This Dissertation

entitled

Search for lepton flavor violating decays of the Higgs boson

typeset with NDdiss2$_\varepsilon$ v3.0 (2005/07/27) on February 21, 2018 for

Fanbo Meng

This LaTeX 2$_\varepsilon$ classfile conforms to the University of Notre Dame style guidelines established in Spring 2004. However it is still possible to generate a non-conformant document if the instructions in the class file documentation are not followed!

> Be sure to refer to the published Graduate School guidelines at `http://graduateschool.nd.edu` as well. Those guidelines override everything mentioned about formatting in the documentation for this NDdiss2$_\varepsilon$ class file.

It is YOUR responsibility to ensure that the Chapter titles and Table caption titles are put in CAPS LETTERS. This classfile does *NOT* do that!

This page can be disabled by specifying the "`noinfo`" option to the class invocation. (i.e.,`\documentclass[...,noinfo]{nddiss2e}`)

This page is *NOT* part of the dissertation/thesis, but MUST be turned in to the proofreader(s) or the reviwer(s)!

NDdiss2$_\varepsilon$ documentation can be found at these locations:

`http://www.gsu.nd.edu`
`http://graduateschool.nd.edu`

Search for lepton flavor violating decays of the Higgs boson

A Dissertation

Submitted to the Graduate School

of the University of Notre Dame

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

by

Fanbo Meng,

_____

Colin Jessop, Director

_____

, Director

Graduate Program in Physics

Notre Dame, Indiana

February 2018

Search for lepton flavor violating decays of the Higgs boson

Abstract

by

Fanbo Meng

this going to put with the abstract, a summary of the whole analysis

NEW DEDICATION NAME

To be written

CONTENTS

iii

vi

TABLES

PREFACE

probably I will write some preface, but it depends $H \to \mu\tau_h$

ACKNOWLEDGMENTS

this part is going to be filled with acknowledge

CHAPTER 1

Introduction

1.1 Pre-LHC era on lepton Flavor violation(LFV) search and results

1.2 Overview of CMS Run1 and Run2 direct search on LFV

CHAPTER 2

Theory

2.1 Standard model

2.2 LFV in beyond standard model theories

CHAPTER 3

LHC and CMS experiment

3.1 LHC accelerator

3.2 CMS experiment

3.2.1 Tracker

3.2.2 ECAL

3.2.3 HCAL

3.2.4 Muon System

3.2.5 Trigger

probably I have to put something there

3.2.5.1 Level 1 Trigger

probably I have to put something there2

3.2.5.2 High Level Trigger

CHAPTER 4

Datasets

4.1 Datasets used in LFV analysis

4.2 Event reconstruction

4.2.1 Muon reconstruction and selection criteria

With the PF muon as the input, muons used in the analysis are further catego-
rized into different identification, isolation categories. Muon ID suggested in CMS
RunII, for the LHC data 2016, running period BCDEF, ICHEP medium muon ID
is applied(table 4.2), for running period G and H, also the monte Carlo samples,
standard medium muon ID(table 4.3)are applied to achieve the best performance
for muon identification.

4.2.2 Tau lepton reconstruction

In Run I CMS experiment, tau lepton are constructed with hadrons plus
strips(HPS) algorithm. In general, HPS starts with PF jets which are recon-
structed with $anti - k_T$, as the initial seeds. π_0 components from the τ hadronic
decays are first constructed and combined with the charge hadrons parts, to iden-

MC simulations	Cross section
/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	4954.0 pb (LO)
/DY1JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1012.5 pb (LO)
/DY2JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	332.8 pb (LO)
/DY3JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	101.8 pb (LO)
/DY4JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	54.8 pb (LO)
/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	18610 pb (LO)
/TT_TuneCUETP8M1_13TeV-powheg-pythia8	831.76 pb
/WJetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	50380 pb (LO)
/W1JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	9644.5 pb (LO)
/W2JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	3144.5 pb (LO)
/W3JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	954.8 pb (LO)
/W4JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/	485.6 pb (LO)
/WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8	3.05 pb
/WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.71 pb
/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595 pb
/ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.85 pb
/ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.85 pb
/ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	80.95 pb
/ST_t-channel_top_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	136.02 pb
/WWTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	1.212 pb
/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.22 pb
/VVTo2L2Nu_13TeV_amcatnloFXFX_madspin_pythia8	11.95 pb
/WGToLNuG_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	489.0 pb
/WGstarToLNuMuMu_012Jets_13TeV-madgraph	2.793 pb
/WGstarToLNuEE_012Jets_13TeV-madgraph	3.526 pb
/GluGluHToTauTau_M125_13TeV_powheg_pythia8	48.58 pb X 0.0627
/VBFHToTauTau_M125_13TeV_powheg_pythia8	3.782 pb X 0.0627
/GluGluHToWWTo2L2Nu_M125_13TeV_powheg_pythia8	48.58 pb X 0.0227
/VBFHToWWTo2L2Nu_M125_13TeV_powheg_pythia8	3.782 pb X 0.0227
/ZHToTauTau_M125_13TeV_powheg_pythia8	0.8839 pb X 0.0627
/WminusHToTauTau_M125_13TeV_powheg_pythia8	1.373 pb X 0.0627
/WplusHToTauTau_M125_13TeV_powheg_pythia8	1.373 pb X 0.0627
/ttHJetToTT_M125_13TeV_amcatnloFXFX_madspin_pythia8	0.507 pb X 0.0627

TABLE 4.1

TABLE 4.2

Muon ID used in the analysis, for the LHC data 2016, running period
BCDEF.

ICHEP mediumID description	Technical description
Loose muon ID	PFLoose Muon
Fraction of valid tracker hits	> 0.49
1.Good Global muon	Global muon
	Normalized global-track $\chi^2 < 3$
	Tracker-Standalone position match < 12
	kick finder < 20
	Segment compatibility > 0.303
2. Tight segment compatibility	Segment compatibility > 0.451

TABLE 4.3

Muon ID used in the analysis, for the LHC data 2016, running period G and H, also the monte Carlo samples.

Standard mediumID description	Technical description
Loose muon ID	PFLoose Muon
Fraction of valid tracker hits	> 0.8
1.Good Global muon	Global muon
	Normalized global-track $\chi^2 < 3$
	Tracker-Standalone position match < 12
	kick finder < 20
	Segment compatibility > 0.303
2. Tight segment compatibility	Segment compatibility > 0.451

tify different τ decay modes and calculate τ four-momentum and other quantities [2].

Photon conversions and the bremsstrahlung of electron/positron when traveling inside the CMS detector are well treated by the HPS algorithm. These phenomenons broaden the signature of the tau decay. With PF jets as input, the algorithm constructs strips out of electromagnetic particles and starts by taking the strip in which contains the most energetic electromagnetic particle as the center one. With the center strip, a window of the size $\Delta\eta = 0.05$ and $\Delta\phi = 0.2$ is taken. Within this window, if other charged particles are found, they are associated with the strip. The position of the strip is taken and four momentum of the strip is calculated. This procedure is repeated, until no strips can be constructed. The selected strips are required to have $P_T^{strip} > 1 GeV$. The following decay topologies are taking into account by HPS:

- one charged particle without any strip, $h^{\pm}$ and the case when π^0 is not energetic enough to form a strip

- one charged particle plus one strip

- one charged particle plus two strips

- three charged partibles.

All of the charged hadrons and strips are required to be contained in the $\Delta R = 2.8/P_T^{\tau_h}$ core, where the $P_T^{\tau_h}$ is the reconstructed τ_h transverse momentum and ΔR is defined as $\Delta R = \sqrt{(\Delta\phi^2 + \Delta\eta^2)}$. The τ_h candidate is also required to match the direction of the seed PF jet within $\Delta R = 0.1$. Assuming all of the charged hadrons to be pions and taking in the associated strips, the HPS

8

algorithm requires that different decay topologies meet the intermediate meson mass as listed in Table. 4.4.

The cut based τ_h isolation discriminant required that the PF charged particles and photons to be considered in the isolation variable have $p_T > 0.5$ GeV and within an isolation cone $\Delta R = 0.5$ in τ_h direction. The particles that constituent τ_h are excluded from the summation. The effect of charged particle from pileup is eliminated by considering on the charged particle oriented from the τ_h production vertex with in $D_z = 0.2$ cm and $\Delta r = 0.03$ cm. The effect of pileup on the isolation of the photons on the strips is estimated by summing the charged particles that are not oriented from τ_h decay primary vertex, within $\Delta R = 0.8$ cm in the direction of and have the impact parameter $D_z > 0.2$ cm. Then a factor $\Delta\beta$ is multiplied to the p_T sum. The isolation variable is defined as in Equation. 4.1.

$$I_\tau = \sum p_T^{\text{charged}}(d_z < 0.2 \text{ cm}) + \max(0, \sum p_T^\gamma - \Delta\beta \sum p_T^{\text{charged}}(d_z > 0.2 \text{ cm}))$$

(4.1)

Tight, loose, medium working points(WP) for the tau isolation discriminants. The exact of the energy selection is suggested by the study of QCD dijet events, by requiring the I_τ in Equation. eq:taucutiso1 to have different values. The loose cut brings in approximate 1% of fake τ from jets.

In CMS Run II, Tau reconstruction algorithm HPS has been improved [3]. The major improvement lies in Dynamic strip instead of fix size strip. Tau decay products can also affect the isolation. Charged pions in tau decay products experience nuclear interaction with tracker materials, which can results in low P_T secondary particles. Photons from the neutral pion decay can also go through

TABLE 4.4

Dominant hadronic τ lepton decays branching fractions and the associated intermediate resonance. The h stands for both π and K. The table is symmetric under charge conjugation.

Decay mode	Resonance	Mass (MeV/c^2)	Branching fraction(%)
$\tau^- \to h^- v_\tau$			11.6%
$\tau^- \to h^- \pi^0 v_\tau$	ρ^-	770	26.0%
$\tau^- \to h^- \pi^0 \pi^0 v_\tau$	α_1^-	1200	9.5%
$\tau^- \to h^- h^+ h^- v_\tau$	α_1^-	1200	9.8%
$\tau^- \to h^- h^+ h^- \pi^0 v_\tau$			4.8%

pair production into e^+e-, which further spread because of bremsstrahlung and the magnetic field. Broadening the strip is need in these cases in order to better cover the tau decay production. On the other hand, if the tau is boosted, high P_T decay products tends to be more concentrate and smaller strip size will be better. Similar to RunI tau reconstruction, the algorithm starts with hightest p_T charged particle as seeds for the strip. Starting from the seed strip, a window in η and ϕ direction is set.

$$\delta\eta = f(P_T^\gamma) + f(P_T^{strip}) \qquad f(P_T) = 0.2 \cdot P_T^{-0.66}$$
$$\delta\phi = g(P_T^\gamma) + g(P_T^{strip}) \qquad g(P_T) = 0.35 \cdot P_T^{-0.71}$$

The window is determined from single τ gun MC simulation. 95% of the decay product will be covered in that range. The upward and downward limits for η is 0.15 and 0.05, for ϕ the range is 0.3 and 0.05. The position of strip is set as p_T weighted average against all of the objects.

$$\eta_{strip} = \frac{1}{P_T^{strip}} \cdot \sum P_T^\gamma \cdot \eta_\gamma$$

$$\phi_{strip} = \frac{1}{P_T^{strip}} \cdot \sum P_T^\gamma \cdot \phi_\gamma$$

Construct the strip until no seed strip can be found. After the construction of the τ lepton, for different decay mode, m_τ is required to lie in different mass windows [4]. The conditions of different hadronic decay mode mass window are listed in the Table. 4.5. With respect to RunI, the difference in mass window is δm, which originates from dynamic clustering. δm is calculated as:

$$\delta m = \sqrt{\left(\frac{\partial m_\tau}{\partial \eta_{\text{strip}}} \cdot f(p_T^{\text{strip}})\right)^2 + \left(\frac{\partial m_\tau}{\partial \phi_{\text{strip}}} \cdot g(p_T^{\text{strip}})\right)^2}$$

with:

$$\frac{\partial m_\tau}{\partial \eta_{strip}} = \frac{P_z^{strip} \cdot E_\tau - E_{strip} \cdot P_z^\tau}{m_\tau}$$

$$\frac{\partial m_\tau}{\partial \phi_{\text{strip}}} = \frac{-(P_y^\tau - P_y^{\text{strip}}) \cdot P_x^{\text{strip}} + (P_x^\tau - P_x^{\text{strip}}) \cdot P_y^{\text{strip}}}{m_\tau}$$

11

TABLE 4.5

τ hadronic decay mode hypothesis signatures compatibility tests. m_τ is required to be in the mass window

Decay mode	Mass window
$\tau^- \to h^- \pi^0 v_\tau$	$0.3 - \delta m_\tau < m_\tau < 1.3 \cdot \sqrt{p_T/100} + \delta m_\tau$
$\tau^- \to h^- \pi^0 \pi^0 v_\tau$	$0.4 - \delta m_\tau < m_\tau < 1.2 \cdot \sqrt{p_T/100} + \delta m_\tau$
$\tau^- \to h^- h^+ h^- v_\tau$	$0.8 - \delta m_\tau < m_\tau < 1.5 + \delta m_\tau$

In current algorithm, $\tau^- \to h^- h^+ h^- v_\tau$ is not included, because of the jets contamination. This hadronic τ decay mode composed of 4.8% of total branching fraction. The $h^- \pi^0$ and $h^- \pi^0 \pi^0$ are analyzed together, which is referred as $h^- \pi^0$.

The analysis with 2016 datasets, MVA based τ isolation criteria is used, which keeps high identification efficiency while maintains relatively low fake rate compared with cut based discriminator. A Boosted Decision Tree(BDT) has been used in the training of the isolation variable. With BDT, the isolation variable shows a good distinguishing power against jets. Various variables have been used as BDT inputs. The variables are isolation variable(I_τ), impact parameter from highest p_T track of τ_h candidate, τ_h decay mode information, shape variables like ΔR, $\Delta \eta$, $\tau-$lifetime information and photon electron multiplicity, more of the exact variables used are discussed in [3, 4]. BDT uses these variables to distinguishing τ_h dacay($H \to \tau\tau$) from jets, which can be the decay products from quarks and gluons(QCD MC).

BDT method also used in the tau discriminating again electron training. The algorithm utilizes the variables that sensitive to the energy deposit in Ecal and Hcal, the electron bremsstrahlung, overall particle multiplicity and difference in electromagnetic and hadronic showers. Detail list of variable can be find in [3, 4].

Tau signals can be faked by muons, especially in τ_h decay mode $h^{\pm}$. Tau against muon cut based muon discriminant is set by checking if there are signals in the muon system within $\Delta R = 0.3$ of the τ_h direction or if the energy sum from Ecal and Hcal is less than 20% of the total τ energy. If less than two hits are found in the muon system, then it passes the loose working point. If no hits are found in the muon system, then this is the tight working point.

4.3   Electron reconstruction

4.4   Event simulation

CHAPTER 5

LFV event selection

In both 8 TeV $H \to e\tau_h$ and 13 TeV $H \to \mu\tau_h$ analyses, events are selected in several steps. A loose selection selects on the different IDs, energy, geometry parameters of the analysis related objects and followed by a tighter set of selection criteria in which selection requirements are placed on the kinematics variables and fits on variable M_{col} in both $H \to e\tau_h$ and $H \to \mu\tau_h$ analysis. This selection sequence is referred as M_{col} fit analysis. In $H \to \mu\tau_h$, there is another selection sequence. After placing the loose selection, a multivariate analysis with Boosted decision tree (BDT) is exploited and fits on the BDT discriminator. This is referred as the BDT fit analysis and provides more sensitive results in H $\to \mu\tau_h$ analysis.

5.1 $H \to \mu\tau_h$

5.1.1   Loose selection

Tau leptons from signal events decay hadronically. SM higgs is much heavier than its decay products μ and τ. μ and τ leptons are expected to have high P_T. The decay products from signal events are boosted, a cut on $\Delta R = \sqrt{(\Delta phi^2 + \Delta eta^2)}$, $\Delta R > 0.3$ is applied. μ and τ candidates are required to have opposite sign of charges as Higgs has no electric charge. Further, events with additional μ and

14

τ that pass a loose selection, events with jets that are identified by the combined secondary vertex(CSVv2) b-tagging algorithm [5] as a b quark jets are discarded. The trigger HLT_IsoMu24 or HLT_IsoTkMu24 used in the analysis selects isolated muons that have energy higher than 24 GeV at HLT level. An further P_T cut on reconstructed μ, $P_T > 26$ GeV and $|\eta| < 2.4$ are required. Muons are required to pass the recommended Medium muon ID and tight cut based isolation $I_{rel}^{\mu} < 0.15$.

Hadronic taus are required to have $P_T > 30$ GeV,$|\eta| < 2.3$, passing old tau decay mode finding, the MVA based tight tau isolation ID and tau discriminators against electrons and muons.

Events in the analysis are divided into four categories based on the number of jets. In 2-jets category, it is furthered divided into 2 categories, 2-jet gluon gluon fusion higgs production(ggH) category and 2-jet vector boson fusion(VBF) category based on the value of 2 jets invariant mass(M_{jj}) . The 0-jet category enhances ggH production mode. In 1-jet category, the dominant signal production mode is also ggH, but with a boosted jet associated with the production and VBF higgs production also contributes this category. In the 2-jet ggH category, signal evens mainly come from ggH, while in 2-jet VBF, VBF higgs is the dominant production channel. The following is a more detailed list of the selection condition in each categories.

0-jet: No events have jets pass the loose PF ID and with jet $P_T > 30$ GeV, $|\eta| < 4.7$.

1-jet: Events with one jet passes losse PF ID and jet $P_T > 30$ GeV, $|\eta| < 4.7$.

2-jets ggH: Events have two jets passing loose PF ID, $P_T > 30$ GeV, $|\eta| < 4.7$ and a requirement on the invariant mass of the two jets, $M_{jj} < 550$ GeV.

2 jets VBF: Events with two jets pass loose PF ID. Jets $P_T > 30$ GeV, $|\eta| < 4.7$ and

M$_{jj}$ > 550 GeV are required.

The threshold on M$_{jj}$ has been optimized to give the best expected exclusion limits.

5.1.2 Cut-based analysis

After the loose selection and the categorization, a further cut-based selection is applied. Variables that can help distinguish signal from background are P_T^μ, P_T^τ and $M_T(\tau_h)$. The lepton P_T variables are very powerful background discriminant variables, but it will also cause the problem that signal picks under backgrounds. Leptons from signal process incline to have higher P_T values, by cutting tighter on the lepton P_T, more background events can be removed. However this will also reshape some of the backgrounds, making them peak closer under the signal so that signal processes will be affected more by the background statistics fluctuation. In the $H \to e\tau_h$ analysis, the effect of cutting tighter on lepton P_T will be shown. In $H \to \mu\tau_h$ search, lepton P_T variables are kept at loose values and optimizes on other variables to achieve better signal significance.

Optimization is proceeded with only Monte Carlo samples so as not to double use data. In H $\to \mu\tau_h$ channel, the variables tuned are M$_{jj}$ and $M_T(\tau_h)$. Cuts have been optimized to have the most stringent expected limits with the Asimov dataset. If looser values of the cuts give the same expected limits as the tighter ones, then the looser cut values are chosen to have more statistics. Examples of optimization by checking the limits are shown in Figure. 5.1 and 5.2, in which the optimization of $M_T(\tau_h)$ and M$_{jj}$ are shown. A full summary of the cuts optimized for H $\to \mu\tau_h$ is shown in Table. 5.1
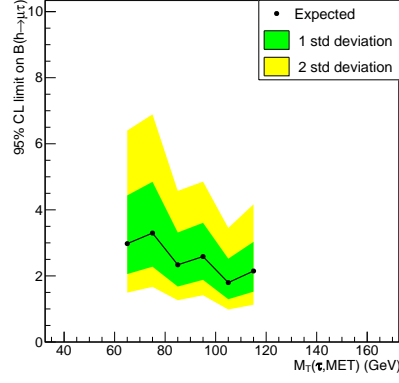
(a) 0 jet
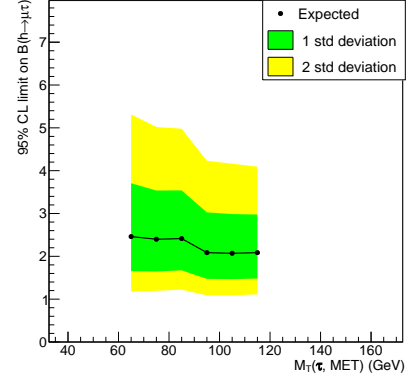
(b) 1 jet
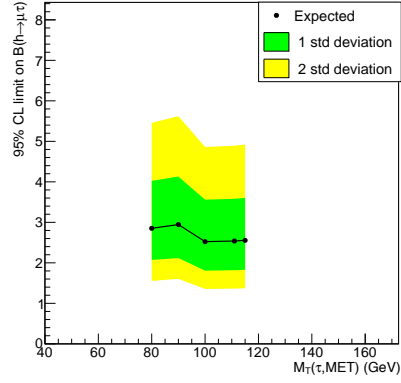
(c) 2 jets, gg-enriched
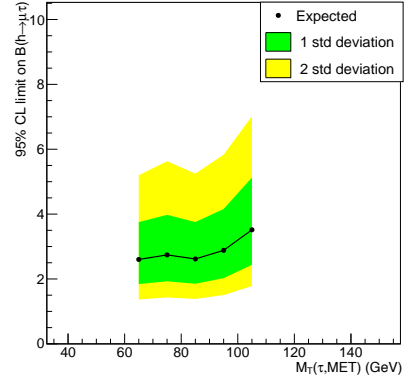
(d) 2 jets, VBF-enriched

Figure 5.1. Expected limits based on an Asimov dataset as a function of $M_T(\tau, MET)$ for the different categories.

TABLE 5.1

Selection criteria in each category with the optimization of the $H \to \mu\tau_h$

analysis

0-jet category

- $p_T^\mu > 26$ GeV, $p_T^\tau > 30$ GeV

- $M_T(\tau) < 105$ GeV

- No jets with $p_T^{jet} > 30$ GeV, $|\eta| < 4.7$, LooseID

1-jet category

- $p_T^\mu > 26$ GeV, $p_T^\tau > 30$ GeV

- $M_T(\tau) < 105$ GeV

- One jet with $p_T^{jet} > 30$ GeV, $|\eta| < 4.7$, LooseID

2-jet, gg-enriched category

- $p_T^\mu > 26$ GeV, $p_T^\tau > 30$ GeV

- $M_T(\tau) < 105$ GeV

- $p_T^{jet1} > 30$ GeV, $p_T^{jet2} > 30$ GeV $|\eta_{jet1}| < 4.7$, $|\eta_{jet2}| < 4.7$, LooseID

- $M_{jj} < 550$ GeV

- Two jets with $p_T^{jet} > 30$ GeV, $|\eta| < 4.7$, LooseID

2-jet, vbf-enriched category

- $p_T^\mu > 26$ GeV, $p_T^\tau > 30$ GeV

- $M_T(\tau) < 85$ GeV

- $p_T^{jet1} > 30$ GeV, $p_T^{jet2} > 30$ GeV $|\eta_{jet1}| < 4.7$, $|\eta_{jet2}| < 4.7$, LooseID

- $M_{jj} > 550$ GeV

- Two jets with $p_T^{jet} > 30$ GeV, $|\eta| < 4.7$, LooseID
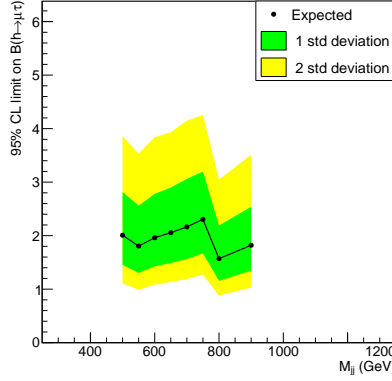
Figure 5.2. Expected limits based on an Asimov dataset as a function of M_{jj} for the 2 jet categories.

5.1.3 Multivariate analysis

A Boost decision trees(BDT) method is chosen as the multivariate analysis method used in $H \rightarrow \mu\tau_h$ search. It provides better sensitivity compared with the cut-based analysis. In this analysis, the BDT is provided by the TMVA package [10]. BDT method takes in signal and background datasets with a selected set of input variables. Input variables are the ones that show distinguishing power between signal and background. The training output is a weight file, which contains a list of weights to indicate in percentage how likely an event is signal like with a give set of input variable values from that event. A more detail description of the BDT method is available in section 7.1. In this analysis, signal and background events are required to pass the loose selection criteria. All of the categories are combined. The signal events from gluon gluon fusion and vector boson fusion higgs production mode are mix by the weight with respect to their production cross section. The background sample used in the training is misidentified lepton background from the like sign region(Region II as in table **). The list of

19

BDT input variables is following and the distribution of the variables are shown in Fig. 5.3.

- Transverse mass between the τ_h and E_T^{miss}, $M_T(\tau_h)$.

- Missing transverse energy, E_T^{miss}.

- Pseudorapidity difference between the μ and the τ_h candidate, $\Delta\eta(\mu, \tau_h)$.

- Azimutal angle between the μ and the τ_h, $\Delta\phi(\mu, \tau_h)$.

- Azimutal angle between the τ_h and the E_T^{miss}, $\Delta\phi(\tau_h, \mathrm{E}_T^{miss})$.

- Collinear mass, M_{col}.

- Muon p_T.

- τ_h p_T.

Fig. 5.4 shows BDT input variables in the signal and background training samples. The chosen input variables show low correlation in both samples. Fig. 5.5 shows the TMVA overtraining checks. Training samples are divided into two groups, one for training and one for testing. After the training, the algorithm applies the training output to the testing sample to check if they are in good matching.

5.2 $H \to e\tau_h$

5.2.1 Loose selection

In $H \to e\tau_h$ channel, the trigger used is HLT_Ele27_WP80, which applies an electron p_T cut at 27GeV at HLT level. A further cut on electron $P_T > 30GeV$
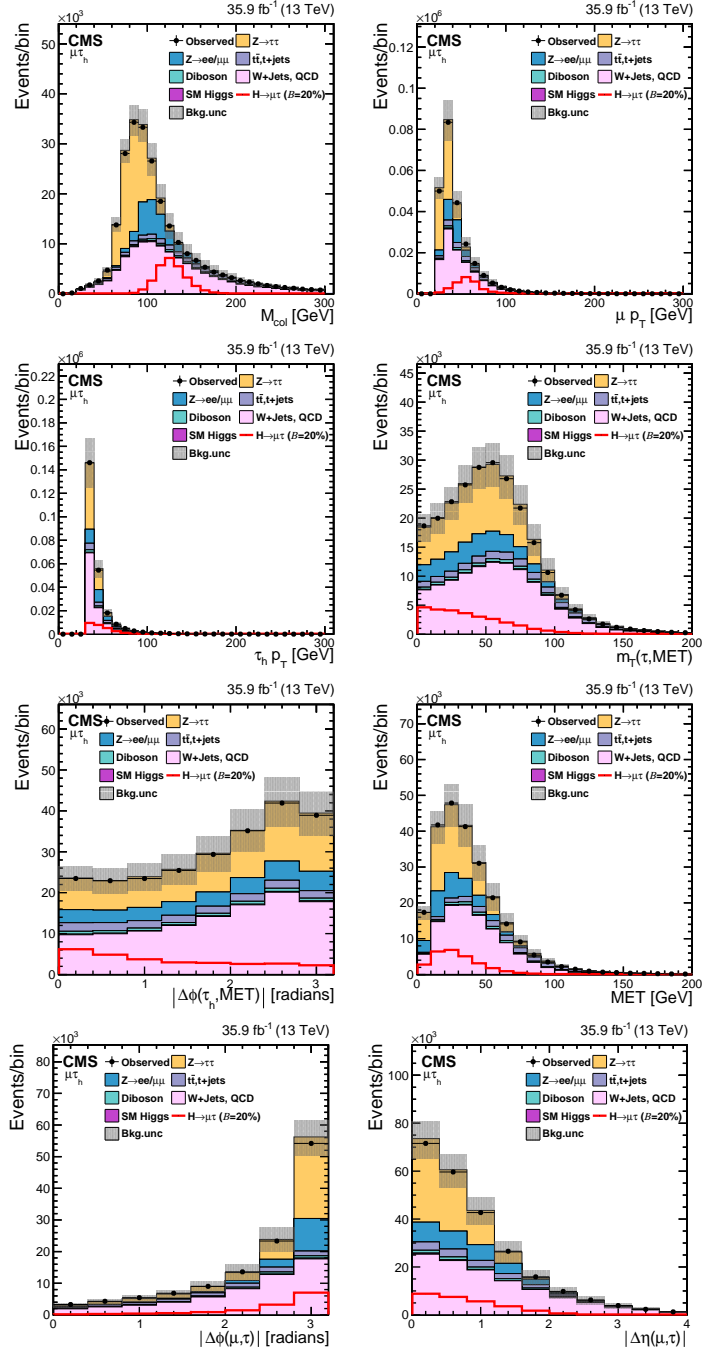
Figure 5.3. Distributions of the input variables to the BDT for the
$H \to \mu\tau_h$ channel.

(a) Signal sample variables correlation (b) Background sample variable correlation
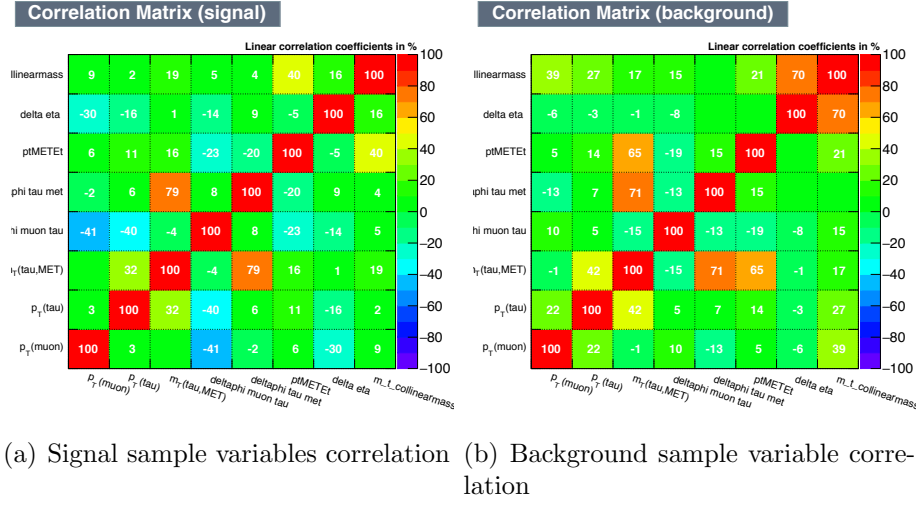
Figure 5.4. Expected limits based on an Asimov dataset as a function of $M_T(\tau, MET)$ for the different categories.
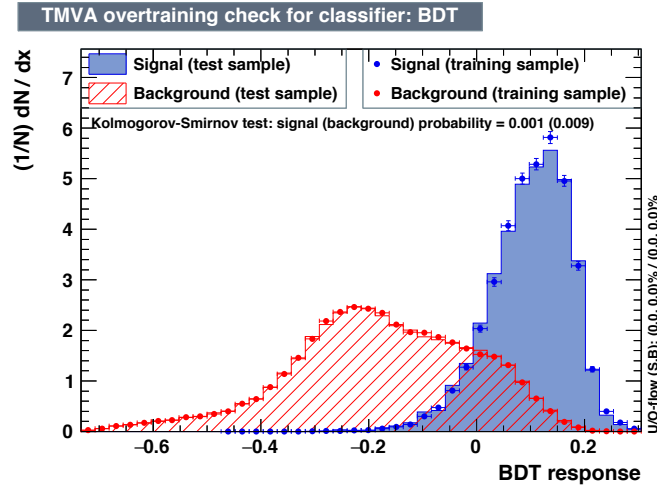
Figure 5.5. Overtraining checking for the BDT training in the TMVA package.

is applied. Electrons are required to have $|\eta_e < 2.3|$ and $D_z < 0.2cm$. D_z is the longitudinal impact parameter that shows the displacement between primary vertex and track path. Electrons are also requirement to pass the MVA based tight ID and cut based PF tight isolation $PFiso < 0.1$. Tau candidates are required to have $P_T > 30GeV$, pseudorapidity in the range $|\eta^\tau < 2.3|$ and the longitudinal impact parameter $D_z < 0.2cm$. Tau isolation used is the cut based tight tau isolation. In addition, tau candidates pass the tau Decay mode finding and tau discriminator against electrons and muons. The analysis also requires no extra isolated electrons with $p_T > 10GeV$ and extra taus with $p_T > 20GeV$. Tau and electron candidates are required to have opposite sign of charges and separate with an $\Delta R > 0.4$ from any jets in the events with $p_T > 30GeV$. All of the requirements contribute to the selections of good qualities candidates. The datasets are binned into three categories according to the number of jets in the events:

0-jet: No events have jets pass the loose PF ID and with jet $P_T > 30GeV$, $|\eta| < 4.7$. This category enhances the gluon-gluon fusion contribution.

1-jet: Events with one jet passes loose PF ID and jet $P_T > 30GeV$, $|\eta| < 4.7$. This category enhances the gluon-gluon fusion production with initial state radiation.

2 jets: Events with two jets pass loose PF ID and with jet $P_T > 30$ GeV and $|\eta| < 4.7$, This category contains both Higgs production mode and with an enhancement in VBF production mode.

With the preselection and binning of the jets numbers, the M_{col} distribution of H $\rightarrow e\tau_h$ in different categories are shown in Fig. 5.6
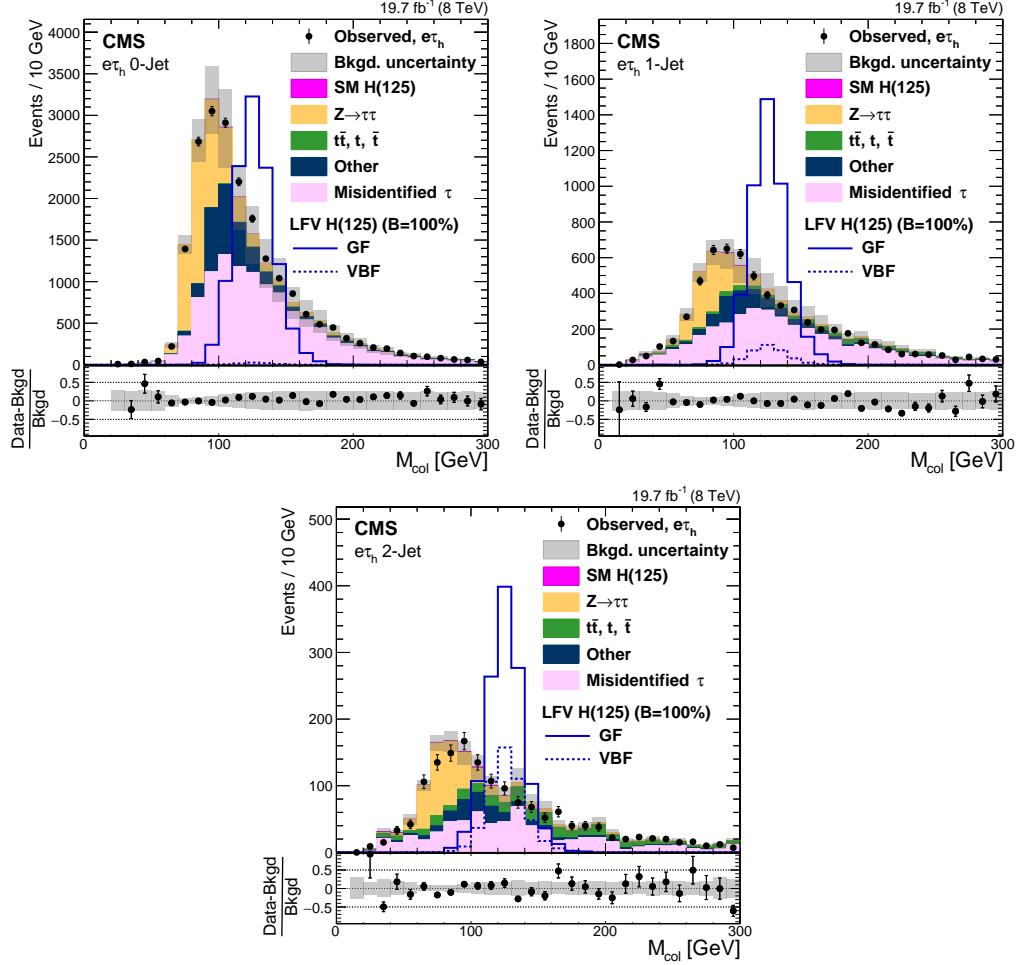
Figure 5.6. With loose selection conditions, the comparison of the observed collinear mass distributions with background from prediction. The shaded grey bands indicate the total background uncertainty. The open histograms correspond to the expected signal distributions for $\mathcal{B}(H \to e\tau_h) = 100\%$ in the 0-jet, 1-jet and 2-jet categories, respectively.

5.2.2 Cut-based analysis

A set of kinematics variables are defined and set to further select signal events. In H $\to e\tau_h$ channel, similar to the H $\to \mu\tau_h$, muon and tau leptons from the signal events are highly boosted, so as p_T variables plays an important role in distinguishing from background events and the separation in ϕ direction is bigger between μ and τ for signal events. Missing energy from the τ neutrino in hadronic decay is one of the characters in signal events, so $M_T(\tau_h)$ is an useful variable. In two jets category, M_{jj} also used as a cut variable. The cuts have been optimized to have the most stringent expected limits with the Asimov dataset. The detailed cuts used for H $\to e\tau_h$ is shown in table. 5.2.

TABLE 5.2

Selection criteria for each event category after cut optimization, for the

$H \to e\tau_h$ channel

0-jet category

- $p_T^e > 45 GeV$, $p_T^\tau > 30 GeV$, $\Delta\phi_{e\tau} > 2.3$

- $M_T(\tau) < 70 GeV$

- No jets with $p_T^{jet} > 30 GeV$, $|\eta| < 4.7$, LooseID

1-jet category

- $p_T^e > 35 GeV$, $p_T^\tau > 40 GeV$

- One jet with $p_T^{jet} > 30 GeV$, $|\eta| < 4.7$, LooseID

2-jet category

- $p_T^e > 35 GeV$, $p_T^\tau > 30 GeV$

- $M_T(\tau) < 50 GeV$

- $p_T^{jet1} > 30 GeV$, $p_T^{jet2} > 30 GeV$ $|\eta_{jet1}| < 4.7$, $|\eta_{jet2}| < 4.7$, LooseID

- $\Delta\eta(jet1, jet2) > 2.3$

- $M_{jj} > 400 GeV$

- Two jets with $p_T^{jet} > 30 GeV$, $|\eta| < 4.7$, LooseID

CHAPTER 6

LFV analysis background estimation

6.0.1 Background estimates for H $\to \mu\tau_h$

6.0.1.1 Misidentified Leptons

The misidentified lepton backgrounds are estimated with full data driven method from the collision data. In H $\to \mu\tau_h$ channel, the misidentification τ and μ lepton rate is obtained from independent Z + jets data sets and then apply the rate to an control region that orthogonal to the signal region to estimate the misidentified lepton backgrounds. The control regions for the estimation and validation of the misidentified lepton background are shown in Table. 6.1. Region I is the signal region with full selection applied. Region III is the background enriched region, in which one or two leptons have loose isolation but not tight isolation. Region II is the same to Region I besides the requirement that μ and τ lepton pairs have the same sign of charge, so as Region IV to Region III. Region III is used to estimate the misidentified background in Region I. Region II and IV are used in the validation of the method.

$Z \to \mu\mu$ is used in Z + jets data sets to estimate both misidentified τ and μ lepton rates. Z bosons are selected in the invariant mass window $70 < M_{ll} < 110$ GeV and trigger HLT_IsoMu24 or HLT_IsoTkMu24 is used in the selection. Two muons that are composed of Z bosons are required to have $p_T > 26$ GeV, $|\eta| < 2.4$,

TABLE 6.1

Definition of the samples used to estimate the misidentified lepton
background.

Region I	Region II
$\ell_1^{\pm}$(isolated)	$\ell_1^{\pm}$(isolated)
$\ell_2^{\mp}$(isolated)	$\ell_2^{\pm}$(isolated)

Region III	Region IV
$\ell_1^{\pm}$(isolated)	$\ell_1^{\pm}$(isolated)
$\ell_2^{\mp}$(not-isolated)	$\ell_2^{\pm}$(not-isolated)

cut based tight muon isolation($I_{rel}^{\mu} < 0.15$) and passing the muon medium ID. In the Z+jets samples, with the selected Z boson in an event, the remaining jets are checked if they pass τ or μ selection.

To estimate the misidentified τ lepton rate, The jets in the sample Z + jets are checked if they pass the τ isolation ID. Misidentified τ lepton ratio $\tau(f_\tau)$ is calculated as in Equation . 6.1, together with one related ratio f_τ. In Equation . 6.1, ℓ stands for leptons in general.

$$\ell(f_\ell) = \frac{f_\ell}{1 - f_\ell} \tag{6.1}$$

$$f_\ell = \frac{N_\ell(Z + jets\ \ell\ tight\ Iso)}{N_\tau(Z + jets\ \ell\ loose\ Iso)} \tag{6.2}$$

f_τ is the ratio between number of jets that pass tight τ MVA isolation ID and number of jets that pass very loose τ MVA isolation ID. The contribution from Diboson sample is subtracted with simulation, in which the third leptons are genuine leptons. With the ratio f_τ , misidentified rate $\tau(f_\tau)$(Equation. 6.2)

is applied as a weight to Region III. When estimating misidentified *tau* leptons, Region III is defined in the same selection condition as the signal region Region I , but τ leptons are required to pass the very loose MVA isolation and not pass the tight MVA isolation. The misidentified τ lepton rate $\tau(f_\tau)$ shows dependence on tau p_T, tau decay mode and ECAL geometry. Thus $\tau(f_\tau)$ is applied in term of tau p_T, tau decay modes and ECAL barrel and endcap seperately. The distribution of f_τ is shown in Fig. 6.1, which shows a comparison of the estimated f_τ between data and simulated Z+jets samples. The distribution of f_τ vs $|\eta|$ are also shown, which show less dependency.

The estimation of μ misidentified rate the same as the τ misidentified rate, but checks if the jets in Z + jets samples pass the muon isolation. As in Equation. 6.1, f_μ is defined as the ratio between tight cut based muon isolation($I_{rel}^\mu < 0.15$) and loose cut based muon isolation($I_{rel}^\mu < 0.25$). The misidentified μ lepton rate $\mu(f_\mu)$ is calculated by Equation. 6.2. $\mu(f_\mu)$ is a weight to apply to Region III. In the case of estimating the misidentified μ, Region III is defined as Regin I but with μ passing the loose cut based muon isolation and not passing tight cut based muon isolation. The ratio f_μ is shown in Figure. 6.2 in term of muon p_T and $|\eta|$ distribution.

There are overlap between the estimated misidentified muon and tau lepton when both of the leptons are fake ones. The overlap events are estimated as in Equation. 6.4. In the equation, the N(Region III) stands for the events in Region III. In this case, Region III is defined as muon leptons passing the lose isolation, but not passing the tight isolation, while tau leptons pass the very loose isolation but not pass the tight isolation. The final misidentified lepton background obtained by adding the misidentified muon and tau background and deducts the overlap

29

(a) EB tau decay mode 0   (b) EB tau decay mode 1   (c) EB tau decay mode 10

(d) EE tau decay mode 0   (e) EE tau decay mode 1   (f) EE tau decay mode 10

(g) tau decay mode 0 vs $|\eta|$ (h) tau decay mode 0 vs $|\eta|$ (i) tau decay mode 0 vs $|\eta|$
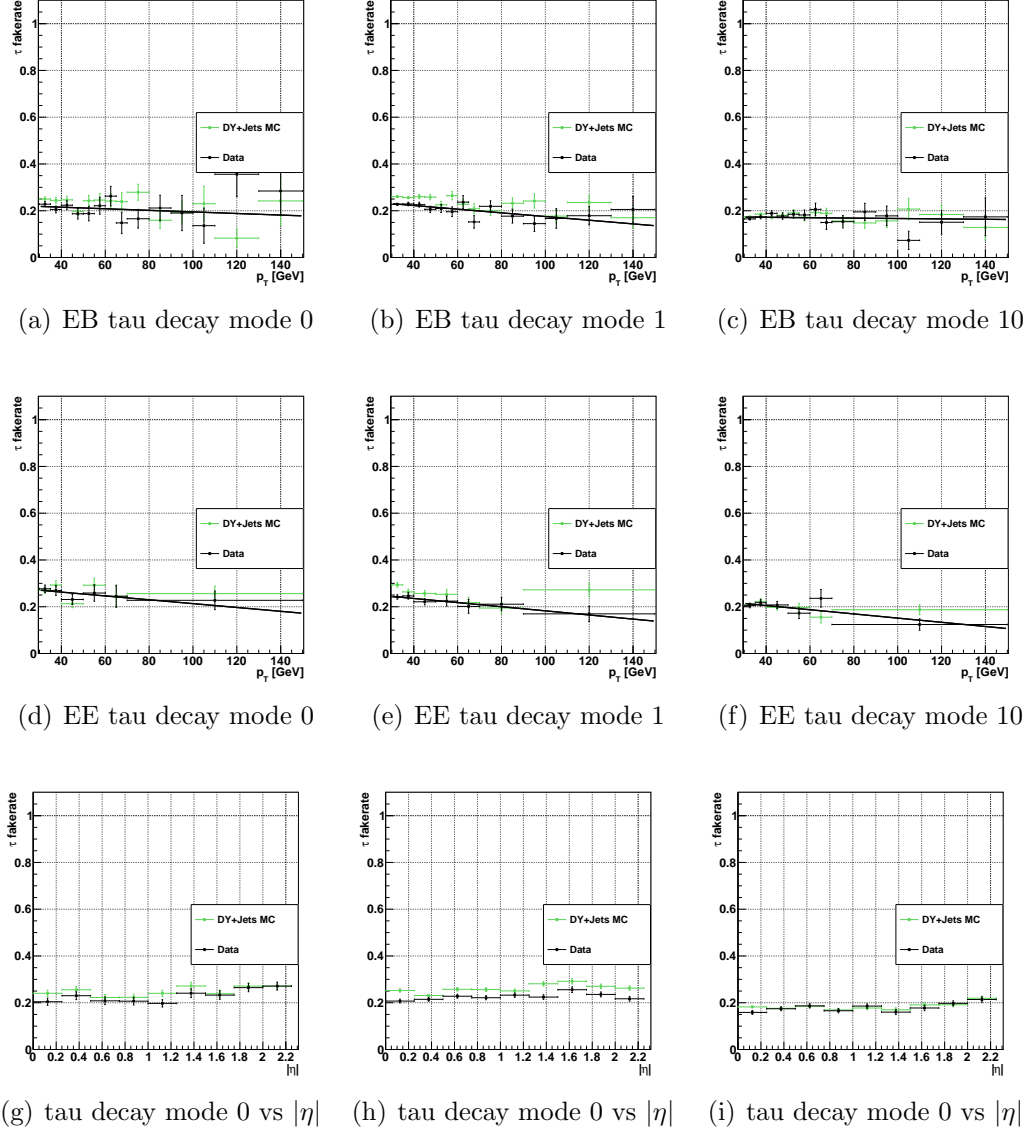
Figure 6.1. f_τ ratio for the τ fake rate calculation shows in term of tau decay modes, p_T, ECAL barrel, ECAL endcap, and $|\eta|$

(a) f_μ vs p_T (b) f_μ vs $|\eta|$
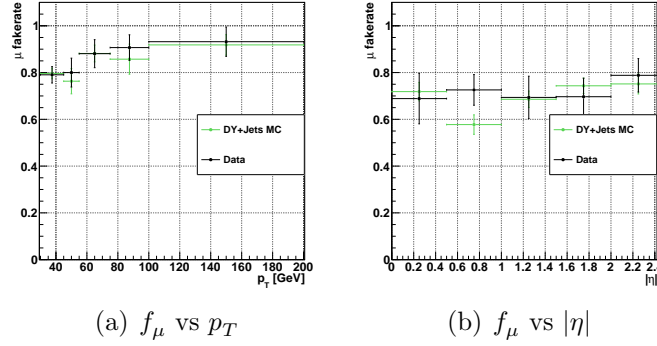
Figure 6.2. f_τ ratio for the μ fake rate calculation shows in term of muon p_T and $|\eta|$.

events between these two.

$$N(\text{overlap}) = \frac{f_\tau \cdot f_\tau}{(1 - f_\tau) \cdot (1 - f_\mu)} \cdot N(\text{Region III}) \qquad (6.3)$$

$$(6.4)$$

In H $\rightarrow \mu\tau_h$ channel, the validation of the misidentified lepton background with full data driven method is performed in a like-sign sample and W+jets enriched control region. The like-sign sample validation requires the events pass the loose selection, the same as the signal region, but with inverted charge. Misidentified background is enriched with the requirement of inverting charge. In the W+jets enriched control region, events that pass the loose selection are further selected with $M_T(\tau) > 60$ GeV and $M_T(\mu) > 80$ GeV. In both of the control regions, misidentified background is the dominant one, as shown in Fig. 6.3. With full data driven method, applied with the $\tau(f_\tau)$ as weight, the like-sign and W+jets

control regions show good data and MC agreement.
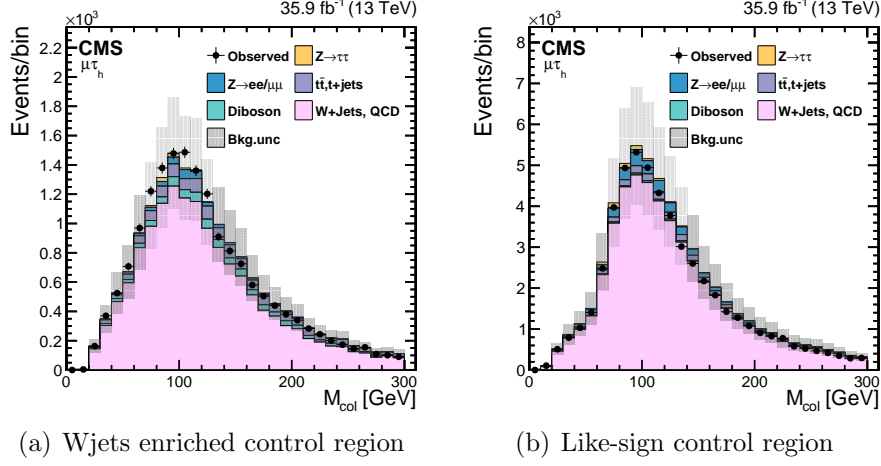
Figure 6.3. The validation of full data driven method for the misidentified lepton background. 30% of uncertainty is assumed in both cases.

6.0.1.2 Other backgrounds

Besides the misidentified background, which is composed by W+jets and QCD multi-jets background and estimated with data driven method, $Z \rightarrow \tau\tau$ and $Z \rightarrow \mu\mu$ are also major background in this analysis. In $Z \rightarrow \tau\tau$, the muons arise from a τ decay, while in $Z \rightarrow \mu\mu$, one of the μ can be misidentified as a τ. In other smaller backgrounds, the $\mu\tau$ pairs can be produced from the weak decays of quarks and vector bosons. These smaller backgrounds include standard model Higgs production($H \rightarrow \tau\tau, WW$),$t\bar{t}$,single top quark, di-boson(WW,WZ,ZZ). A full list

of background used is in the Table. 4.1. These backgrounds are all estimated with MC simulation.

In the $Z \to \mu\mu$ background, an additional scale factor is applied to the tau leptons that do not come from real taus. This is checked in the MC generator level. The scale factor is applied in term of $|\eta|$ as shown in Table. 6.2.

TABLE 6.2

Scale factor to the events in which a muon is misidentified as a tau

$|\eta|$ ranges	Correction values
0-0.4	1.263
0.4-0.8	1.364
0.8-1.2	0.854
1.2-1.7	1.712
1.7-2.3	2.324

6.0.2 Background estimates for H $\to e\tau_h$

6.0.2.1 Misidentified Leptons

Misidentified lepton background is a very important one in H $\to e\tau_h$ channel. This background mainly raises from W+jets and QCD multijets. A background

33

event passes the signal selection, may have one or both leptons are misidentified, which are faked by other objects. In H $\rightarrow e\tau_h$ 8 TeV analysis, both misidentified electron and tau leptons are considered.

Similar to H $\rightarrow \mu\tau_h$ channel, the misidentified background in H $\rightarrow e\tau_h$ is estimated with full data driven method, with independent data sets Z+jets, in which Z bosons decay into two electrons. Data sets are divided into four control regions for this estimation as in Table. 6.1. Region II is used to extra misidentified background in signal region, Region I, while Region II and Region IV are used for the validation of the fake rate. In the Z+jets data set selection, trigger HLT_Ele27_WP80 is used. The two electrons in Z+jets are required to have $p_T > 30$ GeV, $|\eta| < 2.3$, tight electron MVA ID and cut based tight PF isolation($I_{rel}^e < 0.1$). The invariant mass formed by the two electrons are required to be in the Z boson mass window. Then the jets in the events are checked if they pass τ or e selection. Tau leptons should have $p_T > 30$ GeV, $|\eta| < 2.3$, passing tau decay mode finder and discriminator against elections and muons. Electrons pass the same conditions as the elections used in forming the Z boson mass peak, but without the requirement of isolation. The isolation conditions are required in different conditions in the fake ratio estimation. Tau leptons use the cut based tight or loose isolation, while loose or tight MVA based electron isolation have the value $I_{rel}^e < 0.1$ and $I_{rel}^e < 0.2$ respectively. After the selection, the fake ratio f_τ is defined as the number of tau leptons passing the tight isolation divided the number of tau leptons passing the loose isolation. Similarly the fake ratio f_e is defined as the number of elections passing the loose isolation divided by the number of electron passing the tight isolation as shown in Equation. 6.2. With the fake ratio f_τ and f_e, the electron and tau fake rate $\tau(f_\tau)$ and $e(f_e)$ are estimated according to Equation. 6.1. The

fake ratio for tau leptons and electrons are shown in Figure. 6.4.

The tau and electron fake ratio is applied in term of $|\eta|$. In the estimation, the background source of three prompt leptons from Diboson samples are deducted. The tau and electron misidentified backgrounds are estimated separately, the overlap events in which both leptons are misidentified objects are estimated with Equation. 6.4. In the equation, N(Region III) in H $\rightarrow e\tau_h$ channel satisfies the condition both leptons are loose isolated but not tight isolated. The validation of full data driven method is done with Region II and Region IV. The distribution of the M_{col} is checked in the same sign region of the two leptons, shown in Figure. 6.5, in which observed data agrees well with the events from MC and misidentified background from data driven method.

6.0.2.2 $Z \rightarrow \tau\tau$

$Z \rightarrow \tau\tau$ in another significant background in H $\rightarrow e\tau_h$ channel. In the case one of the tau leptons decay to a electron and the other hadronically decay. The background is estimated with the embedding technique[1], which starts by selecting the $Z \rightarrow \mu\mu$ events from collision data. The muons pass loose muon selection and then are replaced by simulated τ leptons and reconstructed with PF algorithm. The key advantage of this technique is that the major topologies like missing energy, jet multiplicity, underlying events are directly from data. Besides tau, the other event records and detector responds are from real environment. When added to the background, the embedded data is normalized to MC prediction. Compared with MC prediction, the visible mass peak from embedding sample shifts 2%. This is mainly because of the difference in final-state radiation of photons between muon and tau and is corrected for. Identification and isolation

(a) f_τ vs p_T
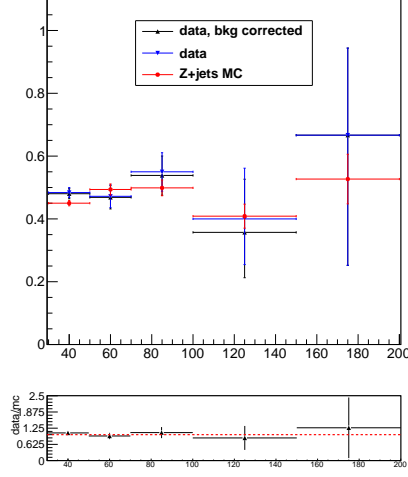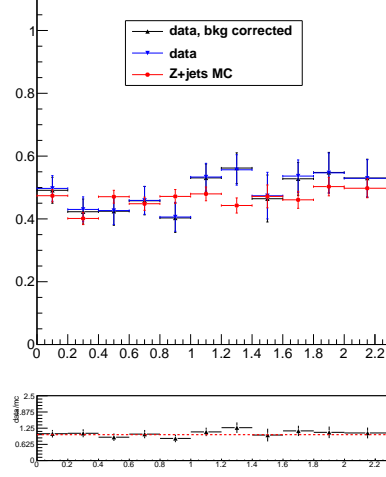
(b) f_τ vs $|\eta|$
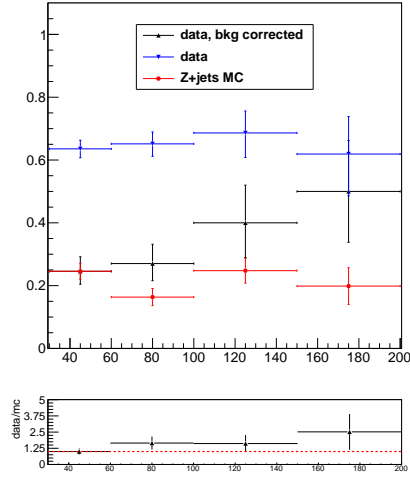
(c) f_e vs p_T
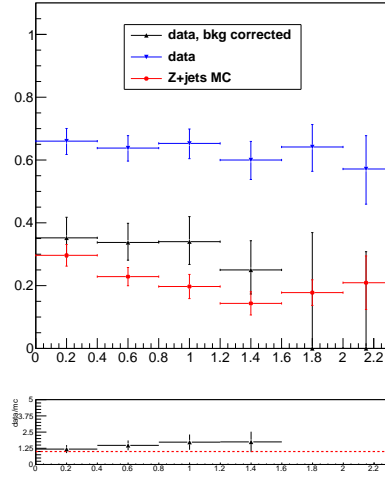
(d) f_e vs $|\eta|$

Figure 6.4. f_τ and f_τ ratio for the fake rate calculation shows in term of p_T and $|\eta|$
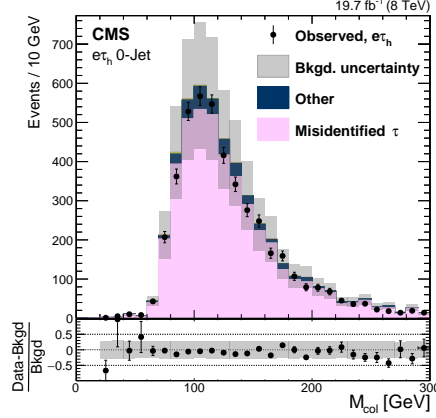
Figure 6.5. Data driven method validation in Region II with e and τ leptons in same sign of charge

correction for this shift to embedded sample are obtained by comparing with MC simulations.

6.0.2.3 Other backgrounds

$t\bar{t}$ is one of the main background. Leptonic decay W bosons can produce opposite sign lepton pair and missing energy. The shape of this background is obtained from MC samples and normalized to $t\bar{t}$ control region sample. The evens in $t\bar{t}$ control samples are required to pass the 2-jet base selection in H $\rightarrow e\tau_h$ channel and at lease one the jets is identified as a b jet.

The other backgrounds presented in the samples are Z+jets in which one of the electrons is misidentified as a τ lepton. Di-boson(WW,WZ,ZZ) events, single Top, SM Higgs boson production($H \rightarrow \tau\tau$) and W$\gamma^{(\star)}$. These backgrounds are all estimated from MC simulation.

CHAPTER 7

Signal extraction and systematics

7.1 Boost decision trees method

7.2 statistical methods

7.3 Systematics

7.3.1 Systematics used in $H \to \mu\tau_h$

TABLE 7.1

1111

Systematic uncertainty	$H \to \mu\tau_h$	$H \to \mu\tau_e$	$H \to e\tau_h$	$H \to e\tau_\mu$
Muon trigger/identification/isolation	2%	2%	–	2%
Electron trigger/identification/isolation	–	2%	2%	2%
Hadronic tau lepton efficiency	5%	–	5%	–
b tagging veto	2.0–4.5%	2.0–4.5%	–	2.0–4.5%
$Z \to \mu\mu, ee$ + jets background	–	10%⊕5%	–	10%⊕5%
$Z \to \tau\tau$ + jets background	10%⊕5%	10%⊕5%	10%⊕5%	10%⊕5%
W + jets background	–	10%	–	10%
QCD multijet background	–	30%	–	30%
WW, ZZ background	5%⊕5%	5%⊕5%	5%⊕5%	5%⊕5%
$t\bar{t}$ background	10%⊕5%	10%⊕5%	10%⊕5%	10%⊕5%
$W\gamma$ background	–	10%⊕5%	–	10%⊕5%
Single top quark background	5%⊕5%	5%⊕5%	5%⊕5%	5%⊕5%
$\mu \to \tau_h$ background	25%	–	–	–
$e \to \tau_h$ background	–	–	12%	–
Jet $\to \tau_h, \mu, e$ background	30%⊕10%	–	30%⊕10%	–
Jet energy scale	3–20%	3–20%	3-20%	3–20%
τ_h energy scale	1.2%	–	1.2%	–
$\mu, e \to \tau_h$ energy scale	1.5%	–	3%	–
e energy scale	–	$0.1 - 0.5\%$	$0.1 - 0.5\%$	$0.1 - 0.5\%$
μ energy scale	0.2%	0.2%	–	0.2%
Unclustered energy scale	$\pm 1\sigma$	$\pm 1\sigma$	$\pm 1\sigma$	$\pm 1\sigma$
Renorm./fact. scales (g gH) [?]		3.9%		
Renorm./fact. scales (VBF and VH) [39]		0.4%		
PDF + α_s (g gH) [?]		3.2%		
PDF + α_s (VBF and VH) [?]		2.1%		

CHAPTER 8

LFV Higgs decay searching results

Results can be extracted after applying the selection criteria and with the full consideration of the relevant systematics. In both $H \to \mu\tau_h$ and $H \to e\tau_h$ analysis, a maximum likelihood fit is performed to derive the expected and observed limits. Each category in the analysis is fitted separately and then combined. Systematic uncertainties are used as nuisance parameter in the fittings. The upper limits on signal branching faction with CL_s criterion are set with asymptotic formula.

8.0.1 Results in $H \to \mu\tau_h$ search

In Lepton flavour violation $H \to \mu\tau_h$ search at central of mass energy of 13 TeV, the analysis is performed in two parallel searching method, the M_{col} fit analysis and BDT fit analysis. After the selection in M_{col} fit analysis and adjusting all of distribution by the fit, the distribution of the signal and background discriminant variable M_{col} is shown in Figure. 8.3. In the BDT fit analysis, in which the BDT discriminator is used as the variable to distinguish signal from background. After fitting to all of the distributions, the observed data versers the expected backgrounds is shown in Figure. 8.2. No excess over the backgrounds is observed in both of the methods. The expected and observed median upper limits at 95% CL and the best fit branching fraction in the $H \to \mu\tau_h$ search with the M_{col} fit

analysis is shown in Table. 8.1 and the result of the BDT fit analysis is shown in Table. 8.2.

The search in $H \to \mu\tau$ with CMS 13 TeV, 36 fb^{-1} data is motivated by the search CMS 8 TeV 19.7 fb^{-1} $H \to \mu\tau$, in which a significance of 2.4 standard deviation excess of data with respect to the SM background-only hypothesis was observed. This thesis is written in H $\to \mu\tau_h$ search, together with another search $H \to \mu\tau_e$ [6], where the tau lepton decays electronically to an electron, instead of tau hadronically decay, gives the most recent and update check for the best fit branching ratio $H \to \mu\tau$. This combined result is shown with BDT fit analysis and M_{col} fit analysis separately in Figure. 8.3. No evidence of excess is found in this combined result. Compared with the 8 TeV result, the new upper limits on the signal branching fraction improved a factor of 5. The BDT fit analysis is more sensitive to the M_{col} fit analysis. The new and tighter limit result comes from the BDT fit analysis.

The constrains on Yukawa couplings, $|Y_{\mu\tau}|, |Y_{\tau\mu}|$, can be derived from the limits on branching ratio $\mathcal{B}(H \to \mu\tau)$, which is the combined result between H $\to \mu\tau_h$ search and $H \to \mu\tau_e$ search. Lepton flavour violating decay can rise from the tree level by assuming the Yukawa interactions. In $H \to \mu\tau$ decays, the Yukawa interactions is related to the decay width $\Gamma(H \to \ell^\alpha \ell^\beta)$, ℓ^α, ℓ^β can be μ or τ and $\ell^\alpha \neq \ell^\beta$. Γ in terms of the Yukawa couplings is given by:

$$\Gamma(H \to \ell^\alpha \ell^\beta) = \frac{m_H}{8\pi}\left(|Y_{\ell^\beta \ell^\alpha}|^2 + |Y_{\ell^\alpha \ell^\beta}|^2\right), \tag{8.1}$$

41

Figure 8.1. M_{col} distribution after the selection. The signal and backgrounds in the plots have been normalized to the best fit values. The gray bands shows the total uncertainties in each bin and the signal is plotted with the branching ratio of 5% of the higgs decay branching ratio for visualization purpose.

(a) 0 jet

(b) 1 jet

(c) 2 jets, gg-enriched

(d) 2 jets, VBF-enriched

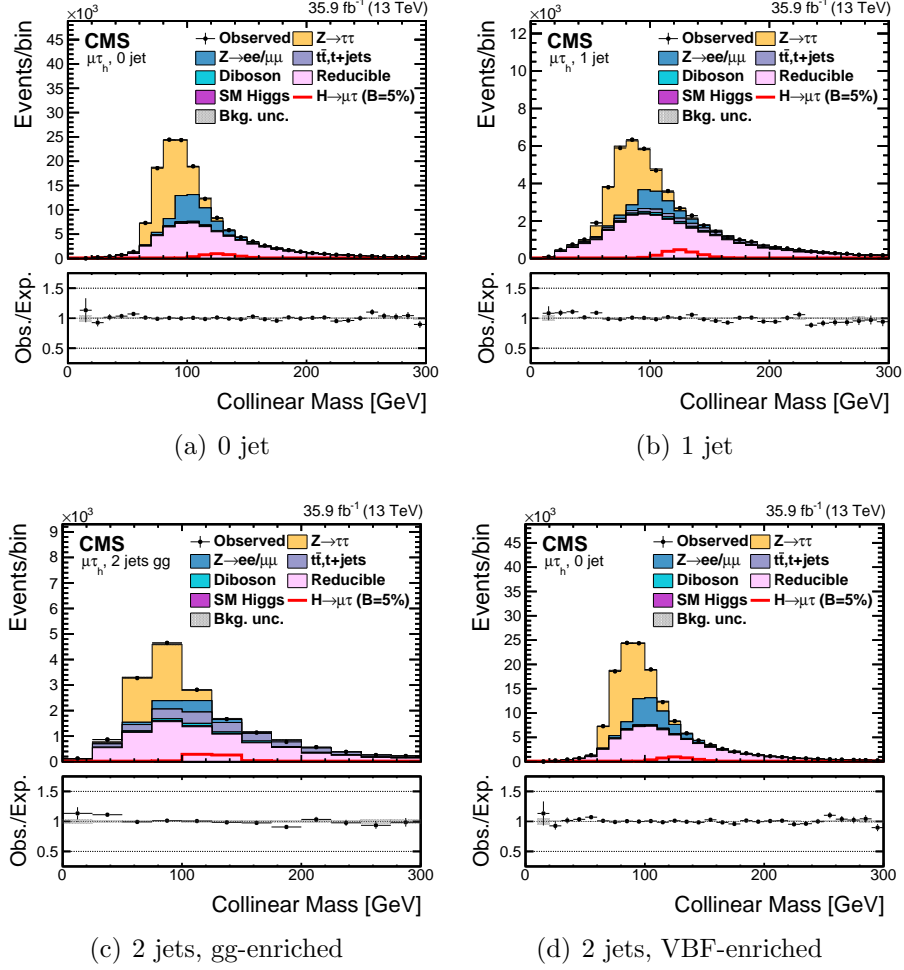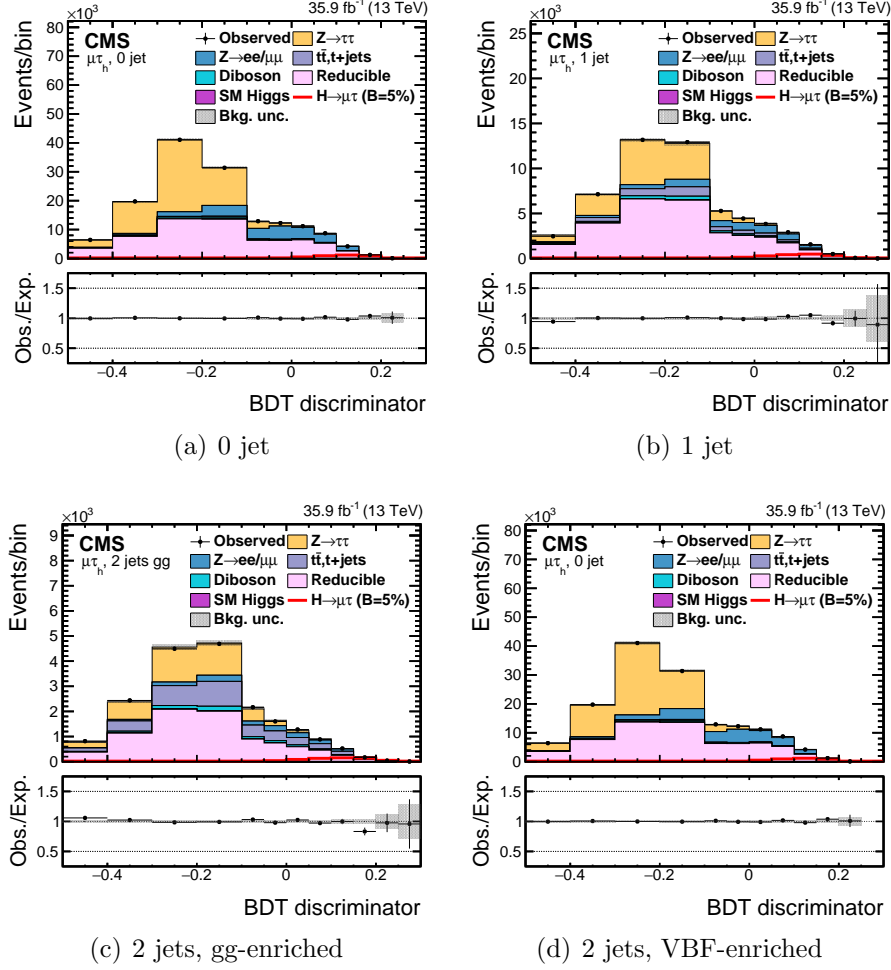Figure 8.2. BDT discriminator distribution after the selection. The signal and backgrounds in the plots has been normalized to the best fit values. The gray bands shows the total uncertainties in each bin and the signal is plotted with the branching ratio of 5% of the higgs decay branching ratio for visualization purpose.

43

TABLE 8.1

M_{col} fit analysis expected and observed upper limits at 95% CL and the best fit branching fractions in each of the categories.

Expected limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_h$	< 1.14	< 1.26	< 2.12	< 1.41	< 0.71
Observed limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_h$	< 1.04	< 1.74	< 1.65	< 1.30	< 0.66
Best fit branching fractions (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_h$	-0.30 ± 0.45	0.68 ± 0.56	-1.23 ± 1.04	-0.23 ± 0.66	-0.08 ± 0.34

TABLE 8.2

BDT fit analysis expected and observed upper limits at 95% CL and the
best fit branching fractions in each of the categories.

	Expected limits (%)				
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_h$	<0.43	<0.56	<0.94	<0.58	<0.29
	Observed limits (%)				
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_h$	<0.51	<0.53	<0.56	<0.51	<0.27
	Best fit branching fractions (%)				
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_h$	0.12 ± 0.20	-0.05 ± 0.25	-0.72 ± 0.43	-0.22 ± 0.31	-0.04 ± 0.14

(a) M_{col} fit analysis   (b) BDT fit analysis
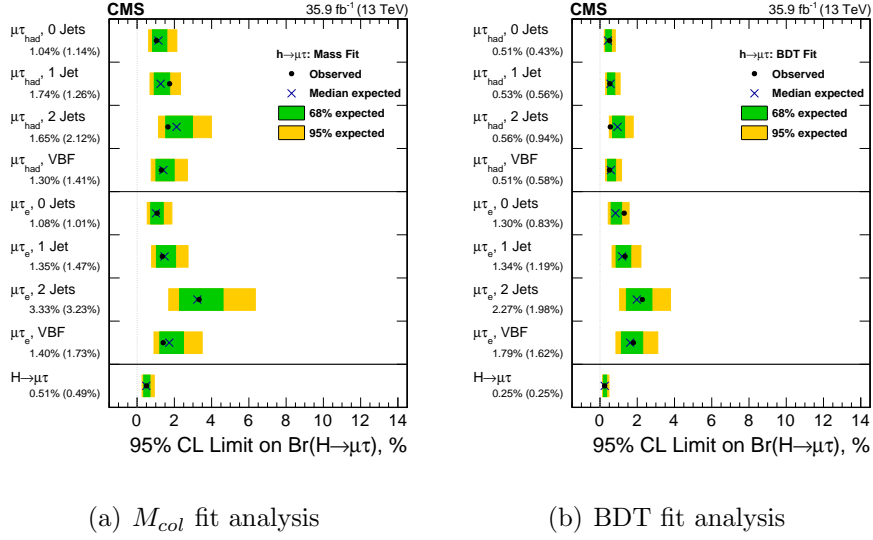
Figure 8.3. The expected and observed 95% CL limits on branching ratio of total $H \to \mu\tau$ which shows the results from both $H \to \mu\tau_h$ and $H \to \mu\tau_e$ in the M_{col} fit analysis and BDT fit analysis.

and the branching fraction by:

$$\mathcal{B}(H \to \ell^\alpha \ell^\beta) = \frac{\Gamma(H \to \ell^\alpha \ell^\beta)}{\Gamma(H \to \ell^\alpha \ell^\beta) + \Gamma_{\mathrm{SM}}}. \qquad (8.2)$$

The 125 GeV Standard Model Higgs decay width is taken as $\Gamma_{\mathrm{SM}} = 4.1\mathrm{MeV}$ [7].In BDT fit analysis, the 95% CL upper limit on the Yukawa coupling is $\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}^2|} < 1.43 \times 10^{-3}$ and in M_{col} fit analysis, $\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 2.05 \times 10^{-3}$. The upper limit on Yukawa couplings of BDT fit analysis is shown in Figure. 8.4. The comparison with the previous experimental search and theoretical naturalness results is also shown in the plot.
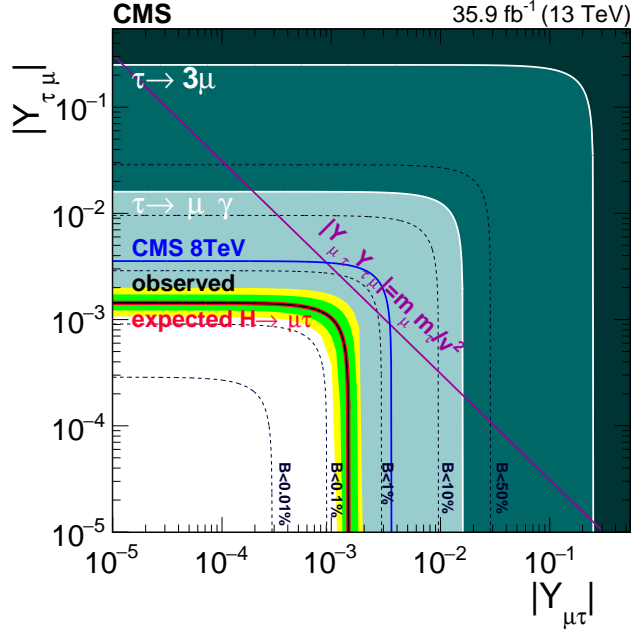
Figure 8.4. Upper limit of flavour violating Yukawa coupling $|Y_{\mu\tau}|, |Y_{\tau\mu}|$ from the BDT fit analysis. The red solid line represents the expected limit and the black solid line is the observed limit. 68% and 95% range of containing the observed limit is shown in the green and yellow band. The result from $\tau \to 3\mu$ is shown in dark green [8, 9, 12] and the one from $\tau \to \mu\gamma$ is shown in lighter green [8, 12]. The theoretical naturalness limit $Y_{ij}Y_{ji} \le m_i m_j/v^2$ is shown by purple diagonal line [8].

8.0.2 Results in H → $e\tau_h$ search

The search of H → $e\tau_h$ is performed in M_{col} fit analysis with CMS 8 TeV 19.7 fb^{-1}. The analysis is binned with 3 categories. After applying the selection cuts and adjusting the signal and background processes by the fit, the fitting variable M_{col} distribution is shown in Figure. 8.5. The number of event yields of both signal and background processes are shown in Table. 8.3. The yields are normalized to the integrated luminosity of 19.7 fb^{-1} and the signal branching ratio is assumed to be 0.69% of the SM higgs production cross section. This thesis is focused and written on the analysis H → $e\tau_h$ which is one part of the whole search on lepton flavour violating higgs decay to $e\tau$. Another closely related search is H → $e\tau_\mu$ [11]. The results of each categories and combined ones of the two analysis H → $e\tau_h$ and H → $e\tau_\mu$ on the expected and observed upper signal branching fraction $\mathcal{B}(H → e\tau)$ limits is shown in Table. 8.4 and summarized in Figure. 8.6. The combined upper limit on $\mathcal{B}(H → e\tau) < 0.69$ is the first direct search on this channel and the tightest limit by the time the result was published. Yukawa coupling between electron and tau leptons is related to decay width Γ and branching fraction $\mathcal{B}$ as shown in Equation. 8.1 and 8.2. The 95% CL upper limit on Yukawa coupling from the combined result in H → $e\tau$ is $\sqrt{|Y_{e\tau}|^2 + |Y_{\tau e}|^2} < 2.4 \times 10^{-3}$. The limit on $e\tau$ Yukawa coupling from this search and other previous search results are shown in Figure. 8.0.2.
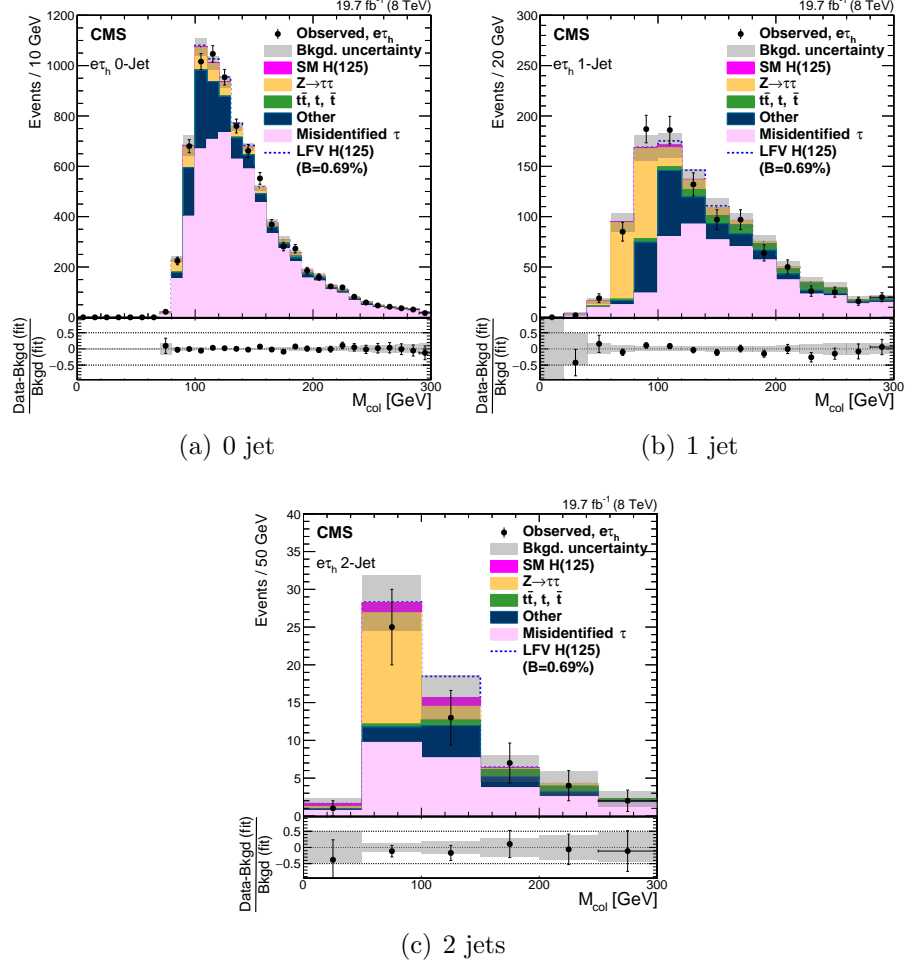
(a) 0 jet

(b) 1 jet

(c) 2 jets

Figure 8.5. M_{col} distribution after the selection. The signal and backgrounds in the plots have been normalized to the best fit values. The gray bands shows the total uncertainties in each bin and the signal is plotted with the branching ratio of 0.69% of the higgs decay branching ratio for visualization purpose.

TABLE 8.3: Events yields of the signal and backgrounds are shown in the mass range 100 GeV $< M_{col} <$ 150 GeV. The branching ratio of H $\rightarrow e\tau_h$ is assumed to be 0.69% of the SM higgs production cross section. The numbers in the table are normalized to the integrated luminosity of 19.7 fb^{-1}.

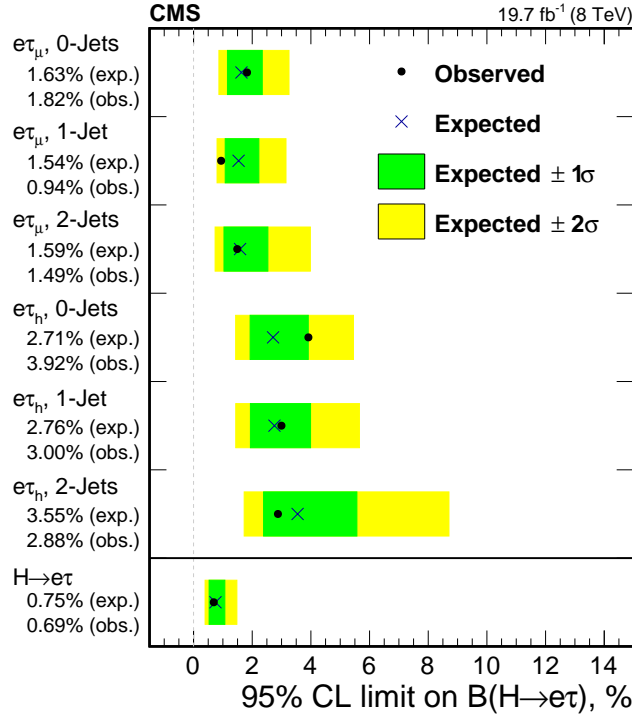| Jet category:	0-jet	1-jet	2-jet
Misidentified leptons	3366±25	223±11	8.7± 2.2
$Z \rightarrow ee, \mu\mu$	714±30	85±4	3.2± 0.2
$Z \rightarrow \tau\tau$	270±10	32±3	1.6± 0.3
$t\bar{t}, t, \bar{t}$	10±2	13±2	0.5± 0.2
ZZ, WZ, WW	53±2	6±1	0.3± 0.1
SM H background	12±1	3±1	1.0± 0.1
Sum of background	4425±28	363±11	15.3±2.3
Observed	4438	375	13
LFV H signal	61±4	15±1	2.8±0.5

Figure 8.6. 95% CL upper limit on branching ratio on $H \rightarrow e\tau$ with both the results from H $\rightarrow e\tau_h$ and H $\rightarrow e\tau_\mu$ and the combined.

50

TABLE 8.4

M_{col} fit analysis expected and observed upper limits at 95% CL and the best fit branching fractions in each of the categories of both H $\to e\tau_h$ and $H \to e\tau_\mu$ analysises. The asymmetric one standard-deviation uncertainties around the expected limits are shown in parentheses.

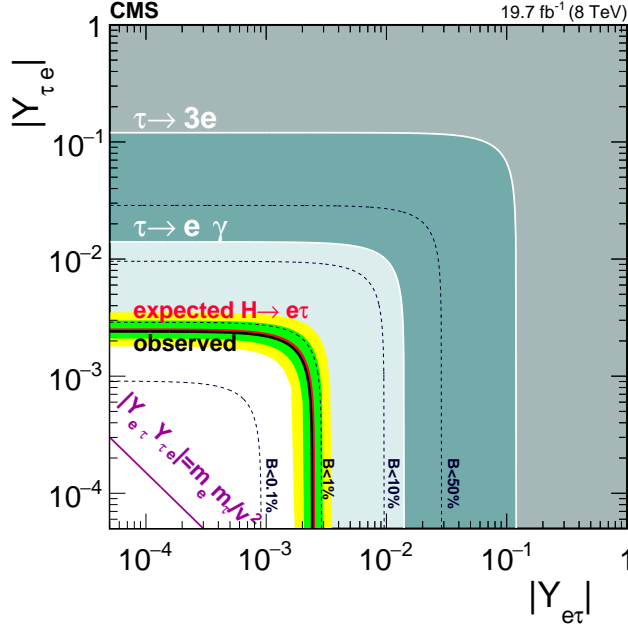|	0-jet	1-jet	2-jet
Expected limits at 95% CL (%)			
$e\tau_\mu$	$<1.63 \left(^{+0.66}_{-0.44}\right)$	$<1.54 \left(^{+0.71}_{-0.47}\right)$	$<1.59 \left(^{+0.93}_{-0.55}\right)$
$e\tau_h$	$<2.71 \left(^{+1.05}_{-0.75}\right)$	$<2.76 \left(^{+1.07}_{-0.77}\right)$	$<3.55 \left(^{+1.38}_{-0.99}\right)$
$e\tau$	$<0.75 \left(^{+0.32}_{-0.22}\right)$		
Observed limits at 95% CL (%)			
$e\tau_\mu$	<1.83	<0.94	<1.49
$e\tau_h$	<3.92	<3.00	<2.88
$e\tau$	<0.69		

Figure 8.7. Upper limit of flavour violating Yukawa coupling $|Y_{e\tau}|, |Y_{\tau e}|$ from the combined $H \to e\tau$ result. The red solid line represents the expected limit and the black solid line is the observed limit. 68% and 95% range of containing the observed limit is shown in the green and yellow band. The result from $\tau \to 3e$ is shown in gray, the one from $\tau \to e\gamma$ is shown in dark green and presented analysis is in light blue. The theoretical naturalness limit $Y_{ij}Y_{ji} \leq m_i m_j/v^2$ is shown by purple diagonal line [8]

.

CHAPTER 9

Conclusion

BIBLIOGRAPHY

1. S. Chatrchyan et al. Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons. 2014.

2. C. Collaboration". Performance of τ-lepton reconstruction and identification in CMS. *JINST*, 7(arXiv:1109.6034. CMS-TAU-11-001. CERN-PH-EP-2011-137):P01001. 33 p, Sep 2011. URL `http://cds.cern.ch/record/1385560`.

3. C. Collaboration. Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2. 2016.

4. C. Collaboration. Reconstruction and identification of tau lepton decays to hadrons and tau neutrino at CMS. *JINST 11 (2016) P01019*, 11, 2016. doi: 10.1088/1748-0221/11/01/P01019.

5. C. Collaboration. Identification of b quark jets at the CMS Experiment in the LHC Run 2. Technical Report CMS-PAS-BTV-15-001, CERN, Geneva, 2016. URL `https://cds.cern.ch/record/2138504`.

6. C. Collaboration. Search for lepton flavour violating decays of the Higgs boson to and e in proton-proton collisions at $\sqrt{s}$= 13 TeV. *JHEP*, 2017.

7. A. Denner, S. Heinemeyer, I. Puljak, D. Rebuzzi, and M. Spira. Standard model Higgs-boson branching ratios with uncertainties. *Eur. Phys. J. C*, 71: 1753, 2011. doi: 10.1140/epjc/s10052-011-1753-8.

8. R. Harnik, J. Kopp, and J. Zupan. Flavor violating higgs decays. *JHEP*, 03: 26, 2013. doi: 10.1007/JHEP03(2013)026.

9. K. Hayasaka et al. Search for lepton flavor violating τ decays into three leptons with 719 million produced $\tau^+\tau^-$ pairs. *Phys. Lett. B*, 687:139, 2010. doi: 10.1016/j.physletb.2010.03.037.

10. A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, M. Backes, T. Carli, O. Cohen, A. Christov, D. Dannheim, K. Danielowski, S. Henrot-Versille, M. Jachowski, K. Kraszewski, A. Krasznahorkay, Jr., M. Kruk, Y. Mahalalel, R. Ospanov, X. Prudent, A. Robert, D. Schouten, F. Tegenfeldt, A. Voigt, K. Voss, M. Wolter, and A. Zemla. TMVA - Toolkit for Multivariate Data Analysis. *ArXiv Physics e-prints*, Mar. 2007.

11. V. Khachatryan et al. Search for lepton flavour violating decays of the higgs boson to eτ and eμ in proton-proton collisions at $\sqrt{s} = 8$ TeV. *Phys. Lett. B*, 763:472, 2016. doi: 10.1016/j.physletb.2016.09.062.

12. Particle Data Group, C. Patrignani, et al. Review of Particle Physics. *Chin. Phys. C*, 40:100001, 2016. doi: 10.1088/1674-1137/40/10/100001.