Search for the Standard Model Higgs Boson in $H \rightarrow \tau^+\tau^-$ Final State in Proton-Antiproton Collisions at CDF

Yuki SAKURAI
Master Program in Physics
February/7/2012

Major in Pure and Applied Physics
Graduate School of
Advanced Science and Engineering
at the
Waseda University

Abstract

We present a search for the Standard Model Higgs Boson in $\tau^+\tau^-$ final state in proton and anti proton collisions at the CDF, using integrated luminosity $9.6fb^{-1}$. Signal processes in this search are four Higgs production processes: W/Z Boson Associated Production ($q\bar{q}' \rightarrow WH$, $q\bar{q} \rightarrow ZH$) , Vector Boson Fusion($qq' \rightarrow qHq'$) and Gluon Fusion($gg \rightarrow H$), followed by $H \rightarrow \tau\tau$.

We select two types of events as candidates by requiring one isolated electron/muon and one hadronic τ, and one isolated electron and one isolated muon from H$\rightarrow$ $\tau_{e/\mu}\tau_h$ and H$\rightarrow$ $\tau_e\tau_\mu$, respectively. The Multivariate analysis which discriminates Higgs signals from backgrounds using Support Vector Machine is performed in order to maximize search sensitivity. We can not see clear excess of the Higgs signal process, therefore we extract a 95% confidence level upper limit on the Standard Model Higgs Boson cross section($XH \rightarrow \tau\tau + X$). The expected and observed limit are evaluated from $M_H = 100GeV/c^2$ to $M_H = 150GeV/c^2$ in $5GeV/c^2$ step. The observed(expected) limit on assumption of $M_H = 115GeV/c^2$ is 8.3(10.6).

Acknowledgments

I would like to thank all people who supported and cooperated with me for my Master course thesis.

First of all, I would like to express my gratitude to my boss, Associate Professor.Kohei Yorita. He has taught me from fundamental to philosophy for physics, moreover lifestyle of physical researcher. He gave me many chance and place to be active including international collaboration, and also gave me many constructive advice for my study. I have became fond to study physics and decide to become physicist thanks to him. I will greatly appreciate his further guidance and encouragement during doctor course.

I would also like to thank staffs in Kohei Yorita laboratory, Dr.Koji Ebina, Dr.Naoki Kimura, Dr.Junji Naganoma. Dr.Koji Ebina often gave me technical advice and shared good times in Waseda university and the Fermilab. Dr.Naoki Kimura and Dr.Junji Naganoma also gave me grate advice for my study from their experience of CDF analysis. I would like to continue to study physics with them.

I am deeply grateful to the same age members of the Kohei Yorita Laboratory, Hayato Okamoto, Hideyoshi Tiba, Shohei Kamatsuka, Takashi Mitani, Yusi Nagasaka, Yujiro Funakoshi. I think the times spending and researching physics with them were irreplaceable for me. Also, I would like thank to other member of the laboratory, first master course students and bachelor students.

I would like to thank the CDF collaboration members. At first, I was afraid to spend and study at Fermilab, but I had very good time because the CDF members are very kind to me during my staying. I learnt many things from them. The Higgs Discovery Group members, especially Eric B. James, Craig Group who are convener of this group are very encouraged me and gave many constructive advice for my study.

And also I would like to thank the CDF Japan Group members, Professor.Shinihong Kim, Professor Fumihiko Ukegawa, Dr.Hideki Miyake, Dr.Masakazu Kurata, Dr.Yoshikazu Nagai. They supported and gave useful advice to me, and shared good times at Fermilab.

Finally, I would like to express my gratitude to my family, father, mother and sister. They are always supported me for my life and studying. I can not thank you too much.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Standard Model of Particle Physics

Our current best understanding on particle physics is given by a theory called Standard Model. The Standard Model is a theory integrating quantum chromodynamics(QCD) [1], Glashow Weinberg Salam theory. [2], [3], [4] The Standard Model is expressed by following gauge symmetry group:

$$SU(3) \otimes SU(2)_L \otimes U(1)_Y, \tag{1.1}$$

where $SU(3)$ expressed QCD, which describes strong interactions, and $SU(2)_L \otimes U(1)_Y$ expressed Glashow Salam Weinberg theory, which describes and unifies electromagnetic and weak interactions.

1.1.1 Elementary Particles

Elementary particles of the Standard Model theory consists of a two types of particles:fermions, which consists of normal matter, and bosons which mediate the forth between fermions.

fermion

Elementary particles of spin $\frac{1}{2}$ are categorized into fermions. Furthermore, fermions are categorized into quarks and leptons. The quarks have six types(flavors): up(u), down(d), charm(c), strange(s), bottom(b), top(t). The leptons also have six types: electron(e), muon(μ), tau(τ) and neutrinos respectively corresponding to them: ν_e, ν_μ, ν_τ. These fermions are classified by three generation and characterized with weak isospin (I) and weak hyper-charge (Y) by the following relation:

$$Q = I_3 + \frac{1}{2}Y, \tag{1.2}$$

where Q is the electric charge. The quarks are also characterized by color charge, which is three type: red, green and blue. In nature, white(colorless) composite particles can only exit: mesons which are composite particles of two quarks and baryons which are composite particles of three quarks. The properties of fermions are summarized in Table 1.1 and Table 1.2.

Bosons

Elementary particles of spin 1 are categorized into bosons. The bosons mediate the interactions of between fermions, and they are four types corresponding to the kind of interactions. The

Fermions (spin=$\frac{1}{2}$)				
Quarks				
Particle	Generation	Charge	Isospin	Mass[GeV/c^2]
up(u)	1^{st}	$+\frac{2}{3}$	$+\frac{1}{2}$	$1.7 - 3.1 \times 10^{-3}$
down(d)		$-\frac{1}{3}$	$-\frac{1}{2}$	$4.1 - 5.7 \times 10^{-3}$
charm(c)	2^{nd}	$+\frac{2}{3}$	$+\frac{1}{2}$	$0.8 - 1.3 \times 10^{-1}$
strange(s)		$-\frac{1}{3}$	$-\frac{1}{2}$	$1.2 - 1.3$
top(t)	3^{rd}	$+\frac{2}{3}$	$+\frac{1}{2}$	~ 172.5
bottom(b)		$-\frac{1}{3}$	$-\frac{1}{2}$	~ 4.2

Table 1.1: The properties of quarks.

Fermions (spin=$\frac{1}{2}$)				
Leptons				
Particle	Generation	Charge	Isospin	Mass[MeV/c^2]
electron(e^-)	1^{st}	-1	$+\frac{1}{2}$	0.511
electron neutrino(ν_e)		0	$-\frac{1}{2}$	$\leq 2 \times 10^{-6}$
muon(μ^-)	2^{nd}	-1	$+\frac{1}{2}$	105.6
muon neutrino(ν_μ)		0	$-\frac{1}{2}$	$\leq 2 \times 10^{-6}$
tau(τ^-)	3^{rd}	-1	$+\frac{1}{2}$	1776.8
tau neutrino(ν_τ)		0	$-\frac{1}{2}$	$\leq 2 \times 10^{-6}$

Table 1.2: The properties of leptons.

photons(γ), W bosons($W^{\pm}$) and Z bosons(Z^0) mediate electroweak interaction of quarks and leptons and the gluons mediate strong interaction of quarks. The properties of bosons are shown in Table 1.3 In the gauge theory which is in the Standard Model, the bosons are required to be mass-less for maintaining gauge symmetry. But we experimentally know that W and Z bosons have mass. In the Standard Model, this problem is solved by the Higgs mechanism which is cause of the electroweak symmetry breaking , which is discussed in Section 1.2.

Bosons (spin=1)			
Particle	Interaction	Charge	Mass[GeV/c^2]
photon(γ)	Electromagnetic	0	0
W bosons($W^{\pm}$)	Weak	±1	80.4
Z boson(Z^0)	Weak	0	91.2
gluon(g)	Strong	0	0

Table 1.3: The properties of bosons

1.2 The Standard Model Higgs Boson

In the $SU(3) \otimes SU(2)_L \otimes U(1)_Y$ gauge symmetry, particles can not have these masses, but we know the W and Z bosons and fermions have these masses, experimentally. The Standard Model includes the Higgs mechanism to solve this inconsistency.

1.2.1 The electroweak theory

The fermions follow $SU(2)_L \otimes U(1)_Y$ gauge symmetry in the electroweak interaction. The Lagrangian of the electroweak interaction is defined as

$$\mathcal{L} = \mathcal{L}_{bosons} + \mathcal{L}_{fermions}. \tag{1.3}$$

The Lagrangian of bosons term, $\mathcal{L}_{bosons}$ is

$$\mathcal{L}_{bosons} = -\frac{1}{4} W_i^{\mu\nu} W^i_{\mu\nu} + \frac{1}{4} B^{\mu\nu} B_{\mu\nu}, \tag{1.4}$$

where B_μ is the filed corresponding to $U(1)_Y$ and W^i is the filed corresponding to the $SU(2)_L$ group. $B^{\mu\nu}$ is described as

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.5}$$

and $W^{\mu\nu}$ is described as

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon_{ijk} W^j_\mu W^k_\nu. \tag{1.6}$$

The Lagrangian of fermions term, $\mathcal{L}_{fermions}$ is

$$\mathcal{L}_{fermions} = \bar{\psi_L} i\gamma^\mu D_{\mu L} \psi_L + \bar{\psi_R} i\gamma^\mu D_{\mu R} \psi_R, \tag{1.7}$$

where ψ_L and ψ_R is the wave function of the right-handed and left-handed fermions, and the co-variant differentiation terms($D_{\mu L}$,$D_{\mu R}$) is expressed by

$$D_{\mu L} = \partial_\mu - ig_W T_i W_\mu^i + \frac{1}{2} g_B Y B_\mu \tag{1.8}$$

$$D_{\mu R} = \partial_\mu + + \frac{1}{2} g_B Y B_\mu, \tag{1.9}$$

where Y is the generator corresponding to $U(1)_Y$ and T_i is the generator corresponding to $SU(2)_L$ and defined as

$$T_i = \frac{1}{2}\sigma_i, \tag{1.10}$$

where is σ_i is 2×2 Pauli matrix.

This Lagrangian is invariant under local $SU(2)_L \times U(1)_Y$ transformation. Since the gauge fields B and W^3 combine with same current, they are mixed by a certain ratio. The mixing angle (θ_W) is called the Weinberg angle.

In the interaction term in the electroweak Lagrangian, the gauge bosons are expressed by

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2) \tag{1.11}$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \tag{1.12}$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu, \tag{1.13}$$

where $W_\mu^\pm$,Z_μ and A_μ represent the W boson, the Z boson and the photon, respectively. However, this description can not give the masses of the fermions and the bosons. Therefore, the Higgs mechanism [5], [6] is included in the Standard Model.

1.2.2 The Higgs mechanism

The Higgs mechanism is cause of the electroweak symmetry breaking, and the W and Z bosons and fermions have these masses by the result of it. In order to give the mass of the bosons and the fermions, the scalar field is added into the Lagrangian of the electroweak interaction, which is called the Higgs field. The gauge boson mediating the Higgs field is called the Higgs boson.

At first, the Higgs field(Φ) is defined as

$$\Phi = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.14}$$

and the Lagrangian corresponding to the Higgs field is defined as

$$\mathcal{L}_\phi = \frac{1}{2}(D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \tag{1.15}$$

where $V(\phi)$ is the potential term and defined as

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2, \tag{1.16}$$

where λ is self coupling constant.

If $\mu^2 \leq 0$, $V(\phi)$ is symmetry in $(\phi_3, \phi_4, V(\phi))$ space and takes minimum value 0 at (ϕ_3, ϕ_4) = (0,0). If $\mu^2 \geq 0$, $V(\phi)$ is non-symmetry and takes non-zero value as the minimum potential value. In this case, when a certain vacuum state taking the minimum potential, the electroweak symmetry is spontaneously broken.

If the vacuum state is selected as

$$\phi_1 = 0, \phi_2 = 0, \phi_4 = 0, \tag{1.17}$$

the Higgs field is described as

$$\Phi = \begin{pmatrix} 0 \\ \nu + h(x) \end{pmatrix} \tag{1.18}$$

The Lagrangian of the Higgs field,$\mathcal{L}_\phi$ becomes

$$\begin{aligned} \mathcal{L}_\phi &= \frac{1}{2}(\partial_\mu h)^2 + \frac{1}{4}g_W^2 W_\mu W^\mu(\nu + h)^2 \\ &+ \frac{1}{8}(\nu\sqrt{g_W^2 + g_B^2})Z_\mu Z^\mu(\nu + h)^2 - V(\frac{1}{2}(\nu + h)^2) \end{aligned}$$

Since the first term is identified as the mass term from the Higgs field, its mass is represented by

$$M_H = \sqrt{2\lambda}\nu. \tag{1.19}$$

The second and third term of the Lagrangian are identified as the mass terms of the W and Z boson. Therefore, the Higgs mechanism gives the masses of the W and Z boson, but not the photon. Their masses are represented by

$$M_W = \frac{1}{2}g_W\nu \tag{1.20}$$

$$M_Z = \frac{1}{2}\nu\sqrt{g_W^2 + g_B^2}. \tag{1.21}$$

Since the value of these masses are not predicted by the Standard Model, these values should be measured by experiments.

In the Standard Model Theory, the Higgs mechanism is also give the masses of fermions. The interaction of between the Higgs field and the fermions are represented by the Yukawa interaction, which is Lorentz invariant. The Lagrangian of this interaction is described as

$$\mathcal{L}_{YUKAWA} = -f(\bar{\psi}\psi)\phi + h.c.,$$

where f is the coupling constant of between the Higgs boson and the fermions, and h.c. is Hermite conjugate. When introducing the electroweak symmetry breaking, the Lagrangian becomes

$$\mathcal{L}_{YUKAWA} = -M_f\bar{\psi}\psi - \frac{M_f}{\nu}\bar{\psi}\psi\phi.$$

Therefore, the fermions are given these masses by the Higgs mechanism. The masses of fermions are represented by

$$M_f = \frac{f\nu}{\sqrt{2}}$$

However, the f of each fermion which is called Yukawa coupling constant is not predicted by the Standard Model. These values are also measured by experiments if the Higgs mechanism describes nature correctly.

1.3 The Standard Model Higgs Search

In this paper, we describe about search for the Standard Model(SM) Higgs boson in $\tau^+\tau-$ final state in Tevatron/CDF experiment. The highest production process of the SM Higgs boson is gluon fusion process ($ig \rightarrow H$). Even in addition, there are some Higgs production processes: W/Z boson associated process($q\bar{q'} \rightarrow AH/q\bar{q} \rightarrow ZH$) and vector boson fusion($qq' \rightarrow qHq'$) and so on. The SM Higgs boson has different decay mode depending on its mass point. The theoretical Higgs boson production cross section as a function of the Higgs boson mass is shown in Figure 1.1, and the Higgs boson branching ratio as a function of its mass is shown in Figure 1.2. [7] [8]

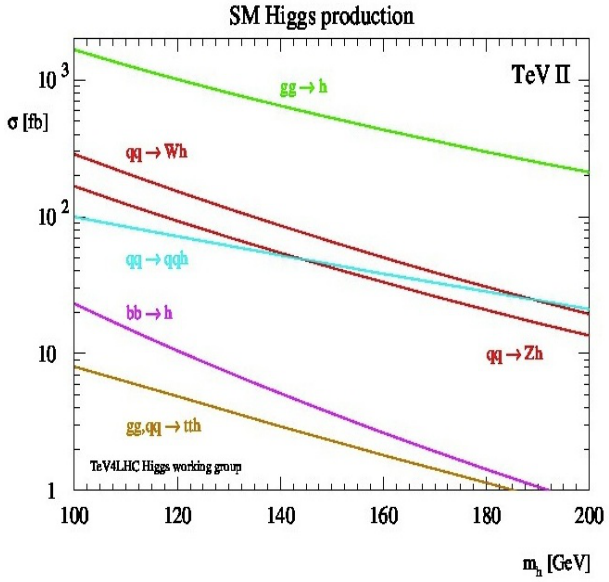

Figure 1.1: Theoretical Standard Model Higgs boson production cross section at Tevatron energy scale.

At the Tevatron/CDF experiment, the SM Higgs boson is searched for a long time (∼10year). In high mass region($M_H \geq 135GeV/c^2$), main decay modes of the SM Higgs boson is W boson pair. The process of gg→ $H \rightarrow WW$ is the main channel for high mass region and have shown strong performance of search(or exclude) for the SM Higgs. [9] In the low mass region($100GeV/c^2 \leq M_H \leq 135GeV/c^2$), on the other hand, the main decay mode is b quark pair. However the process of gg→ $H \rightarrow b\bar{b}$ have the highest probability, this channel is suffered from large QCD background. Therefore, main channel for low mass Higgs boson is W/Z boson associated production with the Higgs boson decaying into $b\bar{b}$ and W/Z boson decaying into leptonically. [10] These searches are difficult for background modeling and estimation of $b\bar{b}$ process due to poor resolution jet energy measurement compared with electrons and muons. Therefore in low mass region, several sub-channels are analyzed to obtain further sensitivity. The resent result of the SM Higgs boson search at Tevatron experiment is shown in Figure 1.3. [11]

Considering this situation, the $H \rightarrow \tau\tau$ process is powerful channel to improve search sensitivity in low mass region. Because the branching ratio of the SM Higgs to τ pair is second highest and this process is possible to consider four Higgs production process: gluon fusion and W/Z boson associated and vector boson fusion. And the main background of this process is Electroweak process, such as $Z \rightarrow \tau\tau$, which we can well estimating production rate and modeling kinematic

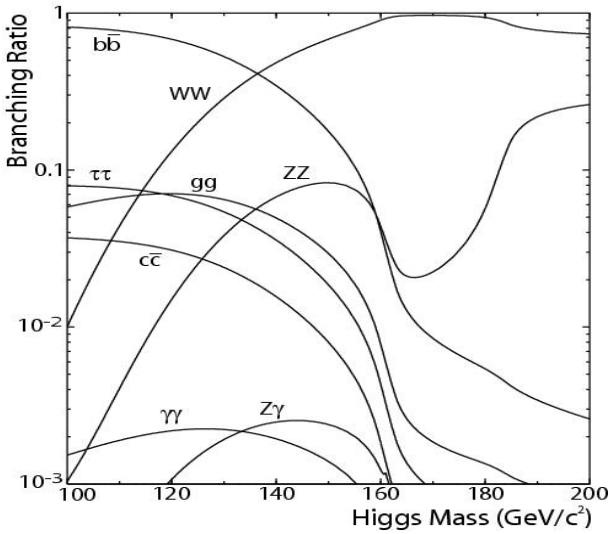

Figure 1.2: The Standard Model Higgs boson branching ratio for each Higgs boson mass.

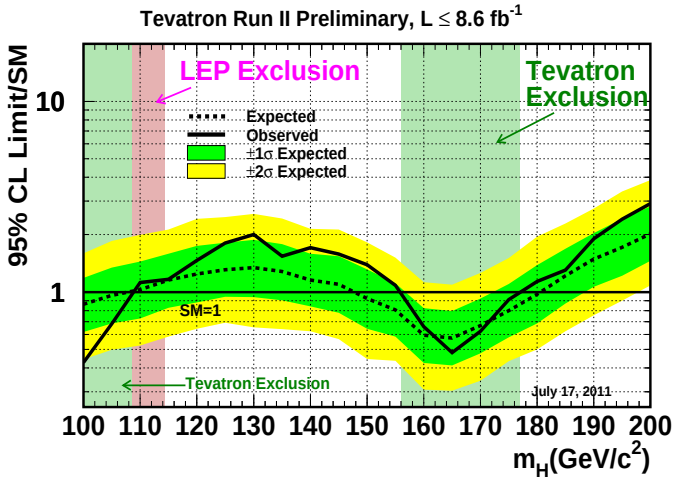

Figure 1.3: The resent result of the search for the Standard Model Higgs boson at Tevatron experiment.

shape compared with QCD process.

As further motivation, if the SM Higgs boson exist, indirect search of the SM Higgs boson suggest that the SM Higgs exists in low mass region ($M_H = 92GeV^{+34}_{-26}$ at 68% confidence level) [12]. Therefore there is importance to improve search sensitivity in low mass region using τ not only b.

Furthermore, there is the most important motivation in this analysis, related to the Large Hadron Collider(LHC). In the LHC, there are two experiment searching for the SM Higgs boson: A Toroidal LHC Apparatus(ATLAS) and Compact Muon Solenoid(CMS). The resent results of the SM Higgs boson search at ATLAS and CMS are shown in Figure 1.4. [13], [14] The left plot is the all channel combination result in low mass region at ATLAS experiment and the right plot is the combination result of the $H \to \gamma\gamma$ and $H \to ZZ \to 4l$ in low mass region.

We can see interesting signs of the exist of the SM Higgs boson around 125GeV/c^2 is observed in both ATLAS and CMS experiment. Those signs are 3.6σ at 126GeV/c^2 in the ATLAS result and 2.6σ at 124GeV/c^2 in CMS experiment. But if we consider *look-elsewhere-effect*(LEE), this signs became 2.5σ and 1.9σ for ATLAS and CMS experiment,respectively. Since we can not declare the exist of the SM Higgs boson from these results, the mutual evaluation of this interesting signs by Tevatron and LHC is very important Because the accelerators which are able to search the SM Higgs boson are only Tevatron and LHC.

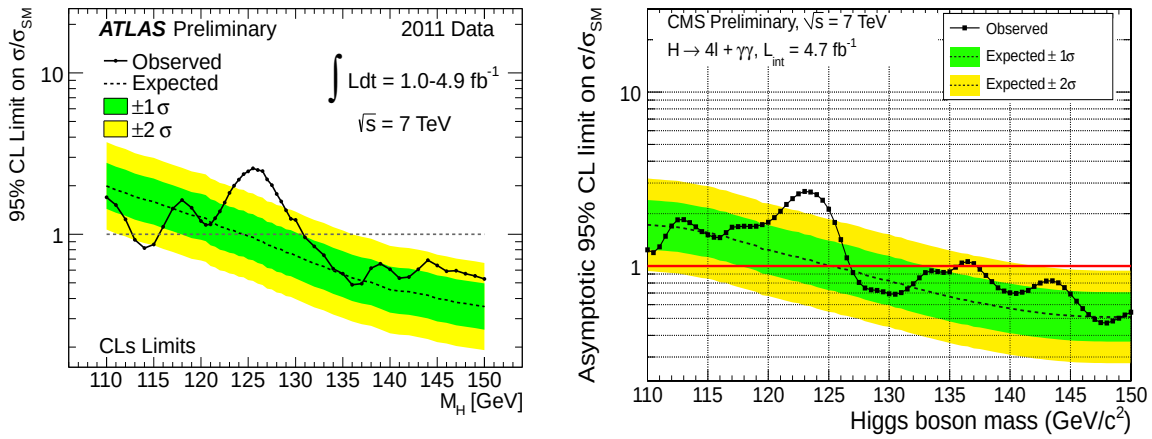

Figure 1.4: The resent result of the search for the Standard Model Higgs boson in low mass region at ATLAS and CMS experiments.

There is also the motivation in analysis methodology. The vector boson fusion(VBF) production cross section becomes larger than the W/Z boson associate production since the colliding particles both proton in the LHC. [15] In this process, $H \to b\bar{b}$ mode is hard to search because of large QCD background. Therefore, the one of main channel of the low mass SH Higgs boson search is $VBF \to \tau\tau + qq'$. Since this channel is included in this analysis, the improvement of this analysis methodology connects the improvement of the SM Higgs search at both the Tevatron and the LHC experiments.

In this analysis, we focus on that final state is τ pair. The categorization of the signal events into 0/1/$\geq$ 2 jet bins improves search sensitivity, because the background components and the signal processes are different depending on the number of jets.

- The W boson associated processes: Since the branching ratio of the W bosons decaying to

hadrons is 67.6% and there is a possibility to generate quarks from initial state radiation, almost events are categorized into $\geq$ 1jet bin channel.

- the Z boson associated processes: Similarly, since the branching ratio of the Z bosons decaying to hadrons is 69.9% and there is a possibility to generate quarks from initial state radiation, almost events are categorized into $\geq$ 1jet bin channel.
- The vector boson fusion process: Since this process have naturally two quarks at tree diagram level, almost events are categorized into $\geq$ 1jet bin channel.
- The gluon fusion process: Since this process has no jets at tree level, almost events are categorized into 0jet bin channel. But this process have large fraction of initial state radiation and the highest production cross section, events are also categorized into $\geq$ 1jet bin channel.

The Feynman diagrams of signal processes are shown in Figure 1.5, Figure 1.6, Figure 1.7 and Figure 1.8.

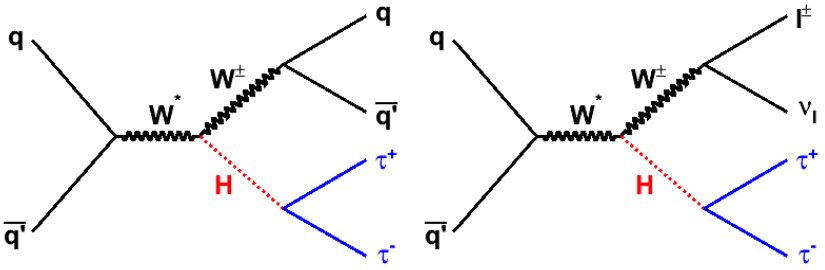

Figure 1.5: Feynman diagram for W boson associated process.

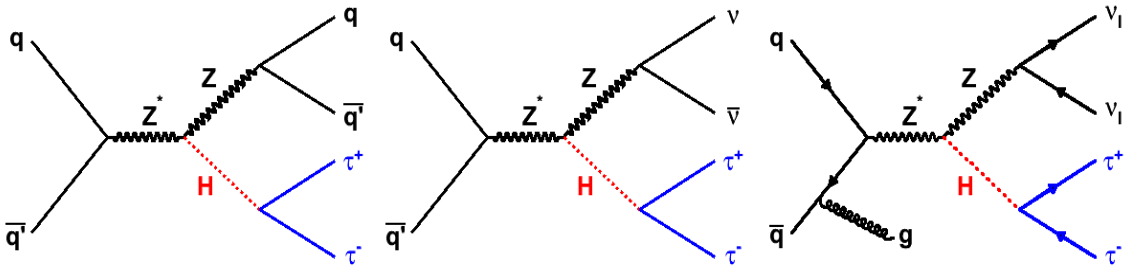

Figure 1.6: Feynman diagram for Z boson associated process.

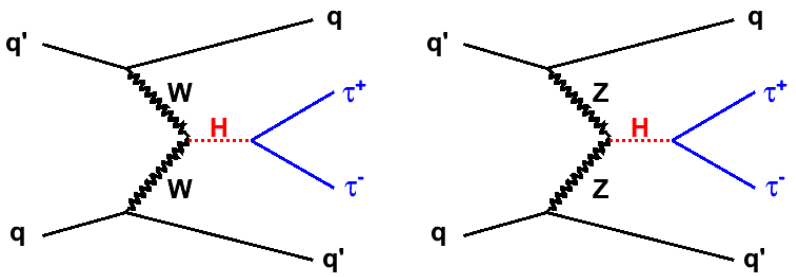

Figure 1.7: Feynman diagram for vector boson fusion process.

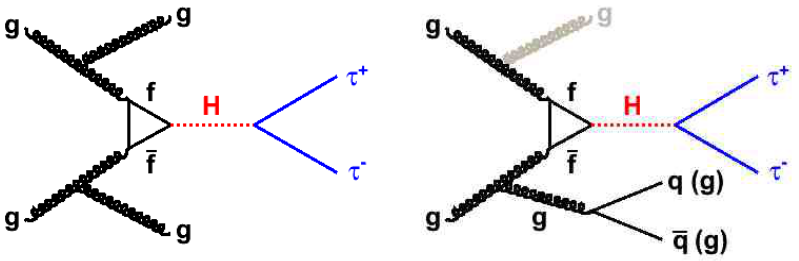

Figure 1.8: Feynman diagram for gluon fusion process with initial state radiation, which final states are $\tau\tau$ + jets.

We select event by requiring two τ which comes from the Higgs boson. In ditau decay, we consider two decay mode as signal.

(1) One of two τ decays to leptonically which is observed as electron or muon and the other one decay to hadronically which is observed as hadronic τ. (BR is 46%)
(2) Both of two τ decay to leptonically which are observed as one electron and one muon. (BR is 6%)

The first mode has a high branching ratio and the requirement of electron or muon suppresses the fake event such as QCD and W+jets. The second mode has a small branching ratio, but is very clean event because the drell-yan and other background events do not mimic this final state easily. Therefore, the final states of this analysis are e/μ + τ_{had} and e + μ with or without jets.

Chapter 2

Experimental Apparatus

In this paper, we analyze the data collected by Collider Detector at Fermilab(CDF) in collisions of proton and anti-proton delivered by Tevatron accelerator. We describe a brief summary of fermilab accelerator complex and detector configuration in this chapter.

2.1 The Fermilab Accelerator Complex

The Tevatron is a proton-antiproton synchrotron collider located at Fermi National Accelerator Laboratory in Batavia of the illinois state (USA). The proton and anti-proton beams are the result of acceleration chain with several stage. The Tevatron is the last state in accelerator chain to increase energy of protons and anti-protons and to collide them. A overview of Fermilab Accelerator Chain is shown in Figure 2.1.

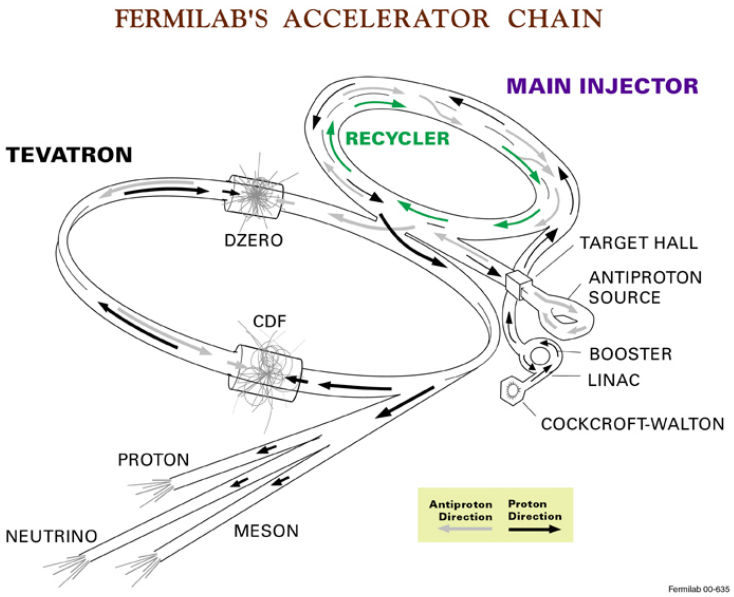

Figure 2.1: A overview of Fermilab Accelerator Chain

2.1.1 Proton Production

The First step of the accelerator chain is proton production extracted from very pure Hydrogen gas, which is moved between two electrodes and ionized into electrons and H^+ ions. After the H^+ ions strike a cathode made of cesium and release electrons, H^- ions are produced and sent to a Cockcroft-Walton preaccelerator.

In Cockcroft-Walton preaccelerator, H^- ions are accelerated to an energy of 750GeV. Then, H^- beams are transported to next step, Linear accelerator(Linac).

Linac accelerates H^- beams to an energy of 400MeV using Radio-Frequency(RF) cavities, when H^- ions are grouped into bunches by RF cavities. After, the 400MeV H^- ions are injected into Booster, when they pass through carbon foil and protons are produced by stripping off electrons. The Fermilab's Cockcroft-Walton preaccelerator and Linac are shown in Figure **??**.

The Booster is a proton synchrotron, approximately 150m in diameter, that accelerates protons to an energy of 8GeV. The proton beams are divided into 84 benches spaced by 18.9ns, each of them consisting of $\sim 6 \times 10^{10}$ proton, and they are transported to the Main Injector.

Figure 2.2: The Fermilab's Cockcroft-Walton preaccelerator and Linac

2.1.2 The Main Injector

The Main Injector is a 3km circular synchrotron with 18 accelerating RF cavities and conventional magnet. It has two operation mode: *Accumulation Mode* and *Collider Mode*.

In *Accumulation Mode*, it produces 120GeV protons from 8GeV, that an energy of 120GeV is optimal to produce antiprotons, and they are transported to nickel target for antiproton production.

When *Collider Mode* is set after antiprotons are injected, it accelerates protons and antiprotons to an energy of 150GeV. Finally, protons and antiprotons are injected into Tevatron.

2.1.3 Antiproton Production

The 120 GeV protons from the Main Injector are send to the nickel target line and collided with them. Secondary particles from this collision are collected and focused with a cylindrical lithium lens. Then, 8GeV antiprotons which have negative charge are selected by 1.5T dipole magnet. One antiproton is collected by $\sim 10^5$ protons. The 8GeV antiprotons are transported to the Main Injector and accelerate to 150GeV in *Collider Mode*.

2.1.4 Tevatron

The Tevatron is a large synchrotron with a 1km radius, which accelerates protons and antiprotons from the Main Injector from 150GeV to 980GeV. The protons and antiprotons in Tevatron are kept in the same beam-pipe with opposite direction since electrostatic separators which produce strong electric field reduce non-interesting interaction except at the collision points. The beams are steered by super-conduction dipole magnets and quadrupole magnets, and are cooled by 4.2Kelvin liquid helium. There are three trains which are groups of 12 bunches of protons and antiprotons and an abort gap between them in the Tevatron. The proton and antiproton beams are brought to two collision points,B0 and D0, where two detectors are installed in them, the Collider Detector at Felmilab(CDF) and D0.

Figure 2.3: The aerial view of the Tevatron accelerator.

The collision rate of protons and antiprotons (R) is described by the instantaneous luminosity (L), given by following formula:

$$R = \sigma \times L, \tag{2.1}$$

where σ is the cross section of the process which has interest, and the instantaneous luminosity is calculated by following formula:

$$L = \frac{f_r N_B N_p N_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}^2})} F \frac{\sigma_l}{\beta^*}, \tag{2.2}$$

where f_r is the revolution frequency, N_B is the number of bunches, N_p and $N_{\bar{p}}$ are the number of protons and antiprotons per one bunch, σ_p and $\sigma_{\bar{p}}$ are the size of proton and antiproton beams, F is a form factor which depends on the beam spared (σ_l) and the beta function (β^*). The time integrated luminosity(L_{int}), used as data quantity, is expressed by:

$$L_{int} = \int L dt. \tag{2.3}$$

In the Tevatron, a span of collision from beginning to end is called as "store". The instantaneous luminosity of the store gradually decreases since the protons and antiprotons in beams are consumed by interactions and beam scattering. After about 20 hours from the beginning of collision, the instantaneous luminosity goes too low to collide. Then, the store is ended and the

current beams are dumped, the Tevatron prepare for injecting new beams. The peak instantaneous luminosity as a function of store is shown in Figure 2.4.

The maximum peak instantaneous luminosity at the beginning of the store was recorded $\sim 4.3 \times 10^{32}[cm^{-2}s^{-1}]$ at CDF. The integrated luminosity as a function of store is shown in Figure 2.5. The total value of the integrated luminosity delivered by the Tevatron is about $12fb^{-1}$, and acquired at CDF is about $10fb^{-1}$, until the termination of the Tevatron.

A delivered integrated luminosity and a acquired luminosity are not same because of detector troubles, the time of raising up high voltage and beam condition. The efficiency of the CDF data taking as a function of store is shown in Figure 2.6. The average efficiency is $\sim 80\%$ is reached by the CDF.

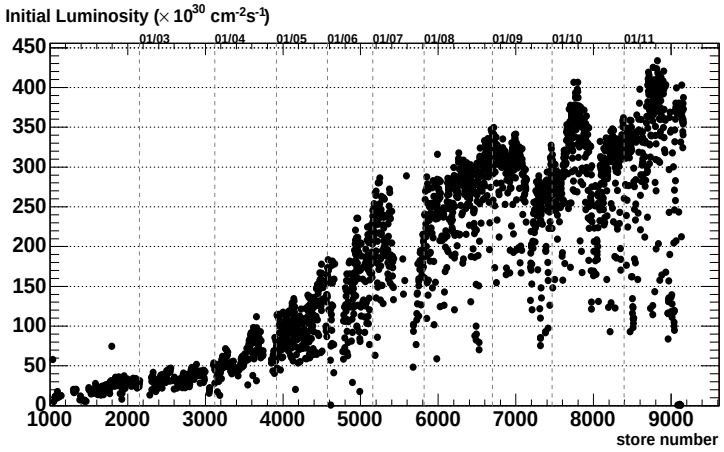

Figure 2.4: The initial instantaneous luminosity

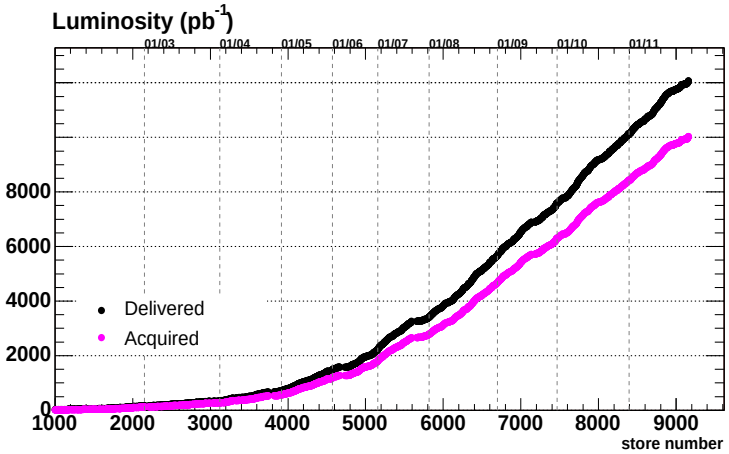

Figure 2.5: The integrated luminosity

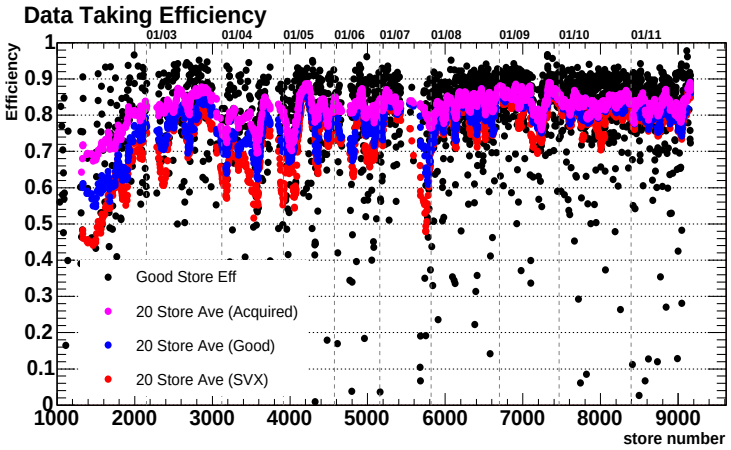

Figure 2.6: The efficiency of the CDF data taking

2.2 Collider Detector at Fermilab (CDF)

The CDF Ⅱ Detector is general purpose detector which is installed at one of collision point of the Tevatron. It was operated from the beginning of the Tevatron Run Ⅱ in 2001, and terminated in 2011, which is the same time of the Tevatron termination.

The CDF is consisted of Silicon Tracking Detector, Central Outer Tracker, Electromagnetic Calorimeters, Hadron Calorimeters and Muon Detectors in order from inner to outer. A cutaway of CDF Ⅱ detector is shown in Figure 2.7.

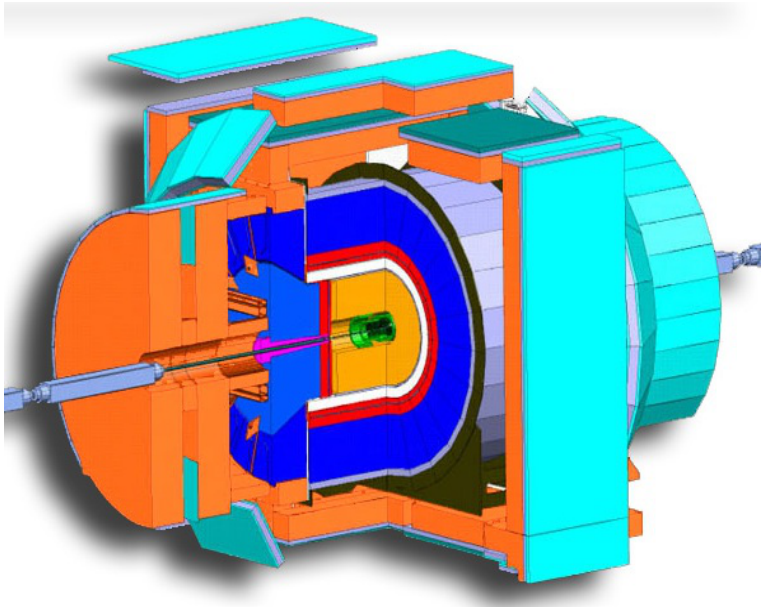

Figure 2.7: A cutaway of the CDF Ⅱ detector

The $\eta - \phi$ plane is used to identify the locations of particles, where ϕ is the azimuthal angle and η is the pseudo-rapidity expressed as following:

$$\eta = -ln(\frac{\theta}{2}), \tag{2.4}$$

where θ is the polar angle expressed as following:

$$\theta = \frac{1}{2}ln(\frac{|\vec{p}| + p_z}{|p| - p_z}), \tag{2.5}$$

where $|\vec{p}|$ is the particle's momentum and p_z is the particle's momentum of the beam axis.

2.2.1 Tracking System

The tracking system which is innermost part of the CDF IIdetector is consisted of Silicon Tracking Detectors, Central Outer Tracker. These detectors are located inside B=1.4T solenoide magnetic field and provide three-dimensional particle tracking and their momenta.

Especially, Silicon Tracking Detectors are also used to detect secondary vertexes from weak decays from bottom quarks or charm quarks, which is useful for b-tagging and B hadron physics. The tracking system in the CDF is shown in the Figure

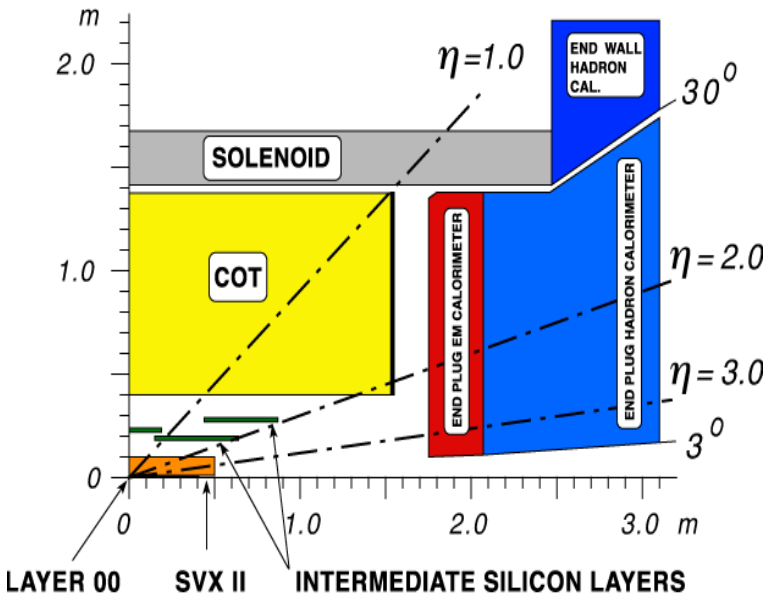

Figure 2.8: The Tracking System in the CDF

Silicon Tracking Detectors

The Silicon Tracking Detectors are consisted of the Layer 00(L00), the Secondary Vertex Detector(SVX) and Intermediate Silicon Layers(ISL). [16], [17], [18] In this analysis, Silicon Tracking Detectors are only used for increasing track quality.

The Layer 00(L00) The L00 is the single-side silicon micro strep detector placed directly on the surface of beam-pipe, providing a coverage of $|\eta| < 4.0$ and a full coverage of ϕ. It improves the resolution of the impact parameter and the position of secondary vertex.

The Secondary Vertex Detector The SVX is the main part of the silicon detectors and consists of five layers of double-side silicon strip, providing a coverage of $|\eta| < 2.0$ and a full coverage of ϕ. The purpose of the SVM is to provide tracking information with high resolution and to reconstruct secondary vertex.

The Intermediate Silicon Layers(ISL) The ISL is a outermost part of the silicon detectors and consists of single layer ($|\eta| \leq 1.0$) and two layer ($0.1 \leq |\eta| \leq 2.0$). It improve track resolution by connecting tracks between the SVX and the COT.

Central Outer Tracker(COT)

The COT is the cylindrical drift chamber which is installed in central region ($|\eta| < 1.0$) at outside of Silicon Tracking Detectors. [19] The COT is constructed from eight *super-layers*, which are placed in alternating the axial and stereo *super-layer*: the axial layers are parallel to the z-axis(beam line) and the stereo layers are attached to $\pm 2^\circ$ angle with respect to the z-axis.

When charged particles are passing through their chambers, the gasses in the COT is ionized and create electrons. Their electrons are drifted by electric field to the wires and create signal pulses of the position of charged particles. We can reconstruct the tracks of charged particles from connecting with detected their positions of super-layers, and calculate momenta from the curvature of tracks. The position resolution of the COT is 140μm and transverse momentum resolution is:

$$\frac{\sigma_{P_T}}{P_T^2} = 0.0015[GeV^{-1} \cdot c]. \tag{2.6}$$

2.2.2 Calorimeters

Calorimeters of the CDF is installed at outside of tracking system and measure the energy of both charged particles and neutral particles. They are consisted of ElectroMagnetic(EM) calorimeters and Hadron calorimeters, and their coverage is $|\eta| < 3.6$. [20] Both calorimeters are scintillator sampling detectors segmented in $\eta - \phi$ sections which is called towers to measure information of the particle position. Each tower consists of alternating layers of materials, a lead(Pb) is used in the electromagnetic calorimeters and a iron(Fe) is used in the hadronic calorimeters, and plastic scintillators for shower sampling. The light from the scintillator plates is read out by photo-multiplier tubes through wavelength shifters and light guides. The cross section of the CDF calorimeter system is shown in Figure 2.9.

The EM calorimeters are consisted of the Central ElectroMagnetic calorimeter(CEM) and the Plug ElectroMagnetic calorimeter(PEM). They measure the energies of electromagnetic particles like electron and γ. They include the Central Electromagnetic Shower maximum detector (CES) and the Plug Electromagnetic Shower maximum detector (PES) to improve the position resolution of the EM calorimeters.

The Hadron calorimeters are consisted of the Central HAdron calorimeter(CHA), the Wall HAdron calorimeter(WHA) and the Plug HAdron calorimeter(PHA). They measure the energy of hadron particles like charged pions and kaons and so on.

The Central Electromagnetic Calorimeter (CEM)

The CEM is segmented into 24 towers in ϕ direction and 10 towers in η direction, whose total thickness is $18X_0$ (X_0 is radiation length). The energy resolution of the CEM is:

$$\frac{\sigma_E}{E} = \frac{13.5\%}{\sqrt{E_T}} \oplus 2\%. \tag{2.7}$$

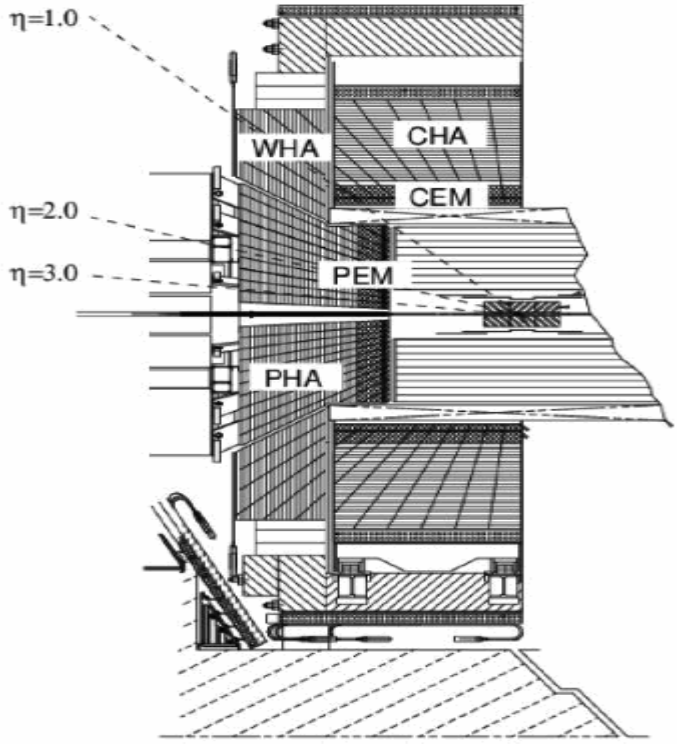

Figure 2.9: The cross section of the CDF calorimeter system

The Central Electromagnetic Shower Maximum Detector (CES)

The CES located at $6X_0$ depth in the CEM is a proportional strip and wire gas chamber. It measures the position of particles and the shape of electromagnetic showers in $\eta - phi$ plan, by measuring the charge deposition of the electromagnetic showers.

The Plug Electromagnetic Calorimeter (PEM)

The PEM is segmented into 24 and 48 towers for inner and outer parts in ϕ direction and 12 towers in η direction, whose total thickness is $21X0$, and it covers $1.1 \leq |\eta| \leq 3.6$ region. The energy resolution of the PEM is:

$$\frac{\sigma_E}{E} = \frac{14.4\%}{\sqrt{E_T}} \oplus 0.7\%. \tag{2.8}$$

The Plug Electromagnetic Shower Maximum Detector (PES)

The PES located at $6X_0$ depth in the PEM is similar detector with the CES, and it covers $1.1 \leq |\eta| \leq 3.6$ region.

The Central Hadron Calorimeter (CHA)

a The CHA is made of of 32 layers of alternating iron and scintillator, and it covers $|\eta| \leq 1.1$. The energy resolution of the CHA is:

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E_T}} \oplus 3\%. \tag{2.9}$$

The Wall Hadron Calorimeter (WHA)

The WHA is made of of 15 layers of alternating iron and scintillator, and it covers the gap between central and plug region. The energy resolution of the WHA is:

$$\frac{\sigma_E}{E} = \frac{75\%}{\sqrt{E_T}} \oplus 4\%. \tag{2.10}$$

The Plug Hadron Calorimeter (PHA)

The PHA is made of 23 layers and the coverage is $1.1 \leq |\eta| \leq 3.6$. The energy resolution of the PHA is:

$$\frac{\sigma_E}{E} = \frac{80\%}{\sqrt{E_T}} \oplus 5\%. \tag{2.11}$$

2.2.3 Muon Detectors

Muon detectors are installed in outermost part of the CDF since muons are minimum ionizing particles (MIP) at the collision energy of the Tevatron. [21]

Muon detectors of the CDF are consist of four part: Central Muon Detectors (CMU), Central Muon Upgrade Detectors (CMP), Central Muon Extension Detectors (CMX) and Intermediate Muon Detectors (IMU), which have the same principle of detection are constructed of drift chambers and scintillation counters. Their detectors are installed at different radial distance from the beam axis and detect muons in $|\eta| \leq 1.5$ region. Muon detectors coverage is shown in Figure 2.10

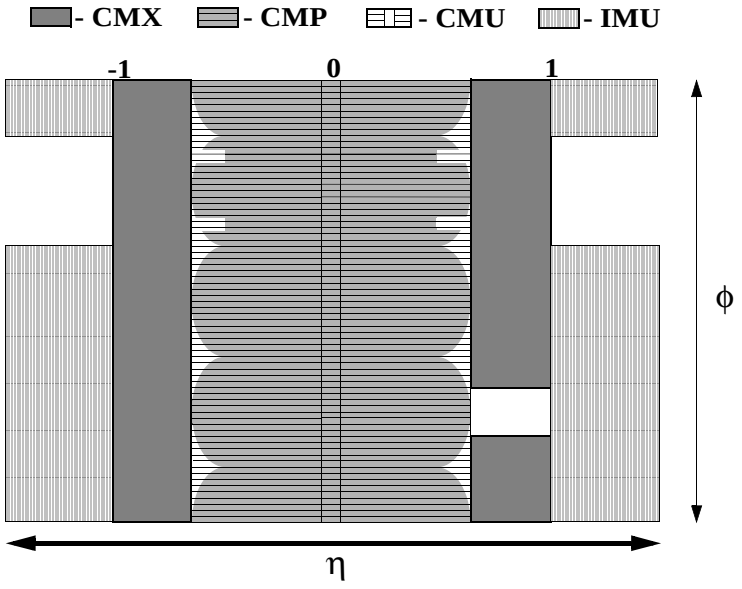

Figure 2.10: The coverage of muon detectors in $\eta - \phi$ plane

CMU

The CMU detector is located outside of CHA and is consisted of four drift tube layers sectioned by wedge matching the CHA towers. It detects muons with Pt $\leq$ 1.4GeV and covers the central region ($\eta \leq 0.6$).

CMP

The CMP detector, is consisted of four drift tube layers sectioned, located same pseudo-rapidity range of the CMU detector and separated from the CMU by 60cm steel outside magnetic yoke. It detects muon with Pt $\leq$ 2.0GeV and the covers central region ($\eta \leq 0.6$). A layer of scintillators (CSP) is installed on the outer surface of the CMP detector, which provides timing information.

CMX

The CMP detector is consisted of four or eight drift tube layers sectioned depending on the polar angle. It detects muon with Pt $\leq$ 1.4GeV and covers central region ($0.6 \leq \eta \leq 1.0$). Two layer of scintillators (CSX) are installed on the inner and outer surface of the CMP detector, which provides timing information.

IMU

The IMU covers the region of $1.0 \leq |\eta| \leq< 1.$ and is consisted of the Barrel Muon Detector (BMU), the Barrel scintillators (BSU) and the Toroid Scintillator Upgrade (TSU). It is not used in this analysis.

2.2.4 Trigger System

In the Tevatron, there are three trains which are groups of 12 bunches of protons and antiprotons. The space of each bunch in trains is 396ns, and the abort gap which is the space of each train is $2.6\mu s$. Therefore, the average bunch crossing rate is 1.7MHz. (If there is no abort gap, this rate is 2.5MHz.) Since this rate is too high to record all events, events are selected by on-line electronics, which is called the trigger system. [22] The trigger system of the CDF is designed to record only physically interesting events efficiently. There are three level in the trigger system: level1, level2 and level3. The event rate is reduced by this three level respectively. The trigger system of the CDF is shown in Figure 2.11.

Level 1

The level 1 trigger(L1) is designed hardware to find physics objects based on a subset of the detector information, and makes a decision based on simple counting of these objects such as electron and muon and so on. The event rate is reduced to $\sim$40kHz by the L1 trigger. The GLOBAL L1 trigger makes decision from L1 CAL, L1 MUON and L1 TRACK decision.

The L1 TRACK makes decision by the Extremely Fast Tracker(XFT). The XFT reconstructs tracks using the hits information from COT axial four layers. It also provides track information to L1 CAL and L1 MUON. L1 CAL makes decision using the cluster energy in the calorimeters and XFT track information. L1 MUON makes decision using the CMU detector and XFT information for muon candidates.

Level 2

The level 2 trigger(L2) is designed custom hardware to do a limited event reconstruction which can be processed in programmable processors. The event rate is reduced to $\sim$400Hz by the L2 trigger. The GLOBAL L2 trigger makes decision from L2 CAL, XCES, Silicon Vertex Tracker(SVT) and Muon information.

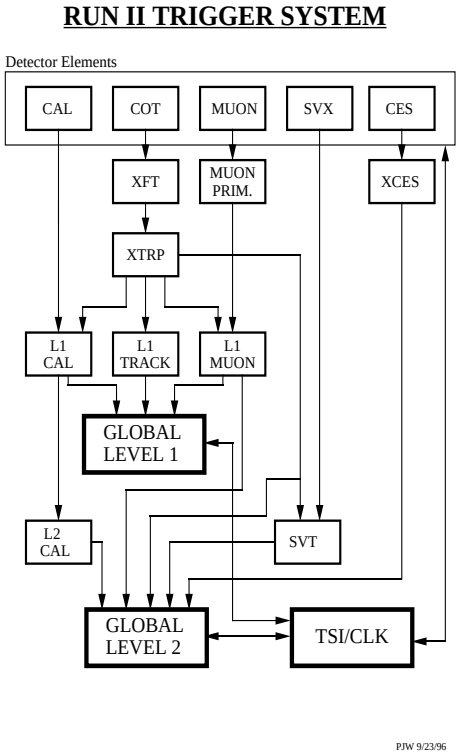

Figure 2.11: The block diagram of Level 1 and Level2 trigger system in the CDF.

L2 CAL is used to trigger cluster objects by applying some cuts. In L2 CAL, cluster defined as:

- seed tower E_T more than 3GeV
- towers E_T more than 1GeV , which is within $\Delta R = \sqrt{(}\Delta\eta^2 + \Delta\phi^2) = 0.7$ around seed tower

The XCES provides the information of the position resolution of electron and photon shower to GLOBAL L2 trigger. The SVT combine the SVX and L1 XFT track information and this system improves the resolution of track parameters; transverse angular(ϕ), transverse momentum(P_T) and impact parameter(d). These information is used to GLOBAL L2 trigger decision. The Muon information combined information from the muon detectors and the L1 XFT track information is also used to GLOBAL L2 decision.

Level 3

The level 3 trigger(L3) uses the full detector resolutions to fully reconstruct events in a few hundred processor farm. The L3 computer farm reduces the event rate to $\sim$100Hz and these events are written to tape.

Chapter 3

Event Reconstruction

The details of Lepton identification, jet energy correction/selection and missing transverse energy reconstruction are described in this chapter.

3.1 Electron Identification

In this analysis, we use electrons detected at CEM(central region: $|\eta| < 1.1$). Several cuts are applied to identify electron and they increase purity of electrons. The details of electron identification cuts are summarized in Table 3.1 [23]. The variables in Table 3.1 are defined as below:

- E_T: The transeverse energy of the candidate cluster.
- P_T: The transverse mormentum of the candidate track.
- E_{had}/E_{em}: The ratio between the energy falled at hadron calorimeter and the energy falled at electromagnetic calorimeter.
- L_{shr}: The lateral shower profile for candidate cluster in the CEM.
- E/P: The ratio between the cluster energy and track mormentum.
- $|z_0|$: The distance between the z-axis position of track and the center of the CDF detector.
- $\Delta|z_0|$(vertex,track): The distance between primary vertex and candidate track.
- $\Delta|z|_{CES}, \Delta|x|_{CES}$: The distance between the postion of CES mesurments and the postion at CES detector which extrapolated by candidate track.
- χ^2_{CES}: The χ^2 probabilty given by comparison between the shower shape measured by CES and the shower profile previously measeured with test beam electrons
- COT Hits: The number of hits at COT superlayers.
- Isolation: The ratio of the total transvarese mormentum in a cone of $R = \sqrt{\eta^2 + \phi^2} = 0.4$ around the candidate cluster to the energy of the candidate cluster.
- Photon Conversion: Electrons from photon conversion by interaction of photons and detector materials.

Central Electron Identification	
Variable	Cut
E_T	$\geq 10.0GeV$
P_T	$\geq 8.0GeV/c$
E_{had}/E_{em}	$\leq 0.055 + 0.00045 \times E$
L_{shr} COT Axial Hits	≥ 3
	≤ 0.2
E/P	≤ 2.0 unless $P_T > 50GeV/c$
$\|z_0\|$	$\leq 60.0cm$
$\Delta z_0(vertex, lepton)$	$\leq 5.0cm$
$\Delta\|z\|_{CES}$	$\leq 3.0cm$
$Q_{track} \times \Delta x_{CES}$	$\geq -3.0cm$, $\leq 1.5cm$
χ^2_{CES}	≤ 10.0
COT Axial Hits	≥ 3
COT Stereo Hits	≥ 2
COT Total Hits	≥ 5
Isolation	≤ 0.1
Fiducial	Yes
Conversion	Veto

Table 3.1: Central Electron (CEM) Identification

3.2 Muon Identification

In this analysis, the two types of muons detected at CMUP and CMX are considerd, and these total coverage is $|\eta| < 1.0$. The CMUP means that muons are detected at both CMU and CMP detector. Since muons detected at only CMU have a large fraction of fake muons, which they are acutually hadrons passing through hadron calorimeter, we also require that muons are detected at CMP to supress fake muons. The muon identification cuts for CMUP and CMX are different, and the details of them are summarized in Table 3.2 [23]. The variables in Table 3.2 are defined as below:

- d_0: The impact parameter of candidate track.
- $|\Delta x|_{CMU,CMP,CMX}$: The distance between the postion of CMU,CMP,CMX mesurment and the postion at CMU,CMP,CMX detector which extrapolated by candidate track. item Cosmic Ray: When cosmic ray through the detector, it looks like back-to-back tracks with the identical d_0. Therfore, we can detected cosmic ray muons depending on d_0.

Central Muon Identification	
Variable	Cut
P_T	$\geq 10.0 GeV/c$
E_{had}	$\leq 6 + max(0, 0.0115 \times (P - 100))$
E_{em}	$\leq 2 + max(0, 0.0028 \times (P - 100))$
$\|z_0\|$	$\leq 60.0cm$
Δz_0(vertex,track)	$\leq 5.0cm$
$\|d_0\|_{track}$ with Silicon Hits	$\leq 0.02cm$
$\|d_0\|_{track}$ without Silicon Hits	$\leq 0.2cm$
COT Axial Hits	≥ 3
COT Stereo Hits	≥ 2
COT Total Hits	≥ 5
$\|\Delta x\|_{CMU}$	$\leq 7.0cm$ for CMUP
$\|\Delta x\|_{CMP}$	$\leq 5.0cm$ for CMUP
$\|\Delta x\|_{CMX}$	$\leq 6.0cm$ for CMX
COT radius (ρ)	$\geq 140.0cm$ for CMX
Isolation	≤ 0.1
Fiducial	Yes
Cosmic	Veto

Table 3.2: Central Muon (CMUP,CMX) Identification

3.3 Tau Identification

Theτ lepton is short lived particle (mean life time is $290.6 \times 10^{-15}[sec]$) compared with other leptons, we can detect τ only after its decay with weak interaction. This decay is categorized to 2

types by W boson decay.

One is leptonic decay ($\tau^{\pm} \to e^{\pm} + \nu_e + \nu_\tau$ / $\tau^{\pm} \to \mu^{\pm} + \nu_\mu + \nu_\tau$) which B.R is $\sim 17\%$ for each lepton type, and this case is identified by electron or muon.

The other is hadronic decay ($\tau^{\pm} \to \nu_\tau + h^{\pm}$ / $\tau^{\pm} \to \nu_\tau + h^{\pm}h^{\mp}h^{\pm}$) which B.R is $\sim 65\%$, where $h^{\pm}$ means $\pi^{\pm}$ or $K^{\pm}$ or other charged hadron and also this decay associate with ≥ 0 π^0 or K^0 or other neutral hadron. In this case, two categories are considered as:

- **1 prong** τ : the number of the charged hadron is one. (B.R. = 50%)
- **3 prong** τ : the number of the charged hadron is three. (B.R. = 15%)

Since the branching ratio to five or more then charged hadron is too rare (B.R.$\leq 10^{-3}$), they are neglected in this case. The feynman diagram of leptonic and hadronic τ decay is shown in Figure 3.1.

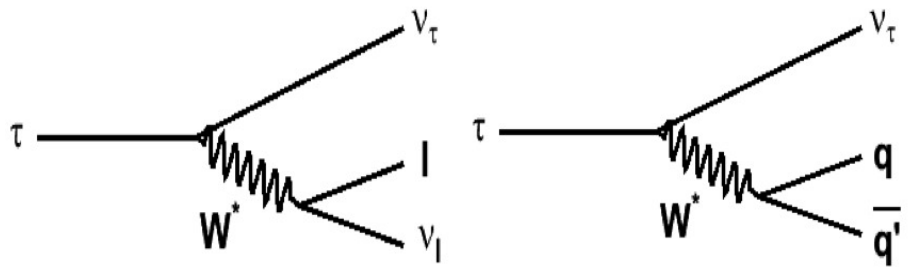

Figure 3.1: Feynman diagram of leptonic and hadronic τ decay.

The hadronic τ is identified by cut base, and the variables and cut values are summarized in Table 3.3. The variables in Table 3.3 are defined as below:

- Seed Tower E_T: The transverse energy of seed tower in the candidate cluster.
- Sholuder Tower E_T: The transverse energy of sholuder tower, where sholuder tower is adjacent towers around the seed tower in the candidate cluster.
- N^{tower}: The number of towers in the candidate cluster.
- Cluster E_T: The transverse energy of candidate cluster.
- Seed Track P_T: The transverse mormentum of the seed tower.
- Sholuder Track P_T: The transverse mormentum of the shoulder tower.
- θ_{sig}: 3 dimensional angle with respect to the seed track direction, which definies the signal con.
- θ_{iso}: 3 dimensional angle with respect to the seed track direction, which definies the isolation con.
- N^{track}_{signal}: The number of tracks in the signal cone.
- $|Q^{track}|$: Sum of track's charge in the signal cone.

- Tracks P_T in isolation cone: Sum of the transverse mormentum of trakcs in isolation cone.
- $\pi^0 s$ E_T in isolation cone: Sum of the transverse energy of π_0s in isolation cone.
- $\Delta z_{sholuder\ track}$:
- $|z_{CES}^{seed\ track}|$: The distance between the postion of seed track by CES mesurments and the postion of seed track at CES detector which extrapolated by track infomation.
- visible P_T $(tracks\ +\ \pi^0 s)$: Sum of the transverse mormentum of visible particle (charged particle and π_0).
- visible $Mass$ $(tracks\ +\ \pi^0 s)$: The mass of visible particles.
- ξ': This variable is applied for suppressing electron fake. The definition is following formula:

$$\xi' = \frac{E_{total}}{\sum |\vec{p}|}(0.95 - \frac{E_{em}}{E_{total}}), \tag{3.1}$$

  where $\sum |\vec{p}|$ is the sum of the scalar momentum of tracks in signal cone and E_{total} is the total energy detected at electromagnetic and hadron calorimeters ,and E_{em} is the energy detected at electromagnetic calorimeter.

The definition of the signal and isolation cone for tracks and π_0s is shown in Figure 3.2 [**?**].

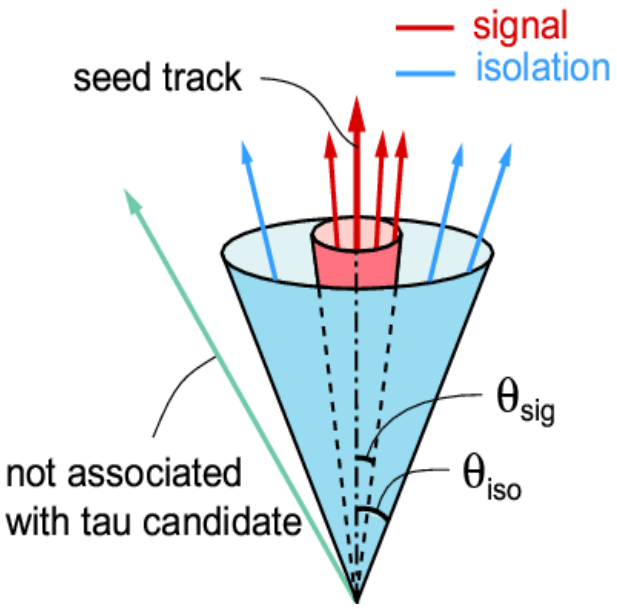

Figure 3.2: The definitaion of the signal and isolation cone for tracks and π_0s.

3.4 Jet Identification

The final state of quarks and gluons is called jet, which is a group of particles. In this analysis, jets are defined as a group of particles within a cone which radius is $\Delta R = \sqrt{\eta^2 + \phi^2} \leq 0.4$. The jet direction is defined as the direction of the center of jet cones, and the energy of jets is

Central Muon Identification	
Variable	Cut
η	≤ 1.0
Seed Tower E_T	$\geq 6.0GeV$
Sholuder Tower E_T	$\geq 1.0GeV$
N^{tower}	≤ 6
Cluster E_T	≥ 9.0
Seed Track P_T	$\geq 6.0GeV$
Shared Track P_T	$\geq 1.0GeV$
COT Axial Hits	≥ 3
COT Stereo Hits	≥ 2
COT Total Hits	≥ 5
θ_{sig}	min(0.17, $\frac{5.0GeV}{E^{cluster}}$)rad
θ_{iso}	0.52rad
N^{track}_{signal}	$= 1 \; or \; 3$
$\lvert Q^{track} \rvert$	$= 1$
Tracks P_T in isolation cone	$\geq 2GeV$, no track $P_T \geq 1.5GeV$
$\pi^0 s$ E_T in isolation cone	$\geq 1GeV$
$\Delta z_{shared\ track}$	$\leq 5.0cm$
$\lvert z^{seed\ track}_{CES} \rvert$	$\geq 9.0cm$, $\leq 230.0cm$
ξ'	≥ 0.1
visible P_T $(tracks \; + \; \pi^0 s)$	$\geq 15(20)GeV/c$ for 1(3) prong
visible $Mass$ $(tracks \; + \; \pi^0 s)$	$\geq 1.8(2.2)GeV/c^2$ for 1(3) prong

Table 3.3: Hadronic τ Identification

sum of energy falled at electromagnetic and hadron calorimeter within $\Delta R \leq 0.4$. But the raw energy of jets are not correctly reconstructed original quarks and gluons due to instrumental and physical reason. Therefore, we apply several energy correction to estimate the energy of quarks and gluons [24].

Relative Correction

The first step of this jet energy correction is the relative correction, which is corresponding to calorimeters. The calorimeters are segmented in several towers corresponding to η, and they have different resolution of the energy mesurement. These effects are corrected depending on η and P_T.

Dijet events are used to measure the scale factor of this correction. Since these events are balance in transverse plane, the transverse energy of two jets should be zero. The energy of jet in plug or forwad region, which called *prove jet*, is scaled using the energy of the jet in $0.2 \leq \eta \leq 0.6$, which called *trigger jet*. beacuse the central calorimeters the CEM and the CHA are the best understood calorimeters in the CDF. In addition, the resolution of the central calorimeters has dependence for the transverse mormentaum of jet. The transverse moementum balancing function $f_{balance}$ is defined as

$$f_{balance} = \frac{P_T^{prove} - P_T^{trigger}}{P_T^{prove} + P_T^{trigger}/2}, \tag{3.2}$$

where P_T^{prove} and $P_T^{trigger}$ are the transverse mormenta of the prove jet and trigger jet. Then, the correction factor β is defined as

$$\beta = \frac{2 + \langle f_{balance} \rangle}{2 - \langle f_{balance} \rangle}, \tag{3.3}$$

The correction factors β as a function of η for data and simulation(PYTHIA and Herwig dijet samples) are shown in Figure 3.3.

Multiple Interaction Correction

In collision of proton and antiproton, there are not only one interaction point. The interactions increase deppending on the instantaneous luminosity, and these minimum bias events affect the jet cluster. The multiple interaction correction subtracts these energy for each jet, and the factor of this correction is a linear function of the number of reconstructed vertices in the event, and they are shown in Figure 3.4. This correction factor is defined by measuring the transeverse mormentum of the jet cluster in minimum bias events.

Absolute Correction

The jet energy measured by the calorimeters must be corrected for any non-linearity and energy loss in the un-instrumented regions of each calorimeter. The absolute correction aims to transform the jet energy into the energy of particle jet. This correction depends on the jet fragmentation properties, and it is decided depending on how well the responses of the calorimeters to the single particle is modeled and on how well the the transverse mormentum of the particles in jet cluster is modeled. The absolute correction factor as a function of the transeverse mormentum is shown in Figure 3.5.

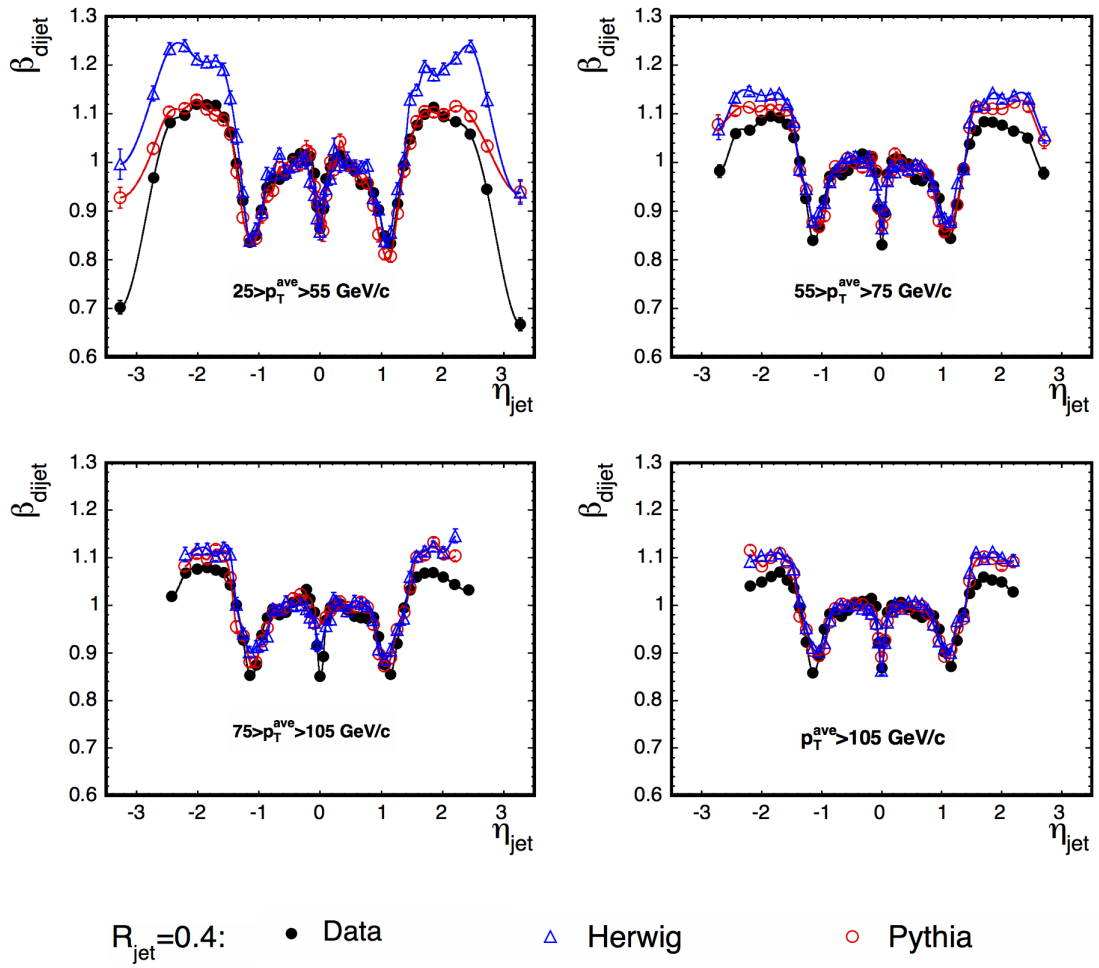

Figure 3.3: The relative correction factor as a function of pseudorapidity.

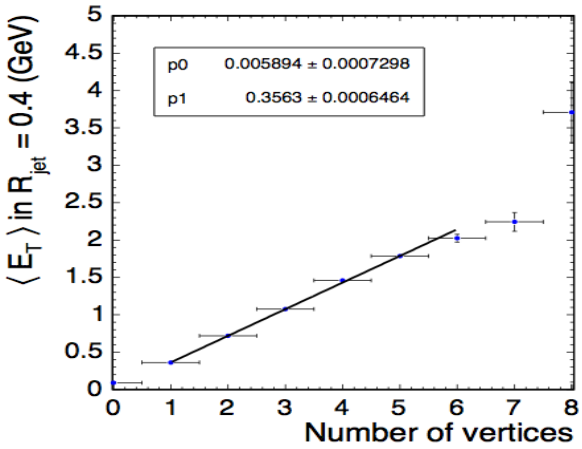

Figure 3.4: The multiple interaction correction factor as a function of the number of vertexx.

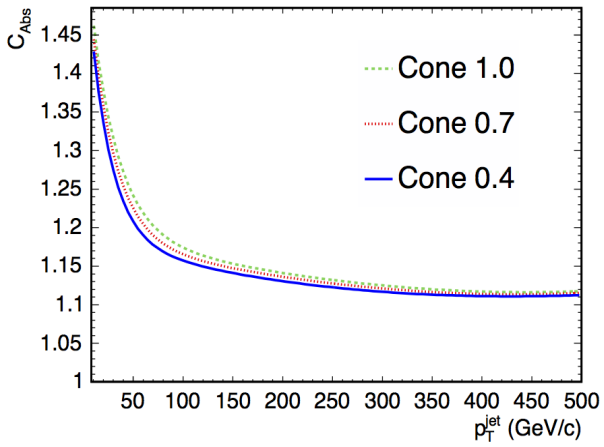

Figure 3.5: The absolute correction factor as a function of the transeverse mormentum.

Underlying Event Correction

The underlying event correction subtracts the energy of spectator quarks and particles from initial state radiation from the energy of the jet cluster. These correction factors are measured by the same procedure of the multiple interaction correction.

Out Of Cone Correction

The jet cluster is not included all the energy of the parton. This out of cone correction is applied to reconstruct the parton energy from the particle energy. The correction factors are decided by Monte Calro simulation and have dependence on the transverse mormentum of jets.

Finally, the jet energy correction is defined as

$$P_T^{parton}(R) = \left[P_T^{jet}(R) \times f_{Rel} - f_{Multi}(R)\right] \times f_{Abs}(R) - f_{UE}(R) + f_{OOC}(R), \qquad (3.4)$$

where f_{Rel} is the relative correction factor, f_{Multi} is the multiple interation correction, f_{Abs} is the absolute correction, f_{UE} is the underlying event correction, f_{OOC} is the out of cone correction and $R = \sqrt{\eta^2 + \phi^2}$ is the jet cluster size. The total systematic uncertainty of this correction is defined as adding in quadrature each correction. The total systematic uncertainty as a function of the transverse mormentum of the corrected jet is shown in Figure 3.6, and this is also discussed in 8.

In this analysis, we reconstruct jets with $\Delta R = 0.4, E_T \geq 15GeV$ and $|\eta| \leq 2.5$ corrected by relative correction, multiple interaction correction and absolute correction. Jets are removed when they match to identified lepton with $\Delta R \leq 0.2$. Signal region classified depending on the number of jets because signal sources and background compornents are respectivery different.

3.5 Missing Transverse Energy

Neutorinos are not detected with CDF detector, but we can indirectly identify them by reconstructing unbalanced energy in transverse($\eta - \phi$) plane. Missing transverse energy ($\not\!\!E_T$) is corrected by

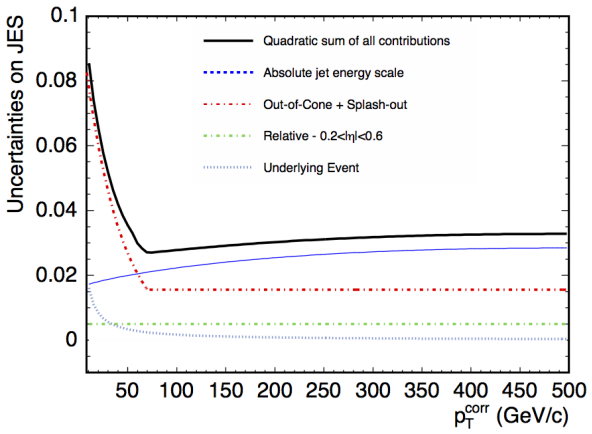

Figure 3.6: The systematic uncertainty as a function of the transeverse mormentum of the corrected jet.

the Z position of primary vertex, muon object and all jets with $E_T \geq 10GeV$ and $|\eta| \leq 2.5$. We do not use missing transverse energy in event selection but use as input variables in discriminant analysis.

Chapter 4

Event Selection

In this chapter, the event selection for $H \rightarrow \tau\tau$ search is described. First, the detaset for this analysis is described such as trigger definition for data taking and the integrated luminosity corresponding to trigger type. Next, the event selection cuts are described. The minimum selection is applied for this analysis to maximize search sensitivity, since a multivariate technique is used to discriminate signal events from background events.

4.1 Trigger Requirement

The events passing the "Lepton + Track" trigger are used in this analysis.. The trigger system of CDF detector is already described in Section **??**. The Lepton + Track trigger is generic name of three triggers: TAU_ELECTRON8_TRACK5_ISO, TAU_CMUP8_TRACK5_ISO and TAU_CMX8_TRACK5_ISO. These triggers are specially design to accept one lepton candidate with $P_T \geq 8GeV/c$ and one isolated track with $P_T \geq 5GeV/c$, which is a candidate of electron/muon or hadronic τ.

4.1.1 Trigger Definition

TAU_ELECTRON8_TRACK5_ISO

The TAU_ELECTRON8_TRACK5_ISO is design to accept one electron candidate detected at CEM and one electron/muon or hadronic τ candidate track. This trigger is consisted of

L1_CEM8_PT8
L2_CEM8_PT8_CES3_&_TRK5_DPHI10
L3_ELECTRON8_TRSCK5_ISO

Their requirements are summarized in Table 4.1.

TAU_CMUP8_TRACK5_ISO

The TAU_CMUP8_TRACK5_ISO is design to accept one muon candidate detected at CMUP and one lepton or hadronic τ candidate track. This trigger is consisted of

L1_CMUP6_PT4
L2_CMUP6_PT8
L3_CMUP8_TRSCK5_ISO

Their requirements are summarized in Table 4.2.

Trigger Name	Requirement
Level 1:	Electron
L1_CEM8_PT8	seed tower $E_T \geq$ 8GeV in $\|\eta\| \leq 1.1$
	$E^{HAD}/E^{EM} \leq 1.25$
	XFT track matched with seed tower = 3 or 4
	XFT track matched with seed tower $P_T \geq$ 8.34GeV/c
Level 2:	Electron
L2_CEM8_PT8_CES3_&_TRK5_DPHI10	seed $E_T \geq$ 8GeV in $\|\eta\| \leq 1.1$
	shoulder $E_T \geq$ 5GeV
	$E^{HAD}/E^{EM} \leq 1.25$
	XFT Hits matched with seed tower $\geq$ 3 or 4
	XFT track matched with seed tower $P_T \geq$ 8.34GeV/c
	CES measured $E \geq$ 3GeV
	Isolated track
	second XFT Hits = 4
	second XFT track $P_T \geq$ 5.18GeV/c
	Electron & Track
	$\Delta\phi$(electron,track) $\geq 10^\circ$
Level 3:	Electron
L3_ELECTRON8_TRSCK5_ISO	$E_T \geq$ 8GeV
	$P_T \geq$ 8GeV/c
	$\chi^2_{strip} \leq 20$
	$\|z_{CES}\| \leq$ 8cm
	Track
	$P_T \geq$ 5GeV/c
	$\|\eta\| \leq 1.5$
	No track with $P_T \geq 1.5$GeV/c, $\|z_0\| \leq$ 1.5cm
	and $0.175 \leq \Delta R \leq 0.524$
	Electron & Track
	$\|z_e - z_{track}\| \geq$ 15cm
	$\Delta R(e, track) \geq$ 0.175

Table 4.1: Requirement of the TAU_ELECTRON8_TRACK5_ISO trigger

Trigger Name	Requirement
Level 1:	Muon
L1_CMUP6_PT4	CMU stub$P_T \geq$ 6GeV/c
	CMP Hits $\geq$2
	XFT Hits = 4
	XFT track $P_T \geq$ 4.09GeV/c
Level 2:	Muon
L2_CMUP6_PT8	XFT Hits = 4
	XFT track $P_T \geq$ 8.34GeV/c
Level 3:	Electron
L3_CMUP8_TRSCK5_ISO	$E_T \geq$ 8GeV
	$P_T \geq$ 8GeV/c
	$\lvert x_{CMP}\rvert \leq$ 20cm
	$\lvert x_{CMU}\rvert \leq$ 15cm
	Track
	$P_T \geq$ 5GeV/c
	$\lvert\eta\rvert \leq 1.5$
	No track with $P_T \geq 1.5$GeV/c, $\lvert z_0\rvert \leq$ 1.5cm
	and $0.175 \leq \Delta R \leq 0.524$
	Muon & Track
	$\lvert z_\mu - z_{track}\rvert \geq$ 15cm
	$\Delta R(\mu, track) \geq$ 0.175

Table 4.2: Requirement of the TAU_CMUP8_TRACK5_ISO trigger

TAU_CMX8_TRACK5_ISO

The TAU_CMX8_TRACK5_ISO is design to accept one muon candidate detected at CMX and one lepton or hadronic τ candidate track. This trigger is consisted of

L1_CMX6_PT8_CSX
L2_AUTO_L1_CMX6_PT8_CSX
L3_CMX8_TRSCK5_ISO

Their requirements are summarized in Table 4.3. Since this trigger has no requirements in level 2 trigger, the events passed level 1 trigger are accepted automatically and send to level3 trigger.

Trigger Name	Requirement
Level 1:	Muon
L1_CMX6_PT8_CSX	CMX stub$P_T \geq$ 6GeV/c
	CSX information
	XFT Hits = 4
	XFT track $P_T \geq$ 8.34GeV/c
Level 2:	Muon
L2_AUTO_L1_CMX6_PT8_CSX	No requirement
Level 3:	Muon
L3_CMX8_TRSCK5_ISO	$E_T \geq$ 8GeV
	$P_T \geq$ 8GeV/c
	$\|x_{CMX}\| \leq$ 30cm
	Track
	$P_T \geq$ 5GeV/c
	$\|\eta\| \leq 1.5$
	No track with $P_T \geq 1.5$GeV/c, $\|z_0\| \leq$ 1.5cm
	and $0.175 \leq \Delta R \leq 0.524$
	Muon & Track
	$\|z_\mu - z_{track}\| \geq$ 15cm
	$\Delta R(\mu, track) \geq$ 0.175

Table 4.3: Requirement of the TAU_CMX8_TRACK5_ISO trigger

4.1.2 Luminosity

We use CDF Run IIdata collected by the Lepton + Track trigger from February 2002 to September 2011. The integrated luminosities using this analysis are different depending on trigger type, and these values are summarized in Table 4.4.

Trigger Type	$\int Ldt\ [Fr^{-1}]$
TAU_ELECTRON8_TRACK5_ISO	9571.80
TAU_CMUP8_TRACK5_ISO	9553.70
TAU_CMX8_TRACK5_ISO	8743.53

Table 4.4: The integrated luminosities for each trigger type.

4.2 Event Selection Cuts

Since final state of this analysis is $\tau\tau$, we can possibly use four decay channels depending on τ decay. The $\tau\tau$ decay channels and these branching ratios are shown in Table 4.5. In the $\tau_{had}+\tau_{had}$ mode, large QCD backgrounds are expected due to high probability of a jet fake to a hadronic τ. Also in the $\tau_{e/\mu} + \tau_{e/\mu}$ mode, large Drell-Yan (Z/γ^*+jets) backgrounds are expected and the branching ratio is very small (each 3%). On the other hand, in the $\tau_{e/\mu} + \tau_{had}$ mode, QCD backgrounds are rejected by having one isolated electron or muon, and this mode have the highest branching ratio(46%). In the $\tau_e + \tau_\mu$ mode, although the branching ratio is as small as $\tau_{e/\mu} + \tau_{e/\mu}$ mode, this mode is very clear channel because the final state of this mode is hard to mimic by QCD and Drell-Yan backgrounds. For these reasons, the $\tau_{had} + \tau_{e/\mu}$ mode and the $\tau_e + \tau_\mu$ mode are considered as signal channels in this analysis.

$\tau\tau$ decay channel	B.R	Comment
$\tau_{had} + \tau_{had}$	42%	QCD dominant
$\tau_{e/\mu} + \tau_{e/\mu}$	3%	Drell-Yan dominant
$\tau_{e/\mu} + \tau_{had}$	46%	Golden channel
$\tau_e + \tau_\mu$	6%	Clean channel

Table 4.5: $\tau\tau$ decay channel

Several cuts are applied to clean up events for this search and these cuts are different between e/μ + τ_{had} channel and e + μ channel. These cuts are summarized in Table 4.6 [25]. First selections are physical object requirements(electron, muon, hadronic τ and jet). The details of particle identification are already discussed in Chapter 3.

We further require that electron/muon and hadronic τ or electron and muon have to be the opposite sign (OS). This is based on the fact that τ pair should be the opposite sign if those come from the Higgs boson which is neutral particle.

In e/μ + τ_{had} channel, the additional cut are applied to suppress $Z \rightarrow ee/\mu\mu$ event (Z boson veto). Since Drell-Yan cross section is relatively large and the final state is actually opposite sign, we try to further reject the $Z \rightarrow ee/\mu\mu$ background. This veto is applied to following events:

- 1prong hadronic τ
- $E_{had}/P \leq 0.4$
- invariant mass of lepton and hadronic τ is between $70GeV/c^2$ and $110GeV/c^2$ that means consistent with invariant mass of Z boson.

e/μ + τ_{had} channel	e + μ channel
1 electron/muon with $P_T \geq 10GeV/c$	1 electron with $P_T \geq 10GeV/c$
1 hadronic τ with $P_T \geq 15(20)GeV/c$	1 muon with $P_T \geq 10GeV/c$
$\geq$ 0jet with $E_T \geq 15GeV$, $\|\eta\| \leq 2.5$	$\geq$ 0jet with $E_T \geq 15GeV$, $\|\eta\| \leq 2.5$
Opposite Charge ($Q_{e/\mu} \times Q_{\tau_{had}} = -1$)	Opp-site Charge ($Q_e \times Q_\mu = -1$)
Z $\rightarrow$ ee/$\mu\mu$ veto	

Table 4.6: Event Selection

Chapter 5

Background Estimation

5.1 Monte Carlo Base Estimation

The number of expected events for background using Monte Carlo (MC) sample is given by the following formula:

$$N^i = \sigma^i \times Br^i \times A^i \times \epsilon_{trig} \times \epsilon_{ID} \times \epsilon_{vtx} \times \int Ldt, \tag{5.1}$$

where i denotes each background process.

σ^i: The production cross section for i-th process

Br^i: The branching ratio for a process i if needed.

A^i: The acceptance of i-th process. It is calculated by following formula:

$$A^i = \frac{N_{select}}{N_{generate}(|Z_0| < 60cm)}, \tag{5.2}$$

where $N_{generate}$ is the number of generated events for i-th process with a primary vertex within 60cm from center of CDF detector. And N_{select} is the number of events which pass all event selection cuts.

ϵ_{trig}: The trigger efficiency for Lepton + Track trigger.

ϵ_{ID}: The e/μ and hadronic τ identification efficiency. In e/μ + τ_{had} channel, we assign $\epsilon_{e/muID} \times \epsilon_{\tau ID}$ as ϵ_{ID} and in e + μ channel, we assign $\epsilon_{eID} \times \epsilon_{\mu ID}$ as ϵ_{ID}.

ϵ_{vtx}: The efficiency of the z vertex position cut.

$\int Ldt$: The integrated luminosity for each trigger type, which described in Section 4.1.2, $\sim 9.6 fb^{-1}$ data is used in this analysis.

This MC based estimation is used to estimate the number of background events: $Z/\gamma*$+jets, $t\bar{t}$ and Diboson(WW/WZ/ZZ). The cross sections for $t\bar{t}(M_{top} = 172.5 GeV/c^2)$, WW , WZ and ZZ are 7.04pb, 10.92pb , 3.57pb and 3.25pb, respectively. In Drell-Yan process, we use 18 samples for each $Z \rightarrow ee/\mu\mu$ and 15 samples for $Z \rightarrow \tau\tau$, that these samples have different mass range and different parton multiplicity. Therefore, we have to weight these cross sections with Data/MC scale factor(K-factor), which described in Section 5.2.1. The systematic uncertainties related to this calculation is summarized in Section 8.

5.2 Drell-Yan estimation

In this section, we describe how to estimate Drell-Yan events, especially about data/MC scale factor and small correction for $e/\mu \rightarrow \tau_{had}$.

5.2.1 Data/MC scale factor for Drell-Yan

For the estimation of Drell-Yan process, we use ALPGEN Monte Calro sample [26], which can generate each parton multiplicity. Therefore, we need to add together these samples with proper theoretical cross section. However we can see good agreements in kinematic shapes between data and MC, different event rate is observed (the number of events in data is always larger than in MC) because ALPGEN reproduces only the event of Leading Order(LO) level. Data events include the events of more higher orders, that most effective order is Next to Leading Order(NLO).

In order to deal with this discrepancies, other analyses in CDF apply NLO scale factor(1.4 ± 0.2) to MC. But this analysis needs to check this scale factor for each jet multiplicity beacuse jet definition in this analysis is relatively looser than other analyses and high instantaneous luminosity data may affect to this scale factor.

Jet definition in this analysis: $E_T \geq 15GeV$, $|\eta| \leq 2.5$

Standard jet definition in CDF: $E_T \geq 20GeV$, $|\eta| \leq 2.0$

In order to check Data/MC scale factor, the Drell-Yan control region is evaluated. We use events which collected with "High P_T Lepton" trigger (ELECTRON_CENTRAL_18, MUON_CMUP18, MUON_CMX18). This trigger requires one electron/muon with $P_T \geq 18GeV/c$. Thw events passing this trigger are required opposite charged electrons(muons) with $E_T(P_T) \geq 20GeV\,(GeV/c)$ and invariant mass of dd-lepton is consistent with Z boson mass region ($70 \leq M_Z \leq 110GeV/c^2$).

The identifications of electron and muon are same as 3.1 and 3.2 except for $E_T(P_T)$ cut (electron:$E_T \geq 20GeV$, muon:$P_T \geq 20GeV/c$). Drell-Yan control region definition is summarized in Table 5.1. QCD fake background($jet \rightarrow e/\mu$) and other physics backgrounds are subtructed, but these are negligible after these selection. Thus, we directly compare data with MC to evaluate Drell-Yan data/MC scale factor.

Drell-Yan Control Region	
Luminosity	$9.1fb^{-1}$
Trigger	High P_T Lepton
MC sample	$Z \rightarrow ee/\mu\mu$ (ALPGEN)
Event Selection	2 electron/muon with $P_T \geq 20GeV/c$
	Opposite Charge ($Q_e \times Q_e = -1$)
	$70 \leq M(dilepton) \leq 110GeV/c^2$

Table 5.1: Drell-Yan control region definition to check data/MC scale factor.

We expect the number of event for data and MC after event selection cuts and calculate data/MC scale factor, this result is shown in Table 5.2. The invariant mass of dilepton distributions of dilepton for each jet bin are shown in Figure 5.1 and Figure 5.2 which are before and after applying these factors.

The data/MC scale factors which we evaluate are consistent with the NLO scale factor(1.4±0.2), but these fractions have differential behavior for each jet bin. Therefore, we decide to apply these differential scale factors for each jet bin instead of 1.4±0.2. The systematic error for these factors are assigned as 14.8%, which is same as the error of the NLO scale factor.

	= 0jet	= 1jet	≥ 2jet	Total
data	188642	39510	10626	238778
MC	149159.0	28018.7	6640.4	183818
data/MC	1.27	1.41	1.60	1.32

Table 5.2: data/MC scale factor for Drell-Yan process

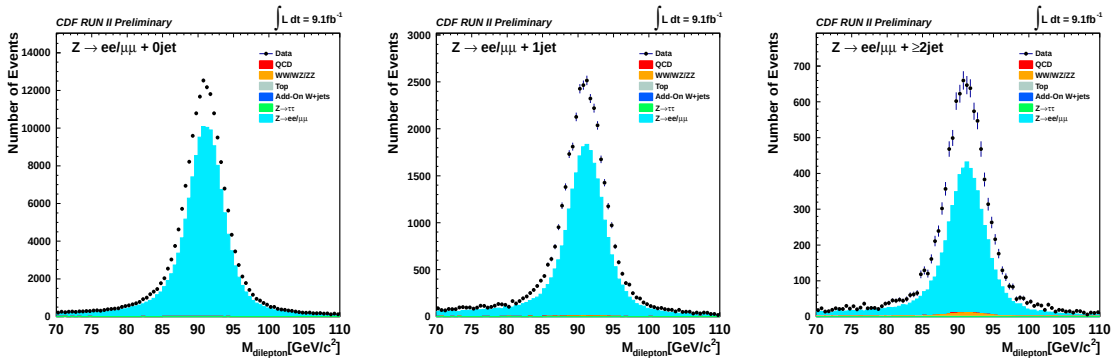

Figure 5.1: The invariant mass of dilepton distributions BEFORE applying data/MC scale factors. From left: 0jet bin, 1jet bin and ≥2jet bin distributions.

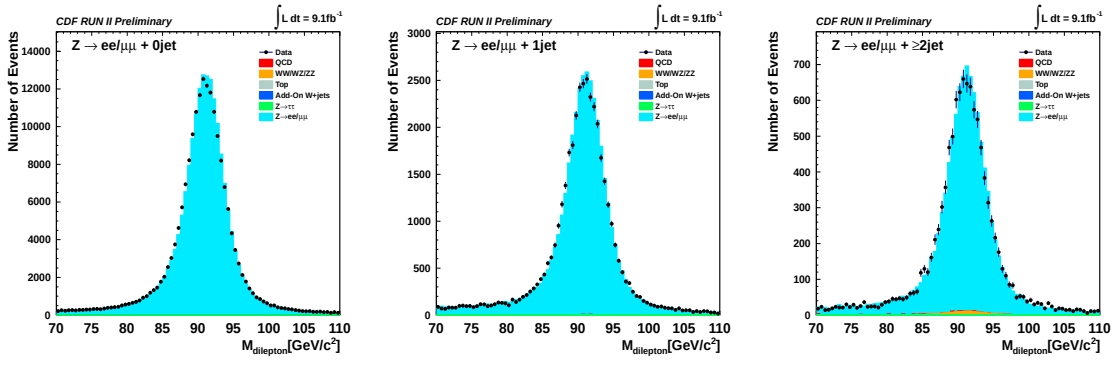

Figure 5.2: The invariant mass of dilepton distributions AFTER applying data/MC scale factors. From left: 0jet bin, 1jet bin and ≥2jet bin distributions.

5.2.2 Small Correction for $e/\mu \rightarrow \tau_{had}$

The $Z \rightarrow ee/\mu\mu$+jets event have two fake components to signal events in e/μ + τ_{had} channel. One is $jet \rightarrow \tau_{had}$ fake where jets are produced by initial state radiation. In this case, we expect same

ratio for opposite sign and same sign event between e/μ and τ_{had} which is faked by jet. Therefore, we estimate $jet \rightarrow tau_{had}$ fake event by same sign data, more detail described in Section 5.3.1. Another one is e/$\mu \rightarrow \tau_{had}$ fake where e/μ comes from Z boson. We expect that all e/μ and τ_{had} which is faked by e/μ enter opposite sign category and estimate the number of this fake events by ALPGEN MC. In this situation, we have to avoid double counting of e/μ fake and jet fake in MC sample. We check this two fakes by the generator level information: matching between generator level e/μ and reconstructed hadronic τ. If ΔR(lepton^{gen},τ_{had}^{reco}) < 0.2, we assign this events as e/$\mu \rightarrow \tau_{had}$, and if it failed, jet$\rightarrow \tau_{had}$. We just remove opposite sign events if tagged as jet$\rightarrow \tau_{had}$ in MC sample to avoid double counting between MC prediction and same sign data.

5.3 jet $\rightarrow \tau_{had}$/lepton estimation

The main sources of jet$\rightarrow \tau_{had}$ fake event are QCD,γ+jets and W+jets.

QCD: If one jet reconstructed as hadronic τ and another one fakes electron or muon in e/μ + τ_{had} channel and if one jet fakes electron and another fakes muon, these final state is same as signal event. However jet$\rightarrow$e/μ fake rate is not quite high, QCD process have high order cross section ($\sim$Ab). Therefore, QCD process is most significant background of jet fake event.

γ+jets: This process almost relate to electron + hadronic τ channel. If one jet reconstructed as hadronic τ and one electron is observed from γ conversion, these event have possibility to pass event selection. In electron + muon channel, event rate is small considering production cross section and jet$\rightarrow \mu$ fake rate.

W+jets: If W boson decay leptonically and one jet reconstructed as hadronic τ or e/μ, it represent final state of signal event.

In order to estimate the event rates and the kinematic shapes, we use data driven estimation method, which is same sign data method ($Q_{e/\mu} \times Q_{\tau_{had}/\mu} = 1$). The feynman diagrams of QCD and W+jets processes are shown in Figure 5.3.

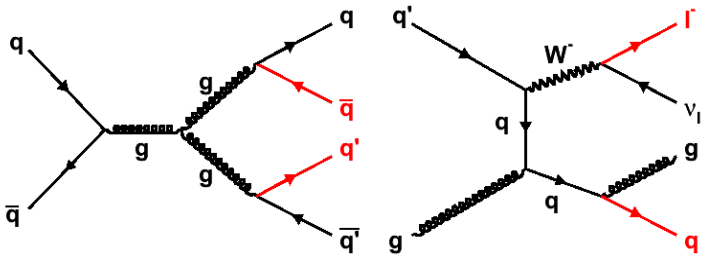

Figure 5.3: Feynman diagram of QCD and W+jets process.

5.3.1 Same Sign Data Method

Since these jet fake events are hard to represent with Monte Carlo simulation, we directly use same sign data($Q_{lep1} \times Q_{lep2} = 1$) to estimate and model jet fake events, where $lep1$ and $lep2$ represent e/μ and τ_{had} or e and μ, respectively. In the same sign data estimation, these is one assumption for QCD and γ+jets processes: in final state of these processes, there is no charge correlation which is expected between the charge of reconstructed e/μ,τ_{had} or e,μ. In another way, the number of

opposite sign(OS) events should be equal to the number of same sign(SS) events. In order to check the assumption of $N_{OS} = N_{SS}$ in QCD, we compare isolation fraction of e/μ in the opposite sign evens and the same sign events.

QCD Control Region	
Luminosity	$9.6fb^{-1}$
Trigger	Lepton + Track
Event Selection	1 electron/muon with $P_T \geq 10GeV/c$ w/o isolation cut
	1 hadronic τ with $P_T \geq 15(20)GeV/c$
	Drell-Yan veto
	Missing $E_T \leq 10GeV$

Table 5.3: QCD control region to check lepton isolation fraction of opposite sign and same sign.

The QCD control region is defined as Table 5.3 for this comparison: At first, we require one e or μ without isolation cut and one hadronic τ. Also the Drell-Yan veto and Missing E_T cut are applied to maximize QCD compornent and to suppress the contribution of other physics processes. The result of isolation fraction comparison is shown in Figure 5.4, where the result in opposite sign and same sign events are show by blue and red line. The ratios of isolation of opposite sign and same sign for each jet bin are shown in 5.5, where the ratio of 0/1/≥2jet bin is shown by black, red and blue, respectively. And also we compare some kinematic distributions of lepton isolation ≥ 0.1 in QCD control region. These distributions are shown in Figure 5.6 and 5.7. We can see almost same fraction of opposite sign and same sign in lepton isolation ≥ 0.1. Therefore, the same sign data is suitable to estimate the event rate of opposite sign and model kinematic shapes.

5.3.2 additional W+jets estimation

In W+jets process, if W boson decays into e or μ, and jets are reconstructed as hadronic τ or e/μ, these events represent same final state as signal event in e/μ + τ_{had} channel and in e + μ channel. However the W+jets events contributions are already estimated by same sign data, this process has a charge correlation in e/μ from W boson and outgoing quark which have $\pm\frac{1}{3}$ or $\pm\frac{2}{3}$ as its charge. There are more opposite sign events than same sign events($N_{OS} \geq N_{SS}$) in this process.

Therefore, we should not directly use same sign data for W+jets estimation without estimating additional contribution of this background. We have to estimate $N_{add-on} = N_{OS} - N_{SS}$ and data/MC scale factor(K-Factor). In order to perform this estimation, we evaluate W+jets control region like Table 5.4. The data events collected by Lepton + Track trigger with one e/μ and one loose τ are used for this estimation. The loose τ category is defined to increase jet fake contribution as following τ_{had} identification cuts: seed track $P_T \geq 4.5GeV/c$, visible $P_T \geq 15.0GeV/c$, visible Mass ≤ 5.0, and other cuts are same as Table 3.3. Since τ_{had} identification is loosened, we expect to have large QCD and W+jets fake events. Then in order to maximize W+jets contribution, the Drell-Yan veto, Missing E_T and transverse mass of lepton and $\not{E}_T$are applied following Table 5.4. These cuts almost kill QCD fake events and other physics backgrounds are negligible.

At first, we evaluate data/MC scale factor for opposite sign events and same sign events. Since this W+jets control region is not significant suffered by other processes, we obtain data/MC scale factor by following formula:

$$SF = \frac{N_{data} - N_{other}}{N_{w+jets}}, \tag{5.3}$$

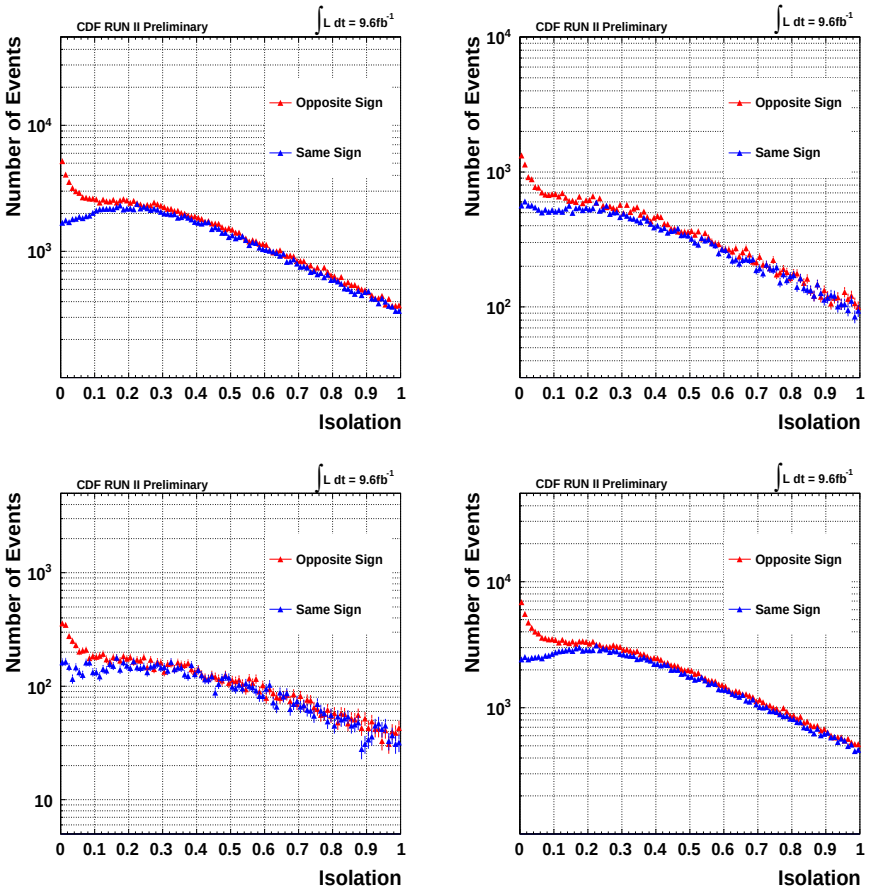

Figure 5.4: The lepton isolation fraction fo each jet bin.

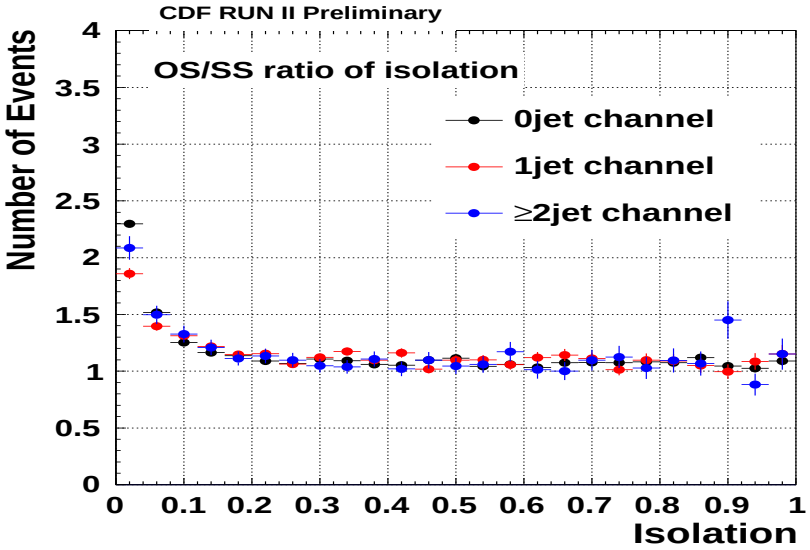

Figure 5.5: The lepton isolation ratio of opposite sign per same sign.

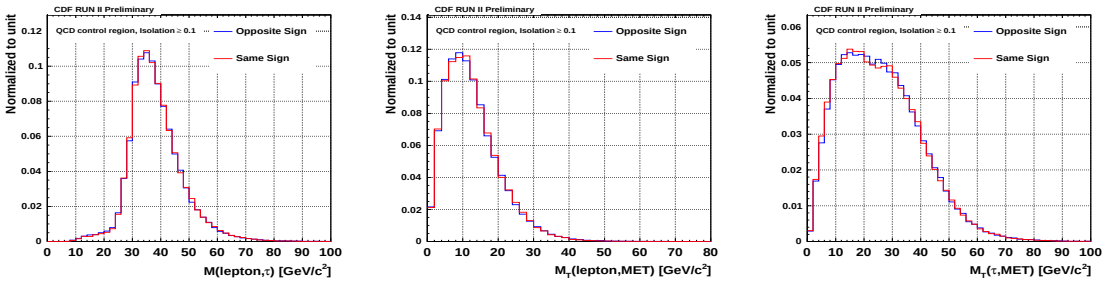

Figure 5.6: The kinematic distributions of isolation ≥ 0.1 in QCD control region: the invariant mass of lepton and τ_{had}, the transverse mass of lepton and $\not\!\!E_T$ and the transverse mass of τ_{had} and $\not\!\!E_T$.

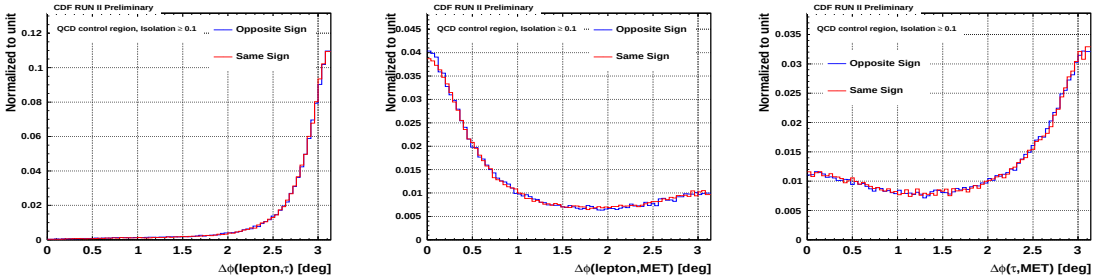

Figure 5.7: The kinematic distributions of isolation ≥ 0.1 in QCD control region: the $\Delta\phi$ of lepton and τ_{had}, the $\Delta\phi$ of lepton and $\not\!\!E_T$ and the $\Delta\phi$ of τ_{had} and $\not\!\!E_T$.

where N_{data} is the number of data and N_{w+jets} is the expected number of W+jets events and N_{other} is sum of the number of events which is come from $t\bar{t}$, Drell-Yan and Diboson processes. From this calculation, we obtain 1.32 in opposite sign events and 1.33 in same sign events as data/MC scale factor for W+jets events. These values are both consistent with NLO data/MC scale factor(1.4 ± 0.2). The several kinematic distributions of W+jet control region before and after data/MC scale factor in opposite sign and same sign events are shown in Figure 5.8 and Figure 5.9. Next, we try to evaluate OS/SS asymmetry scale factor(A_{sym}) to estimate additional contribution of W+jets events:

$$A_{sym} = N_{OS}/N_{SS}, \tag{5.4}$$

where N_{OS} is the number of opposite sign data (or W+jets event) and N_{SS} is the number of same sign data (or W+jets event). We obtain A_{sym} in data and in MC (to check) by this formula. The OS/SS scale factor and the ratio of data and MC are shown in Figure 5.10. In this figure, the scale factors in data and ALPGEN MC are shown in brack and red line, respectively, and the ratio of data/MC is shown in blue line.

We can see good agreement for charge correlation of W+jets event in data and MC. Finally, we estimate the number of additional W+jets event(N_{add-on}) by calculating following formula;

$$N_{add-on} = \frac{N_{OS}(1 - A_{sym})}{A_{sym}}, \tag{5.5}$$

where N_{OS} is the expected number of W+jets events by MC which is after data/MC scale factor is applied. In this procedure, systematic uncertainty for N_{add-on} is only comes from data/MC scale factor because we directly obtain A_{aim} from data. We assign NLO data/MC scale factor error (1.4 ± 0.2) as systematic uncertainty.

W+jets Control Region	
Luminosity	$9.6 fb^{-1}$
Trigger	Lepton + Track
Event Selection	1 e/μ with $P_T \geq 10 GeV/c$
	1 loose hadronic τ with $P_T \geq 15 GeV/c$
	Drell-Yan veto
	Missing E_T $\geq$ 25GeV for =0jet
	$\geq$ 30GeV for =1jet
	$\geq$ 35GeV for $\geq$2jet
	M_T(e/μ,$\not\!\!E_T$) $\geq 40 GeV$

Table 5.4: W+jets control region to estimate additional W+jets contribution.

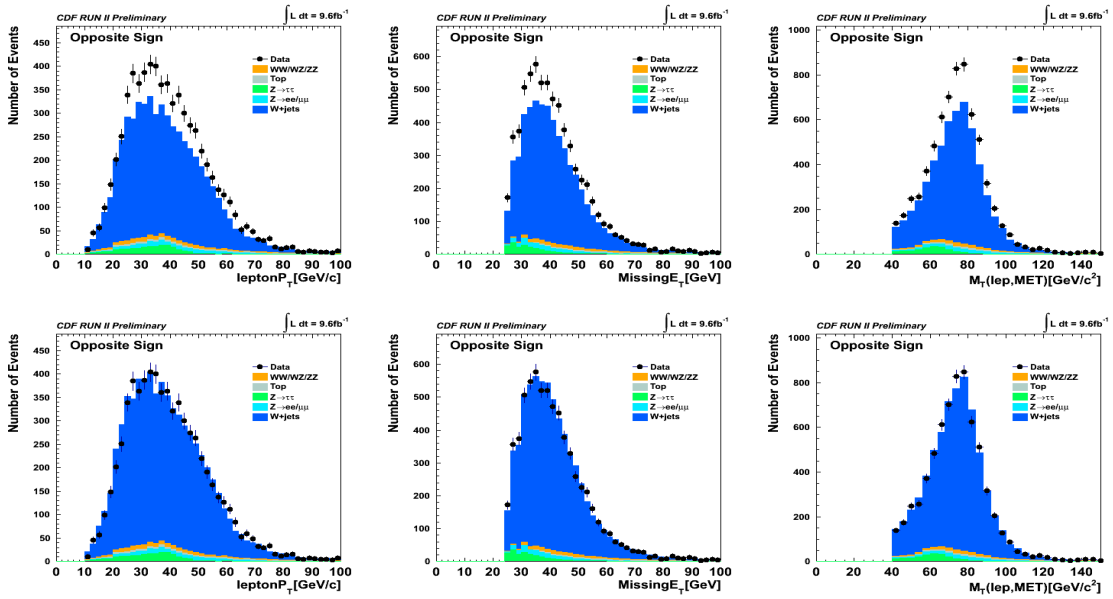

Figure 5.8: The kinematical distributions in W+jets control region of opposite sign events: lepton P_T, Missing E_T and the transverse mass of lepton and $\not\!\!E_T$.

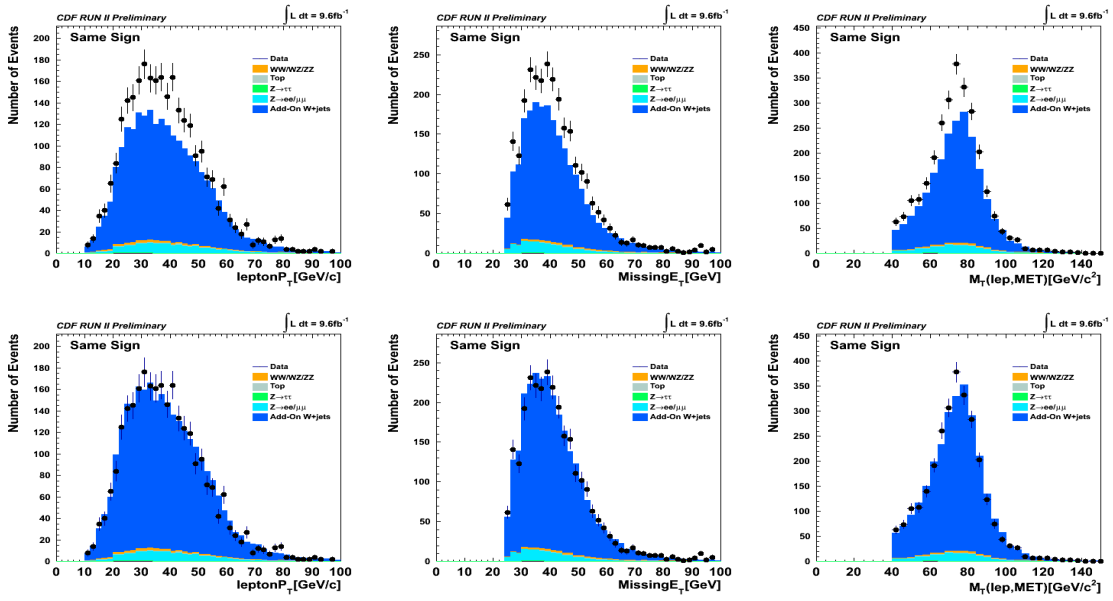

Figure 5.9: The kinematical distributions in W+jets control region of same sign events: lepton P_T, Missing E_T and the transverse mass of lepton and $\not\!\!E_T$.

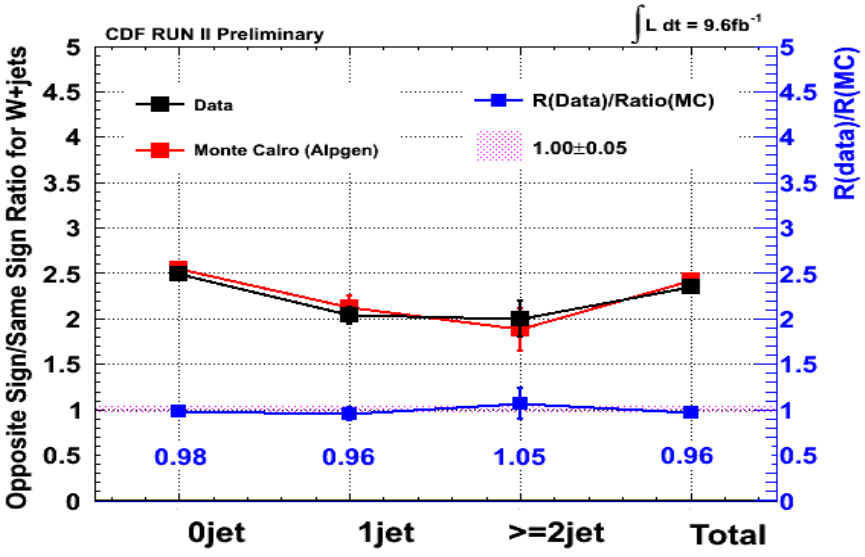

Figure 5.10: The ratio of the number of opposite sign event and same sign event in data and MC as a function of the number of jet.

Chapter 6

Event Expectation

This chapter describes the expectation of the number of events for each signal and background process.

6.1 Background Expectation

After all estimation and correction, the number of background expectation of each process for e/μ + τ_{had} channel and e + μ channel is shown in Table 6.1 and 6.2. The numbers of observed and predicted events as a function of the number of jets are shown in Figure 6.1, the left and right plots are for e/μ + τ_{had} channel and e + μ channel, respectively.

The kinematic distributions for each channel are summarized in Appendix.

H $\rightarrow$ $\tau\tau$ (e/μ + τ_{had} channel)			
Source	0 jet	1 jet	$\geq$2 jet
$Z \rightarrow \tau\tau$	14187.1 ± 2440.2	2663.9 ± 522.1	685.3 ± 174.8
$Z \rightarrow ee/\mu\mu$	823.3 ± 143.3	204.7 ± 38.9	56.8 ± 15.9
$t\bar{t}$	0.5 ± 0.1	14.0 ± 2.7	158.3 ± 1.0
WW/WZ/ZZ	123.1 ± 11.0	109.2 ± 11.1	51.8 ± 7.7
jet$\rightarrow \tau_{had}$	21704.0 ± 147.3	6510.0 ± 80.7	1742.0 ± 41.7
Add-on. W+jets	2522.6 ± 383.4	842.3 ± 136.5	178.2 ± 43.3
Total Background	39360.7 ± 2478.7	10344.1 ± 547.2	2872.4 ± 186.8
Observed ($9.6fb^{-1}$)	39311	10365	2947

Table 6.1: The number of background prediction and observed data for e/μ + τ_{had} channel with 9.6fb^{-1}.

6.2 Signal Expectation

The production cross section and branching ratio of the Higgs boson for each mass point are summarized in Table 6.3 [7] [8]. We calculate signal acceptances by applying event selection cuts and expect the number of events with $9.6fb^{-1}$ for e/μ + τ_{had} channel and for e + μ channel,

H $\rightarrow$ $\tau\tau$ (e + μ channel)			
Source	0 jet	1 jet	$\geq$2 jet
$Z \rightarrow \tau\tau$	1576.8 ± 269.6	302.9 ± 57.9	81.4 ± 20.6
$Z \rightarrow ee/\mu\mu$	90.2 ± 15.9	29.3 ± 5.2	12.4 ± 2.8
$t\bar{t}$	0.8 ± 0.2	19.5 ± 3.5	143.8 ± 18.3
WW/WZ/ZZ	176.7 ± 18.0	47.5 ± 5.6	15.1 ± 2.6
jet$\rightarrow e/\mu$	212.0 ± 14.6	57.0 ± 7.5	23.0 ± 4.8
Add-on. W+jets	39.8 ± 5.9	7.9 ± 1.2	0.9 ± 0.2
Total Background	2096.3 ± 271.2	464.1 ± 59.0	276.6 ± 28.2
Observed $(9.6 fb^{-1})$	2040	442	292

Table 6.2: The number of background prediction and observed data for e + μ channel with 9.6fb^{-1}.

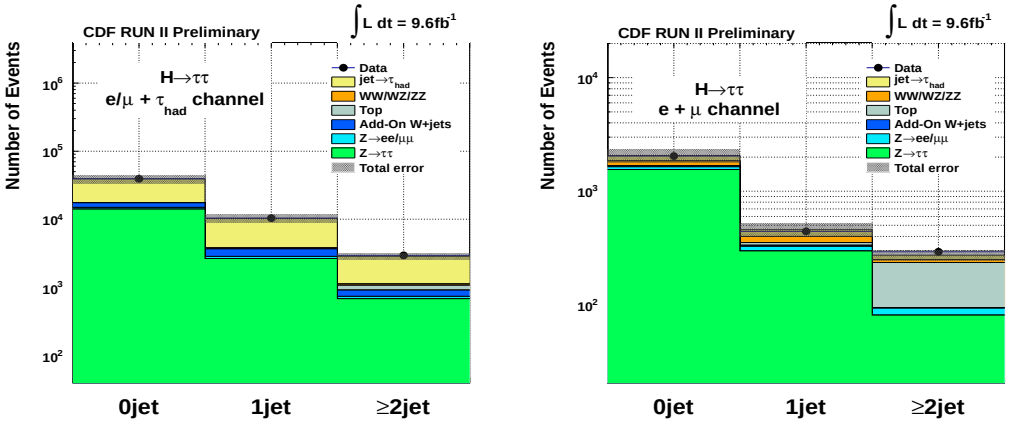

Figure 6.1: The numbers of observed and predicted events as a function of number of jets.

separately. The number of expected events for each signal process is summarized in Table 6.4, Table 6.5, Table 6.6 and Table 6.7, and the number of total signal events for each signal channel is summarized in Table 6.8. Also, the number of expected signal events as a function of the Higgs boson mass for each signal process and the number of total events is shown in Figure 6.2, the left and right plots are for e/μ + τ_{had} channel and e + μ channel, respectively.

Since we expect 18.54 events and 2.33 events at $M_H = 115 GeV/c^2$ in e/μ + τ_{had} channel and in e + μ channel, the total number of signal events is 20.87 at $M_H = 115 GeV/c^2$ in this analysis.

$M_H(GeV/c^2)$	σ (pb)				B.R.(%)
	WH	ZH	VBF	ggH	
100	291.9	169.8	100.1	1821.8	7.92
105	248.4	145.9	92.3	1584.7	7.82
110	212.0	125.7	85.1	1385.0	7.62
115	174.5	103.9	78.6	1215.9	7.29
120	150.1	90.2	72.7	1072.3	6.79
125	129.5	78.5	67.1	949.3	6.12
130	112.0	68.5	62.1	842.9	5.31
135	97.2	60.0	57.5	750.8	4.40
140	84.6	52.7	53.2	670.6	3.47
145	73.7	46.3	49.4	600.6	2.59
150	64.4	40.8	45.8	539.1	1.78

Table 6.3: Production cross section for the Higgs signal and branching ratio of H$\longrightarrow \tau\tau$.

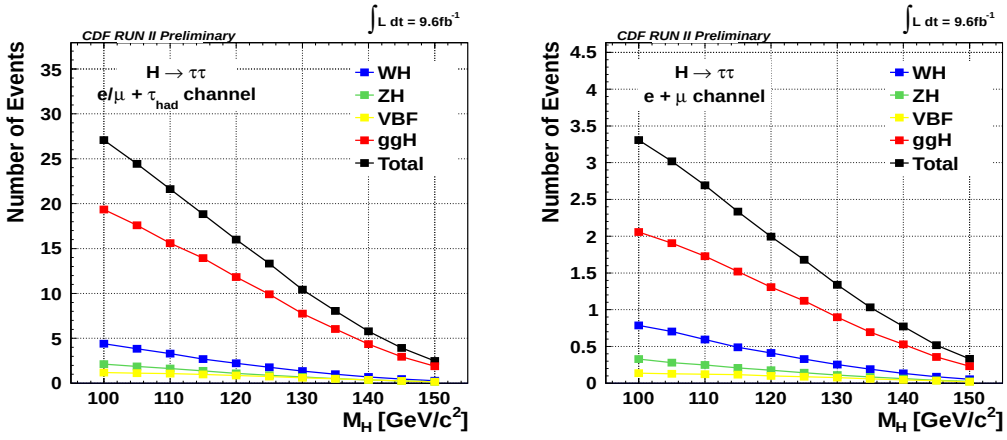

Figure 6.2: The number of expected signal events as a function of the Higgs boson mass.

WH	Number of expected signal event					
	e/μ + τ_{had}			e + μ		
$M_H(GeV/c^2)$	= 0jet	= 1jet	$\geq$ 2jet	= 0jet	= 1jet	$\geq$ 2jet
100	0.79	1.48	1.93	0.24	0.28	0.26
105	0.69	1.30	1.64	0.20	0.26	0.24
110	0.54	1.09	1.45	0.17	0.22	0.21
115	0.46	0.90	1.22	0.14	0.18	0.17
120	0.37	0.77	1.02	0.11	0.15	0.15
125	0.29	0.61	0.82	0.09	0.12	0.12
130	0.22	0.47	0.63	0.07	0.09	0.09
135	0.15	0.34	0.46	0.05	0.07	0.07
140	0.10	0.24	0.33	0.04	0.05	0.05
145	0.07	0.16	0.22	0.02	0.03	0.03
150	0.04	0.10	0.14	0.01	0.02	0.02

Table 6.4: The number of expected events for WH process at each Higgs mass point with the number of jet in final state.

ZH	Number of expected signal events					
	e/μ + τ_{had}			e + μ		
$M_H(GeV/c^2)$	= 0jet	= 1jet	$\geq$ 2jet	= 0jet	= 1jet	$\geq$ 2jet
100	0.41	0.42	1.15	0.07	0.09	0.17
105	0.37	0.37	1.01	0.06	0.08	0.14
110	0.31	0.33	0.89	0.05	0.07	0.13
115	0.26	0.27	0.76	0.04	0.06	0.11
120	0.21	0.22	0.62	0.04	0.05	0.09
125	0.17	0.17	0.50	0.03	0.04	0.07
130	0.13	0.14	0.40	0.02	0.03	0.06
135	0.10	0.10	0.30	0.02	0.02	0.04
140	0.07	0.07	0.21	0.01	0.02	0.03
145	0.05	0.05	0.14	0.01	0.01	0.02
150	0.03	0.03	0.09	0.01	0.01	0.01

Table 6.5: The number of expected events for ZH process at each Higgs mass point with the number of jet in final state.

VBF	Number of expected signal events					
	e/μ + τ_{had}			e + μ		
$M_H(GeV/c^2)$	= 0jet	= 1jet	$\geq$ 2jet	= 0jet	= 1jet	$\geq$ 2jet
100	0.05	0.33	0.79	0.01	0.04	0.09
105	0.05	0.32	0.75	0.01	0.03	0.09
110	0.05	0.30	0.70	0.00	0.03	0.08
115	0.04	0.28	0.64	0.01	0.03	0.08
120	0.04	0.24	0.57	0.00	0.03	0.07
125	0.03	0.21	0.49	0.00	0.02	0.06
130	0.03	0.17	0.40	0.00	0.02	0.05
135	0.02	0.14	0.32	0.00	0.02	0.04
140	0.02	0.10	0.24	0.00	0.01	0.03
145	0.01	0.07	0.17	0.00	0.01	0.02
150	0.01	0.05	0.11	0.00	0.01	0.01

Table 6.6: The number of expected events for VBF process at each Higgs mass point with the number of jet in final state.

ggH	Number of expected signal events					
	e/μ + τ_{had}			e + μ		
$M_H(GeV/c^2)$	= 0jet	= 1jet	$\geq$ 2jet	= 0jet	= 1jet	$\geq$ 2jet
100	12.69	4.76	1.67	1.33	0.52	0.20
105	11.59	4.46	1.49	1.21	0.52	0.18
110	10.01	4.04	1.36	1.10	0.47	0.16
115	8.92	3.57	1.20	0.98	0.40	0.14
120	7.40	3.09	1.18	0.83	0.36	0.11
125	6.19	2.60	0.95	0.70	0.30	0.12
130	4.71	2.06	0.77	0.56	0.24	0.10
135	3.61	1.63	0.58	0.42	0.20	0.08
140	2.60	1.18	0.44	0.33	0.14	0.06
145	1.74	0.82	0.30	0.21	0.11	0.04
150	1.10	0.51	0.20	0.14	0.07	0.03

Table 6.7: The number of expected events for ggH process at each Higgs mass point with the number of jet in final state.

Total	Number of expected signal events					
	e/μ + τ_{had}			e + μ		
$M_H(GeV/c^2)$	= 0jet	= 1jet	$\geq$ 2jet	= 0jet	= 1jet	$\geq$ 2jet
100	13.94	6.99	5.53	1.65	0.93	0.72
105	12.69	6.45	4.89	1.49	0.89	0.64
110	10.92	5.75	4.40	1.33	0.79	0.58
115	9.69	5.02	3.83	1.17	0.67	0.49
120	8.02	4.32	3.39	0.99	0.59	0.42
125	6.69	3.60	2.76	0.83	0.48	0.37
130	5.09	2.84	2.20	0.65	0.39	0.30
135	3.88	2.22	1.65	0.49	0.31	0.23
140	2.79	1.60	1.21	0.38	0.22	0.17
145	1.86	1.10	0.83	0.24	0.16	0.11
150	1.18	0.68	0.53	0.16	0.10	0.08

Table 6.8: The total number of expected signal events at each Higgs mass point with the number of jet in final state.

Chapter 7

Analysis Optimization

We expect that there are the large number of background because event selection in this analysis is the minimum requirement to maximize the number of signal events. However, we have many kinematic variables to discriminate signal from background. The discriminant analysis enhances the search sensitivity in this analysis. We choose the Support Vector Machine(SVM) [27] which is one of multivariate analysis for discriminant analysis. In this chapter, we describe input variables for SVM, training result and how to create final discriminant.

7.1 Support Vector Machine (SVM)

The Support Vector Machine is one of the method of multivariate analysis. The concept of SVM is very simple, that is decision plane to separate signal and background. The SVM arranges each event(vector) in hyperspace which have dimension of the same number as input variable and pulls hyperplane where margin becomes maximum between hyperplane and support vector, where support vector is the most closest vector to boundary.

Since it is difficult to separate signal and background perfectly, we use the soft margin method that allow some vectors to misclassified for maximizing margin. In this case, it is desired to minimize the sum of the distance between misclassified vectors and hyperplane. Therefore, we need to optimize cost parameter(C), which decide about which allow to misclassified.

For the non-linear case, the kernel function is applied to hyperplane in SVM. However there are some kind of kernel function(linear, polynomial, sigmoid and so on), we use Gaussian function as kernel function like following formula:

$$K(x_i, x_j) = exp(-\gamma||x_i - x_j||^2), \gamma > 0, \tag{7.1}$$

where γ is the width of Gaussian function and this is the kernel parameter which is optimized. The Gaussian kernel allow one to apply any discriminating shape in the input space.

Thus, we use the SVM to discriminate Higgs signal from background by optimizing cost parameter C which is comes from the soft-margin method and kernel parameter γ which is comes from kernel function. The example of non-linear mapping of the SVM is shown in Figure 7.1.

7.2 Strategy of Discriminant Analysis

We define 6 signal channels because these background components are respectively different.

e/μ + τ_{had} + 0jet channel

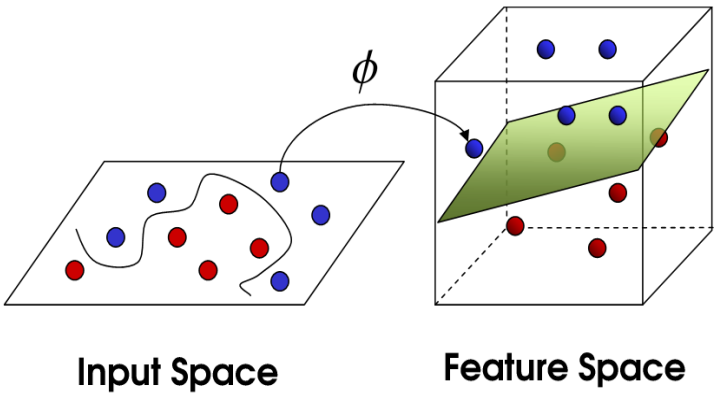

Figure 7.1: The example of non-linear mapping of the SVM.

e/μ + τ_{had} + 1jet channel
e/μ + τ_{had} + $\geq$ 2jet channel
e + μ + 0jet channel
e + μ + 1jet channel
e + μ + $\geq$ 2jet channel

We prepare four classifiers for each 0jet, 1jet and $\geq$2jet in e/μ + τ_{had} channel and two classifiers for 0jet, 1jet and $\geq$ 2jet in e + μ channel, and these classifiers are trained separately. As signal process, the H$\rightarrow \tau\tau$ process is defined, which is mixture of WH,ZH,VBF and ggH by the ratio of the number of expected events, and as background, one of the process is defined, which is dominant source or has different kinematics compared with signal. The definitions of each classifier are summarized in Table 7.1 and Table 7.2, and the feynman diagrams of background using training are shown in Figure 7.2. In Figure 7.2, the diagrams show from top left: Z$\rightarrow \tau\tau$, QCD and W+jets, from bottom left: $t\bar{t}$ and diboson(WZ).

H $\rightarrow \tau\tau$ (e/μ + τ_{had} channel)		
0 jet	1 jet	$\geq$ 2jet
$H \rightarrow \tau\tau$ vs. Z$\rightarrow \tau\tau$	$H \rightarrow \tau\tau$ vs. Z$\rightarrow \tau\tau$	$H \rightarrow \tau\tau$ vs. Z$\rightarrow \tau\tau$
$H \rightarrow \tau\tau$ vs. QCD	$H \rightarrow \tau\tau$ vs. QCD	$H \rightarrow \tau\tau$ vs. QCD
$H \rightarrow \tau\tau$ vs. W+jets	$H \rightarrow \tau\tau$ vs. W+jets	$H \rightarrow \tau\tau$ vs. W+jets
$H \rightarrow \tau\tau$ vs. Diboson	$H \rightarrow \tau\tau$ vs. Diboson	$H \rightarrow \tau\tau$ vs. $t\bar{t}$

Table 7.1: The training category for each signal channel for e/μ + τ_{had} channel.

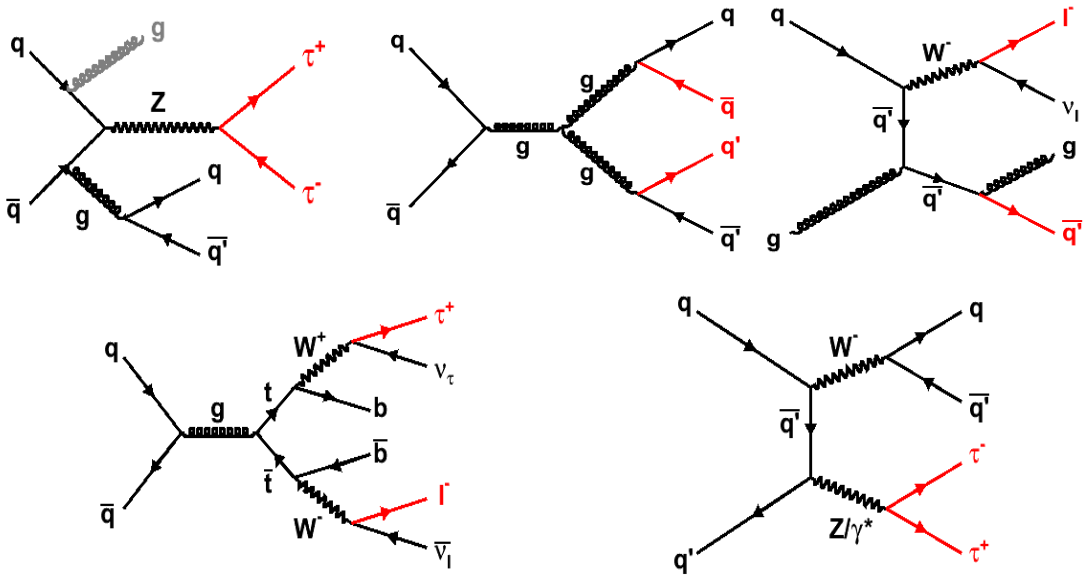

Figure 7.2: The example of feynman diagrams of each background for training.

H $\rightarrow$ $\tau\tau$ (e + μ channel)		
0 jet	1 jet	$\geq$ 2jet
$H \rightarrow \tau\tau$ vs. Z$\rightarrow \tau\tau$	$H \rightarrow \tau\tau$ vs. Z$\rightarrow \tau\tau$	$H \rightarrow \tau\tau$ vs. Z$\rightarrow \tau\tau$
$H \rightarrow \tau\tau$ vs. Diboson	$H \rightarrow \tau\tau$ vs. Diboson	$H \rightarrow \tau\tau$ vs. $t\bar{t}$

Table 7.2: The training category for each signal channel for e + μ channel.

7.3 Training

7.3.1 Training Samples

For the signal training samples, the MC samples of $H \rightarrow \tau\tau$ are used for each process: WH,ZH,VBF and ggH. Each classifier is trained by following Higgs mass points: $110GeV/c^2$, $115GeV/c^2$, $120GeV/c^2$, $130GeV/c^2$, $140GeV/c^2$, $150GeV/c^2$. Each signal sample are mixed by weighting the number of expected events at each mass point.

For the background training samples, the MC samples are used for $Z/\gamma* \rightarrow \tau\tau$, W+jets, $t\bar{t}$, and Diboson. There are 3 independent processes in diboson process: WW, WZ, ZZ. These processes are mixed by weighting the expected number of events for the diboson training sample. For the QCD training sample, real data events are used, which are in isolation side band region. This means lepton(electron or muon) isolation is more than 0.4, that these events are rejected by event selection, and they are almost QCD jet fake events.

7.3.2 Input Variables

Each classifier for a channel are trained by different input variables because these background contributions are different.

e/μ + τ_{had} channels

In this channels, there are 24 possible candidates of input variables as following:

e/μ P_T: The transverse momentum of e/μ (electron or muon).

e/μ η: The pseudo-rapidity of e/μ.

τ visible P_T: The transverse momentum of reconstructed hadronic τ, where visible means tracks or π^0s.

τ seed track P_T: The transverse momentum of track in signal cone.

visible Mass: The invariant mass of hadronic τ, where visible means tracks and π^0s.

1^{st} jet E_T: The transverse energy of leading jet, where leading means having highest energy.

2^{nd} jet E_T: The transverse energy of second jet. (only can use $\geq$ 2jet channel)

1^{st} jet η: The pseudo-rapidity of the leading jet.

2^{nd} jet η: The pseudo-rapidity of the second jet. (only can use $\geq$ 2jet channel)

$\not\!\!E_T$: The missing transverse energy.

H_T: $\sqrt{\sum E_T}$, where E_T is the sum of transverse energy of all objects: $\not\!\!E_T$, jets, e/μ, τ_{had}.

$\not\!\!E_T$significance: $\not\!\!E_T/\sqrt{\sum E_T}$.

$M_T(e/\mu, \not\!\!E_T)$: The transverse mass of e/μ and $\not\!\!E_T$.

$M_T(\tau, \not\!\!E_T)$: The transverse mass of hadronic τ and $\not\!\!E_T$.

$M(e/\mu, \tau)$: The invariant mass of e/μ and hadronic τ.

$M(e/\mu, \tau, \not{E}_T)$: The invariant mass of e/μ, hadronic τ and $\not{E}_T$after collinear approximation.

$M(jet1, jet2)$: The invariant mass of leading and second jet. (only can use $\geq$ 2jet channel)

$\Delta\phi(e/\mu, \not{E}_T)$: The angular between e/μ and $\not{E}_T$.

$\Delta\phi(\tau, \not{E}_T)$: The angular between hadronic τ and $\not{E}_T$.

$\Delta\phi(e/\mu, \tau)$: The angular between e/μ and hadronic τ.

$\Delta\phi(e/\mu, \tau, \not{E}_T)$: The angular between e/μ, hadronic τ and $\not{E}_T$.

$\Delta R(e/\mu, \not{E}_T)$: The distance in $\eta - \phi$ plane between e/μ and $\not{E}_T$.

$\Delta R(\tau, \not{E}_T)$: The distance in $\eta - \phi$ plane between hadronic τ and $\not{E}_T$.

$\Delta R(e/\mu, \tau)$: The distance in $\eta - \phi$ plane between e/μ and hadronic τ.

$\Delta R(e/\mu, \tau, \not{E}_T)$: The distance in $\eta - \phi$ plane between e/μ, hadronic τ and $\not{E}_T$.

e + μ channels

In this channels, there are 23 candidates of the input variables as following:

electron P_T: The transverse momentum of electron.

muon P_T: The transverse momentum of muon.

electron η: The pseudo-rapidity of electron.

muon η: The pseudo-rapidity of muon.

1^{st} jet E_T: The transverse energy of leading jet.

2^{nd} jet E_T: The transverse energy of leading jet. (only can use $\geq$ 2jet channel)

1^{st} jet η: The pseudo-rapidity of the leading jet.

2^{nd} jet η: The pseudo-rapidity of the second jet. (only can use $\geq$ 2jet channel)

$\not{E}_T$: The missing transverse energy.

H_T: $\sqrt{\sum E_T}$, where E_T is the sum of transverse energy of all objects: $\not{E}_T$, jets, electron, muon.

$\not{E}_T$significance: $\not{E}_T/\sqrt{\sum E_T}$.

$M_T(e, \not{E}_T)$: The transverse mass of electron and $\not{E}_T$.

$M_T(\mu, \not{E}_T)$: The transverse mass of muon and $\not{E}_T$.

$M(e, \mu)$: The invariant mass of electron and muon.

$M(e, \mu, \not{E}_T)$: The invariant mass of electron, muon and $\not{E}_T$after collinear approximation.

$M(jet1, jet2)$: The invariant mass of leading and second jet. (only can use $\geq$ 2jet channel)

$\Delta\phi(e, \not{E}_T)$: The angular between electron and $\not{E}_T$.

$\Delta\phi(\mu, \not{E}_T)$: The angular between muon and $\not{E}_T$.

$\Delta\phi(e, \mu)$: The angular between electron and muon.

$\Delta\phi(e, \mu, \not{E}_T)$: The angular between electron, muon and $\not{E}_T$.

$\Delta R(e, \not{E}_T)$: The distance in $\eta - \phi$ plane between electron and $\not{E}_T$.

$\Delta R(\mu, \not{E}_T)$: The distance in $\eta - \phi$ plane between muon and $\not{E}_T$.

$\Delta R(e, \mu)$: The distance in $\eta - \phi$ plane between electron and muon.

$\Delta R(e, \mu, \not{E}_T)$: The distance in $\eta - \phi$ plane between electron, muon and $\not{E}_T$.

At first, we check the separation power of each variable to choose input variables. Here, the definition of separation power is following formula:

$$S = \frac{1}{2} \int \frac{(f_S(x) - f_B(x))^2}{f_S(x) + f_B(x)}, \tag{7.2}$$

where S is separation power and $f_s(x)$ and $f_B(x)$ are the vertical axis value of signal and background at point x. Next, we iterate to input one variable at a time to the sequence whose separation power is high and to calculate separation power of training result. Input variables for each classifier are defined when separation power of training result is highest or goes to almost flat.

The input variables with these separation power ranking and the number of input variable for each classifier are shown in Table 7.3, Table 7.4, Table 7.5 and Table 7.6 for e/μ + τ_{had} + 0/1/$\geq$2jet and e + μ channel, respectively.

7.3.3 Training Results

The input variables for the training are already defined in Section 7.3.2. Each classifier is trained using these input variables and suitable SVM parameters. The performance plots of training result for each channel are shown in Figure 7.3, Figure 7.4,7.5,7.8,7.7 and 7.8. Each figure shows training result and test result for signal(blue) and background(red) with dot points which show response of test samples.

7.4 SVM Response Distributions

Since four or two classifiers for each channel are trained, the total eighteen training response distributions are obtained. They are shown in Figure 7.9, Figure 7.10, Figure 7.11, Figure 7.9, Figure 7.13 and Figure 7.14 for $e/\mu + \tau_{had} + 0/1/geq$2jet channel and $e + \mu + 0/1/geq$2jet channel, respectively In these figures, a red histogram shows signal ($H \rightarrow \tau\tau$) distribution which is scaled by 100 times, and other histograms show background distributions. In each distributions, the trained background distributions are well discriminated from signal distributions, but other several background distributions are not discriminated. Therefore, these SVM responses should be collected to one or more dimensional distribution as a final discriminant.

Input Variables for e/μ + τ_{had} + 0jet channel				
Variable	$Z \rightarrow \tau\tau$	QCD	W+jets	Diboson
1. e/μ P_T	4	3	-	-
2. e/μ η	-	-	-	-
3. τ_{had} visible P_T	2	-	5	-
4. τ_{had} seed track P_T	-	5	6	-
5. τ_{had} visible Mass	-	6	7	-
6. 1^{st} Jet E_T	-	-	-	-
7. 1^{st} Jet η	-	-	-	-
8. 2^{nd} Jet E_T	-	-	-	-
9. 2^{nd} Jet η	-	-	-	-
10.$\not\!\!E_T$	-	-	3	3
11.$\not\!\!E_T$Significance	-	-	-	-
12.H_T	1	2	-	4
13.$M_T(e/\mu, \not\!\!E_T)$	-	-	1	1
14.$M_T(\tau_{had}, \not\!\!E_T)$	6	4	-	-
15.$M(e/\mu, \tau_{had})$	3	-	-	-
16.$M(e/\mu, \tau_{had}, \not\!\!E_T)$	5	1	4	5
17.$M(1^{st}jet, 2^{nd}jet)$	-	-	-	-
18.$\Delta\phi(e/\mu, \not\!\!E_T)$	-	-	-	-
19.$\Delta\phi(\tau_{had}, \not\!\!E_T)$	-	8	8	6
20.$\Delta\phi(e/\mu, \tau_{had})$	-	-	2	2
21.$\Delta\phi(e/\mu, \tau_{had}, \not\!\!E_T)$	-	-	-	-
22.$\Delta R(e/\mu, \not\!\!E_T)$	-	-	-	-
23.$\Delta R(\tau_{had}, \not\!\!E_T)$	-	-	-	-
24.$\Delta R(e/\mu, \tau_{had})$	-	-	-	-
25.$\Delta R(e/\mu, \tau_{had}, \not\!\!E_T)$	-	7	-	7
N_{Input}	6	8	8	7

Table 7.3: The List of input variables for e/μ + τ_{had} + 0jet channel.

Input Variables for e/μ + τ_{had} + 1jet channel				
Variable	$Z \to \tau\tau$	$jet \to \tau_{had}$	W+jets	Diboson
1. e/μ P_T	-	3	-	-
2. e/μ η	-	-	-	-
3. τ visible P_T	5	6	6	-
4. τ seed track P_T	-	-	-	5
5. τ visible Mass	-	7	5	4
6. 1^{st} Jet E_T	8	-	-	-
7. 1^{st} Jet η	7	-	-	-
8. 2^{nd} Jet E_T	-	-	-	-
9. 2^{nd} Jet η	-	-	-	-
10.$\not\!\!E_T$	-	-	7	-
11.$\not\!\!E_T$Significance	-	-	-	-
12.H_T	1	2	8	-
13.$M_T(e/\mu, \not\!\!E_T)$	-	8	1	1
14.$M_T(\tau, \not\!\!E_T)$	4	4	9	-
15.$M(e/\mu, \tau)$	2	1	-	8
16.$M(e/\mu, \tau, \not\!\!E_T)$	3	5	4	6
17.$M(1^{st}jet, 2^{nd}jet)$	-	-	-	-
18.$\Delta\phi(e/\mu, \not\!\!E_T)$	-	-	-	2
19.$\Delta\phi(\tau, \not\!\!E_T)$	-	-	-	-
20.$\Delta\phi(e/\mu, \tau)$	-	-	2	3
21.$\Delta\phi(e/\mu, \tau, \not\!\!E_T)$	-	-	-	-
22.$\Delta R(e/\mu, \not\!\!E_T)$	-	-	-	-
23.$\Delta R(\tau, \not\!\!E_T)$	-	-	-	-
24.$\Delta R(e/\mu, \tau)$	-	-	-	-
25.$\Delta R(e/\mu, \tau, \not\!\!E_T)$	6	-	3	7
N_{Input}	8	8	9	8

Table 7.4: The List of input variables for e/μ + τ_{had} + 1jet channel.

Input Variables for e/μ + τ_{had} + $\geq$ 2 jet channel				
Variable	$Z \rightarrow \tau\tau$	QCD	W+jets	$t\bar{t}$
1. e/μ P_T	7	4	7	4
2. e/μ η	-	-	-	-
3. τ visible P_T	5	7	-	-
4. τ seed track P_T	-	-	6	8
5. τ visible Mass	-	8	5	-
6. 1^{st} Jet E_T	-	-	9	-
7. 1^{st} Jet η	-	-	-	-
8. 2^{nd} Jet E_T	-	-	-	-
9. 2^{nd} Jet η	-	-	-	-
10.$\not\!\!E_T$	-	5	-	-
11.$\not\!\!E_T$Significance	-	-	-	-
12.H_T	2	3	-	3
13.$M_T(e/\mu, \not\!\!E_T)$	8	-	2	1
14.$M_T(\tau, \not\!\!E_T)$	6	6	10	5
15.$M(e/\mu, \tau)$	3	2	8	-
16.$M(e/\mu, \tau, \not\!\!E_T)$	1	1	1	2
17.$M(1^{st}jet, 2^{nd}jet)$	4	-	-	7
18.$\Delta\phi(e/\mu, \not\!\!E_T)$	-	-	3	-
19.$\Delta\phi(\tau, \not\!\!E_T)$	-	-	-	-
20.$\Delta\phi(e/\mu, \tau)$	-	-	-	9
21.$\Delta\phi(e/\mu, \tau, \not\!\!E_T)$	-	-	-	-
22.$\Delta R(e/\mu, \not\!\!E_T)$	-	-	4	-
23.$\Delta R(\tau, \not\!\!E_T)$	-	-	-	-
24.$\Delta R(e/\mu, \tau)$	-	-	-	-
25.$\Delta R(e/\mu, \tau, \not\!\!E_T)$	-	-	-	6
N_{Input}	8	8	10	9

Table 7.5: The List of input variables for e/μ + τ_{had} + $\geq$2jet channel.

Input Variables for e + μ channel						
	$= 0jet$		$= 1jet$		$\geq 2jet$	
Variable	$Z \to \tau\tau$	Diboson	$Z \to \tau\tau$	Diboson	$Z \to \tau\tau$	$t\bar{t}$
1. electron P_T	-	-	9	6	-	8
2. electron η	7	-	-	-	7	-
3. μ P_T	2	-	5	8	3	9
4. μ η	-	-	-	-	-	-
5. 1^{st} Jet E_T	-	-	-	9	-	-
6. 1^{st} Jet η	-	-	-	-	-	-
7. 2^{nd} Jet E_T	-	-	-	-	-	-
8. 2^{nd} Jet η	-	-	-	-	-	-
9. $\not\!\!E_T$	-	-	-	7	-	-
10.$\not\!\!E_T$Significance	-	-	-	-	-	-
11.H_T	1	2	1	1	1	6
12.$M_T(e, \not\!\!E_T)$	4	-	8	5	5	3
13.$M_T(\mu, \not\!\!E_T)$	-	4	2	2	-	2
14.$M(e, \mu)$	-	-	4	3	-	5
15.$M(e, \mu, \not\!\!E_T)$	6	3	3	4	2	1
16.$M(1^{st}jet, 2^{nd}jet)$	-	-	-	-	-	-
17.$\Delta\phi(e, \not\!\!E_T)$	-	6	-	-	-	10
18.$\Delta\phi(\mu, \not\!\!E_T)$	-	5	-	-	-	7
19.$\Delta\phi(e, \mu)$	-	1	6	-	-	-
20.$\Delta\phi(e, \mu, \not\!\!E_T)$	-	-	-	-	-	-
21.$\Delta R(e, \not\!\!E_T)$	-	-	-	-	-	-
22.$\Delta R(\mu, \not\!\!E_T)$	-	-	-	-	-	-
23.$\Delta R(e, \mu)$	3	-	-	-	4	-
24.$\Delta R(e, \mu, \not\!\!E_T)$	5	-	7	-	6	4
N_{Input}	7	6	9	9	7	10

Table 7.6: The List of input variables for e + μ + 0/1/$\geq$2jet channel.

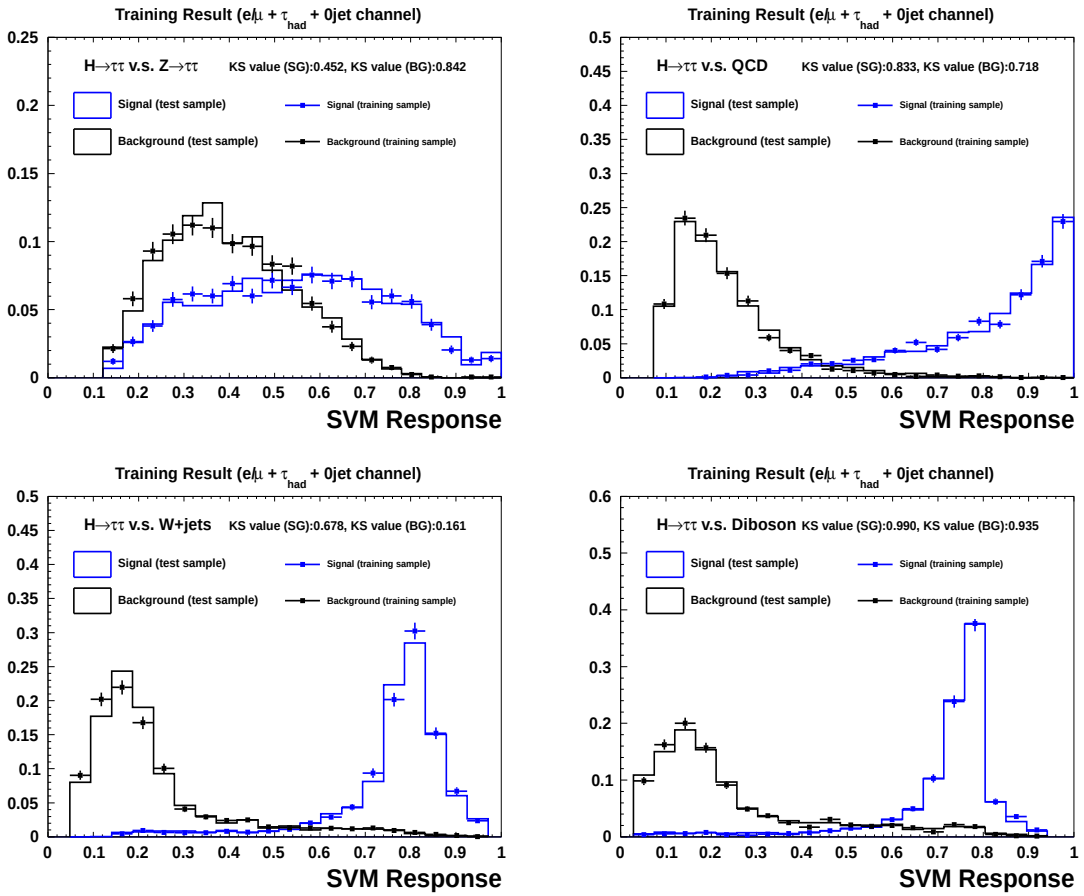

Figure 7.3: Training result of four classifiers for e/μ + 0jet channel.

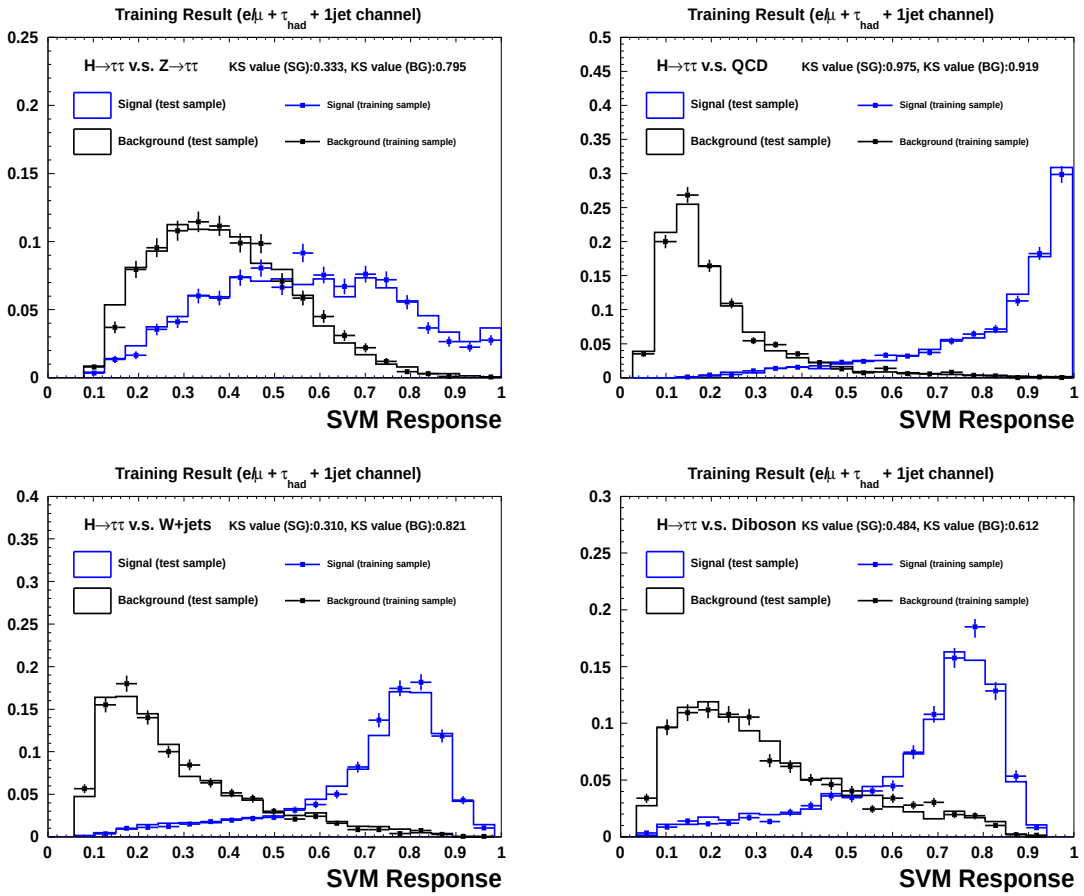

Figure 7.4: Training result of four classifiers for e/μ + 1jet channel.

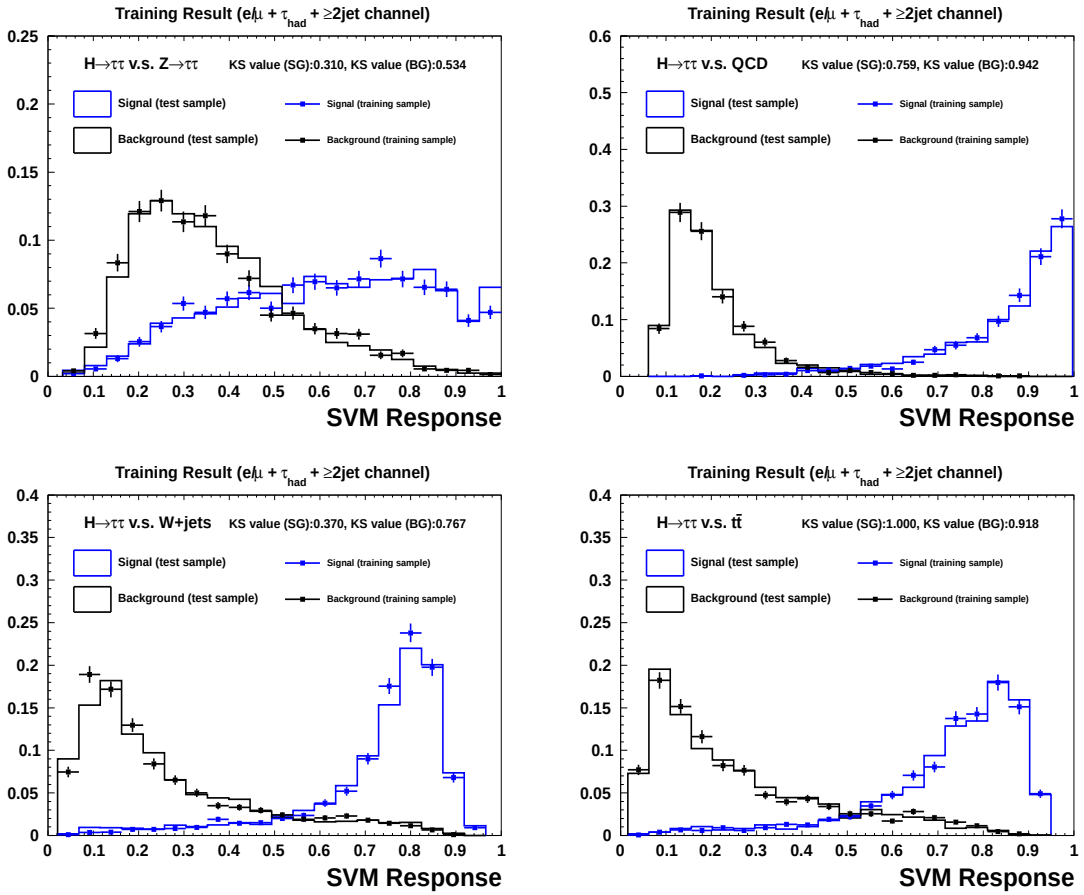

Figure 7.5: Training result of four classifiers for e/μ + $\geq$2jet channel.

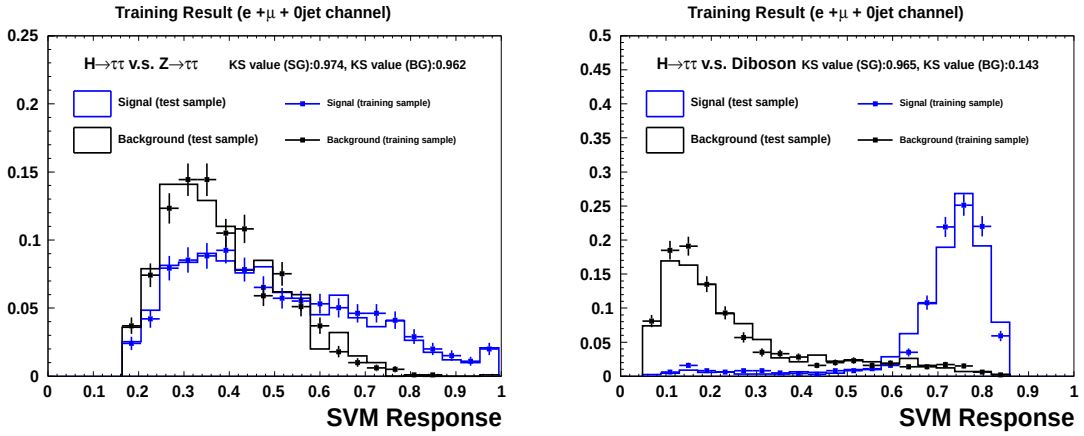

Figure 7.6: Training result of two classifiers for e + μ + 0jet channel.

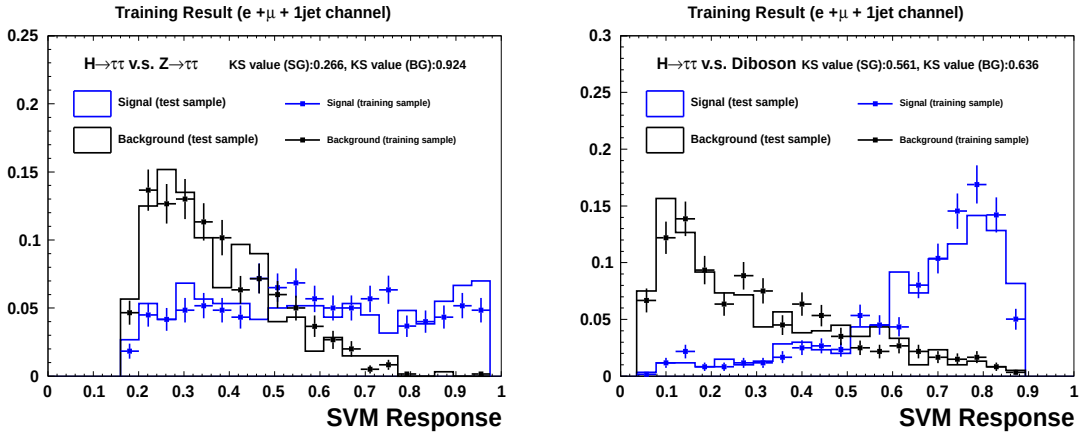

Figure 7.7: Training result of two classifiers for e + μ + 1jet channel.

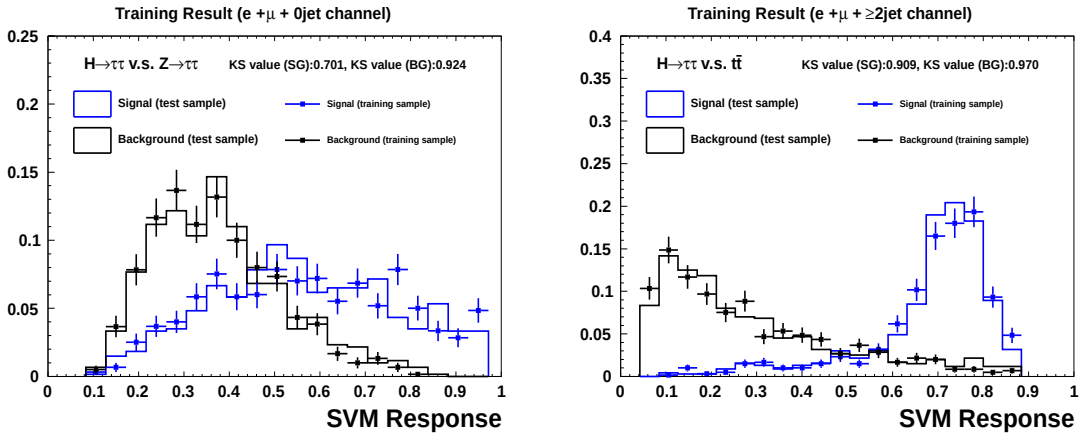

Figure 7.8: Training result of two classifiers for e + μ + $\geq$2jet channel.

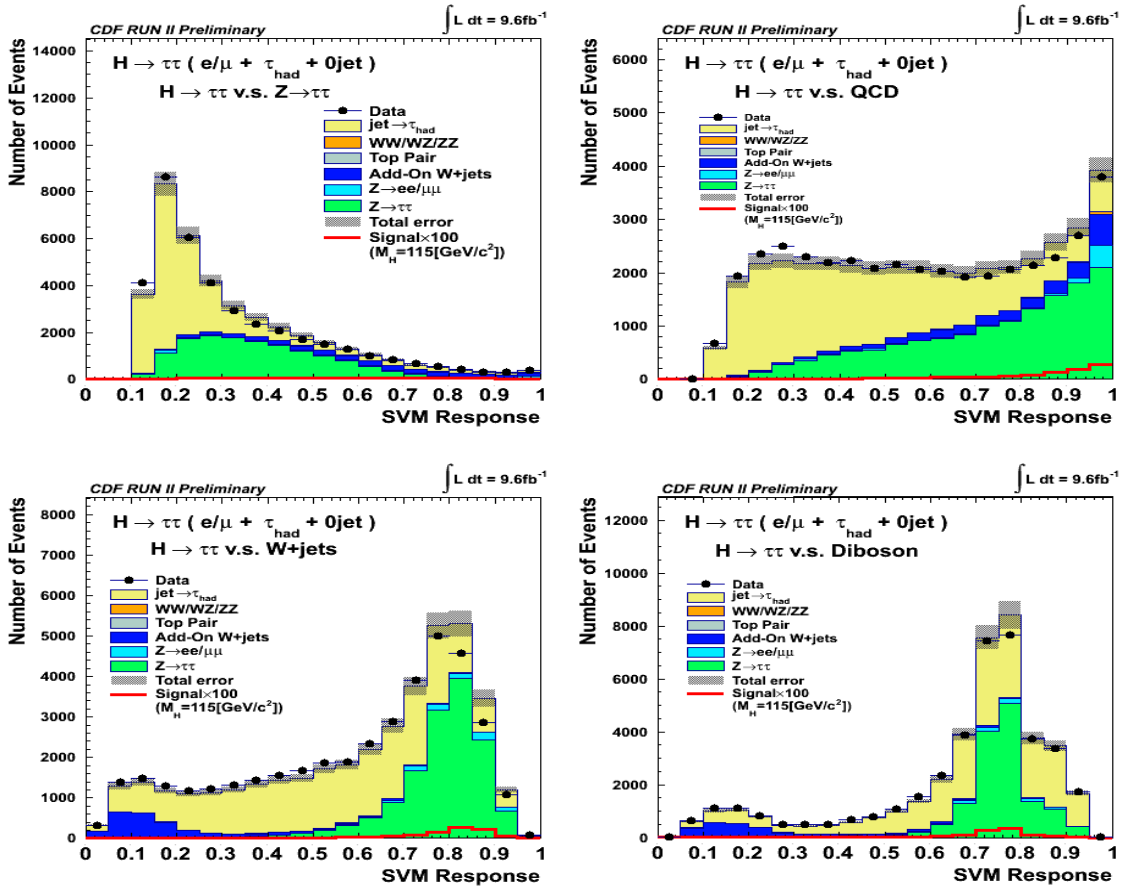

Figure 7.9: The SVM responses for e/μ + τ_{had} + 0jet channel.

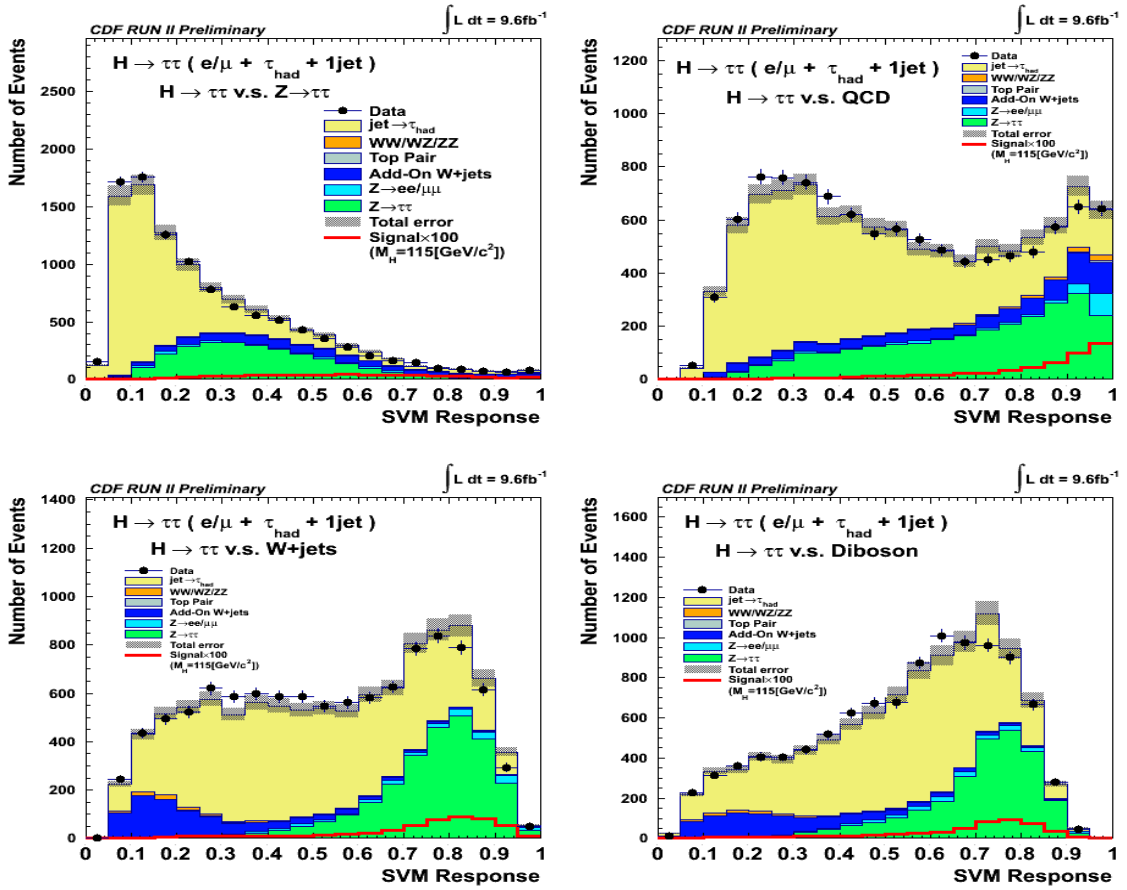

Figure 7.10: The SVM responses for e/μ + τ_{had} + 1jet channel.

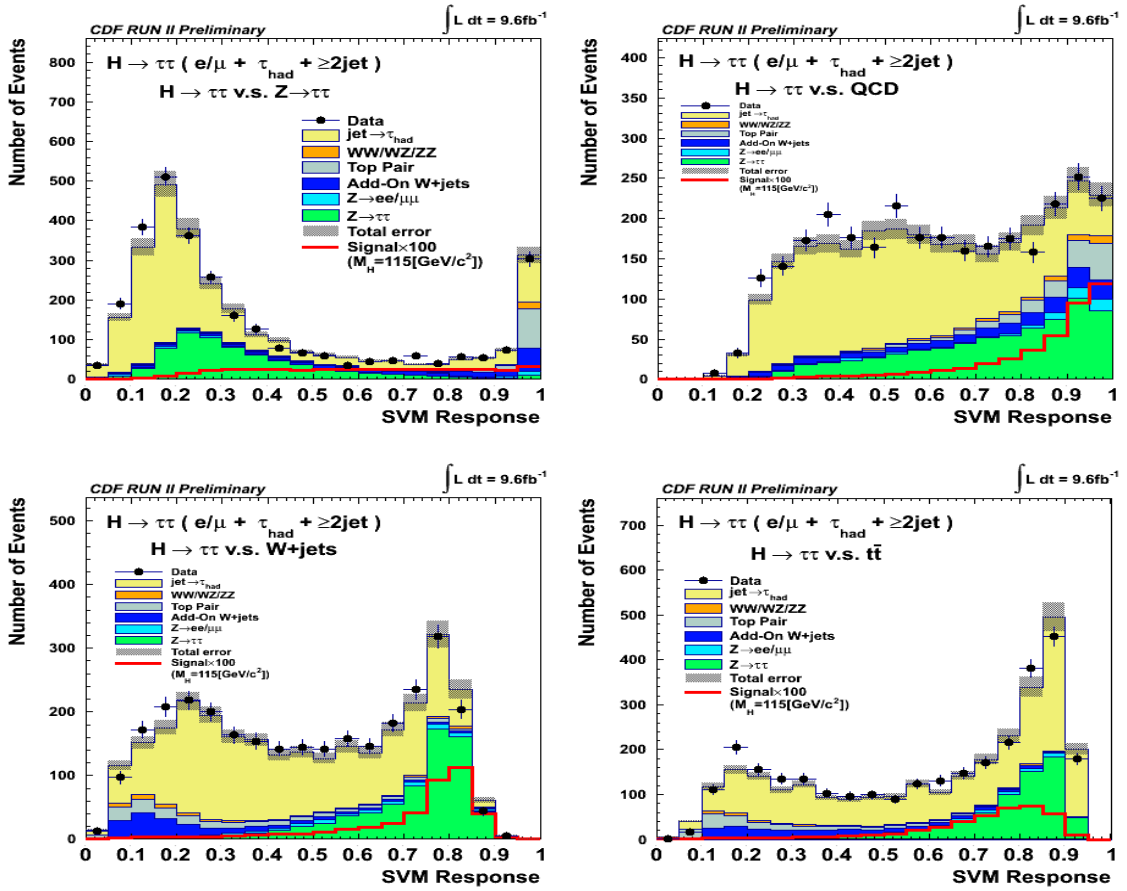

Figure 7.11: The SVM responses for e/μ + τ_{had} + $\geq$2jet channel.

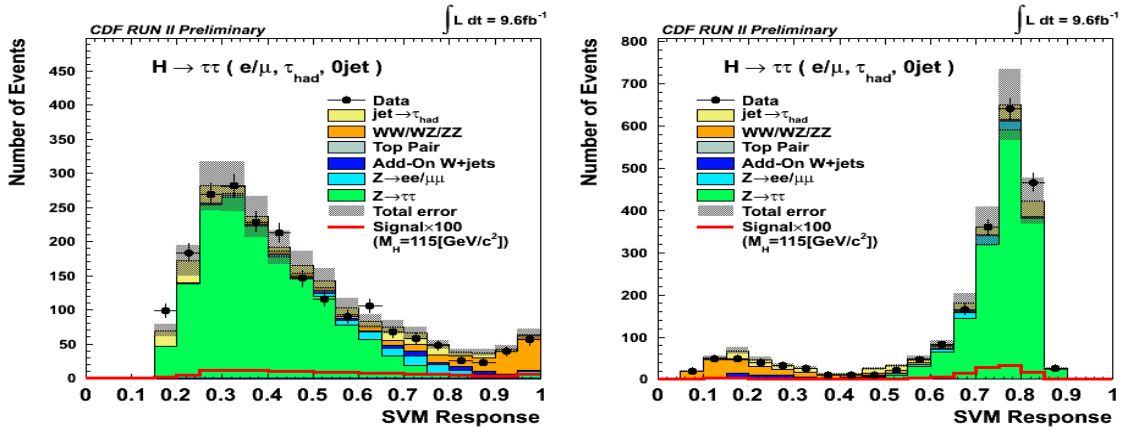

Figure 7.12: The SVM responses for e + μ + 0jet channel.

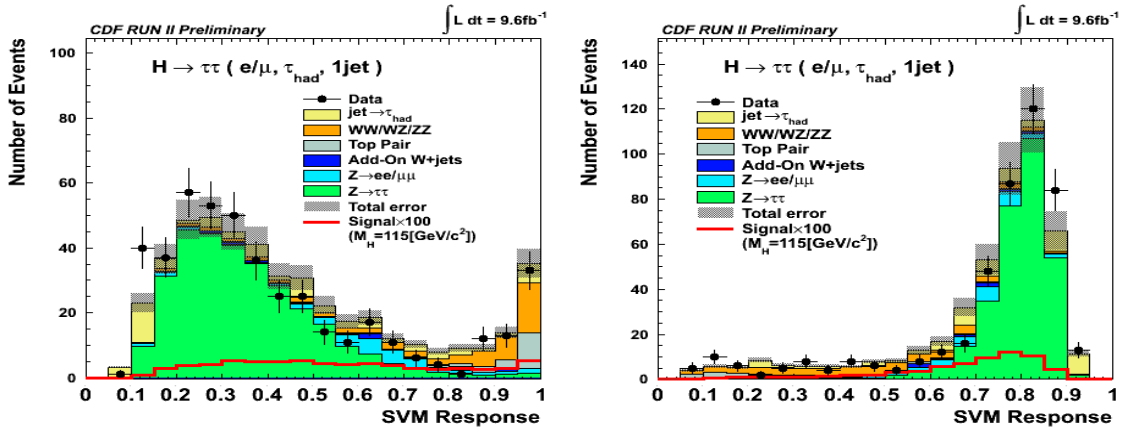

Figure 7.13: The SVM responses for e + μ + 1jet channel.

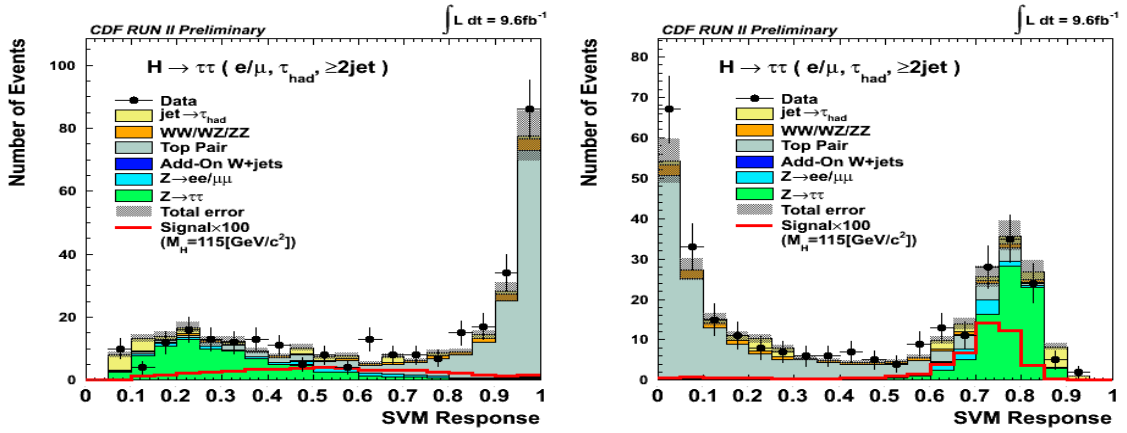

Figure 7.14: The SVM responses for e + μ + $\geq$2jet channel.

7.5 Final Discriminants

For final discriminant, we select a 1D distribution rather than multi-dimensional one. Therefore, for collecting four or two responses to one discriminant value for each event, several methods can be considered as following:

- Taking the **Minimum** response in all responses.
- Taking the **Maximum** response in all responses.
- Taking the **Average** of all responses.
- Taking the **Harmonic Average** of all responses.
- Taking the **Square Root Sum** of all responses.

We create the discriminant distributions using these methods and check the maximum significance value in all bins for each method. The taking MINIMUM response is the best method to maximize the sensitivity of this analysis. Therefore, the final discriminant distributions are obtained by this method. The final discriminant distributions are shown in Figure **??**. In these figures, signal histograms(red) are also scaled by 100 times and the Higgs boson mass point is $115 GeV/c^2$.

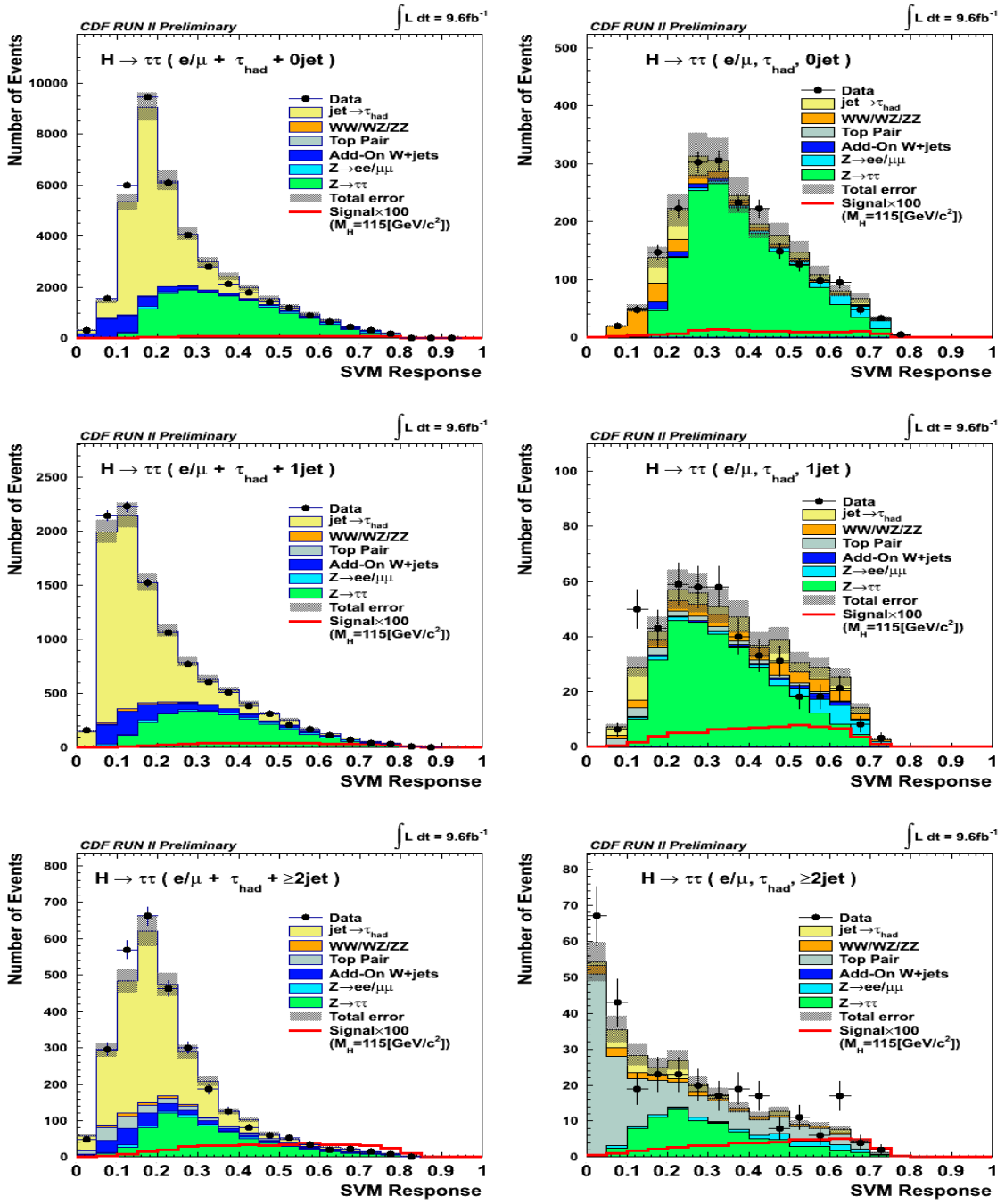

Figure 7.15: The final discriminant distribution for e/μ + τ_{had} + 0/1/$\geq$2jet channel and e + μ + 0/1/$\geq$2jet channel.

Chapter 8

Systematic Uncertainties

There are two kind of systematic uncertainties: rate systematic uncertainties and shape systematic uncertainties. First one affect event expectation in this analysis and the other affect shapes of final SVM responses. In this section, we summarized rate systematic uncertainties and describe shape systematic uncertainties.

8.1 Rate systematic uncertainties

We consider the source of rate systematic uncertainties as following:

Luminosity: 5.9% is assigned as systematic error for MC estimation.

z-vertex cut: 0.5% is assigned as systematic error for MC estimation, which is related data/MC scale factor for z-vertex position cut ($|z_0| \leq 50cm$).

Cross Section: We assign the corresponding cross section systematic uncertainties based on CDF previous measurement or theoretical uncertainty to each signal and background processes. All Higgs boson production cross section uncertainties are quoted from [**?**]. Especially, for $gg \rightarrow H$ cross section, we assign different systematic uncertainty depending on the number of jet because jets come from initial state radiation in this process and these uncertainties are propagate by themselves , more details are described in [**?**].

Trigger efficiency: This systematic uncertainty is related with trigger type. We assign 0.3%, 1.0% and 3.0% for each trigger.

Lepton ID scale factor: This uncertainty is related with data/MC scale factor for lepton identification depending on lepton type of the final state [23]. We assign 2.4%, 2.6% and 3.0% for electron, muon and hadronic tau identification.

Jet Energy Scale (JES): To obtain this uncertainties, jets in all MC simulation samples are applied the jet energy corrections which are shifted to $\pm 1\sigma$ from nominal correction, and we calculate the acceptance and take deviation between this acceptance and the nominal acceptance as systematic uncertainty [24]. This uncertainty is considered both of rate and shape effect, shape uncertainty is described in next section.

Drell-Yan / Additional W+jets scale factor: For Drell-Yan and additional W+jets estimation, 14.8% systematic uncertainty are assigned for data/MC scale factor, that is already mentioned in Section 5.2.1 and 5.3.2.

Same sign data: For the estimation of jet fake event, we directly use same sign data. Since there are no systematic uncertainty, we assign only statistical error for same sign data.

Parton distribution function: This systematic uncertainty is related with theoretical model of parton distribution function(PDF). In this analysis, all signal MC simulations are used CTEQ5L PDF model [28]. Therefore, this uncertainties are taken difference between the nominal acceptance and the acceptance shifted by three contributions:

- weighting parameters of CTEQ5L PDF model
- using different PDF model (MRST72 [29])
- using different strong coupling constant

Finally, we assign 1.2%, 0.9%, 2.2% and 4.9% for each WH, ZH, VBF and ggH processes [25].

Initial State Radiation / Final State Radiation (ISR/FSR): We consider radiation effect from initial state and final state as systematic uncertainties. To estimate this effect, we generate two MC samples generated with more or less radiation at M_H = 120GeV/c^2, and we take a half of the acceptance difference between nominal samples and more or less samples as systematic uncertainty.

In table 8.1 and 8.2, we summarized all systematic uncertainties for each background and signal process in e/μ+ τ_{had} channel and also 8.3 and 8.2 show summary of systematic uncertainties in e + μ channel.

8.1.1 Shape systematic uncertainties

We verify that how the uncertainty of jet energy scale affects the final SVM response for each signal and background process. We compare the minimum SVM response distribution between nominal shape and shifted $\pm$ 1σ JES shapes.

$$\text{Difference} = \frac{(\text{nominal}) - (\pm 1\sigma \text{ shifted})}{\text{nominal}} \tag{8.1}$$

We can see shape difference especially in following processes:

- e/μ + τ_{had} + 0jet channel: $Z \rightarrow \tau\tau, WH, ZH, VBF, ggH$
- e/μ + τ_{had} + 1jet channel: $Z \rightarrow \tau\tau, WH, ZH, VBF, ggH$
- e/μ + τ_{had} + $\geq$ 2jet channel: $Z \rightarrow \tau\tau, WH, ZH, VBF, ggH$

Therefore, we add above list shapes as shape systematics in cross section limit calculation.

The shape comparisons of these processes in each channel are shown in Figure 8.1, Figure 8.2 and Figure 8.3. In these figures, top plots show each nominal shape and shifted $\pm 1\sigma$ JES shape, while bottom plots show relative difference between nominal and $\pm 1\sigma$ shifted like below formula:

The other processes have no strong dependence beyond statistical fluctuation, we do not use these shapes as weight function. The shape comparisons of these processes are all in Appendix.

Systematic Uncertainties (%)							
H $\rightarrow$ $\tau\tau$ (e/μ + τ_{had} channel)							
Source		$Z \rightarrow \tau\tau$	$Z \rightarrow ll$	$t\bar{t}$	Diboson	$jet \rightarrow \tau_{had}$	Add.Wjet
Luminosity		5.9	5.9	5.9	5.9	-	-
$\lvert Z_{vertex} \rvert$		0.5	0.5	0.5	0.5	-	-
σ/Norm.	$= 0jet$	14.8	14.8	10.0	6.0	0.6	14.8
	$= 1jet$	14.8	14.8	10.0	6.0	1.2	14.8
	$\geq 2jet$	14.8	14.8	10.0	6.0	2.4	14.8
ele Trigger		0.3	0.3	0.3	0.3	-	-
μ Trigger		1.0	1.0	1.0	1.0	-	-
track Trigger		3.0	3.0	3.0	3.0	-	-
ele ID		2.4	2.4	2.4	2.4	-	-
μ ID		2.6	2.6	2.6	2.6	-	-
τ ID		3.0	3.0	3.0	3.0	-	-
JES	$= 0jet$	2.5	3.3	18.0	14.4	-	3.4
	$= 1jet$	9.7	8.5	14.7	0.6	-	6.7
	$\geq 2jet$	18.9	22.3	1.4	10.8	-	19.3
Total	$= 0jet$	17.2	17.4	22.2	11.0	0.6	15.2
	$= 1jet$	19.6	19.0	19.6	10.2	1.2	16.2
	$\geq 2jet$	25.5	28.0	13.0	14.8	2.4	23.4

Table 8.1: Summary of rate systematic uncertainties of background sources in e/μ + τ_{had} channel.

Systematic Uncertainties (%)					
H $\rightarrow$ $\tau\tau$ (e/μ + τ_{had} channel)					
Source		WH	ZH	VBF	ggH
Luminosity		5.9	5.9	5.9	5.9
$\lvert Z_{vertex}\rvert$		0.5	0.5	0.5	0.5
σ/Norm.	$= 0jet$	5.0	5.0	10.0	13.4
	$= 1jet$	5.0	5.0	10.0	23.5
	$\geq 2jet$	5.0	5.0	10.0	33.0
PDF		1.2	0.9	2.2	4.9
ele Trigger		0.3	0.3	0.3	0.3
μ Trigger		1.0	1.0	1.0	1.0
track Trigger		3.0	3.0	3.0	3.0
ele ID		2.4	2.4	2.4	2.4
μ ID		2.6	2.6	2.6	2.6
τ ID		3.0	3.0	3.0	3.0
JES	$= 0jet$	9.4	4.9	16.3	2.3
	$= 1jet$	2.3	6.4	7.2	4.8
	$\geq 2jet$	4.7	3.5	3.7	13.3
IF-SR	$= 0jet$	0.7	2.5	11.5	2.1
	$= 1jet$	3.2	6.0	5.7	6.8
	$\geq 2jet$	4.0	3.5	1.7	21.4
Total	$= 0jet$	13.5	11.1	23.9	16.7
	$= 1jet$	10.4	13.0	16.0	26.7
	$\geq 2jet$	11.5	10.7	13.7	42.6

Table 8.2: Summary of rate systematic uncertainties of signal sources in e/μ + τ_{had} channel.

Systematic Uncertainties (%)							
H $\rightarrow$ $\tau\tau$ (e + μ channel)							
Source		$Z \rightarrow \tau\tau$	$Z \rightarrow ll$	$t\bar{t}$	Diboson	$jet \rightarrow e/\mu$	Add.Wjet
Luminosity		5.9	5.9	5.9	5.9	-	-
$\lvert Z_{vertex} \rvert$		0.5	0.5	0.5	0.5	-	-
σ/Norm	$= 0jet$	14.8	14.8	10.0	6.0	6.9	14.8
	$= 1jet$	14.8	14.8	10.0	6.0	13.2	14.8
	$\geq 2jet$	14.8	14.8	10.0	6.0	20.9	14.8
ele Trigger		0.3	0.3	0.3	0.3	-	-
μ Trigger		1.0	1.0	1.0	1.0	-	-
track Trigger		3.0	3.0	3.0	3.0	-	-
ele ID		2.4	2.4	2.4	2.4	-	-
μ ID		2.6	2.6	2.6	2.6	-	-
JES	$= 0jet$	2.4	2.9	25.2	3.1	-	0.6
	$= 1jet$	9.2	4.4	12.5	6.7	-	3.0
	$\geq 2jet$	18.9	14.3	1.9	14.3	-	10.2
Total	$= 0jet$	17.0	17.6	28.2	10.2	7.3	14.8
	$= 1jet$	19.1	17.9	17.7	11.8	17.3	15.1
	$\geq 2jet$	25.8	22.5	12.7	17.3	22.3	18.0

Table 8.3: Summary of rate systematic uncertainties of background sources in e + μ channel.

Systematic Uncertainties (%)					
H $\rightarrow$ $\tau\tau$ (e + μ channel)					
Source		WH	ZH	VBF	ggH
Luminosity		5.9	5.9	5.9	5.9
$\lvert Z_{vertex}\rvert$		0.5	0.5	0.5	0.5
σ/Norm	$= 0jet$	5.0	5.0	10.0	13.4
	$= 1jet$	5.0	5.0	10.0	23.5
	$\geq 2jet$	5.0	5.0	10.0	33.0
PDF		1.2	0.9	2.2	4.9
ele Trigger		0.3	0.3	0.3	0.3
μ Trigger		1.0	1.0	1.0	1.0
track Trigger		3.0	3.0	3.0	3.0
ele ID		2.4	2.4	2.4	2.4
μ ID		2.6	2.6	2.6	2.6
JES	$= 0jet$	6.5	6.0	3.6	4.1
	$= 1jet$	0.7	4.2	3.6	5.4
	$\geq 2jet$	6.0	4.8	19.1	13.6
IFR/FSR	$= 0jet$	3.6	4.8	20.6	6.7
	$= 1jet$	2.7	2.6	4.0	7.6
	$\geq 2jet$	5.0	1.8	0.2	9.8
Total	$= 0jet$	11.8	11.9	27.3	18.0
	$= 1jet$	9.6	10.4	15.0	26.8
	$\geq 2jet$	12.0	10.5	13.3	38.1

Table 8.4: Summary of rate systematic uncertainties of signal sources in e + μ channel.

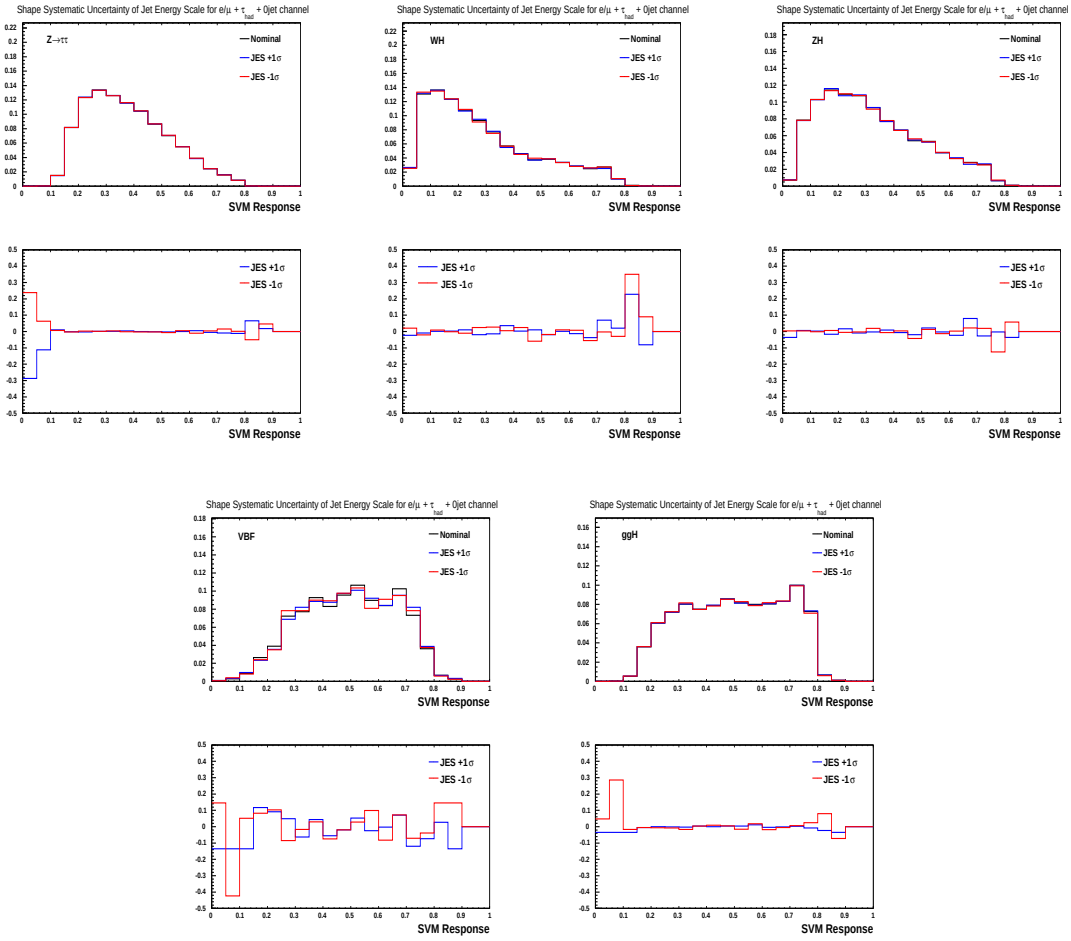

Figure 8.1: The shape comparison for e/μ + τ_{had} + 0jet channel.

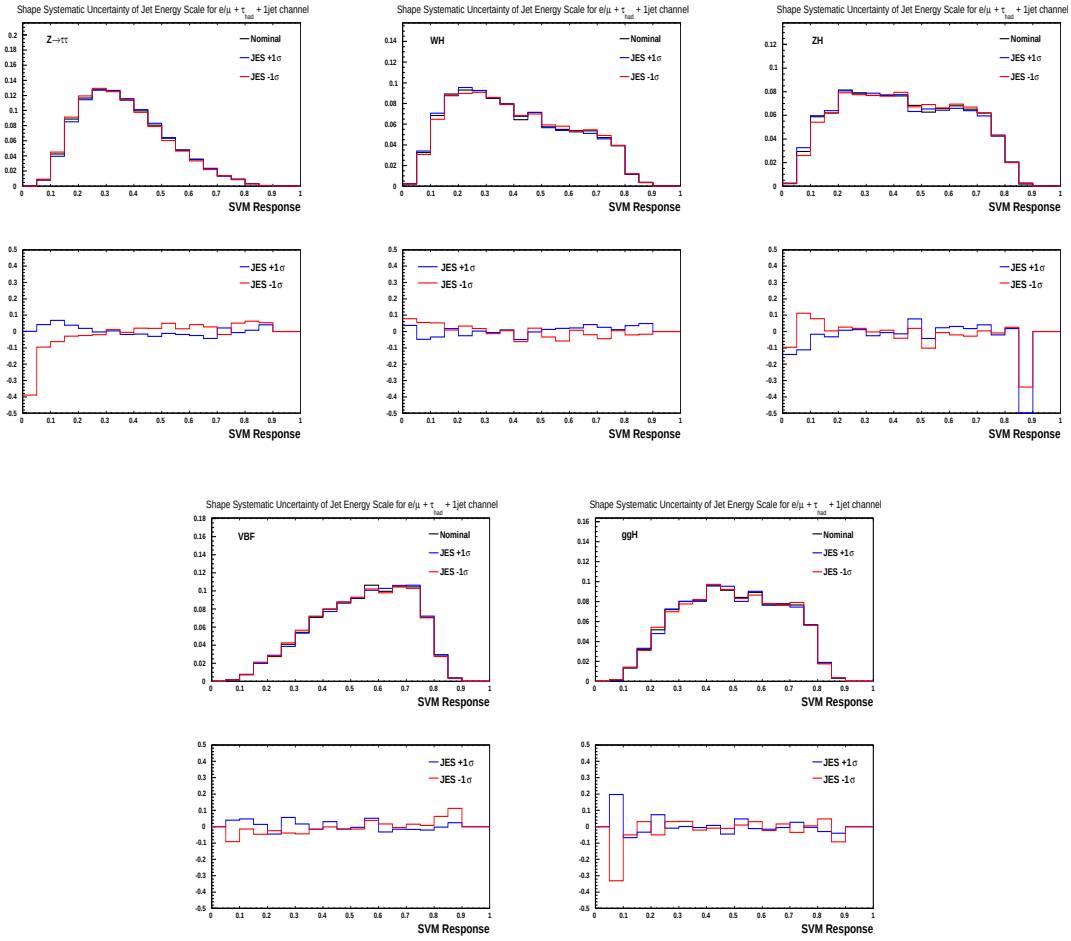

Figure 8.2: The shape comparison for e/μ + τ_{had} + 1jet channel.

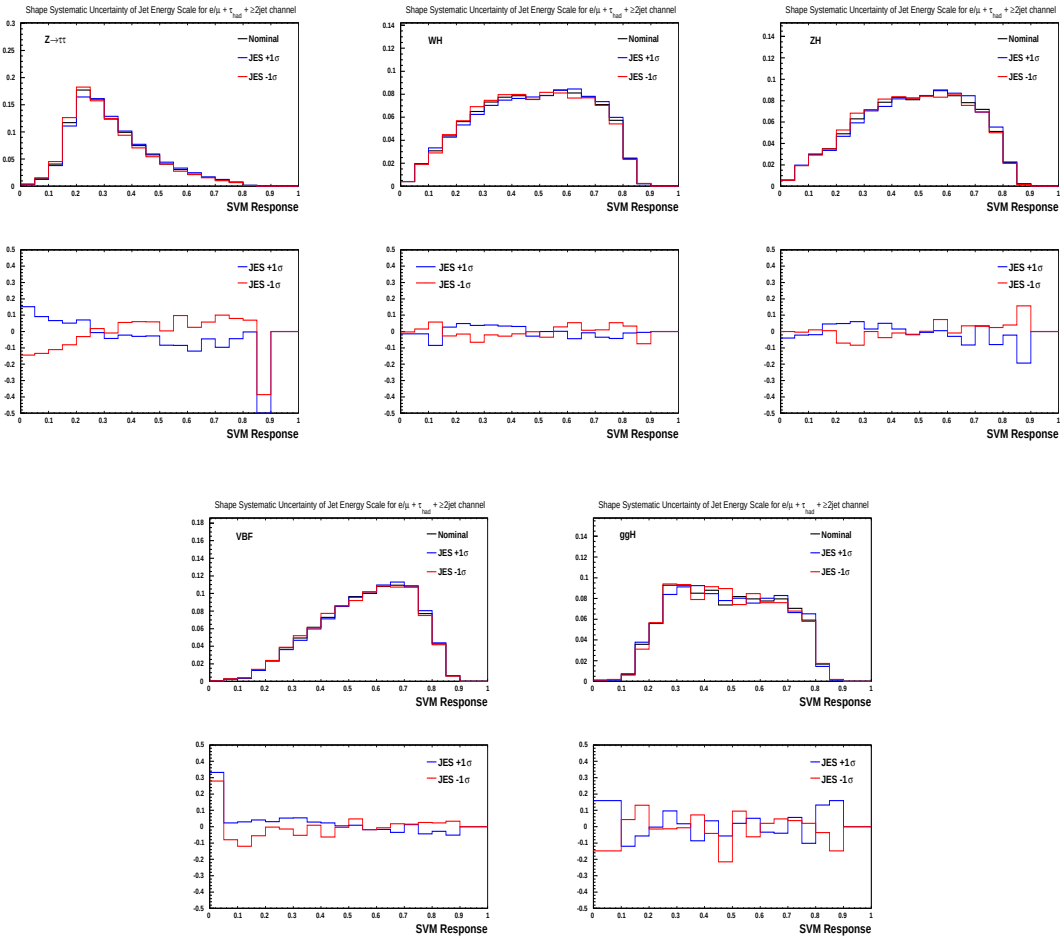

Figure 8.3: The shape comparison for e/μ + τ_{had} + 2jet channel.

Chapter 9

Result

From the expected sensitivity, it is already known to be quite hard to claim an evidence of the SM Higgs and in fact there is no clear excess in the signal region. Given the fact that our background models nicely agree with data in background rich region, we set 95% C.L limit for the SM Higgs cross section in this note. Using final discriminant, we perform binned maximum likelihood method to obtain 95% C.L limit. For each bin, the expected number of event(μ_i) is evaluated as below:

$$\mu_i = \sum_{k=1}^{N_{bkg}} f_i^k \cdot N^k + \sum_{k=1}^{N_{sig}} f_i^l \cdot (\epsilon^l \cdot \sigma^l \cdot \int Ldt), \tag{9.1}$$

, where k represents the kind of background($Z \rightarrow \tau\tau$, $Z \rightarrow ll$, $t\bar{t}$, Diboson, add-on W+jets) and l represents the kind of signal(WH, ZH, VBF, ggH), and N_{bkg} and N_{sig} are the number of kind of background process($N_{bkg} = 5$) and signal process($N_{sig} = 4$). For each k and l, f_i^k and f_i^l represent the expected fraction in i-th bin. In second term, ϵ^l is the detection efficiency (Acceptance, trigger efficiency, lepton identification efficiency, z-vertex efficiency) for signal l and $\int Ldt$ is the integrated luminosity. And finally, σ^l is production cross section of signal l, which is unknown parameter for likelihood fit.

Then, we define likelihood function as below:

$$L(\frac{\sigma}{\sigma_{SM}}) = \int \dots \int \prod_{i=1}^{N_{bin}} \frac{\mu_i^{N_i}}{N_i!} e^{-\mu_i} \prod_{k=1}^{N_{bkg}} G(N^k, \Delta^k dN^k) \prod_{l=1}^{N_{sig}} G(N^l, \Delta^l dN^l), \tag{9.2}$$

, where Δ^k and Δ^l are the systematic uncertainty of each processes, and G represents Gaussian fraction, which is fluctuated by the expected systematic uncertainty. N_i is the number of observed event in i-th bin. For each signal production cross section(σ^l), we assume that these are 100% correlated. Therefore ratios between parameters of σ^l and the SM production cross section(σ_{SM}) are same, we consider one parameter ($\frac{\sigma}{\sigma_{SM}}$) in the likelihood.

Here, we define four likelihoods for each signal channel. Then, four likelihoods are collected by taking product of each likelihood like following formula:

$$L = L_1 \times L_2 \times L_3 \times L_4. \tag{9.3}$$

We evaluate the expected 95% C.L limit from the maximum binned likelihood by pseudo-experiment. When we get the number of event of pseudo data, each background source is fluctuated by Gaussian fraction with systematic uncertainties. Then, the number of event for each bin in one pseudo-experiment is extracted by Poisson distribution [30].

e/μ + τ_{had} + 0 jet channel						
Higgs Mass	Expected Limit/σ_{SM}					Observed Limit/σ_{SM}
GeV/c^2	-2σ	-1σ	Median	+1σ	+2σ	
100.0	33.6	52.0	73.9	107.7	154.1	74.8
105.0	24.7	37.7	52.1	75.9	101.6	47.2
110.0	18.5	29.6	42.8	61.9	87.7	40.3
115.0	15.5	25.4	36.8	54.6	74.6	27.4
120.0	17.5	25.1	36.4	53.0	76.2	31.7
125.0	18.8	26.1	38.4	55.0	77.1	26.3
130.0	20.0	28.3	41.2	59.7	87.6	29.0
135.0	20.0	29.7	44.3	65.0	89.2	31.0
140.0	22.6	35.4	52.0	75.6	111.5	44.3
145.0	34.0	49.8	75.3	110.1	155.2	44.5
150.0	49.7	72.8	106.1	155.6	226.5	62.9

Table 9.1: Expected and Observed limit for e/μ + τ_{had} + 0jet channel.

The 95% C.L cross section limits for e/μ + τ_{had} + 0/1/$\geq$2jet channel are shown in Table 9.1,Table 9.2 and Table 9.1, and also shown in Figure 9.1, Figure 9.2 and Figure 9.3. The 95% C.L cross section limits for e + μ + 0/1/$\geq$2jet channel are shown in Table 9.4,Table 9.5 and Table 9.4, and also shown in Figure 9.4, Figure 9.5 and Figure 9.6.

Finally, all channel combined limit is summarized in Table 9.7 and Figure 9.7 as a result of this analysis.

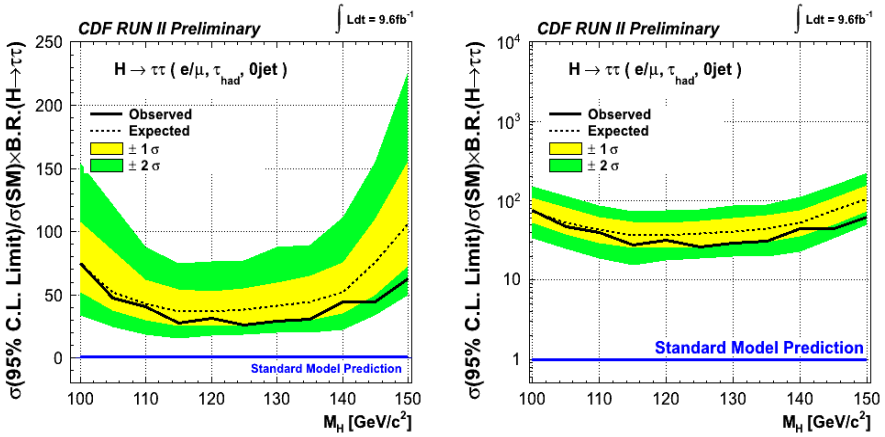

Figure 9.1: Expected and Observed limit for e/μ + τ_{had} + 0jet channel.

e/μ + τ_{had} + 1 jet channel						
Higgs Mass GeV/c^2	Expected Limit/σ_{SM} -2σ	-1σ	Median	+1σ	+2σ	Observed Limit/σ_{SM}
100.0	33.6	52.0	73.9	107.7	154.1	74.8
105.0	24.7	37.7	52.1	75.9	101.6	47.2
110.0	18.5	29.6	42.8	61.9	87.7	40.3
115.0	15.5	25.4	36.8	54.6	74.6	27.4
120.0	17.5	25.1	36.4	53.0	76.2	31.7
125.0	18.8	26.1	38.4	55.0	77.1	26.3
130.0	20.0	28.3	41.2	59.7	87.6	29.0
135.0	20.0	29.7	44.3	65.0	89.2	31.0
140.0	22.6	35.4	52.0	75.6	111.5	44.3
145.0	34.0	49.8	75.3	110.1	155.2	44.5
150.0	49.7	72.8	106.1	155.6	226.5	62.9

Table 9.2: Expected and Observed limit for e/μ + τ_{had} + 1jet channel.

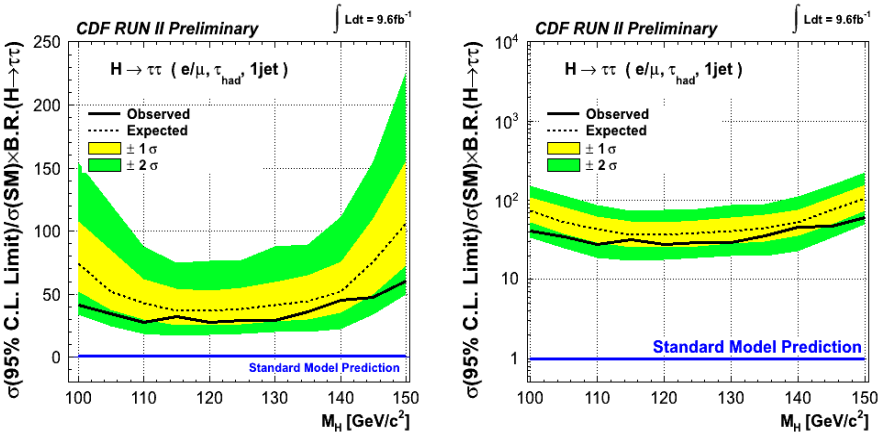

Figure 9.2: Expected and Observed limit for e/μ + τ_{had} + 1jet channel.

e/μ + τ_{had} + $\geq$2 jet channel						
Higgs Mass GeV/c^2	Expected Limit/σ_{SM}					Observed Limit/σ_{SM}
	-2σ	-1σ	Median	+1σ	+2σ	
100.0	10.9	15.9	23.2	33.8	50.6	13.5
105.0	10.1	14.0	20.4	29.8	43.8	11.9
110.0	9.0	11.9	17.0	24.7	35.9	12.4
115.0	7.6	11.0	16.3	23.5	33.8	19.4
120.0	8.1	11.1	15.4	22.2	32.7	14.6
125.0	9.5	12.5	17.9	26.1	36.9	15.9
130.0	9.9	13.3	18.8	27.2	37.1	17.7
135.0	12.7	17.9	24.6	35.0	48.7	28.3
140.0	14.9	21.1	29.7	42.3	62.6	33.9
145.0	21.1	28.6	40.6	59.1	81.8	43.0
150.0	30.1	41.7	60.8	87.9	121.9	64.2

Table 9.3: Expected and Observed limit for e/μ + τ_{had} + 2jet channel.

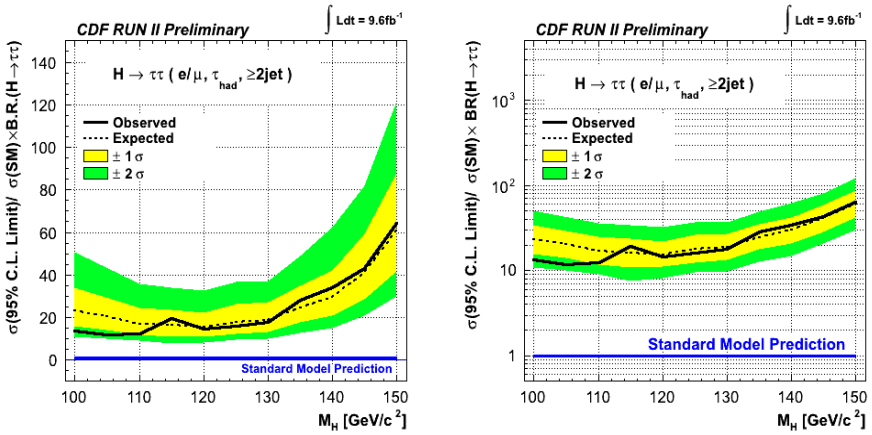

Figure 9.3: Expected and Observed limit for e/μ + τ_{had} + $\geq$2jet channel.

e + μ + 0 jet channel						
Higgs Mass	Expected Limit/σ_{SM}					Observed Limit/σ_{SM}
GeV/c^2	-2σ	-1σ	Median	+1σ	+2σ	
100.0	154.3	231.5	337.9	496.2	688.9	245.8
105.0	135.7	194.0	275.1	387.2	553.7	157.5
110.0	100.3	138.9	200.4	290.3	410.8	136.6
115.0	74.3	102.9	149.0	222.8	302.3	81.1
120.0	72.2	100.8	141.0	208.9	283.3	106.4
125.0	78.4	111.6	162.0	224.9	305.3	110.0
130.0	86.4	118.9	167.4	240.7	335.1	110.5
135.0	91.6	124.8	175.0	253.7	353.0	109.5
140.0	104.0	144.0	201.0	285.9	392.6	122.3
145.0	146.5	194.3	272.7	387.6	535.1	197.5
150.0	209.8	278.2	387.3	543.4	730.7	267.5

Table 9.4: Expected and Observed limit for e + μ + 0jet channel.

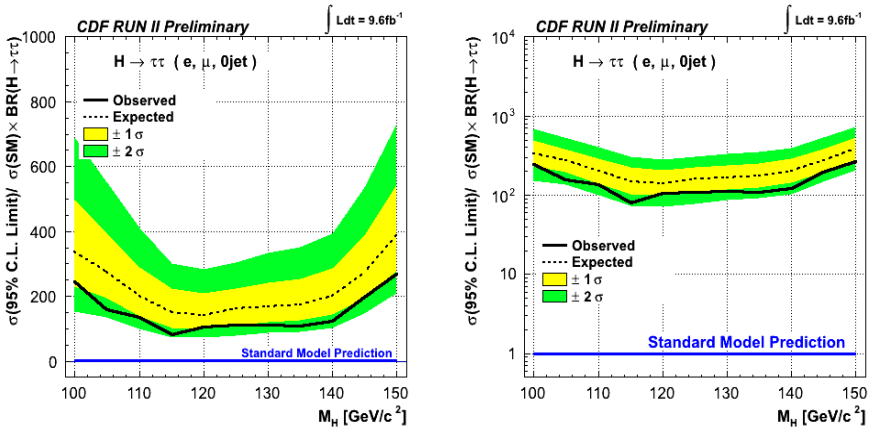

Figure 9.4: Expected and Observed limit for e + μ + 1jet channel.

e + μ + 1 jet channel						
Higgs Mass GeV/c^2	Expected Limit/σ_{SM}					Observed Limit/σ_{SM}
	-2σ	-1σ	Median	+1σ	+2σ	
100.0	53.8	74.9	106.8	153.2	216.3	80.6
105.0	52.1	72.4	103.8	147.5	205.6	82.9
110.0	58.6	79.9	114.6	162.1	233.9	93.0
115.0	63.0	83.5	118.0	172.1	249.0	84.0
120.0	60.4	82.0	116.0	168.3	233.5	93.4
125.0	93.4	127.1	181.8	253.9	354.9	113.0
130.0	120.7	156.7	223.0	321.7	446.6	126.9
135.0	113.9	148.5	207.9	290.1	402.3	137.1
140.0	153.9	197.0	272.3	386.0	559.4	178.8
145.0	205.3	272.0	378.9	554.3	730.5	259.3
150.0	329.4	428.8	580.2	771.0	926.6	389.5

Table 9.5: Expected and Observed limit for e + μ + 1jet channel.

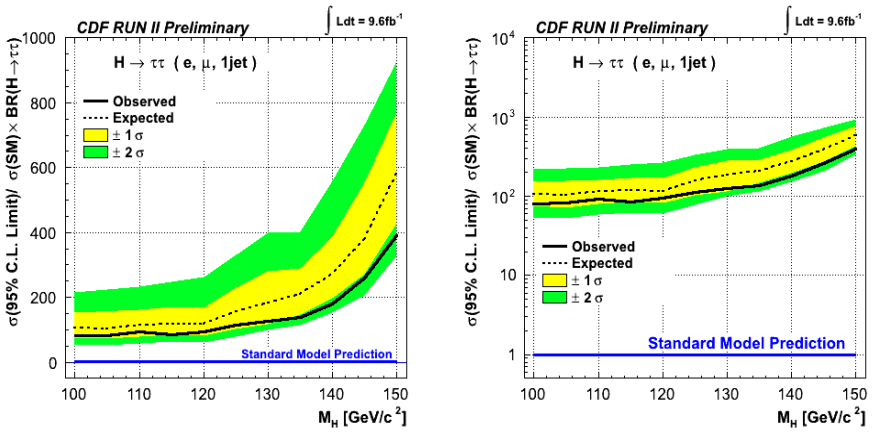

Figure 9.5: Expected and Observed limit for e/μ + τ_{had} + 1jet channel.

e + μ + $\geq$2 jet channel						
Higgs Mass GeV/c^2	Expected Limit/σ_{SM}					Observed Limit/σ_{SM}
	-2σ	-1σ	Median	+1σ	+2σ	
100.0	30.6	44.1	62.7	91.5	127.7	84.1
105.0	31.7	45.3	65.6	92.0	129.0	87.9
110.0	34.3	46.1	66.3	97.6	137.9	98.2
115.0	32.6	44.3	62.6	92.2	128.1	93.3
120.0	36.3	50.4	70.8	103.5	146.0	100.6
125.0	39.1	53.2	73.4	103.7	155.5	122.8
130.0	45.3	60.4	87.3	127.1	174.3	138.7
135.0	56.0	76.2	106.9	152.7	213.6	153.1
140.0	69.6	96.0	135.0	198.9	272.0	197.7
145.0	106.0	144.4	202.1	288.9	409.3	281.2
150.0	160.1	213.1	305.0	434.8	611.9	425.6

Table 9.6: Expected and Observed limit for e + μ + $\geq$2jet channel.

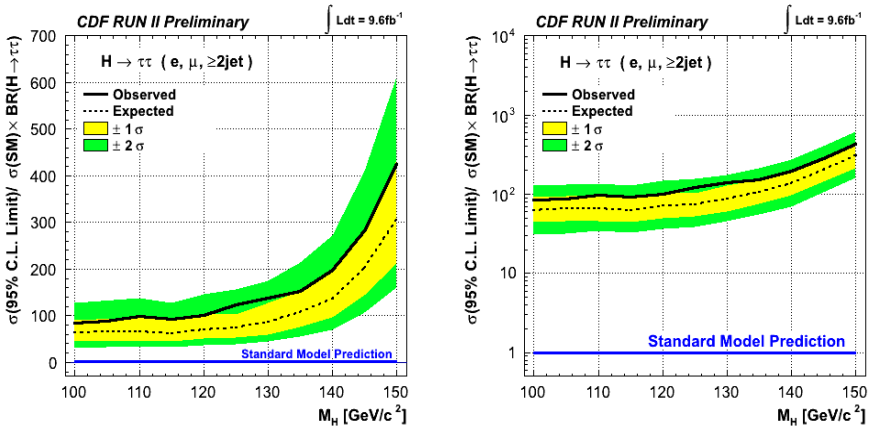

Figure 9.6: Expected and Observed limit for e/μ + τ_{had} + $\geq$2jet channel.

$H \rightarrow \tau\tau$						
Higgs Mass GeV/c^2	Expected Limit/σ_{SM}					Observed Limit/σ_{SM}
	-2σ	-1σ	Median	+1σ	+2σ	
100.0	3.9	7.3	13.7	24.5	42.3	14.4
105.0	4.0	6.9	12.4	21.7	37.0	7.4
110.0	4.0	6.7	11.6	20.0	33.1	6.9
115.0	3.6	6.0	10.6	18.6	32.1	8.3
120.0	4.3	6.8	11.2	19.1	31.3	5.9
125.0	5.1	7.9	12.7	20.4	32.1	10.4
130.0	5.6	8.6	14.0	21.8	34.6	6.9
135.0	6.7	10.3	16.6	26.4	41.1	7.0
140.0	8.5	12.9	20.3	32.0	49.6	10.6
145.0	11.5	17.5	27.6	43.6	66.9	17.2
150.0	16.7	25.4	39.6	62.3	95.4	27.3

Table 9.7: Expected and Observed limit for combining all channel.

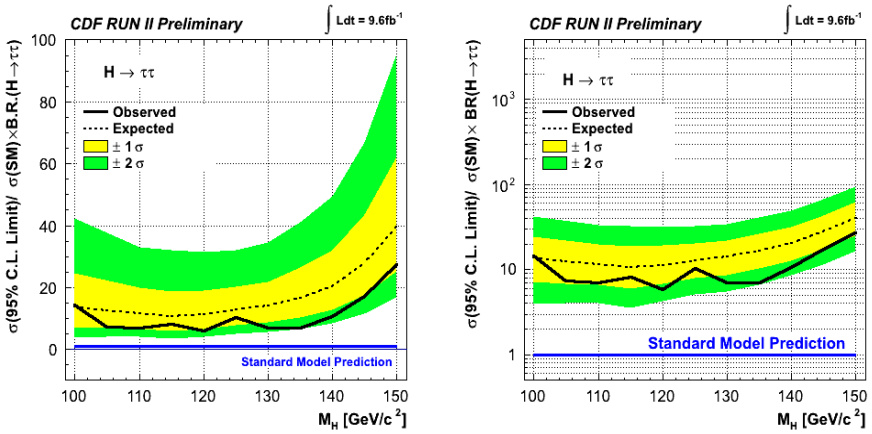

Figure 9.7: Expected and Observed limit for combining all channel.

Chapter 10

Conclusion

We have performed the search for the Standard Model Higgs boson in $\tau\tau$ final state in proton and antiproton collisions at CDF. The search sensitivity in this analysis is optimized by the Support Vector Machine. We could not see any discrepancy in data and background prediction. Therefore, as a result of this analysis, we extract a 95% confidence level upper limit on the Standard Model Higgs boson cross section($XH \rightarrow \tau\tau + X$) at 95% confidence level, these ranges are 14.4 to 27.3 times the SM Higgs boson cross section from $M_H = 100$GeV/c^2 to $M_H = 150$GeV/c^2 in 5GeV/c^2. The observed(expected) 95% confidence level limit times the SM Higgs boson cross section at M$_H$=115GeV/c^2 is

$$\sigma(95\%C.L.limit)/\sigma_{SM}(p\bar{p} \rightarrow XH) \times B.R.(XH \rightarrow \tau\tau + X) < 8.3(10.6). \quad (10.1)$$

In the Tevatron/CDF experiment, there are a lot of the SM Higgs boson search channel such as $WH \rightarrow l\nu b\bar{b}$, $ZH \rightarrow ll/\nu\nu b\bar{b}$, $H \rightarrow WW$ and so on. This result will be combined with other channels to obtain further improvement of the SM Higgs boson discovery sensitivity.

In addition, there is the D0 experiment other than the CDF experiment in the Tevatron, and this experiment has also performed the SM Higgs boson search with many search channel.

The results from two experiments will be combined together, and the discovery sensitivity of the Tevatron SM Higgs boson search result will be further improved.

By this analysis, we are able to get a foothold of discovery or exclude of the SM Higgs boson, in brief the explication of the electroweak symmetry breaking.

Appendix A

Kinematical Distribution

We define e/μ + τ_{had} + 0/1/$\geq$2jet channels and e + μ + 0/1/$\geq$2jet channels as signal region. The significant kinematic variables are shown in below figures. These plots have systematic uncertain band of background prediction, and these systematic uncertainties are described in Section 8.

A.1 e/μ + τ_{had} + 0jet channel

Kinematic variables of e/μ + τ_{had} + 0jet channel are shown in Figure A.1 to Figure A.7.

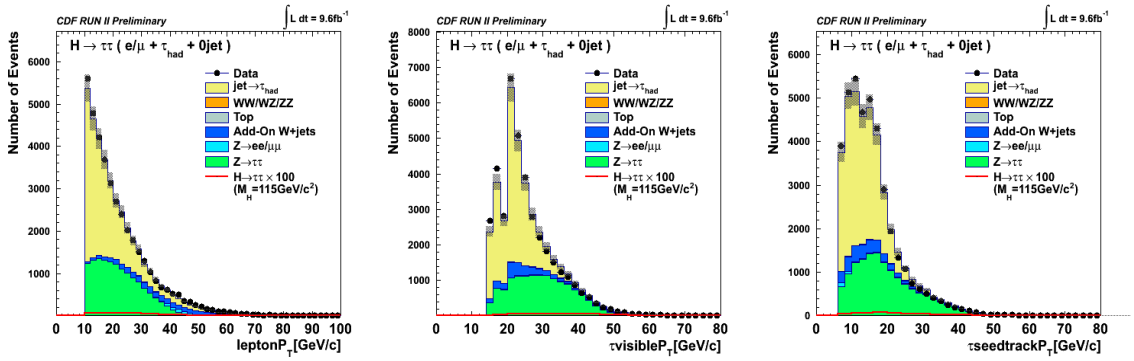

Figure A.1: Lepton P_T, τ_{had} visible P_T and τ_{had} seed track P_T for control region of e/μ + τ_{had} channel

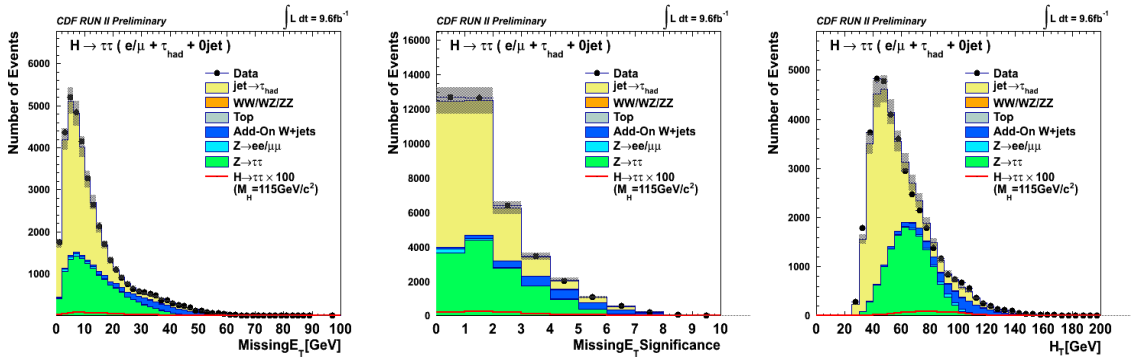

Figure A.2: Mssing E_T, Missing E_T significance and H_T for the control region of e/μ + τ_{had} channel

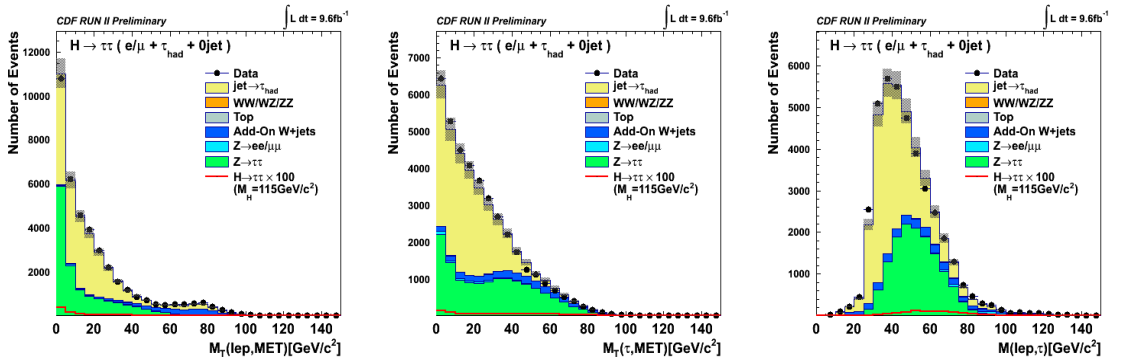

Figure A.3: Transverse mass of lepton and $\not\!\!E_T$, transverse mass of τ_{had} and $\not\!\!E_T$and invariant mass of lepton and τ_{had} for the control region of e/μ + τ_{had} channel

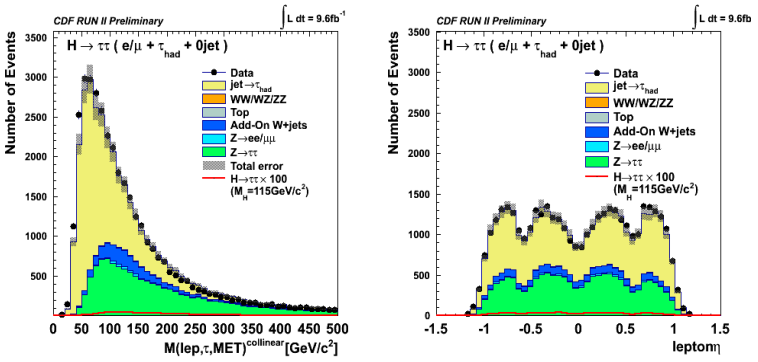

Figure A.4: Invariant Mass of lepton, τ_{had} and $\not\!\!E_T$after collinear approximation and lepton η for the control region of e/μ + τ_{had} channel

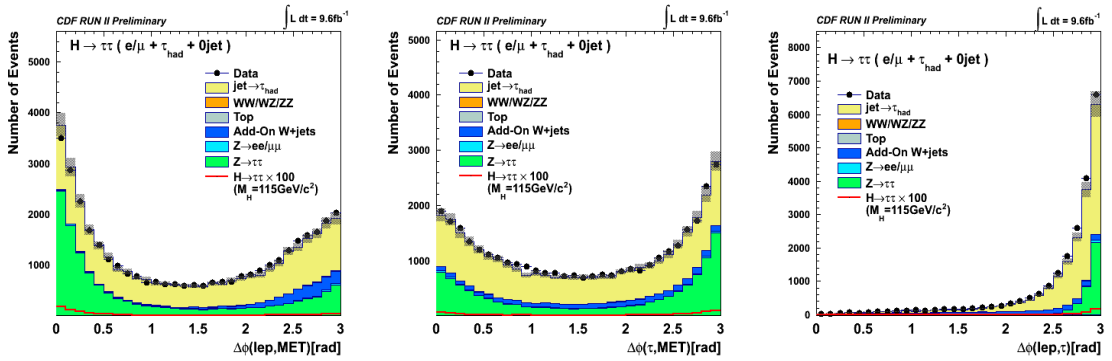

Figure A.5: $\Delta\phi$ between lepton and $\not\!\!E_T$, $\Delta\phi$ between τ_{had} and $\not\!\!E_T$and $\Delta\phi$ between lepton, τ_{had} and $\not\!\!E_T$for the control region of e/μ + τ_{had} channel

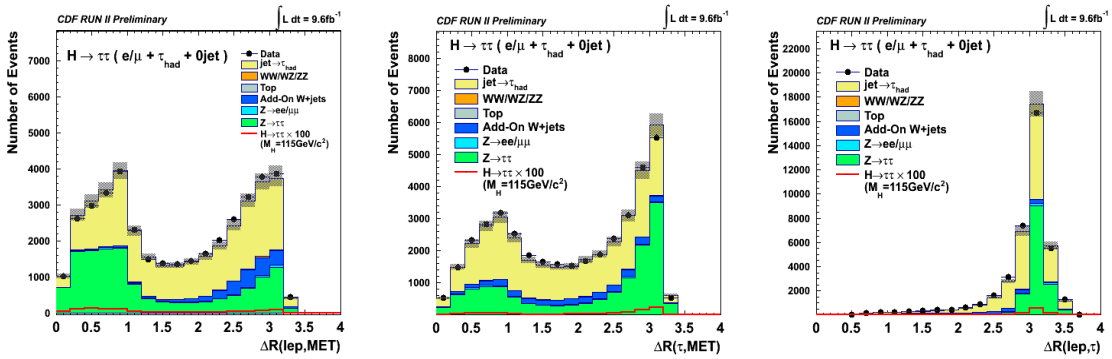

Figure A.6: ΔR between lepton and $\not\!\!E_T$, ΔR between τ_{had} and $\not\!\!E_T$and ΔR between lepton, τ_{had} and $\not\!\!E_T$for the control region of e/μ + τ_{had} channel

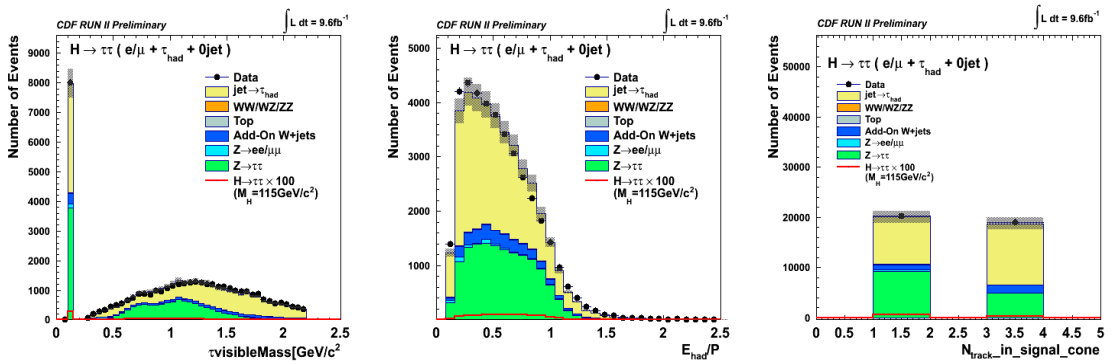

Figure A.7: τ_{had} visible mass, E_{had}/P and the number of tracks in signal cone for the control region of e/μ + τ_{had} channel

A.2 e/μ + τ_{had} + 1jet channel

Kinematic variables for signal region of e/μ + τ_{had} + 1jet channel are shown in Figure A.8 to Figure A.15.

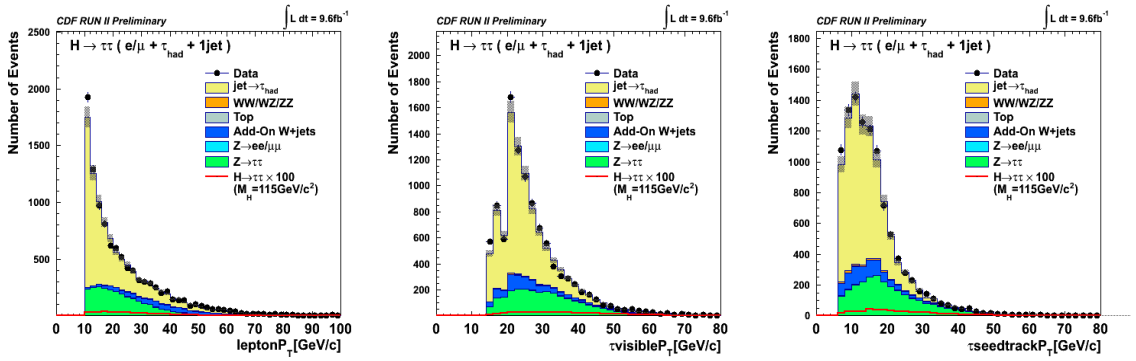

Figure A.8: Lepton P_T, τ_{had} visible P_T and τ_{had} seed track P_T for control region of e/μ + τ_{had} channel

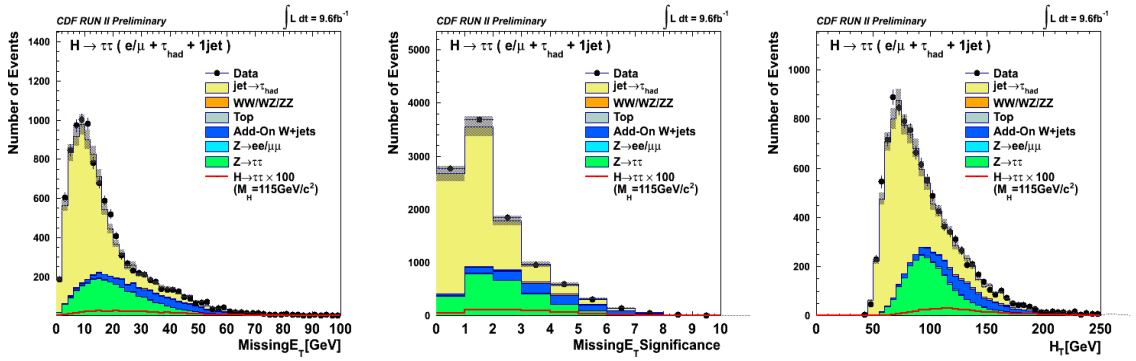

Figure A.9: Mssing E_T, Missing E_T significance and H_T for signal region of e/μ + τ_{had} + 1jet channel

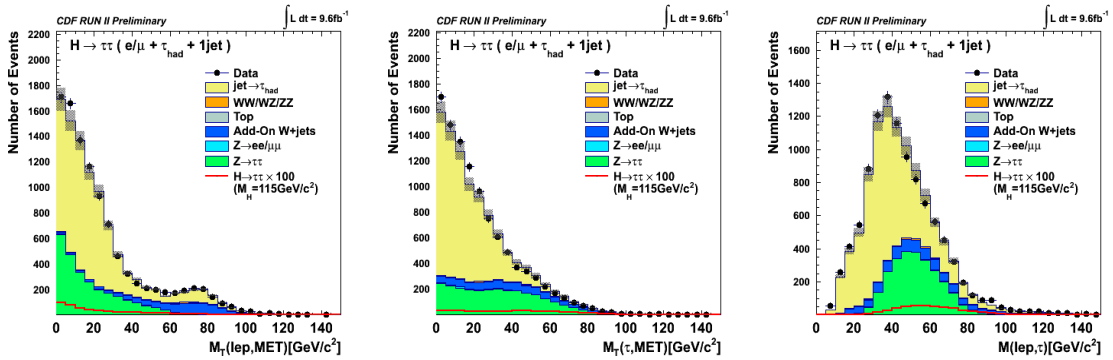

Figure A.10: Transverse mass of lepton and $\not\!\!E_T$, transverse mass of τ_{had} and $\not\!\!E_T$and invariant mass of lepton and τ_{had} for signal region of e/μ + τ_{had} + 1jet channel

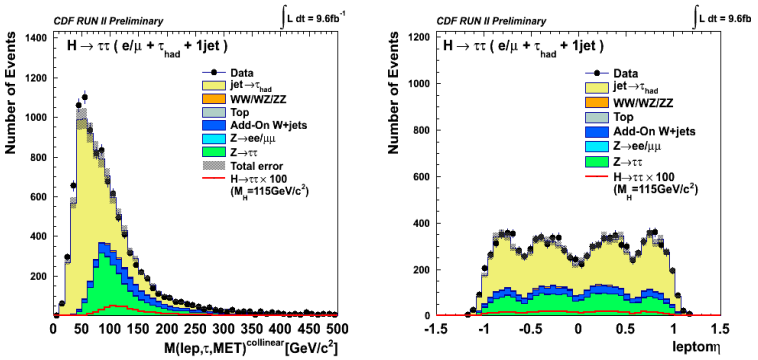

Figure A.11: Invariant Mass of lepton, τ_{had} and $\not\!\!E_T$after collinear approximation and lepton η for signal region of e/μ + τ_{had} + 1jet channel

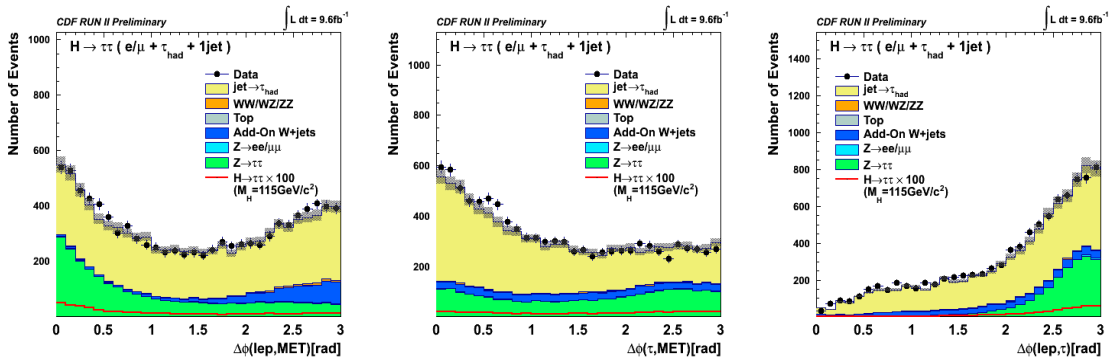

Figure A.12: $\Delta\phi$ between lepton and $\not\!\!E_T$, $\Delta\phi$ between τ_{had} and $\not\!\!E_T$and $\Delta\phi$ between lepton, τ_{had} and $\not\!\!E_T$for signal region of e/μ + τ_{had} + 1jet channel

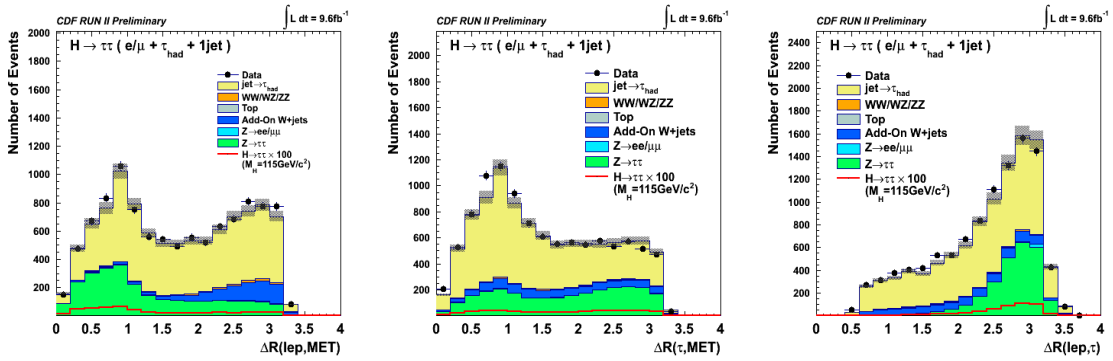

Figure A.13: ΔR between lepton and $\not\!\!E_T$, ΔR between τ_{had} and $\not\!\!E_T$and ΔR between lepton, τ_{had} and $\not\!\!E_T$for signal region of e/μ + τ_{had} + 1jet channel

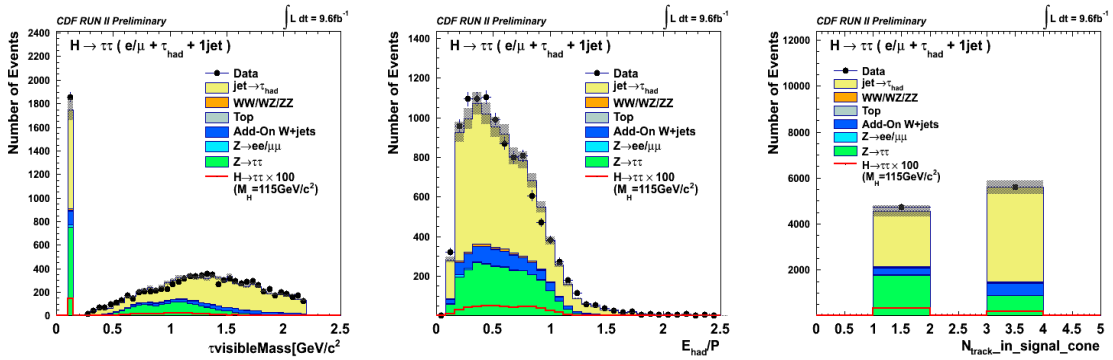

Figure A.14: τ_{had} visible mass, E$_{had}$/P and the number of tracks in signal cone for signal region of e/μ + τ_{had} + 1jet channel

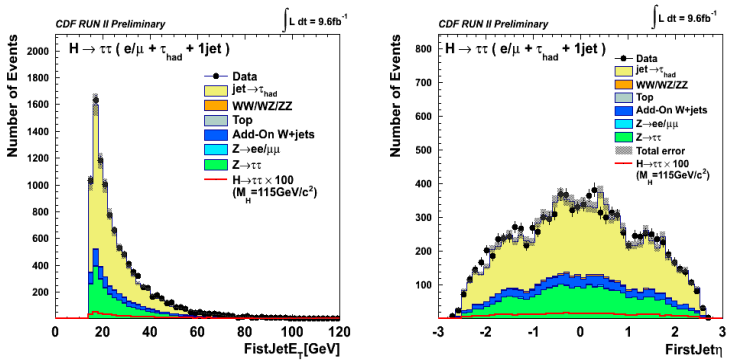

Figure A.15: Leading jet E$_T$ and leading jet η for signal region of e/μ + τ_{had} + 1jet channel

A.3 e/μ + τ_{had} + $\geq$ 2jet channel

Kinematic variables for signal region of e/μ + τ_{had} + $\geq$ 2jet channel are shown in Figure A.16 to Figure A.23.

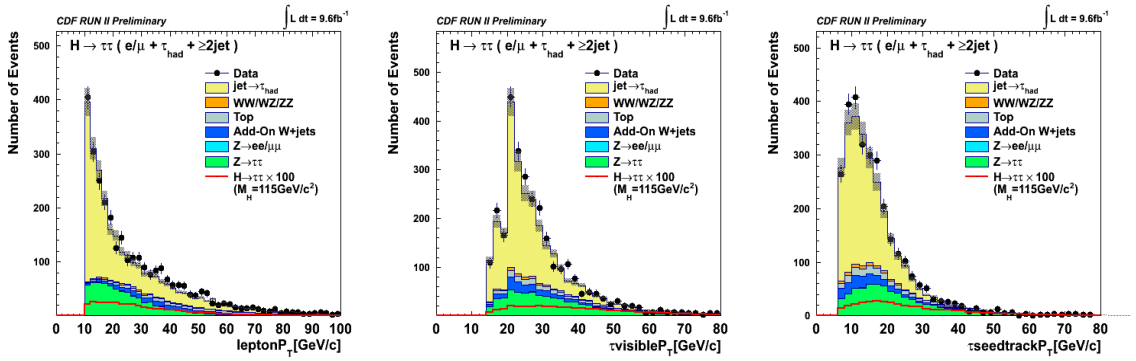

Figure A.16: Lepton P_T, τ_{had} visible P_T and τ_{had} seed track P_T for control region of e/μ + τ_{had} channel

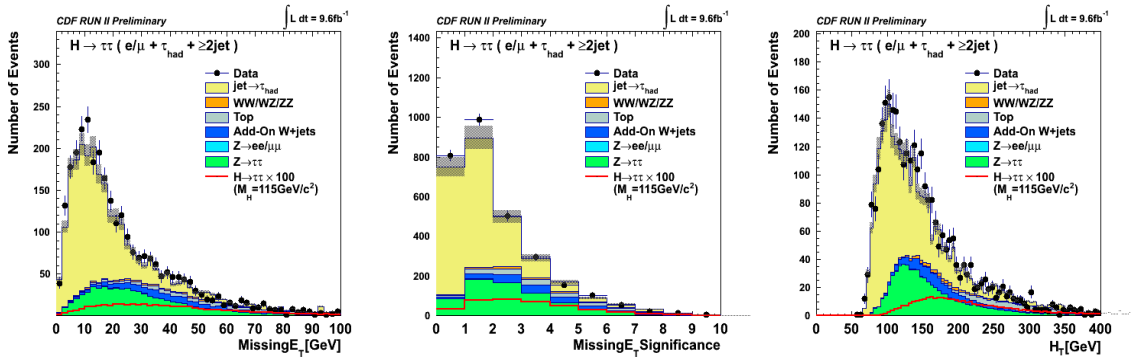

Figure A.17: Mssing E_T, Missing E_T significance and H_T for signal region of e/μ + τ_{had} + $\geq$ 2jet channel

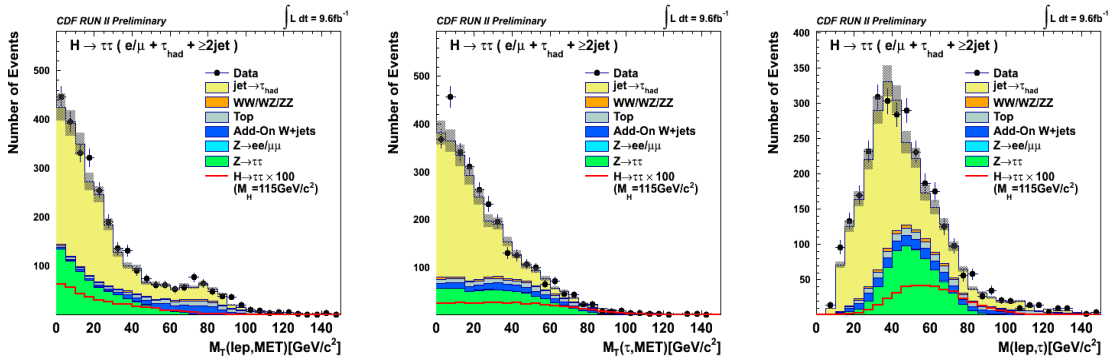

Figure A.18: Transverse mass of lepton and $\not\!\!E_T$, transverse mass of τ_{had} and $\not\!\!E_T$and invariant mass of lepton and τ_{had} for signal region of e/μ + τ_{had} + $\geq$ 2jet channel

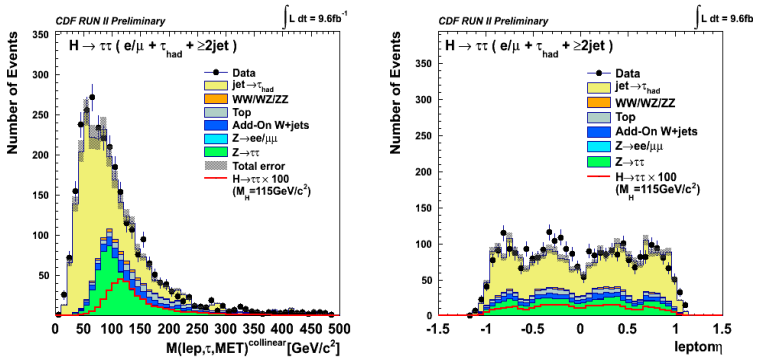

Figure A.19: Invariant Mass of lepton, τ_{had} and $\not\!\!E_T$after collinear approximation and lepton η for signal region of e/μ + τ_{had} + $\geq$ 2jet channel

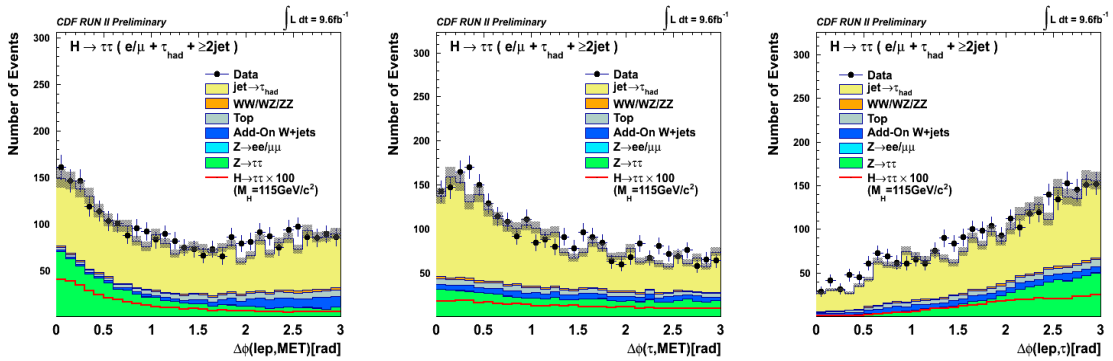

Figure A.20: $\Delta\phi$ between lepton and $\not\!\!E_T$, $\Delta\phi$ between τ_{had} and $\not\!\!E_T$and $\Delta\phi$ between lepton, τ_{had} and $\not\!\!E_T$for signal region of e/μ + τ_{had} + $\geq$ 2jet channel

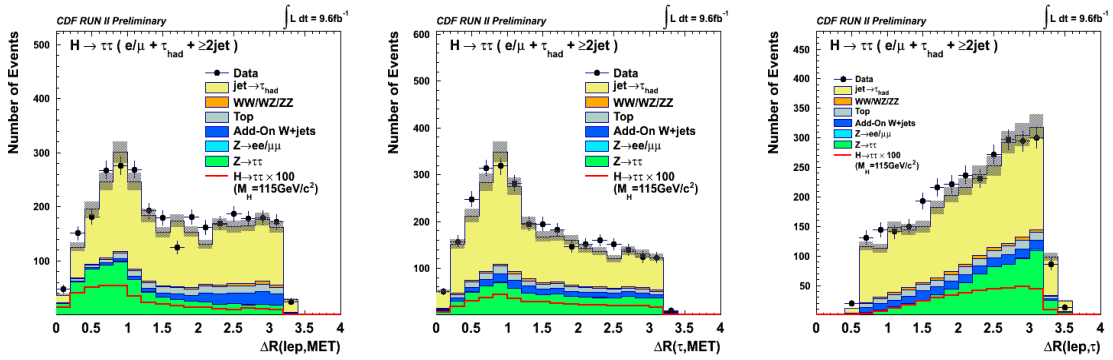

Figure A.21: ΔR between lepton and $\not\!\!E_T$, ΔR between τ_{had} and $\not\!\!E_T$and ΔR between lepton, τ_{had} and $\not\!\!E_T$for signal region of e/μ + τ_{had} + $\geq$ 2jet channel

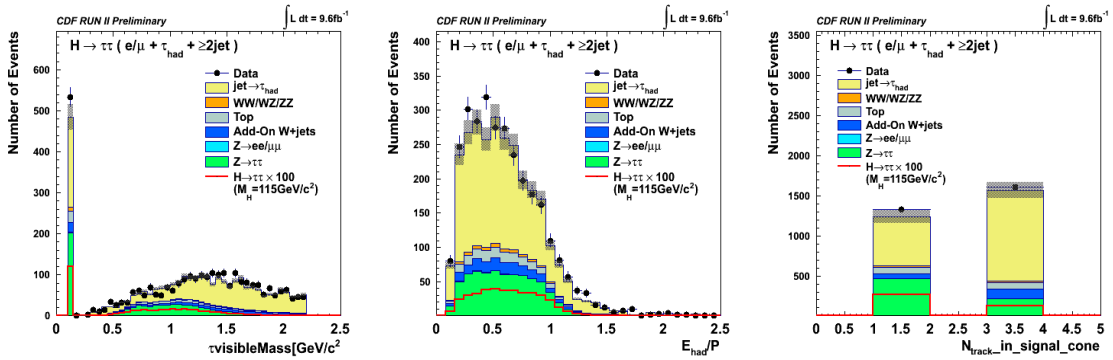

Figure A.22: τ_{had} visible mass, E$_{had}$/P and the number of tracks in signal cone for signal region of e/μ + τ_{had} + $\geq$ 2jet channel

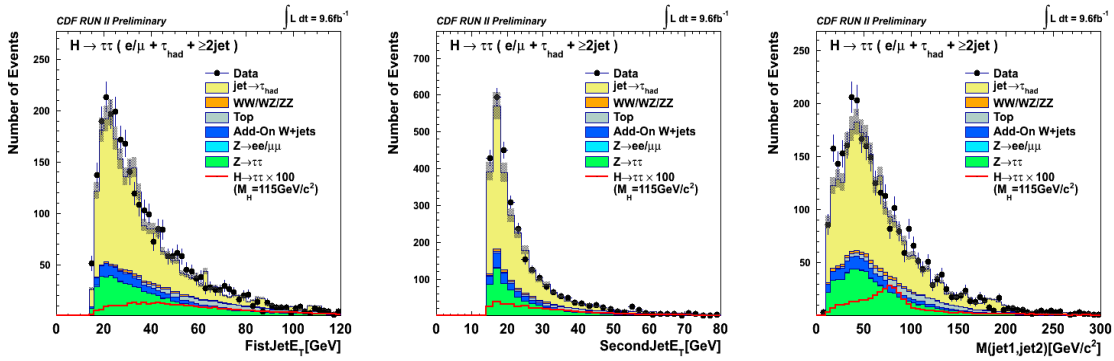

Figure A.23: Leading jet E$_T$, Second jet E$_T$ and invariant mass of di-jet for signal region of e/μ + τ_{had} + $\geq$ 2jet channel

A.4 e + μ + 0jet channel

Kinematic variables for control region of e + μ channel are shown in Figure A.24 to Figure A.28.

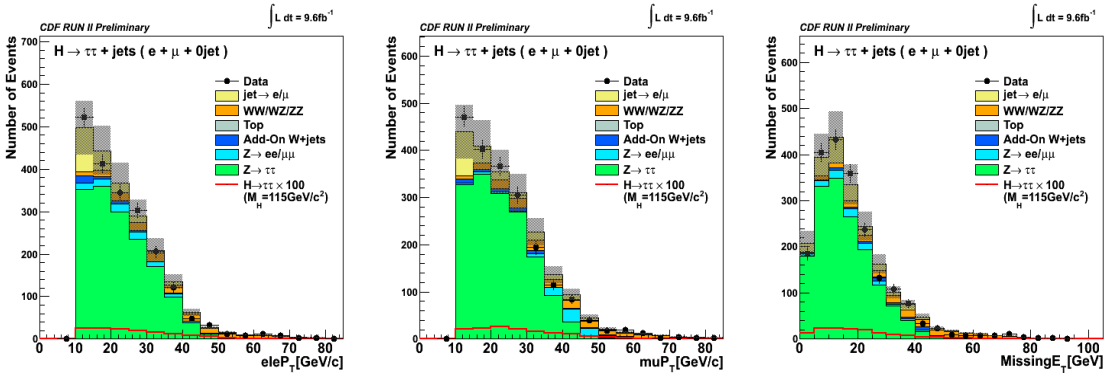

Figure A.24: electron P_T, τ_{had} muon P_T and $\not\!\!E_T$ for control region of e + μ channel

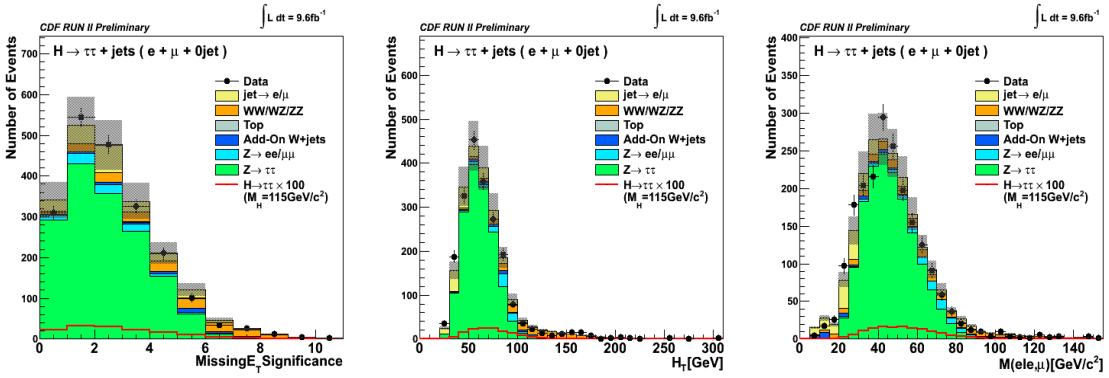

Figure A.25: Missing E_T significance, H_T, invariant mass of electron and μ for control region of e + μ channel

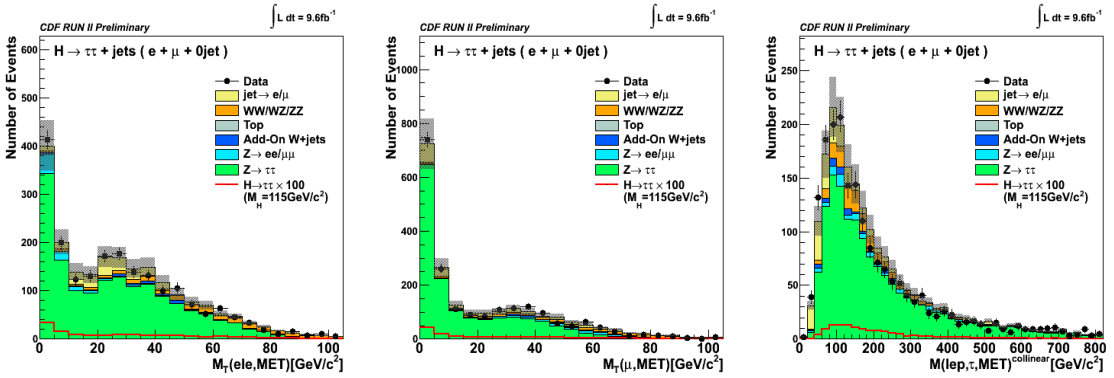

Figure A.26: Transverse mass of electron and $\not\!\!E_T$ and transverse mass of muon and $\not\!\!E_T$ and invariant mass of electron, muon and $\not\!\!E_T$ after collinear approximation for control region of e + μ channel

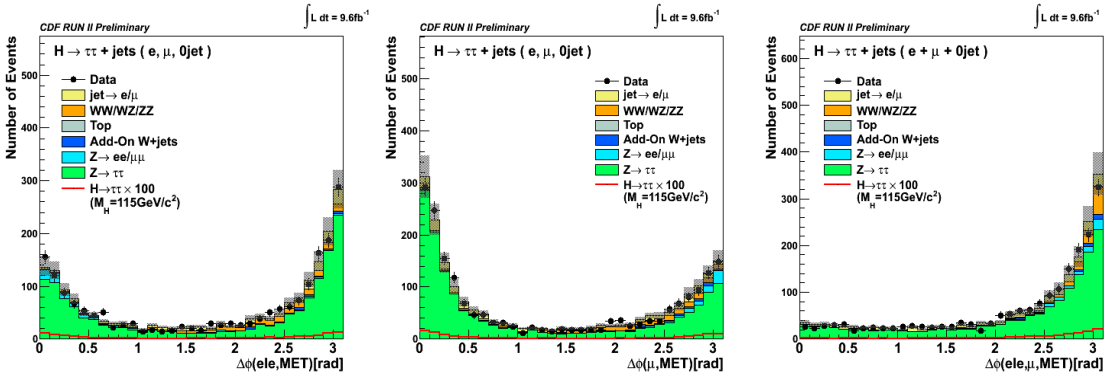

Figure A.27: $\Delta\phi$ between electron and $\not\!\!E_T$, $\Delta\phi$ between muon and $\not\!\!E_T$ and $\Delta\phi$ between electron, muon and $\not\!\!E_T$ for control region of e + μ channel

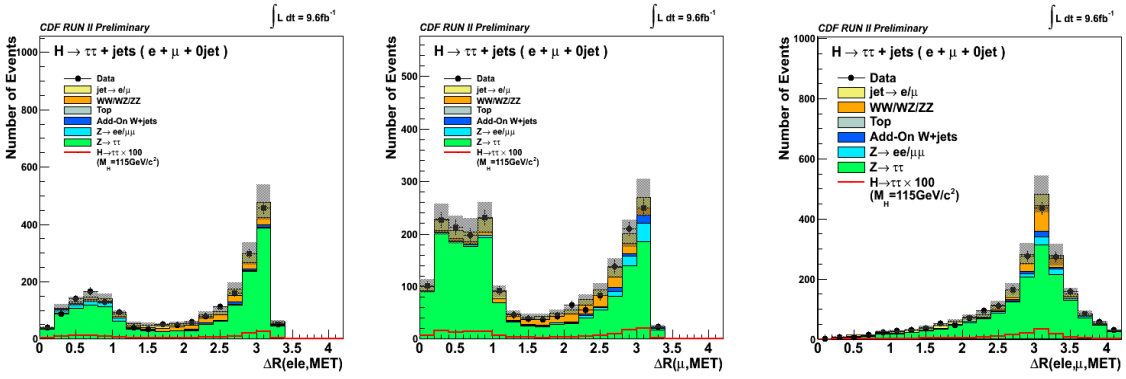

Figure A.28: ΔR between electron and $\not\!\!E_T$, ΔR between muon and $\not\!\!E_T$ and ΔR between electron, muon and $\not\!\!E_T$ for control region of e + μ channel

A.5 e + μ + 1jet channel

Kinematic variables for signal region of e + μ + 1jet channel are shown in Figure A.24 to Figure **??**.

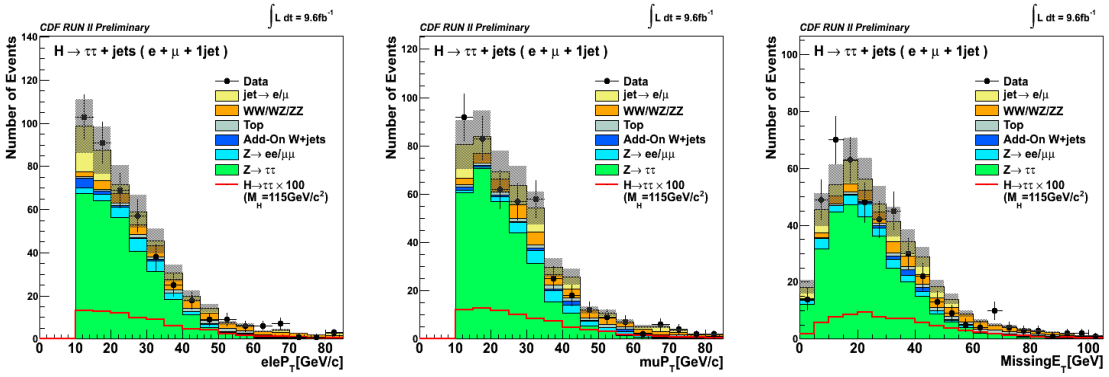

Figure A.29: electron P_T, τ_{had} muon P_T and $\not\!\!E_T$ for signal region of e + μ + 1jet channel

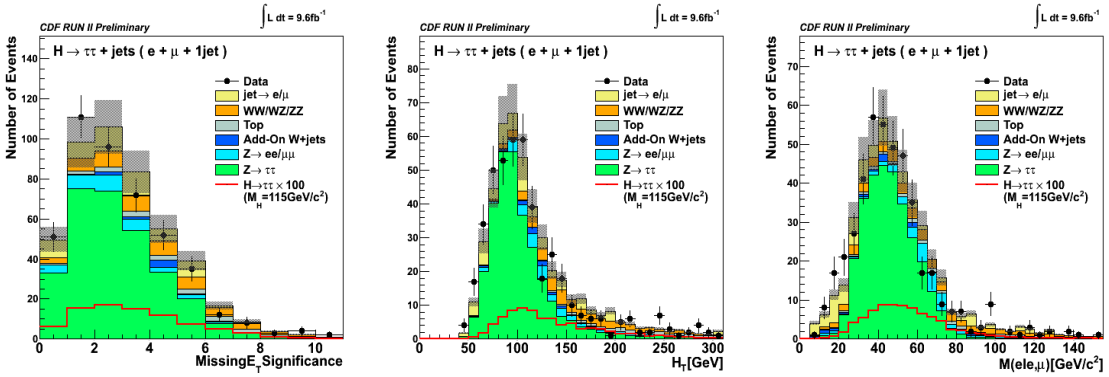

Figure A.30: Missing E_T significance, H_T, invariant mass of electron and μ for signal region of e + μ + 1jet channel

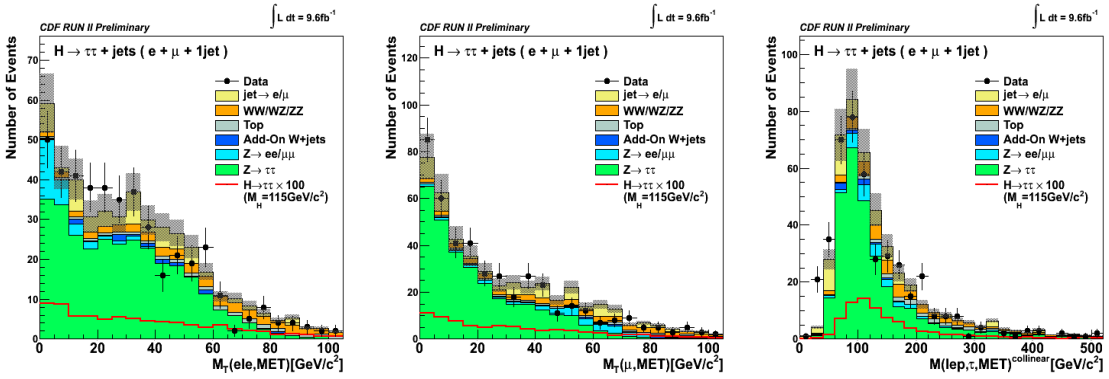

Figure A.31: Transverse mass of electron and $\not\!\!E_T$ and transverse mass of muon and $\not\!\!E_T$ and invariant mass of electron, muon and $\not\!\!E_T$ after colinear approximation for signal region of e + μ + 1jet channel

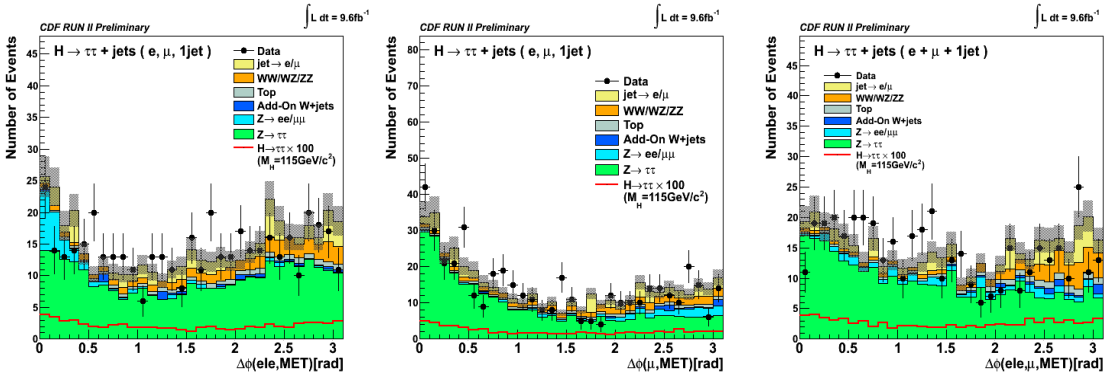

Figure A.32: $\Delta\phi$ between electron and $\not\!\!E_T$, $\Delta\phi$ between muon and $\not\!\!E_T$ and $\Delta\phi$ between electron, muon and $\not\!\!E_T$ for signal region of e + μ + 1jet channel

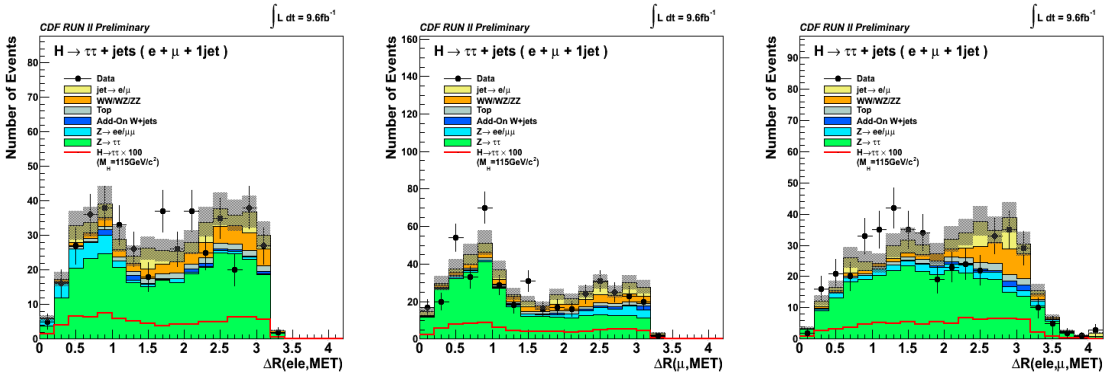

Figure A.33: ΔR between electron and $\not\!\!E_T$, ΔR between muon and $\not\!\!E_T$ and ΔR between electron, muon and $\not\!\!E_T$ for signal region of e + μ + 1jet channel

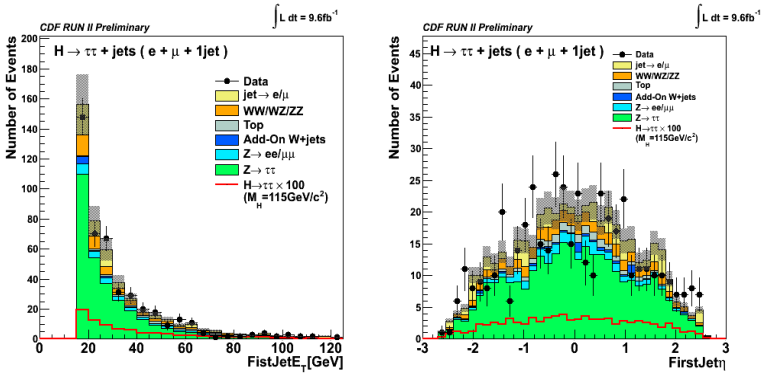

Figure A.34: Leading jet E_T and leading jet η for signal region of e + μ + 1jet channel

A.6 e + μ + $\geq$ 2jet channel

Kinematic variables for signal region of e + μ + $\geq$ 2jet channel are shown in Figure A.35 to Figure A.41.

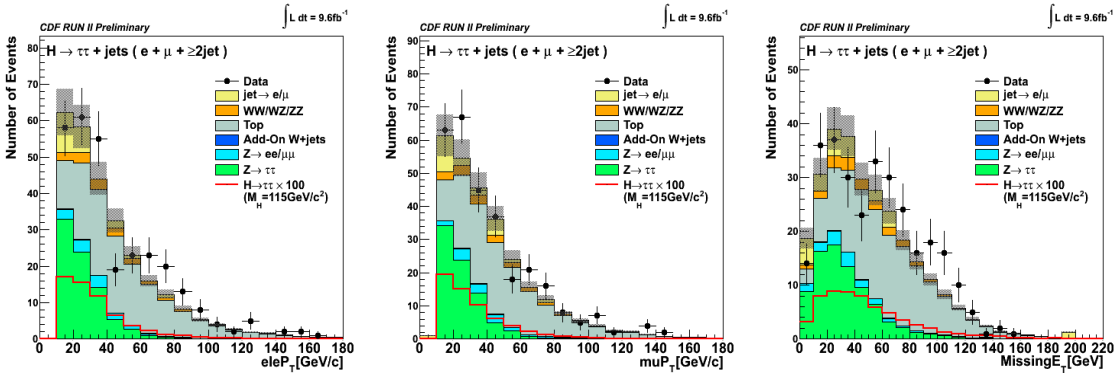

Figure A.35: electron P_T, τ_{had} muon P_T and $\not\!\!E_T$ for signal region of e + μ + $\geq$ 2jet channel

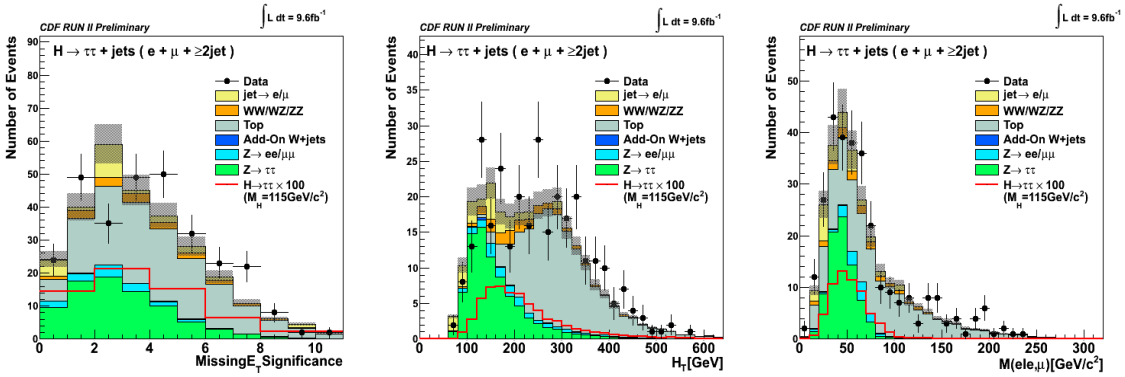

Figure A.36: Missing E_T significance, H_T, invariant mass of electron and μ for signal region of e + μ + $\geq$ 2jet channel

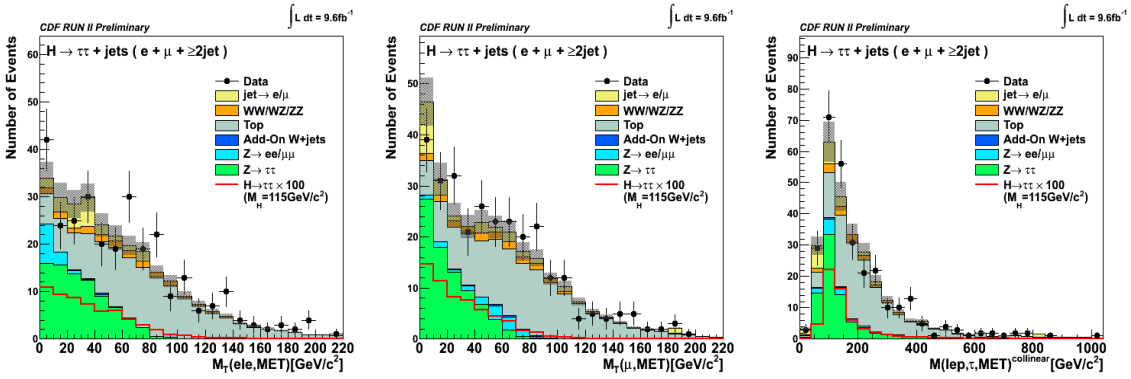

Figure A.37: Transverse mass of electron and $\not\!\!E_T$ and transverse mass of muon and $\not\!\!E_T$ and invariant mass of electron, muon and $\not\!\!E_T$ after colinear approximation for signal region of e + μ + $\geq$ 2jet channel

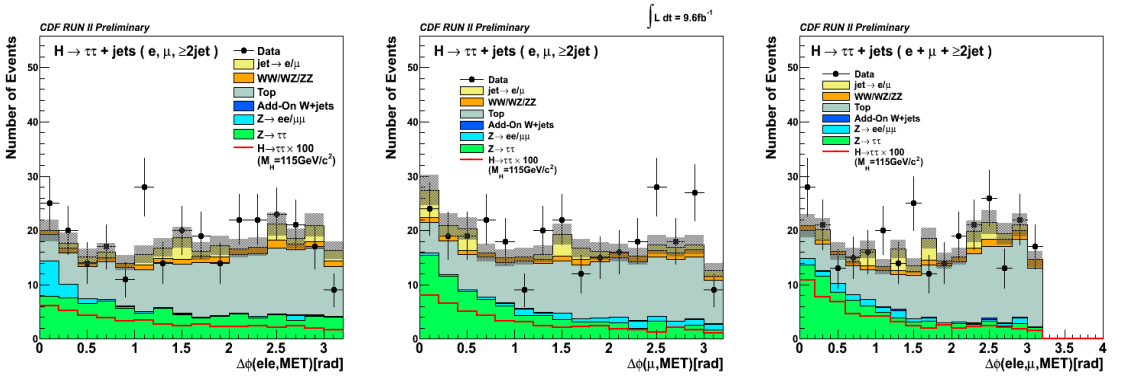

Figure A.38: $\Delta\phi$ between electron and $\not\!\!E_T$, $\Delta\phi$ between muon and $\not\!\!E_T$ and $\Delta\phi$ between electron, muon and $\not\!\!E_T$ for signal region of e + μ + $\geq$ 2jet channel

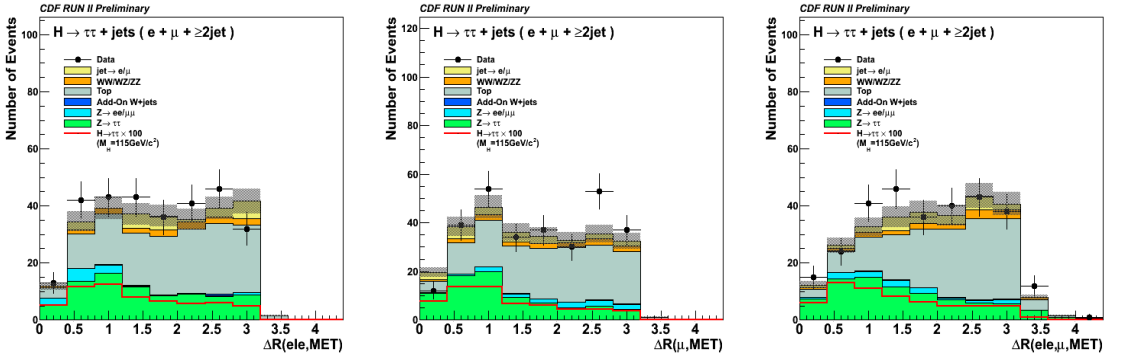

Figure A.39: ΔR between electron and $\not\!\!E_T$, ΔR between muon and $\not\!\!E_T$ and ΔR between electron, muon and $\not\!\!E_T$ for signal region of e + μ + $\geq$ 2jet channel

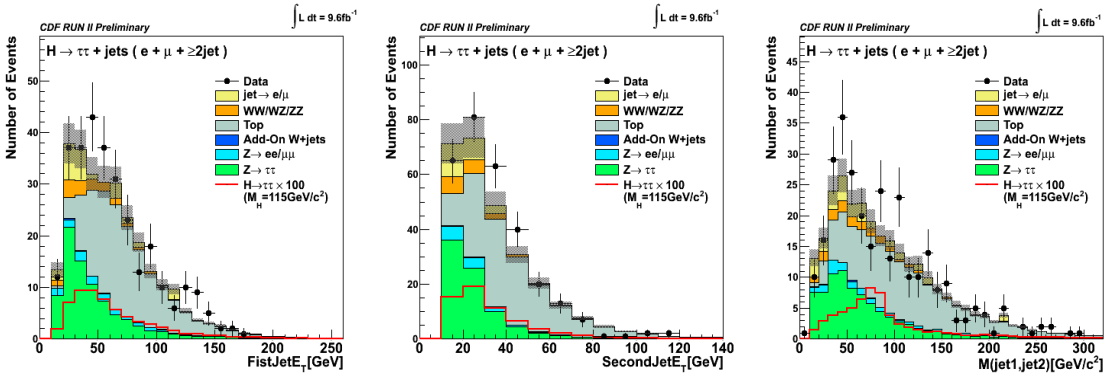

Figure A.40: Leading jet E_T, second jet E_T and invariant mass of dijet for signal region of e + μ + $\geq$ 2jet channel

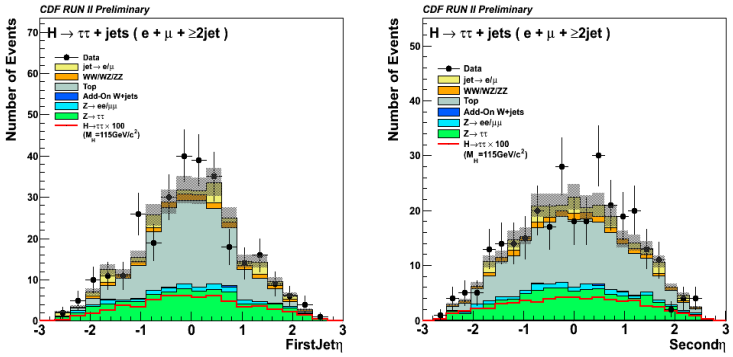

Figure A.41: Leading jet η and second jet η for signal region of e + μ + $\geq$ 2jet channel

Bibliography

[1] Guido Altarelli. Partons in Quantum Chromodynamics. *Phys. Rept.*, 81:1, 1982.

[2] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

[3] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[4] Abdus Salam and John Clive Ward. Electromagnetic and weak interactions. *Phys. Lett.*, 13:168–171, 1964.

[5] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, 1964.

[6] Peter W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12:132–133, 1964.

[7] Julien Baglio and Abdelhak Djouadi. Predictions for higgs production at the tevatron and the associated uncertainties. *JHEP*, 1010:064, 2010.

[8] A. Denner, S. Heinemeyer, I. Puljak, D. Rebuzzi, and M. Spira. Standard model Higgs-boson branching ratios with uncertainties. *EUROPEAN PHYSICAL JOURNAL C*, 71(9), 2011.

[9] T.Aaltonen, et al. Inclusive Search for Standard Model Higgs Boson Production in the WW Decay Channel using the CDF II Detector. *Phys. Rev. Lett.*, 104:061803, 2010.

[10] T.Aaltonen et al. Search for Standard Model Higgs Boson Production in Association with a W Boson at CDF. 2011.

[11] the CDF and D0 collaborations and the Tevatron New Phenomena and Higgs Working group. Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.6 fb-1 of Data. 2011.

[12] R. Barate et al. Search for the standard model Higgs boson at LEP. *Phys. Lett.*, B565:61–75, 2003.

[13] ATLAS collaboration. Combination of Higgs Boson Searches with up to 4.9 fb-1 of pp Collisions Data Taken at a center-of-mass energy of 7 TeV with the ATLAS Experiment at the LHC. *ATLAS CONF*, 2011-163, 2011.

[14] CMS collaboration. Combination of CMS searches for a Standard Model Higgs boson. *CMS PAS*, HIG-11-03, 2011.

[15] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. *CERN-2011-002*, CERN, Geneva, 2011.

[16] C. S. Hill et al. L00: Operational Experience and Performance of the CDFII Silicon Detector. *Nucl. Instrum. Meth.*, A530:1–6, 2004.

[17] A. Sill et al. SVX-II: CDF Run II Silicon Tracking Projects. *Nucl. Instrum. Meth.*, A447:1–8, 2000.

[18] T.Affolde et al. ISL: Intermediate Silicon Layers Detector for the CDF Experiment. *Nucl. Instrum. Meth.*, A453:84–88, 2000.

[19] T.Affolde et al. COT Central Outer Tracker. *Nucl. Instrum. Meth.*, A526:249, 2004.

[20] S. Kuhlmann et al. The CDF Calorimetry Upgrade for Run IIb. *Nucl. Instrum. Meth.*, A518:39–41, 2004.

[21] A. Artikov et al. Design and construction of new central and forward muon counters for CDF II. *Nucl. Instrum. Meth.*, A538:358–371, 2002.

[22] R. Downing et al. Track Extrapolation and Distribution for the CDF-II Trigger System. *Nucl. Instrum. Meth.*, A570:36–50, 2007.

[23] A. Abulencia et al. Measurements of inclusive W and Z cross sections in p anti-p collisions at s**(1/2) = 1.96-TeV. *J.Phys.G*, G34:2457–2544, 2007.

[24] A. Bhatti et al. Determination of the jet energy scale at the collider detector at Fermilab. *Nucl. Instrum. Meth.*, A566:375–412, 2006.

[25] Y-K.Kim K.Yorita. Search for the standard model higgs boson using tau. *CDF/ANAL/EXOTIC/CDFR/9179*, 2008.

[26] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.

[27] Christopher J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2:121–167, 1998.

[28] Martin M. Block, Loyal Durand, and Douglas W. McKay. Analytic treatment of leading-order parton evolution equations: theory and tests. *PHYS.REV.D*, 79:014031, 2009.

[29] A.D. Martin, W.J. Stirling, R.S. Thorne, and G. Watt. Update of parton distributions at NNLO. *Phys.Lett.*, B652:292–299, 2007.

[30] Thomas Junk. Confidence Level Computation for Combining Searches with Small Statistics. *Nucl. Instrum. Meth.*, A434:435–443, 1999.