Stony Brook University

Search for the Standard Model Higgs boson at DØ in the $\mu + \tau$(hadrons) + 2 jets final state

A Dissertation Presented

by

Wanyu Ye

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Physics

Stony Brook University

December 2012

Stony Brook University

The Graduate School

Wanyu Ye

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of this dissertation.

Paul D. Grannis, Thesis Advisor

Professor, Department of Physics and Astronomy

Robert L. McCarthy, Chairperson of Defense

Professor, Department of Physics and Astronomy

Jacobus Verbaarschot

Professor, Program Director, Department of Physics and Astronomy

Serban Protopopescu

Physicist, Omega group, Brookhaven National Laboratory

This dissertation is accepted by the Graduate School

Charles Taber

Interim Dean of the Graduate School

Abstract of the Dissertation

Search for the Standard Model Higgs boson at DØ in the μ + τ(hadrons) + 2 jets final state

by

Wanyu Ye

Doctor of Philosophy

in

Physics

Stony Brook University

2012

The Standard Model has been a successful theory in various aspects. It predicted and led to discovery of many new particles, including the Higgs boson recently found, the last missing piece of the Standard Model. The Higgs mechanism allows the vector bosons and fermions to be massive via the electroweak symmetry breaking. This dissertation presents the search of the Standard Model Higgs through the decay products: one muon, one hadronically decaying tau, and two or more jets using the full 9.7 fb^{-1} of Tevatron collider Run II data set collected in the DØ detector at Fermilab. The main production channels are gluon-gluon fusion, vector boson fusion, and Higgs production associated with a W/Z boson. No evidence of the Standard Model Higgs boson is observed in these channels with hypothesized Higgs mass between 105 GeV and 150 GeV, but the data do not exclude it either. We set the upper limits on the ratio of the 95% CL exclusion to the SM Higgs cross section. Combining with other analyses in Tevatron, the Higgs mass is ruled out at 95% confidence level between 147 and 180 GeV, and a 2.9 σ excess of events indicates a Higgs boson possibly lies in the mass range from 115 to 140 GeV.

Dedicated to my beloved family and friends.

Two things fill the mind with ever new and increasing admiration and awe, the more often and steadily we reflect upon them: the starry heavens above me and the moral law within me.

- Immanuel Kant

Contents

List of Figures

List of Tables

List of Terms and Abbreviations

- Cross section: a measure of the likelihood for particle interactions. It has the units of length squared.
- KS test: The Kolmogorov-Smirnov test (KS test) is a nonparametric test for the equality of continuous, one-dimensional probability distributions to compare how different two samples are. 0 represents no resemblance and 1 represents maximal resemblance.
- Luminosity: measures the ability of a particle accelerator producing the required number of interactions,

$$\frac{dN}{dt} = \mathcal{L}\sigma, \tag{1}$$

  where $\frac{dN}{dt}$ is the number of events per unit time, $\mathcal{L}$ is the luminosity and σ is the cross section. The unit of the luminosity is typically $\mathrm{cm}^{-2}\mathrm{s}^{-1}$.
- Integrated luminosity: the integral of the luminosity with respect to time,

$$\mathcal{L}_{int} = \int_0^T \mathcal{L}(t')dt' = \frac{N}{\sigma}. \tag{2}$$

  It is directly related to the number of events of interest N and the cross section σ. The integrated luminosity is usually taken as a measure of data collected in accelerators. The typical unit is fb^{-1}.
- Bremsstrahlung: means "braking radiation" in German. It refers to the electromagnetic radiation produced by the deceleration of a charged particle when deflected by another charged particle, e.g. an electron deflected by a uranium nucleus. The lost kinetic energy transforms into photons due to energy conservation. The photon loses energy further by pair production ($\gamma \rightarrow e^+e^-$), and the secondary electrons bremsstrahlung too, resulting in a shower of particles.
- Beam halo: a small fraction of particles can acquire enough transverse energy from the repulsive forces within the beam to form a halo.
- Compton scattering: an inelastic scattering of a photon by a free charged particle, usually electron. It results in a decrease of energy, or increase of wavelength in the photon. Part of the energy of the photon is transferred to the scattering electron.
- p_T: the transverse momentum in the plane perpendicular to the beam direction. It is convenient to use the (p_T,η,z) coordinates in the high energy collisions due to their particular topology. The pseudorapidity η is defined below and z is the longitudinal direction along the beam pipe.

- ϕ: the azimuthal angle in the transverse plane perpendicular to beam. The angle begins from the upward direction: $\phi = 0°$ is straight up from the floor and $\phi = 180°$ is straight down to the floor.

- y: the rapidity is by definition

$$y = \frac{1}{2} ln \left(\frac{E + p_L}{E - p_L} \right), \tag{3}$$

  where E is the particle energy and p_L is the component of momentum along the beam axis.

- η: the pseudorapidity. In spherical coordinates (r,θ,ϕ), the pseudorapidity is by definition

$$\eta = -ln \left[tan(\frac{\theta}{2}) \right] = \frac{1}{2} ln \left(\frac{|\boldsymbol{p}| + p_L}{|\boldsymbol{p}| - p_L} \right), \tag{4}$$

  where θ equals 0^o in the proton beam direction and 180^o in the antiproton beam direction, $\boldsymbol{p}$ is the particle momentum and p_L is the component of the momentum along the beam axis. The pseudorapidity is a commonly used spatial coordinate describing the angle of a particle relative to the beam axis. In the high energy limit the particle mass is negligible, $|\boldsymbol{p}| \approx E$, and the pseudorapidity becomes the rapidity.

- ΔR: the distance in the phase space (η, ϕ). It is by definition $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

- $\not\!\!E_T$: the missing transverse energy from the invisible particles escaping the detector. The $\not\!\!E_T$ mostly comes from neutrinos. See Sec. 3.7.7 for detailed descriptions.

- H_T: scalar sum of the p_T of all jets with $p_T > 15$ GeV and $|\eta| < 3.4$

- Run I,II: Run I refers to the data collection in Tevatron at at $\sqrt{s} = 1.8$ GeV between 1992 and 1996, RunII includes two epochs, RunIIa & RunIIb and refers to the data collection after a major upgrade at $\sqrt{s} = 1.96$ GeV between 2001 and the shut-down in 2011.

- Multi-jet (MJ) background: assorted QCD processes that produce low-energy jets, which could fake a hadronically decaying τ, while μ comes from many sources. This is an important background in our $\mu\tau jj$ analysis.

- Vertex confirmation: when a jet has at least 2 associated tracks from primary vertex, it is vertex confirmed.

- LO, NLO, NNLO: the leading order cross section refers to the cross section at the tree level in the Feynman diagram. The next-to-leading order (NLO) includes all one-loop corrections and the next-to-next-to-leading order (NNLO) includes all corrections up to two loops.

- k-factor: the factor used to correct the leading order (LO) cross sections to the next-to-leading order (NLO) or higher orders in MC simulation.
- TEST, NULL hypothesis: The TEST hypothesis assumes the data consist of signal plus background. The NULL hypothesis assumes the data contain the background only.

Acknowledgements

First of all, I would like to express my very great appreciation to my advisor, Professor Paul D. Grannis, who offered enormous help in research advising, analysis instructing, dissertation writing and financial support. This dissertation could not have been written without Paul, who not only served as my supervisor but also guided and encouraged me throughout my academic program. My years in the Stony Brook high energy experimental group is the most valuable academic adventure in my life and I owe him a heavy debt of gratitude.

I would also like to offer my special thanks to Subhendu Chakrabarti, who is my instructor, colleague and one of my best friends. He helped me in great patience with physics, coding and technical issues throughout my days in Fermilab and Stony Brook. I would not have made it thus far without his assistance.

Many thanks to Kathryn Tschann-Grimm, Ian Howley, Daniel Boline and Fabrice Couderc. They helped me in different stages of my research and offered numerous useful suggestions. A special thanks to Aurelio Juste: our discussion on the new training strategy eventually led to its first application in the Tevatron, followed by a successful collaboration review and publication.

I would like to acknowledge and extend my heartfelt gratitude to the members of my dissertation and oral exam committee, Paul Grannis, Robert McCarthy, Jacobus Verbaarschot, Serban Protopopescu and John Hobbs. Your generous help allowed me to significantly improve my dissertation and led me where I am. I also wish to thank the Stony Brook high energy group for attending my talks and offering valuable suggestions. Thank you all very much for your advice and contribution.

And most especially thanks to my beloved family and friends. Words alone cannot express how much I owe them for their encouragement, persistent love and support that enabled me to complete this dissertation and the Ph.D. degree.

Thank you all!

Chapter 1

Introduction

This dissertation presents the search for the Standard Model Higgs boson through the decay products: one muon, one hadronically decaying tau, and two or more jets using the full 9.7 fb^{-1} of Tevatron accelerator Run II data collected in the DØ detector at Fermilab. The main production processes are gluon-gluon fusion, vector boson fusion, and Higgs production associated with a W/Z boson. We set upper limits on the ratio of the 95% CL exclusion to the SM Higgs cross section.

The dissertation is constructed as follows. Chapter 2 introduces the Standard Model of particle physics and the Higgs mechanism. The Standard Model boasts tremendous success in many predictions and the Higgs boson has been a missing piece for decades. It is utterly important to fill up the last gap in the theory or the foundation of the mansion could be in question. This is the main motivation of this dissertation and related research topics. Chapter 3 describes the structures of Tevatron accelerator in Fermilab and the DØ detector. The major components and functionalities of the detector are presented at length. Chapter 4 discusses the $\mu\tau jj$ analysis in great detail. We created millions of MC events, applied appropriate event weights and scaled them down to the SM productions we expected. A final sample was selected with great care to maximize the signal significance over the background. We input this sample into the multivariate analysis to model how much an event is background-like or signal-like.

We added new promising variables for the multivariate classifier from other analyses and developed the Missing Mass Calculator (MMC) technique [1] in Monte Carlo simulation and implemented it in data analysis to improve the Higgs mass resolution and accuracy. We solved an inhomogeneous kinematic equation set with 7 unknowns and 4 equations analytically in Mathematica, and were able to reconstruct the Higgs mass from the $\mu + \tau$(hadrons) final state with neutrino information. The extra 3 unknowns were scanned in grids of a 3 dimensional phase space. We found all possible solutions, applied a likelihood function determined by PYTHIA Monte Carlo simulation, and identified the most probable solution. Each event had one best mass and a number of all valid solutions. We stored these two for all events as machine learning variables to distinguish the Higgs boson and the background. They ranked well among all variables in the depository. The technique was originally designed solely for the $H \to \tau\tau$ process, but we generalized it to the $H \to WW^{(*)}$ process.

The result from the multivariate analysis was then applied to data events to compute the final discriminant, which we used to evaluate the Higgs mass limits between the hypothesized mass points 105 and 150 GeV.

Chapter 2

The Standard Model and the Higgs mechanism

2.1 Overview

The Standard Model (SM) successfully describes the kinematics and interactions of most fundamental particles and is the best microscopic underlying theory of nature by far. All fundamental particles fall into two categories: bosons and fermions. Bosons are particles with integer spins and follow Bose-Einstein statistics, while fermions are particles with half-integer spins and follow Fermi-Dirac statistics. Figure 2.1 lists the known elementary particles in nature and Fig. 2.2 summarizes their categories and the theories beneath. Together with the recently discovered Higgs boson, they form a complete picture for the known building blocks of nature and obey the rules depicted by SM.

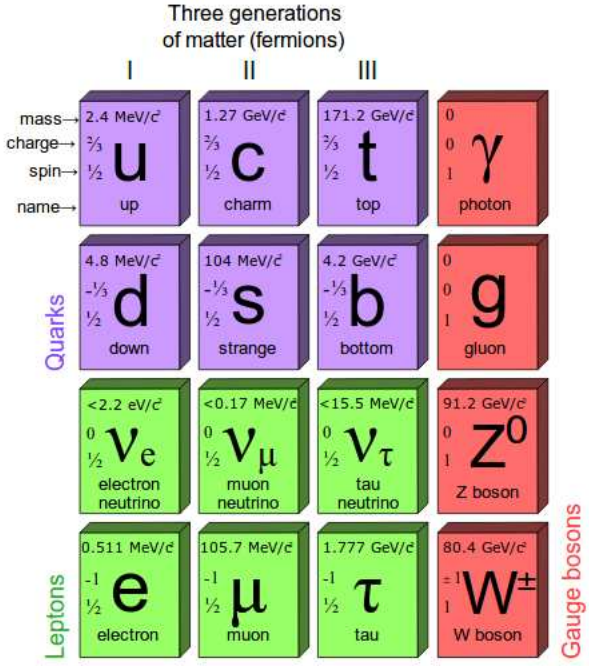


Figure 2.1: The fundamental particles described in the Standard Model. The mass, charge, spin and name are given in all blocks.

The Standard Model Lagrangian can be summarized as the following,

$$\mathcal{L}_{\mathcal{SM}} = \mathcal{L}_{\mathcal{YM}}(\mathcal{G}) + \mathcal{L}(\mathcal{G},\mathcal{F}) + \mathcal{L}(\mathcal{F},\mathcal{H}) + \mathcal{L}(\mathcal{G},\mathcal{H}) - \mathcal{V}(\mathcal{H}), \tag{2.1}$$

where $\mathcal{L}_{\mathcal{YM}}(\mathcal{G})$ is the Yang-Mills Lagrangian with the gauge group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, $\mathcal{L}(\mathcal{G},\mathcal{F})$ involves the couplings between all fundamental fermions and the gauge fields, $\mathcal{L}(\mathcal{F},\mathcal{H})$ contains the Yukawa couplings between the fermions and the Higgs boson, $\mathcal{L}(\mathcal{G},\mathcal{H})$ contains the couplings of the gauge fields and the Higgs boson, and V(H) gives the Higgs potential. The Lagrangian is invariant under continuous groups of local transformations, i.e. each term must be gauge invariant

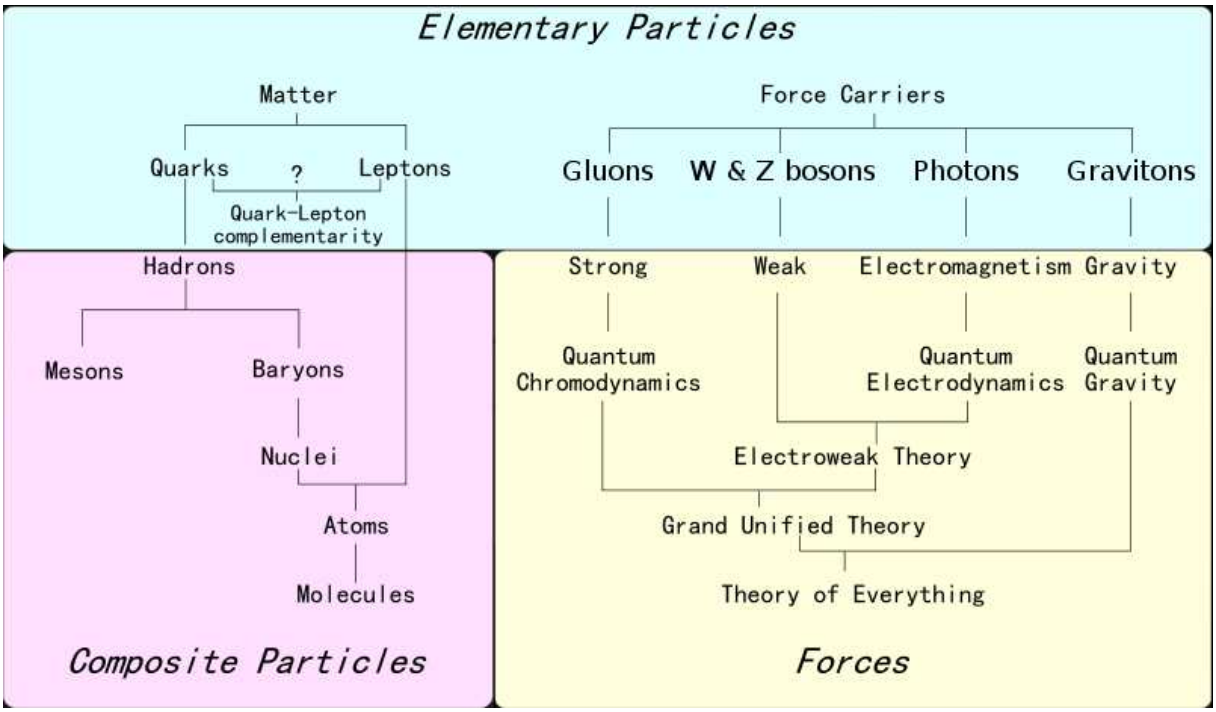


Figure 2.2: The categories of elementary particles and the theories involved.

under all defining gauge groups. This is an essential requirement in writing down the Lagrangian as the charge conservation is guaranteed if and only if the Lagrangian preserves the gauge invariance.

The Yang-Mills Lagrangian $\mathcal{L}_{\mathcal{YM}}(\mathcal{G})$ is a gauge invariant term involving three fundamental forces in nature: strong, weak and electromagnetic. Table 2.1 summarizes their properties. All force mediators are bosons and each force features charges corresponding to its gauge group. The charge is the source of all interactions related to that gauge group. The charge for the strong interaction is the color from the $SU(3)_C$ group, where massless gluons serve as the force carrier. The charges for the original $SU(2)_L \otimes U(1)_Y$ gauge group materialize into the weak and electric charges as observed in nature after spontaneous symmetry breaking (SSB).

The second through fourth terms in Eq. (2.1), $\mathcal{L}(\mathcal{G},\mathcal{F})$, $\mathcal{L}(\mathcal{F},\mathcal{H})$ and $\mathcal{L}(\mathcal{G},\mathcal{H})$ give interactions between the force mediators, fundamental fermions and the Higgs boson. After spontaneous symmetry breaking, these terms generate masses for all fermions and vector bosons. The vector bosons $W^{\pm}$, Z, leptons (except neutrinos) and quarks obtain masses during this process in similar ways, while the photon and gluon remain massless. The strong force or the $SU(3)_C$ group does not participate in the spontaneous symmetry breaking. The Standard model assumes the right-handed neutrino does not exist. It is therefore impossible to write a gauge invariant

force	strong	weak	electromagnetic
gauge group before SSB	$SU(3)_C$	$SU(2)_L \otimes U(1)_Y$	$SU(2)_L \otimes U(1)_Y$
mediators	gluons	$W^{\pm}$, Z	photons
# gauge bosons	8	3	1
charge before SSB	color	isospin	hypercharge
range (m)	10^{-15}	10^{-18}	10^{11}
relative strength	10^{13}	1	10^{11}
long-distance behavior	1 within range, 0 otherwise	$\frac{1}{r}e^{-m_{W,Z}r}$	$\frac{1}{r^2}$

Table 2.1: The three fundamental forces described in the Standard Model.

Yukawa coupling between left-handed, right-handed neutrinos and the Higgs boson. The Higgs mechanism hence does not provide masses for neutrinos.

The last term, $-\mathcal{V}(\mathcal{H})$, is the Higgs potential containing a mass term and quartic term. Giving a "wrong sign" for the mass term triggers the spontaneous symmetry breaking. The motivation to give a ad-hoc minus sign to the mass term is the gauge invariance requirement for the fermionic mass terms, if any. The fermions can not obtain mass from a trivial mass term like $m\bar{\psi}\psi$. It can be proven that such form of mass term fails to preserve the gauge invariance of the Lagrangian in Eq. (2.1) and is therefore forbidden. They have to acquire masses via other means, and a sign change in the Higgs mass term serves this purpose well. We discuss the topic at length in the next section.

The Standard Model has been a tremendously successful theory and predicted the existence of many novel elementary particles. It foresaw the existence and masses for the W and Z boson, explained the flavor conservation of neutral currents, and argued the need for a Higgs boson to generate mass for all fermions. The recently discovered Higgs-like particle in LHC, potentially the last missing piece of SM, signifies another great triumph of the model after decades of wait among the high energy community. However, further work needs to be done to verify the newly found particle indeed agrees with the SM descriptions for the Higgs boson in various production and decay channels.

2.2 The Higgs Mechanism

The Higgs mechanism is by far the most appropriate gauge-invariant approach to introduce masses for fundamental fermions and gauge bosons. It is by no means the most "natural" - the introduction of such a scalar particle inevitably brings in a hierarchy problem, i.e. fine-tuning in the loop order between Higgs-fermion and Higgs-boson couplings is required to keep the Higgs mass convergent. But we expect new physics to kick in and play an important role around 10 TeV of energy scale or so. Little Higgs [2] [3], minimal supersymmetric Standard Model (MSSM) [4], etc. address the hierarchy problem in various ways. This dessertation only focuses on

the Higgs mechanism within the SM framework.

Let us consider the last two terms $\mathcal{L}(\mathcal{G},\mathcal{H})-\mathcal{V}(\mathcal{H})$ in Eq. (2.1) first. These two terms can be written explicitly as [10]:

$$\begin{aligned}\mathcal{L}_{scalar}=\mathcal{L}(\mathcal{G},\mathcal{H})-\mathcal{V}(\mathcal{H}) &= (\mathcal{D}^{\mu}\phi)^{\dagger}(\mathcal{D}_{\mu}\phi)-\mathcal{V}(\phi^{\dagger}\phi), &(2.2)\\ \mathcal{V}(\phi^{\dagger}\phi) &= \mu^2(\phi^{\dagger}\phi)+\lambda(\phi^{\dagger}\phi)^2, &(2.3)\end{aligned}$$

where $\lambda>0$. The covariant derivative is defined as

$$\begin{aligned}\mathcal{D}_{\mu} &= \partial_{\mu}+\frac{ig'}{2}\mathcal{A}_{\mu}Y+\frac{ig}{2}\boldsymbol{\tau}\cdot\boldsymbol{b_{\mu}}, &(2.4)\\ & &(2.5)\end{aligned}$$

where $\boldsymbol{\tau}$ is the Pauli matrix vector or the generators of the $SU(2)_L$ group, Y is the hypercharge or the generator of the $U(1)_Y$ group, $\boldsymbol{b_{\mu}}$, $\mathcal{A}_{\mu}$ are the gauge fields of the $SU(2)_L$, $U(1)_Y$ groups, respectively, and

$$\phi\equiv\begin{pmatrix}\phi^{+}\\ \phi^{0}\end{pmatrix} \tag{2.6}$$

is the complex doublet of the Higgs field, which transforms as an $SU(2)_L$ doublet and carries weak charge $Y_{\phi}=+1$. As is mentioned in the last section, we want to explore a way to assign the fermions and vector bosons mass by breaking the $SU(2)_L\otimes U(1)_Y$ symmetry. Note Eq. (2.3) can be arranged as

$$\begin{aligned}\mathcal{V}(|\phi|^2) &= \mu^2|\phi|^2+\lambda|\phi|^4\\ &= \lambda(|\phi|^2+\frac{\mu^2}{2\lambda})^2-\frac{\mu^4}{4\lambda}. &(2.7)\end{aligned}$$

When $\mu^2>0$, Eq. (2.7) corresponds to a parabolic Higgs potential centered at the origin, as is shown in Fig. 2.3 (a). Now assume $\mu^2<0$ and choose the vacuum expectation value (VEV) of the scalar field

$$\langle\phi\rangle_0=\begin{pmatrix}0\\ v/\sqrt{2}\end{pmatrix}, \tag{2.8}$$

where $v=\sqrt{-\mu^2/\lambda}$. The old symmetry around the origin is broken by this modification and the Higgs potential now develops two new local minima at $\pm v$, as is shown in Fig. 2.3 (b). Given the VEV in Eq. (2.8), it can be proven that the four generators $\boldsymbol{\tau}$, $\mathcal{A}_{\mu}$ breaks this symmetry, while the generator corresponding to the electric charge preserves the $U(1)_{EM}$ symmetry. The photon therefore remains massless and the gauge bosons will acquire mass as we shall see soon. Expand the Lagrangian in Eq. (2.2) around the minimum of the Higgs potential,

$$\phi = exp\frac{i\boldsymbol{\zeta}\cdot\boldsymbol{\tau}}{2v}\begin{pmatrix}0\\(v+\eta)/\sqrt{2}\end{pmatrix}, \tag{2.9}$$

, where η is the quantum field of the Higgs boson around the VEV, and transform the fields according to the U-gauge,

$$\begin{aligned}\phi &\rightarrow \phi' = exp\frac{-i\boldsymbol{\zeta}\cdot\boldsymbol{\tau}}{2v}\phi = \begin{pmatrix}0\\(v+\eta)/\sqrt{2}\end{pmatrix},\\ \boldsymbol{\tau}\cdot\boldsymbol{b_\mu} &\rightarrow \tau\cdot b'_\mu,\\ \mathcal{A}_\mu &\rightarrow \mathcal{A}_\mu,\end{aligned} \tag{2.10}$$

the Lagrangian in Eq. (2.2) now becomes

$$\begin{aligned}\mathcal{L}_{scalar} = \frac{1}{2}(\partial^\mu\eta)(\partial_\mu\eta) - \mu^2\eta^2 + \frac{v^2}{8}[(g^2|W^+_\mu|^2 + |W^-_\mu|^2) + (g^2+g'^2)Z^2_\mu] \\ + interaction\ terms,\end{aligned} \tag{2.11}$$

where the charged gauge fields are

$$W^\pm_\mu \equiv \frac{b^1_\mu \mp ib^2_\mu}{\sqrt{2}}, \tag{2.12}$$

and the orthogonal neutral gauge fields for the Z^0 boson and photon are

$$Z_\mu = \frac{-g'\mathcal{A}_\mu + gb^3_\mu}{\sqrt{g^2+g'^2}}, \tag{2.13}$$

$$A_\mu = \frac{g\mathcal{A}_\mu + g'b^3_\mu}{\sqrt{g^2+g'^2}}. \tag{2.14}$$

We see from Eq. (2.11) that the Higgs boson acquired a mass $M_H = \sqrt{-2\mu^2}$, after rendering $\mu^2 < 0$ and rewriting the mass term (2nd term) as $M_H^2\eta^2/2$. Figure 2.3 shows the Higgs potential in Eq. (2.3) before and after the spontaneous symmetry breaking. The three vector bosons $W^\pm$, Z^0 explicitly obtain masses $gv/2$ and $\sqrt{g^2+g'^2}v/2$, respectively after spontaneous symmetry breaking, while the photon gauge field A remains massless.

Similarly, the interactions between fermions and the Higgs boson endow the fermions masses in the third term of Eq. (2.1). It can be proven after diagonalization in the quark sector that

$$\begin{aligned}\mathcal{L}(\mathcal{F},\mathcal{H}) = \quad & (-m_u\bar{u}u - m_c\bar{c}c - m_t\bar{t}t \\ & -m_d\bar{d}d - m_s\bar{s}s - m_b\bar{b}b \\ & -m_e\bar{e}e - m_\mu\bar{\mu}\mu - m_\tau\bar{\tau}\tau)(1+\eta/v),\end{aligned} \tag{2.15}$$

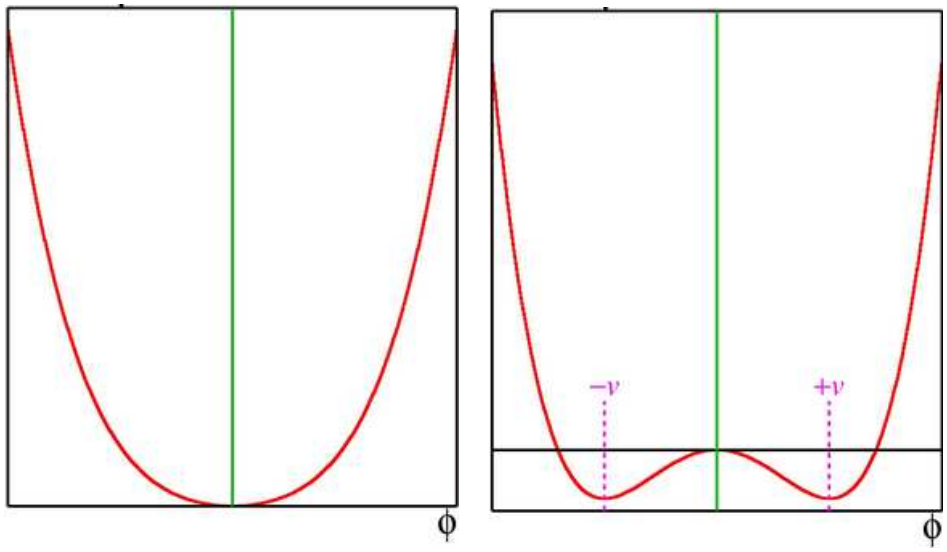


Figure 2.3: The Higgs potential $V(\phi^\dagger\phi)$ (a) before SSB; (b) after SSB.

with all masses defined properly. The gauge bosons and fermions hence derive masses according to their couplings with the Higgs boson while preserving the gauge invariance of the Lagrangian.

The Higgs mechanism breaks the electroweak symmetry and introduces masses for the gauge bosons and fermions, but this is not the only reason we need the Higgs boson and related interactions. Consider the process $e^+e^- \to W^+W^-$, the Feynman diagrams at the leading order are listed in Fig. 2.4. It can be proven that the sum of scattering amplitudes of the process via an intermediate photon, Z^0 boson and electron neutrino diverges, and the scattering amplitude via the Higgs boson diverges too, but the sum of all four gives a finite amplitude. This phenomenon is typical in interactions where the Higgs boson serves as an intermediate state: the Higgs boson or a Higgs-like particle is required in such circumstances, otherwise the sum of all scattering amplitudes would have been divergent. This requirement is called "perturbative unitarity".

2.3 The Higgs mass constraints

There are many good reasons to expect a relatively light Higgs mass [11]. An upper bound of the Higgs mass can be given by a *triviality* argument. Note the Higgs potential before and after expansion around the VEV, Eq. (2.2), (2.3) and Eq. (2.11) indicates the quartic coupling constant

$$\lambda = \frac{M_H^2}{2v^2}, \tag{2.16}$$

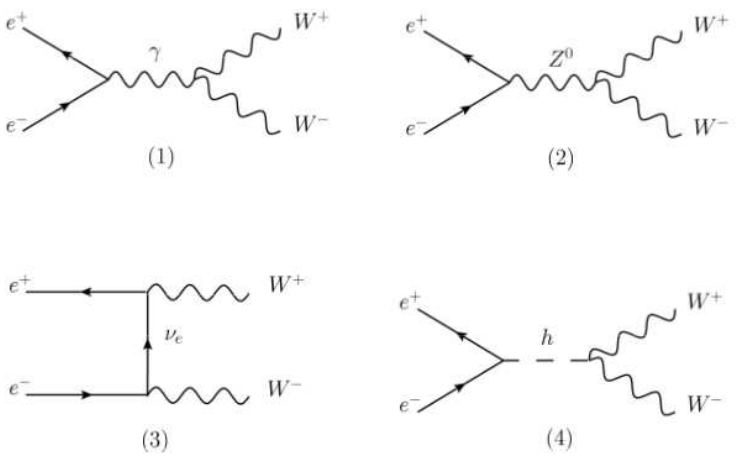


Figure 2.4: Leading order contribution to the process $e^+e^- \to W^+W^-$ through an intermediate (1) photon, (2) Z^0 boson, (3) electron neutrino, (4) Higgs boson.

assuming the theory has no gauge boson or fermions. It can be proven that λ runs with the effective energy scale Q at one-loop level due to the self-interaction of the Higgs field,

$$\frac{d\lambda}{dt} = \frac{3\lambda^2}{4\pi^2}, \tag{2.17}$$

where $t \equiv log(Q^2/Q_0^2)$ and Q_0 is any reference scale, often taken as VEV in the SM. Eq. (2.17) gives

$$\lambda(Q) = \frac{\lambda(Q_0)}{[1 - \frac{3\lambda(Q_0)}{4\pi^2} log\frac{Q^2}{Q_0^2}]}. \tag{2.18}$$

The name *triviality* refers to the case $\lambda(Q) \to 0$ when $Q \to 0$: the theory becomes *trivial* when the quatic coupling term in the Higgs potential Eq. (2.3) vanishes. Assuming $Q = \Lambda$ is the energy scale when the new physics comes in, $Q_0 = v$ as was discussed, we therefore have an upper bound for the Higgs mass from the constraint $\lambda > 0$,

$$M_H^2 < \frac{8\pi^2 v^2}{3log(\Lambda^2/v^2)}. \tag{2.19}$$

Requiring no new physics before energy scale $\Lambda = 10^{16}$ GeV gives a Higgs mass upper bound $M_H < 160$ GeV. A weaker requirement $\Lambda = 3$ TeV gives $M_H < 600$ GeV. This argument assumes the one-loop evolution Eq. (2.17) is valid at all energies, but higher order or non-perturbative corrections must be taken into account for large λ.

Another observation from Eq. (2.18) is $\lambda(Q)$ eventually becomes infinite as Q increases from 0. A Landau pole appears when the coupling constant λ becomes infinite at a finite energy scale. To tackle this mathematical inconsistency, the Higgs mass must be constrained too. See Fig. 2.5 for the upper bound to prevent the Landau pole.

For a theory with gauge bosons and fermions, we need to recalculate the running of the coupling coefficient $\lambda(Q)$ and include the heavy quark impact in particular, as the Higgs coupling to fermions is proportional to $M_H^2 \mathcal{M}_f^2$. The top quark is the heaviest and most relevant, and for $\mathcal{M}_t = 175$ GeV the Higgs mass upper limit is $M_H < 170$ GeV. Otherwise some new physics must kick in below the unification level and take over. A similar argument in lattice gauge theory requires $M_H < 640$ GeV.

A lower bound of the Higgs mass can be inferred by the requirement of vacuum stability, i.e. the spontaneous symmetry breaking does occur, $V(v) < V(0)$. It is essentially equivalent to the requirement that $\lambda > 0$ at all energies. If the SM is valid up to 10^{16} GeV, it can be proven that $M_H > 130.5 + 2.1(m_{\rm t} - 174)$ GeV. If the SM is only valid up to 1 TeV, then $M_H > 71 + 0.74(m_{\rm t} - 174)$ GeV. Figure 2.5 shows the theoretical Higgs mass bounds from the Landau pole and vacuum stability requirement.

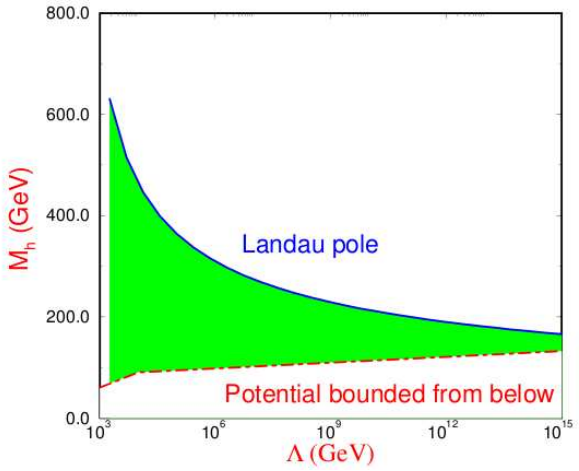


Figure 2.5: The Higgs mass bound from the Landau pole (upper solid line) and vacuum stability requirement (lower dashed line) [11].

The Higgs mass is also constrained by its radiative corrections at the loop level. The fourth term in Eq. (2.1) indicates the Higgs boson interacts with the gauge bosons, which contribute to the Higgs mass at loop levels in the form

$$g^2 \left(log \frac{M_H}{M_{\rm W}} + g^2 log \frac{M_H^2}{M_{\rm W}^2} + \ ... \right). \tag{2.20}$$

More generally, a "screening theorem" [12] states that the radiative corrections from the gauge bosons always give logarithmic contributions to the Higgs mass. The precision measurements for the W and Z boson mass therefore provide further restrictions on M_H. A few measurements on the electroweak observables at LEP, SLC and Tevatron determined another upper bound for the Higgs mass, $M_H < 280$ GeV, with 95% CL level. This result assumes the minimal Standard Model and does not include the direct Higgs search exclusions. If new physics like MSSM plays a role at the loop level, the electroweak constraint could be circumvented. The precision measurement on the electroweak observables, the direct measurement of $M_{\rm W}$ and M_t at Tevatron, and the ν scattering experiments indicate another Higgs mass upper bound $M_H < 280$ GeV at 95% CL level due to the radiative corrections.

We presented various reasons the SM Higgs boson is believed to be light or new physics would have played a role at certain energy scale. This is the most important motivation to hunt for the Higgs boson at a relatively low mass region. The Tevatron at Fermilab collides high energy proton-antiproton beams at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV and focuses on the Higgs search in the mass region 100-200 GeV. This dissertation focuses on the limits across hypothesized Higgs masses 105-150 GeV, in 5 GeV increment.

Chapter 3

Tevatron and the DØ detector

3.1 Introduction

The Tevatron is a circular particle accelerator located in Fermi National Accelerator Laboratory (Fermilab), Batavia, Illinois. It is a synchrotron and collides proton and antiproton beams in a 6.28 km ring at a center of mass energy $\sqrt{s} = 1.96$ TeV, or about 1 TeV per beam, hence the name. The accelerator was proposed in 1983, began running and eventually shut down in September, 2011 due to budget cuts and the completion of LHC in CERN. During its two major runs, Tevatron made a variety of important achievements, including the discovery of the top quark in 1995 and its mass measurement, detailed analysis of gauge boson couplings, discovery of the charged Ξ_b baryon, detection of matter-antimatter asymmetry in the B meson oscillation, greatly improved limits on the SM Higgs boson and new phenomena, etc. Two multi-purpose detectors, DØ and CDF are established at collision points $D0$ and $B0$ along the 4-mile ring.

As is shown in Fig. 3.1, protons and antiprotons go through multiple stages before actual collisions:

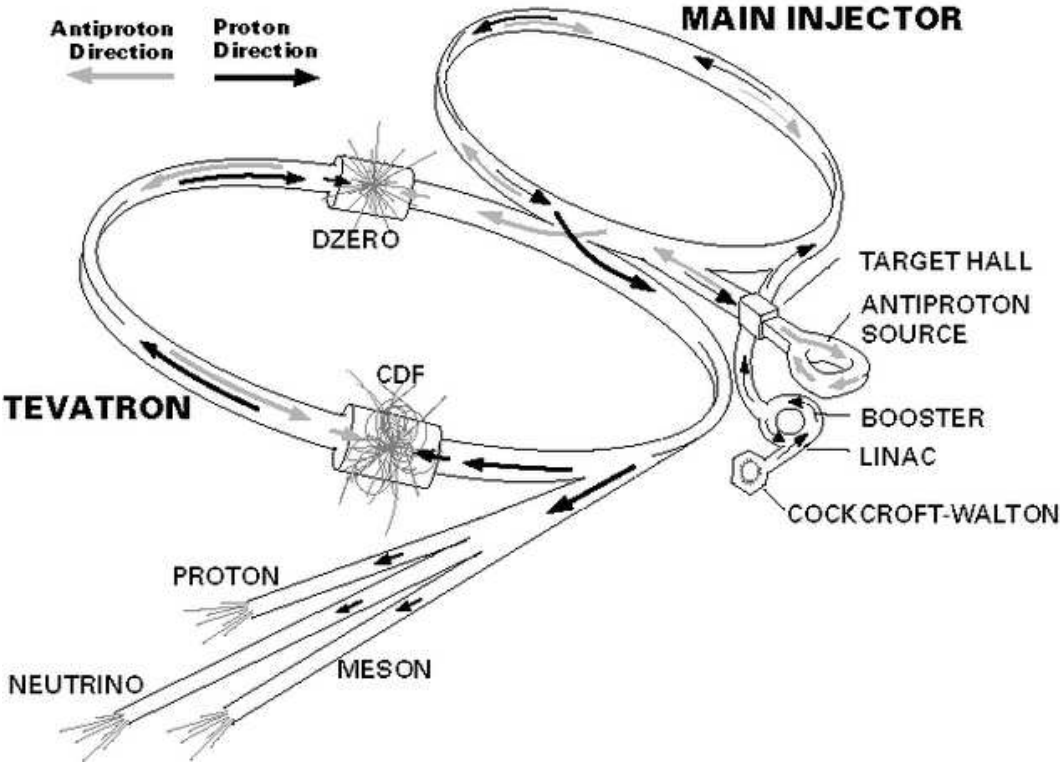


Figure 3.1: The scheme of the Tevatron and assisting accelerators.

1. **Cockcroft-Walton preacceleration**: the hydrogen gas enters the magnetron chamber (see Fig. 3.2), ironizes into H^- and gets accelerated by an

electric field in the Cockcroft-Walton accelerator to 750 keV. The H^- is then bunched by a series of radio frequency (RF) cavities and brought to the next stage.

2. **Linac**: a 500-foot linear accelerator, Linac provides a few sections of alternating electric fields further boosts the H^- ion energy to 400 MeV. The negatively charged ions are pulled from the negative end to positive, shielded while the fields alternates and accelerated again, as shown in Fig. 3.3. A hollow pipe vacuum chamber holds these electric fields and the length of the field sections varies with the distance along the pipe, typical with shorter segments near the source and longer segments toward the destination.

3. **Booster**: the H^- ions pass through a carbon foil which strips their electrons. The remaining protons are then injected to a synchrotron, the Booster, with a storage ring of 151 m diameter. The Booster keeps accelerating the proton bunches with 18 RF cavities for about 20,000 revolutions until 8 GeV of energy is reached.

4. **Main Injector**: the 8 GeV protons are delivered to a bigger synchrontron, the Main Injector (MI) and accelerated to 150 GeV. The MI also delivers 120 GeV protons for $\bar{p}$ productions in the antiproton source, and boosts the antiprotons to 150 GeV.

5. **Antiproton production**: The MI delivers the 120 GeV protons on a nickel target every 1.47 seconds to produce $\bar{p}$ together with many other secondary particles. A lithium lens then removes the negative secondary particles and a pulsed dipole bending magnet steers the antiprotons into the debuncher and accumulator. The antiprotons are cooled down first in the Debuncher, a triangular synchrotron, then transfered to the Accumulator (the second triangular synchrotron), stacked there for about 12 hours until there are enough antiprotons for another MI injection. A shot setup refers to extracting the antiprotons from the Accumulator, sending them to the Recycler and injecting them to the MI. The Recycler, as well as the Debuncher, uses stochastic cooling to create compact bright bunches of antiprotons.

6. **Tevatron and collisions**: After the shot setup, the p and $\bar{p}$ beams are squeezed by quadruple magnets into a small transverse area of $5 \times 10^{-5}\text{cm}^2$ and steered to the geometric center of the DØ and CDF detectors. Each beam consists of 36 bunches and each bunch contains about 10^{11} particles. The Tevatron then accelerates both proton and antiproton beams from 150 GeV to 980 GeV in a 1 km diameter vacuum storage ring. It is done with over 1000 superconducting magnets with 4 Tesla of magnetic field and 8 RF cavities: the magnetic field keeps the $p\bar{p}$ beams circulating in the ring and the RF cavities provide periodic acceleration to the particles every time the bunch passes. The

two beams collide at center of mass energy $\sqrt{s} = 1.96$ TeV, with a record peak luminosity as high as $431 \times 10^{30}\text{cm}^{-2}\text{s}^{-1}$ at the DØ detector.

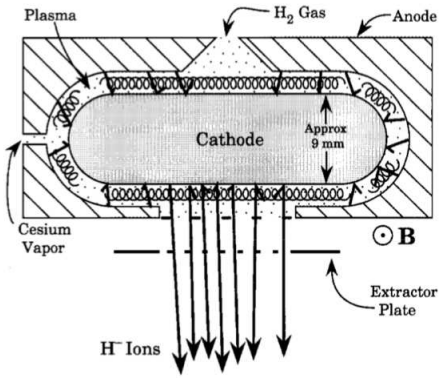


Figure 3.2: The magnetron chamber used to ironize the hydrogen gas. The injected H_2 gas turns into plasma and eventually hits the cathode and picks up two electrons.

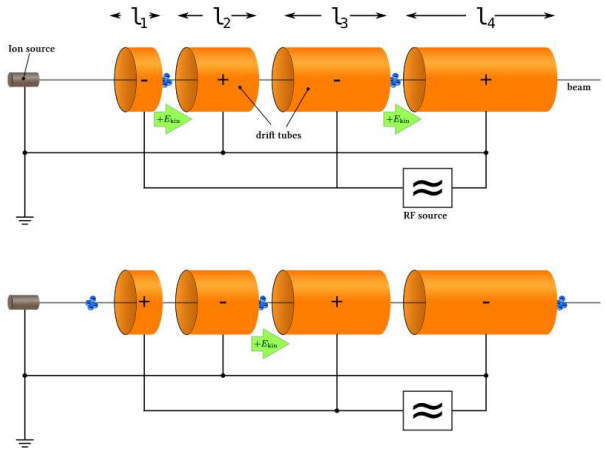


Figure 3.3: The linear accelerator (Linac) applies alternating electric fields to provide further acceleration to H^-.

The DØ detector is a 5,500-ton multi-purpose detector for the study of short-distance phenomena in high energy $p\bar{p}$ collisions. It is 20 meters long, 13 meters high and almost covers full 4π in solid angle. The detector focuses on the detection of electrons, muons, jets and missing transverse momentum. Since inception in 1992, the DØ experiment successfully ran and collected data at $\sqrt{s} = 1.8$ TeV (Run I, 1992-1996), and $\sqrt{s} = 1.96$ TeV (Run II, 2001-2011). Major upgrades between the two runs were performed and a silicon microstrip tracker, a scintillating-fiber tracker, a 1.9 T solenoidal magnet, a preshower detector and a forward proton detector were

added. Also improved were the tracking and triggering technologies. See Fig. 3.4 for the structure of the detector. The detector was established at the collision point $D0$ and named after it conveniently. However, it is highly suspected that the designer of the DØ detector, Dr. Paul Grannis's unusual affection for his neighbor's dog eventually led to the name.

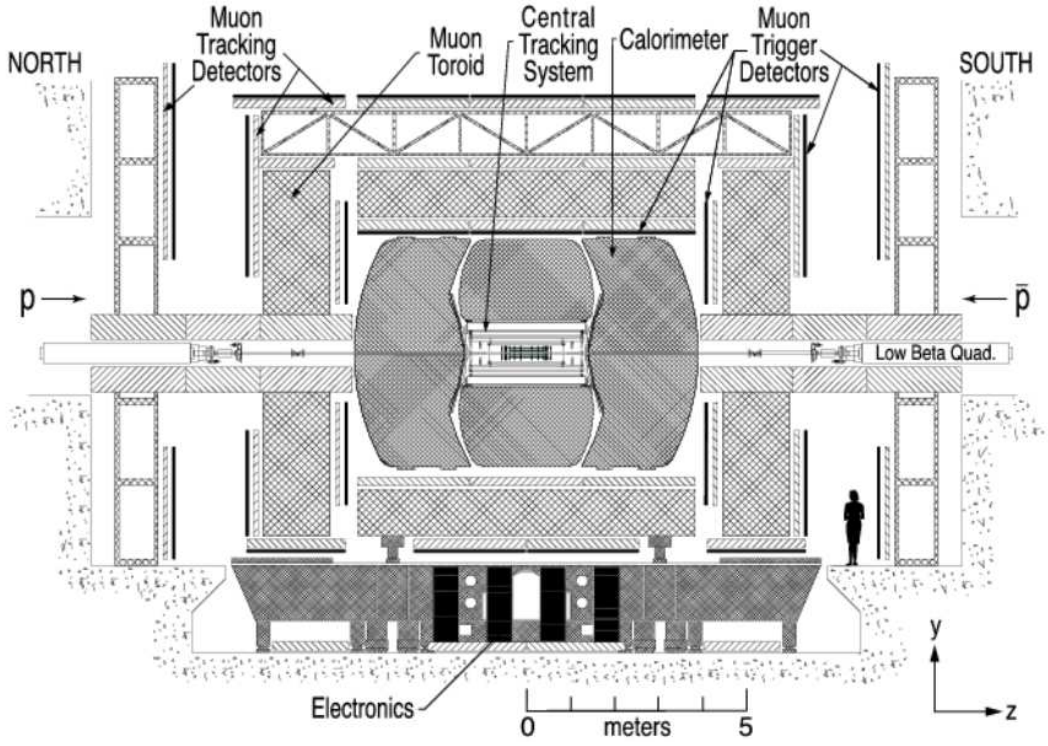


Figure 3.4: The side view of the DØ detector from inside the Tevatron ring.

The DØ detector records the information necessary to evaluate the momentum, energy, electric charge, etc. for emerging particles from interesting events. It measures elementary particles in three distinct layers:

1. **The central tracking system**: identifies the traces of energy and maps the flight paths of all charged particles. It is surrounded by a 2T solenoid and provides tracking information to the trigger.

2. **The calorimetry**: measures the energy showers of electrons, photons and hadrons by absorption. It includes the electromagnetic and hadronic sections.

3. **The muon system**: measures the 4-momenta of the muons.

We expand each of the three layers in greater detail in the following sections. This dissertation is based on Run II data and only describes the upgraded DØ detector.

3.2 The central tracking and preshowering systems

The central tracking system consists of the Silicon Microstrip Tracker (SMT), the Central Fiber Tracker (CFT), and a surrounding solenoidal magnet. A cross-sectional view of the tracking system is shown in Fig. 3.5. This is the first layer outside the DØ beryllium beam pipe, which is 2.37 m long, 38.1 mm in outer diameter and 0.508 mm thick. The SMT and CFT locate the primary interaction vertex with 35 μm resolution along the beam line. The b-quark tagging can be done with a resolutoin as good as 15 μm for particles with $p_T < 10$ GeV at $|\eta| = 0$. High resolution in central tracking allows precise measurements on lepton p_T and jet transverse energy.

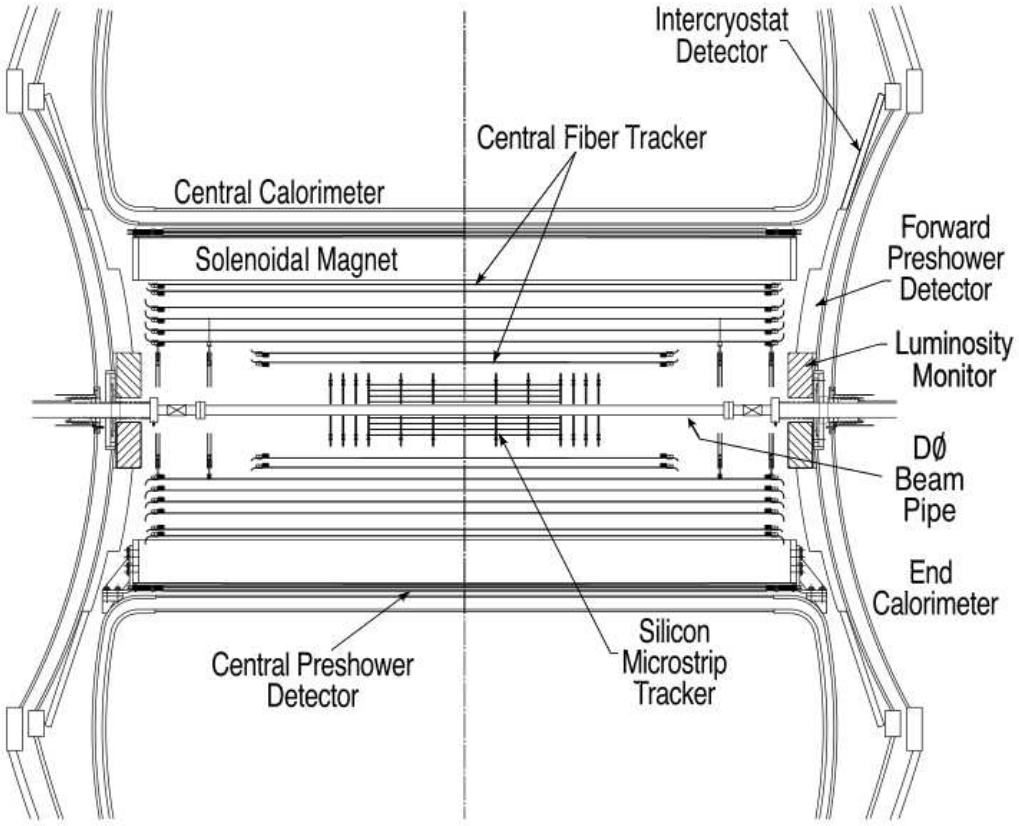


Figure 3.5: The cross-sectional view of the DØ central tracking system.

Silicon Microstrip Tracker (SMT)

The SMT matches both tracking and triggering information to the calorimeters and the muon chambers in almost full 4π solid angle coverage. The interaction region of high energy collisions is about 60 cm long. To make sure the tracks are in general orthogonal to the detector surfaces for all η, barrel modules interspersed with disks were deployed in the center and disk detectors were placed in the forward regions. See Fig. 3.6 for a demonstration for the SMT structure. The particles go through many layers this way and leave multiple hits in the detector, which helps to identify the tracks in high precision. There are six barrels in the central region, each contains four silicon readout layers. Each barrel is capped with a disk of 12 double-sided wedge detectors, the "F-disks" at high z, where the z coordinate begins from the center of the detector and goes along the beam pipe. These F-disks are symmetric about the center of SMT. Two large disks, the "H-disks", are located in the far forward region at each side. The barrel modules primarily focus on the measurement of $r-\phi$ coordinate and the disk modules measure both $r-\phi$ and $r-z$. The $r-\phi$ coordinate refers to the radial coordinate in the transverse plane with respect to the beam pipe. The high-η region is therefore covered mainly by the disk detectors and the low-η region by the barrel detectors and the central fiber tracker (CFT).

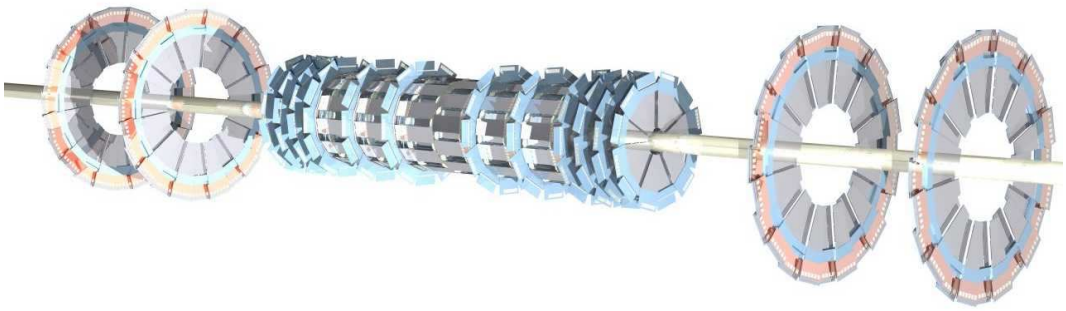


Figure 3.6: The barrel and disk modules in the SMT.

Central Fiber Tracker (CFT)

The CFT is located in the radial region $20-52$ cm from the beam pipe, between the SMT (see Fig. 3.6) and the solenoid. It contains scintillating fibers mounted on eight concentric carbon fiber cylinders; a schematic view is provided in Figs. 3.7, 3.8. Every cylinder consists of a doublet layer of fibers in axial direction (axial layer) and a doublet layer of fibers with a stereo angle in ϕ of 3° (stereo layer). The inner two layers are 1.66 m long and the outer six layers are 2.52 m long, covering a pseudorapidity range up to 1.7 and fully accomodate the SMT H-disks inside. The CFT measures both the $r-\phi$ and $r-z$ coordinates, with a resolution about 100

μm, or 2×10^{-4} in ϕ and 1 cm in z direction. The charged particles ionize and pass through the fiber, causing scintillating light that travels toward both ends of the fiber. An aluminum mirror is located at one end and reflects the light back, while on the other end, a wavelength shifting waveguide transmits the light to the Visible Light Photon Counter (VLPC) and converts the signal into electronic signal for further processing. VLPCs are also applied in the preshower detectors.

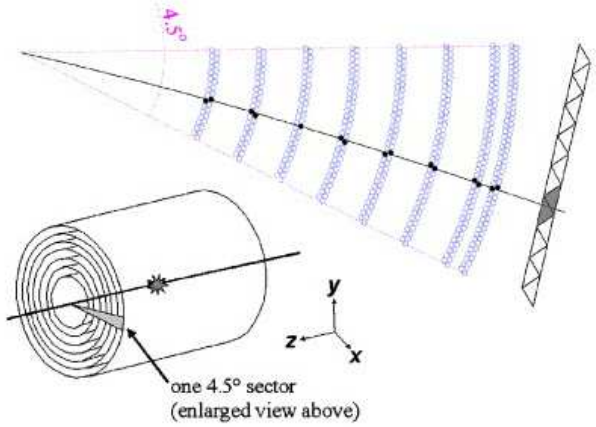


Figure 3.7: A schematic view of the 8 axial double layer in CFT and the CPS axial layer.

Forward/central preshower systems (FPS/CPS)

The forward and central preshower detectors (FPS and CPS) serve as part of both tracking and calorimetry systems. These detectors are implemented in the pseudo-rapidity ranges $|\eta| < 1.3$ and $1.5 < |\eta| < 2.5$, respectively between the solenoid and the calorimeter, see Fig. 3.5 and Fig. 3.8. They are designed to correct the ionization energy and showering of particles in the solenoid and upstream material before they enter the calorimeters. They are capable of fast energy and position measurements and aid both the tracking and calorimetry systems. The particles produce showers and create scintillating light in the preshower detectors, which allows more precise measurement of their positions. The information is then included in the Level 1 trigger.

Figure 3.9 demonstrates the structures of FPS and CPS scintillator units. Both detectors are made from triangular interleaved strips of scintillators. There is no dead space in between and most tracks traverse through multiple strips, allowing inter-strip interpolation and improved position measurements. A wavelength-shifting fiber is installed in the center of each strip, which collects and carries the light to the end of the DØ detector. The light is then transmitted via clear fibers to VLPC cassettes for readout.

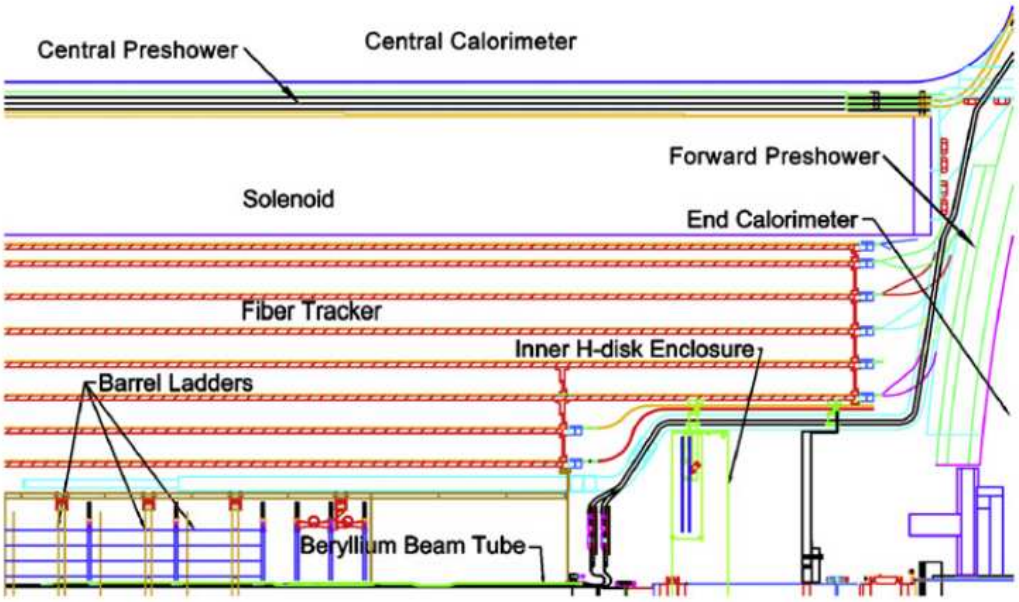


Figure 3.8: A cross-sectional view of the CFT, CPS, FPS and the solenoid in a quarter of the DØ detector.

The CPS consists of three concentric cylindrical layers of 1280 strips each, oriented along the z axis. The innermost layer is axial while the outer two (u and v) layers are stationed at stereo angles of 24° to the innermost. The two FPS detectors are mounted on the spherical heads of the north and south end calorimeters. Each of them comprises stations of two layers of scintillating strips at different z. The upstream layers (those closer to the interaction point), or the minimum ionizing particle (MIP) layers, cover $1.65 < |\eta| < 2.5$ together with the absorber, and the downstream layers behind the absorber, or the shower layers, cover $1.5 < |\eta| < 2.5$. An absorber separates these two layers. Figure 3.10 shows four FPS measuring planes: MIP u and v and shower u and v, with absorber in between. Charged particles traverse through the MIP layer and leave track information in terms of η, ϕ and z. Electrons shower in the absorber and leave about 3-strip-wide clusters of energy in the shower layer. The signals are then matched to the MIP layer. Photons generally do not interact with the MIP layer but do shower in the shower layer. Heavier charged particles do not necessarily shower in the absorber, but typically produce a second MIP signal in the shower layer.

3.3 The calorimetry

The DØ calorimetry system contains three uranium / liquid-argon calorimeters and an Inter-cryostat detector. They are designed to measure the energy deposition

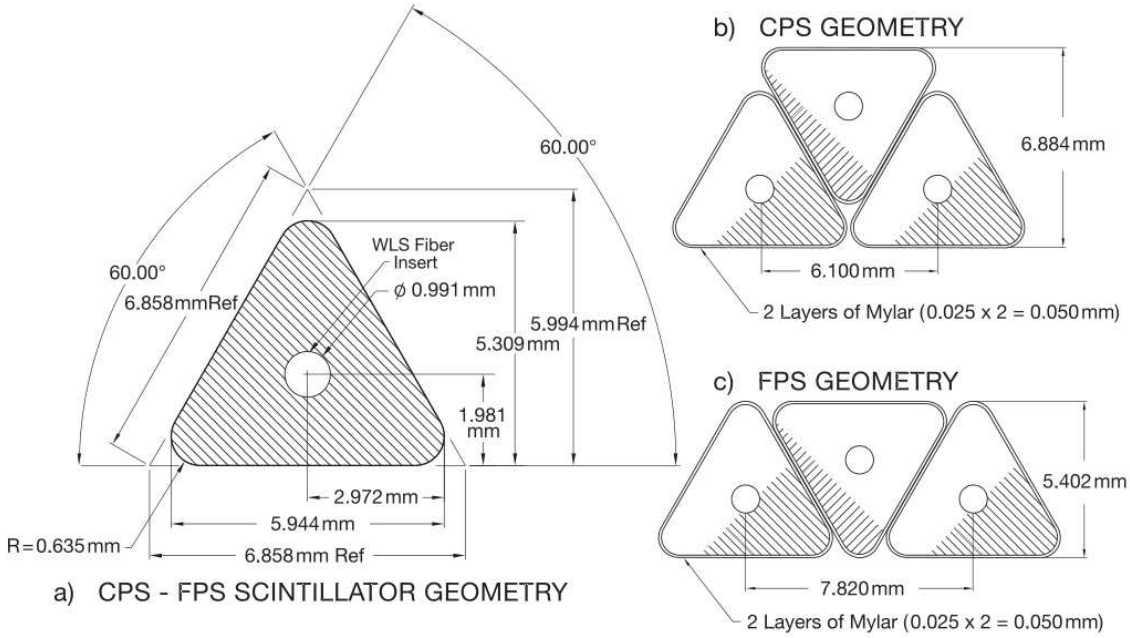


Figure 3.9: The cross-sectional layout of the FPS and CPS scintillator strips with dimensions.

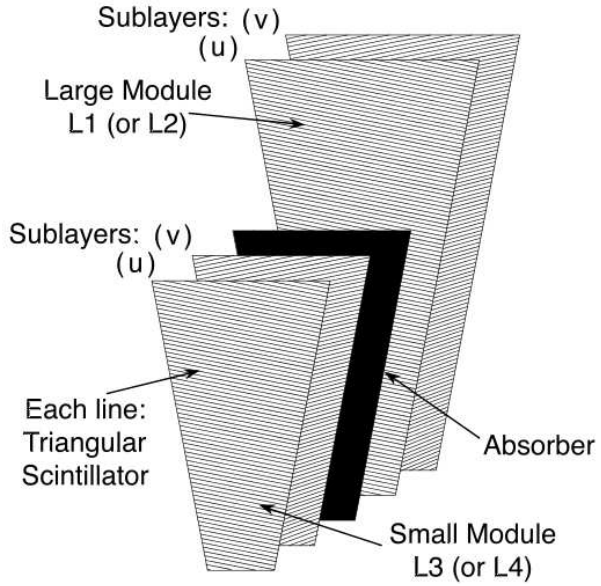


Figure 3.10: The $u - v$ MIP, shower layers and the lead-stainless-steel absorber in the FPS module.

of electrons, photons, jets, assist particle idenfications and measure the transverse energy balance per event. This is the second layer of detector outside the central tracking system (see Fig. 3.11).

Calorimeters

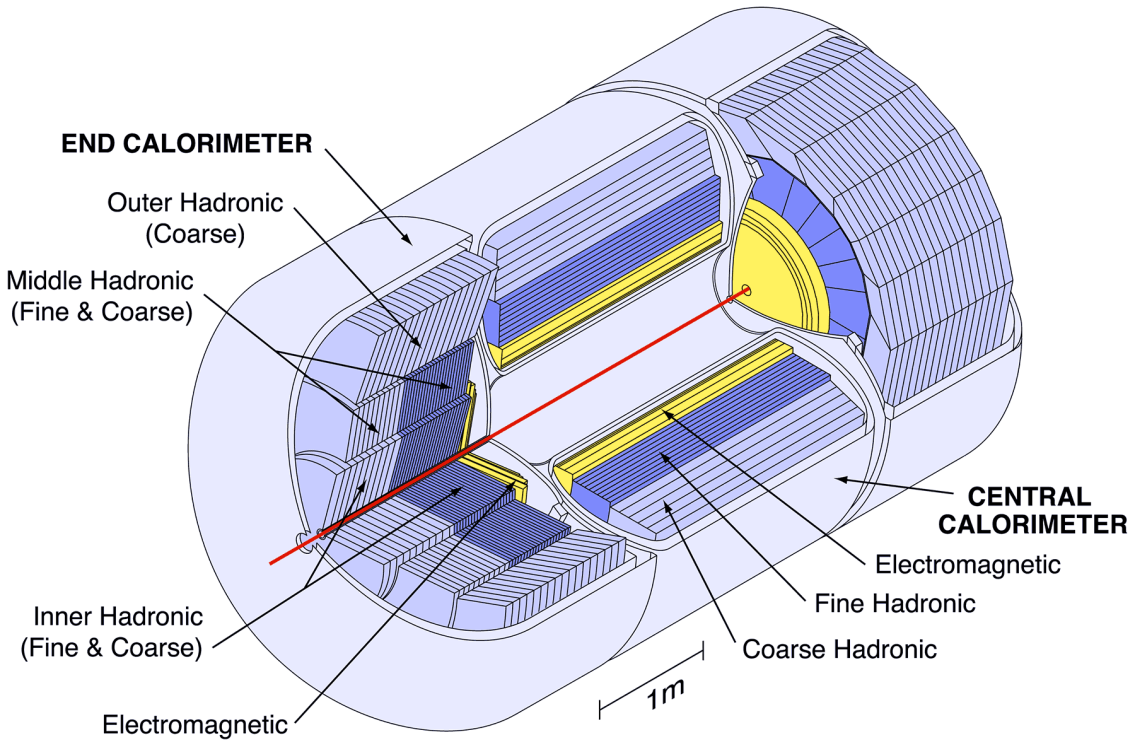


Figure 3.11: The DØ central and two end calorimeters.

The central calorimeter (CC) is barrel-shaped and covers $|\eta| < 1$, the end calorimeters in north (ECN) and south (ECS) extend the pseudorapidity coverage up to $|\eta| \approx 4$. Figure 3.11 shows an isometric view of the calorimetry system in the DØ detector and Fig. 3.12 shows the coverage diagram. Each carries liquid-argon (LAr) as the active medium and contains two sections: the electromagnetic (EM) section is located closer to the interaction point, followed by the fine / coarse hadronic (FH/CH) sections. Independent cryostats surround the calorimeters and maintain about 80 K temperature. Different absorber plates are applied in these sections. The EM section uses thin plates (3 or 4 mm) made from nearly pure depleted uranium, the FH section uses 6 mm plates made from uranium-niobium alloy, and the CH section uses 46.5 mm copper (in CC) or stainless steel plates (in EC).

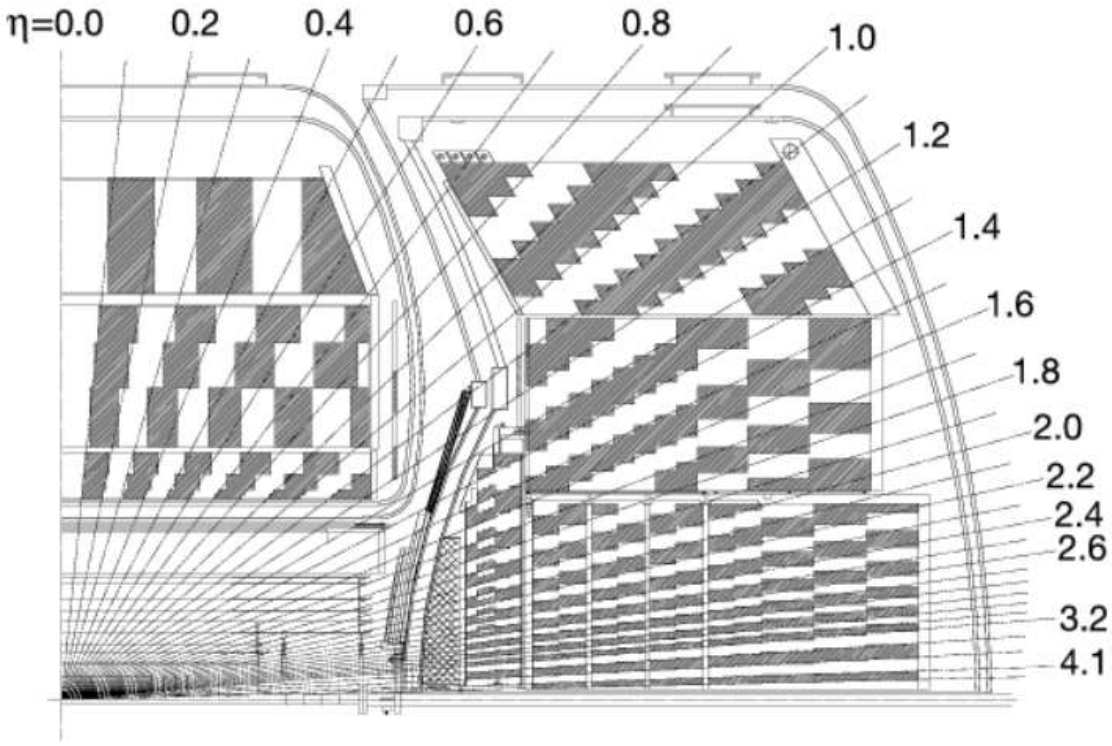


Figure 3.12: The cross-sectional view of quarter DØ detector with transverse and longitudinal segmentation pattern. Shaded areas are readout cells (towers) and rays indicate the pseudorapidity intervals from the center.

Each of the three calorimeter parts applies a radial tower segmentation. Four electromagnetic layers form the inner most calorimeter section, surrounded by three fine hadronic layers and one coarse hadronic layers, see Fig. 3.12 for a cross-sectional view. The calorimeter readout cells form radial projective towers. The segmentation of the first, second and fourth layer of the EM calorimeter and the fine, coarse hadronic calorimeters in the $\eta - \phi$ space is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, comparable to the transverse size of the showers. The third layer of the EM calorimeter is typically located close to the maximum of the shower development; it is therefore designed twice as fine in both η and ϕ to allow more precise measurement of the position.

The schematic view of a unit cell in the calorimeters is shown in Fig. 3.13. Each cell contains layers of absorber plates and copper pads, coated with resistive electrode surfaces that sustains 2.0 kV of high voltage in between. The 2.3 mm wide gap between the plates are filled with LAr. The absorber layers are designed to stop the electrons and photons in the EM section and hadrons in the FH/CH section. Charged particles pass through the medium and leave traces by ionization and bremsstrahlung. Energy deposited in the absorber ionizes atoms in LAr, and the electric field between the plates drifts these ions and electrons to the pads. The

copper pads then read out a charge proportional to the energy absorbed in the shower. The electron drift time is about 450 ns and the Tevatron bunch crossing time is 396 ns, which results in certain energy pile-up in the detector. A baseline subtraction is therefore performed in the calorimeter readout and the signal from the previous crossings is removed.

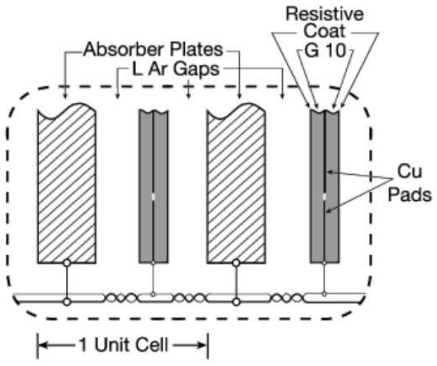


Figure 3.13: The unit cell of the calorimeter.

Inter-cryostat detector (ICD)

Three independent cryostats provides cooling for the three calorimeters, and the gaps in between them lead to incomplete pseudorapidity coverage in the region $0.8 < |\eta| < 1.4$, as is seen in Fig. 3.12. Unsampled materials in this regions further degrade the energy resolution. An additional detector, the scintillator tile Inter-cryostat Detector (ICD) is therefore installed in this region to address the problem (see Fig. 3.5). The ICD provides coverage within $1.1 < |\eta| < 1.4$ and consists of a single layer of 384 scintillating tiles sized $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, matching the unit cell size of the calorimeters. Readout cells called massless gaps (MG) fill the area $0.7 < |\eta| < 1.7$, before the first layer of uranium and in between the central and end calorimeters. These MG cells are constructed and segmented in the same way as the regular calorimeter cells. They offer missing information due to unsampled materials in the region.

3.4 The muon system

The muons are difficult to identify with the central tracking and calorimetry systems due to their heavy mass: they do not bremsstrahlung as easily as the electrons. They leave MIP signals in the central tracking, but little energy deposition in the calorimeters. A third layer of the detector, the muon system is therefore established to identify and measure the muon tracks, see Fig. 3.4 for their locations. It consists

of proportional/mini drift tubes (PDTs/MDTs), solid iron toroidal magnets and scintillation counters. Figure 3.14 shows the cut-away view of various components in the muon system. The central drift chambers in the Wide Angle Muon System (WAMUS) cover the pseudorapidity $|\eta| < 1.0$, and the two Forward Angle Muon Systems (FAMUS) extend the coverage to $|\eta| \approx 2.0$.

There are three muon detection layers: A, B and C layers, see Fig. 3.15 for an exploded view. The A layer is located inside the central toroidal magnet and B, C layers are outside. The central drift chambers use PDTs and the forward drift chambers use MDTs. All drift tubes are filled with gas to be ionized by passing charged particles. The PDTs carry three or four decks of drift cells in each layer. Each cell is 10.1 cm across, and the chambers are typically 24 cells long and contain 3 or 4 cells wide. The MDT cells are 1 cm long each. They are designed to record the electron drift time (usually 10 cm/μs), signal pulse arrival time difference between the hit cell and the readout partner, and the charge deposition on the vernier pads. The time difference is then used to determine the radial distance from the hit cell.

One central and two forward toroidal magnets in between the A and B layers curve the muons and measure their p_T and charge for an independent determination of the central tracking information. Two kinds of scintillating counters are installed to facilitate the triggering and identification of particles. The $A\phi$ scintillating counters are implemented inside the toroid magnets and the trigger scintillation counters are mounted outside. These counters collect scintillation light from the charged particles, transmit it to the wavelength shifting tubes and convert it to electronic signal outputs in photomultipliers. These signals provide additional information in the muon positions, in particular the ϕ coordinates.

3.5 The luminosity system

The luminosity monitor (LM) measures the luminosity from $p\bar{p}$ collisions in the DØ interaction region. It also measures the beam halo rates and the z coordinates of the interaction vertices. The detector comprises two arrays of 24 scintillation counters with photomultiplier tubes (PMT) readout at $z = \pm 140$ cm; see Fig. 3.16 for schematic views showing LM locations and the array structure. The arrays are located by the end calorimeters and between the beam pipe and the forward preshower detector, see Fig. 3.5 for their relative locations. The counters in the array are 15 cm long and cover $2.7 < |\eta| < 4.4$.

The luminosity ($\mathcal{L}$) is computed with the average number of inelastic collisions per beam crossing ($\bar{N}_{LM}$), the beam crossing frequency (f), and the effective cross section taking the acceptance and efficiency of the LM detector into consideration (σ_{LM}),

$$\mathcal{L} = \frac{f\bar{N}_{LM}}{\sigma_{LM}}. \tag{3.1}$$

Note $\bar{N}_{LM} > 1$ typically, which requires multiple $p\bar{p}$ collision processing within a

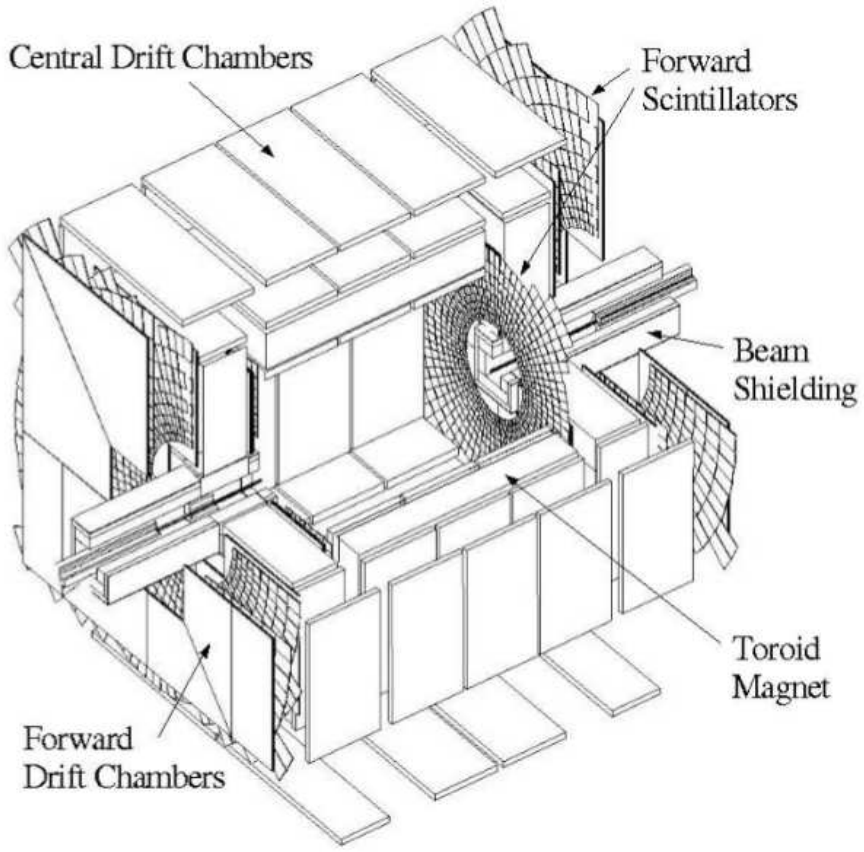


Figure 3.14: The DØ muon systems.

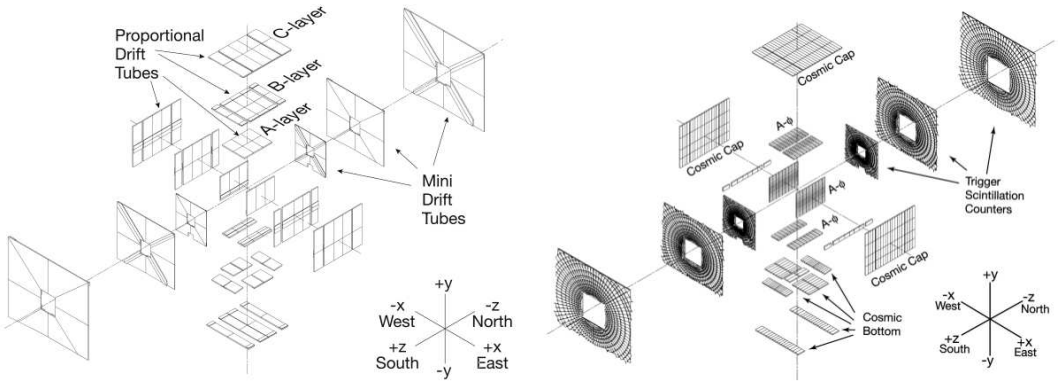


Figure 3.15: The exploded view of the DØ muon system layers. Left: proportional and mini drift tubes; right: scintillators.

single beam crossing. This can be estimated with Poisson statistics from the fraction of beam crossings with no collisions.

The beam halo backgrounds are removed by precise time-of-flight measurements of particles traveling at small angles with respect to the beams. Assuming the particles hitting the LM stem from the $p\bar{p}$ collisions, the z coordinate of the interaction vertex (z_v) can be estimated from the times of flight (t_-, t_+) when the LMs located at $z = \pm 140$ cm pick up the signal,

$$z_v = \frac{c}{2}(t_- - t_+). \tag{3.2}$$

Almost all z_v for $p\bar{p}$ collisions ranges within $|z_v| = 100$ cm, while the beam halo particles typically hit a region around $|z_v| \approx 140$ cm. The $|z_v| < 100$ cm cut therefore rules out the beam halo effect satisfactorily.

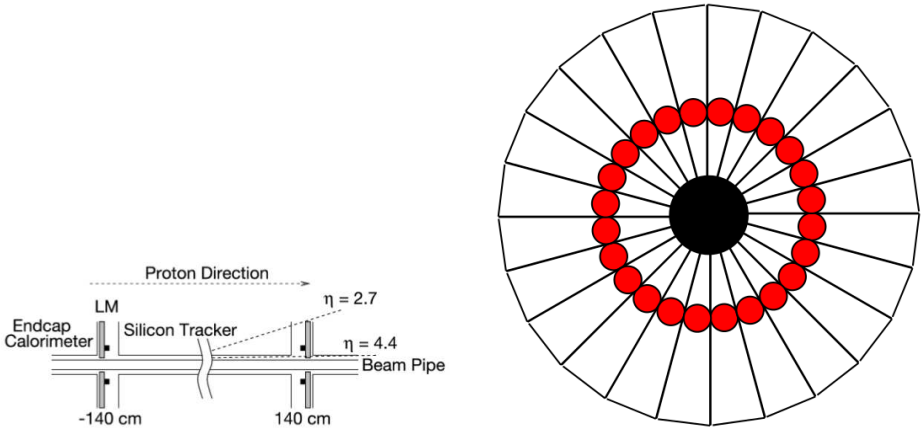


Figure 3.16: The luminosity monitor. (a) Schematic view of LM locations. (b) Schematic view of LM counters and the locations of the PMTs (solid dots) in each array.

3.6 The trigger system

The $p\bar{p}$ collisions happen in the DØ detector over 2.5 million times per second, most of which are of no interest for practical physics analyses. One needs to decide which events to keep in real-time due to constraints in the storage and speed of electronic response. The triggering framework serves this purpose. Three distinct trigger levels supply filters for event processing, with each subsequent level examining fewer events with greater detail and complexity, see Fig. 3.17 for an overview of the DØ trigger and data acquisition (DAQ) systems. The COOR package coordinates the trigger system via an online host. Figure 3.18 shows the functional blocks within L1 and L2 triggering.

1. The L1 triggers are implemented in specialized hardware to examine a huge pool of events for interesting features. The calorimeter trigger (L1Cal) uses fast readout of energy depositions in the calorimeter towers shown in Fig. 3.12. The central track trigger (L1CTT) and muon trigger (L1Muon) compare the hits in CFT and muon chambers with 128 predefined patterns to identify interesting events. All events are pipelined before entering the L1 and thus make no contribution to the deadtime. The processing rate in L1 is about 2 kHz.

2. The L2 triggers recognize simple physical objects with preprocessors in each detector sub-system. A global processor, L2Global, selects events based on L1 results and additional script-controlled criteria. The preprocessors handle events after L1 from tracking, preshower, calorimeter, and muon systems. The calorimeter preprocessor (L2Cal) builds simple jet and electron candidates with clustering algorithms. The L2CTT further sorts the L1CTT tracks by their transverse momentum. The L2Muon combines wire and scintillator hits and forms muon objects with track quality and p_T to improve the muon identifications. This information is then gathered in L2Global for L2 trigger test, as shown in Figure 3.18. L2Global also collects and uses the correlations among the individual detectors. The L2 accept rate is about 1 kHz.

3. The L3 triggers generate complex candidate physical objects or relations between them using fast versions of object-specific offline reconstruction algorithms to further filter events from L2. The objects, including electrons, jets, muons, missing transverse energy, track, vertex, etc. are defined precisely and reconstructed to a certain level. The remaining events are then recorded in the tape for offline analyses. The L3 accept rate is about 50 Hz only.

This dissertation focuses on the $\mu\tau jj$ final state of Higgs decays and uses a composite trigger called "Single Muon OR" - an OR of various muon triggers requiring a muon of medium quality in the event. The SingleMuonOR trigger imposes slightly different criteria for RunIIa and RunIIb data: RunIIa certification requires $p_T^\mu > 12$ GeV, $|\eta^\mu| < 2.0$, while RunIIb imposes tighter requirements $p_T^\mu > 15$ GeV, $|\eta^\mu| < 1.6$. See Sec. 4.3 for further details.

3.7 Offline object identification and reconstruction

3.7.1 Track reconstruction

All charged particles pass through the solenoidal magnetic field in the central tracking system and leave tracks. The medium in SMT and CFT ionizes as a result and clustered hits are recorded. Two algorithms are performed upon these clusters to fully reconstruct the trajectories.

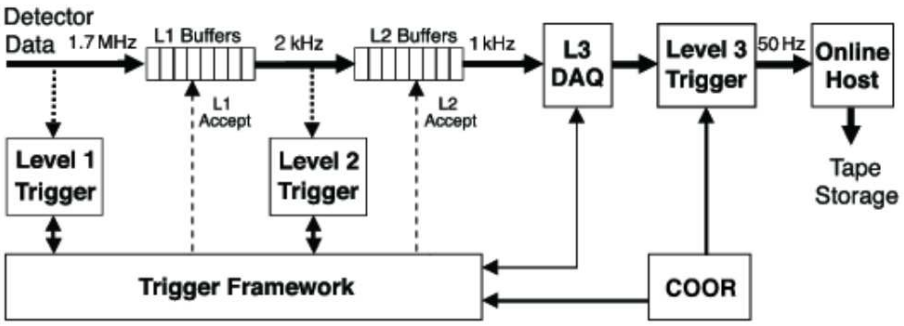


Figure 3.17: The DØ trigger and data acquisition systems.

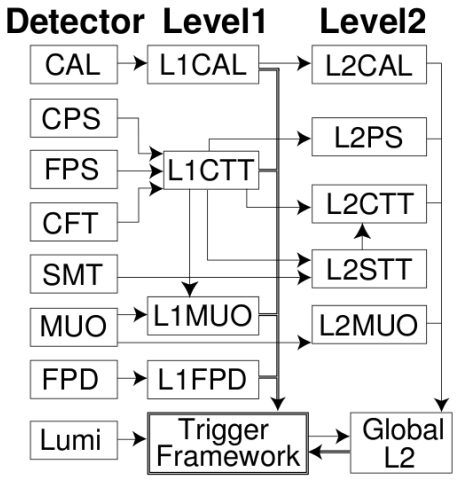


Figure 3.18: The flow of trigger-related data in DØ L1 and L2 trigger systems.

1. **The AA algorithm.** The Alternative Algorithm (AA) forms an initial track hypothesis with a combination of hits in the SMT. Starting from any combination of 3 or more hits, AA tracks the sequence of hits in the next SMT or CFT. In order to qualify, the hits are required to meet certain criteria, e.g. the axial angle between different hits, χ^2 quality threshhold, etc. The spatial coordinates of the hit cluster are then recorded and associated with the track hypothesis. If multiple tracks meet the same criteria, the hypothesis is split and a new track is formed upon each valid combination.

2. **The HTF algorithm.** The Histogram Track Finder (HTF) algorithm characterizes each track with three parameters in the plane perpendicular to the magnetic field: ρ, the radius of the hit; $d0$, the distance of closest approach (DCA) to the origin; ϕ, the azimuthal angle of the track at DCA. A track produces hit clusters in multiple locations and every two of them correspond to the same point in the $\rho - \phi$ plane. Filling each pair of hits from a track candi-

date produces a peak in a two-dimensional $\rho - \phi$ histogram, which correspond to the best estimate for the track of concern.

The AA algorithm performs better for tracks with low p_T and high impact parameters. It extends the seed clusters of initial hits to the rest of the tracking system, creates a large number of candidates and eliminates poor track candidates with well-defined criteria. The HTF is on the other hand good at high p_T track identification. The peak for a true track in the $\rho - \phi$ histogram is more pronounced and easier to identify when the track carries high p_T.

3.7.2 Vertex reconstruction

The vertex is the interaction point of the high energy collision. The primary vertex (PV) is the original point of $p\bar{p}$ collision and the secondary vertex (SV) represents the subsequent decay location. A proper PV and SV identification is essential to distinguish tracks from primary or secondary interactions, identify the $\not\!\!E_T$ and jet flavor, and separate the objects from overlapping events. The vertex reconstruction consists of two steps, vertex fitting and vertex finding [13]. Vertex fitting estimates the positions of the target vertex and the track momentum from this point, and vertex finding performs pattern recognition to identify which tracks come from the same point.

3.7.3 Electron and photon reconstructions

The EM objects, electrons and photons, are identified by tracks ($e^{\pm}$) in the central tracking system and energy depositions (both $e^{\pm}$ and γ) in the EM calorimeter in the form of showers. Electrons and positrons develop shower via bremsstrahlung, while photons shower via electron pair production. The subsequent particles interact further and lead to more showers with fractions of initial energy, until the remaining energy is too low to multiply. This is when the shower maximizes; ionization (electrons) or Compton scattering (photons) takes over from this point.

The reconstruction of EM objects begins with seed energy clusters with minimum transverse energy $E_T = 0.5$ GeV. A sum over cone $\Delta R < 0.4$ is then performed around the seeds to construct calorimeter clusters. The object is accepted if over 90% of total energy is deposited in the EM calorimeter and the energy in cone is larger than a threshhold $E_T > 1.5$ GeV. Isolated clusters tend to be of EM nature, in contrast to clusters resulting from the hadronic showers. The longitudinal showers also develop differently between EM and hadronic objects. The electrons are charged and their identifications require track match in the central tracking. A χ^2 fit with location differences in the tracking and EM calorimetry systems is performed for the purpose. An additional criterion, a likelihood determined by a machine learning algorithm, is adopted to provide further discrimination between electron candidates and the backgrounds. Objects without an associated track are more likely attributed to photons.

3.7.4 Muon reconstruction

The central tracking system, calorimeters and the muon chambers all provide necessary information to reconstuct muons. The tracking detectors give precise measurement on muon trajectory and momentum, and the muon chambers unambiguously identify muon candidates. The wire and scintillator hits in A, B and C layers of the muon system (Sec. 3.4) are matched to the central tracks with different quality grades. Muons with scintillator hit times over 10 ns more than expected from collisions are categorized as cosmic rays and rejected.

We summarize the muon type, quality parameters and meanings in Table 3.1. The type is represented by *nseg* and the quality can be *loose*, *medium* or *tight*, depending on *nseg* and the hit pattern in the muon system. The sign of *nseg* indicates if the muon candidate is matched to a central track successfuly, and $|nseg|$ indicates if the local muon (a muon reconstructed in the muon system) is made up of hits inside the toroidal magnet (A-layer only hits), outside the toroid (B or C-layer only hits), or both cases.

nseg	segment	central track match	MTC match criterion
3	A+BC	muon to central or central to muon	$\Delta\eta$, $\Delta\phi$ between MTC and central track extrapolated to CAL
2	BC only	central to muon	as above
1	A only	central to muon	as above
0	muon hit or MTC	central to muon and CAL	as above
-1	A only	no match	$\Delta\eta$, $\Delta\phi$ between MTC and A layer segment
-2	BC only	no match	$\Delta\eta$, $\Delta\phi$ between MTC and BC layer segment
-3	A+BC	no match	$\Delta\eta$, $\Delta\phi$ between MTC and local muon track at A layer if fit converged, otherwise A segment position

Table 3.1: The muon type and quality definitions.

- **tight muons.** Only $|nseg = 3|$ muons can be tight. A muon is tight if it has:
  - ⋆ at least two A layer wire hits
  - ⋆ an A layer scintillator hit
  - ⋆ at least three BC layer wire hits
  - ⋆ at least one BC scintillator hit
  - ⋆ a converged fit within the muon system ($\chi^2_{loc} > 0$)

- **$|nseg| = 3$.** $|nseg| = 3$ muons are muons with an A and a BC segments matched or not with a central track. A $|nseg = 3|$ muon is medium if it has:
  - ⋆ at least two A layer wire hits
  - ⋆ an A layer scintillator hit
  - ⋆ at least two BC layer wire hits
  - ⋆ at least one BC scintillator hit (except for central muons with less than four BC wire hits).
  - ⋆ medium/loose: A $|nseg = 3|$ medium muon meets all above requirements. A $|nseg = 3|$ loose muon is defined the same as a medium muon but allows one of the above tests to fail, with the A wire and scintillator requirement treated as one test and requiring always at least one scintillator.
- **$nseg = 2$.** $nseg = 2$ muons are muons with a BC segment matched with a central track. A $nseg = 2$ muon is loose if it has:
  - ⋆ at least one BC scintillator hit
  - ⋆ at least two BC layer wire hits.
  - ⋆ medium/loose: A $nseg = 2$ muon is defined as medium if it fullfills the above requirements and if it is located in the bottom part of the detector (octant 5 and 6 with $|\eta_{detector}| < 1.6$). Otherwise the muon is loose. The bottom part of the muon chamber is poorly instrumented, see Sec. 3.7.4 and Fig. 3.15 for further details. The criteria for muons detected there are hence different.
- **$nseg = 1$.** $nseg = 1$ muons are muons with an A segment matched with a central track. A $nseg = 1$ muon is loose if it has:
  - ⋆ an A layer scintillator hit
  - ⋆ at least two A layer wire hits.
  - ⋆ loose/medium: A $nseg = 1$ muon is defined as medium if it fullfills the above requirements and if it is located in the bottom part of the detector (octant 5 and 6 with $|\eta_{detector}| < 1.6$) due to poor instrumentation there. Low momentum $nseg = 1$ muons are also defined as medium. Otherwise the muon is loose.

Three types of muon track quality in the central tracking system are defined as follows:

- loose track: DCA<0.2 cm if track has no SMT hit; DCA<0.02 cm if track has SMT hits.
- medium track: it is a loose track; the χ^2_{dof} of the track fit fulfills $\chi^2_{dof} < 4$.

- tight track: it is a medium track; There is at least one SMT hit attached to the track.

We describe our muon selections in Sec. 4.5.1 in more detail. Our $\mu\tau jj$ analysis adopts the $nseg = 3$ muons with medium quality and medium track.

3.7.5 Tau reconstruction

The tau lepton liftime is 290.6×10^{-15} s and travels about 87.11 μm before decay on an average. This decay length is well within the beampipe and requires tau identification via its decay products. The branching ratios for leptonic and hadronic tau decays are 35% and 65%, respectively. The leptonic decay products, electrons/muons with two neutrinos, can not be distinguished from those from primary interactions, and the tau reconstruction hence refers to the hadronically decaying tau reconstruction, which could be one- or three-prong, with charged (neutral) pions and one neutrino; see Sec. 4.5.2 for a full description. Information from central tracking system, and both EM and hadronic sections of calorimeters, is gathered to detect a tau lepton. The neutral pion is detected in the EM calorimeter via its energy deposition from immediate decay to two photons $\pi^0 \to 2\gamma$. A charged pion requires track matching between the central tracking and the hadronic calorimeter. To be more specific, the following criteria are required to identify a tau [18]:

- **A hadronic cluster.** Energy in the hadronic calorimeter identified with simple cone algorithm with cone $\Delta R < 0.3$. All clusters are required to be isolated from any other clusters by $\Delta R = 0.5$ at least.

- **A EM sub-cluster.** Energy in the EM calorimeter identified with a nearest neighbor algorithm with a seed in the 3rd EM layer with a cone $\Delta R < 0.3$. The EM towers exceeding a minimum transverse energy of 0.5 GeV are considered seeds. If found, EM cells in other layers and preshower hits are added to the cluster. The energy threshold for an EM sub-cluster to be associated with a tau is 800 MeV.

- **Tracks.** Up to three tracks are associated with tau and any attempt to include more tracks has a good chance to bring in underlying event influence and lead to misconstruction. All tracks in the cone $\Delta R < 0.5$ are collected and sorted in decreasing p_T. Only tracks with $p_T > 1.5$ GeV are considered. After the first track is found, up to two more tracks with DCA within 2 cm of the z coordinate of first track are considered. The second track is included if the invariant mass of the first and second is less than 1.1 GeV. The third track is included if the invariant mass of all tracks is less than 1.7 GeV and the sum of the charges in these tracks is either $+1$ or -1.

More descriptions on type associated tau identification can be found in Sec. 4.5.2.

3.7.6 Jet reconstruction

A jet is a large amount of hadronic energy identified in a small angular region in the hadronic calorimeter. Color confinement does not allow the existence of colored particles alone. The quarks and gluons therefore always form sprays of hadrons right after the collision in the detector. These hadrons leave traces in the central tracking system and showers in the colorimeters. It is difficult to measure the energy of the initial partons, since they soon create multiple bound states, and one needs to add up all energy deposition in a cone to learn about the initial state. This is achieved by the cone algorithm in DØ. The cone algorithm comprises three steps:

1. Collect and sort the preclusters in the calorimeter towers by the transverse energy E_T in cone $\Delta R = 0.3$. Clusters with $E_T > 1$ GeV and more than one tower are considered and act as seeds. The η, ϕ and E_T of a cluster are defined as

$$\begin{array}{rcl} \eta & = & \sum_{\boldsymbol{i}} E_T^i \eta^i \\ \phi & = & \sum_{\boldsymbol{i}} E_T^i \phi^i \\ E_T & = & \sum_{\boldsymbol{i}} E_T^i sin\theta_i, \end{array} \tag{3.3}$$

   where i runs over all towers in the cluster and η, ϕ are measured with respect to the primary vertex.

2. Produce the jet candidates. Starting from the cluster with highest E_T in step 1, sum over all clusters using Eq. (3.3) in cone $\Delta R = 0.5$, where i stands for cluster number this time. A new direction of the jet is calculated, and the procedure iterates until all jet centers are found and stable. These jet candidates are called "proto-jets".

3. Merge or split each pair of proto-jets. Each pair of proto-jets with distance larger than one single cone but smaller than double the cone size are checked. If the overlap contains over 50% of the lower energy proto-jet, the two jets are merged, otherwise they are kept separate. All jets are required to have $E_T > 6$ GeV.

Further requirements are applied to the jet candidates identified by the cone algorithm. The energy fraction deposited in the EM calorimeter (EMF) must stay within 0.05 <EMF< 0.95 to exclude the jets from the EM objects. The coarse hadronic calorimeter suffers a higher noise level and the energy fraction deposited there (CHF) is required to be CHF< 0.4. There are also requirements to remove jets clustered in one or several hot cells. Hot cells are related to detector problems like hardware failure, abnormal electronic noise, etc. Their energy is typically large (> 1 GeV).

Jet energy scale (JES)

After the successful identification of jets, the jet energy scale is applied to match the measured jet energy in the calorimeter towers (calorimeter jet) to the final state particle jet - a composite of multiple stable particles (mainly hadrons). This procedure is necessary because many factors could distort the calorimeter jet energy measurement, including non-linear calorimeter responses, un-instrumented detector regions, energy radiated outside the cone algorithm, etc. For each MC simulated jet located in the calorimeter, a spatially matched particle jet in data is sought to calculate the correction factor f_{corr} [19] [20],

$$f_{corr} = \frac{\sum_{\boldsymbol{i}} R^i_{cal,data} E_i}{\sum_{\boldsymbol{i}} R^i_{cal,MC} E_i}, \tag{3.4}$$

where R^i_{cal} is the response function from MC or data for each particle type (γ, $e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, etc.), i runs over all particles in jet and E_i is the true energy of the MC particle. The MC single particle responses $R^i_{cal,MC}$ are the sum energy of the calorimeter cells contained in the $\Delta R = 0.5$ cone around the measured particle divided by its MC truth energy, parametrized as functions of the particle 4-momenta. It is measured from MC samples with single primary vertex and zero calorimeter noise. The EM object and muon responses in data, $R^i_{cal,data}$, are the same with MC. The hadron data responses introduce three more parameters to parametrize the difference between data and MC, based on selected γ+jets and di-jets events. The tuning is performed so that the ratio of the transverse momenta for the jet and the photon, $p^{jet}_{T,corr}/p^{\gamma}_T$, is the same between MC and data. The matching between the particle jet and calorimeter jet is based on their angular distance ΔR: the closest particle jet to the calorimeter jet in ΔR is considered the one making the calorimeter jet. The correction factor is then applied to correct the measured energy in jet,

$$E^{corr} = f_{corr}(E^{meas} - E_{offset}) + E_{offset}, \tag{3.5}$$

where E^{corr} is the corrected jet energy, E^{meas} is the jet energy deposition measured in the calorimeters, and E_{offset} is the energy offset for noise, pile-up and multiple interactions.

Jet energy resolution (JER)

The differences in data and MC exist in many aspects. Aside from the calorimeter responses, the jet resolution, jet reconstruction efficiencies and identification efficiencies are different between data and MC and addressed in a "JSSR" (Jet shifting, smearing and removal) method [21]. The JSSR method recalibrates, smears, and selectively discards MC simulated jets to match the jet behavior observed in real data. It uses the transverse momentum imbalance in photon+jets and Z+jets events to parametrize correction functions,

$$\Delta S = \frac{p_T^{jet} - p_T^{\gamma/Z}}{p_T^{\gamma/Z}}. \tag{3.6}$$

The purpose of JSSR is to match the quantity ΔS between data and MC. This function can be fitted by the product of a Gaussian and a error function, or just the Gaussian. The jet energy resolution is then corrected by the widths of the Gaussians in MC and data. The jet reconstruction and identification efficiencies can be corrected by the error function, which also serves as the turn-on curve in MC and data. There is a low-p_T bias in the ΔS distribution, which indicates jet reconstruction and identification inefficiency. This can be deconvoluted by fitting Gaussians to ΔS distrubution multiplied by the turn-on curves. See Fig. 3.19 for an example of ΔS distribution at low and high p_T^{γ}.

The results of these corrections are then applied to the MC simulated jets. They are oversmeared and shifted by corresponding pametrizations aimed to match MC and data. A simple cut on jet p_T at 15 GeV is imposed in both samples, where the reconstruction and identification efficiencies reach their plateau. See Fig. 3.20 for the effects of JSSR.

3.7.7 $\not\!\!E_T$ reconstruction

The missing transverse energy $\not\!\!E_T$ comes predominantly from invisible neutrinos, which carry the weak charge only and hardly interact with the medium in the detector. It is not truly measurable, but rather inferred from the momentum conservation laws. Note $\not\!\!E_T$ is a vector with x, y coordinates in the transverse plane,

$$\begin{aligned} \not\!\!E_{Tx} &= -\textstyle\sum_{\boldsymbol{i}} E_i cos\theta_i \\ \not\!\!E_{Ty} &= -\textstyle\sum_{\boldsymbol{i}} E_i sin\theta_i \\ \not\!\!E_T &= \sqrt{\not\!\!E_{Tx}^2 + \not\!\!E_{Ty}^2} \end{aligned} \tag{3.7}$$

where i runs over all EM and fine hadronic calorimeter cells (including ICD), and θ is the angle with respect to the x axis in the transverse plane of detector. The coarse hadronic calorimeter cells are not included due to their high noise levels. The energy correction to physics objects, in particular the jets (see last section), is taken into account in Eq. (3.7), i.e. the measured energy from the calorimeter cells are replaced by the corrected energy from Eq. (3.5). The muon only deposits a few GeVs of energy in the calorimeter, which is independent of its momentum. The muon p_T is therefore included directly in the $\not\!\!E_T$ calculation, while the small calorimeter energy deposition associated to the muon is removed.

The $\not\!\!E_T$ could be spurious since many sources potentially contribute to the missing energy, including measurement resolution, instrumental defects, reconstruction inefficiencies, improper pattern recognition, etc. A quality measure call the $\not\!\!E_T$ significance, $\mathcal{S}$ [39], is defined to evaluate how likely the $\not\!\!E_T$ stems from the detector resolution or the true neutrinos. It is computed by a convolution of the objects

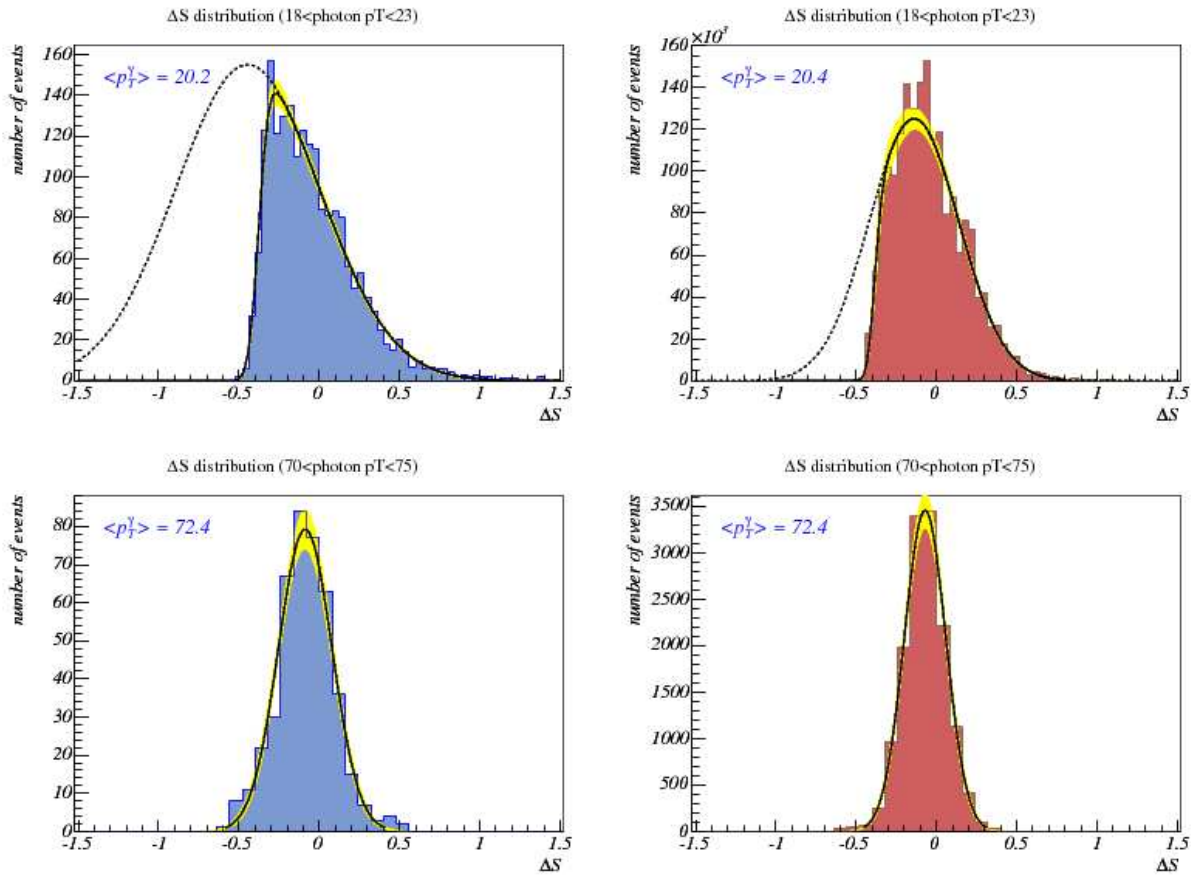


Figure 3.19: ΔS distribution in the γ+jets sample with $18 < p_T^\gamma < 23$ GeV (top) and $70 < p_T^\gamma < 75$ (bottom), for data (left) and MC simulation (right). Solid lines are the fit for the product of the error function and the Gaussian (top), and the Gaussian only (bottom). Statistical uncertainties are shown in yellow bands. The dashed curves are the extrapolations of the Gaussians in the regions affected by the turn-on (error function).

(jets, leptons, and unclustered energy) resolution functions in the event. Note $\mathcal{S}$ is typically between 0 and 1 for events with mismeasured objects, and around 6 for events with true neutrinos. See Fig. 4.10 (g) for an example.

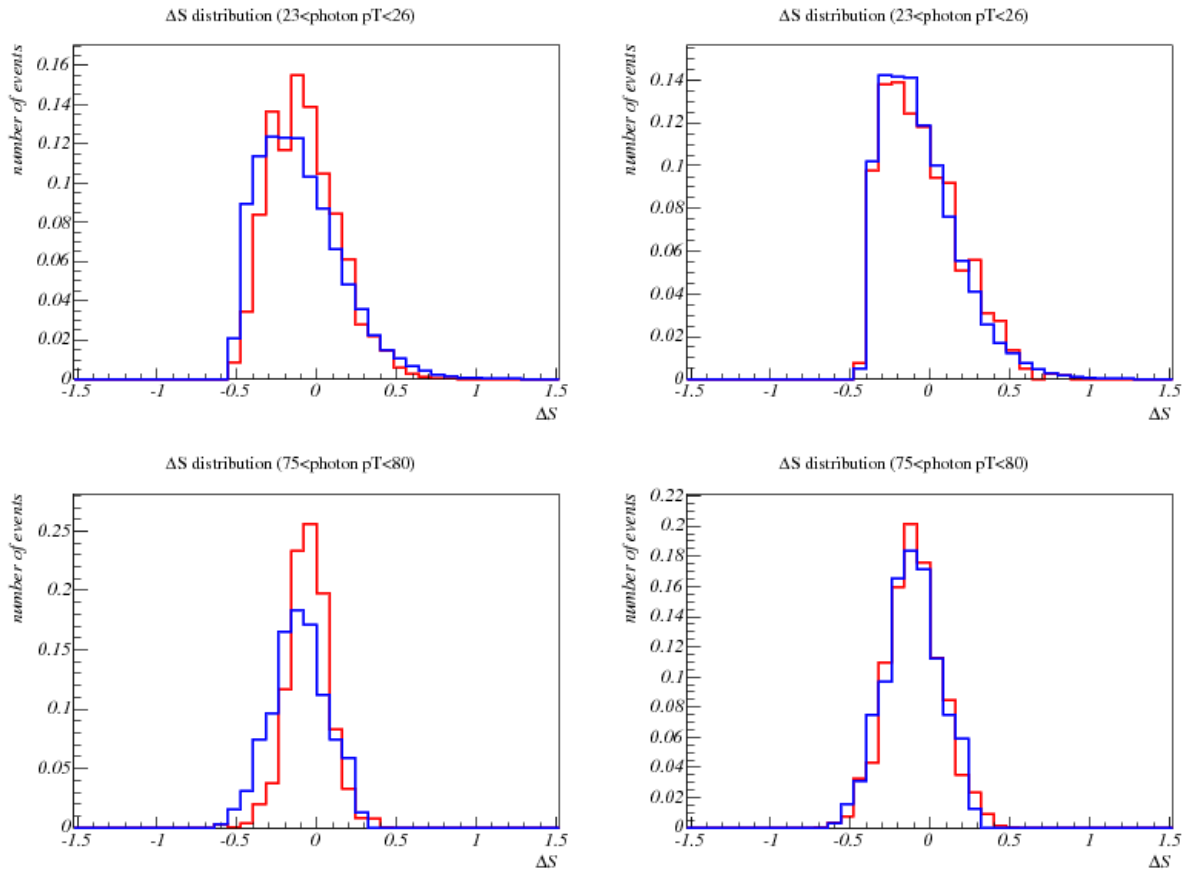


Figure 3.20: ΔS distribution in the γ+jets sample with $23 < p_T^\gamma < 26$ GeV (top) and $75 < p_T^\gamma < 80$ (bottom), for data (blue) and MC simulation (red), before (left) and after (right) JSSR.

Chapter 4

Search for the Standard Model Higgs boson at DØ in the μ + τ(hadrons) + 2 jets final state

4.1 Introduction

This chapter presents the search of the Standard Model Higgs boson through major production channels with final state μ, τ and at least two jets. We assumed a hypothesized Higgs mass between 105 and 150 GeV in 5 GeV intervals, conducted the search at each mass point, and gave a quantitave measure on the mass limit;in terms of the ratio of 95% confidence level cross section to the SM cross section. The production processes considered are:

$$q\overline{q} \to H(\to b\bar{b})Z(\to \tau^+\tau^-) \quad \text{(denoted HZ)} \tag{4.1}$$

$$q\overline{q} \to Z(\to q\overline{q})H \quad \text{(ZH)} \tag{4.2}$$

$$q\overline{q} \to W(\to q\overline{q}')H \quad \text{(WH)} \tag{4.3}$$

$$gg \to H + (\geq)2 \text{ jets} \quad \text{(gluon gluon fusion, GGF)} \tag{4.4}$$

$$q\overline{q'} \to q\overline{q'}H \quad \text{(vector boson fusion, VBF)} \tag{4.5}$$

Figure 4.1 shows the Feynman diagrams of main production channels: vector boson + Higgs associated production, gluon-gluon fusion and vector boson fusion. Figure 4.2 gives the production rates of various channels for the Tevatron, as a function of hypothesized Higgs mass. Our analysis focuses on the channels (4.2)-(4.5) associated with 2 or more jets, which further constrains the production rate. Figure 4.3 illustrates the Higgs decay branching ratio in the mass range $M_H = 105$ - 150 GeV, within the Tevatron detection capability. We focus on the $H \to \tau\tau$ and the $H \to WW^{(*)}$ decays:

- $H \to \tau\tau$: one tau lepton decays into μ + neutrinos, and the other decays hadronically.
- $H \to WW^{(*)}$: one W boson decays to muon + neutrino, and the other decays to a tau lepton + neutrino (followed by a hadronic tau decay); or both W bosons decay into tau leptons + neutrinos, followed by muonic and hadronic decays, respectively.

A significant fraction of signals (VH, GGF) involve light quark jets only and so we do not require b-tagging of jets.

The major backgrounds in our analysis are W + jets (where W decays directly or indirectly to muon + neutrino and jets fake a τ), $t\bar{t}$ (to two W's and 2 jets, with subsequent W decays to μ or τ with neutrino(s)), Z + jets ($Z \to \mu\mu$ and jets fake a τ, or $Z \to \tau\tau$ with subsequent τ decays), diboson (which decays like $H \to WW^{(*)}$), and the multi-jet (MJ) background (QCD processes that produce

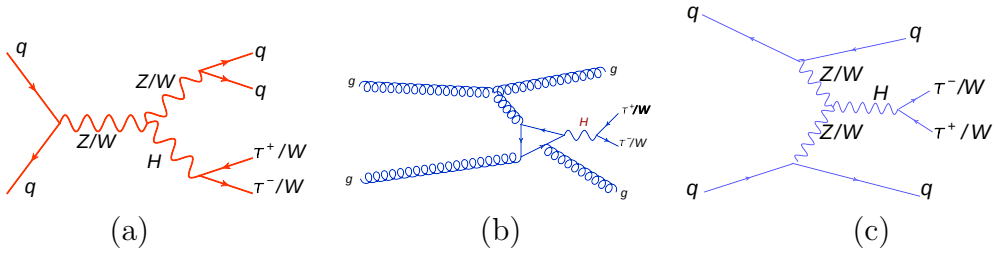


Figure 4.1: Leading order Feyman diagrams for Higgs production processes: (a) VH associated production; (b) gluon gluon fusion; (c) vector boson fusion.

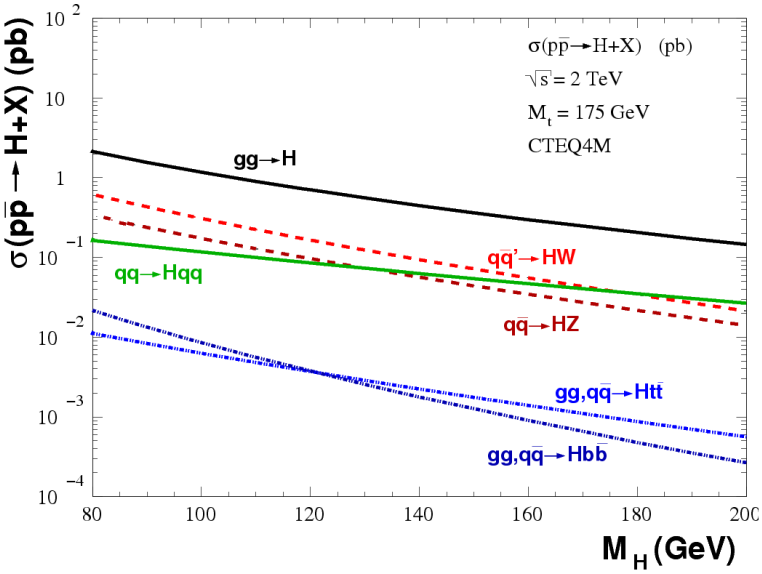


Figure 4.2: The Higgs production rates through major channels at Tevatron, as a function of M_H [40].

jets, which could fake a hadronically decaying τ, while μ comes from any available production channels). Figure 4.4 shows the legends we applied when plotting various Higgs signals and stacked backgrounds throughout the analysis.

Further details on the production rates for VH ($V = W$ or Z), GGF and VBF can be found in Refs. [22], [23] and [24] respectively. A general overview of the Higgs production processes can be found in [25].

4.2 Data set

We base our research on full Tevatron RunII data with 9.7 fb^{-1} of integrated luminosity. Given the accelerator shut-down in September 2011, this dissertation represents our best understanding of the SM Higgs in the $\mu\tau jj$ channel in DØ and

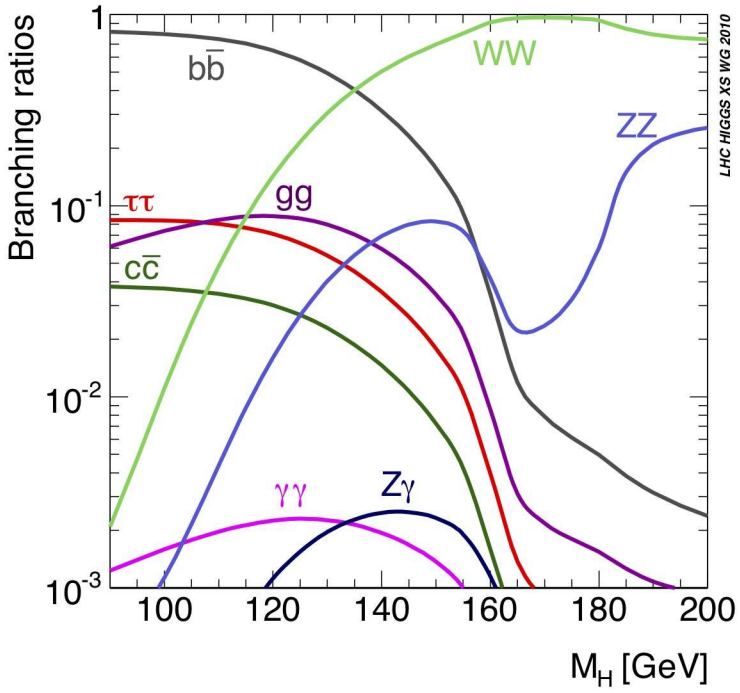


Figure 4.3: The Higgs decay major branching ratios at Tevatron, as a function of M_H [41].

the $H \to \tau\tau$ search in Tevatron. We followed the criteria recommended by the DØ Common Sample Group. Each event is required to have at least one "loose" quality muon with p_T of at least 8 GeV. We removed runs labeled bad by the CFT, SMT, CAL or MUON quality groups. Bad flagged events and data within bad luminosity blocks according to the CAL group were also rejected.

4.3 Triggers

We applied the SingleMuonOR trigger in the $\mu\tau jj$ analysis. The SingleMuonOR trigger is an OR of various muon triggers requiring a muon of medium quality in the event (see Sec. 3.6 for an overview of the trigger system and Sec. 3.7.4 for definitions of muon qualities). To increase event yield, we collect events from all active triggers (ALLTRIG) and adjust the trigger efficiency of MC events to data by the ratio of event numbers in data and MC. This was done in the RunIIa and IIb epochs separately. We applied the signal sample selections (see Sec. 4.5) to all samples and plotted the ALLTRIG/SingleMuonOR trigger ratio dependence on a few important kinematic variables. The MJ background is subtracted from both ALLTRIG and SingleMuonOR after selections, since the scaling is desired for MC

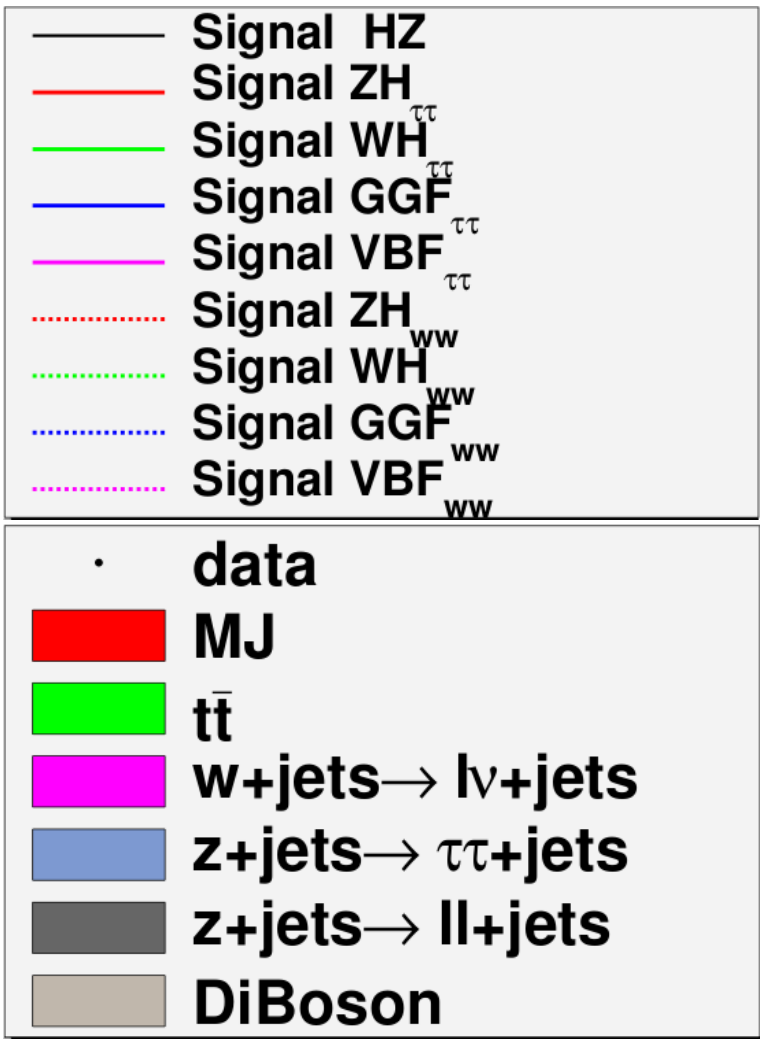


Figure 4.4: The legends for various Higgs signals and backgrounds throughout the analysis.

events only. Appendix A gives a detailed description on the fitting procedure and the kinematic distributions. We observed no prominant dependence in the RunIIa ratio on any kinematic variables, a constant fit was hence adopted to describe the scaling for each τ type. The RunIIb ratio shows a parabolic dependence on muon η.

The SingleMuonOR trigger efficiency in the $\mu\tau jj$ analysis is determined by a tag and probe analysis on the $Z \to \mu\mu$ events as a function of muon p_T, η and ϕ. The yield enhancement is about 40% from the SingleMuonOR trigger to ALLTRIG.

Other available trigger combinations include Muon plus Jets, and Muon plus Tau ORs. They were certified toward the end of the final DØ limit combination. Although a 20% improvement in event yield is expected due to inclusion of multiple triggers in the OR combination, we reluctantly gave them up in order to converge before the deadline for publication.

4.4 Monte Carlo simulation

We generated all signals and backgrounds (other than MJ background) to investigate their kinematic variable distributions. It is of extreme importance to study the difference between the signals and backgrounds from pure MC samples and apply the knowledge to 9.7 fb^{-1} data. The signal samples were created with PYTHIA 6 [26]. 10 mass points of signals were generaged between $M_H = 105$ and 150 GeV with a 5 GeV increment. The cross sections of these signals were normalized to next-to-next-to-leading (NNLO) order using CTEQ6l parton distributions functions [27]. The decay branching ratios of the SM Higgs were computed with the HDECAY program [28], and TAUOLA [30] is used to compute the branching ratio for the τ decay.

A variety of processes that potentially mimic a Higgs signal, the backgrounds, were also generated by MC simulation. The $t\bar{t}$ and W/Z+jets MC background samples were generated using ALPGEN [29], a leading order (LO) parton level generator with subsequent PYTHIA parton shower and hadronizations. The diboson (WW, WZ and ZZ) background events were created with PYTHIA. The MJ background is estimated with data as described in Sections 4.6. TAUOLA is used to simulate the tau polarization and decay. We then passed all generated events through the full GEANT3 [31] detector simulation, digitization (D0SIM), and reconstruction (D0RECO).

PYTHIA calculates the cross sections of all MC processes at the tree level only. We corrected the leading order (LO) cross sections to the next-to-leading order (NLO) by the k-factor in all samples we generated [32] [33]. Table 4.1 gives the k-factors we applied. We did not consider the heavy flavor factors as the majority of processes under consideration involve light quarks only. The cross-sections of WW, WZ and ZZ are normalized to NLO [34], where the k-factors quoted are for Run2b1; other epochs have up to a few percent changes.

The MC simulations do not match the data perfectly and a number of corrections were applied to bring them into better agreement. An event in the data sample can have more than one high energy interactions (ZB overlay). To mimic the ZB overlay in the MC simulation, we reweighted the luminosity profile by the ratio of data to normalized MC events as a function of run range. The beam profile was also reweighted as a function of instantaneous luminosity to correct distribution of the longitudinal position of the primary vertex using data [35]. The p_T distribution of the observed Z boson is not well represented in MC, so we used the Z+jets ALPGEN+PYTHIA samples to apply an n_{jet}-dependent weight to each MC event [36]. Similarly, the MC W + jets samples were reweighted too [32]. We also applied a reweight as a function of leading and sub-leading jet pseudorapidities η_1, η_2 to all MC events. We took the Higgs boson cross sections for the VH, GGF and VBF processes from Higgs Group standard values [37].

processes				
$t\bar{t}$	lepton + jets	dilepton		
	1.39	1.45		
Wjets	1.266			
Zjets, $Z \to \mu\mu$	$M_{\mu\mu}$=15-75	$M_{\mu\mu}$=75-130	$M_{\mu\mu}$=130-260	$M_{\mu\mu}$=260-1960
	1.256	1.264	1.33	1.256
Zjets, $Z \to \tau\tau$	$M_{\tau\tau}$=15-75	$M_{\tau\tau}$=75-130	$M_{\tau\tau}$=130-260	$M_{\tau\tau}$=260-1960
	1.275	1.29	1.281	1.3

Table 4.1: The k-factors for MC $t\bar{t}$, Wjets and Zjets backgrounds to correct the cross sections from LO to NLO.

4.5 Object selection

We present the details of our $\mu\tau jj$ object selection criteria in this section. The convention on the kinematic variables for the leading (with largest p_T) and subleading (with second largest p_T) jets is to add 1 or 2 after a variable, i.e. p_{T1}, p_{T2} refer to the leading, subleading jet p_T's, η_1, η_2 refer to the leading, subleading jet psudorapidities, etc. To force othogonality with the analysis on the $H \to WW \to e\mu$ search, we allow no events with an electron of $p_T^e > 15$ GeV.

4.5.1 Muon selection

The muon chamber identifies and reconstructs muons by certain hit patterns in the scintillation counters and the proportional drift tubes, see Sec. 3.7.4 for details. The successful muon candidates are also required to match to a central track. The good muons meet the following criteria:

- Require a muon of medium (Med_Nseg3) quality, with $p_T^\mu > 12$ GeV, $|\eta^\mu| < 2.0$ in RunIIa, and $p_T^\mu > 15$ GeV and $|\eta^\mu| < 1.6$ in RunIIb. RunIIa does not require vertex confirmation and results in poor modeling in events with low jet energy; $H_T > 80$ GeV is hence required for RunIIa in addition. See Sec. 3.7.4 for definitions of muon qualities.

- Matched to a track of TrackNewMedium with $p_T > 15$ GeV for RunIIb; we require $p_T > 12$ GeV in RunIIa.

- $|\eta_{\rm det}| < 1.6$ for RunIIb, $|\eta_{\rm det}| < 2.0$ for RunIIa to match respective trigger requirements.

- Muons in the poorly instrumented bottom region of the muon system are removed. See Fig. 3.15 for the layout of the muon system; the PDT's and MDT's are not instrumented in the bottom of the muon chamber and do not supply precise muon information in the bottom region.

- Required to satisfy the following (called "NP tight") isolation requirements based on calorimeter energy and track momenta around the muon:

    - The sum of transverse energies of the calorimeter cells in a hollow cone around the muon 4-momenta is computed as:

$$E_{\rm iso} = \sum_{{\rm cells},i} E_T^i < 2.5 \text{ GeV} \quad \text{in } 0.1 < \Delta R < 0.4, \tag{4.6}$$

      where $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ is the distance in azimuth ϕ and pseudorapidity η;

- The sum of the transverse momenta of all tracks within a cone of $\Delta R = 0.5$ around the muon, excluding the muon track itself, is computed as:

$$p_T^{\text{iso}} = \sum_{\text{tracks},i} p_T^i < 2.5 \text{ GeV} \tag{4.7}$$

The isolation requirements are altered below for the special multijet background samples.)

- No additional loose muon (without the isolation requirement in Eq. (4.6) and Eq. (4.7)) with $p_T > 10$ GeV and $|\eta| < 2.0$ to ensure othogonality with the $Z(\mu\mu) + H$ and $\mu\mu + \not{E}_T$ searches.

- Muons with a scintillator hit time more than 10 ns different from expectation in typical collisions are attributed to cosmic rays and rejected.

The muon selection efficiencies between data and MC samples are not always consistent and require certain corrections to the MC. We performed the tag-and-probe method on both data and Monte Carlo Z $\rightarrow \mu^+\mu^-$ samples to compute the correction factors.

4.5.2 Tau selection

The Tau ID group defined the τ identification upon reconstruction. There are three types of τ decays into hadrons:

- **Type 1** (based upon $\tau^\pm \rightarrow \pi^\pm\nu$): Calorimeter cluster, with one associated track and no EM sub-cluster.

- **Type 2** (based upon $\tau^\pm \rightarrow \rho^\pm\nu \rightarrow \pi^\pm\pi^0\nu$): Calorimeter cluster, with one associated track and at least one EM sub-cluster.

- **Type 3** (based upon $\tau^\pm \rightarrow a_1^\pm\nu \rightarrow \pi^\pm\pi^\pm\pi^\mp(\pi^0 s)\nu$): Calorimeter cluster, with more than one associated track and with or without EM sub-cluster. Candidates with two opposite sign tracks, for which the tau charge is ambiguous, are rejected.

Here the calorimeter cluster refers to the hadron cluster found by the simple cone algorithm with core cone size $\Delta R = 0.3$ and isolation cone size $\Delta R = 0.5$ (see Sec. 4.5.3). The EM sub-cluster is identified by a nearest neighbor algorithm seeded in the EM3 layer of the calorimeter and extended to the nearest EM cells in other layers. Up to three tracks are reconstructed in cone $\Delta R = 0.5$ with $p_T > 1.5$ GeV and invariant mass consistent with τ-mass.

All τ candidates must satisfy the following criteria to be selected:

- The transverse momentum of tau leptons, $p_T^\tau >$(12.5, 12.5 or 15) GeV for type (1, 2 or 3) τ's.

- The tau lepton pseudorapidity in the calorimeter with respect to the detector coordinates, $|\eta^{\tau}_{\rm det}| < 2.0$;
- The track transverse momentum, p_T^{trk} >(7, 5) GeV for type (1, 2) τ's. Type 3 τ's must have at least one track with $p_T > 5$ GeV and the sum of all track p_T's $\Sigma p_T^i > 7$ GeV.
- p_T^{trk}/E_T^{τ} >(0.65, 0.5 or 0.5) for type (1, 2, or 3) τ's.
- The sum of charges for all tracks associated with each type 3 τ must be non-zero. This is to guarantee all events have the same or opposite charged lepton and tau. The tau lepton must have opposite sign to the muon in the Higgs signals from $H \to \tau\tau$ or $H \to WW^{(*)}$ process.
- The p20 τ identification Neural Network output NN_{τ} >(0.92, 0.9, or 0.91) for type (1, 2 or 3) τ's.
- If two or more taus pass the signal sample selection, we chose the one with highest p_T^{τ}.

Jets misidentified as taus significantly degrade the quality of tau selection. In order to separate the tau events and the jets faking taus, a Neural Network, $\rm NN_{\tau}$ is developed to model the likelihood corresponding events are indeed real tau candidate events [38]. Multiple variables are used to distinguish the tau signals and backgrounds:

- $(E^{EM1} + E^{EM2})/E^{\tau}$, where E^{EM1}, E^{EM2} are energy depositions in the first and second layer of the EM calorimeter, E^{τ} is the energy of the cluster in the cone $\Delta R < 0.5$.
- $\sum p_T^{trk} / \sum p_T^{\tau_{trk}}$. $\sum p_T^{trk}$ is the p_T of all tracks in the cone $\Delta R < 0.5$, $\sum p_T^{\tau_{trk}}$ is the sum of the transverse momenta of tracks associated with the tau candidate.
- Fine hadronic fraction, fraction of E_T in the hadronic part of the calorimeter.
- $E_T^{\tau}/(E_T^{\tau} + \sum p_T^{trk})$. E_T^{τ} is the energy deposition of the cluster.
- $\sqrt{(\Delta\phi/sin\Theta)^2 + (\Delta\eta)^2}/\pi$, where the differences are between the vector sum of τ track directions and the vector sum of the EM cluster. Used for tau types 1 and 3.
- Transverse energy of the leading EM subcluster divided by the transverse energy in the layer 3 of the calorimeter in the cone $\Delta R < 0.5$. Only used for tau type 2.

- $profile = (E_T^1 + E_T^2)/E_T^\tau$. E_T^1 and E_T^2 are the transverse energies of the two most energetic calorimeter towers. A modified profile parameter, $profile2 = profile/(0.67 + 0.22|\eta_{det}|)$, is used to remove η dependence of $profile$ when $|\eta_{det} > 1.5|$.

- E_T^{em}/E_T^τ. E_T^{em} is the transverse energy of the EM cluster. Used for tau types 1 and 2 only.

- Transverse energy of the leading τ track divided by the transverse energy of the τ.

- $\mathcal{I}_{cal} = (E_T^\tau - E_{core}^\tau)/E_{core}^\tau$, calorimeter isolation. E_T^τ is the cluster energy in the $\Delta R < 0.5$ cone, and E_{core}^τ is the cluster energy in the $\Delta R < 0.3$ cone. A modified parameter $\mathcal{I}_{cal2} = \mathcal{I}_{cal}/(1.5|\eta_{det}| - 0.5)$ is used when $|\eta_{det}| > 1$ to remove the η dependence in $\mathcal{I}_{cal}$.

- $\sqrt{\sum_{i=1}^{n}[(\Delta\phi_i)^2 + (\Delta\eta_i)^2]E_{T_i}/E_T}$, the RMS of the shower. This is the width of the calorimeter cluster of the tau.

4.5.3 Jet selection

The jets are reconstructed in the calorimeter by cell energy deposition with the RunII cone algorithm, where the cone size is set to be $\Delta R = 0.5$. The jet energy is corrected with the certified jet energy scale (JES). MC jets are corrected with the "jet shifting, smearing, and removal algorithm"(JSSR) from the DØ Jet Energy Scale Group. See Sec. 3.7.6 for more details. All jets within $\Delta R = 0.5$ of a τ candidate are removed to avoid object overcounting. A good jet is required to meet the following criteria after these corrections:

- Jet detector $|\eta| < 3.4$.

- Leading jet $p_T > 20$ GeV, subleading jet $p_T > 15$ GeV.

- Jets in RunIIb are required to be vertex confirmed, i.e. has at least 2 associated tracks from the primary vertex. RunIIa jets do not have this requirement.

4.5.4 Missing transverse energy

A large amount of energy is invisible in typical high energy processes due to production of undetectable neutrinos. Nevertheless, the transverse momentum of the neutrinos can be inferred by momentum conservation in the transverse plane. The missing transverse energy ($\not\!\!E_T$) is the vector sum of transverse energies from all calorimeter cells after all jet energy corrections introduced in the last section. Also see Sec. 3.7.7 for the muon corrections and the definition of the $\not\!\!E_T$ significance, which offers an event-by-event assessment of the likelihood the observed $\not\!\!E_T$ is physical,

given the reconstructed content of the event and known measurement resolutions. No specific cut is required for the $\not\!\!E_T$ in the $\mu\tau jj$ analysis.

4.6 Multi-jet background estimation

The multi-jet (MJ) background, contributed by various QCD processes, is the largest background in our analysis. The $p\bar{p}$ collisions produce abundant events with low-energy jets, and a jet could fake a muon or a hadronically decaying τ candidate. It is difficult to simulate these MJ events with standard Monte Carlo methods due to their complex nature, and we estimate them with data.

The MJ estimation decomposes into two steps: (1) Obtain the number of MJ background events from the signal sample. (2) Estimate the shape of the MJ background with a MJ enriched sample and normalize the sum to the number obtained in step (1). We describe the definitions of the samples involved in the following sections.

4.6.1 MJ-enriched sample definitions and methods

The signal sample, in which the muon and tau has opposite sign (OS), is the nominal sample we used to conduct the analysis; the selection criteria were depicted in detail in Sec. 4.5. In order to estimate the MJ background, we reverse the sign requirement, produce a same-sign (SS) sample, and subtract the MC SM background to estimate the MJ events. The MC SM background refers to the Monte Carlo SM backgrounds, i.e. $t\bar{t}$, Vjets and di-boson backgrounds (everything other than MJ). Note the MJ background is created from data, not MC. These two sets combined give the total background, and we want them as orthogonal as possible. The MJ background sample is slightly contaminated by the MC SM backgrounds (2-3%) with same selections, and we hence subtract the latter from the former to ensure orthogonality. The SS and OS MJ backgrounds are expected to be similar, because the QCD processes do not have a sign dependence in muon or tau. In reality we are forced to adopt a variant of this approach, since the statistics of the SS signal sample is fairly limited and does not provide enough events for an effective Higgs-background discrimination. We construct an orthogonal sample, a "MJ control sample", to enrich the event statistics. We denote the event number in this sample by letter M, and the event number from the signal sample by letter N in the following discussions.

The MJ control sample differs from the signal sample in that we reverse one of the selections of the muons ("bad μ, good τ"). The bad muon does not meet the isolation requirements in Eq. (4.6) and Eq. (4.7) and otherwise is the same as the muon selection. All other selections remain the same. This MJ sample gives us similar jet-faking-tau probability and allows us to add the tau Neural Network output, NN_τ as one of the distinguishing variables. We examined another sample that reversed both muon and tau selections before ("bad μ, bad τ") and decided to follow the "bad μ, good τ" sample since the additional NN_τ ranks reasonably well and serves as a good discriminator between the Higgs and background events. Bad τ's refer to the taus with Neural Network output $0.3 < \tau_{NN} < 0.9$, which is orthogonal to the signal sample tau selection in Sec. 4.5.2. The MJ enriched sample

has very limited SM background contamination and is estimated to be about 97% pure. It contains 24,965 events, whereas the signal sample contains only 990 events.

We take the MJ background shape from the MJ control sample and normalize the total number of events to the SS signal sample. Three scale factors, the ρ_i factors, are used to estimate the ratio of MJ events between the OS and SS sample,

$$N^{\rm MJ}_{\rm OS,i} = \rho_i (N^{\rm data}_{\rm SS,i} - N^{\rm SM}_{\rm SS,i}),\ i = \text{tau type } 1,2,3,$$

with

$$\rho_i = (M^{\rm data}_{\rm OS,i} - M^{\rm SM}_{\rm OS,i})/(M^{\rm data}_{\rm SS,i} - M^{\rm SM}_{\rm SS,i}) \quad ,$$

where we assumed (1) the ρ_i factors obtained from the MJ control sample is proper for the signal sample and (2) the kinematic distributions in the MJ control sample is similar to those of the MJ events in the OS signal sample.

We fit these three constant scale factors ρ_i separately in each hadronic tau type and apply them to formulate the final distribution of MJ background events. A reweighting method, suggested by the WH group, is taken to improve the modeling of the jet related variables.

4.6.2 MJ scaling factors

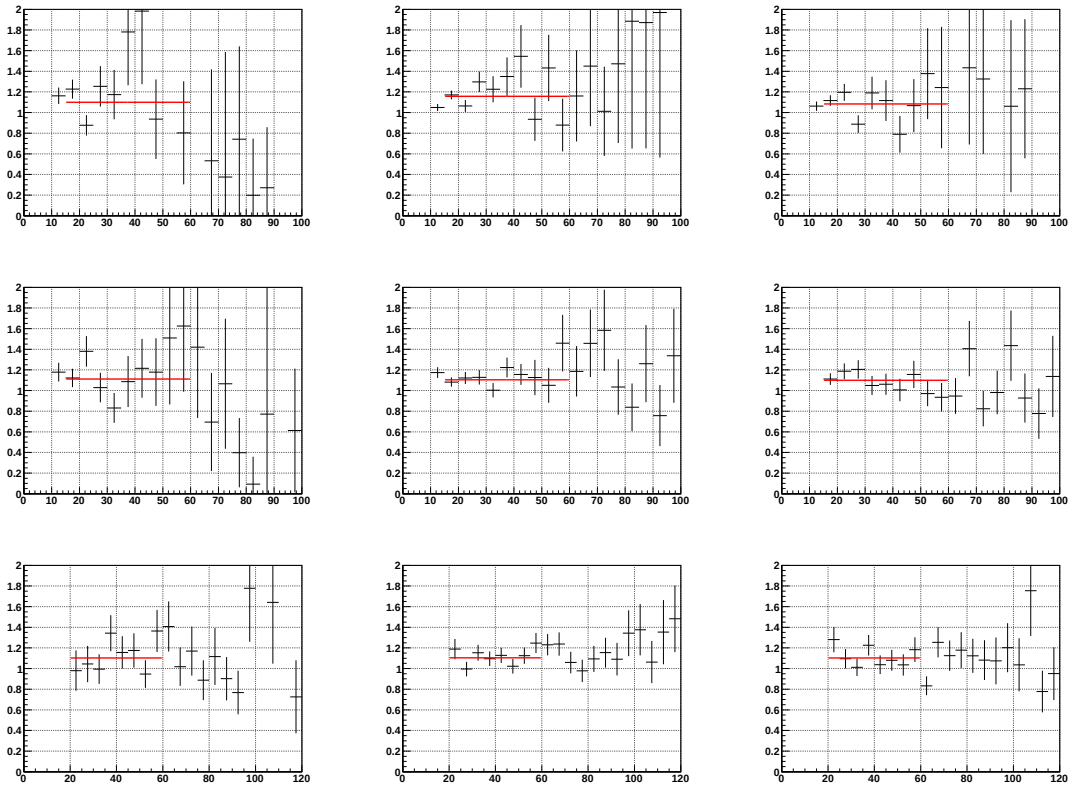
We discuss in this section the determination of the MJ scaling (ρ_i) factors. The ρ factors are used to normalize the MJ background between the signal sample and the MJ control sample. It is possible that they depend on various kinematic variables like p_T^{μ}, p_T^{τ}, $p_T^{\rm jet1}$, etc. for each tau type. We therefore plot their distributions on a few fundamental kinematic variables to investigate such potential dependences in Fig. 4.5. Figure 4.6 shows the dependences on η_{μ}, η_{τ} and $\eta_{\rm jet1}$. A selection $15 < p_T^{\mu} < 60$ GeV, $15 < p_T^{\tau} < 60$ GeV, $20 < p_T^{\rm jet1} < 60$ GeV and $|\eta_{\mu}| < 1.5$, $|\eta_{\tau}| < 1.5$, $|\eta_{\rm jet1}| < 1.5$ are applied in these fittings. The fitted ρ values are shown in Table 4.2 for the MJ control sample, RunIIa and RunIIb. We have observed no strong dependence on the kinematic variables and therefore adopted a constant fit for each tau type. The dependence on H_T is noted and addressed in Sec. 4.6.3.

4.6.3 MJ shape correction between signal and control samples

We took the shape of the MJ control sample and normalize the total number of events to that of the MJ SS sample to enrich the MJ statistics. We estimated the MJ background in the OS sample from the SS sample, since the QCD processes are typically charge blind. It is natural to question if the MJ shapes of the SS signal and MJ Control samples are similar. The pseudorapidities and angular correlation variables are indeed consistent between these two samples, however there is certain discrepancy between the jet p_T related variables, e.g. H_T, p_{T1}, etc. We therefore

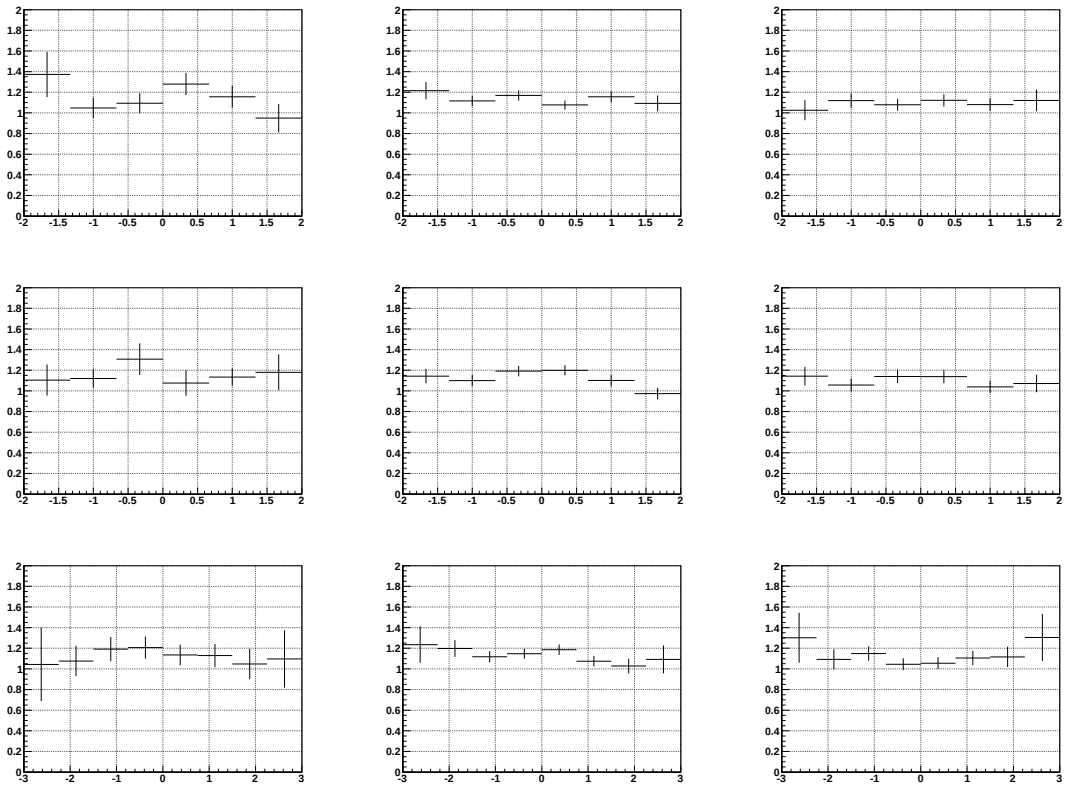
Sample	Run		τ type 1	τ type 2	τ type 3
MJ Control	IIa	N_{evnts}	656	2670	1653
		purity	0.980	0.971	0.985
		ρ	0.94±0.10	1.20±0.05	1.02±0.06
MJ Control	IIb	N_{evnts}	2554	11006	6426
		purity	0.981	0.973	0.988
		ρ	1.10±0.06	1.16±0.03	1.08±0.04

Table 4.2: The number of events, purity (fraction of events estimated to be MJ) and scale factors, by tau type, for MJ Control samples in the $\mu\tau jj$ analysis.



Type 1 τ Type 2 τ Type 3 τ

Figure 4.5: Distributions of the ratio of OS to SS events in the $\mu\tau jj$ MJ enriched sample, as function of p_T^{μ} (top), p_T^{τ} (middle) and p_T^{jet1} (bottom) for different τ types.



Type 1 τ Type 2 τ Type 3 τ

Figure 4.6: Distributions of the ratio of OS to SS events in the $\mu\tau jj$ MJ enriched sample, as function of η_μ (top), η_τ (middle) and η_{jet1} (bottom) for different τ types.

reweighted the MJ shapes of the control sample with the correction factor as function of H_T. The correction was obtained by a fit function $f(H_T) = A + Bexp(-C \times H_T)$, where the best fit values are $A = 0.067$, $B = 13.25$, $C = 0.032$. Figure 4.7 shows the dependences of the ratio of shapes for SS/MJ samples after the reweight, which demonstrates significant improvement on the modeling of jet p_T related distributions.

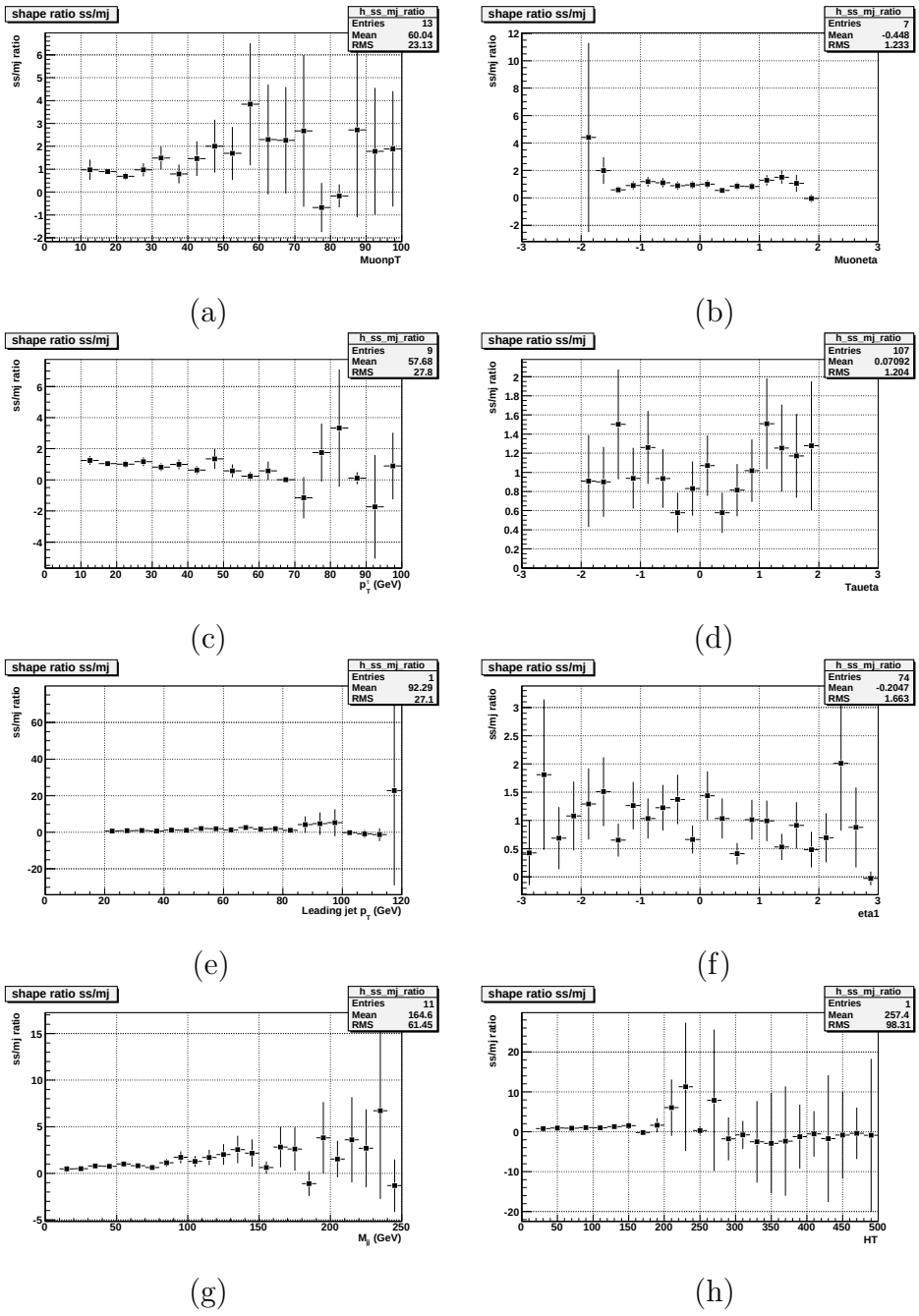


Figure 4.7: Ratios of kinematic distributions for the $\mu\tau jj$ analysis using the SS signal sample, summed over tau types, relative to using the MJ enriched sample to obtain the shapes of the MJ background. (a) p_T^{μ}; (b) η_{μ}; (c) p_T^{τ}; (d) η_{τ}; (e) p_T^{jet1}; (f) η_{jet1}; (g) m_{jj}; (h) H_T. These plots are made after the H_T correction discussed in the text.

4.7 The event yields upon preselections

The data, SM and MJ background event yields after selections discussed in the last section are shown in Table 4.3 by tau types. The MC event yields for individual signals in two subsamples are shown in Tables 4.4, 4.5 and Fig. 4.8 for hypothesized Higgs mass 105-150 GeV. See Sec. 4.10.2 for details of the subsample definition and separation.

τ type	$t\bar{t}$	W+jets	$Z_{\mu\mu}$+jets	$Z_{\tau\tau}$+jets	di-boson	MJ	Σbkgd	Data
type 1	15.3	10.2	4.4	37.1	2.3	39.1	108.4	119
type 2	121.3	65.2	29.3	241.8	14.5	135.4	607.5	684
type 3	20.2	39.1	4.4	54.5	3.2	50.6	172.1	187
All	156.9	114.5	38.1	333.4	20.0	225.1	888.0	990

Table 4.3: The RunIIa and RunIIb combined event yields for the expected SM and MJ backgrounds in the $\mu\tau jj$ analysis, for both the total and individual tau types.

m_H	HZ	$ZH_{\tau\tau}$	$WH_{\tau\tau}$	$GGF_{\tau\tau}$	$VBF_{\tau\tau}$	ZH_{WW}	WH_{WW}	GGF_{WW}	VBF_{WW}	Total
105	0.138	0.431	0.662	0.607	0.351	0.006	0.006	0.010	0.001	2.212
110	0.120	0.392	0.622	0.537	0.351	0.013	0.010	0.007	0.002	2.054
115	0.111	0.360	0.538	0.570	0.327	0.020	0.020	0.006	0.005	1.957
120	0.092	0.300	0.451	0.463	0.287	0.027	0.031	0.011	0.009	1.673
125	0.072	0.255	0.388	0.436	0.254	0.033	0.046	0.014	0.027	1.525
130	0.058	0.210	0.291	0.353	0.213	0.061	0.061	0.020	0.021	1.287
135	0.042	0.152	0.222	0.290	0.172	0.073	0.109	0.022	0.048	1.130
140	0.030	0.111	0.164	0.216	0.131	0.095	0.107	0.041	0.034	0.927
145	0.020	0.079	0.105	0.160	0.091	0.095	0.127	0.048	0.038	0.763
150	0.013	0.045	0.069	0.093	0.061	0.092	0.146	0.059	0.044	0.621

Table 4.4: Number of events for each signal/decay channel expected after selections in the $\mu\tau jj$ $\tau\tau$ subsample as a function of Higgs boson mass.

m_H	HZ	$ZH_{\tau\tau}$	$WH_{\tau\tau}$	$GGF_{\tau\tau}$	$VBF_{\tau\tau}$	ZH_{WW}	WH_{WW}	GGF_{WW}	VBF_{WW}	Total
105	0.054	0.039	0.044	0.053	0.020	0.020	0.034	0.086	0.013	0.364
110	0.047	0.026	0.033	0.026	0.014	0.040	0.066	0.088	0.026	0.366
115	0.038	0.019	0.027	0.026	0.012	0.074	0.129	0.074	0.043	0.442
120	0.032	0.015	0.018	0.014	0.010	0.122	0.192	0.142	0.089	0.634
125	0.026	0.011	0.015	0.013	0.009	0.156	0.322	0.161	0.116	0.828
130	0.021	0.010	0.012	0.013	0.007	0.212	0.411	0.247	0.175	1.108
135	0.016	0.006	0.007	0.009	0.007	0.256	0.521	0.287	0.193	1.302
140	0.010	0.005	0.007	0.009	0.005	0.326	0.570	0.331	0.298	1.561
145	0.008	0.004	0.005	0.007	0.004	0.371	0.683	0.543	0.327	1.951
150	0.004	0.003	0.003	0.004	0.003	0.408	0.659	0.557	0.392	2.034

Table 4.5: Number of events for each signal/decay channel expected after selections in the $\mu\tau jj$ WW subsample as a function of Higgs boson mass.

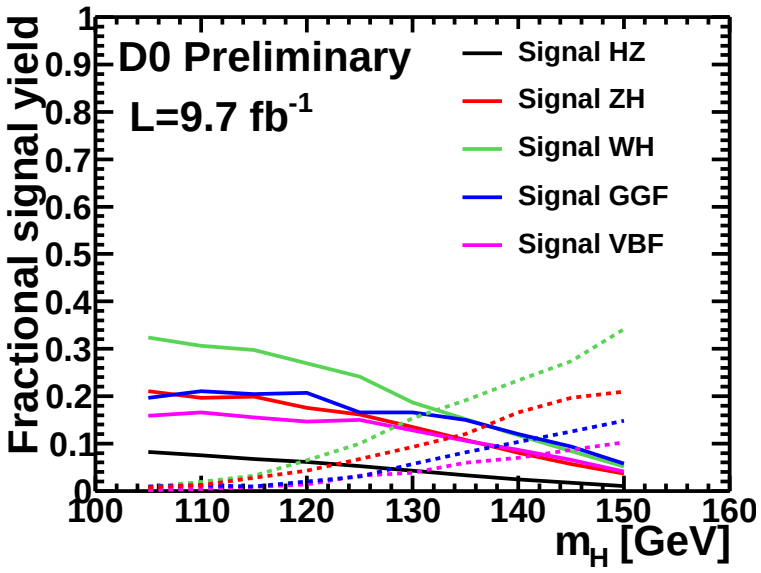


Figure 4.8: Fractional yields for signals in the $\mu\tau jj$ sample as a function of Higgs boson mass. Solid lines: $H \to \tau\tau$ signals, dashed lines: $H \to WW^{(*)}$ signals.

4.8 BDT input variables

4.8.1 Introduction

The DØ detector has collected 9.7 fb^{-1} of data in RunIIa and RunIIb1-4 epochs. The majority of these events are backgrounds (around 1000 after selections), and the number of authentic Higgs signal events in the $\mu\tau jj$ channel is estimated to be only around 2 or 3 by Monte Carlo simulation. It is difficult to identify these Higgs events by a cut on any single variable.

A methodology employing correlations among variables, the multivariate analysis, has been adopted as standard to synthesize all information from various kinematic variables. It employs boosted decision trees (BDT) to train the Higgs signals against the backgrounds, see Sec. 4.10 for a full description. Multiple variables are taken to supply information needed to serve the purpose. We identify the variables that seem most promising to distinguish various Higgs and background events in the data pool, plot the signal, background and data distributions on each of these variables, reject the poorly modeled ones, and take the rest as input into the training. The resulting output distribution, the final discriminant (FD), is input into a standardized program (COLLIE) to determine the Higgs mass limits as multiple of Standard Model Higgs production cross section at 95% confidence level.

We searched various DØ notes and papers to decide on the most appropriate variables for the Higgs signal-background discrimination. In particular, we checked the $H \rightarrow W^+W^- \rightarrow l^+\nu l^-\bar{\nu}$, $ZH \rightarrow l^+l^-b\bar{b}$ and $\nu\bar{\nu}b\bar{b}$, $WH \rightarrow l\nu b\bar{b}$ analyses, obtained a list of variables with best discrimination power, and adopted those with similar potential in our analysis. A CDF paper on the Missing Mass Calculator (MMC) [1] has been very helpful in identifying new variables too. We also adopted the spin correlation variables as advised in [14].

The variables we used fall into 3 major categories:

- The fundamental kinematic variables. Object p_Ts, pseudorapidities, etc.
- Derivatives of the first type, e.g. dijet mass, transverse mass of objects, etc.
- The variables from the MMC technique, e.g. the MMC $\tau\tau$ mass, number of solutions, etc.

A general observation is that the more variables, the better the discrimination power and Higgs mass exclusion. It is equally important to make sure the variables implemented is well modeled. We define a figure of merit ξ, taken as the mean of the Kolmogorov-Smirnov test value and the χ^2 probability, as the criterion of goodness in modeling. In general, we only adopt kinematic variables with ξ greater than 0.25. We raised the criterion somewhat and did not find conclusive evidence that there is a better requirement.

Substantial amount of work was devoted to the MMC algorithms. We will present further details in Sec. 4.9.

4.8.2 Typical kinematic variables

The variables we considered include:

1. p_T^ℓ: p_T of the lepton candidate.
2. p_T^{j1}: p_T of the leading jet candidate.
3. $\not\!\!E_T$: missing transverse energy.
4. $M_{\tau\tau}$: invariant mass of the τ_ℓ – $\tau_{\rm had}$ system. The $\not\!\!E_T$ is apportioned to the ℓ and τ as follows. The $\not\!\!E_T$ is projected on the axis, $\hat{A}$, taken as the direction $\vec{p}_T^{\,\ell} - \vec{p}_T^{\,\tau}$. to give projection MET$_A$. The $\not\!\!E_T$ associated with the ℓ is $(1 +$ MET$_A)/2$ and that associated with the τ is $(1 -$ MET$_A)/2$. The neutrino p_z is estimated by assuming the p_z/p_T is the same for neutrino and the associated visible products of either tau.
5. $M_{\rm jj}$: invariant mass of the two candidate jets.
6. $\Delta R_{\rm jj}$: $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ is the distance in azimuth ϕ and pseudorapidity η between the two leading jets.
7. $M_{\rm T}^\ell$: transverse mass calculated from the p_T^ℓ and $\not\!\!E_T$. $M_{\rm T}^\ell = \sqrt{2E_T^\ell \not\!\!E_T - 2\mathbf{p}_\mathbf{T}^\ell \cdot \not\!\!\mathbf{E}_\mathbf{T}}$.
8. $M_{\rm T}^\tau$: transverse mass calculated from the p_T^τ and $\not\!\!E_T$. $M_{\rm T}^\tau = \sqrt{2E_T^\tau \not\!\!E_T - 2\mathbf{p}_\mathbf{T}^\tau \cdot \not\!\!\mathbf{E}_\mathbf{T}}$.
9. H_T: scalar sum of the p_T of all jets with $p_T > 15$ GeV and $|\eta| < 3.4$.
10. S_T: the scalar sum of the p_T of the lepton candidate, the tau candidate, the two candidate jets and of the event missing transverse energy, $S_T = p_T^\ell + p_T^\tau + p_T^{jet1} + p_T^{jet2} + \not\!\!E_T$.
11. V_T: the magnitude of the vector sum of the p_T's of the lepton candidate, the tau candidate, the two candidate jets and the event missing transverse energy, $V_T = |\overrightarrow{p_T^\ell} + \overrightarrow{p_T^\tau} + \overrightarrow{p_T^{jet1}} + \overrightarrow{p_T^{jet2}} + \not\!\!E_T|$
12. $A(\not\!\!E_T, \not\!H_T)$: Asymmetry between $\not\!\!E_T$ and $\not\!H_T$, $(\not\!\!E_T - \not\!H_T)/(\not\!\!E_T + \not\!H_T)$, where $\not\!H_T$ is the missing H_T, defined as the magnitude of the vector sum of all jet p_T's.
13. $\Delta\phi(\not\!\!E_T, \not\!T_T)$: the azimuthal angle difference between $\not\!\!E_T$ and the missing transverse momentum, $\not\!T_T$, calculated as the negative of the vector sum of the p_T of all tracks with at least 8 CFT hits and a DCA to the primary vertex of less than 2 mm.

14. $\min \Delta\phi(\not{E}_T, jets)$: the minimum azimuthal angle difference between the $\not{E}_T$ and any good jet p_T.

15. $\mathcal{S}$: the missing E_T 'signficance' [39]. See Sec.3.7.7 for more details.

16. $\Delta\eta$(jj): the absolute value of the pseudorapidity difference between the two leading jets.

17. p_T^τ: the transverse momentum of the hadronic tau candidate.

18. $\mathcal{C}$: Centrality $= H_T/H_E$, where H_E is the scalar sum of the energies of all jets with $p_T > 15$ GeV and $|\eta| < 2.5$.

19. $\Delta\phi_{\ell\tau}$: the azimuthal angle difference between the lepton and tau.

20. $\mathcal{A}$: Aplanarity [42], constructed from the momentum tensor of the ℓ, τ and jet candidates.

21. $\cos\theta^*$: the spin correlation. The quantity θ^* is the angle between the leading+subleading jet system and the proton beam in the lab frame.

22. $\Delta\phi(\ell\tau, j_1 j_2)$: the azimuthal angle difference between the ℓ-τ system and the two leading jets system in the 3-dimensional space.

23. $M_{\mathrm{T,min}}(\not{E}_T, \ell/\tau)$: the minimum of transverse mass between the $\not{E}_T$ and ℓ or τ.

24. $\Delta\phi_{min}(\not{E}_T, \ell/\tau)$: the minimum of azimuthal angle difference between the $\not{E}_T$ and ℓ or τ.

25. $\Delta\phi_{max}(\not{E}_T, \ell/\tau)$: the maximum of azimuthal angles difference between the $\not{E}_T$ and ℓ or τ.

26. $M_{\mathrm{T}}(\not{E}_T, \ell + \tau)$: the transverse mass between the $\not{E}_T$ and the ℓ-τ system.

27. $p_T(\ell, \tau, \not{E}_T)$: p_T of the ℓ, τ and $\not{E}_T$.

28. $M(\ell, \tau, j_1, j_2)$: the four-body invariant mass of the two leading jets and two leading leptons.

29. $\Delta\phi_{max}(\not{E}_T, j_1/j_2)$: the maximum of azimuthal angle difference between the $\not{E}_T$ and leading or subleading jet p_T.

30. $\not{H}_T/H_T$.

31. $\Delta\phi(\ell, j_1)$: the azimuthal angle difference between the lepton and the leading jet.

32. $M_{\tau\tau,\ combined}$: the invariant MMC $\tau\tau$ mass when there is solution, and the $\tau\tau$ mass by colliear approximation ($M_{\tau\tau}$) when there is none. See Sec.4.9 for a full description of the MMC technique.

33. $N_{\tau\tau\ soln.}$: the number of physical $\tau\tau$ solutions in the event.

34. $M_{WW,\ combined}$: the invariant MMC WW mass when there is solution, and WW mass estimated as $M_{\rm T}(\not\!\!E_T, \ell{+}\tau) \times {\rm p}(\ell{+}\tau)\ /\ p_T(\ell{+}\tau)$ when there is none. One W boson is on-shell while the other is off-shell most of the time, and the MMC solutions vary depending on which W is on-shell. We therefore assumed both possibilities, solved for the combined masses separately and averaged them. The MMC $\tau\tau$ mass does not bear the burden, since the tauons are light and almost always produced on-shell.

35. $N_{WW\ soln.}$: the number of physical WW solutions in the event. The average was taken for the same reason in $M_{WW,\ combined}$.

36. τ_{NN}: the τ Neural Net output.

Variables 1-20 were considered in earlier publications, the rest were added in this analysis to further enhance the discrimination power. A few of them failed our ξ >0.25 standard and were dropped along the way. The final variables in our analysis are shown in Table 4.6. The distributions of these variables in data, background and Higgs signal events are shown in Figs. 4.9-4.12. The plots evaluate a total of 9.7 fb^{-1} of data collected in both RunIIa and RunIIb epochs. Note the error bars on data points only involve the statistical uncertainty in data. In general, the data yield is somewhat higher than the backgrounds, and we estimate the discrepancy for events with type 1, 2, 3 taus to be 0.80σ, 1.42σ, 0.81σ, respectively, taking into account all statistical and flat systematic uncertainties (see Table 4.12). Note there are further shape-dependent uncertainties too, see Sec. 4.11 for a full description. Taking all uncertainties into consideration, the data and backgrounds agree reasonably well.

Due to the importance and significant amount of work involved in the MMC variables 32-35, we open a new section for a more thorough discussion.

Variable no.	Variables used
1	p_T^{ℓ}
2	p_T^{j1}
5	M_{jj}:
6	ΔR_{jj}
7	M_{T}^{ℓ}
8	M_{T}^{τ}
9	H_T
10	S_T
11	V_T
12	$A(\not{E}_T, \not{H}_T)$
13	$\Delta\phi(\not{E}_T, \not{T}_T)$
14	$\min \Delta\phi(\not{E}_T, jets)$
15	$\mathcal{S}$
16	$\Delta\eta(\mathrm{jj})$
17	p_T^{τ}
21	$\cos\theta^*$
22	$\Delta\phi(\ell\tau, j_1 j_2)$
23	$M_{\mathrm{T,min}}(\not{E}_T, \ell/\tau)$
24	$\Delta\phi_{min}(\not{E}_T, \ell/\tau)$
25	$\Delta\phi_{max}(\not{E}_T, \ell/\tau)$
26	$M_{\mathrm{T}}(\not{E}_T, \ell + \tau)$
27	$p_T(\ell, \tau, \not{E}_T)$
28	$M(\ell, \tau, j_1, j_2)$
29	$\Delta\phi_{max}(\not{E}_T, j_1/j_2)$
30	$\not{H}_T/H_T$
31	$\Delta\phi(\ell, j_1)$
32	$M_{\tau\tau,\ combined}$
33	$N_{\tau\tau\ soln.}$
34	$M_{WW,\ combined}$
35	$N_{WW\ soln.}$
36	τ_{NN}

Table 4.6: Variables used for the $\mu\tau jj$ analyses in BDT training.

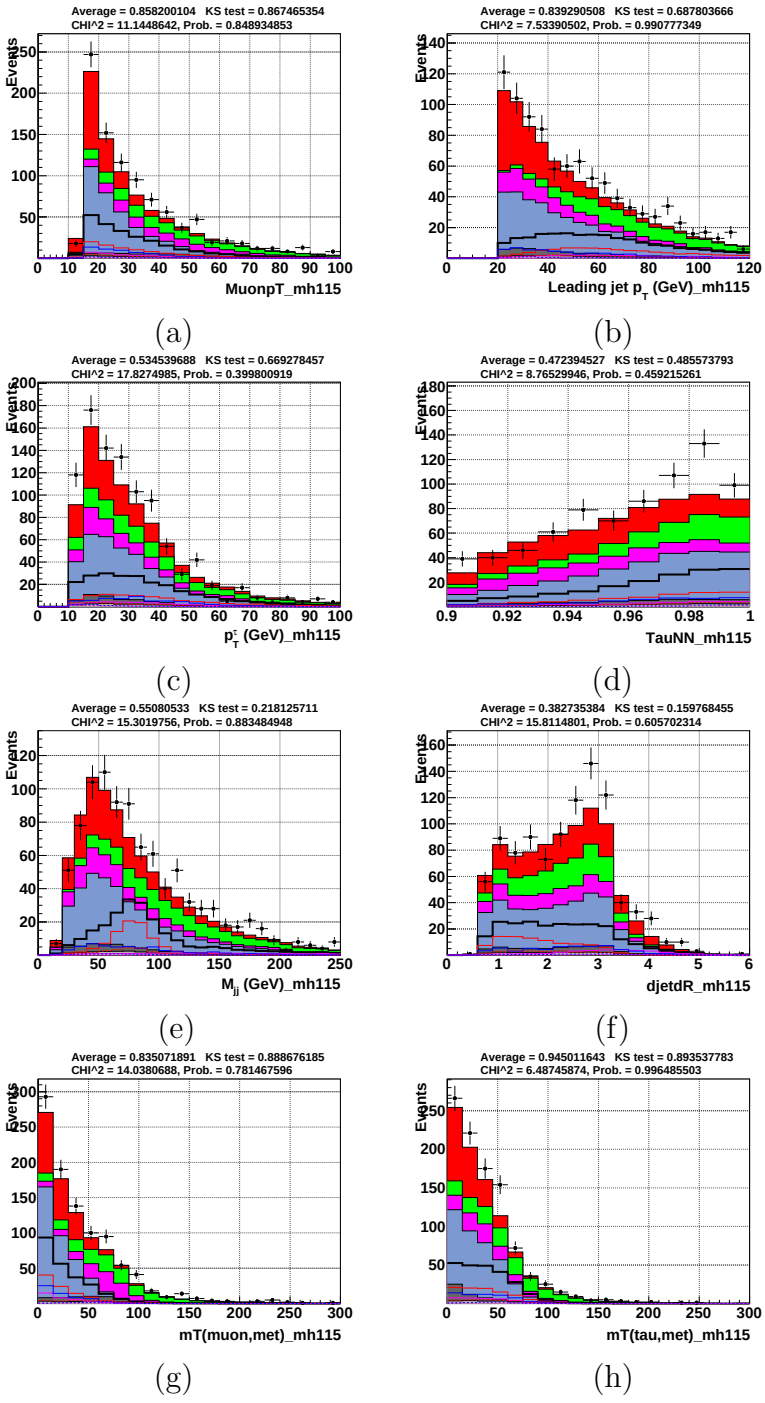


Figure 4.9: Data - MC comparison for BDT input variables: (a) p_T^{ℓ}, (b) p_T^{j1}, (c) p_T^{τ}, (d) τ_{NN}, (e) $M_{\rm jj}$, (f) $\Delta R_{\rm jj}$, (g) $M_{\rm T}^{\ell}$, (h) $M_{\rm T}^{\tau}$. Signal curves are for m_H=115 GeV multiplied by 100.

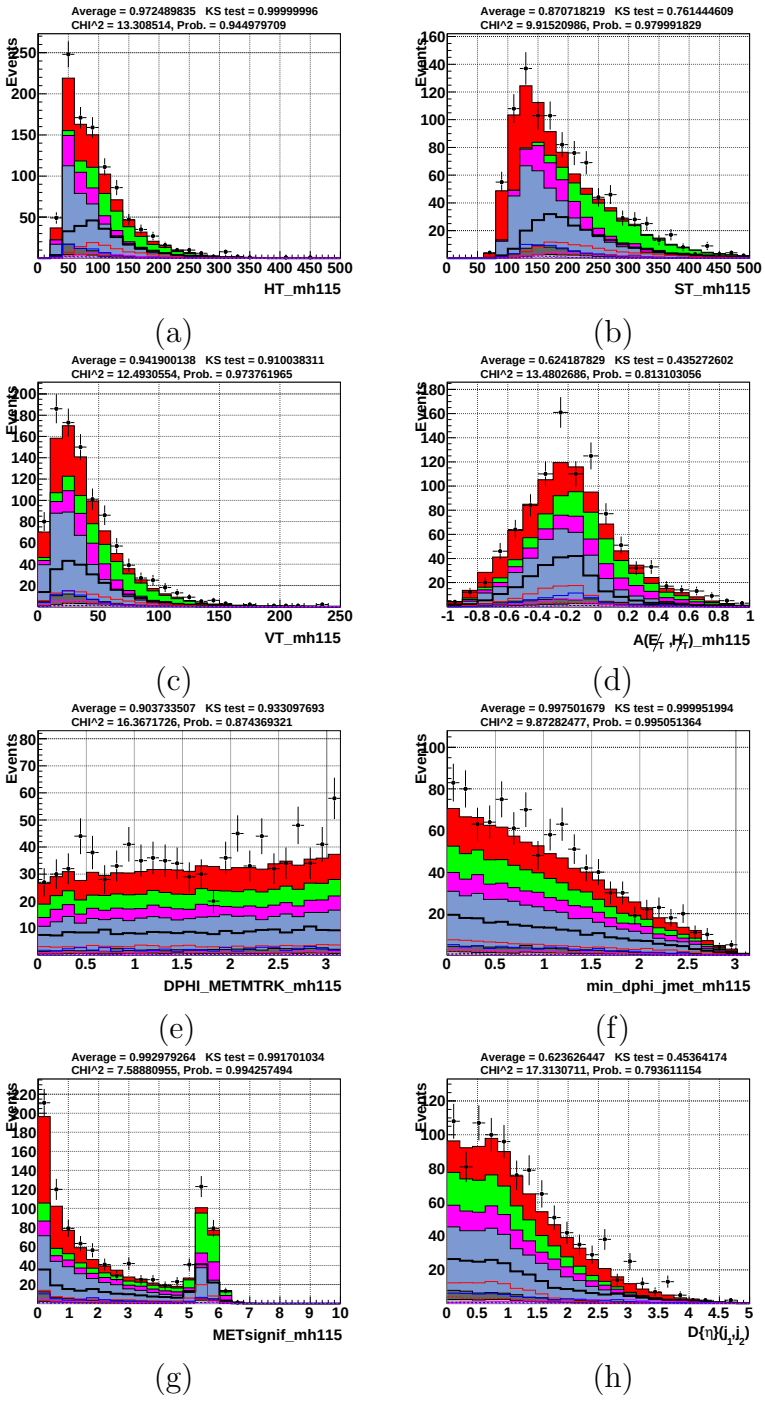


Figure 4.10: Data - MC comparison for BDT input variables: (a) H_T, (b) S_T, (c) V_T, (d) $A(\not\!\!E_T, \not\!\!H_T)$, (e) $\Delta\phi(\not\!\!E_T, T_T)$, (f) $\min\Delta\phi(\not\!\!E_T, jets)$, (g) $\mathcal{S}$, (h) $\Delta\eta$(jj). Signal curves are for m_H=115 GeV multiplied by 100.

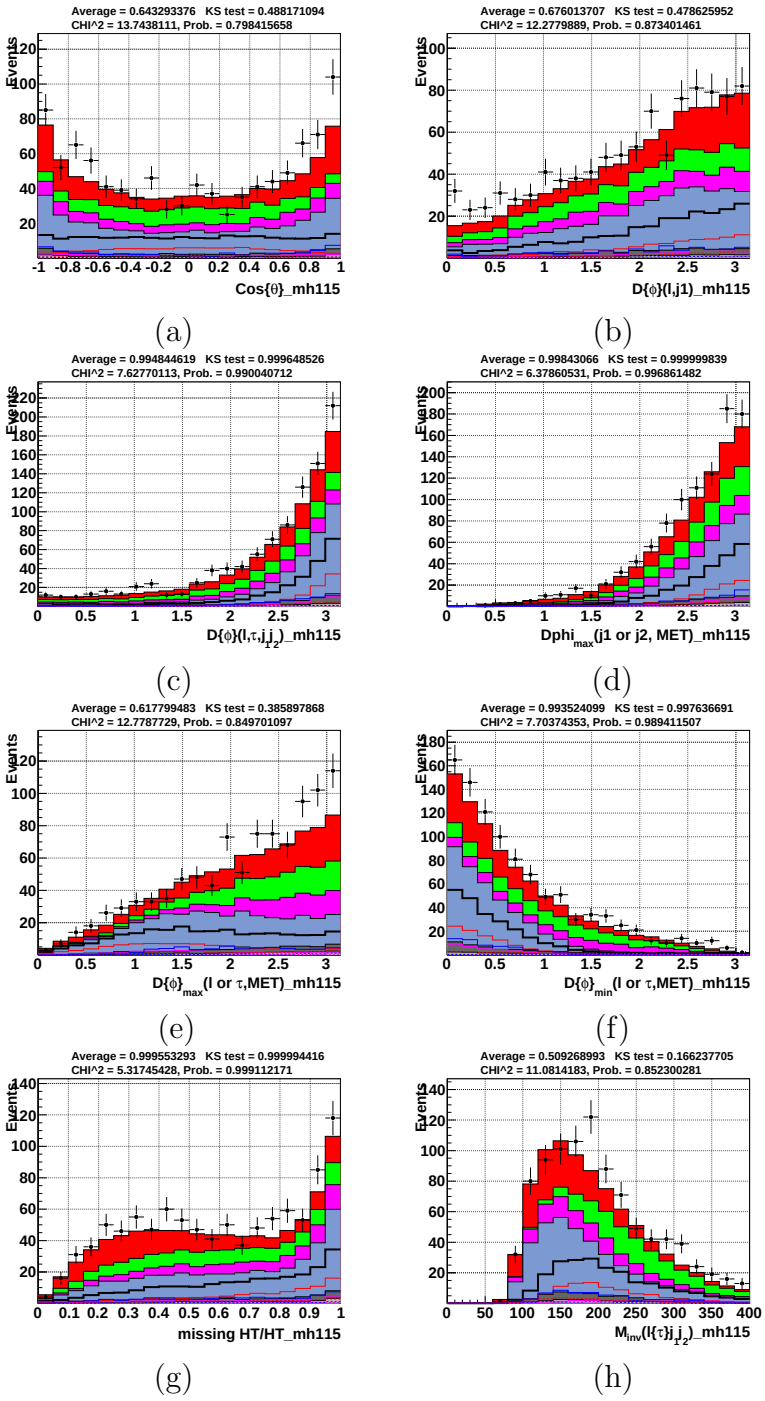


Figure 4.11: Data - MC comparison for BDT input variables: (a) $\cos\theta^*$, (b) $\min\Delta\phi(\not{E}_T, jets)$, (c) $\Delta\phi(\ell\tau, j_1j_2)$, (d) $\Delta\phi_{max}(\not{E}_T, j_1/j_2)$, (e) $\Delta\phi_{max}(\not{E}_T, \ell/\tau)$, (f) $\Delta\phi_{min}(\not{E}_T, \ell/\tau)$, (g) $\not{H}_T/H_T$, (h) $M(\ell, \tau, j_1, j_2)$. Signal curves are for m_H=115 GeV multiplied by 100.

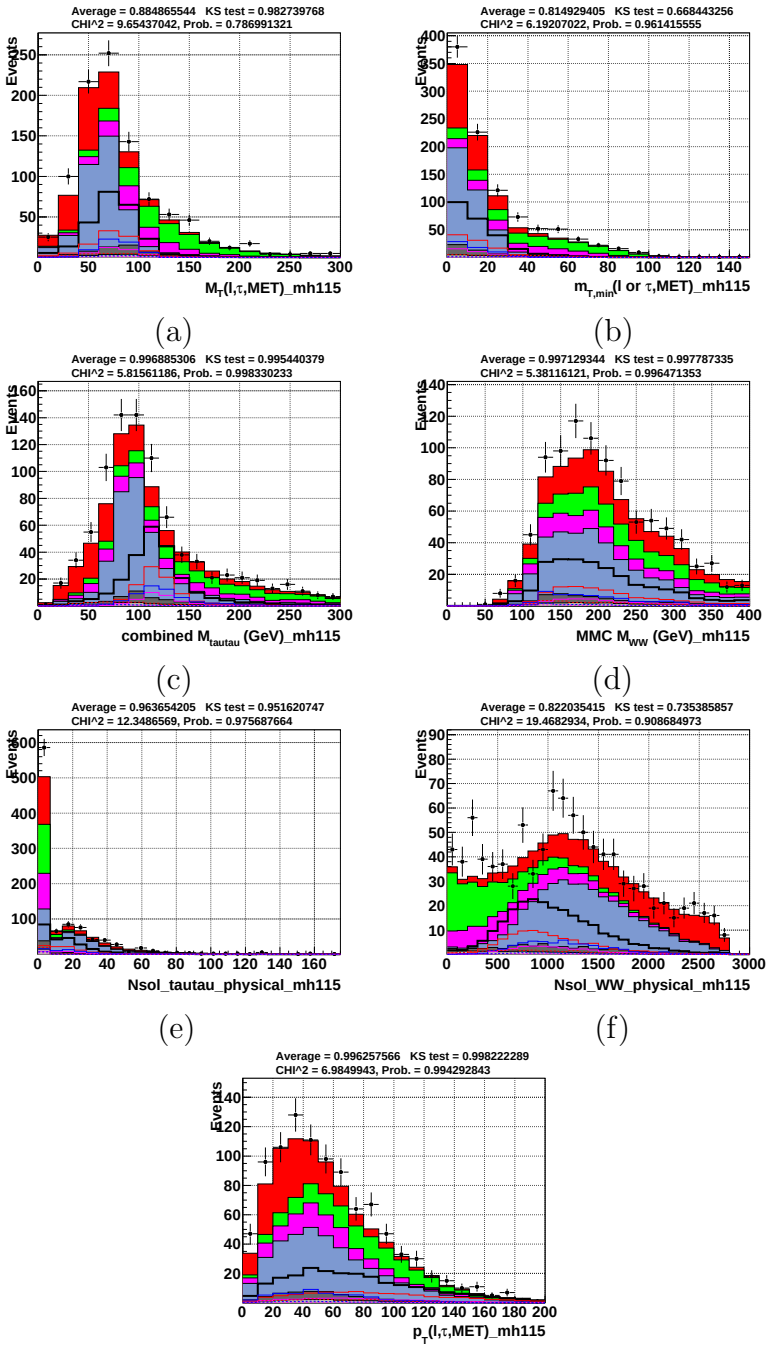


Figure 4.12: Data - MC comparison for BDT input variables: (a) $M_{\mathrm{T}}(\not\!\!E_T, \ell + \tau)$, (b) $M_{\mathrm{T,min}}(\not\!\!E_T, \ell/\tau)$, (c) $M_{\tau\tau,\ combined}$, (d) $M_{WW,\ combined}$, (e) $N_{\tau\tau\ soln.}$, (f) $N_{WW\ soln.}$, (g) $p_T(\ell, \tau, \not\!\!E_T)$Signal curves are for m_H=115 GeV multiplied by 100.

4.9 The variables from the Missing Mass Calculator (MMC)

4.9.1 Introduction and motivations

A. Elagin, et al. proposed a novel technique to reconstruct the mass for resonance decay to $\tau^+\tau^-$ [1]. The mass reconstruction of resonance decay to a pair of τ leptons has always been a challenge to modern experimental particle physics. A τ lepton decays to multiple neutrinos in different patterns, $\tau \to l\nu\bar{\nu}$ or $\tau \to hadrons + \nu$ (one- or three- plong), see Sec. 4.5.2 for a full description. These invisible neutrinos compose a majority of missing transverse energy ($\not{E}_T$) collected in the detector. The Higgs boson decay almost always involves multiple neutrinos. The rich kinematic structure of these neutrino clusters hinders the the Higgs mass calculation in forms of di-tau or di-W invariant mass and degrades the reconstruction resolution and accuracy significantly.

In past, there have been two typical methods to evaluate the di-tau resonance mass: (1) Take the 4-momentum of the visible products only (jets, leptons) and reconstruct the $\tau^+\tau^-$ mass partially. (2) Assume the neutrinos and the visible products are 100% aligned (collinear approximation). Method (1) derives a partial mass lower than the true mass, while method (2) is only accurate for a small fraction of events and serves merely as an approximation for the majority.

Collinear approximation

Since the MMC technique is a generalization of the collinear approximation, let us consider how the $\tau^+\tau^-$ invariant mass is reconstructed in the collinear approximation first. Two basic assumptions must be made: (1) The neutrino(s) from each decaying τ, leptonically or hadronically, is almost collinear with the visible decay products. (2) All $\not{E}_T$ purely comes from neutrinos. We can then decompose $\not{E}_T$ in x and y components,

$$\begin{aligned} \not{E}_{T_x} &= p_{mis1} sin\theta_{vis1} cos\phi_{vis1} + p_{mis2} sin\theta_{vis2} cos\phi_{vis2} \\ \not{E}_{T_y} &= p_{mis1} sin\theta_{vis1} sin\phi_{vis1} + p_{mis2} sin\theta_{vis2} sin\phi_{vis2}, \end{aligned} \tag{4.8}$$

where p_{mis1}, p_{mis2} are combined momenta of invisible neutrino clusters from each τ and θ_{vis1}, θ_{vis2}, ϕ_{vis1}, ϕ_{vis2} are their polar and azimuthal angles. Since neutrinos are assumed to align with visible products, we take $\phi_{mis1,2} = \phi_{vis1,2}$, $\theta_{mis1,2} = \theta_{mis1,2}$. This is an equation set with 2 unknowns and is fully solvable. The invariant mass of the $\tau^+\tau^-$ system will be $M_{\tau\tau} = m_{vis}/\sqrt{x_1 x_2}$, where m_{vis} is the reconstructed visible $\tau^+\tau^-$ mass and $x_{1,2} = p_{vis1,2}/(p_{vis1,2} + p_{mis1,2})$. The solutions to the equations are:

$$\begin{aligned} p_{Tmis1} &= |(\not{E}_{T_y} cos\phi_{vis2} - \not{E}_{T_x} sin\phi_{vis2})/denom| \\ p_{Tmis2} &= |(\not{E}_{T_x} sin\phi_{vis1} - \not{E}_{T_y} cos\phi_{vis1})/denom|, \end{aligned} \tag{4.9}$$

where $denom = sin(\phi_{vis1} - \phi_{vis2})$. Note this denominator diverges when the two visible clusters are back-to-back or highly boosted, $|\phi_{vis1} - \phi_{vis2}| \rightarrow 0, \pi$. Our former analysis adopted collinear approximation to reconstruct the $\tau^+\tau^-$ mass. It was one of the old 17 input variables [7] [8]. To avoid the divergence behavior, we excluded the solutions with $denom < 0.01$ and replaced those solutions with another approximation,

$$\begin{aligned} p_{Tmis1} &= |(\not\!\!E_T E_{vis2}/(E_{vis1} + E_{vis2})| \\ p_{Tmis2} &= |(\not\!\!E_T E_{vis1}/(E_{vis1} + E_{vis2})|, \end{aligned} \tag{4.10}$$

which come from two approximate but reasonable assumptions,

$$\begin{aligned} p_{Tmis1} + p_{Tmis2} &= \not\!\!E_T \\ p_{Tmis1} E_{vis1} &= p_{Tmis2} E_{vis2}. \end{aligned} \tag{4.11}$$

The first assumption is merely an assertion that $\not\!\!E_T$ consists of neutrinos only, which is true most of the time. When *denom* is small, the two neutrino clusters most probably travel along the same direction at high energy ($\phi_{mis1} - \phi_{mis2} \approx 0$), hence the estimate is made with magnitudes only to simplify the problem. The second assumption assumes the p_T of neutrino cluster is inversely proportional to the corresponding visible cluster: both come from a τ lepton, so the more energetic the visible cluster is, the less energetic the invisible neutrino cluster is.

Another drawback of the collinear approximation is its limited sensitivity to the $\not\!\!E_T$ resolution. As a result, a long tail toward the high mass region degrades the resolution of the reconstructed $\tau^+\tau^-$ mass.

The MMC technique

The authors suggested in [1] a new algorithm, the Missing Mass Calculator (MMC), to tackle the problem in collinear approximation. They loosened the assumptions of collinear approximation and admitted that neutrino(s) could go anywhere in the 3-dimensional space. Note the likelihood to observe neutrino(s) with certain 4-momentum can be determined by MC simulation. The 4-momentum with the best likelihood is then assumed to be the 4-momentum of the neutrino(s) per event basis. It is argued that the MMC improves both the accuracy and resolution of the $Z \rightarrow \tau\tau$ and $H \rightarrow \tau\tau$ mass peak. We verified the claim successfully in both MC simulation and the analysis with DØ data. Moreover, we generalized the MMC technique to the $H \rightarrow WW^{(*)}$ mass reconstruction, where the previous collinear approximation method compeletely fails to provide a good approximation to the Higgs mass.

The Missing Mass Calculator (MMC) technique allows a full reconstruction of the event topology with the $\tau^+\tau^-$ or W^+W^- kinematics. Assuming perfect detector resolution and no neutrino from other than the $\tau^+\tau^-$ decay, the following equation set describes the full event topology,

$$
\begin{aligned}
\not{E}_{T_x} &= p_{mis1} sin\theta_{mis1} cos\phi_{mis1} + p_{mis2} sin\theta_{mis2} cos\phi_{mis2} \\
\not{E}_{T_y} &= p_{mis1} sin\theta_{mis1} sin\phi_{mis1} + p_{mis2} sin\theta_{mis2} sin\phi_{mis2} \\
M_{\tau 1}^2 &= m_{mis1}^2 + m_{vis1}^2 + 2\sqrt{p_{vis1}^2 + m_{vis1}^2}\sqrt{p_{mis1}^2 + m_{mis1}^2} - 2p_{vis1}p_{mis1}cos\Delta\theta_{\nu m1} \\
M_{\tau 2}^2 &= m_{mis2}^2 + m_{vis2}^2 + 2\sqrt{p_{vis2}^2 + m_{vis2}^2}\sqrt{p_{mis2}^2 + m_{mis2}^2} - 2p_{vis2}p_{mis2}cos\Delta\theta_{\nu m2}.
\end{aligned}
\tag{4.12}
$$

The subscripts "vis", "mis" indicate the visible or invisible components of the decay product, and p, m, ϕ, θ are the momenta, invariant masses, azimuthal and polar angles, and $\Delta\theta_{\nu m1,2}$ are the angles between vectors p_{vis} and p_{mis}, which can be rewritten in terms of the other variables. Note the first two equations are Eq. (4.8) with $\theta_{vis1,2}, \phi_{vis1,2}$ recovered to $\theta_{mis1,2}, \phi_{mis1,2}$: the 4-momenta of neutrinos are defined independently, they are not collinear with the visible decay products anymore.

Equation (4.12) features 8 unknowns labeled with "mis": $p_{vis1,2}$, $m_{vis1,2}$, $\theta_{mis1,2}$, $\phi_{mis1,2}$. It is nevertheless possible to simplify the equation set and render it solvable. The $\mu\tau jj$ analysis investigates two major Higgs decay products: $H \to \tau\tau$ and $H \to WW$. In the $H \to \tau\tau$ case, one τ decays to muon + 2 neutrinos, while the other decays hadronically. The latter always produces one neutrino only; we therefore set m_{vis1} zero without loss of generality. The $H \to WW$ case is slightly more complex. There are two major possibilities:

- Channel (1): One W boson decays leptonically to a muon and a muon neutrino, while the other decays to a neutrino and a τ lepton with subsequent hadronic decay.
- Channel (2): Both W bosons decay to τ leptons and neutrinos, with subsequent leptonic and hadronic τ decays as the $H \to \tau\tau$ case.

The leptonic W decay branching ratios to 3 types of leptons are about the same, and the hadronic decay branching ratio of τ is about 4 times of the leptonic decay into either lepton [16]. There are other complications in the $H \to WW \to \mu\tau + X$ production too. The vector boson + Higgs associated production gives rise to a third W or Z, which subsequently decays into muons or taus, but this is only one of the three major Higgs production channels, and the $H \to WW$ production mainly affects searches for $M_H > 135$ GeV. We therefore ignored channel (2) and VH complications and optimized our Higgs searching strategy for channel (1) only. Here the invisible product of one W (decays to muon + neutrino) is again massless, and we can set m_{vis1} zero once again. This is why the MMC technique, originally designed for $\tau\tau$ invariant mass reconstruction, could also work for the WW mass reconstruction. The only change is to replace M_τ with M_W in Eq. (4.12).

The analytic solutions to Eq. (4.12) for four unknowns, p_{mis1}, p_{mis2}, θ_{mis1}, θ_{mis2}, can be sought with Mathematica. These solutions have highly complex structures, and we first define quantities $p1s1$, $p2s2$, $chunk1$, $chunk2$ in Eq. (4.13), (4.14) to simplify the problem:

$$
\begin{aligned}
&p1s1 = p_{mis1} sin\theta_{mis1} \\
&= (\not{E}_{T_y} * cos\phi_{mis2} - \not{E}_{T_x} * sin\phi_{mis2})/(cos\phi_{mis2} * sin\phi_{mis1} - cos\phi_{mis1} * sin\phi_{mis2}), \\
&p2s2 = p_{mis2} sin\theta_{mis2} \\
&= (-\not{E}_{T_y} * cos\phi_{mis1} + \not{E}_{T_x} * sin\phi_{mis1})/(cos\phi_{mis2} * sin\phi_{mis1} - cos\phi_{mis1} * sin\phi_{mis2}),
\end{aligned}
\tag{4.13}
$$

$$
\begin{aligned}
chunk1 &= p_{x,vis1} * p1s1 * cos\phi_{mis1} + p_{y,vis1} * p1s1 * sin\phi_{mis1}, \\
chunk2 &= p_{x,vis2} * p2s2 * cos\phi_{mis2} + p_{y,vis2} * p2s2 * sin\phi_{mis2}.
\end{aligned}
\tag{4.14}
$$

We can derive $\theta_{mis,i}$ with the first two equations in Eq. (4.13). The problem is hence converted to solving for $p_{mis,i}$. With the aid of $chunk1$ and $chunk2$ definitions in Eq. (4.14), $p_{mis,i}$ are given in Fig. 4.13. Note p_{mis1} has two solutions and p_{mis2} has four (labeled 1st - 4th in subscript), where $p_{mis2,1st} = -p_{mis2,2nd}$ and $p_{mis2,3rd} = -p_{mis2,4th}$. Only real and positive solutions are physical, in which case we have 4 positive combinations for (p_{mis1}, p_{mis2}) in each event without loss of generality: ($p_{mis1,1st}$, $|p_{mis2,1st}|$), ($p_{mis1,1st}$, $|p_{mis2,3rd}|$), ($p_{mis1,2nd}$, $|p_{mis2,1st}|$), ($p_{mis1,2nd}$, $|p_{mis2,3rd}|$). If we relax the sign requirement of $cos\theta_{mis}$, there will be even more solution choices: we merely obtained $sin\theta_{mis}$ in Eq. (4.12), but θ_{mis} could always stay in two quadrants and $cos\theta_{mis}$ can be either positive or negative. For simplicity, we decided to take the most probable scenario: the neutrino clump goes along a similar direction of visible tau or W decay products, i.e. same signs are taken for $cos\theta_{vis}$ and $cos\theta_{mis}$. All these physical solutions were filled in the histogram in each event for further probability determination upon a ΔR distribution. Total number of solutions was also stored for reasons soon to be clarified.

Now that there are 7 unknowns in these solutions, let us move one step forward and impose further constraints to reduce the number of unknowns. Consider a 3-dimensional phase space scan (ϕ_{mis1}, ϕ_{mis2}, m_{mis2}). Each grid point gives rise to a fully-solvable event topology. All grid sizes shown later were based on these three dimensions. Clearly not all such solutions are equally likely. It is however possible to determine the most probable event kinematics in the solution space and calculate the corresponding $\tau\tau$ or the WW mass. In the case there are no solutions, we take the collinear approximation mass for the di-tau pair and construct a reasonable estimate for the di-W pair.

We compute such probability as a distribution function of $\Delta R_{1,2}$ for each solution per event basis.

$$
\Delta R_i = \sqrt{(\eta_{vis,i} - \eta_{mis,i})^2 + (\phi_{vis,i} - \phi_{mis,i})^2},\ i = 1,2\ for\ two\ taus, \tag{4.15}
$$

where ΔR is the distance in the angular space between the visible and invisible tau decay products. Three distinct hadronic τ decays are considered separately.

$$\left\{\left\{p1 \to -\left(-2\,\text{chunk1}\,\text{evis1} + \text{evis1}\,\text{mvis1}^2 - \text{evis1}\,m\tau^2 + \sqrt{\text{pvis1z}^2\left(\left(2\,\text{chunk1} - \text{mvis1}^2 + m\tau^2\right)^2 + 4\,\text{p1s1}^2\left(-\text{evis1}^2 + \text{pvis1z}^2\right)\right)}\right) \Big/ \left(2\left(\text{evis1}^2 - \text{pvis1z}^2\right)\right)\right\}, \left\{p1 \to \left(2\,\text{chunk1}\,\text{evis1} - \text{evis1}\,\text{mvis1}^2 + \text{evis1}\,m\tau^2 + \sqrt{\text{pvis1z}^2\left(\left(2\,\text{chunk1} - \text{mvis1}^2 + m\tau^2\right)^2 + 4\,\text{p1s1}^2\left(-\text{evis1}^2 + \text{pvis1z}^2\right)\right)}\right) \Big/ \left(2\left(\text{evis1}^2 - \text{pvis1z}^2\right)\right)\right\}\right\}$$

$$\left\{\left\{p2 \to -\frac{1}{2}\sqrt{\left(\left(-4\,\text{evis2}^4\,\text{m2}^2 + \left(-2\,\text{chunk2} + \text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2\text{pvis2z}^2 + 4\,\text{p2s2}^2\,\text{pvis2z}^4 + \text{evis2}^2\left(\left(-2\,\text{chunk2} + \text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2 + 4\left(\text{m2} - \text{p2s2}\right)\left(\text{m2} + \text{p2s2}\right)\text{pvis2z}^2\right) - 2\sqrt{\left(\text{evis2}^2\left(-2\,\text{chunk2} + \text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2\text{pvis2z}^2\left(4\,\text{chunk2}^2 - 4\,\text{chunk2}\left(\text{m2}^2 + \text{mvis2}^2 - m\tau^2\right) + \left(\text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2 - 4\,\text{evis2}^2\left(\text{m2}^2 + \text{p2s2}^2\right) + 4\left(\text{m2}^2 + \text{p2s2}^2\right)\text{pvis2z}^2\right)\right)\Big/\left(\text{evis2}^2 - \text{pvis2z}^2\right)^2\right)\right\},$$

$$\left\{p2 \to \frac{1}{2}\sqrt{\left(\left(-4\,\text{evis2}^4\,\text{m2}^2 + \left(-2\,\text{chunk2} + \text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2\text{pvis2z}^2 + 4\,\text{p2s2}^2\,\text{pvis2z}^4 + \text{evis2}^2\left(\left(-2\,\text{chunk2} + \text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2 + 4\left(\text{m2} - \text{p2s2}\right)\left(\text{m2} + \text{p2s2}\right)\text{pvis2z}^2\right) - 2\sqrt{\left(\text{evis2}^2\left(-2\,\text{chunk2} + \text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2\text{pvis2z}^2\left(4\,\text{chunk2}^2 - 4\,\text{chunk2}\left(\text{m2}^2 + \text{mvis2}^2 - m\tau^2\right) + \left(\text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2 - 4\,\text{evis2}^2\left(\text{m2}^2 + \text{p2s2}^2\right) + 4\left(\text{m2}^2 + \text{p2s2}^2\right)\text{pvis2z}^2\right)\right)\Big/\left(\text{evis2}^2 - \text{pvis2z}^2\right)^2\right)\right\},$$

$$\left\{p2 \to -\frac{1}{2}\sqrt{\left(\left(-4\,\text{evis2}^4\,\text{m2}^2 + \left(-2\,\text{chunk2} + \text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2\text{pvis2z}^2 + 4\,\text{p2s2}^2\,\text{pvis2z}^4 + \text{evis2}^2\left(\left(-2\,\text{chunk2} + \text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2 + 4\left(\text{m2} - \text{p2s2}\right)\left(\text{m2} + \text{p2s2}\right)\text{pvis2z}^2\right) + 2\sqrt{\left(\text{evis2}^2\left(-2\,\text{chunk2} + \text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2\text{pvis2z}^2\left(4\,\text{chunk2}^2 - 4\,\text{chunk2}\left(\text{m2}^2 + \text{mvis2}^2 - m\tau^2\right) + \left(\text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2 - 4\,\text{evis2}^2\left(\text{m2}^2 + \text{p2s2}^2\right) + 4\left(\text{m2}^2 + \text{p2s2}^2\right)\text{pvis2z}^2\right)\right)\Big/\left(\text{evis2}^2 - \text{pvis2z}^2\right)^2\right)\right\},$$

$$\left\{p2 \to \frac{1}{2}\sqrt{\left(\left(-4\,\text{evis2}^4\,\text{m2}^2 + \left(-2\,\text{chunk2} + \text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2\text{pvis2z}^2 + 4\,\text{p2s2}^2\,\text{pvis2z}^4 + \text{evis2}^2\left(\left(-2\,\text{chunk2} + \text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2 + 4\left(\text{m2} - \text{p2s2}\right)\left(\text{m2} + \text{p2s2}\right)\text{pvis2z}^2\right) + 2\sqrt{\left(\text{evis2}^2\left(-2\,\text{chunk2} + \text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2\text{pvis2z}^2\left(4\,\text{chunk2}^2 - 4\,\text{chunk2}\left(\text{m2}^2 + \text{mvis2}^2 - m\tau^2\right) + \left(\text{m2}^2 + \text{mvis2}^2 - m\tau^2\right)^2 - 4\,\text{evis2}^2\left(\text{m2}^2 + \text{p2s2}^2\right) + 4\left(\text{m2}^2 + \text{p2s2}^2\right)\text{pvis2z}^2\right)\right)\Big/\left(\text{evis2}^2 - \text{pvis2z}^2\right)^2\right)\right\}\right\}$$

Figure 4.13: The MMC analytic solutions. Top: two solutions for p_{mis1}. Bottom: four solutions for p_{mis2}.

4.9.2 The MC simulations

We built the Monte-Carlo simulation in Fortran to determine the ΔR distribution between the visible and invisible decay products of τ/W, see Eq. 4.15 for the definition. The ΔR distribution depends on the τ momentum, decay type and polarization. There are one leptonic and three hadronic τ decays, whose ΔR distributions vary; we hence implemented the τ type recognition on a per-event basis (see Table 4.7 for details) and derived the ΔR distribution for each tau decay type separately. We neglected the polarization of τ leptons for simplicity. Both $Z \to \tau\tau$ and $H \to \tau\tau$ processes were generated with the same selections described in Sec. 4.5. The general procedures go as follows.

First, we scanned all grids on the phase space (ϕ_{mis1}, ϕ_{mis2}, m_{mis2}), found all physical solutions (real and positive solutions to Eq. (4.12)) in each event, calculated the invariant mass per event basis and plot the distribution by τ types. All ϕ's range from 0 to 2π and the invisible mass ranges between (0,1.8). Figure 4.14 shows an example with 10K $Z \to \tau\tau$ events, grid size (0.03, 0.03, 0.2).

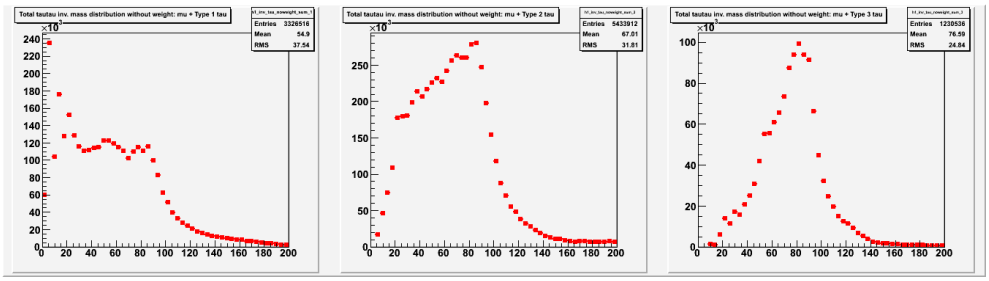


Figure 4.14: $Z \to \tau\tau \to \mu + hadrons$ invariant mass distribution from solutions of 10K events for type 1 (left), 2 (middle), 3 (right) hadronically decaying τ. All physical solutions are filled without histogram weight. Grid size (0.03, 0.03, 0.2), M_Z=91.2GeV.

Second, we determine the ΔR probability of every single physical solution in each event according to the ΔR probability distribution and the initial τ or W momentum. We generate sufficient MC events, apply the same selections as the analysis and compute ΔR for τ's or W's by Eq. (4.15) per event. To take into account the influence of the τ or W momentum, we divide the momentum in bins with 5 GeV increment, and plot the ΔR distribution per τ type. See Appendix B for the distributions for $\tau\tau$ and WW decays. We assume the original τ or W momentum in a certain range, e.g. (10,100) GeV, divide it in regions of 5 GeV and plot the ΔR distribution in all events generated in each region. The momentum regions for leptonic and type 1,2,3 hadronic τ decays are (10,100), (12.5,102.5), (12.5,102.5), (15,105) GeV, respectively. These ranges are tailored to include our selections muon $p_T > 12$ (RunIIa) or 15 (RunIIb) GeV, type 1,2,3 τ $p_T > 12.5$, 12.5, 15 GeV. The fit function to these ΔR distributions is a linear combination of the Gaussian and Landau functions,

$$p_2 * f_{Landau}(x, p_0, p_1) + g_{Gaussian}(x, p_3, p_4, p_5), \tag{4.16}$$

where p_0 thru p_5 are fit parameters. The fit ranges for ΔR of τ and W are (0,0.8) and (0,7), respectively. Table 4.7 summarizes the differences in the MMC formulation between the processes $H \to \tau\tau$ and $H \to WW$. The likelihood, or the weight we apply to each solution depends on the τ/W momentum and the ΔR distribution,

$$L = P(\Delta R_1, p_{\tau 1}) \times P(\Delta R_2, p_{\tau 2}), \tag{4.17}$$

where $P(\Delta R_i, p_{\tau_i})$, i=1,2 is the final probability function for each single τ/W decay.

items	$H \to \tau\tau$	$H \to WW^{(*)}$
ΔR range	$[0, 0.8]$	[0.0.7]
single τ, W mass (GeV)	1.78	on-shell: 80.4 off-shell: varies
neutrino cluster mass range (GeV)	$(0, 1.8)$	$(0, 85)$
MC sample type recognition		
Type 0 (leptonic)	2 neutrinos. $mmis$ scanned in $[0.2, 1.8]$ GeV	1 or 3 neutrinos, muon from W ($\frac{6}{7}$) or τ ($\frac{1}{7}$). Assumed $mmis = 0$ for simplicity.
Type 1	1 neutrino, 1 hadron, no photon. $mmis = 0$	2 neutrinos, 1 hadron, no photon. $mmis$ scanned in $[0, 80]$ GeV
Type 2	1 neutrino, 1 hadron, 2 photons. $mmis = 0$	2 neutrinos, 1 hadron, 2 photons. $mmis$ scanned in $[0, 80]$ GeV
Type 3	1 neutrino, 3 hadrons. $mmis = 0$	2 neutrinos, 3 hadrons. $mmis$ scanned in $[0, 80]$ GeV

Table 4.7: The differences in $H \to \tau\tau$ and $H \to WW$ processes when formulating the MMC technique.

Third, we weight the raw histograms in step 1 with ΔR probability of step 2 and get the most probable mass per event basis. The corresponding mass at the maximum of such distributions serves as the final reconstructed MMC $\tau\tau$ or WW mass. Figure 4.15 shows distribution of physical solutions after weights for 3 types of hadronically decaying τ. We recovered the $M_Z = 91.2$ GeV peak in all cases.

Now we are ready to compare the $\tau^+\tau^-$ mass spectrum with the collinear approximation. Figure 4.16, 4.17 shows the $Z \to \tau^+\tau^-$, $H \to \tau^+\tau^-$ MMC and collinear approximation masses. We expect different shapes in the Z and Higgs decays due to the different event topology: the former has spin 1 and the differential cross section is polar angle dependent, while the latter carries no spin and decays isotropically in the center-of-mass frame. The closer MMC and collinear approximation peaks in $H \to \tau\tau$ might result from more collinear τ decays.

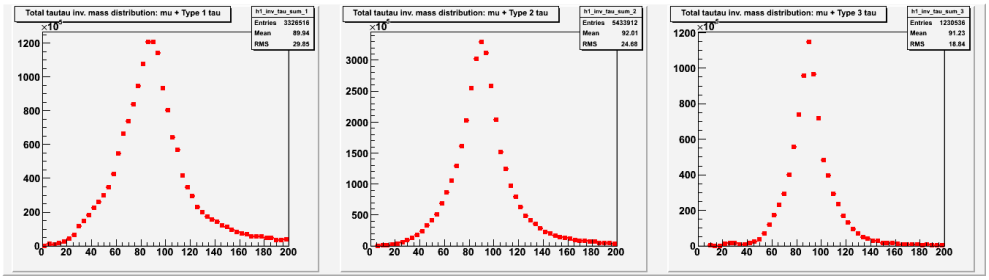


Figure 4.15: $Z \to \tau\tau \to \mu + hadrons$ invariant mass distribution from solutions of 10K events for type 1 (left), 2 (middle), 3 (right) hadronically decaying τ. All physical solutions are filled with histogram weight. Grid size (0.03, 0.03, 0.2), M_Z=91.2GeV.

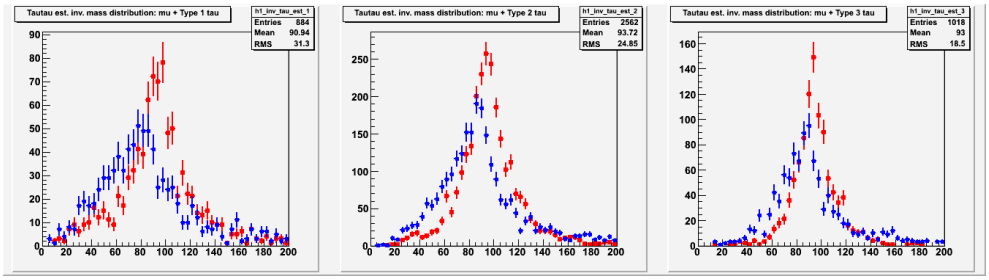


Figure 4.16: $Z \to \tau\tau \to \mu + hadrons$ invariant mass spectrum for type 1 (left), 2 (middle), 3 (right) hadronically decaying τ. Red: MMC, blue: collinear approximation. Grid size (0.03, 0.03, 0.2), $M_Z = 91.2$ GeV.

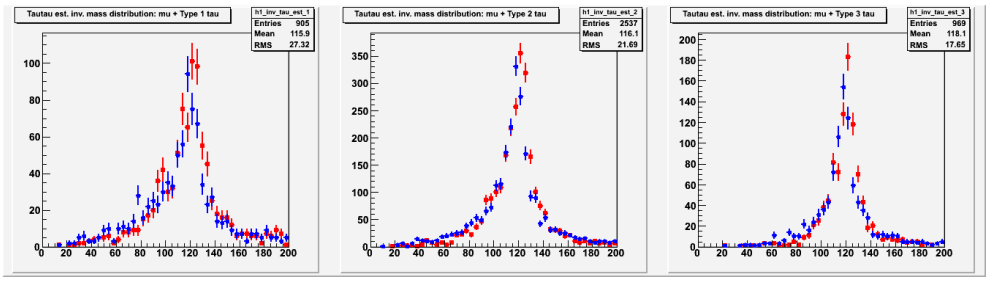
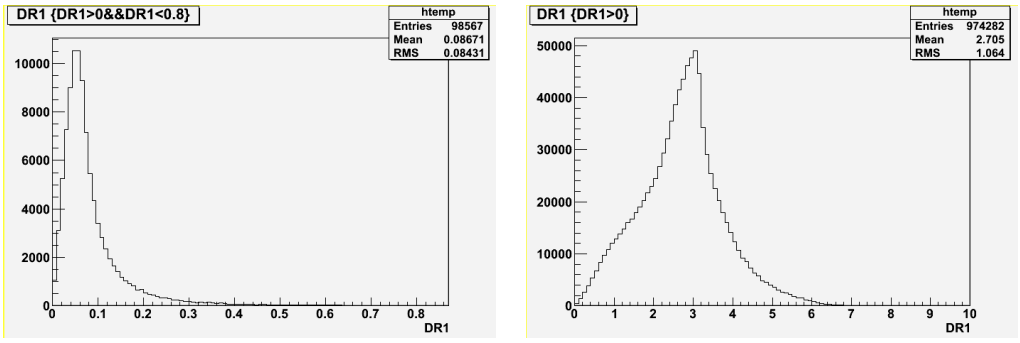


Figure 4.17: $H \to \tau\tau \to \mu + hadrons$ invariant mass spectrum for type 1 (left), 2 (middle), 3 (right) hadronically decaying τ. Red: MMC, blue: collinear approximation. Grid size (0.03, 0.03, 0.2), assuming $M_H = 120$ GeV.

The MMC $\tau\tau$ mass distribution is sharper, more accurate than the collinear approximation and free of tail in the high energy end as well. The improvement in $Z \to \tau\tau$ is more pronounced than $H \to \tau\tau$, which has to do with their distinct spins and Standard Model couplings. We hence recovered the conclusion in [1].

We went one step further and investigated the quality of complex solution distributions with Eq. (4.12). We considered the complex phase space when $\not{E}_T$ is involved, and computed the modulus or the real part of the complex parameters for distribution assessment and possible improvement over real phase space parameters. We did not observe such improvement from the use of complex solutions in our study. The solution efficiency improved, but the resolution got worse and computation time was large. We tried to cut off solutions with larger imaginary parts, but observed no prominent improvements either.

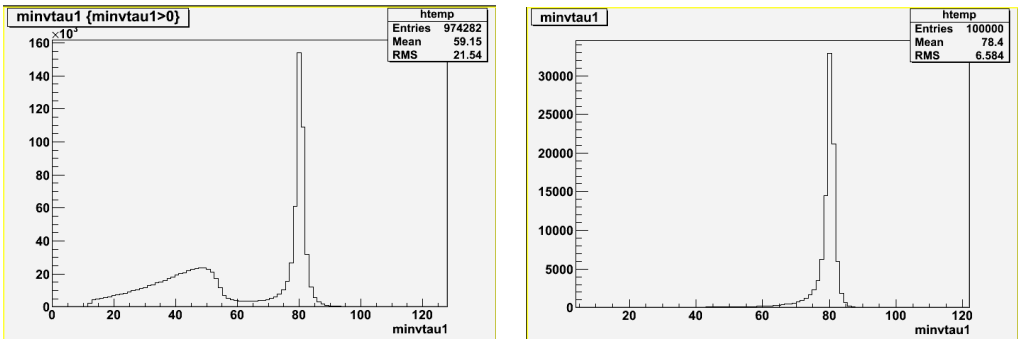
The story on the MMC WW reconstruction is quite similar. We substituted the τ masses in Eq. (4.12) with the W masses, set the mass for one of the missing cluster zero, scanned the same 3-dimensional space and solved the same equation set per event basis for the most likely WW mass. The likelihood is determined by the W decay ΔR distribution in Appendix B. One important difference is the $H \to WW^{(*)}$ decay produces one W boson on-shell, and the other on or off-shell, depending on the hypothesized Higgs mass. Figure 4.19 demonstrates the mass distribution for two W bosons at hypothesized Higgs mass 135 and 165 GeV. The complete differences between the MMC $\tau\tau$ and WW algorithms are summarized in Table 4.7. Further details can be found in Figs. 4.18-4.20.



(a) (b)

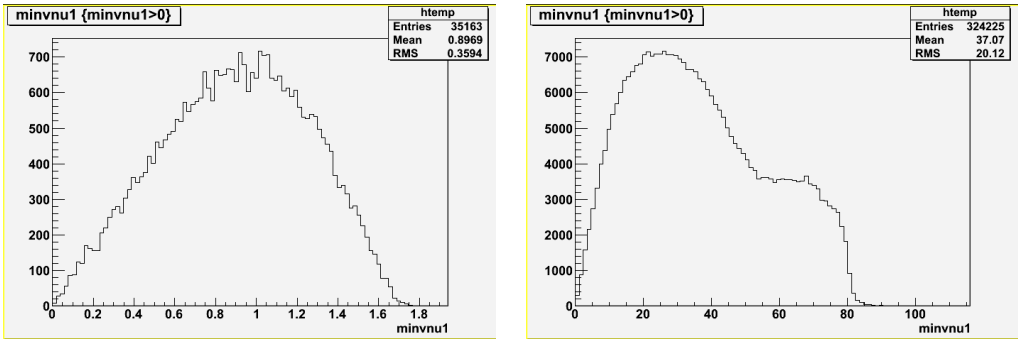
Figure 4.18: ΔR distribution for (a) $H \to \tau\tau$ in range [0,0.8]; (b) $H \to WW^{(*)}\Delta R$ distribution in range [0,7].

Contrary to the $\tau\tau$ MMC mass reconstruction, the invisible W decay products could be quite massive and requires a full scan all the way up to the W mass, see Fig. 4.20. We investigated the grid size (ϕ_{mis1}, ϕ_{mis2}, m_{mis2}) = (0.07, 0.07, 5) as a compromise between computing time and better resolution. Two distributions were computed in parallel and shared the same solution efficiency:



(a) (b)

Figure 4.19: $H \to WW^{(*)}$ mass distribution for two W bosons, assuming (a) $M_H = 135$ GeV; (b) $M_H = 165$ GeV. We expect one W on-shell, one off-shell when $M_H < 160.8GeV$, and both on-shell otherwise. Note there is further complication in the VH production or when a faked W boson is present in the background.



(a) (b)

Figure 4.20: The neutrino cluster mass distribution to be scanned for (a) τ: (0, 1.8) GeV; (b) W: (0, 85) GeV.

- (1). the MMC weighted mass: as in the $\tau\tau$ mass scan - fill and weight each entry with the ΔR likelihood in the histogram, get the peak location.
- (2). the MMC profile mass: find out the most probable WW mass by a "profile plot" in Root language - fill the ΔR likelihood in vertical axis and the invariant mass in horizontal axis per event basis, get the peak location.

Both approaches are sensible and the difference between them is subtle. In practice, we chose the one with better resolution and accuracy. In the $\tau\tau$ case the two methods perform almost identically and we took the MMC weighted mass, whereas in the WW case they show certain deviation and we took the MMC profile mass in the end. See Figs. 4.21-4.23 for an evolution of the two reconstruction algorithms across different hypothesized Higgs masses. Both weighted and profile masses provide

better peaks and resolution than collinear approximation. Remember the W boson is heavy and capable of supplying a significant amount of energy to the neutrinos along any direction. The collinear approximation clearly does not describe the WW invariant mass well. The profile algorithm turns out to represent the Higgs mass better and shifts toward the high mass correctly as M_H increases. The MMC WW mass reconstruction is however less accurate than $\tau\tau$ in general, owing to missing information in the much more energetic invisible clusters that the ΔR distribution fails to recover.

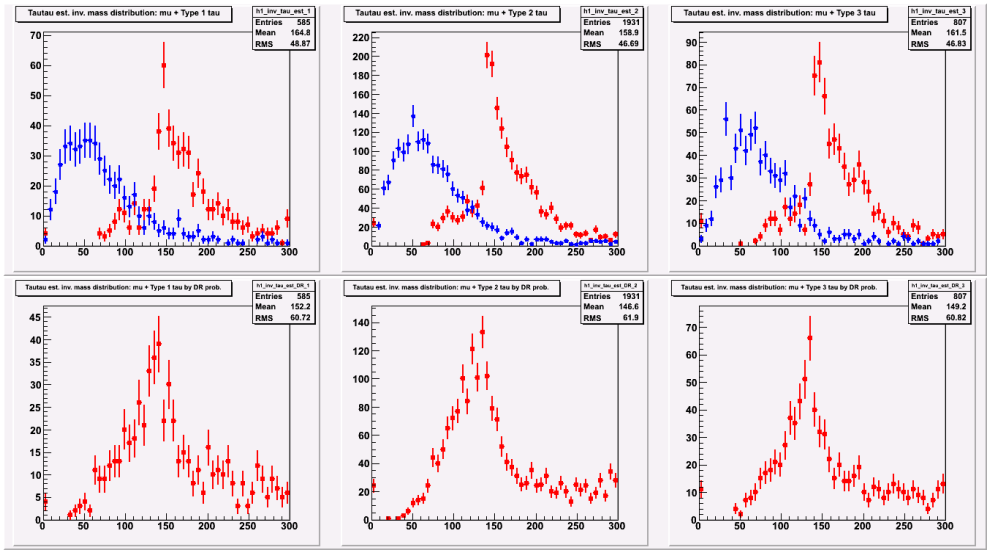


Figure 4.21: $H \to WW^* \to \mu + hadrons$ invariant mass spectrum for type 1 (left), 2 (middle), 3 (right) hadronically decaying τ. Top, red: MMC weighted mass, blue: collinear approximation. Bottom: MMC profile mass. Grid size (0.07, 0.07, 5), assuming $M_H = 135$ GeV.

4.9.3 The MMC $\tau\tau$ and WW variables with DØ data

The MMC technique has proven to work in the MC simulation. It provides better resolution and accuracy in the Z and Higgs mass reconstruction than the traditional collinear approximation. We observed similar behavior when we transplanted the algorithm to real DØ data. Figure 4.24 shows the event distribution with MMC mass only, leaving the events with no physical solutions in the first bin. Figure 4.25 gives the $\tau\tau$ mass with traditional collinear approximation determination, while Fig. 4.26 supplies a combined mass of the two: whenever there is a MMC solution, the MMC mass is taken; otherwise the collinear approximation mass is taken. It is noted that the MMC combined mass has better resolution in the Higgs signals or backgrounds with $\tau\tau$ final state ($Z \to \tau\tau$), particularly in the low mass region when $M_H = 115$

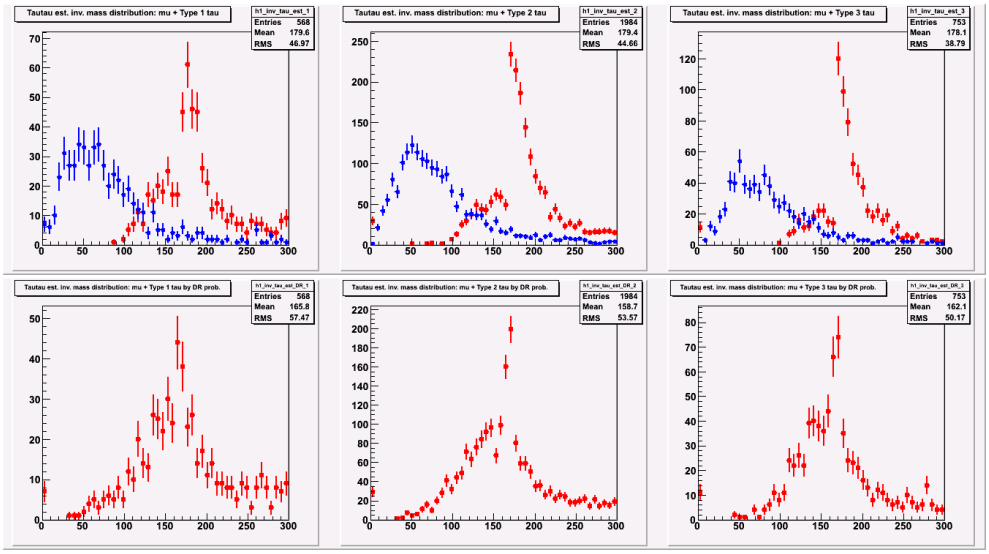


Figure 4.22: $H \to WW \to \mu + hadrons$ invariant mass spectrum for type 1 (left), 2 (middle), 3 (right) hadronically decaying τ. Top, red: MMC weighted mass, blue: collinear approximation. Bottom: MMC profile mass. Grid size (0.07, 0.07, 5), assuming $M_H = 165$ GeV.

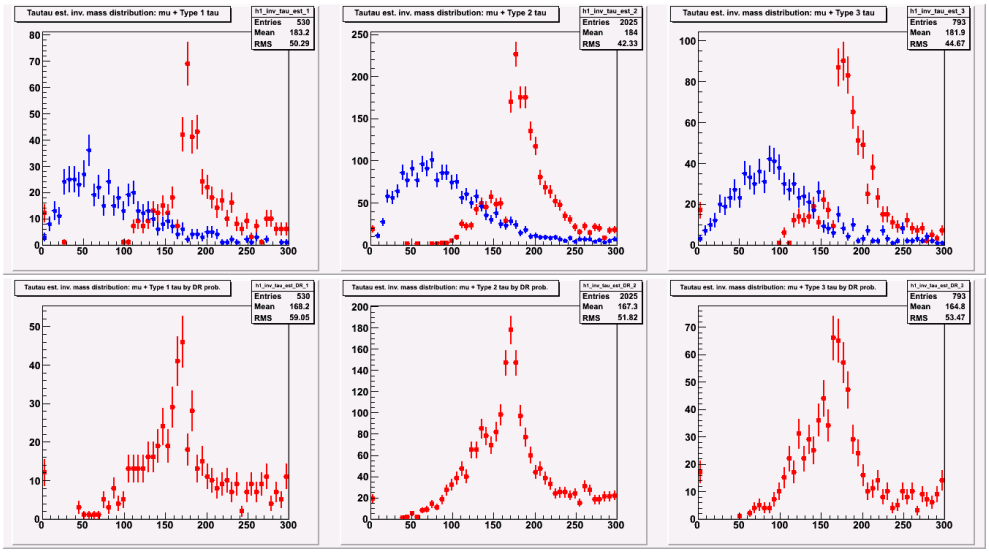


Figure 4.23: $H \to WW \to \mu + hadrons$ invariant mass spectrum for type 1 (left), 2 (middle), 3 (right) hadronically decaying τ. Top, red: MMC weighted mass, blue: collinear approximation. Bottom: MMC profile mass. Grid size (0.07, 0.07, 5), assuming M_H=180GeV.

GeV is assumed, since the $H \to \tau\tau$ decay dominates the $\mu\tau jj$ analysis in the low mass region.

Finer grid size typically leads to better mass resolution. However, the computation time is proportional to the number of grid points taken. The $\mu\tau jj$ analysis took the grid sizes (0.05, 0.05, 0.3) for $H \to \tau\tau$ and (0.15, 0.15, 10) in the phase space (ϕ_{mis1}, ϕ_{mis2}, m_{mis2}) for $H \to WW^{(*)}$ mass reconstruction as a compromise of computing time and resolution - the heaviest jobs took over 60 hours already and the computing arrays in DØ impose 72 hours of stringent upper limit for all jobs. A test on a small sample revealed the $\tau\tau$ solution efficiency improved only 10% with finer grid (0.03, 0.03, 0.2) while the computing time quadruples. We believe our grid choice is a reasonable maximization of our computation power at hand.

There is one more complication in the MMC WW analysis with real data: we do not know if the second W boson is on- or off-shell and from the Higgs signal or the background. Remember we need two τ or W masses in Eq. (4.12). This was not a problem in the $\tau\tau$ MC simulation since we just worked on either $Z \to \tau\tau$ or $H \to \tau\tau$: taus are light, a Z boson or a Higgs boson with $M_H > 100$ GeV does not produce taus off-shell, and it is safe to assume both taus of mass 1.777 GeV. While we assume a hypothesized Higgs mass in the analysis, the final states involving W's are complex. Consider $WH \to WWW$ for example, how do we decide the off-shell W mass? There are several possibilities:

M_H/GeV	115	125	135	145	155	165	180
M_{W}/GeV	29	38	48	61	73	80.4	80.4

Table 4.8: Locations of the second W mass peak assuming M_H between 115 and 180 GeV.

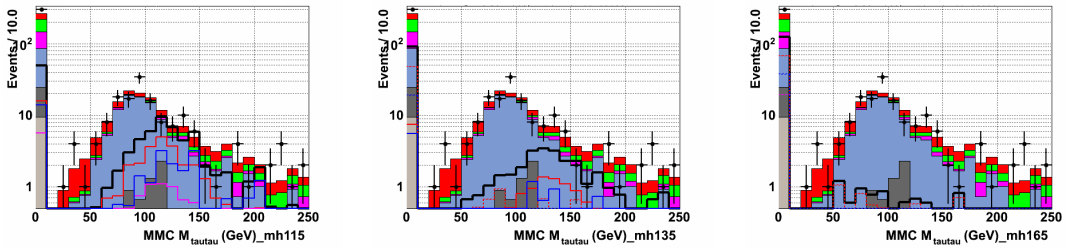
- (1). Simply assume a mass for the second W. Table 4.8 shows the mass peaks for the second W boson assuming different Higgs boson masses in the MC simulation $H \to WW^{(*)}$. This does not require further computation and is adopted in our analysis later. We took 66 GeV at first, since we did not know the correct Higgs mass peak and this value gave best discrimination between the Higgs signals and the backgrounds. It corresponds to a Higgs boson of mass 150 GeV if it were from the Higgs decay. We changed it to 38 GeV later after LHC announced the discovery of a Higgs-like particle with mass around 125 GeV: this is the most probable off-shell W mass peak for a 125 GeV Higgs and will maximize our sensitivity at this mass point. Note the other W boson in $H \to WW^{(*)}$ almost always goes on-shell. Nevertheless, we still don't know which W is off- or on-shell and therefore took the average of the two scenarios in computing the combined WW mass and number of physical MMC WW solutions.

- (2). Set the off-shell W mass differently at each mass point by Table 4.8. This requires to generate the whole set of background samples and conduct an independent multivariate analysis per mass point, since the definitions of all variables used must be identical in the same training. This is clearly the most accurate method, but great power comes with great costs: it is computationally heavy and highly susceptible to errors.

- (3). Develop a BDT or the like to figure it out. This sounds like a fun project for a new graduate student, but our manpower was truly limited at the moment.

- (4). Elevate the 3D scan to 4D: scan the second W mass in [0,80] GeV too. This sounds sensible and promising but again takes a lot more computing time, scanning an extra dimension with N grids cost N multiple of time. We did explore the direction with coarsened grids and did not find a more competitive WW mass resolution.

We went for the first option as discussed.

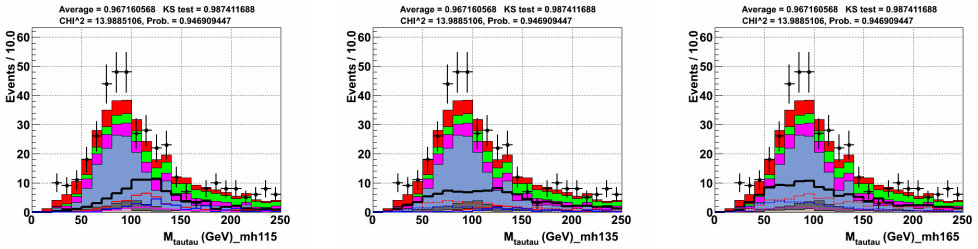


(a) (b) (c)

Figure 4.24: The 2b1-3 MMC $\tau\tau$ mass distribution assuming: (a) $M_H = 115$ GeV; (b) $M_H = 135$ GeV, ; (c) $M_H = 165$ GeV. KS=0.97, χ^2 prob.=0.76, ξ=0.86, grid size (0.05, 0.05, 0.3).

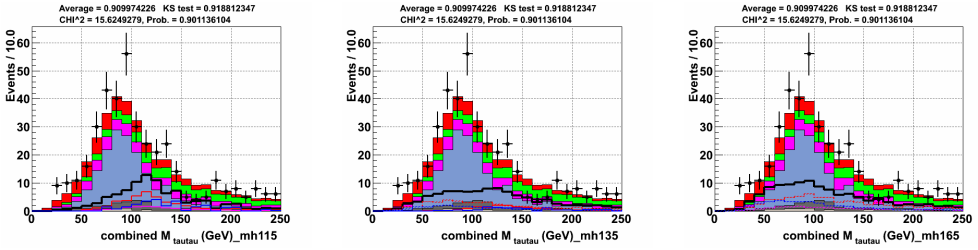
When the MMC technique fails to find a mass, a sensible estimate is more preferable than none. For the $H \to WW^{(*)}$ case this can not be given by the collinear approximation as in the $H \to \tau\tau$ MMC reconstruction. The collinear approximation assumes collinearity of the neutrino and visible decay products, but this is by no means true in the $H \to WW$ decay, which dominates in the high mass region. It is clear from Fig. 4.25 that the collinear approximation fails to address a high Higgs mass distribution, as was noted in the MC simulation in Figs. 4.21-4.23 too. We hence reconstruct and estimate the high mass Higgs in a totally different approach for events with $N_{WW\ soln.} = 0$:

$$m_{WW,est} = m_T(l, \tau, \not{E}_T) * (p_{vis1} + p_{vis2})/(p_{Tvis1} + p_{Tvis2}), \tag{4.18}$$



(a) (b) (c)

Figure 4.25: The 2b1-3 collinear approximation $\tau\tau$ mass distribution assuming: (a) $M_H = 115$ GeV; (b) $M_H = 135$ GeV; (c) $M_H = 165$ GeV. Grid size (0.05, 0.05, 0.3).



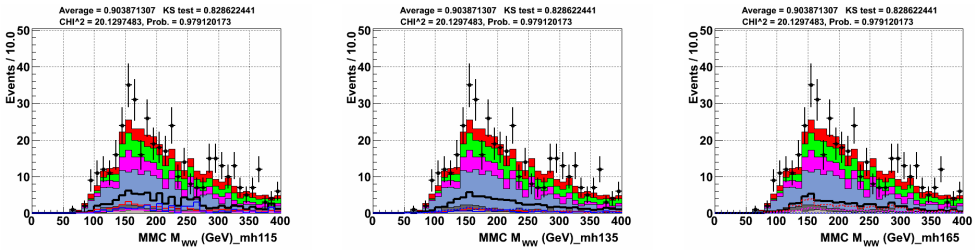
(a) (b) (c)

Figure 4.26: The 2b1-3 combined MMC $\tau\tau$ mass distribution assuming: (a) $M_H = 115$ GeV; (b) $M_H = 135$ GeV; (c) $M_H = 165$ GeV. When there is no MMC solution, the mass from collinear approximation is taken instead. Grid size (0.05, 0.05, 0.3).

i.e. to multiply the transverse mass of the μ, τ, $\not\!\!E_T$ by a factor to get the full invariant WW mass.

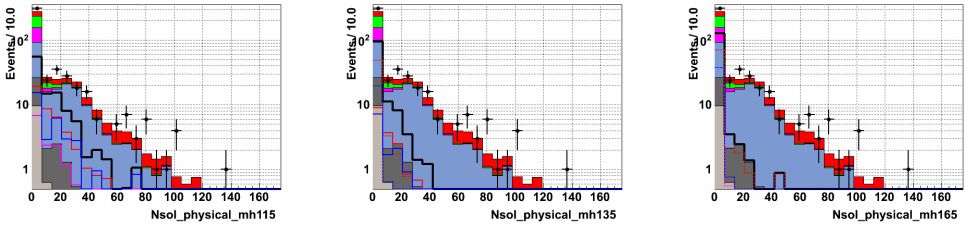
In reality there are few events without a solution: the W boson is heavy and produces much more $\not\!\!E_T$ than a τ lepton, thus loosens the constraints on the first two equations in Eq. (4.12). It is therefore much easier to find a solution in $H \to WW^{(*)}$ than in the $H \to \tau\tau$ process. This is the rule of thumb to understand why the distribution for the number of physical WW solutions, Fig. 4.29, is much higher than those for $\tau\tau$ solutions in Fig. 4.28. Figure 4.27 shows the combined MMC WW profile mass distribution.

Though not obvious in the beginning, the numbers of physical $\tau\tau$ and WW solutions also turn out to be interesting variables to distinguish the Higgs signals and the backgrounds. From Fig. 4.28 and Fig. 4.29 we see the signal and background shapes are quite different. The rule of thumb as stated, more $\not\!\!E_T$ leads to fewer number of solutions, explains why "unphysical" processes (the processes where Eq. (4.12) does not apply, e.g. backgrounds except the $t\bar{t}$ background or the low mass Higgs signal, which carries less $\not\!\!E_T$) have more physical WW solutions than the physical processes (processes with valid WW final state, e.g. the $t\bar{t}$ background or the high



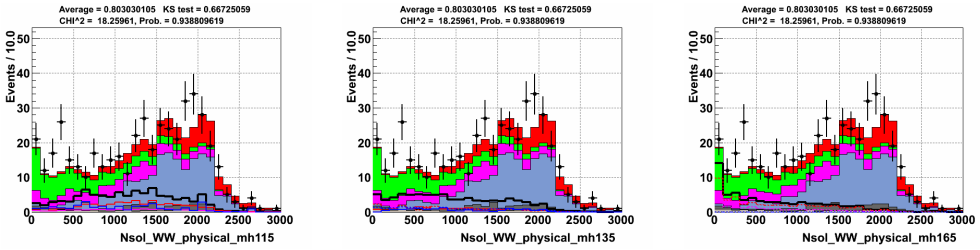
(a) (b) (c)

Figure 4.27: The 2b1-3 combined MMC WW mass distribution assuming: (a) $M_H = 115$ GeV; (b) $M_H = 135$ GeV; (c) $M_H = 165$ GeV. When there is no MMC solution, the mass from estimate $m_{WW,est}$ is taken instead. Grid size (0.15, 0.15, 10).



(a) (b) (c)

Figure 4.28: The 2b1-3 event distribution for the number of MMC $\tau\tau$ solutions assuming: (a) $M_H = 115$ GeV; (b) $M_H = 135$ GeV; (c) $M_H = 165$ GeV. KS=1.00, χ^2 prob.=0.55, ξ=0.78, grid size (0.05, 0.05, 3).



(a) (b) (c)

Figure 4.29: The 2b1-3 event distribution for the number of MMC WW solutions assuming: (a) $M_H = 115$ GeV; (b) $M_H = 135$ GeV; (c) $M_H = 165$ GeV. Grid size (0.15, 0.15, 10).

mass Higgs signal, which carries more $\not\!\!E_T$). A similar argument applies to the distribution for the number of MMC $\tau\tau$ solutions. Here the $Z_{\tau\tau}$ and low mass Higgs signals are well described by Eq. (4.12), deemed physical and tend to carry less $\not\!\!E_T$ or have more solutions. Table 4.9 shows the solution efficiency, or the percentage of events carrying solutions in various Higgs signals and backgrounds.

category	solution efficiency
$t\bar{t}$	22.6%
W+jets	14.8%
$Z_{\mu\mu}$+jets	37.5%
$Z_{\tau\tau}$+jets	66.2%
di-boson	23.2%
MJ	25.4%
all bkgd	40.5%
Data	41.4%
signal @ mH=115 GeV	62.5%
signal @ mH=135 GeV	33.6%
signal @ mH=165 GeV	10.6%

Table 4.9: The solution efficiencies for various MC Higgs signals and backgrounds in the MMC $H \to \tau\tau$ number of physical solutions.

The MMC numbers of solutions rank among the top of the input variables in low and high mass regions, respectively, see Table 4.10. Two conclusions can be drawn from Table 4.10:

- When there is $\tau\tau$ or WW final state (either in the Higgs signals or backgrounds), the ranking of the combined mass improves a bit and number of physical solutions ranks well. The MMC masses do not always rank well, but do supply useful information toward better discrimination.
- The more different the event topology is in Higgs signals and backgrounds, the better the rankings of new variables.

signal	bgnd	mass region	$M_{\tau\tau}$	$M_{\tau\tau}^{combined}$	$N_{\tau\tau}^{soln.}$	$M_{WW}^{combinead}$	$N_{WW}^{soln.}$
$VH\tau\tau$	$t\bar{t}+wjets$	low	28	27	1	17	12
$VH\tau\tau$	$zjets$	low	18	17	19	20	21
$GGF\tau\tau$	MJ	intermediate	23	21	14	11	13
$GGFWW$	MJ	intermediate	26	26	27	22	7
$GGFWW$	$zjets$	high	24	27	7	20	1
$GGFWW$	$t\bar{t}+wjets$	high	30	30	29	25	2

Table 4.10: The rankings of the MMC variables in trainings to distinguish the single channel Higgs signal and single background in low (M_H=105-120 GeV), intermediate (M_H=125-135 GeV) and high (M_H=140-200 GeV) mass regions.

4.10 Multivariate analysis

4.10.1 Overview

The Higgs events are typically hidden in the vast ocean of backgrounds. A search of the SM Higgs is not any easier than looking for a needle in the haystack. None of the kinematic variables is able to identify a Higgs event alone. We need an effective way to synthesize the information from the well-modeled variables and identify the Higgs-like events. The key to our Higgs hunt is application of a machine learning technique, the Boosted Decision Tree (BDT) [6], to model the likelihood of events being the Higgs boson or background in a large data ensemble. The training is based on a huge pool of events from Monte Carlo (MC) simulation. We created millions of MC events, applied appropriate event weights and scaled them down to the SM productions we expected after selections. The final sample was carefully selected to maximize the signal significance over the background. We then input these selected MC events into the training algorithm to create a BDT output ranging between -1 and 1, signifying how much an event is background-like or signal-like.

We grew 400 trees in total, each of which consists of at most 15 nodes. Each node is a decision on an appropriate variable to maximize the separation of the Higgs signal and background events. A Gini Index, $G = 2sb/(s+b)^2$, is defined as the quantitative measure for this maximization, where s is the number of signal events and b is the number of background events in the pool. To determine the best decision, we loop over all variables, perform 20 uniform cuts on each variable per node basis and take the cut that maximizes the reduction of Gini Index after splitting the parent to a pair of daughter nodes. 60% of events are randomly picked in each tree to minimize the effects of finite statistics. We set the learning rate of the algorithm to 0.3, which slows the reweight of misclassified events during boosting. The lower the learning rate is, the more trees need to be grown, and the better accuracy the prediction has in difficult settings. The gradient boosting refers to multiplying the weight of each misclassified event by a mild factor and renormalizing all events. Every time a tree is grown, we review the goodness of Higgs signal-background separation, give a higher weight to events that were misclassified and roll it over to the optimization of the successive recursive tree. Adaboost BDT or Neural Network could do the job reasonably well, but a previous study shows the gradient Boosted Decision Trees (gBDT) outperformed the other two by 10 - 15% and is free of instability resulting from variable orders or inclusion of new variables. Also, poor discriminating variables do not compromise the classifier performance, as long as they are well modeled.

We used well-modeled kinematic variable distributions as inputs in the training. It is equally important to make sure variables adopted are good discriminators between the Higgs signals and the backgrounds. One of our important endeavors was to add novel variables to this depository. We adapted promising variables from other analyses and tailored them to our needs. More importantly, we developed unique

variables that aided our own analysis, featuring the Higgs decay to two tau leptons or W bosons.

We invested significant time and resources to optimize the multivariate training strategy. Our efforts in pursuit of best Higgs limits include:

- Trained single signal against single background in three representative mass regions (super BDT). It was applied in [7] and worked well, though some fluctuations in limits were noted in adjacent hypothesized Higgs mass points.
- Trained all signals against all backgrounds at each mass point (single BDT). Fluctuations are limited and limits are similar.
- Trained single signal against single background at each mass point, and
  - took the mean, maximal or minimal of all BDT's as the final discriminant (smart BDT), per bin or per event basis. These approaches yielded poor limits.
  - took the weighted mean of all BDT's as the final discriminant (smart BDT). The weights could be (a) signal $\times$ background yields in the particular pair, (b) goodness of signal vs. background discrimination in the pair, or (c) product of (a) and (b). These approaches yielded poor limits.
  - used all single signal / single background BDT outputs as input to the second BDT and trained all signals against all backgrounds again (combined BDT). This approach improved the limits significantly (by 20%) over all previous methods, but displayed strong fluctuations across some adjacent mass points.
- Trained all signals against all backgrounds for all mass points together (global BDT). The limits were somewhat compromised by this generic training, whereas we performed the training per mass point before. However there is little fluctuation across all hypothesized Higgs mass points. This is the method we adopted in the final paper [5].

4.10.2 Separating the $\tau\tau$-like and WW-like subsamples

We separated orthogonal $\tau\tau$ and WW subsamples to facilitate the DØ combination in the low and high mass regions, where the SM Higgs $\mu\tau jj$ channel decays are predominantly through two taus and W bosons, respectively. The separation also allows us to exploit the different kinematics of $\tau\tau$ / WW background and signals. We refer to these two subsamples as the "T" and "W" subsamples for convenience in this thesis. The $\mu\tau jj$ samples between hypothesized Higgs mass 105 and 150 GeV consist of both $\tau\tau$ and WW channels. The separation is done for $M_H \leq 150$ GeV. We trained the BDT, BDT_{TW}, with all $\tau\tau$ Higgs events against all WW Higgs events at M_H=125 GeV and applied the results to all samples at all mass points. Figures

4.30-4.32 show the BDT output based on the full 9.7 fb^{-1} of data. Note the shapes of Higgs signals and backgrounds stay about the same between adjacent mass points. Since the Higgs mass is determined around 125 GeV, we took the local minima of the signal BDT output 0.3 as the boundary for T and W subsamples in hypothesized Higgs mass 105-150 GeV. The corresponding signal purities at 125 GeV are 92.1% and 91.2%, respectively, indicating the training and the cut successfully separates the two types of signals.

It is possible to conduct an independent training per mass point and perform the cut on different BDT's separately. We attempted to do so and found severe fluctuation in the LLR and limits throughout the low mass region. It could be that these separate BDT's vary across different mass points and bring in fluctuating numbers of events in the final samples after a uniform cut at 0.3. We could manually tune the cut at each mass point to reduce such fluctuations in event yields, but such fine-tuning seems ad-hoc, whereas a global cut on the same BDT at M_H=125 GeV is well-defined and free of significant fluctuations. The latter is what we applied eventually to separate the T and W subsamples.

The modeling of the T and W subsamples is good upon the separation $BDT_{TW} = 0.3$ at M_H=125 GeV. Figure 4.33 shows a few basic kinematic variable distributions at hypothesized $M_H = 125$ GeV. The same corrections we performed in the combined sample (trigger ratio, MJ/ss shape, jet reweight corrections, etc.) worked perfectly well in each subsample. We noted the figure of merit ξ for $M_{\rm T}^{\ell}$ fell below 0.25 in the W subsample. We therefore dropped it in the subsequent BDT trainings.

The T, W subsample separation turned out to improve the Higgs mass limits. The unseparated sample with single BDT training gave 13.2 times of expected limit at M_H=125 GeV, while the T, W subsample trained separately and combined gave 11.0, about 17% of improvement over the old scheme.

4.10.3 The global BDT training: distinguishing the Higgs signals and the backgrounds

We distinguished the Higgs and background events in the T and W subsamples by two global trainings, referred as "global BDT" for the rest of the thesis. We trained in each subsample all Higgs events at the full range of hypothesized Higgs masses M_H=105 - 150 GeV against all background events. Since a different Higgs mass is assumed and merged in the algorithm, it is necessary to include certain information about the Higgs mass, and the most straightforward solution was adopted: we added a new variable, the Higgs mass M_H, to our input variable list. This addition requires the M_H information present in all samples, and we duplicated 10 times of the data and background sample for each hypothesized Higgs mass of interest. An alternative approach is to assign a random M_H for each data, background event. The advantage of the first treatment is the relative sizes of the Higgs, background samples remain the same as the single mass training in global BDT. Empirical experience shows this will enhance the stability of the training algorithm.

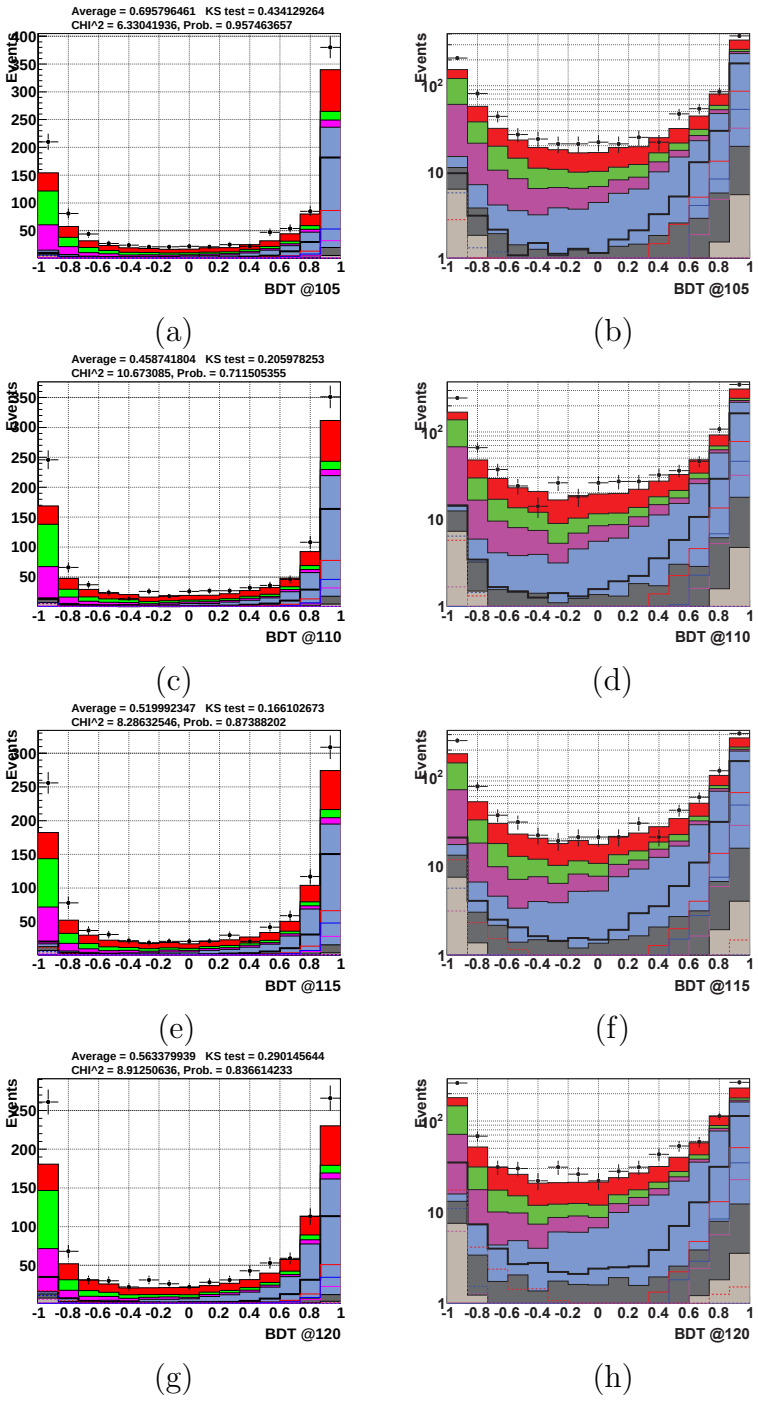


Figure 4.30: BDT trained to distinguish the $\tau\tau$-like and WW-like events in signals, backgrounds and data. Left: linear scale, right: log scale, assuming: (a), (b) $M_H = 105$ GeV, (c), (d) $M_H = 110$ GeV, (e), (f) $M_H = 115$ GeV, (g), (h) $M_H = 120$ GeV. All signals are multiplied by 100.

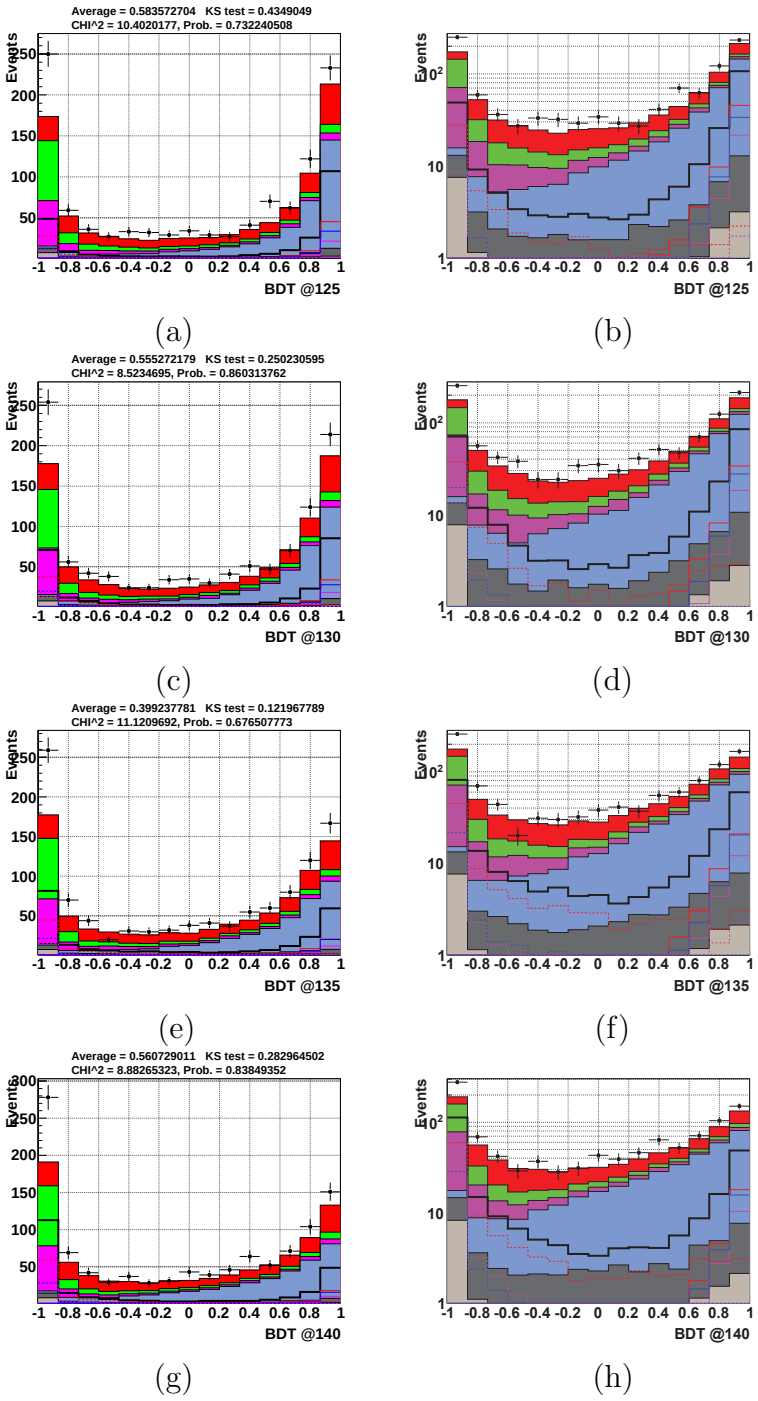


Figure 4.31: BDT trained to distinguish the $\tau\tau$-like and WW-like events in signals, backgrounds and data. Left: linear scale, right: log scale, assuming: (a), (b) $M_H = 125$ GeV, (c), (d) $M_H = 130$ GeV, (e), (f) $M_H = 135$ GeV, (g), (h) $M_H = 140$ GeV. All signals are multiplied by 100.

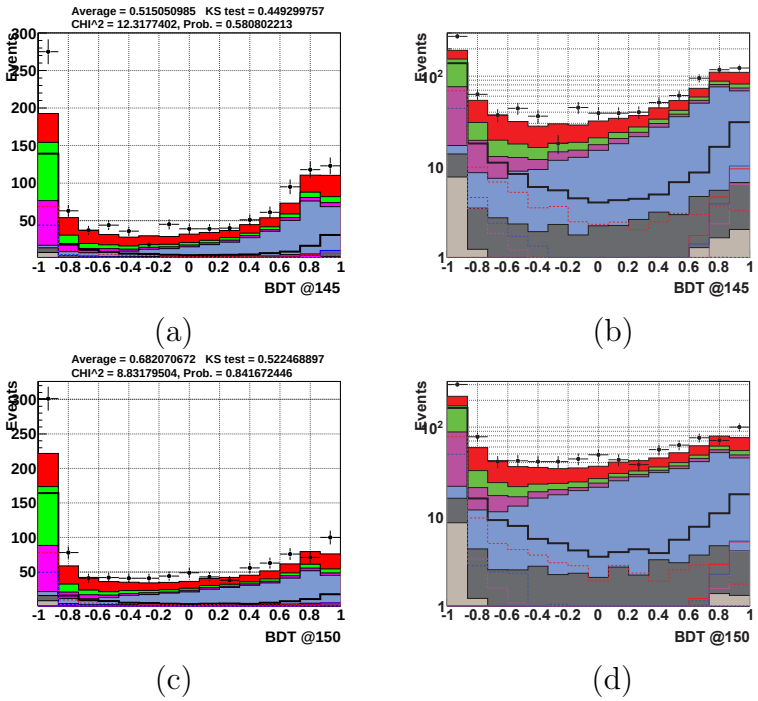


Figure 4.32: BDT trained to distinguish the $\tau\tau$-like and WW-like events in signals, backgrounds and data. Left: linear scale, right: log scale, assuming: (a), (b) $M_H = 145$ GeV, (c), (d) $M_H = 150$ GeV. All signals are multiplied by 100.

We applied an additional event weight at each mass point and scaled the M_H distribution of the backgrounds to the signals. It is important to make sure the added M_H distrubutions in the signal and background samples are the same. As a dicriminative variable, the Higgs mass must be restricted to supply auxiliary information in the global BDT training only. Any discrepancy of the M_H distributions between the signal and background will be recognized and used to distinguish the Higgs signal and background events, which introduces undue bias in the BDT training. We therefore scaled the background event yields to the signals at each mass point according to Tables 4.4, 4.5 in Sec. 4.7. See Fig. 4.34 for comparisons of M_H distributions before and after the scaling in each subsample. Technically, each event weight is multiplied by the ratio of signal and background yields before scaling at each mass point.

All variables in Table 4.6 except M_T^l are used as discussed in the last section in the training of BDT_{TW}. We loop over 20 uniform cuts in all variables and determine the best cut at each decision node, and a ranking representing the discriminating power can be extracted by the frequency each variable is adopted. The variable ranking at $M_H = 125$ GeV is given in Table 4.11 for T and W subsamples separately. Note M_H is ranked the 18th and 20th in two subsamples as designed: it offers the Higgs mass information in the global BDT training, but does not serve as a good discriminator

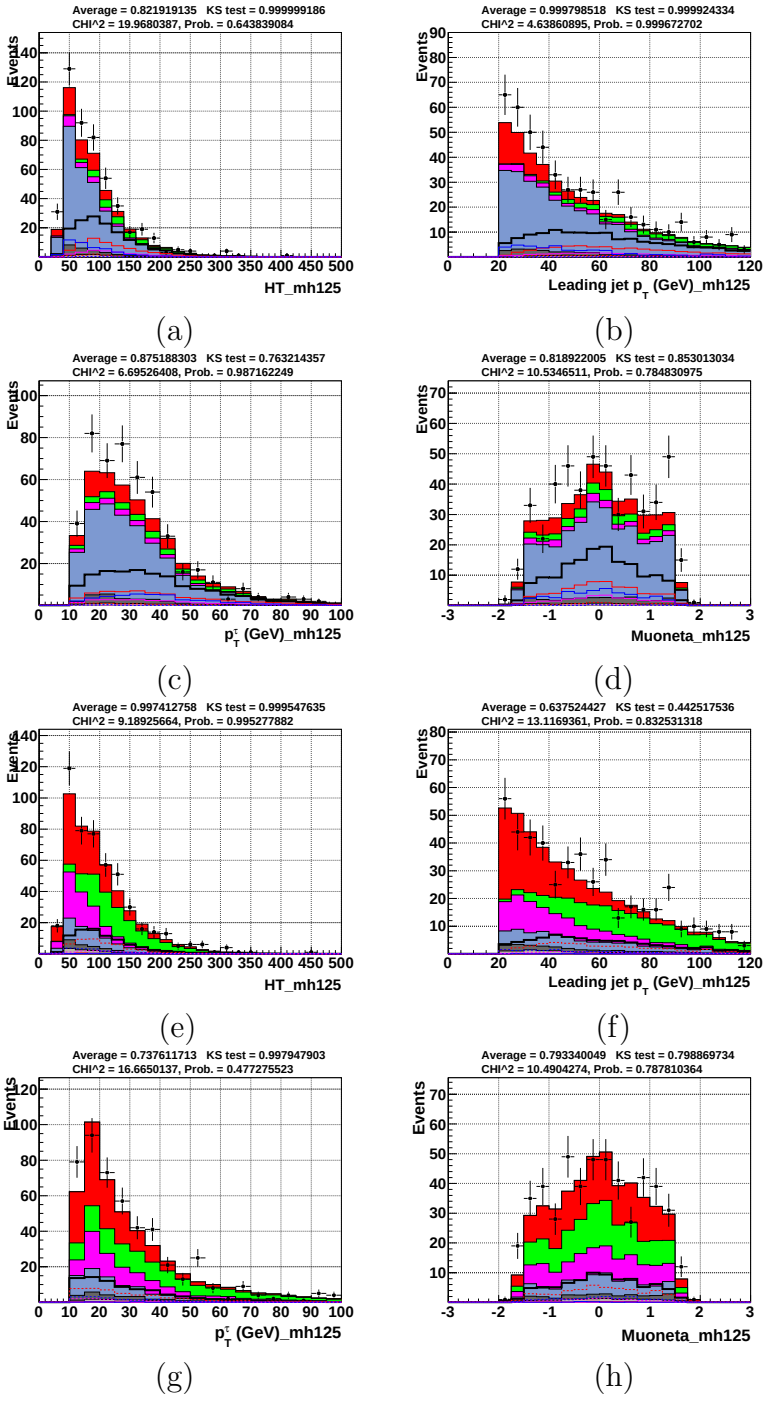


Figure 4.33: The Modeling of typical kinematic variables in the $\tau\tau$ and WW subsamples upon the BDT_{TW} cut at 0.3, assuming $M_H = 125$ GeV. (a) $\tau\tau$ subsample, HT. (b) $\tau\tau$ subsample, leading jet pT. (c) $\tau\tau$ subsample, Tau pT. (d) $\tau\tau$ subsample, muon eta. (e) WW subsample, HT. (f) WW subsample, leading jet pT. (g) WW subsample, Tau pT. (h) WW subsample, muon eta. All signals are multiplied by 100.

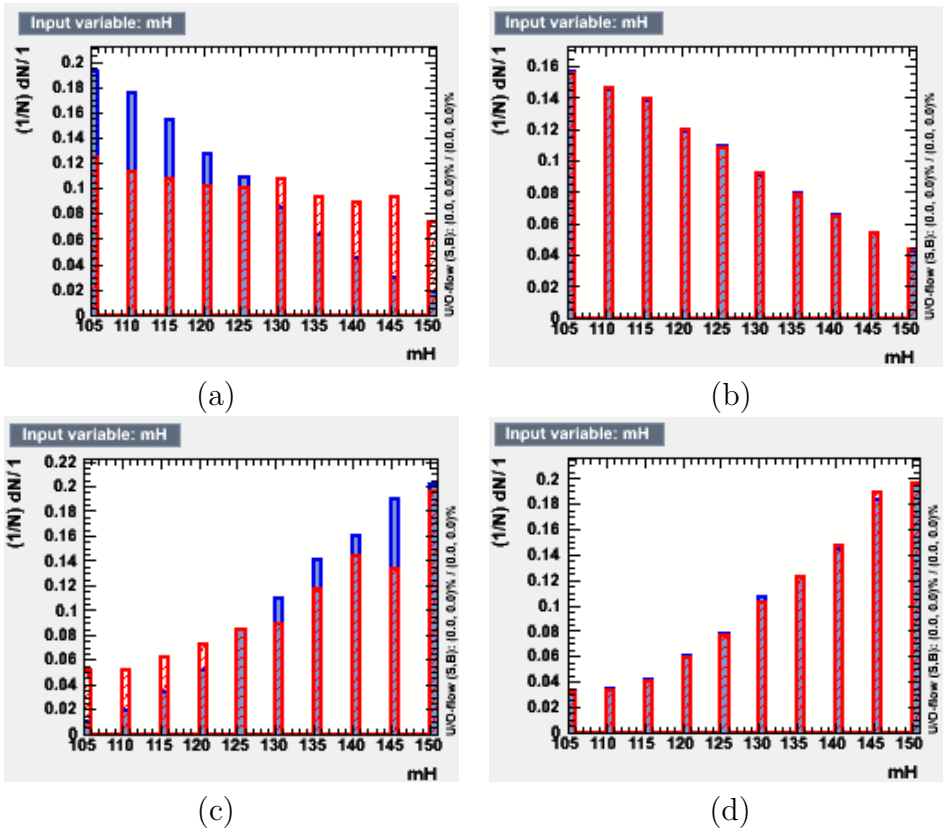


Figure 4.34: The M_H distributions before and after the event yield scale between the backgrounds and the signals. Left: before scale, right: after. Top: the T subsample, bottom: the W subsample. Red: M_H distribution for sum of all signals. Blue: M_H distribution for sum of all backgrounds. All distributions are normalized to 1.

after the scaling in Fig. 4.34.

The global BDT on all hypothesized Higgs masses gives a set of training parameters indicating which variable and what cut are applied at each node, weighted appropriately according to the effectiveness of the decision tree. 400 such trees store all necessary parameters to distinguish the blended Higgs signals with $M_H = 105 - 150$ GeV and the backgrounds. We then applied this global training result to all samples at each mass point for the final discriminant (FD). Since the training includes the information of Higgs mass M_H, the data sample must contain this additional variable too, and we duplicate it ten times and assign the mass points 105-150 to each of them in 5 GeV increment. Each event in data, signal and background samples goes through the BDT_{TW} and is weighted by associated kinematic variables. The BDT output, or the distribution of the final discriminant can then be determined in all samples. The FD distribution is a quantitative measure on how much each event is signal-like or background-like. Figures. 4.35-4.44 show these distributions for hypothesized Higgs masses 105 - 150 GeV.

ranking	$\tau\tau$ subsample	WW subsample
1	M_{jj}	τ_{NN}
2	$M_{\mathrm{T}}(\not{E}_T, \ell+\tau)$	H_T
3	p_T^{τ}	M_{jj}
4	ΔR_{jj}	$\Delta\eta(\mathrm{jj})$
5	S_T	$p_T(\ell, \tau, \not{E}_T)$
6	$M_{\mathrm{T,min}}(\not{E}_T, \ell/\tau)$	$\Delta\phi_{min}(\not{E}_T, \ell/\tau)$
7	M_{T}^{τ}	$M(\ell, \tau, j_1, j_2)$
8	$M_{\tau\tau,\ combined}$	ΔR_{jj}
9	$\cos\theta^*$	$M_{\mathrm{T,min}}(\not{E}_T, \ell/\tau)$
10	τ_{NN}	$M_{WW,\ combined}$
11	$\mathcal{S}$	$M_{\mathrm{T}}(\not{E}_T, \ell+\tau)$
12	$N_{WW\ soln.}$	M_{T}^{τ}
13	$M_{WW,\ combined}$	S_T
14	$M(\ell, \tau, j_1, j_2)$	$N_{\tau\tau\ soln.}$
15	$\Delta\eta(\mathrm{jj})$	p_T^{j1}
16	$N_{\tau\tau\ soln.}$	$\not{H}_T/H_T$
17	p_T^{j1}	$\Delta\phi(\ell, j_1)$
18	M_H	$\Delta\phi_{max}(\not{E}_T, \ell/\tau)$
19	M_{T}^{ℓ}	$\mathcal{S}$
20	$\Delta\phi(\not{E}_T, \not{T}_T)$	M_H
21	$p_T(\ell, \tau, \not{E}_T)$	$N_{WW\ soln.}$
22	$\Delta\phi_{min}(\not{E}_T, \ell/\tau)$	$A(\not{E}_T, \not{H}_T)$
23	$\not{H}_T/H_T$	V_T
24	$\Delta\phi(\ell, j_1)$	p_T^{τ}
25	H_T	$\cos\theta^*$
26	$\Delta\phi_{max}(\not{E}_T, \ell/\tau)$	M_{T}^{ℓ}
27	$\Delta\phi(\ell\tau, j_1 j_2)$	$\Delta\phi(\ell\tau, j_1 j_2)$
28	V_T	$\min\Delta\phi(\not{E}_T, jets)$
29	$\min\Delta\phi(\not{E}_T, jets)$	$\Delta\phi(\not{E}_T, \not{T}_T)$
30	$A(\not{E}_T, \not{H}_T)$	$\Delta\phi_{max}(\not{E}_T, \ell/\tau)$
31	$\Delta\phi_{max}(\not{E}_T, \ell/\tau)$	$M_{\tau\tau,\ combined}$

Table 4.11: The ranking of all input variables used in training BDT_{TW} at $M_H = 125$ GeV. Note M_{T}^{ℓ} is taken in the $\tau\tau$ and WW subsample separation but dropped here due to poor modeling.

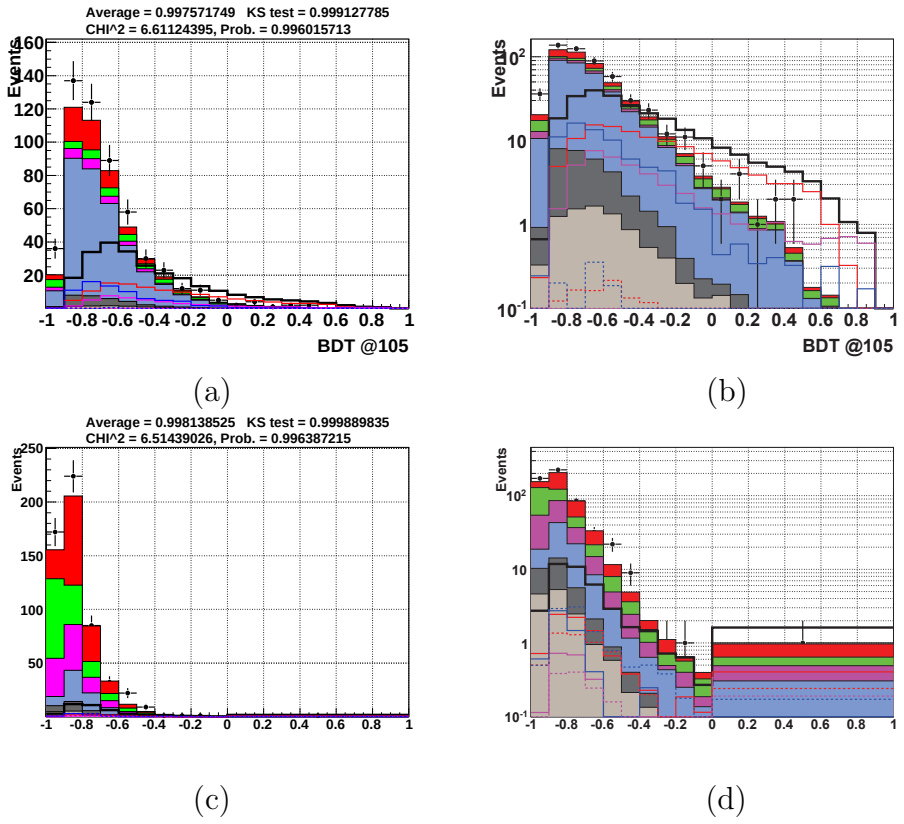


Figure 4.35: BDT outputs for the $\boldsymbol{\mu\tau jj}$ analysis subsamples for $\boldsymbol{m_H = 105}$ GeV. (a), (b): The T subsample outputs, linear and log scales; (c), (d): the W subsample outputs, linear and log scales. The signals are shown multiplied by 100. The linear and log plots share the same figure of merit.

The global BDT method is new and we are the first in Tevatron to apply it to the high energy experimental analysis with real data. It tames the fluctuations across adjacent mass points effectively and offers a promising solution for analyses that suffered from unstable LLR and limit fluctuations (see Sec. 4.12 for more details). We thank Dr. Aurelio Juste for useful discussions in developing this new training strategy.

The sum of MC background events in high BDT bins is sometimes limited and leads to unwanted statistical fluctuations. Starting from the highest bins, we checked the raw number of events in finer intervals, merged those with a few events and made sure all bins have at least 20 raw events. This 20-event cutoff was verified to give stable limits. This criterion leads to 20 uniform bins with size 0.1 in the T subsample. The W subsample has insufficient events and strong fluctuation above BDT$=0$ and we decided to merge all bins with BDT>0 to tame the fluctuation.

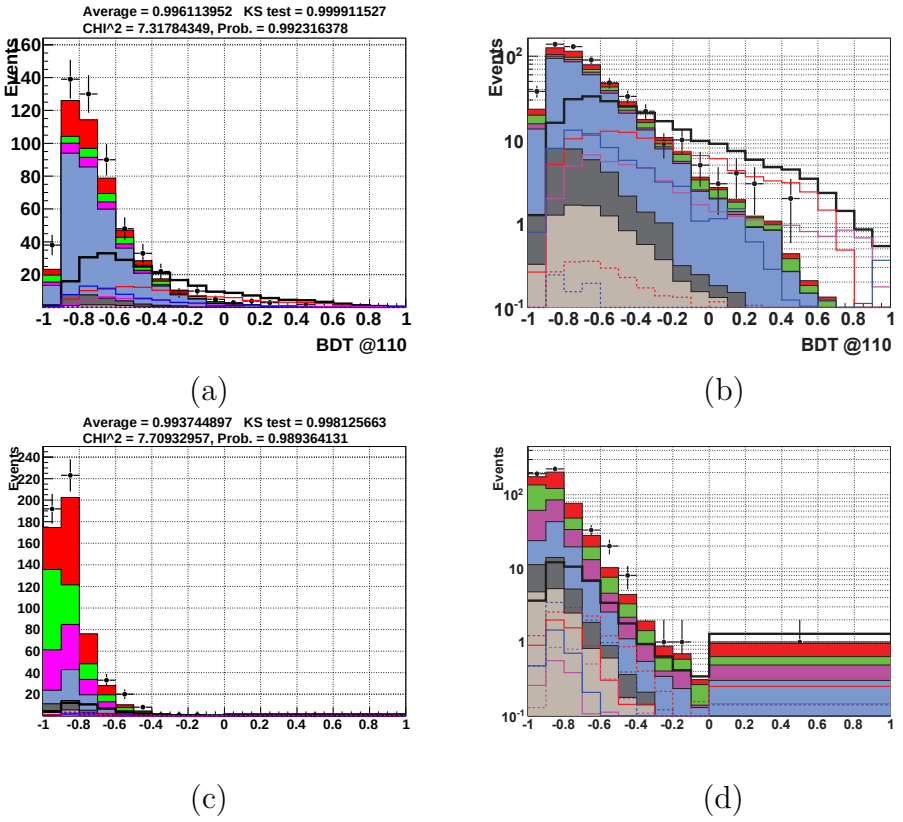


Figure 4.36: BDT outputs for the $\mu\tau jj$ analysis subsamples for $m_H = 110$ GeV. (a), (b): The T subsample outputs, linear and log scales; (c), (d): the W subsample outputs, linear and log scales. The signals are shown multiplied by 100. The linear and log plots share the same figure of merit.

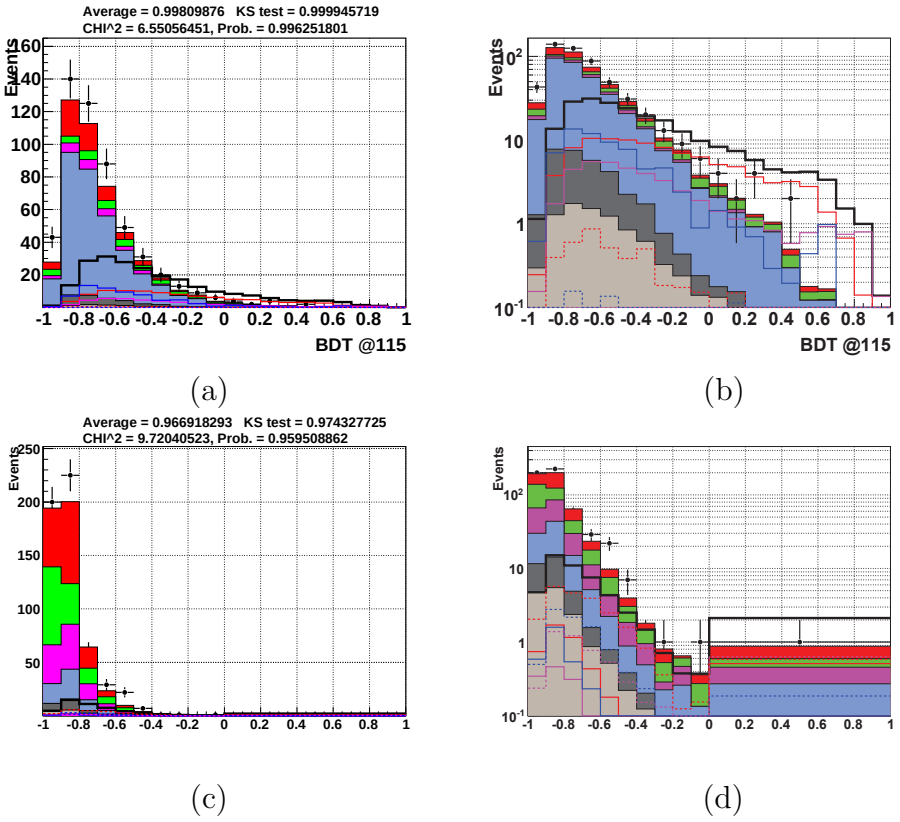


Figure 4.37: BDT outputs for the $\mu\tau jj$ analysis subsamples for $m_H = 115$ GeV. (a), (b): The T subsample outputs, linear and log scales; (c), (d): the W subsample outputs, linear and log scales. The signals are shown multiplied by 100. The linear and log plots share the same figure of merit.

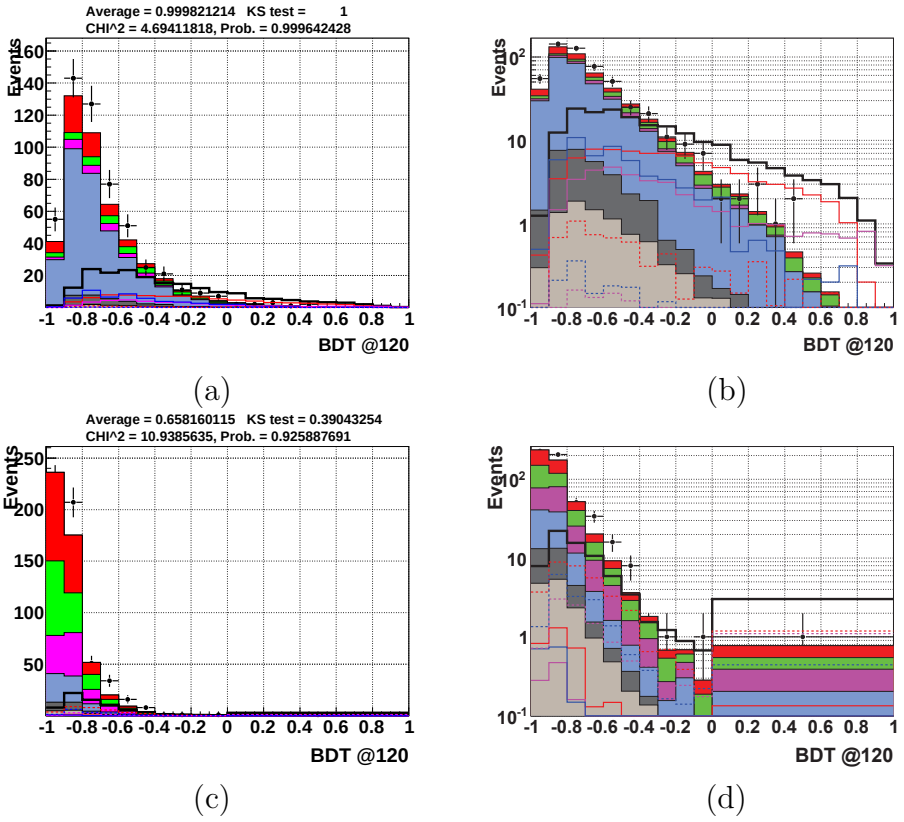


Figure 4.38: BDT outputs for the $\mu\tau jj$ analysis subsamples for $m_H = 120$ GeV. (a), (b): The T subsample outputs, linear and log scales; (c), (d): the W subsample outputs, linear and log scales. The signals are shown multiplied by 100. The linear and log plots share the same figure of merit.

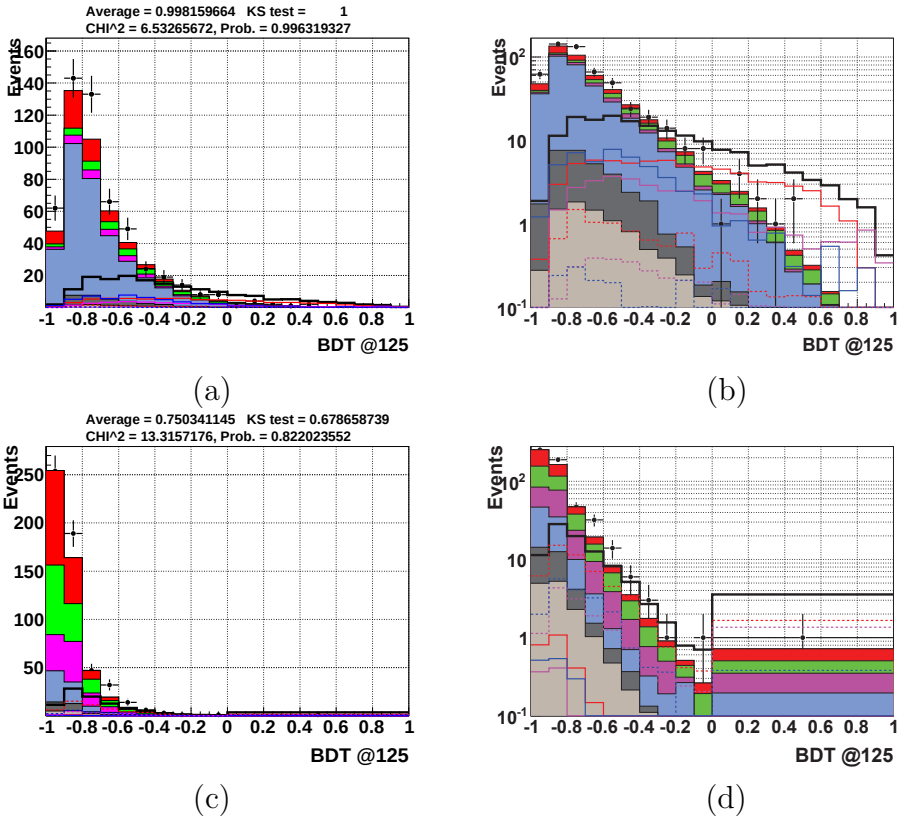


Figure 4.39: BDT outputs for the $\mu\tau jj$ analysis subsamples for $m_H = 125$ GeV. (a), (b): The T subsample outputs, linear and log scales; (c), (d): the W subsample outputs, linear and log scales. The signals are shown multiplied by 100. The linear and log plots share the same figure of merit.

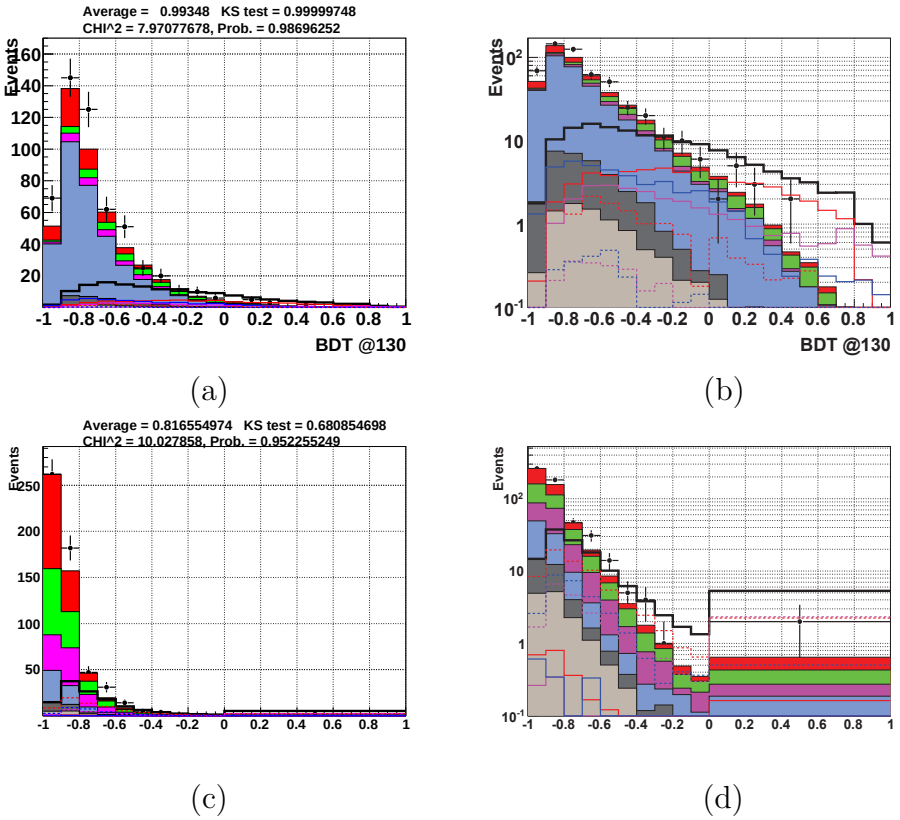


Figure 4.40: BDT outputs for the $\mu\tau jj$ analysis subsamples for $m_H = 130$ GeV. (a), (b): The T subsample outputs, linear and log scales; (c), (d): the W subsample outputs, linear and log scales. The signals are shown multiplied by 100. The linear and log plots share the same figure of merit.

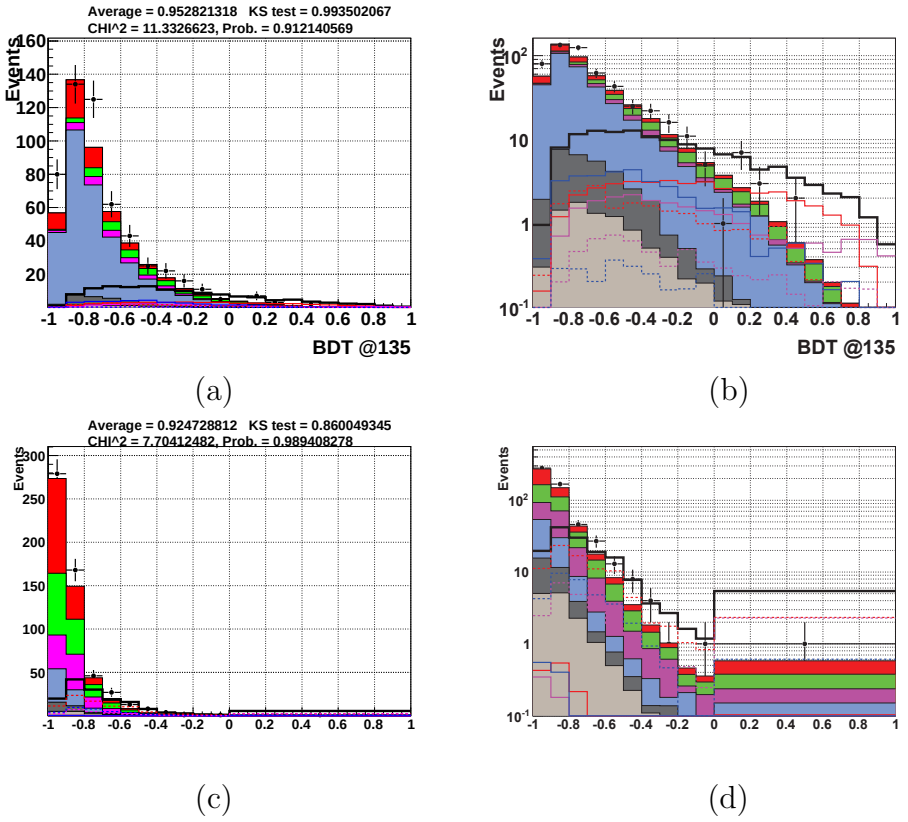


Figure 4.41: BDT outputs for the $\mu\tau jj$ analysis subsamples for $m_H = 135$ GeV. (a), (b): The T subsample outputs, linear and log scales; (c), (d): the W subsample outputs, linear and log scales. The signals are shown multiplied by 100. The linear and log plots share the same figure of merit.

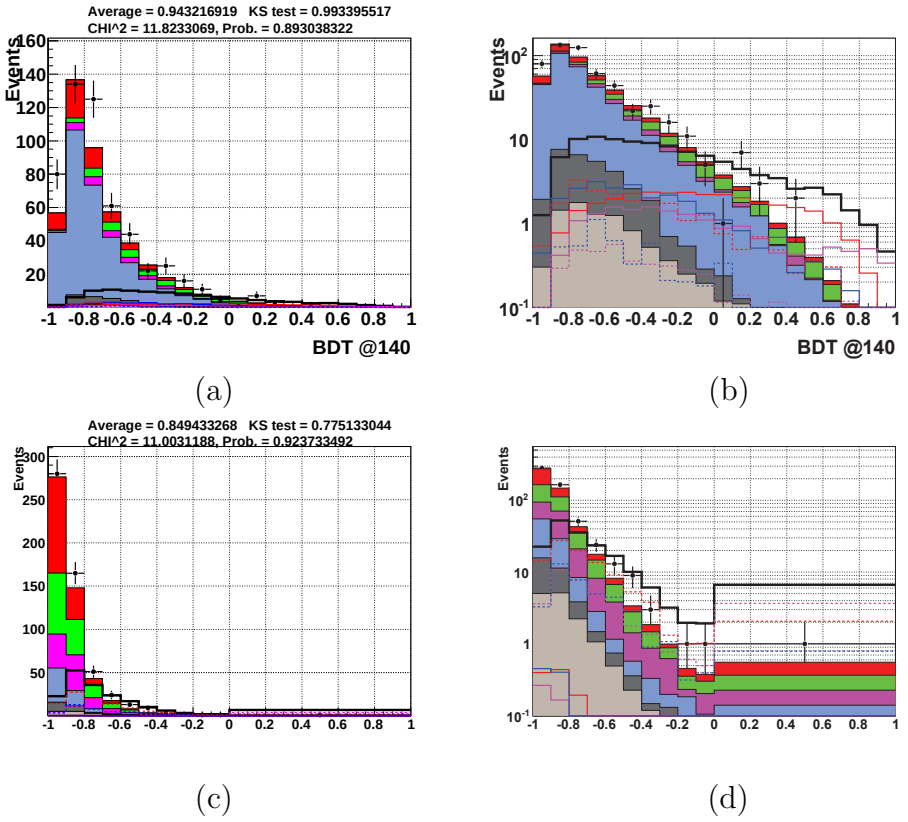


Figure 4.42: BDT outputs for the $\mu\tau jj$ analysis subsamples for $m_H = 140$ GeV. (a), (b): The T subsample outputs, linear and log scales; (c), (d): the W subsample outputs, linear and log scales. The signals are shown multiplied by 100. The linear and log plots share the same figure of merit.

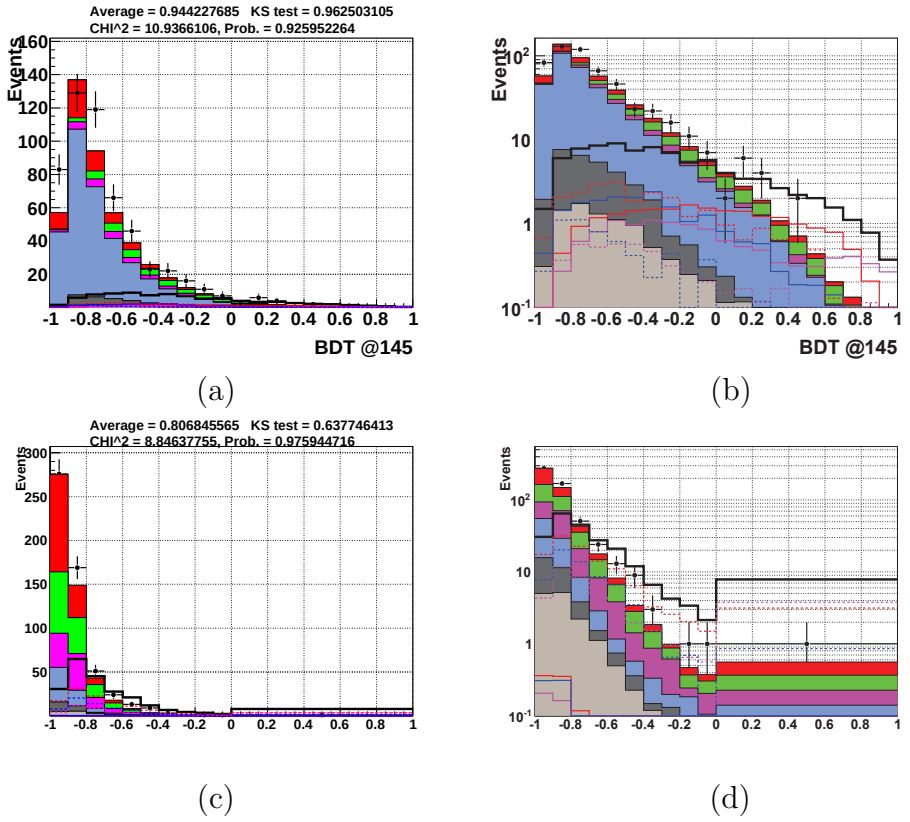


Figure 4.43: BDT outputs for the $\mu\tau jj$ analysis subsamples for $m_H = 145$ GeV. (a), (b): The T subsample outputs, linear and log scales; (c), (d): the W subsample outputs, linear and log scales. The signals are shown multiplied by 100. The linear and log plots share the same figure of merit.

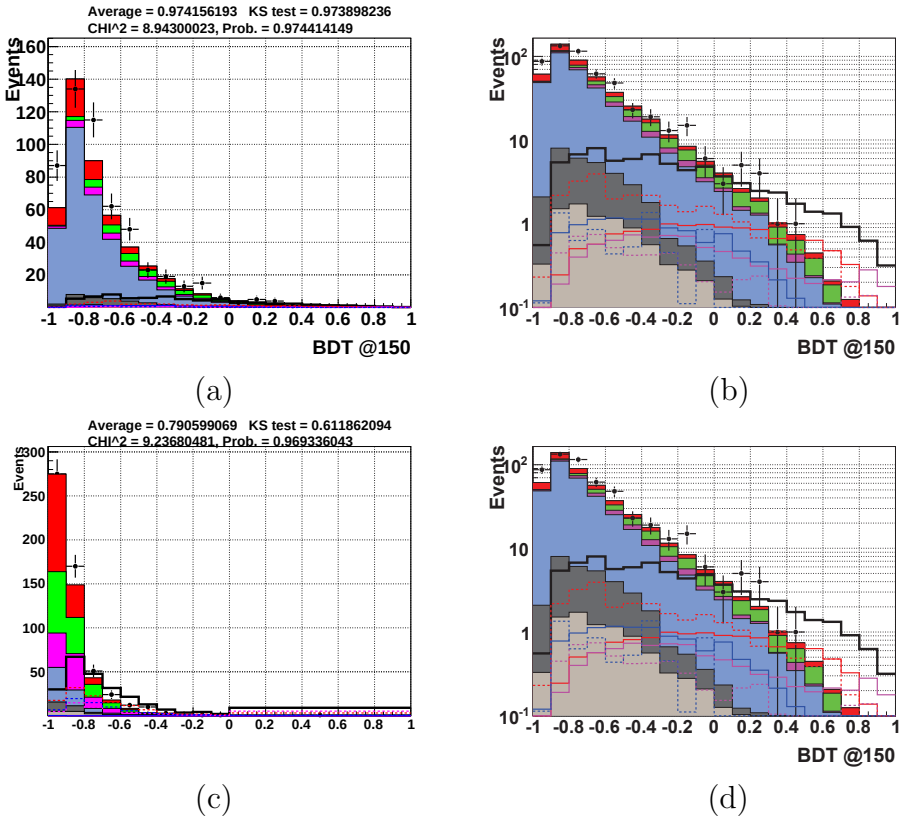


Figure 4.44: BDT outputs for the $\mu\tau jj$ analysis subsamples for $m_H = 150$ GeV. (a), (b): The T subsample outputs, linear and log scales; (c), (d): the W subsample outputs, linear and log scales. The signals are shown multiplied by 100. The linear and log plots share the same figure of merit.

4.11 Systematic uncertainties

A few generic methods apply when we evaluate the various systematic uncertainties. Most typically we vary each uncertainty source by ± 1 standard deviation and compare the BDT output to the nominal value obtained from the signal sample using the baseline BDT training weights. The resulting difference is taken as the systematic uncertainty from each source. In some cases we follow the standard DØ /Higgs Group determinations, or compare special control samples. If the systematic variation carries a shape dependence in the BDT distribution, we transmit such dependence to the limit setting program. Tables 4.12 lists the sources and sizes of systematic uncertainties.

Source	type	Uncertainty (%)
Luminosity (DØ specific)	flat	4.1
Luminosity (Tevatron common)	flat	4.6
μ ID, track match, iso.	flat	2.9
SingleμOR trigger	flat	5
AllTrigger/SingleμOR relative eff.	flat	7
τ energy correction	flat	9.8
τ track efficiency	flat	1.4
τ selection by type	flat	5.5, 4.0, 6
W/Z+light flavor XS	flat	6.0
$t\bar{t}$, single top XS	flat	7.0
diboson XS	flat	6.0
VH signal XS	flat	6.2
VBF signal XS	flat	4.9
GGF signal XS normalization	flat	33
GGF signal XS PDF	flat	29
jet vetex confirmation	shape	2-11 (T) 1-12 (W)
Jet ID/reco eff.	shape	1-12 (T) 1-12 (W)
Jet E resolution.	shape	2-11 (T) 1-10 (W)
JES	shape	2-10 (T) 1-11 (W)
PDF	flat	1.6 (sig), 2.0 (bknd)
MJ normalization	flat	5.3
MJ shape	shape	15 -20

Table 4.12: Systematic uncertainties (in percent) for the $\mu\tau jj$ analysis. The range of jet shape uncertainties is shown separately for the T and W subsamples.

The "flat" uncertainties only affects the overall normalization of the BDT output. They include:

- The uncertainty on the integrated luminosity is considered to be flat with value of 6.1% [43]. This uncertainty can be decomposed into a component unique to DØ of 4.1% and a 4.6% component owing to uncertainty in the luminosity cross section which is fully correlated with CDF.

- The uncertainty on the efficiencies of μ ID, track match, and isolation is taken to be flat with value of 2.9% [44]. This inclues a 2% additional systematic added in quadrature to the muon ID systematic to take into account muons selected with $p_T^{\mu} < 20$ GeV.

- The uncertainty on the single muon OR trigger efficiency is taken to be flat with value of 5% [44]. The scaling from the single muon OR to the inclusive triggers has a 7% uncertainty leading to an overall uncertainty on the trigger of 8.6%.

- The signal cross section uncertainties are taken from the standard TeVNPHWG and Higgs group accords [37] as shown in Table 4.12.

- The uncertainty due to the tau energy scale correction is considered to be flat and is taken as an average (weighted by our relative yields for different tau types) of the change in yield of taus when changing the tau energy scale by $\pm 1\sigma$ for that tau type, using the tau-type-specific uncertainties from the Tau ID group p20 determination [45]. The uncertainty is 9.8%.

- The uncertainty due to the tau track efficiency is taken to be flat with a value of 1.4%.

- The uncertainty for the tau selection, background subtraction and variations in efficiency with tau momentum is taken as the weighted average over tau types, (5.5%, 4.0% and 6%) respectively for tau types (1, 2, 3) [46] (weighted averge = 4.6%).

- The uncertainty due to the PDFs is obtained from the CTEQ pdf eigenvector set which gives the largest change in BDT shapes, for signals and backgrounds separately. We take flat uncertainties of 1.6% for signals and 2% for backgrounds.

- The normalization uncertainty on the MJ background is estimated from (a) the uncertainty on the ρ parameters (Table 4.2) giving the expected OS/SS ratio for each tau type in the MJ enriched sample and (b) the statistical uncertainties on the raw SM MC events that are subtracted from the SS signal data to obtain the MJ yields; and (c) the statistical uncertainties on the SS data yield. The weighted average for the three tau types is used. We do not include uncertainties on the SM cross sections, which are accounted for separately.

The "shape" uncertainties refer to the shape dependent uncertainties. We vary the contributing source (jet energy scale, jet identification, jet energy resolution, jet vertex confirmation) by $-\sigma$, go through the same procedure of the signal sample BDT determination at hypothesized Higgs mass M_H=125 GeV for T and W

subsamples separately, and take the change induced as the shape systematic uncertainties. We determined the jet-related uncertainties with RunIIb2 data only, and the same uncertainty profile is applied to M_H=105 - 150 GeV in signal sample, as no strong dependence on the Higgs mass or the run epochs is observed in our previous study. Practically we only vary -1σ and symmetrize it to give $\pm 1\sigma$ shapes. A previous study shows this symmetrization hardly changes the Higgs mass limits, and the change in the shapes of systematics is only minimal. We coarsen the binning by merging 3 or 4 adjacent bins to reduce the statistical fluctuation. The binning in each uncertainty histogram matchs the binning in the corrsponding BDT output. The following items summarize the shape uncertainties we evaluated:

- The uncertainties due to the jet-id and jet reconstruction efficiency, jet energy resolution, jet energy scale and vertex confirmation (Run 2b only) are estimated by varying the appropriate factors by -1σ and symmetrized to give the $+1\sigma$ values. The fractional uncertainties in BDT output relative to the baseline are shown in Appendix C.

- The MJ background shape uncertainty is obtained using the difference between our baseline sample using the MJ enriched control sample described in Sec. 4.6 and the MJ background estimated in the SS signal sample after subtraction of the SM MC backgrounds. The fractional uncertainties in the BDT output relative to the baseline are shown in Appendix D.

4.12 Calculating the limits

We compute the cross section upper limits with the modified frequentist method as implemented in COLLIE [17], a DØ standard software specializing in such calculations. The algorithm calculates a negative log-likelihood (LLR) as Eq. (4.19),

$$\chi = -2ln(Q) = -2ln(\frac{e^{-(s+b)}(s+b)^d}{d!}\frac{e^{-b}b^d}{d!}), \tag{4.19}$$

where s, b are expected numbers of signal, background estimations and d is the observed number of data in the actual experiment. An ensemble of psudoexperiments is prepared by fluctuating the numbers of signals and backgrounds with a Poisson distribution in each BDT bin, assuming the data consist of signal plus background (LLR_{s+b}, the TEST hypothesis), or the background only (LLR_b, the NULL hypothesis). The test statistic is a value derived from data to measure the probabilities of these two hypothesis. It is by definition $Q = \mathrm{LLR}_{s+b}/\mathrm{LLR}_b$, the ratio of signal plus background hypothesis and background-only hypothesis and summed over all signal channels and BDT output bins.

The confidence level (CL) represents how the observed data are consistent with certain hypothesis. We define a few quantitative measures to aid the evaluations of the Higgs mass limits:

- CL_{s+b}: gives the probability that both signal and background exist in data, or the fraction of signal+background pseudo experiments more background-like than observed in data, i.e. the TEST hypothesis.

- CL_b: gives the probability only the background is present in data, or the fraction of background-only pseudo experiments more background-like than observed data, i.e. the NULL hypothesis.

- CL_s: The signal confidence level, $CL_s = 1 - CL_{s+b}/CL_b$, denotes how likely the SM Higgs boson does not exist. A traditional Frequentist hypothesis test relies solely on CL_{s+b}, but the corresponding exclusion could be pathological when the data fluctuate down significantly below the background prediction. This could result in high observed LLR values and strong exclusion, but in reality it is not true due to lack of statistics. The CL_s definition, or the modified-Frequentist approach, is designed to make the limit evaluation more robust to potential mis-modeling. Large downward fluctuations in the data or poor background models can generate exclusions that may not be reproducible with larger statistics or a modified background model, and the division aims to cancel such coherent fluctuation to some level.

The Higgs expected and observed limits presented later are computed as ratio of SM Higgs cross section so as to exclude the Higgs at 95% confidence level. COLLIE scales up the SM Higgs cross section by this ratio until 95% CL exclusion is reached.

Figure 4.45 shows an example of the LLR distribution COLLIE creates. The observed data level (vertical solid line) is located at the median of the background-only LLR, indicating a full agreement between data and the background-only model, but in reality they do not need be. A comparison between the TEST, NULL and observed test statistics reveals the information on the agreement between the observed data and the background models. The median LLR value for the NULL hypothesis ($d_i = b_i$) is positive and the median for the TEST hypothesis ($d_i = s_i + b_i$) is negative. Neglecting terms with higher orders than $(\frac{s}{b})^2$, these two median values should be symmetric about the origin.

The information from the LLR values in Fig. 4.45 is expanded to a two-dimensional plot with the x axis some model parameters (typically the Higgs mass points) and the y axis the LLRs. The signal rate and therefore LLR evolve as a function of the model parameter, see Fig. 4.46 for an example. Shown in the figure are the median values of the NULL (black dashed line), TEST (red dashed line) and observed (black solid line) LLRs, the $\pm 1\sigma$ (green band) and the $\pm 2\sigma$ (yellow band) of the NULL hypothesis.

3 types of limits are computed in the COLLIE algorithm:

- (1). The CLFast limit, as the limit without systematic uncertainties. It can be computed quickly and is a good diagnostic figure of merit to evaluate if certain training method enhances the limits.

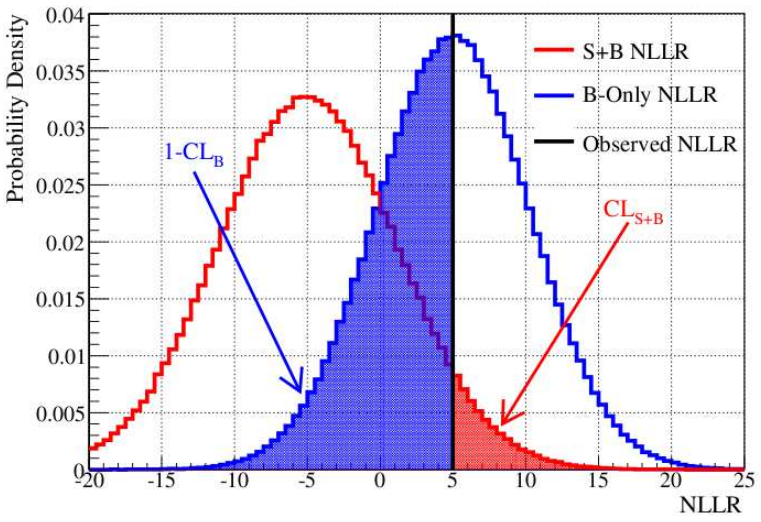


Figure 4.45: The example LLR distributions for the TEST (red) and NULL (blue) hypotheses. The shaded red area represents CL_{s+b} and shaded blue area represents $1-CL_b$.

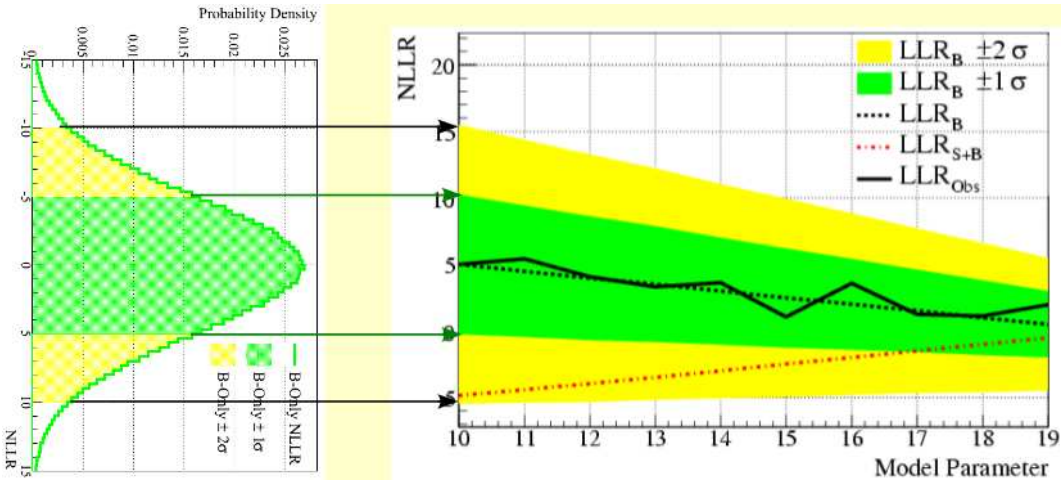


Figure 4.46: The evolution of the LLR values as functions of some model parameter. Note this plot does not come from the $\mu\tau jj$ analysis and is for illustration only. The NULL (black dashed line), TEST (red dashed line) and observed (black solid line) LLRs, the $\pm1\sigma$ (green band) and the $\pm2\sigma$ (yellow band) of the NULL hypothesis are shown in the figure.

- (2). The AWW limits, or CLFastApproximation. Systematic uncertainties are taken into account in this scenario for better understanding of signal+background and background-only hypotheses consistency with data, but the MC statistical uncertainties are not included. It is built on the assumption of Gaussian uncertainties in the systematics.

- (3). The CLFit2 limits, the full estimation with the MC statistics and systematic uncertainties. Each hypothesis (signal+background or background-only) is fitted for LLR determination. This is the most robust limit evaluation but the slowest.

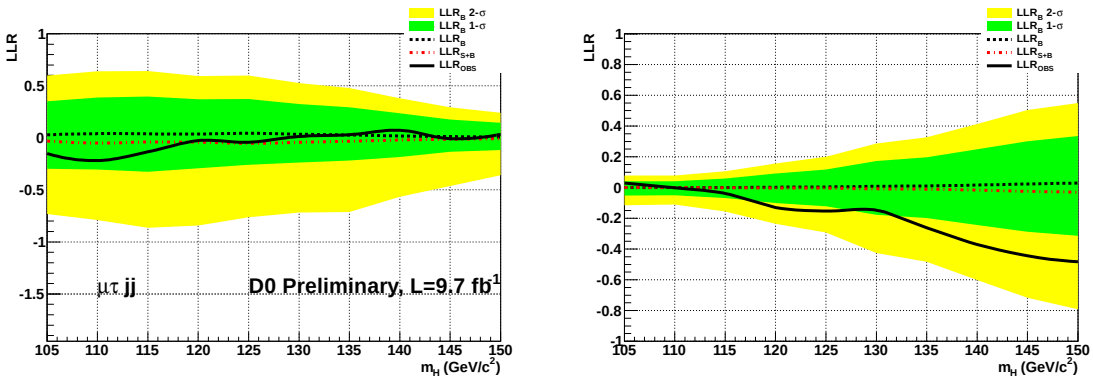
Both the AWW and CLFit2 limit computations employ a profile likehood algorithm [15]. They rewrite the test statistic in Eq.(4.19) to address the degrading effects of uncertainties by introducing the "nuisance parameters", which are the physics parameters that generate all systematic uncertainties,

$$\chi = -2ln(Q(data|\theta_0, \theta_1) = -2ln\left(\frac{P(data|H_1,\hat{\theta_1})}{P(data|H_0,\hat{\theta_0})}\right), \tag{4.20}$$

where θ_1 is the set of nuisance parameters for the TEST hypothesis H_1, θ_0 is the set of nuisance parameters for the NULL hypothesis H_0, and $\hat{\theta_i}$ represents the nuisance parameters that maximize H_i. Consequently $P(data|H_i, \hat{\theta_i})$ are the maximized probabilities that the psudo-data contain signal + background ($i = 1$, TEST hypothesis) and background only ($i = 0$, NULL hypothesis). The likelihood now becomes a function of signal, bkgd, data, and nuisance parameters. The AWW and CLFit2 algorithms fit both the NULL and TEST hypotheses, maximize their probabilities independently and construct the test statistic with systematic uncertainties. The AWW toggles on a fast approximation flag in the uncertainty estimation and serves as a faster evaluation for the limits.

We fed T and W subsample BDT outputs separately into COLLIE to determine the Higgs expected and observed LLR and limits. Figure 4.47 shows the expected LLR for the signal plus background and background-only hypotheses as functions of M_H for the two subsamples, and Fig. 4.48 shows the LLR for two subsamples combined. The green and yellow bands illustrate the $\pm 1\sigma$ and $\pm 2\sigma$ ranges for the expected background-only LLR. The solid line gives the observed LLR determined by data. The 95% Higgs exclusion limits relative to the SM Higgs cross sections are listed in Tables 4.13, with limit plots in Figs. 4.49, 4.50.

A companion analysis of the Higgs boson search based on the $e\tau jj$ final state [9] adopts similar selections and multivariate analysis. The combined $e\tau jj$, $\mu\tau jj$ CLFit2 limits are shown in Fig. 4.51 separately for the T and W subsamples. When the T and W limits are combined again, we obtain the final limits for the search of the SM Higgs with the $\ell\tau jj$ final state in Fig. 4.52.



(a) (b)

Figure 4.47: LLR for the $\mu\tau jj$ analysis for the (a) T subsample and (b) the W subsample as a function of Higgs mass, for expected background only (black dotted line), expected with signal + background (red dotted line) hypotheses, and the observed values (solid black line). The $\pm 1\sigma$ and $\pm 2\sigma$ variations from the expected background only hypothesis are shown in green and yellow bands respectively.

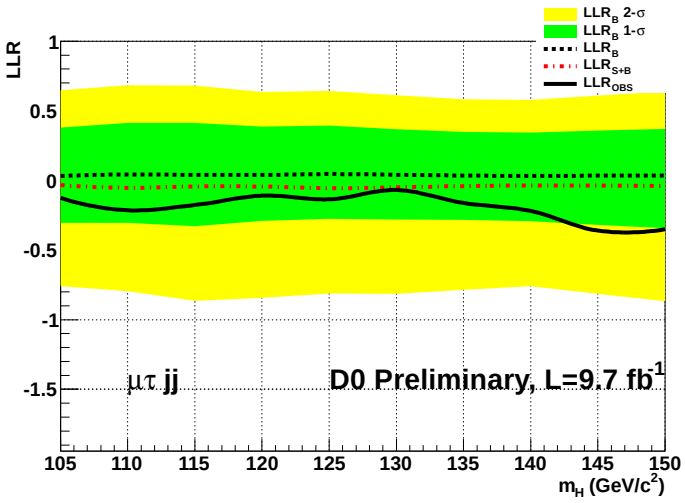
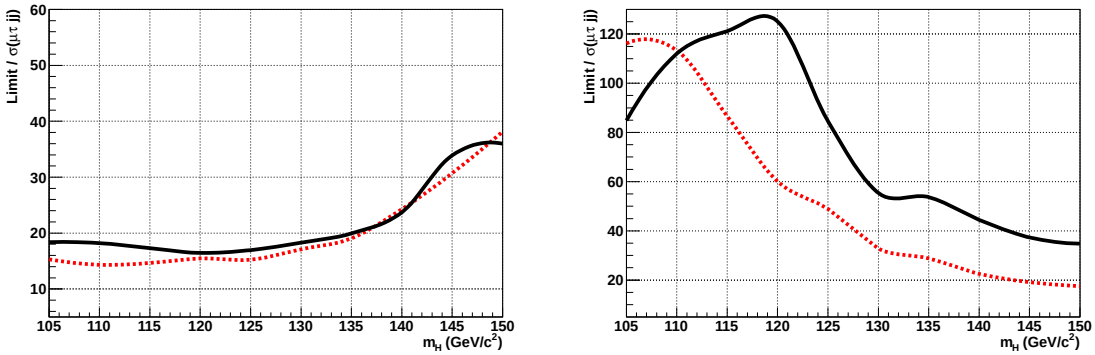


Figure 4.48: LLR for the $\mu\tau jj$ analysis for the combined T and W subsamples as a function of Higgs mass, for expected background only (black dotted line), expected with signal + background (red dotted line) hypotheses, and the observed values (solid black line). The $\pm 1\sigma$ and $\pm 2\sigma$ variations from the expected background only hypothesis are shown in green and yellow bands respectively.

	CLfast						CLfit2					
	$\tau\tau$		WW		$\tau\tau + WW$		$\tau\tau$		WW		$\tau\tau + WW$	
mass	exp.	obs.	exp.	obs.	exp.	obs.	exp.	obs.	exp.	obs.	exp.	obs
105	10.8	23.5	71.6	124.4	10.7	24.7	15.3	18.4	116.2	84.9	14.1	16.3
110	10.6	23.3	71.4	137.3	10.5	24.8	14.3	18.2	113.3	112.0	13.9	16.8
115	10.9	23.2	54.4	128.1	10.4	25.8	14.6	17.3	86.6	121.2	13.8	17.2
120	11.7	23.0	36.3	100.4	11.4	28.5	15.5	16.5	60.1	125.1	14.6	17.6
125	12.3	23.1	27.7	73.6	11.0	28.8	15.3	17.0	48.9	84.6	14.4	18.4
130	14.0	25.2	19.7	52.7	10.9	29.9	17.1	18.3	32.8	55.4	15.3	17.4
135	15.5	27.7	17.2	46.0	10.9	30.8	19.1	19.9	28.8	53.7	15.6	21.0
140	19.4	33.7	13.8	38.7	10.8	31.5	24.3	23.7	22.5	44.5	15.9	22.8
145	24.1	47.9	11.2	31.6	9.9	30.3	30.7	33.9	19.1	37.3	15.1	23.7
150	30.0	53.7	10.4	29.4	9.7	28.9	38.1	36.0	17.5	34.8	14.9	22.2

Table 4.13: The ratio of the expected and observed 95% C.L. limits to the SM Higgs cross section expectations for the $\mu\tau jj$ analysis.



(a) (b)

Figure 4.49: 95% upper cross section limits for the $\mu\tau jj$ analysis as a function of Higgs mass: (a) T subsample limits, (b) W subsample limits. Black solid line: observed limits, red dashed line: expected limits.

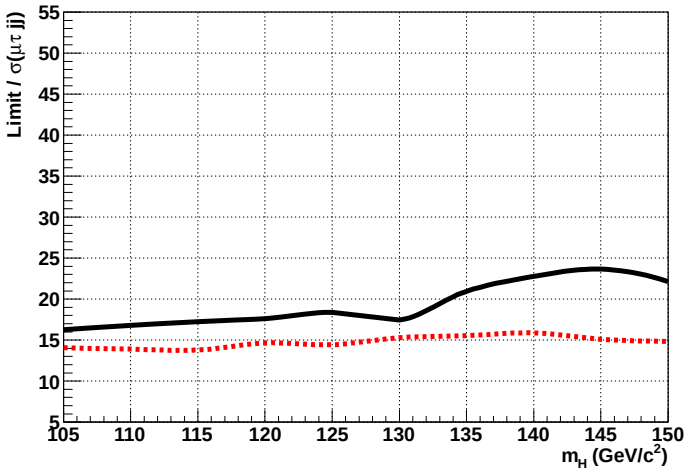
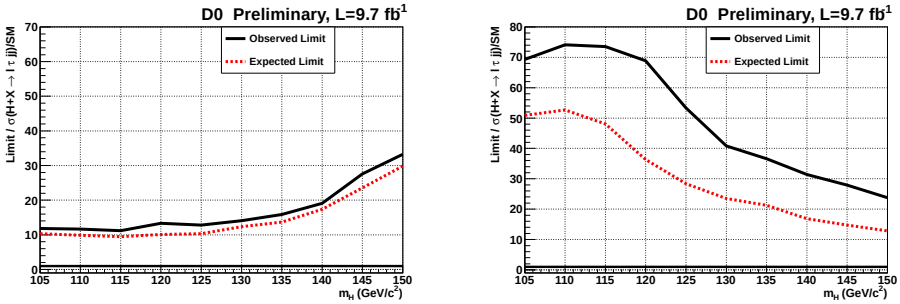


Figure 4.50: 95% upper cross section combined limits for the $\mu\tau jj$ analysis as a function of Higgs mass. Black solid line: observed limits, red dashed line: expected limits.



(a) (b)

Figure 4.51: 95% upper cross section limits from CLfit2 for the $\mu\tau jj$ and $e\tau jj$ analyses combined for (a) T subsample, (b) W subsample.

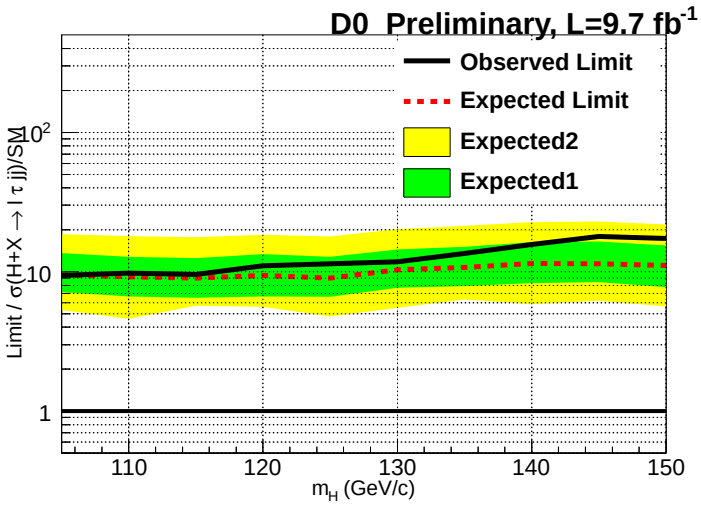


Figure 4.52: 95% upper cross section limits from CLfit2 for the $\mu\tau jj$ and $e\tau jj$ analyses combined.

Chapter 5

Conclusions

We presented the search for the SM Higgs boson through the $\mu\tau jj$ channel using the full 9.7 fb^{-1} of Tevatron RunII data set collected in the DØ detector at Fermilab. No evidence of the Standard Model Higgs boson is observed with hypothesized Higgs mass between 105 GeV and 150 GeV, but the data do not exclude it either. We gave the upper limits on the ratio of the 95% CL exclusion to the SM Higgs cross section.

The $\mu\tau jj$ channel is one of many channels being searched in DØ for the SM Higgs boson. Important Higgs decay channels include $H \to b\bar{b}$, $H \to \gamma\gamma$, $H \to \tau\tau$ and $H \to WW^{(*)}$, see Sec. 4.1 and Fig. 4.3 for details and [47] for an overview. The $H \to b\bar{b}$ process gives the best limit in the low mass region and drives the Tevatron combination. The $H \to \gamma\gamma$ process is quite rare, but hadron collisions in the Tevatron do not produce many di-photon background events either and the EM calorimeter is good at detecting photons. This is an unusually clean channel and eventually led to the discovery of the SM-like Higgs boson in LHC, though it does not outperform the other channels in Tevatron due to lower event yields compared to the LHC. Figure 5.1 summarizes the Higgs mass inferred by all searching channels at $M_H = 125$ GeV in Tevatron. Figure 5.3 [47] shows the Tevatron combined LLR and limits from both DØ and CDF detectors.

The $\mu\tau jj$ analysis separates the T and W subsamples and hence contributes to both the $H \to \tau\tau$ and $H \to WW^{(*)}$ decay channels, see Fig. 5.2 [47]. Comparing the combined $H \to \tau\tau$, $H \to WW^{(*)}$ with the $\mu\tau jj$ subsample LLRs and limits in Figs. 4.47-4.50, we see the search of the Higgs boson with $\mu\tau jj$ final state contributes in both the low and high mass regions with the T and W subsamples, respectively, as is expected from their branching ratio dependence on the Higgs test mass. In general, this dissertation offers a uniform limit improvement over all Higgs mass points ranging from 105 to 150 GeV. Our analysis presents the best Tevatron limits on the $H \to \tau\tau$ channel.

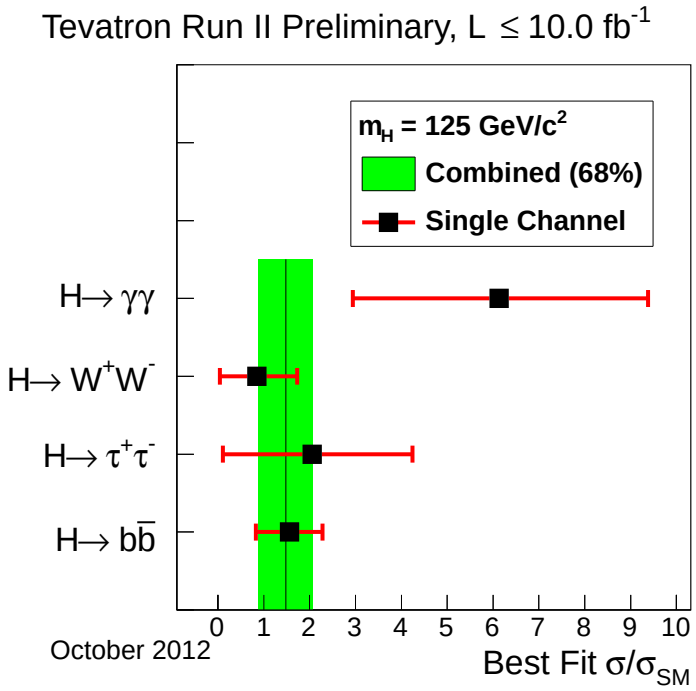


Figure 5.1: The Higgs mass indicated by $H \to b\bar{b}$, $H \to \gamma\gamma$, $H \to \tau\tau$ and $H \to WW^{(*)}$ processes in Tevatron combination.

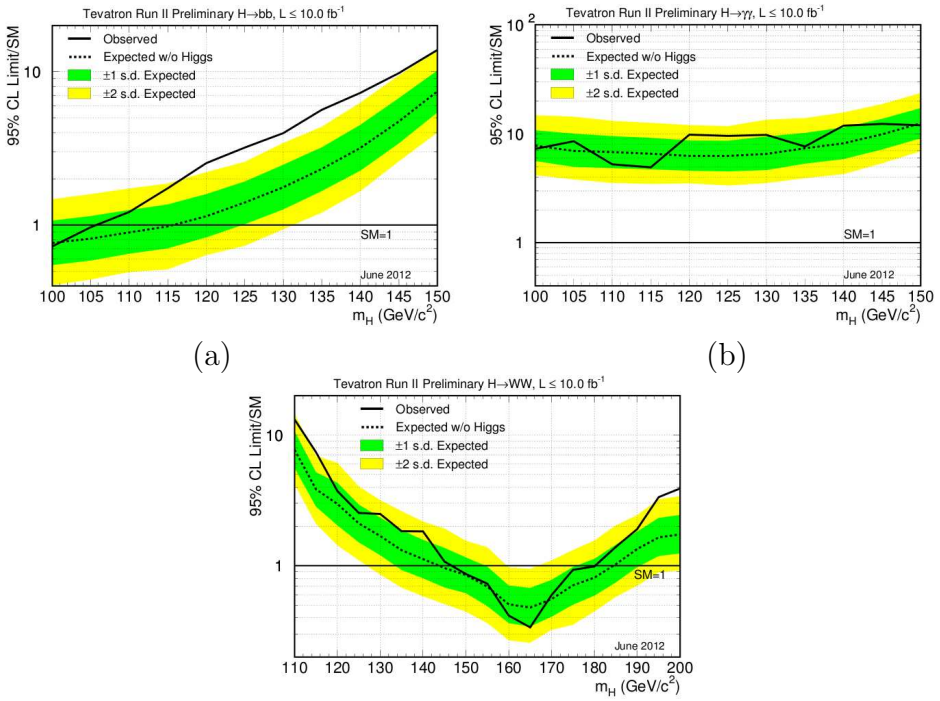


Figure 5.2: The Tevatron Higgs mass limits for individual channels in multiple of the SM prediction for test Higgs masses 105-200 GeV, updated to June 2012. (a) $H \to b\bar{b}$, (b) $H \to \gamma\gamma$, (c) $H \to WW^{(*)}$.

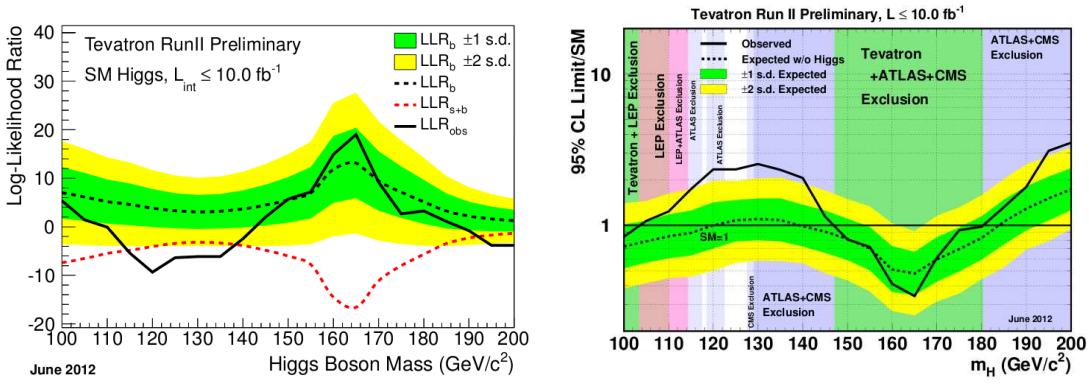


Figure 5.3: The Tevatron combined results for test Higgs masses 105-200 GeV updated to June 2012. (a) LLR, (b) limits in multiple of the SM prediction.

Bibliography

[1] A. Elagin, P. Murat, A. Pranko, A. Safonov, "Mass Reconstruction Technique for Resonances Decaying to $\tau\tau$", Nucl. Instrum. Methods Phys. Res. **A** 654, 481 (2011).

[2] Nima Arkani-Hamed, Andrew G. Cohen, Howard Georgi, "Electroweak symmetry breaking from dimensional deconstruction", Physics Letters **B** 513, 232 (2001).

[3] Nima Arkani-Hamed, Andrew G. Cohen, Thomas Gregoire, Jay G. Wacker, "Phenomenology of electroweak symmetry breaking from theory space", JHEP 0208, 020 (2002).

[4] S. Dimopoulos, H. Georgi. "Softly Broken Supersymmetry and SU(5)", Nuclear Physics **B** 193, 150 (1981).

[5] V.M. Abazov *et al.* (DØ Collaboration), "Search for the Higgs boson in lepton, tau and jets final states", submitted to Phys. Rev. **D**. arXiv(hep-ex): 1211.6993 (2012).

[6] A. Hoecker *et al.* PoS ACAT:040 (2007).

[7] V.M. Abazov *et al.* (DØ Collaboration), "Search for the standard model Higgs boson in tau lepton pair final states", Phys. Lett. **B** 714, 237 (2012).

[8] Subhendu Chakrabarti, Paul Grannis, Katy Tschann-Grimm, Ian Howley, Wanyu Ye, DØ Note 6019.

[9] Subhendu Chakrabarti, Paul Grannis, Ian Howley, Wanyu Ye, DØ Note 6272.

[10] C. Quigg, Gauge Theories of the Strong, Weak, and Electromagnetic Interactions, the Benjamin/Cummings Publishing Company, Inc. (1983).

[11] S. Dawson, "Introduction to Electroweak Symmetry Breaking", BNL-HET-99/1, arXiv:hep-ph/9901280.

[12] M. Veltman, Acta. Phys. Pol. **B**8 (1977) 475.

[13] C. Tully, A. Schwartzman, "Primary Vertex Reconstruction by Means of Adaptive Vertex Fitting", DØ Note 4918.

[14] S. Parke and S. Veseli, "Spin correlation of WH and Wbb", Phys. Rev. **D** 60, 093003.

[15] Wade Fisher, "Calculating Limits for Combined Analyses", DØ Note 4975.

[16] J. Beringer *et al.* (Particle Data Group), Phys. Rev. **D** 86, 010001 (2012).

[17] Wade Fisher, "Collie: A Confidence Level Limit Evaluator", DØ Note 5595.

[18] D. Chakraborty *et al.* "Reconstruction of tau leptons in hadronic final states at DØ in Run 2", DØ Note 4210.

[19] S. Atkins, *et al.* "Correction For the MC-Data Difference in the Jet Response at DØ", DØ Note 6143.

[20] S. Atkins, *et al.* "Correction For the MC-Data Difference in the Jet Response at for DØ Run IIB", DØ Note 6144.

[21] Nikola Makovec and Jean-Francois Grivaz, "Shifting, Smearing and Removing Simulated Jets", DØ Note 4914.

[22] O. Brein, A. Djouadi and R. Harlander, Phys. Lett. **B** 579, 149 (2004); M.L. Ciccolini, S. Dittmaier and M. Krämer, Phys. Rev. **D** 68, 073003 (2003).

[23] C. Anastasiou, R. B̃oughezal and F. Petriello, J. High Energy Phys. 04, 003 (2009); D. de Florian and M. Grazzini, Phys. Lett. **B** 674, 291 (2009).

[24] E.L. Berger and J. Campbell, Phys. Rev. **D** 90, 073011 (2004).

[25] K.A. Assmagan *et al.*, "The Higgs Working Group: Summary Report 2003", arXiv:hep-ph/0406152.

[26] T. Sjöstrand *et al.*, Computer Phys. Commun. 135 (2001) 238;

[27] J. Pumplin *et al.*, J. High Energy Phys. 07 (2002) 12.

[28] A. Djouadi *et al.*, Comput. Phys. Commun. 108, (1998) 56-74.

[29] MLM matching procedure, http://www-d0.fnal.gov/computing/MonteCarlo/generators/common_alpgen.html

[30] S. Jadach *et al.*, Comput. Phys. Commun. 76 (1993) 361.

[31] R. Brun and F. Carminati, CERN Program Library Long Writeup W5013, 1993 (unpublished).

[32] K. Melnikov and F. Petriello, Phys. Rev. D 74, 114017 (2006).

[33] N. Kidonakis, Phys. Rev. D 74, 114012 (2006); U. Langenfeld, S. Moch, and P. Uwer, Phys. Rev. D 80, 114012 (2006).

[34] J. Campbell and K. Ellis, "MCFM, *Monte-Carlo for FeMtobarn processes*", http://mcfm.fnal.gov/.

[35] https://plone4.fnal.gov/P1/DØWiki/comp/caf/caffaq/LumiReWeight.

[36] V.M. Abazov et al. (DØ Collaboration), Phys. Rev. Lett. 100, 102002 (2008).

[37] T. Aaltonen *et al.* (CDF and D0 Collaborations), Phys. Rev. Lett. 104, 061802 (2010).

[38] P. Svoisky S. Protopopescu, "Tau identification with neural networks for p17 data", DØ Note 5094.

[39] A. Schwartzman, DØ Note 4254.

[40] U. Aglietti *et al.*, "Tevatron-for-LHC Report: Higgs", FERMILAB-CONF-06-467-E-T, arXiv:hep-ph/0612172.

[41] S. Dittmaier *et al.* (LHC Higgs Cross Section Working Group), "Handbook of LHC Higgs Cross Sections", CERN-2011-002, arXiv:1101.0593.

[42] Greenlee H.B. (DØ Collaboration), FERMILAB-Conf-93/363-E

[43] http://www-d0.fnal.gov/d0dist/dist/packages/lm_access/devel/doc

[44] The separate contributions are discussed in DØ Note 6025, and the final systematic is available in DØ Note 6083.

[45] http://www-d0.hef.kun.nl//askArchive.php?base=agenda&categ=a09917&id=a09917s1t1/transparencies .

[46] http://www-d0.hef.kun.nl//askArchive.php?base=agenda&categ=a10724&id=a10724s4t2/transparencies (p. 19).

[47] CDF Collaboration, DØ Collaboration, "Updated Combination of CDF and D0 Searches for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data", FERMILAB-CONF-12-318-E, CDF Note 10884, D0 Note 6348, arXiv(hep-ex):1207.0449.

Appendix A

Alltrigger to SingleMuonOR ratio scale functions

We adopted the all-trigger acceptance in data to increase the event yield. The yield improvement over the SingleMuonOR data is estimated to be about 40%. The MC samples were however passed through the SingleMuonOR trigger, and we scaled the MC signals and backgrounds by a trigger weight determined from the data. The trigger weight is determined as the ratio of the all-trigger data subtracting the MJ background and SingleMuonOR data subtracting the MJ background, all after full selections. The subtraction of the MJ background is aimed to deduct unnecessary correlation with irrelevant processes - we are only interested in the scaling of pure MC events.

We fit the ratio as a constant $f_i(p_T^\mu) = C_i$ by τ type in RunIIa with the muon pT, and a parabolic function of muon η, $f_i(\eta_\mu) = A_i - B_i\eta_\mu^2$ in RunIIb, per τ type i=1,2,3. Eq. (A.1), (A.2) give their fit parameters and uncertainties, respectively. The RunIIb fit features 2 parameters and their uncertainties are correlated. We decorrelated them by a matrix transformation and computed the uncertainty for the fit function. The red line is the fit result and the green lines indicate the $\pm 1\sigma$ uncertainty. Figures A.1-A.12 show the dependence of the ratio and the fit in basic kinematic variables including the p_T and η of μ, τ, jet1, jet2. We also defined a few variables closely related to jets and plotted the dependence on $H_T = \Sigma|p_T|$ jets, $T_T = H_T + p_T^\tau$, $S_T = H_T + p_T^\tau + p_T^\mu$, and $Q_T = H_T + p_T^\tau + p_T^\mu + \not{E}_T$. The original plan was to investigate the limits with 0 and 1 jet also too, though this dissertation and the analysis only explored the final state with 2 or more jets ($\mu\tau jj$).

$$\begin{aligned} type1 &: C_1 = 1.31 \pm 0.42; \\ type2 &: C_2 = 1.18 \pm 0.11; \\ type3 &: C_3 = 1.16 \pm 0.21. \end{aligned} \tag{A.1}$$

$$\begin{aligned} type1 &: A_1 = 1.35 \pm 0.22, B_1 = 0.24 \pm 0.19; \\ type2 &: A_2 = 1.49 \pm 0.11, B_2 = 0.14 \pm 0.09; \\ type3 &: A_3 = 1.50 \pm 0.23, B_3 = 0.40 \pm 0.14. \end{aligned} \tag{A.2}$$

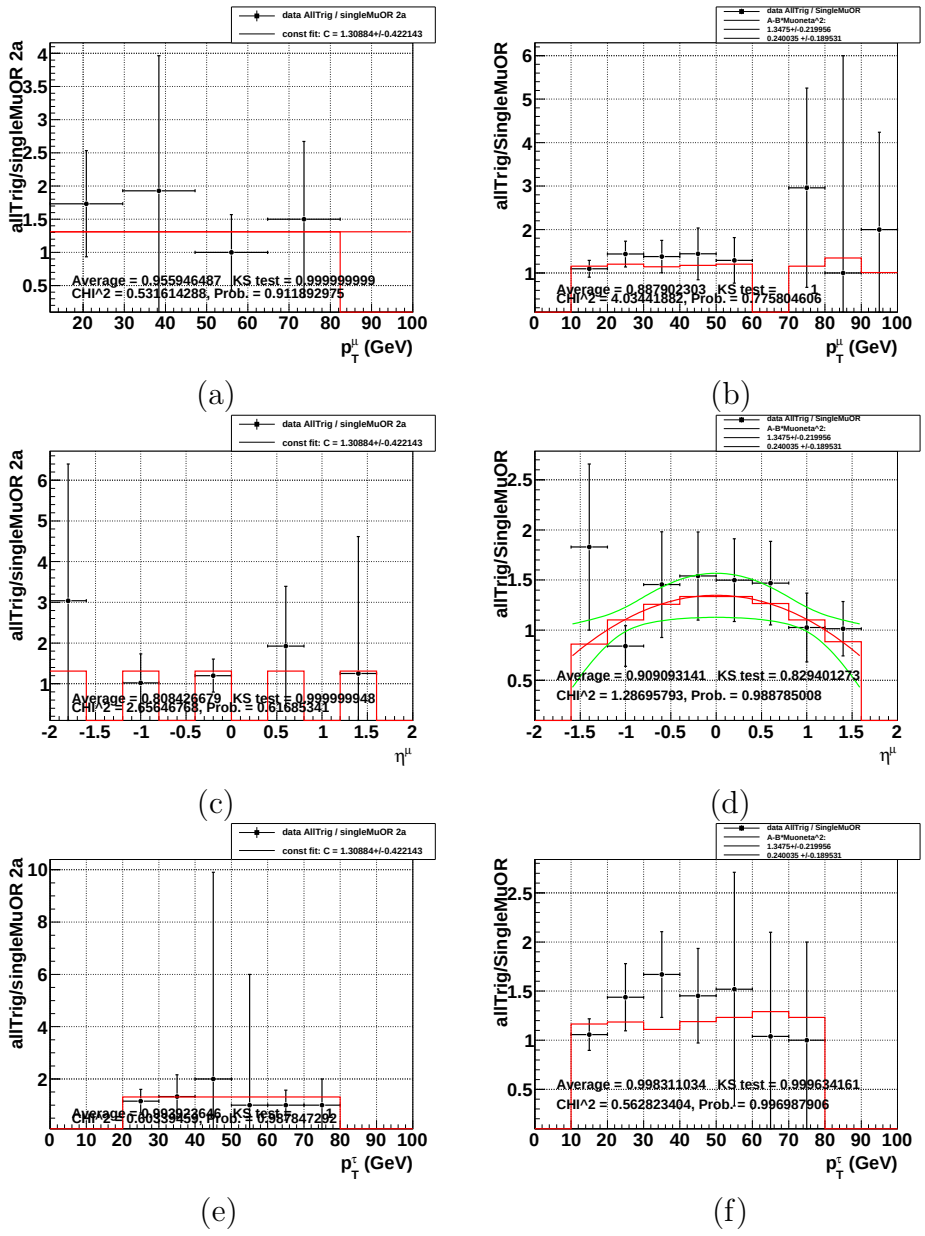


Figure A.1: Ratio of AllTrigger events to SingleMuOR events in the inclusive $\mu\tau$ selection for type 1 τ's : (a) p_T^μ, 2a; (b) p_T^μ, 2b; (c) η_μ, 2a; (d) η_μ, 2b; (e) p_T^τ, 2a; (f) p_T^τ, 2b. The red line in (d) is the parabolic η_μ fit and the green lines indicate the $\pm 1\sigma$ uncertainty.

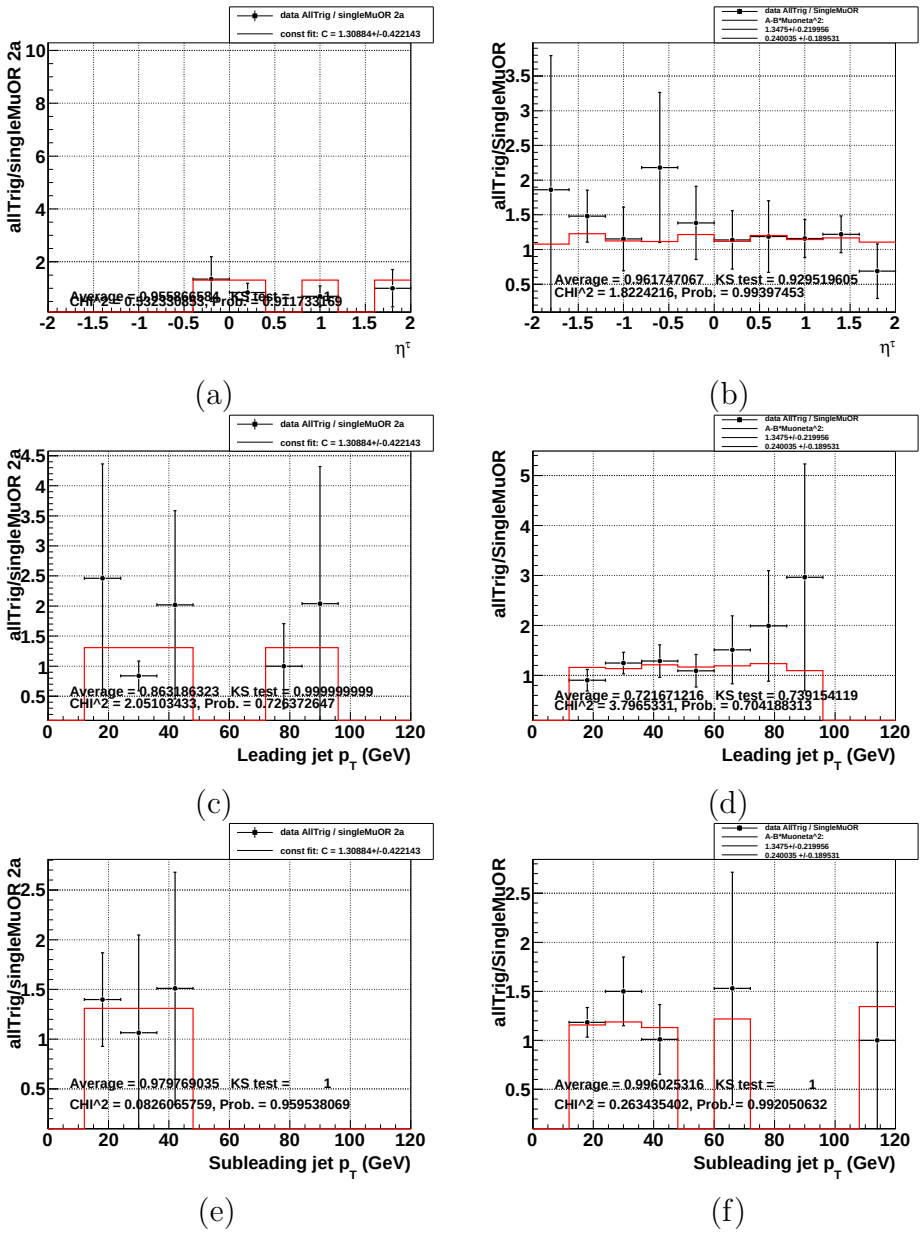


Figure A.2: Ratio of AllTrigger events to SingleMuOR events in the inclusive $\mu\tau$ selection for type 1 τ's : (a) η_τ, 2a; (b) η_τ, 2b; (c) p_T^{jet1}, 2a; (d) p_T^{jet1}, 2b; (e) p_T^{jet2}, 2a; (f) p_T^{jet2}, 2b.

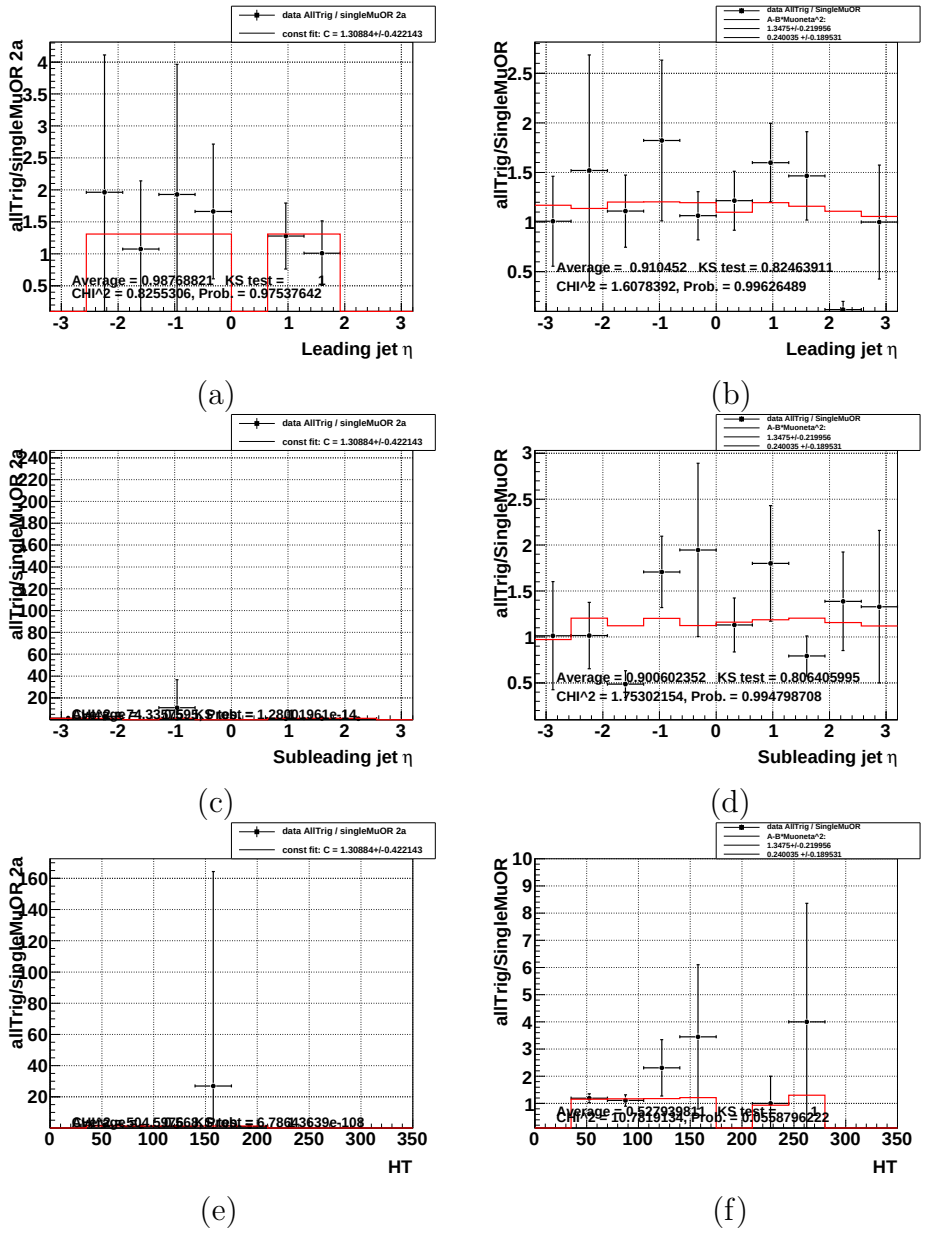


Figure A.3: Ratio of AllTrigger events to SingleMuOR events in the inclusive $\mu\tau$ selection for type 1 τ's : (a) η_{jet1}, 2a; (b) η_{jet1}, 2b; (c) η_{jet2}, 2a; (d) η_{jet2}, 2b; (e) H_T, 2a; (f) H_T, 2b.

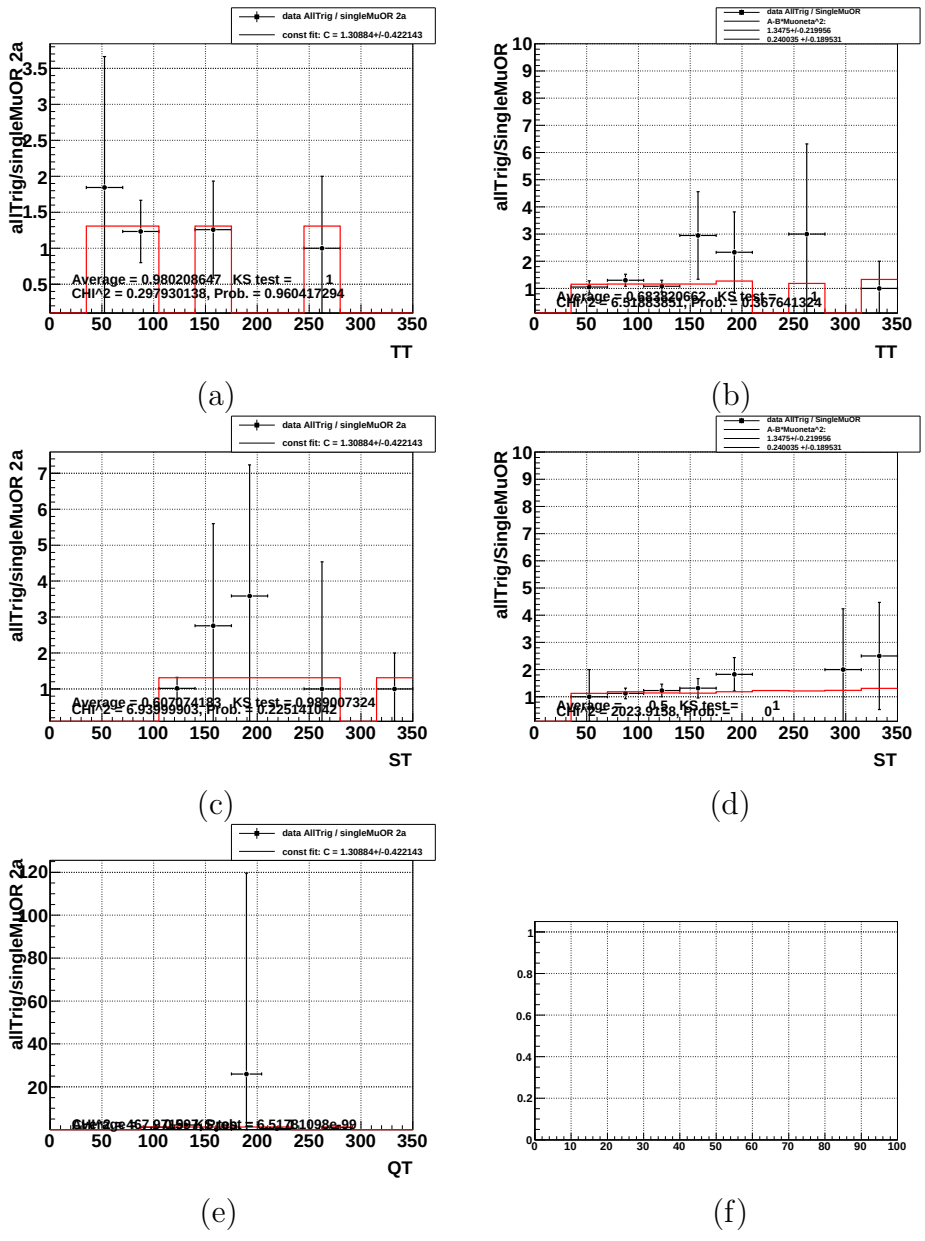


Figure A.4: Ratio of AllTrigger events to SingleMuOR events in the inclusive $\mu\tau$ selection for type 1 τ's : (a) T_T, 2a; (b) T_T, 2b; (c) S_T, 2a; (d) S_T, 2b; (e) Q_T, 2a.

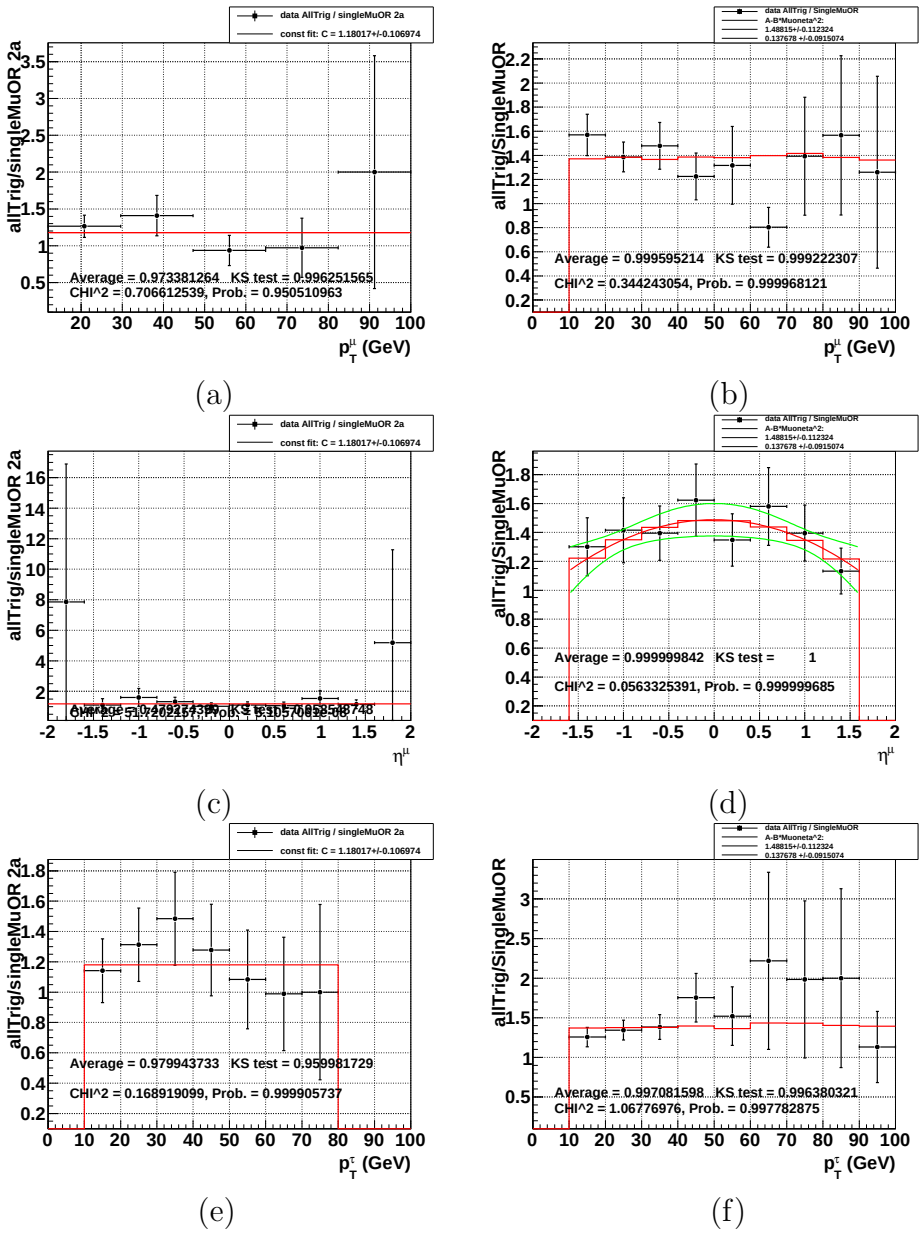


Figure A.5: Ratio of AllTrigger events to SingleMuOR events in the inclusive $\mu\tau$ selection for type 2 τ's : (a) p_T^μ, 2a; (b) p_T^μ, 2b; (c) η_μ, 2a; (d) η_μ, 2b; (e) p_T^τ, 2a; (f) p_T^τ, 2b. The red line in (d) is the parabolic η_μ fit and the green lines indicate the $\pm 1\sigma$ uncertainty.

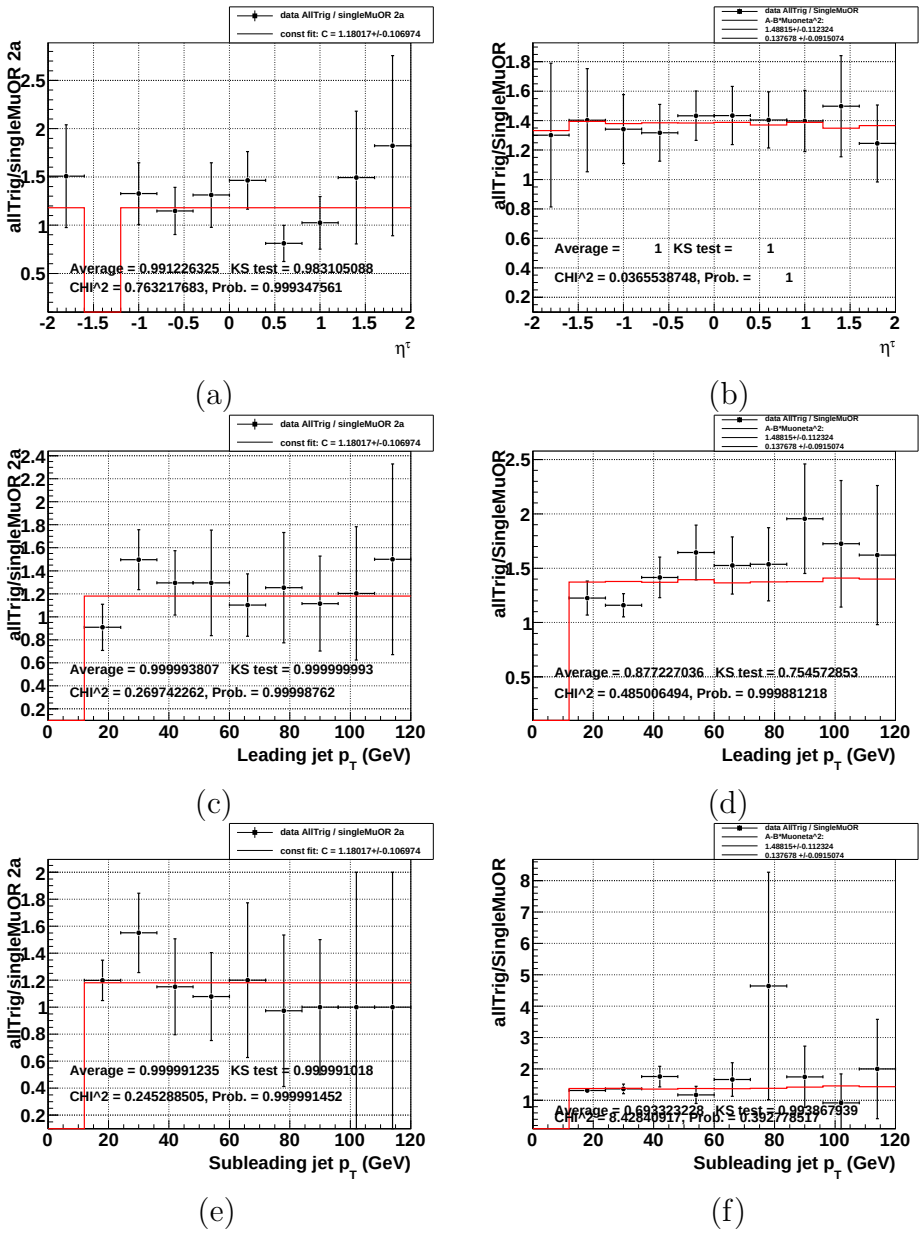


Figure A.6: Ratio of AllTrigger events to SingleMuOR events in the inclusive $\mu\tau$ selection for type 2 τ's : (a) η_τ, 2a; (b) η_τ, 2b; (c) p_T^{jet1}, 2a; (d) p_T^{jet1}, 2b; (e) p_T^{jet2}, 2a; (f) p_T^{jet2}, 2b.

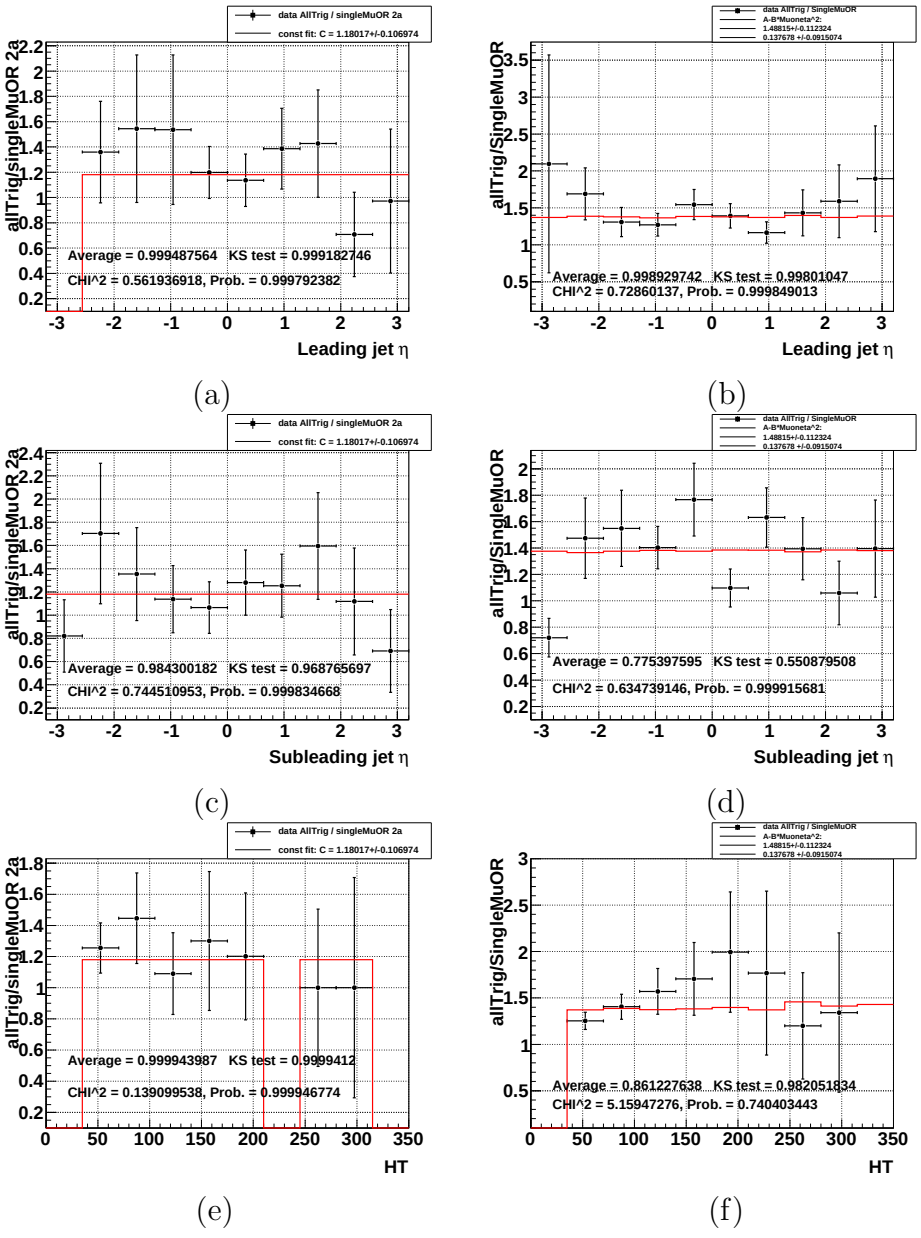


Figure A.7: Ratio of AllTrigger events to SingleMuOR events in the inclusive $\mu\tau$ selection for type 2 τ's : (a) η_{jet1}, 2a; (b) η_{jet1}, 2b; (c) η_{jet2}, 2a; (d) η_{jet2}, 2b; (e) H_T, 2a; (f) H_T, 2b.

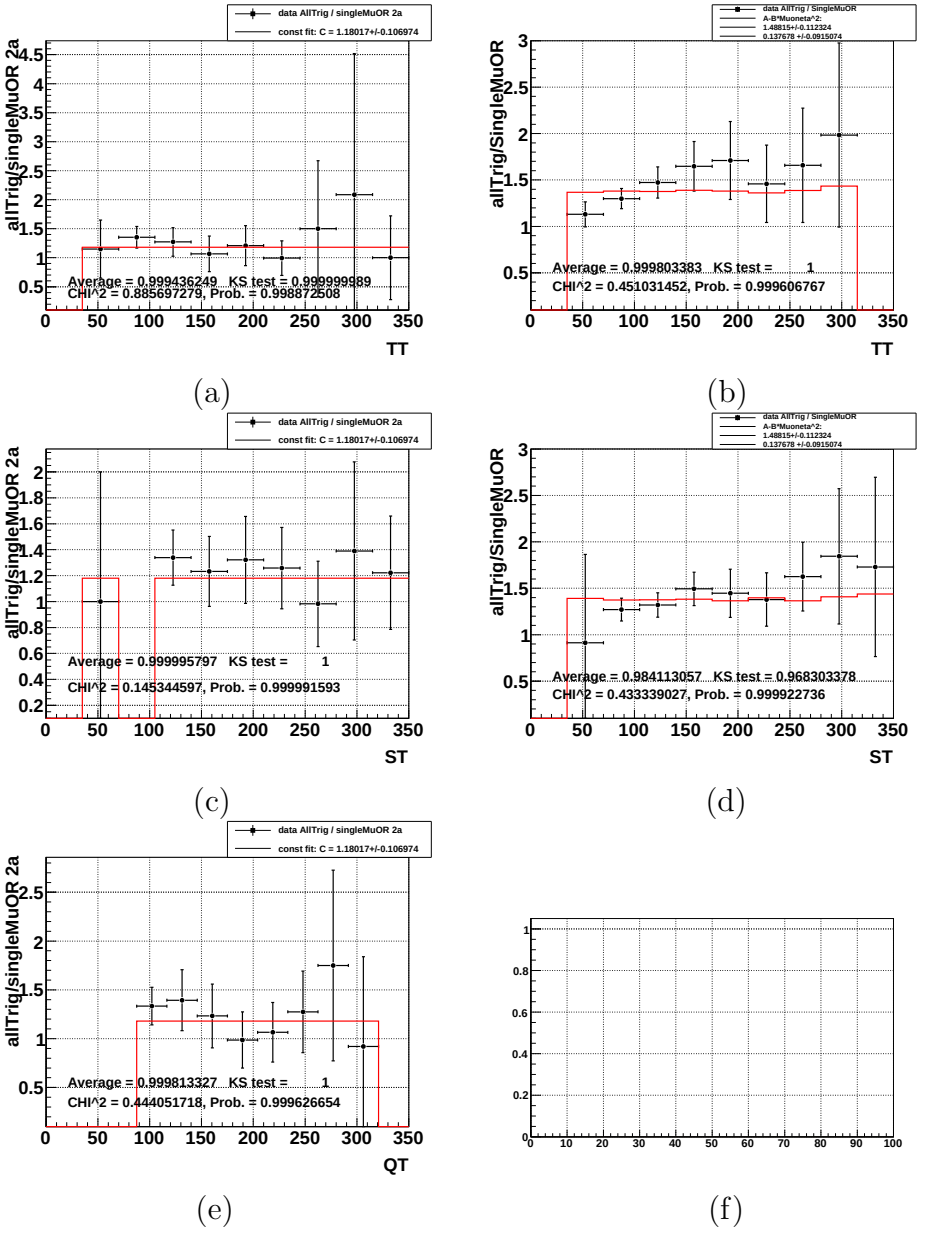


Figure A.8: Ratio of AllTrigger events to SingleMuOR events in the inclusive $\mu\tau$ selection for type 2 τ's : (a) T_T, 2a; (b) T_T, 2b; (c) S_T, 2a; (d) S_T, 2b; (e) Q_T, 2a.

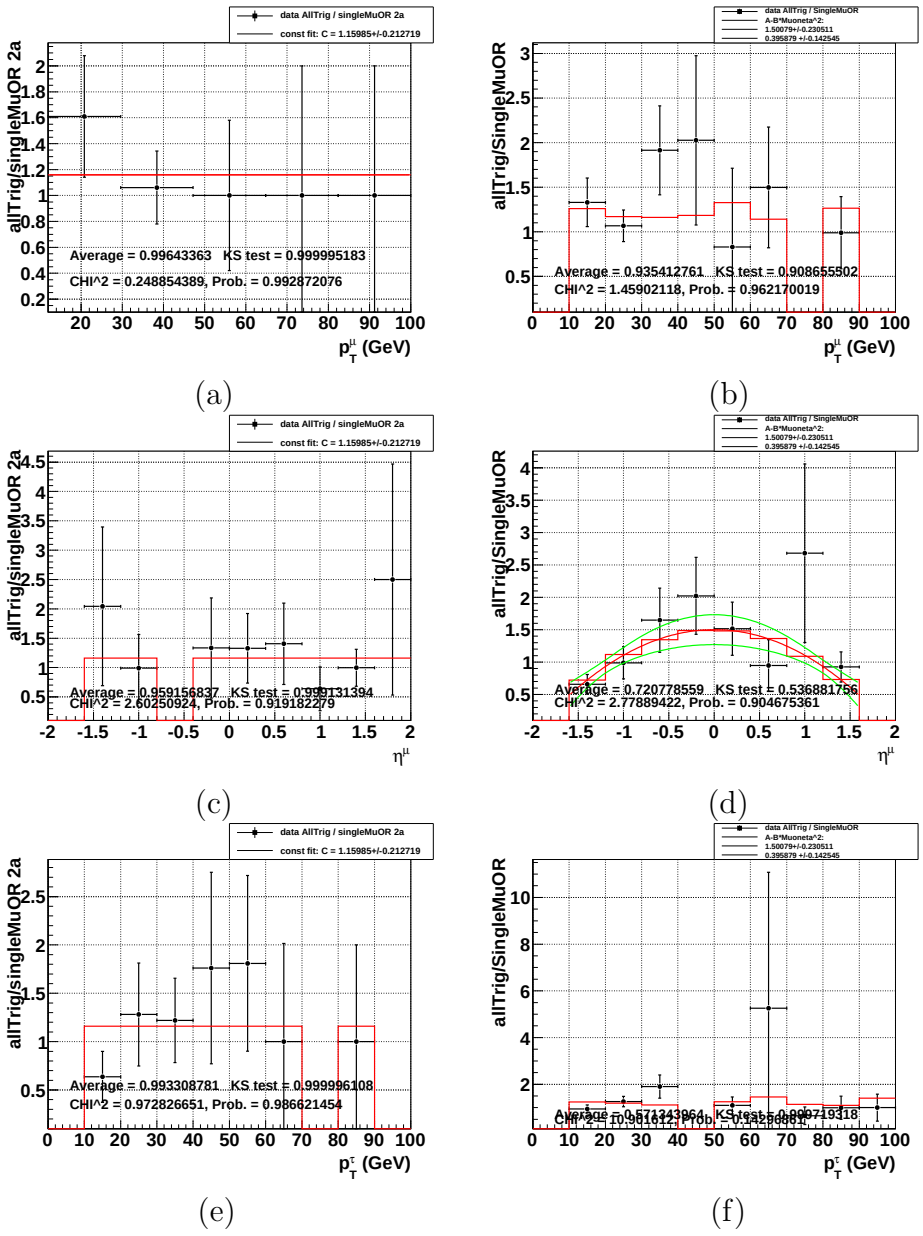


Figure A.9: Ratio of AllTrigger events to SingleMuOR events in the inclusive $\mu\tau$ selection for type 3 τ's : (a) p_T^μ, 2a; (b) p_T^μ, 2b; (c) η_μ, 2a; (d) η_μ, 2b; (e) p_T^τ, 2a; (f) p_T^τ, 2b. The red line in (d) is the parabolic η_μ fit and the green lines indicate the $\pm 1\sigma$ uncertainty.

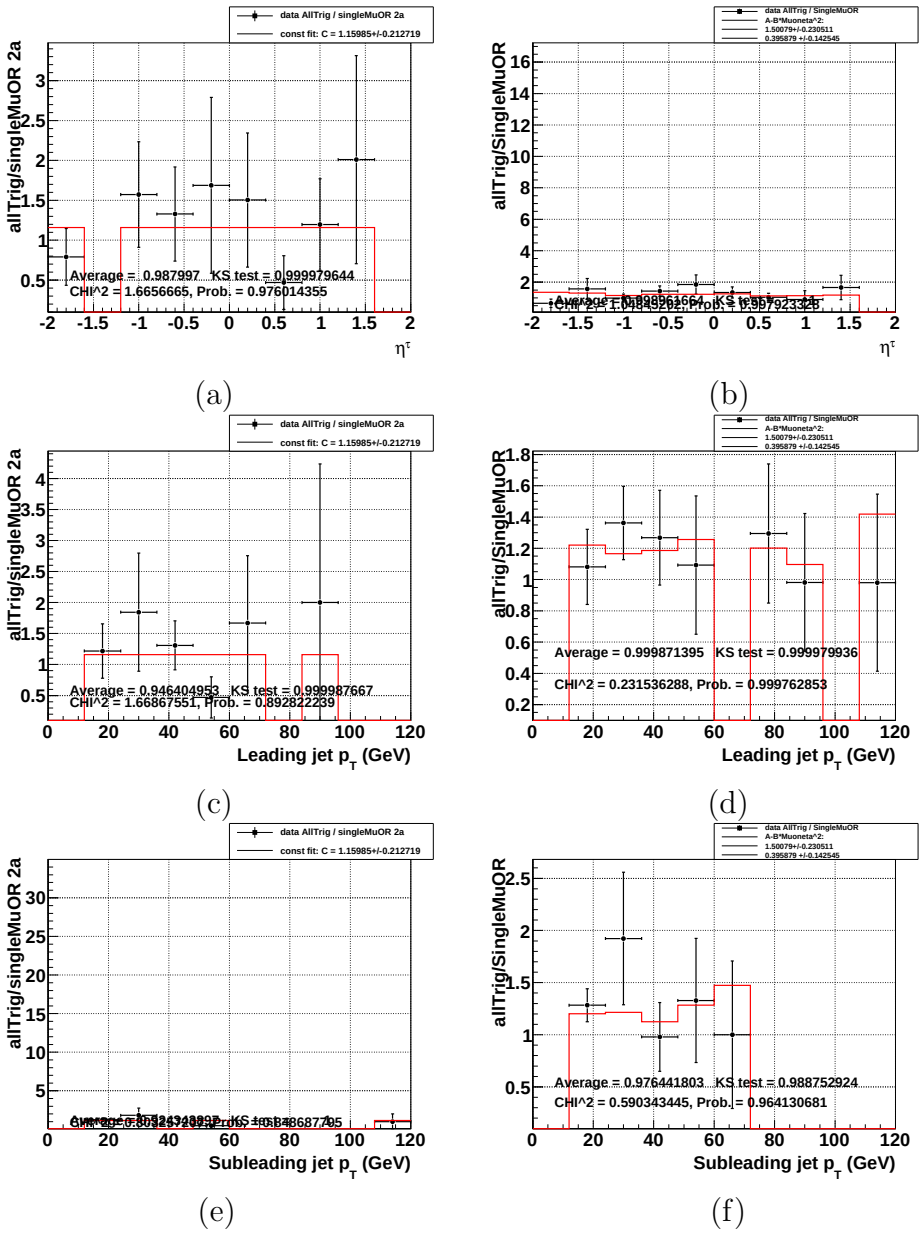


Figure A.10: Ratio of AllTrigger events to SingleMuOR events in the inclusive $\mu\tau$ selection for type 3 τ's : (a) η_τ, 2a; (b) η_τ, 2b; (c) p_T^{jet1}, 2a; (d) p_T^{jet1}, 2b; (e) p_T^{jet2}, 2a; (f) p_T^{jet2}, 2b.

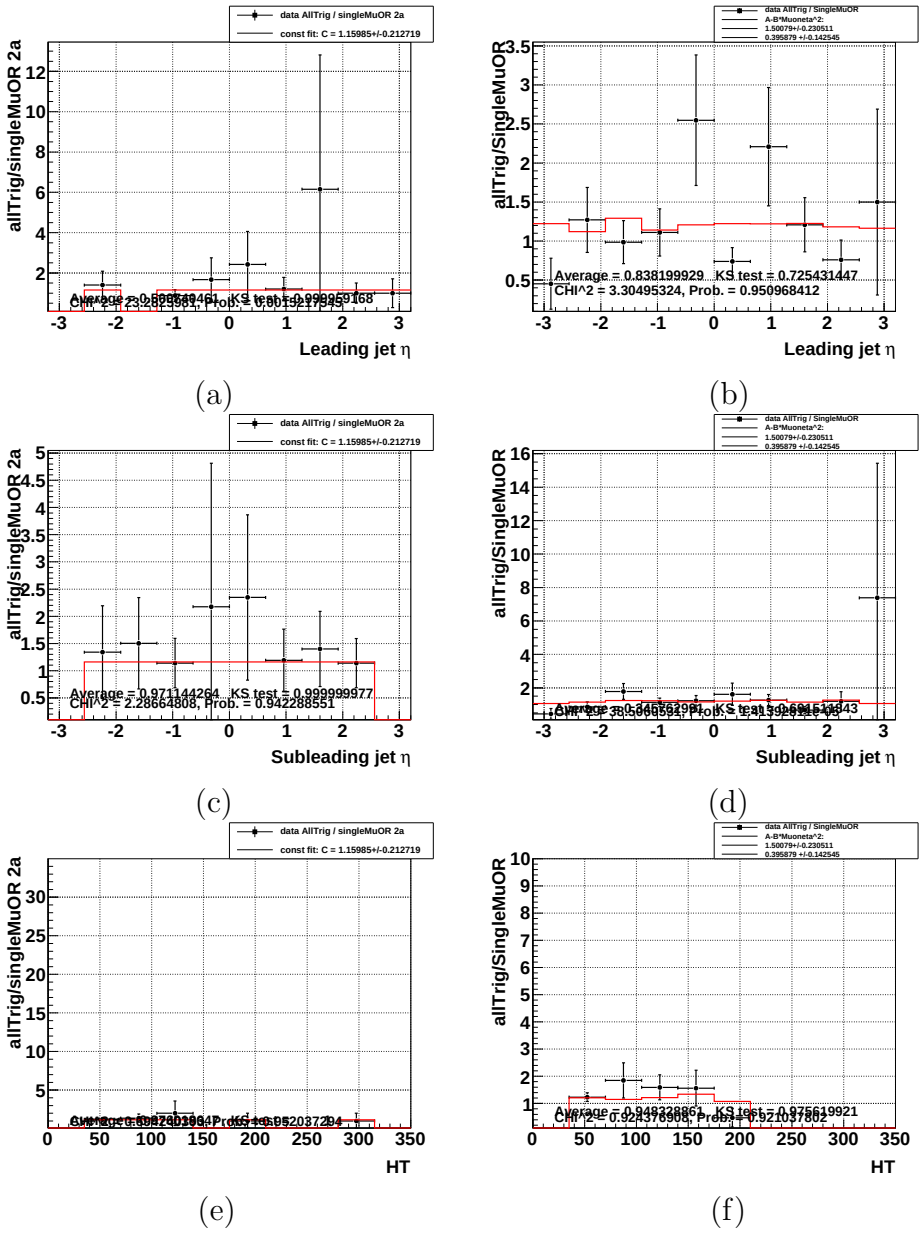


Figure A.11: Ratio of AllTrigger events to SingleMuOR events in the inclusive $\mu\tau$ selection for type 3 τ's : (a) η_{jet1}, 2a; (b) η_{jet1}, 2b; (c) η_{jet2}, 2a; (d) η_{jet2}, 2b; (e) H_T, 2a; (f) H_T, 2b.

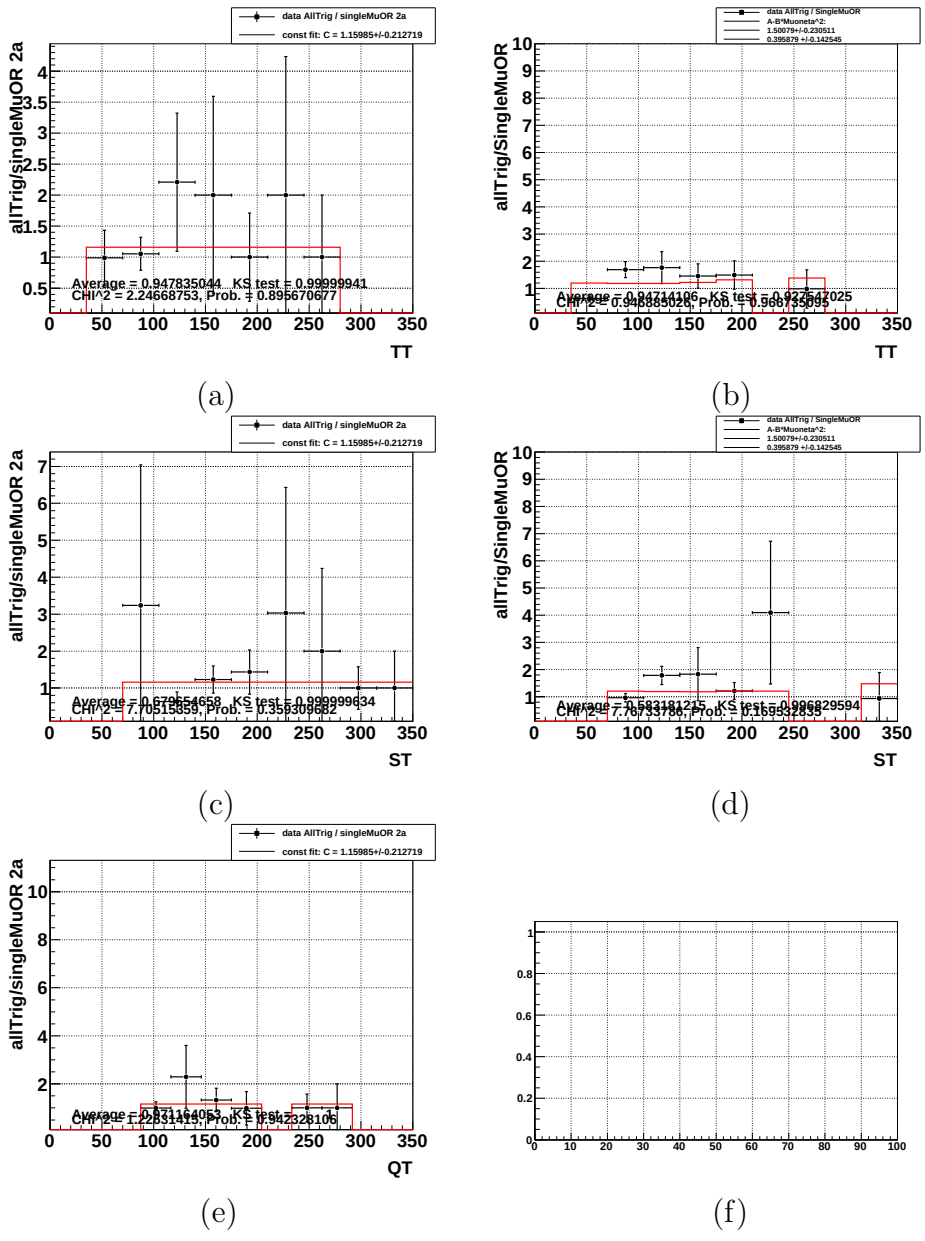


Figure A.12: Ratio of AllTrigger events to SingleMuOR events in the inclusive $\mu\tau$ selection for type 3 τ's : (a) T_T, 2a; (b) T_T, 2b; (c) S_T, 2a; (d) S_T, 2b; (e) Q_T, 2a.

Appendix B

ΔR distribution for $H \to \tau\tau$, $WW \to \mu$ + type 1,2,3 hadrons

The ΔR distribution, or the distance in the angular space between the visible and invisible decay products (see definition in Eq. 4.15), is computed for leptonic and three types of hadronic tau decays to determine the most likely solution in the MMC technique. We generate sufficient $H \to \tau\tau$ and $H \to WW^{(*)}$ MC events, apply the same selections as the data analysis and compute ΔR for each event. The ΔR distribution is then fitted in each 5 GeV bin of the τ or W MC momentum, ranging from 10 to 100 GeV for the leptonic τ/W decay, 12.5 to 102.5 GeV for type 1, 2 hadronic tau decay and 15 to 105 GeV for type 3 hadronic tau decay. The fit function is given in Eq. (4.16). See Sec. 4.9.2 for a full discussion.

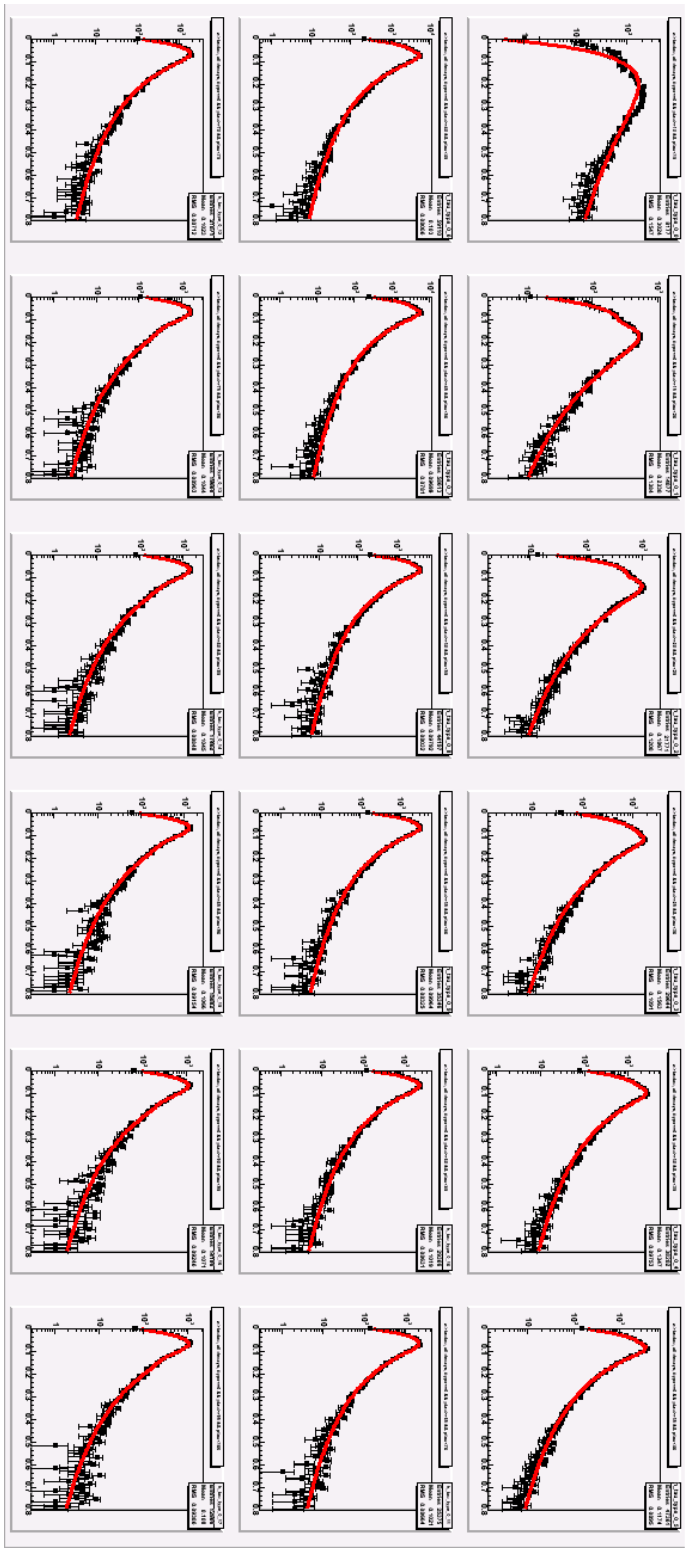


Figure B.1: ΔR distribution for $H \to \tau\tau \to \mu$+leptons, assuming the τ momentum within (10,100) GeV. Each plot represents the ΔR distribution in 5 GeV interval, i.e. (10,15), (15,20), ..., (95,100) from left to right and top to bottom.

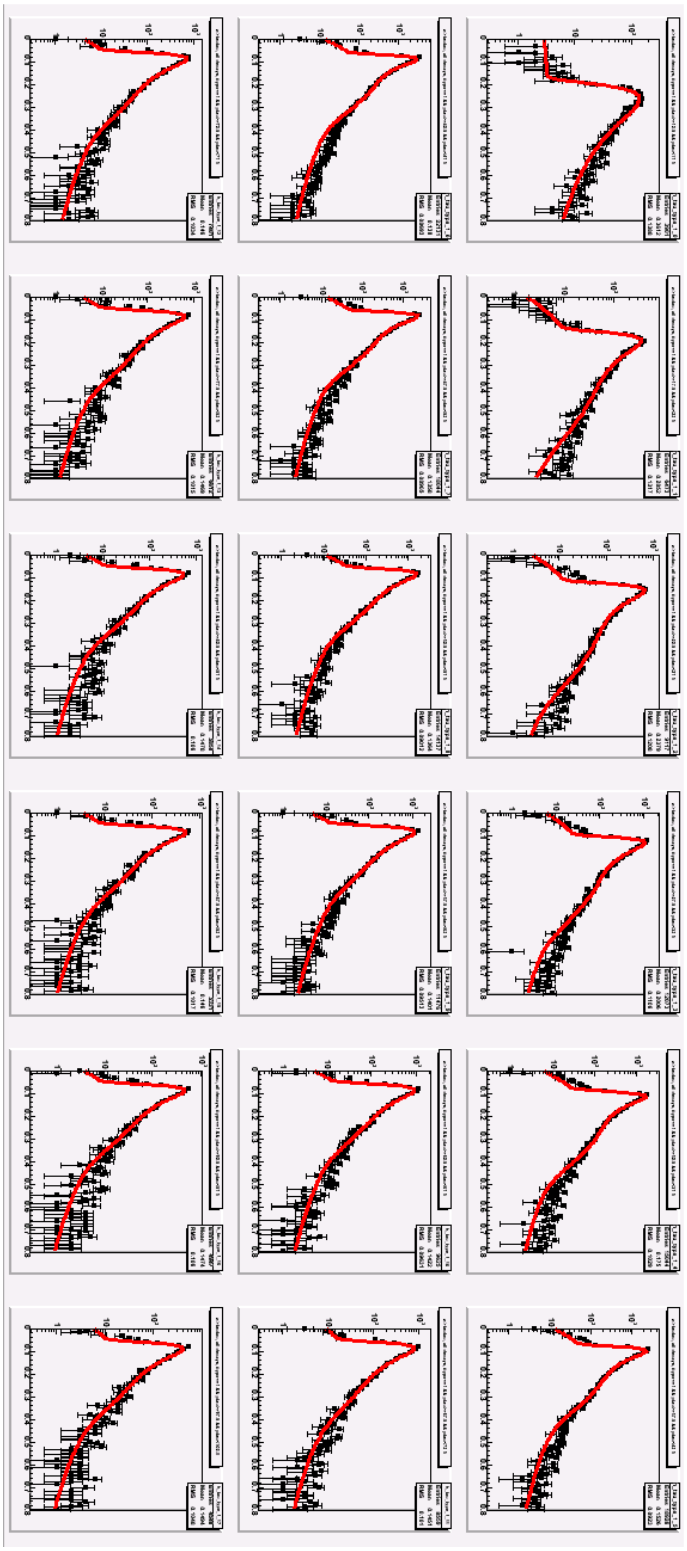


Figure B.2: ΔR distribution for $H \to \tau\tau \to \mu$+type 1 hadrons, assuming the τ momentum within (12.5,100) GeV. Each plot represents the ΔR distribution in 5 GeV interval, i.e. (12.5,17.5), (17.5,22.5), ..., (97.5,102.5) from left to right and top to bottom.

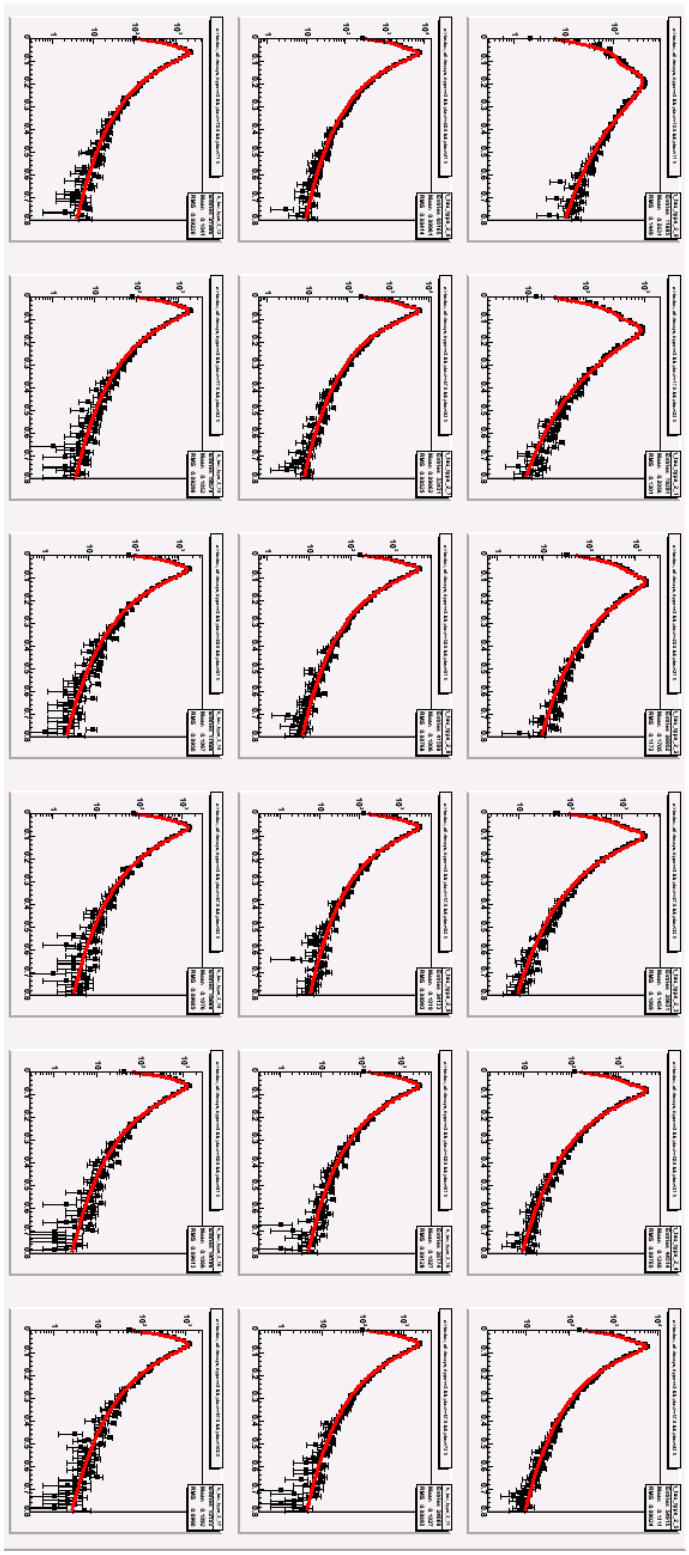


Figure B.3: ΔR distribution for $H \to \tau\tau \to \mu$+type 2 hadrons, assuming the τ momentum within (12.5,100) GeV. Each plot represents the ΔR distribution in 5 GeV interval, i.e. (12.5,17.5), (17.5,22.5), ..., (97.5,102.5) from left to right and top to bottom.

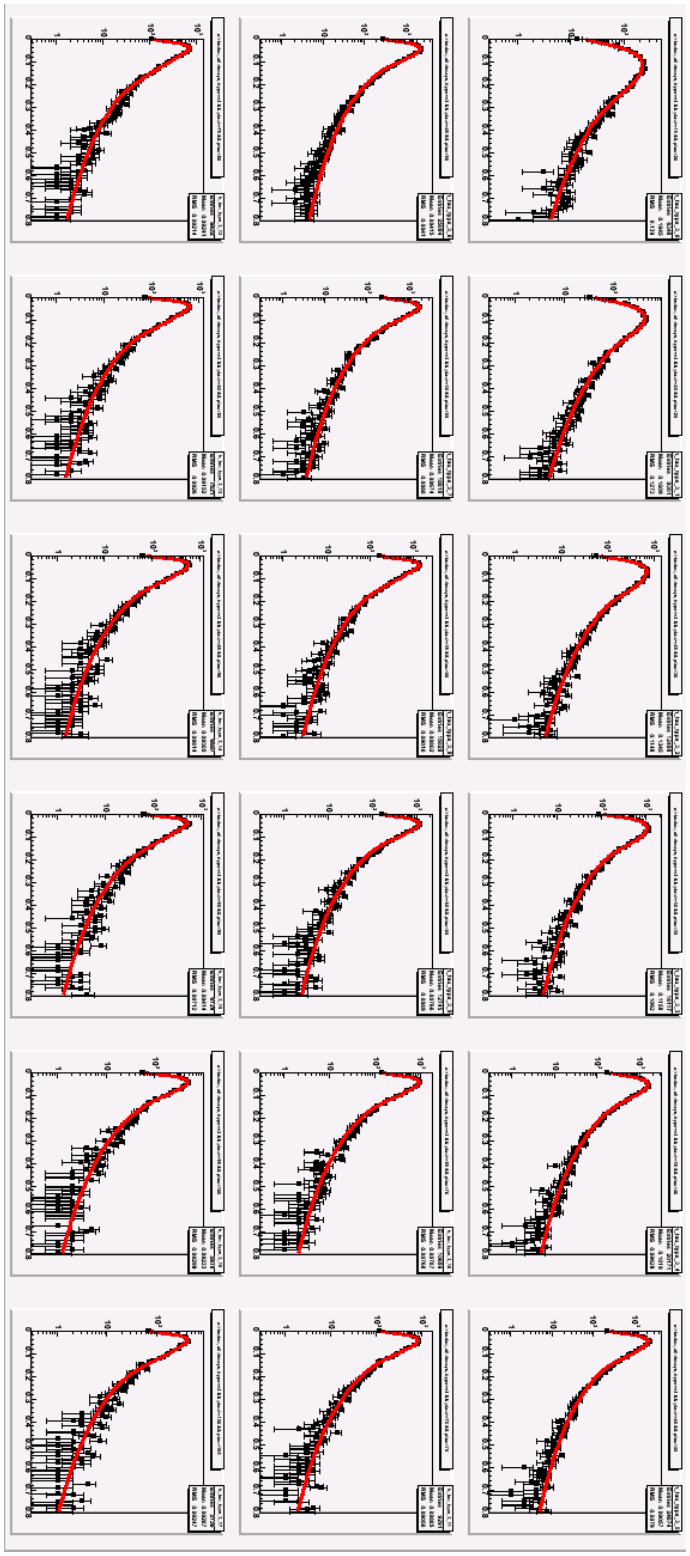


Figure B.4: ΔR distribution for $H \to \tau\tau \to \mu$+type 3 hadrons, assuming the τ momentum within (15,105) GeV. Each plot represents the ΔR distribution in 5 GeV interval, i.e. (15,20), (20,25), ..., (100,105) from left to right and top to bottom.

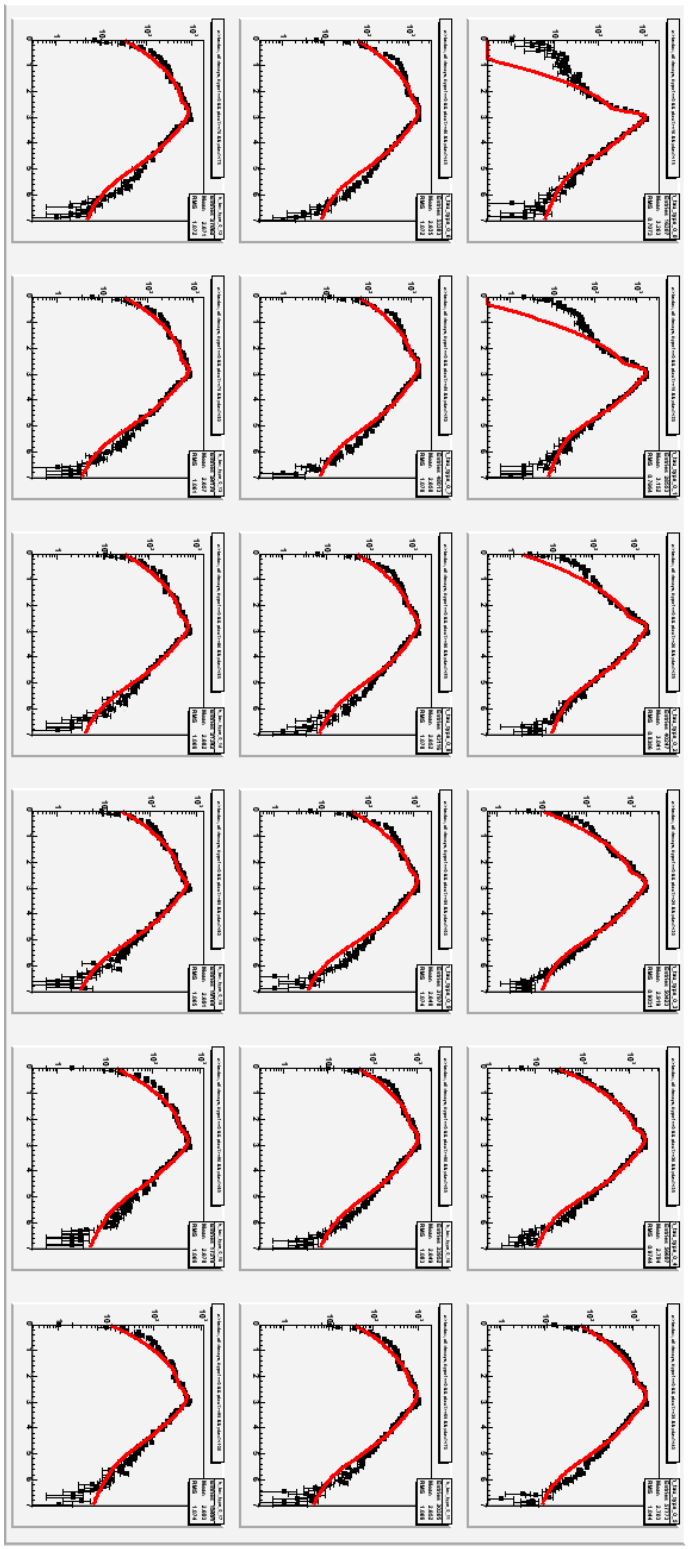


Figure B.5: ΔR distribution for $H \to WW \to \mu$+leptons, assuming the W momentum within (10,100) GeV. Each plot represents the ΔR distribution in 5 GeV interval, i.e. (10,15), (15,20), ..., (95,100) from left to right and top to bottom.

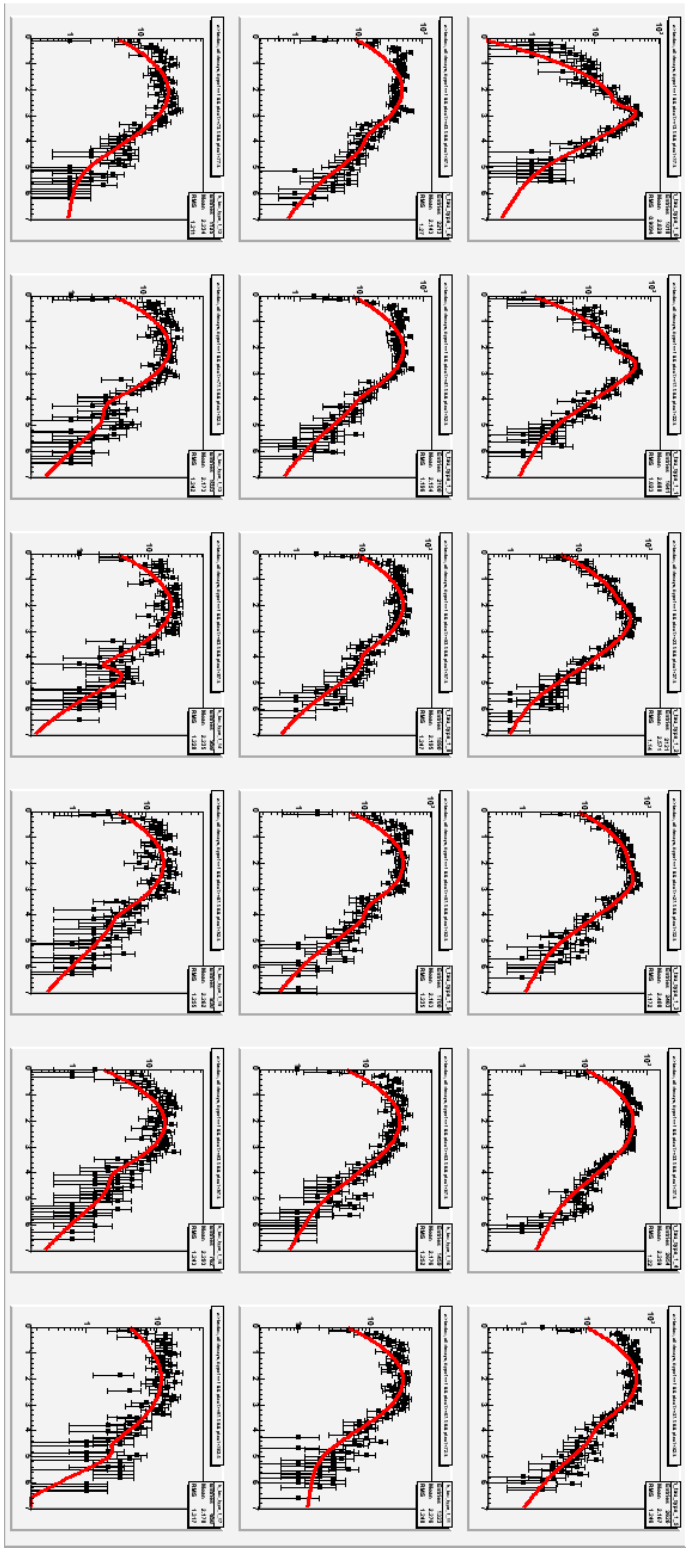


Figure B.6: ΔR distribution for $H \rightarrow WW \rightarrow \mu$+type 1 hadrons, assuming the W momentum within (12.5,100) GeV. Each plot represents the ΔR distribution in 5 GeV interval, i.e. (12.5,17.5), (17.5,22.5), ..., (97.5,102.5) from left to right and top to bottom.

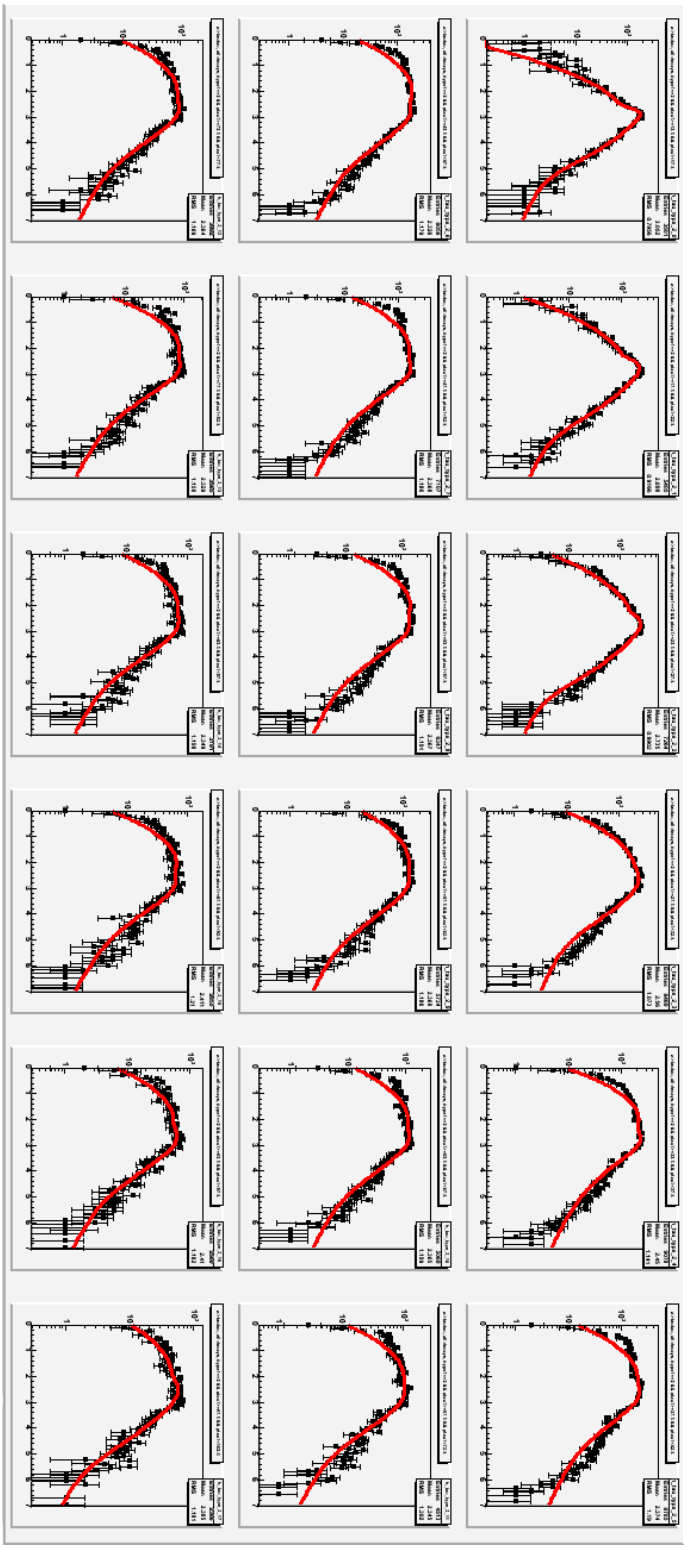


Figure B.7: ΔR distribution for $H \to WW \to \mu$+type 2 hadrons, assuming the W momentum within (12.5,100) GeV. Each plot represents the ΔR distribution in 5 GeV interval, i.e. (12.5,17.5), (17.5,22.5), ..., (97.5,102.5) from left to right and top to bottom.

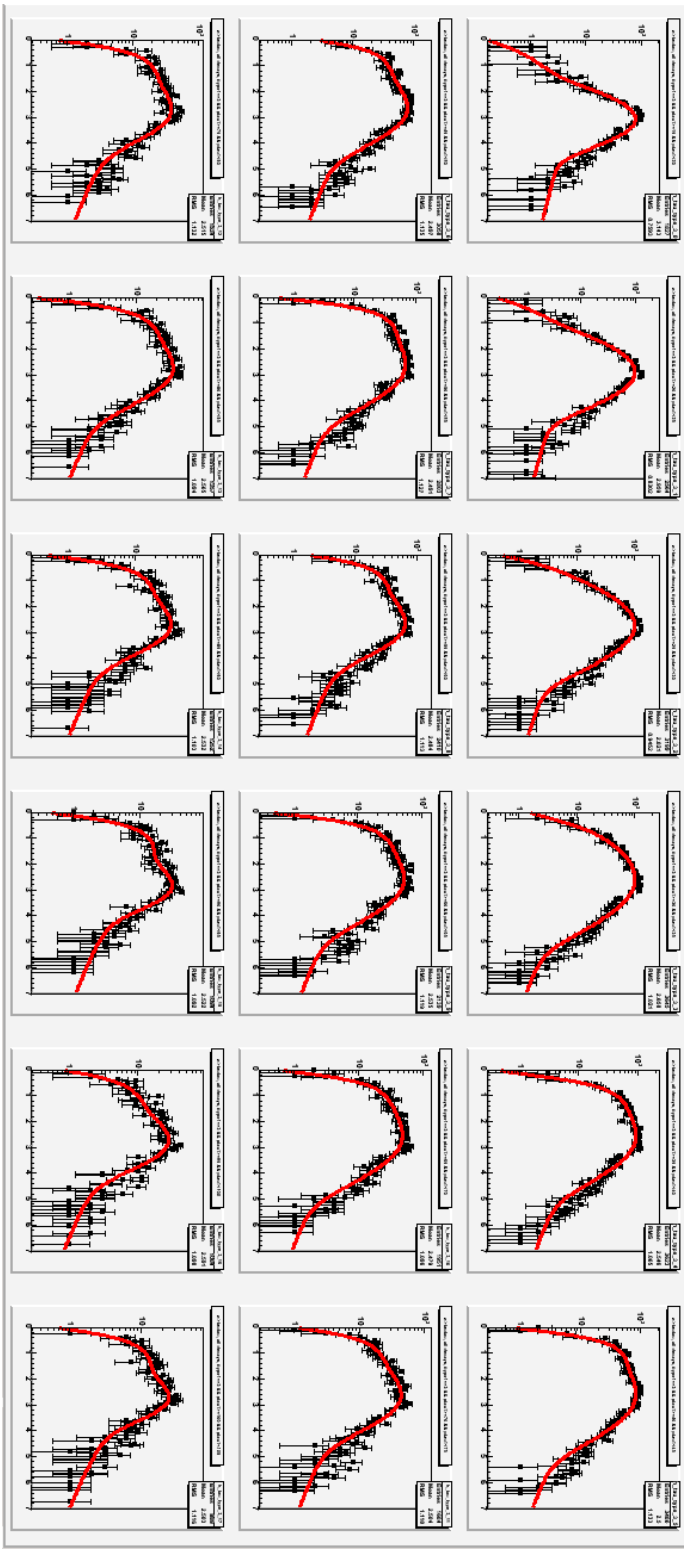


Figure B.8: ΔR distribution for $H \rightarrow WW \rightarrow \mu$+type 3 hadrons, assuming the τ momentum within (15,105) GeV. Each plot represents the ΔR distribution in 5 GeV interval, i.e. (15,20), (20,25), ..., (100,105) from left to right and top to bottom.

Appendix C

$\mu\tau jj$ channel shape variations for jet modeling

We show the jet related systematics (JetID, JER, JES and the vertex confirmation) in this appendix. Adjacent bins with strong statistical fluctuations are identified and merged. We vary each systematic source by -1σ and propagate it through the final BDTs for the modified shapes. The -1σ shape changes are then symmetrized to $\pm 1\sigma$ relative systematic uncertainties. We extract the uncertainties from the T and W subsamples at $m_H = 125$ GeV and applied them for $m_H \leq 150$ GeV in both subsamples, respectively.

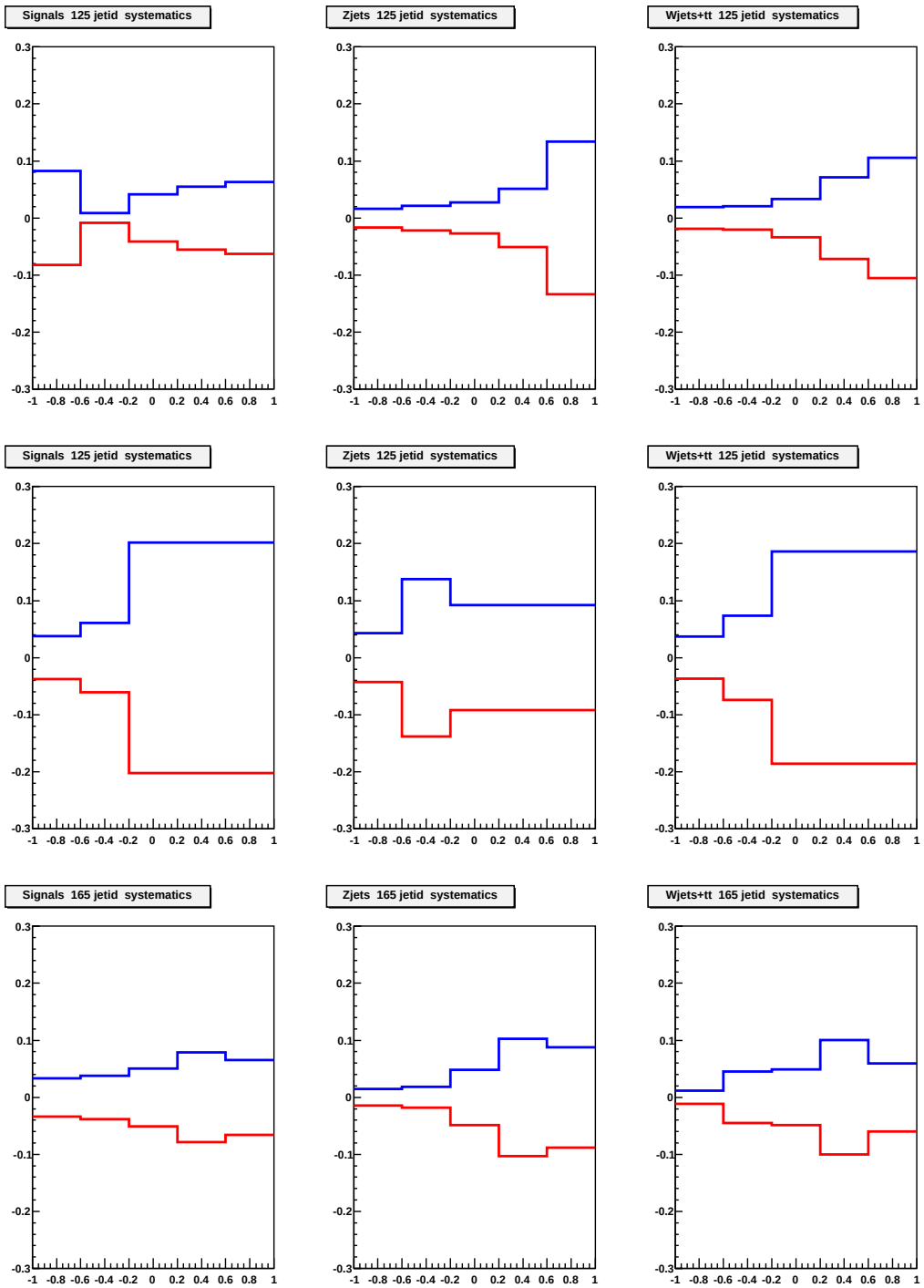


Figure C.1: Normalized $\mu\tau jj$ BDT outputs for nominal and $\pm 1\sigma$ changes of the **jet ID and reconstruction efficiencies**, for the T subsample at 125 GeV (top row), the W subsample at 125 GeV (middle row). The bottom row should be ignored (the high mass W subsample). The signals, W+jets, $t\bar{t}$ and Z+jets backgrounds are shown in the left, middle and right rows respectively.

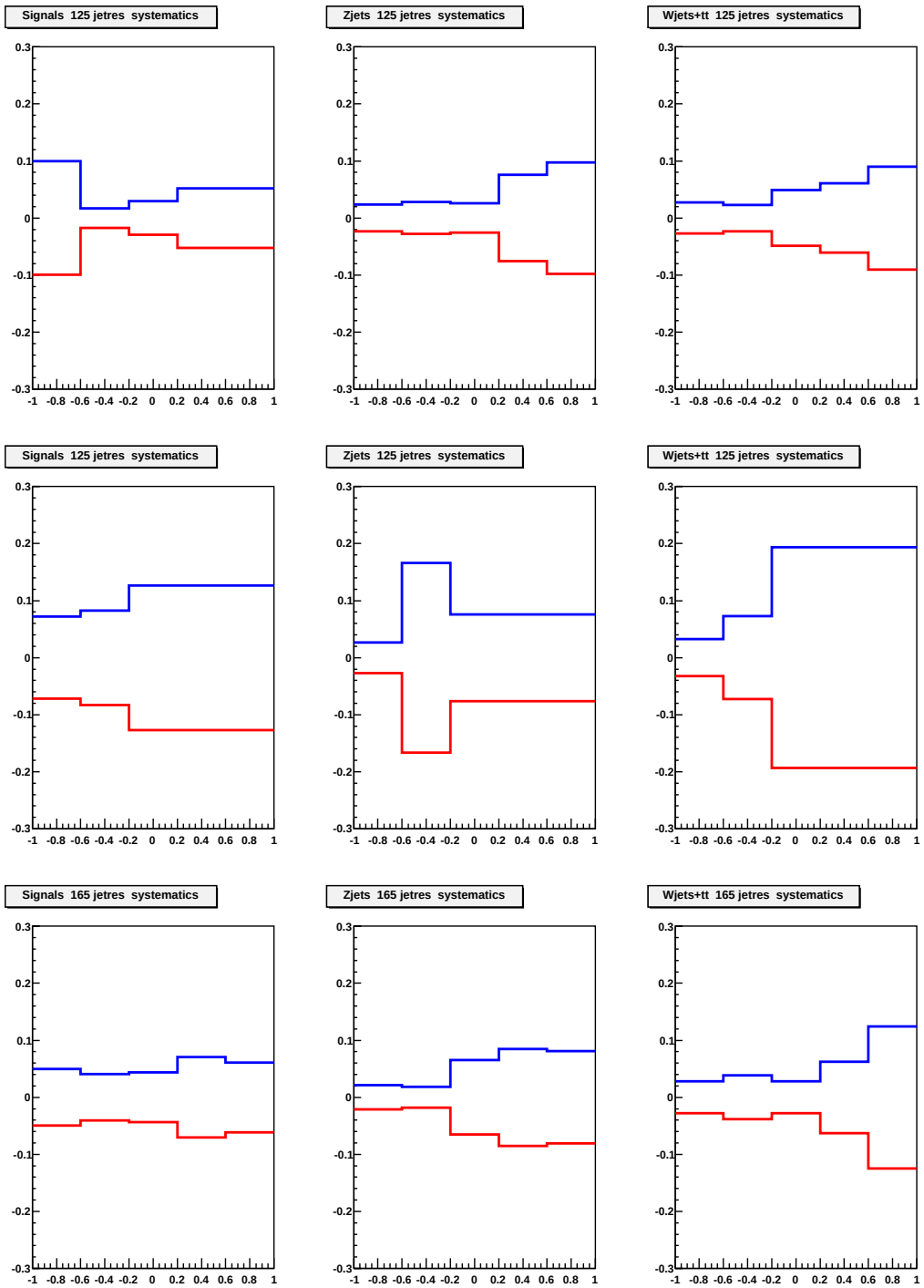


Figure C.2: Normalized $\mu\tau jj$ BDT outputs for nominal and $\pm 1\sigma$ changes of the **jet energy resolution**, for the T subsample at 125 GeV (top row), the W subsample at 125 GeV (middle row). The bottom row should be ignored (the high mass W subsample). The signals, W+jets, $t\bar{t}$ and Z+jets backgrounds are shown in the left, middle and right rows respectively.

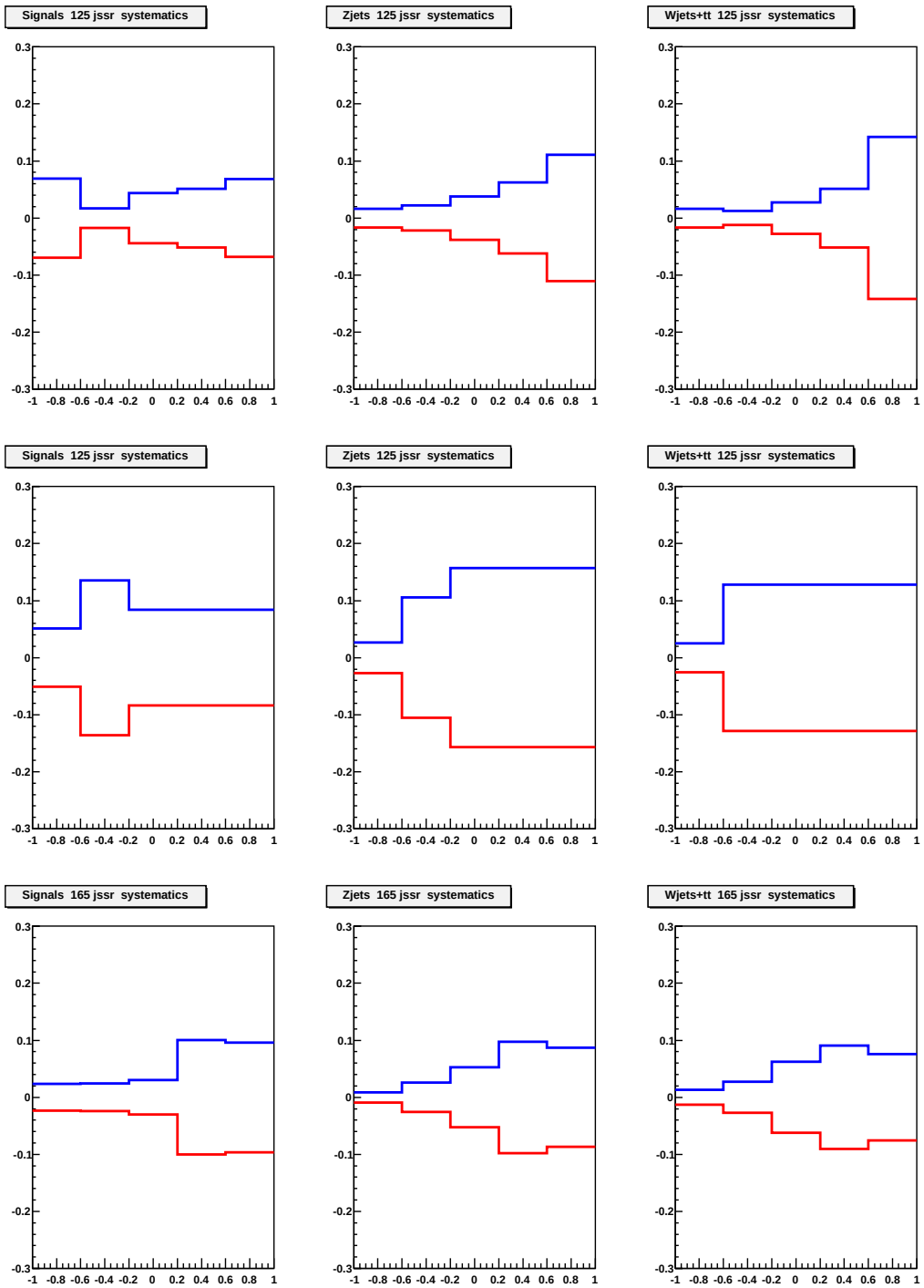


Figure C.3: Normalized $\mu\tau jj$ BDT outputs for nominal and $\pm 1\sigma$ changes of the **jet energy scale**, for the T subsample at 125 GeV (top row), the W subsample at 125 GeV (middle row). The bottom row should be ignored (the high mass W subsample). The signals, W+jets, $t\bar{t}$ and Z+jets backgrounds are shown in the left, middle and right rows respectively.

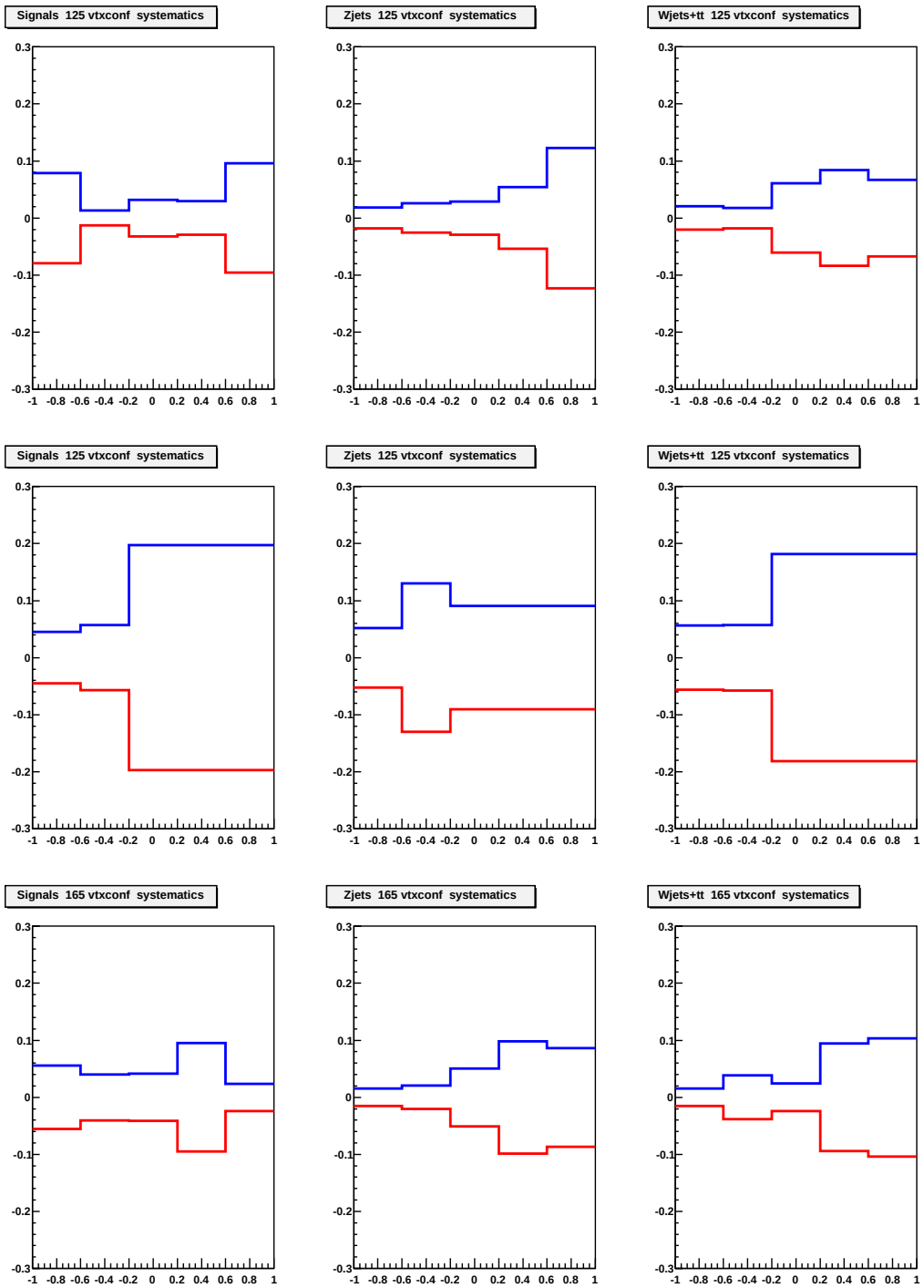


Figure C.4: Normalized $\mu\tau jj$ BDT outputs for nominal and $\pm 1\sigma$ changes of the **vertex confirmation**, for the T subsample at 125 GeV (top row), the W subsample at 125 GeV (middle row). The bottom row should be ignored (the high mass W subsample). The signals, W+jets, $t\bar{t}$ and Z+jets backgrounds are shown in the left, middle and right rows respectively.

Appendix D

Shape variations for $\mu\tau jj$ multijet background modelling

We use the "good τ bad muon" sample as discussed in Sec. 4.6 to estimate the systematic uncertainty of the MJ background. We compare the final discriminant shapes for the signal and control sample MJ backgrounds and take the difference divided by the control sample MJ as the MJ systematic uncertainty. This is a fractional uncertainty ranging between -1 and 1. The systematic uncertainties for T and W subsamples are determined separately at $M_H = 125$ GeV and applied in the whole mass range. We symmetrize the $-\sigma$ difference to $\pm\sigma$ as the jet related uncertainties in Appendix C.

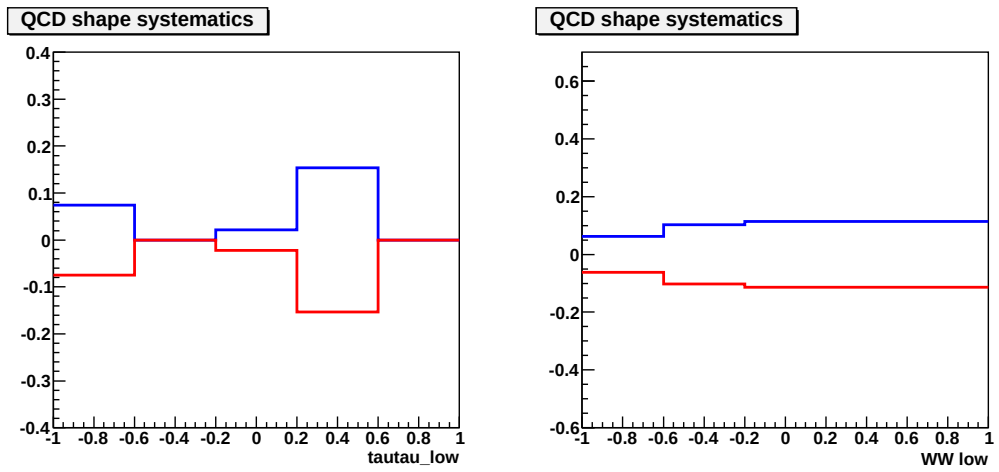


Figure D.1: Fractional uncertainties of $\mu\tau jj$ multijet backgrounds taken from the nominal same-sign estimated signal sample and the MJ enriched sample for the T subsample, $m_H = 125$ (left) and the W subsample, $m_H = 125$ (right).