SEARCH FOR LEPTON FLAVOR VIOLATING DECAYS OF THE HIGGS 
BOSON IN THE $\mu\tau_{\text{hadronic}}$ CHANNEL

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

Nathan M. Kellams

Colin Jessop, Director

Graduate Program in Physics
Notre Dame, Indiana
September 2016
SEARCH FOR LEPTON FLAVOR VIOLATING DECAYS OF THE HIGGS
BOSON IN THE $\mu\tau_{hadronic}$ CHANNEL

Abstract

by

Nathan M. Kellams

A search for the lepton flavor violating decay of the Higgs boson to a muon and a hadronically decaying tau lepton is presented. The search uses 19.7 fb$^{-1}$ of $\sqrt{s} = 8$ TeV proton-proton collision data collected by the Compact Muon Solenoid. Limits on the branching fraction are combined with those from a similar search with an electronically decaying tau to set a 95% CL limit on $B(H \rightarrow \mu\tau) < 1.51\%$. The branching fraction limit is transformed to a limit on Yukawa couplings $\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 3.6 \times 10^{-3}$. These limits are an order of magnitude stronger than existing constraints set by indirect methods. A slight excess of signal events is observed in multiple categories with a combined significance of 2.4 $\sigma$ corresponding to a best fit branching fraction of $B(H \rightarrow \mu\tau) = 0.84^{+0.39}_{-0.37}\%$. 
## CONTENTS

**FIGURES** ................................................................. iv

**TABLES** ................................................................. vi

**CHAPTER 1: INTRODUCTION** .............................................. 1

**CHAPTER 2: THEORY** ....................................................... 3
  2.1 Standard Model ......................................................... 3
    2.1.1 Higgs Production at the LHC ...................................... 8
    2.1.2 Tau Lepton Decays .................................................. 8
  2.2 Lepton Flavor Violation .............................................. 10
    2.2.1 Examples of Models with Lepton Flavor Violation ................. 12
    2.2.2 Previous Constraints .............................................. 13

**CHAPTER 3: COLLIDER AND DETECTOR** ................................. 17
  3.1 Large Hadron Collider ............................................... 17
  3.2 Compact Muon Solenoid ................................................ 20
    3.2.1 Superconducting Solenoid ......................................... 23
    3.2.2 Tracker ............................................................. 23
    3.2.3 Electromagnetic Calorimeter ..................................... 25
    3.2.4 Hadronic Calorimeter ............................................. 28
    3.2.5 Muon System ....................................................... 31

**CHAPTER 4: RECONSTRUCTION AND SELECTION** ....................... 35
  4.1 Trigger Selection ..................................................... 35
  4.2 Pileup ................................................................. 37
  4.3 Particle Flow ......................................................... 38
  4.4 Objects ............................................................. 39
    4.4.1 Jets ............................................................. 39
    4.4.2 Missing Transverse Energy ..................................... 41
    4.4.3 Electrons ......................................................... 42
    4.4.4 Muons ............................................................ 44
    4.4.5 Hadronically Decaying Taus ..................................... 46
  4.5 Event Selection ...................................................... 50
    4.5.1 Preselection ..................................................... 51
2.1 Predicted SM Higgs production cross sections for various processes and decay mode branching fractions. ................................................. 9
2.2 Dominant Higgs production processes at the LHC, gluon fusion (left) and vector boson fusion (right) .................................................. 9
2.3 Tau decay to a neutrino and two fermions ..................................... 11
2.4 Diagram showing the LFV process $H \rightarrow \mu\tau_{had}$ .......................... 14
2.5 Diagrams for $\tau \rightarrow \mu\gamma$ with one LFV vertex ........................ 14
2.6 Diagram for $\tau \rightarrow 3\mu$ ............................................................. 15
2.7 Diagram for muonium-antimuonium oscillation via LFV ................. 15
2.8 Diagram showing a contribution to muon dipole moments .................. 15
2.9 Diagram showing $\mu \rightarrow e$ transitions in nuclei .............................. 16

3.1 CERN accelerator complex .......................................................... 18
3.2 Integrated luminosity of CMS pp collision data ................................. 19
3.3 Cut away view of the CMS detector and its subsystems ..................... 21
3.4 Wedge slice in the transverse plane showing the interactions of different particles with the subsystems ................................. 22
3.5 Pixel tracker layout ................................................................. 24
3.6 Tracker layout .......................................................................... 25
3.7 Tracker resolution estimates for transverse momentum (top), transverse impact parameter (left) and longitudinal impact parameter (right) using muons of 1, 10, and 100 GeV ........................................ 26
3.8 ECAL layout of one quadrant showing $\eta$ values of region boundaries . 27
3.9 Electron energy resolution from the ECAL as a function of energy ................................. 29
3.10 HCAL components HB, HE, HO, and HF layout shown with dashed lines of $\eta$ ................................................................. 30
3.11 Reconstructed jet transverse energy resolution vs simulated jet transverse energy for barrel, endcap, and forward regions ..................... 31
3.12 Layout of the muon systems DT, CSC, and RPC shown with dashed lines in $\eta$ ................................................................. 33
3.13 Muon resolution using muon system only, tracker only, and both combined for separate $\eta$ regions ................................................ 34
4.1 Trigger efficiency of HLT_IsoMu24_eta2p1 measured vs muon transverse momentum for separate pseudorapidity regions ........................................ 37
4.2 Reconstruction efficiency for PF muons vs $p_T$ (top) and vs $\eta$ (bottom) ........................................ 47
4.3 Reconstruction efficiency for HPS hadronic taus vs $p_T$ ........................................ 49

5.1 Comparison of visible mass and collinear mass for simulated $H \rightarrow \mu\tau_{\text{had}}$ signal events after preselection in VBF category ........................................ 59

6.1 Collinear mass for the $Z \rightarrow \tau\tau$ control region ........................................ 65
6.2 Fake rate $f$ vs the jet $p_T$ (Left) and vs $\tau\eta$ (Right) ........................................ 68
6.3 Like sign control region showing good agreement between estimated misidentified lepton background and data ........................................ 69
6.4 Comparison studies for fake rates measured from $Z \rightarrow \mu\mu+X$ to QCD enhanced opposite sign and QCD enhanced same sign vs $p_T$ (left) and tau decay mode (right) ........................................ 70
6.5 Same sign control region for different definitions of fake rates. Very tight over tight (top left), very tight over medium (top right), very tight over loose (bottom left), and very tight over very loose (bottom right) ........................................ 72
6.6 Collinear mass after preselection requirements for 0-Jet, 1-jet, and 2-jet ........................................ 74

7.1 Shifted distributions for jet energy scale (left) and tau energy scale (right) for the 0 jet category. Inset plots are ratios of shifted to unshifted. 82

8.1 Collinear mass after full selection requirements for 0-Jet, 1-jet, and 2-jet after fitting to observed data ........................................ 84
8.2 Upper limits on Yukawa couplings for the LFV analysis expected (red) with 1 and 2 $\sigma$ error bands (yellow/green bands) and observed limit (black). Also shown are existing limits from tau decay searches (white), a reinterpretation of an LHC $H \rightarrow \tau\tau$ search (yellow line), branching fraction contours (black dashed) and the discussed “naturalness” condition (orange). ........................................ 89
TABLES

2.1 TAU DECAY BRANCHING FRACTIONS AND RESONANCES . . 11
4.1 SELECTION EFFICIENCIES FOR SIGNAL REGION . . . . . . 56
6.1 EVENT YIELDS IN SIGNAL REGION AT PRESELECTION . . . . 75
7.1 SYSTEMATIC UNCERTAINTIES INDEPENDENT OF $M_{coll}$ . . 78
7.2 THEORETICAL UNCERTAINTIES INDEPENDENT OF $M_{coll}$ . . 80
7.3 SYSTEMATIC UNCERTAINTIES DEPENDENT ON $M_{coll}$ . . . . 82
8.1 EVENT YIELDS IN THE SIGNAL REGION AFTER FULL SELEC-
TION . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 85
8.2 EXPECTED UPPER LIMITS, OBSERVED UPPER LIMITS, AND
BEST FITS FOR BRANCHING FRACTIONS . . . . . . . . . . . . . 86
CHAPTER 1

INTRODUCTION

One of the main goals of the Large Hadron Collider (LHC) is to search for the Higgs boson, a fundamental particle predicted to exist since the 1960’s. The Higgs boson is an integral part of the Standard Model of particle physics (SM) which describes the electromagnetic, weak nuclear, and strong nuclear interactions of particles. Despite years of searching, it was the only predicted SM particle that had yet to be discovered until the Compact Muon Solenoid (CMS) and A Toroidal LHC ApparatuS (ATLAS) experiments at the LHC jointly announced proof of the Higgs boson’s existence [2–4]. Having discovered this new particle, the focus shifts now to characterizing its properties. It must be determined if the discovered Higgs boson matches precisely that which was predicted for the SM, or if it has unexplained behaviors that call for physics beyond the Standard Model (BSM). One such possible behavior would be that if its decays do not conserve lepton flavor. The only way to determine if such behavior occurs is to search for it.

This dissertation will begin with a brief overview of the Standard Model of particle physics and the role of the Higgs boson. It will discuss the theoretical motivation for searching for lepton flavor violating (LFV) Higgs boson decays and possible models that include such decays as well as existing constraints. Brief descriptions of the Large Hadron Collider and Compact Muon Solenoid are given. From there, a search for the lepton flavor violating decay of the Higgs boson into a muon plus a tau lepton which in turn decays hadronically is described in detail. This description includes the reconstruction and selection criteria for physics objects such as particles, methods
to establish the presence of a signal, estimations of background processes, studies of systematic uncertainties, and a presentation of the results for this search as well as their inclusion in a combination with another measurement of the same Higgs decay but where the tau decays electronically. The combination allows an exclusion limits and best fit value of the branching fraction of $H \rightarrow \mu \tau$. 
2.1 Standard Model

The Standard Model (SM) of particle physics has a long history of development and consistently explains and predicts the behavior of fundamental particles. Fundamental particles are those which cannot be further broken down into parts. In the SM, fundamental particles with half-integer intrinsic angular momentum, or spin, known as fermions form the building blocks of matter. These come in two types, leptons and quarks. The leptons are further divided into two separate types, those that carry electric charge: the electron, muon, and tau and the electrically neutral leptons called neutrinos. There is one neutrino for each flavor of charged lepton so they can be organized:

\[
\text{leptons} : \begin{pmatrix} e \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}
\]  

where the three columns correspond to the three flavors of lepton: electron, muon, and tau.

Similar to the leptons there are six different quarks: up, down, charm, strange, top, and bottom. These are either up-type, having electric charge \(2/3\), or down-type which have electric charge \(-1/3\). There are three generations of quarks giving a
similar organization as the leptons:

\[
\text{quarks : } \begin{pmatrix} u \\ d \\ c \\ s \\ t \\ b \end{pmatrix}
\] (2.2)

where the columns correspond to first, second, and third generations while the top row is up-type quarks and the bottom row is down-type quarks.

To describe the interactions of these particles, the SM utilizes gauge theory. A simple example of a gauge comes from classical electromagnetism. Recall that the field strength tensor \( F_{\mu\nu} \), which quantifies the electric and magnetic fields, is given by

\[
F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}
\] (2.3)

where \( A \) is the electromagnetic four-potential. Under the transformation

\[
A'_{\mu} \rightarrow A_{\mu} - \partial_{\mu}\chi
\] (2.4)

the transformed four-potential \( A' \) will yield the same field strength tensor as \( A \) as long as \( \chi \) is differentiable and continuous. Thus, a gauge represents a degree of freedom that does not produce an observable effect.

Gauge theory deduces the interactions between the SM particles by requiring the Lagrangian density be invariant under gauge transformations of the particle wave functions. This reveals a gauge symmetry which underlies the different interactions. For the SM, this symmetry is \( SU(3)_C \times SU(2)_L \times U(1)_Y \) where the group \( SU(3)_C \) corresponds to the strong nuclear interactions while the group \( SU(2)_L \times U(1)_Y \) corresponds to the electroweak interactions. The generators of these symmetry groups each have a corresponding vector field which enters the Lagrangian to ensure invariance under the gauge transformations. Quanta of these vector fields, or gauge fields, are the gauge bosons known as gluons, \( W^\pm \) and \( Z \), and photons which correspond
to the strong interaction, the weak interaction, and the electromagnetic interaction respectively. The gauge bosons are thus particles of the SM and are said to mediate the force of their respective interactions.

The above only accounts for gauge bosons that are massless. This is fine for the photon and gluon which are indeed massless, but it is well established that the $W^\pm$ and $Z$ bosons are massive. In order to give the appropriate gauge bosons masses, spontaneous electroweak symmetry breaking (EWSB) is used in a process that is called the Higgs mechanism. Masses of gauge bosons arise when fields have non-zero expectations values for the vacuum ground state. This causes the corresponding symmetries of the Lagrangian to not be shared by the vacuum. Thus the vacuum state will have the $SU(3)_C$ symmetry corresponding to the massless gluon while the electroweak symmetry $SU(2)_L \times U(1)_Y$ must be broken into a $U(1)_{EM}$ corresponding to the massless photon.

The construction of a non-zero vacuum expectation value starts with the introduction of scalar Higgs field $\Phi$ with symmetry $SU(2) \times U(1)$ and can be separated into real and imaginary components and written as a doublet in $SU(2)$

$$\Phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^+_1(x) + i\phi^+_2(x) \\ \phi^0_1(x) + i\phi^0_2(x) \end{pmatrix}$$

(2.5)

The four degrees of freedom in this scalar field can then be counted as $\phi^+_1$, $\phi^+_2$, $\phi^0_1$ and $\phi^0_2$. Three of these four degrees of freedom are taken up to form the longitudinal polarizations of the three massive bosons since these were not considered prior to this point since massless bosons cannot have such polarization. In order to be bounded and have a non-zero vacuum expectation, the Lagrangian for the Higgs field must, at minimum, have the form:

$$V(\Phi^\dagger \Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2$$

(2.6)
where $\mu^2$ and $\lambda$ are positive constants. By convention, an orientation is chosen along the $\phi^0_1$ for the vacuum expectation value which gives a minimum at

$$\Phi_{\text{min}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad \text{where} \quad v = \sqrt{\frac{\mu^2}{\lambda}} \quad (2.7)$$

Assigning a charge of $+1/2$ under $U(1)$ symmetry for the scalar field in combination with it being a doublet in $SU(2)$ gives a gauge transformation of

$$\Phi \rightarrow e^{ia^a \tau^a} e^{i\beta/2} \Phi \quad (2.8)$$

where $\tau^a = \sigma^a/2$ ($a = 1, 2, 3$) are the Pauli matrices which are the generators for $SU(2)$. Because the vacuum expectation value takes the form in Equation 2.7, this transformation with $\alpha^1 = \alpha^2 = 0$ and $\alpha^3 = \beta$ will leave $\Phi_{\text{min}}$ invariant. This indicates the vacuum has retained one symmetry and thus one of the bosons will remain massless. This is a necessary part of the mechanism for the photon. The gluon will also remain massless as the $SU(3)_C$ symmetry of the strong interaction is not being broken.

The covariant derivative of $\Phi$ required for gauge invariance of the Lagrangian is

$$D_\mu \Phi = (\partial_\mu - igA^a_\mu \tau^a - i\frac{1}{2}g'B_\mu)\Phi \quad (2.9)$$

where $A^a_\mu \tau^a$ and $g$ are the gauge field and coupling constant corresponding to the $SU(2)$ symmetry while $B_\mu$ and $g'$ are that of the $U(1)$ symmetry. The mass terms for the gauge bosons can be seen in the kinetic term of the Lagrangian which is the square of this covariant derivative, evaluated at the vacuum expectation value from
Equation 2.7 Evaluating the $A_{\mu}^a \tau^a$ product, the relevant terms of the Lagrangian are

$$\Delta \mathcal{L} = \frac{1}{2} v^2 [g^2 (A_{\mu}^1)^2 + g^2 (A_{\mu}^2)^2 + (-gA_{\mu}^3 + g'B_{\mu})^2]$$  \hspace{1cm} (2.10)$$

This allows the identification of three massive bosons and one massless boson.

$$W_{\mu}^\pm = \frac{1}{\sqrt{2}} (A_{\mu}^1 \mp i A_{\mu}^2) \text{ with mass : } m_W = g \frac{v}{2}$$

$$Z_{\mu}^0 = \frac{1}{\sqrt{g^2 + g'^2}} (gA_{\mu}^3 - g'B_{\mu}) \text{ with mass : } m_Z = \sqrt{g^2 + g'^2 \frac{v^2}{2}}$$

$$A_{\mu} = \frac{1}{\sqrt{g^2 + g'^2}} (g'A_{\mu}^3 + gB_{\mu}) \text{ with mass : } m_\gamma = 0$$

The EWSB of the Higgs mechanism naturally produces mass terms for fermions. Similar to the process for the vector bosons, substituting the vacuum expectation value in these forms gives the masses. These are of the form $m_f = \lambda_f \frac{v}{\sqrt{2}}$ where $\lambda_f$ is the corresponding Yukawa coupling.

Since the Higgs is self-interacting, it should also have a mass. Excitations of the Higgs field can parameterized by expanding about the vacuum expectation value of Equation 2.7 in the form:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}$$  \hspace{1cm} (2.11)$$

Putting this into the potential terms of Equation 2.6 and gathering terms in powers of $H$ gives

$$V(H) = -\frac{1}{4} \mu^2 v^2 + \mu^2 H^2 + \lambda v H^3 + \frac{1}{4} \lambda H^4$$  \hspace{1cm} (2.12)$$

The $H^2$ term allows us to write the mass of the Higgs $m_H = \mu \sqrt{2}$ which can be rewritten using the expression for $v$ in Equation 2.7 to be $m_H = v \sqrt{2\lambda}$. Additionally the $H^3$ and $H^4$ terms give the three point and four point self interaction vertices. The same expansion introduced to the fermion mass terms mentioned above give the
couplings of the fermions to $H$ in the form

$$\mathcal{L}_f = m_f \bar{f} f (1 + \frac{H}{v})$$  \hspace{1cm} (2.13)

### 2.1.1 Higgs Production at the LHC

The Fermi constant $G_F$ can be expressed as

$$G_F = \frac{g^2 \sqrt{2}}{8m_W^2}$$  \hspace{1cm} (2.14)

and is precisely measured in muon decay experiments. Using the expression above for the $W$ mass, the vacuum expectation value can be rewritten solely in terms of this known constant

$$v = (G_F \sqrt{2})^{-1/2} = 246 \text{GeV}$$  \hspace{1cm} (2.15)

Using the value of $v$, the production cross section for Higgs at the LHC can be theoretically predicted as well as the branching fractions to SM decay modes. These are shown in Figure 2.1. Because of the high operating energy of the LHC, the content of the colliding protons are dominated by gluons. This causes the gluon fusion process to be the dominant way of producing Higgs bosons. The second most common way of producing Higgs bosons at the LHC is the vector boson fusion (VBF) process. Because the VBF process has two quarks in the final state that tend to have high energy, the resultant high momentum jets are useful in selecting these types of events. The diagrams for these two processes are shown in Figure 2.2.

### 2.1.2 Tau Lepton Decays

As discussed above, the Higgs boson couples to fermions proportionally to their masses. Therefore the tau is a good decay product of a Higgs boson to look for since it significantly heavier than the other leptons and does not hadronize like quarks.
Figure 2.1. Predicted SM Higgs production cross sections for various processes and decay mode branching fractions.

Figure 2.2: Dominant Higgs production processes at the LHC, gluon fusion (left) and vector boson fusion (right)
However, the tau has a very short lifetime of $\tau = 2.91 \times 10^{-13}$ s and thus will decay after traveling a very short distance from the interaction point. Considering relativistic effects, the distance the tau will travel in its lifetime is given by

$$d = \frac{p\tau}{m} c$$

where $p$ is the tau momentum, $\tau$ the tau lifetime, $m$ the tau mass and $c$ the speed of light. Selection criteria in this analysis require minimum tau momentum of around 40 GeV. This corresponds to a distance traveled of less than 2 mm. This means that unlike leptons or muons, only the decay products of taus are directly observed by the detector. The tau objects can be reconstructed from these products and the resolution of the CMS tracker is capable of separating the tau’s point of decay from the proton interaction point.

Tau decays produce a tau neutrino and two other fermions as shown in Figure 2.3. These may be one of the lighter leptons plus its associated neutrino or two quarks. In the later case, the quarks will produce a variety of final states during hadronization. As shown in Table 2.1, hadronic modes account for the majority of tau decays and are thus a desirable choice for searches. These hadronic modes most typically consist of one or three charged pions and some number of neutral pions. The neutral pions will subsequently decay into two photons. The presence of a resonance structure in these modes may also be exploited during reconstruction.

2.2 Lepton Flavor Violation

In 2012, both the CMS and ATLAS experiments announced an observation in the search for the Standard Model Higgs boson [2,4]. This was achieved through combining evidence across several expected decay modes of the Higgs boson. However,
Figure 2.3: Tau decay to a neutrino and two fermions

Table 2.1

<table>
<thead>
<tr>
<th>Final State</th>
<th>Resonance</th>
<th>Branching Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu^-\nu_\mu\bar{\nu}_\tau$</td>
<td>-</td>
<td>17.41%</td>
</tr>
<tr>
<td>$e^-\nu_e\bar{\nu}_\tau$</td>
<td>-</td>
<td>17.83%</td>
</tr>
<tr>
<td>$\pi^-\nu_\tau$</td>
<td>-</td>
<td>10.83%</td>
</tr>
<tr>
<td>$\pi^-\pi^0\nu_\tau$</td>
<td>$\rho(770)$</td>
<td>25.52%</td>
</tr>
<tr>
<td>$\pi^-\pi^0\pi^0\nu_\tau$</td>
<td>$a_1(1260)$</td>
<td>9.30%</td>
</tr>
<tr>
<td>$\pi^-\pi^+\pi^-\nu_\tau$</td>
<td>$a_1(1260)$</td>
<td>9.31%</td>
</tr>
<tr>
<td>$\pi^-\pi^+\pi^-\pi^0\nu_\tau$</td>
<td>$a_1(1260)$</td>
<td>4.61%</td>
</tr>
</tbody>
</table>

this evidence for the SM decay modes does not preclude the existence of decay modes which are unexpected. Evidence of behaviors not explained by the SM would require modification to the model, forming what is often referred to as Beyond the Standard Model (BSM) theories.

While the SM has been successful in describing and predicting the interactions of matter as described above, it falls short of its ideal goal of being a theory of everything. Obvious flaws include the absence of the gravitational interaction and
lack of dark matter, the existence of which is indirectly observed. The existence of flaws in the SM leads research to not only confirm its predictions but to consistently test its limitations in hopes of finding insight into the causes of its shortcomings.

In the description of the SM above, neutrinos remain massless. Unlike the charged leptons, there are no right-handed neutrinos which would be necessary to have mass terms analogous to the other fermions. As a consequence, the lepton flavor is conserved in SM interactions. Experimentally, this conservation has held true except in the case of observed neutrino oscillations. These oscillations of the flavor of a neutrino require a non-zero neutrino mass and show that lepton flavor is not a strictly conserved quantity. Knowing that it is not strictly conserved creates a need to verify conservation in various interactions, especially those involving newly discovered particles such as the Higgs boson whose properties are yet to be fully determined.

2.2.1 Examples of Models with Lepton Flavor Violation

Many BSM theories feature possible lepton flavor violating (LFV) interactions of the Higgs boson. The Minimal Supersymmetric Standard Model (MSSM) is an example of a BSM theory that may include such interactions \cite{5, 6}. The MSSM introduces a new boson partner to every SM fermion and vice versa while keeping the introduction of new particles and interactions to a minimum. LFV terms in the Lagrangian are a direct result of the MSSM and typically pose a concern given there is no evidence of LFV. Such a model must be careful to be consistent with observed LFV constraints.

Supersymmetry models necessitate more than one Higgs doublet. More generally, models which introduce a second Higgs doublet (2HDM) also give rise to LFV interactions \cite{5}. Often these models separate the Higgs doublets that up-type quarks, down-type quarks, and charged leptons interact with to avoid flavor-changing neutral currents (FCNCs) which violate lepton flavor.
Composite Higgs models theorize that the Higgs boson of the SM is actually a bound state of other new particles. Because different flavors of fermions have large discrepancies in mass, and thus large discrepancies in couplings to the Higgs, LFV is introduced as the fermions access the components of the composite Higgs differently \[7-9\]. There are also partially composite Higgs models such as Randall-Sundrum models which add an additional dimension to space-time motivated by discrepancy in strength of SM particle interactions and gravity. This extra dimension creates higher mass leptons which our familiar leptons mix with. This mixing of different masses causes the couplings to the Higgs boson to be non-diagonal in flavor space and thus lead to LFV \[10\].

2.2.2 Previous Constraints

The LHC is the first collider where a direct measurement of LFV Higgs decays is possible but a variety of constraints exist on the couplings via indirect methods. The defining characteristic of LFV interactions is a three point vertex between the Higgs boson and two different flavor charged leptons. For the direct search presented in this analysis, the specific process of \(H \rightarrow \mu\tau_{\text{had}}\) is shown in Figure 2.4. Indirect measurements of the same interaction must have the same vertex. One such measurement is that of \(B(\tau \rightarrow \mu\gamma)\) shown in Figure 2.5. The other LFV interactions \(H \rightarrow e\tau\) and \(H \rightarrow e\mu\) have analogous constraints from the processes \(\tau \rightarrow e\gamma\) and \(\mu \rightarrow e\gamma\) respectively. Since experimental constraints on \(B(\mu \rightarrow e\gamma)\) are three orders of magnitude stronger than those on \(B(\tau \rightarrow \mu\gamma)\) and \(B(\tau \rightarrow e\gamma)\), the branching fraction \(B(H \rightarrow e\mu)\) is already constrained to \(\mathcal{O}(10^{-8})\). Meanwhile, \(B(H \rightarrow \mu\tau)\) and \(B(H \rightarrow e\tau)\) are only constrained to be \(\mathcal{O}(10\%)\). While not as stringent as the above indirect constraints, other processes are capable of placing limits on relevant couplings. These include the decays \(\tau \rightarrow 3\mu\), \(\tau \rightarrow e\mu\mu\), and \(\mu \rightarrow 3e\) (Figure 2.6).
Figure 2.4: Diagram showing the LFV process $H \rightarrow \mu \tau_{had}$

Figure 2.5: Diagrams for $\tau \rightarrow \mu \gamma$ with one LFV vertex

muonium-antimuonium oscillations (Figure 2.7), muon dipole moment contributions (Figure 2.8), and $\mu \rightarrow e$ transitions in nuclei (Figure 2.9) [1].
Figure 2.6: Diagram for $\tau \to 3\mu$

Figure 2.7: Diagram for muonium-antimuonium oscillation via LFV

Figure 2.8: Diagram showing a contribution to muon dipole moments
Figure 2.9: Diagram showing $\mu \rightarrow e$ transitions in nuclei
CHAPTER 3

COLLIDER AND DETECTOR

3.1 Large Hadron Collider

The Large Hadron Collider (LHC) is located underground on the French-Swiss border near Geneva, Switzerland. It is operated as part of the European Organization for Nuclear Research (CERN) which has 22 member countries and employs over 2,000 staff members with thousands more science and engineering professionals and students involved in projects. The LHC consists of two rings that store protons that travel the 26.7 km circumference in opposite directions. The protons are kept in the ring by an 8.3 T magnetic field and are separated into clusters known as bunches. The counter rotating paths are made to intersect at each of the four LHC experiments: CMS, ATLAS, LHCb, and ALICE. The designed center of mass energy for these collisions is 14 TeV, but was delayed in reaching this energy due to an incident in 2008 caused by a faulty electrical connection that damaged superconducting magnets in the ring. Data was taken at center of mass energy of 7 TeV in 2010 and 2011, 8 TeV in 2012, and 13 TeV from 2015 to the present.

Since the LHC is not designed to accelerate protons from rest, to reach these energies, other accelerators must be used to accelerate protons in steps. These accelerators predate the LHC and were once or still are part of separate research programs. The protons originate from a tank of hydrogen gas which is ionized to produce protons. These are then fed into the linear accelerator Linac2 which accelerates them to 50 MeV and feeds them to the Proton Synchrotron Booster. There they are accelerated to 1.4 GeV, fed to the Proton Synchrotron, accelerated to 25 GeV, fed to
the Super Proton Synchrotron, accelerated to 450 GeV and then finally injected into the LHC where they are accelerated to their final energy. A map of the accelerator complex is shown in Figure 3.1.

![CERN accelerator complex](image)

**Figure 3.1. CERN accelerator complex**

The center of mass energy is an important parameter for a collider as it determines how great of mass of particle the collider is capable of producing. Additionally, the cross section of many sought after signals such as Higgs boson production increase with higher energy. The other centrally important parameter for a collider is the instantaneous luminosity $\mathcal{L}$. It is related to the number of events that will occur of
a given process by

\[ N_{\text{expected}} = \sigma_{\text{process}} \int L \, dt \]  

(3.1)

where \( \sigma_{\text{process}} \) is the cross section of the process, and the time integrated instantaneous luminosity represents the total amount of data taken. Since the LHC uses beams of protons that are bunched the instantaneous luminosity is given by

\[ L = f \frac{n_1 n_2}{4\pi \sigma_x \sigma_y} \]  

(3.2)

where \( f \) is the bunch crossing frequency, \( n_1 \) and \( n_2 \) are the number of protons in the colliding bunches, and \( \sigma_x \) and \( \sigma_y \) are the rms beam size in the transverse directions. The LHC recently reached its designed luminosity goal of \( 10^{34} \text{ cm}^{-2}\text{s}^{-1} \). The integrated luminosities recorded by CMS at various energies are shown in Figure 3.2.

Figure 3.2. Integrated luminosity of CMS pp collision data
3.2 Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is designed to detect all products of proton-proton collisions that are possible. This makes it a general purpose detector that is useful in measuring any processes that come from these high energy collisions. To accomplish this, the detector consists of several subsystems that ensure the different types of final state particles give unique signatures in how they interact with combinations of subsystems. These subsystems, described in the following sections, are the tracker, electromagnetic calorimeter (ECAL), hadronic calorimeter (HCAL), and muon system. More detailed descriptions can be found elsewhere \cite{11,12}. The subsystems are cylindrical in shape and are arranged concentrically in the listed order from the collision point out. This arrangement is shown in Figure\ref{fig:3.3}. The signatures left in the detector by different types of particles are illustrated in Figure\ref{fig:3.4}. The coordinate system utilized by CMS places the origin at the nominal collision point of the intersecting beams. In Cartesian coordinates, the positive $x$ direction points toward the center of the LHC ring, positive $y$ points vertically upward, and positive $z$ points along the beam which is roughly west. The cylindrical design of the detector lends itself to using cylindrical coordinates. CMS defines the polar angle $\theta \geq 0$ to be the angle from the positive $z$ axis making the negative $z$ axis at $\theta = \pi$. The azimuthal angle lies in the $x-y$ plane, is chosen to have the range $-\pi/2 \leq \phi \leq \pi/2$ and is measured from the positive $x$ axis. Typically pseudorapidity $\eta$ is used rather than the polar angle $\theta$. It is defined as

$$\eta = -\ln \tan \left( \frac{\theta}{2} \right)$$

This quantity is zero in the transverse plane and infinite along the $z$ axis. It is useful because differences in $\eta$ are Lorentz invariant to boosts in the $z$ direction if the
Figure 3.3. Cut away view of the CMS detector and its subsystems
Figure 3.4. Wedge slice in the transverse plane showing the interactions of different particles with the subsystems
particles are massless. Since it is typical at CMS that stable particles have $E \gg m$, the invariance approximately holds for particles with nonzero mass as well.

3.2.1 Superconducting Solenoid

CMS uses a cylindrical superconducting niobium titanium magnet to bend the trajectories of charged particles in the transverse plane. The inner-diameter of the magnet is 6 m and it resides between the HCAL and muon system. It creates a 3.8 T magnetic field aligned along the longitudinal direction. A steel return yoke is layered with the muon system and is used to guide the field lines back along the outside of the solenoid in the opposite direction. This field is used in conjunction with the tracker and calorimeters to measure transverse momentum. For a charged particle, a radius of curvature in the transverse plane can be measured by fitting the track. The radius of curvature is given by

$$r_T = \frac{p_T}{qB}$$  \hspace{1cm} (3.4)$$

where $p_T$ is the transverse momentum, $q$ is the charge, and $B$ is the longitudinal magnetic field. The sign of the charge can be determined by the direction of curvature while the magnitude is usually assumed to be unity. With a known magnetic field and measured radius, this gives the transverse momentum. Strong magnetic fields are desirable because as particle momentum increases, the radius of curvature increases eventually becoming large enough to be difficult to determine within the confines of the detector.

3.2.2 Tracker

The tracker is the subsystem closest to the interaction point and is divided into two separate components. Nearer the collision point is the pixel tracker. 65 million
pixels 100 μm × 150 μm in size provide the fine position resolution necessary in this region where track density is the highest. Also because of the high particle flux experienced by these pixels and high frequency of bunch crossings, the pixels must be radiation hard and have a fast response time. The choice of silicon for a detector material satisfies these requirements. Pixels are arranged as in Figure 3.5 with three concentric cylindrical layers called the barrel and two disk layers on each end of the barrel called endcaps.

![Figure 3.5. Pixel tracker layout](image)

Outside the pixel tracker, tracks are less dense and a silicon strip tracker is used. The strip tracker consists of the Tracker Inner Barrel (TIB), the Tracker Outer Barrel (TOB), the Tracker Inner Disks (TID) and the Tracker End Cap (TEC) which all consist of multiple layers of strips. The orientation of these components is shown
in Figure 3.6. Adjacent layers of strips are oriented in different directions so the position of the charged particle is located by the intersection of two active strips. The combination of the pixel and strip trackers give CMS the excellent track resolution necessary for identifying secondary vertices such as those from tau decays. Estimated resolutions for transverse momentum and impact parameters are shown in 3.7.

3.2.3 Electromagnetic Calorimeter

Like the tracker components, the Electromagnetic Calorimeter (ECAL) is divided into a cylindrical barrel region and flat endcap regions. Both regions consist of scintillating lead tungstate crystals with attached avalanche photodiodes for barrel crystals and vacuum phototriodes in the endcaps to collect the scintillated light. Lead
Figure 3.7. Tracker resolution estimates for transverse momentum (top), transverse impact parameter (left) and longitudinal impact parameter (right) using muons of 1, 10, and 100 GeV.
tungstate is chosen for its fast response, radiation hardness, short radiation length, and small Moliere radius. Although considered radiation hard, the intense radiation environment inside the detector does damage the crystals. However, this damage is monitored and can be corrected for in measurements. The radiation length of the crystals is important to ensure that electrons and photons typically deposit all of their energy in the ECAL. The Moliere radius characterizes the width of electromagnetic showers in the crystals and smaller radius leads to better position resolution. The crystal face toward the collision point is a 22 mm square while the length extends away from the collision point and is 23 cm. Crystals are intentionally slightly off directly pointing toward the collision point to avoid the cracks between crystals interfering with measurements. One quadrant of a cross sectional layout of the ECAL is shown in Figure 3.8.

Figure 3.8. ECAL layout of one quadrant showing η values of region boundaries
The ECAL is designed to measure the energy of electromagnetically interacting particles, namely the electron and photon. These particles deposit their energy into the calorimeter by showering. The process of showering is a combination of bremsstrahlung radiation and electron pair production. An electron traveling through the crystal can release photons via bremsstrahlung radiation. These photons can create electron pairs which can in turn release more photons. This chain reaction continues until the photons do not have enough energy to pair produce. Thus the photons created are proportional to the energy of the initial electron which is measured by photo sensors at the back of the crystal. Similarly an incident photon starts this process by first pair producing, leading to bremsstrahlung radiation etc. A plot of energy resolution of electrons is shown in Figure 3.9.

The ECAL also has a preshower detector that sits directly in front of the endcap crystals. In this region of the detector, neutral pions decay into two photons that are very close together. It is often important to differentiate between neutral pions and single photons so preshower has a finer position granularity than the ECAL crystals. It uses strips of silicon in a grid similar to the strip tracker layered with lead to ensure initiation of a shower.

3.2.4 Hadronic Calorimeter

The goal of the hadronic calorimeter (HCAL) is to measure the energy of everything that makes it through the ECAL except muons, typically hadrons. With muons accounted for by the muon system, the HCAL is hermetic, meaning effectively no energy escapes it. This is necessary to be able to use missing transverse energy to infer the presence of invisible particles such as neutrinos.

The HCAL consists of four distinct parts. The barrel (HB) and endcap (HE) sections form the hermetic portion of the detector with no avoidable gaps in coverage. A forward hadronic calorimeter (HF) is used to cover the range $3 < |\eta| < 5$ and
Figure 3.9. Electron energy resolution from the ECAL as a function of energy
is mainly used to measure instantaneous luminosity. To detect particles that may have punched through the HB, an outer portion (HO) is placed outside the magnet solenoid. The layout of these components is shown in Figure 3.10.

Figure 3.10. HCAL components HB, HE, HO, and HF layout shown with dashed lines of $\eta$

The HCAL is a sampling style calorimeter meaning it does not collect all of the energy deposited for measurement as the ECAL does. Hadrons would pass through far too much scintillating material to have a design like the ECAL. Instead the scintillators are layered with solid brass whose mass forces hadrons to shower in shorter lengths. The scintillation layers consist of scintillating plastic plates embedded with fibers that carry the scillation light to photodetectors. The total energy deposited
in the calorimeter can be extrapolated from the energy collected by sampling the shower with the layers of scintillator. The ability to balance transverse momentum heavily depends on quality measurements from the HCAL of all various jets created in collisions. Thus, jet energy resolution is an important quantity. Reconstructed jet energy resolution in simulation is shown in Figure 3.11.

Figure 3.11. Reconstructed jet transverse energy resolution vs simulated jet transverse energy for barrel, endcap, and forward regions

3.2.5 Muon System

Precise and robust muon measurement is a central feature for CMS. Muons occur as a signature to many interesting processes including Higgs decays. One of the main reasons they are so attractive is that they deposit little energy when passing through matter and so are able to pass through the tracker, calorimeters, and magnet. This
is very atypical of other particles so it is easier to identify a muon with a high degree of certainty.

The muon system has three major components. In each design, a passing muon ionizes a gas. The loose electrons then travel to a positive anode creating an electrical signal. In the barrel region of $|\eta| < 1.2$, drift tubes (DT) are used since the magnetic field is uniform and largely contained within the return yoke. Large non-uniform magnetic fields in the endcap regions adversely affect DTs so cathode strip chambers (CSC) are used in the region $0.9 < |\eta| < 2.4$. CSCs also benefit from their closely spaced anode wires giving a fast response time in this region with greater muon rates. The third component is mainly used for triggering on muons and utilizes resistive plate chambers (RPC). The very fast response times required for triggering are accomplished by gaps between plates being very small so that the electrons from the ionized gas travel a very short and consistent distance. The layout of the muon system is shown in Figure 3.12.

The muon system functions much like the tracker in that as the muons pass through, subsequent signals form a path along the muon trajectory. This extra track can be matched with the muon’s track from the tracker to form what is called a global muon track. Requiring a global track instead of just a tracker track improves momentum resolution in high energy muons but not greatly as can be seen in Figure 3.13. The greater benefit in identifying muons in the muon system is separating them from other types of particles.
Figure 3.12. Layout of the muon systems DT, CSC, and RPC shown with dashed lines in $\eta$.
Figure 3.13. Muon resolution using muon system only, tracker only, and both combined for separate $\eta$ regions.
4.1 Trigger Selection

It is infeasible to record data from every collision that happens in the detector, as it would require data bandwidths much higher than current technology allows. To work around this, CMS uses a multi-level trigger system to select the most interesting events at a rate which is manageable to record. With 50 ns bunch spacing, collisions happen at a rate of 20 MHz. The first step in filtering which collisions to record is the Level-1 Trigger. The decisions it can make must be simple to keep up with the very high incoming rate. Data from the detector is stored in small memory buffers on the detector electronics until the Level-1 Trigger is ready to accept it so it must have low latency to avoid overflowing the buffers. The Level-1 Trigger uses custom electronics to make fast calculations that would be slower with general purpose processors. These electronics consist of custom circuits specially produced known as Application Specific Integrated Circuits (ASICs) as well as reprogrammable circuits like Field Programmable Gate Arrays (FPGAs) and Programmable Logic Devices (PLDs). They may also include Random Access Memory (RAM). The flexibility of the FPGAs allows better optimization often at a cost advantage. Operations that are known not to be changing use ASICs.

The Level-1 Trigger reduces its input rate of 20 MHz down to an output rate of approximately 100 kHz. The second step is called the High Level Trigger and it receives the output of the Level-1. Between the two levels is a latency buffer which
is for the short term storage of data to help smooth out the randomness of events passed on by the Level-1. If the Level-1 happens to accept too many events in a given time window, the buffer overflows and the data is lost.

Unlike the Level-1, the High Level Trigger is purely software which is run on a processor farm. With the reduced rate coming from the Level-1, the High Level Trigger has more time to make decisions and thus can utilize more information and do more complex calculations than the Level-1. The High Level Trigger is designed to output a rate of 100 Hz which is low enough for all the information of the events to be permanently recorded [13, 14].

This analysis utilizes the trigger of a single muon with minimum transverse momentum of $p_T^{\mu} > 24 \text{ GeV}$ and pseudorapidity range $|\eta^{\mu}| < 2.1 \text{ GeV}$. These thresholds are chosen so that the trigger maintains a high efficiency as described below. This trigger path, known as HLT_IsoMu24_eta2p1 was chosen because muons are easier to identify than hadronically decaying taus and because it is unprescaled. Prescaled triggers only record a set percentage of events which pass their criteria which leads to reduced recorded luminosity. Conversely, unprescaled triggers record every event which passes the trigger’s criteria. It is necessary to prescale some triggers so as to fit all the desired types of triggers in the available bandwidth that can be recorded.

It is possible that a muon in the detector which should be recognized by the trigger criteria is not. This means that muon trigger has some efficiency which must be measured. This is done with a tag and probe method utilizing $Z \rightarrow \mu\mu$ decays in data. Tight identification criteria are applied on one muon, the tag, while loose identification is applied on the second muon, the probe. Because of the Z resonance, both objects are very likely to be real muons regardless of whether the probe is recognized by the trigger. Thus, the percentage of probe muons which satisfy the trigger is then a measure of the trigger’s efficiency. We measure this efficiency against
the muon’s transverse momentum in different pseudorapidity regions as shown in Figure 4.1.

4.2 Pileup

Since the processes being investigated at the LHC are generally rare, it is beneficial to have high luminosity so rarer processes may be observed more frequently. Increased luminosity, however, leads to more proton collisions per bunch crossing which create additional objects in the detector that obscure the objects from a collision of interest. The collision point within the detector with the highest sum of $p_T^2$ of all tracks coming from that vertex is called the primary vertex and is assumed to be the origin of the hard-scattering process. Any effects that originate outside the primary vertex are called pileup. Pileup mostly comes from additional collisions within the bunch crossing which is called in-time pileup. There is also out-of-time pileup which contributes less and comes from energy signatures in the calorimeters that are
from previous or subsequent bunch crossings \cite{15}. Pileup mostly affects jet energy and isolation values for different objects. Corrections for these effects are discussed below for relevant objects.

Not every event has the same number of pileup vertices and simulations may have different distributions of number of pileup vertices in an event compared to observed data. To correct for this, a new distribution is formed that is the ratio of observed over simulated number of vertices. Then, for any given simulation event, the value of the ratio is looked up for the number of pileup vertices and used as a weight for the event. This ensures that the distribution of the number of pileup events in simulation matches what is observed. This process is referred to as pileup reweighting.

4.3 Particle Flow

The CMS experiment utilizes the Particle Flow (PF) reconstruction technique to enhance the performance for jets, taus, and missing transverse energy \cite{16}. This technique focuses on using information from every subsystem to reconstruct and identify all stable particles from each event. The goal is to have each object within the detector assigned to a unique particle with optimally determined type, direction and energy. First, charged particle tracks and clusters within the calorimeters are identified. These elements from single subsystems are then matched by their locations in the detector to form “blocks.” An algorithm is then performed on each block to determine the particle or particles that most likely caused the elements. This algorithm starts with identifying reconstructed global muons which become PF muons if its momentum is within three standard deviations of the muon measured in the tracker only. If successful the track is removed from the block. Next, electrons are treated in a similar fashion but become PF electrons if their tracks, extrapolated all the way to the ECAL, have a corresponding expected energy deposit in the ECAL. After the muons and electrons are taken out, the remaining tracks in the block are
subJECTED TO TIGHTER CRITERIA TO CUT DOWN FAKE TRACKS. THE REMAINING ELEMENTS OF THE
BLOCK ARE TYPICALLY HADRONS AND PHOTONS. ANY REMAINING TRACKS ARE ATTRIBUTED TO THE
MOST COMMON CHARGED HADRON, THE PION. LASTLY, ANY REMAINING ECAL AND HCAL
ENERGY DEPOSITS ARE ATTRIBUTED TO PHOTONS AND NEUTRAL PIONS RESPECTIVELY.

THIS GIVES A LIST OF ALL MUONS, ELECTRONS, PHOTONS, NEUTRAL HADRONS, AND CHARGED
HADRONS IN THE EVENT. GROUPS OF PARTICLES CAN THEN BE CLUSTERED INTO JETS FOR USE IN
ANALYSES WHICH WILL BE MORE ACCURATE THAN JETS FORMED FROM JUST CALORIMETRY CLUSTERING
AS CALORIMETRY ONLY JETS OFTEN DO NOT HAVE ANY INFORMATION ON HOW TO SHARE ENERGY
BETWEEN TWO ADJACENT JETS. THIS IS ALSO AN IMPROVEMENT TO IDENTIFYING HADRONICALLY
decaying taus by allowing the identification of each component of the different hadronic
decay modes which all belong to a Particle Flow jet. Identification of hadronically
decaying taus is further discussed in Section 4.4.5. Additionally, identifying all the
particles in an event in this way improves the accuracy with which missing transverse
energy is calculated.

4.4 Objects

4.4.1 Jets

Jets are clusters of particles that have a common origin and thus have trajectories
that roughly form a cone shape. Because common Quantum Chromodynamic (QCD)
processes often produce jets, they are abundant in hadron collisions. Particle Flow
gives a list of individual particles, so in order to have jets these particles must be
clustered together. We use jets clustered with the anti-$k_t$ jet clustering algorithm
with a distance parameter of $R = 0.5$ [17]. This algorithm, like others, clusters
separate objects together based on a weighted distance to high $p_T$ objects called
seeds. Unlike other algorithms, this uses a negative power of transverse momentum
to weight the distance between objects, hence “anti-$k_t$.” This ensures that softer
objects do not largely affect the shape of the jet cone and tend to be clustered with the seeds rather than to themselves. Additionally, nearby seeds will share surrounding energy proportional to their momentum or be combined into a single jet based on the distance parameter \( R \).

Jets, whether simulated or in observed data need to be corrected for several effects that cause measured jet energy to differ than that of true energy of particles which should comprise the jet. Firstly, during clustering, energy coming from a pileup interaction may overlap a jet and be clustered together. For PF jets this is corrected using a hybrid jet area method which utilizes an average \( p_T \) density per unit area \( \rho \) which represents the soft jet activity from pileup as well as noise and uninteresting parts of the hard scatter known as the underlying event \[18\]. The density is calculated for each event and is the median of the quantity \( p_T/A \) calculated for every jet in the event where \( A \) is the area of the jet. Since there are a large number of soft jets, taking the median makes the density insensitive to the presence of high \( p_T \) jets. This density does not account for a non-uniform energy response with respect to \( \eta \) which must be calculated separately from averaging zero-bias events binned in \( \eta \).

After the pileup correction, a correction factor is applied to account for invisible particles that will be in the jet. This correction factor uses Monte Carlo information to compare reconstructed jet \( p_T \) to that of the generated particles of the jet. The ratio of the two momenta, binned in jet \( p_T \) and \( \eta \), can be then used to correct the energy.

Lastly, corrections for residual \( \eta \) and \( p_T \) dependence typically on the order of a few percent are considered for observed data. The dependence on \( \eta \) is measured using a QCD dijet sample in which two jets that are back to back in azimuthal angle are selected. Since these two jets will fall in different \( \eta \) regions and should have a sum of zero momentum in the transverse plane, they can be compared to determine a relative difference in response based on \( \eta \) region. This relative response is then used
to correct the jet energy to be flat versus $\eta$. In order to correct for the response versus jet $p_T$, samples of $Z + jets$ and $\gamma + jets$ are used. These are chosen because the $Z/\gamma$ energy can be accurately measured as well as single jet recoiling from them. Again exploiting conservation of momentum in the transverse plane, the jet energy response can be calculated as a function of $p_T$.

For this analysis, jets are only used explicitly for categorization of events according to the number of jets. The zero, one, and two jet categories correspond roughly to the Higgs production mechanisms gluon fusion, associated production, and vector boson fusion respectively. In order to qualify as a jet for use in categorization, the jet must have $p_T > 30$ GeV, have $|\eta| < 4.7$ and pass loose identification working point. Identification for jets uses a set of criteria based on several quantities. The loose working point requires:

- $p_T > 10$ GeV
- fraction of charged hadrons $> 0.0$
- fraction of neutral hadrons $< 1.0$
- at least one constituent particle
- fraction of energy in ECAL $< 1.0$

The constraints involving charged portions of the jet are only applied to the region $|\eta| < 2.4$ which is the region of tracker coverage.

4.4.2 Missing Transverse Energy

Missing transverse energy, $E_T^{miss}$, represents neutrinos which are invisible to the detector. Because the incoming protons which will collide have nearly zero momentum in the transverse plane, the sum of all outgoing particle momentum should also be zero in this plane by conservation of momentum. Utilizing Particle Flow described above, we have a list all the particles in the event and their momenta. This makes the definition of PF $E_T^{miss}$ simply:
\[ \vec{E}_{T}^{\text{miss}} = - \sum \vec{p}_{T}^{PF} \] (4.1)

Since this quantity involves summing the momenta of all particles including the many jets from pileup interactions, the jet energy corrections described above must be propagated to all the particles which can be clustered as jets.

4.4.3 Electrons

Although our signal does not include any electrons, their reconstruction is still relevant to the analysis. To look for events such as \( H \rightarrow \mu \tau_{\text{had}} \), it is desirable to veto any events which a hard electron also comes from the primary vertex. In order to do so, it is necessary to have well identified electrons. Electron reconstruction requires matching an electron track to an energy cluster in the ECAL [19]. This is complicated by the fact that electrons can radiate a significant amount of their energy in the form of bremsstrahlung radiation. As an electron passes through the material of the tracker, interactions with atoms may cause the release of bremsstrahlung photons. Because of the magnetic field of the solenoid, the electron’s path is curved causing these photons to spread out in the \( \phi \) direction. Depending on the region of the detector, and thus how much material the electron must pass through to get to the ECAL, these photons can carry a majority of the electron’s original energy. Therefore, the bremsstrahlung photons are clustered together with electron energy signature. Additionally, the bremsstrahlung photons may also convert to \( e^+e^- \) pairs within the tracker. These too must be clustered together with the original electron’s energy deposit to ensure accurate electron energy reconstruction. Since the electron is radiating away energy, the radius of curvature shrinks as it travels through the tracker. This must accounted for when reconstructing the electron tracks.

Several factors that degrade energy and momentum reconstruction for electrons
must be accounted for. A multi-variate analysis (MVA), trained on simulated electrons, uses a collection of input variables to address these issues simultaneously. Energy leakage out of the ECAL cluster into other crystals, gaps between components, or into the region between barrel and endcap calorimeters is addressed by input variables for shower shapes, cluster positions, number of clusters attributed to the electron, and comparison between seed crystal position and cluster position. Energy leakage into the HCAL is accounted for by utilizing the ratio of energy in the ECAL cluster to the energy in the HCAL at the same \( \eta \phi \) location known as H/E. Energy from pileup interactions is considered by including the number of pileup vertices and the average energy density from pileup \( \rho \). This is the same density discussed for jet object.

Similar to the jet objects, electrons have a set of recommended identification working points in place for use in CMS analyses. Identification for electrons is the output of a boosted decision tree (BDT) algorithm \[20\] that combines several variables including geometric, kinematic, tracks, isolation, and shower shape parameters. A BDT is a type of multivariate analysis that utilizes machine learning to map out conditions input variables to draw conclusions about a target variable typically called a discriminator. In this case, the discriminator is a variable ranging from 0 to 1 where values closer to one indicate more certainly an electron. A tight working point is established for the discriminator and is called CicTightIso.

The performance of electron reconstruction is monitored by using several \( e^+e^- \) resonances including \( J/\psi, \psi(2S), \Upsilon(1S) \), and combined \( \Upsilon(2S) + \Upsilon(3S) \). For the electron veto used in the LFV search, we select the tight working point be satisfied by an isolated electron in order to veto the event. How isolated one object is from others is quantified by comparing the energy attributed to the object of interest to the surrounding energy. Surrounding energy here is defined by anything that lands within a cone centered about the object of interest, in this case the electron. It is
typical to use relative isolations which are the ratio of energy in the cone to that of the electron. Objects that originate from a pileup vertex outside the cone may still deposit their energy inside the cone. Since charged objects leave tracks, Particle Flow can be used to exclude those that originate from pileup interactions from the sum of energy surrounding the electron. However, since neutral particles do not leave any tracks their contribution is estimated to be half the charged contribution. The subtraction of the neutral contribution is known as a $\Delta \beta$ correction. This gives a final isolation quantity of:

$$I_{\Delta \beta_{\text{rel}}} = \frac{\sum p_{T}^{\pm, PV} + \max(0, \sum E_{T}^{0, PV} + \sum E_{T}^{\gamma, PV} - 0.5 \sum p_{T}^{\pm, PU})}{p_{T}^{e}}$$

(4.2)

where $\pm$ and 0 superscripts indicate charged and neutral quantities respectively and PV and PU indicate interactions from the primary vertex and pileup vertices respectively. The max function is included to ensure that the total corrected neutral contribution is positive. The size of the isolation cone is chosen to be $\Delta R < 0.4$. To be deemed isolated for the purpose of the LFV electron veto, $I_{\Delta \beta_{\text{rel}}} < 0.3$.

4.4.4 Muons

Being a charged particle, muons leave tracks in the tracker. Additionally they also leave muon tracks in the muon system located outside the solenoid. Two reconstruction methods are used in CMS for muons based their use of these objects. One approach is called Global Muon reconstruction and relies on both systems [21]. It requires a tracker track and a muon system track be matched together by comparing parameters of the two tracks which are propagated into common regions. A global muon track can then be fitted by combining hits from both the tracker track and the muon system track. The other method is Tracker Muon reconstruction which does not use the muon system as heavily. This method extrapolates tracker tracks
to the muon system accounting for magnetic field, expected energy loss in materials, and Coulomb scatterings. If a single segment in the muon system matches this extrapolated tracker track, the object is considered Tracker Muon.

The tight identification working point for muons is recommended for most analyses and satisfies all criteria of both reconstruction methods. To satisfy tight identification, a muon must

- be recognized by Particle Flow
- be reconstructed as Global Muon
- have Global Muon track fit $\chi^2/ndof < 10$
- have at least one muon chamber included in Global Muon track fit
- be matched to at least two muon stations
- have impact parameters to primary vertex of $d_{xy} < 2$ mm and $d_z < 5$ mm
- have at least one hit in the pixel tracker
- have hits in at least 5 layers of the tracker

In text, this collection of conditions is sometimes referred to as PFIDTight

Muon energy measurement is degraded by inaccurate energy loss modeling and alignment effects. The muon energy scale is monitored and corrected using different methods depending on their $p_T$. For muons with $p_T < 100$ GeV, two methods are both used to determine correction factors. The first is MuScleFit (Muon momentum Scale calibration Fit), which compares tight identification muons from $Z$ boson decays and compares to the same from simulation. This comparison yields a relative oscillation in $\phi$ which points to slight misalignments of the detector. The oscillation is fit with a function that can then be used to correct for the effect. The second method is SIDRA (Sim Driven Analysis), which introduces parameterized energy shifts and resolution distortions to simulated $Z \rightarrow \mu\mu$ decays until the modified simulation matches observed data. The parameters can then be determined through fits and used
to correct the energy scale. Both of these methods are used and any difference in correction factors are used to study the uncertainty of the corrections.

Isolation is an important quantity for muons in this analysis. Muons from decays from bosons such as W, Z, and H will be isolated from other objects while muons from hadronic and heavy flavor decays will tend to be accompanied by various other objects and thus not isolated. A relative $\Delta \beta$ isolation is used similarly to as was defined for electrons above:

$$I^{\Delta \beta}_{rel} = \frac{\sum p_{T}^\pm, PV T + \max(0,0.0, \sum E_{T}^{0,PV} + \sum E_{T}^{\gamma,PV} - 0.5 \sum p_{T}^\pm,PU)}{p_{T}^\mu}$$

(4.3)

where $\pm$ and 0 superscripts indicate charged and neutral quantities respectively and PV and PU indicate interactions from the primary vertex and pileup vertices respectively. The max function is included to ensure that the total corrected neutral contribution is positive and size of the isolation cone is again $\Delta R < 0.4$. In text this value is sometimes written as RelPFIsoDB.

An efficiency for reconstructing muons can be measured using the tag and probe method as described for trigger efficiencies. These efficiencies are shown in Figure 4.2.

4.4.5 Hadronically Decaying Taus

As calculated in Section 2.1.2 tau leptons relevant to this analysis only travel about 2 mm before decaying and therefore do not directly land in the detector. This analysis uses hadronic decay modes of the tau which account for nearly two thirds of all tau decays. Of the hadronic decays, the most common consist of one or three charged pions in addition to a neutrino and possible neutral pions as well. The neutral pions themselves will also decay almost exclusively to two photons. As discussed in the electron section these photons may convert to $e^+e^-$ pairs which may radiate additional bremsstrahlung photons which spread out in $\phi$ direction due to
Figure 4.2. Reconstruction efficiency for PF muons vs $p_T$ (top) and vs $\eta$ (bottom)
the magnetic fields acceleration of the electrons. Because Particle Flow gives such
a complete event reconstruction, the particle content of hadronic tau decays can be
seen instead of just seeing a jet cluster.

For this analysis the Hadrons Plus Strips (HPS) reconstruction method is used [22].
To account for the spread in $\phi$ of the $\pi^0$, this algorithm matches electrons and pho-
tons into strips. The algorithm starts by forming strips centered about a PF jet,
then iteratively reforming the strip to include nearby electrons and photons. Com-
bining the strips with reconstructed charged hadrons within the PF jet, four classes
of hadronic decays can be defined:

- Single hadron
- Single hadron + one strip
- Single hadron + two strips
- Three hadrons

These classifications do not have a one to one correspondence with decay modes
because it is possible that the neutral pions are not reconstructed as strips because
they have too low energy. All hadrons and strips are required to be within a cone of
$\Delta R < 2.8 \text{ GeV}/p_T^{\tau}$ and be within $\Delta R < 0.1$ of the original PF jet. Since the taus
travel a small distance before decaying, charged decay product tracks will originate
from a secondary vertex near the primary vertex. Because of the fine resolution in
the pixel tracker, these secondary vertices can be seen for taus and the decays with
three charged hadrons are required to share a secondary vertex. With these criteria
in place, candidates are reconstructed as specific decay mode hypotheses. Since pions
are much more common than kaons for tau decays, all hadron masses are assumed to
be pion masses. Additionally, modes are required to be consistent with intermediate
resonances $\rho$ and $a_1$ where appropriate. The allowed invariant mass windows are:

- $\pi^0$: 50 - 200 MeV
Figure 4.3. Reconstruction efficiency for HPS hadronic taus vs $p_T$

- $\rho$: 0.3 - 1.3 GeV
- $a_1$: 0.8 - 1.5 GeV

In cases where a $\tau_{\text{had}}$ candidate is consistent with more than one decay mode hypothesis, the one giving the highest $p_T^{\tau_{\text{had}}}$ is chosen. The efficiency of tau reconstruction using HPS working points is estimated from generated hadronic tau decays in simulated data. These efficiencies are shown in Figure 4.3.

A $\Delta\beta$ corrected absolute isolation is used and is defined by

$$I^{\Delta\beta} = \sum p_T^{\pm, PV} + \max(0, 0, \sum E_T^{0, PV} + \sum E_T^{\gamma, PV} - 0.5 \sum p_T^{\pm, PU})$$ (4.4)
working point is set for the isolation variable and in conjunction with the 3 tracker
hit requirement is referred to as a binary value tightIsoHits3.

Because hadron and photon energy reconstruction is precise, it is expected that
$\tau_{\text{had}}$ energy reconstructed be precise as well. However, small correction factors on
the order of a couple percent are determined through comparing observed $Z \rightarrow \mu \tau_{\text{h}}$
data to simulation. This process is very sensitive to the $\tau_{\text{had}}$ energy scale because the
muon energy reconstruction is also very precise.

It is possible that a muon be erroneously reconstructed as a hadronic tau decay.
To account for this, a discrimination is applied against $\tau_{\text{had}}$’s that meet a set of tight
muon criteria. The muon criteria require a $\tau_{\text{had}}$ candidate to have a muon segment be
matched to the leading track of the tau decay. A muon segment is a small collection
of hits in the muon chambers forming a short muon track stub. Since it is possible for
high $p_T$ $\tau_{\text{had}}$’s to have hadronic showers that leak into the muon chambers it is also
required that there be a hit in the last two stations of the muon chambers. Finally, it
is required that the candidate deposit less than 20% of its energy in the calorimeters.
If these criteria are met it is assumed that the candidate is a muon and not a $\tau_{\text{had}}$ and
thus thrown out. These requirements are collectively called AntiMuonTight2.

The similarities between electron reconstruction accounting for bremsstrahlung
radiation and hadronic tau decays with strips can cause an electron to be recon-
structed as a hadronic tau instead. An MVA discriminator is used identify $\tau_{\text{had}}$ can-
didates that are likely to be electrons mistakenly reconstructed as hadronic tau de-
cays. Candidates identified as such are thrown out during selection of the $\tau_{\text{had}}$’s. The
working point used for this identification is called AntiElectronLoose.

4.5 Event Selection

Datasets consist of recorded events that passed a designated trigger as discussed
in Section 4.1. After reconstruction of objects in the recorded data, selection criteria
tailored to individual searches are applied. In this analysis, this is done in three distinct steps. First, a set of criteria termed preselection is applied. Then, events are separated into categories. Lastly, final selection criteria are applied to individual categories.

4.5.1 Preselection

The goal of the preselection criteria is to keep only events that have the final state particles corresponding to the signal process and that these objects be well defined. At this stage there is no effort to reduce background processes with the same final state particles and the distributions can be used as high statistic tests of predictive modeling. For this analysis, the relevant final state is one $\mu$ and one $\tau_{\text{had}}$ of opposite charges with no additional hard leptons from the PV. Because we are looking for a Higgs boson of mass 125 GeV decaying into two leptons which are much lighter than the Higgs, it is expected that these prompt leptons have fairly high momentum. Thus, preselection criteria includes momentum cuts of $p_T > 30$ GeV for both leptons. Additionally $\eta$ values are constrained to be within the fiducial region of the detector. For muons, criteria for the isolation values and identification working points as discussed in Section 4.4.4 are applied. Hadronic taus similarly have isolation and identification criteria in addition to anti-e and anti-$\mu$ discrimination as defined in Section 4.4.5. An explicit requirement that the HPS algorithm identifies a particular decay mode for the tau is also applied sometimes called DecayFinding. A list of the preselection criteria is included for reference:

- **muon**
  - $p_T^{\mu} > 30$ GeV
  - $|\eta|^{\mu} < 2.1$
  - **Isolation:** RelPF IsoDB $< 0.12$
  - **ID:** PFID Tight
• tau
  - $p_T^{\tau} > 30$ GeV
  - $|\eta|^{\tau} < 2.3$
  - $dZ^{\tau} < 0.2$
  - Isolation: tightIsoHits3
  - ID: PFTau (HPS) with tau DecayFinding
  - Veto: AntiElectronLoose, AntiMuonTight2

• event
  - muon and tau opposite charge.
  - $\Delta R_{\mu-jet} > 0.4$ for any jet with $p_T > 30$ GeV
  - $\Delta R_{\tau-jet} > 0.4$ for any jet with $p_T > 30$ GeV
  - Veto: extra muons with RelPFIsolDB < 0.15 and $p_T > 5$ GeV
  - Veto: extra electrons with CicTightIso and $p_T > 10$ GeV
  - Veto: extra taus with $p_T > 20$ GeV

4.5.2 Categorization

After preselection, events are separated into categories based on how many additional jets are in the event. The categories are: 0 jets, 1 jet, and 2 or more jets. The presence of energetic jets can alter the topology of the rest of the event and thus optimal background reducing selection is aided by separating the different categories. These categories roughly correspond to the three major Higgs boson productions methods at the LHC which are gluon fusion, associated production, and vector boson fusion (VBF). In order to be considered for categorization a jet must have $p_T > 30$ GeV, $|\eta| < 4.7$ and pass the loose identification working point discussed in Section 4.4.1.

4.5.3 Final Selection and Optimization

The goal of the final selection criteria is to make any signal as prominent over the backgrounds as possible. To do so, a set of variables that offer possible separation
of signal events and background events are collected. These include the $p_T$ of the $\mu$ and the $\tau_{\text{had}}$, $\Delta \phi$ between the two leptons, and between the leptons and the $E_T^{\text{miss}}$. Because the LFV signal is a two body decay it is expected that the $\mu$ and the $\tau_{\text{had}}$ be back to back in $\phi$. Additionally, the only neutrino in the signal process is from the hadronic tau decay and so the $E_T^{\text{miss}}$ is expected to be near the visible portion of the $\tau_{\text{had}}$ and opposite the $\mu$. Another variable considered is the transverse mass of either lepton with the $E_T^{\text{miss}}$. The transverse mass of objects of a and b is defined by

$$M_T^2(a, b) = (E_{T,a} + E_{T,b})^2 - (\vec{p}_{T,1} + \vec{p}_{T,2})^2$$ (4.5)

This quantity is especially useful when one of the objects is the $E_T^{\text{miss}}$ for which a traditional invariant mass cannot be formed as only the transverse component is known. A lepton that shares a parent particle with a neutrino, such as the $\tau_{\text{had}}$ in our signal, will have a very different $M_T(\ell, E_T^{\text{miss}})$ distribution than a lepton that is not.

The VBF category has additional variables considered due to the two high energy jets produced alongside the Higgs boson in addition to the usual requirements for jets used in categorization. These variables are the $\eta$ separation of the two jets, their invariant mass and the presence of additional jets between the two VBF jets.

The set of variables is put through an optimization process to maximize the significance which is defined as $S/\sqrt{S + B}$ where S is the number of simulated signal events and B is the number of estimated background events which occur in the signal region of the signal variable defined in Chapter 5. The optimization of the cut values for considered variables starts at preselection with the additional variables only cut on their loosest possible value i.e. $\Delta \phi_{\mu, \tau_{\text{had}}} > 0.0$ and the significance computed. Then, for one variable, the cut is tightened incrementally over a wide range, while holding all other cuts at the same value. The significance is computed at each increment giving a significance distribution versus that variable. The variable is then returned.
to its starting value, and the same is done for each variable considered. Then, the cut value of the single variable that produces the highest gain in significance is chosen to be implemented. The process starts over again in the same way with the one exception of the starting cut value for the one variable chosen in the previous round is no longer its loosest value but the value chosen. During successive rounds all variables are tested even those that have had chosen values in previous rounds. This is because the optimal value for a variable may change as values for other variables are chosen. The process is repeated until the significance no longer increases, indicating an optimized set of cuts. Because of the differences in the categories topologies, this is done individually for each category. The final selection criteria is listed here:

**0 Jet Category:**
- $p_T^\mu > 40 \text{ GeV}$
- $p_T^\tau > 35 \text{ GeV}$
- $M_T = \sqrt{2p_T^\tau E_T^{miss}(1 - \cos \Delta \phi)} < 50 \text{ GeV}$
- $\Delta \phi_{\mu - \tau_{had}} > 2.7$
- **Veto:** jets with $p_T^{jet} > 30 \text{ GeV}$, $|\eta| < 4.7$, and LooseID

**1 Jet Category:**
- $p_T^\mu > 35 \text{ GeV}$
- $p_T^\tau > 40 \text{ GeV}$
- $M_T = \sqrt{2p_T^\tau E_T^{miss}(1 - \cos \Delta \phi)} < 35 \text{ GeV}$
- One and only one jet with $p_T^{jet} > 30 \text{ GeV}$, $|\eta| < 4.7$, and LooseID

**2 Jet Category (VBF):**
- $p_T^\mu > 30 \text{ GeV}$
- $p_T^\tau > 40 \text{ GeV}$
- $M_T = \sqrt{2p_T^\tau E_T^{miss}(1 - \cos \Delta \phi)} < 35 \text{ GeV}$
• Two jets with:
  
  - $p_T^{jet1} > 30$ GeV and $p_T^{jet2} > 30$ GeV
  - $|\eta_{jet1}| < 4.7$ and $|\eta_{jet2}| < 4.7$
  - $\Delta(\eta_{jet1-jet2}) > 3.5$
  - $M_{jet1-jet2} > 550$ GeV
  - **Veto:** central jets with:
    
    * $p_T^{jet} > 30$ GeV
    * $|\eta_{jet1}| < |\eta_{jet}| < |\eta_{jet2}|

The yields given by preselection of the various background and signal processes shown in Table 6.1 can be compared to those of the final selection in Table 8.1 to calculate the selection efficiencies for different processes in different categories. These efficiencies are shown in Table 4.1.
### TABLE 4.1

SELECTION EFFICIENCIES FOR SIGNAL REGION

<table>
<thead>
<tr>
<th>Category</th>
<th>0-jet eff. (%)</th>
<th>1-jet eff. (%)</th>
<th>2-jet eff. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fakes</td>
<td>9.1</td>
<td>4.4</td>
<td>0.0</td>
</tr>
<tr>
<td>$Z \rightarrow \tau \tau$</td>
<td>1.3</td>
<td>1.3</td>
<td>0.0</td>
</tr>
<tr>
<td>$ZZ, WW$</td>
<td>4.8</td>
<td>3.1</td>
<td>0.2</td>
</tr>
<tr>
<td>$Z \rightarrow ee$ or $\mu \mu$</td>
<td>3.2</td>
<td>2.2</td>
<td>0.2</td>
</tr>
<tr>
<td>$t \bar{t}$</td>
<td>1.6</td>
<td>1.7</td>
<td>0.0</td>
</tr>
<tr>
<td>$t, \bar{t}$</td>
<td>11.9</td>
<td>21.2</td>
<td>0.1</td>
</tr>
<tr>
<td>SM Higgs Background</td>
<td>6.4</td>
<td>8.5</td>
<td>6.5</td>
</tr>
<tr>
<td>Sum of Backgrounds</td>
<td>5.6</td>
<td>3.3</td>
<td>0.0</td>
</tr>
<tr>
<td>LFV Higgs Signal</td>
<td>36.4</td>
<td>27.6</td>
<td>5.1</td>
</tr>
</tbody>
</table>
5.1 Signal Variable

The existence or absence of a signal is established from the distribution of a single variable. When searching for a particle that decays, the invariant mass of the daughter particles is a common choice for this signal variable. This is because signal events form a peak which reveals or confirms the mass of the sought after particle. For decays in which all final state particles are visible, this peak in the invariant mass distribution should be fairly sharp and centered about the mass of the parent particle. However, decays that include invisible products peak below the parent mass and are of decreased sharpness as the portion of energy that is invisible varies. This flattening increases the difficulty of identifying the presence of a signal.

For the process $H \rightarrow \mu \tau_{had}$ the invisible energy from the tau decay is typically small so the invariant mass of the muon and visible portion of the tau could be used, but some additional information can be used to sharpen the peak.

The hadronic decay of a tau lepton only includes one neutrino which is the only invisible particle in the signal process $H \rightarrow \mu \tau_{had}$. This means that in an idealized detector with no underlying event, the $E_T^{miss}$ would exactly reflect the transverse momentum of this neutrino. In a real experiment, the $E_T^{miss}$ also includes contributions from other sources such as invisible particles from processes other than the signal process and general mismeasurement. However, a single energetic neutrino such as that in our signal process will dominate compared to these other sources. Therefore,
an improvement over the visible mass $M_{\text{vis}}$ would be to include the $E_T^{\text{miss}}$ into the invariant mass since it is predominately the neutrino momentum.

An even better solution is known as the collinear mass $M_{\text{coll}}$. Since the tau lepton in our signal process carries energy much greater than its mass, its decay products are highly boosted in the direction the tau was traveling. Thus the decay products will be collinear to very good approximation. Any component of the $E_T^{\text{miss}}$ that is not collinear to the direction of the visible tau decay products is then not likely to be due to the signal neutrino but rather the other effects mentioned above and are best discounted when forming a mass for use as a signal variable. When forming the collinear mass, this is done by approximating the neutrino transverse momentum as the component of the $E_T^{\text{miss}}$ that lies along the direction of the visible tau products:

$$p_T^\nu = E_T^{\text{miss}} \cdot \hat{p}_{T \text{vis}}$$

Since we have no information on the longitudinal component of the neutrino momentum, it is useful to form the fraction of the tau momentum carried by the visible decay products

$$x_{\text{vis}} = \frac{|\vec{p}_{T \text{vis}}|}{|\vec{p}_{T \text{vis}}| + |\vec{p}_{T \nu}|}$$

This fraction is then used to scale visible tau four momentum to a total momentum which accounts for invisible products

$$(E_{\tau_{\text{total}}}, \vec{p}_{\tau_{\text{total}}}) = \left(\frac{E_{\tau_{\text{vis}}} \cdot \vec{p}_{\tau_{\text{vis}}}}{x_{\text{vis}}}, \frac{\vec{p}_{\tau_{\text{vis}}}}{x_{\text{vis}}}\right)$$

Since the mass of the parent particle is much greater than the products, we can neglect the product masses giving the collinear mass as a function only of the visible fraction

$$M_{\text{coll}} = \frac{M_{\text{vis}}}{\sqrt{x_{\text{vis}}}}$$
Unlike the visible mass, $M_{coll}$ is approximately equal to the mass of the Higgs boson. Additionally, it is a noticeably narrower peak which is desirable. A comparison of $M_{vis}$ and $M_{coll}$ is shown in Figure 5.1.

![Figure 5.1. Comparison of visible mass and collinear mass for simulated $H \rightarrow \mu\tau_{had}$ signal events after preselection in VBF category](image)

The analysis is performed blind, meaning that the development of the methods and selections are done without knowledge of the observed data in the signal region. The signal region is the area of the signal variable where signal is expected to appear: $100 \text{ GeV} \leq M_{coll} \leq 150 \text{ GeV}$. 
5.2 Statistical Methods

The main goal of this analysis is to establish whether or not the signal process \( H \rightarrow \mu \tau \) exists. An excess of events beyond what is expected from background contributions is indicative of possible existence of the signal process. Since particle interactions are quantum mechanical in nature and thus have a degree of randomness, in addition to experimental uncertainties, it is possible that the process exists but an excess, if any, is not observed. If an excess is observed, the amount of signal is quantified by the best fit of the branching fraction \( B(H \rightarrow \mu \tau) \). Additionally, the \( p \)-value, which represents the likelihood the excess is due to fluctuations, is computed. If no excess is observed, upper limits are set on the branching fraction at 95% confidence level (CL).

In general, a confidence level quantifies how compatible a set of data is with a hypothesis. By convention, a 95% CL is taken as a requirement for ruling out a signal at or above a certain value known as an exclusion limit, but this is an interpretation of the compatibility not an exact consequence. The incompatibility with the signal plus background hypothesis \((CL_{s+b})\) has been used for exclusion limits in the past but is unsuitable when the number of signal events is small compared to expected background. Being incompatible with the background only hypothesis as quantified by \( CL_b \) is evidence for something besides expected background but is insufficient for testing signal hypotheses. The prescribed method for computing confidence levels at CMS is via the \( CL_s \) method which utilizes both of these concepts [23].

The following sections describe how limits are computed using this method. The parameter \( \mu \) appears as a signal strength modifier. It exists due to wanting to set limits with predicted cross sections and branching fractions as set values across different decay modes. As there is no predicted branching fractions for LFV decays of the Higgs boson, the signal strength becomes synonymous with the branching fraction
and no combination of different decay modes of the Higgs occurs. Event yields for signal and background are denoted by $s$ and $b$ respectively.

5.2.1 Systematic Uncertainties

As covered in Chapter[7] numerous sources of uncertainties are taken into account when performing a measurement. In order for the likelihood functions used in the $CL_s$ method to have a useful, factorized form, uncertainties must be 100% correlated, anti-correlated or completely uncorrelated. If an uncertainty is partially correlated it must be separated into correlated and uncorrelated components or estimated as one or the other, whichever is more conservative. Quantities which have uncertainties are treated as a set of nuisance parameters denoted by $\theta$. Because of the uncertainties in these quantities, each component of the set $\theta$ has a distribution with a measured default value $\tilde{\theta}$. As a probability distribution function (pdf), this is denoted as $\rho(\theta|\tilde{\theta})$. If instead, we had a pdf of $\tilde{\theta}$, the set of uncertainties can be used in a fit to constrain the measurement of interest[24]. To do so, Bayes’ theorem is used to relate the two different pdfs:

$$\rho(\theta|\tilde{\theta}) \sim p(\tilde{\theta} | \theta) \cdot \pi_\theta(\theta) \quad (5.5)$$

where the function $\pi_\theta(\theta)$ is a prior taken to be a flat distribution representing no prior knowledge of distribution of $\theta$. The set of nuisance parameters enters the likelihood functions by being a parameter of the signal and background event yields which are now denoted as $s(\theta)$ and $b(\theta)$ respectively.

5.2.2 Observed Limits

To calculate $CL_s$ for a given signal strength $\mu$, a likelihood function $L(data|\mu, \theta)$ must be constructed. We used a binned likelihood in which each bin of the signal variable follows Poisson statistics. The likelihood function is then simply the product
over bins
\[
\mathcal{L}(\text{data}|\mu, \theta) = p(\tilde{\theta}|\theta) \cdot \prod_i \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)}
\] (5.6)

where \( n_i \) is the observed number of events in the \( i \)th bin, \( s_i \) and \( b_i \) are the signal and background event yields of the \( i \)th bin and understood to still be functions of \( \theta \), and \( p(\tilde{\theta}|\theta) \) is the pdf from Equation 5.5. The yields \( s_i \) and \( b_i \) are obtained by a simultaneous fit to the data.

A test statistic \( \tilde{q}_\mu \) is formed as the log ratio of the likelihood function over its maximum:
\[
\tilde{q}_\mu = -2 \ln \frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})}, \quad \text{where: } 0 \leq \hat{\mu} \leq \mu
\] (5.7)

where \( \hat{\mu} \) and \( \hat{\theta} \) are the values of \( \mu \) and \( \theta \) which maximize the likelihood and where \( \hat{\theta}_\mu \) maximizes the likelihood given signal strength \( \mu \). Because signal events can only be in addition to background and never somehow decrease the amount of background, the signal strengths are required to be greater than zero. The condition \( \hat{\mu} \leq \mu \) is enforced to ensure one-sided limits such that a particular signal strength is not excluded because the observed event yield is too high [24].

Observed data is used to evaluate \( \tilde{q}_\mu^{\text{obs}} \), \( \hat{\theta}_\mu^{\text{obs}} \) for signal plus background hypothesis and \( \hat{\theta}_0^{\text{obs}} \) for the background only hypothesis i.e. \( \mu = 0 \). The later two are used to generate toy Monte Carlo pseudo-data. These generated datasets are each used to calculate a \( \tilde{q}_\mu \) thus creating the pdfs \( f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{\text{obs}}) \) and \( f(\tilde{q}_\mu|0, \hat{\theta}_0^{\text{obs}}) \). Using these pdfs, the p-values for signal+background hypothesis (\( p_\mu \)) and background only hypothesis (\( p_b \)) are formed:

\[
p_\mu = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}}|\text{signal+background}) = \int_{\tilde{q}_\mu^{\text{obs}}}^\infty f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{\text{obs}}) d\tilde{q}_\mu,
\]
\[
1 - p_b = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}}|\text{background only}) = \int_{\tilde{q}_0^{\text{obs}}}^\infty f(\tilde{q}_\mu|0, \hat{\theta}_0^{\text{obs}}) d\tilde{q}_\mu
\]

The p-values \( p_\mu \) and \( p_b \) as confidence levels would be \( CL_{s+b} \) and \( CL_b \) respectively but
we define

\[ CL_s(\mu) = \frac{p_\mu}{1 - p_b} \]  

(5.8)

Finally \( \mu \) is iterated over until \( CL_s = 0.05 \) and it is said that the amount of signal given by that \( \mu \) (or the corresponding branching fraction in the LFV case) is excluded at 95% confidence level. This value of \( \mu \) is denoted as \( \mu^{95\%} \) [24].

5.2.3 Expected Limits

Expected limits are the exclusion limit expected for the background only hypothesis independent of observed data. This is an important quantity because it is calculated while the analysis is still blinded in the signal region of the signal variable and used to gauge the sensitivity of the measurement without experimenter bias. These limits are calculated by generating many pseudo-datasets with the background only hypothesis to use in place of observed data. Because there are many datasets, a \( \mu^{95\%} \) can be calculated for each forming a distribution. The median expected value and its \( \pm 1\sigma \) and \( \pm 2\sigma \) bands are then calculated by integrating over this distribution until the appropriate quantiles are reached [24].
CHAPTER 6

BACKGROUND ESTIMATION

The strategy for this analysis is to estimate each background individually, preferring data-driven techniques for the more significant backgrounds while Monte Carlo simulations are used for the less significant backgrounds. The largest backgrounds are $Z \to \tau\tau$ and those from misidentified hadronic tau decays which come mainly from $W + jets$ and QCD multijet events.

6.1 $Z \to \tau\tau$

The process $Z \to \tau\tau$ becomes a significant background when one tau decays to a muon plus neutrinos and the other tau decays hadronically. The particle flow embedding technique is used to estimate this background [25]. This technique uses a sample of $Z \to \mu\mu$ selected from collision data. The muons in the data are replaced with simulated taus and their decays while the rest of the event is preserved. This is an improvement over complete simulation of the event because characteristics of the event such as topology, extra jets, missing energy, and the underlying event are taken directly from data and thus more accurate. This gives the shape of the background distribution while the normalization is obtained from Monte Carlo expectation.

To validate this method we look to a control region with enhanced $Z \to \tau\tau$ background formed by adding the criteria $\Delta_{\mu-\tau} R < 0.2$ to the baseline selection. The quantity $\Delta_{\mu-\tau} R$ is the distance between two objects in $\eta\phi$ space given by:

$$\Delta R_{\mu-\tau} = \sqrt{(\phi_{\mu} - \phi_{\tau})^2 + (\eta_{\mu} - \eta_{\tau})^2} \quad (6.1)$$
Requiring that the mu and the tau be close together enhances $Z \rightarrow \tau\tau$ background contribution because any momentum the Z boson has, will result in its decay products being closer together. Other significant background sources do not have both leptons coming from the same parent particle and thus are not enhanced by this condition. Figure 6.1 shows good agreement in this region indicating the method is valid.

Figure 6.1: Collinear mass for the $Z \rightarrow \tau\tau$ control region

6.2 Misidentified Lepton

One of the most dominant backgrounds for $H \rightarrow \mu\tau_{had}$ is that which arises from misidentified leptons. Our selection criteria of high momentum and tight identificatio-
tion for the muon, coupled with their clean signatures means that, in this analysis, the muon is rarely a misidentified lepton. However, hadronically decaying taus are easier to misidentify because they appear in the detector as a jet of particles. Many processes can produce jets which may fit the criteria of our hadronically decaying taus when in reality they are not. Therefore, the misidentified lepton background consists of events in which there was a real muon and an object which was not a hadronically decaying tau but that was misidentified as one. The majority of our misidentified lepton background comes from the $W + jets$ process where the $W$ decays muonically and the jet is misidentified as a hadronically decaying tau. This can be seen with Monte Carlo simulated data sets of $W + jets$. Other sources also contribute to the misidentified lepton background including $Z \rightarrow \mu\mu + jet$ in which one of the muons does not meet our veto criteria for additional muons and a jet is misidentified as a hadronically decaying tau or $t\bar{t} + jets$ events where one of the $W$ bosons from a top decay decays muonically while a jet is misidentified as a hadronic tau decay. Another source of this background is QCD jet events, simulations for which were not accurate enough for the purpose of estimating this background. We determined that a data driven technique would be best for estimating this background.

In order to get an accurate data driven estimate of the misidentified leptons background, we use what is known as the fake rate method \cite{26}. This method uses an independent sample to measure the rate at which objects which are not the particle of interest are misidentified as such. In this analysis we are concerned only with the rate at which hadronic tau decays are misidentified. To measure this rate, we need a data sample in which we can pick out objects that are identified as hadronic tau decays but that we are reasonably certain are not, in fact, hadronic tau decays. This is done utilizing a sample of $Z \rightarrow \mu\mu + X$ where $X$ is the object that might be misidentified as a hadronic tau decay. The object $X$ is required to pass loose tau criteria so that it is possible that $X$ has a chance to pass the tighter selection criteria.
for the taus of the analysis. Additionally, both muons are required to pass tight criteria and the two are required to have a invariant mass near the Z boson mass to give a reasonably pure sample of $Z \rightarrow \mu\mu + X$. Since there is no such process $Z \rightarrow \mu\mu\tau$, if X passes the analysis criteria for a hadronic tau decay, we can be reasonably sure it was misidentified. Thus, the fraction of X which pass the analysis criteria, $f$, represents the probability that any object which passes loose tau criteria will be misidentified as a hadronic tau decay as defined in the analysis:

$$ f = \frac{N_{\text{events}}(Z \rightarrow \mu\mu + X, X = \text{tight } \tau \text{ isolation})}{N_{\text{events}}(Z \rightarrow \mu\mu + X, X = \text{loose } \tau \text{ isolation})} \quad (6.2) $$

The rate at which hadronic tau decays are misidentified is not necessarily constant with respect to transverse momentum or eta. This rate can be binned in these variables and be used to estimate backgrounds in those bins. However, this divides the number of events used to measure the rate amongst the bins increasing the error on the rate measurements. As can be seen in Figure 6.2, the rate for this analysis is consistent with being constant with respect to these variables so we do not apply the rate as functions of these variables.

Knowing the misidentification rate as measured in $Z \rightarrow \mu\mu + X$, it can then be applied to the data set used to look for a signal. By taking the region of data that fulfills all analysis criteria except has loose tau isolation instead of tight and applying a factor based on the fake rate, we estimate the number of events in which the tau will pass the tight criteria through being misidentified.
Figure 6.2. Fake rate $f$ vs the jet $p_T$ (Left) and vs $\tau\eta$ (Right)

To ensure we are estimating the background correctly with this method, we use a control region of same sign muons and taus. This region is chosen because it is heavily dominated by fake taus since no other process produces these two leptons in same sign while other leptons are being vetoed. The control region shows good agreement of the estimated misidentified lepton background to the data as seen in Figure 6.3 indicating that the background estimation method is accurate. The misidentification rate is

\[
    f = \frac{N_{\text{tight}}}{N_{\text{loose}}} \\
    1 - f = \frac{N_{\text{NOT-tight}}}{N_{\text{loose}}} \\
    N_{\text{loose}} = \frac{N_{\text{NOT-tight}}}{1 - f} \\
    f = \frac{N_{\text{tight}}(1 - f)}{N_{\text{NOT-tight}}} \\
    N_{\text{tight}} = N_{\text{NOT-tight}} \cdot \frac{f}{1 - f}
\]
assumed to be the same regardless of the process which produced the misidentified object. To test this, we compare our fake rate measured from the electroweak process \( Z \rightarrow \mu\mu + X \) to one measured in a sample with enhanced QCD events. To do this we change the analysis selection criteria to require an anti-isolated muon, while the rest of the selection remains the same. This gives a measured rate from predominantly QCD processes which we further split into opposite sign and same sign categories depending on whether the muon and the tau-like object match in electric charge. All three rate measurements were found to be close as in Figure 6.4 with any slight differences being covered by recommended uncertainties. Since there are multiple different hadronic decay modes for a tau, the misidentification rate was also measured vs the decay mode. Figure 6.4 shows that the different modes are consistent with one another.

![Figure 6.3. Like sign control region showing good agreement between estimated misidentified lepton background and data](image)

Figure 6.3. Like sign control region showing good agreement between estimated misidentified lepton background and data
Figure 6.4. Comparison studies for fake rates measured from $Z \rightarrow \mu\mu + X$ to QCD enhanced opposite sign and QCD enhanced same sign vs $p_T$ (left) and tau decay mode (right)

Although the misidentified leptons background satisfactorily models the process as evidenced by the control regions and background fits to data, additional effort was put forth to try to increase the accuracy of this prediction and more importantly find evidence that large systematic error associated with this background could be reduced while remaining conservative. At the heart of the fake rate method is the definition of a fake. This boils down to the ratio formed to define the rate. The method described above is concerned with an object which passes loose tau criteria faking a tau as defined by the tight criteria. The definitions of loose and tight here are not immutable. In general, these two working points are somewhat arbitrary and are simply chosen to represent something which vaguely resembles a tau in the loose case, and something we feel certain enough about to label a tau in the tight case. However, several working points are available. In order of increasingly stringent re-
requirements these are very loose, loose, medium, tight, very tight, and very very tight. A rate can be defined as the ratio of any working point over one that is looser. This raises the question if the typical tight over loose gives the most accurate background prediction. To investigate this question, fake rates of the various combinations of working points were calculated and the agreement in the same sign control region compared. Figure 6.5 shows such a comparison, keeping the numerator of the ratio at very tight and varying the denominator. In general, more significant changes occur when changing the denominator as changing the numerator typically just scales distribution up or down. Discrepancies between data and the modeled background are clear from the accompanying ratio plots. Looser denominators introduce a leftward shift in the distribution. The ratio which produces a background that best fits the data is the one whose ratio plot lies most nearly on 0 and is flat. It should be noted that the working points used here were part of a redefinition not used in the analysis outside of this study so that the “tight” and “loose” seen here do not correspond with the working points of the same name used outside of this study. Ultimately it was determined that some improvement in agreement could be gained by changing the fake rate definition but not significant enough to alter the publication which was already in the review process. However, a similar study could be useful in upcoming future analyses of this channel.

6.3 Other Backgrounds

Other, less significant backgrounds are modeled using Monte Carlo data for both shape and normalization. One such background is $Z + jets$ not including $Z \rightarrow \tau \tau$ which is estimated separately. $t \bar{t} + jets$ enters here by the tops decaying to W bosons which then may decay to leptons. Similarly, $WW + jets$ provides a small background contribution as well as single top production. Standard Model Higgs decays also provide a small background primarily through the $H \rightarrow \tau \tau$ channel where one tau
Figure 6.5. Same sign control region for different definitions of fake rates. Very tight over tight (top left), very tight over medium (top right), very tight over loose (bottom left), and very tight over very loose (bottom right)
decays to a muon and the other hadronically. There is, however, significant kinematic
differences between the LFV signal and the Standard Model decay due to difference
in the number and distribution of neutrinos.

6.4 Preselection Control Plots

The collinear mass distributions for each category are shown in Figure 6.6. Here
the total estimated background distribution can be compared to data and seen to
have good agreement. This agreement is a measure of validity of our background
estimation. The branching ratio of $H \rightarrow \mu \tau$ is set to 1 here so that the signal may be
visible. Table 6.1 shows the event yields in the signal region $100 \text{ GeV} < M_{coll} < 150$
GeV normalized to an integrated luminosity of 19.7 fb$^{-1}$ and LFV branching fraction
set to 10%.
Figure 6.6. Collinear mass after preselection requirements for 0-Jet, 1-jet, and 2-jet
## TABLE 6.1

**EVENT YIELDS IN SIGNAL REGION AT PRESELECTION**

<table>
<thead>
<tr>
<th>Category</th>
<th>0-jet</th>
<th>1-jet</th>
<th>2-jet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fakes</td>
<td>$21734 \pm 6531$</td>
<td>$9532 \pm 2864$</td>
<td>$6218 \pm 1867$</td>
</tr>
<tr>
<td>$Z \rightarrow \tau\tau$</td>
<td>$14768 \pm 445$</td>
<td>$3749 \pm 115$</td>
<td>$1773 \pm 55$</td>
</tr>
<tr>
<td>$ZZ,WW$</td>
<td>$976 \pm 147$</td>
<td>$459 \pm 69$</td>
<td>$162 \pm 25$</td>
</tr>
<tr>
<td>$Z \rightarrow ee$ or $\mu\mu$</td>
<td>$1564 \pm 182$</td>
<td>$124 \pm 31$</td>
<td>$63 \pm 11$</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>$92 \pm 9.8$</td>
<td>$847 \pm 85$</td>
<td>$3528 \pm 354$</td>
</tr>
<tr>
<td>$t,\bar{t}$</td>
<td>$95 \pm 12$</td>
<td>$473 \pm 50.$</td>
<td>$505 \pm 53$</td>
</tr>
<tr>
<td>SM Higgs</td>
<td>$143 \pm 15$</td>
<td>$100. \pm 7.8$</td>
<td>$29 \pm 5.9$</td>
</tr>
<tr>
<td>Sum of Bg.</td>
<td>$39373 \pm 6551$</td>
<td>$15283 \pm 2869$</td>
<td>$12278 \pm 1902$</td>
</tr>
<tr>
<td>LFV Higgs</td>
<td>$1913 \pm 256$</td>
<td>$1075 \pm 134$</td>
<td>$585 \pm 128$</td>
</tr>
<tr>
<td>Data</td>
<td>$38376 \pm 196$</td>
<td>$14687 \pm 121$</td>
<td>$11942 \pm 109$</td>
</tr>
</tbody>
</table>
Numerous sources of systematic uncertainties must be considered to make a final measurement. These sources of errors come in two varieties; those which only affect the scale of the background distributions and those which change the shape of the background distribution in addition to the scale.

7.1 Normalization Uncertainties

Systematic uncertainties which only affect the scale of a background distribution are referred to as normalization uncertainties. A list of normalization uncertainties is provided in Table 7.1. These are independent of the signal variable $M_{\text{coll}}$ and treated as correlated between categories except where uncorrelated portions are marked by an asterisk. For example, backgrounds which are modeled using Monte Carlo simulated data use an experimentally measured value for their production cross section to determine how many such background events would occur in an amount of data equal to that which is recorded. This measurement, like all measured quantities, comes with an associated uncertainty. Since the cross section determines the rate at which the process will produce background events, it only affects the normalization or scale of the background distribution in the signal variable.

The efficiency for muon and hadronic taus to pass trigger, identification, and isolation criteria is studied via the tag and probe method as described in Section 4.1. These efficiencies come with uncertainties which are determined through the same method. For muons this is done via tag and probe in observed $Z \rightarrow \mu\mu$ data [27, 28].
Similarly for hadronically decaying taus, this is done with the tag and probe method in observed $Z \rightarrow \tau\tau$ data [22].

Background distributions are normalized to the amount of observed data that is recorded, known as the integrated luminosity. The luminosity is determined via the pixel cluster counting method which utilizes zero-bias events and counts clusters in the pixel tracker which are proportional to the number of collisions in the event. The method is calibrated with van der Meer scans which are a process of scanning the proton beams past each other in the transverse plane to determine the overlap of the colliding beams [29]. The uncertainty in the luminosity measurement is predominately from fitting of the proton bunch shapes. To estimate this uncertainty, the fits from pixel cluster counting method are compared to another method of measuring the luminosity. Energy deposited in the forward hadron calorimeter is also proportional to the number of collisions in an event and thus can be used for comparison [30].

The uncertainty for the misidentified lepton background is estimated as the maximum difference between misidentification rates using data samples that enhance different sources such as electroweak versus QCD processes [26]. Rates are also calculated in simulated datasets for comparison. The misidentification rates shown in Figures 6.2 and 6.4 show this variance and very conservative estimate for the uncertainty is used that would encompass all these measured rates.

The uncertainty for the $Z \rightarrow \tau\tau$ background comes in two parts. The first is from the cross section uncertainty as described for other backgrounds above. The second is a portion that is uncorrelated between categories that arises from biases in $\tau_{had}$ efficiency when embedding simulated tau leptons in place of identified muons [31].

In addition to the experimental normalization uncertainties described above, there are also theoretical uncertainties associated with the Higgs boson production cross section. These uncertainties vary for different production mechanisms for the Higgs
# TABLE 7.1

**SYSTEMATIC UNCERTAINTIES INDEPENDENT OF $M_{\text{coll}}$**

<table>
<thead>
<tr>
<th>Uncertainty</th>
<th>0-jet</th>
<th>1-jet</th>
<th>2-jet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muon Trigger/ID/Isolation</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>Hadronic Tau efficiency</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>Luminosity</td>
<td>2.6%</td>
<td>2.6%</td>
<td>2.6%</td>
</tr>
<tr>
<td>$Z \rightarrow \tau\tau$</td>
<td>3+5*%</td>
<td>3+5*%</td>
<td>3+10*%</td>
</tr>
<tr>
<td>$Z \rightarrow \mu\mu, ee$</td>
<td>30%</td>
<td>30%</td>
<td>30%</td>
</tr>
<tr>
<td>misidentified leptons</td>
<td>30+10*%</td>
<td>30%</td>
<td>30%</td>
</tr>
<tr>
<td>$WW, ZZ + jets$</td>
<td>15%</td>
<td>15%</td>
<td>65%</td>
</tr>
<tr>
<td>$t\bar{t} + jets$</td>
<td>10%</td>
<td>10%</td>
<td>10+33*%</td>
</tr>
<tr>
<td>Single Top</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
</tbody>
</table>
as well as the jet category. They affect signal events and SM Higgs events equally and are treated as correlated between the two sources. A list of theoretical uncertainties is given in Table 7.2. Theory uncertainties independent of $M_{coll}$ are treated as fully correlated. These are also treated as fully correlated between categories except where negative values indicate fully anti-correlated quantities caused by category migration.

One theoretical uncertainty comes from the parton distribution function (PDF) which describes the probability density for a parton to have a given fraction $x$ of a proton’s momentum. This is important at a proton collider because the energy available to a process is not the collision energy of the protons but rather the energy of the interacting partons. Since QCD is non-perturbative, PDFs are not calculated but determined experimentally. The use of PDFs and the calculation of their uncertainties in the LFV analysis follows the prescribed method from the PDF4LHC group [32]. This method uses the three independent PDF sets CT10, MSTW, and NNPDF over a range of parameters [33]. The envelope of the three sets and their respective uncertainties is taken as the 1σ band to determine the uncertainty value for our analysis.

The effects of varying the renormalization scale $\mu_R$ and factorization scale $\mu_F$ are also taken as a theoretical uncertainty. These scales have a default value of $\mu_R = \mu_F = m_H/2$ and are varied up by a factor of two and down by a factor of 1/2 but constrained by the condition $0.5 < \mu_F/\mu_R < 2$ [33]. Under these variations, half the difference between the maximum and minimum simulated yields as a percent of the default scale is taken as the uncertainty.

When partons are emitted in a collision they will shower, giving rise to a jet. Calculating the showering of partons for simulated events is inexact and depends on many parameters used by the chosen model. At CMS, the simulation software PYTHIA is used and the parameters used for modeling parton showers are tuned such that various simulated processes match with what is observed [33]. This set of
tuned parameters is known as the \( Z2^* \) tune. The ATLAS experiment goes through a similar tuning process and has its own set of tuned parameters known as AUET2. There is no physical reason why partons would shower differently at ATLAS versus CMS so the difference in yields as a percent of the default CMS tuned value is used as a theoretical uncertainty due to parton shower modeling.

### TABLE 7.2

THEORETICAL UNCERTAINTIES INDEPENDENT OF \( M_{\text{coll}} \)

<table>
<thead>
<tr>
<th>Uncertainty</th>
<th>Gluon Fusion</th>
<th>Vector Boson Fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-jet</td>
<td>1-jet</td>
</tr>
<tr>
<td>Parton Density Function</td>
<td>+9.7%</td>
<td>+9.7%</td>
</tr>
<tr>
<td>Renormalization Scale</td>
<td>+8%</td>
<td>+10%</td>
</tr>
<tr>
<td>Parton Shower</td>
<td>+4%</td>
<td>-5%</td>
</tr>
</tbody>
</table>

7.2 Shape Uncertainties

Uncertainties which depend on the value of the signal variable \( M_{\text{coll}} \) affect not only the normalization but the shape of the background distribution in this variable as well. These are known as shape uncertainties and are listed in Table 7.3.

Corrections to a jets energy are described in Section 4.4.1 and come with associated uncertainties. The total uncertainty in the jet energy scale has many contributing sources which are each estimated individually. These include varying the \( \rho \) value.
and jet energy resolution as well as considering photon energy resolution and effects from parton showering models and many other factors [18]. To determine the effect of the jet energy scale uncertainty on the shape of distributions, all jets are varied both up and down by one standard deviation. The effects of scaling the jets up and down are then propagated separately through the various parts of the analysis until the post-selection distributions are obtained. These shifted distributions are then used as template by which distributions are allowed to change shape during fitting.

A similar process is used for two other energy scales. The tau energy scale is determined through comparing observed $Z \rightarrow \tau\tau$ data to simulated data. An uncertainty is assigned based on the agreement of the two. Jets below 10 GeV and Particle Flow candidates that do not get clustered are referred to as unclustered energy and have a similar uncertainty. Similar to the jet energy scale corrections these get propagated through the analysis to form additional shape templates. Examples of the effects of shifting the jet energy and tau energy scales for signal simulation compared to their unshifted shapes are shown in Figure 7.1.

Possible shape distortions to the misidentified lepton background are also considered. As shown in Figure 6.2 the fake rate vs $p_T$ is relatively flat. However, any dependence would affect the shape of the fakes background. Therefore, the fake rate vs $p_T$ is fit with a linear function to obtain the uncertainty of the slope. The slope is then varied up and down one standard deviation, and propagated through the analysis to produce shape templates for the misidentified lepton background.

Finally, the capacity for statistical uncertainties to change the shape of distributions is considered. It is possible that statistical fluctuations in separate bins of the signal variable could compound to affect the shape of distributions. This is deemed a bin-by-bin uncertainty and is accounted for by creating shape templates by simply varying each bin in a distribution up and down by one standard deviation.
Figure 7.1. Shifted distributions for jet energy scale (left) and tau energy scale (right) for the 0 jet category. Inset plots are ratios of shifted to unshifted.

### TABLE 7.3

SYSTEMATIC UNCERTAINTIES DEPENDENT ON $M_{\text{coll}}$

<table>
<thead>
<tr>
<th>Uncertainty</th>
<th>$H \rightarrow \mu\tau_{\text{had}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hadronic Tau Energy Scale</td>
<td>3%</td>
</tr>
<tr>
<td>Jet Energy Scale</td>
<td>3-7%</td>
</tr>
<tr>
<td>Unclustered Energy Scale</td>
<td>10%</td>
</tr>
</tbody>
</table>
CHAPTER 8

RESULTS

8.1 Limit Setting and Best Fit Branching Fractions

Event yields for each category after the final selection criteria as listed in 4.5.3 have been applied are shown in Table 8.1. Plots of the signal variable $M_{coll}$ after a simultaneous fit of all backgrounds and signal to the observed data, are shown in Figure 8.1. Following the limit setting procedure described in Chapter 5, 95% CL upper limits are set on the branching fraction $B(H \rightarrow \mu \tau_{had})$ for each category. Expected and observed limits are listed in Table 8.2. These limits are combined with limits for $B(H \rightarrow \mu \tau_{e})$ measured in a very similar analysis preformed by the same analysis group [34] to set limits on $B(H \rightarrow \mu \tau)$ independent of the tau decay mode. Note that the possible final state $H \rightarrow \mu \tau_{\mu}$ is not measured due to its lack of sensitivity to the LFV process caused by its similarity to lepton flavor conserving processes. Given that a slight excess is observed in the combined measurement, a best fit value of the branching fraction is also shown in Table 8.2.

8.2 Limits on Yukawa Couplings

Theoretical models often describe an interaction in terms of the Yukawa couplings which represent the strength of an interaction given by a term in the Lagrangian. Therefore, it is useful to transform the limits on the branching fraction $B(H \rightarrow \mu \tau)$ into limits on the Yukawa couplings $Y_{\mu \tau}, Y_{\tau \mu}$. The branching fraction of a particular
Figure 8.1. Collinear mass after full selection requirements for 0-Jet, 1-jet, and 2-jet after fitting to observed data.
### TABLE 8.1

**EVENT YIELDS IN THE SIGNAL REGION AFTER FULL SELECTION**

<table>
<thead>
<tr>
<th>Category</th>
<th>0-jet</th>
<th>1-jet</th>
<th>2-jet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fakes</td>
<td>1858.1 ± 558.8</td>
<td>362.9 ± 110.0</td>
<td>0.5 ± 0.5</td>
</tr>
<tr>
<td>$Z \rightarrow \tau\tau$</td>
<td>198.8 ± 11.0</td>
<td>50.5 ± 3.5</td>
<td>0.4 ± 0.2</td>
</tr>
<tr>
<td>$ZZ, WW$</td>
<td>47.0 ± 8.0</td>
<td>14.6 ± 2.6</td>
<td>0.3 ± 0.2</td>
</tr>
<tr>
<td>$Z \rightarrow ee$ or $\mu\mu$</td>
<td>94.5 ± 25.2</td>
<td>17.6 ± 6.7</td>
<td>0.1 ± 0.1</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>2.5 ± 0.6</td>
<td>24.3 ± 3.2</td>
<td>0.7 ± 0.3</td>
</tr>
<tr>
<td>$t, \bar{t}$</td>
<td>2.7 ± 1.2</td>
<td>19.9 ± 3.9</td>
<td>0.4 ± 0.5</td>
</tr>
<tr>
<td>SM Higgs Background</td>
<td>7.0 ± 1.3</td>
<td>4.9 ± 0.7</td>
<td>1.9 ± 0.7</td>
</tr>
<tr>
<td>Sum of Backgrounds</td>
<td>2210.4 ± 559.6</td>
<td>494.7 ± 110.4</td>
<td>4.3 ± 1.1</td>
</tr>
<tr>
<td>LFV Higgs Signal</td>
<td>69.7 ± 17.0</td>
<td>29.7 ± 6.7</td>
<td>3.0 ± 1.0</td>
</tr>
<tr>
<td>Data</td>
<td>2255.0 ± 47.5</td>
<td>506.0 ± 22.5</td>
<td>8.0 ± 2.8</td>
</tr>
</tbody>
</table>
## TABLE 8.2

**EXPECTED UPPER LIMITS, OBSERVED UPPER LIMITS, AND BEST FITS FOR BRANCHING FRACTIONS**

<table>
<thead>
<tr>
<th>Expected Limits</th>
<th>0-Jet (%)</th>
<th>1-Jet (%)</th>
<th>2-Jets (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu\tau_{\text{had}}$</td>
<td>$&lt;2.34 \pm 1.19$</td>
<td>$&lt;2.07 \pm 1.06$</td>
<td>$&lt;2.31 \pm 1.18$</td>
</tr>
<tr>
<td>$\mu\tau_{e}$</td>
<td>$&lt;1.32 \pm 0.67$</td>
<td>$&lt;1.66 \pm 0.85$</td>
<td>$&lt;3.77 \pm 1.92$</td>
</tr>
<tr>
<td>$\mu\tau$</td>
<td>$&lt;0.75 \pm 0.38$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observed Limits</th>
<th>0-Jet (%)</th>
<th>1-Jet (%)</th>
<th>2-Jets (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu\tau_{\text{had}}$</td>
<td>$&lt;2.61$</td>
<td>$&lt;2.22$</td>
<td>$&lt;3.68$</td>
</tr>
<tr>
<td>$\mu\tau_{e}$</td>
<td>$&lt;2.04$</td>
<td>$&lt;2.38$</td>
<td>$&lt;3.84$</td>
</tr>
<tr>
<td>$\mu\tau$</td>
<td>$&lt;1.51$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Best Fit Branching Fractions</th>
<th>0-Jet (%)</th>
<th>1-Jet (%)</th>
<th>2-Jets (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu\tau_{\text{had}}$</td>
<td>$0.41_{-1.22}^{+1.20}$</td>
<td>$0.21_{-1.09}^{+1.03}$</td>
<td>$1.48_{-0.93}^{+1.16}$</td>
</tr>
<tr>
<td>$\mu\tau_{e}$</td>
<td>$0.87_{-0.62}^{+0.66}$</td>
<td>$0.81_{-0.78}^{+0.85}$</td>
<td>$0.05_{-0.97}^{+1.58}$</td>
</tr>
<tr>
<td>$\mu\tau$</td>
<td>$0.84_{-0.37}^{+0.39}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
decay mode is given by the ratio of the decay width for that mode over the total decay width of the parent. Assuming that $H \rightarrow \mu \tau$ is the only non-SM process, its branching fraction is given by:

$$B(H \rightarrow \mu \tau) = \frac{\Gamma(H \rightarrow \mu \tau)}{\Gamma(H \rightarrow \mu \tau) + \Gamma_{SM}}$$  \hspace{1cm} (8.1)$$

Where $\Gamma_{SM}$ is the SM Higgs decay width and is has a value of 4.1 MeV for a Higgs mass ($m_h$) of 125 GeV. Following a prescription that turns the Yukawa couplings of leptons to the Higgs field into matrices that are not necessarily diagonal with respect to lepton flavor $\mathbb{H}$, the decay width of the LFV process can be described in terms of the relevant couplings:

$$\Gamma(H \rightarrow \mu \tau) = \frac{m_h}{8\pi} \sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2}$$  \hspace{1cm} (8.2)$$

Substituting this expression into the above definition of the branching fraction and rearranging gives an expression of the two couplings in terms of the branching fraction $B$ and $m_h$:

$$\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} = \sqrt{\frac{8\pi \cdot B}{m_h(1 - B)}}$$  \hspace{1cm} (8.3)$$

For a limit on the branching fraction of $B(H \rightarrow \mu \tau) < 1.51\%$, the corresponding limit on the Yukawa couplings is:

$$\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 3.6 \times 10^{-3}$$  \hspace{1cm} (8.4)$$

Yukawa couplings for SM decays of the Higgs boson follow the relationship:

$$|Y_{\ell\ell}|^2 = \frac{m_{\ell}^2}{v^2}$$  \hspace{1cm} (8.5)$$
where $\ell$ represents a flavor of lepton, $m_\ell$ the lepton mass, and $v$ the vacuum expectation value. It is considered more natural if the Yukawa couplings for an LFV process are relatively smaller than those of the SM processes. This means that a good benchmark for sensitivity for this analysis is to probe couplings that fit the condition:

$$|Y_{\mu\tau}Y_{\tau\mu}| < \frac{m_\mu m_\tau}{v^2}$$

(8.6)

Figure 8.2 shows the observed limit reaching into this region but room for improvement still exists. Additionally it shows that previous limits do not significantly reach into this region.

8.3 Conclusions

The presented analysis sets limits on $B(H \to \mu\tau_{had})$ and in combination with limits on $B(H \to \mu\tau_e)$ is used to set the first ever limits on the branching fraction and Yukawa couplings of the LFV process $H \to \mu\tau$ by direct measurement. At 95% CL we observe an upper limit of $B(H \to \mu\tau) < 1.51\%$. This limit is approximately an order of magnitude more stringent than previously existing limits from indirect measurements. A slight excess is observed in data that is consistent with a $H \to \mu\tau$ signal of 2.4 $\sigma$ significance corresponding to a $p$-value of 0.010. Although this excess is not significant enough to be considered evidence of lepton flavor violation, it is a hint that such physics beyond the Standard Model may exist and should surely be continued to be investigated as CMS continues to collect additional data at an even higher center of mass energy.
Figure 8.2. Upper limits on Yukawa couplings for the LFV analysis expected (red) with 1 and 2 $\sigma$ error bands (yellow/green bands) and observed limit (black). Also shown are existing limits from tau decay searches (white), a reinterpretation of an LHC $H \rightarrow \tau\tau$ search [1] (yellow line), ranching fraction contours (black dashed) and the discussed “naturalness” condition (orange).
BIBLIOGRAPHY


33. A. Apryan et al. Theoretical uncertainty for the higgs production via vbf and gluon fusion process. CMS Analysis Note 2013/262, 2013.