation to generate events at the four-vector level;

2. simulate the D0 detector response to final state particles;

3. overlay minimum bias data events (i.e. superimpose noise and pileup events);

4. reconstruct the MC events with the same software used in data reconstruction.

QCD-multijet events are not well-modeled by MC and hence must be derived from data. Other physics sources needed in this analysis are: single top [50], W+jets, Z+jets [51], $t\bar{t}$ [50] and dibosons (WW, WZ, and ZZ), and are simulated by MC. Generation of the MC samples is done by the D0 MC production group and defined by the D0 common sample group. Here are their brief descriptions:

Signal events: The single top MC events [50] are generated with the SINGLETOP in COM-PHEP [52] MC event generator. SINGLETOP produces events whose kinematic distributions match those from NLO calculations. The top quark mass is 170 GeV, the PDF set is CTEQ6M [53], and the scales are m_t^2 for the s-channel and $(m_t/2)^2$ for the t-channel. The top quarks and the W bosons from the top quark decays are decayed in CompHEP-SINGLETOP to ensure the spins are properly transferred. PYTHIA [54] version 6.409 was used to add the underlying event and initial- and final-state radiation. TAUOLA [55] (version 2.5) was used to decay τ leptons, and EVTGEN [56] to decay b hadrons.

W+jets, Z+jets, and $t\bar{t}$: These were generated using ALPGEN [57] version 2.11. This version includes a jet-matching algorithm following the MLM prescription [58]. The matching algorithm ensures that each jet is generated by ALPGEN at the parton level and not filled in by PYTHIA, thus avoiding regions of ΔR and transverse momentum space for the radiated jets that used to be double-counted. For the W+jets and Z+jets samples the events with heavy flavor jets added by PYTHIA are also removed so as not to duplicate the phase space of those generated already by ALPGEN [59]. For the $t\bar{t}$ samples [50], the top quark mass is 170 GeV, the scale used was $m_t^2 + \sum p_T^2(\text{jets})$, and the PDF set used was CTEQ6L1. For the W+jets events, the PDF was the same as for the $t\bar{t}$ events. The scale was $m_W^2 + \sum m_T^2$, where m_T is the transverse mass defined as $m_T^2 = m^2 + p_T^2$ and the sum $\sum m_T^2$ extends to all final state partons (including

the heavy quarks, excluding the W decay products) [57]. The $W+$ light-parton (lp) jets samples have to have parton-level cuts on the light partons to avoid divergences in the cross section. These cuts were $p_T(\text{lp}) > 8$ GeV and $\Delta R(\text{lp, lp}) > 0.4$ for all massless partons (including the charm partons in these samples). For the W +heavy-flavor samples, there are no p_T or ΔR cuts on the b or c partons, but additional light partons have the $p_T(\text{lp}) > 8$ GeV and $\Delta R(\text{lp, lp}) > 0.4$ applied.

Dibosons WW, WZ, and ZZ: These were generated using PYTHIA, with inclusive decays. They form only a small fraction of the total background.

The single top and Z+jets samples have the decays into electrons, muons, and τs as separate samples, whereas the W+jets and $t\bar{t}$ samples have them generated together in combined samples with approximately one third of each present (according to the branching fractions). All the MC event samples are processed, firstly, through the GEANT [60] simulation of the D0 detector, "DØgstar". "DØSim" then does electronics simulation and pileup of any additional minimum bias interactions. At last, the simulated events are reconstructed in the same way as data.

Table 4.2 shows the cross sections, branching fractions, and initial numbers of events of the Monte Carlo samples. The cross sections for single top [61] and $t\bar{t}$ pairs [62] are for 170 GeV top mass. The W+jets and Z+jets cross sections are from ALPGEN.

Table 4.2: Monte Carlo event samples with cross sections, branching fractions, and initial numbers of events. The symbol ℓ stands for electron plus muon plus tau decays.

Event Type	The Monte Carlo Event Sets			
	Cross Section [pb]	Branching Fraction	No. of p17 Events	No. of p20 Events
Signals				
$tb \rightarrow \ell$+jets	1.12 ± 0.06	0.3240 ± 0.0032	0.6M	0.8M
$tqb \rightarrow \ell$+jets	2.34 ± 0.14	0.3240 ± 0.0032	0.5M	0.8M
Signal total	$\mathbf{3.46 \pm 0.21}$	$\mathbf{0.3240 \pm 0.0032}$	**1.1M**	**1.6M**
Backgrounds				
$t\bar{t} \rightarrow \ell$+jets	7.91 ± 0.71	0.4380 ± 0.0044	2.7M	1.8M
$t\bar{t} \rightarrow \ell\ell$	7.91 ± 0.71	0.1050 ± 0.0010	1.4M	2.7M
$t\bar{t} \rightarrow$alljets	7.91 ± 0.71	0.4570 ± 0.0044	1.3M	1.8M
Top pairs total	$\mathbf{7.91 \pm 0.71}$	$\mathbf{1.0000 \pm 0.0000}$	**5.4M**	**6.4M**
$Wb\bar{b} \rightarrow \ell\nu bb$	93.8	0.3240 ± 0.0032	2.7M	3.0M
$Wc\bar{c} \rightarrow \ell\nu cc$	266	0.3240 ± 0.0032	2.7M	2.7M
$Wjj \rightarrow \ell\nu jj$	$24,844$	0.3240 ± 0.0032	35.2M	66.5M
W+jets total	$\mathbf{25,205}$	$\mathbf{0.3240 \pm 0.0032}$	**40.7M**	**72.2M**
$Zb\bar{b} \rightarrow \ell\ell bb$	43.0	0.10098 ± 0.00006	5.5M	5.0M
$Zc\bar{c} \rightarrow \ell\ell cc$	114	0.10098 ± 0.00006	4.9M	5.7M
$Zjj \rightarrow \ell\ell jj$	$7,466$	0.10098 ± 0.00006	34.4M	30.4M
Z+jets total	$\mathbf{7,624}$	$\mathbf{0.10098 \pm 0.00002}$	**44.9M**	**41.2M**
Diboson	**17.1**	$\mathbf{1.0 \pm 0.0}$	**3.7M**	**1.9M**

Chapter 5

Object Identification and Event Selections

In accelerator-based particle physics, reconstructed objects in each event are a key component by which physics quantities are measured or new physics is discovered. One of the final-state particles in this analysis is a tau lepton, so it plays an important and special role. Tau reconstruction and tau identification optimization will be presented in Chapter 7. In this chapter, general object identifications and event selections are introduced. Selection efficiencies and event yield estimates for different sources are also presented.

5.1 Object Identification

Reconstruction and identification of some general objects are presented below. Electrons and muons are not key objects in this analysis. However, in order to avoid sharing events with other single top searches, the events containing isolated electrons and muons are vetoed. The result in the tau+jets channel will be combined with that in the two other leptonic channels.

5.1.1 Electrons

Electron objects are reconstructed by clustering shower energy depositions in the D0 electromagnetic calorimeter. Only electrons within the central calorimeter with $|\eta^{\text{det}}| < 1.1$

(CC) are considered. The tight isolated electron is defined as follows:

- Electron $p_T > 15$ GeV

- f_{em} is defined as the ratio of the cluster energy deposited in the EM section of the calorimeter over E_{total}; it must be greater than 90%

- An electron should be isolated. Isolation is defined as ratio of energy in a halo around the EM cluster over the total energy, i.e., $\frac{E_{total}(R<0.4)-E_{EM}(R<0.2)}{E_{EM}(R<0.2)}$, where R is the cone radius, $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$; it has to be less than 0.15.

- χ^2 of the 7×7 H-matrix < 50. [1]

- The energy deposition must be matched with $\chi^2 > 0$ to a charged particle track with $p_T^{trk} > 5$ GeV and z(track, primary vertex)< 1 cm

- Based on seven electron property variables, a likelihood discriminant is created to separate real electrons from W/Z boson decays from jets with a large EM fraction. The EM-likelihood $\mathcal{L}$ is required to be larger than 0.85.

5.1.2 Muons

Muon reconstruction and identification are based on information from the 3-layer muon detector system and central tracks. Good muons are defined by the following criteria:

- Muon $p_T > 15$ GeV

- Muon $|\eta| < 2.0$

- Muon quality: medium[2], with hits on 3 muon layers.

- Distance in z between the muon track and the primary vertex: z(track,PV)<1 cm.

---

[1]H-matrix, a matrix that is used for shower shape analysis, is a measure of similarity between the candidate and a real electron shower. The χ^2 is computed as $\chi^2 = \Sigma_{i,j=1}^{7}(x'_i - \bar{x}_i)H_{ij}(x'_j - \bar{x}_j)$ where $x_{i,j}$ refers to electron shower shape variables [63].

[2]The muon quality means that the muon reconstructed in the muon system (called a local muon) is matched to a track in the central tracking system

– Muon must be isolated from any jets by the cut $\Delta R(\mu, \text{jet}) > 0.5$

– Calorimeter isolation $|\frac{\Sigma^{\text{cells}} E_T}{p_T(\mu)}| < 0.12$ in $0.1 < \Delta R(\text{cells}, \text{muon cal-track}) < 0.4$, and track isolation $|\frac{\Sigma^{\text{tracks}} p_T}{p_T(\mu)}| < 0.12$ in $\Delta R(\text{track}, \text{muon track}) < 0.5$ cone [64].

5.1.3 Jets

Jet objects are reconstructed by the Improved Legacy Cone Algorithm (ILCA) [65] and the jets with a cone size $R = 0.5$ are used in this analysis (called JCCB). Before applying identification cuts, the Jet Energy Scale (JES) is used to correct raw jet energy to the energy in a particle-jet level. For Monte Carlo samples, Jet Shifting, Smearing and Removing (JSSR) is also applied to make simulated jets realistic (data-like). Good jet objects are identified by the following cuts:

– Highest p_T jet (Jet1) $p_T \geqslant 25$ GeV, the second highest p_T jet (Jet2) $p_T \geqslant 20$ GeV, other jets $p_T \geqslant 15$ GeV

– Highest p_T jet $|\eta| < 2.5$, other jets $|\eta| < 3.4$

– Fraction of energy deposited in the EM calorimeter over the total energy must be $0.05 < f_{\text{em}} < 0.95$

– Fraction of energy deposited in the coarse hadronic section of the calorimeter has to be less than 0.4

– Removal of jets which overlap in space with EM and tau objects:

$$\Delta R(\text{jet}, \text{EM}/\text{tau}) > 0.5.$$

5.1.4 *b*-jets

There is at least one b quark in the final state of single top quark decays, so one or two b jets should be present in each event. A neural network (NN) b-jet tagger developed by the b-jet identification (B-ID) group is used to identify jets from b-, c-quarks [66]. Taggable jets[3] in

---

[3]A taggable jet is defined as a b jet candidate which matches a track-jet cluster within a cone radius 0.5. The track-jet cluster must contain at least two tracks with at least one SMT hit and $p_T > 1.0$ (0.5) GeV for the first(second) track [67].
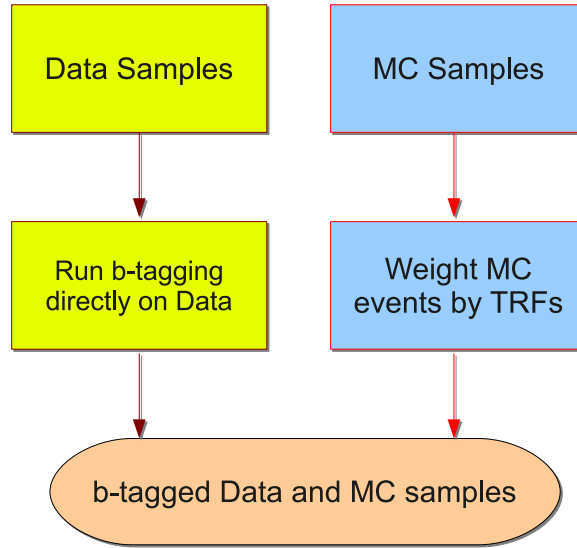
Figure 5.1: The two approaches to apply *b*-tagging on data and MC samples.

data samples are directly tagged. However, the tracking in the D0 MC simulation is overly optimistic in both the quality and number of tracks found, which results in an overestimated *b*-tagging efficiency in MC by 10–20% compared to data [66]. Thus, taggable jets in MC samples are given weights using *b*-, *c*-quark and fake-jet tag rate functions (TRF). The two procedures are shown in Fig. 5.1.

In the NN algorithm, a NN is trained on the outputs of three *b*-jet identification algorithms: Secondary VerTex (SVT), Jet Lifetime Impact Parameter probability (JLIP), and Counting Signed Impact Parameter (CSIP) [66, 68]. A typical *b* jet is illustrated in Fig. 5.2. Its characteristics include: a displaced secondary vertex (on which the SVT algorithm is based), and displaced tracks with large impact parameters (on which the JLIP and CSIP algorithms are based).

If the NN output for a jet in data samples is larger than 0.775, the jet is tagged as a *b* jet. This operating point corresponds to a *b*-tagging efficiency of ≈40% and a light-quark fake rate of ≈0.4%.

As mentioned above, for *b*-tagging on MC samples, *b*-, *c*-jet and fake-jet TRFs are measured in data to weight MC *b*-, *c*-quark and light-quark (gluon) events respectively. A TRF value for each taggable jet represents the probability (p_i) of that jet being tagged as a *b*-jet. The basic idea of weighting is to list all taggable jet permutations per event and, accord-
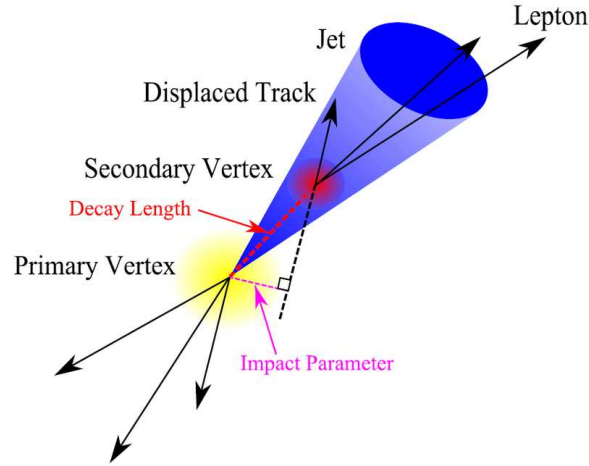
Figure 5.2: Illustration of a typical *b* jet. A displaced secondary vertex and displaced tracks with large impact parameters are used to identify *b* jets (after [66]).

ing to each jet's TRF value, to calculate the event probability for each permutation. This probability can be applied to the event as a weight. The event probability for permutation is formulated by

$$P_k = \prod_{\text{tagged}} p_i \prod_{\text{non-tagged}} (1 - p_j) \tag{5.1}$$

.

5.1.5   Missing Transverse Energy

Due to conservation of momentum in collisions, a neutrino's momentum can be inferred from measuring missing transverse energy. In practice the raw $\displaystyle{\not}E_T$ is obtained by adding up vectorially the transverse energies in all cells of the EM and fine hadronic calorimeters. For the coarse hadronic calorimeter, only cells belonging to a "good" jet[4] are added. Starting from this raw reconstructed quantity, energy correction propagation from tau, muon, electron and jet energy corrections are also applied.

5.1.6   Primary Vertex

A primary vertex is defined as the location where the hard scattering interaction takes place. It provides important information in discriminating reconstructed physics objects

---

[4]A "good" jet in D0 has had quality cuts applied.

coming from the $p\bar{p}$ collision. At D0 it is reconstructed by means of an Adaptive Vertex Fitting algorithm [69]. The reconstruction process of a primary vertex consists of three steps: (i) select tracks by requiring track $p_T >0.5$ GeV and two or more hits registered in SMT, and assign the tracks to a vertex to which the extrapolated paths of the tracks point. (ii) Given a fitting result χ_i^2 for a track i to a vertex, a weight is assigned to the track as $\frac{1}{1+\exp\left((\chi_i^2-\chi_{\text{cutoff}}^2)/8\right)}$ where χ_{cutoff}^2 is set to 16. The procedure is repeated until the difference of weights between two iterations is less than 10^{-4}. Thus a list of potential vertices from the hard scattering is obtained. (iii) select the vertex by a minimum-bias probability selection algorithm [70] to discriminate the vertex of the hard scattering from those of minimum bias events. In the selection chain, z of primary vertex $|z_{PV}| < 60$ cm is required.

5.2 Event Selection

Based on the signature: one isolated tau lepton, $\not{E}_T$, and at least 1 b-jet, the following criteria are applied to create the pre-selected sample:

- Good data quality to make sure all subdetectors are working properly and removal of duplicated events. This procedure is basically to remove bad runs, blocks, or events defined by the D0 data quality group

- Trigger requirements that have been presented in Section 4.2

- Good primary vertex. $|z_{\text{PV}}| < 60$ cm with at least three tracks associated.

- Missing transverse energy: $20 < \not{E}_T < 200$ GeV

- Tau and jet identifications.

- Veto electrons and muons in order to combine the result in this channel with other channels.

- One tau lepton per event

- Two or three good jets after removal of any jet which is matched to the good tau in the event.

In addition, cuts are applied on two topological variables, $H_{\text{AllJets}-\text{BTaggedJet}}$ (Scalar sum of jets' energy for all jets except the leading b-tagged jet) and $\Delta\phi(\text{Jet2},\not{E}_T)$ in order to suppress QCD-multijets. These cuts are applied only on the Run IIa dataset[5]. Table 5.1 shows cut values used.

Table 5.1: Summary of the cuts $H_{\text{AllJets}-\text{BTaggedJet}}$ (Cut1) and $\Delta\phi(\text{Jet2},\not{E}_T)$ (Cut2) to suppress QCD-multijet events.

	1-prong tau				3-prong tau			
	1 tag		2 tags		1 tag		2 tags	
	>Cut1	>Cut2	>Cut1	>Cut2	>Cut1	>Cut2	>Cut1	>Cut2
2 jets	50 GeV	0.3	0 GeV	0.0	50 GeV	0.6	0 GeV	0.0
3 jets	100 GeV	0.3	60 GeV	0.0	100 GeV	0.6	80 GeV	0.0

5.3 Efficiency of Selections and Yield

Tables 5.2 and 5.3 show selection efficiencies on Run IIa and Run IIb data samples, respectively. Object and event selection criteria, numbers of events and cumulative and relative passing rates are also listed in the tables. Tables 5.4 and 5.5 show similar selection information for signal MC samples.

The final dataset should consist of several background components: QCD-multijet events, $t\bar{t}$, W+jets, Z+jets and diboson samples. MC samples generated at D0 are used to simulate $t\bar{t}$, W+jets and Z+jets and diboson sources (see more details about the samples in Section 4.3), however, QCD-multijet events are not well-modeled by MC and hence must be modeled from data. So an approach is developed to derive a QCD-multijet sample from a non-b-tagged dataset, which is discussed in Chapter 8. Table 5.6 shows a summary of yields after all selections and b-jet tagging [6]. Tables E.1 – E.4 in Appendix E show similar yields but in different analysis channels. Yield values in the tables have been rounded for

---

[5]As seen before, Run IIb L1 has tighter requirements than Run IIa to kill more QCD-multijet events. Thus we apply the cuts only to Run IIa.

[6]"selections" here also includes a final boosted decision tree cut to exclude the events used for determining QCD-multijet normalization factors. See Section 8.6 for the QCD-multijet normalization discussion.

Table 5.2: Efficiencies of the different selections on skimmed Run IIa dataset. The 1st column shows the selection criteria used, numbers in the 2nd column are numbers of events passing the selection in the 1st column, the 3rd column is the cumulative passing event rate and, the 4th column shows relative event passing rates.

Selection Criteria	Passed Events	Cumulative Rate	Relative Rate
Initial number	58059422	100.00%	100.00%
Removal of duplicated events	58058780	100.00%	100.00%
Data quality selection	42965720	74.00%	74.00%
Tau jets triggers	20090025	34.60%	46.76%
Tight tau kinematic cuts and ID	590157	1.02%	2.94%
Remove matched jet with tau from jets	590157	1.02%	100.00%
Jet selection	590157	1.02%	100.00%
Single top jet selection	413386	0.71%	70.05%
Vertex selection	400934	0.69%	96.99%
Veto of good electrons	383037	0.66%	95.54%
Veto of good muons	381903	0.66%	99.70%
$\not{E}_T$ cut	252141	0.43%	66.02%

clarity although all calculations are done with full-precision values. Pie charts in Fig. 5.3 illustrate fractions of the different components of the 1-*b*-tagged datasets of tau type 1+2.
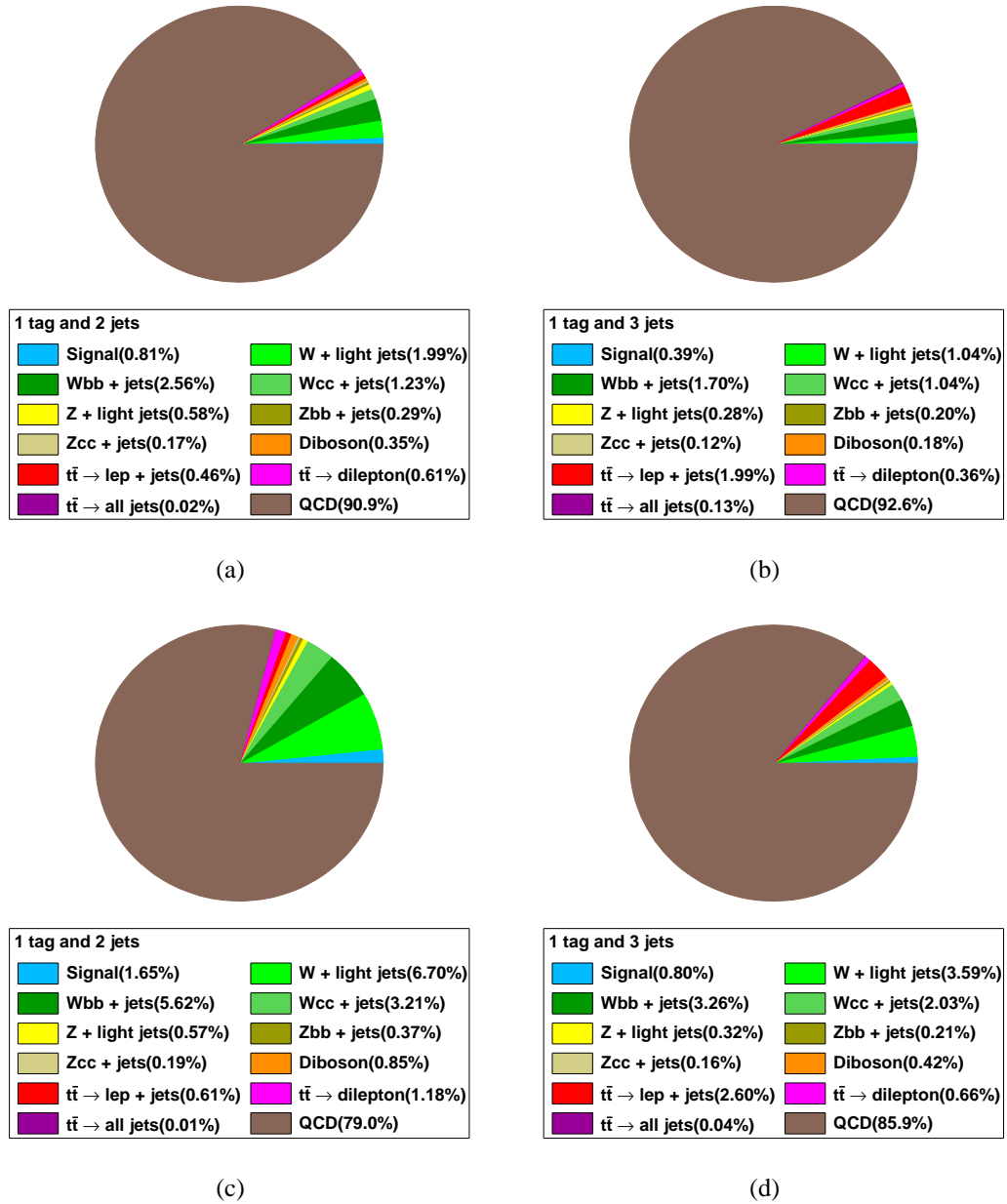
Figure 5.3: Pie charts of 1-*b*-jet data source components in tau type 1+2 for (a) Run IIa, 2 jets; (b) Run IIa, 3 jets; (c) Run IIb, 2 jets; (d) Run IIb, 3 jets.

Table 5.3: Efficiencies of the different selections on skimmed Run IIb dataset. The 1st column shows the selection criteria used, numbers in the 2nd column are numbers of events passing the selection in the 1st column, the 3rd column is the cumulative passing event rate and, the 4th column shows relative event passing rates.

Selection Criteria	Passed Events	Cumulative Rate	Relative Rate
Initial number	107275891	100.00%	100.00%
Removal of duplicated events	107275891	100.00%	100.00%
Data quality selection	86338171	80.48%	80.48%
Tau jets triggers (OR)	39214344	36.55%	45.42%
Tight tau kinematic cuts and ID	1069833	1.00%	2.73%
Remove matched jet with tau from jets	1069833	1.00%	100.00%
Jet selection	1069833	1.00%	100.00%
Single top jet selection	638299	0.60%	59.66%
Cut on number of good jets	577009	0.54%	90.40%
Vertex selection	559251	0.52%	96.92%
Veto of good electrons	465270	0.43%	83.20%
Veto of good muons	462584	0.43%	99.42%
$\not{E}_T$ selection	347121	0.32%	75.04%

Selection Criteria	$tb \to e\nu bb$		$tqb \to e\nu bqb$		$tb \to \tau\nu bb$		$tqb \to \tau\nu bqb$	
	N_{evt}	Rate	N_{evt}	Rate	N_{evt}	Rate	N_{evt}	Rate
Initial number	200000	100.0%	200000	100.0%	200000	100.0%	175000	100.0%
Removal of events with zero luminosity	199302	99.7%	199051	99.5%	199328	99.7%	174393	99.7%
Removal of duplicated events	157099	78.5%	163069	81.5%	161900	80.9%	167050	95.5%
Data quality	147614	73.8%	151568	75.8%	150182	75.1%	155628	88.9%
Smearing of objects	147614	73.8%	151568	75.8%	150182	75.1%	155628	88.9%
Tau kinematic cuts and ID	92152	46.1%	93927	47.0%	43329	21.7%	44595	25.5%
Remove matched jet with tau from jets	79642	39.8%	77645	38.8%	37529	18.8%	36746	21.0%
Jet selection	79642	39.8%	77645	38.8%	37529	18.8%	36746	21.0%
Primary vertex selection	79481	39.7%	77450	38.7%	37364	18.7%	36570	20.9%
Veto of good electrons	25492	12.7%	25689	12.8%	33379	16.7%	32805	18.7%
Veto of good muons	25489	12.7%	25689	12.8%	33233	16.6%	32650	18.7%
$\not{E}_T$ cut	23194	11.6%	23299	11.6%	29480	14.7%	28862	16.5%
Tau jets trigger simulation	11188	5.6%	10467	5.2%	14376	7.2%	12859	7.3%

Table 5.4: Numbers of events and cumulative rate of different selections on the Run IIa single top MC samples.

Selection Criteria	$tb \to e\nu bb$		$tqb \to e\nu bqb$		$tb \to \tau\nu bb$		$tqb \to \tau\nu bqb$	
	N_{evt}	Rate	N_{evt}	Rate	N_{evt}	Rate	N_{evt}	Rate
Initial number	271797	100.0%	274546	100.0%	275085	100.0%	273896	100.0%
Removal of events with zero luminosity	270307	99.5%	273331	99.6%	273459	99.4%	272775	99.6%
Removal of duplicated events	218794	80.5%	266272	97.0%	201899	73.4%	264518	96.6%
Data quality	210541	77.5%	256431	93.4%	194284	70.6%	254528	92.9%
Smearing of objects	210541	77.5%	256431	93.4%	194284	70.6%	254528	92.9%
Tau kinematic cuts and ID	125822	46.3%	151594	55.2%	51196	18.6%	67344	24.6%
Remove matched jet with tau from jets	109381	40.2%	126710	46.2%	44618	16.2%	56302	20.6%
Jet selection	109381	40.2%	126710	46.2%	44618	16.2%	56302	20.6%
Primary vertex selection	109029	40.1%	126262	46.0%	44406	16.1%	55990	20.4%
Veto of good electrons	36290	13.4%	42938	15.6%	39684	14.4%	50442	18.4%
Veto of good muons	36283	13.3%	42933	15.6%	39478	14.4%	50151	18.3%
$\not{E}_T$ cut	32963	12.1%	38883	14.2%	34918	12.7%	44129	16.1%
Tau jets trigger simulation	16778	6.2%	19794	7.2%	13876	5.0%	16942	6.2%

Table 5.5: Numbers of events and cumulative rate of different selections on the Run IIb single top MC samples.

Table 5.6: Yields with uncertainty after all selections in all analysis channels combined. The fractions of different background components are also shown in percent.

Source	Events	
Signals		
tb+tqb signal	72±12	
Backgrounds		
W+jets	679±104	(19%)
Z+jets	60±10	(2%)
Dibosons	37±6	(1%)
$t\bar{t}$	231±44	(6%)
QCD-multijets	2627±98	(72%)
Background Sum	3633±153	
Data	3845	

Chapter 6

Boosted Decision Trees

In data mining, classification and prediction models are powerful tools to dig out hidden useful information to make intelligent decisions [71]. Many such techniques, such as neural networks, actually play important roles in particle physics by discriminating signal events from large backgrounds. In this analysis, a technique called Boosted Decision Trees (BDT) is employed to develop a new tau ID optimized for single top quark decays in Chapter 7. And later this technique is also used to perform our ultimate single top signal-background separation to be shown in Chapter 9.

In this chapter, particle physics jargon (instead of a computer science language) is used to explain the BDT concept and working algorithm used in this analysis. Section 6.1 presents what decision trees are. Section 6.2 introduces a technique called boosting, which enhances the performance of decision trees. Section 6.3 talks about one transformation of BDT output.

6.1 Decision Trees

A decision tree is a popular predictive classification technique to explore hidden knowledge in data by making a flow-chart decision tree using sample learning. Its output is used to label classifications or as a descriptive means for calculating conditional probabilities [71].

Ususally, particle physicists are concerned with only two rough classes of samples: signal and background. A binary decision tree can be introduced and used to do this classification. Since the classification can be visualized by a simple binary tree structure, in this

respect, a decision tree is similar to a rectangular-cuts analysis [72].
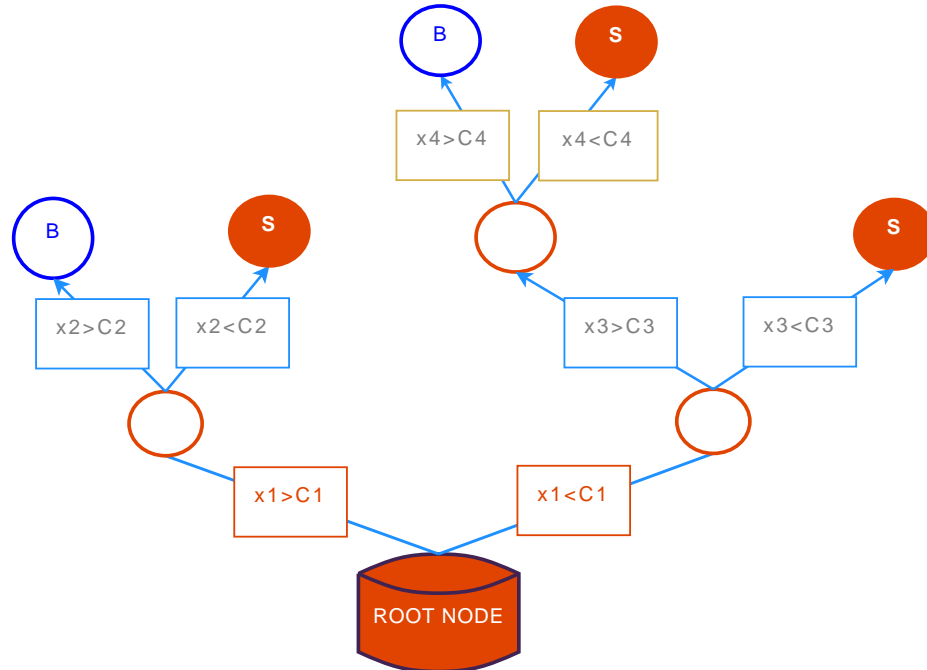
Figure 6.1: Illustration of growing of a decision tree.

A decision tree growing process (also called building, learning or training) is graphically demonstrated in Figure 6.1 which is described blow:

Input

- Training and testing samples. The training sample is used for machine learning while the testing sample is for evaluating DT performance

- Topological and kinematic variables for each candidate event

- Splitting criteria which determine the best partitions of the data into individual classes: signal and background.

Output A decision tree

Growing Procedure

1. Normalize the signal training sample to the background training sample

2. Start with the *root node*[1] including all training events

3. The variable x_1 is selected by splitting criteria to divide the whole sample to two classes by the cut value: C_1. Use majority voting to determine the class of split samples

4. Continue to split the resulting nodes from the last step by selected variables x_2, x_3 and x_4

5. When some stop criterion is satisfied, stop splitting and return either a binary bit ± 1 (signal and background) or the signal purity of the leaf contents

Thus a resulting decision tree is built with many nodes at different depths. Each splitting node should have a splitting test and a voting result stored so that it makes a decision when a specific event passes it in later BDT output calculations.

The separation algorithm used in splitting nodes in growing a tree plays a very important role in performance of the resulting decision trees. Software called the Toolkit for Multivariate Data Analysis in ROOT (TMVA, `version 3.9.6`) [73] is employed in this analysis.

A Gini index approach implemented in TMVA is the separation criterion that measures the impurity of a class-labeled training sample D in this analysis. It is defined as

$$\text{Gini}(D) = p \cdot (1 - p) \tag{6.1}$$

for a binary decision tree where p is the probability that a node belongs to class signal or background. Then for each variable, each of the possible binary splits is considered and the subset that gives the minimum Gini index is selected. This minimum Gini index approach maximizes the difference in impurity between the mother node and the two daughters.

6.2 Boosted Decision Trees

Decision trees have broad applications in social science. Although the performance of decision trees is outstanding, a shortcoming is their instability due to statistical fluctuations when the tree structure is derived from the training sample [72]. A small change in the

---

[1]"node" means a group of events.

training sets may yield large variation in the classification. A commonly used technique called boosting can overcome this shortcoming by building a "forest" of decision trees and making a decision on a majority vote based on each tree in the forest. Boosted decision trees were recently used by the MiniBooNE experiment [74, 75] and the single top quark production measurement of the D0 experiment [12, 13] and the CDF experiment [14].

Boosting is a general technique which is not limited to decision trees only but can be applied to any weak classifier. The most popular boosting model is called the adaptive boosting (*AdaBoost*) in which misclassified events during the training of a tree are given higher event weights in the next cycle of tree training [71]. The initial decision tree is trained starting with the original event weights. Misclassified events, are then given higher weights by multiplying by a common boost weight α defined as

$$\alpha = \left(\frac{1 - \text{error}}{\text{error}} \right)^{\beta} \tag{6.2}$$

where "error" is the misclassification error rate of the previous tree and β is a parameter needing optimization (usually it is set as 0.5). In the mean time, the entire event sample is normalized back to the sum of weights in the original tree set.

If $h_i(x)$ is the output of the ith decision tree given input variables x, then the output of all boosted decision trees is calculated by

$$y_{\text{BDT}}(x) = \sum_{i \in \text{Forest}} \ln(\alpha_i) \cdot h_i(x) \tag{6.3}$$

where the sum is conducted over all trees of the forest. The optimal number of trees in a forest is analysis-dependent. A forest which is too large wastes computing resources and may also suffer from worsening performance as it becomes too specialized on the training sample (overtraining).

6.3 BDT Output and Transformation

Once a forest of trees is built, classification on an independent sample of interest should be done tree by tree and Eq. 6.3 is used to calculate a final output of the forest. This step is the calculation of BDT outputs. BDT have a flow-chart structure. When calculating an output of each tree, each event accepts decisions made by the tests stored at each node, hence

traces down the tree structure until it reaches the last node. The signal-purity value at the stop node is the BDT output of the event. The BDT output value is utilized to quantify the classification of the event. The BDT output covers the range of [0,1]. But due to boosting, the output is generally pushed to the middle of the range. Signal-like events should have BDT scores close to 1 while background should be close to 0. This signal purity distribution is used in the BDT tau ID study in Chapter 7.

Instead of using the original BDT output value (signal purity) above, sometimes it is necessary to transform the output to avoid the problems caused by sparse population in extreme signal or/and background regions and limited statistics. Thus, after the transformation, the BDT output values will spread over the whole range [0, 1]. In the final single top discrimination, a transformed output called BDT probability output is used. It is obtained by applying a monotonic transformation function to original BDT outputs[2]. The basic idea is illustrated in Fig. 6.2. A transformation function is empirically defined as

$$\text{Transformation Function} = \frac{\text{BDT}_{\text{signal}}}{\text{BDT}_{\text{signal}} + \text{BDT}_{\text{background}}} \tag{6.4}$$

where $\text{BDT}_{\text{signal}}$ and $\text{BDT}_{\text{background}}$ are original BDT purity distributions normalized with each other that are shown in (a) of Fig. 6.2. Thus a transformation histogram is obtained as shown in (b) of the same figure. However, if this histogram is used to transform original BDT distributions, then the BDT probability distributions obtained are very spiky due to granularity as shown in (c). So a procedure is necessary in which the transformation histogram is fit by means of an error function. The fitted function is used to replace the histogram. Thus new transformed smooth distributions are obtained as shown in (d) of the figure.

---

[2]It should be emphasized that since the transformation function is required to be monotonic, it can neither improve nor degrade the final sensitivity of BDT, and it just serves as a tool re-distributing the BDT outputs, hence making calculation easier.
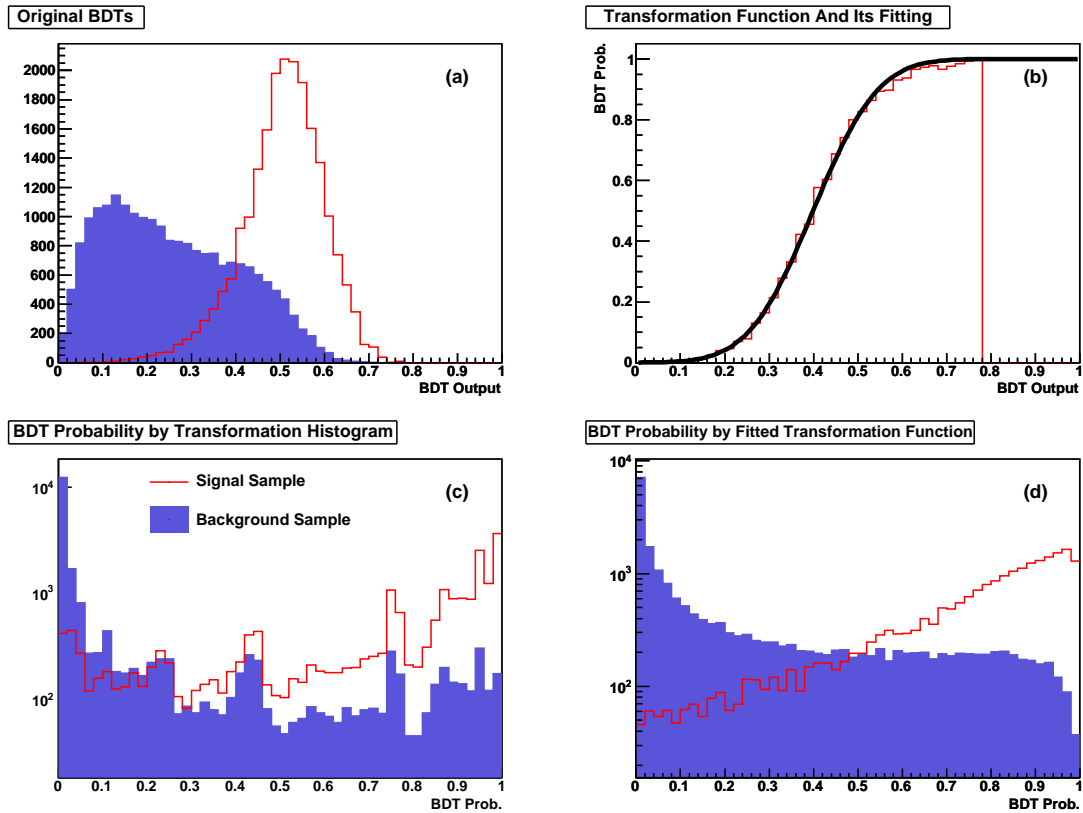
Figure 6.2: Illustration of BDT purity output transformation to BDT probability output. (a) Original BDT purity output distributions normalized; (b) Transformation histogram and fitting function; (c) BDT probability output distributions obtained by transformation histogram and (d) BDT probability output distributions obtained by transformation function.

Chapter 7

Optimization of Tau Identification

In this analysis, the tau lepton is one of the important objects appearing in the final state. The tau lepton is the heaviest lepton, with mass 1.78 GeV/c^2. Its lifetime ($c\tau$) is about 87 μm [5], which indicates a tau decays before reaching any active elements of the D0 detector. At D0, hadronic tau candidates are reconstructed by means of detector signatures of decay daughters, such as tracks, EM clusters, and hadronic clusters. Tau objects at detector level appear as narrow-cone jets and hence tau samples are heavily contaminated by gluon or quark jets. For this reason, the analysis needs a high-performance tau identification (ID). At D0, there is a standard tau ID available provided by the tau ID group which uses several tau kinematic cuts and a multivariate variable cut. However, the multivariate variable is trained on $Z \to \tau\tau$ signal and background samples so that it may not be optimized for taus originally from single top quarks. In this chapter, a study on a high-performance tau ID optimization for single top quarks by Boosted Decision Trees (BDT) is introduced.

In the chapter, Section 7.2 discusses hadronic tau reconstruction. Section 7.3 briefly presents the D0 standard Neural Network (NN) tau ID. In Section 7.4, optimization of a tau ID is introduced including motivation, methodology and related results. Section 7.5 introduces the estimation of systematic uncertainty of the optimized BDT tau ID. At last, a summary is given in Section 7.6.

7.1 Hadronic Taus and Jets

Hadronic tau objects appear as narrow-cone jets in the detector because:

- Most hadronically decaying taus decay to final states with one or three charged pions (1-prong or 3-prong). Thus the number of daughter hadrons from a tau decay is less than a number of hadrons produced in a regular gluon or quark jet. Track multiplicity is also low compared to gluon and quark jets.

- In a tau decay, the daughter hadrons are boosted and are more collimated. Hadrons from gluons and quarks are more widely spread than those in a tau cluster.

Although tau samples are heavily contaminated by gluon or quark jets, one can use the different features above to discriminate a tau from background jets. The more hadrons in an object cone, the more difficult the discrimination, so identifying 3-prong tau jets is more difficult than 1-prong.

7.2 Hadronic Tau Reconstruction at D0

7.2.1 Tau reconstruction and types

A hadronic tau candidate is a collection of the following objects [76]:

Calorimeter cluster

Constructed by means of a simple cone algorithm from all the towers with energy above a threshold (>0.05 GeV) around a seed tower ($E_{T\,\text{seed}} >1$ GeV) within a cone radius 0.5. The cluster should have a width *rms* < 0.25 where *rms* width is the root square sum of the E_T weighted η-ϕ distance of all calorimeter towers with respect to the tau axis (i.e. $rms = \sqrt{\sum_{i=1}^{n} \left(\Delta\eta_i^2 + \Delta\phi_i^2 \right) \frac{E_{Ti}}{E_T}}$ where i is the index of calorimeter towers and $E_T = \sum_i E_{Ti}$).

Electromagnetic calorimeter sub-clusters

Found by a regular nearest neighbor clustering algorithm in the 3rd layer of the electromagnetic calorimeter where EM showers are expected to reach their profile maximum. If such subclusters are found, then EM cells in other layers and preshower hits are attached to them with their E_T as weights. The objects are used to reconstruct taus with neutral decay daughters like π^0 or γ radiation.

Associated tracks

Any tau candidate must have at least one associated track found by the following track-matching procedures:

1. Sort in decreasing p_T all the track candidates that are in a cone of 0.5 about the centroid of the calorimeter cluster

2. Associate the highest p_T track with $p_T > 1.5$ GeV to the calorimeter cluster

3. Up to 2 more tracks will be considered if their z is within 2 cm of that of the first track

4. The second track will be added if its invariant mass with the first track is less than 1.1 GeV

5. The third track will be added if its invariant mass with the first 2 tracks is less than 1.7 GeV and the sum of the three track charges is $+1$ or -1

Motivated by hadronic tau decay modes, it is convenient to classify hadronic reconstructed tau candidates in three types:

Type 1 One calorimeter cluster and one associated track

Type 2 One calorimeter cluster, one associated track and at least one EM subcluster

Type 3 More than one associated track and wide calorimeter cluster with or without an EM subcluster

Tau type 1 and 2 correspond to 1-prong tau decay modes $\tau^{\pm} \to \pi^{\pm}\nu$ and $\tau^{\pm} \to \rho^{\pm}\nu$ while tau type 3 is for 3-prong decay modes such as $\tau^{\pm} \to \pi^{\pm}\pi^{\mp}\pi^{\pm}(\pi^0)\nu$.

7.2.2 Tau property variables

There are $\sim$40 property variables calculated in a reconstructed tau object. The variables can be classified in four categories. Some examples are:

Kinematic variables

tau energy, momentum, pseudo-rapidity η, azimuthal angle ϕ *etc.*

Shower shape variables

– Tau cluster width rms $= \sqrt{\sum_{i=1}^{n} [(\Delta\phi_i)^2 + (\Delta\eta_i)^2] \cdot \frac{E_{Ti}}{E_T}}$

– tau profile $= \frac{E_{T1} + E_{T2}}{E_T^\tau}$ where E_{T1} and E_{T2} are the transverse energies of the two most energetic calorimeter towers in a tau object.

– emf and fhf: EM and hadronic fractions of tau energy deposited in the EM and hadronic calorimeters respectively.

– prf3 $= \frac{E_T^{\text{leading EM subcluster}}}{E_T^{EM3}}$, a ratio of E_T of leading EM sub-cluster over E_T deposited in the 3rd EM layer.

Isolation variables

– caliso $= \frac{E_T^\tau - E_{T\text{core}}^\tau}{E_{T\text{core}}^\tau}$ where E_T^τ and $E_{T\text{core}}^\tau$ are tau E_T within a cone $R < 0.5$ and $R < 0.3$, respectively.

– trkiso $= \frac{\sum p_T^{\text{trk}}}{\sum p_T^{\tau\text{trk}}}$, where $\sum p_T^{\text{trk}}$ is sum of p_T of non-tau-associated tracks within a cone size 0.5 and $\sum p_T^{\tau\text{trk}}$ is the sum over all tau-associated tracks' p_T.

– EM12isof $= \frac{E^{EM1} + E^{EM2}}{E^\tau}$ where E^{EM1} and E^{EM2} are energies deposited in the 1st and 2nd layers of the EM calorimeter.

Calorimeter-track correlation variables

– ET_o_sum $= \frac{E_T^\tau}{E_T^\tau + \sum p_T^{\tau\text{trk}}}$ where E_T^τ is the tau calorimeter cluster transverse energy.

– dalpha $= \sqrt{(\Delta\phi/\sin\theta)^2 + (\Delta\eta)^2}/\pi$ where the angle differences, $\Delta\phi$ and $\Delta\eta$, are between the vector sum of tau tracks and the vector sum of EM-clusters, θ is an azimuthal angle of the centroid of the vector sum of EM-clusters.

7.3 Standard D0 Tau ID

7.3.1 Introduction

No single tau variable defined above can serve alone to separate taus from the fake tau backgrounds. So, a multivariate technique called neural networks (NN) is used in the standard tau ID. The standard tau ID consists of two parts: kinematic cuts and a NN cut. The kinematic cuts provide a very loose selection before applying the NN cut. The NN study is done in the context of measurement of $\sigma(p\bar{p} \to Z) \cdot \text{Br}(Z \to \tau\tau)$ [76]. In the measurement, the data candidates of interest are from Z bosons decaying to $\tau(\to \mu\nu_\nu\nu_\tau)\tau$, isolated μ-τ pairs. So, the multivariate technique is trained on $Z \to \tau(\to \mu\nu_\nu\nu_\tau)\tau$ signal and background samples although it serves as the standard tau ID at D0.

7.3.2 NN Training and Testing

The tau identification group defines tau signal and background samples, and trains and tests NNs on these samples for three tau types [77]. Tau signal samples are defined by $Z \to \tau\tau$ MC samples with Z boson mass altered to 130–250 GeV for training and Z boson mass = 60–130 GeV for testing. The background training sample is defined as the data events having an anti-isolated μ–τ pair in which the muon is located within a jet cone by $\Delta R_{\mu,\text{jet}} < 0.5$, i.e., anti-isolated, and this anti-isolated muon and the tau candidate are back-to-back, $|\Delta\phi(\mu,\tau)| > 2.5$. Such events are likely to be QCD $b\bar{b}$ events. The background testing sample is based on isolated μ–τ pair data events in which μ and τ have the same charge sign. The standard tau ID NN training uses about 10 selected variables depending on different tau types.[1]

7.3.3 NN Output

If the NN output for a tau candidate is close to 0, it is assumed to be a gluon jet or a quark jet. A tau candidate with NN output near 1 is assumed to be a real tau. Thus, applying a NN output cut can reduce the jet background. For example, by placing a cut of 0.9 on NN

---

[1]In addition to NNs which are used to remove jet background, there is also a NN available to remove electrons [77], which is not discussed here.

output of tau candidates with transverse energy between 20 and 40 GeV, the jet background is reduced by a factor of 50 while the total efficiency is kept to 70% [78]. The NN tau ID is used as a standard tau ID at D0, however, no analysis conducted a tau ID optimization study such as the one discussed below.

7.4 BDT Tau ID Optimization

7.4.1 Motivation

Ideally, a general tau ID should work well for all tau-related studies. However, tau identification can be affected by several factors, such as the event environment. Generally, the busier the event is, the harder it is to identify the tau. The standard D0 tau ID is derived partially from a $Z \rightarrow \tau\tau$ MC sample which is a relatively clean environment. If the standard tau ID is applied to busier events, for example, single top events with extra jets compared to $Z \rightarrow \tau\tau$ events, the tau identification efficiency may be reduced due to the change of event topology. Different physics processes may also have different tau kinematics causing tau properties such as profiles, and isolations to vary. This means that the standard ID may not be optimal for every process. In addition, from the background side, the standard NN tau ID is able to optimally discriminate real taus from its specific training background but may not do so from other backgrounds. For example, it will be shown that it is not optimal for single top backgrounds.

The loss of efficiency of the standard tau ID for single top quark events is visible in the NN output distributions shown in Fig. 7.1. In the figure, tau signal samples are single top MC samples in s- and t-channels and $Z \rightarrow \tau\tau$ samples with Z mass 130-250 GeV and 60-130 GeV [2]. Fake tau samples are a tau trigger skimmed data sample to be defined later, and the anti-isolated μ–τ data sample defined above. A high NN output indicates a high probability to be a true tau while a low NN output means the object is likely a fake tau. As expected, the NN distribution of $Z \rightarrow \tau\tau$ MC events is peaked at 1.0. However, the NN outputs for tau candidates from single top quark events also have a small peak at low NN output, which indicates that the standard tau NN ID identifies some true taus as fakes. The

---

[2]The two samples are the tau signal samples used by the standard D0 tau NN ID samples. One with m_Z=130-250 GeV serves as a training sample while that with m_Z=60-130 GeV is the test sample.
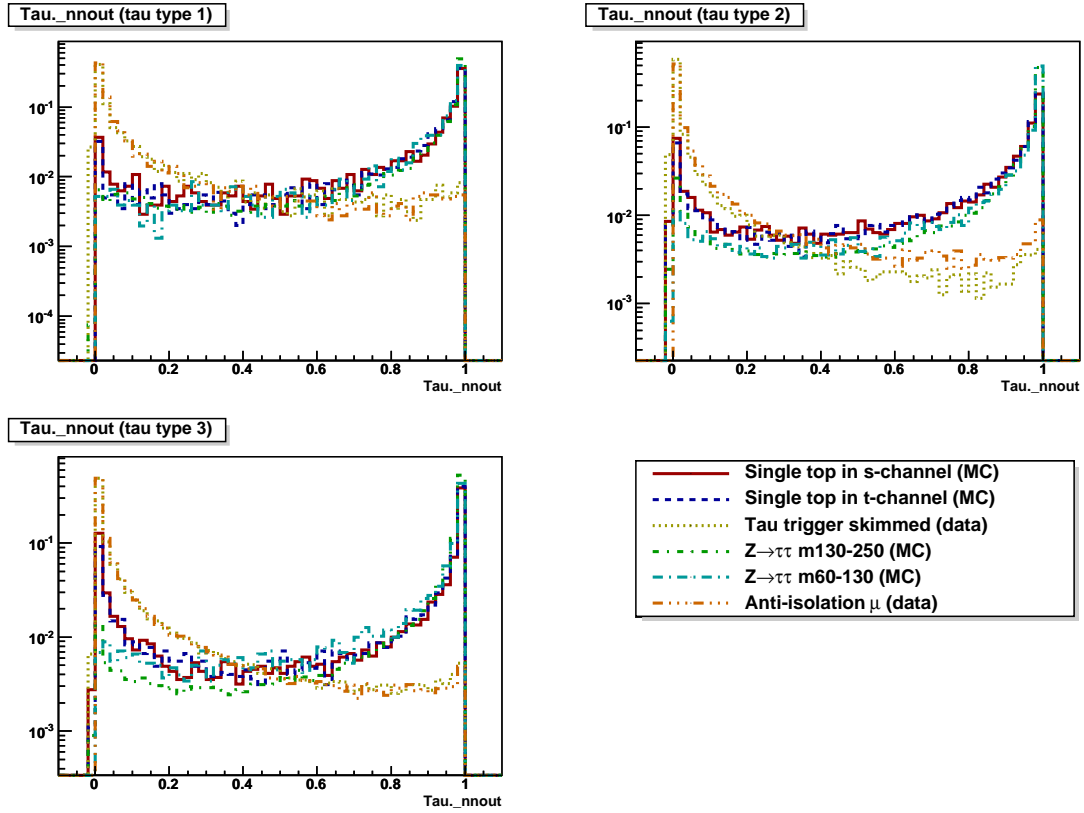
Figure 7.1: Comparison of NN output between different samples.

standard tau ID is not optimized for the single top analysis.

7.4.2 Hadronic Tau ID Optimization Strategy and Results

As seen from the discussion above, the definition of training and testing samples is crucial for tau ID. The different samples may be characteristic of different tau kinematics of interest. In addition, multivariate techniques and multivariate training variables also play important roles. Discussions below will focus on these three aspects.

Training and Testing Samples

In the optimization study, the tau signal sample is made from reconstructed taus matched to true-generated MC visible taus in single top quark tau channel events with either a spatial

distance $R_{\text{reco. }\tau,\text{true }\tau} < 0.5$ or the least spatial distance. A general dataset is skimmed by data quality and tau-jet trigger requirements[3]. Reconstructed taus in the skimmed dataset are treated as a fake tau background sample with very little real tau contamination. In order to mimic the cuts used in the realistic analysis, both the tau signal and background samples also have the following kinematic cuts applied:

Cut	Tau type 1	Tau type 2	Tau type 3
$|\eta| <$	3.0	3.0	3.0
$E_T >$	10 GeV	5 GeV	10 GeV
leading $p_T^{\text{trk}} >$	7 GeV	5 GeV	5 GeV
$\Sigma_{\text{trk}} \, p_T^{\text{trk}} >$	–	–	7 GeV

Then each of the signal and background samples is split into two equal-sized subsamples for training and testing purposes.

Neural network tau identification (NN) and boosted decision trees (BDT)

NN is the default multivariate technique used to derive the official D0 tau ID by the D0 tau ID group [77]. Chapter 6 has provided an introduction to the BDT technique. Compared to NN, BDT has some advantages as follows:

– BDT employs a binary tree technique, which makes fast training possible. With similar performance to other multivariate techniques such as NN, BDT runs faster [72]. The BDT training time depends on several factors, for example, number of training variables used, sample size and BDT boosting tree cycles.

– There is no special requirement on the range of BDT input training variables, while for NN, the range of the variables must be adjusted to [0, 1].

– Boosting is available in BDT to improve stability and training performance.

– BDT performance is insensitive to variable correlations. This means that adding well-modeled training variables never degrades performance.

---

[3]"tau-jet triggers" here are online-level jet triggers since most taus can also be identified by jet trigger algorithm. Most tau candidates in the tau-jet skimmed sample ($> 98\%$) are fake taus: real jets or fake jets.

In order to justify which technique has a better performance in this tau optimization, a comparison between BDT and NN is conducted given exactly the same conditions for both techniques. In the comparison, the BDT algorithm (see Section 6 for more about the algorithm) used is provided by the TMVA package[4]. The BDT training parameters are 50 tree cycles (1 original tree + 49 boosting trees), $\beta = 0.5$, $N_{\text{min. leaves}} = 100$. The NN algorithm used is the one used by the D0 tau ID group. Standard D0 tau ID samples and variables [77] matching the ones used by the D0 tau ID group are used for both BDT and NN training. Table 7.1 lists variables (defined in Section 7.2 and Appendix C) used in the comparison. Fig. 7.2 shows the NN and BDT output distributions. From Fig. 7.3, it is concluded that NN and BDT have comparable performance for tau types 1 and 3 when trained using exactly the same training samples and variables.However, the BDT is much better than the NN for tau type 2. So, the BDT is selected to be the default multivariate technique in the tau ID optimization for this thesis.

(a)   (b)
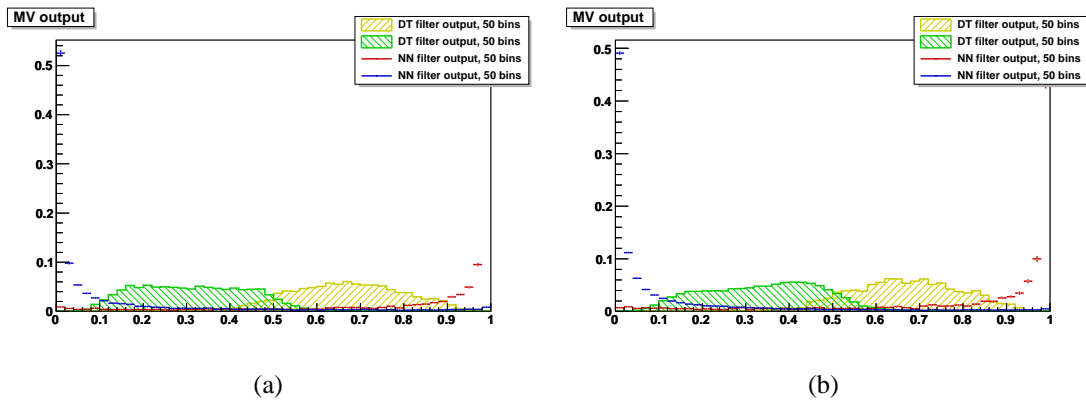
Figure 7.2: BDT and NN ouput distributions (a) tau type 2 (b) tau type 3. The shaded distribution in yellow is the BDT output of the signal sample while one in green is that of the background sample; The distribution in red is the NN output of the signal sample while one in blue is that of the background sample.

---

[4]The version of the package TMVA used is 3.9.2 with modification on the node splitting algorithm used by the BDT package developed by SFU.
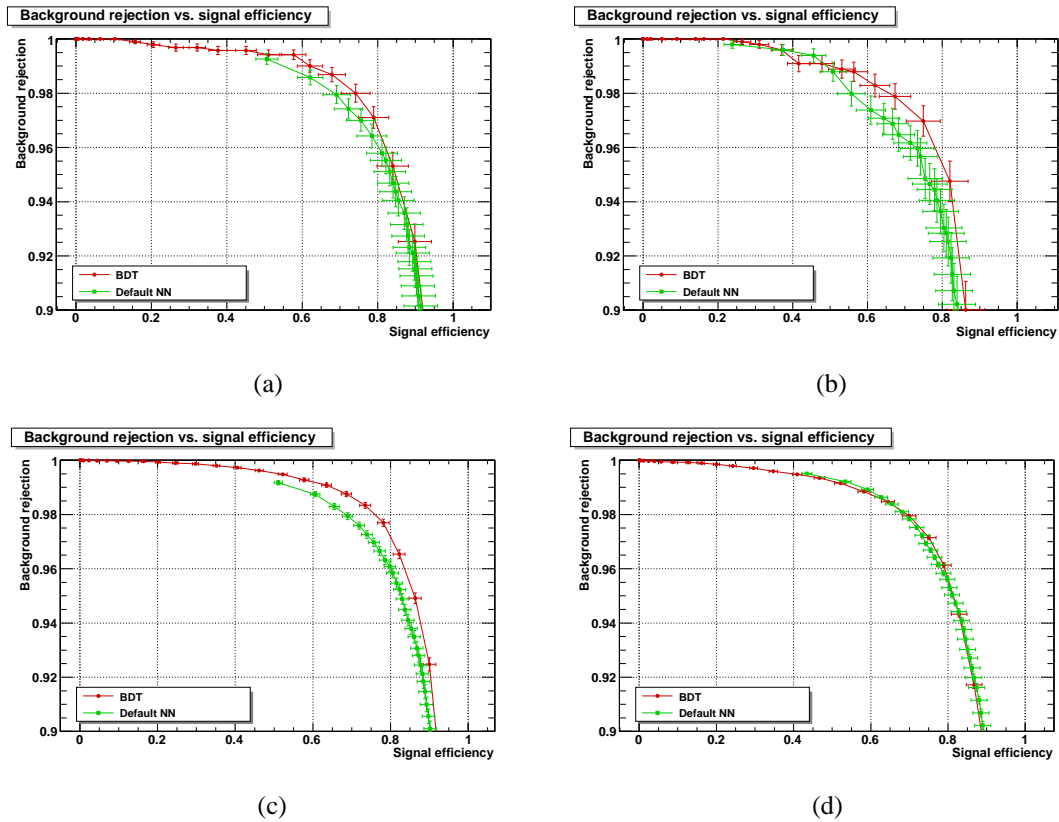
Figure 7.3: Comparison of background rejection vs. signal efficiency between NN and BDT based on the same training samples and training variables. (a) tau type 1 except ICD region (b) tau type 1 ICD region (c) tau type 2 (d) tau type 3

Table 7.1: Variables used per tau type in the standard tau NN. Performance comparison between BDT and NN is based on those variables (see Section 7.2 and Appendix C).

Variable	Tau type		
	1	2	3
EM12isof	✓		
iso	✓	✓	✓
profile	✓	✓	✓
prf3		✓	
trkiso	✓	✓	✓
rms	✓	✓	✓
ET_o_sum	✓	✓	✓
fhf	✓	✓	✓
etad	✓	✓	✓
dalpha		✓	✓
emET_o_ET		✓	✓
ett1_o_ETiso			✓

Training variables

A tau object reconstructed by D0 software has ~40 associated discriminating variables (properties). However, not all of them will be used in tau ID training since there are some strategies for selecting training variables:

- Training variables must be well modeled

- Training variables should not be strongly influenced by overall event kinematics

- Variables should have some distinguishing power between taus and other types of jets

- Select as few as possible in order to decrease multivariate technique training time and reduce complexity.

In order to reduce complexity without affecting performance, a following "tear-down" procedure is used to reduce variables given the facts that:

– extra variables may not significantly improve performance

– adding new variables to a variable list may change the ranked order of existing variables

In the procedure, all variables are gradually reduced using the following steps:

1. All available tau variables are used. The list is called `reduction-level-0`.

2. Train a BDT on the entire variable list above. After training, rank variables in the list by discriminating power, pick the first N_1 variables as `reduction-level-1`.

3. Then go on training a BDT on the `reduction-level-1` list and rank a new list in discriminating power, pick the first N_2 variables as `reduction-level-2`.

4. Go on reducing variables in the list until the number of variables reaches 10 at most.

5. Compare performance of the BDT achieved using these variable lists to select which level is the one with acceptable performance and few variables included

The procedure above is done for the three tau types individually. A variable change map of tau type 2 is shown in Figs. 7.4 to illustrate how the procedure works. On the map, each variable column corresponds to a variable level. After each training, the five least powerful variables in gray are removed from the list. Arrows indicate ranking "flow" direction of each variable. Graphs in Fig. 7.5 show performance using different training variable lists. Performance using level-1 to 4 variables is almost same within statistical uncertainty. The level-4 list contains 20 final training variables and is used in the realistic analysis. As a comparison, the NN default performance curves are also shown in the same graphs. The new trained BDT for tau type 2 and 3 have much better performance than the default NN. Due to the limited statistics of training samples, the difference in performance for tau type 1 can be explained as statistical fluctuation.

Tau properties vary with tau energy because the taus in different energy ranges have different shower shapes. Since tau energy is not used as one of training variables and the BDT may not be able to distinguish the tendency of slight changes in τ shower shape with energy, it may be worthwhile to split training samples in tau transverse energy bins to grow separate trees and apply these trees back to the samples in corresponding E_T. Graphs in

profile	profile	profile	profile	profile	profile
rms	rms	rms	rms	rms	rms
empt/ET	dalpha	profile2	profile2	profile2	prf3
prf3	trkiso	e1e2/ET	e1e2/ET	dalpha	e1e2/ET
e1e2/ET	e1e2/ET	prf3	empt/ET	trkiso	dalpha
dalpha	profile2	trkiso	dalpha	e1e2/ET	ett1/ET
emcl_et2/ET	empt/ET	ET_o_sum	emcl_et2/ET	empt/ET	emcl_et2/ET
trkiso	ettr/ET	dalpha	ET_o_sum	prf3	empt/ET
ET_o_sum	ett1/ET	ett1/ET	ett1/ET	rms2	EM12_Et/ET
emcl_et1/ET	prf3	emcl_et2/ET	trkiso	EM12_Et/ET	rms2
ett1/ET	fhf	Et_iso/ET	EM4isof	ett1/ET	profile2
rms2	rms2	empt/ET	prf3	iso2	trkiso
iso2	ET_o_sum	ettr/ET	iso2	emcl_et2/ET	iso2
Et_iso/ET	EM4isof	tzDCA	ettr/ET	EM12_Et_iso/ET	EM12_Et_iso/ET
chf	tzDCA	rms2	rms2	ettr/ET	ettr/ET
icdf/ET	emcl_et2/ET	emcl_f12	EM3_Et/ET	ET_o_sum	
EM12isof	emcl_f12	EM12_Et_iso/ET	EM12_Et_iso/ET	EM4isof	
EM4isof	EM12_Et/ET	EM3_Et/ET	EM12_Et/ET	EM3_Et/ET	
profile2	Et_iso/ET	EM12_Et/ET	emcl_f12	emcl_f12	
fhf	EM3_Et_iso/ET	icdf/ET	icdf/ET	icdf/ET	
EM12_Et/ET	EM3isof	EM3isof	Et_iso/ET		
ettr/ET	hot	EM4isof	tzDCA		
tzDCA	emcl_et1/ET	chf	EM3isof		
EM3_Et_iso/ET	icdf/ET	emcl_f3	chf		
emf	EM12_Et_iso/ET	iso2	emcl_f3		
EM12_Et_iso/ET	iso2	fhf			
emcl_f12	emcl_f3	EM3_Et_iso/ET			
hot	chf	hot			
EM4_Et/ET	EM4_Et/ET	emcl_et1/ET			
EM3_Et/ET	EM3_Et/ET	EM4_Et/ET			
emcl_f3	EM12isof				
EM3isof	emf				
EM4_Et_iso/ET	EM4_Et_iso/ET				
emcl_f4	emcl_f4				
EM4f	EM4f				

Variable Ranking Change Map (Tau type 2, NoCut)

Figure 7.4: Change map of variable ranking for tau type 2 (for the variable definitions, see Section 7.2 and Appendix C).

Fig. 7.6 show this effect for tau type 1, 2 and 3 respectively. Performance based on E_T-split training samples is slightly better than without splitting for all τ types. The E_T-split BDT are used in the final analysis.

In the standard tau NN ID, a regular cut on NN output of each tau type is recommended by the D0 tau ID group: NN>0.9 for tau type 1 and 2, and NN>0.95 for tau type 3 (tau type 3 has a larger background from jets, so a tighter cut is needed than tau type 1 and 2). Table 7.2 shows given a NN cut (column 1), that corresponds to the NN background rejection rate (column 2) and the NN signal efficiency (column 3). If the BDT background rejection rate is fixed to the same value as the NN, the BDT signal efficiency (column 4) and the equivalent BDT cut (column 5) are calculated using signal and background efficiency curves of NN and BDT. The last column "BDT cut" means the equivalent BDT cut to achieve the given background rejection rate. Of course, with respect to NN and BDT performance, the conclusion is consistent with the graphs shown above. For example, if NN cut = 0.9 for tau type 2, the NN background rejection rate is about 98.3%, the NN's

signal efficiency is 51% while the BDT's is ∼70%, a more than 40% gain. Fig. 7.7 shows the curves of equivalent BDT cuts given NN cuts on which three cut points corresponding to NN=0.9, 0.9, 0.95 have been marked by red dots.

Table 7.2: Signal efficiency comparison between NN and BDT given the NN cut and the same background rejection rate for tau type 2

NN cut	NN Background Rejection	NN Signal Eff.	BDT Signal Eff.	BDT cut
0.05	74.6%	90.2%	97.4%	0.452
0.10	81.9%	87.4%	95.8%	0.479
0.15	85.6%	85.6%	94.6%	0.495
0.20	88.1%	83.8%	93.5%	0.509
0.25	89.7%	82.2%	92.8%	0.518
0.30	91.2%	80.8%	91.6%	0.532
0.35	92.3%	79.4%	90.4%	0.544
0.40	93.2%	78.1%	89.4%	0.554
0.45	94.0%	76.6%	88.2%	0.567
0.50	94.5%	75.0%	87.4%	0.574
0.55	95.0%	73.2%	86.5%	0.582
0.60	95.5%	71.6%	85.3%	0.592
0.65	96.0%	69.5%	84.2%	0.601
0.70	96.4%	67.2%	83.0%	0.611
0.75	96.8%	64.5%	81.6%	0.622
0.80	97.3%	61.0%	79.6%	0.637
0.85	97.7%	56.8%	76.5%	0.657
0.90	98.3%	50.6%	69.0%	0.696
0.95	99.0%	39.5%	60.5%	0.733

Fig. 7.8 is a summary plot of signal efficiency ratio of BDT over NN in terms of background rejection efficiency in percent for τ type 1, 2 and 3. The higher the background rejection efficiency is, the more the gain of the signal efficiency. The gain for tau type 2 goes as high as 50% while type 1 and 3 improve by 15–20% at the very highest rejection.
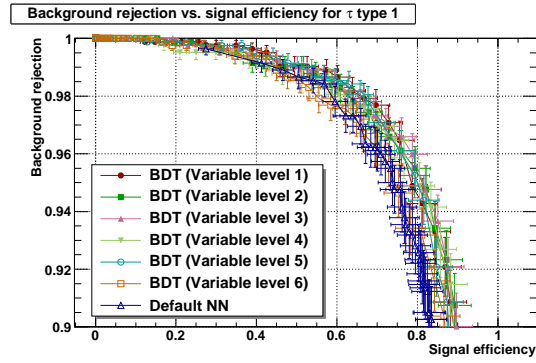
(a)
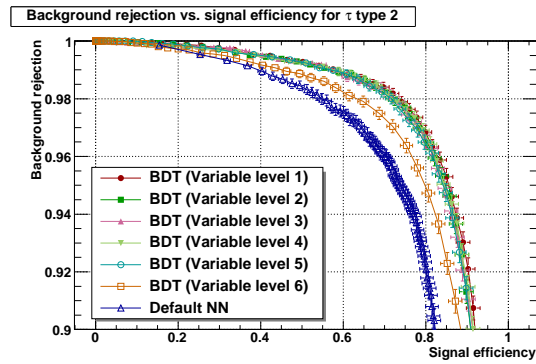
(b)
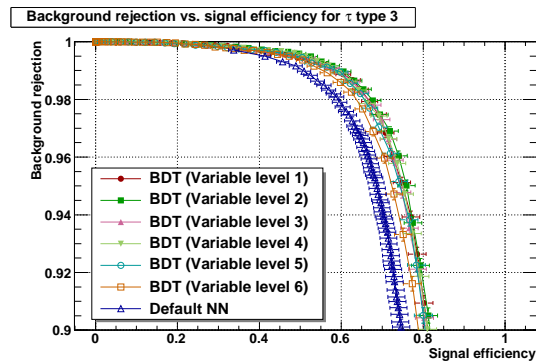
(c)

Figure 7.5: Comparison of background rejection rate vs. signal efficiency with different variable levels. (a) τ type 1, (b) τ type 2 and (c) τ type 3.
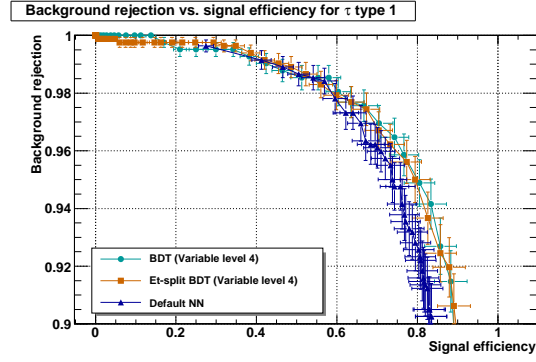
(a)
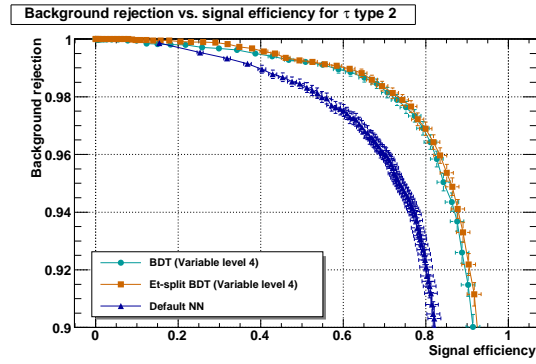
(b)
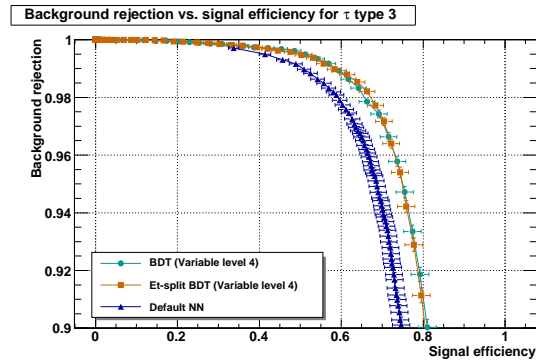
(c)

Figure 7.6: Comparison of background rejection rate vs. signal efficiency for the no-E_T-split and E_T-split cases. (a) τ type 1, (b) τ type 2 and (c) τ type 3.
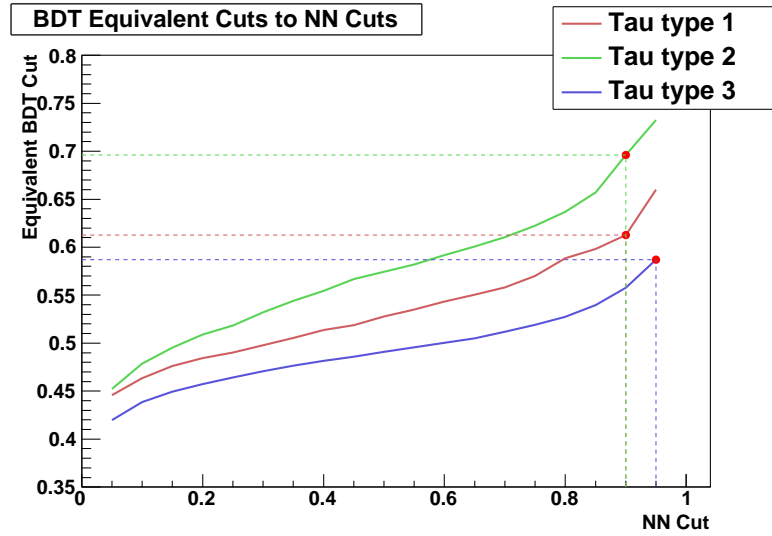
Figure 7.7: Map between NN cuts and equivalent BDT cuts.
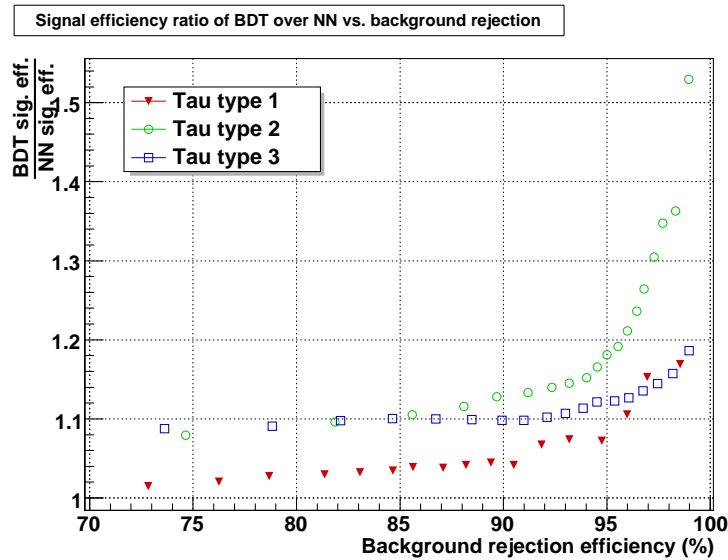
Figure 7.8: Summary plot of signal efficiency ratio of BDT over NN in terms of background rejection efficiency in percent.

Application of BDT to electron signal channels

Electrons leave tracks and deposit energy in a narrow cone in the calorimeters, so they can also be found by tau reconstruction algorithms with high reconstruction efficiency, and most of them are labelled as tau type 2. Single top quark events of interest in the analysis have channels with the final state of electron plus jet(s) which are from the single top electron channel and tau channel decaying to electrons[5]. In fact, these electron events provide an important part of signal sensitivity. It is necessary to check the new tau ID performance when it is applied to the MC samples of $tb \rightarrow e\nu bb$ and $tqb \rightarrow e\nu bqb$. A basic cross check is to apply the new tau ID to reconstructed MC τs matched with true electrons with $\Delta R < 0.5$ to calculate performance diagnostic curves. Fig. 7.9 compares the performance of the default NN with BDT for tau type 2 based on the single top MC samples $tb \rightarrow e\nu bb$ and $tqb \rightarrow e\nu bqb$ (as signal) and a tau-jet trigger skimmed dataset (as background). It is concluded that the newly derived tau BDT ID again performs better than the standard tau NN ID. For the NN output cut 0.9, signal efficiency enhancement is $\sim 24\%$.
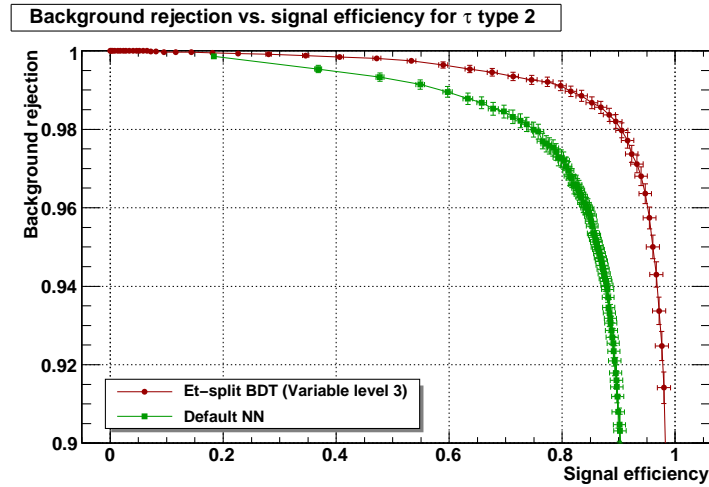
Figure 7.9: Performance comparison between default NN with BDT for tau type 2 based on the MC samples $tb \rightarrow e\nu bb$ and $tqb \rightarrow e\nu bqb$ (as signal) and QCD-multijet skimmed data (as background).

---

[5]The electron events are not present in the single-top-quark-to-electron channel analysis because the electrons do not pass the stringent criteria for that analysis.

7.5 Estimation of Systematic Uncertainty of Tau ID

7.5.1 Uncertainty Derived from Fluctuation of Input Variables

For each BDT input variable, the predicted distributions of signal and background events are divided in a number of bins. The number of events in each bin is varied according to the statistical error (assuming a Gaussian distribution) and the fluctuation is propagated to its BDT output distribution. Thus when applying a BDT cut, there is an uncertainty originating from the choice of the cut [79].

The basic steps to estimate the uncertainty are:

– Generate pseudo-experimental datasets for one input training variable. For the input variable, this is done by sampling the expected number in each bin of the variable distribution from a Gaussian distribution to form a new distribution. The mean of the Gaussian distribution is calculated as a difference between the number of data and background events in each bin. Then a new BDT output distribution is obtained by propagation from the generated spectrum of that input variable by reweighting. For one input variable, 50 pseudo-experimental datasets (corresponding to 50 weighted BDT output distributions) are generated.

– Calculate the ratio of the number of events passing a particular BDT cut on the BDT output distributions generated above to the number without the cut. The standard deviation of the ratio distribution is taken as a measure of the fluctuation uncertainty for this variable.

– Repeat the steps above for all input variables.

– The overall systematic uncertainty from all input variables is calculated as the square root of a quadratic sum of the standard deviations of all input variables.

The samples used for studying single top quarks contain few tau events and are not appropriate for this approach. The data sample used here was skimmed by applying single muon triggers, muon and tau kinematic cuts and data quality definitions for identifying a tau+muon pair. The muon and tau must be back-to-back, i.e., $\Delta\phi(\text{muon}, \text{tau}) > 2.5$. Background MC samples include $W \to \mu\nu$ and $Z \to \mu\mu$ and the signal MC sample is $Z \to \tau\tau$.
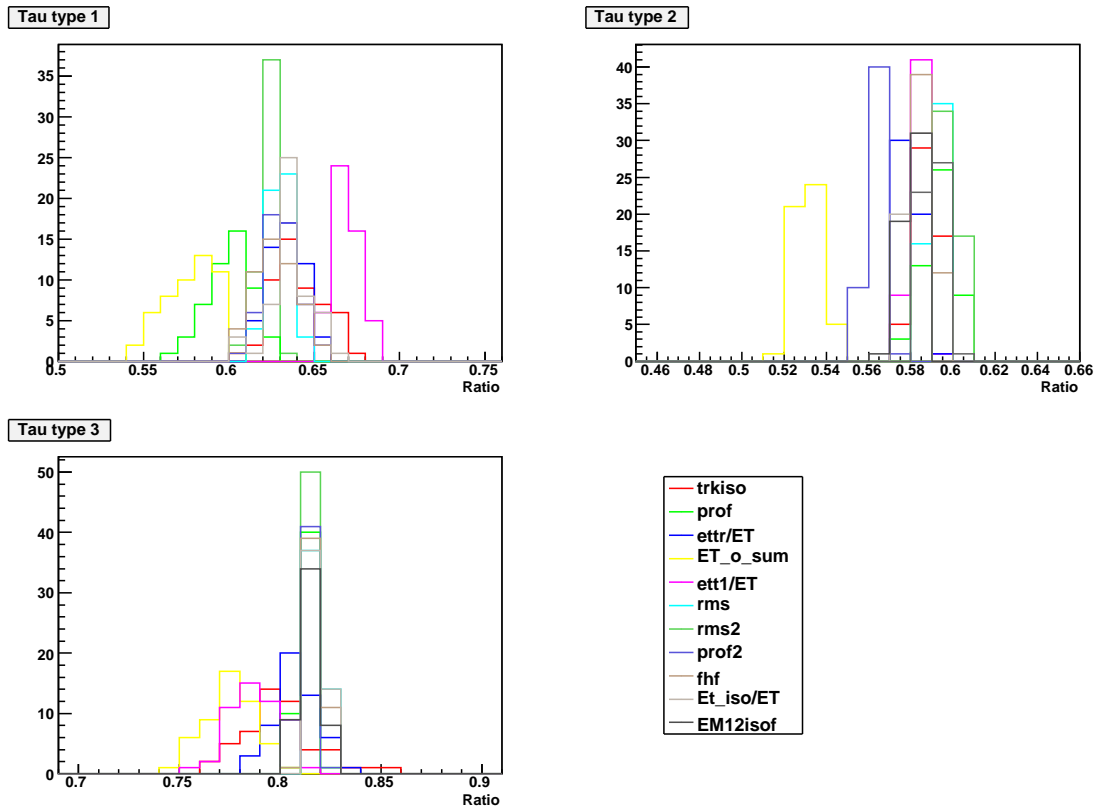
Figure 7.10: Example distributions of the ratio of the number of events which pass the equivalent NN cut, 0.90, to the total number of events before the cut. Each distribution corresponds to one ensemble of 50 pseudo-datasets fluctuating an individual variable.

Each sample was weighted by the trigger probabilityafter applying the same kinematic cuts as the data sample.

Figure 7.10 shows the ratio distributions for different BDT input variables from which a systematic uncertainty value is derived. Each entry of a ratio distribution corresponds to a pseudo-experimental dataset. The square root of the quadrature sum of the standard deviations of the ratio distributions is counted as a systematic uncertainty from fluctuation. Table 7.3 lists systematic uncertainties in tau types given different BDT cuts[6].

---

[6]Those BDT cuts are equivalent to NN cuts 0.7, 0.8, 0.85, 0.9 and 0.95 respectively. "Equivalent" means that the same background rejection rate is given.

Table 7.3: Systematic uncertainty values in tau types given different BDT cuts

Equivalent NN cut	Tau type 1		Tau type 2		Tau type 3	
	BDT cut	Uncertainty	BDT cut	Uncertainty	BDT cut	Uncertainty
0.70	0.554	0.57%	0.608	0.53%	0.518	0.44%
0.80	0.598	1.07%	0.643	0.85%	0.533	0.65%
0.85	0.630	1.63%	0.665	1.01%	0.542	0.99%
0.90	0.652	1.81%	0.697	1.22%	0.559	0.99%
0.95	0.707	2.55%	0.744	1.80%	0.598	1.60%

7.5.2   Uncertainty of BDT Output Efficiency

Another important uncertainty is from the different tau ID efficiency between data and MC. Ideally, a correction to the tau ID output probability density function (PDF) should be derived andapplied to MC samples to fix the MC tau ID efficiency. However, since the BDT tau ID is based on the single top sample and its yield is very small, there is no direct approach available to derive such a correction in our case. The samples enriched in $Z \rightarrow \tau\tau$ events used in the D0 standard NN tau ID study are also employed here. Fig. 7.11 (a) shows a comparison plot of BDT outputs of $Z \rightarrow \tau\tau$ enriched samples for τ type 2 between data and background while Fig. 7.11 (b) is a ratio of PDF of data over PDF of MC for $Z \rightarrow \tau\tau$ events, and the straight line and grey band are the fit line and its 1σ confidence band. The band is almost overlapping with 1.0 and it is reasonable to assign a 10% uncertainty to replace tau ID correction.

This uncertainty is summed in quadrature with the statistical fluctuation uncertainty mentioned above and the overall tau ID uncertainty is quoted as 11%.

7.6   Summary

This chapter discusses the study of tau ID optimization for single top search in the τ+jets channel. The two multivariate techniques NN and BDT are compared. Given a NN background rejection rate of $\sim$98%, the NN signal efficiency is 51% while the BDT efficiency is 70% for tau type 2, a relative gain of 40%. For tau type 1 and 3, a relative gain of $\sim$18%

(a)   (b)
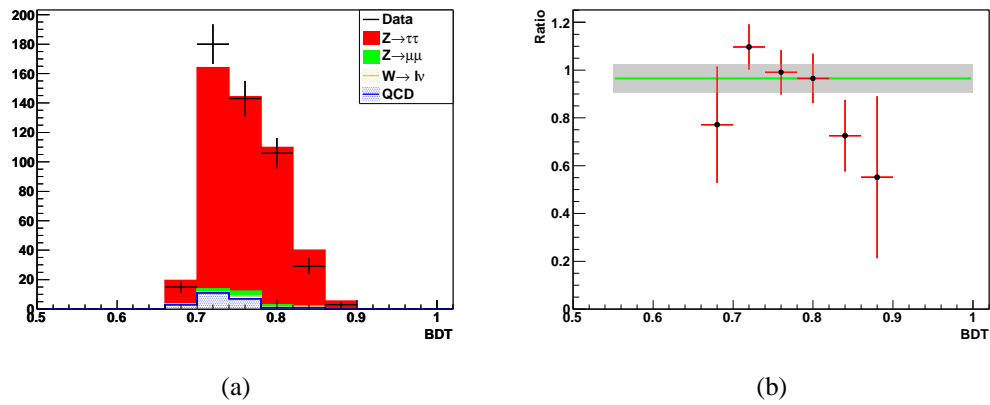
Figure 7.11: $Z \to \tau\tau$ enriched samples for τ type 2. (a) comparison plot of BDT outputs between data and background; (b) Ratio of PDF of data over one of MC for $Z \to \tau\tau$ events. In (b), the straight line and grey band is the fit line and its 1σ confidence band.

can be achieved.

Chapter 8

Background Modeling

A high energy physics analysis aiming to search for small signals needs a very precise background model that is descriptive of the final analysis sample. The quality of such an analog between the final analysis sample and the modeled background sample can be estimated by comparing various topological variables in those samples. If those variable distributions are well-matched to each other, then the two samples are compatible in many dimensions and the model is suitable.

In this chapter, Section 8.1 provides an outline of the background modeling. QCD-multijet tag rate functions play a very important role in modeling the background. Sections 8.2, 8.3 and 8.4 present measurement, normalization and application of QCD-multijet tag rate functions. Section 8.5 talks about non-QCD-multijet contamination removal aiming at achieving a pure QCD-multijet sample. Section 8.6 presents a QCD-multijet normalization based on a QCD-multijets-enriched sample. At last, quality checking of the model is given in Section 8.7 by comparing the final analysis sample and the background sample.

8.1   Outline of the Background Modeling

One important signature of single top quarks in any channel is the presence of at least one b quark in final state. Signal-to-background is therefore enhanced by requiring the presence of b quark jet(s). At an operating point $NN > 0.775$, b-jets are tagged with approximately 40% average efficiency and requiring 1 or 2 b-tags leads to a $\sim$55% yield for single top signal. At this operating point, the mistag rate for light-quarks and gluons is $\sim$0.4%. Of

course, b quarks can also be produced in strong interactions like $qq(gg) \rightarrow b\bar{b}$ production, so the search still suffers from both fake-b and real-b backgrounds. The cross section for $b\bar{b}$ is as much as 3 orders of magnitude smaller than light jet QCD-multijet events, so light-quark mistags are expected to dominate.

As addressed previously, a tau in a final state is likely to be reconstructed as a narrow-cone jet and, conversely, a narrow jet may be reconstructed as a fake tau by the tau reconstruction algorithm. For this reason, the main background to single top quarks in the tau+jets channel is QCD-multijet events. This is unlike the other leptonic channels of single top quarks where W+jets events are the main background. At D0, QCD-multijet events are not simulated well by the MC generation, thus they have to be derived from data. For other minor background sources such as W+jets, Z+jets, $t\bar{t}$ and diboson, there are corresponding MC samples available to use. They have been described in Chapter 4. Our background model is formulated as

$$\langle \text{Background} \rangle = \langle \text{QCD-multijets} \rangle + \langle W + \text{jets} \rangle + \langle Z + \text{jets} \rangle + \langle t\bar{t} \rangle + \langle \text{diboson} \rangle. \quad (8.1)$$

Fig. 8.1 shows examples of leading order Feynman diagram for a typical QCD-multijet event and a W+jets event.
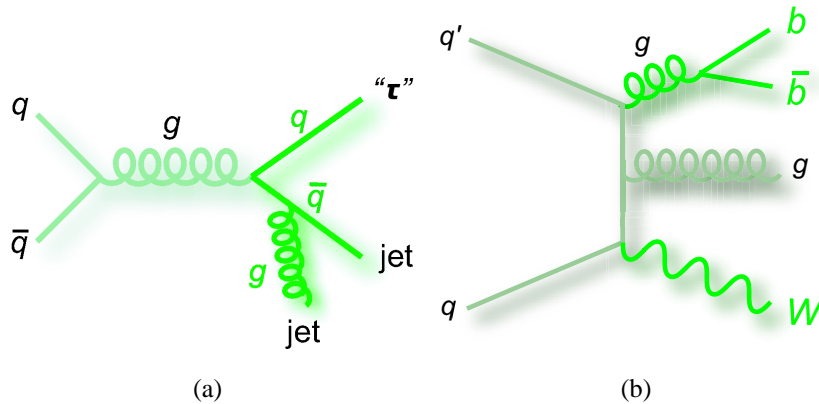
Figure 8.1: Examples of leading order Feynman diagram for (a) a QCD-multijet event and (b) a W+jets event.

The principal steps in the background model can be summarized as:

1. Derive a tag rate function (TRF) to describe the probability to b-tag any individual jet in the sample.

2. Apply this TRF to the data and MC samples that have no *b*-tagged jets.

3. Using simulated MC events for other physics sources, subtract them from the TRFed data sample to get "pure QCD-multijets".

4. Combine the derived background sample, pure QCD-multijets, with simulations of other background sources: $t\bar{t}$, W+jets, Z+jets, dibosons.

The corresponding flow chart is shown in Fig. 8.2 and each step above has been labeled with Arabic numbers in the figure.

In Step 1, the ratio of the number of *b*-tagged jets in our data sample to the total number of jets is defined as a tag rate: the average probability that a jet is identified as a *b* jet. The tag rate is measured as a function of jet p_T, η and multiplicity. More information can be found in Sections 8.2 and 8.3.

In Step 2, these TRFs are applied to those data and MC events that have no *b*-tagged jets. The TRFed data sample is kinematically similar to our analysis sample, but there is no overlap since at least one *b*-tagged jet is required in our analysis sample.

In Step 3, physics background sources such as $t\bar{t}$, W+jets, Z+jets and dibosons are removed. In this procedure, the contaminations of $t\bar{t}$, W+jets, Z+jets and diboson are subtracted from the 0-*b*-tagged TRFed QCD-multijet sample. Other background sources are modeled through simulations. See Section 8.5 for this removal. A similar procedure is used to ensure that any small single top signal contamination in the background data sample is also subtracted. For signal contamination removal, see Section 9.3.4.

In Step 4, the QCD-multijet events after contamination removal are normalized to data in a QCD-multijets-enriched region, as defined by the background-dominated region of the multivariate discriminant described below. The normalization will be discussed in Section 8.6. This data-derived QCD-multijet sample is then combined with the *b*-tagged MC samples to make the background model.

At the end of the background modeling procedure, approximately 150 topological variables are investigated to confirm that data and the background model are in good agreement.Since single top quark events represent only a small fraction ($\approx$2%), and are spread throughout each distribution, signal will not significantly modify this agreement.
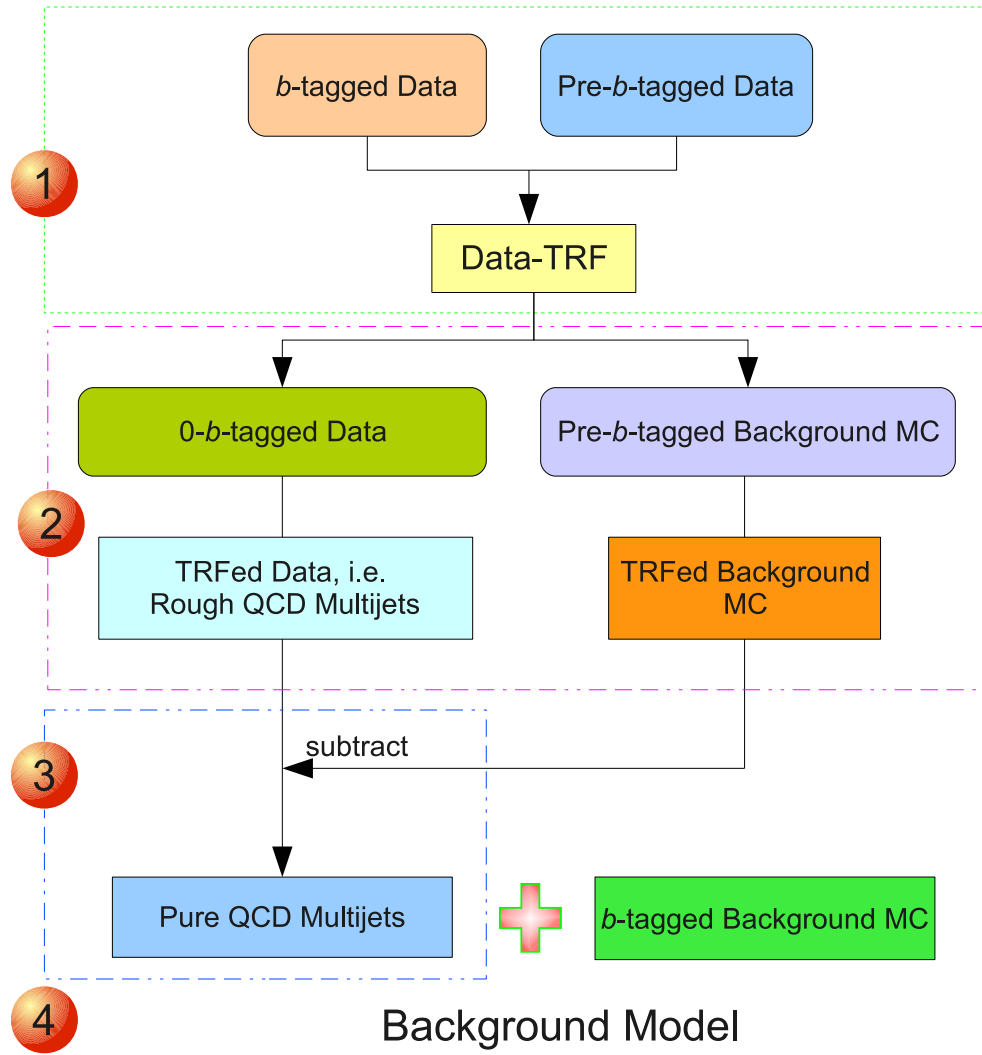
Figure 8.2: Flow chart to illustrate outline of the background modeling

8.2 Measurement of Tag Rate Functions from Data

A tag rate function is the probability that a taggable jet is identified as a *b*-jet by a neural network *b*-tagging algorithm. It is derived as a function of jet p_T and pseudorapidity η.

The TRFed QCD-multijet background prediction is comprised of the two basic steps as follows:

1. derivation of TRFs from data and

2. their application to the 0-*b*-tagged data sample

Data TRFs are derived based on the *b*-tagged and pre-*b*-tagged data samples that are selected by the object identifications and event selections listed in Tables 5.2 and 5.3. Data TRFs are then applied to a sample without *b*-tagged jets (called the 0-*b*-tagged sample), thus making sure that the data-TRFed multijet background is orthogonal to the final analysis data sample.

The definition of a tag rate is:

$$\text{TRF} = \frac{\text{Number of } b\text{-tagged jets in } b\text{-tagged sample}}{\text{Number of all taggable jets in pre-tagged sample}} \tag{8.2}$$

where the ratio is a ratio of **numbers of objects** instead of numbers of events. It is parameterized in terms of jet pseudorapidity η and transverse momentum p_T. Due to statistical limitations in some p_T-η bins seen from Fig.8.3, however, $\text{TRF}(p_T, \eta)$ can not be well measured. So, TRFs by $N \cdot \text{TRF}(p_T) \cdot \text{TRF}(\eta)$ (where N is a normalization factor) are measured. This works because the TRFs in terms of the jet p_T and the jet η are not highly correlated as seen in Fig.8.4

The data TRFs are also parameterized in jet multiplicity bins and in data-taking period bins (Run IIa and Run IIb). In order to take into account the TRF dependence on the energy scale of events, the TRFs are also measured in H_T bins, *i.e.* $H_T < 100$ GeV and $H_T > 100$ GeV samples. The TRFs in terms of jet p_T are fit using an empirical function:

$$\text{TRF}(p_T) = \frac{p_0}{2} \left(1 + p_1 \cdot \text{Erf}(\frac{p_T - p_2}{p_3 \cdot p_T^2}) \right) \tag{8.3}$$

on the RunIIa $H_T < 100$ GeV sample where p_0, p_1, p_2 and p_3 are fitting parameters, and Erf() is an error function. However, Eq. 8.3 doesn't give a good fit for the other datasets.
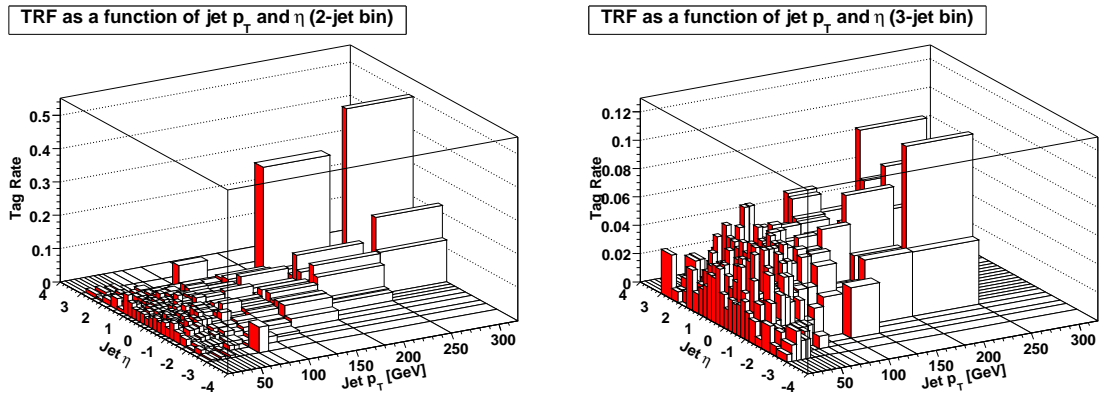
Figure 8.3: 2-D tag rate lego graph as a function of jet p_T and η. Left: 2-jet bin; Right: 3-jet bin. These plots show that some regions have too few events for a 2D TRF.
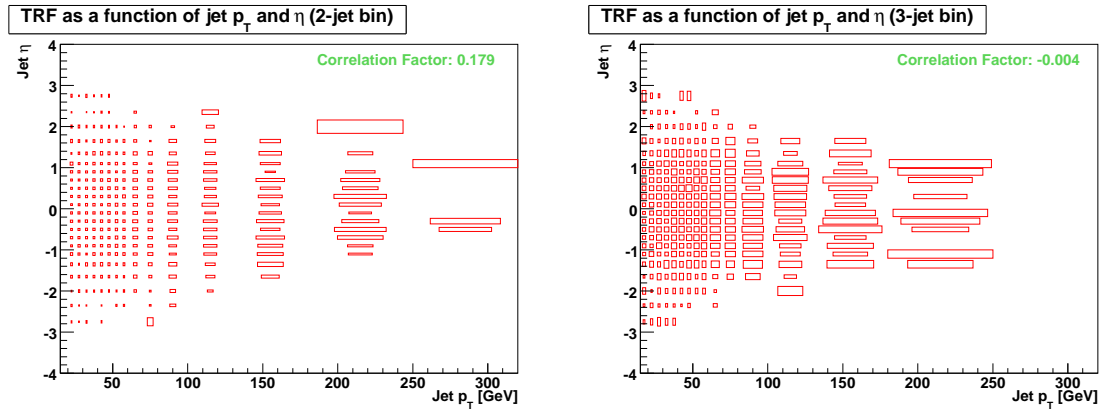
Figure 8.4: 2D Tag rate graphs as a function of jet p_T and η. Left: 2-jet bin; Right: 3-jet bin. Correlation factors shown on the graphs indicate that the TRFs in terms of the jet p_T and the jet η have low correlation.

So another empirical function

$$\text{TRF}(p_T) = p_0 + p_1 \log p_T + p_2 \log p_T{}^2 + p_3 \exp\left(-\frac{(p_T - p_4)^2}{2p_5^2}\right) \tag{8.4}$$

is used to do fitting on other samples. The last term is a Gaussian function with mean= p_4, $\sigma = p_5$.

The tag rate points in terms of jet η are fit using another empirical symmetric function:

$$\text{TRF}(\eta) = p_0 + p_1\eta^2 + p_2\eta^4 + p_3\eta^6 + p_4\eta^6 e^{p_5 \cdot |\eta|} \tag{8.5}$$

where p_is are fitting parameters. Figs. 8.5 – 8.12 show TRFs and their fitting results. The gray bands in the graphs are 68% confidence level bands by which the systematic uncertainty on TRF measurements is estimated.
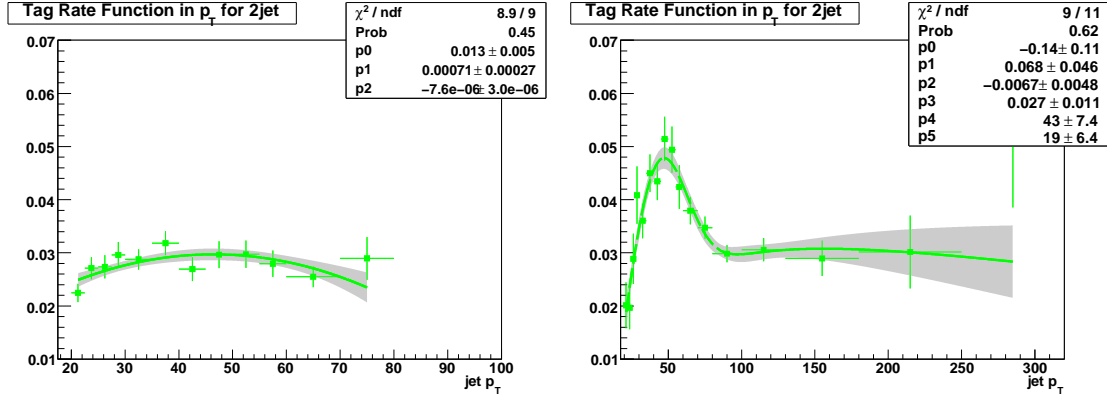
Figure 8.5: Tag rate curves and their fitting lines with 1 σ confidence band in terms of jet p_T in the 2-jet bin based on the Run IIa data sample. a) $H_T < 100$ GeV, b) $H_T > 100$ GeV

8.3 TRF Normalization

Since a product of $\text{TRF}(p_T)$ and $\text{TRF}(\eta)$ is used to represent $\text{TRF}(p_T, \eta)$, a normalization factor N is needed so that $\text{TRF}(p_T, \eta) = N \cdot \text{TRF}(p_T) \cdot \text{TRF}(\eta)$.
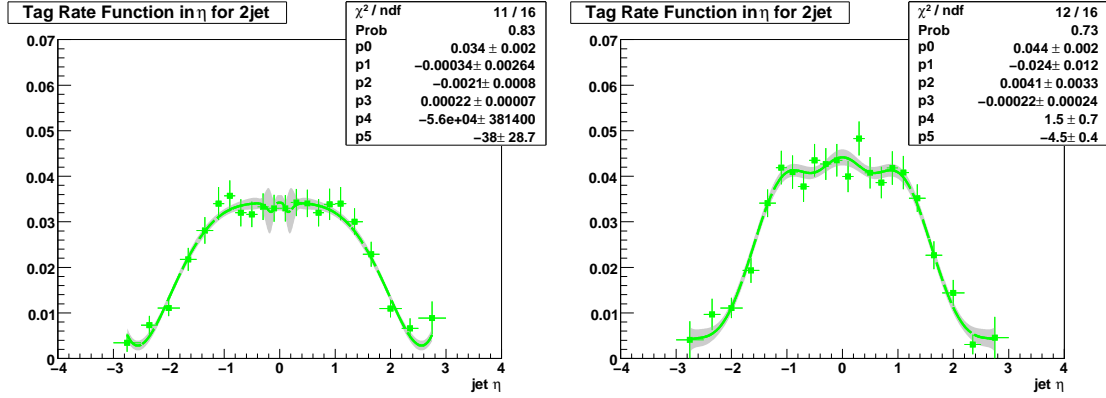
Figure 8.6: Tag rate curves and their fitting lines with 1 σ confidence band in terms of jet η in the 2-jet bin based on the Run IIa data sample. a) $H_T < 100$ GeV, b) $H_T > 100$ GeV
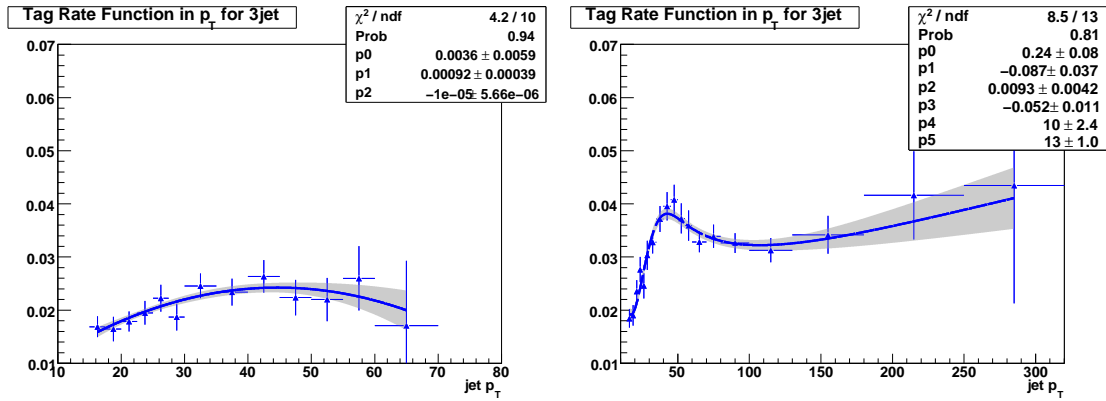
Figure 8.7: Tag rate curves and their fitting lines with 1 σ confidence band in terms of jet p_T in the 3-jet bin based on the Run IIa data sample. a) $H_T < 100$ GeV, b) $H_T > 100$ GeV

8.3.1 Normalizing in Each Jet Multiplicity Bin

The simplest normalization approach is

$$N_{\text{tagged jets}} = \sum_{evt} \left[\sum_i N \cdot f_i \right], \tag{8.6}$$

where the index i is the taggable jet index per event, $N \cdot f_i = N \cdot TRF(p_T) \cdot TRF(\eta)$ is the probablity of a taggable jet with p_T and η being tagged. N is derived separately in each jet multiplicity bin used in the analysis. For simplicity, only the case in one jet multiplicity bin
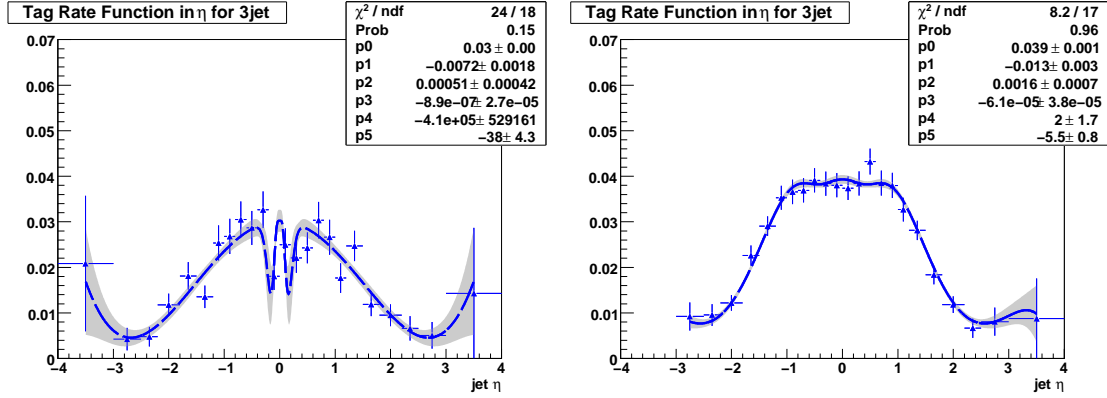
Figure 8.8: Tag rate curves and their fitting lines with 1 σ confidence band in terms of jet η in the 3-jet bin based on the Run IIa data sample. a) $H_T < 100$ GeV, b) $H_T > 100$ GeV
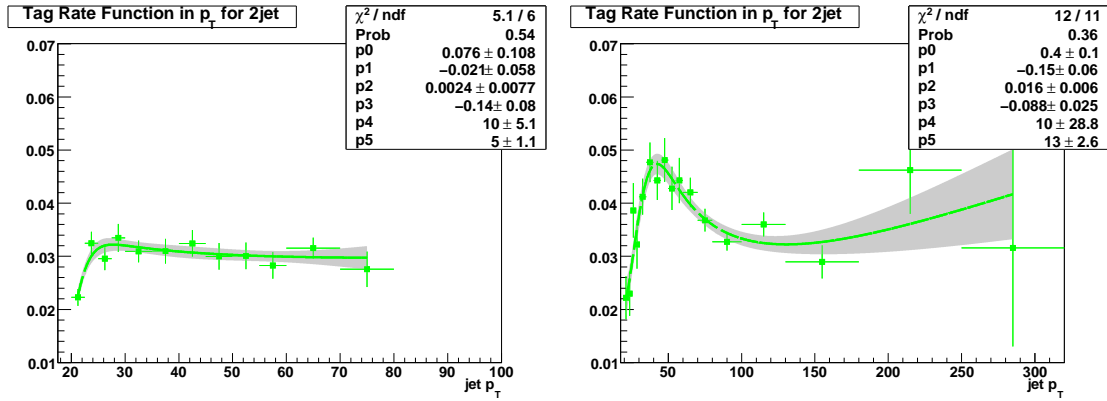
Figure 8.9: Tag rate curves and their fitting lines with 1 σ confidence band in terms of jet p_T in the 2-jet bin based on the Run IIb data sample. a) $H_T < 100$ GeV, b) $H_T > 100$ GeV

is discussed here. Using the normalization factor calculated by the simple approach above, there are two outstanding problems:

– The number of b-tagged jets does not match the data, as shown in Fig. 8.13.

– For a sample with 2 b-tagged jets per event, the predicted sample does a poor job of modeling variables which relate these 2 b-tagged jets, such as ΔR, $\Delta\phi$ *etc*. This arises because correlations between 2 b-tagged jets are not considered in the normalization. The correlation is due to the presence of b jets, originating from gluon splitting $g \to b\bar{b}$, in the sample. The higher the momentum of the gluon, the closer the 2 b jets will
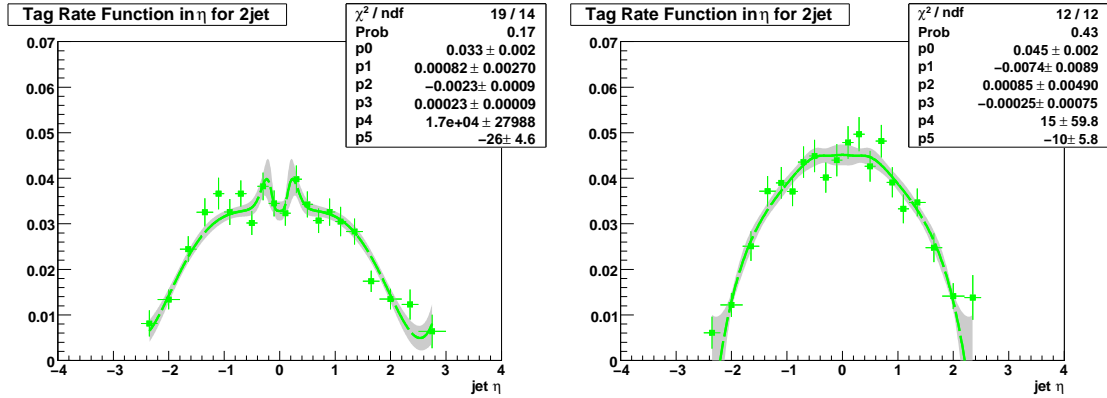
Figure 8.10: Tag rate curves and their fitting lines with 1 σ confidence band in terms of jet η in the 2-jet bin based on the Run IIb data sample. a) $H_T < 100$ GeV, b) $H_T > 100$ GeV
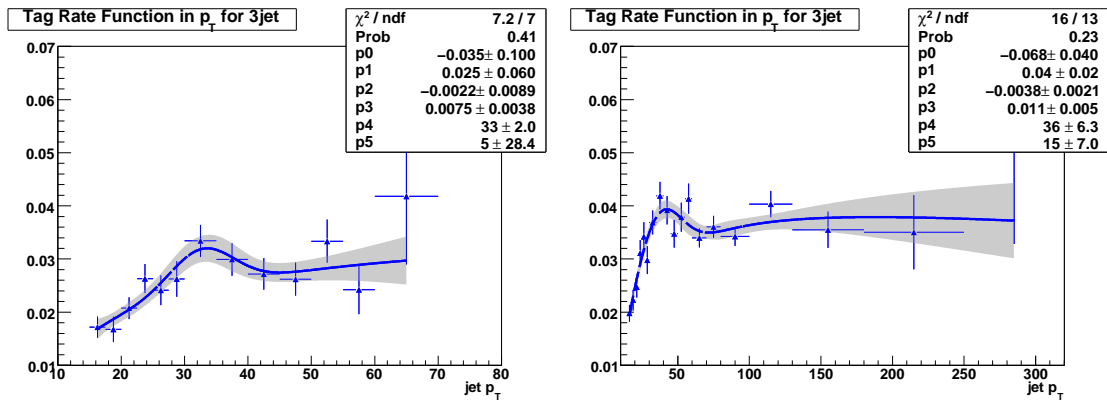
Figure 8.11: Tag rate curves and their fitting lines with 1 σ confidence band in terms of jet p_T in the 3-jet bin based on the Run IIb data sample. a) $H_T < 100$ GeV, b) $H_T > 100$ GeV

be. One approach to minimize this is to set a cut on ΔR to reduce contamination from the QCD $b\bar{b}$ events, however, this would cause a loss of signal events.

The two problems above can be fixed by solving two equations considering correlations between 2 b-tagged jets, which is discussed in the next section.
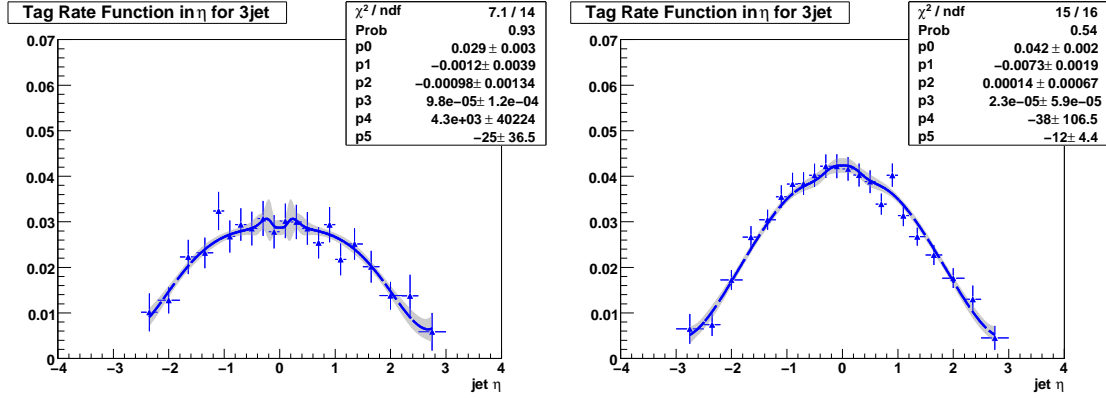
Figure 8.12: Tag rate curves and their fitting lines with 1 σ confidence band in terms of jet η in the 3-jet bin based on the Run IIb data sample. a) $H_T < 100$ GeV, b) $H_T > 100$ GeV

8.3.2 Separate Normalization by Number of b-tags

Due to the correlations discussed above, it is necessary to derive separate TRFs in the 1-b-tagged jet bin and the 2-b-tagged jet bin. Due to the limited sample size, however, TRF curves for the sample cannot be precisely measured in the 2-b-tagged-jet sample. However, we assume that TRF of the 2-tagged sample has the same shape as that of the 1-tagged sample, but a different normalization factor. The tag probability of one event with only **one** b-tagged jet is $\sum_i P_i \prod_j (1 - P_j)$ where i and j are taggable jet index per event while the tag probability for one event with only **two** b-tagged jets is $\sum_{i \neq j} P_i P_j \prod_{k \neq i, j} (1 - P_k)$ where i, j and k are taggable jet index per event. Since events with three or more b-tagged jets are rejected in the analysis, the total number of b-tagged **events** can be written as

$$
\begin{aligned}
N_{\text{tagged evts}} &= \langle \text{1-tagged-jet events} \rangle + \langle \text{2-b-tagged-jet events} \rangle \\
&= \sum_{evt} \sum_i \left[N_1 f_i \prod_{j \neq i} (1 - N_1 f_j) \right] \\
&+ \sum_{evt} \sum_{i \neq j} \left[\rho(\Delta R_{ij}) \cdot N_2 f_i \cdot N_2 f_j \prod_{k \neq i, j} (1 - N_2 f_k) \right]
\end{aligned}
\tag{8.7}
$$

where N_1 and N_2 are normalization factors in the single b-tagged jet bin and the double b-tagged jet bin, $f_i = TRF(p_T) \cdot TRF(\eta)$, $\rho(\Delta R_{ij})$ is a weighting function reflecting correlation between two b-tagged jets per event, which will be discussed later. Similarly, the
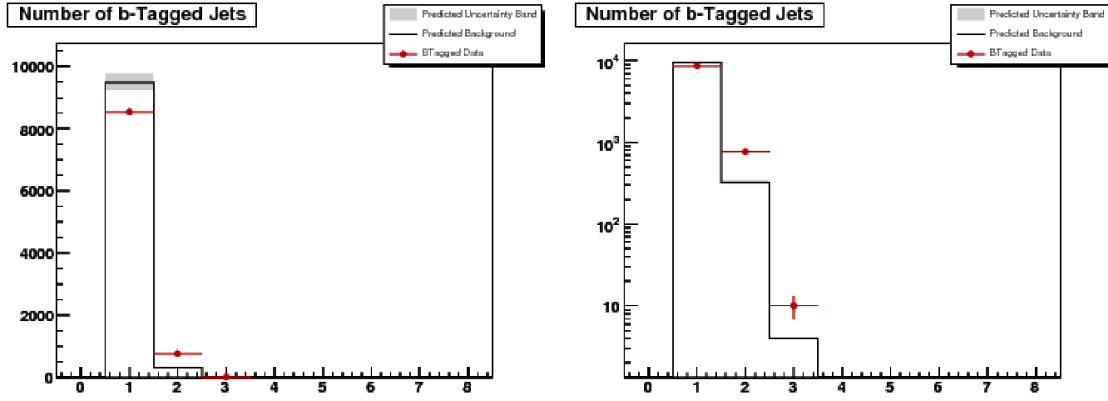
Figure 8.13: Number of *b*-tagged jets with one universal normalization factor. Left: linear scale; right: log scale.

total number of *b*-**tagged jets** can be written as

$$
\begin{aligned}
N_{\text{tagged jets}} &= \langle \text{b-tagged jets in 1-tagged-jet events} \rangle + \langle \text{b-tagged jets in 2-tagged-jet events} \rangle \\
&= \sum_{evt} \sum_i \left[N_1 f_i \prod_{j \neq i} (1 - N_1 f_j) \right] \\
&\quad + 2 \left\{ \sum_{evt} \sum_{i \neq j} \left[\rho(\Delta R_{ij}) \cdot N_2 f_i \cdot N_2 f_j \prod_{k \neq i,j} (1 - N_2 f_k) \right] \right\}
\end{aligned}
\tag{8.8}
$$

The first term is a number of *b*-tagged jets in single *b*-tagged events while the 2nd term is the number of *b*-tagged jets in double *b*-tagged events (the factor 2 in the front of the second term means 2 *b*-tagged jets per event). The only difference between these two equations above is the factor 2 so that it is easy to isolate N_1 and N_2. So

$$
2N_{\text{tagged events}} - N_{\text{tagged jets}} = \sum_{evt} \sum_i \left[N_1 f_i \prod_{j \neq i} (1 - N_1 f_j) \right]
\tag{8.9}
$$

and

$$
N_{\text{tagged jets}} - N_{\text{tagged events}} = \sum_{evt} \sum_{i \neq j} \left[\rho(\Delta R_{ij}) \cdot N_2 f_i \cdot N_2 f_j \prod_{k \neq i,j} (1 - N_2 f_k) \right]
\tag{8.10}
$$

where $N_{\text{tagged jets}}$, $N_{\text{tagged events}}$ and $f_{i,j,k}$ are all known. A tricky issue is how to calculate values of N_1 and N_2 from Eq.8.9 and 8.10, and how to take the correlation term $\rho(\Delta R_{ij})$ into

account in the predicted 2-b-tagged sample. Here iterations have to be used. For example, first fix N_1 in the bracket $(1 - N_1 f_i)$ as 0 and determine another N_1 in the Eq.(8.9), then in the 2nd loop fix N_1 in the bracket $(1 - N_1 f_i)$ as one determined in the 1st loop to determine a new N_1, thus after several iterations, N_1 is close to one asymptotic value. It is noticed that the relative change in N_1 is less than 0.1% after 5 iterations. Normalization factors in different jet multiplicity bins and different tau type bins are listed in Table 8.1.

Table 8.1: Normalization factors in different jet multiplicity bins and different tau type bins. The normalization factors' precision is shown with 3-4 significant figures while in practice 4 digits after the decimal are used within each iteration and the relative uncertainties on the factors are less than 0.1%.

H_T bin	Reco bin	Normalization Factor	2 jet	3 jet
$H_T < 100$	p17	N_1	36.6	50.0
		N_2 (Tau12)	79.6	98.0
		N_2 (Tau3)	83.0	100.0
	p20	N_1	34.0	40.0
		N_2 (Tau12)	69.2	73.1
		N_2 (Tau3)	78.5	96.2
$H_T > 100$	p17	N_1	27.2	29.6
		N_2 (Tau12)	46.7	46.3
		N_2 (Tau3)	36.5	47.3
	p20	N_1	26.0	27.8
		N_2 (Tau12)	46.0	50.1
		N_2 (Tau3)	33.4	44.1

8.3.3 Correlation Between 2 b-tagged Jets

One source of 2-b-tagged-jet events in the TRFed QCD-multijet sample is $g \to b\bar{b}$. Since the gluons have large momentum and no real mass, both b quarks will be boosted in the direction of the original gluons. Therefore, background events with two real b-jets often have a small ΔR between the jets. So, a weighting function is derived in terms of ΔR in
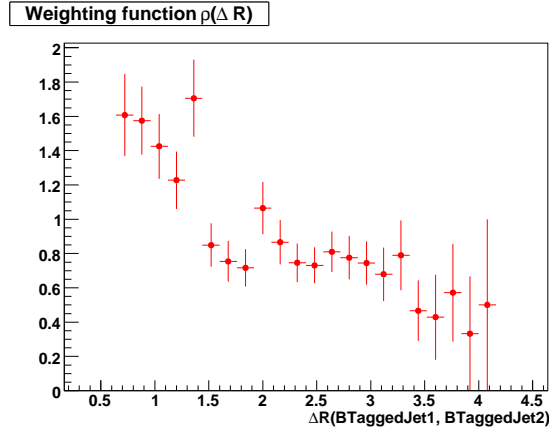
Figure 8.14: Weighting function $\rho(\Delta R)$ showing angular correlations between two b-tagged jets.

order to take this correlation into account.

The weighting function $\rho(\Delta R)$ is defined as

$$\rho(\Delta R) = \frac{\Delta R \text{ distribution in the b-tagged sample}}{\Delta R \text{ distribution in the predicted sample}} \tag{8.11}$$

(see Fig.8.11) Values from this function are used to weight the events with 2 b-tagged jets. Using this approach, reasonable normalization in both 1-b-tagged and 2-b-tagged samples is obtained, as shown in Fig. 8.15.

8.4   Application of TRFs

As indicated above, a data-TRF value for each taggable jet represents the *average* probability (p_i) of that jet being tagged as a b-jet. In order to avoid loss in MC statistics, a similar procedure by means of event permutations presented in Section 5.1.4 is used to apply data-TRFs to the zero-b-tagged sample. The Eq. 5.1 is still valid. The two differences from the standard MC b-tagging are:

– The data-TRFs imply the *average* probabilities of jets being tagged in specific b-tagged and pre-b-tagged samples in this analysis.
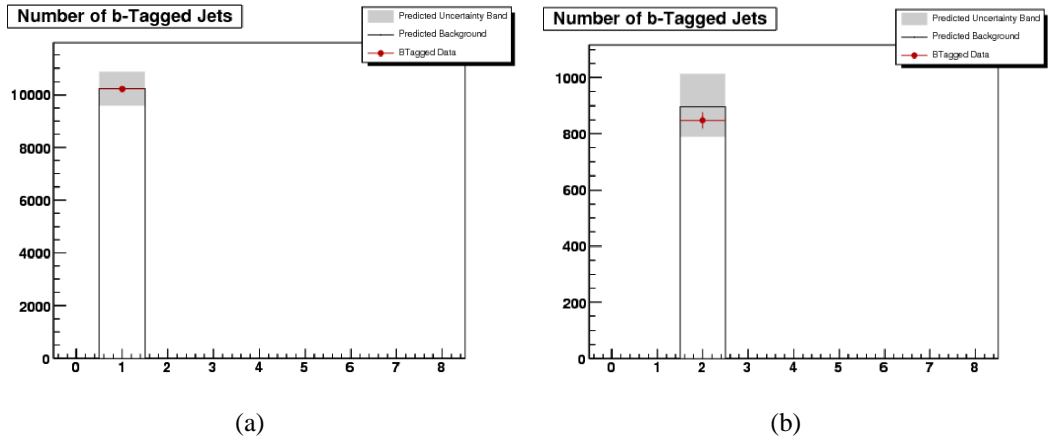
Figure 8.15: Comparison of the number of *b*-tagged jets between TRFed multijets and data for (a) a 1-*b*-tagged-jet sample and (b) a 2-*b*-tagged-jet sample

  – The data-TRFs should be applied to the zero-*b*-tagged sample while the standard TRFs should be applied only to MC samples.

Thus the number of predicted *b*-tagged jets should be

$$\text{predicted } N_{\text{bjet}} = \sum_{\text{Evt}} \left[\sum_{\text{perm}} n_{\text{bjets}} P_k \right] \tag{8.12}$$

where n_{bjets} is the number of taggable jets labeled as *b*-jets per permutation and P_k is given by Eq. 5.1.

The *b*-tagged sample is compared with the data-TRFed predicted sample mentioned above and they are in good agreement (see Appendix B for comparison plots between the *b*-tagged sample and its data TRFed predicted sample in the most sensitive analysis bin: Run IIb, tau type 1+2 and 2 jets one of which is *b*-tagged).

8.5 Removal of Non-QCD-multijets Contamination from Multijet Model

As mentioned above, there is a contamination of non-QCD-multijet events in the background data sample because the 0-*b*-tagged data sample is not a pure QCD-multijet sample.

When data TRFs are applied to the 0-*b*-tagged data sample, some non-QCD-multijet events also have been tagged by data TRFs. So, in order to get a pure QCD-multijet data sample in background, any contamination from, for example, W+jets should be subtracted. Otherwise some events in the model will be double-counted once the MC samples are added to the background model. This correction is realized by subtracting data-TRFed MC samples from the data-TRFed QCD-multijet sample. These MC sources are: W+jets, Z+jets, $t\bar{t}$ and di-boson. The background sample becomes a combination of several data and MC samples represented as

$$\langle \text{Background} \rangle = \langle \text{PureQCD-multijets} \rangle + \langle \text{BTaggedMCs} \rangle \tag{8.13}$$

$$= \langle \text{TRFedQCD-multijets} \rangle - \langle \text{TRFedMCs} \rangle + \langle \text{BTaggedMCs} \rangle. \tag{8.14}$$

However, it is technically impossible to perform a direct subtraction between a data-TRFed QCD-multijet sample and data-TRFed MC samples to get a pure QCD-multijet sample since they belong to different samples. So the subtraction has to be done by reweighting the TRFed QCD-multijet sample by means of a ratio weighting function. The ratio weighting function is defined as

$$\text{ratio} = \frac{\langle \text{PureQCD-multijets} \rangle}{\langle \text{TRFedQCD-multijets} \rangle} \tag{8.15}$$

$$= \frac{\langle \text{TRFedQCD-multijets} \rangle - \langle \text{TRFedMCs} \rangle}{\langle \text{TRFedQCD-multijets} \rangle} \tag{8.16}$$

where the function *ratio* is parameterized in leading jet p_T, tau lepton p_T and $\not{E}_T$. As an example, Fig. 8.16 shows comparison plots of these variables plus W transverse mass M_T^W before and after reweighting on the sample $\langle \text{TRFedQCD-multijets} \rangle$ in the most sensitive channel: τ type 1 and 2, 1 *b*-tagged jet and total 2 jets. Subtraction of $\langle \text{TRFedMC} \rangle$ samples for $\not{E}_T$ from $\langle \text{TRFedQCD-multijets} \rangle$ suppresses the high $\not{E}_T$ region, hence the scaled QCD-multijet sample becomes pure. For leading jet p_T and τ p_T, the subtraction effect appears in the low and intermediate regions that are non-QCD-multijet contamination regions. Also as expected, the effect suppresses the high M_T^W region where there are more non-QCD-multijet events.

Thus the background model becomes

$$\langle \text{Background} \rangle = \text{ratio}(p_T^{\text{leading jet}}, p_T^{\tau}, \not{E}_T) \times \langle \text{TRFedQCD-multijets} \rangle + \langle \text{BTaggedMCs} \rangle \tag{8.17}$$
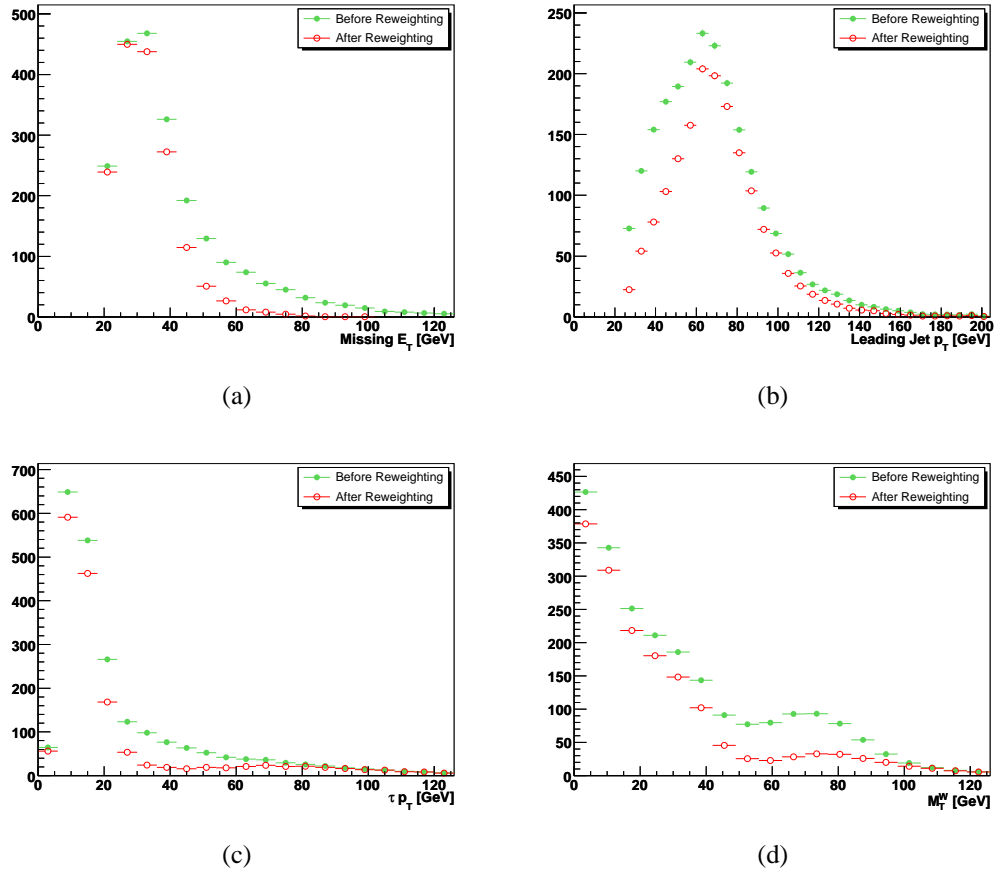
Figure 8.16: Comparison plots of (a) $\not{E}_T$ (b) leading jet p_T (c) τ p_T before and after QCD-multijet reweighting (d) W transverse mass.

Comparison plots of topological variables between data and background based on the model above will be shown below. A systematic error introduced by scaling the TRFed QCD-multijet sample will be discussed in Chapter 9.2.

8.6 QCD-multijet Normalization

Before making a cross section calculation, it is necessary to normalize the pure QCD-multijet sample obtained in last section. In this procedure, the sample of events with BDT output lower than 0.2 is assumed to be enriched in QCD-multijet events. Eq. 8.18 below is

used to calculate QCD-multijet normalization factors:

$$R_{\text{QCD-multijets}} = \frac{\langle \text{\# of data} \rangle - \langle \text{\# of background MC} \rangle}{\langle \text{\# of QCD-multijets} \rangle} \qquad (8.18)$$

where $\langle \text{\# of Sample} \rangle$ ("Sample" is data, background MC or QCD-multijets) is the number of weighted event entries in the BDT output region [0.0, 0.2]. Later the orthogonal BDT output region [0.2, 1.0] is used to determine cross section values. Table 8.2 lists the normalization factors and statistical uncertainties calculated. The table also lists the factors that are determined in the BDT region [0.2, 0.5], indicating that the factors calculated in the two regions [0.0, 0.2] and [0.2, 0.5] are consistent within statistical uncertainties and that the normalization approach is insensitive to the actual BDT region used. The statistical uncertainty of $R_{\text{QCD-multijets}}$ will be used as a systematic uncertainty.

Table 8.2: QCD-multijet normalization factors and their relative statistical uncertainties in different channels. The factors are calculated in two BDT regions [0.0, 0.2] and [0.2, 0.5]. The factors in [0.0, 0.2] are used in the QCD-multijet scaling. Those in [0.2, 0.5] indicate that the factors determined in the two regions are same within uncertainties and the QCD-multijet normalization is insensitive to the actual value of the cut selected.

	Run IIa				Run IIb			
	[0.0, 0.2]		[0.2, 0.5]		[0.0, 0.2]		[0.2, 0.5]	
	SF	RelErr	SF	RelErr	SF	RelErr	SF	RelErr
τ type 1+2 / 1 tag / 2 jets	1.0	3.9%	1.0	9.5%	1.0	2.9%	1.1	8.3%
τ type 1+2 / 1 tag / 3 jets	1.0	3.8%	1.3	10.9%	1.0	2.8%	1.1	8.1%
τ type 1+2 / 2 tag / 2 jets	0.8	14.5%	1.8	77.2%	0.8	13.4%	1.0	69.6%
τ type 1+2 / 2 tag / 3 jets	1.0	13.2%	2.0	71.5%	1.0	10.2%	1.3	51.5%
τ type 3 / 1 tag / 2 jets	1.0	3.8%	1.1	12.1%	1.0	2.5%	1.2	8.2%
τ type 3 / 1 tag / 3 jets	1.0	3.6%	1.1	11.7%	1.0	2.5%	0.9	8.2%
τ type 3 / 2 tag / 2 jets	1.0	13.3%	0.9	50.6%	0.9	13.5%	2.7	91.8%
τ type 3 / 2 tag / 3 jets	1.0	11.2%	0.5	25.5%	0.9	10.2%	0.9	31.6%

8.7 Comparison Plots Between Data and Background

There are around 150 event topological variables available to check the quality of the background model in the analysis. Comparison plots of the 20 top-ranked variables (by BDT discriminating power) between data and background in the most sensitive channel (Run IIb, τ type 1+2, 1 tag and 2 jets) are shown in Section 8.7.1. Meanings of these variables can be found in Table 9.2 in Chapter 9.3.2. Figure 8.17 shows the color scheme used in the comparison plots in this analysis.

DATA
tb→ eν bb
tb→ $\tau\nu$ bb
tqb→ eν bqb
tqb→ $\tau\nu$ bqb
W + light jets
Wcc + jets
Wbb + jets
Zlp
Zcc + jets
Zbb + jets
diboson
$t\bar{t}$ → lep + jets
$t\bar{t}$ → dilepton
$t\bar{t}$ → all jets
QCD−multijets
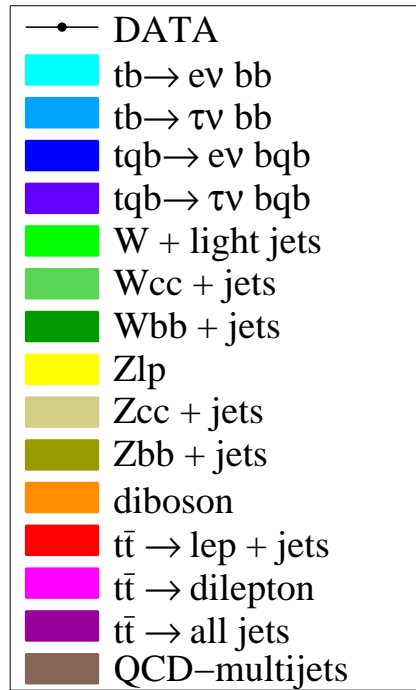
Figure 8.17: Legends of comparison plots.

In summary, these comparison plots are in a good agreement between data and background, which indicates that the multijet tag background model matches the data. The Kolmogorov-Smirnov (KS) test is a numerical tool to understand how well the background sample models the data sample [80]. The KS test is used here to estimate compatibility in shapes of variables between distributions of data and background samples. The closer the KS value is to 1, the better the agreement between data and background is. A cut on

KS values (KS>0.1) is made to select the topological variables used in BDT training later. These KS values are calculated based on histograms with a large number of bins to avoid the effect of histogram binning on the KS calculation.

However, due to the fact that the background is QCD-multijets dominated and the background sources modeled with MC represent only a small fraction of the total background, it is hard to claim that the backgrounds are properly modeled from these tests alone. Therefore an important cross check on our background samples is performed later by measuring the $t\bar{t}$ cross section. If the measured $t\bar{t}$ cross section is consistent with the current measurement by D0, then it will provide strong supporting evidence that the background model is correct. Detailed studies can be found in Section 9.4.

8.7.1 Comparison Plots in the Bin: Run IIb, Type 1+2 Tau, 1 b Tag, 2 Jets
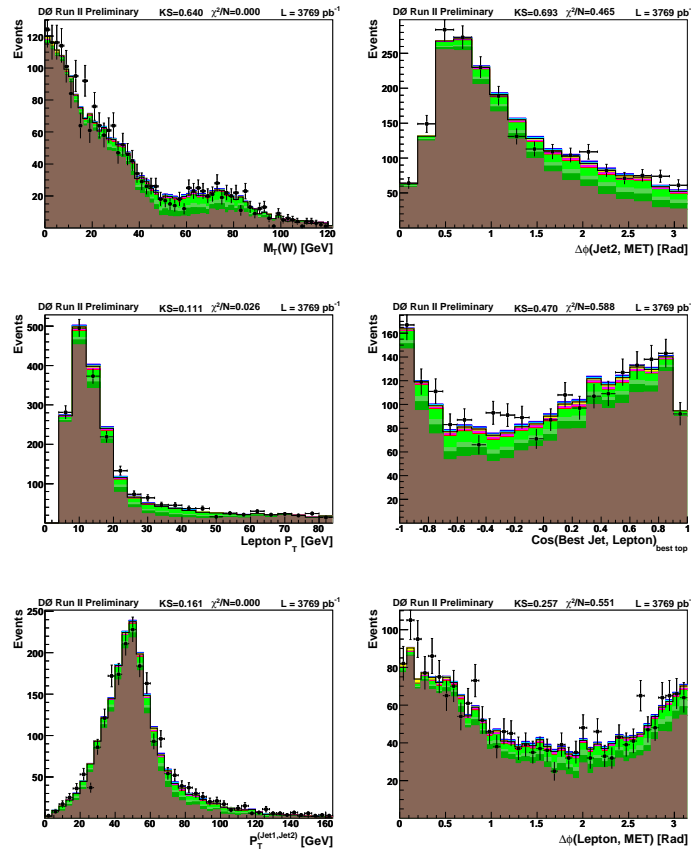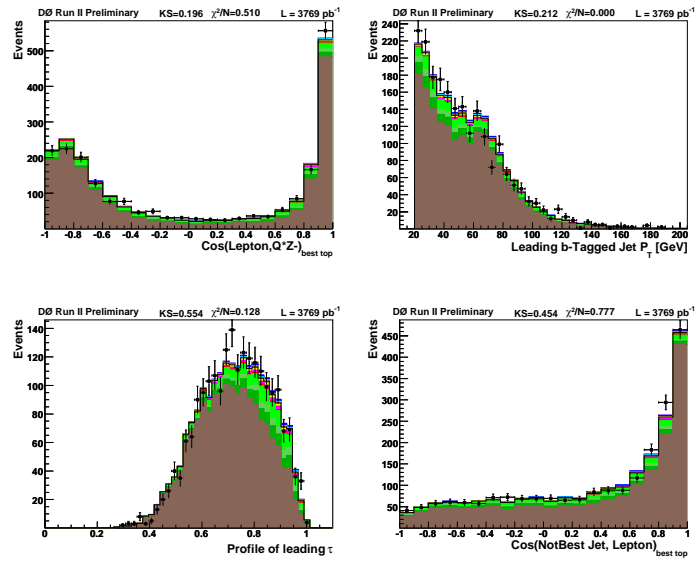
Chapter 9

Cross Section Measurement

In particle physics, the experimental cross section can be calculated by

$$\sigma = \frac{N_{\text{Observed}} - N_{\text{Background}}}{\int \mathcal{L}dt \cdot \varepsilon} \tag{9.1}$$

where N_{Observed} is the number of observed events, $N_{\text{Background}}$ is the number of background events predicted from data measurements or calculated from theory, $\mathcal{L}$ is the instantaneous luminosity and $\int \mathcal{L}dt$ is the integrated luminosity determined by the accelerator, trigger prescale *etc.*, and ε is a product of various acceptances and efficiencies from the analysis.

In the tau+jets channel alone there is not enough signal sensitivity to make a measurement of the single top cross section. In this case, setting a limit in this channel alone can be achieved instead and measurement of the cross section should be in combination with other channels. A multivariate technique can serve as a powerful tool to generate a good variable to discriminate small signals from background. All searches for single top production at D0 use a statistical approach based on Bayes' theorem to set limits or measure cross section values [81]. In the analysis, the same statistical method is employed to derive the limit of single top production in the τ+jets channel and to measure a cross section.

In this chapter, first of all, an overview of the cross section measurement methodology used is given in Section 9.1. An introduction of systematic uncertainties is given in Section 9.2. Section 9.3 describes in detail the multivariate analysis using boosted decision trees. Section 9.4 discusses a cross check on the background model by measuring the $t\bar{t}$ cross section. Section 9.5 and Section 9.6 show expected and measured cross section results. In Section 9.7 the results of the electron/muon+jets channels and the tau+jets channel

are combined.

9.1 Cross Section Measurement Methodology

9.1.1 Bayesian Statistical Approach

The basic idea of the Bayesian statistical approach is to calculate a posterior density function given an observed count and then to derive related information, such as cross section, experimental sensitivity *etc*. The idea can be formulated by the modified Bayes' theorem:

$$P(\text{theory}|\text{data}) \propto P(\text{data}|\text{theory})P(\text{theory}) \tag{9.2}$$

where 'theory' stands for some hypothesis (i.e., a model that helps to understand data) and 'data' is the experimental result. The posterior probability, $P(\text{theory}|\text{data})$, is a subjective probability given the data. $P(\text{data}|\text{theory})$ is the likelihood of observing the data, given the theory. The prior probability of the theory, $P(\text{theory})$, is interpreted as how believable the theory is.

Given a mean event yield d predicted by one model, the probability to observe D events is described by the Poisson distribution:

$$P(\text{data}|\text{theory}) = p(D|d) = \frac{\mathrm{e}^{-d}d^D}{\Gamma(D+1)} \tag{9.3}$$

where Γ is the Gamma function. The mean yield d comprises of the signal and N sources of background:

$$d = a \cdot \sigma + \sum_{i=1}^{N} b_i \tag{9.4}$$

where a is the effective luminosity, a product of the signal acceptance and the integrated luminosity, σ is the signal cross section, and b_i the expected number of events for background source i.

Then the posterior probability density can be computed by

$$P(\text{theory}|\text{data}) = p(\sigma|D) = \iint p(\sigma, \mathbf{a}, \mathbf{b}|D)d\mathbf{a}d\mathbf{b} = \frac{1}{\mathcal{N}} \iint L(D|\sigma, \mathbf{a}, \mathbf{b})\pi(\sigma, \mathbf{a}, \mathbf{b})d\mathbf{a}d\mathbf{b} \tag{9.5}$$

where $\mathbf{a}$ is a vector of the effective luminosities and $\mathbf{b}$ is a vector of the background yields in multi-bins, $\mathcal{N}$ is a normalization factor by $\int_0^{\sigma_{max}} p(\sigma|D)d\sigma = 1$ (σ_{max} to be defined later), $L(D|\sigma, \mathbf{a}, \mathbf{b})$ is the likelihood of measuring D given σ, $\mathbf{a}$ and $\mathbf{b}$, $\pi(\sigma, \mathbf{a}, \mathbf{b})$ is a prior probability density in terms of σ, $\mathbf{a}$ and $\mathbf{b}$.

It is conventional that the signal cross section σ is uncorrelated with $\mathbf{a}$ and $\mathbf{b}$, thus the prior function can be expressed as

$$P(\text{theory}) = \pi(\sigma, \mathbf{a}, \mathbf{b}) = \pi(\mathbf{a}, \mathbf{b}) \cdot \pi(\sigma) \tag{9.6}$$

The prior density of the cross section $\pi(\sigma)$ is assumed to be flat in terms of σ:

$$\pi(\sigma) = \begin{cases} \frac{1}{\sigma_{max}}, & 0 < \sigma < \sigma_{max} \\ 0, & \text{otherwise} \end{cases} \tag{9.7}$$

where σ_{max} is chosen above which any preferred value for the signal cross section is ignored. In this analysis, $\sigma_{max} = 30$ pb. Eq. 9.5 therefore becomes

$$p(\sigma|D) = \frac{1}{\mathcal{N} \cdot \sigma_{max}} \iint L(D|\sigma, \mathbf{a}, \mathbf{b})\pi(\mathbf{a}, \mathbf{b})d\mathbf{a}d\mathbf{b}. \tag{9.8}$$

The prior density $\pi(\mathbf{a}, \mathbf{b})$ encodes knowledge of the effective signal luminosities and background yields. The integration in Eq. 9.8 is performed numerically by means of MC importance sampling. Randomly sampling a large number of points $(\mathbf{a}_k, \mathbf{b}_k)$ from the prior density $\pi(\mathbf{a}, \mathbf{b})$, Eq. 9.8 becomes

$$p(\sigma|D) \propto \frac{1}{K}\sum_{k=1}^{K} L(\mathbf{D}|\sigma, \mathbf{a}_k, \mathbf{b}_k). \tag{9.9}$$

Effects of systematic uncertainties are taken into account during the generation of the samples by direct sampling as follows. For a systematic *isys* of one source, given shifted systematic distributions $y_{isys}^{\pm}$ and nominal distribution y, a yield shift Δy_{isys} of this systematic for this source can be obtained. It is obtained by sampling a shift $g(0,1)_{isys}$ from a Gaussian distribution with mean 0 and width 1:

$$\Delta y_{isys} = s^{\pm} \times g(0,1) \times |y_{isys}^{\pm} - y| \tag{9.10}$$

where $s^{\pm}$ is a scale factor, that is normally 1. If only shape uncertainty is considered, $s^{\pm} = \frac{\Sigma y}{\Sigma y_{isys}^{\pm}}$ where the sum goes over all of the bins of the systematic histograms $y_{isys}^{\pm}$ for

this source. If $g(0,1) > 0$, (+) is taken, otherwise (−) is taken. Over all of the systematic uncertainties, the overall shifted yield can be formulated as:

$$y' = y + \sum_{isys} \Delta y_{isys}. \tag{9.11}$$

An upper limit σ_{CL} at confidence level (CL) can be obtained by solving:

$$\int_0^{\sigma_{CL}} p(\sigma|\mathbf{D}) = CL. \tag{9.12}$$

A cross section measurement is the peak location of the posterior density $p(\sigma|\mathbf{D})$.

9.1.2 Sensitivity Estimation

Three parameters to estimate experimental sensitivity are discussed below. They are the ratio of peak to width, the Bayes factor significance and the Bayes ratio significance. The larger those parameters, the higher the experimental sensitivity.

Ratio of peak to lower half-width

Half of the interval around the peak location covering 68% of the whole area is considered as an estimate of the width of the posterior density distribution. The ratio of the peak location to the lower half-width is treated as one of parameters to estimate experimental sensitivity. This is a concept similar to the number of standard deviations.

Bayes Factor Significance

Given two hypotheses H_0 and H_1, where the null hypothesis H_0 is the background-only model and the alternative hypothesis H_1 is the signal+background model, the Bayes factor B_{10} is defined as

$$B_{10} = \frac{p(\mathbf{D}|H_1)}{p(\mathbf{D}|H_0)} \tag{9.13}$$

Its significance is calculated by $\sqrt{2\log B_{10}}$. This quantity is only valid for the expected case since it is based on a specific signal model.

Bayes Ratio Significance

A Bayes ratio can be defined as

$$B_r = \frac{p(\hat{\sigma}|\mathbf{D})}{p(\sigma=0|\mathbf{D})} \tag{9.14}$$

where $\hat{\sigma}$ is the peak location of a posterior density function. Its significance is $\sqrt{2\log B_r}$. The optimal analysis is assumed to be the one with the largest expected Bayes ratio significance.

9.1.3 Pseudo-data Ensembles and Linearity Test

In order to validate the approach outlined above, a linearity test that checks potential measurement biases is performed using pseudo-data ensembles. Pseudo-data sets are generated from a pool of weighted signal and background events. Each pseudo-data set is then analyzed in the same way as real data. Different signal cross sections can be used as inputs to generate different ensembles and the linearity of the approach can be checked.

Five ensembles of pseudo-datasets were generated from background and signal model events. Each ensemble is comprised of around 2000 pseudo-experiments with all systematic uncertainties considered. In this procedure, the five inputsignal cross sections are 2.0, 3.46, 6.0, 8.0 and 10.0 pb. The output cross section values are measured in the same way as for real data. Linearity implies consistency between input and output cross section values. Graphs in Fig. 9.2 show the output cross section distributions of all ensembles separately and Fig. 9.1 shows a good linear fit through their peak locations. The test indicates there is not a linearity problem in this approach.

9.2 Systematic Uncertainties

Systematic uncertainties are taken into account in the cross section calculations in two ways: as a normalization uncertainty on background samples and as a shape uncertainty on the distributions of the background samples and expected signal samples.

Table 9.1 summarizes all sources of systematics uncertainties and their relative uncertainties. Each of the sources is described below. Detailed tables of uncertainties for each individual analysis channel are listed in Appendix F.
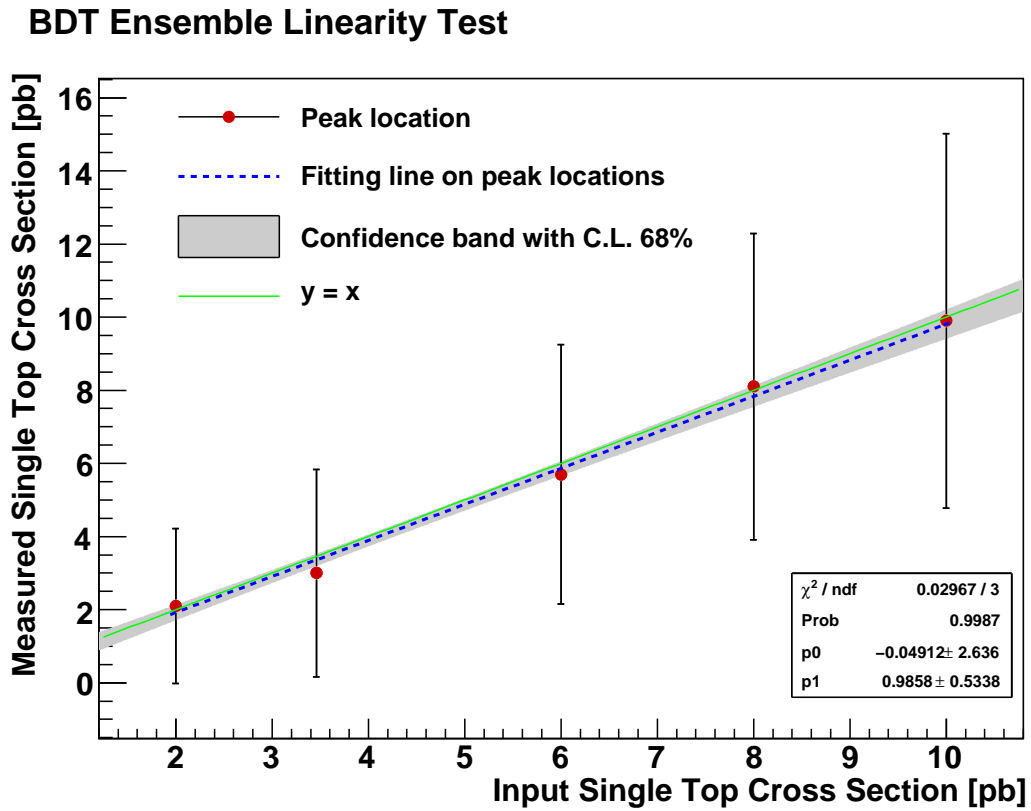
Figure 9.1: Linear fit through the peak locations of output cross section distributions (Fig. 9.2) given different input single top cross section values. The linear fitting covers the range from 2.0 to 10.0 pb.
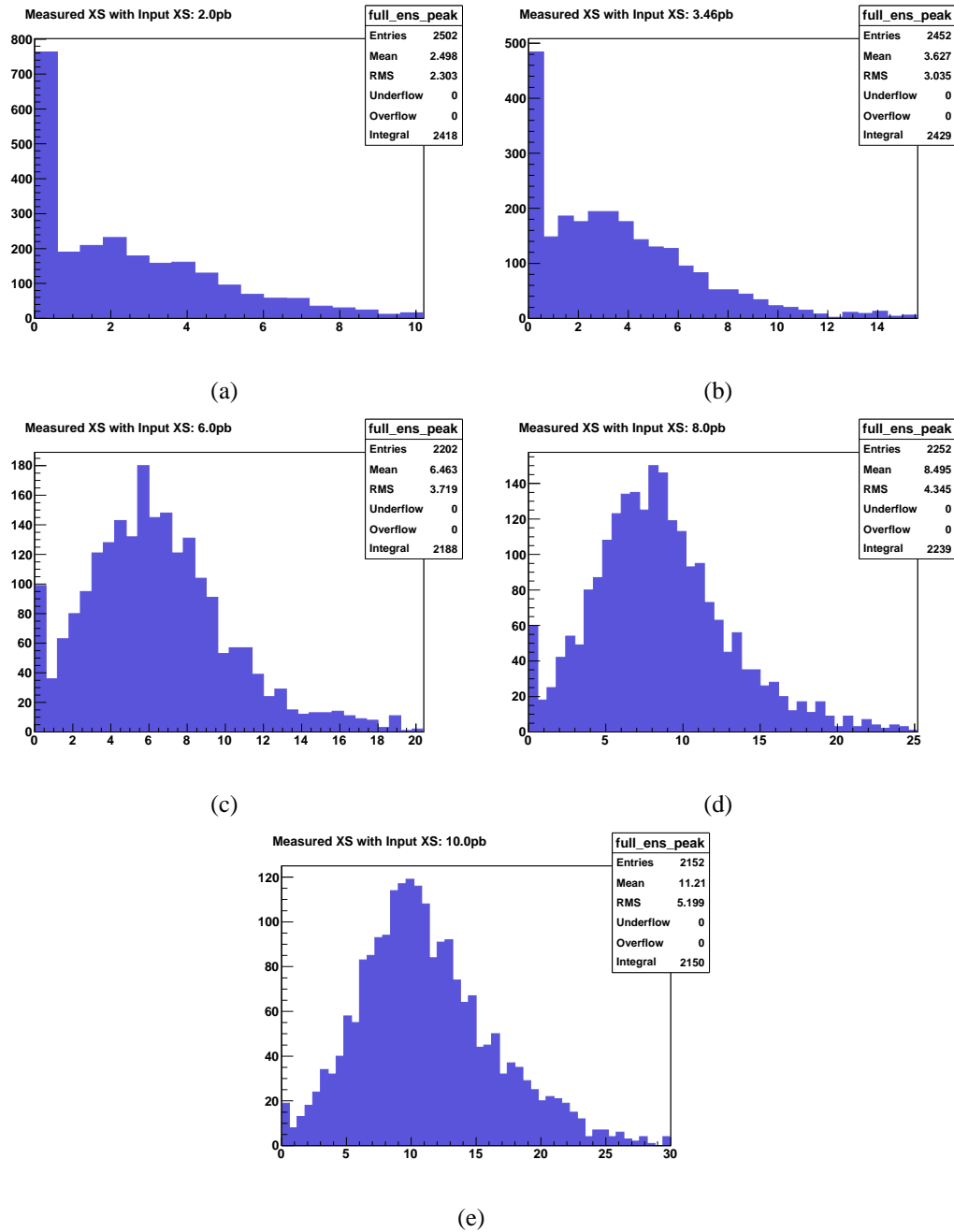
Figure 9.2: Measured single top cross section in ensembles with different amounts of input single top: (a) 2.0 pb (b) 3.46 pb (c) 6.0 pb (d) 8.0 pb (e) 10.0 pb.

– **Integrated luminosity**

The 6.1% uncertainty on the luminosity estimate that comes from the uncertainties on the measured inelastic $p\bar{p}$ cross section affects the signal, $t\bar{t}$, Z+jets, diboson and on $p\bar{p}$ yields.

– **Theoretical cross sections**

The uncertainty on the cross sections for signal, Z+jets, diboson and $t\bar{t}$ includes the theoretical uncertainty for all and the uncertainty from the top quark mass for $t\bar{t}$. The values used are 11.2% (tb), 7.4% (tqb), 3.6% (Z+jets), 5.8% (diboson) and 12.7% ($t\bar{t}$).

– **Trigger efficiency**

An uncertainty of 5.5% to the trigger efficiency is assigned. The uncertainty arises as 5% from the multijet trigger parameterization (not taking into account taus) and a 2% (estimated) uncertainty added in quadrature for the difference in tau and jet turn-on curves.

– **Instantaneous luminosity reweighting**

The instantaneous luminosity distributions of all MC samples are reweighted to match Run IIa or Run IIb data distributions as appropriate. The initial distributions are from the minimum bias data overlaid on the MC events to simulate the underlying events, and are generally at too low values for later data-taking conditions. The uncertainty on this reweighting is 1.0%.

– **Primary vertex modeling and selection**

The distribution of the primary vertices along the beamline in MC is reweighted to match that in data. The uncertainty on this reweighting is 0.05% (negligible). The uncertainty on the difference in primary vertex selection efficiency between data and MC is 1.4%.

– **Tau lepton reconstruction and identification efficiency**

An 11.0% uncertainty is assigned as the uncertainty on the tau lepton reconstruction and identification efficiency. A detailed description of the determination can be found

in Section 7.5. It is applied to all MC samples except W+jets[1].

– **Tau energy scale**

A tau energy scale uncertainty occurs in applying tau energy correction by taking associated tracks into account for low energy tau candidates. A value of (1.0-1.5)% is applied to all MC samples except W+jets.

– **Jet fragmentation**

A systematic uncertainty covers the differences between the jet fragmentation models of PYTHIA (used for the ALPGEN samples in the analysis) and HERWIG. The resulting uncertainty of 5% is obtained by comparing the $t\bar{t}$ acceptance of the two models and is applied to all MC samples.

– **b-jet fragmentation**

The uncertainty on the modeling of b jet fragmentation is 2.0%. The value is determined by the difference between fragmentation parameterizations measured by SLD vs. LEP data.

– **Initial-State Radiation (ISR) and Final-State radiation (FSR)**

The values are determined using $t\bar{t}$ samples by comparing results using PYTHIA with ISR and FSR parameters varied up and down.

– **Jet reconstruction and identification**

The efficiency with which jets are reconstructed and identified has an uncertainty of 1%. This is estimated by taking the difference of the jet η and jet multiplicity between the data and MC.

– **Jet energy scale and jet energy resolution**

A flat uncertainty of 4%-14% is assigned to the jet energy scale uncertainty since in some regions statistics and smoothness are lacking to determine the shape of JES samples. For the same reason, a flat uncertainty of 4% is also set for the JES in all MC samples.

---

[1]W+jets samples are applied scale factors so that they are normalized to data, which serves as a constraint. Thus with the constraint, no other systematics are considered.

- **W+jets and Z+jets heavy-flavor scale factor**

  This is treated in the same way as in the electron/muon studies [82]. An uncertainty of 13.7% is assigned to the heavy-flavor scale factors for $Wb\bar{b}$, $Wc\bar{c}$, $Zb\bar{b}$ and $Zc\bar{c}$.

- ALPGEN **reweighting on** W**+jets sample**

  Due to a known issue about W+jets MC samples that certain variables of the leading log ALPGEN Monte Carlo disagree with data, the discrepancies of the W+jets events have to be fixed. Since W+jets samples are not the dominant background source, the reweighting functions derived in electron/muon+jets channels are applied and a shape uncertainty from the reweighting functions is included.

- **Sample statistics**

  The MC and data samples that are used to estimate the signal and background shapes are limited in size. The background sample statistics are taken into account for each sample in each bin of the final discriminant distribution.

- **Non-QCD-multijet contamination removal**

  Reweighting functions are parameterized to remove non-QCD-multijet contamination discussed in Section 8.5. The systematic uncertainty arises originally from the measurement of the *ratio* functions (Eq.8.15) and is estimated by uncertainties on the reweighting functions (bin errors of the histograms). Systematic samples are generated by shifting the reweighting function up and down. It affects the normalization and shape of the QCD-multijet sample.

- **Monte Carlo tag rate functions and taggability**

  The uncertainty associated with the b-tagging tag-rate functions is evaluated by adding the taggability and the tag rate components of the uncertainty in quadrature. The TRF uncertainties originate from several sources: statistical errors of MC event sets; the assumed fraction of heavy flavor in the multijets MC events for the mistag rate determination; and the parameterizations. For some channels in which TRF systematics cannot be well determined, we assigned a flat uncertainty of 4%-14%.

- **QCD-multijet tag rate function shape-changing**

  As mentioned in Section 8, an important part of the background model is based on

the TRF measurement. QCD-multijet TRF uncertainties are estimated by the TRF measurement. TRF fitting curves are shifted by ± 1 σ to generate new TRF-shifted samples that later will be applied to the BDT outputs. Since we use the QCD-multijet normalization as a constraint, this systematic only affects shape.

– **QCD-multijet normalization**
An explicit QCD-multijet normalization has to be performed before making any cross section calculation. Normalization factors are calculated by the equation:

$$N_{\text{multijets}} = \frac{\langle \text{\# of data}\rangle - \langle \text{\# of bkg MC}\rangle}{\langle \text{\# of multijets}\rangle} \tag{9.15}$$

where $\langle \text{\# of Sample}\rangle$ ("Sample" is data, bkg MC or QCD-multijets) is the weighted number of events in the BDT probability region 0.0 - 0.2.The statistical uncertainty of $N_{\text{multijets}}$ is used as a systematic uncertainty. The factors are shown in Table 8.2.

– *W*+**jets to data**
In the background modeling, a pre-*b*-tagged QCD-multijet sample is not available to do a normalization of the sum of *W*+jets and QCD-multijet backgrounds to pretagged data, so the weighted average normalization scale factors of *W*+jets in electron/muon channels are used. The difference of the scale factors between the electron and muon channels is assigned as the normalization uncertainty.

– **Signal contamination removal**
Signal contamination removal will be discussed in Section 9.3.4. Uncertainty on this is estimated by bin content errors of removal reweighting functions.

The relative values of the systematic uncertainties above are not the only factors that reflect how important they are in experimental sensitivity. The contribution from each systematic effect to the expected cross section uncertainty is estimated by considering one systematic source at a time. Among all the systematic uncertainties, the five most important ones are: the *W*+jets scale factors, the tau lepton reconstruction and identification efficiency, the **W**+jets heavy-flavor scale factor, the tag rate function, and the integrated luminosity.

Table 9.1: A summary of the relative systematic uncertainties for each of the correction factors or normalizations. The uncertainty shown is the error on the correction or the efficiency, before it has been applied to the MC or data samples.

Relative Systematic Uncertainties	
Components for Normalization	
Integrated luminosity	6.1%
$t\bar{t}$ cross section	12.7%
Z+jets cross section	3.6%
Diboson cross sections	5.8%
Trigger efficiency	5.5%
Instantaneous luminosity reweighting	1.0%
Primary vertex modeling and selection	1.4%
Tau reconstruction and identification efficiency	11.0%
Tau energy scale	(1.0–1.5)%
Jet fragmentation	5.0%
b-jet fragmentation	2.0%
ISR/FSR	(0.6–8.0)%
Jet reconstruction and identification	1.0%
Jet energy scale and resolution	(4.0-14.0)%
W+jets and Z+jets heavy-flavor fraction	13.7%
W+jets to data	(7.0–15.0)%
Multijet normalization	(3.0–7.0)%
MC statistics	(0.5–16)%
Components for Shape	
Non-multijet contamination removal	—
Alpgen Reweighting on W+jets sample	—
Tag-rate functions	—
Signal Contamination Removal	—
(not shape for Z+jets, $W\ell p$ or dibosons)	

9.3 BDT Analysis and Sample Preparation

9.3.1 BDT Parameter Selection

The BDT algorithm used in the analysis has been discussed in Section 6. In signal discrimination, the number of boosting cycles (number of booted trees) is fixed to 50, the Ada boost parameter is set to 0.20, the minimal leaf size is limited to 100 and the criterion for impurity measure is the Gini index.

9.3.2 Discriminating Variables

44-80 variables are selected from $\sim$150 topological variables to train decision trees. The selection criteria are (1) the variable has to be well modeled. In the analysis, the selected discriminant variables have KS test values larger than 0.1 between data and the background model. (2) most of the tau properties are excluded since they are used for the new τ ID optimization.

As mentioned before, analysis samples are split into different analysis channels according to data reconstruction version (Run IIa and Run IIb), tau type (1+2 and 3), number of tags (1 tag and 2 tags) and jet multiplicity (2 and 3 jets), thus there are 16 channels in total. Table 9.2 shows a summary list collecting the variables appearing in each channel. However, not all variables listed in Table 9.2 are included in the BDT training of every channel. Thus different channels have different training variable lists (each list has 50–80 variables). Tables G.1 to G.16 show the tables listing the top 20 variables ranked in discriminating power and their KS values in each channel. The comparison plots of these top topological variables between data and background are shown in Section 8 and Appendix D. Then all the variables are classified in four categories:

Object kinematics and properties

Variables describing object kinematics are transverse momentum (p_T) and pseudo-rapidity (η) of individual objects (jet, tau and$\not{E}_T$) per event. Jet objects are sorted decreasingly in jet p_T. In variable names, the highest-p_T jet is called "jet1", the second highest-p_T jet is called "jet2" *etc*. The leading b-tagged jet is called "btaggedjet1". When a jet in an event is combined with a reconstructed W boson [2] and results in an

---

[2]A W boson on the transverse plane is reconstructed by a visible tau and$\not{E}_T$. By constraining the transverse

invariant mass closest to 170 GeV, such a jet is called the best jet of the event and the reconstructed top is called the best top. Only the two leading jets in p_T in the event are considered. The light quark jet in this case is whichever of the two leading jets was not chosen as the best jet. Some tau properties: isolation, profile, RMS, EMF (definitions of the tau properties can be found in Section 7.2), and tau identification BDT output are also included in this class. $Q(\tau) \times \eta$ is a product of tau charge and tau η taking advantage of CP symmetry in the *t*-channel production.

Event kinematics

These variables are calculated from all objects or a subset of objects in an event. **Aplanarity** defines how the reconstructed W boson and all the jets in the event are placed with respect to a plane. If the aplanarity is zero, it indicates that the system of W and all jets spans only one plane. The single top quark signal events tend to be more aplanar than background events [83]. **Sphericity** defines how spherically the W and jets are situated in the event. The objects in signal events tend to be spherical and the events have a higher Sphericity value [83]. $M(\text{objects})$ and $M_T(\text{objects})$ represent the mass and transverse mass [5] of a subset of objects. $H(\text{objects})$ and $H_T(\text{objects})$ are defined as:

$$H(\text{objects}) = \sum_{\text{objects}} E \tag{9.16}$$

$$H_T(\text{objects}) = \sum_{\text{objects}} p_T. \tag{9.17}$$

For example, $H_T(\text{alljets-bestjet})$ means H_T of all jets excluding the best jet. **Centrality** is defined as $H_T(\text{alljets})/H(\text{alljets})$. $\sqrt{\hat{s}}$ is the invariant mass of all basic objects (tau, $\not{E}_T$ and jets) in the event. $\sum p_{T\text{DCAcut}}^{\text{trk}}$ and $\sum p_T^{\text{trk}}$ are transverse momentum of vectorial sum of all tracks with and without a cut on Distance of Closest Approach (DCA) to the primary vertex in one event. The cut requires: $\frac{\text{DCA}}{\sigma_{\text{DCA}}} < 3$.

Top quark reconstruction

The four-vectors of the reconstructed W boson using the neutrino p_Z constraint and one of the jets in the event are added to reconstruct a top quark. The leading *b*-tagged

---

mass of the system to the W mass, two neutrino p_Z solutions can be calculated and the one with the smaller absolute value is used.

jet and the best jet defined above correspond to 2 different top quarks, hence 2 top quark mass variables: $M(W, \text{tag1})$ and $M(W, \text{best1})$ respectively.

Angular correlation

Angular correlation variables are either ΔR or $\Delta \phi$ (and cosine of $\Delta \phi$) between different objects in the different reference frames [84, 85]. Frames include CM frame and best-top frame, which are labelled by the subscripts in Table 9.2.

9.3.3 Sample Preparation for BDT Training

A conventional approach is used to split the final analysis samples into three equal-size subsamples: one third used for training the BDT, one third for testing and estimating the trained BDT performance, one third for measurement of the cross section.

9.3.4 Signal Contamination Removal (SCR)

Since the *b*-tagging efficiency of single top events is not 100%, there are some signal events left in the 0-*b*-tagged data sample from which the multijets sample is derived. These should be removed to avoid underestimating the final cross section. One way is just like the removal of non-QCD-multijet background MC sources discussed in Section 8, i.e., to apply data TRF on the single top MC sample to get a data-TRFed single top sample and subtract the latter from the multijet sample. However, this approach is not realistic for two reasons:

– the amount of contamination in the multijet sample is unknown since the cross section of single top production is unknown.

– the contamination is too small to be reflected in a non-QCD-multijet reweighting function.

Table 9.2: All discrimination variables used in the BDT analyses in 4 categories: object kinematics, event kinematics, top quark reconstruction and angular variables.

Object kinematics

$p_T(\text{best1})$	$p_T^{\text{notbest}}(\text{jet1})$	$p_T^{\text{notbest}}(\text{jet2})$
$\eta^{\text{detector}}(\text{jet1})$	$\eta(\text{jet1})$	$p_T(\text{jet1})$
$\eta^{\text{detector}}(\text{jet2})$	$\eta(\text{jet2})$	$p_T(\text{jet2})$
$\eta(\text{jet3})$	$p_T(\text{jet3})$	$\eta(\text{jet4})$
$p_T(\text{jet4})$	$p_T(\text{btaggedjet1})$	$p_T(\text{untaggedjet1})$
$p_T(\text{untaggedjet2})$	$\eta^{\text{detector}}(\tau)$	$\eta(\tau)$
$p_T(\tau)$	$\text{BDT}(\tau)$	$\text{Iso}(\tau)$
$\text{RMS}(\tau)$	$\text{Prof}(\tau)$	$\text{EMF}(\tau)$
$\not{E}_T$	$z_{\text{primary vertex}}$	$Q(\tau) \times \eta$

Event kinematics

Aplanarity(W,alljets)	Centrality(alljets)	Sphericity(W,alljets)
$\sqrt{\hat{s}}$	$H_T(\text{alljets}+\tau+\not{E}_T)$	$H_T(\text{alljets-btaggedjet})$
$H_T(\text{alljets-bestjet})$	$H_T(\text{alljets})$	$H_T(\text{jet1+jet2}+\tau+\not{E}_T)$
$H_T(\text{jet1+jet2})$	$H_T(\tau+\not{E}_T)$	$H(\text{alljets-btaggedjet})$
$H(\text{alljets-bestjet})$	$H(\text{alljets})$	$H(\text{jet1+jet2})$
$M(\text{alljets-btaggedjet})$	$M(\text{alljets-bestjet})$	$M(\text{alljets})$
$M(\text{jet1+jet2}+W)$	$M(\text{jet1+jet2})$	$p_T(W)$
$M_T(W)$	$\sum p_T^{\text{trk}}{}_{\text{DCAcut}}$	$\sum p_T^{\text{trk}}$
$p_T(\text{alljets-btaggedjet})$	$p_T(\text{alljets-bestjet})$	$p_T(\text{jet1+jet2})$
$M_T(\text{jet1,jet2})$		

Top quark reconstruction

$M(W,\text{tag1})$ ("*b*-tagged" top mass)		$M(W,\text{best1})$ ("best" top mass)

Angular correlations

$\cos\phi(\text{tag1,alljets})_{\text{alljets}}$	$\cos\phi(\text{tag1},\tau)_{\text{btaggedtop}}$	$\cos\phi(\text{tag1},\tau)_{\text{lab}}$
$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	$\cos\phi(\text{best1},\tau)_{\text{lab}}$	$\cos\phi(\text{best1,notbest})_{\text{besttop}}$
$\cos\phi(\text{jet1,alljets})_{\text{alljets}}$	$\cos\phi(\text{jet1},\tau)_{\text{btaggedtop}}$	$\cos\phi(\text{jet1},\tau)_{\text{lab}}$
$\cos\phi(\text{jet2,alljets})_{\text{alljets}}$	$\cos\phi(\text{jet2},\tau)_{\text{btaggedtop}}$	$\cos\phi(\text{jet2},\tau)_{\text{lab}}$
$\cos\phi(\tau_{\text{btaggedtop}},\text{btaggedtop}_{\text{CMFrame}})$	$\cos\phi(\tau_{\text{besttop}},\text{besttop}_{\text{CMFrame}})$	$\cos\phi(\tau,Q \times z)$
$\cos\phi(\text{notbest,alljets})_{\text{alljets}}$	$\cos\phi(\text{notbest},\tau)_{\text{besttop}}$	$\cos\phi(\text{untaggedjet1,alljets})_{\text{alljets}}$
$\cos\phi(\text{untaggedjet1},\tau)_{\text{btaggedtop}}$	$\Delta R(\text{jet1,jet2})$	$\Delta R(\tau,\text{jet1})$
$\Delta R(\tau,\text{jet2})$	$\Delta R^{min}(\tau,\text{jets})$	$\Delta R^{min}(\text{alljets})$
$\Delta\phi(\text{jet1,jet2})$	$\Delta\phi(\text{jet1}\not{E}_T)$	$\Delta\phi(\text{jet2}\not{E}_T)$
$\Delta\phi(\tau,\text{jet1})$	$\Delta\phi(\tau,\text{jet2})$	$\Delta\phi(\tau\not{E}_T)$

So another approach is employed. The equation below is an expression of the relationship between these samples:

$$\langle \text{Data} \rangle = \underbrace{\langle \text{Bkg} \rangle - \langle \text{ST Contam} \rangle} + \langle \text{ST} \rangle \tag{9.18}$$

$$= \langle \text{Bkg} \rangle + \underbrace{\langle \text{ST} \rangle - \langle \text{ST Contam} \rangle} \tag{9.19}$$

$$= \langle \text{Bkg} \rangle + \frac{\langle \text{ST} \rangle - \langle \text{ST Contam} \rangle}{\langle \text{ST} \rangle} \cdot \langle \text{ST} \rangle \tag{9.20}$$

$$= \langle \text{Bkg} \rangle + R(y) \cdot \langle \text{ST} \rangle \tag{9.21}$$

where $\langle \text{Data} \rangle$ is the b-tagged sample, $\langle \text{Bkg} \rangle$ is the background sample in Eq. 8.13 introduced in the section 8, $\langle \text{ST} \rangle$ and $\langle \text{ST Contam} \rangle$ are the b-tagged single top quark sample and its contamination in the background sample, $R(y) = \frac{\langle \text{ST} \rangle - \langle \text{ST Contam} \rangle}{\langle \text{ST} \rangle}$ is a weighting function used to estimate the relative correction in terms of a BDT variable y. The TRFed 0-b-tagged signal MC sample is used to estimate $\langle \text{ST Contam} \rangle$. Figure 9.3 shows an example in the most sensitive channel (Run IIb, τ type 1+2, 1 tag and 2 jets) of the derivation of the signal contamination removal (SCR) correction. Bin content errors of the weighting functions are treated as one of the systematic uncertainties, called a "SCR" uncertainty.Its relative uncertainty on the final cross section is ∼1%.
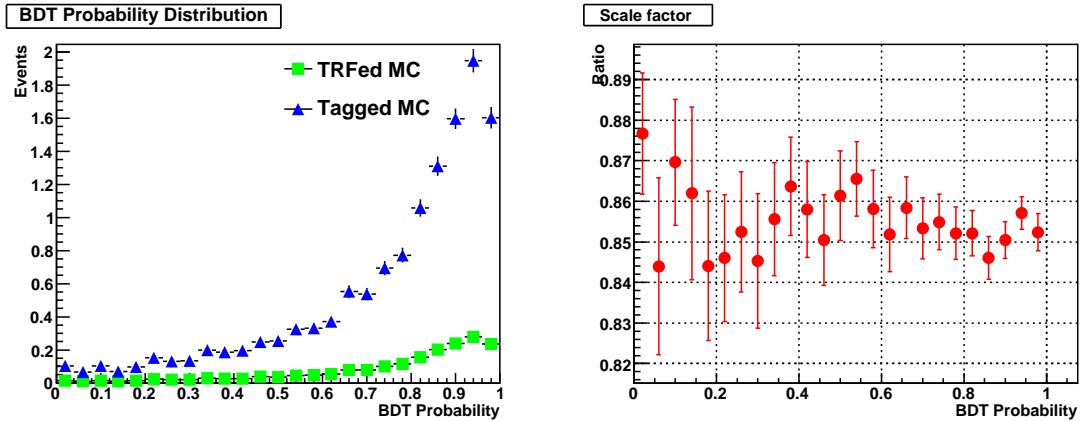
Figure 9.3: Left: BDT output distributions of the sample $\langle \text{ST} \rangle$ (labelled as Tagged MC) and the sample $\langle \text{ST Contam} \rangle$ (labelled as TRFed MC) for s-channel of single top. Right: Illustration of signal contamination removal (SCR) correction. It shows the weighting function $R(y) = \frac{\langle \text{ST} \rangle - \langle \text{ST Contam} \rangle}{\langle \text{ST} \rangle}$ in terms of BDT output.

9.4 Cross Check by Measurement of $t\bar{t}$ Cross Section

9.4.1 Motivation

As seen in all the background agreement plots, the background model is dominated by QCD-multijet events. It is useful to perform a "benchmark" measurement which is independent of single top quarks but shares the same non-QCD-multijets backgrounds. In order to validate such a background model, a cross check is made by measuring the $t\bar{t}$ cross section. The $t\bar{t}$ sample is one of the components in the background model. If the background is properly modeled, then the measured $t\bar{t}$ cross section should be consistent with theoretical expectations within uncertainties. In this cross-check, SM-predicted single top samples are treated as one of our background components while the $t\bar{t}$ sample is treated as signal. The following points make the cross check practical:

- The single top cross section predicted bythe SM is smaller by a factor of $\sim$2 than that of $t\bar{t}$

- The $t\bar{t}$ cross section is well known.

- $t\bar{t}$ samples have different BDT outputs from single top samples, thus their signal regions are located in different places, which means that the single top quark cross section is not so important for this check.

9.4.2 Results of Cross Check

All training/testing/measurement samples in this $t\bar{t}$ cross check are the same as the ones used in the single top cross section measurement which has been mentioned above. Since the $t\bar{t}$ cross section is measured in the check, BDT are trained on $t\bar{t}$ samples as signal and other MC sources plus the QCD-multijet sample as background using the same BDT training parameters and variable lists shown in Table 9.2. The QCD-multijet sample is normalized using the approach discussed in Section 8.6. Then those BDT are applied to estimate BDT probability values of measurement samples event by event. The same statistical software as is used for the single top quark analysis is used to calculate the $t\bar{t}$ cross section in different channels and different combinations of these channels. The all-16-channel-combined value, $10.0^{+2.3}_{-1.6}$ pb, is consistent with the results in most channels.

More importantly, it is also consistent with the $t\bar{t}$ Standard-Model theoretical cross section, $7.91^{+0.61}_{-0.56}$ pb[3], and a recent D0 experimental result, $8.18^{+0.98}_{-0.87}$ pb [86], within 1.5 σ. The consistency provides strong support for the background model.

Table 9.3: Observed $t\bar{t}$ cross section measurements with all systematic uncertainties taken into account, for many combinations of analysis channels.

	Observed Cross Section Measurements						
	1,2tags+2,3jets		tau12,3+2,3jets		tau12,3+1,2tags		all
	Tau12	Tau3	1 tag	2 tags	2 jets	3 jets	channels
Run IIa	$7.9^{+2.7}_{-2.2}$	$12.5^{+5.1}_{-4.1}$	$9.0^{+3.2}_{-2.6}$	$8.9^{+3.5}_{-2.7}$	$6.7^{+4.8}_{-4.0}$	$9.8^{+2.9}_{-2.4}$	$8.8^{+2.6}_{-2.1}$
Run IIb	$10.2^{+2.6}_{-2.0}$	$13.1^{+4.1}_{-3.2}$	$9.6^{+2.6}_{-2.0}$	$16.4^{+7.2}_{-4.5}$	$9.9^{+3.8}_{-3.0}$	$11.5^{+2.8}_{-2.2}$	$11.2^{+2.4}_{-2.1}$
Run IIa+b	$9.0^{+2.1}_{-1.6}$	$12.8^{+3.4}_{-2.8}$	$9.1^{+2.2}_{-1.7}$	$12.2^{+3.2}_{-2.5}$	$8.6^{+3.0}_{-2.5}$	$10.9^{+2.3}_{-1.9}$	$10.0^{+2.3}_{-1.6}$

9.5 Expected Results

In this section results of the expected cross section calculations and corresponding sensitivity estimates are presented. The expectation calculation here means given the Standard-Model cross section of 3.46 pb [18], what cross section of single top is expected to be measured and what is the expected uncertainty. The inputs to the Bayesian software are the decision tree discriminating histograms from Figs. H.1–H.8 in Appendix H. However, real data have to be replaced with background plus the Standard-Model amount of single top in order to perform such an expectation calculation. Table 9.4 shows the expected upper limits at 95% confidence level (CL). Table 9.5 shows the expected cross section measurements with all systematic uncertainties taken into account. Most of the expected cross section values are consistent with the Standard Model. However, some values present a large deviation but with a very large systematic uncertainty. The more the channels are combined, the smaller the uncertainty is on the measurement. Table 9.6 shows the ratio of the posterior peak position over the lower half-width. In addition, the Bayes factor significance and the Bayes ratio significance are shown in Tables 9.7 and 9.8. These two quantities give consistent results on estimates of the expected experimental sensitivity with results obtained

---

[3]The calculation is a next-to-next-to-leading-order calculation for a top quark mass of 170 GeV [62].

frompseudo-experiment datasets. Figure 9.4 shows the posterior density distribution for the expected cross section measurement for all 16 channels combined.

Table 9.4: Expected 95% C.L. upper limit values in pb with all systematic uncertainties taken into account, for many combinations of analysis channels.

	Expected 95% C.L. Upper Limits						
	1,2tags+2,3jets		tau12,3+2,3jets		tau12,3+1,2tags		all
	Tau12	Tau3	1 tag	2 tags	2 jets	3 jets	channels
Run IIa	14.6	27.6	15.5	23.1	15.0	25.0	14.7
Run IIb	9.2	19.7	9.2	19.2	9.3	18.0	9.0
Run IIa+b	8.6	19.4	8.7	16.1	8.8	17.1	8.5

Table 9.5: Expected cross section measurements with all systematic uncertainties taken into account, for many combinations of analysis channels.

	Expected Cross Section Measurements						
	1,2tags+2,3jets		tau12,3+2,3jets		tau12,3+1,2tags		all
	Tau12	Tau3	1 tag	2 tags	2 jets	3 jets	channels
Run IIa	$3.7^{+4.1}_{-3.6}$	$6.5^{+-7.5}_{-6.5}$	$3.8^{+4.4}_{-3.8}$	$3.5^{+8.6}_{-3.5}$	$3.8^{+4.2}_{-3.7}$	$3.8^{+10.2}_{-3.8}$	$3.9^{+4.1}_{-3.7}$
Run IIb	$3.6^{+2.7}_{-2.3}$	$3.5^{+6.5}_{-3.5}$	$3.6^{+2.7}_{-2.3}$	$3.6^{+5.7}_{-3.6}$	$3.6^{+2.7}_{-2.3}$	$3.9^{+5.4}_{-3.9}$	$3.7^{+2.6}_{-2.2}$
Run IIa+b	$3.6^{+2.5}_{-2.1}$	$3.9^{+6.1}_{-3.9}$	$3.6^{+2.5}_{-2.1}$	$3.6^{+4.6}_{-3.6}$	$3.6^{+2.5}_{-2.1}$	$3.9^{+5.0}_{-3.9}$	$3.7^{+2.4}_{-2.1}$

Table 9.6: Expected posterior peak over half-width with all systematic uncertainties taken into account, for many combinations of analysis channels.

	Expected Posterior Peak Over Half-Width						
	1,2tags+2,3jets		tau12,3+2,3jets		tau12,3+1,2tags		all
	Tau12	Tau3	1 tag	2 tags	2 jets	3 jets	channels
Run IIa	1.0	1.0	1.0	1.0	1.0	1.0	1.1
Run IIb	1.6	1.0	1.6	1.0	1.6	1.0	1.6
Run IIa+b	1.7	1.0	1.7	1.0	1.7	1.0	1.8

Table 9.7: Expected Bayes factor significances with all systematic uncertainties taken into account, for many combinations of analysis channels.

	Expected Bayes Factor Significance						
	1,2tags+2,3jets		tau12,3+2,3jets		tau12,3+1,2tags		all
	Tau12	Tau3	1 tag	2 tags	2 jets	3 jets	channels
Run IIa	0.8	0.3	0.8	0.4	0.8	0.4	0.9
Run IIb	1.6	0.5	1.6	0.7	1.6	0.7	1.7
Run IIa+b	1.8	0.6	1.7	0.8	1.8	0.8	1.9

Table 9.8: Expected Bayes ratio significances with all systematic uncertainties taken into account, for many combinations of analysis channels.

	Expected Bayes Ratio Significance						
	1,2tags+2,3jets		tau12,3+2,3jets		tau12,3+1,2tags		all
	Tau12	Tau3	1 tag	2 tags	2 jets	3 jets	channels
Run IIa	0.8	0.3	0.8	0.4	0.8	0.4	0.9
Run IIb	1.6	0.5	1.6	0.7	1.6	0.7	1.7
Run IIa+b	1.8	0.6	1.7	0.8	1.8	0.8	1.9

Table 9.9: Expected upper limit, cross section, peak over half-width, Bayes factor significance and Bayes ratio significance, with all systematic uncertainties considered for the all 16 analysis channels.

Expected Results in Individual Channels					
Channels	Upper limit	$\sigma \pm \Delta\sigma$	P/HW	BFS	BRS
Run IIa / τ type 1+2 / 1 tag / 2 jets	16.1	$3.8^{+4.7}_{-3.8}$	1.0	0.7	0.8
Run IIa / τ type 1+2 / 1 tag / 3 jets	26.1	$3.2^{+12.1}_{-3.2}$	1.0	0.2	0.2
Run IIa / τ type 1+2 / 2 tags / 2 jets	24.7	$3.2^{+10.2}_{-3.2}$	1.0	0.3	0.3
Run IIa / τ type 1+2 / 2 tags / 3 jets	27.2	$2.9^{+14.1}_{-2.9}$	1.0	0.2	0.2
Run IIa / τ type 3 / 1 tag / 2 jets	28.0	$6.9^{+-7.9}_{-6.9}$	1.0	0.2	0.3
Run IIa / τ type 3 / 1 tag / 3 jets	28.2	$8.0^{+-9.0}_{-8.0}$	1.0	0.1	0.2
Run IIa / τ type 3 / 2 tags / 2 jets	28.1	$6.8^{+-7.8}_{-6.8}$	1.0	0.2	0.2
Run IIa / τ type 3 / 2 tags / 3 jets	28.2	$3.8^{+-4.8}_{-3.8}$	1.0	0.1	0.1
Run IIb / τ type 1+2 / 1 tag / 2 jets	9.7	$3.6^{+2.8}_{-2.4}$	1.5	1.5	1.5
Run IIb / τ type 1+2 / 1 tag / 3 jets	19.2	$4.0^{+5.9}_{-4.0}$	1.0	0.6	0.7
Run IIb / τ type 1+2 / 2 tags / 2 jets	20.5	$3.4^{+6.6}_{-3.4}$	1.0	0.6	0.6
Run IIb / τ type 1+2 / 2 tags / 3 jets	25.2	$3.4^{+10.4}_{-3.4}$	1.0	0.4	0.4
Run IIb / τ type 3 / 1 tag / 2 jets	20.9	$3.4^{+7.4}_{-3.4}$	1.0	0.5	0.5
Run IIb / τ type 3 / 1 tag / 3 jets	27.4	$3.1^{+14.4}_{-3.1}$	1.0	0.2	0.2
Run IIb / τ type 3 / 2 tags / 2 jets	27.3	$3.4^{+13.8}_{-3.4}$	1.0	0.2	0.2
Run IIb / τ type 3 / 2 tags / 3 jets	28.0	$3.2^{+-4.2}_{-3.2}$	1.0	0.1	0.1
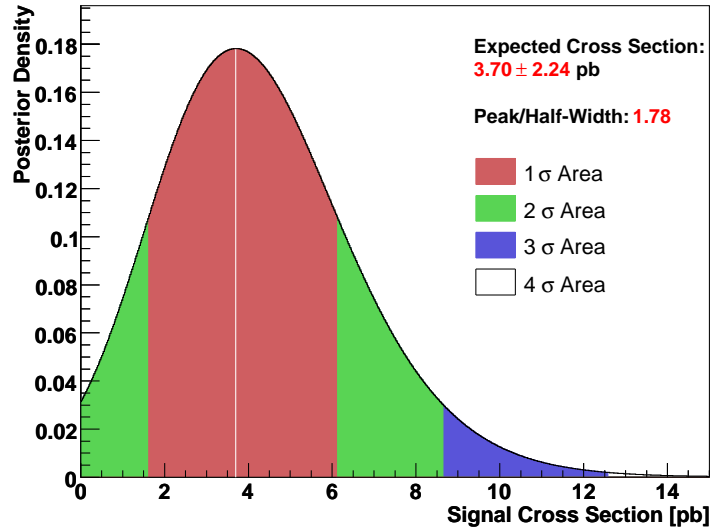
Figure 9.4: Posterior density distribution for the expected cross section measurement for all 16 channels combined (Run IIa–Run IIb, tau type 1,2,3, 1-2 tags, 2-3 jets). All systematic uncertainties are taken into account. The input theoretical cross section is 3.46 pb.

9.6 Observed Results

The observed cross section calculations and corresponding significance estimates are presented in this section. The BDT probability distributions used for this measurement are shown in Figs. H.1–H.8 in Appendix H. Figure 9.5 shows the comparison plots with different regions summed over all 16 channels (the plots in Figure 9.5 are only for illustration purposes). The comparison plots between data and background look compatible. The histograms in Figs. H.1–H.8, along those shifted by systematic effects, are used to calculate the limit or the cross sections of single top production.

(a)   (b)
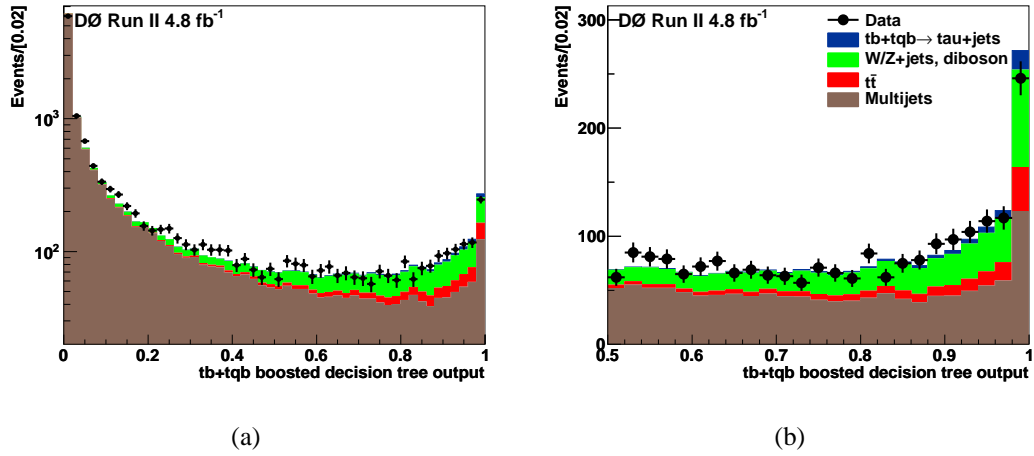
Figure 9.5: BDT probability output distributions combined from all channels with different regions. (a) the region is [0.0, 1.0] with a log scale, (b) the region is [0.5, 1.0] with a linear scale.

Table 9.10 shows the upper limit values at 95% CL and Table 9.11 shows the observed cross section measurements with all systematic uncertainties taken into account in certain combined channels. Table 9.12 shows the ratio of the posterior peak position over the lower half-width. In addition, the Bayes ratio significance is shown in Table 9.13[4]. Fig. 9.6 shows the posterior density distribution with 1–4 σ area shaded for observed cross section measurement for all 16 channels combined (Run IIa–Run IIb, tau type 1,2,3, 1-2 tags, 2-3 jets).

Table 9.10: Observed 95% C.L. upper limit values in pb with all systematic uncertainties taken into account, for many combinations of analysis channels.

	Observed 95% C.L. Upper Limits						
	1,2tags+2,3jets		tau12,3+2,3jets		tau12,3+1,2tags		all
	Tau12	Tau3	1 tag	2 tags	2 jets	3 jets	channels
Run IIa	12.7	28.6	15.8	23.1	17.9	19.9	13.3
Run IIb	11.4	7.1	9.2	19.5	9.5	11.7	8.1
Run IIa+b	9.6	7.7	8.6	14.4	9.0	9.0	7.3

---

[4]It doesn't make sense to calculate the Bayes factor significance for a measurement in real data.

Table 9.11: Observed cross section measurements with all systematic uncertainties taken into account, for many combinations of analysis channels.

	Observed Cross Section Measurements						
	1,2tags+2,3jets		tau12,3+2,3jets		tau12,3+1,2tags		all
	Tau12	Tau3	1 tag	2 tags	2 jets	3 jets	channels
Run IIa	$2.4^{+4.1}_{-2.4}$	$24.7^{+-25.7}_{-22.6}$	$4.3^{+4.5}_{-4.0}$	$3.6^{+8.6}_{-3.6}$	$6.1^{+5.2}_{-4.5}$	$0.0^{+8.9}_{-0.0}$	$3.3^{+3.8}_{-3.3}$
Run IIb	$5.6^{+2.9}_{-2.4}$	$0.0^{+2.8}_{-0.0}$	$3.7^{+2.6}_{-2.2}$	$6.7^{+5.9}_{-4.7}$	$4.4^{+2.6}_{-2.2}$	$0.0^{+5.4}_{-0.0}$	$3.6^{+2.3}_{-2.0}$
Run IIa+b	$4.9^{+2.4}_{-2.1}$	$0.0^{+3.1}_{-0.0}$	$3.8^{+2.4}_{-2.1}$	$4.8^{+4.2}_{-3.7}$	$4.4^{+2.4}_{-2.0}$	$0.0^{+4.0}_{-0.0}$	$3.4^{+2.0}_{-1.8}$

Table 9.12: Observed posterior peak over half-width with all systematic uncertainties taken into account, for many combinations of analysis channels.

	Observed Posterior Peak Over Half-Width						
	1,2tags+2,3jets		tau12,3+2,3jets		tau12,3+1,2tags		all
	Tau12	Tau3	1 tag	2 tags	2 jets	3 jets	channels
Run IIa	1.0	1.1	1.1	1.0	1.3	0.0	1.0
Run IIb	2.3	0.0	1.7	1.4	2.0	0.0	1.8
Run IIa+b	2.3	0.0	1.8	1.3	2.1	0.0	1.9

Table 9.13: Observed Bayes ratio significances with all systematic uncertainties taken into account, for many combinations of analysis channels.

	Observed Bayes Ratio Significance						
	1,2tags+2,3jets		tau12,3+2,3jets		tau12,3+1,2tags		all
	Tau12	Tau3	1 tag	2 tags	2 jets	3 jets	channels
Run IIa	0.5	1.3	0.9	0.4	1.3	0.0	0.8
Run IIb	2.6	0.0	1.7	1.4	2.1	0.0	1.9
Run IIa+b	2.6	0.0	1.9	1.2	2.3	0.0	1.9

Table 9.14: Observed upper limit, cross section, peak over half-width and Bayes ratio significance, with all systematic uncertainties considered for the all 16 analysis channels.

Observed Results in Individual Channels				
Channels	Upper limit	$\sigma \pm \Delta\sigma$	P/HW	BRS
Run IIa / τ type 1+2 / 1 tag / 2 jets	18.7	$5.7^{+5.4}_{-4.7}$	1.2	1.1
Run IIa / τ type 1+2 / 1 tag / 3 jets	24.1	$0.0^{+12.4}_{-0.0}$	0.0	0.0
Run IIa / τ type 1+2 / 2 tags / 2 jets	26.3	$6.2^{+9.6}_{-6.2}$	1.0	0.6
Run IIa / τ type 1+2 / 2 tags / 3 jets	25.8	$0.0^{+14.0}_{-0.0}$	0.0	0.0
Run IIa / τ type 3 / 1 tag / 2 jets	28.5	$27.2^{+-28.2}_{-27.2}$	1.0	1.0
Run IIa / τ type 3 / 1 tag / 3 jets	28.2	$0.0^{+-1.0}_{-0.0}$	0.0	0.0
Run IIa / τ type 3 / 2 tags / 2 jets	28.5	$29.8^{+-30.8}_{-29.8}$	1.0	1.0
Run IIa / τ type 3 / 2 tags / 3 jets	28.3	$21.0^{+-22.0}_{-21.0}$	1.0	0.4
Run IIb / τ type 1+2 / 1 tag / 2 jets	13.3	$6.3^{+3.4}_{-2.8}$	2.3	2.6
Run IIb / τ type 1+2 / 1 tag / 3 jets	15.5	$0.0^{+7.1}_{-0.0}$	0.0	0.0
Run IIb / τ type 1+2 / 2 tags / 2 jets	25.2	$9.6^{+8.0}_{-6.1}$	1.6	1.6
Run IIb / τ type 1+2 / 2 tags / 3 jets	25.8	$4.8^{+9.9}_{-4.8}$	1.0	0.5
Run IIb / τ type 3 / 1 tag / 2 jets	8.2	$0.0^{+3.2}_{-0.0}$	0.0	0.0
Run IIb / τ type 3 / 1 tag / 3 jets	25.3	$0.0^{+13.2}_{-0.0}$	0.0	0.0
Run IIb / τ type 3 / 2 tags / 2 jets	27.9	$10.7^{+-11.7}_{-10.7}$	1.0	0.6
Run IIb / τ type 3 / 2 tags / 3 jets	28.5	$29.8^{+-30.8}_{-29.8}$	1.0	1.0

In this analysis, due to the low experimental sensitivity, the *s*- and *t*-channels are measured together. It would be interesting to determine their cross sections separately since the two processes will be sensitive to different types of new physics. For this reason, limitsfor treating them in two dimensions are also set. The two-dimensional (2D) limits are set by the same Bayesian approach described in Section 9.1 but with one more signal channel. The full description can be found in Ref. [81]. Fig. 9.7 shows results of 2D limit setting: (a) in the figure is *s*-, *t*-channel and *s+t* combined observed posterior density probabilities while (b) presents the contour of posterior density probabilities of the *s* channel versus the *t* channel. In (b), the measured peak in black is consistent with the Standard-Model prediction.
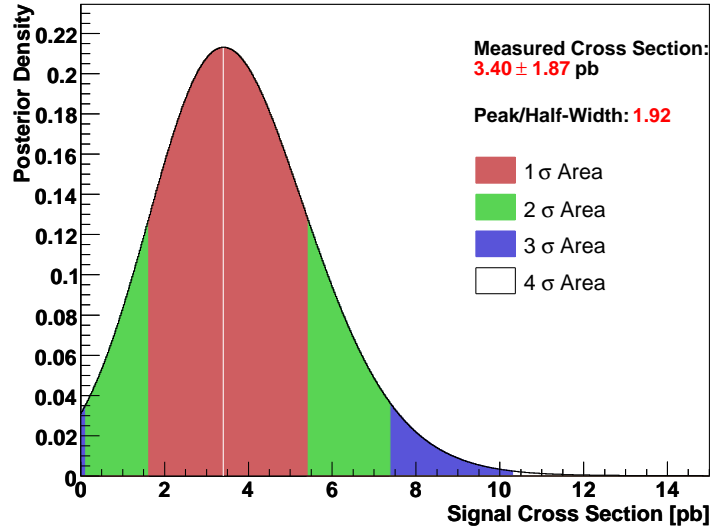
Figure 9.6: Posterior density distribution for the observed cross section measurement for all 16 channels combined (Run IIa–Run IIb, tau type 1,2,3, 1-2 tags, 2-3 jets). All systematic uncertainties are taken into account.
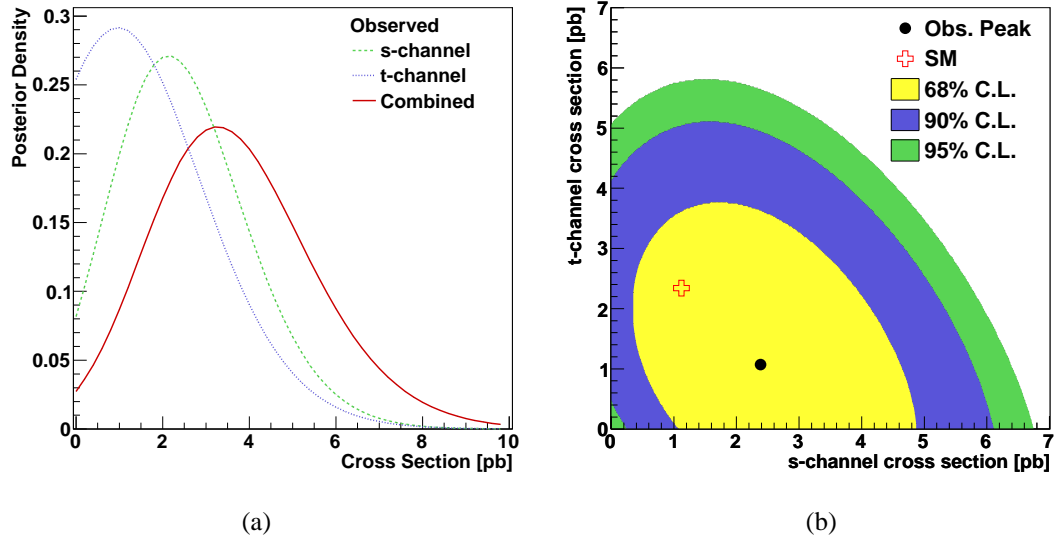
Figure 9.7: Results of 2D limit setting. (a) *s*-, *t*-channel and *s*+*t* combined observed posterior density probabilities; (b) Contour of posterior density probabilities of *s* channel versus *t* channel. The Standard Model prediction is within 1 σ band.

9.7 Combination with e/μ+jets Channels

9.7.1 Combined Cross Section

As the sample in the tau+jets channel is orthogonal to that in electron/muon+jets channels [13], it is possible to combine the results. A cross section of $3.94^{+0.88}_{-0.88}$ pb is measured in the electron/muon+jets channels. In the combination, the tau lepton and the electron/muon+jets channels are treated as two independent channels using the same Bayesian approach implemented to combine different tau analysis channels above. The combined expected and observed posterior densities are calculated and the combined measured cross section is

$$\sigma(p\bar{p} \rightarrow tb+X, tqb+X) = 3.84^{+0.89}_{-0.83} \text{ pb} \tag{9.22}$$

with a ratio of the peak of the expected posterior density to its width 4.7 compared to 4.5 in electron and muon combined.

9.7.2 $|V_{tb}|$ Measurement

As mentioned before, the coupling $|V_{tb}|$ between b and t quarks is sensitive to new physics beyond the SM.

The Bayesian posterior density functions for $|V_{tb}f_1^L|^2$ and $|V_{tb}|^2$ are shown in Fig. 9.8. The latter corresponding to $f_1^L = 1$ in the former is obtained by restricting the prior to be in the region [0,1].

By the peak of the posterior density function of $|V_{tb}f_1^L|^2$ shown in Fig. 9.8 (a), $|V_{tb}f_1^L|^2$ is measured to be $1.12^{+0.27}_{-0.25}$ from which a $|V_{tb}f_1^L|$ value of $1.06^{+0.13}_{-0.12}$ is extracted. Thus the V-A coupling strength $|V_{tb}f_1^L|$ of 1.06 is measured above the SM expectation.

If the prior is restricted to the region [0,1] and $f_1^L = 1$, $|V_{tb}^2| = 1.00^{+0.00}_{-0.22}$ is measured and hence $|V_{tb}| = 1.00^{+0.00}_{-0.12}$. Finally, from the posterior density function of $|V_{tb}|^2$ shown in (b) of Fig. 9.8, a lower limit of 0.72 on $|V_{tb}|^2$ at 95% confidence level is calculated corresponding to a lower limit of $V_{tb} = 0.85$. For data in e/mu+jets channels, $|V_{tb}| > 0.78$ [13].

(a)

(b)
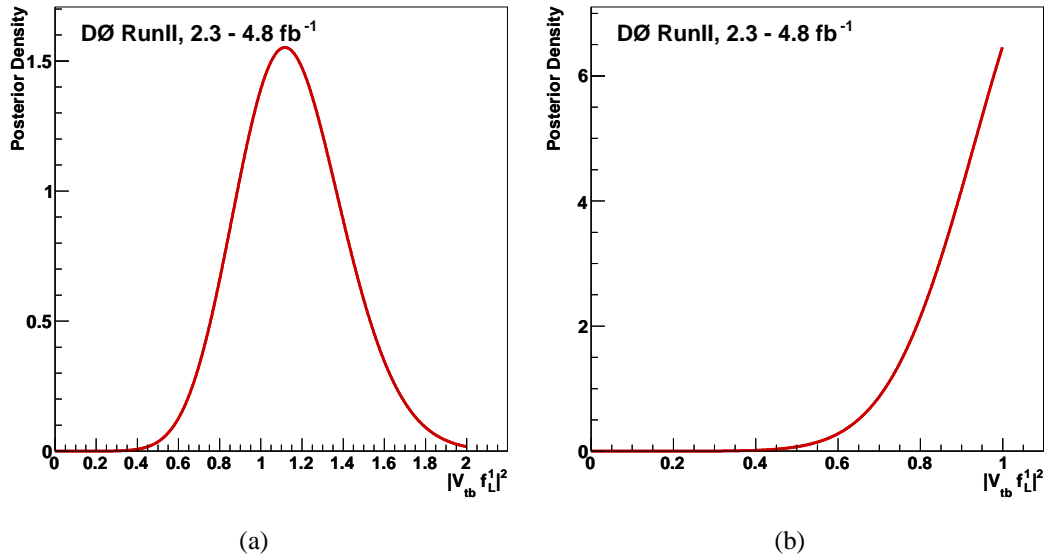
Figure 9.8: Posterior probability density function for (a) $|V_{tb}f_1^L|^2$ and (b) $|V_{tb}|^2$ with a prior in the region [0,1].

Chapter 10

Summary

This thesis presents a detailed analysis of the measurement of the rare electroweak single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV at the Tevatron. The analysis uses 4.8 fb^{-1} of data collected by the D0 detector system during the Tevatron Run II period. This is the first effort to search for single top quark production in the tau+jets channel. In the analysis, an optimization of tau lepton identification is conducted using boosted decision trees and a new tau ID is designed and optimized for this channel. Due to the dominant QCD-multijet background, a novel multijet tag rate background model, by means of b-jet tagging rate functions and available MC sources, is carefully developed to understand the final analysis data. The $t\bar{t}$ cross section based on the final data sample is also measured to cross check and guarantee the quality of the background model. Boosted decision trees are used to make a variable to discriminate single top signals from backgrounds. By a Bayesian statistical approach, an upper limit of 7.3 pb at 95% confidence level is obtained. The observed cross section of single top quark production is measured to be $3.4^{+2.0}_{-1.8}$ pb with a 1.9 standard deviation significance. This is consistent with the Standard-Model prediction of 3.46 pb and also with the result in the electron/muon+jets channels within 1 σ uncertainty.

Since the data sample selected in the tau+jets channel is orthogonal to that in the electron/muon+jets channels, a combination is performed by the same Bayesian approach. The combined observed cross section is measured as

$$\sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 3.84^{+0.89}_{-0.83} \text{ pb}$$

with an expected sensitivity of 4.7 compared to 4.5 in the electron/muon+jet channels alone. Assuming $f_1^L = 1$, a lower limit of $|V_{tb}|$ is also measured to be larger than 0.85 at 95% confidence level. As a summary, Fig. 10.1 shows several recent measurements of single top quark production compared to the theoretical SM prediction, 3.46±0.18 pb, calculated for a top mass of 170 GeV [18].
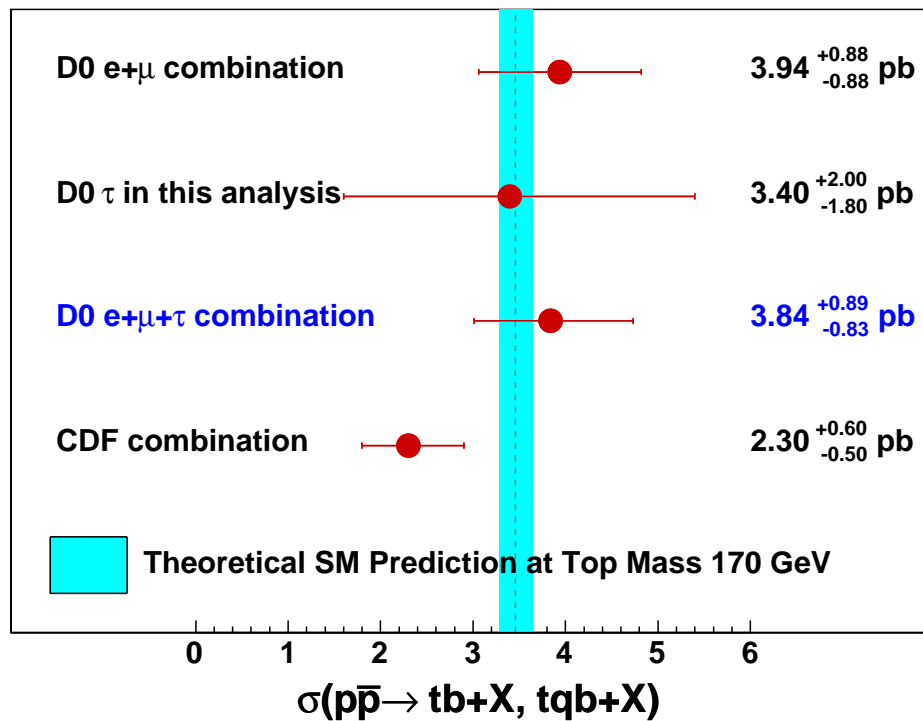
Figure 10.1: Summary plot of several recent measurements of single top quark production. The theoretical SM prediction [18] at a top mass of 170 GeV is included as a shaded band. The "D0 e+μ combination" result is taken from [13] while "CDF combination" results come from [14].

Appendix A

Event Display

In this appendix, three single top quark candidate events with tau types 1, 2 and 3 are displayed in Figs. A.1, A.2 and A.3 respectively. These events have large boosted decision tree outputs indicating that they are likely to be single top quark signal. Each event contains one tau lepton candidate, $\not{E}_T$ and two jets (one of which is b-tagged). Each event is displayed in a calorimeter η-ϕ space lego plot, a transverse view and longitudinal view.
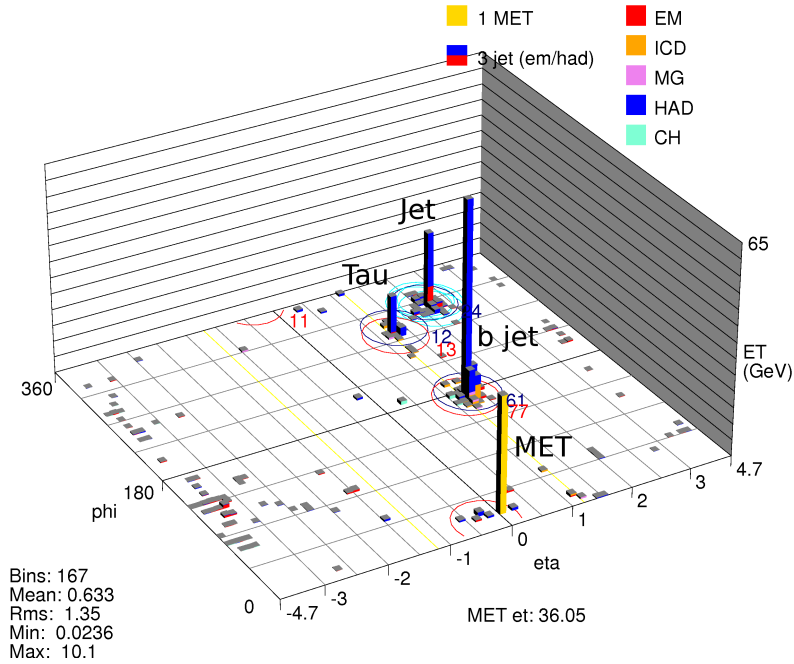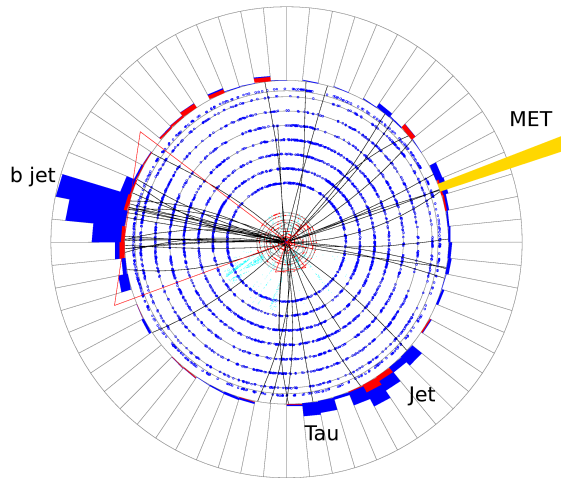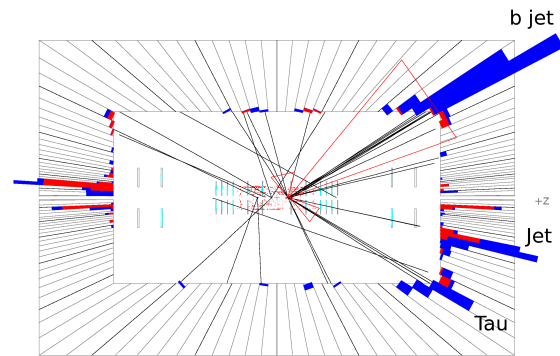
Figure A.1: Event display of a signal candidate event with a type 1 tau and two jets one of which is b-tagged. The output of the boosted decision tree to discriminate signal for this event is 0.974. (a) is a calorimeter $\eta - \phi$ space lego plot, (b) is a X-Y transverse view and (c) is a longitudinal side view. Reconstructed tracks are shown as black lines while tower energy deposits in the EM and hadronic calorimeter are shown as red and blue bars and $\not\!\!E_T$ as yellow bars. In the event, the type 1 tau appears as a narrow jet and later it is identified as a hadronic tau by the tau ID.
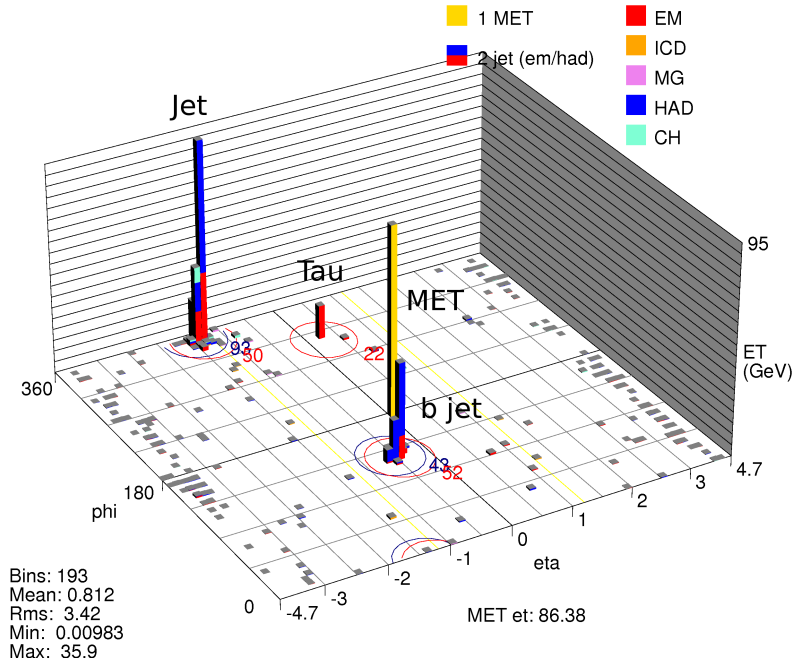
(a)

(b)
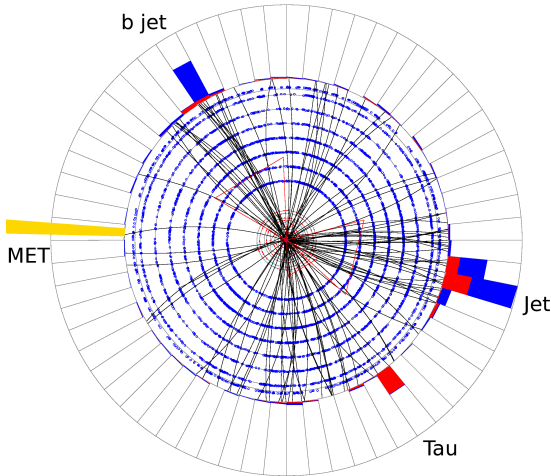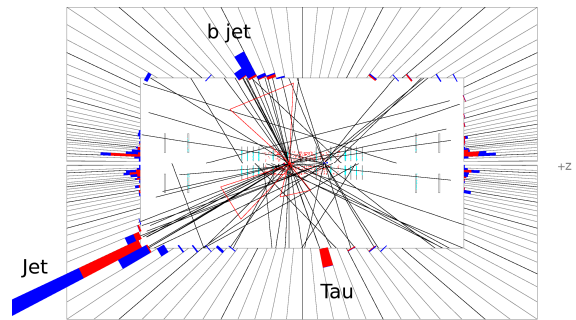
(c)

Figure A.2: Event display of a signal candidate event with a type 2 tau and two jets one of which is *b*-tagged. The output of the boosted decision tree to discriminate signal for this event is 0.975. (a) is a calorimeter $\eta - \phi$ space lego plot, (b) is a *X-Y* transverse view and (c) is a longitudinal side view. Reconstructed tracks are shown as black lines while tower energy deposits in the EM and hadronic calorimeter are shown as red and blue bars and $\not{E}_T$ as yellow bars. In the event, this type 2 tau deposits most of its energy in the EM section of the calorimeter due to two photons from π^0 decay.
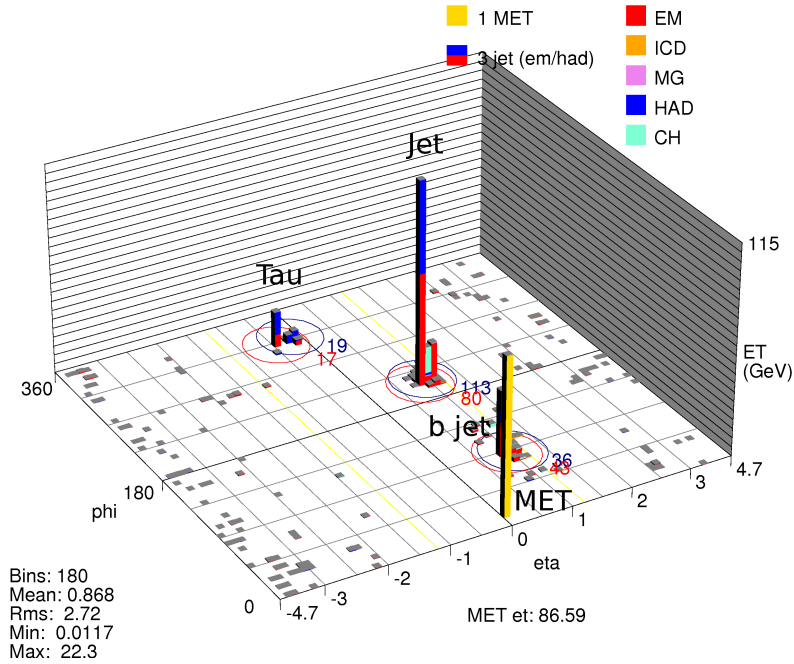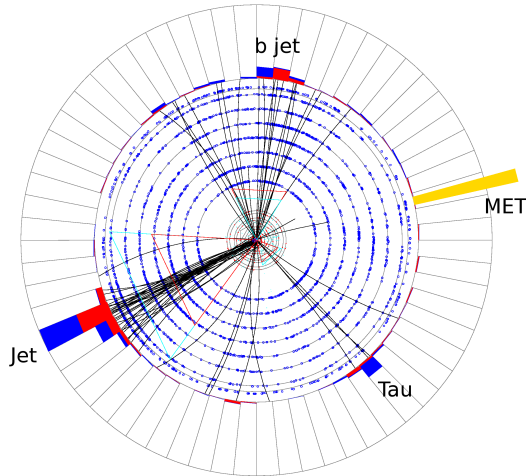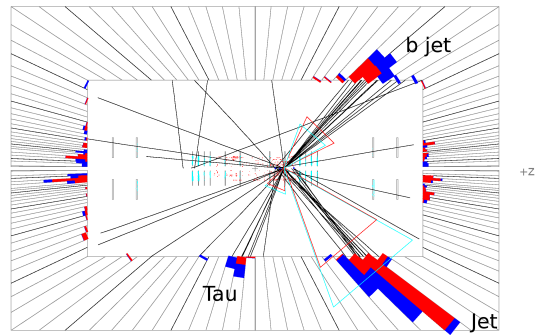
Figure A.3: Event display of a signal candidate event with a type 3 tau and two jets one of which is *b*-tagged. The output of the boosted decision tree to discriminate signal for this event is 0.983. (a) is a calorimeter $\eta - \phi$ space lego plot, (b) is a *X-Y* transverse view and (c) is a longitudinal side view. Reconstructed tracks are shown as black lines while tower energy deposits in the EM and hadronic calorimeter are shown as red and blue bars and $\not{E}_T$ as yellow bars. In the event, the type 3 tau appears as a narrow jet in the calorimeter with which three tracks are associated.

Appendix B

b-tagged Sample and Its Data TRFed Sample

In Chapter 8, it is mentioned that the QCD-multijet event sample is the dominant background component; an approach by measuring data-TRF is therefore employed to estimate a rough background sample first. Plots below show comparisons of some variables (explanations of the variables can be found in Table 9.2) in the most sensitive channel (Run IIb, tau types 1 and 2, 1 b-tag and 2 jets) between the b-tagged sample and its data-TRFed predicted sample (i.e. an event-permuted sample). The comparison shows that between the two samples are compatible within the 1 σ uncertainty band, which indicates that the data-TRFed predicted sample can serve as a base background sample from which physics processes like $t\bar{t}$, W+jets, Z+jets, dibosons should be subtracted. The band is obtained by shifting the data-TRF up and down by 1 σ and is used to evaluate the systematic shape on the QCD-multijet TRF. For each row below, the left plot is in a linear scale while the right one is in a log scale.
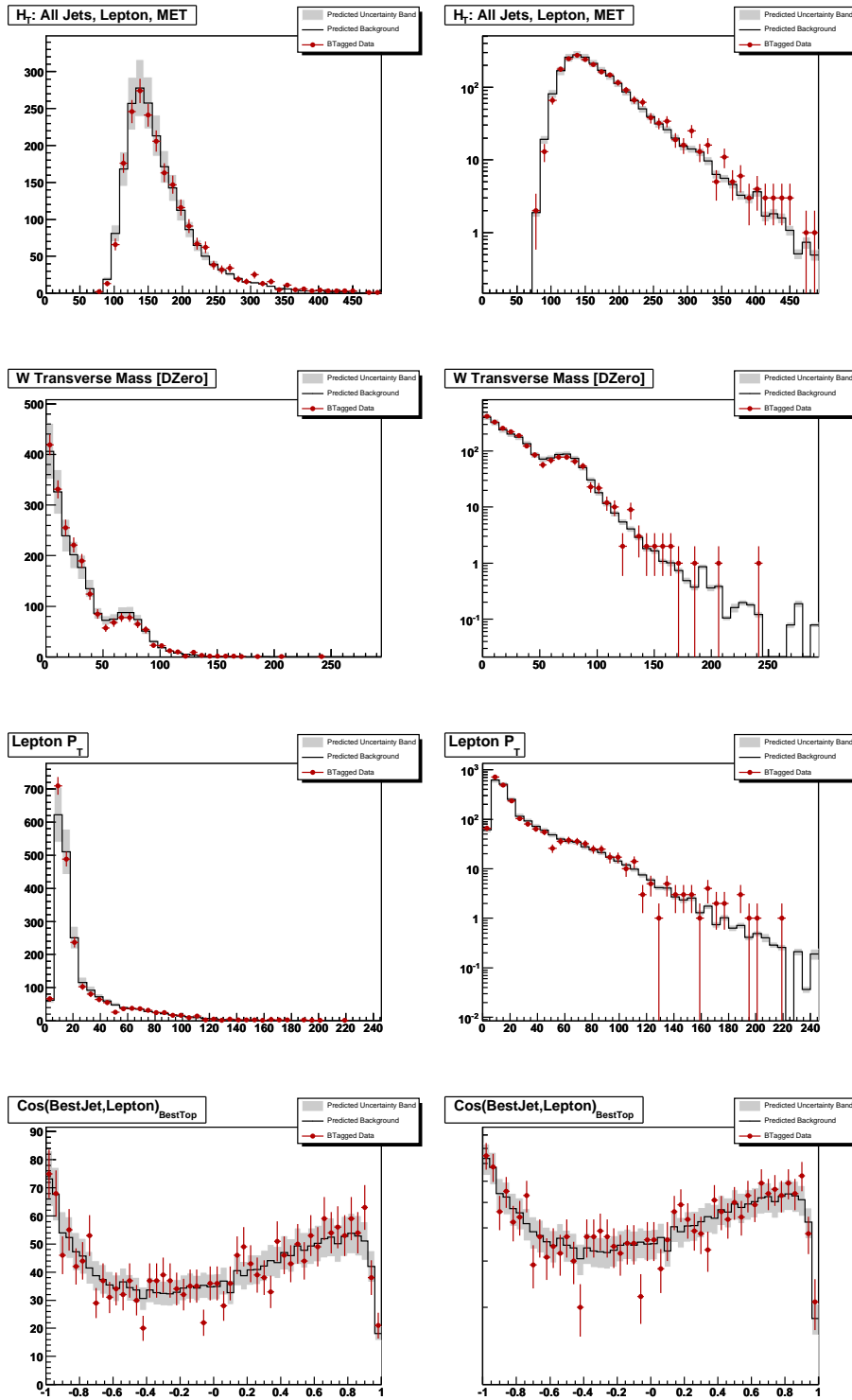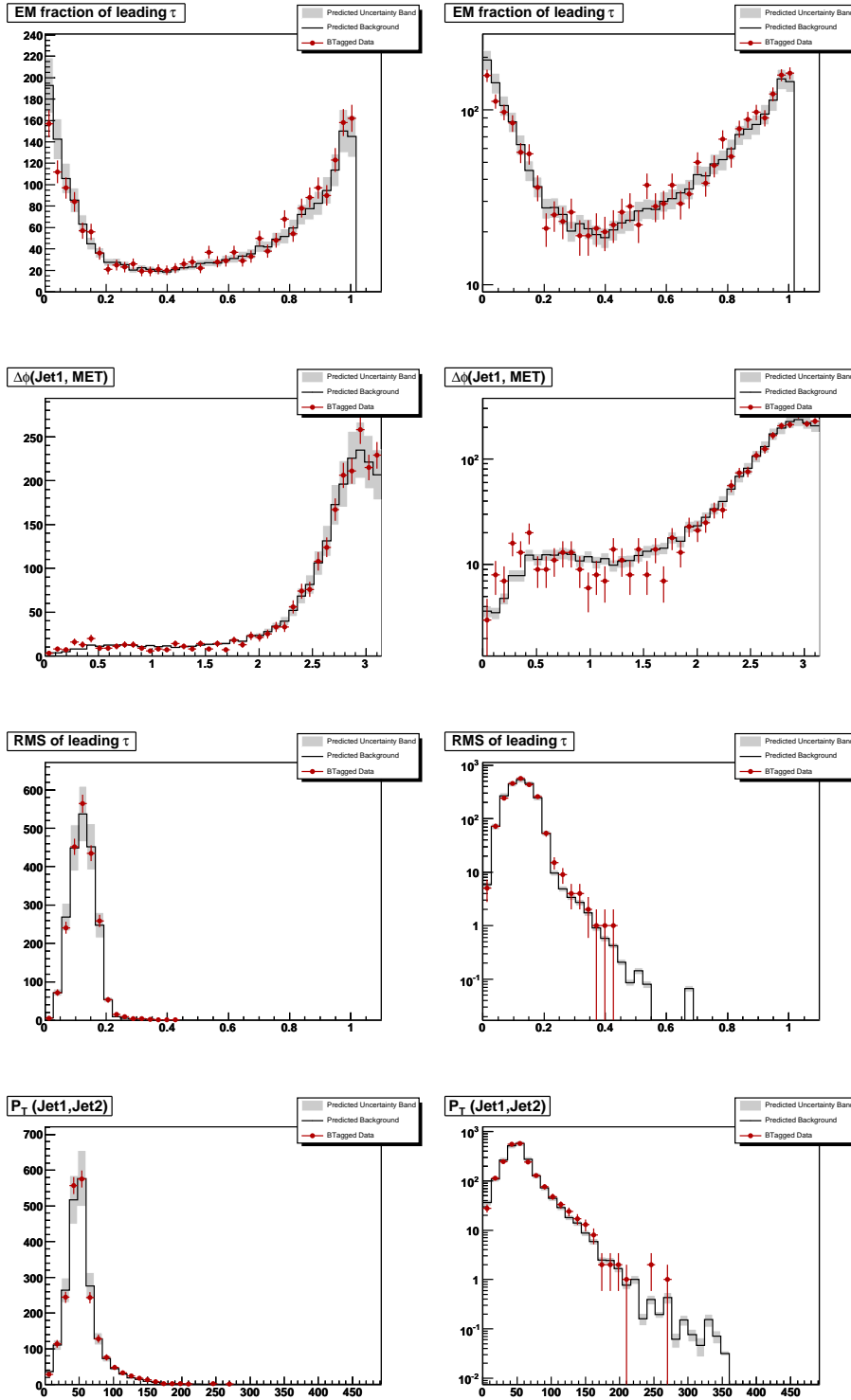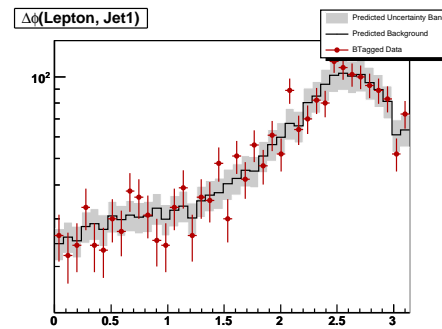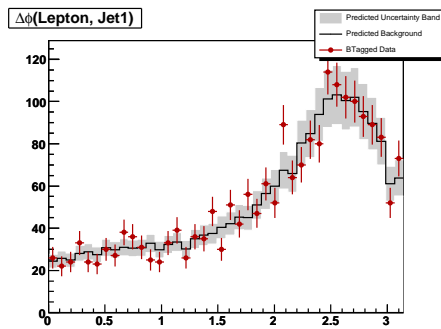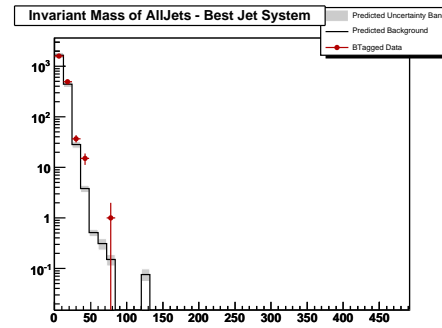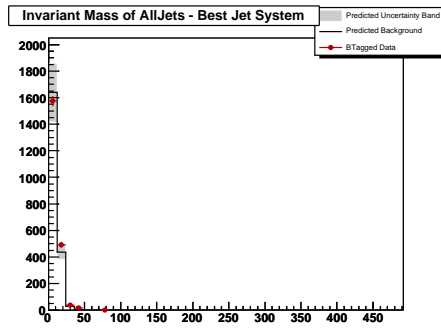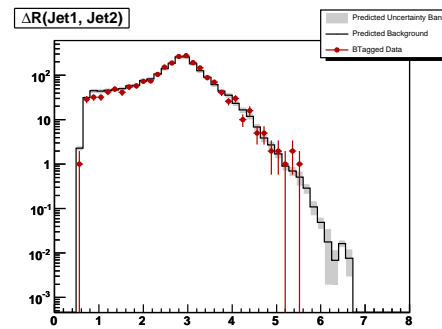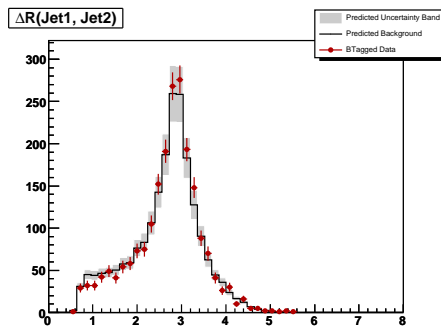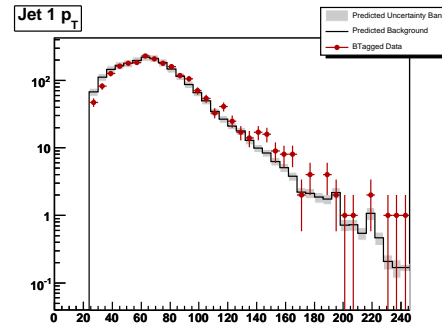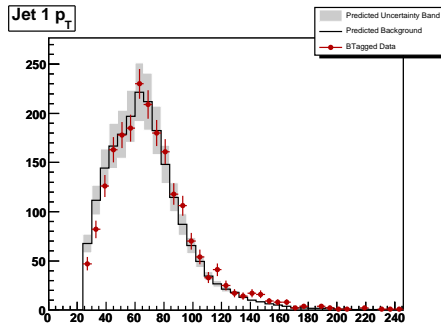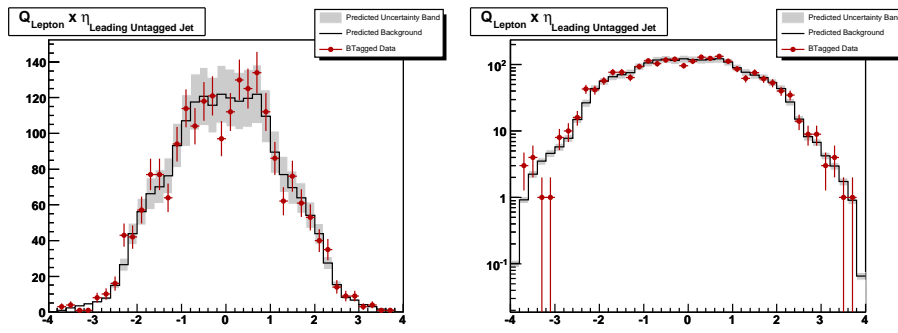
Appendix C

Extra Tau Variables and Distributions

In Section 7.2, hadronic tau reconstruction at D0 has been discussed and definitions of some tau variables are given. In this appendix, definitions of some extra variables with high power for real tau discrimination from background jets are given. Distributions of those variables are also shown below. The comparison samples on which the distributions are based are:

- Single top in s- and t-channel MC samples (used as a BDT tau ID training/testing signal sample)

- Tau trigger skimmed sample (used as a BDT tau ID training/testing background sample, the definition is presented in Section 7.4.2.)

- $Z \rightarrow \tau\tau$ MC sample with Z mass 130–250 GeV (used as a NN tau ID training signal sample)

- $Z \rightarrow \tau\tau$ MC sample with Z mass 60–130 GeV (used as a NN tau ID testing signal sample)

- Anti-isolation μ-τ pair (used as a NN tau ID training background sample)

The definitions of the samples used in the BDT tau ID are presented in Section 7.4.2 while those of the samples in NN tau ID are available in Section 7.3.2.

C.1 More Tau Variables

dalpha

$$\frac{\sqrt{(\frac{\Delta\phi}{\sin\theta})^2 + (\Delta\eta)^2}}{\pi} \tag{C.1}$$

where $\Delta\phi$ and $\Delta\eta$ are differences between the vector sums of tau tracks and of all EM sub-clusters, θ is an azimuthal angle of the centroid of the vector sum of EM sub-clusters

e1e2 $\sqrt{E_T^{\tau_{\text{trk}}} \cdot E_T^{\text{EM}}}$ where $E_T^{\tau_{\text{trk}}}$ is the sum over all tau-associated tracks' E_T, E_T^{EM} is E_T of the sum over EM sub-cluster(s). For a system of tau-associated tracks and EM sub-clusters, its mass is $\pi \times$e1e2$\times$dalpha

prf3 a ratio of E_T of the highest p_T EM sub-cluster over E_T deposited in the EM section layer 3 of the calorimeter within a cone $R < 0.5$

profile2 If $|\eta_{\text{detector}}| > 1.5$, profile2 = profile$\times$(0.67+0.22$\times|\eta_{\text{detector}}|$), else profile2=profile

emET_o_ET a ratio of E_T of EM sub-clusters over tau E_T

ett1_o_ETiso a ratio of p_T of the highest p_T tau track over $E_T^{\tau\text{iso}}$ which is the tau E_T within a cone $R < 0.7$

rms2 rms2=rms/(1.0+0.29$\times|\eta_{\text{det}}|$) where η_{det} is τ detector η

ettr E_T of tracks except the first two highest p_T tracks

emcl_et1 E_T of the highest p_T EM-cluster

emcl_et2 E_T of the secondly highest p_T EM-cluster

emcl_f12 Fraction of EM sub-cluster energy deposited in the EM section layer 1 and layer 2

emcl_f4 Fraction of EM sub-cluster energy deposited in the EM section layer 4

ett1 E_T of the highest p_T track

iso2 If $|\eta_{\text{detector}}| > 1.0$, iso2 $=$ iso$/(1.5|\eta_{\text{detector}}| - 0.5)$, otherwise iso2=iso

Et_iso tau E_T within a cone $R < 0.7$

chf Fraction of energy deposited in the coarse hadronic calorimeter section

icdf Fraction of energy deposited in the inter-cryostat detector (ICD)

EM12_Et E_T of energy deposited in the EM section layer 1 and layer 2 within a cone $R < 0.5$

EM3_Et E_T of energy deposited in the EM section layer 3 within a cone $R < 0.5$

EM4_Et E_T of energy deposited in the EM section layer 4 within a cone $R < 0.5$

EM12_Et_iso E_T of energy deposited in the EM section layer 1 and layer 2 within a cone $R < 0.7$

EM3_Et_iso E_T of energy deposited in the EM section layer 3 within a cone $R < 0.7$

EM4_Et_iso E_T of energy deposited in the EM section layer 4 within a cone $R < 0.7$

EM4f Fraction of tau energy deposited in the EM section layer 4 within a cone $R < 0.5$

EM4isof Fraction of tau energy deposited in the EM section layer 4 within a cone $R < 0.7$

tzDCA z of the highest p_T track at DCA

hot a ratio of E_T of the hottest cell over that of the secondly hottest cell

C.2   Some Tau Variable Distributions
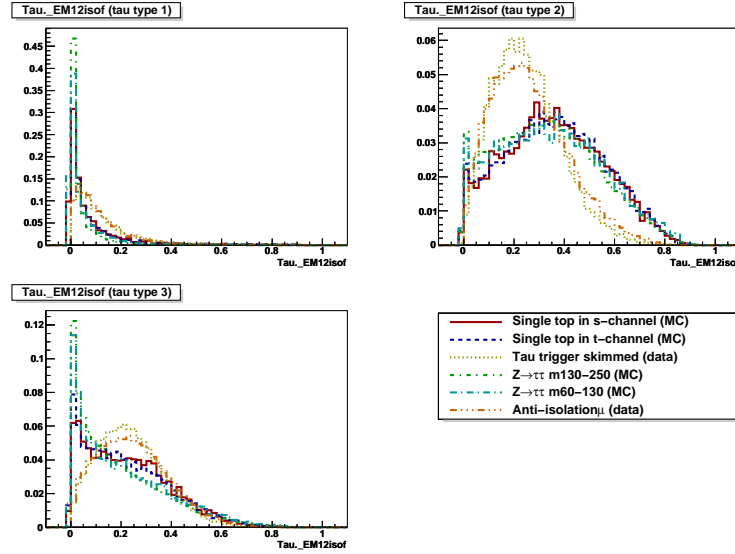
Figure C.1: Comparison between different samples used in tau ID study for the variable EM12isof.
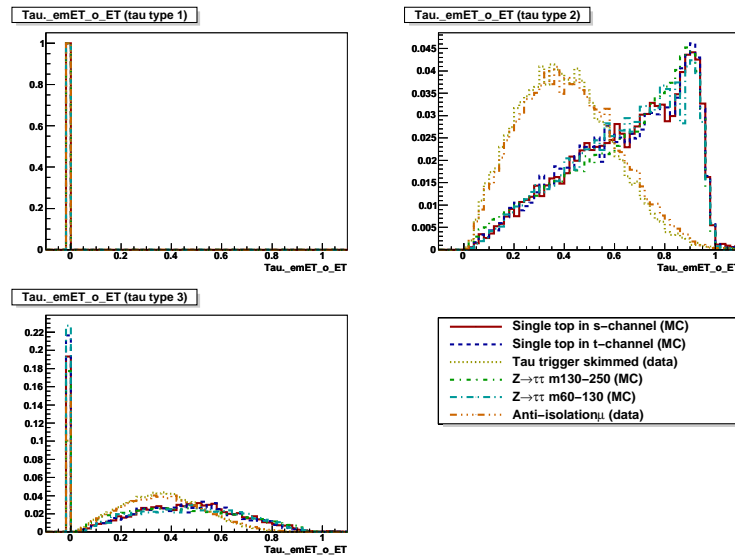
Figure C.2: Comparison between different samples used in tau ID study for the variable emET_o_ET.
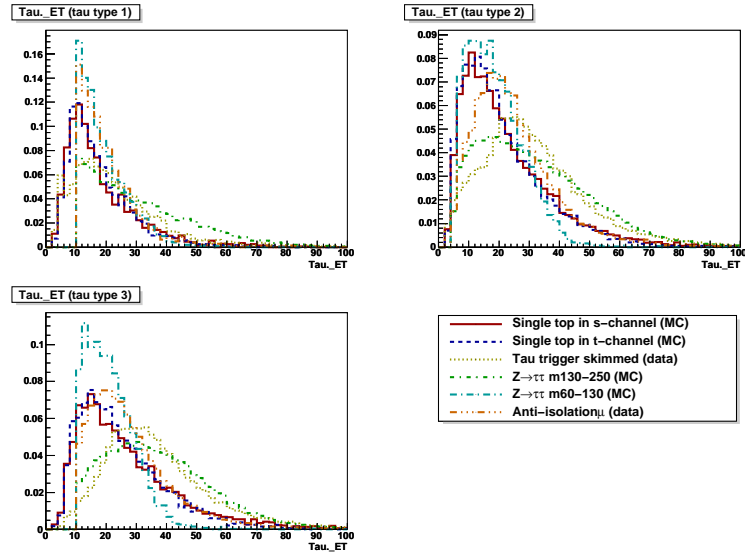
Figure C.3: Comparison between different samples for the variables E_T.
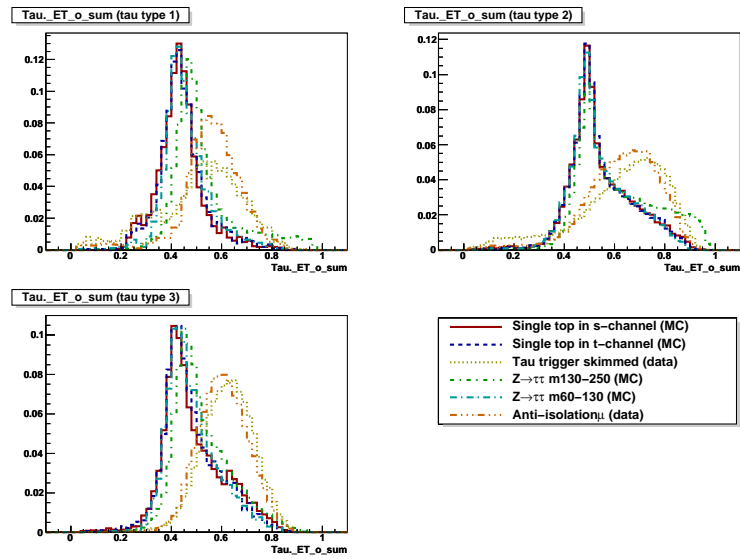
Figure C.4: Comparison between different samples used in tau ID study for the variable ET_o_sum.
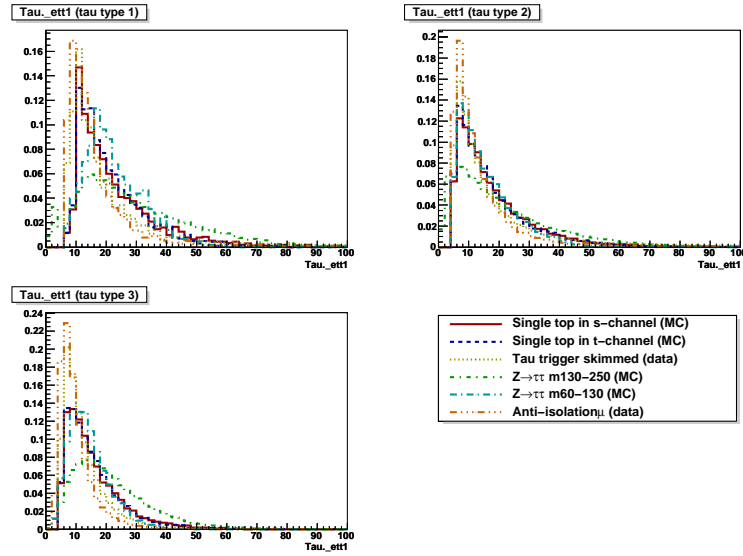
Figure C.5: Comparison between different samples used in tau ID study for leading track E_T of tau (ett1).
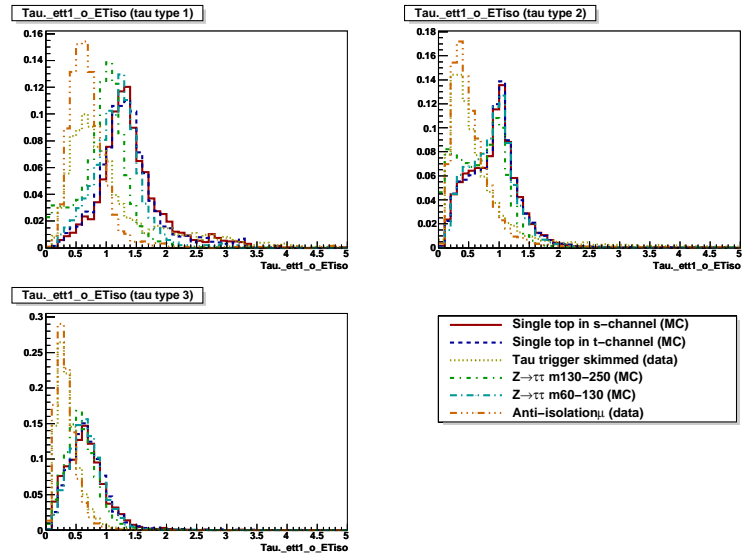
Figure C.6: Comparison between different samples used in tau ID study for the variable ett1_o_ETiso.
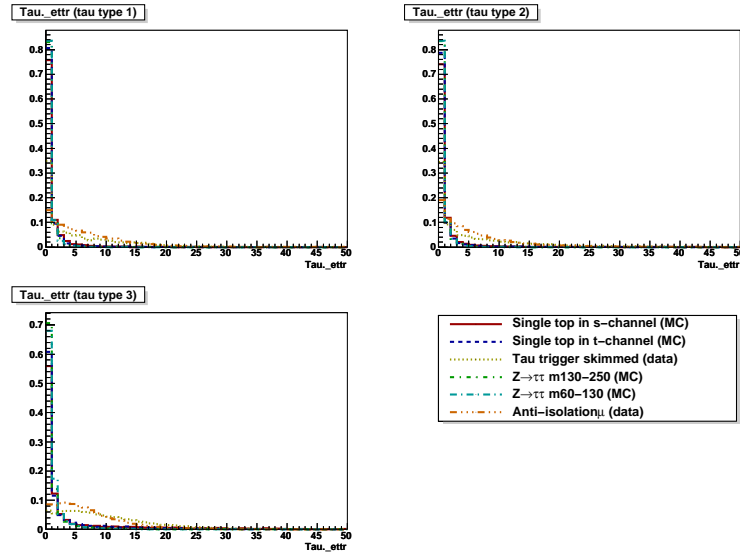
Figure C.7: Comparison between different samples used in tau ID study for E_T of tracks except the first 2 leading tracks.
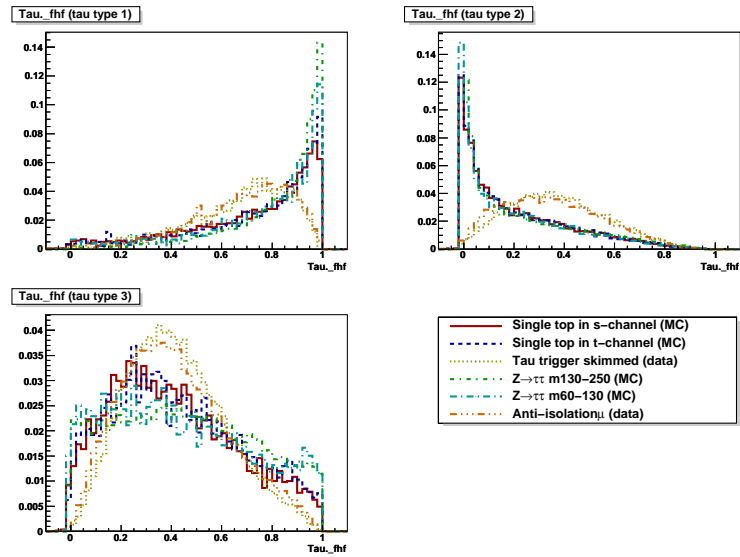
Figure C.8: Comparison between different samples used in tau ID study for the variable fhf.
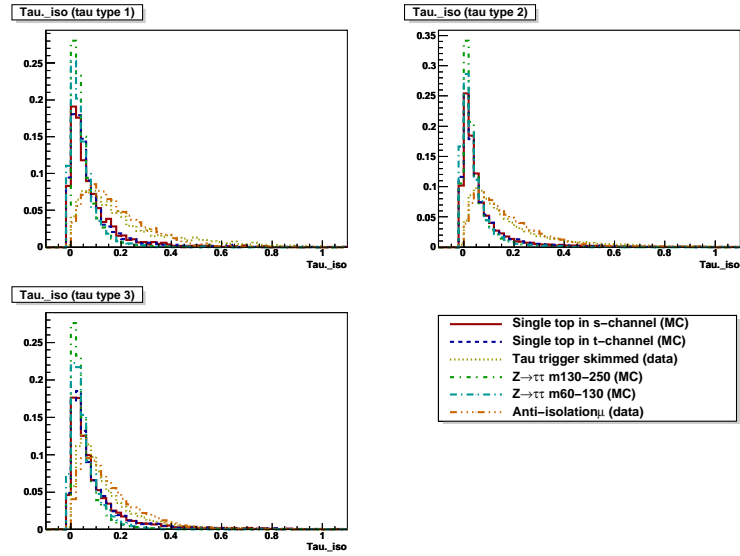
Figure C.9: Comparison between different samples used in tau ID study for the variable iso.
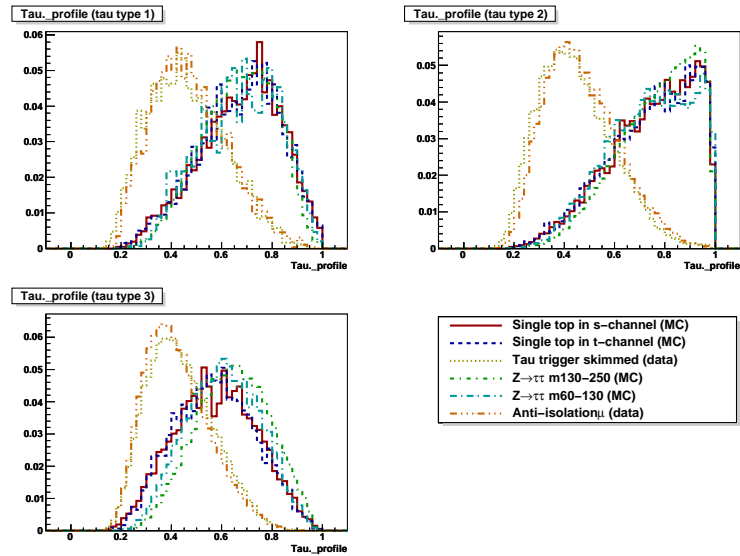
Figure C.10: Comparison between different samples used in tau ID study for the variable profile.
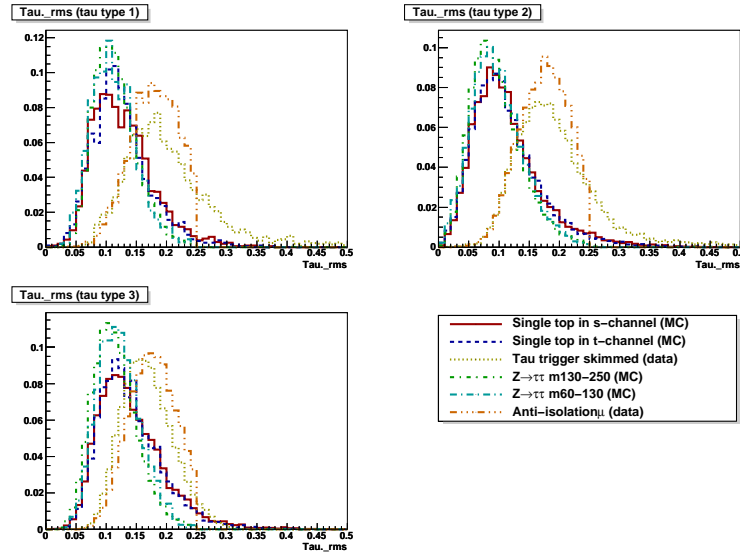
Figure C.11: Comparison between different samples used in tau ID study for the variable rms.
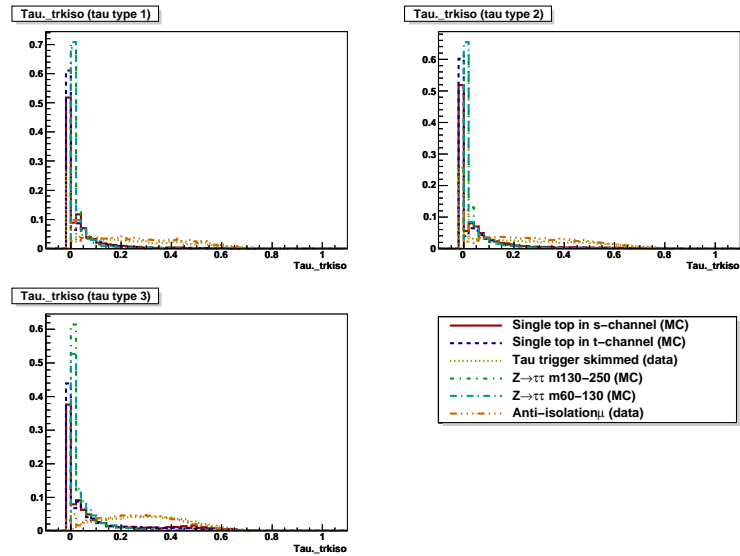
Figure C.12: Comparison between different samples used in tau ID study for the variable trkiso.

Appendix D

Discriminant Variables

D.1 Comparison Plots in the Bin: Run IIa, τ type 1+2, 1 tag, 2 jets
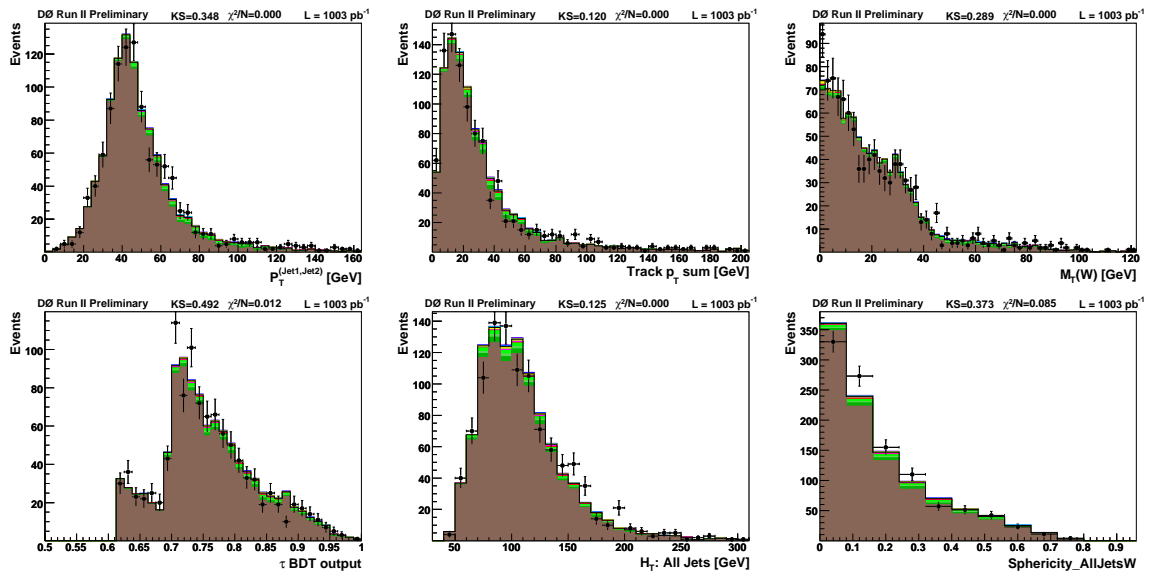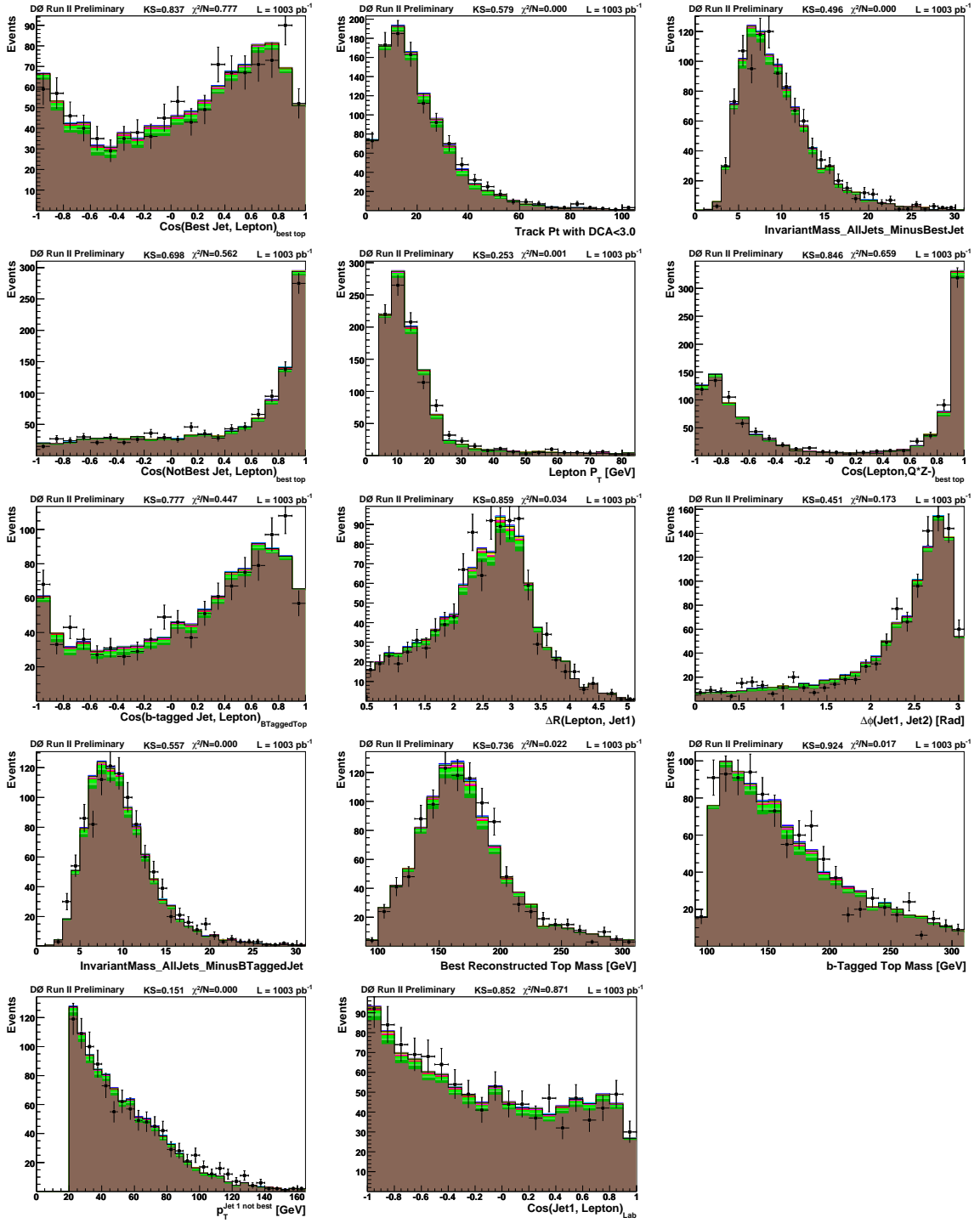
D.2 Comparison Plots in the Bin: Run IIa, τ type 1+2, 1 tag, 3 jets
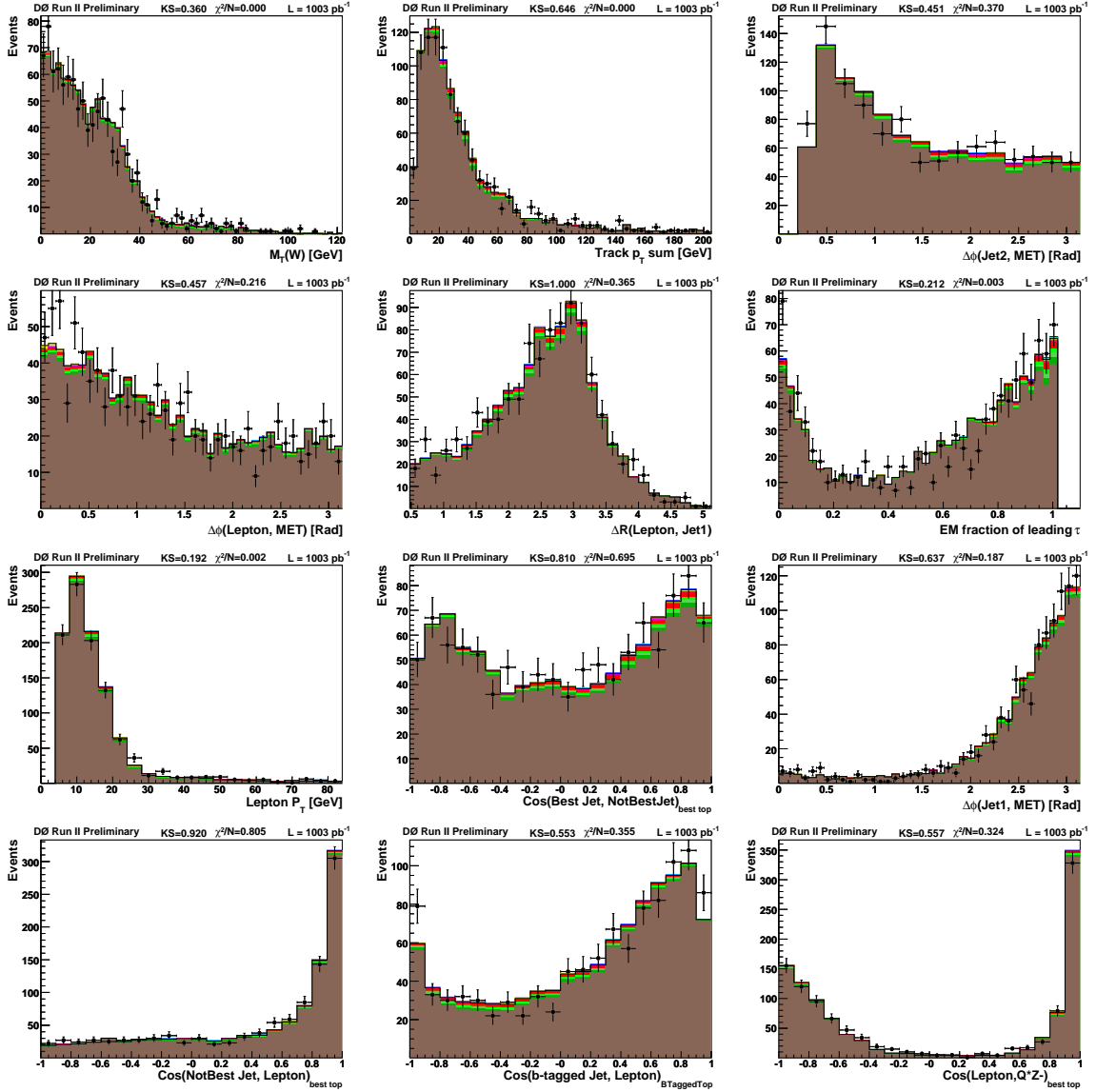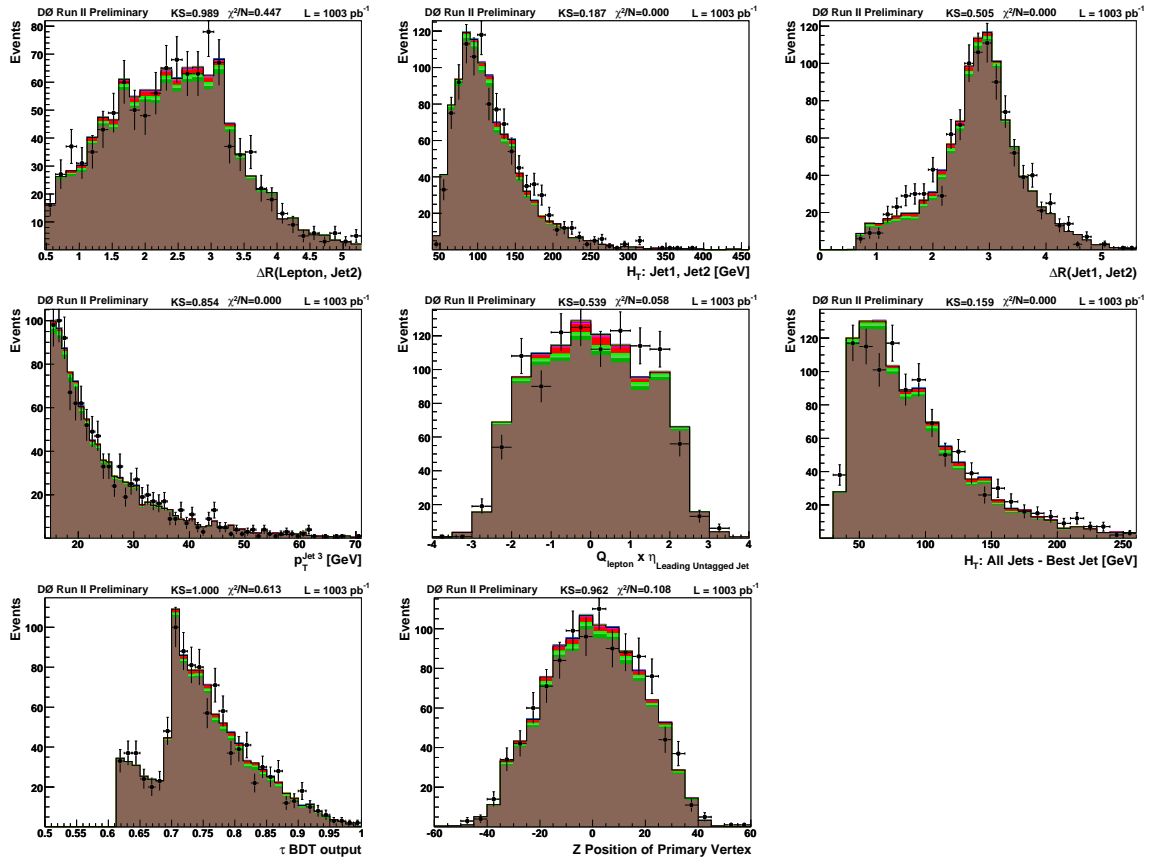
D.3 Comparison Plots in the Bin: Run IIa, τ type 1+2, 2 tags, 2 jets
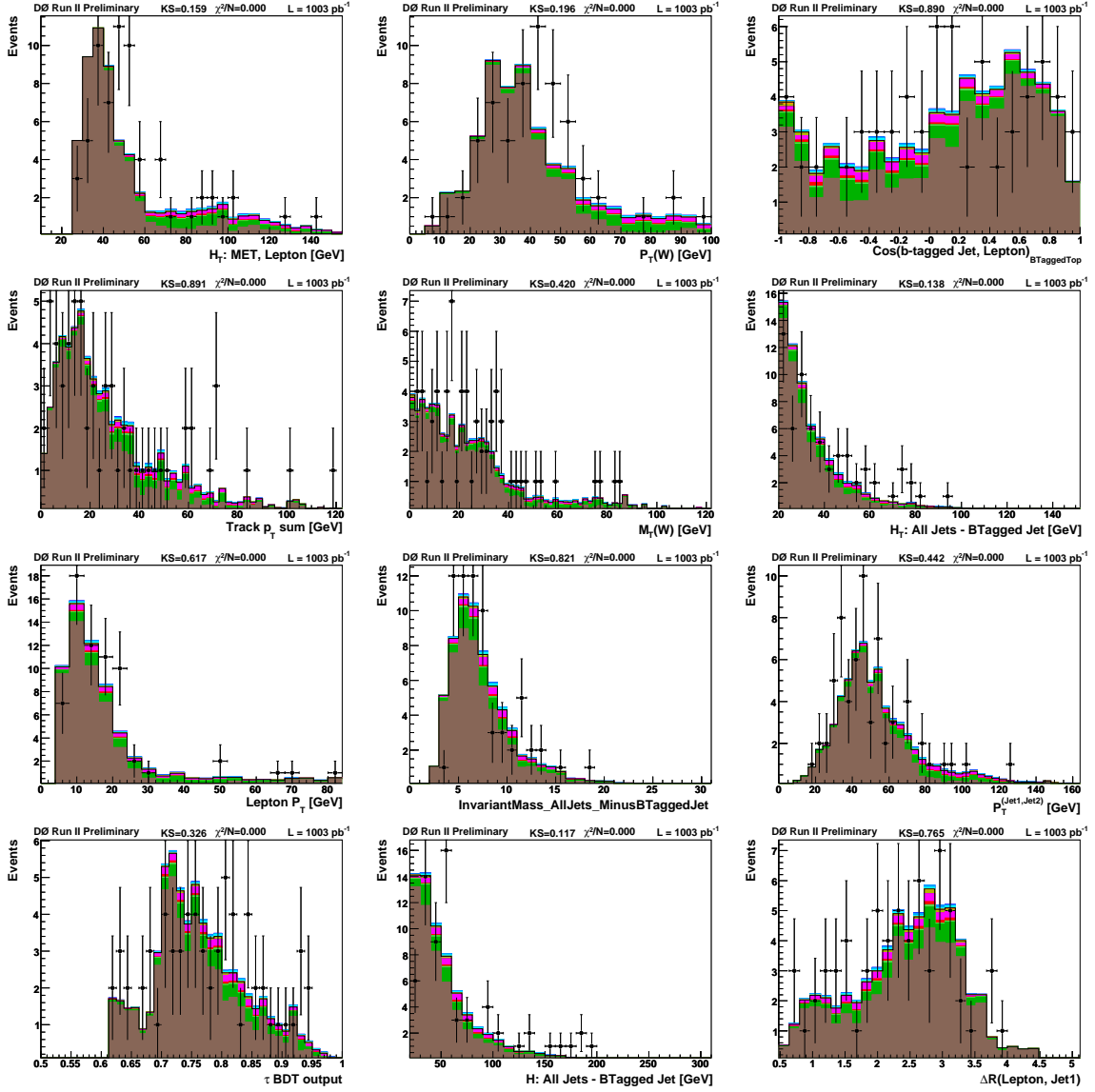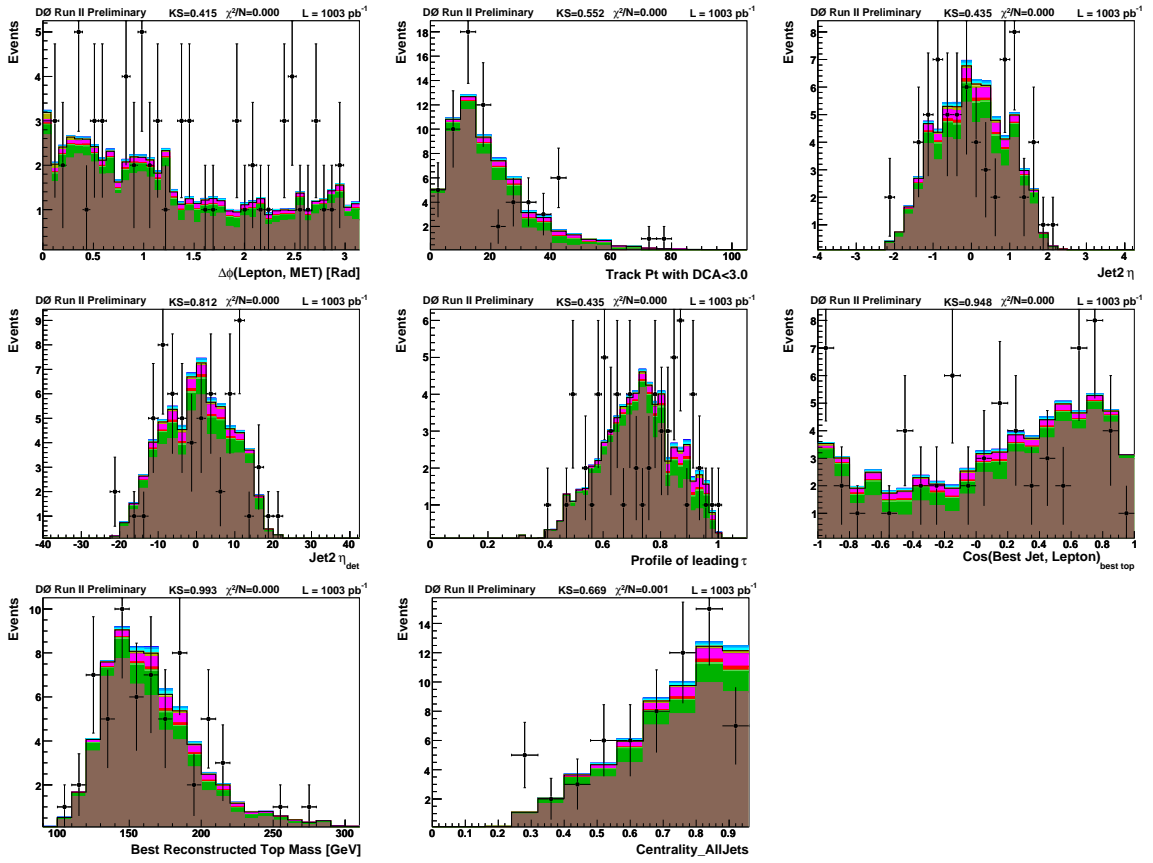
D.4 Comparison Plots in the Bin: Run IIa, τ type 1+2, 2 tags, 3 jets
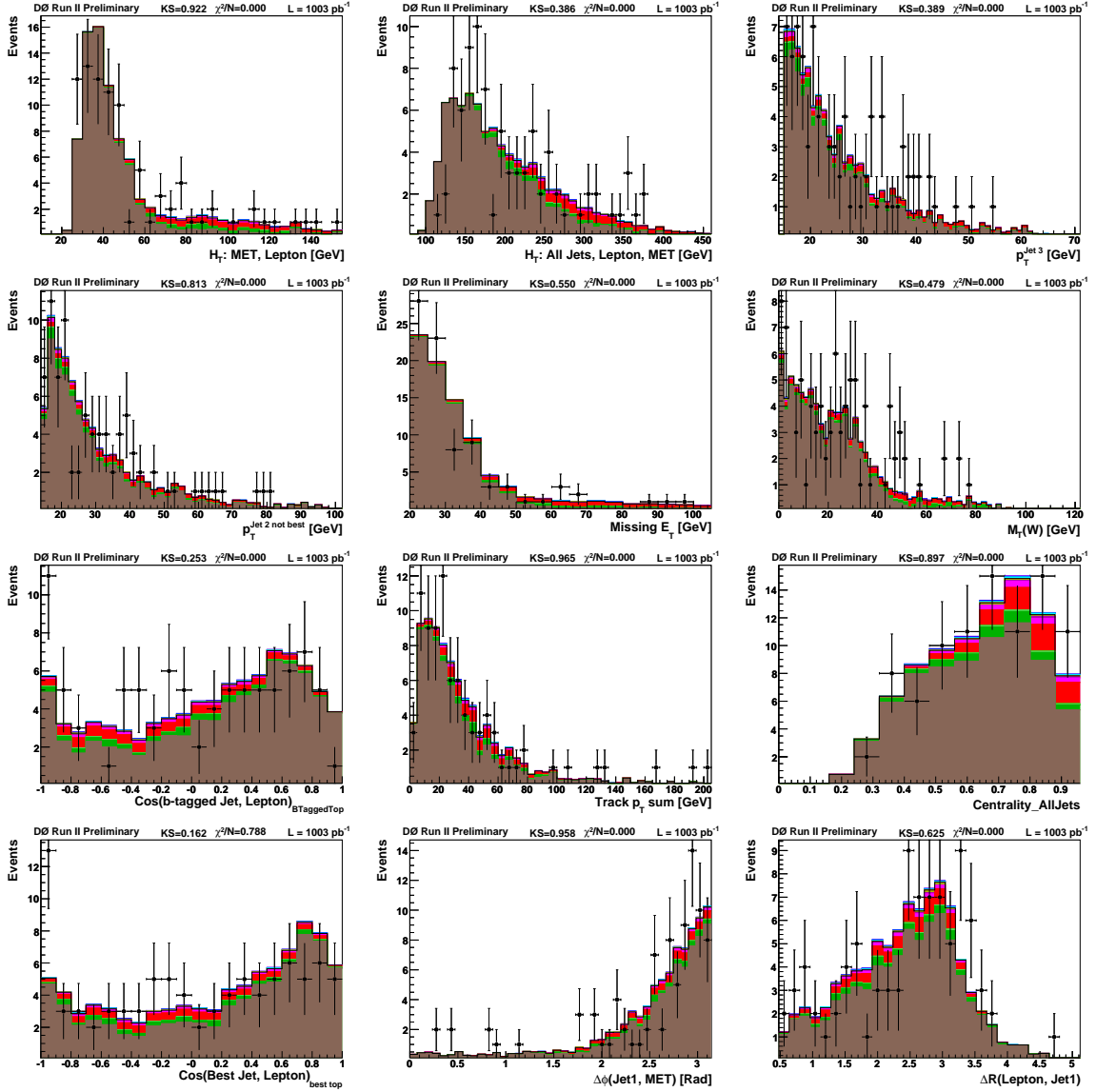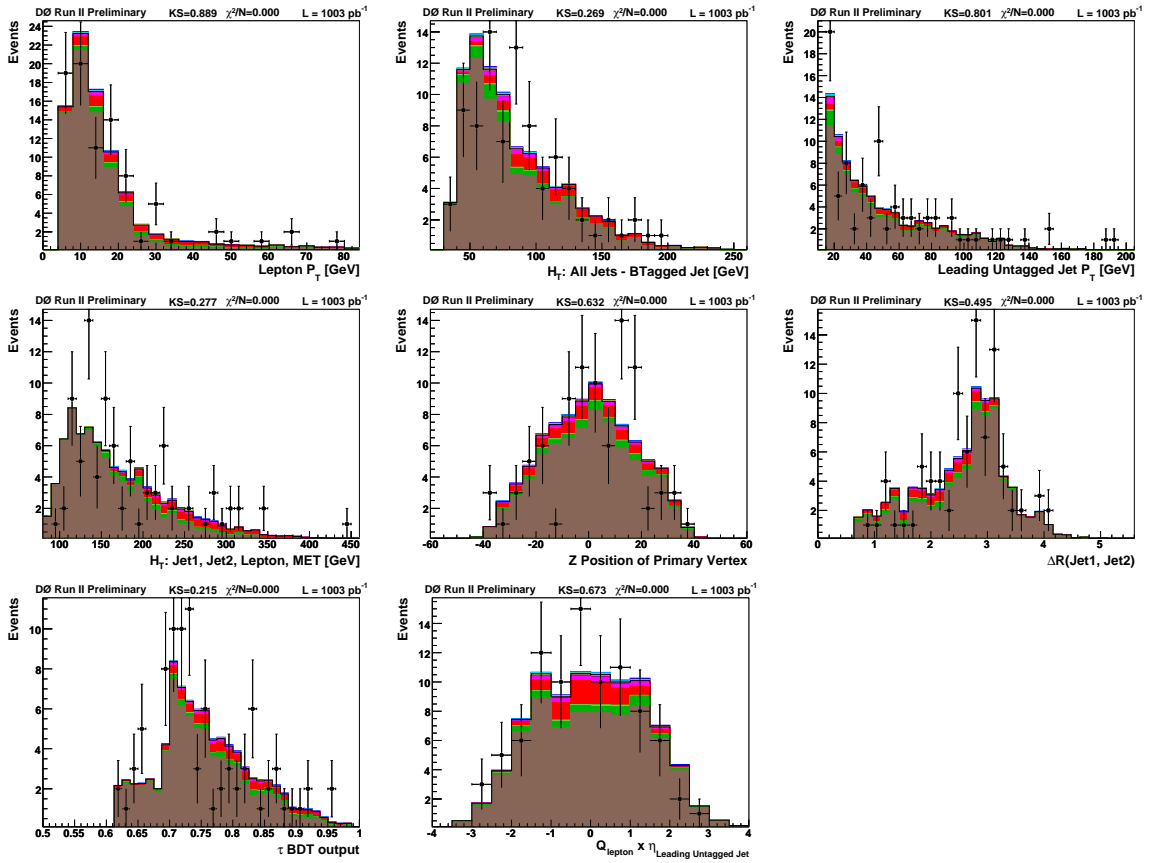
D.5 Comparison Plots in the Bin: Run IIa, τ type 3, 1 tag, 2 jets
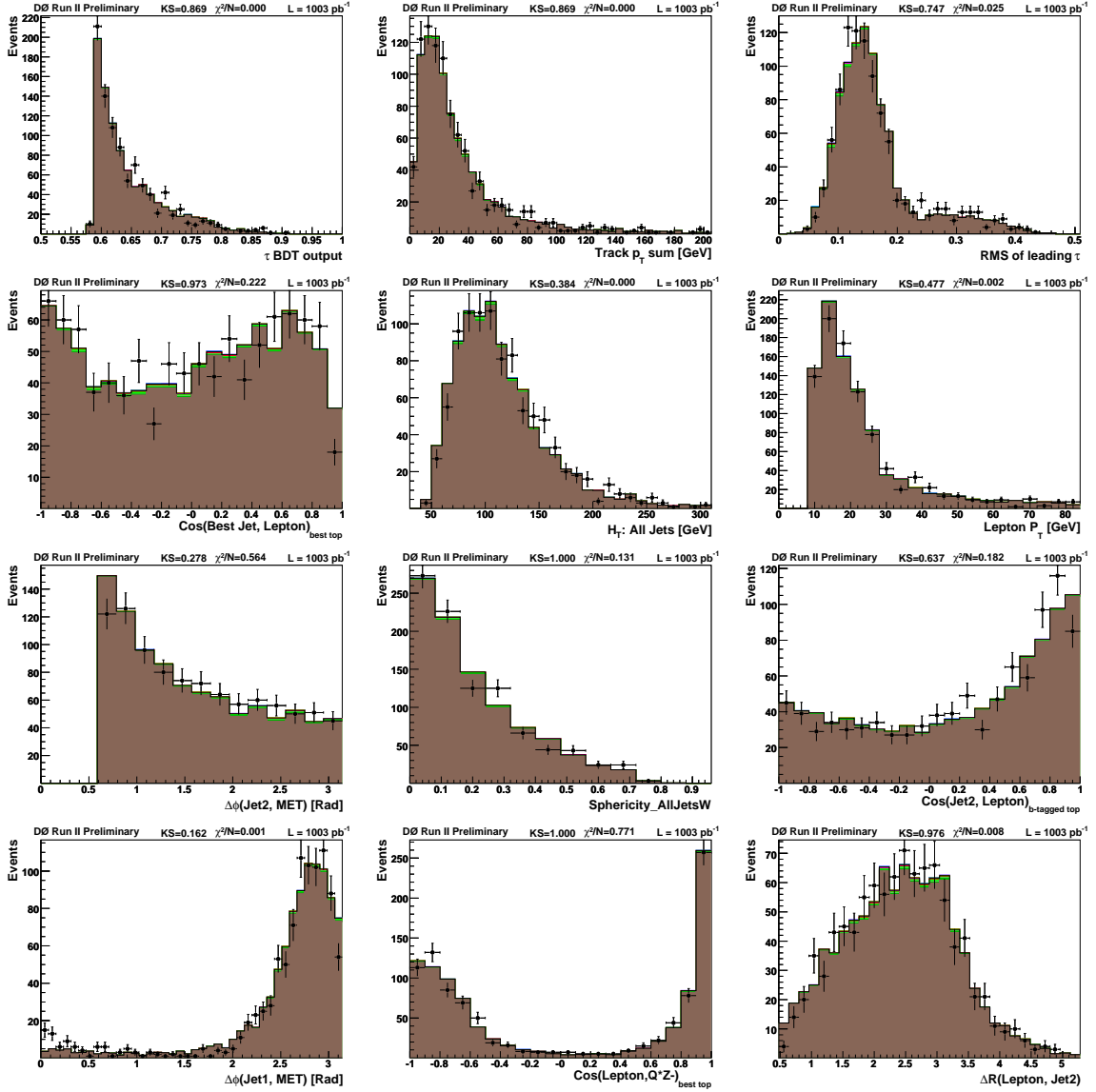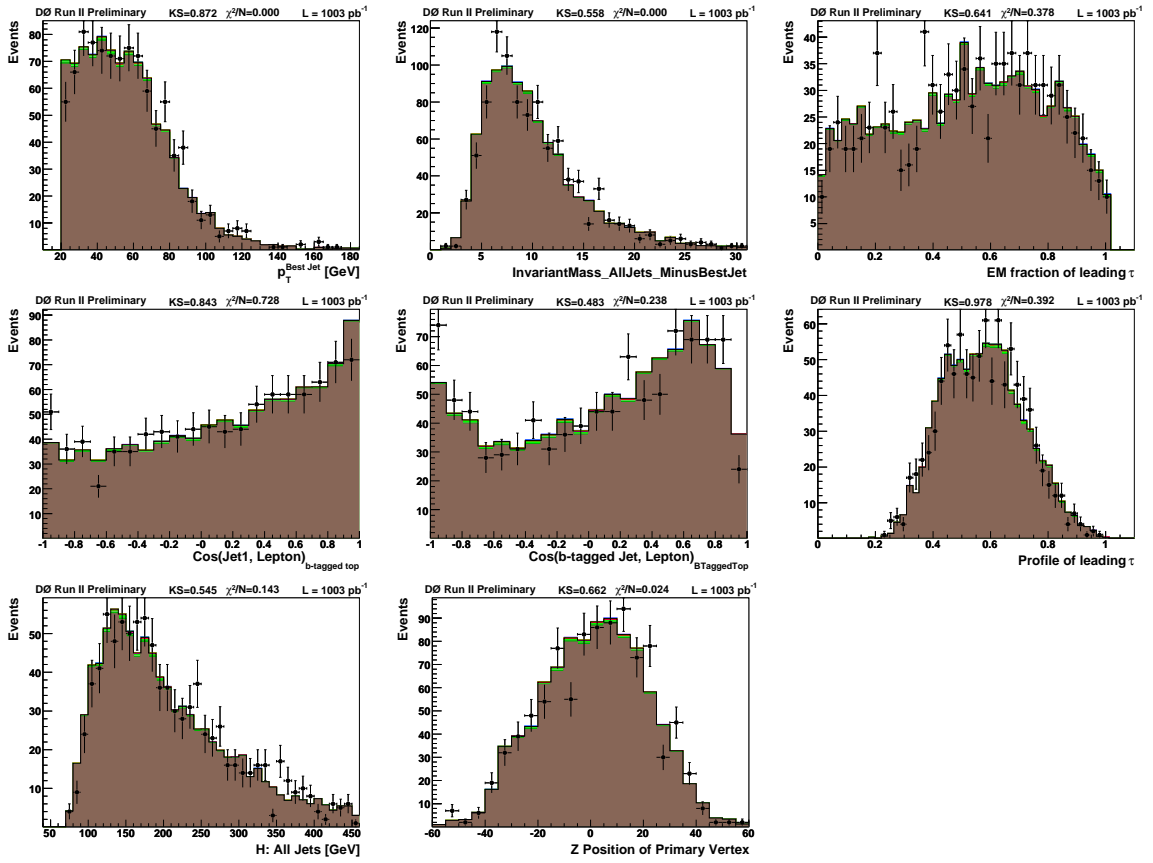
D.6 Comparison Plots in the Bin: Run IIa, τ type 3, 1 tag, 3 jets
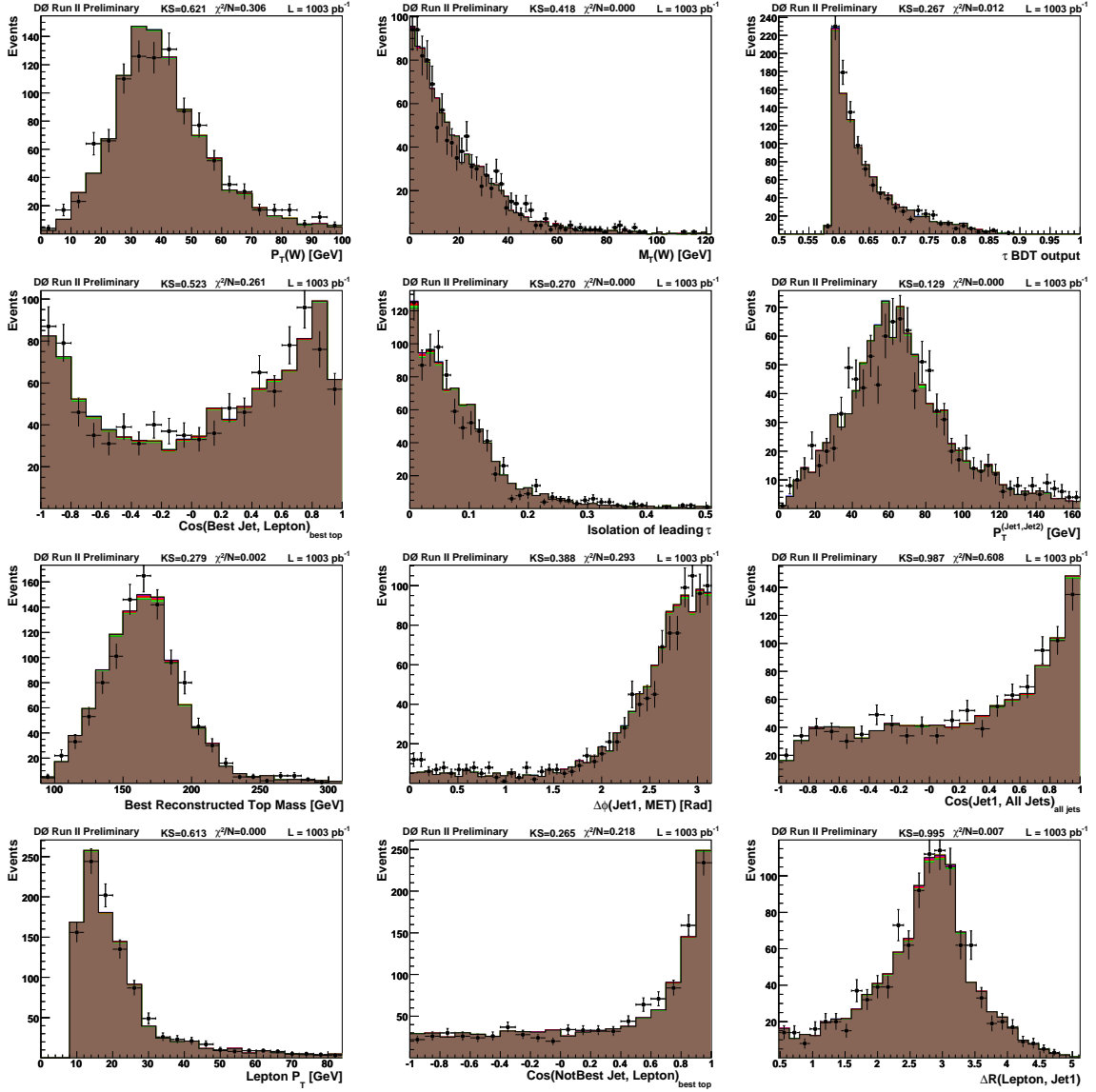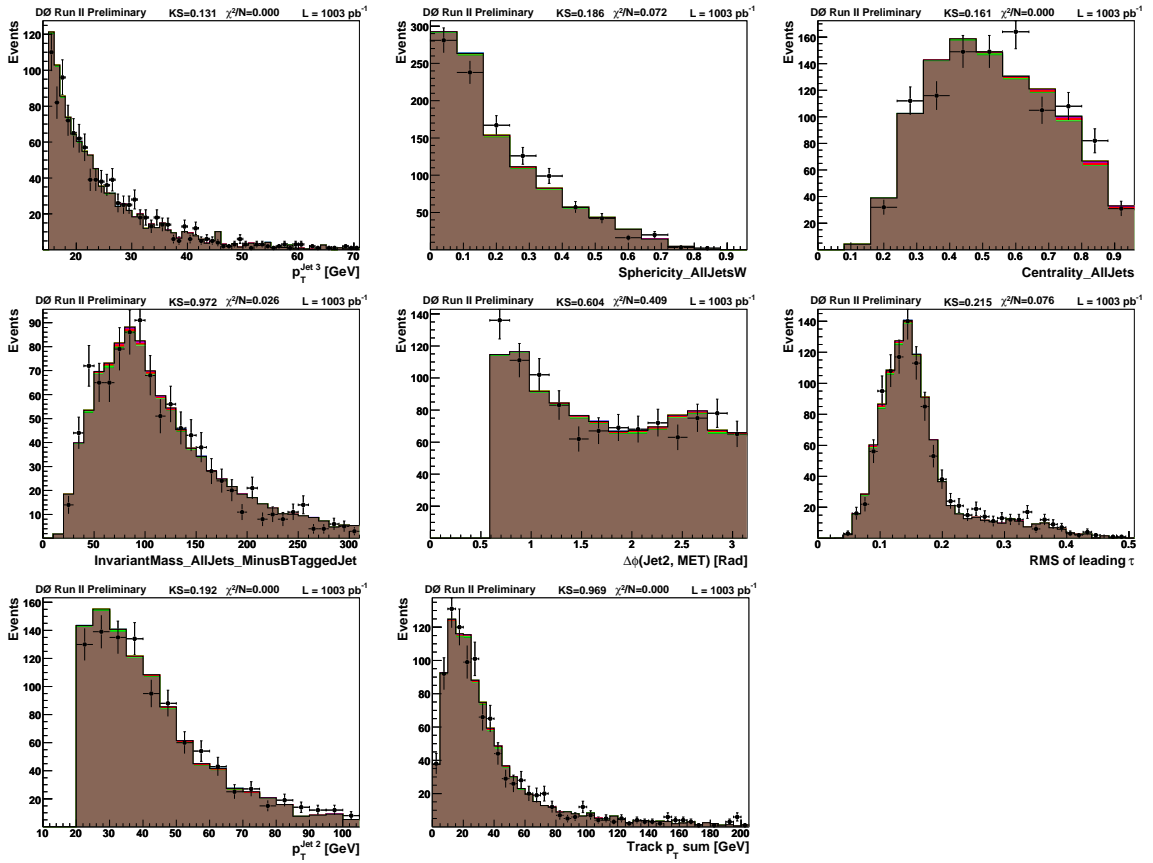
D.7 Comparison Plots in the Bin: Run IIa, τ type 3, 2 tags, 2 jets
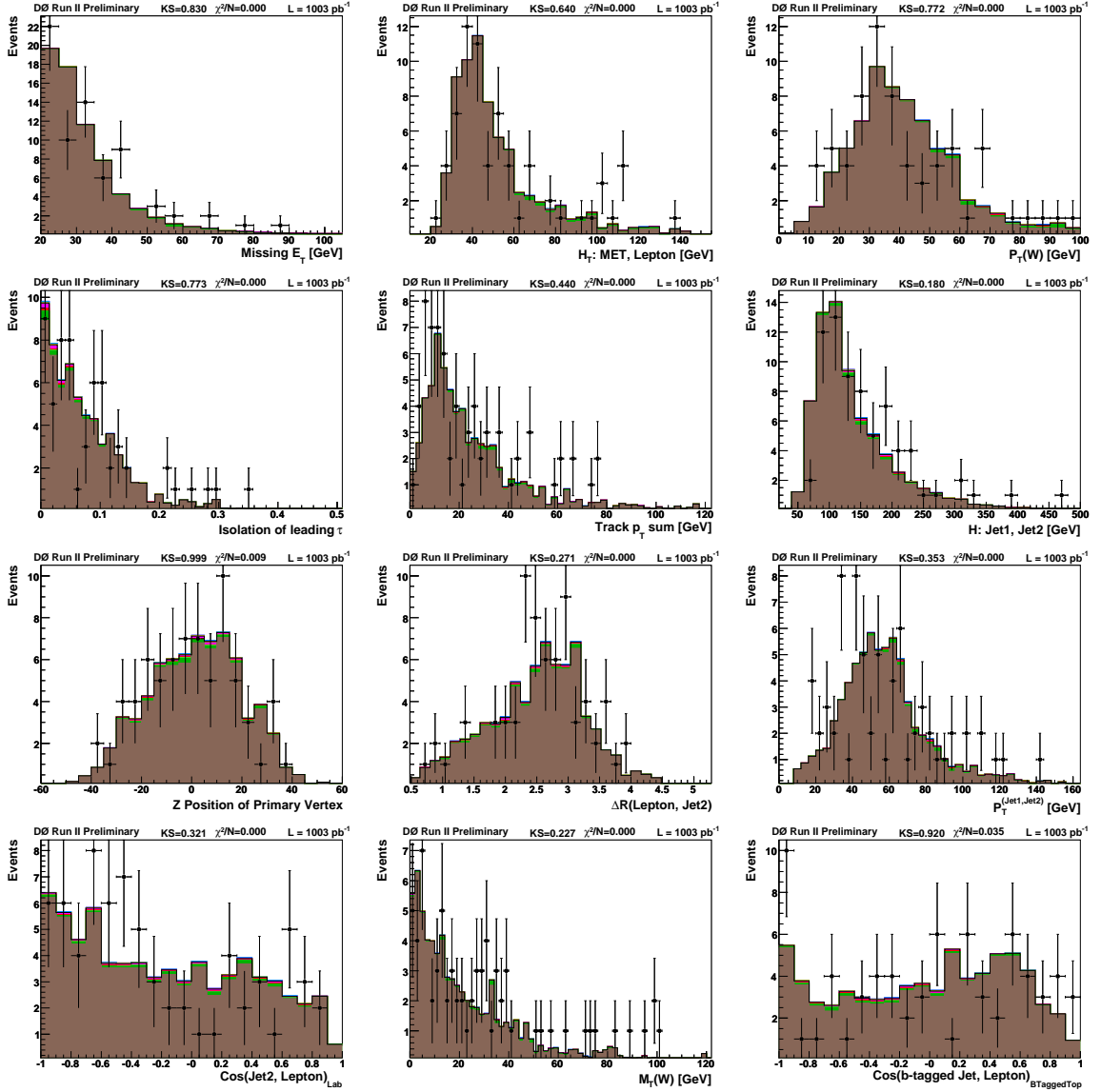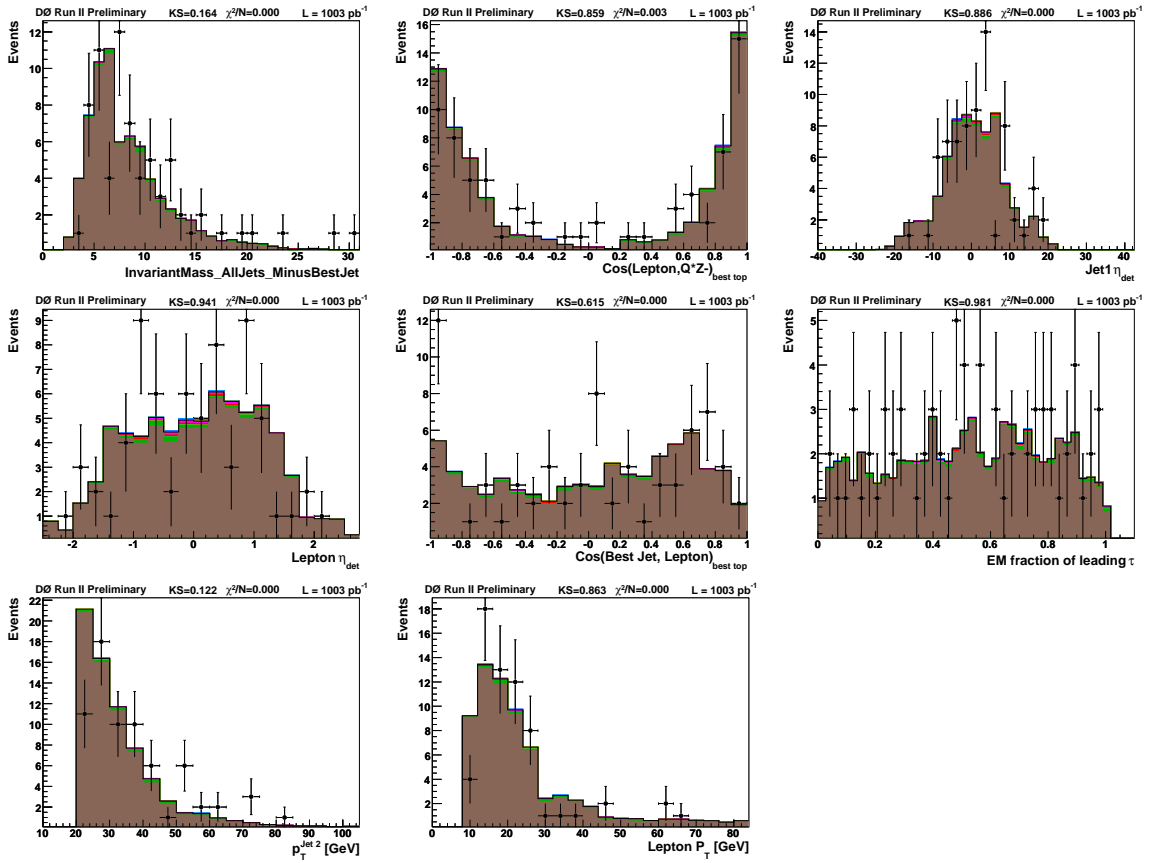
D.8 Comparison Plots in the Bin: Run IIa, τ type 3, 2 tags, 3 jets
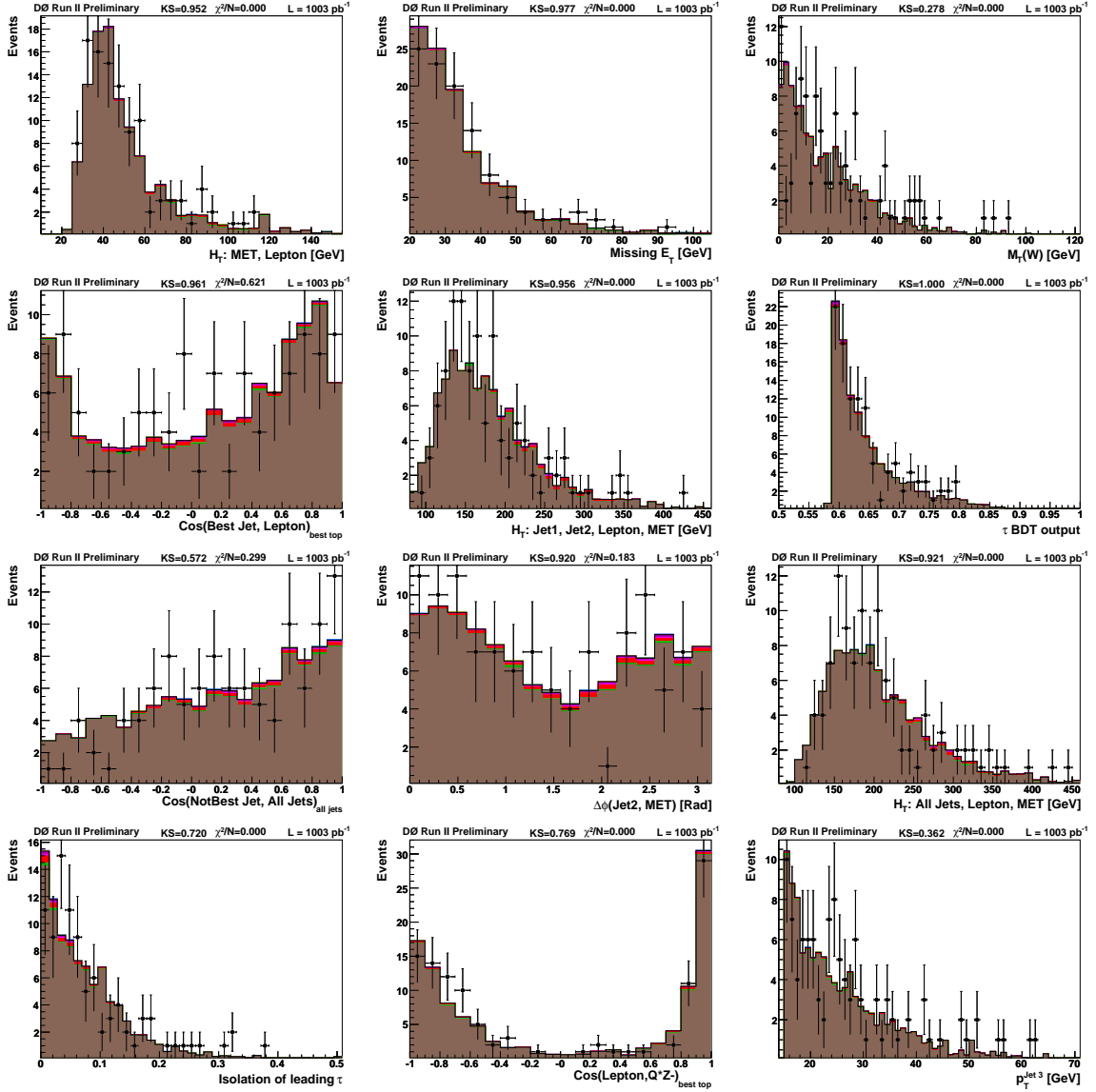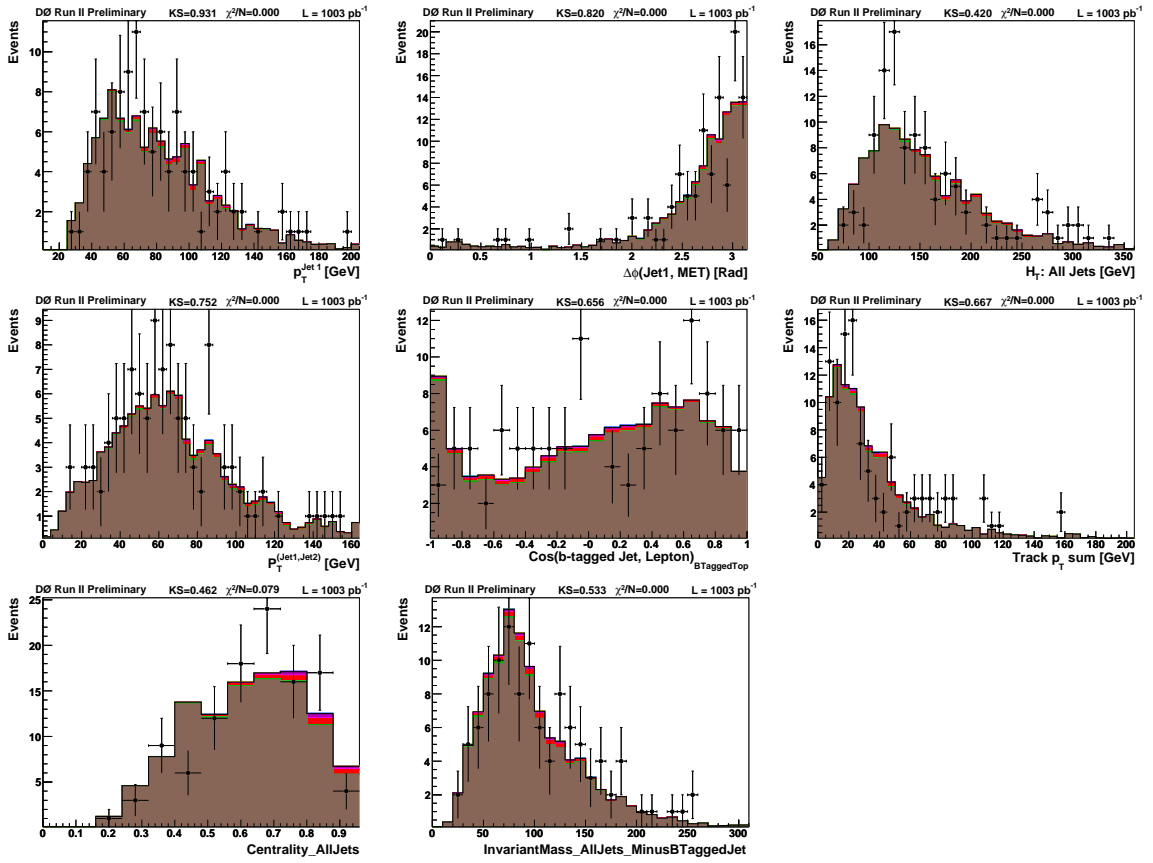
D.9 Comparison Plots in the Bin: Run IIb, τ type 1+2, 1 tag, 3 jets
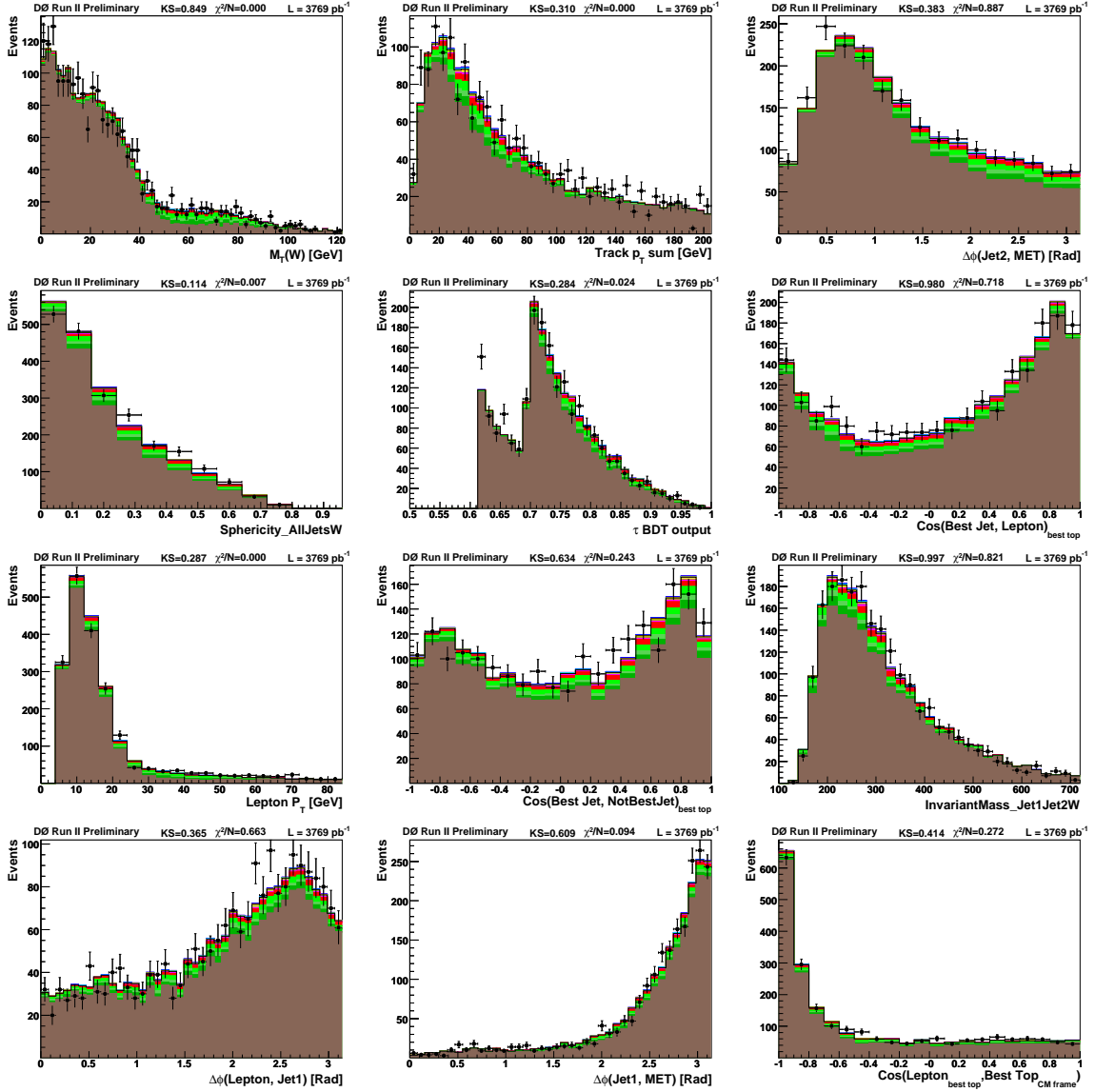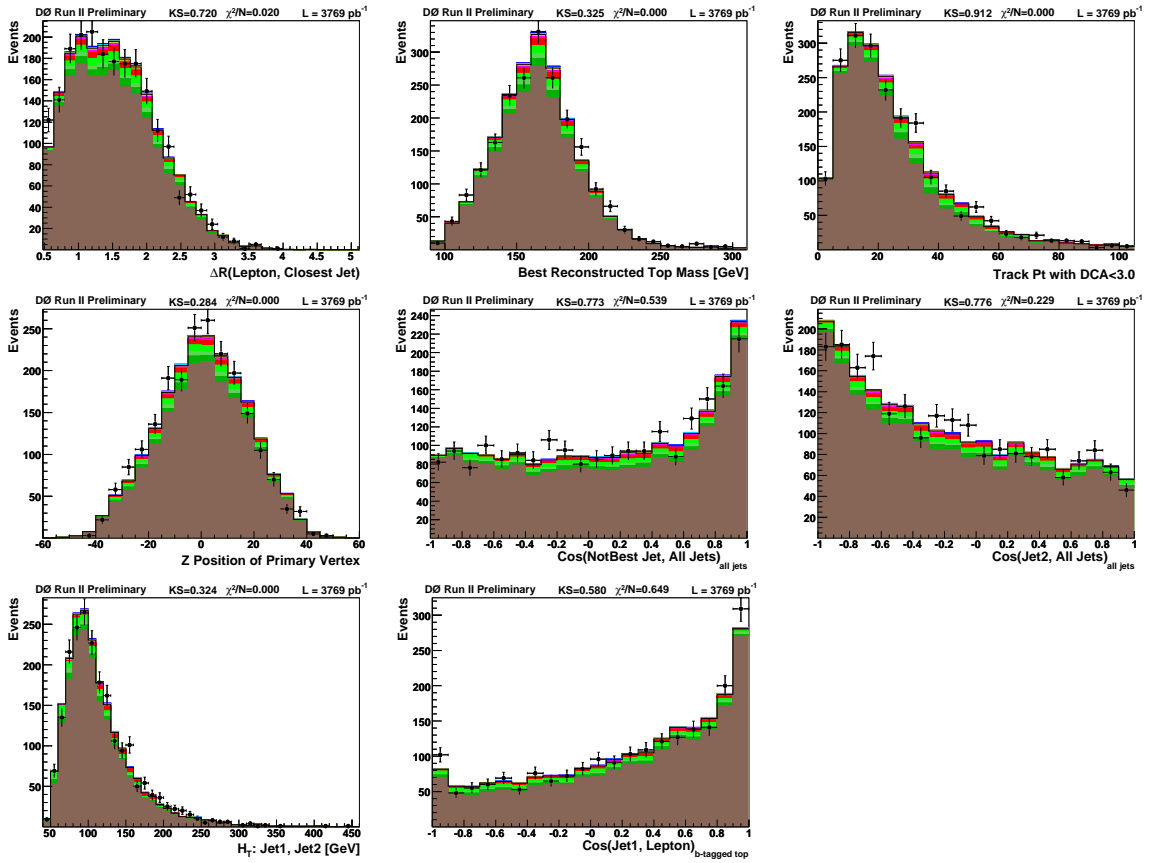
D.10 Comparison Plots in the Bin: Run IIb, τ type 1+2, 2 tags, 2 jets
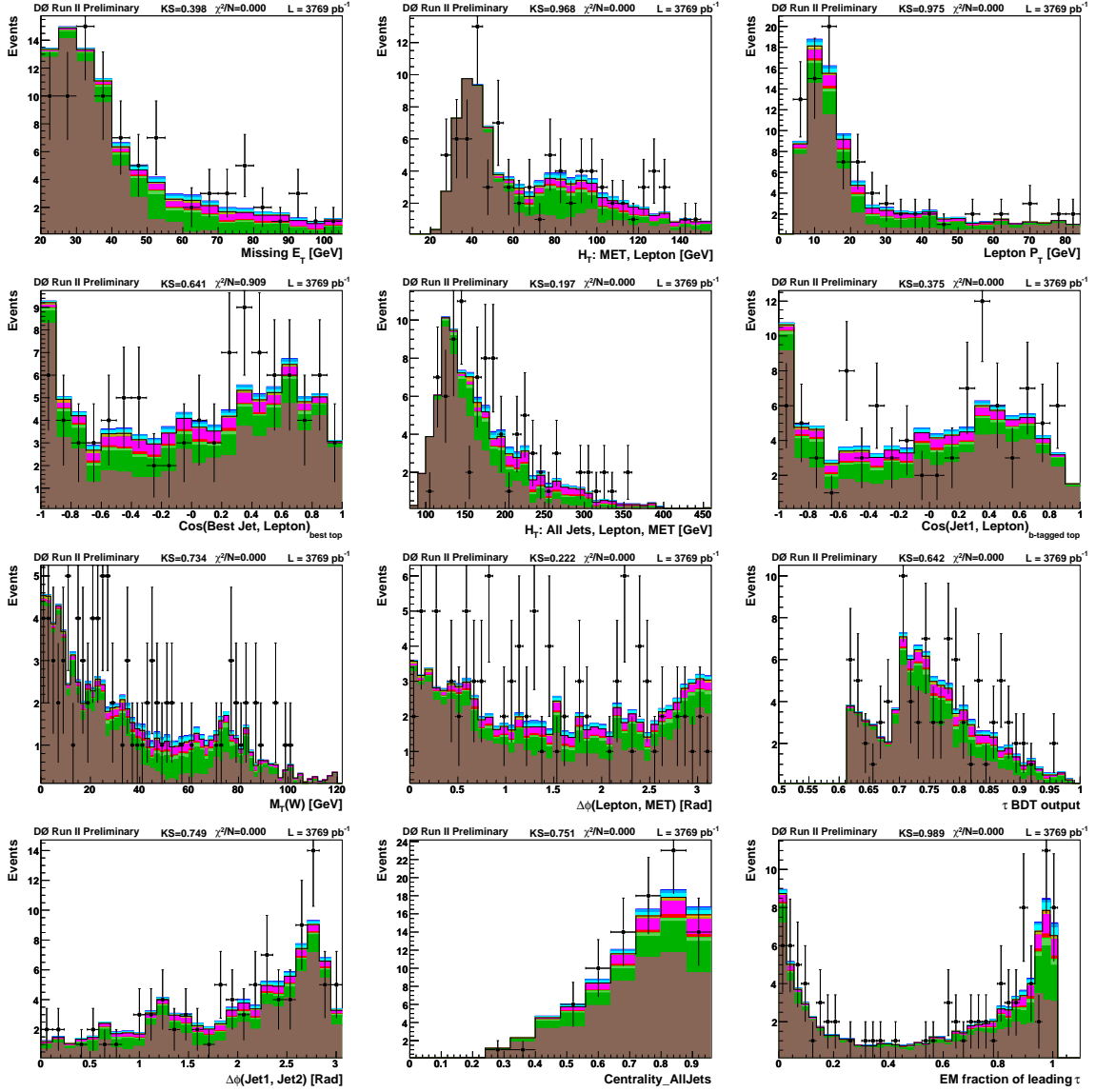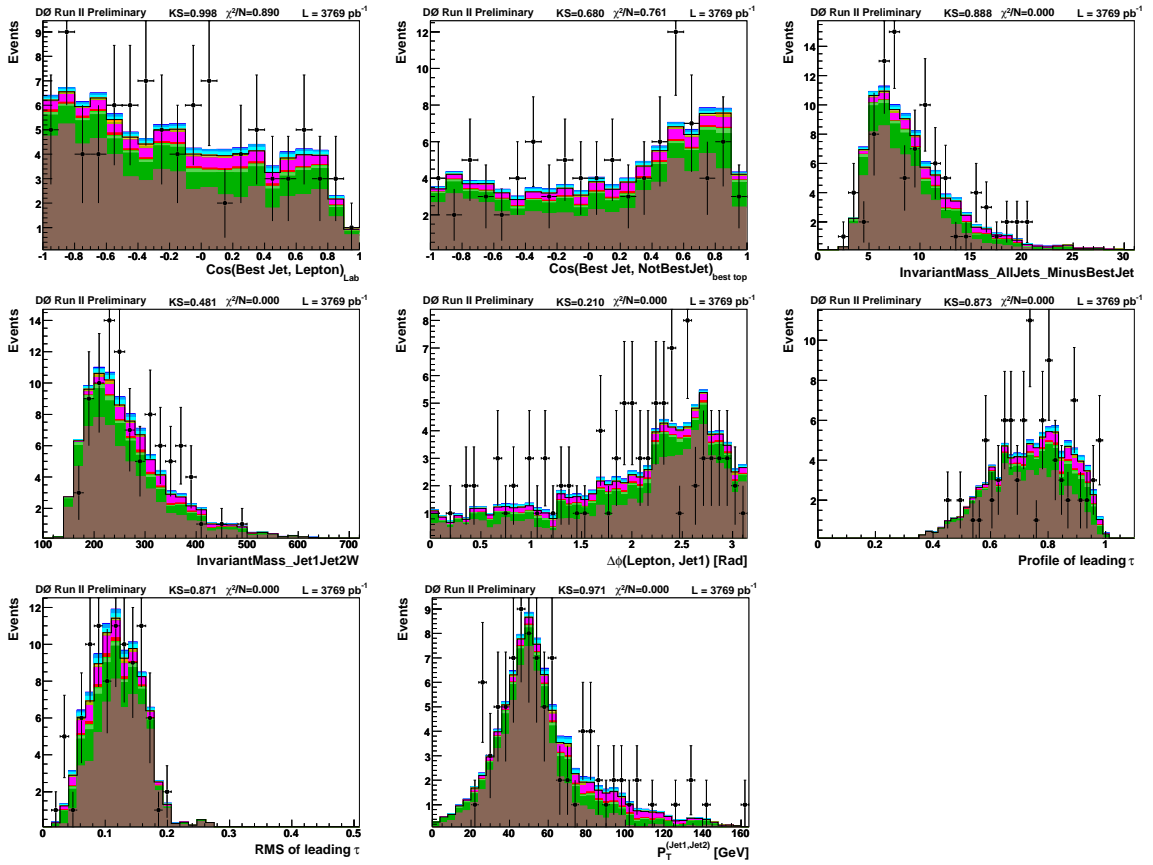
D.11 Comparison Plots in the Bin: Run IIb, τ type 1+2, 2 tags, 3 jets
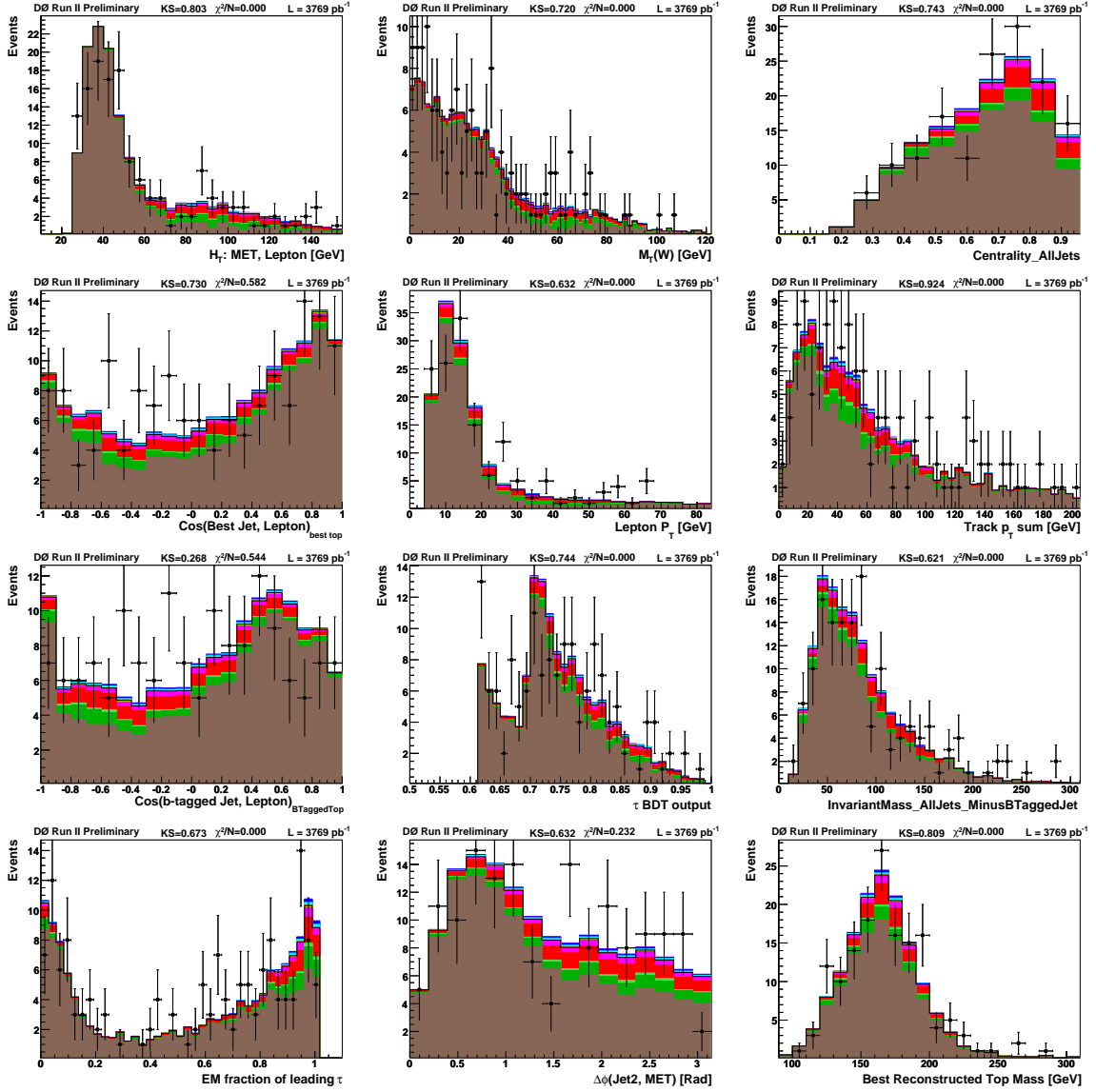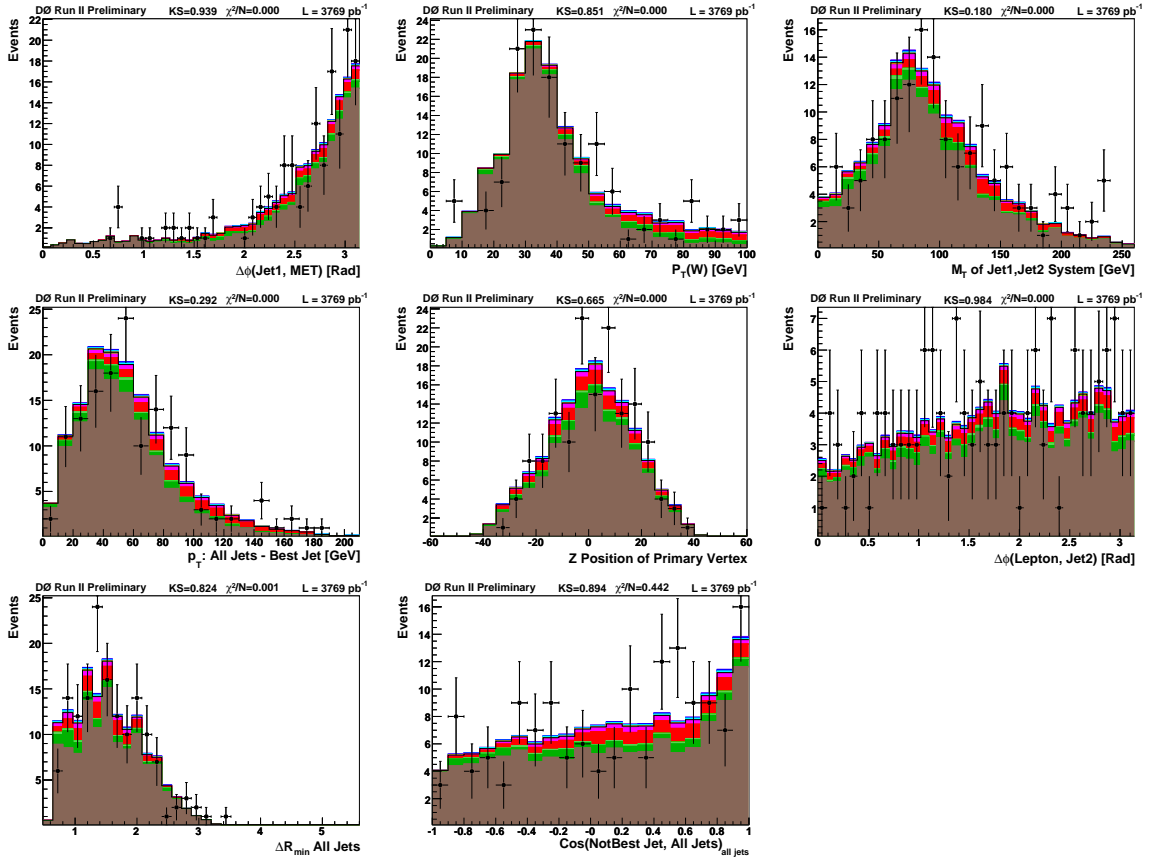
D.12 Comparison Plots in the Bin: Run IIb, τ type 3, 1 tag, 2 jets
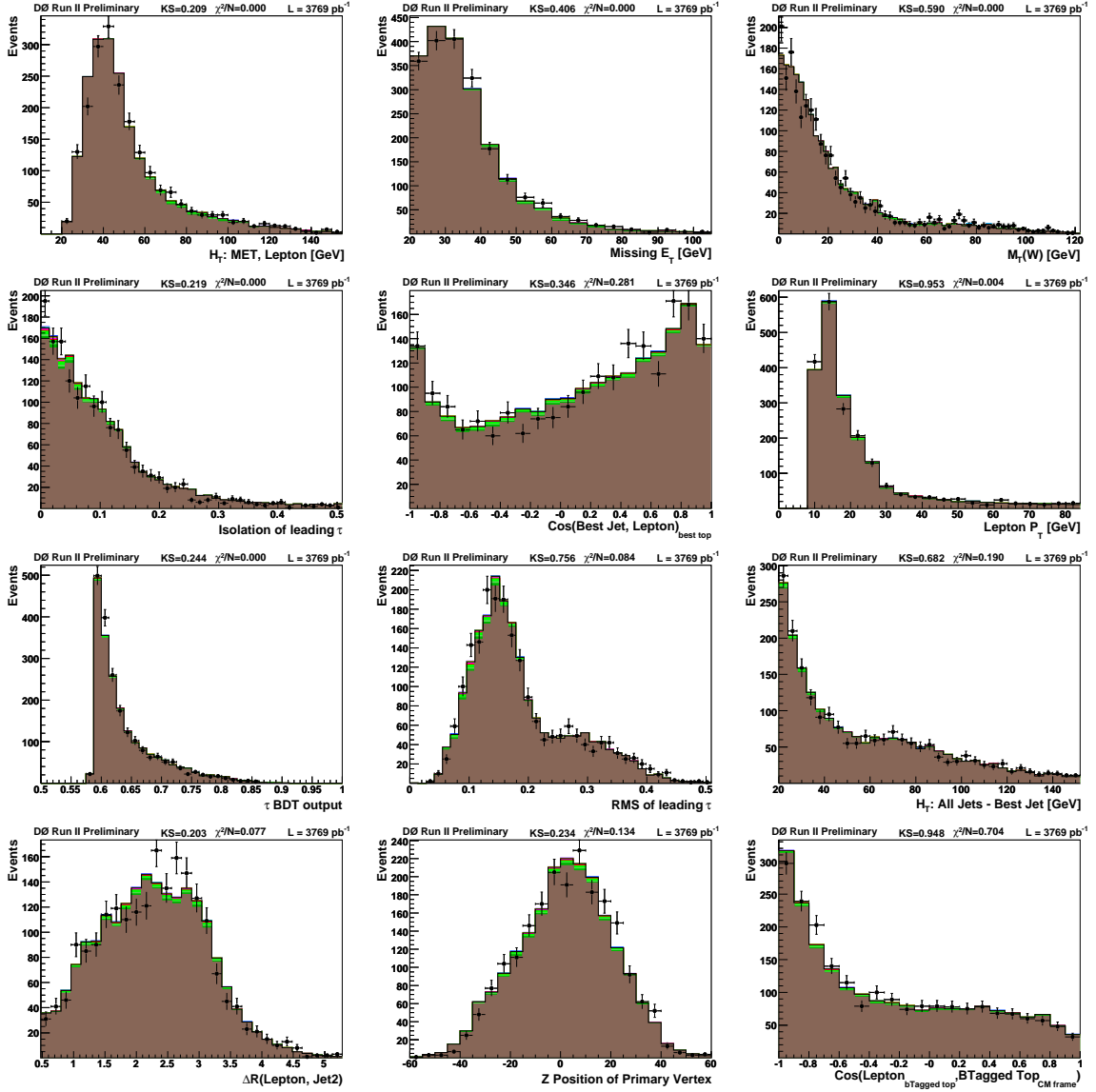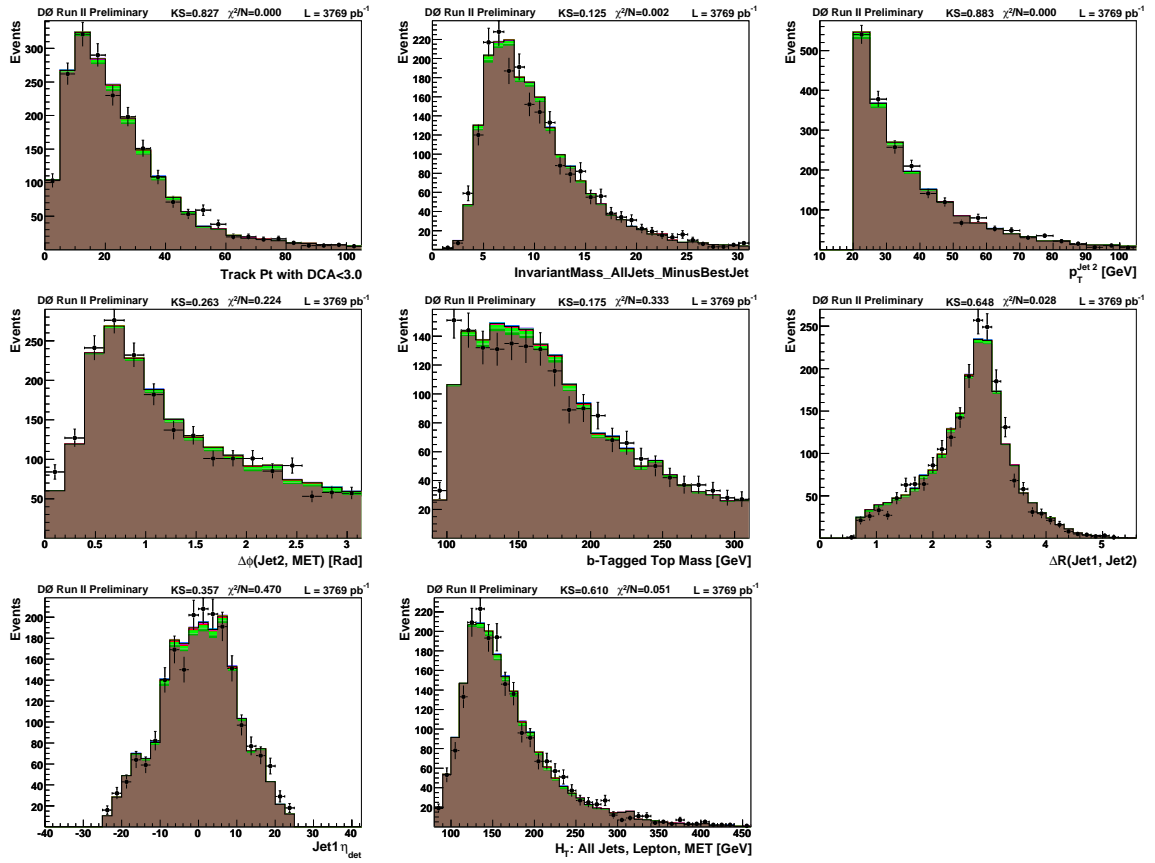
D.13 Comparison Plots in the Bin: Run IIb, τ type 3, 1 tag, 3 jets
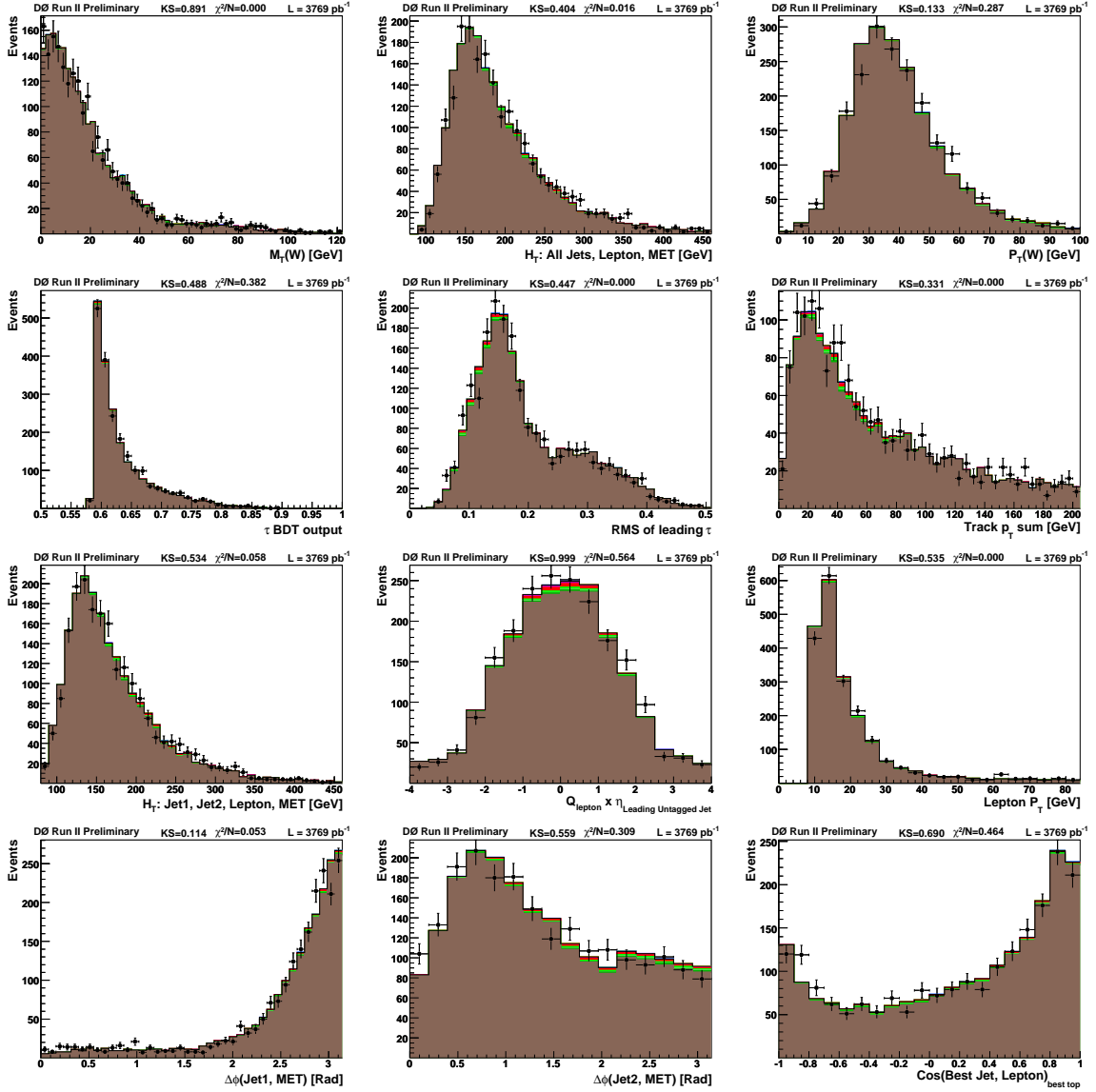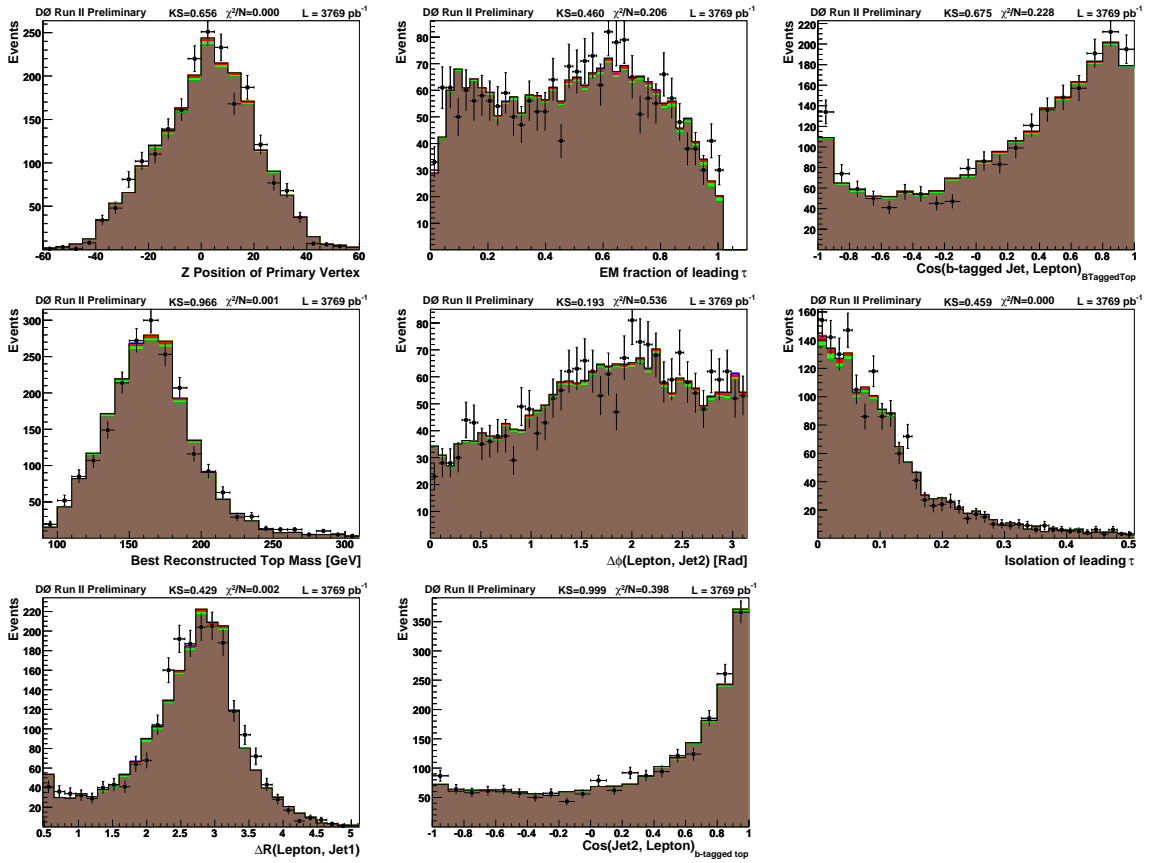
D.14 Comparison Plots in the Bin: Run IIb, τ type 3, 2 tags, 2 jets
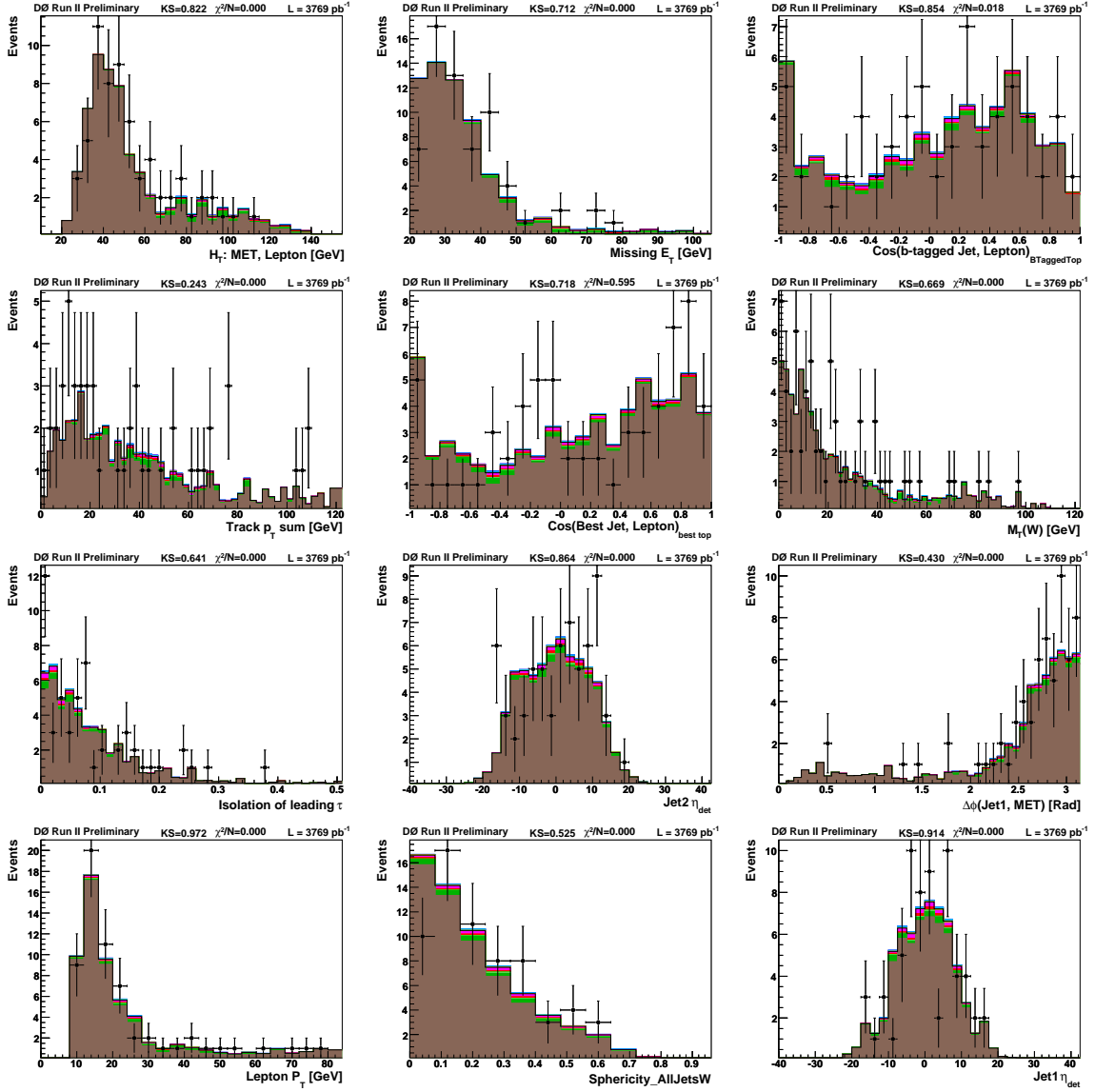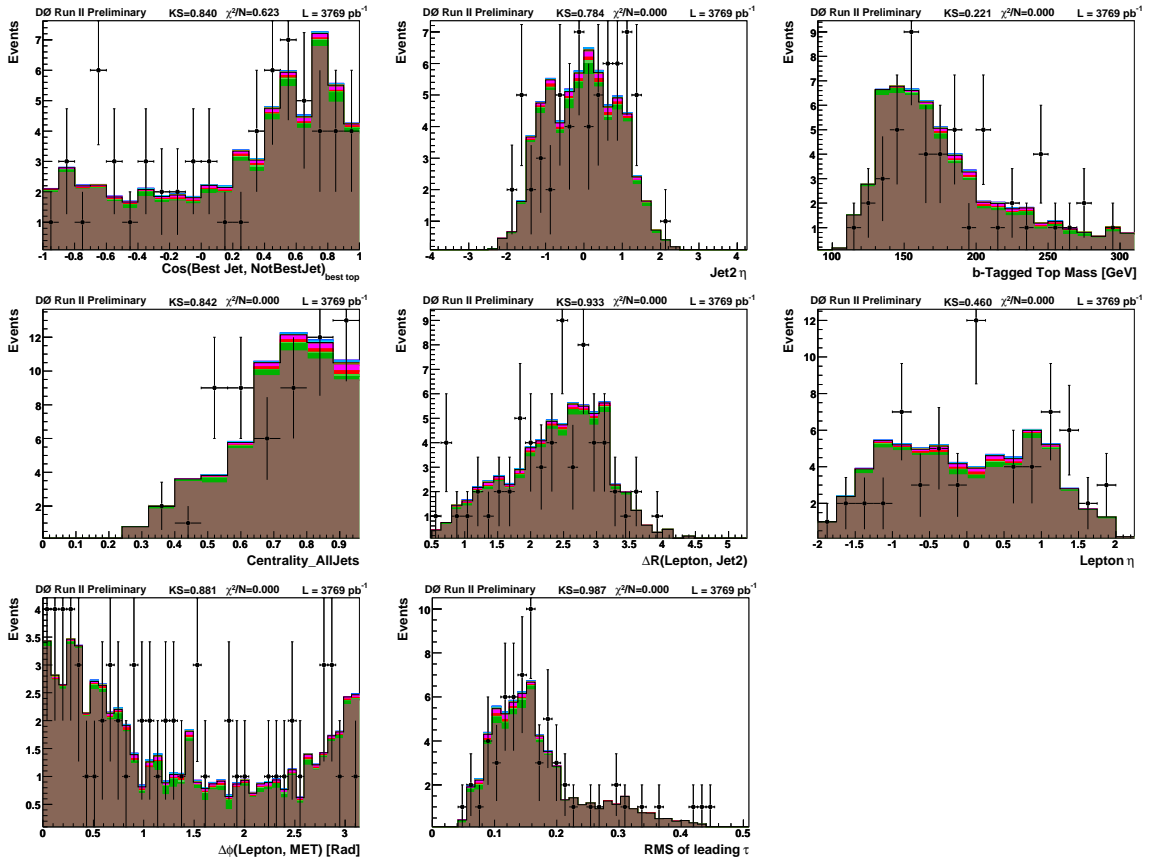
D.15 Comparison Plots in the Bin: Run IIb, τ type 3, 2 tags, 3 jets
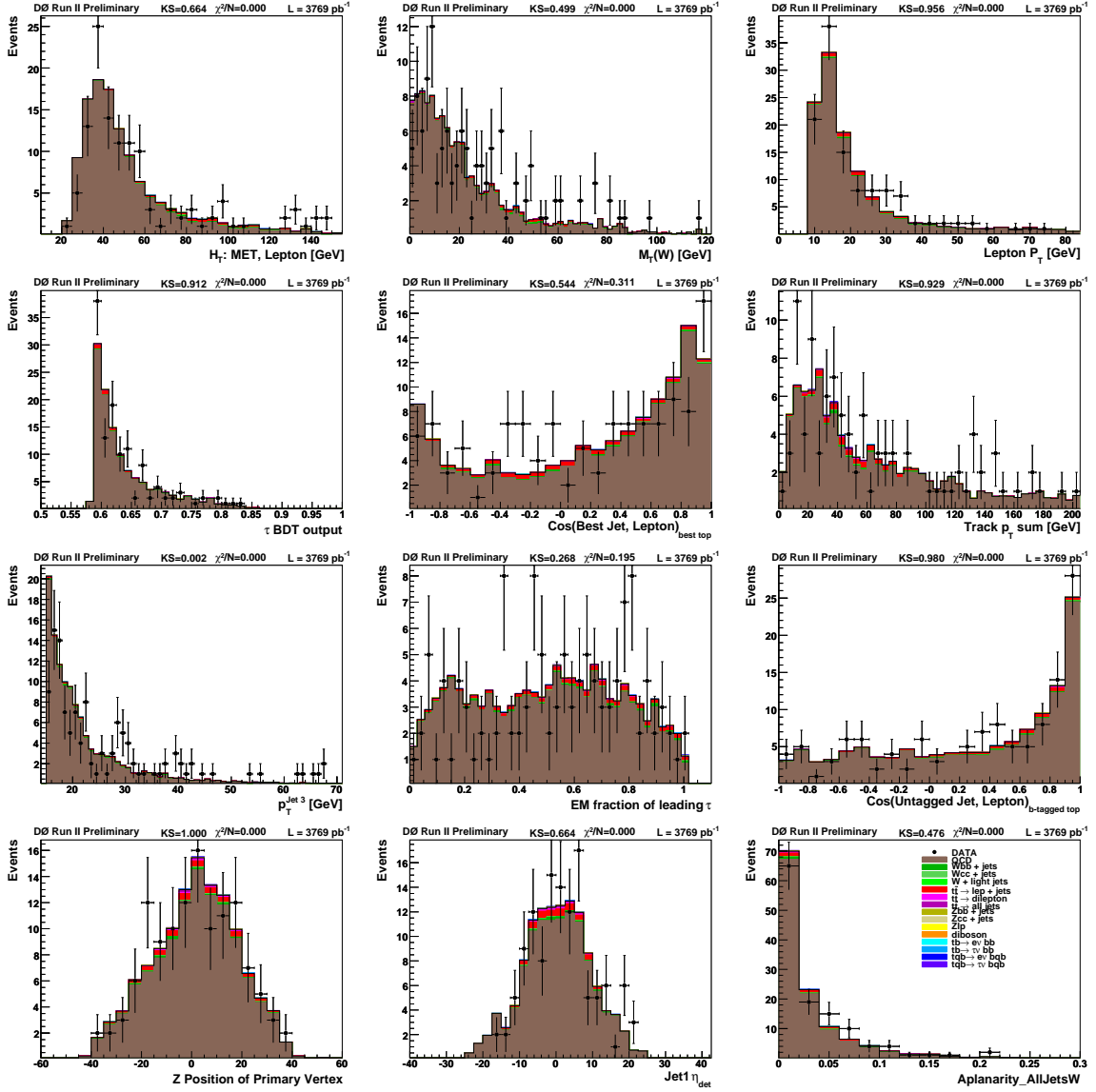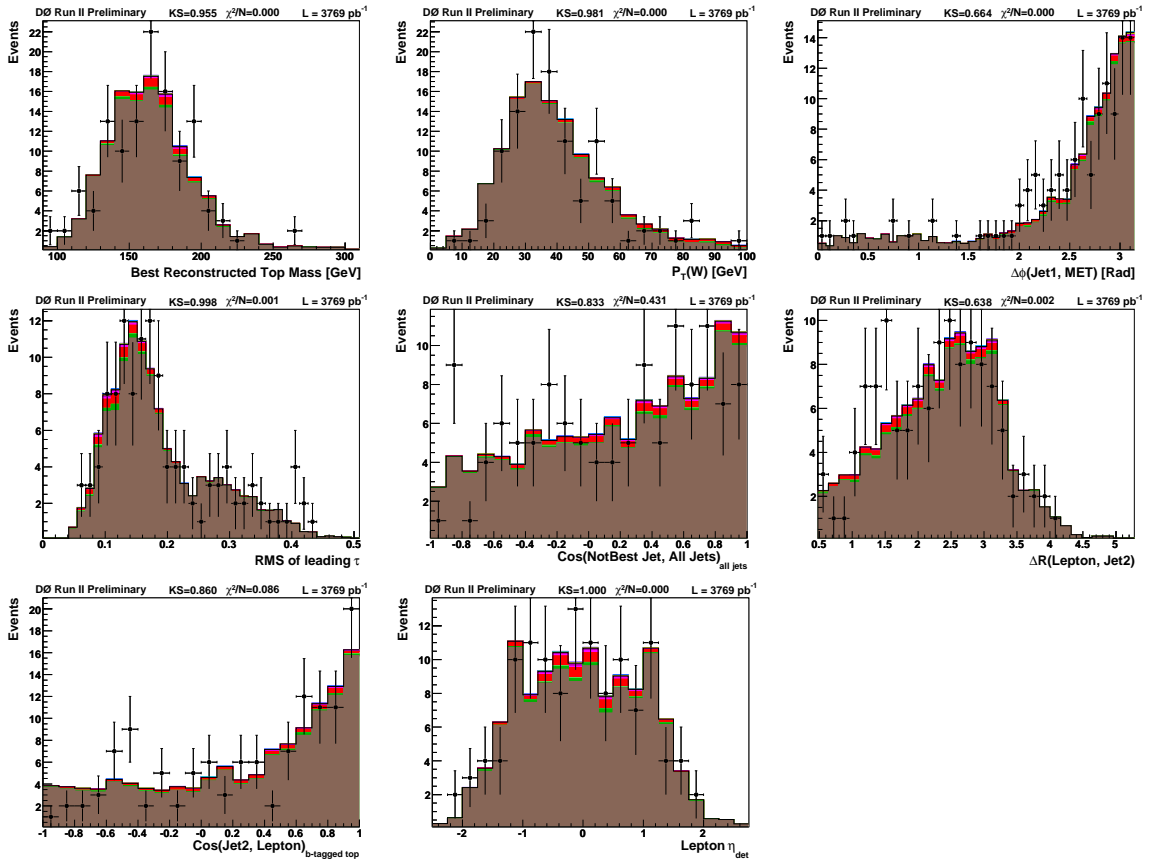
Appendix E

Event Yields

Table E.1: Yields with uncertainty after selection for Run IIa data with 1 b-tag jet. The fraction of each background component is also listed in percent.

| | tau type 1+2 | | | | tau type 3 | | | |
	2 jets		3 jets		2 jets		3 jets	
Signals								
tb+tqb	7.2±1.3		3.2±0.6		1.4±0.3		0.7±0.1	
Backgrounds								
W+2b	22.8±4.7	(6.9%)	16.6±2.6	(5.1%)	4.8±1.0	(2.0%)	3.2±0.5	(1.5%)
W+2c	11.7±2.4	(3.6%)	9.8±1.5	(3.0%)	2.3±0.5	(0.9%)	2.0±0.3	(0.9%)
W+light jet	17.7±2.7	(5.4%)	10.3±0.7	(3.2%)	3.8±0.6	(1.5%)	2.1±0.2	(1.0%)
Z+2b	2.5±0.6	(0.8%)	1.7±0.4	(0.5%)	0.9±0.2	(0.4%)	0.5±0.1	(0.2%)
Z+2c	1.4±0.3	(0.4%)	0.9±0.2	(0.3%)	0.5±0.1	(0.2%)	0.6±0.1	(0.3%)
Z+light jet	4.4±0.9	(1.3%)	2.2±0.5	(0.7%)	1.3±0.3	(0.5%)	0.6±0.1	(0.3%)
Dibosons	3.1±0.5	(0.9%)	1.7±0.3	(0.5%)	0.8±0.1	(0.3%)	0.6±0.1	(0.3%)
$t\bar{t}$	10.9±2.2	(3.3%)	23.9±4.7	(7.3%)	3.8±0.8	(1.6%)	9.6±1.9	(4.3%)
Multijets	254.4±11.8	(77.3%)	259.5±10.5	(79.4%)	228.8±9.0	(92.6%)	203.1±7.6	(91.4%)
Background Sum	329.0±15.6		326.7±12.2		247.1±9.0		222.3±7.6	
Data	325		379		264		240	

Table E.2: Yields with uncertainty after selection for Run IIa data with 2 *b*-tagged jets. The fraction of each background component is also listed in percent.

	tau type 1+2				tau type 3			
	2 jets		3 jets		2 jets		3 jets	
Signals								
tb+tqb	1.43±0.31		0.92±0.21		0.28±0.06		0.19±0.04	
Backgrounds								
W+2b	6.36±1.31	(33.7%)	3.88±0.60	(14.9%)	1.36±0.28	(10.1%)	0.66±0.10	(2.6%)
W+2c	0.48±0.10	(2.5%)	0.50±0.08	(1.9%)	0.10±0.02	(0.7%)	0.06±0.01	(0.2%)
W+light jet	0.08±0.01	(0.4%)	0.06±0.00	(0.2%)	0.02±0.00	(0.2%)	0.02±0.00	(0.1%)
Z+2b	0.56±0.14	(3.0%)	0.38±0.10	(1.5%)	0.20±0.05	(1.5%)	0.11±0.03	(0.4%)
Z+2c	0.06±0.02	(0.3%)	0.08±0.02	(0.3%)	0.02±0.00	(0.1%)	0.04±0.01	(0.2%)
Z+light jet	0.03±0.01	(0.2%)	0.02±0.00	(0.1%)	0.01±0.00	(0.1%)	0.00±0.00	(0.0%)
Dibosons	0.39±0.09	(2.1%)	0.19±0.04	(0.7%)	0.07±0.02	(0.5%)	0.05±0.01	(0.2%)
$t\bar{t}$	3.80±0.90	(20.1%)	8.33±1.94	(32.0%)	1.25±0.31	(9.3%)	3.04±0.72	(11.9%)
Multijets	7.11±1.14	(37.7%)	12.58±1.76	(48.3%)	10.42±1.42	(77.5%)	21.53±2.51	(84.4%)
Background Sum	18.87±2.11		26.01±2.68		13.44±1.43		25.51±2.43	
Data	29		30		16		22	

Table E.3: Yields with uncertainty after selection for Run IIb data with 1 *b*-tag jet. The fraction of each background component is also listed in percent.

	tau type 1+2				tau type 3			
	2 jets		3 jets		2 jets		3 jets	
Signals								
tb+tqb	29.2±5.4		13.4±2.6		5.1±1.0		2.3±0.5	
Backgrounds								
W+2b	102.6±20.4	(13.2%)	61.0±10.9	(8.6%)	17.8±3.5	(4.0%)	10.8±1.9	(2.8%)
W+2c	60.3±12.0	(7.8%)	38.0±6.8	(5.4%)	11.0±2.2	(2.5%)	7.5±1.3	(2.0%)
W+light jet	119.8±16.9	(15.4%)	64.8±7.2	(9.2%)	23.0±3.2	(5.1%)	12.2±1.4	(3.2%)
Z+2b	6.3±1.5	(0.8%)	3.4±0.8	(0.5%)	2.8±0.7	(0.6%)	1.8±0.4	(0.5%)
Z+2c	2.8±0.7	(0.4%)	2.9±0.7	(0.4%)	1.7±0.4	(0.4%)	1.3±0.3	(0.3%)
Z+light jet	6.9±1.4	(0.9%)	4.8±1.0	(0.7%)	2.8±0.6	(0.6%)	1.4±0.3	(0.4%)
Dibosons	15.7±2.8	(2.0%)	7.5±1.4	(1.1%)	3.4±0.6	(0.8%)	1.7±0.3	(0.4%)
$t\bar{t}$	34.7±7.5	(4.5%)	65.9±13.6	(9.3%)	11.5±2.5	(2.6%)	23.9±4.9	(6.3%)
Multijets	427.6±16.2	(55.1%)	457.5±15.9	(64.8%)	373.5±9.8	(83.5%)	319.3±8.5	(84.1%)
Background Sum	776.6±49.9		705.8±31.7		447.4±13.3		379.8±10.6	
Data	810		702		463		415	

Table E.4: Yields with uncertainty after selection for Run IIb data with 2 *b*-tagged jets. The fraction of each background component is also listed in percent.

	tau type 1+2				tau type 3			
	2 jets		3 jets		2 jets		3 jets	
Signals								
tb+tqb	3.48±0.88		2.47±0.64		0.59±0.15		0.40±0.11	
Backgrounds								
W+2b	13.79±2.75	(43.2%)	8.49±1.52	(18.7%)	2.05±0.41	(17.6%)	1.27±0.23	(5.0%)
W+2c	1.48±0.30	(4.6%)	1.25±0.22	(2.7%)	0.26±0.05	(2.3%)	0.28±0.05	(1.1%)
W+light jet	0.42±0.06	(1.3%)	0.34±0.04	(0.8%)	0.08±0.01	(0.7%)	0.07±0.01	(0.3%)
Z+2b	0.77±0.23	(2.4%)	0.50±0.14	(1.1%)	0.22±0.06	(1.9%)	0.15±0.04	(0.6%)
Z+2c	0.07±0.02	(0.2%)	0.11±0.03	(0.3%)	0.03±0.01	(0.3%)	0.05±0.01	(0.2%)
Z+light jet	0.04±0.01	(0.1%)	0.04±0.01	(0.1%)	0.02±0.00	(0.1%)	0.01±0.00	(0.0%)
Dibosons	0.87±0.20	(2.7%)	0.43±0.10	(0.9%)	0.14±0.03	(1.2%)	0.10±0.02	(0.4%)
tt̄	7.24±2.01	(22.7%)	15.67±4.11	(34.5%)	2.02±0.56	(17.3%)	5.05±1.32	(19.6%)
Multijets	7.27±1.25	(22.8%)	18.52±2.03	(40.8%)	6.82±1.01	(58.6%)	18.72±1.95	(72.8%)
Background Sum	31.95±4.09		45.36±5.20		11.64±1.28		25.69±2.37	
Data	47		50		19		34	

Appendix F

Flat Systematic Uncertainty Tables

Table F.1: Flat systematic percentage errors for channel (Run IIa, τ type 1+2, 1 tag, 2 jets).

| | Flat systematic percentage errors for channel (Run IIa, τ type 1+2, 1 tag, 2 jets) | | | | | | | | | | | |
	Multijets	Wbb	Wcc	Wlp	Zbb	Zcc	Zlp	$t\bar{t}$	dibosons	tb	tqb	$tb+tqb$
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	3.0	0.6	0.6	0.6	0.6
Jet Energy Scale	–	–	–	–	4.0	4.0	10.0	–	4.0	4.0	4.0	4.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.7	0.7	0.7	0.7	0.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	3.9	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale W+Jets to Data	–	15.0	15.0	15.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	4.0	8.0	4.0	–	5.0	4.0	4.0	4.0
Tau energy scale	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.2: Flat systematic percentage errors for channel (Run IIa, τ type 1+2, 1 tag, 3 jets).

	Flat systematic percentage errors for channel (Run IIa, τ type 1+2, 1 tag, 3 jets)											
	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	*tt̄*	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	2.8	5.2	5.2	5.2	5.2
Jet Energy Scale	–	–	–	–	4.0	4.0	10.0	–	4.0	4.0	4.0	4.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.1	3.7	3.7	3.7	3.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	3.8	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	7.0	7.0	7.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	4.0	8.0	4.0	–	5.0	4.0	4.0	4.0
Tau energy scale	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.3: Flat systematic percentage errors for channel (Run IIa, τ type 1+2, 2 tags, 2 jets).

	Flat systematic percentage errors for channel (Run IIa, τ type 1+2, 2 tags, 2 jets)											
	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	*tt̄*	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	3.0	0.6	0.6	0.6	0.6
Jet Energy Scale	–	–	–	–	6.0	6.0	10.0	–	6.0	4.0	4.0	4.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.7	0.7	0.7	0.7	0.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	14.5	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	15.0	15.0	15.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	10.0	14.0	–	–	14.0	12.0	12.0	12.0
Tau energy scale	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.4: Flat systematic percentage errors for channel (Run IIa, τ type 1+2, 2 tags, 3 jets).

	Flat systematic percentage errors for channel (Run IIa, τ type 1+2, 2 tags, 3 jets)											
	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	*t̄t*	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	2.8	5.2	5.2	5.2	5.2
Jet Energy Scale	–	–	–	–	6.0	6.0	10.0	–	6.0	4.0	4.0	4.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.1	3.7	3.7	3.7	3.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	13.2	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	7.0	7.0	7.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	10.0	14.0	–	–	14.0	12.0	12.0	12.0
Tau energy scale	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.5: Flat systematic percentage errors for channel (Run IIa, τ type 3, 1 tag, 2 jets).

	Flat systematic percentage errors for channel (Run IIa, τ type 3, 1 tag, 2 jets)											
	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	*t̄t*	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	3.0	0.6	0.6	0.6	0.6
Jet Energy Scale	–	–	–	–	6.0	6.0	10.0	4.0	6.0	4.0	4.0	4.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.7	0.7	0.7	0.7	0.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	3.8	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	15.0	15.0	15.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	4.0	8.0	10.0	–	5.0	4.0	4.0	4.0
Tau energy scale	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.6: Flat systematic percentage errors for channel (Run IIa, τ type 3, 1 tag, 3 jets).

	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	*tt̄*	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	2.8	5.2	5.2	5.2	5.2
Jet Energy Scale	–	–	–	–	6.0	6.0	10.0	4.0	6.0	4.0	4.0	4.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.1	3.7	3.7	3.7	3.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	3.6	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	7.0	7.0	7.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	4.0	8.0	10.0	–	5.0	4.0	4.0	4.0
Tau energy scale	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.7: Flat systematic percentage errors for channel (Run IIa, τ type 3, 2 tags, 2 jets).

	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	*tt̄*	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	3.0	0.6	0.6	0.6	0.6
Jet Energy Scale	–	–	–	–	6.0	6.0	10.0	4.0	6.0	4.0	4.0	4.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.7	0.7	0.7	0.7	0.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	13.3	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	15.0	15.0	15.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	12.0	14.0	15.0	–	14.0	12.0	12.0	12.0
Tau energy scale	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.8: Flat systematic percentage errors for channel (Run IIa, τ type 3, 2 tags, 3 jets).

	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	*tt̄*	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	2.8	5.2	5.2	5.2	5.2
Jet Energy Scale	–	–	–	–	6.0	6.0	10.0	4.0	6.0	4.0	4.0	4.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.1	3.7	3.7	3.7	3.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	11.2	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	7.0	7.0	7.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	12.0	14.0	15.0	–	14.0	12.0	12.0	12.0
Tau energy scale	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.9: Flat systematic percentage errors for channel (Run IIb, τ type 1+2, 1 tag, 2 jets).

	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	*tt̄*	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	3.0	0.6	0.6	0.6	0.6
Jet Energy Scale	–	–	–	–	4.0	4.0	8.0	–	4.0	3.0	3.0	3.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.7	0.7	0.7	0.7	0.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	2.9	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	14.0	14.0	14.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	–	–	–	–	–	6.0	6.0	6.0
Tau energy scale	–	–	–	–	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.10: Flat systematic percentage errors for channel (Run IIb, τ type 1+2, 1 tag, 3 jets).

	Flat systematic percentage errors for channel (Run IIb, τ type 1+2, 1 tag, 3 jets)											
	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	$t\bar{t}$	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	2.8	5.2	5.2	5.2	5.2
Jet Energy Scale	–	–	–	–	4.0	4.0	8.0	–	4.0	3.0	3.0	3.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.1	3.7	3.7	3.7	3.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	2.8	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	11.0	11.0	11.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	–	–	–	–	–	6.0	6.0	6.0
Tau energy scale	–	–	–	–	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.11: Flat systematic percentage errors for channel (Run IIb, τ type 1+2, 2 tags, 2 jets).

	Flat systematic percentage errors for channel (Run IIb, τ type 1+2, 2 tags, 2 jets)											
	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	$t\bar{t}$	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	3.0	0.6	0.6	0.6	0.6
Jet Energy Scale	–	–	–	–	6.0	6.0	8.0	–	6.0	3.0	3.0	3.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.7	0.7	0.7	0.7	0.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	13.4	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	14.0	14.0	14.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	–	–	–	–	–	18.0	18.0	18.0
Tau energy scale	–	–	–	–	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.12: Flat systematic percentage errors for channel (Run IIb, τ type 1+2, 2 tags, 3 jets).

Flat systematic percentage errors for channel (Run IIb, τ type 1+2, 2 tags, 3 jets)											
	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	*tt̄*	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	2.8	5.2	5.2	5.2	5.2
Jet Energy Scale	–	–	–	–	6.0	6.0	8.0	–	6.0	3.0	3.0	3.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.1	3.7	3.7	3.7	3.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	10.2	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	11.0	11.0	11.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	–	–	–	–	–	18.0	18.0	18.0
Tau energy scale	–	–	–	–	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.13: Flat systematic percentage errors for channel (Run IIb, τ type 3, 1 tag, 2 jets).

Flat systematic percentage errors for channel (Run IIb, τ type 3, 1 tag, 2 jets)											
	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	*tt̄*	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	3.0	0.6	0.6	0.6	0.6
Jet Energy Scale	–	–	–	–	6.0	6.0	8.0	–	6.0	3.0	3.0	3.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.7	0.7	0.7	0.7	0.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	2.5	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	14.0	14.0	14.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	–	–	–	–	–	–	–	6.0
Tau energy scale	–	–	–	–	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.14: Flat systematic percentage errors for channel (Run IIb, τ type 3, 1 tag, 3 jets).

Flat systematic percentage errors for channel (Run IIb, τ type 3, 1 tag, 3 jets)											
	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	*tt̄*	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	2.8	5.2	5.2	5.2	5.2
Jet Energy Scale	–	–	–	–	6.0	6.0	8.0	–	6.0	3.0	3.0	3.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.1	3.7	3.7	3.7	3.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	2.5	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	11.0	11.0	11.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	–	–	–	–	–	–	–	6.0
Tau energy scale	–	–	–	–	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.15: Flat systematic percentage errors for channel (Run IIb, τ type 3, 2 tags, 2 jets).

Flat systematic percentage errors for channel (Run IIb, τ type 3, 2 tags, 2 jets)											
	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	*tt̄*	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	3.0	0.6	0.6	0.6	0.6
Jet Energy Scale	–	–	–	–	6.0	6.0	8.0	–	6.0	3.0	3.0	3.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.7	0.7	0.7	0.7	0.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	13.5	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	14.0	14.0	14.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	–	–	–	–	–	–	–	18.0
Tau energy scale	–	–	–	–	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table F.16: Flat systematic percentage errors for channel (Run IIb, τ type 3, 2 tags, 3 jets).

	Flat systematic percentage errors for channel (Run IIb, τ type 3, 2 tags, 3 jets)											
	Multijets	*Wbb*	*Wcc*	*Wlp*	*Zbb*	*Zcc*	*Zlp*	$t\bar{t}$	dibosons	*tb*	*tqb*	*tb+tqb*
Branching frac.	–	–	–	–	–	–	–	1.5	–	1.5	1.5	1.5
ISR/FSR	–	–	–	–	8.0	8.0	8.0	2.8	5.2	5.2	5.2	5.2
Jet Energy Scale	–	–	–	–	6.0	6.0	8.0	–	6.0	3.0	3.0	3.0
Jet frag.	–	–	–	–	4.0	4.0	4.0	0.1	3.7	3.7	3.7	3.7
Jet ID	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jet res.	–	–	–	–	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Tau ID	–	–	–	–	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Luminosity	–	–	–	–	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.1
PDF	–	–	–	–	–	–	–	–	–	3.0	3.0	3.0
Prim. vertex	–	–	–	–	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Multijet Normalization	10.2	–	–	–	–	–	–	–	–	–	–	–
H.F. (Wjets)	–	13.7	13.7	–	–	–	–	–	–	–	–	–
H.F. (Zjets)	–	–	–	–	13.7	13.7	–	–	–	–	–	–
Scale *W*+Jets to Data	–	11.0	11.0	11.0	–	–	–	–	–	–	–	–
Tag Rate Fun.	–	–	–	–	–	–	–	–	–	–	–	18.0
Tau energy scale	–	–	–	–	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
Trigger	–	–	–	–	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Xsect.	–	–	–	–	3.6	3.6	3.6	12.7	6.0	11.2	7.4	8.4
b-jet frag.	–	–	–	–	2.0	–	–	2.0	–	2.0	2.0	2.0
Lumi. rewtg.	–	–	–	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Appendix G

Ranked BDT Training Variables

Table G.1: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 1+2, 1 tag and 2 jets bin for RunIIa data.

Ranking	Variable	KS	Ranking	Variable	KS
1	p_T(jet1+jet2)	0.35	11	$p_T(\tau)$	0.25
2	$\sum p_T^{\text{trk}}$	0.12	12	$\cos\phi(\tau, Q \times z)$	0.85
3	$M_T(W)$	0.29	13	$\cos\phi(\text{tag1},\tau)_{\text{btaggedtop}}$	0.78
4	BDT(τ)	0.49	14	$\Delta R(\tau,\text{jet1})$	0.86
5	H_T(alljets)	0.12	15	$\Delta\phi(\text{jet1},\text{jet2})$	0.45
6	Sphericity(W,alljets)	0.37	16	M(alljets-btaggedjet)	0.56
7	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.84	17	$M(W,\text{best1})$ ("best" top mass)	0.74
8	$\sum p_T^{\text{trk}}{}_{\text{DCAcut}}$	0.58	18	$M(W,\text{tag1})$ ("b-tagged" top mass)	0.92
9	M(alljets-testjet)	0.50	19	p_T^{notbest}(jet1)	0.15
10	$\cos\phi(\text{notbest},\tau)_{\text{besttop}}$	0.70	20	$\cos\phi(\text{jet1},\tau)_{\text{lab}}$	0.85

Table G.2: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 1+2, 1 tag and 3 jets bin for RunIIa data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$M_T(W)$	0.36	11	$\cos\phi(\text{tag1},\tau)_{\text{btaggedtop}}$	0.55
2	$\sum p_T^{\text{trk}}$	0.65	12	$\cos\phi(\tau, Q \times z)$	0.56
3	$\Delta\phi(\text{jet2},\not{E}_T)$	0.45	13	$\Delta R(\tau,\text{jet2})$	0.99
4	$\Delta\phi(\tau,\not{E}_T)$	0.46	14	H_T(jet1+jet2)	0.19
5	$\Delta R(\tau,\text{jet1})$	1.00	15	$\Delta R(\text{jet1},\text{jet2})$	0.50
6	EMF(τ)	0.21	16	p_T(jet3)	0.85
7	$p_T(\tau)$	0.19	17	$Q(\tau) \times \eta$	0.54
8	$\cos\phi(\text{best1},\text{notbest})_{\text{besttop}}$	0.81	18	H_T(alljets-bestjet)	0.16
9	$\Delta\phi(\text{jet1},\not{E}_T)$	0.64	19	BDT(τ)	1.00
10	$\cos\phi(\text{notbest},\tau)_{\text{besttop}}$	0.92	20	$z_{\text{primary vertex}}$	0.96

Table G.3: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 1+2, 2 tags and 2 jets bin for RunIIa data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$H_T(\tau + \not{E}_T)$	0.16	11	H(alljets-btaggedjet)	0.12
2	$p_T(W)$	0.20	12	$\Delta R(\tau,\text{jet1})$	0.76
3	$\cos\phi(\text{tag1},\tau)_{\text{btaggedtop}}$	0.89	13	$\Delta\phi(\tau\not{E}_T)$	0.41
4	$\sum p_T^{\text{trk}}$	0.89	14	$\sum p_T^{\text{trk}}{}_{\text{DCAcut}}$	0.55
5	$M_T(W)$	0.42	15	$\eta(\text{jet2})$	0.44
6	H_T(alljets-btaggedjet)	0.14	16	$\eta^{\text{detector}}(\text{jet2})$	0.81
7	$p_T(\tau)$	0.62	17	$\text{Prof}(\tau)$	0.44
8	M(alljets-btaggedjet)	0.82	18	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.95
9	$p_T(\text{jet1}+\text{jet2})$	0.44	19	$M(W,\text{best1})$ ("best" top mass)	0.99
10	$\text{BDT}(\tau)$	0.33	20	Centrality(alljets)	0.67

Table G.4: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 1+2, 2 tags and 3 jets bin for RunIIa data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$H_T(\tau + \not{E}_T)$	0.92	11	$\Delta\phi(\text{jet1}\not{E}_T)$	0.96
2	$H_T(\text{alljets}+\tau+\not{E}_T)$	0.39	12	$\Delta R(\tau,\text{jet1})$	0.62
3	$p_T(\text{jet3})$	0.39	13	$p_T(\tau)$	0.89
4	$p_T^{\text{notbest}}(\text{jet2})$	0.81	14	H_T(alljets-btaggedjet)	0.27
5	$\not{E}_T$	0.55	15	$p_T(\text{untaggedjet1})$	0.80
6	$M_T(W)$	0.48	16	$H_T(\text{jet1}+\text{jet2}+\tau+\not{E}_T)$	0.28
7	$\cos\phi(\text{tag1},\tau)_{\text{btaggedtop}}$	0.25	17	$z_{\text{primary vertex}}$	0.63
8	$\sum p_T^{\text{trk}}$	0.96	18	$\Delta R(\text{jet1},\text{jet2})$	0.50
9	Centrality(alljets)	0.90	19	$\text{BDT}(\tau)$	0.22
10	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.16	20	$Q(\tau) \times \eta$	0.67

Table G.5: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 3, 1 tag and 2 jets bin for RunIIa data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$\text{BDT}(\tau)$	0.87	11	$\cos\phi(\tau,Q \times z)$	1.00
2	$\sum p_T^{\text{trk}}$	0.87	12	$\Delta R(\tau,\text{jet2})$	0.98
3	$\text{RMS}(\tau)$	0.75	13	$p_T(\text{best1})$	0.87
4	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.97	14	M(alljets-testjet)	0.56
5	H_T(alljets)	0.38	15	$\text{EMF}(\tau)$	0.64
6	$p_T(\tau)$	0.48	16	$\cos\phi(\text{jet1},\tau)_{\text{btaggedtop}}$	0.84
7	$\Delta\phi(\text{jet2}\not{E}_T)$	0.28	17	$\cos\phi(\text{tag1},\tau)_{\text{btaggedtop}}$	0.48
8	Sphericity(W,alljets)	1.00	18	$\text{Prof}(\tau)$	0.98
9	$\cos\phi(\text{jet2},\tau)_{\text{btaggedtop}}$	0.64	19	H(alljets)	0.55
10	$\Delta\phi(\text{jet1}\not{E}_T)$	0.16	20	$z_{\text{primary vertex}}$	0.66

Table G.6: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 3, 1 tag and 3 jets bin for RunIIa data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$p_T(W)$	0.62	11	$\cos\phi(\text{notbest},\tau)_{\text{besttop}}$	0.26
2	$M_T(W)$	0.42	12	$\Delta R(\tau,\text{jet1})$	1.00
3	$\text{BDT}(\tau)$	0.27	13	$p_T(\text{jet3})$	0.13
4	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.52	14	Sphericity(W,alljets)	0.19
5	$\text{Iso}(\tau)$	0.27	15	Centrality(alljets)	0.16
6	$p_T(\text{jet1+jet2})$	0.13	16	M(alljets-btaggedjet)	0.97
7	$M(W,\text{best1})$ ("best" top mass)	0.28	17	$\Delta\phi(\text{jet2},\not{E}_T)$	0.60
8	$\Delta\phi(\text{jet1},\not{E}_T)$	0.39	18	$\text{RMS}(\tau)$	0.22
9	$\cos\phi(\text{jet1},\text{alljets})_{\text{alljets}}$	0.99	19	$p_T(\text{jet2})$	0.19
10	$p_T(\tau)$	0.61	20	$\sum p_T^{\text{trk}}$	0.97

Table G.7: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 3, 2 tags and 2 jets bin for RunIIa data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$\not{E}_T$	0.83	11	$M_T(W)$	0.23
2	$H_T(\tau+\not{E}_T)$	0.64	12	$\cos\phi(\text{tag1},\tau)_{\text{btaggedtop}}$	0.92
3	$p_T(W)$	0.77	13	M(alljets-testjet)	0.16
4	$\text{Iso}(\tau)$	0.77	14	$\cos\phi(\tau,Q \times z)$	0.86
5	$\sum p_T^{\text{trk}}$	0.44	15	$\eta^{\text{detector}}(\text{jet1})$	0.89
6	$H(\text{jet1+jet2})$	0.18	16	$\eta^{\text{detector}}(\tau)$	0.94
7	$z_{\text{primary vertex}}$	1.00	17	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.62
8	$\Delta R(\tau,\text{jet2})$	0.27	18	$\text{EMF}(\tau)$	0.98
9	$p_T(\text{jet1+jet2})$	0.35	19	$p_T(\text{jet2})$	0.12
10	$\cos\phi(\text{jet2},\tau)_{\text{lab}}$	0.32	20	$p_T(\tau)$	0.86

Table G.8: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 3, 2 tags and 3 jets bin for RunIIa data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$H_T(\tau+\not{E}_T)$	0.95	11	$\cos\phi(\tau,Q\times z)$	0.77
2	$\not{E}_T$	0.98	12	$p_T(\text{jet3})$	0.36
3	$M_T(W)$	0.28	13	$p_T(\text{jet1})$	0.93
4	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.96	14	$\Delta\phi(\text{jet1}\not{E}_T)$	0.82
5	$H_T(\text{jet1+jet2}+\tau+\not{E}_T)$	0.96	15	$H_T(\text{alljets})$	0.42
6	$\text{BDT}(\tau)$	1.00	16	$p_T(\text{jet1+jet2})$	0.75
7	$\cos\phi(\text{notbest,alljets})_{\text{alljets}}$	0.57	17	$\cos\phi(\text{tag1},\tau)_{\text{btaggedtop}}$	0.66
8	$\Delta\phi(\text{jet2}\not{E}_T)$	0.92	18	$\sum p_T^{\text{trk}}$	0.67
9	$H_T(\text{alljets}+\tau+\not{E}_T)$	0.92	19	Centrality(alljets)	0.46
10	$\text{Iso}(\tau)$	0.72	20	$M(\text{alljets-btaggedjet})$	0.53

Table G.9: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 1+2, 1 tag and 2 jets bin for RunIIb data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$M_T(W)$	0.64	11	$p_T(\text{best1})$	0.47
2	$\Delta\phi(\text{jet2}\not{E}_T)$	0.69	12	$Q(\tau)\times\eta$	0.20
3	$p_T(\tau)$	0.11	13	$z_{\text{primary vertex}}$	1.00
4	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.47	14	$\Delta R(\tau,\text{jet2})$	0.65
5	$p_T(\text{jet1+jet2})$	0.16	15	$\Delta R(\tau,\text{jet1})$	0.98
6	$\Delta\phi(\tau\not{E}_T)$	0.26	16	$\text{RMS}(\tau)$	0.92
7	$\cos\phi(\text{tag1},\tau)_{\text{btaggedtop}}$	0.50	17	$\cos\phi(\tau,Q\times z)$	0.20
8	$\Delta\phi(\text{jet1}\not{E}_T)$	0.47	18	$p_T(\text{btaggedjet1})$	0.21
9	$\sum p_T^{\text{trk}}{}_{\text{DCAcut}}$	0.67	19	$\text{Prof}(\tau)$	0.55
10	$M(W,\text{best1})$ ("best" top mass)	0.75	20	$\cos\phi(\text{notbest},\tau)_{\text{besttop}}$	0.45

Table G.10: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 1+2, 1 tag and 3 jets bin for RunIIb data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$M_T(W)$	0.85	11	$\Delta\phi(\text{jet1}\not{E}_T)$	0.61
2	$\sum p_T^{\text{trk}}$	0.31	12	$\cos\phi(\tau_{\text{besttop}},\text{besttop}_{\text{CMFrame}})$	0.41
3	$\Delta\phi(\text{jet2}\not{E}_T)$	0.38	13	$\Delta R^{min}(\tau,\text{jets})$	0.72
4	Sphericity(W,alljets)	0.11	14	$M(W,\text{best1})$ ("best" top mass)	0.32
5	$\text{BDT}(\tau)$	0.28	15	$\sum p_T^{\text{trk}}{}_{\text{DCAcut}}$	0.91
6	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.98	16	$z_{\text{primary vertex}}$	0.28
7	$p_T(\tau)$	0.29	17	$\cos\phi(\text{notbest,alljets})_{\text{alljets}}$	0.77
8	$\cos\phi(\text{best1,notbest})_{\text{besttop}}$	0.63	18	$\cos\phi(\text{jet2,alljets})_{\text{alljets}}$	0.78
9	$M(\text{jet1+jet2}+W)$	1.00	19	$H_T(\text{jet1+jet2})$	0.32
10	$\Delta\phi(\tau,\text{jet1})$	0.37	20	$\cos\phi(\text{jet1},\tau)_{\text{btaggedtop}}$	0.58

Table G.11: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 1+2, 2 tags and 2 jets bin for RunIIb data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$\not{E}_T$	0.40	11	Centrality(alljets)	0.75
2	$H_T(\tau+\not{E}_T)$	0.97	12	EMF(τ)	0.99
3	$p_T(\tau)$	0.98	13	$\cos\phi(\text{best1},\tau)_{\text{lab}}$	1.00
4	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.64	14	$\cos\phi(\text{best1},\text{notbest})_{\text{besttop}}$	0.68
5	$H_T(\text{alljets}+\tau+\not{E}_T)$	0.20	15	M(alljets-testjet)	0.89
6	$\cos\phi(\text{jet1},\tau)_{\text{btaggedtop}}$	0.38	16	M(jet1+jet2+W)	0.48
7	$M_T(W)$	0.73	17	$\Delta\phi(\tau,\text{jet1})$	0.21
8	$\Delta\phi(\tau,\not{E}_T)$	0.22	18	Prof(τ)	0.87
9	BDT(τ)	0.64	19	RMS(τ)	0.87
10	$\Delta\phi(\text{jet1},\text{jet2})$	0.75	20	p_T(jet1+jet2)	0.97

Table G.12: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 1+2, 2 tags and 3 jets bin for RunIIb data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$H_T(\tau+\not{E}_T)$	0.80	11	$\Delta\phi(\text{jet2},\not{E}_T)$	0.63
2	$M_T(W)$	0.72	12	$M(W,\text{best1})$ ("best" top mass)	0.81
3	Centrality(alljets)	0.74	13	$\Delta\phi(\text{jet1},\not{E}_T)$	0.94
4	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.73	14	$p_T(W)$	0.85
5	$p_T(\tau)$	0.63	15	M_T(jet1,jet2)	0.18
6	$\sum p_T^{\text{trk}}$	0.92	16	p_T(alljets-bestjet)	0.29
7	$\cos\phi(\text{tag1},\tau)_{\text{btaggedtop}}$	0.27	17	$z_{\text{primary vertex}}$	0.67
8	BDT(τ)	0.74	18	$\Delta\phi(\tau,\text{jet2})$	0.98
9	M(alljets-btaggedjet)	0.62	19	ΔR^{min}(alljets)	0.82
10	EMF(τ)	0.67	20	$\cos\phi(\text{notbest},\text{alljets})_{\text{alljets}}$	0.89

Table G.13: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 3, 1 tag and 2 jets bin for RunIIb data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$H_T(\tau+\not{E}_T)$	0.21	11	$z_{\text{primary vertex}}$	0.23
2	$\not{E}_T$	0.41	12	$\cos\phi(\tau_{\text{btaggedtop}},\text{btaggedtop}_{\text{CMFrame}})$	0.95
3	$M_T(W)$	0.59	13	$\sum p_T^{\text{trk}}{}_{\text{DCAcut}}$	0.83
4	Iso(τ)	0.22	14	M(alljets-testjet)	0.12
5	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.35	15	p_T(jet2)	0.88
6	$p_T(\tau)$	0.95	16	$\Delta\phi(\text{jet2},\not{E}_T)$	0.26
7	BDT(τ)	0.24	17	$M(W,\text{tag1})$ ("b-tagged" top mass)	0.17
8	RMS(τ)	0.76	18	ΔR(jet1,jet2)	0.65
9	H_T(alljets-bestjet)	0.68	19	η^{detector}(jet1)	0.36
10	$\Delta R(\tau,\text{jet2})$	0.20	20	$H_T(\text{alljets}+\tau+\not{E}_T)$	0.61

Table G.14: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 3, 1 tag and 3 jets bin for RunIIb data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$M_T(W)$	0.89	11	$\Delta\phi(\text{jet2},\not{E}_T)$	0.56
2	$H_T(\text{alljets}+\tau+\not{E}_T)$	0.40	12	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.69
3	$p_T(W)$	0.13	13	$z_{\text{primary vertex}}$	0.66
4	$\text{BDT}(\tau)$	0.49	14	$\text{EMF}(\tau)$	0.46
5	$\text{RMS}(\tau)$	0.45	15	$\cos\phi(\text{tag1},\tau)_{\text{btaggedtop}}$	0.68
6	$\sum p_T^{\text{trk}}$	0.33	16	$M(W,\text{best1})$ ("best" top mass)	0.97
7	$H_T(\text{jet1}+\text{jet2}+\tau+\not{E}_T)$	0.53	17	$\Delta\phi(\tau,\text{jet2})$	0.19
8	$Q(\tau)\times\eta$	1.00	18	$\text{Iso}(\tau)$	0.46
9	$p_T(\tau)$	0.54	19	$\Delta R(\tau,\text{jet1})$	0.43
10	$\Delta\phi(\text{jet1},\not{E}_T)$	0.11	20	$\cos\phi(\text{jet2},\tau)_{\text{btaggedtop}}$	1.00

Table G.15: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 3, 2 tags and 2 jets bin for RunIIb data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$H_T(\tau+\not{E}_T)$	0.82	11	Sphericity$(W,\text{alljets})$	0.52
2	$\not{E}_T$	0.71	12	$\eta^{\text{detector}}(\text{jet1})$	0.91
3	$\cos\phi(\text{tag1},\tau)_{\text{btaggedtop}}$	0.85	13	$\cos\phi(\text{best1},\text{notbest})_{\text{besttop}}$	0.84
4	$\sum p_T^{\text{trk}}$	0.24	14	$\eta(\text{jet2})$	0.78
5	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.72	15	$M(W,\text{tag1})$ ("b-tagged" top mass)	0.22
6	$M_T(W)$	0.67	16	Centrality(alljets)	0.84
7	$\text{Iso}(\tau)$	0.64	17	$\Delta R(\tau,\text{jet2})$	0.93
8	$\eta^{\text{detector}}(\text{jet2})$	0.86	18	$\eta(\tau)$	0.46
9	$\Delta\phi(\text{jet1},\not{E}_T)$	0.43	19	$\Delta\phi(\tau,\not{E}_T)$	0.88
10	$p_T(\tau)$	0.97	20	$\text{RMS}(\tau)$	0.99

Table G.16: Ranked discriminant variables used in the BDT analyses and their KS values in the τ type 3, 2 tags and 3 jets bin for RunIIb data.

Ranking	Variable	KS	Ranking	Variable	KS
1	$H_T(\tau+\not{E}_T)$	0.66	11	$\eta^{\text{detector}}(\text{jet1})$	0.66
2	$M_T(W)$	0.50	12	Aplanarity$(W,\text{alljets})$	0.48
3	$p_T(\tau)$	0.96	13	$M(W,\text{best1})$ ("best" top mass)	0.95
4	$\text{BDT}(\tau)$	0.91	14	$p_T(W)$	0.98
5	$\cos\phi(\text{best1},\tau)_{\text{besttop}}$	0.54	15	$\Delta\phi(\text{jet1},\not{E}_T)$	0.66
6	$\sum p_T^{\text{trk}}$	0.93	16	$\text{RMS}(\tau)$	1.00
7	$p_T(\text{jet3})$	0.00	17	$\cos\phi(\text{notbest},\text{alljets})_{\text{alljets}}$	0.83
8	$\text{EMF}(\tau)$	0.27	18	$\Delta R(\tau,\text{jet2})$	0.64
9	$\cos\phi(\text{untaggedjet1},\tau)_{\text{btaggedtop}}$	0.98	19	$\cos\phi(\text{jet2},\tau)_{\text{btaggedtop}}$	0.86
10	$z_{\text{primary vertex}}$	1.00	20	$\eta^{\text{detector}}(\tau)$	1.00

Appendix H

Comparison Plots of BDT Probability Distributions
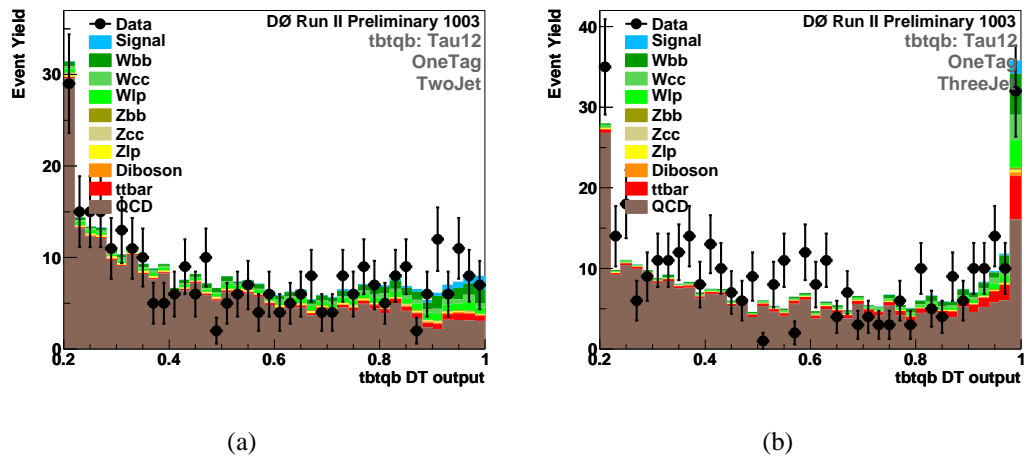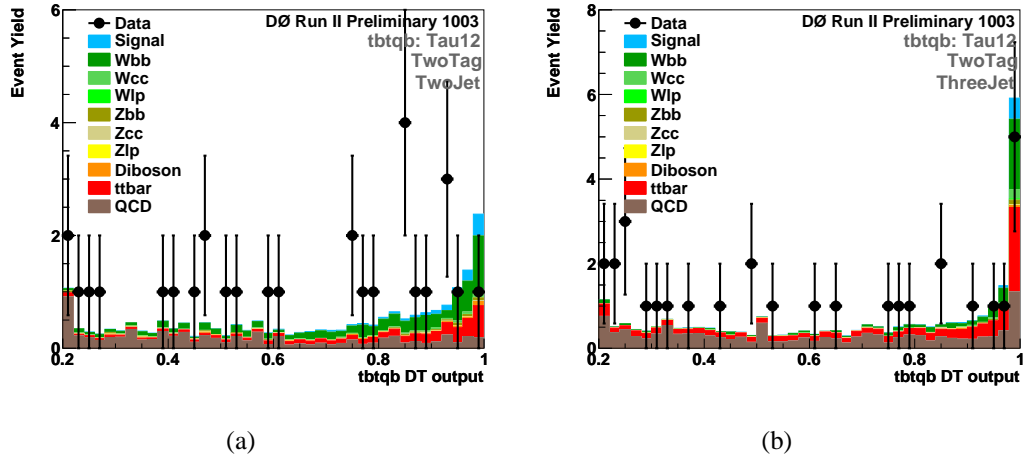

<div align="center">(a)　　　　　　　　　　　　　　(b)</div>

Figure H.1: Decision tree probability distributions for the channels: 1 *b*-tagged jet, τ type 1 and 2 and total (a) 2 jets and (b) 3 jets for Run IIa data

(a)   (b)

Figure H.2: Decision tree probability distributions for the channels: 2 *b*-tagged jets, τ type 1 and 2 and total (a) 2 jets and (b) 3 jets for Run IIa data
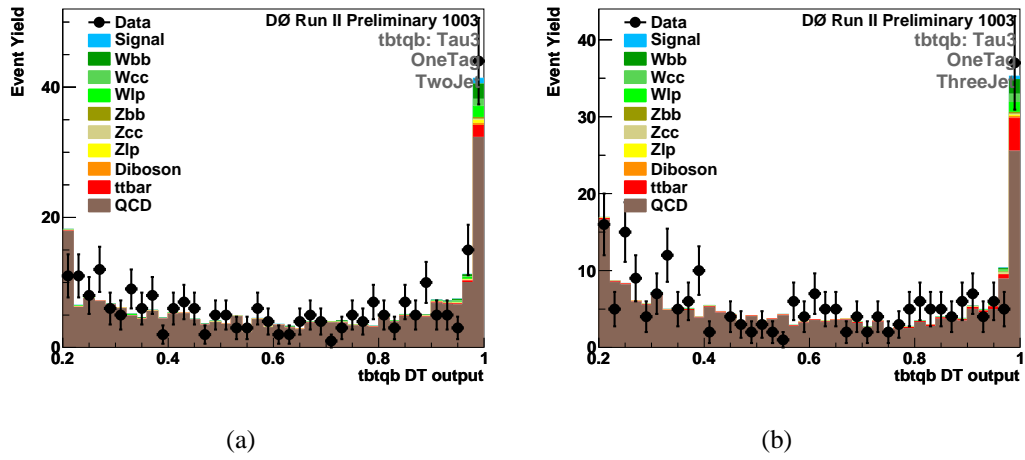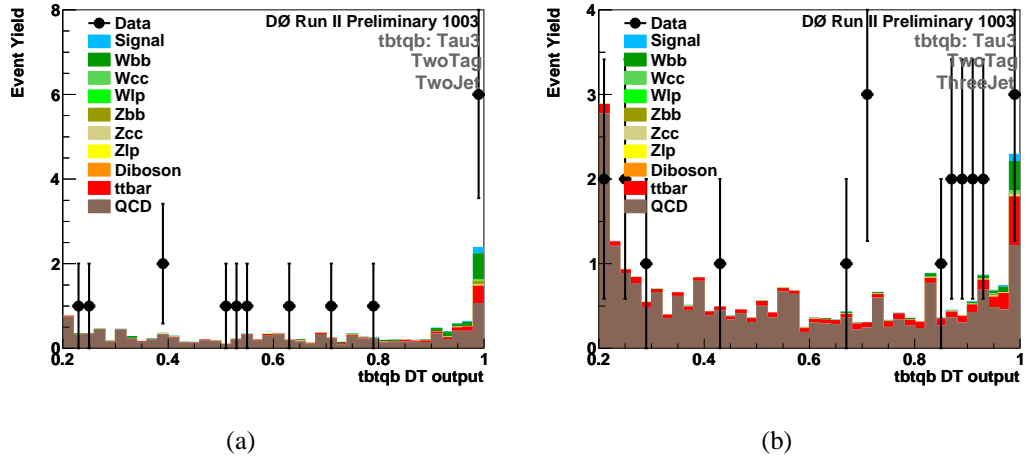

(a)   (b)

Figure H.3: Decision tree probability distributions for the channels: 1 *b*-tagged jet, τ type 3 and total (a) 2 jets and (b) 3 jets for Run IIa data

Figure H.4: Decision tree probability distributions for the channels: 2 *b*-tagged jets, τ type 3 and total (a) 2 jets and (b) 3 jets for Run IIa data
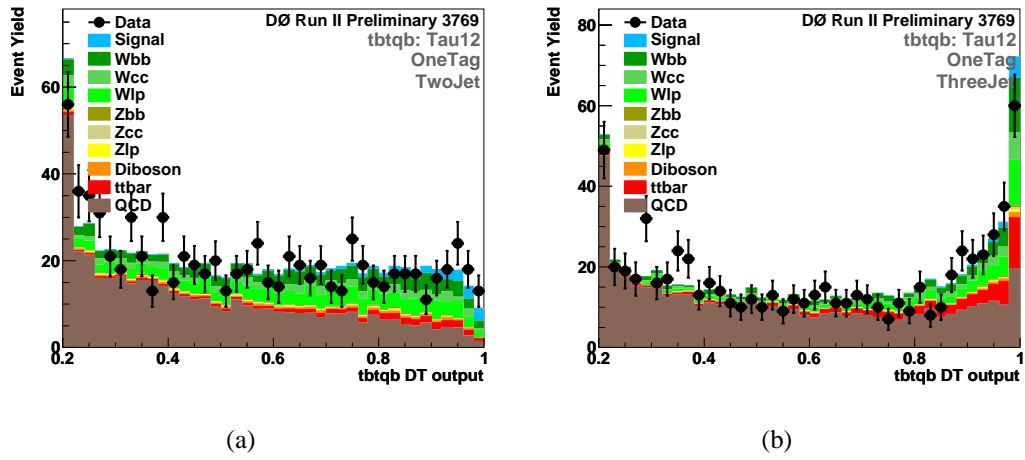

Figure H.5: Decision tree probability distributions for the channels: 1 *b*-tagged jet, τ type 1 and 2 and total (a) 2 jets and (b) 3 jets for Run IIb data
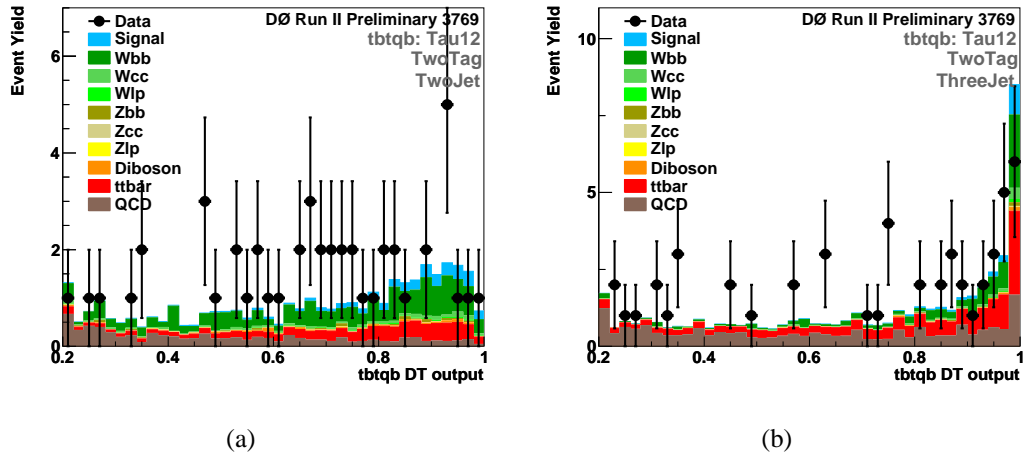
(a)   (b)

Figure H.6: Decision tree probability distributions for the channels: 2 *b*-tagged jets, τ type 1 and 2 and total (a) 2 jets and (b) 3 jets for Run IIb data
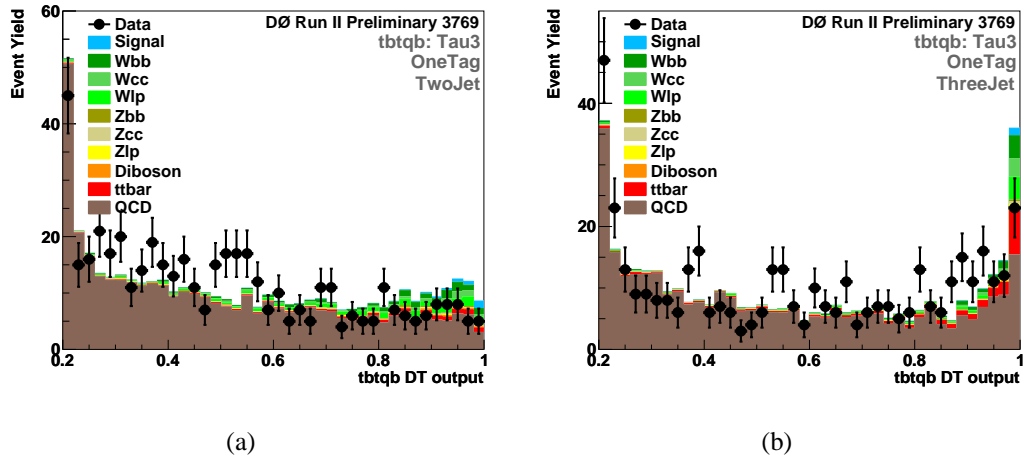

(a)   (b)

Figure H.7: Decision tree probability distributions for the channels: 1 *b*-tagged jet, τ type 3 and total (a) 2 jets and (b) 3 jets for Run IIb data
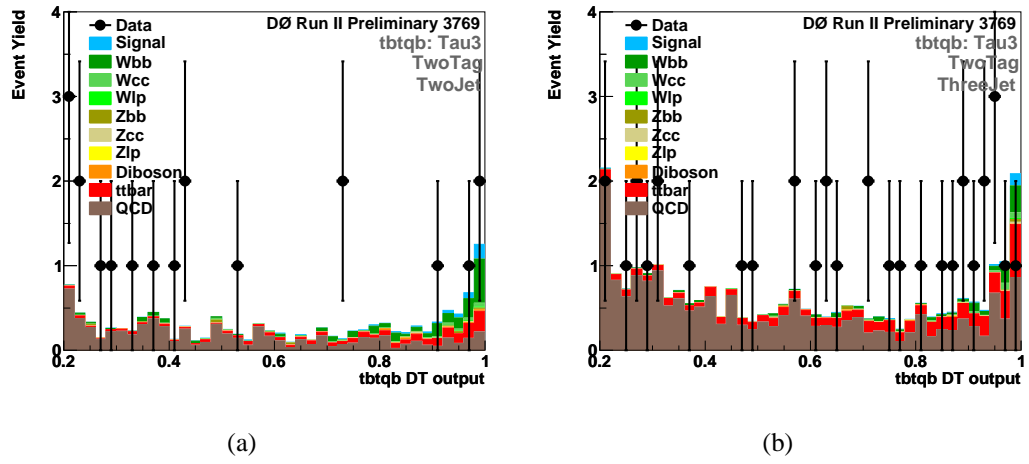
(a)   (b)

Figure H.8: Decision tree probability distributions for the channels: 2 *b*-tagged jets, τ type 3 and total (a) 2 jets and (b) 3 jets for Run IIb data

Bibliography

[1] C. Seife. What is the universe made of? *Science*, **309**:78, 2005.

[2] D. Griffiths. *Introduction of Elementary Particles*. John Wiley & Sons, Inc., 1987.

[3] O. Nachtmann. *Elementary Particle Physics: Concepts and Phenomena*. Springer-Verlag, 1989.

[4] D.H. Perkins. *Introduction to High Energy Physics*. Cambridge University Press, 2000.

[5] C. Amsler *et al.* (Particle Data Group). The review of particle physics. *Physics Letters*, **B667**:1, 2008.

[6] A. Lahiri and P.B. Pal. *A First Book of Quantum Field Theory*. Alpha Science Internatinal Ltd., 2005.

[7] R.C. Fernow. *Introduction to experimental particle physics*. Cambridge University Press, 1986.

[8] C. Campagnari. The discovery of the top quark. *Rev. Mod. Phys.*, **69**:137, 1997.

[9] A. Quadt. Top quark physics at hadron colliders. *Eur. Phys. J.*, **48**:835, 2006.

[10] R. Demina and E.J. Thomson. Top quark properties and interactions. *Annu. Rev. Nucl. Part. Sci.*, **58**:125, 2008.

[11] Marc-Andre Pleier. Review of top quark properties measurements at the Tevatron. FERMILAB-PUB-08-495-E, 2008.

[12] V.M. Abazov *et al.* (D0 collaboration). Evidence for production of single top quarks and first direct measurement of $|V_{tb}|$. *Phys. Rev. Lett.*, **98**:181802, 2007.

[13] V.M. Abazov *et al.* (D0 Collaboration). Observation of single top quark production. *Phys. Rev. Lett.*, **103**:092001, 2009.

[14] T. Aaltonen *et al.* (CDF Collaboration). First observation of electroweak single top quark production. *Phys. Rev. Lett.*, **103**:092002, 2009.

[15] Z.Sullivan. Understanding single-top-quark production and jets at hadron colliders. *Phys. Rev.*, **D70**:114012, 2004.

[16] Q.H. Cao, Reinhard Schwienhorst, and C.P. Yuan. Next-to-leading order corrections to single top quark production and decay at the Tevatron I: *s*-channel process. *Phys. Rev.*, **D71**:054023, 2005.

[17] Q.H. Cao, Reinhard Schwienhorst, and Jorge A. Benitez *et al.* Next-to-leading order corrections to single top quark production and decay at the Tevatron II: *t*-channel process. *Phys. Rev.*, **D72**:094027, 2005.

[18] N. Kidonakis. Single top quark production at the Fermilab Tevatron: Threshold resummation and finite-order soft gluon corrections. *Physical Review D*, **74**:114012, 2006.

[19] V.M. Abazov *et al.* (D0 collaboration). Evidence for production of single top quarks. *Phys. Rev.*, **D78**:012005, 2008.

[20] T. Aaltonen *et al.* (CDF Collaboration). Measurement of the single top quark production cross section at CDF. *Phys. Rev. Lett.*, **101**:252001, 2008.

[21] V.M. Abazov *et al.* (D0 collaboration). Search for anomalous *Wtb* couplings in single top quark production. *Phys. Rev. Lett.*, **101**:221801, 2008.

[22] V.M. Abazov *et al.* (D0 collaboration). Search for anomalous top-quark couplings with the D0 detector. *Phys. Rev. Lett.*, **102**:092002, 2009.

[23] V.M. Abazov *et al.* (D0 collaboration). Search for W' boson resonances decaying to a top quark and a bottom quark. *Phys. Rev. Lett.*, **100**:211803, 2008.

[24] V.M. Abazov *et al.* (D0 collaboration). Measurement of the t-channel single top quark production cross section. Fermilab-Pub-09/372-E, 2009.

[25] T.M.P. Tait. Single top quark production as a window to physics beyond the Standard Model. *Physical Review D*, **63**:014018, 2000.

[26] V.M. Abazov *et al.* (D0 collaboration). Search for production of single top quarks via *tcg* and *tug* flavor-changing-neutral-current couplings. *Phys. Rev. Lett.*, **99**:191802, 2007.

[27] G.D. Lafferty. Prospects for tau physics. *Nuclear Physics (Proc. Suppl.)*, **B189**:358, 2009.

[28] Dan Green. *High p_T Physics at Hadron Colliders*. Cambridge University Press, 2005.

[29] R.R. Wilson. The Tevatron. Technical Report TM-763, catalogue #0102.000, Fermilab, 1978.

[30] Assembled and edited by F.T. Cole, M.R. Donaldson and D.A. Edwards *et al.* A report on the design of the Fermi National Accelerator Laboratory superconducting accelerator. Technical report, Fermilab, 1979.

[31] Operated by Universities Research Association Inc. Design report Tevatron 1 production. Technical report, Fermilab, 1984.

[32] Operated by Universities Research Association Inc. Conceptual design report: Proton-proton collider upgrade (main injector, new Tevatron). Technical report, Fermilab, 1988.

[33] http://www-bd.fnal.gov/public/index.html.

[34] http://www-bd.fnal.gov/public/proton.html.

[35] http://www-bd.fnal.gov/public/maininj.html.

[36] http://www.fnal.gov/pub/news04/update_archive/update_11-10.html.

[37] http://www-bd.fnal.gov/public/tevatron.html.

[38] http://www-bd.fnal.gov/public/antiproton.html.

[39] V.M. Abazov *et al.* (D0 collaboration). The upgraded D0 detector. *Nucl. Instrum. Meth. Phys. Res.*, **A565**:463, 2006.

[40] Editors: U. Baur, R.K. Ellis, and D. Zeppenfeld. QCD and weak boson physics in Run II. Fermilab-Pub-00/297, 2000.

[41] C. Grupen. *Particle Detectors*. Cambridge University Press, 1996.

[42] R. Fruhwirth, M. Regler, and R.H. Bock *et al.* *Data Analysis Techniques for High Energy Physics*. Cambridge University Press, 2000.

[43] S. Abachi *et al.* (D0 collaboration). The D0 detector. *Nucl. Instrum. Methods Phys. Res.*, **A338**:185, 1994.

[44] B.S. Acharya, B. Baldin, and S. Banerjee *et al.* Scintillation counters for the D0 muon upgrade. *Nucl. Instrum. Meth. in Physics Research A*, **401**:45, 1997.

[45] B. Quinn. Upgrade of the D0 detector: The Tevatron beyond 2 fb^{-1}. *International Journal of Modern Physics A*, **20**:3793, 2005.

[46] D0 Layer 0 silicon group. D0 layer 0 silicon upgrade. D0 note #4415 (unpublished), 2004.

[47] M. Abolins, J. Ban, and J. Bystricky *et al.* The Run IIb trigger upgrade for the D0 experiment. *IEEE Transactions on nuclear science*, **51**:340, 2004.

[48] K. Peters, P. Rich, and S. Soldner-Rembold. Search for SM Higgs boson production in the decay $WH \rightarrow \tau_H \tau_\tau bb$ with the D0 detector at $\sqrt{s} = 1.96$ Tev. D0 note #5669 (unpublished).

[49] M. Petteni, G. Davies, and T. Christoudias. Search for $ZH \rightarrow \nu\bar{\nu}b\bar{b}$ at D0. D0 note #5352 (unpublished), 2007.

[50] U. Husemann. Monte Carlo simulations for top pair and single top production at the Tevatron. In *International Workshop on Top-Quark Physics*, La Biodola, Isola d'Elba, Italy, May 2008.

[51] Amnon Harel. Monte Carlo for top background at the Tevatron. In *International Workshop on Top-Quark Physics*, La Biodola, Isola d'Elba, Italy, May 2008.

[52] E.E. Boos *et al.* Method for simulating electroweak top-quark production events in the NLO approximation: Singletop generator. *Phys. Atom. Nucl.*, **69**:1317, 2006. We use version 4.2p1.

[53] J. Pumplin *et al.* New generation of parton distributions with uncertainties from global QCD analysis. *J. High Energy Phys.*, **0207**:012, 2002. We used versions CTEQ6M (signals) and CTEQ6L1 (backgrounds).

[54] S. Mrenna T. Sjöstrand and P. Skands. Pythia 6.4 physics and manual. *J. High Energy Phys.*, **0608**:026, 2006. We used version 6.409.

[55] S. Jadach *et al.* The tau decay library TAUOLA: version 2.4. *Comput. Phys. Commun.*, **76**:361, 1993. We used version 2.5.

[56] D.J. Lange. The EVTGEN particle decay simulation package. *Nucl. Instrum. Meth. A*, **462**:152, 2001. We used a D0 version similar to version 00-14-05.

[57] M.L. Mangano *et al.* ALPGEN, a generator for hard multiparton processes in hadronic collisions. *J. High Energy Phys.*, **0307**:001, 2003. We used ALPGEN version 2.11.

[58] S. Höche *et al.* Matching parton showers and matrix elements. arxiv:hep-ph/0602031, 2006.

[59] D. Gillberg. Heavy flavour removal and determination of weighting factors for ALPGEN W+jets Monte Carlo. D0 note #5129 (unpublished), 2006.

[60] R. Brun and F. Carminati. Geant: Detector description and simulation tool. CERN Program Library Long Writeup, Report No. W5013, 1993.

[61] Z. Sullivan. Understanding single-top-quark production and jets at hadron colliders. *Phys. Rev. D*, **70**:114012, 2004.

[62] N. Kidonakis and R. Vogt. Next-to-next-to-leading order soft gluon corrections in top quark hadroproduction. *Phys. Rev. D*, **68**:114014, 2003.

[63] G. Steinbruck. *Measurement of the Angular Distribution of Electrons from W Boson decays at D0*. PhD thesis, University of Oklahoma, 1999.

[64] P. Calfayan, T. Gadfort, and G. Hesketh *et al.* Muon identification certification for p17 data. D0 note #5157 (unpublished).

[65] G. C. Blazey *et al.* Run II jet physics. FERMILAB-PUB-00-297, 2000.

[66] T. Scanlon. *b-Tagging and the search for neutral supersymmetric Higgs bosons at D0*. PhD thesis, Imperial College London, 2006.

[67] The D0 collaboration. Search for $ZH(\to l^+l^-b\bar{b})$ in $p\bar{p}$ collisions at $\sqrt{s}$=1960 Gev. D0 note #5482 (unpublished), 2007.

[68] X. Wu, A. Clark, and M. Campanelli, editors. *b-tagging at D0*, Proceedings of the 1st Hadron Collider Physics Symposium, Les Diablerets, Switzerland, July 4-9, 2005 2005. Springer.

[69] A. Schwartzman and C. Tully. Primary vertex reconstruction by means of adaptive vertex fitting. D0 note #4918 (unpublished).

[70] A. Schwartzman and M. Narain. Probabilistic primary vertex selection. D0 note #4042 (unpublished), 2002.

[71] J. Han and M. Kamber. *Data mining: concepts and techniques*. Elsevier, 2006.

[72] A. Hoecker, P. Speckmayer, and J. Stelzer *et al.* *TMVA 4 (Toolkit for Multivariate Data Analysis with ROOT) Users Guide*, 2009.

[73] http://tmva.sourceforge.net/.

[74] B.P. Roe, H.J. Yang, and J. Zhu *et al.* Boosted decision trees as an alternative to artificial neural networks for particle identification. *Nucl. Instrum. Methods Phys. Res., Sect. A*, **543**:577, 2005.

[75] H.J. Yang, B.P. Roe, and J. Zhu *et al.* Studies of boosted decision trees for miniboone particle identification. *Nucl. Instrum. Methods Phys. Res., Sect. A*, **555**:370, 2005.

[76] V.M. Abazov *et al.* (D0 collaboration). Measurement of $\sigma(p\bar{p} \to Z) \cdot \mathrm{Br}(Z \to \tau\tau)$ at $\sqrt{s}$=1.96 Tev. *Phys. Rev.*, **D71**:072004, 2005.

[77] S. Protopopescu and P. Svoisky. Tau identification with neural networks for p17 data. D0 note #5094 (unpublished).

[78] C. Galea for the D0 collaboration. Tau identification at D0. Fermilab-conf-06-462-E, 2006.

[79] S. Protopopescu and P. Svoisky. Calculation of tau identification neural networks systematic uncertainty for p17 data. D0 note #5408 (unpublished).

[80] F. James. *Statistical methods in experimental physics*. World Scientific, 2006.

[81] S. Jain, H.B. Prosper, and R. Schwienhorst. Statistical methods implemented in the package "top_statistics". D0 note #5817 (unpublished).

[82] E. Aguilo *et al.* (The D0 Single Top Group). Single top quark production in 2.3 fb^{-1} of data – signal and background modeling and event selction. D0 note #5810 (unpublished).

[83] Freya Blekman. *Top quark pair production in proton antiproton collisions*. PhD thesis, University of Amsterdam, 2005.

[84] Edward Boos, Lev Dudko, and Thorsten Ohl. Complete calculations of $Wb\bar{b}$ and $Wb\bar{b}$+jet production at Tevatron and LHC: Probing anomalous Wtb couplings in single top production. *Eur. Phys. J.*, **C11**:473, 1999.

[85] E.Boos and L.Dudko. Optimized neural networks to search for Higgs boson production at the Tevatron. *Nucl. Instrum. Meth.*, **A502**:486, 2003.

[86] V.M. Abazov *et al.* (D0 collaboration). Combination of $t\bar{t}$ cross section measurements and constraints on the mass of the top quark and its decays into charged Higgs bosons. *Phys. Rev.*, **D 80**:071102(R), 2009.