

This Dissertation  
entitled  
SEARCH FOR LEPTON FLAVOR VIOLATING DECAYS OF  
STANDARD MODEL HIGGS TO A MUON AND A TAU LEPTON  
OR TO AN ELECTRON AND A TAU LEPTON

typeset with  $\text{NDDiss2}_{\mathcal{E}}$  v3.2017.2 (2017/05/09) on November 23, 2020 for

Prasanna Kumar Siddireddy

This L<sup>A</sup>T<sub>E</sub>X 2 <sub>$\mathcal{E}$</sub>  classfile conforms to the University of Notre Dame style guidelines as of Fall 2012. However it is still possible to generate a non-conformant document if the instructions in the class file documentation are not followed!

Be sure to refer to the published Graduate School guidelines at <http://graduateschool.nd.edu> as well. Those guidelines override everything mentioned about formatting in the documentation for this NDDiss2 <sub>$\mathcal{E}$</sub>  class file.

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

**This page is *NOT* part of the dissertation/thesis. It should be disabled before making final, formal submission, but should be included in the version submitted for format check.**

NDDiss2 <sub>$\mathcal{E}$</sub>  documentation can be found at these locations:

<http://graduateschool.nd.edu>  
<https://ctan.org/pkg/nddiss>

SEARCH FOR LEPTON FLAVOR VIOLATING DECAYS OF  
STANDARD MODEL HIGGS TO A MUON AND A TAU LEPTON  
OR TO AN ELECTRON AND A TAU LEPTON

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  
Prasanna Kumar Siddireddy

---

Colin Philip Jessop, Director

Graduate Program in Physics

Notre Dame, Indiana

November 2020

SEARCH FOR LEPTON FLAVOR VIOLATING DECAYS OF  
STANDARD MODEL HIGGS TO A MUON AND A TAU LEPTON  
OR TO AN ELECTRON AND A TAU LEPTON

Abstract

by

Prasanna Kumar Siddireddy

This dissertation presents searches for Lepton Flavor Violating decay of the Standard Model Higgs boson ( $H$ ) into a muon and a tau lepton or to an electron and a tau lepton. Data collected by the CMS detector in 2016-2018, in proton-proton collisions at the LHC, at a center-of-mass energy of 13 TeV was used to perform the search. Observed (expected) upper limits on the branching fraction of  $H$  decaying into a muon and a tau lepton,  $\mathcal{B}(H \rightarrow \mu\tau)$ , was set at 0.15 (0.16) % at 95% CL and branching fraction of  $H$  decaying into an electron and a tau lepton,  $\mathcal{B}(H \rightarrow e\tau)$ , was set at 0.29 (0.19) % at 95% CL. These are the most stringent limits set on these processes till date.

## CONTENTS

Figures . . . . .	v
Tables . . . . .	xi
Chapter 1: Introduction . . . . .	1
Chapter 2: Theoretical Motivation . . . . .	5
2.1 Phenomenological overview . . . . .	5
2.1.1 Quantum electrodynamics . . . . .	7
2.1.2 Electroweak interaction . . . . .	8
2.1.3 Quantum chromodynamics . . . . .	8
2.1.4 The Higgs boson . . . . .	9
2.2 Overview of theoretical building blocks . . . . .	10
2.3 The Standard Model . . . . .	16
2.4 Electroweak symmetry breaking and the Higgs boson . . . . .	18
2.5 Main shortcomings of the Standard Model . . . . .	20
2.6 Lepton flavor violating decays of the Higgs boson . . . . .	20
2.6.1 Introduction of lepton flavor violating decays . . . . .	21
2.6.2 Constraints from low-energy measurements . . . . .	27
2.6.3 Lepton flavor violating decays of the Higgs boson at the LHC	31
Chapter 3: LHC and the CMS detector . . . . .	37
3.1 Introduction . . . . .	37
3.2 The Large Hadron Collider . . . . .	37
3.3 The Compact Muon Solenoid experiment . . . . .	39
3.3.1 The coordinate system of CMS . . . . .	42
3.3.2 Kinematic quantities . . . . .	43
3.3.3 Detector requirements . . . . .	44
3.3.4 Magnet . . . . .	45
3.3.5 Inner Tracking system . . . . .	46
3.3.6 Electromagnetic Calorimeter . . . . .	47
3.3.7 Hadron Calorimeter . . . . .	48
3.3.8 Muon System . . . . .	50
3.3.9 Trigger and data acquisition system . . . . .	52
3.3.10 Luminosity measurement . . . . .	53

Chapter 4: Monte Carlo Event Generation . . . . .	58
4.1 Introduction . . . . .	58
4.2 Perturbative simulation . . . . .	60
4.2.1 Matrix element generators . . . . .	60
4.2.2 Parton shower algorithm . . . . .	61
4.2.3 Matching . . . . .	62
4.3 Hadronisation models . . . . .	64
4.3.1 Fragmentation function . . . . .	64
4.3.2 String model . . . . .	64
4.3.3 Hadron and tau decays . . . . .	65
4.4 Soft hadron-hadron physics modelling . . . . .	66
4.5 Parameter Tuning . . . . .	66
4.6 Monte Carlo generators . . . . .	67
4.7 Detector simulation . . . . .	71
 Chapter 5: Event Reconstruction . . . . .	73
5.1 Particle flow . . . . .	73
5.2 Track and primary vertex reconstruction . . . . .	74
5.2.1 Muon reconstruction . . . . .	76
5.2.2 Electron reconstruction . . . . .	79
5.3 Hadronic tau leptons . . . . .	83
5.3.1 Hadrons Plus Strips . . . . .	83
5.3.2 DeepTau . . . . .	84
5.4 Jet reconstruction . . . . .	87
5.5 Missing transverse energy: $\vec{p}_T^{\text{miss}}$ . . . . .	89
5.6 Relative isolation . . . . .	90
 Chapter 6: Event Selection . . . . .	92
6.1 Introduction . . . . .	92
6.2 Boosted Decision Tree . . . . .	98
6.3 $H \rightarrow \mu\tau_h$ channel . . . . .	100
6.4 $H \rightarrow \mu\tau_e$ channel . . . . .	104
6.5 $H \rightarrow e\tau_h$ channel . . . . .	105
6.6 $H \rightarrow e\tau_\mu$ channel . . . . .	109
 Chapter 7: Background Estimation . . . . .	114
7.1 Introduction . . . . .	114
7.2 Embedding technique . . . . .	116
7.3 Misidentified lepton background . . . . .	118
7.3.1 Fully data-driven approach . . . . .	118
7.3.2 Semi data-driven approach . . . . .	125
7.4 MC Simulation . . . . .	131

Chapter 8: Statistical Methods and Systematic Uncertainties . . . . .	133
8.1 Introduction . . . . .	133
8.2 Statistical methods . . . . .	133
8.3 Systematic Uncertainties . . . . .	138
Chapter 9: Results . . . . .	145
Chapter 10: Conclusion . . . . .	163
Appendix A: SVFit Mass . . . . .	165
A.1 “Classic” SVfit algorithm . . . . .	166
Bibliography . . . . .	169

## FIGURES

2.1	Elementary particles of the Standard Model. . . . .	6
2.2	Fundamental vertices of the Higgs boson. Fermions are denoted $f$ , anti-fermions are denoted $\bar{f}$ , and Gauge bosons of the weak interaction are denoted $V$ . . . . .	10
2.3	Diagrams contributing to the flavor violating decay $\tau \rightarrow \mu\gamma$ , mediated by a Higgs boson with flavor violating Yukawa couplings. . . . .	28
2.4	Constraints on the LFV Yukawa couplings for a 125 GeV Higgs boson. The diagonal Yukawa couplings are approximated by their SM values. Shaded regions show the constraints from low-energy measurements. The thin red dotted lines show the theoretical naturalness limits $ Y_{ji}Y_{ij}  \lesssim \frac{m_i m_j}{v^2}$ . . . . .	33
2.5	Expected and observed 95% CL upper limits by category of a search for LFV $H \rightarrow \mu\tau$ and $H \rightarrow e\tau$ decays with 2016 dataset [18]. . . . .	36
3.1	Cumulative luminosity as a function of time delivered to CMS during stable beams for proton-proton collisions. The luminosity is shown for 2015 (purple), 2016 (orange), 2017 (light blue), and 2018 (dark blue). [30] . . . . .	40
3.2	Schematic view of the CMS detector. . . . .	41
3.3	Coordinate system convention of CMS (left) and the relation between pseudorapidity $\eta$ and polar angle $\theta$ (right) . . . . .	43
3.4	Schematic view of the electromagnetic calorimeter. . . . .	48
3.5	Schematic view of the muon detectors. . . . .	51
4.1	Monte Carlo simulation of an event in proton-proton collisions. . . . .	59
5.1	Track reconstruction efficiencies for single isolated muons as a function of $\eta$ and $p_T$ [60]. . . . .	76
5.2	Efficiency of muon identification as a function of $\eta$ and $p_T$ , for data (black) and simulation (blue) [64]. . . . .	78
5.3	Performance of the BDT-based electron identification algorithm (red dots) compared with results from several working points of cut-based selection for electron candidates in the ECAL barrel (left), and endcaps (right) [34]. . . . .	82

5.4	Performance of tau discrimination against quark and gluon induced jets (left), electrons(middle), and muons(right) for DeepTau and the previously available discriminators . . . . .	86
6.1	Feynman diagrams of lepton-flavor violating Higgs-boson decays. The first row shows diagrams for the Higgs boson coupling to $\mu\tau$ (a,b). Couplings to $e\tau$ (c,d) are shown in the second row. Taus can decay leptonically or hadronically. Feynman diagrams are shown for the leptonic decay of taus (b,d) and the hadronic decay of taus (a,c). . . . .	93
6.2	Illustration of the differences in $d\phi(\ell = \tau \text{ or } e, MET)$ and $p_T^\mu$ spectrums in LFV and SM $H \rightarrow \tau\tau$ processes. . . . .	94
6.3	Illustration of the differences in $d\phi(\ell = \tau \text{ or } \mu, MET)$ and $p_T^e$ spectrums in LFV and SM $H \rightarrow \tau\tau$ processes. . . . .	95
6.4	Estimation of the neutrino momentum $\not{E}_{T,proj}$ by using the component of the missing transverse energy $\not{E}_T$ which is collinear to the visible decay products of tau in the transverse plane. . . . .	97
6.5	Illustration of decision tree. . . . .	98
6.6	Distribution of the input variables to the BDT for the $H \rightarrow \mu\tau_h$ process.	102
6.7	Overtraining check as performed in TMVA for the trained BDT in $H \rightarrow \mu\tau_h$ channel for 2016 (a), 2017 (b), and 2018 (c). . . . .	104
6.8	Distribution of the input variables to the BDT for the $H \rightarrow \mu\tau_e$ process.	106
6.9	Overtraining check as performed in TMVA for the trained BDT in $H \rightarrow \mu\tau_e$ channel for 2016 (a), 2017 (b), and 2018 (c). . . . .	107
6.10	Distribution of the input variables to the BDT for the $H \rightarrow e\tau_h$ process.	108
6.11	Overtraining check as performed in TMVA for the trained BDT in $H \rightarrow e\tau_h$ channel for 2016 (a), 2017 (b), and 2018 (c). . . . .	109
6.12	Distribution of the input variables to the BDT for the $H \rightarrow e\tau_\mu$ process.	112
6.13	Overtraining check as performed in TMVA for the trained BDT in $H \rightarrow e\tau_\mu$ channel for 2016 (a), 2017 (b), and 2018 (c). . . . .	113
7.1	Feynman diagrams of background processes to LFV Higgs boson decays: (a) $H \rightarrow \tau\tau$ , (b) $Z \rightarrow \tau\tau$ , (c) $t\bar{t}$ , (d) Single Top, (e) WW, (f) WZ, (g) ZZ, and (h) $W\gamma^{(*)}$ . . . . .	115
7.2	Schematic of Embedding Technique . . . . .	117
7.3	Distributions of $M_{\text{col}}$ discriminator in the $Z \rightarrow \tau\tau$ control regions for the (a) $H \rightarrow \mu\tau_h$ , (b) $H \rightarrow \mu\tau_e$ , (c) $H \rightarrow e\tau_h$ , and (d) $H \rightarrow e\tau_\mu$ channels.	119
7.4	Distributions of BDT discriminator in the $Z \rightarrow \tau\tau$ control regions for the (a) $H \rightarrow \mu\tau_h$ , (b) $H \rightarrow \mu\tau_e$ , (c) $H \rightarrow e\tau_h$ , and (d) $H \rightarrow e\tau_\mu$ channels.	120
7.5	Signal region (green) contrasted with the control regions used for estimating the misidentified background . . . . .	121

7.6	Fit performed to $\tau_h$ misidentification rates for $H \rightarrow \mu\tau_h$ (a) and $H \rightarrow e\tau_h$ (b) channel as a function of $\tau_h$ $p_T$ for the different years. The misidentification rates used are further parametrized based on $\tau_h$ Decay Mode along with the pseudorapidity of $\tau_h$ . However, here only the inclusive misidentification rates are shown. The misidentification rates are labeled as “tight-to-loose” to clarify that they are calculated as a ratio of the number of events passing the tight WP to the loose WP of DNN discrimination against jets. . . . .	123
7.7	Fit performed to the $\mu$ (a) and $e$ (b) misidentification rates as a function of their $p_T$ for 2016 (Left), 2017 (Center), and 2018 (Right). The misidentification rates are labeled as “tight-to-loose” to clarify that they are calculated as a ratio of the number of events passing the tight isolation to the loose isolation. The hyperbolic tangent function is used for performing the fit. . . . .	124
7.8	Distributions of $M_{col}$ discriminator in the same-sign (Left) and W boson enriched (Right) control regions for the $H \rightarrow \mu\tau_h$ (top) and $H \rightarrow e\tau_h$ (bottom) channels. . . . .	126
7.9	Distributions of BDT discriminator in the same-sign (Left) and W boson enriched (Right) control regions for the $H \rightarrow \mu\tau_h$ (top) and $H \rightarrow e\tau_h$ (bottom) channels. . . . .	127
7.10	QCD OS/SS extrapolation factors in events with 0 Jets (Left), 1 Jet (Center), and 2 Jets (Right) for 2016 (a), 2017 (b), 2018 (c). The line is the best fit, and the shaded region corresponds to the shape uncertainties. . . . .	129
7.11	(a) Corrections of the QCD OS/SS extrapolation factors determined in the region with an anti-isolated muon as a function of the $p_T$ of the electron and the muon, using data collected in 2016, 2017, and 2018. (b) Correction of the QCD OS/SS extrapolation factors to account for the mismodeling introduced by anti-isolating the muon to measure the SFs, using data collected in 2016, 2017, and 2018. . . . .	130
7.12	Distribution of $M_{col}$ discriminator in the muon anti-isolated control regions for the $H \rightarrow \mu\tau_e$ channel. . . . .	131
7.13	Distributions of $M_{col}$ (BDT) discriminator in $t\bar{t}$ enriched control region for $H \rightarrow \mu\tau_e$ and $H \rightarrow e\tau_\mu$ channel. . . . .	132
8.1	Test statistic distributions for ensembles of pseudo-data generated for background-only (blue) and signal-plus-background (red) hypotheses [77]. . . . .	136

8.2 (Left) An example of differential distribution of possible limits on $\mu$ for the background-only hypothesis. (Right) The cumulative probability distribution of the plot on the left with 2.5%, 16%, 50%, 84%, and 97.5% quantiles defines the median expected limit as well as the 68% and 95% bands for the expected value of $\mu$ for the background-only hypothesis. . . . .	138
9.1 BDT discriminator distributions for the observed and estimated background in the $H \rightarrow \mu\tau_h$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow \mu\tau) = 5\%$ . $H \rightarrow \mu\tau_h$ channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.	146
9.2 BDT discriminator distributions for the observed and estimated background in the $H \rightarrow \mu\tau_e$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow \mu\tau) = 5\%$ . $H \rightarrow \mu\tau_e$ channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.	147
9.3 $M_{\text{col}}$ distributions for the observed and estimated background in the $H \rightarrow \mu\tau_h$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow \mu\tau) = 10\%$ . $H \rightarrow \mu\tau_h$ channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature. .	148
9.4 $M_{\text{col}}$ distributions for the observed and estimated background in the $H \rightarrow \mu\tau_e$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow \mu\tau) = 10\%$ . $H \rightarrow \mu\tau_e$ channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature. .	149

9.5	BDT discriminator distributions for the observed and estimated background in the $H \rightarrow e\tau_h$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow e\tau) = 5\%$ . $H \rightarrow e\tau_h$ channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.	150
9.6	BDT discriminator distributions for the observed and estimated background in the $H \rightarrow e\tau_\mu$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow e\tau) = 5\%$ . $H \rightarrow e\tau_\mu$ channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.	151
9.7	$M_{\text{col}}$ distributions for the observed and estimated background in the $H \rightarrow e\tau_h$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow e\tau) = 10\%$ . $H \rightarrow e\tau_h$ channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.	152
9.8	$M_{\text{col}}$ distributions for the observed and estimated background in the $H \rightarrow e\tau_\mu$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow e\tau) = 10\%$ . $H \rightarrow e\tau_\mu$ channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.	153
9.9	Observed (expected) 95% CL upper limits on the $\mathcal{B}(H \rightarrow \mu\tau)$ (left) and $\mathcal{B}(H \rightarrow e\tau)$ (right) for each individual category and combined from the BDT fit analysis.	155
9.10	Observed (expected) 95% CL upper limits on the $\mathcal{B}(H \rightarrow \mu\tau)$ (left) and $\mathcal{B}(H \rightarrow e\tau)$ (right) for each individual category and combined from the $M_{\text{col}}$ fit analysis.	156

9.11 Constraints on the LFV Yukawa couplings, $ Y_{\mu\tau}  -  Y_{\tau\mu} $ (left), and $ Y_{e\tau}  -  Y_{\tau e} $ (right). The expected (red line) and observed (black solid line) limits are derived from the results shown in Figure 9.9. The flavor-diagonal Yukawa couplings are approximated by their SM values. The green hashed region is derived by the CMS direct search presented in this paper. The green (yellow) band indicates the range that is expected to contain 68% (95%) of all observed limit variations from the expected limit. The shaded regions are derived constraints from null searches for $\tau \rightarrow 3\mu$ or $\tau \rightarrow 3e$ (dark green) [89] and $\tau \rightarrow \mu\gamma$ or $\tau \rightarrow e\gamma$ (lighter green) [6]. The blue diagonal line is the theoretical naturalness limit $ Y_{ij}Y_{ji}  \leq m_i m_j / v^2$ . . . . .	157
A.1 Illustration of the variables $\theta_{inv}$ and $\phi_{inv}$ that specify the orientation of the $\mathbf{p}_{inv}$ vector relative to the momentum vector $\mathbf{p}_{vis}$ of the visible $\tau$ decay products. . . . .	166
A.2 Collinear mass vs SVFit mass. . . . .	168

## TABLES

2.1	Constraints on LFV Yukawa couplings from low-energy measurements [6]. . . . .	32
3.1	Integrated luminosity considered for physics analysis at the CMS experiment during Run II . . . . .	39
6.1	Event selection criteria for the kinematic variables for the $H \rightarrow \mu\tau$ channels . . . . .	103
6.2	Event selection criteria for the kinematic variables for the $H \rightarrow e\tau$ channels . . . . .	110
8.1	Systematic uncertainties in the expected event yields for the $H \rightarrow \mu\tau$ channels. . . . .	140
8.2	Systematic uncertainties in the expected event yields for the $H \rightarrow e\tau$ channels. . . . .	141
9.1	Observed and expected upper limits at 95% CL and best fit branching fractions for each individual jet category, and combined, in the $H \rightarrow \mu\tau$ process. . . . .	158
9.2	Observed and expected upper limits at 95% CL and best fit branching fractions for each individual jet category, and combined, in the $H \rightarrow e\tau$ process. . . . .	159
9.3	Observed and expected upper limits at 95% CL and best fit branching fractions for each individual jet category, and combined, in the $H \rightarrow \mu\tau$ process from $M_{\text{col}}$ fit analysis. . . . .	160
9.4	Observed and expected upper limits at 95% CL and best fit branching fractions for each individual jet category, and combined, in the $H \rightarrow e\tau$ process from $M_{\text{col}}$ fit analysis. . . . .	161
9.5	Summary of observed and expected upper limits at 95% CL, best fit branching fractions and corresponding constraints on Yukawa couplings for $H \rightarrow \mu\tau$ and $H \rightarrow e\tau$ processes. . . . .	162

## CHAPTER 1

### INTRODUCTION

The nature of our universe is explored in particle physics by understanding the properties of elementary particles and the fundamental forces between them. The Standard Model (SM) of particle physics has been well-tested, and it is consistent with all known particles and all known interactions, but gravity. The fundamental forces are mediated by the exchange of particles, called gauge bosons. The W and Z bosons mediate the weak interaction. These bosons are heavy, and they acquire mass through the Higgs mechanism. The Higgs mechanism is incorporated in SM by introducing a new scalar field associated with a new particle, the Higgs boson.

Yukawa interactions between the scalar field and the fermion fields give rise to the mass of the fermions. In SM, the mass matrix and the Yukawa interaction matrix are diagonalized simultaneously by choosing a particular eigenstate basis. This corresponds to the SM Higgs boson not having lepton flavor violating (LFV) decays. The SM has been well-tested over a wide range of energies, and despite its success, it is considered an effective theory as it fails to include gravity. It also cannot explain Dark Matter, which is inferred from astrophysical and cosmological observations. It also cannot explain Dark Energy, which is inferred to explain our universe's accelerated expansion.

Flavor physics is an active area of research as we try to understand the reason for observing six quarks and six fermions that can be arranged into three generations. The large difference between matter and anti-matter in our universe cannot be explained completely by the SM's CP-violating effects. Cabibbo-Kobayashi-Maskawa

(CKM) matrix is the observed CP-violation source, a  $3 \times 3$  unitary matrix that can be parameterized by three mixing angles and one complex CP-violating phase [1]. The CKM matrix describes the quarks' mixing because the mass eigenstates are not equal to the weak eigenstates.

Due to the absence of right-handed neutrinos, there is no Dirac mass term for neutrinos in the SM. There is no Majorana mass term due to the exact conservation of baryon and lepton number. However, several experiments have observed neutrino oscillations, which can be explained by nonzero neutrino masses. The mixing of the neutrinos is described by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [1]. An additional source of the CP-violation is the PMNS matrix described by three mixing angles and a phase. Besides, the SM cannot explain the substantial difference between the very small neutrino masses ( $m < 2$  eV) and the masses of the charged leptons and quarks ( $\sim 173$  GeV for the top quark) [1].

One question that needs to be addressed is any relation between quark mixing and neutrino mixing. The Yukawa interactions within the SM cannot be constrained by Gauge invariance. To address the open questions, we need new physics models. Grand Unification Theories (GUTs) are a class of new physics models that try to unify all known interactions but gravity in a single gauge group. At the same time, extra dimensions could link flavor to the geometry of these extra dimensions [2]. Another class of new physics models is simple extensions of the SM with more than one Higgs doublets, and they could add additional sources for CP-violation. These multi-Higgs doublet models can have tree level Higgs-mediated flavor changing neutral currents leading to lepton-flavor violating (LFV) Yukawa couplings [2].

Large Hadron Collider (LHC) [3] was built to discover the Higgs boson and other exotic particles predicted by models like Supersymmetry or potential Dark Matter particles like weakly interacting massive particles. It was designed to run at a center-of-mass energy of  $\sqrt{s} = 14$  TeV and is the most powerful particle collider built and

operational to date. ATLAS and Compact Muon Solenoid (CMS) are two general-purpose detectors designed to detect the result of the particle collisions. In 2012, both experiments discovered the Higgs boson with a mass of 125 GeV [4, 5]. Many precision measurements need to be performed to confirm if it is the SM Higgs boson or a Higgs boson of an SM’s possible extension. Until now, no significant deviations from the SM Higgs boson have been observed.

In this thesis, a search for the Higgs boson’s LFV decays to a muon and a tau ( $\mu\tau$ ) or an electron and a tau ( $e\tau$ ) is presented. In each channel, the tau can further decay either hadronically or leptonically. If the tau lepton decays leptonically, we only consider final states with different lepton flavors to avoid the large Drell-Yan background. Thus, the  $\mu\tau$  channel is further divided into  $\mu\tau_h$  and  $\mu\tau_e$  final states, while the  $e\tau$  channel is further divided into the  $e\tau_h$  and  $e\tau_\mu$  final states. This search is performed with proton-proton collision data collected at a center-of-mass energy of  $\sqrt{s} = 13$  TeV in 2016, 2017, and 2018 corresponding to an integrated luminosity of  $35.9\text{fb}^{-1}$ ,  $41.5\text{fb}^{-1}$ , and  $59.3\text{fb}^{-1}$ , respectively. Thus, the total integrated luminosity analyzed in this search is  $137\text{fb}^{-1}$ , which corresponds to a four times larger dataset than the one used in the previous search.

Apart from the much larger dataset, significant improvements have been made regarding the background estimation techniques and a detailed study of the systematic uncertainties involved in the analysis. A significant portion of the background is estimated using data-driven techniques with limited dependence on the MC simulations. This gives rise to a better description of the event kinematics and an improvement with regards to the corresponding systematics involved. The categorization of the events has been kept the same as the previous search. In contrast, improvements have been made to the classification done with multivariate techniques for discriminating the signal from the background to improve the search’s sensitivity. All these changes collectively gave rise to a factor of two improvements in the sensitivity of the search

and helped to set the most stringent limits set on these LFV Higgs decays to date.

This thesis is structured as follows. An overview of the SM of particle physics, along with a short review on LFV decays of the Higgs boson, is given in Chapter 2. In that chapter, we will see how new physics can introduce LFV Yukawa-couplings and how low-energy measurements constrain them [6]. The experimental setup of LHC and the CMS experiment is discussed in Chapter 3. In that chapter, the various sub-detectors of the CMS experiment, which work in unison, is explained. Chapter 4 will dive into the MC event generation, followed by the event reconstruction description in Chapter 5. The event selection is described in Chapter 6, followed by a detailed explanation of the background estimation in Chapter 7. We will understand the systematic uncertainties in Chapter 8 and the corresponding statistical analysis to obtain the results. We will then discuss the results in Chapter 9, followed by a conclusion. Some studies for the future LFV analysis have been performed, and they are detailed in the Appendix.

## CHAPTER 2

### THEORETICAL MOTIVATION

The constituents of matter and their interactions at the most fundamental level is described by the Standard Model (SM) of particle physics. SM is a renormalizable quantum field theory with a  $SU(3) \times SU(2) \times U(1)$  symmetry structure. SM incorporates electromagnetic, weak, and strong interactions. A brief theoretical overview of the elementary particles, the fundamental interactions, and the Higgs boson is presented in this chapter. The main shortcomings of the SM are listed at the end of the chapter.

#### 2.1 Phenomenological overview

The intrinsic angular momentum (spin) of the elementary particles is used to classify them. Elementary particles with half-integer spin ( $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$ ), are classified as fermions, while particles with integer spin (0, 1, 2, ...) are classified as bosons. Quarks and leptons constitute matter and are fermions. There are six quark flavors: up (u), down (d), strange (s), charm (c), bottom (b), and top (t). There are three different types of charged leptons: the electron (e), the muon ( $\mu$ ), and the tau ( $\tau$ ). Each lepton has its corresponding neutral partner: the electron neutrino ( $\nu_e$ ), the muon neutrino ( $\nu_\mu$ ), and the tau neutrino ( $\nu_\tau$ ). The six flavors of leptons and quarks can be arranged into three generations.

Every fermion has a corresponding anti-particle with the same properties but opposite charges. Elementary particles of the SM can be seen in Figure [? ]. All interactions, except gravitation, are part of the Standard Model and can be described

as quantum fields. Their interactions are mediated by field quanta, the gauge bosons, which have spin 1. Quantum electrodynamics (QED) is the relativistic quantum field theory of electrodynamics. The weak interaction and electromagnetism can be unified to one theory, the electroweak theory. Quantum chromodynamics (QCD) is the theory of strong interaction.

## Standard Model of Elementary Particles

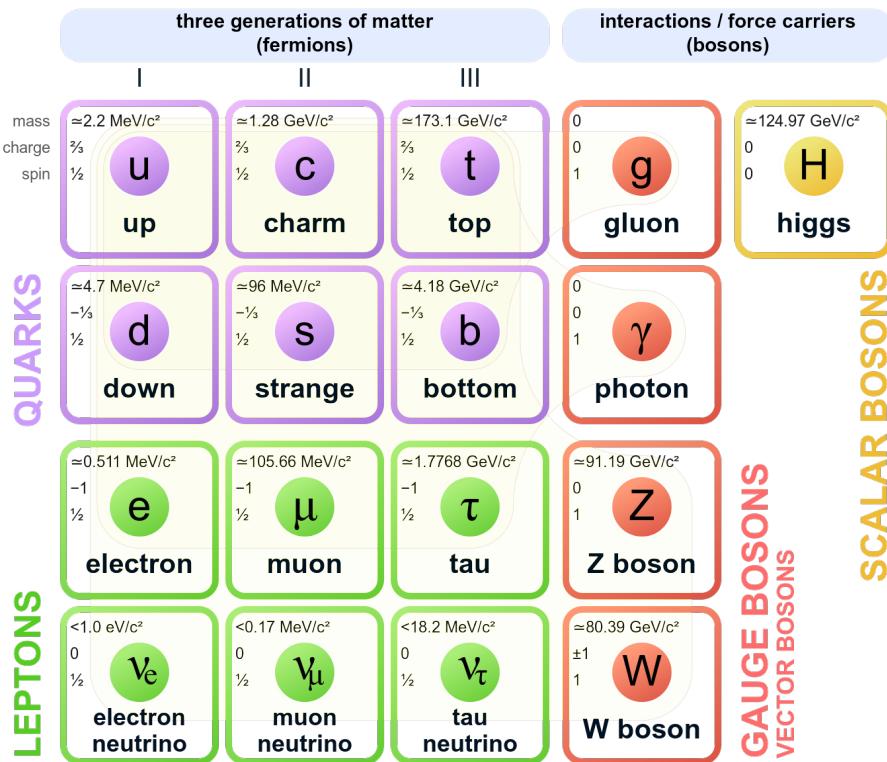


Figure 2.1. Elementary particles of the Standard Model.

### 2.1.1 Quantum electrodynamics

Electromagnetism is described by Quantum electrodynamics (QED). QED interactions are mediated by the photon ( $\gamma$ ), which acts between electrically charged particles. The photon carries no electrical charge and is massless; therefore, electromagnetic interaction has an infinite range. Neutrinos are electrically neutral; thus, they do not interact via electromagnetic interaction. Up-type quarks have a charge of  $+2/3$ , while down-type quarks have a charge of  $-1/3$ . In QED, the elementary process is the emission and absorption of a photon:  $e^- \rightarrow e^- + \gamma$ .

Initial and final-state particles define the physical process. The electromagnetic interaction processes can be described by combining two or more of the fundamental vertices to Feynman diagrams, representing mathematical expressions for calculating the probability amplitudes for a given process. Antiparticles are indicated as arrows going backward in time. Internal lines represent particles, which cannot be observed. The total probability amplitude of a process is proportional to all Feynman diagrams' squared sum, representing the same process. The calculated amplitudes can have infinite contributions. Renormalisation aims to separate the finite and infinite parts of the amplitude. Regularisation modifies the infinite observable by adding a parameter to make it finite. The resulting observable depends on the additional parameter, but it can be computed without divergences. The result is obtained by taking the parameter to its physical limit.

Virtual  $e^+e^-$  pairs can spontaneously be produced around a “bare” electron reducing the observed charge at larger distances. This is referred to as charge screening. The fine-structure constant  $\alpha$  characterizes the strength of the electromagnetic interaction between electrically charged particles. Each vertex introduces a factor proportional to  $\sqrt{\alpha}$ , which depends on the energy scale. The value of  $\alpha$  decreases with energy to a constant value of  $\approx 1/137$  for long-distance interactions. Very precise QED predictions can be obtained using only lower-order diagrams, as the

higher-order diagrams can be neglected due to the small value of  $\alpha$ .

### 2.1.2 Electroweak interaction

The weak interaction is the mechanism of interaction between subatomic particles responsible for the radioactive decay of atoms. It is the only interaction that can change particles' flavor, and all fermions can interact weakly. Neutrinos can only interact via weak interaction. Three gauge bosons mediate the interaction: the  $W^+$ ,  $W^-$ , and  $Z$  bosons. The  $Z$  boson can couple to two fermions of the same flavor, while the  $W$  boson couples to fermions of a different flavor. Also, both bosons can interact with each other. As the  $W$  bosons are electrically charged, they can additionally couple to photons. The weak interaction only acts over very short distances in the order of  $10^{-18}\text{m}$ , because the  $Z$  and  $W$  bosons have large masses, about 91 GeV and 80 GeV, respectively.

### 2.1.3 Quantum chromodynamics

Strong interaction is described by Quantum chromodynamics (QCD), which binds neutrons and protons into atomic nuclei. The strong interaction has a short-range in the order of  $10^{-15}\text{m}$ . QCD interactions are mediated by a massless gauge boson called gluon ( $g$ ), which acts on particles with a color charge. There are three such charges: red, green, and blue ( $r, g, b$ ). Quarks are the only fermions which carry a color charge. A fundamental process of the strong interaction is the process, where a quark emits or absorbs a gluon:  $q \rightarrow q + g$ . The gluons carry color themselves, a color, and an anti-color; therefore, they interact. There are eight types of gluons.

A bare quark is surrounded by a “sea” of virtual quarks and gluons. At shorter distances, corresponding to higher energies, the bare charge can be seen. This corresponds to the phenomenon of asymptotic freedom in which the interaction strength gets weaker with increasing energy and decreasing distance. The potential energy

between two quarks is large enough to create a real quark-antiquark pair from the vacuum at a large distance. This process is known as fragmentation or hadronization. Two separating quarks always hadronize to colorless particles. The observation that only colorless bound states have been observed in nature is referred to as color confinement. The strong interaction binds quarks into composite states called hadrons. Hadrons can be classified into baryons, which are fermions, and mesons, which are bosons. Baryons are composed of three quarks. Mesons are composed of a quark and an anti-quark. The coupling strength of the strong interaction  $\alpha_S$  describes the dependence of the effective charge on the distance between them. In the lowest order,  $\alpha_S(Q)$  is given by:

$$\alpha_S = \frac{6\pi}{(33 - 2n_f) \ln(Q/\Lambda_{\text{QCD}})} \quad (2.1)$$

where  $Q$  is the momentum transfer in a given process,  $n_f$  the number of flavors which can participate in the process, and  $\Lambda_{\text{QCD}}$  corresponds to the energy boundary of hadronization. For  $Q \approx \Lambda_{\text{QCD}}$  quarks and gluons interact strongly and form hadrons, while  $\alpha_S(Q)$  becomes smaller for  $Q \gg \Lambda_{\text{QCD}}$  and quarks and gluons interact with each other only weakly. At low energies, perturbative methods cannot investigate the theory due to the large coupling strength.

#### 2.1.4 The Higgs boson

In the SM, the gauge bosons must have no mass; however,  $W$  and  $Z$  bosons have non-zero mass. They receive their mass by interacting with the Higgs field and via the Higgs mechanism. Also, the fermions receive mass via this mechanism. The Higgs boson ( $H$ ) is the field quanta of the Higgs field and has spin-0. The Higgs boson's fundamental vertices with the fermions, the gauge bosons ( $V$ ) of the weak interaction, and its self-interactions can be seen in Figure [? ]. It couples to two

fermions of the same type, two W or Z bosons, and further Higgs bosons.

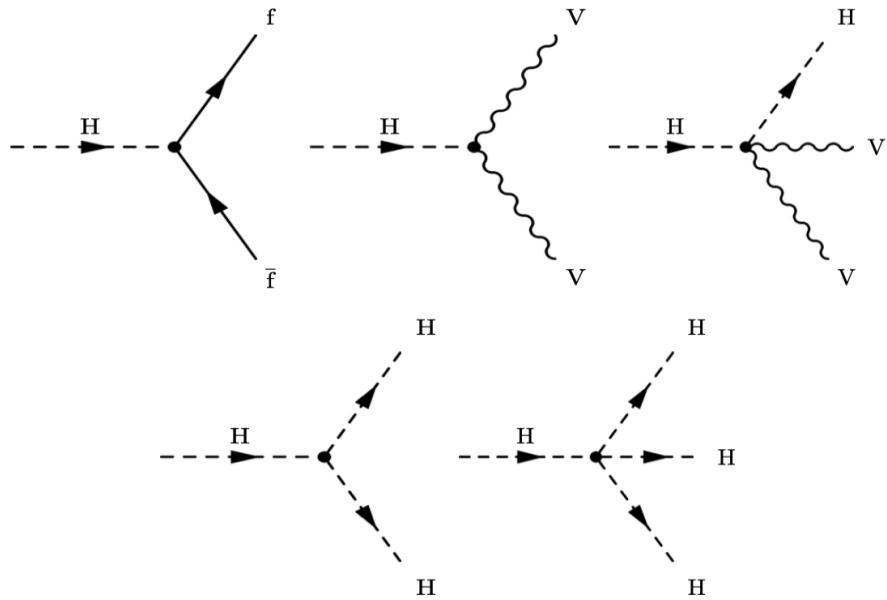


Figure 2.2. Fundamental vertices of the Higgs boson. Fermions are denoted  $f$ , anti-fermions are denoted  $\bar{f}$ , and Gauge bosons of the weak interaction are denoted  $V$ .

## 2.2 Overview of theoretical building blocks

The Standard Model is a renormalizable quantum field theory. It has a  $SU(3) \times SU(2) \times U(1)$  symmetry structure. Understanding the general, abstract details of structures is the aim of group theory. The corresponding Lagrangian of the theory describes the dynamics of a system. A type of quantum field theory is gauge theory, where gauge fields are introduced by requiring symmetry under local transformations. Finally, the idea of spontaneous symmetry breaking is discussed.

Group theory: A group  $(G, \star)$  is a set of objects  $G$  and some operations  $\star$  on those objects. A set of elements  $g_i$  is a group under some operation if the group axioms are satisfied: closure, associativity, the existence of an identity element, and an inverse element. A group is Abelian if the elements of the group commute under the group operation ( $g_i \star g_j = g_j \star g_i$ ). Groups that are not Abelian are called non-Abelian.

The order  $n$  of the group is given by the number of elements in the group. A group can be represented by  $n \times n$  matrices. A special type of groups are Lie groups, which are parameterized by one or more continuous variables. Transformations of elements  $\in \mathbb{C}$  can be described as matrices  $R$ . These matrices are unitary if the Hermitian conjugate is also the inverse of the transformation,  $R^\dagger = (R^*)^T = R^{-1}$ . A set of such  $n \times n$  matrices forms the unitary group  $U(n)$ , which preserves the norm under transformations, thus the probability amplitude. Some of these unitary matrices  $R$  have a determinant  $\det |R| = 1$ . The volume is preserved under such transformations. Such matrices form the special unitary group  $SU(n)$ .

Lagrangian formalism: Its Lagrangian gives the information about the dynamics of a system as a function of generalized coordinates  $q$  and their time derivatives  $\dot{q}$ . In classical physics, the Lagrangian is defined as the kinetic energy  $T$  minus the potential energy  $V$ :

$$L = T(q, \dot{q}) - V(q) \quad (2.2)$$

Fields are functions of spacetime. In field theory, the Lagrangian is replaced by the Lagrangian density  $L$ , which is a function of the fields  $\phi(x^\mu)$  in the system and their derivatives  $\partial_\mu \phi$ . The action  $S$  of the system is defined as:

$$S = \int d^4x \mathcal{L} \quad (2.3)$$

Quantum field theory: Quantum field theory (QFT) is a relativistic quantum the-

ory of interactions. Particles are interpreted as fields, and other particles, the gauge bosons mediate forces between particles. In quantum mechanics, space is treated as an operator, while time is treated as in classical physics, as a parameter. However, space and time have to be treated in the same way for a relativistic theory. This is solved in QFT by quantizing in a new way. The fields are reinterpreted no longer as states but as operators parameterized by the spacetime coordinates. Commutation relations are imposed on the field operators. It is impossible to measure both the field and its time rate of change simultaneously to infinite precision at the same spacetime point. The state with the lowest possible energy level is called a vacuum. Acting on the vacuum with a field operator produces a state with some energy. The Lorentz transformations known from special relativity are elements of the Lorentz group. This group contains two copies of the SU(2) group, where the SU(2) group represents quantum mechanical spin. There are different representations, which are characterized by the value  $j$ , the spin of the particle. Consequently, a representation  $(j, j')$  of the Lorentz group is characterised by two values  $j$  and  $j'$  for each SU(2) subgroup. There are three physically acceptable irreducible representations of the Lorentz group:

- the  $(0, 0)$  scalar representation acting on scalars
- the  $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$  left-handed/right-handed spinor representations acting on spinors
- the  $(\frac{1}{2}, \frac{1}{2})$  vector representation acting on vectors.

There are two types of spinors, left-handed  $\Psi_L$  and right-handed  $\Psi_R$ , which transform the same under rotations, but they transform differently under boosts. Some fundamental properties of these objects are outlined in the following.

Scalar fields are spin-0 fields. Previously, spin was defined as intrinsic angular momentum. Spin is not a rotation in spacetime but the spinor space of SU(2). Complex scalar fields have the form  $\phi = \phi_1 + i\phi_2$ , with two real degrees of freedom.

The fields  $\phi$  and  $\phi^\dagger$  are treated independently. The Lagrangian density is given as a function of both the fields and their derivatives.

Spinor fields  $\Psi$  are spin-1/2 fields and consist of four complex fields  $\Psi_i$ :

$$\Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix}, \Psi^\dagger = (\Psi_1^*, \Psi_2^*, \Psi_3^*, \Psi_4^*) \quad (2.4)$$

The adjoint spinor  $\bar{\Psi} = \Psi^\dagger \gamma^0$  has to be defined to have Lorentz invariant terms in the Lagrangian density. The matrices satisfy the anti-commutation relation (Clifford algebra):

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = -2\eta^{\mu\nu}\not{1} \quad (2.5)$$

with Minkowski metric  $\eta^{\mu\nu}$ . The Lagrangian density is given as a function of the fields  $\Psi, \bar{\Psi}$  and their derivatives. Parity transformations  $\Lambda_P$  are one type of Lorentz transformations, which switch the handedness of a coordinate frame. Left-handed spinors are transformed into right-handed spinors under parity transformations. To be Lorentz invariant, spinor fields  $\Psi$  should have a left-handed and a right-handed component, with the adjoint field of a left-handed spinor field  $\bar{\Psi}_L$  being the same field as the right-handed spinor field  $\Psi_R$ :

$$\bar{\Psi}_L = \Psi_R, \quad \bar{\Psi}_R = \Psi_L \quad (2.6)$$

The projection operators  $P_\pm$  are the projection operators onto the left- and right-handed parts of the field  $\Psi$ :

$$P_+ \Psi = \Psi_R, \quad P_- \Psi = \Psi_L \quad (2.7)$$

Another important transformation is the charge conjugation:  $\Psi \rightarrow C\bar{\Psi}^T = -i\gamma^2\Psi^*$ , which swaps the charge of the field. Vector fields  $A_\mu$  have spin 1. They have four components that transform as a spacetime vector under Lorentz transformations. Lorentz invariant terms are:

$$A_\mu A^\mu, \quad (\partial_\mu A_\nu)(\partial^\mu A^\nu), \quad (\partial_\mu A^\mu)(\partial_\nu A^\nu) \quad (2.8)$$

The fundamental fermions are represented as spinor fields, while gauge bosons are represented as vector fields in QFT. Combining the Lagrangian densities of both fields leads to two fields, which do not interact with each other. An interaction term has to be added, which mixes both fields. Such a term can be introduced using gauge theory.

**Gauge theory:** A type of field theories are gauge theories, which are theories in which the Lagrangian is invariant under a continuous local transformations group. For example, a spinor field can be transformed with the global U(1) transformation  $\Psi \rightarrow e^{i\alpha}\Psi$ , with  $\alpha$  being a real parameter. The transformation is just a phase and acts on the field in the same way at every point in spacetime. A local transformation has a spacetime dependence. In this example of the U(1) transformation,  $\alpha$  would depend on spacetime:  $\alpha = \alpha(x^\mu)$ . The same type's possible transformations are associated with a Lie group, the theory's symmetry group. All elements of a group can be expressed in terms of the group generators.

First, a Lagrangian density is defined for a fermion field with an equation of motion that satisfies the Dirac equation. This Lagrangian density is not invariant under a local gauge transformation. A new vector field  $A_\mu$  has to be introduced, called gauge field, for each group generator  $T_a$  to ensure invariance under the local group transformations, called gauge invariance. The vector field can be expressed with the group generators,  $A_\mu = A_\mu^a T_a$ , and the gauge bosons are the quanta of these fields. The derivative operator  $\partial_\mu$  in the Lagrangian density is replaced by the

covariant derivative  $D_\mu$ , which is  $\partial_\mu$  with a term proportional to a gauge field.

Local transformations can be written in form of  $U = e^{ig\chi_a(x^\mu)T_a}$ , where  $g$  is a real parameter, the coupling constant, and  $\chi_a(x^\mu)$  is a scalar with a spacetime dependence. The covariant derivative is then given by  $D_\mu = \cancel{\partial}^{n \times n} \partial_\mu + igA_\mu^a T_a$ . Interaction terms are introduced by the second term of the covariant derivative. A kinetic term, proportional to the fields' derivatives, can be introduced for including dynamics of the vector field. Vector fields with mass break the local gauge symmetry; therefore, the vector fields are massless, i.e., have no terms proportional to  $A_\mu A^\mu$ .

All fundamental interactions of the SM can be derived from this general principle of local gauge invariance. As the SM W and Z bosons are massive, a mechanism is needed to acquire mass. This mechanism is the Higgs mechanism and combines the principles of gauge theory and spontaneous symmetry breaking.

**Spontaneous symmetry breaking:** Gauge fields can acquire mass in case of broken local symmetry. Breaking a global symmetry always results in massless bosons, called Goldstone bosons [7, 8]. Consider a complex scalar field with a Lagrangian which has a local U(1) symmetry and a potential  $V(\phi^\dagger, \phi)$  with the vacuum  $V_{minimum}$  given by the circle  $|\phi| = v$ . Choosing a single ground state breaks the system's symmetry and is considered “spontaneous” because there are no external means by which this occurs. The theory can be rewritten in terms of the new vacuum, by a change of basis. The field is expanded around the constant vacuum value, writing the fields in terms of fluctuations around the chosen vacuum.

The Higgs mechanism combines the principles of gauge invariance with spontaneous symmetry breaking. For a local U(1) symmetry, the vacuum can be chosen so that the vacuum is real and that  $\phi$  is always real; therefore,  $\phi$  can be expanded as  $\phi = v + h$ , with  $h$  being a real scalar field. Breaking the local U(1) symmetry results in a real scalar  $h$  with a mass, and the field  $A_\mu$ , originating from the local gauge symmetry, acquires mass. The Higgs mechanism introduces mass into a theory

by breaking the local symmetry. A scalar Higgs field  $h$  is introduced, and the gauge field of the broken symmetry acquires mass. In general, gauge fields corresponding to group generators of broken symmetries acquire mass, while those corresponding to generators of unbroken symmetries do not acquire a mass.

### 2.3 The Standard Model

The Standard Model is a non-Abelian gauge theory based on the group  $SU(3)_C \times SU(2)_L \times U(1)_Y$ , where  $C$  stands for color,  $Y$  stands for the hypercharge to distinguish  $U(1)_Y$  from the  $U(1)$  of QED, and  $L$  indicates that the interaction only involves left-handed states. QCD is based on the group  $SU(3)_C$ , which corresponds to rotations in color space. Eight gauge bosons, called gluons, mediate the strong interaction. Fermions are described as a triplet of spinors which differ in color

$$\psi = \begin{pmatrix} \psi_r \\ \psi_b \\ \psi_g \end{pmatrix} \quad (2.9)$$

Group	Operation	Coupling	Vector field
$U(1)_Y$	phase rotation	$g'$	$B_\mu$
$SU(2)_L$	rotation in the weak isospin	$g$	$W_\mu^a$
$SU(3)_C$	rotation in the colour	$g_S$	$G_\mu^a$

(2.10)

There are eight gauge fields  $G_\mu^a$  with the gluons being the field quanta, which mediate the strong interaction. Leptons do not interact under the  $SU(3)$  color group, but quarks do. The electroweak theory is based on the group  $SU(2)_L \times U(1)_Y$ . Only left-handed fields interact under  $SU(2)_L$ ; therefore, only a left-handed neutrino is needed. Left-handed and right-handed fields can interact under the  $U(1)_Y$  group,

therefore charged leptons have to exist in a left-handed and a right-handed state. Left-handed  $SU(2)_L$  doublets

$$\psi_L = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix} \quad (2.11)$$

and right-handed  $SU(2)_L$  singlets  $e_R$  are introduced. There are three gauge fields  $W_\mu^a$  for  $SU(2)_L$ , and one gauge field  $B_\mu$  for  $U(1)_Y$ . Table 2.1 gives an overview of the symmetry groups in the SM. The gauge fields are all massless, but the gauge bosons of the weak theory have mass. Masses can be generated for the W and Z bosons via the Higgs mechanism by breaking the electroweak symmetry. Also, quarks and leptons acquire mass via the Higgs mechanism by applying the mechanism to the corresponding fields.

A Higgs field is introduced, which is a complex scalar  $SU(2)_L \times U(1)_Y$  doublet field. Breaking the vacuum's symmetry via the Higgs mechanism results in a broken theory with two separate forces: the broken weak theory and the unbroken theory of electromagnetism associated with the  $U(1)_{em}$  symmetry of QED. The new fields  $A_\mu$ ,  $W_\mu^\pm$ , and  $Z_\mu^0$  are linear combinations of the vector fields of  $SU(2)_L \times U(1)_Y$ :

$$\begin{aligned} A_\mu &= \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu \\ W_\mu^\pm &= (W_\mu^1 \mp i W_\mu^2)/\sqrt{2} \\ Z_\mu^0 &= -\cos \theta_W W_\mu^3 + \sin \theta_W B_\mu \end{aligned} \quad (2.12)$$

with the weak mixing angle or Weinberg angle  $\theta_W = \tan^{-1}(g'/g)$ ,  $\sin \theta_W = g'/\sqrt{g^2 + g'^2}$  and  $\cos \theta_W = g/\sqrt{g^2 + g'^2}$  in terms of the couplings  $g$  of  $SU(2)_L$  and  $g'$  of  $U(1)_Y$ . The weak mixing angle relates the strength of the weak and electromagnetic interaction, which are mediated by the massive W and Z bosons, and the massless photon which is the field quanta of  $A_\mu$ . Furthermore, the electroweak coupling  $e$  of  $U(1)_{em}$  is given by  $e \equiv g \sin \theta_W$ .

## 2.4 Electroweak symmetry breaking and the Higgs boson

In the SM, the weak gauge bosons acquire mass via the Higgs mechanism. An arbitrary scalar potential

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (2.13)$$

with a self-interacting SU(2) scalar complex field  $\Phi$  is introduced, where  $\mu$  is the mass term of the field. The complex Higgs doublet  $\Phi$  has four degrees of freedom, and it is composed of CP-even and CP-odd components  $\phi^0$  and  $a^0$ , and the complex charged component  $\phi^+$ :

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\phi^+ \\ \phi^0 + ia^0 \end{pmatrix} \quad (2.14)$$

For  $\mu^2 < 0$ , the scalar doublet's neutral component acquires a non-zero vacuum expectation value (VEV)  $v$ . Then  $\phi^0$  can be expanded as  $\phi^0 = v + H$  with:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \quad (2.15)$$

The  $SU(3)_C \times SU(2)_L \times U(1)_Y$  is broken into a  $SU(3)_C \times U(1)_{em}$  symmetry. After symmetry breaking, there are three massless Goldstone bosons, which can be identified with three degrees of freedom of the Higgs field. The Higgs field couples to the W and B fields through the covariant derivative in the kinematic term. As a result, the Goldstone bosons mix with the gauge bosons of the corresponding generators of broken symmetries, and the physics W and Z gauge bosons acquire masses:

$$M_W^2 = \frac{g^2 v^2}{4}, \quad M_Z^2 = \frac{(g'^2 + g^2)v^2}{4} \quad (2.16)$$

There is one remaining degree of freedom, which is the Higgs boson. The mass

of this new scalar (spin-0) particle is given by  $m_H = \sqrt{2\lambda}v$ , with the self-coupling parameter. The sign of  $\mu^2 = -\lambda v^2$  is crucial for the electroweak symmetry. The expectation value of the Higgs field is fixed by the precise measured Fermi coupling GF:  $v = (\sqrt{2}G_F)^{-1/2} \approx 246\text{GeV}$ . The SM also acquire fermion masses through the new interactions between the Higgs field and the fermion, the Yukawa interactions. The charge conjugated Higgs doublet  $\tilde{\Phi}$  is needed to allow interactions with up-type quarks. Then, the Lagrangian of the Yukawa term is given by:

$$\mathcal{L}_{\text{Yukawa}} = -\hat{\lambda}_{d_{ij}} \bar{q}_{L_i} \Phi d_{R_j} - \hat{\lambda}_{u_{ij}} \bar{q}_{L_i} \tilde{\Phi} d_{R_j} - \hat{\lambda}_{\ell_{ij}} \bar{\ell}_{L_i} \Phi e_{R_j} + \text{h.c.} \quad (2.17)$$

where *h.c.* indicates the Hermitian conjugated terms and  $\hat{\lambda}$  are  $3 \times 3$  matrices for the up and down-type quarks, and the charged leptons. Under electroweak symmetry breaking,

the Higgs field acquires a VEV and the fermion mass eigenstate basis is chosen, so that the Higgs interactions are diagonalised:  $\hat{\lambda}_{f_{ij}} \rightarrow \lambda_{f_i} \delta_{ij}$ . The fermions acquire a mass, which is given by  $m_{f_i} = \lambda_{f_i} v / \sqrt{2}$ , with the corresponding Yukawa coupling  $\lambda_f$ . Furthermore, they couple to the Higgs boson

$$\mathcal{L}_{\text{Yukawa}} = -g_{H_{ff}} \bar{\Psi} \Psi H \quad (2.18)$$

with  $g_{H_{ff}} = \frac{m_f}{v}$ . As the masses of the fermions are known, the only free parameter is the SM's Higgs boson mass. The Higgs boson is the most recently discovered fundamental particle. It was discovered by the ATLAS and CMS collaborations in 2012 [9]. Its mass and full width were measured to be  $M_H = 125.09 \pm 0.24\text{GeV}$  and  $\Gamma < 1.7\text{GeV}$ .

## 2.5 Main shortcomings of the Standard Model

The Standard Model is a very successful theory, and the electroweak theory has been verified over a wide range of energies. It was crowned with the discovery of the Higgs boson by the ATLAS and CMS collaborations in 2012. However, the SM is an incomplete theory as it does not incorporate gravitation. In SM's original formulation, neutrinos are strictly massless because there is no Dirac mass term due to the absence of right-handed neutrinos. There is no Majorana mass term due to the conservation of baryon and lepton number. However, neutrino experiments have established the existence of neutrino oscillations, transitions between the different neutrino flavors. This is evidence for neutrino mixing and nonzero neutrino masses. Astrophysical observations imply an unknown type of matter and energy, called dark matter and dark energy, which cannot be explained within the SM. Space observatories like Planck have studied the fluctuations in the spectrum of the relic cosmic microwave background (CMB) and have established the existence of cold dark matter [10]. Also, the CP-violating effects of the SM cannot explain the large difference between matter and anti-matter in the universe. Furthermore, the flavor structure is not explained in the theoretical description of the SM. There is a large number of free parameters, such as masses and mixing parameters. The Higgs mechanism can explain how masses are introduced in the SM, but it cannot explain its values.

## 2.6 Lepton flavor violating decays of the Higgs boson

The Standard Model is considered to be an effective theory up to a certain scale  $\Lambda$ . One shortcoming is that the flavor structure is not incorporated in the theoretical description of the SM. There are three main puzzles in the SM related to flavor [2]:

- The replication of the generation constitutes.
- There is no principle underlying the formation of the Yukawa matrices describing the SM Yukawa interactions.

- The peculiar pattern of fermion masses and mixing cannot be explained.

New physics could explain some of the symmetries and open questions. Grand Unification Theories (GUTs) try to unify all known interactions but gravity in a single gauge group, while extra dimensions could link flavor to this extra dimension's geometry. A simple extension of the Higgs sector with at least two Higgs doublets could add additional CP-violation. A Higgs boson was discovered in 2012, whose properties are compatible with the SM Higgs boson, within experimental uncertainties. Its mass has been measured by the ATLAS and CMS collaborations and found to be  $125.09 \pm 0.24\text{GeV}$ . After the discovery of the Higgs boson, its nature has to be investigated by measuring its properties, including its couplings to charged leptons. Charged leptons acquire mass through Yukawa interactions with the Higgs field, as described in Chapter 2. In the SM, the Yukawa interaction matrix can be diagonalized in the mass basis. Thus, LFV decays of the Higgs boson are forbidden. Extensions of the SM, such as multi-Higgs doublet models, can introduce LFV Yukawa couplings. Multi-Higgs doublet models have in general tree-level Higgs-mediated flavor changing neutral currents leading to LFV Yukawa couplings. LFV Yukawa couplings can also be introduced in composite Higgs models [11] and models with extra dimensions.

In the following, a short review on LFV decays of the Higgs bosons is given, which was the status at the time when the search was started. It is described how new physics can introduce LFV Yukawa-couplings and how low-energy measurements constrain them. Finally, it is described how LFV decays of the Higgs boson can be measured at the LHC and which sensitivity is expected, including observed upper limits on the branching fraction [12].

### 2.6.1 Introduction of lepton flavor violating decays

The Higgs mechanism is responsible for the masses of the gauge bosons and the fermions. First, it is discussed how charged leptons gain mass in the SM. Yukawa

interactions mediate interactions between charged leptons with the Higgs field; therefore, the corresponding Yukawa matrix structure in the mass basis is afterward discussed. Then, it is explained how LFV Yukawa couplings can be introduced considering an effective field theory approach. Finally, an overview is given on the structure of the Yukawa matrix for charged leptons in general. Decay widths and branching fractions of the LFV decay in terms of the corresponding Yukawa couplings are also shown.

Yukawa couplings of charged leptons in the SM Charged leptons acquire a mass via the Higgs mechanism, introduced in Section 2.4. A direct mass term is not gauge invariant, but interactions of the charged leptons  $\ell$  to the scalar Higgs field  $\phi$  can be introduced with the following term:

$$\mathcal{L}_{\text{SM}} = -\lambda_{ij} \bar{L}_L^i \Phi \ell_R^j + \text{h.c.} \quad (2.19)$$

where with  $\bar{L}_L^i = (\bar{\nu}_\ell^i, \bar{\ell}_L^i)$ , are the  $SU(2)_L$  doublets, and  $\ell_R^i$  the weak singlets with the indices  $i, j$  running over generations. After electroweak symmetry breaking, the Higgs doublet can be written in terms of the vacuum expectation value  $v$  and the physical Higgs boson  $H$ :

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \quad (2.20)$$

leading to the following Lagrangian:

$$\begin{aligned} \mathcal{L}_{\text{SM}} &= -\frac{\lambda_{ij}}{\sqrt{2}} (v + H) \bar{\ell}_L^i \ell_R^j + h.c. \\ &= -\frac{\lambda_{ij}}{\sqrt{2}} v [\bar{\ell}_L^i \ell_R^j] - \frac{\lambda_{ij}}{\sqrt{2}} [\bar{\ell}_L^i \ell_R^j H] + h.c. \end{aligned} \quad (2.21)$$

where the term proportional to  $\bar{\ell}_L^i \ell_R^j$  is the mass term of the charged leptons and the term proportional to  $\bar{\ell}_L^i \ell_R^j H$  describes the Yukawa interactions. The mass matrix

$m$ , with  $m_{ij} = \frac{\lambda_{ij}}{\sqrt{2}}v$ , can be diagonalised using the matrices  $V_L$  and  $V_R$ :

$$m = \begin{pmatrix} m_e & 0 & 0 \\ 0 & m_\mu & 0 \\ 0 & 0 & m_\tau \end{pmatrix} = \frac{1}{\sqrt{2}}V_L \lambda V_R^\dagger v = \frac{\lambda^m}{\sqrt{2}}v \quad (2.22)$$

where  $\lambda^m = V_L \lambda V_R^\dagger$  is written in terms of the mass basis. The Yukawa interaction matrix  $Y$  can also be written in terms of the mass basis:

$$Y = V_L \frac{\lambda}{\sqrt{2}} V_R^\dagger = \frac{\lambda^m}{\sqrt{2}} = \frac{m}{v} = \begin{pmatrix} Y_{ee} & 0 & 0 \\ 0 & Y_{\mu\mu} & 0 \\ 0 & 0 & Y_{\tau\tau} \end{pmatrix} \quad (2.23)$$

which is also diagonal. Thus, there are no LFV Yukawa interactions in the SM.

Dimension-6 operators: If the Higgs field is the only field that causes electroweak symmetry breaking and if the particle spectrum is consisting only out of SM particles, including the Higgs boson, up to some energy scale  $\Lambda \gg 200\text{GeV}$ , additional heavy fields can be integrated out, leading to an effective field theory. Terms including higher-dimension operators, can be added to the SM Lagrangian, which has dimension four. An effective Lagrangian can be written including dimension-6 operators

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \sum_{nij} \frac{\alpha_n^{ij}}{\Lambda^2} O_n^{ij} \quad (2.24)$$

where  $i, j$  denote flavor indices,  $n$  runs over the number of independent operators,  $\Lambda$  is the scale of new physics, and  $\alpha_n^{ij}$  are coefficients [13]. This effective Lagrangian describes the effects of new physics at the electroweak scale in a model-independent way. A dimension-5 operator can generate neutrino masses. However, this operator does not generate LFV interactions for the Higgs boson. The dimension-6 operator  $O_{L\phi}^{ij} = (\Phi^\dagger \Phi)(\bar{L}_i \ell_{Rj} \Phi)$  is a Yukawa-type operator that only generates LFV interactions of the Higgs boson. It also contributes to the fermion mass matrices. These

contributions have to be included in the fermion mass matrices. Then, the mass matrices are diagonalized. Finally, the interaction term of the Lagrangian is written in terms of the mass basis. Including the Yukawa-type operator to the Lagrangian leads to the following additional term:

$$\Delta\mathcal{L} = -\frac{\lambda'_{ij}}{\Lambda^2}(\Phi^\dagger\Phi)(\bar{\ell}_L^i\ell_R^j\Phi) + h.c. \quad (2.25)$$

where  $\lambda'_{ij}$  is the coefficient  $\alpha_n^{ij}$  of the dimension-6 operator  $O_{L\phi}^{ij}$ .

Electroweak symmetry breaking leads to the following Lagrangian terms:

$$\begin{aligned} \Delta\mathcal{L} &= -\frac{\lambda'_{ij}}{\Lambda^2}\left(\frac{v+H}{\sqrt{2}}\right)^2\bar{\ell}_L^i\ell_R^j\left(\frac{v+H}{\sqrt{2}}\right) \\ &= -\frac{\lambda'_{ij}}{\Lambda^2}\frac{1}{(\sqrt{2})^3}(v^3 + 3v^2H + 3vH^2 + H^3)\bar{\ell}_L^i\ell_R^j \\ &= -\frac{1}{\sqrt{2}}\frac{\lambda'_{ij}}{2\Lambda^2}(v^3\overline{\ell_L^i\ell_R^j} + 3v^2\overline{\ell_L^i\ell_R^j}H + 3v\bar{\ell}_L^i\ell_R^jH^2 + \bar{\ell}_L^i\ell_R^jH^3) \end{aligned} \quad (2.26)$$

where term proportional to  $\overline{\ell_L^i\ell_R^j}$  leads to an additional mass term of the charged leptons and the term proportional to  $\overline{\ell_L^i\ell_R^j}H$  adds a further Yukawa-interaction term. Thus, the mass term of the effective Lagrangian  $\mathcal{L}_{\text{mass}}$  is the sum of the SM mass term and the additional contribution:

$$\begin{aligned} \mathcal{L}_{\text{mass}} &= \mathcal{L}_{\text{mass,SM}} + \Delta\mathcal{L}_{\text{mass}} \\ &= -\left(\frac{\lambda_{ij}}{\sqrt{2}}v + \frac{1}{\sqrt{2}}\frac{\lambda'_{ij}}{2\Lambda^2}v^3\right)\bar{\ell}_L^i\ell_R^j \\ &= -\frac{1}{\sqrt{2}}\left(\lambda_{ij} + \frac{v^2}{2\Lambda^2}\lambda'_{ij}\right)v\bar{\ell}_L^i\ell_R^j \end{aligned} \quad (2.27)$$

The term of the effective Lagrangian which describes the Yukawa interactions,

$\mathcal{L}_Y$ , is given by:

$$\begin{aligned}\mathcal{L}_Y &= \mathcal{L}_{Y,\text{SM}} + \Delta\mathcal{L}_Y \\ &= -\left(\frac{\lambda_{ij}}{\sqrt{2}} + \frac{1}{\sqrt{2}}\frac{\lambda'_{ij}}{2\Lambda^2}3v^2\right)\bar{\ell}_L^i\ell_R^j H \\ &= -\frac{1}{\sqrt{2}}\left(\lambda_{ij} + 3\frac{v^2}{2\Lambda^2}\lambda'_{ij}\right)\bar{\ell}_L^i\ell_R^j H\end{aligned}\tag{2.28}$$

The mass matrix of the charged leptons can be diagonalized as for the SM using other matrices  $V_L, V_R$ :

$$\sqrt{2}m = V_L[\lambda + \frac{v^2}{2\Lambda^2}\lambda']V_R^\dagger v\tag{2.29}$$

Then, the Yukawa interaction matrix can be written in terms of the mass basis:

$$\begin{aligned}\sqrt{2}Y &= V_L[\lambda + 3\frac{v^2}{2\Lambda^2}\lambda']V_R^\dagger \\ &= \frac{\sqrt{2}m}{v} + V_L2\frac{v^2}{2\Lambda^2}\lambda'V_R^\dagger \\ &= \frac{\sqrt{2}m}{v} + \frac{v^2}{\Lambda^2}V_L\lambda'V_R^\dagger \\ &= \frac{\sqrt{2}m}{v} + \frac{v^2}{\Lambda^2}\hat{\lambda}\end{aligned}\tag{2.30}$$

where  $m$  is the diagonal mass matrix and  $\hat{\lambda} = V_L\lambda'V_R^\dagger$ . The first term is the Yukawa term of the SM, while the second term can introduce non-diagonal terms as  $\hat{\lambda}$  is in principle an arbitrary non-diagonal matrix, leading to Yukawa couplings  $Y_{ij}$  given by:

$$Y_{ij} = \frac{m_i}{v}\delta_{ij} + \frac{v^2}{\sqrt{2}\Lambda^2}\hat{\lambda}_{ij}\tag{2.31}$$

with possible LFV Yukawa interactions.

**LFV Yukawa couplings:** The inclusion of LFV Yukawa interactions was discussed by introducing a dimension-6 operator to the Lagrangian. There are several possible

additional terms to the Lagrangian due to new physics, which prevent the diagonalization of the mass matrix and the Yukawa matrix at the same time. Another example is theories with more than one Higgs doublet, where the scalar fields can mix, leading to a neutral Higgs boson within general LFV Yukawa interactions. Thus, the Yukawa term of the Lagrangian for the charged leptons in the mass basis is after the electroweak symmetry breaking in general given by:

$$\mathcal{L}_{\text{Yukawa}} = -m_i \bar{\ell}_L^i f_R^i - Y_{ij} \bar{\ell}_L^i \ell_R^j H + \text{h.c.} \quad (2.32)$$

where  $Y_{ij}$  are the entries of the Yukawa matrix:

$$Y = \begin{pmatrix} Y_{ee} & Y_{e\mu} & Y_{e\tau} \\ Y_{\mu e} & Y_{\mu\mu} & Y_{\mu\tau} \\ Y_{\tau e} & Y_{\tau\mu} & Y_{\tau\tau} \end{pmatrix} \quad (2.33)$$

which can have LFV Yukawa couplings  $Y_{ij}$ . LFV Yukawa couplings allow LFV decays of the Higgs boson. The decay width of LFV Higgs boson decays is

$$\Gamma(H \rightarrow \ell_i \ell_j) = \frac{m_H}{8\pi} (|Y_{ji}|^2 + |Y_{ij}|^2) \quad (2.34)$$

Using the decay width of the LFV decay, the branching fraction of the LFV Higgs boson decay can be derived when contributions of the other LFV decays are negligible

$$B(H \rightarrow \ell_i \ell_j) = \frac{\Gamma(H \rightarrow \ell_i \ell_j)}{\Gamma(H \rightarrow \ell_i \ell_j) + \Gamma_{\text{SM}}} \quad (2.35)$$

where  $\Gamma_{\text{SM}}$  is the total width of the Higgs boson in the SM. Non-negligible contributions to the total decay width from other LFV Higgs boson decays would have to be added to  $\Gamma_{\text{SM}}$  for estimating the branching fractions.

In the SM, the Higgs boson couples to all particles according to their mass. In the presence of new physics, LFV Yukawa couplings might be related closely to the

fermion mass matrices, reflecting the observed fermion mass hierarchy [14]. Then, the LFV couplings have a hierarchical structure given by  $\Delta_{ij}\sqrt{m_i m_j}$ , where  $i, j$  are the generation indices, and  $\Delta_{ij}$  is related to the mass mixing. Mass corrections to the lepton masses coming from mass mixing are assumed to be small in the so-called natural assumption. SM Yukawa couplings are taken for the flavor diagonal couplings. Thus, LFV branching fractions can be expressed in terms of SM ones:

$$B_{\text{nat}}(H \rightarrow ij) = B_{\text{SM}}(H \rightarrow ii) \cdot \frac{m_j}{m_i} \quad (2.36)$$

with  $\Gamma_{\text{tot}} \approx \Gamma_{\text{SM}}$ ,  $B_{\text{nat}}(H \rightarrow ij)$  being the natural branching fraction and  $B_{\text{SM}}(H \rightarrow ij)$  being the branching fraction for the SM Higgs boson decay to two leptons  $i$ . The LFV branching fractions are expected to be of the same size or smaller than  $B_{\text{nat}}(H \rightarrow ij)$ . Larger total widths lead to a smaller branching fraction. A naturalness assumption referred to as theoretical naturalness limit can also be derived for the Yukawa couplings  $Y_{ij}$ :

$$|Y_{ji} Y_{ij}| \lesssim \frac{m_i m_j}{v^2} \quad (2.37)$$

with  $v$  being the vacuum expectation value of the Higgs boson.

### 2.6.2 Constraints from low-energy measurements

Constraints on LFV decays of the Higgs boson have been derived using low-energy measurements and certain assumptions on the value of flavor diagonal Yukawa couplings. In the following, constraints from several low-energy measurements on LFV Yukawa-couplings are discussed. Diagrams that can contribute to the low-energy measurements are shown. The results of the measurements and assumptions used to derive constraints on the LFV Yukawa-couplings are given. Finally, an overview of all constraints is given.

Constraints from LFV decays  $\ell_i \rightarrow \ell_j \gamma$ : LFV Yukawa couplings also contribute to LFV decays of type  $\ell_i \rightarrow \ell_j \gamma$ , where  $i, j$  are flavor indices with  $i \neq j$ . One-loop diagrams and two-loop diagrams of this process are shown in Figure [? ]. Constraints on the LFV Yukawa couplings of the Higgs boson with  $m_H = 125$  GeV can be derived assuming SM values for  $Y_{\tau\tau}$ ,  $Y_{\mu\mu}$ , and  $Y_{tt}$ .

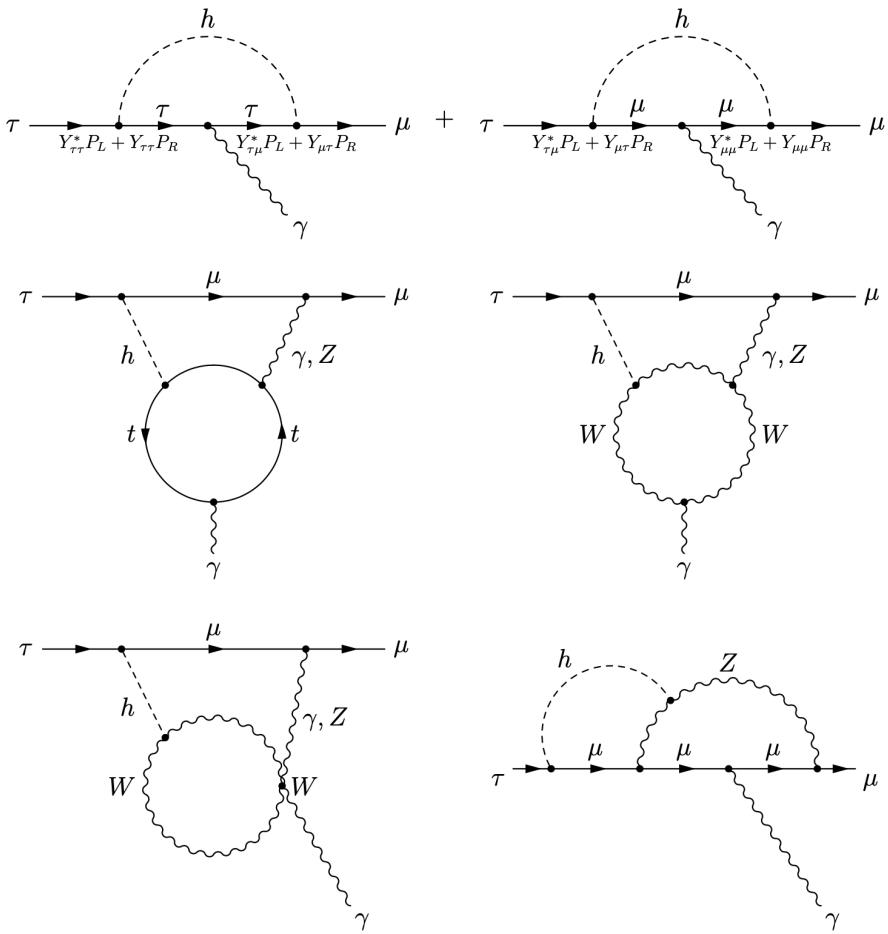


Figure 2.3. Diagrams contributing to the flavor violating decay  $\tau \rightarrow \mu \gamma$ , mediated by a Higgs boson with flavor violating Yukawa couplings.

Experimental upper bounds on the branching fractions  $B(\ell_i \rightarrow \ell_j \gamma)$  used for deriving constraints are summarised in Table [? ]. Constraints of type  $\sqrt{|Y_{ij}|^2 + |Y_{ji}|^2}$  can be derived. A further constraint on  $(|Y_{\tau\mu} Y_{e\tau}|^2 + |Y_{\mu\tau} Y_{\tau e}|^2)^{1/4}$  can be obtained from  $\mu \rightarrow e\gamma$  by setting  $Y_{\mu e}$  and  $Y_{e\mu}$  to zero.

Decay	$B(\tau \rightarrow \mu\gamma)$	$B(\tau \rightarrow e\gamma)$	$B(\mu \rightarrow e\gamma)$	(2.38)
Experimental bound	$< 4.4 \times 10^{-8}$	$< 3.3 \times 10^{-8}$	$< 2.4 \times 10^{-12}$	

Constraints from LFV decays  $\ell_i \rightarrow 3\ell_j$ : LFV decays of type  $\ell_i \rightarrow 3\ell_j$  can constrain the LFV Yukawa couplings. Such decays have been searched for, and the experimental upper bounds on the branching fractions are summarised in Table 3.2. The previous assumptions on the mass of the Higgs boson and flavor diagonal Yukawa couplings were used to derive constraints on the LFV Yukawa couplings. Z boson contribution to the LFV decay  $\tau \rightarrow \mu$ , described in are neglected, as it turns out that their contributions are negligible [15].

Constraints from muonium-antimuonium oscillations: The bound state  $\mu^+ e^-$  is called muonium ( $M$ ) and can oscillate to the  $e^+ \mu^-$  bound state called antimuonium ( $\bar{M}$ ). An upper constraint on the conversion probability  $P(M \rightarrow \bar{M}) < 8.3 \times 10^{-11}$  [16] was derived by the muonium-antimuonium conversion spectrometer (MACS) experiment at PSI. The time-integrated conversion probability depends on the mass splitting between the two mass eigenstates of the mixed  $M - \bar{M}$  system, which depends on  $|Y_{\mu e} + Y_{e\mu}^*|$  [6].

A correction factor  $S_B = 0.35$ , accounting for the splitting of the muonium states in the magnetic field of the detector, was applied on the conversion probability  $P(M \rightarrow \bar{M}) < 8.3 \times 10^{-11}/S_B$  when deriving the constraints on the LFV Yukawa couplings, leading to a weaker upper bound on the conversion probability. The correction factor depends on the conversion operator, and the smallest value was

used to derive the constraints.

Constraints from magnetic and electric dipole moments: The muon's experimental value  $g_\mu - 2$  is more than three standard deviations above the SM prediction. Neglecting terms suppressed by  $m_\mu/m_\tau$  or  $m_\tau/m_H$ , the LFV contribution to  $g_\mu - 2$  due to the one-loop diagram has the form

$$a_\mu \equiv \frac{g_\mu - 2}{2} \propto \text{Re}(Y_{\mu\tau} Y_{\tau\mu}) \quad (2.39)$$

with the discrepancy between measurement and SM prediction having the size

$$\Delta a_\mu \equiv a_\mu^{\text{exp}} - a_\mu^{\text{SM}} = (2.87 \pm 0.63 \pm 0.49) \times 10^{-9} \quad (2.40)$$

where  $a_\mu^{\text{exp}}$  and  $a_\mu^{\text{SM}}$  are the measured and predicted value, respectively. The one-loop diagram can contribute to an electric dipole moment (EDM) of the muon if the LFV Yukawa couplings are complex. If the terms suppressed by  $m_\mu/m_\tau$  or  $m_\tau/m_H$  are neglected, the dependence on the LFV Yukawa couplings of the electric dipole moment  $d_\mu$  is

$$d_\mu \propto -\text{Im}(Y_{\mu\tau} Y_{\tau\mu}) \quad (2.41)$$

Constraints from  $\mu \rightarrow e$  conversion in nuclei LFV Yukawa couplings can contribute to  $\mu \rightarrow e$  conversion in nuclei via tree-level exchange of a Higgs boson and one-loop diagrams with a Higgs boson and a photon exchange. Two-loop contributions are larger than the one-loop ones, as they are only suppressed by the weak gauge coupling or  $Y_{\tau\tau}$ . Thus, they were taken into account for deriving the constraints on the LFV Yukawa couplings assuming SM values for the flavor diagonal Yukawa couplings, as described before. These Yukawa couplings have not been observed at the LHC up to today. Using these assumptions, the LFV decays of type  $\ell_i \rightarrow \ell_j \gamma$  give the strongest constraints.

Constraints on the LFV Yukawa couplings from low-energy measurements are summarised in Table 2.1 and Figure 2.4. They were, in most cases, derived using certain assumptions regarding the flavor diagonal Yukawa couplings. These measurements constrain the branching fractions of LFV Higgs boson decays to  $\mu\tau$  or  $e\tau$  to be  $\lesssim \mathcal{O}(10^{-1})$ , while the constraint for the decay to  $e\mu$  is stronger and  $\lesssim \mathcal{O}(10^{-8})$ . Constraints on the branching fractions were derived under the assumption that only one of them contributes to the Higgs boson's total width, additionally to the SM total width.

### 2.6.3 Lepton flavor violating decays of the Higgs boson at the LHC

A Higgs boson was discovered in 2012, whose properties are compatible with the SM Higgs boson, within experimental uncertainties. Its mass has been measured by the ATLAS and CMS collaborations and found to be  $125.09 \pm 0.24$ GeV. The nature of the discovered Higgs boson has to be investigated by measuring its properties, including its couplings to charged leptons. In the following, SM production-mechanisms of the Higgs boson are discussed, relevant at the LHC. Possible LFV decays of the Higgs boson are presented. The size of the branching fractions on SM Higgs boson decays and the natural LFV branching fractions, Section 3.1, is discussed. Status is given about the upper limits on these decays' branching fraction, which was in July 2014.

Higgs boson production mechanisms: Gluon fusion (GF) and vector-boson fusion (VBF) are the dominant Higgs boson production-mechanisms [17]. Other production mechanisms have a lower cross-section; therefore, they can be neglected for this search.

LFV Yukawa couplings: LFV Yukawa couplings can introduce decays of the Higgs boson to  $\mu\tau$ ,  $e\tau$ , or  $e\mu$ . Decays of the Higgs boson to  $\mu\tau$  or  $e\tau$  can be split into a leptonic channel ( $\tau_e/\tau_\mu$ ) and a hadronic channel ( $\tau_h$ ). The branching fractions of tau

TABLE 2.1

CONSTRAINTS ON LFV YUKAWA COUPLINGS FROM  
LOW-ENERGY MEASUREMENTS [6].

Channel	Coupling	Bound
$\mu \rightarrow e\gamma$	$\sqrt{ Y_{\mu e} ^2 +  Y_{e\mu} ^2}$	$< 3.6 \times 10^{-6}$
$\mu \rightarrow 3e$	$\sqrt{ Y_{\mu e} ^2 +  Y_{e\mu} ^2}$	$\lesssim 3.1 \times 10^{-5}$
electron $g - 2$	$\text{Re}(Y_{e\mu} Y_{\mu e})$	$-0.019 \dots 0.026$
electron EDM	$-\text{Im}(Y_{e\mu} Y_{\mu e})-$	$< 9.8 \times 10^{-8}$
$\mu \rightarrow e$ conversion	$\sqrt{ Y_{\mu e} ^2 +  Y_{e\mu} ^2}$	$< 1.2 \times 10^{-5}$
$M - \bar{M}$ oscillations	$ Y_{\mu e} + Y_{e\mu}^* $	$< 0.079$
$\tau \rightarrow e\gamma$	$\sqrt{ Y_{\tau e} ^2 +  Y_{e\tau} ^2}$	$< 0.014$
$\tau \rightarrow 3e$	$\sqrt{ Y_{\tau e} ^2 +  Y_{e\tau} ^2}$	$\leq 0.12$
electron $g - 2$	$\text{Re}(Y_{e\tau} Y_{\tau e})$	$-2.1 \dots 2.9 \times 10^{-3}$
electron EDM	$-\text{Im}(Y_{e\tau} Y_{\tau e})-$	$< 1.1 \times 10^{-8}$
$\tau \rightarrow \mu\gamma$	$\sqrt{ Y_{\tau\mu} ^2 +  Y_{\mu\tau} ^2}$	$< 0.016$
$\tau \rightarrow 3\mu$	$\sqrt{ Y_{\tau\mu} ^2 +  Y_{\mu\tau} ^2}$	$\lesssim 0.25$
muon $g - 2$	$\text{Re}(Y_{\mu\tau} Y_{\tau\mu})$	$(2.7 \pm 0.75) \times 10^{-3}$
muon EDM	$-\text{Im}(Y_{\mu\tau} Y_{\tau\mu})-$	$-0.8 \dots 1.0$
$\mu \rightarrow e\gamma$	$( Y_{\tau\mu} Y_{e\tau} ^2 +  Y_{\mu\tau} Y_{\tau e} ^2)^{1/4}$	$< 3.4 \times 10^{-4}$

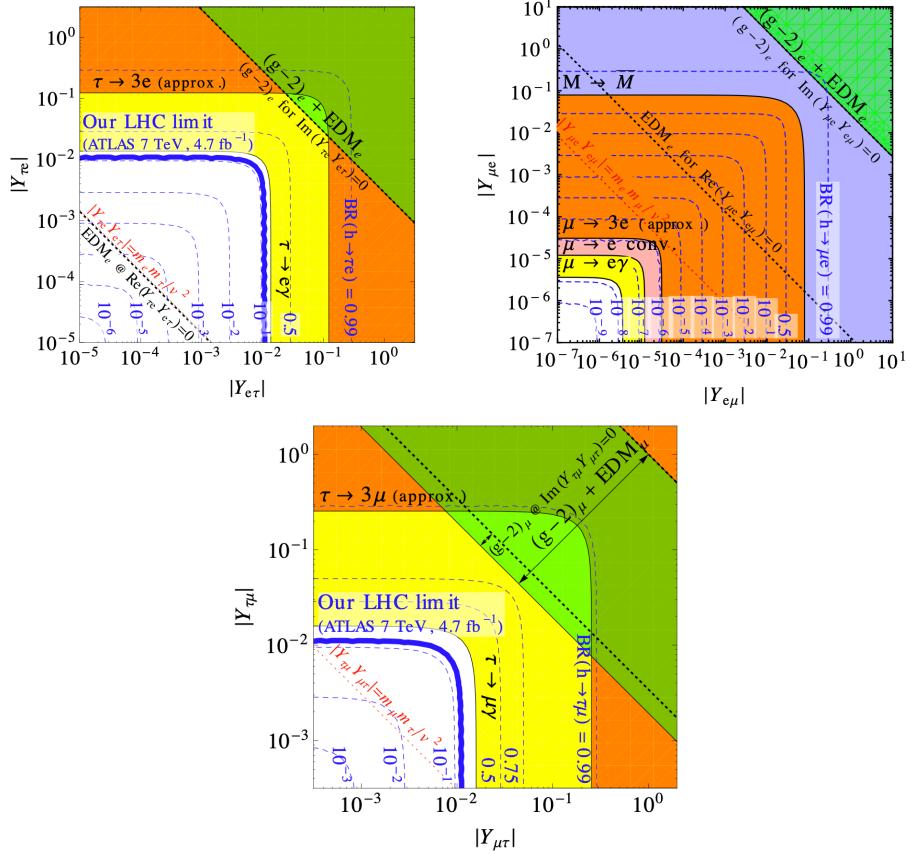


Figure 2.4. Constraints on the LFV Yukawa couplings for a 125 GeV Higgs boson. The diagonal Yukawa couplings are approximated by their SM values. Shaded regions show the constraints from low-energy measurements. The thin red dotted lines show the theoretical naturalness limits  $|Y_{ji} Y_{ij}| \lesssim \frac{m_i m_j}{v^2}$ .

to a muon or an electron are  $\sim 20\%$ , while  $\sim 65\%$  of the taus decay hadronically.

Taus have a short life-time of  $\tau = (290.3 \pm 0.5) \times 10^{-15}\text{s}$  and a decay length of  $c\tau = 87.03\mu\text{m}$  leading to a secondary vertex, which can be resolved within the CMS pixel tracking-detector. Muons, electrons, and hadronic taus can be reconstructed with the CMS detector. Neutrinos from tau decays lead to missing transverse energy  $\cancel{E}_T$ . Feynman diagrams for  $H \rightarrow \mu\tau_e$  (a),  $H \rightarrow e\tau_\mu$  (c), and  $H \rightarrow e\mu$  (e) lead to a signature with an isolated electron and an isolated muon of opposite charge in the CMS detector.

The last decay mode can be distinguished from the other Higgs boson decays by the absence of  $E_T$  due to neutrinos. Decays to  $\mu\tau_e$  or  $e\tau_\mu$  can be distinguished by the small angle between the lepton's direction from the tau decay and the direction of the missing transverse momentum  $\sim E_T$ , due to the neutrinos, for high momenta of the tau. Feynman diagrams for  $H \rightarrow \mu\tau_h$  (b), or  $H \rightarrow e\tau_h$  (d) have signatures with a hadronic tau and an isolated electron or muon, respectively. Their signature includes  $E_T$  as well due to the neutrino of the tau decay.

Tau decays to a lepton with the same flavor as the lepton of the Higgs boson decay. The channels  $e\tau_e$  and  $\mu\tau_\mu$ , are not considered because their signatures are very similar to Z-boson decays, leading to an overwhelming background of di-leptons for such processes.

**Branching fractions:** In the SM, the Higgs boson should couple to all particles according to their mass. The branching fraction for Higgs boson decays to taus is in the order of 6%, and the branching fraction of the Higgs boson to muons is in the order of 0.02%.

The LFV branching fractions are expected to be of the same size or smaller than  $\mathcal{B}_{\text{nat}}(H \rightarrow ij)$  using the natural assumption. Furthermore, LFV branching fractions  $\mathcal{B}_{\text{nat}}(H \rightarrow ij)$  can be expressed in terms of SM ones. Then, the LFV Higgs boson decay to  $\mu\tau$  should have a natural branching fraction in the order of 0.5%, which is

between decays to taus muons. LFV decays of the Higgs boson to  $e\tau$  would have a smaller natural branching fraction in the order of  $0.001\% = 10^{-5}$ . Decays of the Higgs boson to  $e\mu$  would even have a smaller natural branching fraction of the order  $10^{-6}$ .

Upper limits on LFV branching fractions: Low-energy measurements, described in Section 3.2, indirectly constrain branching fractions of the LFV Higgs boson decays to  $\mu\tau$  or  $e\tau$  to be smaller than 10%, which is larger than their natural branching fractions. The constraint for decays to  $e\mu$  is in the order of  $10^{-8}m$ , which is smaller than the natural branching fraction. These constraints were derived, assuming flavor changing neutral currents to be dominated by Higgs boson contributions; therefore, LFV effects could be canceled by other new physics effects leading to weaker limits. Direct searches for LFV decays of the Higgs boson are independent of assumptions on other new physics models.

A search for LFV decays of the Higgs boson to  $\mu\tau$  has been done using the CMS data sample collected in proton-proton collisions at a center-of-mass energy  $\sqrt{s} = 8 \text{ TeV}$ , which was the first direct search for LFV decays of the observed Higgs boson. Constraints on the branching fraction  $\mathcal{B}(H \rightarrow \mu\tau) < 1.51\%$  could be set. Figure 2.5 shows the expected and observed 95% CL upper limits for each category of the search and their combination. The expected limit for the combination is very close to the natural expectation, assuming similar LFV couplings of the Higgs boson for the SM Higgs boson decays. This search has improved the indirect limit on the branching fraction  $\mathcal{B}(H \rightarrow \mu\tau)$  by one order of magnitude. The search was performed in the  $\mu\tau_e$  channel, which has the same signature as the  $e\tau_\mu$  channel of a corresponding search for decays of the Higgs boson to  $e\tau$ , but different kinematics of the leptons. Thus, it is expected that a search for LFV decays of the Higgs boson to  $e\tau$  in the  $e\tau_\mu$  channel is sensitive for branching fractions up to  $O(1\%)$ , which would also improve upon indirect limits by one order of magnitude. The status presented in this section

was when the search for LFV decays of the Higgs boson to  $e\tau_\mu$ , described in this thesis, was started. The ATLAS and the CMS collaborations have published other results from searches of LFV decays of the Higgs boson.

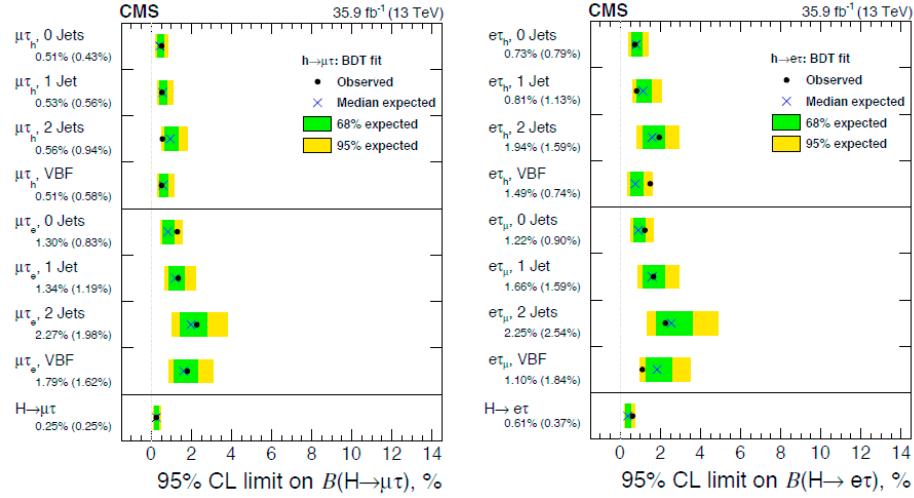


Figure 2.5. Expected and observed 95% CL upper limits by category of a search for LFV  $H \rightarrow \mu\tau$  and  $H \rightarrow e\tau$  decays with 2016 dataset [18].

## CHAPTER 3

### LHC AND THE CMS DETECTOR

#### 3.1 Introduction

The prime goals of the LHC and the Compact Muon Solenoid (CMS) [19] experiment are exploring the physics at the TeV scale and studying the mechanism of electroweak symmetry breaking, including studying the SM Higgs boson properties along with searches for new particles predicted by beyond the SM physics. The physics program's other main interests include studies of the SM top quark properties, electroweak physics, and physics of hadrons containing a charm or bottom quark. Heavy-ion collisions address the physics of strongly interacting matter and the quark-gluon plasma. Hadrons consist of quark and gluons; therefore, two colliding partons' initial energy is not known. In contrast, the collision's energy is known at lepton colliders, where each particle has the same energy. Thus, hadron colliders can explore a wide range of collision energies, while lepton colliders are well suited for precision measurements.

#### 3.2 The Large Hadron Collider

The LHC is a hadron accelerator located at CERN. The design was intended to collide proton beams with a beam energy of 7 TeV leading to a center-of-mass energy  $\sqrt{s}$  of 14 TeV and reach a luminosity of  $10^{34}\text{cm}^{-2}\text{s}^{-1}$ . Lead (Pb) ions can be accelerated up to an energy of 2.8 TeV per nucleon and reach a luminosity of  $10^{27}\text{cm}^{-2}\text{s}^{-1}$ . LHC is the most powerful tool for particle physics research that is

currently available. The 26.7 km tunnel constructed between 1984 and 1989 for the Large Electron-Positron collider (LEP) [20] was reused to install the LHC. There are eight arcs and eight straight sections lying 45-170 m below the surface. Unlike particle-antiparticle colliders that can use a single ring for both beams, the LHC uses two rings with counter-rotating beams. There is less synchrotron radiation owing to the heavier particles being collided at the LHC.

The accelerator complex acts as an injector of the protons and heavy ions. Protons are obtained from hydrogen gas after the electrons are stripped off, and they enter the Linear accelerator 2 (Linac2) [21], where they are accelerated to 50 MeV. They are further accelerated in the Proton Synchrotron Booster (PSB) [22] to 1.4 GeV. They are then injected in the Proton Synchrotron (PS) [23], where they are accelerated to 25 GeV. A bunch train is produced within the PS before extraction. The protons are then accelerated in the Super Proton Synchrotron (SPS) [24] to 450 GeV before injecting in the LHC.

The four interaction points are equipped with particle detectors: the CMS experiment, A Toroidal LHC ApparatuS (ATLAS) experiment [25], A Large Ion Collider Experiment (ALICE) [26], and a Large Hadron Collider beauty (LHCb) [27] experiment. Two further smaller experiments, a Total, Elastic, and diffractive cross section Measurement (TOTEM) [28] and the Large Hadron Collider forward (LHCf) [29] experiment, are located near the CMS interaction point and near the ATLAS interaction point, respectively.

The ATLAS and CMS experiments have both multi-purpose detectors installed. The detectors were build to detect particles from proton-proton (p-p) or heavy-ion (Pb-Pb) collisions. One of the main tasks currently is to study the production and decay of the discovered Higgs boson [9], disentangle its properties, and check if it is the SM Higgs boson or a Higgs boson of an extension of the SM. The other tasks involve high precision tests of QCD, electroweak interactions, and heavy flavor physics.

Precision measurements of production, the couplings, and the spin of the top quark are also pursued. Several searches for supersymmetric particles and Dark Matter are also ongoing.

During the LHC Run II data-taking, a bunch spacing of 25 ns was used, and proton-proton collision data were collected at  $\sqrt{s} = 13$  TeV in 2016, 2017, and 2018. The integrated luminosity delivered to CMS as a function of time is shown in Figure 3.1 for each proton-proton collision data-taking period. CMS does not record the whole delivered data, and only part of the recorded that is considered good is used for physics analysis.

TABLE 3.1

INTEGRATED LUMINOSITY CONSIDERED FOR PHYSICS  
ANALYSIS AT THE CMS EXPERIMENT DURING RUN II

Year	Integrated luminosity
2016	$35.9\text{fb}^{-1}$
2017	$41.5\text{fb}^{-1}$
2018	$59.3\text{fb}^{-1}$

### 3.3 The Compact Muon Solenoid experiment

The layout of the CMS detector is shown in Figure 3.2. As the name “Compact Muon Solenoid” indicates, a superconducting solenoid is the heart of CMS. Considerable bending power for the momentum measurement of charged particles within a

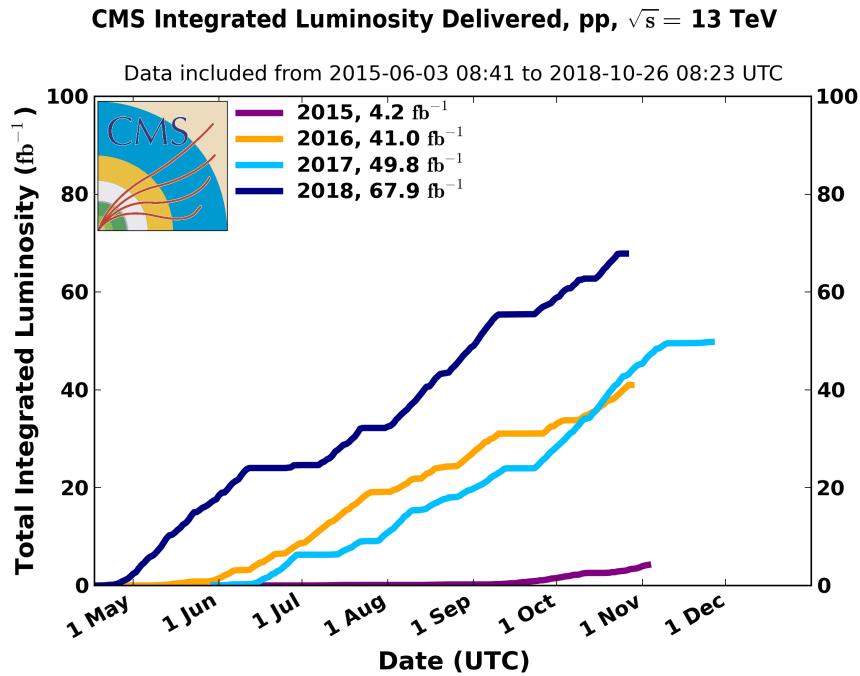


Figure 3.1. Cumulative luminosity as a function of time delivered to CMS during stable beams for proton-proton collisions. The luminosity is shown for 2015 (purple), 2016 (orange), 2017 (light blue), and 2018 (dark blue).

[30]

compact design is achieved using a high magnetic field of 3.8 T. The internal part of the magnetic coil is large enough to accommodate the inner tracking system and the calorimetry. The inner tracking system is composed of a pixel detector close to the interaction region and a silicon strip tracker. The pixel detector can resolve individual vertices and distinguish between vertices from the primary interaction and secondary vertices from the decay of the primary interaction particles. Trajectories are precisely measured with the high granular silicon strip tracker, which can deal with high charged particle multiplicities.

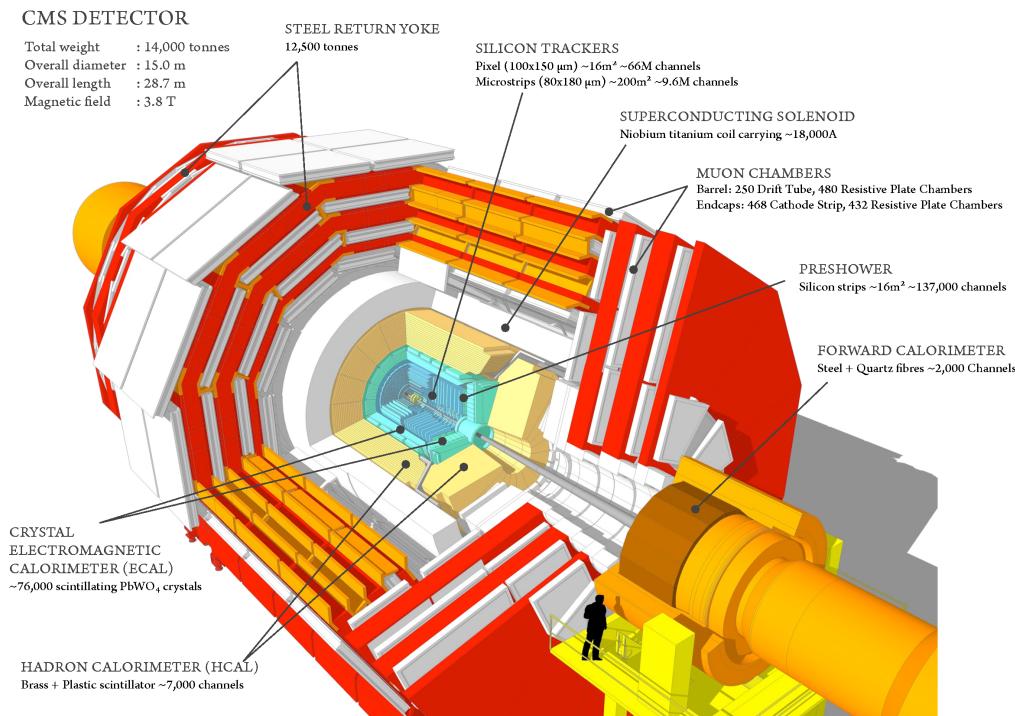


Figure 3.2. Schematic view of the CMS detector.

The energy of the particles is measured in calorimeters. Calorimeter material ini-

tiates electromagnetic (EM) or hadronic showers. For electromagnetic interactions, the characteristic interaction length is the radiation length  $X_0$ , while the characteristic interaction length for hadronic showers is the nuclear interaction length  $\lambda_I$ . The entire volume is sensitive in homogeneous calorimeters, while sampling calorimeters consist of metallic absorber sandwiched or threaded with an active material that generates the signal. CMS has an EM calorimeter (ECAL) made of lead tungstate  $PbWO_4$  in front of a brass/scintillator sampling hadron calorimeter (HCAL). An additional layer of scintillators is outside the coil. The magnet is used as an absorber material. This iron/quartz-fiber calorimeter is referred to as the hadron outer (HO) detector. The muon detectors are sandwiched between the layers of the steel return yoke. Their main task is to trigger on muons and to identify the muons with good momentum resolution.

### 3.3.1 The coordinate system of CMS

CMS uses a right-handed Cartesian coordinate system with its origin at the detector's center, the nominal interaction point. The x-axis points towards the LHC center, while the y-axis points upward towards the surface and perpendicular to the LHC plane. Thus, the z-axis points along the anticlockwise beam-direction. Two angles are defined, where the azimuthal angle  $\phi$  is measured from the x-axis in the x-y plane, and the polar angle  $\theta$  is measured from the z-axis. Pseudorapidity is defined as

$$\eta = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right] \quad (3.1)$$

and describes the angle of a particle relative to the beam axis, illustrated in Figure 3.3. Distances in  $\phi$  and  $\eta$  are denoted  $\Delta\phi$  and  $\Delta\eta$ . These distance measures are used

to define the spatial separation between physics objects by  $\Delta R$  with

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} \quad (3.2)$$

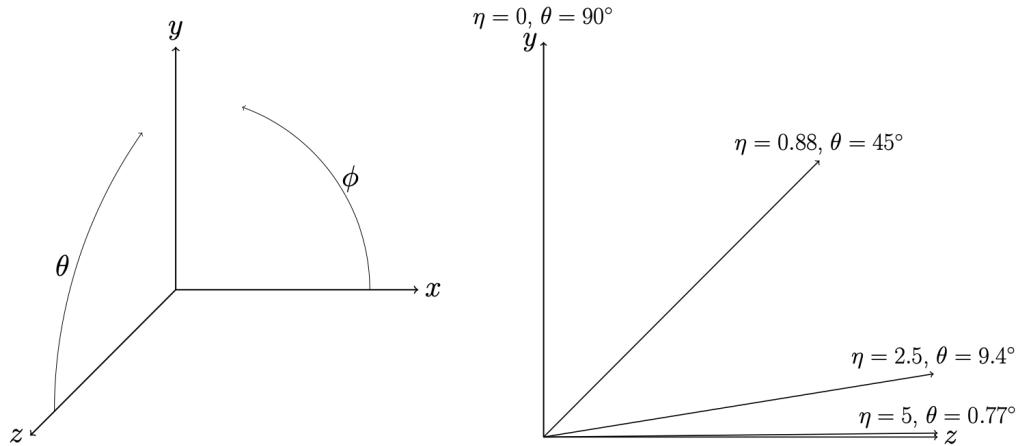


Figure 3.3. Coordinate system convention of CMS (left) and the relation between pseudorapidity  $\eta$  and polar angle  $\theta$  (right)

### 3.3.2 Kinematic quantities

At LHC, the protons carry half of the collision energy  $\sqrt{s}$ . The hard interactions do not take place between the colliding protons but between two of their partons. Each parton carries a fraction of  $x_i$  of the proton momentum. As the parton mass can be neglected with respect to the momentum  $\vec{p}$ , the parton energy is given by its

momentum. In the laboratory frame, the four-momenta of the partons are

$$\mathbf{p}_1 = x_1 \cdot \frac{\sqrt{s}}{2} (1, 0, 0, 1) \quad \mathbf{p}_2 = x_2 \cdot \frac{\sqrt{s}}{2} (1, 0, 0, -1) \quad (3.3)$$

Then, the invariant mass  $M$  of the hard collision is given by

$$\hat{s} \equiv M^2 = (\mathbf{p}_1 + \mathbf{p}_2)^2 = \frac{s}{4} \cdot [(x_1 + x_2)^2 - (x_1 - x_2)^2] = x_1 x_2 s \quad (3.4)$$

where  $\sqrt{\hat{s}}$  denotes the center-of-mass energy of the parton-parton collision. The hard collision products have a total momentum, zero for the x- and y-components but in general non-zero for the z-component. Thus, the hard-collision products' momentum and energy are measured transverse to the beam direction in the x-y plane.  $p_T$  and  $E_T$  denote the transverse momentum and transverse energy respectively, with  $E_T = E \cdot \sin\theta$ . Only weakly interacting particles, such as neutrinos, do not produce a signal in the CMS detector and lead to an imbalance in the observed total transverse momentum in the event, missing transverse energy, denoted as  $\cancel{E}_T$ .

### 3.3.3 Detector requirements

Final states that contain isolated leptons and photons leave clean signatures in the CMS detector. A good muon identification and momentum resolution are needed over a wide range of momenta. The muon charge has to be determined unambiguously. Furthermore, a good momentum resolution and reconstruction efficiency of charged particles are crucial for reconstructing electrons and charged hadrons. Charged-particle momenta are measured using the curvature of their trajectory. Considerable bending power is needed to measure precisely the momentum of particles with large momentum. Additionally, a good electromagnetic energy resolution is needed with wide geometrical coverage. The direction of the photons and correct localization of the primary interaction should be measurable.

Quarks hadronize, and the hadrons can be detected with the CMS detector. Good identification of hadronically decaying taus is needed for the Higgs boson physics. The triggering and identification efficiency of hadronically decaying taus can be improved by a good measurement of the impact parameter of charged-particle tracks and good position measurement of the secondary vertices. This requires pixel detectors close to the interaction region. Hadronic calorimeters with a large hermetic coverage ( $|\eta| < 5$ ) and with a fine lateral segmentation ( $\Delta\eta \times \Delta\phi < 0.1 \times 0.1$ ) are required for a good energy measurement of hadrons and estimation of  $E_T$ .

The collisions are happening at a rate of 40MHz and not all of these events can be stored; therefore, only interesting events have to be selected. The online event selection process, trigger, must reduce the rate to no more than a few hundred events per second. The short time between two bunch crossings, 25 ns, has a major implication on the readout and trigger system design. Multiple proton-proton interactions (pileup) happen in one bunch crossing. The interactions' products overlap and can be wrongly linked. Long response times of detector elements and their electronic signal longer than 25 ns increase this effect. High granularity detectors with good time resolution result in a low occupancy. This requires many detector channels and, therefore, a good synchronization of the electronic detector channels. The large flux of the particles and resulting high radiation levels require radiation-hard detectors and front-end electronics.

### 3.3.4 Magnet

The curvature of the particle trajectory in a magnetic field determines the momenta and sign of charged particles. The most important aspect of muon measurement is the choice of magnetic field configuration. A good momentum resolution of  $\Delta p/p \approx 10\%$  is required for muons with a momentum of 1 TeV. CMS uses a large superconducting solenoid with a magnetic field of 3.8 T. The solenoid is 13 m long and

has an inner diameter of 5.9 m. The superconductor material is Niobium titanium [31]. A high-purity aluminum-stabilized conductor with a four-layer winding is used, which has to withstand outward pressure. The conductor is composed of five coils, and its superconducting wire is cooled using an indirect cooling by thermosyphon.

### 3.3.5 Inner Tracking system

The inner tracking system [32] measures the trajectories of particles up to  $|\eta| < 2.5$ . The particle flux is the highest close to the interaction region. The silicon pixel detector is placed close to the interaction region. This allows for the reconstruction of vertices from heavy flavor hadrons with a b or c quark. Each pixel has a size of  $\approx 100 \times 150\mu\text{m}^2$ . The spatial resolution of the radius and azimuthal angle  $r - \phi$  measurement is about  $10 \mu\text{m}$ , and  $20 \mu\text{m}$  for the z-coordinate measurement. In the intermediate region ( $20 < r < 55 \text{ cm}$ ) and outermost region ( $r > 55 \text{ cm}$ ) silicon microstrips are used with a size of  $10\text{cm} \times 80\mu\text{m}$  (minimum cell size) and a size of  $25\text{cm} \times 180\mu\text{m}$  (maximum cell size), which provide the required granularity. The barrel's tracking volume has a cylinder shape with a length of 5.8 m and a diameter of 2.6 m. The pixel detector consists of three layers at radii of 4, 7, and 11 cm in the barrel.

Additionally, there are ten layers of silicon microstrip detectors. The strip tracker is divided into a Tracker Inner Barrel (TIB) made out of four layers and a Tracker Outer Barrel (TOB) made out of six layers. Each part has two layers which provide a single-point measurement in the  $r - \phi$  and  $r - z$  coordinates. In the TIB, the single-point resolution is  $23 - 34\mu\text{m}$  in the r-direction and  $23\mu\text{m}$  in z, while it is  $35 - 52\mu\text{m}$  in the  $r - \phi$  direction and  $52\mu\text{m}$  in z for the TOB. In the two end caps, there are just two-pixel layers and nine microstrip layers. The strip tracker's endcaps are divided into the Tracker End Cap made of 9 disks and the Tracker Inner Disks (TID), which are made of three small disks, which fill the gap between the TIB and the TEC.

### 3.3.6 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) measures the energy of electrons and photons. A hermetic, homogeneous crystal (ECAL) with a coverage of  $|\eta| < 3$  is used, which has an excellent energy resolution and high granularity; therefore, it also has excellent separation of close clusters [33, 34]. It is used to identify photons, which leave no signal in the inner tracking detector. Lead tungstate ( $PbWO_4$ ) crystals are used, which have a short radiation length of  $X_0 = 0.89$  cm, are radiation hard, and emit blue-green scintillation light with a maximum at 420 nm. The scintillation decay time is short and in the same order as the LHC bunch crossing time. About 80% of the light is emitted within 25 ns. However, the light output depends on the temperature; therefore, a cooling system is needed to preserve the energy resolution. Furthermore, a low light yield requires photodetectors with intrinsic gain that can also operate in a high magnetic field. The crystals allow a compact calorimeter inside the solenoid that is fast, has a fine granularity, and is radiation-resistant. Figure 3.4 illustrates the layout of the CMS ECAL. The calorimeter is divided into a barrel ECAL (EB) which covers the region  $0 < |\eta| < 1.479$  and two ECAL endcaps (EE) which cover the region  $1.479 < |\eta| < 3.0$ . Preshower detectors (ES) are installed in front of each endcap and cover the region  $1.653 < |\eta| < 2.6$ . In the EB, silicon avalanche photodiodes (APDs) are installed to detect the scintillation light. They also respond to temperature changes; therefore, they require a stable temperature.

The barrel section is structured in 36 identical supermodules, 18 in each half barrel, with the crystals being arranged in a grid and covering  $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$ . Vacuum phototriodes (VPTs) are installed in both EE for detecting the scintillation light. The crystals are arranged in units of  $5 \times 5$  crystals termed super crystals (SC). Both the crystals and the SCs, are arranged in a rectangular x-y grid. Neutral pions dominantly decay to two photons. Two closely separated photons can mimic high-energy photons; therefore, a preshower system is installed in front of each

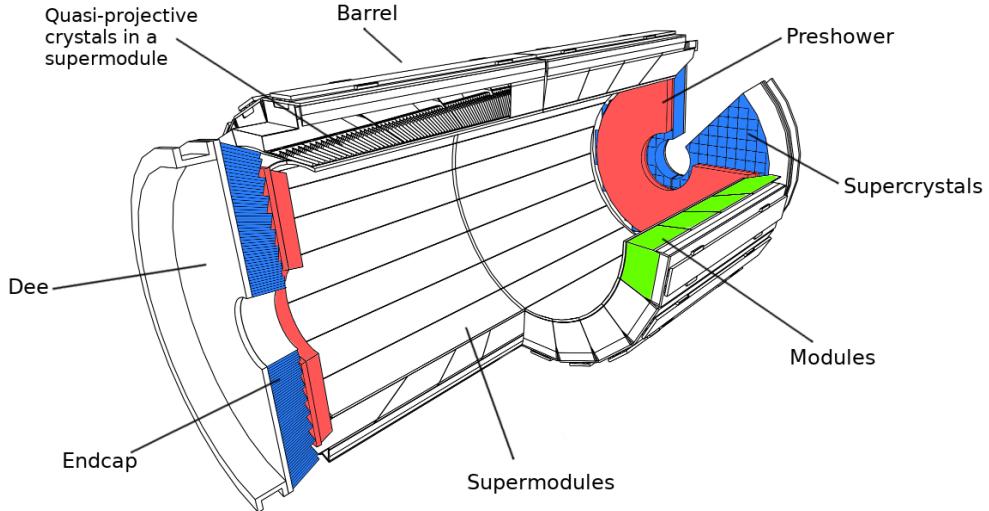


Figure 3.4. Schematic view of the electromagnetic calorimeter.

EE to identify and reject  $\pi^0$  mesons. It also improves the position determination of electrons and photons as it has a higher granularity than the EE. The preshower detector is a sampling calorimeter consisting of two layers of a lead absorber, which initiate the electromagnetic showers from incoming electrons or photons. Each lead radiator is followed by silicon strip sensors measuring the energy deposits and the transverse shower profiles. The strips in both planes of silicon sensors are orthogonal oriented and are placed after a radiation length of  $1 X_0$  and  $2 X_0$  of the lead absorber.

### 3.3.7 Hadron Calorimeter

The energy of hadrons is measured in the hadron calorimeter (HCAL). Most of the HCAL calorimetry is located inside the magnetic coil surrounding the ECAL system. An important requirement on the HCAL design is to minimize the non-Gaussian tails in the energy resolution and to provide a good hermeticity for the determination of  $E_T$ . It is designed to maximize the interaction length of the material within the magnet coil. On the other hand, the amount of space devoted to the active

medium is minimized. The HCAL is a sampling calorimeter which is divided into a hadron barrel (HB) detector, covering the region  $|\eta| < 1.4$  and two hadron endcap (HE) detectors, covering  $1.3 < |\eta| < 3.0$ . Brass is used as an absorber material. Brass has a reasonably short interaction length, is easy to machine, and it is non-magnetic. A tile/fiber technology made of plastic scintillator tiles is used as an active medium. For the innermost and the outermost layer, stainless steel is used for structural strength. The 3.7-mm-thick scintillator plates are sandwiched between the absorber plates. Wavelength-shifting (WLS) fibers, embedded in the scintillator tiles, convert the scintillation light. Then, it is channeled to photodetectors via optical fibers. Multi-channel hybrid photodiodes (HPDs) detect the light. They can operate in high axial magnetic fields. The HB is read out as a single longitudinal sampling with a segmentation  $\Delta\eta \times \Delta\phi = 0.087 \times 0.087 \approx 5^\circ \times 5^\circ$ , which are termed tower. In the HE, the  $\phi$  segmentation is  $5^\circ - 10^\circ$  and the  $\eta$  segmentation is  $0.087 - 0.35$  depending on  $\eta$ .

Additional layers of scintillators are installed outside the coil within the return yoke using the iron as an absorber—the sample energy from hadron showers, which leak through the rear of the calorimeters. The central shower containment and the  $E_T$  resolution of the calorimeter are thus improved. This sampling calorimeter is referred to as Hadron Outer (HO) detector. It is located along with the barrel muon system; therefore, its segmentation closely follows the barrel muon system. It is divided into five sections along with  $\eta$  termed rings. The HO follows the HCAL barrel geometry in  $\eta$  and  $\phi$  and covers the region  $|\eta| < 1.26$ . Two further detectors, which cover the region  $2.9 < |\eta| < 5.0$ , are not shown. They are specialized in measuring energetic forward hadronic showers and ensuring full geometric coverage for the transverse energy measurement. The Hadron Forward (HF) sampling calorimeters use steel as absorber material and quartz fibers as the active medium, which run parallel to the beam. Shower particles in the quartz emit Cherenkov light fibers. This light is

channeled by the fibers to photomultipliers. There are two types of quartz fibers, long ones (1.65 m) and short ones (1.43 m). Neutral components of the hadron showers are preferentially sampled in the HF, leading to narrower and shorter hadronic showers. This is ideally suited for the forward region. The towers in HF have a segmentation of  $0.1 - 0.3$  and a  $\phi$  segmentation of  $10^\circ$ , except in high  $\eta$ -towers, where  $\phi$  segmentation is  $20^\circ$ .

### 3.3.8 Muon System

The Muon system is installed in the magnet return yokes of CMS. Its main tasks are identifying muons, improving the  $p_T$  measurement, and charge-sign determination of high- $p_T$  muons. Additionally, it is used to trigger muons. It is divided into a barrel detector (MB) covering  $|\eta| < 1.2$  and two endcap detectors (ME), covering the region  $0.9 < |\eta| < 2.4$ , defining three regions: the barrel region ( $|\eta| < 0.9$ ), the overlap region ( $0.9 < |\eta| < 1.2$ ), and the endcap region ( $1.2 < |\eta| < 2.4$ ). Three different types of gaseous detectors are used in different radiation environments. Drift tube (DT) chambers are used in the barrel region, where the muon rate, the neutron-induced background rate, and the residual magnetic field are low. In the two endcaps, the muon rate, neutron-induced background, and the magnetic field are high; therefore, cathode strip chambers (CSC) are used in this region. Resistive plate chambers (RPC) are used in both subdetectors covering the region  $|\eta| < 1.6$ . They are operated in avalanche mode to ensure good operation at high rates. The RPCs have double gaps with a width of 2 mm filled with gas. They have a fast response with a good time resolution but a coarser position resolution than the DTs and CSCs; therefore, they are used to identify the correct bunch crossing.

Figure 3.5 gives an overview of the layout of the muon system. The MB is divided into four stations arranged in cylinders interleaved with the iron yoke. Additionally, it is divided into five wheels along the beam direction following the five wheels of the

return yokes. Each chamber consists of 12 layers divided into 3 Super Layers (SL), made out of four DTs layers. Two SL measure the  $r - \phi$  coordinate, while a third SL sandwiched in between them measures the  $z$  coordinate. In the last muon station, there are only two SL to measure the  $r - \phi$  coordinate. Each DT chamber has one or two associated RPCs. The single point resolution of the DTs is  $\approx 200\mu\text{m}$ . In each endcap, the CSCs and RPCs are arranged in four disks. They are divided into three or two concentric rings in the innermost station and other stations. Each CSC measures up to six space coordinates ( $r, \phi, z$ ) and the provided spatial resolution is  $\approx 200\mu\text{m}$ , except for the innermost ring in the first disk, where it is  $\approx 100\mu\text{m}$ . The angular resolution in  $\phi$  is  $\sim 10\text{mrad}$ . Two independent and complementary information sources come from the DTs or CSCs and the RPCs, which feed the trigger system.

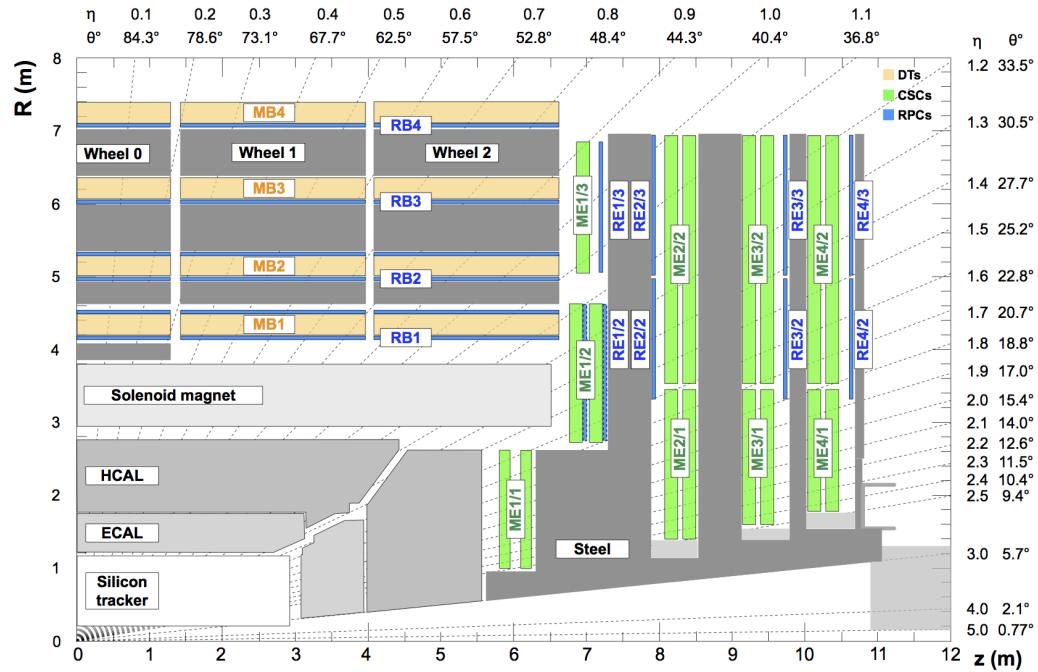


Figure 3.5. Schematic view of the muon detectors.

### 3.3.9 Trigger and data acquisition system

The online event selection process, trigger, reduces the rate to about 1KHz. The CMS experiment's trigger and data acquisition system is divided into four parts, summarised in Figure ?. First, the detector electronics detect signals and pass them to the Level-1 trigger processors. The Level-1 trigger logic is located in a service cavern and selects events of interest. Due to the size of the CMS detector and underground caverns, the transit and time needed for a decision is about  $3.2 \mu s$ . However, the time needed for Level-1 trigger calculations is only  $1 \mu s$ . At this time, the data is held in pipelined memory buffers.

Custom hardware processors form the Level-1 decision to keep or discard data using the calorimetry, muon systems, and correlation information. Trigger primitive objects, such as electrons or muons above a set  $E_{\text{Tor}}$   $p_T$  thresholds are used for the decision taking. Reduced granularity and resolution data are used to form the trigger objects. Sums of  $E_T$  and  $\cancel{E}_T$  are also employed. Only about one out of 1000 interactions are retained. If the data of interaction is kept, the data is transferred after a fixed time interval of about  $3.2 \mu s$  to the front-end readout buffers. The data is further processed and compressed, and placed in dual-port memories. A processor farm is responsible for filtering the events. Each processor runs the same High-Level Trigger (HLT) code, reducing the output rate from a few kHz to a few hundred Hz for mass storage. Using a processor farm for event filtering allows for computer technology evolution and maximizes the flexibility in selecting data and algorithms. Whenever possible, only necessary objects are reconstructed, and only needed information of some subdetectors are used. The idea of partial reconstruction and many virtual trigger levels is the basis for the HLT. First, calorimeter and muon information is used, followed by the use of pixel tracker data. Finally, the full event information is used, including full tracking. The HLT system uses identification and isolation criteria as well as minimal energy or transverse momentum thresholds.

### 3.3.10 Luminosity measurement

Luminosity  $L$  is defined as the ratio of the event rate  $\dot{N}$  to the cross-section  $\sigma$  of a given process, and it is the effective area quantifying the likelihood of a scattering event.

$$\mathcal{L} = \frac{\dot{N}}{\sigma} \quad (3.5)$$

The cross-section is measured in units of area, thus the luminosity is given in units of events per time per area,  $(b.s)^{-1} = 10^{24} \text{cm}^{-2}\text{s}^{-1}$ . Then, the integrated luminosity  $L$  is the integral over the instantaneous luminosity:

$$L = \int \mathcal{L}(t) dt \quad (3.6)$$

For a given process, the number of expected events which are produced is given by the product of the integrated luminosity and the production cross-section  $\sigma_{exp}$ :

$$N_{exp} = L \cdot \sigma_{exp} \quad (3.7)$$

Thus, the integrated luminosity has to be known to estimate the number of events for a given process or measure the production cross-section. Luminosity measurements are used to monitor the LHC performance in real-time, and they provide an overall normalization for physics analyses [35]. A reference process can be used to estimate the cross-section of a given process:

$$\sigma_{exp} = \frac{N_{exp}}{N_{ref}} \cdot \sigma_{ref} \quad (3.8)$$

where  $N_{ref}$  and  $\sigma_{ref}$  are the number of reference events and the reference process's cross-section. Furthermore, the same integrated luminosity has to be used for both processes. Colliders used nowadays employ bunched beams [1]. If two bunches with

$n_1$  and  $n_2$  particles, respectively, collide head-on with frequency  $f$ , the instantaneous luminosity is given by:

$$\mathcal{L} = f \cdot \frac{n_1 n_2}{\sqrt{\epsilon_x \beta_x^* \epsilon_y \beta_y^*}} \quad (3.9)$$

where  $x, y$  are the coordinates transverse to the beam,  $\beta_{x,y}^*$  are the amplitude functions at the interaction point, where the beam optics produces a narrow focus. Emittance  $\epsilon$  is a measure of the beam width defined as  $\epsilon_x \equiv \frac{\sigma_x^2}{\beta_x}$ , with  $\sigma_x$  and  $\sigma_y$  being the root mean square (RMS) of the transverse beam sizes in the horizontal or vertical direction, respectively. A high luminosity can be achieved with a high population of bunches of low emittance colliding at high frequency at locations where the beam optics provide low values of the amplitude functions. As the instantaneous luminosity depends on the beam parameters, it has to be measured when the beam parameter change. A reference process's event rate can estimate the instantaneous luminosity if the cross-section is known. The visible cross-section  $\sigma_{vis}$  is given by:

$$\sigma_{vis} = \sigma(E) \cdot A(t, \mu, \dots) \quad (3.10)$$

where the cross-sections depend on the collision energy  $E$  and the detector acceptance  $A$  depends on time  $t$ , the mean number of interactions per bunch crossing  $\mu$ , and other parameters. Luminometers are independent detectors or parts of the detector for measuring the instantaneous luminosity. Once the calibration constant  $\sigma_{vis}$  has been determined for a luminometer, the luminosity can be estimated using

$$\mathcal{L} = \frac{\dot{N}}{\sigma_{vis}} \quad (3.11)$$

where  $\dot{N}$  is the event rate for this luminometer. Ideally, the luminometer's visible cross-section should not be time-dependent and should not depend on experimental conditions. The visible cross-section has to be determined again when the beam

parameters change. The same detector configuration has to be used during data-taking as during calibration. In summary, there are two important parts of the luminosity measurement. First, the luminometer, which measures the event rate. Secondly, the calibration of the luminometer. In the following, both parts of the luminosity measurement are introduced.

**Luminometer:** CMS uses five detectors to monitor and measure the luminosity based on rate measurements [36–38]. The CMS silicon pixel detector, the DTs in the muon system’s barrel, the forward hadronic calorimeter, the Fast Beam Conditions Monitor (BCM1f), and the Pixel Luminosity Telescope (PLT) are used as luminometers. The luminometer of the pixel detector and the DTs use the standard CMS trigger and data acquisition system, while the PLT, BCM1f, and HF have an independent, fast readout system. However, the silicon pixel detector and the DT have very low occupancy and very good stability over time. In LHC Run I, only the silicon pixel detector and the HF were used as luminometers [39].

The HF has a high rate of acquisition, and it is most sensitive to the electromagnetic component of the hadronic showers. Two methods have been studied to estimate the luminosity. The first one, referred to as zero-counting, counts the hits above the single physical towers’ threshold and averages each tower’s result. The second method exploits the linear relationship between the total transverse energy deposit in the HF and the number of interactions and the luminosity. Both methods require that the mean value of interactions  $\mu$  is proportional to the luminosity. The Pixel Cluster Counting (PCC) method employs a large number of pixels in the CMS detector. A given pixel has an exceedingly small probability of being hit by two different tracks from the same bunch crossing. For  $\mu = 25$ , the fraction of occupied pixels is less than a per mille. It is thus expected that the number of hit pixel clusters is a linear function of the number of interactions per crossing. Thus, the number of hit pixel clusters is a good measure of the instantaneous luminosity, which is given

by:

$$\mathcal{L} = \frac{\langle N_{\text{cluster}} \rangle \cdot f}{\sigma_{\text{vis}}^{\text{PCC}}} \quad (3.12)$$

where  $\langle N_{\text{cluster}} \rangle$  is the mean number of hit pixel clusters,  $f$  the orbit frequency of the LHC, and  $\sigma_{\text{vis}}^{\text{PCC}}$  the visible cross-section of the PCC method. Pixel modules that have not been fully operational during calibration are omitted for the offline luminosity measurement. As the PCC method has a very small dependence on pileup and other experimental conditions, this method is chosen for the precision offline luminosity measurement. The HF measurements have a smaller statistical uncertainty and can be used for cross-checks or studies of the luminosity measurements' systematic uncertainties.

Van der Meer (VdM) scans: The whole cross-section is evaluated using Van der Meer (VdM) scans, which allows measuring the luminosity per colliding bunch pair from machine parameters. A counter system measures the counting rate proportional to the rate of the beam-beam interaction. One of the two beams is displaced, and the luminometer rate is measured as a function of the beam-beam separation resulting in a maximum at zero separation. The VdM scan method is used to assume that the two bunch densities factorize in  $x$  and  $y$ . These scans are performed with a dedicated LHC machine set up. Two beams are scanned through one another in the transverse plane of the detector. The luminosity and the luminosity rate must be determined at the same time to measure the visible cross-section:

$$\sigma_{\text{vis}} = 2\pi \Sigma_x \Sigma_y \langle n \rangle_0 \quad (3.13)$$

where  $\Sigma_x$  and  $\Sigma_y$  are the effective beam widths and  $\langle n \rangle_0 = \frac{1}{2}(R_x + R_y)$  with normalisation rates  $R_x$  and  $R_y$ , which are the amplitudes of the fitted scan curves. Beam Imaging scans are used for studies on the beam shapes. The VdM scan also

determines the transverse and longitudinal interaction point centroids.

Luminosity integration: After the VdM scan, the luminometer's visible cross-section is known, and the instantaneous luminosity can be measured. The integrated luminosity is obtained by summing the luminosities of short time intervals. The convenient minimal time interval to consider for the estimation of luminosity is the luminosity section (LS) corresponding to  $t_{\text{LS}}$  ( $\sim 23$ ). The average number of clusters per event is computed for each LS, and the luminosity for the LS is derived and multiplied by  $t_{\text{LS}}$ . The overall uncertainty on the luminosity measurement was between 2.3% to 2.5% for the three years. This is better than the design goal of a systematic accuracy of 5%.

## CHAPTER 4

### MONTE CARLO EVENT GENERATION

#### 4.1 Introduction

Accurate simulations for signal and backgrounds are needed for searches for new physics. The primary collision along with the decay processes in an event can be described by perturbative quantum field theory. However, perturbative QCD (pQCD) cannot describe the quantum chromodynamics (QCD) bound states. Therefore phenomenological models are needed to describe hadronization.

Event generators are used for generating simulated particle physics events. Event generators factorize the full process of the event simulation into individual tasks. Monte Carlo (MC) methods are used for the probabilistic branching between these individual problems. MC methods are a class of computational algorithms that rely on repeated random sampling to have the same average behaviour in simulation as in collision data. Event signature of beyond standard model particles can be generated to compare its signature to the one of generated background processes.

General-purpose Monte Carlo (GPMC) generators, like PYTHIA [40], provide fully exclusive simulations of high-energy collisions. However, there are also event generators which are specialised on a certain aspect of the event simulation. Perturbative matrix elements for the scattering process are implemented in matrix element generators. Hadronic event generators simulate the initial- and final-state particle showers, hadronization, and soft hadron-hadron physics including composition and substructure of the initial state. An overview of different steps in MC generation for proton-proton collision events can be seen in Figure 4.1.

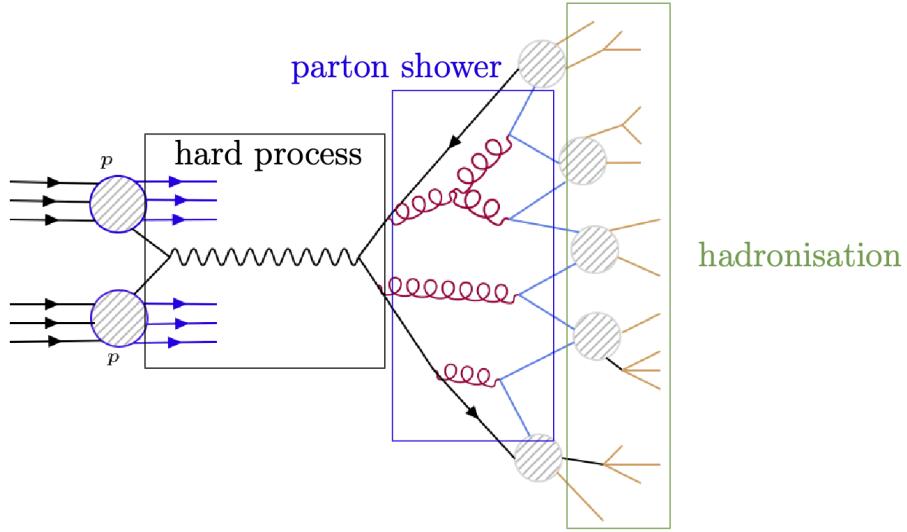


Figure 4.1. Monte Carlo simulation of an event in proton-proton collisions.

In the following sections, first, the simulation of short distance processes is described. In these processes the momentum scales are high enough to use perturbative quantum chromodynamics for the description of the strong interaction between quarks and gluons. Next, quarks fragment into composite particles called hadrons is described using the hadronization models. This is followed by soft hadron-hadron physics modelling and tuning.

First, the simulation of short distances processes is described, where the momentum scales are high enough to use perturbative quantum chromodynamics for the description of the strong interaction between quarks and gluons. Afterwards, hadronisation models are discussed which become important at lower energy scales, where quarks fragment into composite particles called hadrons. Soft hadron-hadron physics modelling, parameters and tuning are introduced. Lastly, the Monte Carlo generators that are important for the current analysis are presented along with the simulation of the detector response.

## 4.2 Perturbative simulation

The primary hard interaction process along with the decay of short-lived particles happen at short-distance scales. The QCD and quantum electrodynamics (QED) radiation at time scales much below  $\frac{1}{\Lambda}$ , where  $\Lambda$  is a typical hadronic scale of a few hundred MeV are also happening at short-distance scales. Soft- and collinear-safe inclusive observables, such as total decay widths or inclusive cross sections, can be computed with pQCD theory for momentum scales much larger than this scale. The final-state collinear splittings and soft emissions give rise to large logarithmically divergent corrections which cancel against virtual corrections in the total cross section. Initial state collinear singularities are factorised into parton density functions (PDFs). Therefore the crosssection for the process remains accurate up to higher order corrections if it is interpreted as inclusive crosssection. If this is not the case then the singularities in QCD can lead to a non-convergence of the fixed-order expansion.

Matrix-element generators generate the matrix element of the hard process and large-angle emissions. The parton showers are generated by hadronic event generators using collinear and soft radiation approximations. Parton-level events are transferred from a hard-process generator to a shower generator, containing a list of particles and the used free parameters, using the Les Houches Event File (LHEF) standard [41].

### 4.2.1 Matrix element generators

Matrix element generators generate the exact matrix elements for the production of the process. They also produce a certain number of additional partons for hard, large-angle emissions. The radiation of extra partons is not included at tree-level accuracy of the hard process. The radiation of an extra parton with tree-level accuracy can be included to provide Next-to-leading order (NLO) corrections along with all NLO virtual corrections. The parton shower algorithms use as input the final-state partons of the hard process and their phase space.

#### 4.2.2 Parton shower algorithm

The parton shower algorithm is used for computing the cross section for a generic hard process. The kinematics of the basic process are first generated, followed by a sequence of independent shower splittings. Sudakov form factors  $\Delta_i(t, t')$  [42], are used for estimating the probability for undergoing a branching before the infrared cut-off for each primary process parton  $i$ . The infrared cut-off is defined by the decay width for an unstable particle or the shower hadronization scale. Sudakov form factor is interpreted as the probability for a splitting not to occur between two scales  $t < t'$ . The parton is either split into two partons or the parton is defined as final parton. Altarelli-Parisi splitting kernels [43] define the probability of the parent parton  $i$ , with energy fraction  $z$ , to decay into two partons  $j$  and  $k$ . All generated partons undergo this procedure recursively and the algorithm stops when no final-state parton undergoes further splitting.

At each splitting vertex we can assign the azimuthal angle  $\phi$  of the splitting process with respect to the incoming parton momentum, the energy fractions  $z$  of the two partons, and an ordering variable for the purpose of ordering the parton splittings. PYTHIA uses the imparted transverse momentum  $p_{\perp}$ . The cross section for the given final state is calculated by assigning a probability to each splitting vertex. Collinear emission and emissions of soft gluons at arbitrary angles are the two sources of infrared singularities in massless field theories like QCD. PYTHIA uses a  $p_{\perp}$ -ordered shower evolution for correctly describing both effects. There is also an angular veto to avoid the particle multiplicity growing too rapidly with energy.

A cut-off for collinear radiation is obtained from quark masses larger than  $\Lambda$ , like  $c$ ,  $b$ , or  $t$  quarks. For angles between both produced partons  $\theta < \theta_0 = \frac{m_q}{E}$ , where  $m_q$  is the quark mass and  $E$  is the energy, the divergent behaviour is regulated. Heavy quarks have less collinear activity than light quarks. Therefore, in the hard process a larger fraction of the momentum acquired is carried by them. A matrix-element

correction method is used by PYTHIA to include the mass effects. PYTHIA does not take the spin correlations into account when generating the parton showers. The initial-state radiation (ISR) has to be taken into account and it induces a nonvanishing transverse momentum of the particles in the matrix element. ISR is taken into account using backwards evolution algorithms.

Photon emission from light charged particles are also added by shower algorithms to account for electromagnetic corrections. For electrons, soft photon emission is especially important. A cut-off for the electromagnetic shower is used to terminate the algorithm. This cut-off is the electron mass in the case of electrons and for quarks, the photon wavelength has to be smaller than the typical hadronic size of about 1 fm. Additional bremsstrahlung can be produced from hadron and tau decays involving charged particles.

The production and decay of particles are treated as being factorisable. The distribution of the mass along with the distributions of its decay products are relevant. A  $\delta$  function at pole mass  $m_0$  or a Breit-Wigner distribution with particle width  $\Gamma$  can be used for mass distribution. Differential decay matrix elements or the pole mass and a uniform phase-space distribution can be used for describing the decay products. The total invariant mass of the decay products is preserved by most parton-shower models to keep the original resonance shape.

#### 4.2.3 Matching

QCD color confinement restricts quarks and gluons from existing as isolated particles. The hadronisation of a quark or a gluon gives rise to hadrons or their decay products. Jets are collimated bunches of these hadrons. The collinear/soft-radiation of an appropriate ( $N + 1$ )-parton final state, generated by a matrix element generator, can give rise to a ( $N + 1$ )-jet event. A ( $N + 1$ )-jet event can also be obtained from an  $N$ -parton final-state with hard, large-angle emission during shower evolu-

tion. A matching has to be done if different generators have been used for generating matrix-elements and parton showers or extra partons have been generated by the hard-process generator.

The exact matrix elements for the production of the basic process and up to  $n$  partons are generated by the matrix element generator. At large angles they are tree-level accurate and match the results of the shower algorithms at small angle. The relative transverse momentum is above a scale  $Q_{cut}$  for each produced parton pair. There are no final states with more partons and no emissions below the scale  $Q_{cut}$ .

The hadronic event generator generates events with less than these  $n$  additional partons and splittings for scales below  $Q_{cut}$ . Splittings are also generated for initial events with  $n$  partons for relative transverse momenta below the scale of smallest pair momentum  $Q_{cut}$ . For fi

xed-order perturbation theory to hold, the matching parameter  $Q_{cut}$  has to be large enough. However, the  $Q_{cut}$  also has to be small enough so that the shower algorithm is accurate for emissions below it. Consistent coupling constant,  $\alpha_S$ , choices between real and virtual corrections have to be used.

To avoid phase-space double-counting as well as unpopulated phase-space regions matching schemes [44] are used. These schemes define which of both paths should be generated for a given event. The choice of the path is optimised to use the best possible approximation to given kinematics. For the production of final state  $X$  and  $n$  jets, a jet measure is defined and all cross sections are calculated. Hard parton samples are then generated with a probability according to the total cross section and the matrix element. A dynamical, kinematics-dependent probability is used to accept or reject the events. The parton shower is finally invoked and no extra jets are produced with it.

## 4.3 Hadronisation models

The hadronisation scale  $Q_{had}$  is by construction equal to the infrared cut-off where the parton shower ends. Coloured partons are transformed into a set of colourless hadrons by GPMCs. This happens at scales with low momentum transfers and at long distances, where non-perturbative effects become important. QCD inspired models which rely on the colour-flow information between partons are used by GPMCs as starting point for hadronisation. The probability distribution for the hadron energy fraction from coloured partons is given by fragmentation functions.

### 4.3.1 Fragmentation function

Fragmentation functions are probability distributions for inclusive hadron spectra. They can be calculated using pQCD and a non-perturbative initial condition obtained by fitting hadron spectra and are de

fined at an arbitrary perturbative scale  $Q$ . More information is included in the MC modelling and is exclusive, while the fragmentation function only describes inclusive spectra [45].

### 4.3.2 String model

String models are based on “linear confinement”. In quenched lattice QCD, at distances greater than a femtometre the potential of the colour-dipole field between a colour charge and an anticharge appears to grow linearly with charge separation. One such model os the Lund model. Parton showering gives rise to color-connected quark-antiquark pairs. As the string grows, the non-perturbative creations of new quark-antiquark pairs are favoured, which break the string in two new strings. Kinks are used for representing intermediate gluons which can also be involved, building a transverse structure in the one-dimensional object, while in

nitely soft gluons are absorbed into the string. When a string breaks they are causally disconnected.

In the Lund model, starting with the leading hadrons the string breaks are generated, containing the endpoint quarks, iterating inwards to the centre of the string, alternating between both sides. In each step, a single on-shell hadron can be split off. After the breakup process the quarks can acquire a transverse momentum  $p_\perp$ .  $m_q$  and  $p_\perp$  are used for exponentially suppressed production of quarks. Strangeness can be generated in the parton shower through perturbative  $g \rightarrow s\bar{s}$  splittings. Baryons can be produced by allowing string breaks to a pairs of diquarks. The relative rate of diquark to quark production can be extracted from collision measurements of ratio of proton to pion.

The produced quarks are assigned to hadron multiplets. As the individual rates are not predicted by the model there are many free parameters. The ratio of vector to pseudoscalar production is another parameter. The fraction  $z$  of the quark longitudinal-momentum carried by the created hadron is estimated using the Lund symmetric fragmentation function. Heavier flavoured hadrons carry a larger momentum fraction  $z$  of the heavy quark [46], therefore their fragmentation functions should be harder than a light hadron.

#### 4.3.3 Hadron and tau decays

Many primary hadrons originating from string breaks are unstable and decay further until a set of particles is obtained which is stable on relevant time scales. The final particle yields and spectra have a significant impacts from the decay modelling. The summary of available experimental measurements are represented in particle summary tables. Charmed hadron decays have been measured at different experiments and all measured branching ratios do not have to sum to unity. However, MC simulations need decay packages with quantified, consistent information, with

all branching fractions adding up to unity. Hence, choices have to be made when adapting summary tables and double counting has to be avoided. The differential distribution of the decay products in phase space also needs to be decided. For a selected class of decays matrix elements can be used. It depends on the generator if additional effects, such as B-meson oscillations, or CP-violating effects, are included.

#### 4.4 Soft hadron-hadron physics modelling

Underlying Event (UE) is the additional activity beyond the basic process and its associated ISR and FSR activity. The dominant part is coming from additional colour exchanges between the beam remnants. Multiple parton-parton interactions (MPI) can produce two or more back-to-back jet pairs with each pair having a small transverse momentum. This is in contrast to jets coming from bremsstrahlung which tend to be aligned with their parent parton. Most MPI are soft and they influence the colour flow and the total scattered energy of the event. This increases the particle multiplicity in the final state and affects the final-state activity. Compared to events with no hard jets, the hard jets appear to sit on top of a higher "pedestal" of underlying activity. This comes from the impact parameter-dependence, since central collisions are more likely to contain at least one hard scattering due to the higher probability of interactions and is called the "jet pedestal" effect. Therefore, the impact-parameter shape is another free parameter.

#### 4.5 Parameter Tuning

The accuracy of the used models is very important for event simulation. The accuracy depends on the inclusiveness of the chosen observables and on the sophistication of the simulation. The models can be improved by improving the theoretical calculations. The precision also depends on the constraints in the free parameters and existing collision data is used to constrain them and is referred to as genera-

tor tuning. MC generators are not tuned beyond the constraints in theoretical and experimental precision due to overfitting. Otherwise it would describe fluctuations or noise instead of the underlying relationship. PDFs for the evolution probability for ISR,  $\alpha_S$ , properties of non-perturbative fragmentation functions, the matching parameter Qcut and the hadronisation scale Qhad are the parameters. Other parameters are the production rates of mesons and baryons, including the ratio of vector to pseudoscalar production, their masses and decay widths. The largest amount of free parameters are from these parameters and the parameters of the decay modelling. If significant changes to decay treatment are made then the hadronisation parameters should be retuned. As discussed for the underlying event activity and the jet-peDESTAL shape, the impact parameter shape of the beam particles is an important parameter. The final state  $y$  the particles and their spectra are influenced by event modelling and the generator tuning. Events generated with different generators or tunes can be different and might not describe the collision data in the entire phase space.

#### 4.6 Monte Carlo generators

GPMC generators like PYTHIA can simulate the full process. However, there are specialized generators which deal with a certain aspect of the event simulation. MadGraph 5 generates the matrix element with leading-order (LO) accuracy [47]. MadGraph 5 is a matrix element generator for processes that involve final states with a large number of jets, heavy flavor quarks, leptons and missing energy. Events from new physics models which are renormalisable or from an effective field theory that can be written in form of a Lagrangian can be generated. The Feynman rules from a given input Lagrangian are derived using FeynRules [48]. The Feynman diagrams are generated and the code is computed which is necessary to evaluate the matrix element at a given phase-space point for a process using the Feynman rules. Valid diagrams are constructed by the diagram generation algorithm which

recursively generates all diagrams in parallel reusing already calculated subdiagrams. Wave functions can be reused when they correspond to identical subdiagrams, but contribute to different diagrams. For each fermion line in a diagram, a fermion flow is defined. Fermions are grouped in pairs with each constituting a fermion line. For states which are charged under SU(3)C, the color coefficients in scattering amplitudes are computed. Multiparton amplitudes are important as they correspond to leading order approximation of multi-jet production. The full amplitude is split into gauge invariant subamplitudes. The matrix element that results contains the full spin correlation and Breit-Wigner effects, but are not valid far from the mass peak.

POWHEG is a framework for implementing next-to-leading order (NLO) matrix-element calculations [49]. It includes NLO virtual corrections and radiation of an extra parton in the matrix element. POWHEG method has been applied to several hard processes, including single top production,  $t\bar{t}$  production and Higgs production via gluon fusion, vector boson fusion, and Higgs boson fusion production associated with a vector boson (Higgs strahlung). It needs the LO matrix-elements and the finite part of the virtual corrections as input from which it finds all the singular regions. Using the Frixione, Kunszt and Signer (FKS) subtraction scheme [50, 51] the real cross section is split into a sum of contributions that are divergent in one singular region. The singular regions are characterised by final-state parton becoming collinear or soft to either an initial-state parton or a final-state parton. The singular regions can be grouped according to their underlying LO-diagram by replacing this parton pair with a single parton of appropriate flavor. Soft and collinear counterterms as well as remnants are built after finding all singular regions. The POWHEG Sudakov form factor are used for generating the radiation. When no radiation is generated, Born-like events are generated. Below a threshold gluon splittings into heavy quark pairs are avoided. The incoming and outgoing partons are assigned color on basis of the underlying LO-process and to the real emitter and the radiated parton. Since the

resonance mass must be preserved by the shower, the information on intermediate resonances is made available to the shower program along with the decay products being specified.

The flavor structure has one more light parton in its final state and colourless and massive coloured particles remain the same at LO and NLO level. The interference of LO and one-loop amplitude, have the same flavor structure as the LO term.

aMC@NLO implements all aspects of NLO computation and its matching with parton showers [52, 53]. NLO calculations can be achieved by combining the computation of one-loop matrix elements and tree-level matrix elements. Tree-level computations are performed using MadGraph and one-loop amplitudes are evaluated with MadLoop [54]. The undesirable large short-scale effects can be removed by MC counterterms. They also set an upper bound for the hardness of each branching. Matched samples, which differ by their final-state multiplicity, can be merged using the FxFx-merging scheme.

MLM matching scheme is a matching algorithm [55, 56] that matches partons from matrix element calculations to jets reconstructed after shower generation. Parton-level events are required to have a separation greater than a minimum value  $R_{jj} > R_{min}$  between them and at least a minimum transverse energy  $E_T^{min}$  for partons. These events do not have any hard-emission veto during shower. Starting with the hardest parton, the jet closest in  $(\eta, \phi)$  is selected and both match if the distance is smaller than  $R_{min}$ . Once a match is found the jet is removed and matching is done with the next parton. If a match is not found then the event is rejected. This is the case for collinear partons or soft partons, which do not lead to an independent jet or are too soft for jet reconstruction.

FxFx merging scheme is an NLO merging procedure [57]. There can be NLO accuracy for exclusive events with  $J$  light jets by the computation based on matrix

elements that have  $J$  and  $(J + 1)$  partons. NLO mergings are more complicated than LO ones. This is because the matrix elements are considered twice, as Born contribution for processes with  $J$  partons, and as real-emission contribution, IR subtraction terms, and the one-loop contributions to processes with  $(J - 1)$  partons. Double counting is avoided by parton-shower dependent MC counterterms and Sudakov reweighting, as Sudakov form factors include virtual corrections and non-emission probabilities. The parton shower dominates at soft scales, while the contributions from the matrix element dominate at hard scales. At intermediate scales a probability function estimates which description is more accurate. There is no emission larger than the matching scale generated by shower. Events are reweighted and a certain amount of events might carry negative weights.

PYTHIA 8 has been developed for multiparticle production in  $e^+e^-$ ,  $pp$  and  $ep$  collisions and simulation of jets. PYTHIA can generate the hard subprocess, initial- and

final-state parton showers, hadronisation, decays and the underlying event. A lot of hard processes have been implemented for generating the matrix elements for final-state and phase space calculation.  $2 \rightarrow 1$  and  $2 \rightarrow 2$  processes can be optimally generated by PYTHIA. Resonance decays with the resonance masses above the b-quark system are implemented. Their branching fractions and partial width can be dynamically calculated as function of their mass. If the spin information is available for resonance decays it leads to properly correlated angular correlations of the resonance decay products, otherwise the resonance decays isotropically.

A multijet structure is added by the parton shower which does not take into account the spin effects. Showers are ordered by their virtuality  $p_\perp$ . The initial-state showering is done by a backward evolution scheme. The shower is traced backwards in time starting from the incoming parton of the hard interaction to find the parton which initiated the shower. PYTHIA 8 uses Dipole showering. Gluon emission is

generated via dipole radiation rather than by splitting partons. FSR are associated with dipoles which are stretched between the "hole" left by an initial-state parton and a final-state parton. The hard scattering subsystem takes the recoil having initial-state radiation unchanged. Lund string model is used for implementing the Hadronization. The process is split into generating the fragmentation and the decays of these hadrons.

#### 4.7 Detector simulation

In detector simulation, the interactions of particles with the detector material and the detector response are simulated. These events can then be reconstructed and analysed. Geant4 [58] is used for detector simulation. It is a toolkit for simulating passage of particles through matter and for simulating particle interactions with matter across a very wide energy range. The user defines the detector geometry and materials. A large number of components with different shape and materials can be included in the geometrical model. Sensitive elements can be defined which record information in the form of hits. Hits are needed to simulate the detector responses called digitisation. The detector's geometrical structure is divided into logical and physical volumes. Logical volumes contain the information of the material and the sensitive detector behaviour. A mixture of different elements and isotopes can be used for the material. Physical volumes carry information about the spatial positioning or placement of the logical volumes.

Particles can interact with the detector material or can decay while they are transported through the geometry. A model can be implemented by electromagnetic and hadronic processes in Geant4 depending on the energy or particle type. Geant4 can handle ionisation described by energy loss and range tables, bremsstrahlung, pair production of electron-positrons from muons, photo-electric effect, pair conversion, annihilation, synchrotron and transition radiation, scintillation, refraction, reflection,

absorption, the Cherenkov effect, and many other processes. Particles with their basic properties, like mass, charge, list of sensitive processes, can be defined. Particles are transported in steps and they are tracked through materials and external electromagnetic fields. Event data is generated during simulation. First, events contain primary vertices and primary particles before processing an event. After processing, hits and digitisations generated by simulation are added. Trajectories of simulated particles can be added optionally for recording of "simulation truth".

## CHAPTER 5

### EVENT RECONSTRUCTION

This chapter begins with the description of the particle-flow algorithm followed by the reconstruction of tracks and vertices, electrons, muons, jets, and other physics objects.

#### 5.1 Particle flow

The global event reconstruction (also called particle-flow (PF) event reconstruction [59]) aims to reconstruct and identify each particle in an event with an optimized combination of information from the various elements of the CMS detector. In this process, the identification of the PF candidate type (photon, electron, muon, charged, and neutral hadrons) plays a vital role in determining particle direction and energy. The PF algorithm links several PF elements that a physics object can give rise to, across different sub-detector layers. The PF elements are tested for their compatibility in the  $\eta - \phi$  plane and are combined to form PF blocks. A predefined sequence of reconstruction and identification algorithms are run in each of these PF blocks. This sequence starts with the reconstruction and identification of muon candidates. PF quality criteria are placed for the muon candidate. The PF elements associated with a muon candidate passing these criteria are removed from the block. The next step in the sequence is to reconstruct and identify electron candidates. The electron candidates are defined as PF electrons if the extrapolated tracks from the tracker have a corresponding energy deposit in the ECAL. The sequence now proceeds with

identifying photons and hadrons. At this stage in the sequence tracks with momentum uncertainty more than the resolution of the calorimeters are removed to reduce fake track identification. All the remaining tracks are associated with charged hadrons, and all the remaining calorimeter energy deposits are associated with photons (ECAL) and hadrons (HCAL). At the end of this sequence, we are left with a list of all electrons, photons, muons, charged hadrons, and neutral hadrons in the event with optimally determined direction, charge, and energy.

## 5.2 Track and primary vertex reconstruction

The hits from the pixel and strip detectors in the tracker are used to reconstruct the tracks of charged particles [60]. Signals above specified thresholds in the pixel and strip channels are clustered to form the hits. The cluster positions and corresponding uncertainties are estimated in a local orthogonal system plane of each sensor. A translation is done between the local coordinate system of these hits to the global coordinate system of the tracks during track reconstruction. Kalman filter [61] based algorithm is used to reconstruct tracks and is called the Combinatorial Track Finder (CTF). Tracks with high  $p_T$  and those that are produced near the interaction region are easiest to find. Track reconstruction uses an iterative procedure with the initial iterations searching for the most accessible tracks. In subsequent iterations, tracks with low  $p_T$  and those produced far from the interaction region are searched. Hits unambiguously assigned to the track in the previous iterations are removed. This reduces the combinatorial complexity in the subsequent iterations. In each iteration, there are four sequential steps.

The first step in the sequence is seed generation. The magnetic field causes the charged particles to follow a helical path, thus requiring five parameters to determine the trajectory. These parameters are extracted using two or three hits in the inner region of the tracker. The seeds are constructed in the inner part owing to the high

granularity of the pixel detectors. The tracks are then constructed outwards. The motivation to use the inner region for seed construction also rests on the fact that particles like pions and electrons interact inelastically with tracker material as they traverse through the tracker to its outer regions.

Kalman filter-based algorithm is then used for track finding. Track parameters are estimated by using the trajectory seeds generated in the previous step. The seed trajectories are extrapolated along the expected path of a charged particle. The track candidates are built using the location and uncertainty of detected hits, also taking into account effects such as Coulomb scattering at successive detector layers. The parameters are updated at each layer. An analytical extrapolation is done that determines which adjacent layers of the detector the trajectory can intersect. A search is performed for silicon modules in these layers that are compatible with the extrapolated trajectory. Mutually exclusive groups are built from all compatible modules in each layer such that no two modules in each group overlap. One of the compatible hits from a group of hits is added to the original track candidate to form new track candidates. The information from the added hits is combined with the original track candidates' extrapolated trajectory to update the trajectory parameters of the new candidates. Figure 5.1 illustrates the reconstruction efficiency of tracks in the case of isolated muons.

The collection of hits from the last step are refitted using a Kalman filter in a phase called track fitting. This provides the best possible estimate of parameters for each track trajectory. There can be several fake tracks that are not associated with any charged particle passing through the tracker. Several quality requirements are applied to the set of reconstructed tracks, which substantially reduces the fake contribution. The quality criterion involves the minimum number of layers the track has hits in, how compatible its origin is with a primary vertex, and how good a fit it yields.

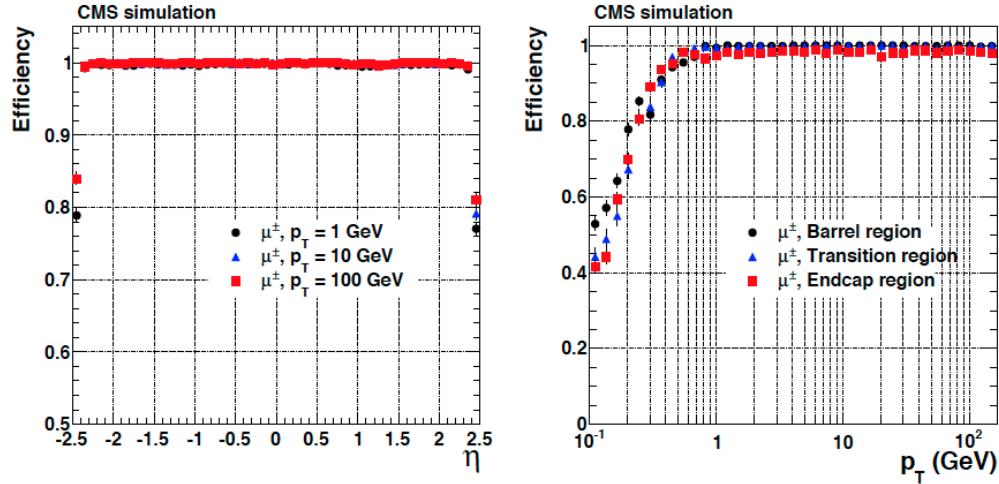


Figure 5.1: Track reconstruction efficiencies for single isolated muons as a function of  $\eta$  and  $p_T$  [60].

The interaction vertices in the proton-proton collisions are reconstructed by selecting tracks produced promptly in the primary interaction region. The chosen tracks are clustered based on their z-coordinates at their point of closest approach to the center of the beam spot. The beam spot represents a 3-D profile of the region where the proton beams collide inside the CMS detector. The adaptive vertex fitter procedure is used for finding the exact positions of the vertices from these clustered candidates [62]. The primary interaction vertex has the most significant sum of squared transverse momenta of tracks originating from it.

### 5.2.1 Muon reconstruction

Muons are reconstructed using the hits in the muon system and tracks from the tracker [63]. The gas in the muon chambers is ionized when muons traverse through them. The ionization is read out by electronics systems that associate these “hits” with well-defined locations in the detector. The hits in the muon chambers are reconstructed independently of track reconstruction in the tracker. Kalman filter is used for reconstructing the hits from the muon system. These tracks are called

*standalone-muon tracks*. Tracker tracks with transverse momentum above 0.5 GeV are propagated to the muon system. Muon tracks are built from these tracks by matching them to segments of hits in DT or CSC. A matching tracker track is called a *tracker muon track*. Standalone-muon tracks can be matched with tracker tracks and combining information from both using a Kalman filter fit. The muon tracks built in such a manner are called *global muon tracks*. Muons leaving hits in several muon stations have a very efficient global muon reconstruction. Muon candidates with low  $p_T$  have an efficient *tracker muon* reconstruction. However, it can cause fake muon tracks due to hadronic particles, which *punch-through* to the innermost muon stations. The *global muon* reconstruction reduces the muon misidentification rate compared to tracker muons. The efficiency for reconstructing a muon is as high as 99% when *tracker muon tracks* and *global muon tracks* are combined. PF algorithm applies the quality criterion for the reconstructed muon candidates. The PF muon candidates used in the analysis were required to satisfy the following set of criterion to be identified as a muon:

- The candidate is reconstructed as a Global Muon along with PF muon identification.
- $\chi^2/\text{ndof}$  of the global-muon track fit  $< 10$ .
- At least one muon-chamber hit included in the global-muon track fit.
- Muon segments in at least two muon stations. This implies that the muon is also an arbitrated tracker muon.
- Its tracker track has transverse impact parameter  $|dxy| < 2 \text{ mm}$  w.r.t. the PV.
- The longitudinal distance of the tracker track wrt. the PV is  $|dz| < 5 \text{ mm}$ .
- Number of pixel hits  $> 0$ .
- Number of tracker layers with hits  $> 5$ .

PF muon identification efficiency is illustrated using a plot from a study performed by the CMS Muon Physics Object group in Figure 5.2. There are differences in the

efficiencies in data and MC simulation. They are corrected using a set of scale-factors applied as a function  $\eta$  and  $p_T$  to adjust the efficiency in simulation to get it to match the efficiency in data.

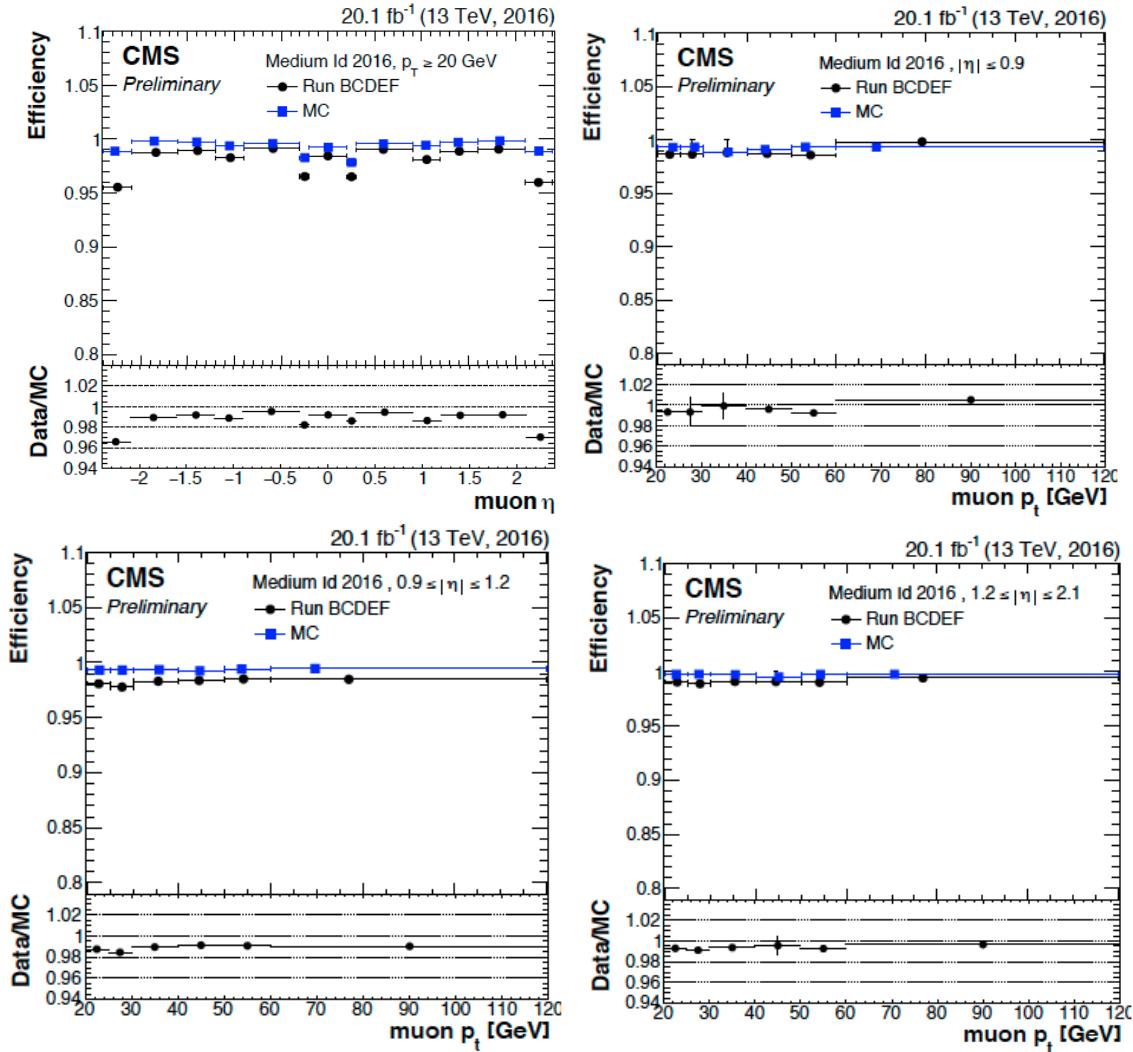


Figure 5.2: Efficiency of muon identification as a function of  $\eta$  and  $p_T$ , for data (black) and simulation (blue) [64].

### 5.2.2 Electron reconstruction

Clusters of energy formed in the ECAL are associated with tracks from the tracker to reconstruct electrons. Electrons radiate bremsstrahlung photons caused by the interaction of electrons with atoms as they pass through the tracker. The radiation depends on the amount of detector material the electron has to cross. The clustering algorithm needs to account for the energy from these bremsstrahlung photon showers to measure the electron's energy. The energy from the bremsstrahlung photons spreads primarily in the  $\phi$  direction, and the spread in  $\eta$  direction is relatively small.

The *hybrid* algorithm is used to cluster the electron energy deposit in the ECAL barrel. It uses the geometry of the ECAL to form clusters that are wide in  $\phi$  direction but are narrow in  $\eta$  direction. A seed crystal contains the most significant amount of energy deposited in the considered region above a 1 GeV threshold. Starting with the seed crystal, 5x1 arrays of crystals are added in  $\eta \times \phi$  around the seed crystals in both directions of  $\phi$  if the energy contained in the arrays is above 0.1 GeV threshold. Contiguous arrays are merged into clusters. An electron supercluster is formed from all such strip clusters which have at least one seed strip with energy above 0.35 GeV threshold. A different clustering algorithm is used in the ECAL endcap owing to the different geometrical arrangements of the crystals. This algorithm is called the 5x5 algorithm. It starts with a seed crystal satisfying the minimum energy requirement of 0.18 GeV. A supercluster is formed by progressively grouping clusters of 5x5 crystals around the seed crystal. The added clusters need to have energy in excess of 1 GeV and to be within  $\pm 0.7$  and  $\pm 0.3$  respectively in  $\eta$  and  $\phi$  around the seed crystal. The energy-weighted mean of the cluster positions is taken as the position of the supercluster. The sum of the energy of all its constituent clusters is its energy. The energy from the preshower is also added to the supercluster. This is implemented by using it's most energetic cluster, and it's maximum distance in  $\phi$  to other clusters, and extrapolating it to the preshower plane to define the spread in the preshower.

A dedicated tracking procedure is used for electron candidates that use information not only from the tracker but also from the ECAL. The first step in electron track reconstruction is seeding. The position and energy of the reconstructed superclusters can be used to constrain the trajectory of the electron through the tracker and the assumption that the electrons originated close to the center of the beam spot. The electron seeds are the hits in the first layers of the trackers compatible with these trajectories. In an alternative approach, tracks constructed by the regular tracking algorithm are extrapolated to the ECAL and matched with a supercluster. The seed collections from these two approaches are merged, leading to an increase in the seeding procedure's overall efficiency. Electron track finding and fitting phases use these seeds. Track finding procedure is adjusted to accommodate tracks that deviate from their expected trajectory because of bremsstrahlung. The penalties assigned to track candidates for passing through a tracker layer without being assigned a hit are similarly adjusted. The Gaussian Sum Filter (GSF) is used for the final track fit. This accounts for the fact that the energy loss of an electron traversing the tracker material is non-Gaussian. As the Kalman filter algorithm assumes Gaussian distribution, the GSF technique deals with this by approximating this non-Gaussian energy-loss distribution as the sum of several Gaussian functions and is found to perform much better than the regular fitting procedure.

The electron candidates are constructed by associating the GSF track produced by the above procedure with a supercluster in the ECAL. A geometrical matching in  $\eta - \phi$  is used for the association for ECAL-seeded candidates. A multivariate (MVA) technique that combines information from supercluster and GSF track is used for tracker-seeded candidates. The charge of the electron is estimated using the GSF track curvature, ECAL supercluster's relative position in  $\phi$  to that of the first hit in the GSF track, and also by using KF tracks that have common hits with the GSF tracks. This combined approach reduces the charge misidentification probability to

1.5%. A combination of tracker and ECAL measurements is used for estimating the momentum of electrons.

Several quality criteria are placed on the reconstructed electron candidates to identify electrons. This helps in suppressing fake sources such as photon conversions, jets misidentified as electrons, etc. Electrons are required to pass an identification variable based on a Boosted Decision Tree (BDT) discriminator, which uses track quality, shower shapes, and kinematic quantities. The following variables are used as input to the BDT:

- Cluster shape variables  $\sigma_{i\eta,i\eta}$  and  $\sigma_{i\phi,i\phi}$ , with  $i\eta$  and  $i\phi$  the integer label of the  $\eta$  and  $\phi$  of a calorimeter cell. The circularity =  $1 - \frac{E_{1\times 5}}{E_{5\times 5}}$ , with  $E_{1\times 5}$  and  $E_{5\times 5}$  the energies in a  $1\times 5$  and a  $5\times 5$  grid around the super cluster seed, respectively.
- Shape variable  $R9 = \frac{E_{3\times 3}}{E_{SC}}$ , with  $E_{3\times 3}$  the energy in a  $3\times 3$  grid of cells around the super cluster seed and  $E_{SC}$  the raw energy of the super cluster.
- The number of valid hits in the track fit, the  $\chi^2$  of the track fit, and the  $\chi^2$  of the GSF Track fit.
- The number of GSF track hits, the number of expected missing inner hits, and the result of the conversion vertex fit.
- The distance  $\Delta\eta$  and  $\Delta\phi$  between the reconstructed supercluster and the associated track at the position of the PV, and the distance in  $\eta$  between the supercluster and the track at the calorimeter surface.
- H/E, the ratio of the hadronic energy over the electromagnetic energy in the supercluster and E/P, the ratio of the supercluster energy over the momentum of the track associated with the electron.
- The ratio of the energy of the electron cluster and the momentum of the associated track, evaluated at the electron cluster, and  $1/E_e - 1/P_e$ , with  $E_e$  the energy of the electron candidate and  $P_e$  its momentum.

Electrons are required to pass an identification variable based on a Boosted Decision Tree (BDT) discriminator, which uses track quality, shower shapes, and kinematic quantities as input. The BDT was trained on a  $Z/\gamma$  Monte Carlo sample generated with MadGraph5, in 3  $\eta$  bins for electrons with  $p_T > 10$  GeV. Instead of using fixed cuts on the MVA score, the cuts are alternatively varied exponentially with

$p_T$  to achieve a more constant efficiency. The tight and loose working points (WPs) are again tuned to give 80% and 90% efficiency, respectively.

The EGamma POG provided two versions of the electron ID, one which includes the electron isolation in the training and one which does not include this. This analysis uses the version of the ID without the isolation included in the training. An additional selection of electron isolation is used. This is done so that in the  $H \rightarrow e\tau_h$  channels, the tight-to-loose method for estimating the misidentified lepton background can be implemented, which uses isolation-based sideband regions. The electrons are also subject to the same impact parameter cuts as the muons: the impact parameters between the electron track (best track) and the PV are restricted as  $|dxy| < 0.045$  cm and  $|dz| < 0.2$  cm to ensure the electron is associated with the PV.

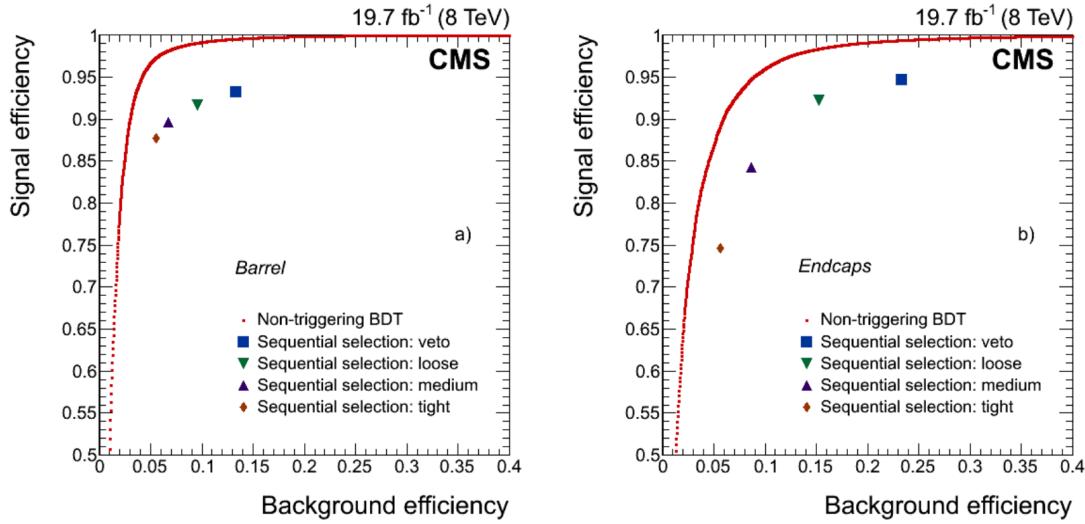


Figure 5.3: Performance of the BDT-based electron identification algorithm (red dots) compared with results from several working points of cut-based selection for electron candidates in the ECAL barrel (left), and endcaps (right) [34].

### 5.3 Hadronic tau leptons

Hadrons plus strips (HPS) algorithm is used for reconstructing the hadronic decays of tau [65]. The Tau POG recommends the use of DeepTau based identification for discrimination against jets, electrons, and muons. A detailed description of the HPS algorithm, followed by the DeepTau Identification algorithm, is given in the following subsections.

#### 5.3.1 Hadrons Plus Strips

Taus in the hadrons plus strips algorithm is seeded by jets clustered with the anti- $k_T$  algorithm with a distance parameter  $\Delta R = 0.4$ . To reconstruct the energy deposits  $\pi^0$  candidates leave in the ECAL, photon and electron constituents of the jet that seeds the  $\tau_h$  reconstruction are clustered into strips. The eor  $\gamma$  (not yet included in a strip) with the highest  $p_T$  is used to build a new strip. The  $\eta$  and  $\phi$  of this candidate determine the initial position of the strip. The next highest  $p_T$  or  $\gamma$  within an  $\eta - \phi$  window centered on the strip location is added to the strip. The position is recomputed as the energy-weighted average of the electron/photon constituents in the strip. This procedure is repeated until there are no more electrons or photons with  $p_T > 0.5$  GeV within the strip window. The  $\Delta\eta$  and  $\Delta\phi$  of the strip vary based on the  $p_T$  or  $E_T$  to be added to the strip. It also depends on the energy, the strip already has, as

$$\Delta\eta = f(p_T^{e/\gamma}) + f(p_T^{strip})$$

$$\Delta\phi = g(p_T^{e/\gamma}) + g(p_T^{strip})$$

where  $p_T^{e/\gamma}$  is the transverse momentum of the candidate to be added to the strip and  $p_T^{strip}$  is the transverse momentum of the strip before merging a new candidate in. In addition, the strip size is bounded as  $0.05 < \Delta\eta < 0.15$ ,  $0.05 < \Delta\phi < 0.3$ . The

functions  $f(p_T)$  and  $g(p_T)$  are defined as

$$f(p_T) = 0.2 \cdot p_T^{-0.66}$$

$$g(p_T) = 0.35 \cdot p_T^{-0.71}$$

If the  $p_T^{strip}$  is at least 2.5 GeV, it is considered as a  $\pi^0$  candidate. Hadronic taus are reconstructed by combining charged particles and strips into different signatures, which are said to be compatible with a specific DM if the set of cuts listed below is satisfied. If a candidate satisfies more than one of the hypotheses, the one that maximizes the  $p_T$  is retained.

The decay modes (DMs) considered for reconstructing taus are:

**One prong, 0  $\pi^0$ :** One charged particle, no strips.

**One prong, 1  $\pi^0$ :** One charged particle + one strip with mass  $0.3 < M_T < 1.3 \cdot \sqrt{p_T/100}$  GeV.  
The mass window upper limit is constrained to lie between 1.3 and 4.2 GeV.

**Three prong, 0  $\pi^0$ :** Three charged particles with mass  $0.8 < M_T < 1.5$  GeV. The tracks are required to originate within  $|dz| < 0.4$  cm of the same vertex.

**Three prong, 1  $\pi^0$ :** Three charged particles and one strip with a total mass  $0.8 < M_T < 1.5$  GeV.

The reconstructed hadronic tau candidates are subject to the impact parameter cuts: the impact parameter between the reconstructed hadronic tau and the PV is restricted as  $|dz| < 0.2$  cm to ensure the hadronic tau is associated with the PV.

### 5.3.2 DeepTau

The application of Machine Learning techniques has been proven to provide superior results for multi-dimensional problems. DeepTau is a new multiclass tau identification algorithm based on a convolutional deep neural network (DNN). DeepTau combines information from the high-level variables attributed to the reconstructed  $\tau_h$  candidates with low-level information from the inner tracker, calorimeters and muon

sub-detectors using PF candidates reconstructed within the  $\tau_h$  signal and isolation cones. DeepTau also takes advantage of using the updated DM definitions.

A balanced mix of  $\tau_e$ ,  $\tau_\mu$ ,  $\tau_h$ , and  $\tau_j$  candidates coming from  $t\bar{t}$ , W + jets, and Z + jets Monte Carlo (MC) simulation is used to perform the training. The  $\tau_h$  has a loose preselection:  $p_T \in [20, 1000]$  GeV,  $|\eta| < 2.3$ , and  $|dz| < 0.2$ , which makes it suitable for the current analysis. The inputs are separated into sets of high-level and low-level features. As high-level inputs, the algorithm takes variables used during tau reconstruction, and one global event variable is the average energy deposition density ( $\rho$ ). For each candidate reconstructed within the tau signal or isolation cones, information of 4-momentum, track quality, relation with the PV, calorimeter clusters, and muon stations are used.

The tau signal and isolation cones define two regions of interest in the vicinity of the tau candidate. Based on the angular distance between the reconstructed tau 4-momentum, all available candidates are split into two  $\eta \times \phi$  grids of  $11 \times 11$  ( $21 \times 21$ ) cells with a cell size of  $0.02 \times 0.02$  ( $0.05 \times 0.05$ ) for the signal (isolation) cone. In cases where there is more than one object of the given type that belong to the same cell, only the object with the highest  $p_T$  is considered as input. Within each cell, the input variables are split into three blocks: e-gamma, muon, hadrons.

The organization of the low-level inputs into two 2D grids allows first processing the local patterns originating from the tau or jet structure. The information obtained is then iteratively combined, covering bigger  $\eta \times \phi$  regions up to the point where the total tau signal or isolation cones are covered. The four outputs of the network represent estimates of the probabilities of the reconstructed tau candidate to be  $\tau_e$ ,  $\tau_\mu$ ,  $\tau_j$ , or a genuine  $\tau_h$ . The performance of tau discrimination against quark and gluon induced jets (left), electrons(middle), and muons(right) for DeepTau and the previously available discriminators can be seen in Figure 5.4.

Seven WPs ranging from Very Very Loose to Very Very Tight, are provided. The

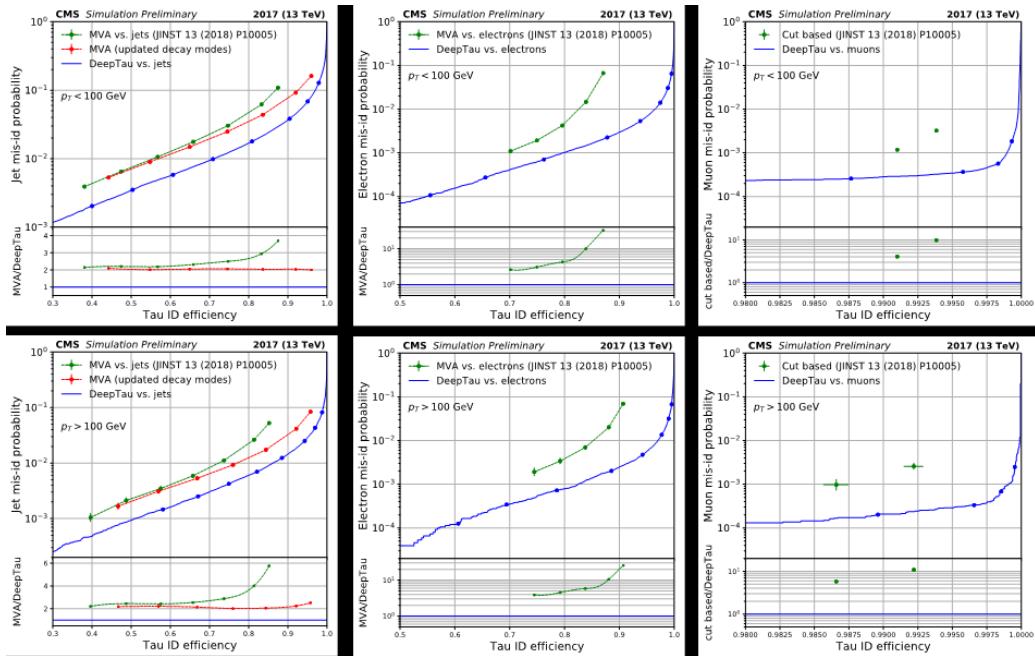


Figure 5.4. Performance of tau discrimination against quark and gluon induced jets (left), electrons(middle), and muons(right) for DeepTau and the previously available discriminators

Tight  $\tau_h$  WP of the DeepTau discrimination against jets is used to provide good tau efficiency and jet rejection. To reduce the  $e \rightarrow \tau_h$  and  $\mu \rightarrow \tau_h$  misidentification, anti-electron, and anti-muon discriminators are used. They are retrieved from the same neural network that is used for the discrimination against jets. The chosen WPs depends on the channel and are specified in the event selection section ???. Eight WPs are provided for the anti-electron discriminator, ranging from Very Very Very Loose to Very Very Tight. Four WPs are provided for the anti-muon discriminator. The choices are made such that efficiency and fake rejection are better than for the WPs chosen in previous analyses for the former discriminators. For anti-muon discriminator, the efficiency hardly changes, so a higher fake rejection is preferred.

#### 5.4 Jet reconstruction

Quarks and gluons hadronize due to color confinement producing a fine spray of particles called jets [66]. Jets are reconstructed using the anti- $k_T$  clustering algorithm. This is a sequential clustering algorithm based on the quantities  $d_{ij}$ , which represents the distance between two entities, and  $d_{iB}$ , which represents the distance of the i-th object from the beam axis.

These distances are defined as:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \quad (5.1)$$

$$d_{iB} = k_{ti}^{2p} \quad (5.2)$$

where  $\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ ,  $k_{ti}$  is the transverse momentum of the i-th entity and R is the radius parameter which is set as 0.4.

In the anti- $k_T$  clustering algorithm,  $p = -1$ , where the parameter p governs the relative power of energy versus geometrical scales.  $d_{ij}$  represents the distance be-

tween all entity pairs present. If the minimum of those distances is smaller than the minimum distance  $d_{iB}$  of any entity from the beam axis, those entities i and j are combined into a single entity. Otherwise, the object closest to the beam axis is considered a jet. It is then removed from the list of entities to be further clustered. The high  $p_T$  particles dominate in the anti- $k_T$  and are clustered first. Softer constituents are subsequently clustered. Before the soft particles cluster among themselves, they cluster with hard particles. As a result of this, a hard particle with no hard neighbors within a distance  $2R$  accumulates all the soft particles within a circle of radius  $R$ . Anti- $k_T$  tries to produce jets with somewhat conical shapes centered around the hardest particles of the event. The boundaries are resilient to the effect of infrared and collinear radiation.

The reconstructed jets' energy differs from their true values as they are complex objects suffering from several effects. Correction factors are applied to calibrate their  $p_T$  and to ensure a uniform response in  $\eta$  [67, 68]. The energy coming from the pileup that has been clustered into the jet needs to be corrected. This is corrected using the *hybrid jet area* method. This method is a combination of the average offset method and the jet area method. The average amount of energy added to the event due to pileup is measured using the zero bias events in the average offset method. This relies on the assumption that averaging over zero bias events makes this measurement insensitive to high  $p_T$  objects. The average offset is measured in bins of  $\eta$ , and the number of pileup vertices ( $N_{PV}$ ) averaged over  $\phi$ . The correction is then given by  $1 - \frac{\langle \text{Offset}(N_{PV}, \eta) \rangle}{p_T^{RAW}}$ , where  $p_T^{RAW}$  is the uncorrected jet  $p_T$ . The other assumption is that every jet contains the same amount of pileup contribution, which is a drawback for this method.

The jet area method calculates corrections on a jet-by-jet basis. The energy density per event is calculated by clustering jets using the  $k_T$  algorithm. The  $k_T$  algorithm favors clustering soft jets as opposed to hard ones. The  $p_T$  is then divided

by jet area, which is defined as the region in  $\eta - \phi$  occupied by soft particles clustered in the jet. The median of this distribution ( $\rho$ ) for an event is expected to be insensitive to hard particles. This  $\rho A_j$  is a good approximation of pileup contribution to the  $i$ -th jet. However, this approach has a drawback because it doesn't take into account the fact that the detector response is  $\eta$  dependent. Thus, the *hybrid jet area* method combines these two methods to calculate a jet-by-jet correction depending on  $\eta$  and  $N_{PV}$ .

The energy of reconstructed jets is corrected with an MC calibration factor to match the generated MC particle jet energy on average. The energy response of reconstructed jets is calibrated to be uniform with respect to  $\eta$  and  $p_T$ . A QCD dijet sample is used to correct the dependence on  $\eta$ . Using jets that are approximately back-to-back in the azimuthal direction but at different  $\eta$  regions of the detector, the difference in response between these two  $\eta$  regions can be ascertained and corrected. Using the same method of measuring residual response in the transverse direction in  $\gamma + jets$  or  $Z + jets$  events, the absolute jet energy scale as a function of  $p_T$  can be made uniform.

## 5.5 Missing transverse energy: $\vec{p}_T^{\text{miss}}$

Neutrinos and other hypothetical particles that are weakly interacting cannot be detected in the CMS detector. The momentum imbalance in the transverse plane can be used to infer their presence. The missing transverse momentum vector  $\vec{p}_T^{\text{miss}}$  is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event, and its magnitude is denoted as  $p_T^{\text{miss}}$  [69]. The  $\vec{p}_T^{\text{miss}}$  is modified to account for corrections to the reconstructed jets' energy scale in the event. Anomalous high- $p_T^{\text{miss}}$  events can be due to a variety of reconstruction failures, detector malfunctions, or non-collision backgrounds. Such events are rejected by dedicated filters that are designed to reject more than 85–90% of the spurious high-

$p_T^{\text{miss}}$  events with a signal efficiency of more than 99.9% [69]. In addition to the event filtering algorithms, the jet identification selection that is imposed requires the neutral hadron energy fraction of a jet to be less than 0.9 rejects more than 99% of the noise jets, independent of jet  $p_T$ , with a negligible mistag rate. Corrections to the  $\vec{p}_T^{\text{miss}}$  are applied to reduce the mismodeling of the simulated  $Z + \text{jets}$ ,  $W + \text{jets}$ , and Higgs boson samples. The corrections are applied to the simulated events on the basis of the vectorial difference of the measured missing transverse momentum and total transverse momentum of neutrinos originating from the decay of the  $Z$ ,  $W$ , or Higgs boson. Their average effect is the reduction of the  $p_T^{\text{miss}}$  obtained from the simulation by a few GeV.

$$\vec{p}_T^{\text{miss}} = -\sum \vec{p}_T \quad (5.3)$$

The  $\vec{p}_T^{\text{miss}}$  plays a vital role in this analysis as it helps gauge the momentum of the neutrinos from the decaying tau lepton. The  $\vec{p}_T^{\text{miss}}$  reconstruction is directly dependent on the reconstruction of all the other objects in the event, from jets to muons to electrons. Consequently, it is sensitive to all the effects that influence the precise reconstruction and calibration of these objects.

## 5.6 Relative isolation

The muon (electron) isolation is measured relative to its  $p_T^\ell (\ell = e, \mu)$ , by summing over the  $p_T$  of PF particles in a cone with  $\Delta R = 0.4(0.3)$  around the lepton:

$$I_{\text{rel}}^\ell = \left( \sum p_T^{\text{PV charged}} + \max \left[ 0, \sum p_T^{\text{neutral}} + \sum p_T^\gamma - p_T^{\text{PU}}(\ell) \right] \right) / p_T^\ell,$$

where  $p_T^{\text{charged}}$ ,  $p_T^{\text{neutral}}$ , and  $p_T^\gamma$  indicate the  $p_T$  of a charged particle, a neutral particle, and a photon within the cone, respectively. The neutral contribution to isolation from pileup,  $p_T^{\text{PU}}(\ell)$ , is estimated from the area of the jet and the median energy

density of the event [70, 71] for the electron or from the sum of transverse momenta of charged hadrons not originating from the primary vertex scaled by a factor of 0.5 for the muons. The charged contribution to isolation from the pileup is rejected by requiring the tracks to originate from the PV.

## CHAPTER 6

### EVENT SELECTION

#### 6.1 Introduction

This chapter summarizes the event selection criteria for the analysis. The signal topology consists of an isolated lepton,  $\mu$  or  $e$ , along with an oppositely charged isolated tau lepton ( $\tau_\mu$ ,  $\tau_e$ , or  $\tau_h$ ). Jets misidentified as electrons or muons are suppressed by imposing isolation requirements. The events are first categorized into  $\mu\tau$  and  $e\tau$  and then further divided into leptonic and hadronic channels based on tau decay mode. Figure 6.1 shows the corresponding Feynman diagrams for the LFV  $H \rightarrow \mu\tau$  and  $H \rightarrow e\tau$  decays.

The final states of this analysis are similar to the  $H \rightarrow \tau\tau$  decay allowed by the SM and since been observed [72]. However, there are some significant kinematic differences. The LFV  $H \rightarrow \mu\tau_h$  and  $H \rightarrow \mu\tau_e$  ( $H \rightarrow e\tau_h$  and  $H \rightarrow e\tau_\mu$ ) decays consist of a muon (an electron) that comes directly from the Higgs and has a hard  $p_T$  spectrum, along with a hadronically decaying tau or a softer electron (muon) that comes from the tau lepton of opposite sign charge, and missing transverse momentum from the tau decay. Also, there are fewer neutrinos in LFV decays, coming from the decay of the single  $\tau$ . The decay products of this highly boosted tau are closely aligned, leading to a narrow separation between the visible decay products of the tau and the  $\vec{p}_T^{\text{miss}}$  in the azimuthal plane. The same is not true in the  $H \rightarrow \tau\tau$  decays. These differences are illustrated pictorially in Figures 6.2 and 6.3.

In each decay mode ( $H \rightarrow e\tau_\mu$ ,  $H \rightarrow e\tau_h$ ,  $H \rightarrow \mu\tau_e$ ,  $H \rightarrow \mu\tau_h$ ), a set of loose selection (preselection) for the respective signature is first defined. The jets in the

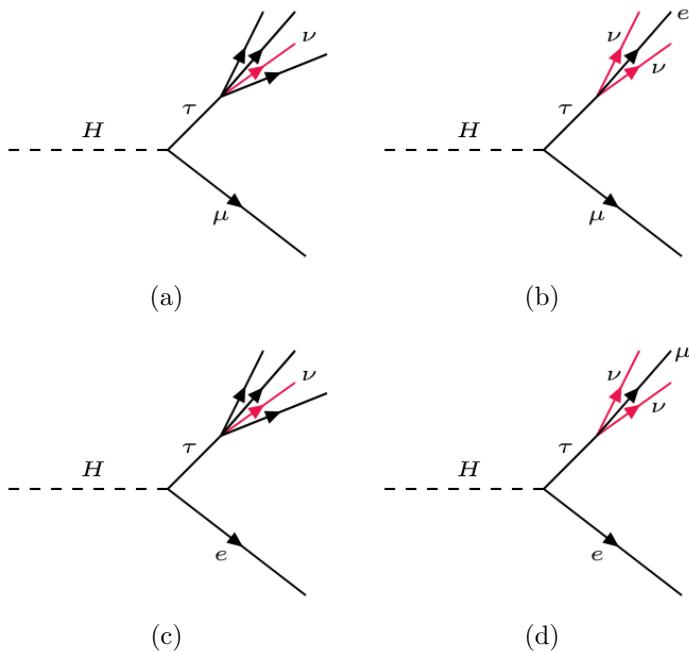


Figure 6.1. Feynman diagrams of lepton-flavor violating Higgs-boson decays. The first row shows diagrams for the Higgs boson coupling to  $\mu\tau$  (a,b). Couplings to  $e\tau$  (c,d) are shown in the second row. Taus can decay leptonically or hadronically. Feynman diagrams are shown for the leptonic decay of taus (b,d) and the hadronic decay of taus (a,c).

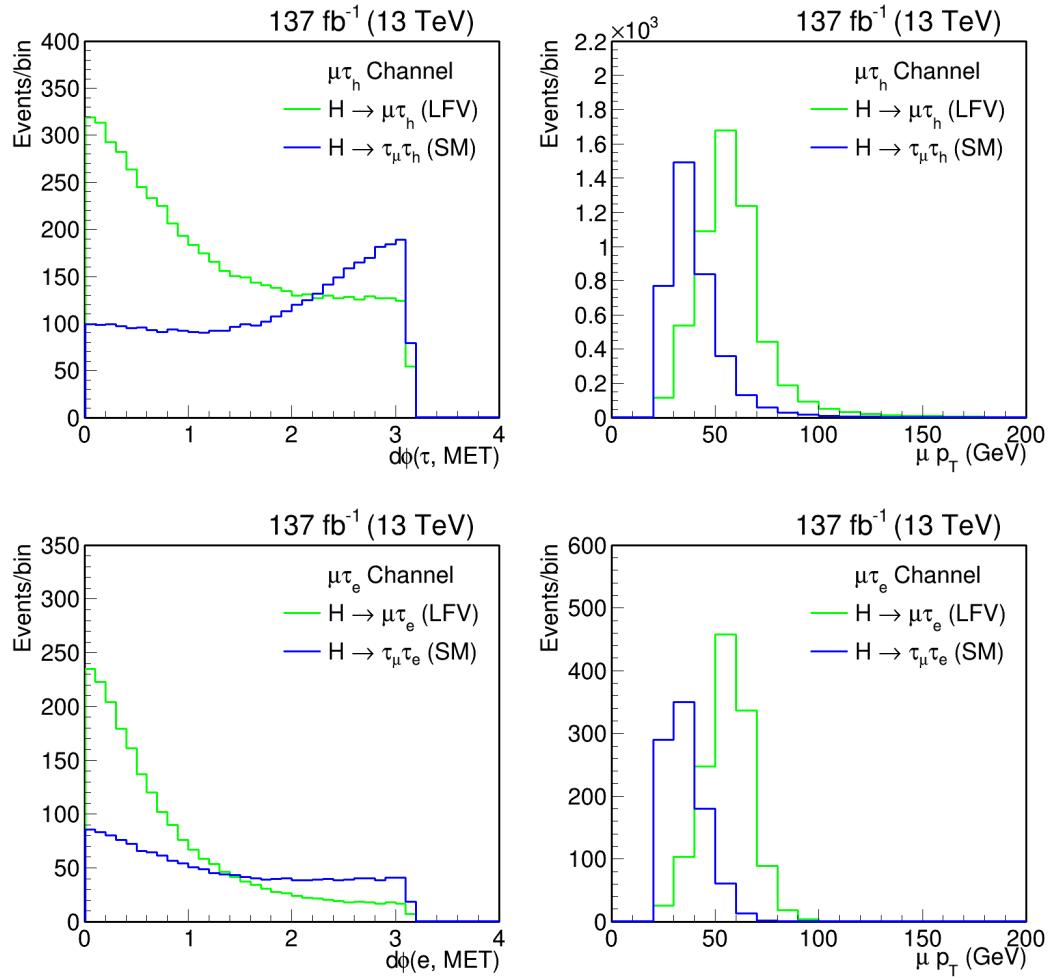


Figure 6.2. Illustration of the differences in  $d\phi(\ell = \tau \text{ or } e, \text{MET})$  and  $p_T^\mu$  spectrums in LFV and SM  $H \rightarrow \tau\tau$  processes.

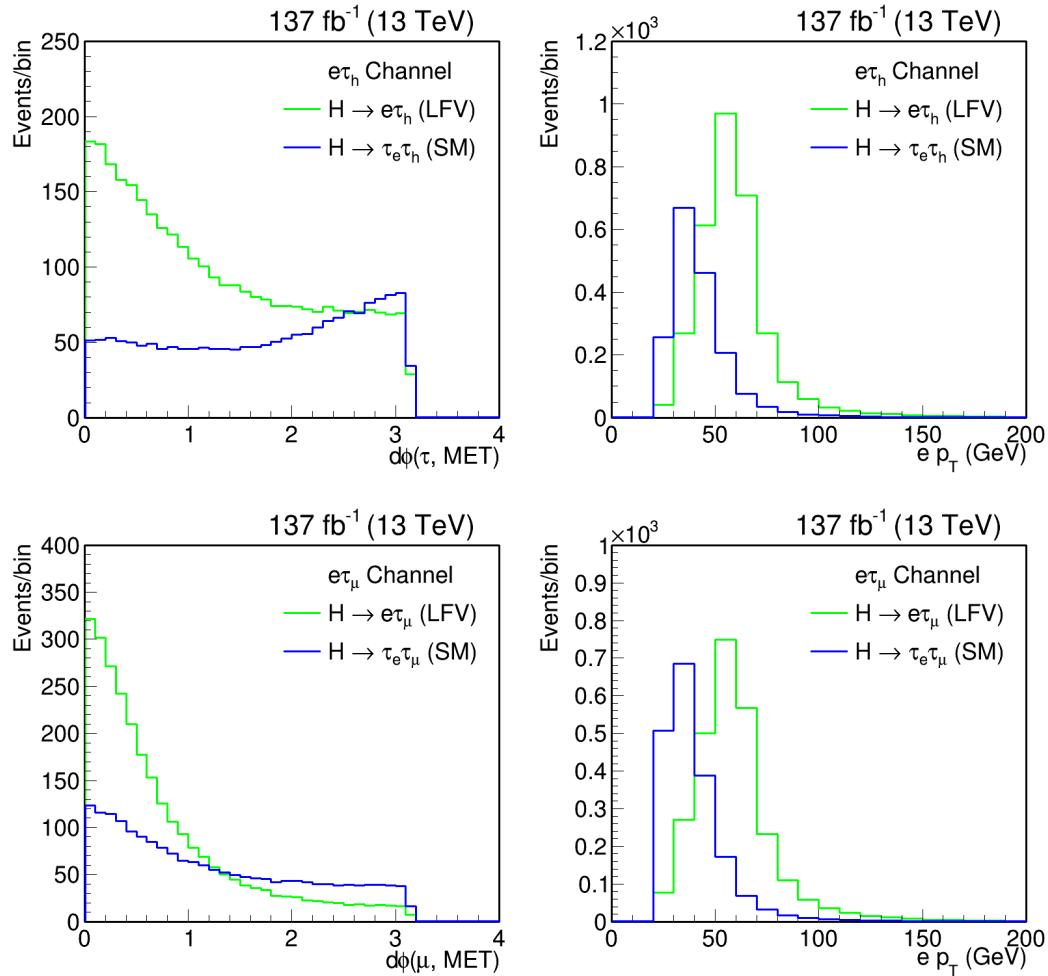


Figure 6.3. Illustration of the differences in  $d\phi(\ell = \tau \text{ or } \mu, \text{MET})$  and  $p_T^e$  spectrums in LFV and SM  $H \rightarrow \tau\tau$  processes.

event are required to have a  $p_T > 30 \text{ GeV}$  and  $|\eta| < 4.7$ . The event in each decay channel is divided into categories based on the number of jets in the event (0-jet, 1-jet, 2-jet) to enhance the contribution of different Higgs boson production mechanisms.

The 0-jet category enhances the Gluon Gluon Fusion (GGF) Higgs production contribution, while the 1-jet category enhances the GGF Higgs production with initial-state radiation. The 2-jet category is further broken into two based on the invariant mass of the two jets ( $M_{jj}$ ). The threshold of 550 (500) GeV on  $M_{jj}$  for  $\mu\tau$  (e $\tau$ ) channels has been optimized to give the best-expected exclusion limits. Events with  $M_{jj} < 550(500)$  GeV enhances GGF Higgs production contribution while  $M_{jj} \geq 550(500)$  GeV enhances VBF Higgs production contribution.

To better discriminate between signal and background events, a Boosted Decision Trees (BDT) discriminator is trained using simulated events, using the TMVA tool of the ROOT analysis package [73]. After applying preselection, a binned likelihood is used to fit the distribution of a BDT discriminator for the signal and the background contributions, and we call this the BDT fit method. The collinear mass ( $M_{\text{col}}$ ) and the transverse mass ( $M_T(\ell)$ ) that are used as input variables to the BDT are defined in the following paragraphs. A brief description of the BDT is given in the next section.

The  $M_{\text{col}}$  provides an estimate of  $m_H$  using the observed decay products of the Higgs boson candidate. It is reconstructed using the collinear approximation based on the observation that, since  $m_H \gg m_\tau$ , the  $\tau$  lepton decay products are highly Lorentz boosted in the direction of the  $\tau$  candidate [74]. The momentum of the neutrino coming from  $\tau$  decay can be approximated to have the same direction as the visible decay products of the  $\tau(\vec{\tau}^{\text{vis}})$ . Figure 6.4 shows the corresponding Feynman diagram for the  $\cancel{E}_T$  projected in the direction of the visible decay products of tau.

The component of the  $\vec{p}_T^{\text{miss}}$  in the direction of the visible  $\tau$  lepton decay products, is used to estimate the transverse component of the neutrino momentum ( $p_T^{\nu, \text{est}}$ ). The

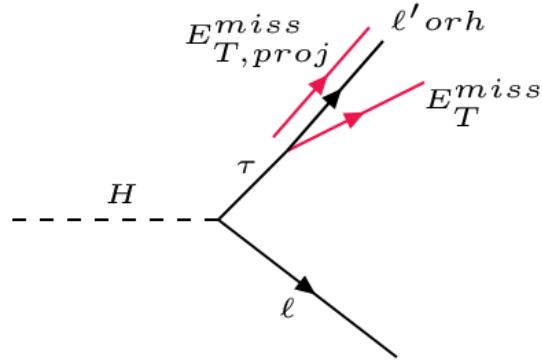


Figure 6.4. Estimation of the neutrino momentum  $\cancel{E}_{T,proj}$  by using the component of the missing transverse energy  $\cancel{E}_T$  which is collinear to the visible decay products of tau in the transverse plane.

collinear mass can then be derived from the visible mass of the  $\mu - \tau$  or  $e - \tau$  system ( $M_{\text{vis}}$ ) as  $M_{\text{col}} = M_{\text{vis}} / \sqrt{x_\tau^{\text{vis}}}$ , where  $x_\tau^{\text{vis}}$  is the fraction of energy carried by the visible decay products of the  $\tau$  ( $x_\tau^{\text{vis}} = p_T^{\tau^{\text{vis}}} / (p_T^{\tau^{\text{vis}}} + p_T^{\nu, \text{est}})$ ), and  $M_{\text{vis}}$  is the invariant mass of the visible decay products.

The  $M_T(\ell)$  is a variable constructed from the lepton momentum and the missing transverse momentum vectors:  $M_T(\ell) = \sqrt{2|\vec{p}_T^\ell||\vec{p}_T^{\text{miss}}|(1 - \cos\Delta\phi_{\ell-\vec{p}_T^{\text{miss}}})}$ , where  $\Delta\phi_{\ell-\vec{p}_T^{\text{miss}}}$  is the angle in the transverse plane between the lepton and the missing transverse momentum, which is used to discriminate the Higgs boson signal candidates from the  $W + \text{jets}$  background.

An alternate analysis has been implemented to cross-check the results obtained from the BDT fit method. This approach involves placing requirements on several kinematic variables and then using the resulting distribution of  $M_{\text{col}}$  as a discriminant for a binned likelihood fit. Henceforth, we call this the  $M_{\text{col}}$  fit method. The BDT and  $M_{\text{col}}$  fit methods were performed blinded in the signal region [75]. The selection criterion described was developed without looking at the data in the region where the signal is expected. This approach is standard in particle physics analysis and

eliminates the experimenter's bias. We use a blinding criteria of  $\frac{s}{\sqrt{s+b}} > 0.2$  for the plots that are shown in this chapter.

## 6.2 Boosted Decision Tree

A decision tree is a tree structure in which there is a condition on an attribute at each internal node. Each branch represents the outcome of this condition, and each leaf node represents a class label. The tree structure is built based on binary splits Figure 6.5. The starting point of the tree structure is called a root node containing all the events we want to classify. A sequence of binary splits is made using conditions on the input variables provided to the classifier. The variable which ensures the separation of the signal and the background is used for each split.

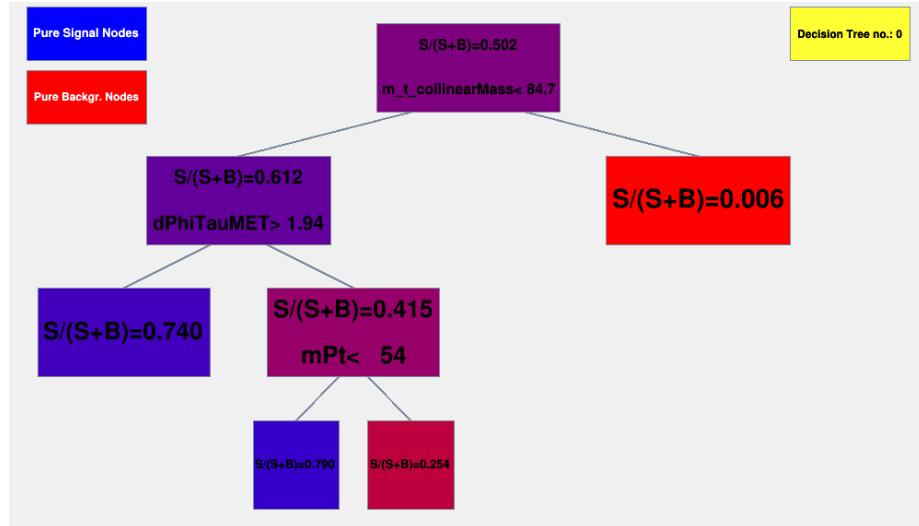


Figure 6.5. Illustration of decision tree.

Gini Index is used as the separation criterion, and it is defined by  $p.(1-p)$ , where

$p$  is the purity of the node. The purity of a node is given by the ratio of signal events to all events in that node. Since the splitting criterion is always a cut on a single variable, the training procedure selects the variable and cut value that optimizes the increase in the Gini Index between the parent node and the sum of the two daughter nodes' indices, weighted by their relative fraction of events. The same variable can thus be used for splitting several nodes, and the splitting is stopped when a predefined depth of the tree, purity of leaf nodes, the minimum number of events in a leaf node is reached. An event which ends up in a leaf node with a majority of signal events is classified as a signal event and vice versa.

A single decision tree is a weak classifier. The performance of weak classifiers can be enhanced using the Boosting technique. This technique works by building classifiers using reweighted training data and then taking a weighted majority vote of the sequence of classifiers thus produced. AdaBoost (adaptive boosting) method was used for boosting. AdaBoost is adaptive in that subsequent classifiers are tweaked in favor of those instances misclassified by previous classifiers. The misclassified event weights depend on the training error of each decision tree. The training error is calculated as

$$\text{err}_m = \frac{\sum_{i=1}^N w_i I(y_i \neq DT_m(x_i))}{\sum_{i=1}^N w_i} \quad (6.1)$$

in which the subscript  $m$  is the tree label and  $w$  is the event weight. The  $y_i$  is the true label for the event, 1 for signal and -1 for background.  $DT_m(x_i)$  is the output of the decision tree. The variable  $I(y_i \neq DT_m(x_i))$  equals 1 if  $y_i \neq DT_m(x_i)$  or 0 otherwise. The weight for event  $i$  is updated using  $\alpha_m$  which is calculated from the training error.  $\beta$  is the learning rate.

$$\alpha_m = \beta \times \ln ((1 - \text{err}_m) / \text{err}_m) \quad (6.2)$$

$$w_i \rightarrow w_i \times e^{\alpha I(y_i \neq DT_m(x_i))} \quad (6.3)$$

By construction, the training error is  $\text{err}_m \leq 0.5$  as the same training events used to classify the output nodes of the previous tree are used to calculate the training error. The learning rate parameter can be used to adjust the step size of each re-weighting. Event weights in each tree are renormalized to keep the summed weights constant. After the boosting and training processes, the final score of each event is  $DT(x)$ . A high score indicates a signal like event while a low score indicates a background like event.

$$DT(x) = \sum_{m=1}^{N_{tree}} \alpha_m DT_m(x) \quad (6.4)$$

This technique also helps in stabilizing the response of the classifiers for fluctuations in the training data. It utilizes a predefined depth of the tree instead of pruning it to avoid overfitting to the training data. All the BDT trainings were done with an ensemble of 850 decision trees, with each tree having a maximum depth of 3. The minimum node size is required to be 2.5%, and the learning rate is set to 0.5. A training to testing split of 50:50 was used.

### 6.3 $H \rightarrow \mu\tau_h$ channel

The first step is to require the events to pass an isolated muon trigger. For 2016 data, this trigger has a muon  $p_T$  threshold of 24 GeV. However, for the 2017 and 2018 data, the trigger with a 24 GeV threshold is prescaled. Prescaling corresponds to collecting one out of every n events to reduce the event rate. We use this trigger in conjunction with the isolated muon trigger with a muon  $p_T$  threshold of 27 GeV.

In addition to the event passing the trigger, the reconstructed leptons corresponding to the trigger have to match the HLT objects within  $\Delta R < 0.5$ .

Next, the preselection begins by requiring an isolated  $\mu$  and an isolated  $\tau_h$  candidates of opposite electric charge and separated by  $\Delta R > 0.5$ . The muon candidate is required to have  $p_T > 26 \text{ GeV}$ ,  $|\eta| < 2.1$  and isolation  $I_{\text{rel}}^\mu < 0.15$ . The  $\tau_h$  candidate is required to have  $p_T > 30 \text{ GeV}$  and  $|\eta| < 2.3$ . Events containing additional electrons, muons, or  $\tau_h$  candidates are vetoed. Events with at least one b jet tagged by DeepCSV algorithm are rejected in order to suppress the  $t\bar{t}$  background.

A BDT is trained after applying preselection criteria. The signal training sample considered is a mixture of simulated GGF and VBF events, weighted according to their respective SM production cross-sections. The misidentified lepton background and  $Z \rightarrow \tau\tau$  background are the dominant backgrounds in this channel. The background used for training the BDT is a data sample of misidentified lepton events with the same charge assignment for both leptons along with the Drell-Yan MC sample with signal selections. The input variables to the BDT are  $p_T^\mu, p_T^{\tau_h}, M_{\text{col}}, \vec{p}_T^{\text{miss}}, M_T(\tau_h, \vec{p}_T^{\text{miss}}), \Delta\eta(\mu, \tau_h), \Delta\phi(\mu, \tau_h)$ , and  $\Delta\phi(\tau_h, \vec{p}_T^{\text{miss}})$ . The distribution of the input variables to the BDT can be seen in Figure 6.6.

The selection on  $\vec{p}_T^{\text{miss}}$  is motivated by the presence of neutrinos in the  $\tau$  lepton decays. The neutrino is expected to be collinear with  $\tau_h$ , which leads to selection on the  $\Delta\phi(\tau_h, \vec{p}_T^{\text{miss}})$  variable. The two leptons are usually in the opposite direction in the azimuthal plane, which leads to selection on the  $\Delta\phi(\mu, \tau_h)$  variable. The BDT discriminator distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow \mu\tau_h$  channel, are shown in results chapter.

In the  $M_{\text{col}}$  fit method, additional selection criteria require  $M_T(\tau_h, \vec{p}_T^{\text{miss}}) < 105 \text{ GeV}$  in the 0-, 1-, and 2-jet GGF categories and  $M_T(\tau_h, \vec{p}_T^{\text{miss}}) < 85 \text{ GeV}$  in the 2-jet VBF category. The  $M_{\text{col}}$  distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow \mu\tau_h$  channel, are shown in results chapter.

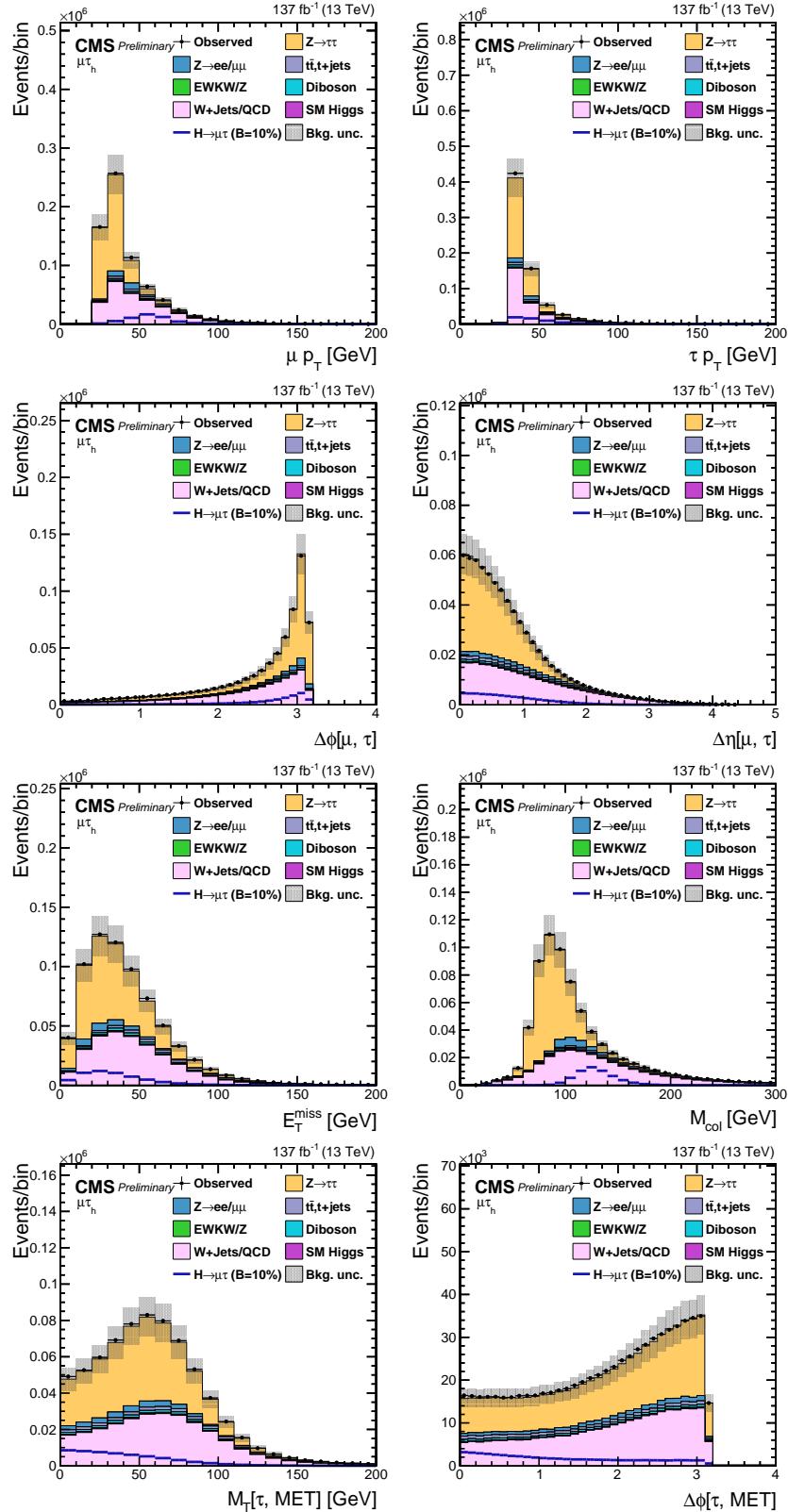


Figure 6.6. Distribution of the input variables to the BDT for the  $H \rightarrow \mu\tau_h$  process.

TABLE 6.1

 EVENT SELECTION CRITERIA FOR THE KINEMATIC VARIABLES  
 FOR THE  $H \rightarrow \mu\tau$  CHANNELS

Variable	$H \rightarrow \mu\tau_h$	$H \rightarrow \mu\tau_e$
$p_T^e$	—	$> 13$
$p_T^\mu$	$> 26$	$> 24$
$p_T^{\tau_h}$	$> 30$	—
$ \eta ^e$	—	$< 2.5$
$ \eta ^\mu$	$< 2.1$	$< 2.4$
$ \eta ^{\tau_h}$	$< 2.3$	—
$I_{\text{rel}}^e$	—	$< 0.1$
$I_{\text{rel}}^\mu$	$< 0.15$	$< 0.15$
$I_{\text{rel}}^{\tau_h}$	DNN $\tau_h$ ID	—
Trigger	$\mu(24)$ (all years)	$e(12)$ and $\mu(23)$ (all years)

$M_{\text{col}}$ fit selection								
	0-jet		1-jet		2-jet			
	GGF	VBF	GGF	VBF	GGF	VBF		
$M_{jj}$	—	—	$< 550$	$\geq 550$	—	—	$< 550$	$\geq 550$
$p_T^\mu$	—	—	—	—	$> 30$	$> 26$	$> 26$	$> 26$
$M_T(\mu)$	—	—	—	—	$> 60$	$> 40$	$> 15$	$> 15$
$M_T(\tau_h)$	$< 105$	$< 105$	$< 105$	$< 85$	—	—	—	—
$\Delta\phi(e, \vec{p}_T^{\text{miss}})$	—	—	—	—	$< 0.7$	$< 0.7$	$< 0.5$	$< 0.3$
$\Delta\phi(e, \mu)$	—	—	—	—	$> 2.5$	$> 1.0$	—	—

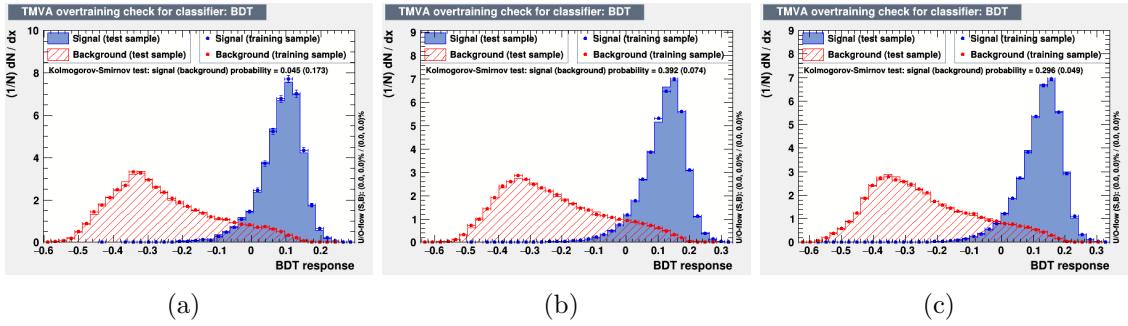


Figure 6.7. Overtraining check as performed in TMVA for the trained BDT in  $H \rightarrow \mu\tau_h$  channel for 2016 (a), 2017 (b), and 2018 (c).

## 6.4 H → $\mu\tau_e$ channel

The events are required to pass the cross-trigger with  $p_T$  thresholds on the muon and the electron. The  $p_T$  threshold on the muon is 23 GeV, and on the electron is 12 GeV. The cross-trigger also places a constraint on the two leptons' longitudinal impact parameter to the primary vertex. However, this constraint is not present in the initial 2016 data samples and 2016 MC samples. In addition to the event passing the trigger, the reconstructed leptons corresponding to the trigger have to match the HLT objects within  $\Delta R < 0.5$ .

The preselection criteria for  $H \rightarrow \mu\tau_e$  channel requires an isolated muon and an isolated electron candidates of opposite charge and separated by  $\Delta R > 0.3$ . The muon candidate is required to have  $p_T > 24 \text{ GeV}$ ,  $|\eta| < 2.4$  and isolation  $I_{\text{rel}}^\mu < 0.15$ . The electron candidate is required to have  $p_T > 13 \text{ GeV}$ ,  $|\eta| < 2.5$  and isolation  $I_{\text{rel}}^e < 0.1$ . The  $p_T$  threshold of the electron and the muon are dictated by the cross-trigger we use for selecting the events of this channel. Events containing additional electrons, muons,  $\tau_h$  candidates or at least one b jet tagged by DeepCSV algorithm are removed.

Similar to the  $H \rightarrow \mu\tau_h$  channel, a BDT is trained after applying preselection

criteria. The signal training is done in a similar way while for background training dominant contributors,  $t\bar{t}$  and  $Z \rightarrow \ell\ell$  ( $\ell = e, \mu, \tau$ ) events are mixed and weighted by their respective production cross-sections. The  $t\bar{t}$  process contributes dominantly for the 2-jet category, with significant contribution to 1-jet category.  $Z \rightarrow \ell\ell$  background processes dominantly contribute the 0- and 1-jet categories. The QCD multijet background has the third-largest contribution, so we use the same sign control region in data as additional background for training. The input variables to the BDT are:  $p_T^\mu, p_T^e, M_{\text{col}}, M_T(\mu, \vec{p}_T^{\text{miss}}), M_T(e, \vec{p}_T^{\text{miss}}), \Delta\phi(e, \mu), \Delta\phi(\mu, \vec{p}_T^{\text{miss}})$ , and  $\Delta\phi(e, \vec{p}_T^{\text{miss}})$ . The distribution of the input variables to the BDT can be seen in Figure 6.8. The BDT discriminator distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow \mu\tau_e$  channel, are shown in results chapter.

In the  $M_{\text{col}}$  fit method, additional selection criteria require a stringent selection on muons,  $p_T > 30 \text{ GeV}$  for 0-jet category and  $p_T > 26 \text{ GeV}$  in rest of the categories. The  $M_T(\mu, \vec{p}_T^{\text{miss}})$  is required to be greater than 60, 40, 15 and 15 GeV for 0-, 1-, 2-jet GGF and VBF categories, respectively, while azimuthal separation between the electron and  $\vec{p}_T^{\text{miss}}$  is required to be less than 0.7, 0.7, 0.5 and 0.3 for 0-, 1-, 2-jet GGF and VBF categories, respectively. For the 0- and 1-jet categories  $\Delta\phi(e, \mu) > 2.5$  and 1.0, respectively. The preselections and the selections for  $H \rightarrow \mu\tau_h$  and  $H \rightarrow \mu\tau_e$  channels in all categories are summarized in Table 6.1. The  $M_{\text{col}}$  distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow \mu\tau_e$  channel, are shown in results chapter.

## 6.5 $H \rightarrow e\tau_h$ channel

The first step is to require the events to pass a single electron trigger. For 2016 data, this trigger has an electron  $p_T$  threshold of 25 GeV. However, for the 2017 and 2018 data, the trigger with the 25 GeV threshold is prescaled. Single-electron triggers with an electron  $p_T$  threshold of 27 GeV, 32 GeV, and 35 GeV are used in conjunction

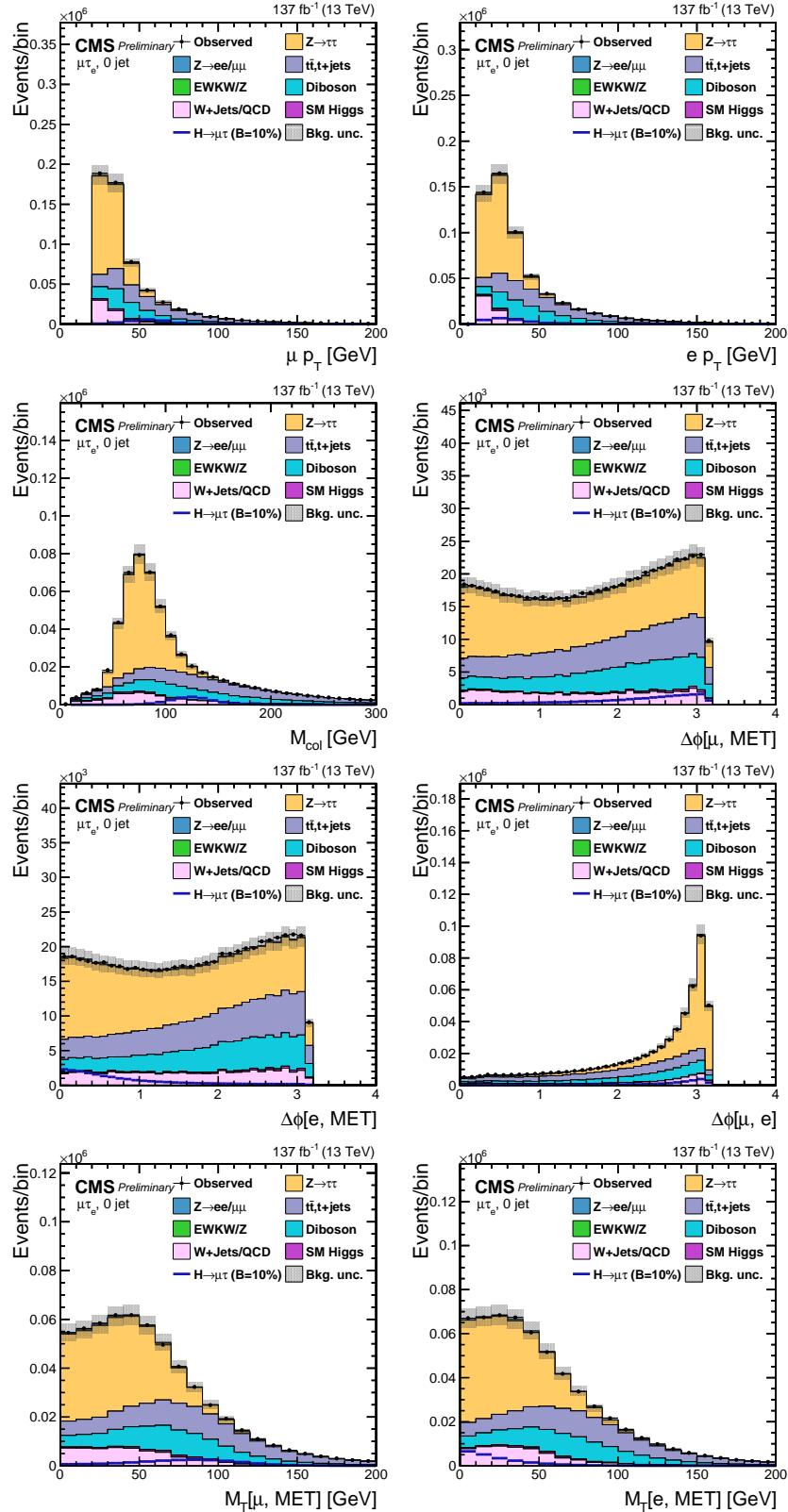


Figure 6.8. Distribution of the input variables to the BDT for the  $H \rightarrow \mu\tau_e$  process.

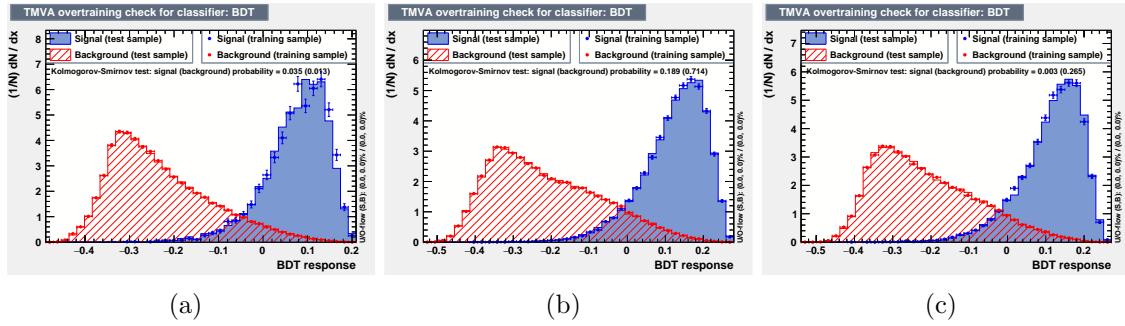


Figure 6.9. Overtraining check as performed in TMVA for the trained BDT in  $H \rightarrow \mu\tau_e$  channel for 2016 (a), 2017 (b), and 2018 (c).

with the cross-trigger with an electron  $p_T$  threshold of 24 GeV and tau  $p_T$  threshold of 30 GeV. In addition to the event passing the trigger, the reconstructed leptons corresponding to the trigger have to match the HLT objects within  $\Delta R < 0.5$ .

The preselection in this channel requires an isolated electron and an isolated  $\tau_h$  candidates of opposite charge and separated by  $\Delta R > 0.5$ . The electron candidate is required to have  $p_T > 27 \text{ GeV}$ ,  $|\eta| < 2.1$  and isolation  $I_{\text{rel}}^\mu < 0.15$ . The  $\tau_h$  candidate is required to have  $p_T > 30 \text{ GeV}$  and  $|\eta| < 2.3$ . Events containing additional electrons, muons, or  $\tau_h$  candidates or at least one b jet tagged by DeepCSV algorithm are removed.

A BDT is trained after applying preselection criteria. The same training samples, as used in the  $H \rightarrow \mu\tau_h$  channel, are considered. The list of input variables to BDT training stays the same, except for the addition of the visible mass,  $M_{\text{vis}}$  variable, and removal of  $\vec{p}_T^{\text{miss}}$ . The  $M_{\text{vis}}$  variable is more useful as the relative composition of the two channels' backgrounds is different. In particular,  $Z \rightarrow ee + \text{jets}$  background contributes more with respect to  $Z \rightarrow \mu\mu + \text{jets}$  background. The distribution of the input variables to the BDT can be seen in Figure 6.10. The BDT discriminator distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow e\tau_h$  channel, are shown in results chapter.

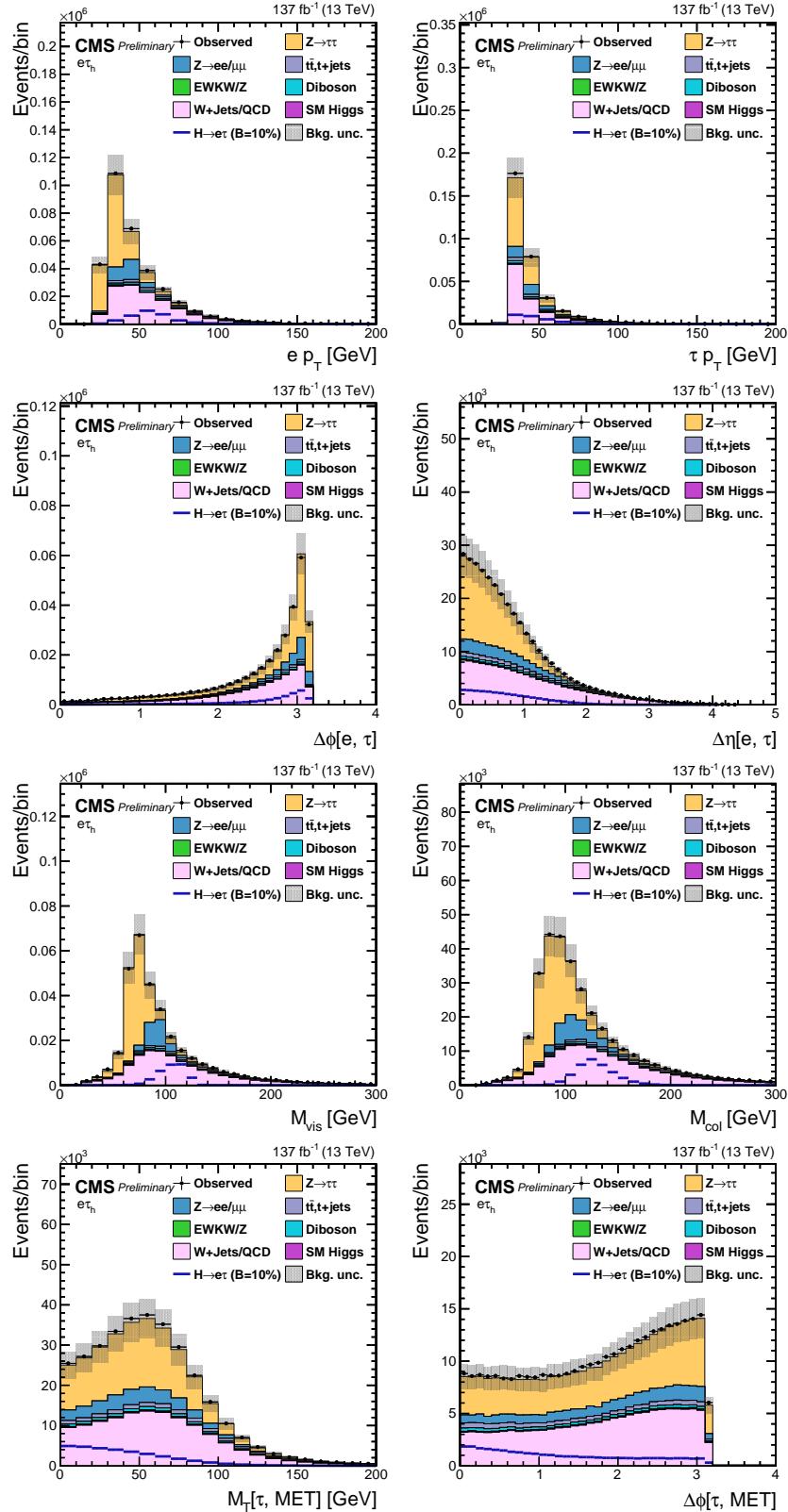


Figure 6.10. Distribution of the input variables to the BDT for the  $H \rightarrow e\tau_h$  process.

In the  $M_{\text{col}}$  fit method, additional selection criteria require  $M_{\text{T}}(\tau_h, \vec{p}_{\text{T}}^{\text{miss}}) < 60 \text{ GeV}$  in all the categories. The  $M_{\text{col}}$  distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow e\tau_h$  channel, are shown in results chapter.

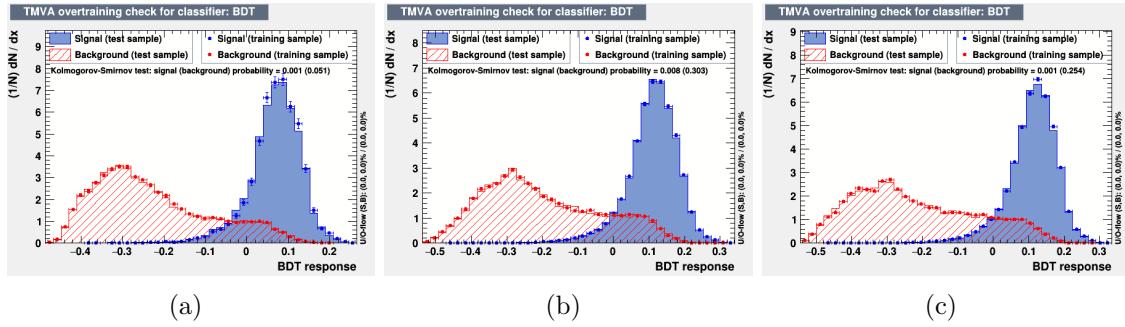


Figure 6.11. Overtraining check as performed in TMVA for the trained BDT in  $H \rightarrow e\tau_h$  channel for 2016 (a), 2017 (b), and 2018 (c).

## 6.6 $H \rightarrow e\tau_\mu$ channel

The events are required to pass the cross-trigger with  $p_{\text{T}}$  thresholds on the electron and the muon. The  $p_{\text{T}}$  threshold on the electron is  $23 \text{ GeV}$ , and on the muon is  $8 \text{ GeV}$ . The cross-trigger also places a constraint on the longitudinal impact parameter of the two leptons to the primary vertex. However, this constraint is not present in the initial 2016 data samples and 2016 MC samples. In addition to the event passing the trigger, the reconstructed leptons corresponding to the trigger have to match the HLT objects within  $\Delta R < 0.5$ .

The preselection criteria for this channel requires an isolated electron and an isolated muon candidates of opposite charge and separated by  $\Delta R > 0.4$ . The electron

TABLE 6.2

 EVENT SELECTION CRITERIA FOR THE KINEMATIC VARIABLES  
 FOR THE  $H \rightarrow e\tau$  CHANNELS

Variable	$H \rightarrow e\tau_h$	$H \rightarrow e\tau_\mu$
$p_T^e$	$> 27$	$> 24$
$p_T^\mu$	—	$> 10$
$p_T^{\tau_h}$	$> 30$	—
$ \eta ^e$	$< 2.1$	$< 2.5$
$ \eta ^\mu$	—	$< 2.4$
$ \eta ^{\tau_h}$	$< 2.3$	—
$I_{\text{rel}}^e$	$< 0.15$	$< 0.1$
$I_{\text{rel}}^\mu$	—	$< 0.15$
$I_{\text{rel}}^{\tau_h}$	DNN $\tau_h$ ID	—
Trigger	e(25) (2016) e(27) (2017) / e(32) (2018) e(24) and $\tau_h$ (30) (2017, 2018)	e(23) and $\mu$ (8) (all years)

$M_{\text{col}}$ fit selection							
	0-jet		1-jet		2-jet		
	ggH	VBF	ggH	VBF	ggH	VBF	
$M_{jj}$	—	—	$< 500$	$\geq 500$	—	—	$< 500$ $\geq 500$
$p_T^e$	—	—	—	—	$> 30$	$> 26$	$> 26$ $> 26$
$M_T(e)$	—	—	—	—	$> 60$	$> 40$	$> 15$ $> 15$
$M_T(\tau_h)$	$< 60$	$< 60$	$< 60$	$< 60$	—	—	—
$\Delta\phi(\mu, \vec{p}_T^{\text{miss}})$	—	—	—	—	$< 0.7$	$< 0.7$	$< 0.5$ $< 0.3$
$\Delta\phi(e, \mu)$	—	—	—	—	$> 2.5$	$> 1.0$	—

candidate is required to have  $p_T > 24 \text{ GeV}$ ,  $|\eta| < 2.4$  and isolation  $I_{\text{rel}}^e < 0.1$ . The muon candidate is required to have  $p_T > 10 \text{ GeV}$ ,  $|\eta| < 2.5$  and isolation  $I_{\text{rel}}^\mu < 0.15$ . The  $p_T$  threshold of the electron and the muon are dictated by the trigger we use for selecting the events of this channel. Events containing additional electrons, muons,  $\tau_h$  candidates or at least one b jet tagged by DeepCSV algorithm are removed. The selections for both  $H \rightarrow e\tau_h$  and  $H \rightarrow e\tau_\mu$  channels in all categories are also summarized in Table 6.2.

The BDT training is done after applying preselection criteria. The same training samples as used in the  $H \rightarrow \mu\tau_e$  channel are considered. The list of input variables to BDT training also stays the same, except for the addition of the visible mass,  $M_{\text{vis}}$  variable, and removal of  $M_T(e, \vec{p}_T^{\text{miss}})$ . The distribution of the input variables to the BDT can be seen in Figure 6.12. The BDT discriminator distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow e\tau_\mu$  channel, are shown in results chapter.

In the  $M_{\text{col}}$  fit method, additional selection criteria require a stringent selection on electrons,  $p_T > 30 \text{ GeV}$  for 0-jet category and  $p_T > 26 \text{ GeV}$  in rest of the categories. The  $M_T(e, \vec{p}_T^{\text{miss}})$  is required to be greater than 60, 40, 15 and 15 GeV for 0-, 1-, 2-jet GGF and VBF categories, respectively, while azimuthal separation between the muon and  $\vec{p}_T^{\text{miss}}$  is required to be less than 0.7, 0.7, 0.5 and 0.3 for 0-, 1-, 2-jet GGF and VBF categories, respectively. For the 0- and 1-jet categories  $\Delta\phi(e, \mu) > 2.5$  and 1.0, respectively. The selections for both  $H \rightarrow e\tau_h$  and  $H \rightarrow e\tau_\mu$  channels in all categories are also summarized in Table 6.2. The  $M_{\text{col}}$  distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow e\tau_\mu$  channel, are shown in results chapter.

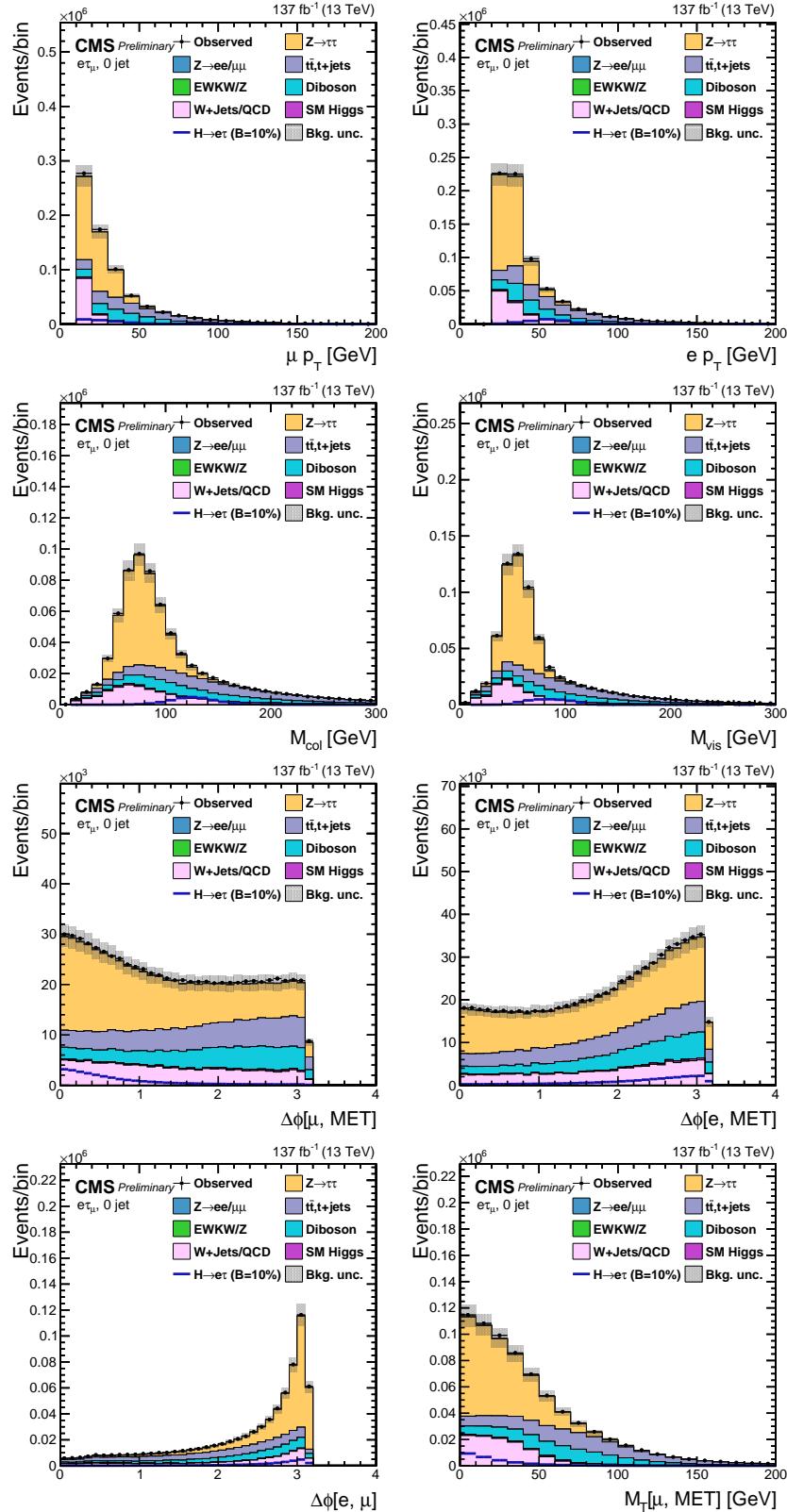


Figure 6.12. Distribution of the input variables to the BDT for the  $H \rightarrow e\tau_\mu$  process.

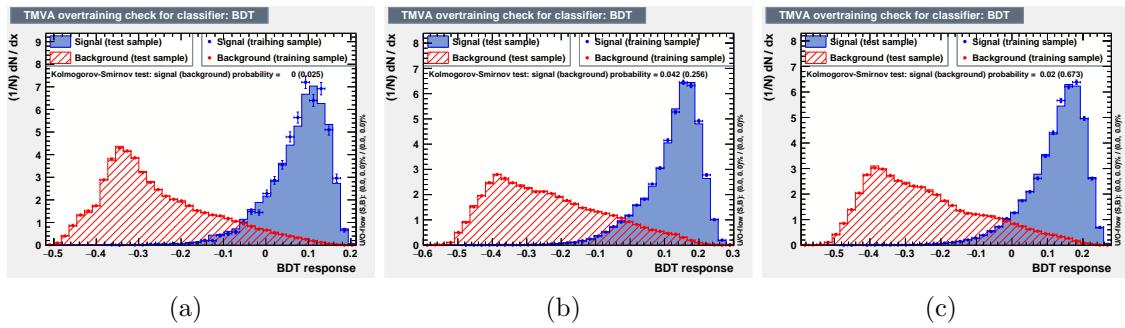


Figure 6.13. Overtraining check as performed in TMVA for the trained BDT in  $H \rightarrow e\tau_\mu$  channel for 2016 (a), 2017 (b), and 2018 (c).

## CHAPTER 7

### BACKGROUND ESTIMATION

#### 7.1 Introduction

The signal is a pair of oppositely charged leptons with different flavors, an isolated lepton,  $e$  or  $\mu$ , accompanied by an isolated  $\tau(\tau_\mu, \tau_e, \text{ or } \tau_h)$  lepton. The dominant contribution for such a signature comes from  $Z \rightarrow \tau\tau$  process, in which the  $\mu$  or  $e$  arises from a  $\tau$  decay. The other dominant contribution comes from  $W + \text{jets}$  and QCD multijets processes, where one or more of the jets are misidentified as leptons. In the leptonic channels, the  $t\bar{t}$  process also has a dominant contribution.

Other contributions come from the processes in which a lepton pair is produced from the weak decays of quarks and vector bosons. These processes include Higgs boson production ( $H \rightarrow \tau\tau, WW, WW, WZ, \text{ and } ZZ$ ). There are non-negligible contributions from processes like  $W\gamma^{(*)} + \text{jets}$ , single top quark production, and  $Z \rightarrow \ell\ell (\ell = e, \mu)$ . Feynman diagrams of background processes to LFV Higgs boson decays are shown in Figure 7.1.

The dominant contributors,  $Z \rightarrow \tau\tau$  and misidentified lepton backgrounds, are estimated from data using either a fully data-driven or semi data-driven approach. All the other backgrounds are estimated from simulated samples. The background estimates are validated in different orthogonal control regions constructed to have enhanced contributions from the dominant backgrounds.

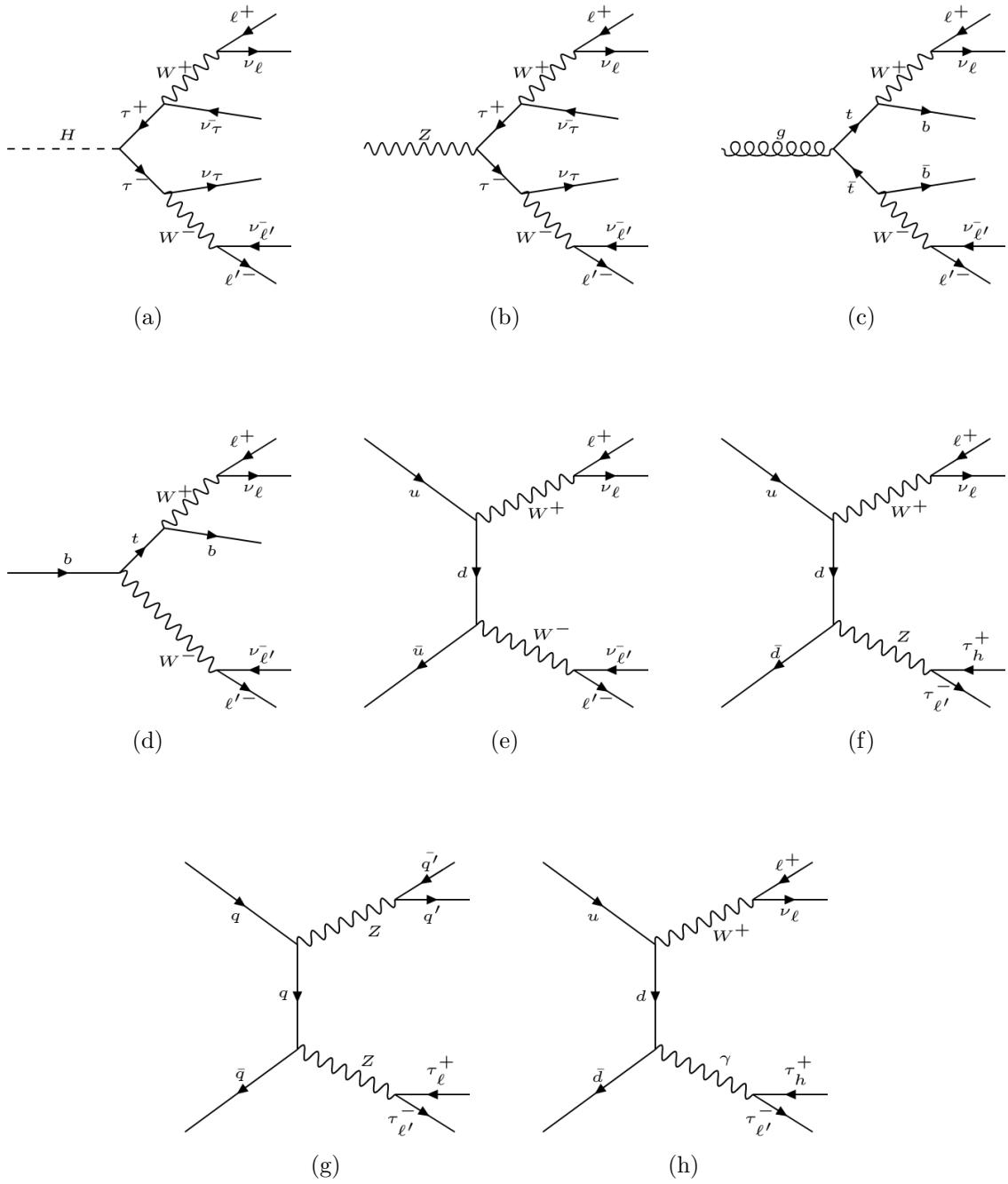


Figure 7.1. Feynman diagrams of background processes to LFV Higgs boson decays: (a)  $H \rightarrow \tau\tau$ , (b)  $Z \rightarrow \tau\tau$ , (c)  $t\bar{t}$ , (d) Single Top, (e)  $WW$ , (f)  $WZ$ , (g)  $ZZ$ , and (h)  $W\gamma^{(*)}$ .

## 7.2 Embedding technique

The  $Z \rightarrow \tau\tau$  background is estimated from data using the embedding technique [76]. The embedding technique allows for an estimation of the genuine  $\tau\tau$  standard model backgrounds from data that minimizes uncertainties arising from a poor event description, with minimal simulation input. Events with a pair of oppositely charged muons are selected in data so that  $Z \rightarrow \mu\mu$  events largely dominate it. These data events are chosen independently of the event selection criteria that are described in Chapter 6.

The muons are removed from the selected events and replaced with simulated  $\tau$  leptons with the same kinematic properties as that of the replaced muon. In that way, a set of hybrid events is obtained that relies on simulation only for the decay of the tau leptons. The description of the underlying event or the production of associated jets is taken entirely from data, and there is no reliance on the simulation. This technique results in a more accurate description of the  $\vec{p}_T^{\text{miss}}$ , jet related variables, and an overall reduction in the systematic uncertainties that arise due to the usage of simulated samples.

Embedded samples cover all backgrounds with two real  $\tau$  leptons decaying semi-hadronically or leptonically. This includes a small fraction of  $t\bar{t}$ , Diboson, and electroweak W/Z events. The events from the  $t\bar{t}$ , Diboson, and electroweak W/Z MC samples where both tau candidates match genuine taus at the generator level are removed to avoid any double counting. A schematic for the embedding technique can be seen in Figure 7.2.

The  $Z \rightarrow \tau\tau$  background is validated by looking at the agreement between observed data and estimated background in a region enriched with  $Z \rightarrow \tau\tau$  events. In  $H \rightarrow \mu\tau_h$  channel, this region is constructed by requiring, in addition to the preselection,  $M_T(\mu) < 40 \text{ GeV}$ ,  $40 \text{ GeV} < M_{\text{vis}}(\mu, \tau) < 80 \text{ GeV}$ , and  $P_\zeta(\mu, \tau) > -25$ .  $P_\zeta$  is the difference of the projections of  $p_T^\ell$  plus MET, and the  $p_T^\ell$  on the axis bisecting the

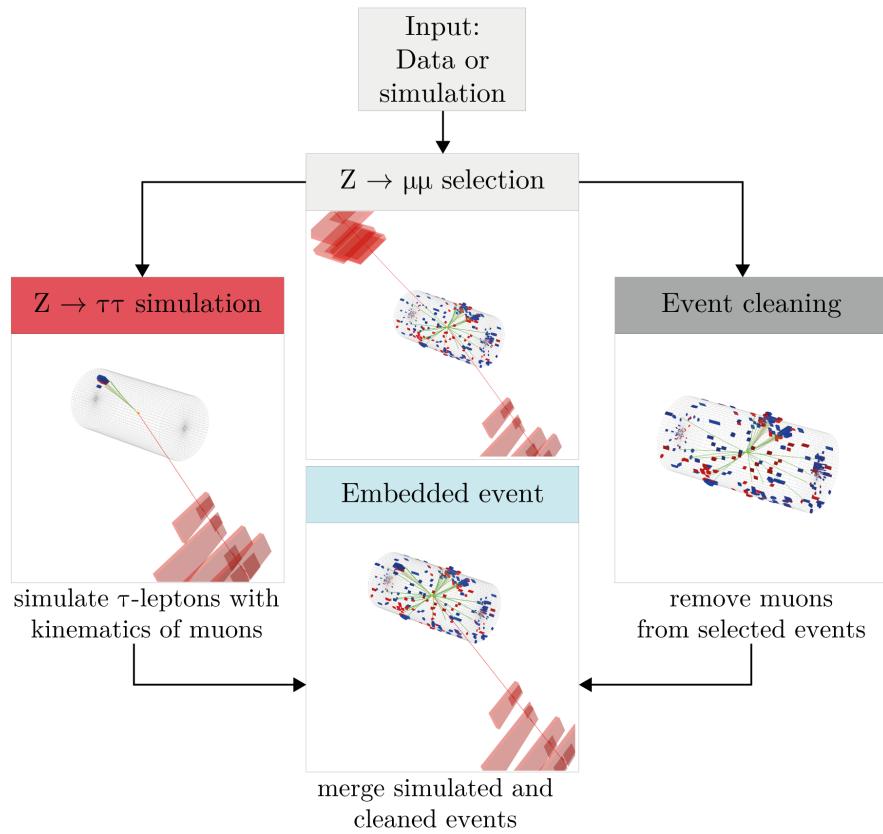


Figure 7.2. Schematic of Embedding Technique

two leptons. In  $H \rightarrow e\tau_h$  channel, the same requirements are placed with the muon variables replaced by corresponding electron variables.

In  $H \rightarrow \mu\tau_e$  channel, this region is constructed by requiring, in addition to the preselection,  $M_T(\mu) < 60 \text{ GeV}$ ,  $30 \text{ GeV} < M_{\text{vis}}(\mu, e) < 70 \text{ GeV}$ , and  $p_T^\mu < 40 \text{ GeV}$ . In  $H \rightarrow e\tau_\mu$  channel, the same requirements are placed with the muon variables replaced by corresponding electron variables and vice versa. Figures 7.3 and 7.4 show the comparison of data with background estimates in the  $Z \rightarrow \tau\tau$  control regions for the  $H \rightarrow \mu\tau$  and  $H \rightarrow e\tau$  channels.

### 7.3 Misidentified lepton background

Misidentified lepton background corresponds to processes where jets are misidentified as leptons. They mostly arise from two sources,  $W + \text{jets}$ , and QCD multijet events. In  $W + \text{jets}$  background events, one of the lepton candidates is from the  $W$  boson decay while the other is a jet misidentified as a lepton. In QCD multijet events, both the lepton candidates are misidentified jets.

In two channels of this analysis ( $\mu\tau_h$  and  $e\tau_h$ ), the contributions from misidentified lepton backgrounds have been estimated using a fully data-driven approach. In the leptonic channels ( $\mu\tau_e$  and  $e\tau_\mu$ ), a semi data-driven approach is adopted. The results from the semi-data driven approach are found consistent with the fully data-driven method and are undertaken due to limited statistics in the leptonic channel.

#### 7.3.1 Fully data-driven approach

The misidentified lepton background is estimated from collision data by defining a control region with the same selection as the signal region, but loosening the isolation requirements on one of the leptons, to enrich the contribution from  $W + \text{jets}$  and QCD multijets. The misidentification rates are evaluated using events with a  $Z$  boson candidate, and at least one jet that can be misidentified as a lepton and then

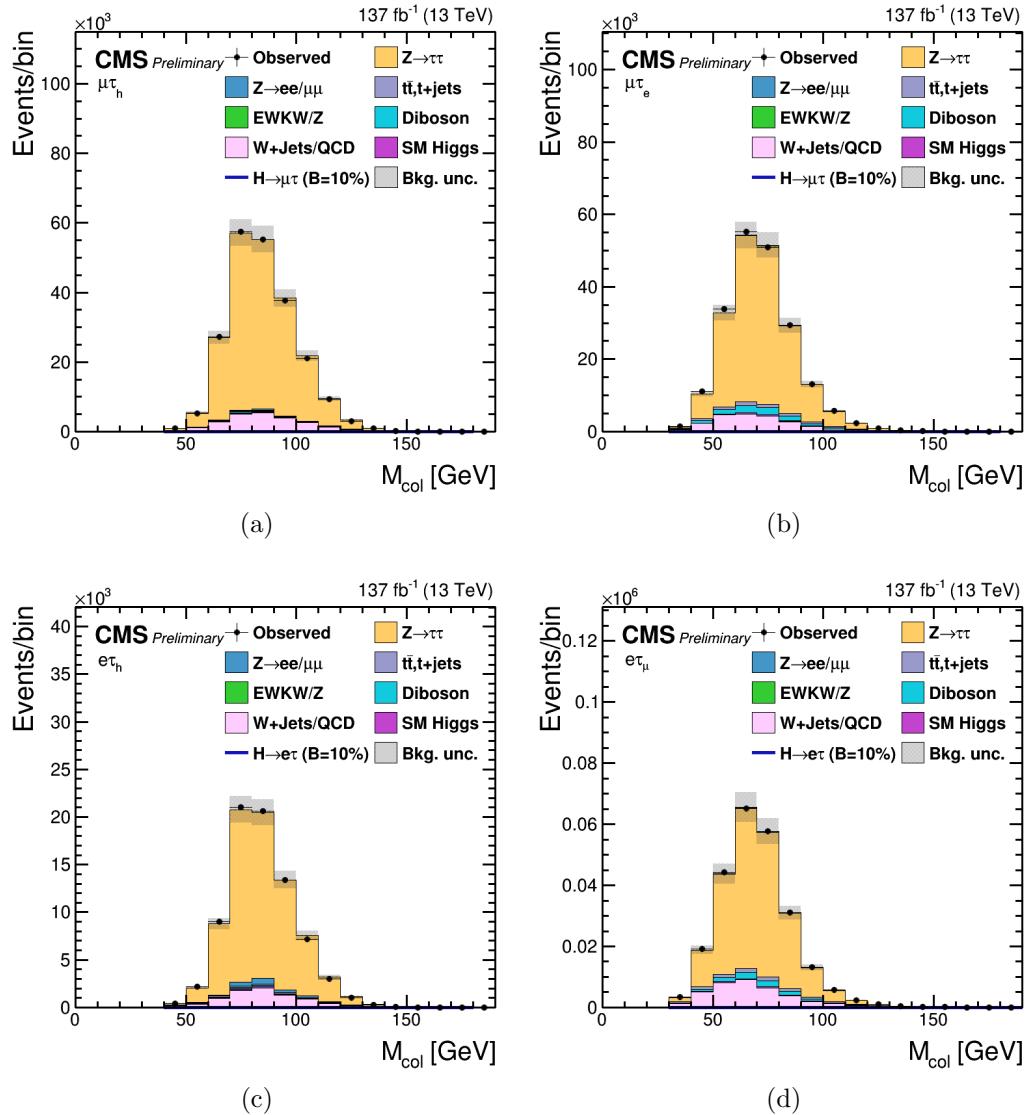


Figure 7.3. Distributions of  $M_{\text{col}}$  discriminator in the  $Z \rightarrow \tau\tau$  control regions for the (a)  $H \rightarrow \mu\tau_h$ , (b)  $H \rightarrow \mu\tau_e$ , (c)  $H \rightarrow e\tau_h$ , and (d)  $H \rightarrow e\tau_\mu$  channels.

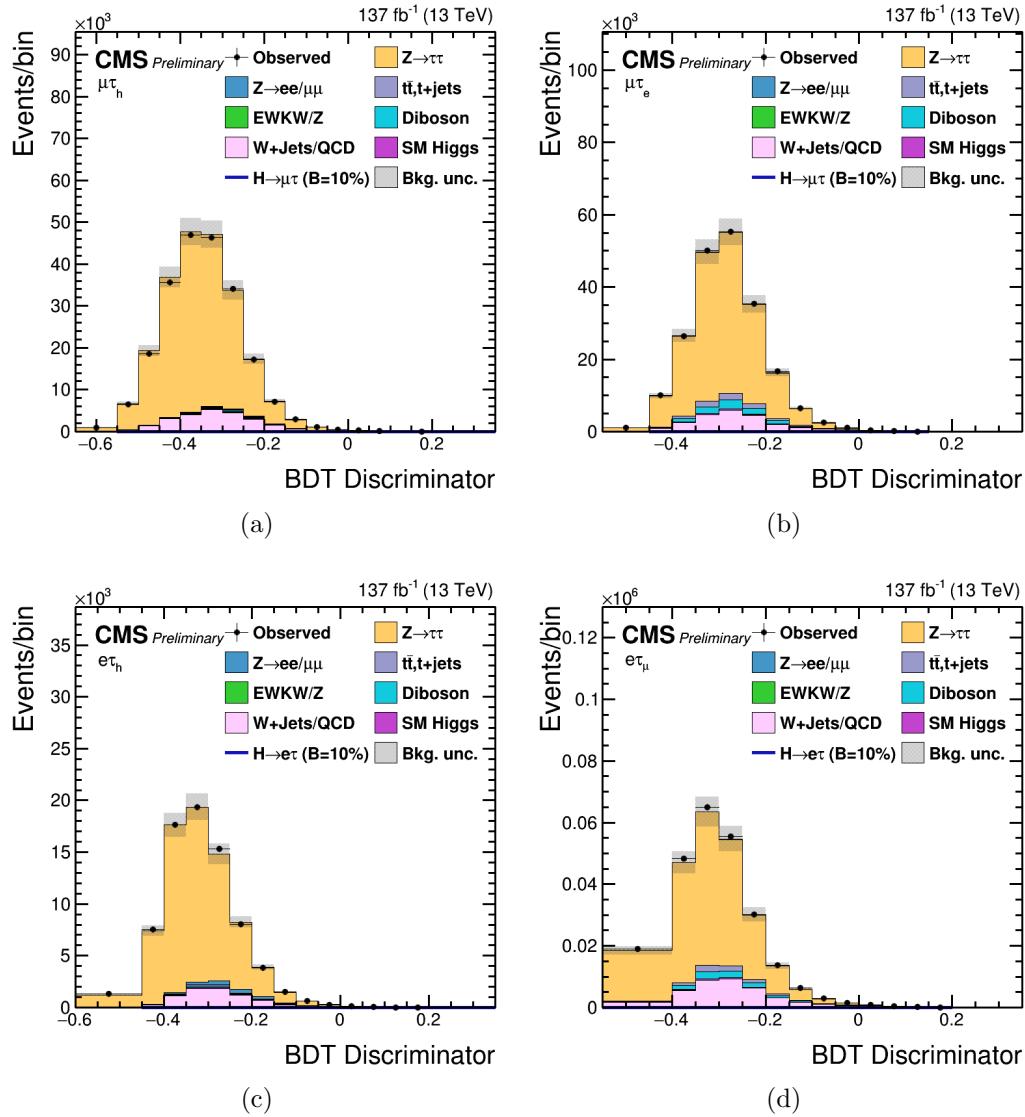


Figure 7.4. Distributions of BDT discriminator in the  $Z \rightarrow \tau\tau$  control regions for the (a)  $H \rightarrow \mu\tau_h$ , (b)  $H \rightarrow \mu\tau_e$ , (c)  $H \rightarrow e\tau_h$ , and (d)  $H \rightarrow e\tau_\mu$  channels.

applied to the control region, to estimate the misidentified background of the signal region. The signal region contrasted with the control regions used for determining the misidentified background can be seen in Figure 7.5.

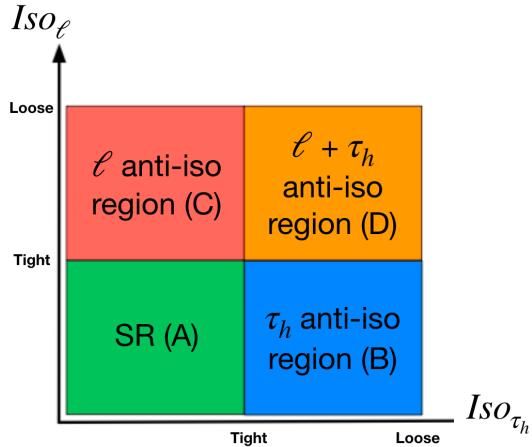


Figure 7.5. Signal region (green) contrasted with the control regions used for estimating the misidentified background

The probabilities with which jets are misidentified as an electron, muon, or hadronically decaying tau are labeled as  $f_e$ ,  $f_\mu$ , and  $f_{\tau_h}$ , respectively. The Z boson candidate is formed using two muons with  $p_T^\mu > 26 \text{ GeV}$  and  $|\eta| < 2.4$  and  $I_{\text{rel}}^\ell < 0.15$  for measuring the jet  $\rightarrow \tau_h, \mu, e$  misidentification rate. The muons are required to be oppositely charged and have their invariant mass ( $M_{\mu\mu}$ ) between 70 and 110 GeV.

The contribution from diboson events, where the jet candidate corresponds to a real lepton, is subtracted using simulation. The jet is required to pass the same lepton identification criteria as used in the signal region. A “signal-like” sample is defined if the jet passes the tight lepton isolation, else a “background-like” sample is

defined if it only passes the looser lepton isolation. These two samples are used to estimate  $f_e$ ,  $f_\mu$ , and  $f_{\tau_h}$  using the following:

$$f_i = \frac{N_i(\text{signal-like})}{N_i(\text{background-like}) + N_i(\text{signal-like})}$$

where  $N_i(\text{signal-like})$  is the number of events with a third lepton candidate that passes the tight lepton isolation, while  $N_i(\text{background-like})$  is the number of events that pass only the looser lepton isolation and index  $i = e, \mu$ , or  $\tau$ . The lepton selection criteria is summarized in Tables 6.1 and 6.2.

To estimate the misidentified  $\mu$  and  $e$  contribution in the background-like category, lepton isolation is required to be  $0.15 < I_{\text{rel}}^\mu < 0.25$  and  $0.15 < I_{\text{rel}}^e < 0.5$ , respectively. The misidentification rate is computed as a function of the lepton  $p_T$ . To estimate the  $\tau_h$  misidentified contribution,  $\tau_h$  candidates are required to pass the loose Working Point (WP) of Deep Neural Network (DNN) discrimination against jets but fail the tight WP used for the signal selection. The  $\tau_h$  misidentification rate shows a  $p_T$  dependence that varies with the  $\tau$  decay mode and  $|\eta|$  and are thus evaluated as a function of  $p_T^\tau$  for the different decay modes and two  $|\eta|$  regions ( $|\eta| < 1.5$  or  $|\eta| > 1.5$ ).

In the  $H \rightarrow e\tau_h$  channel, the  $\tau_h$  misidentification rate is evaluated using events with a  $Z$  boson candidate that is formed using two electrons with  $p_T^e > 27 \text{ GeV}$  and  $|\eta| < 2.5$  and  $I_{\text{rel}}^\ell < 0.15$ . The electrons are required to be oppositely charged and have their invariant mass ( $M_{ee}$ ) between 70 and 110 GeV. The reason for using  $Z \rightarrow ee$  events for evaluating the  $\tau_h$  misidentification rate in  $H \rightarrow e\tau_h$  channel is because the DNN WPs used for discriminating  $\tau_h$  against electrons and muons is different in this channel compared to the  $H \rightarrow \mu\tau_h$  channel. The misidentification rates that are evaluated using this control region are compatible with the measurement in  $Z \rightarrow \mu\mu$  events.

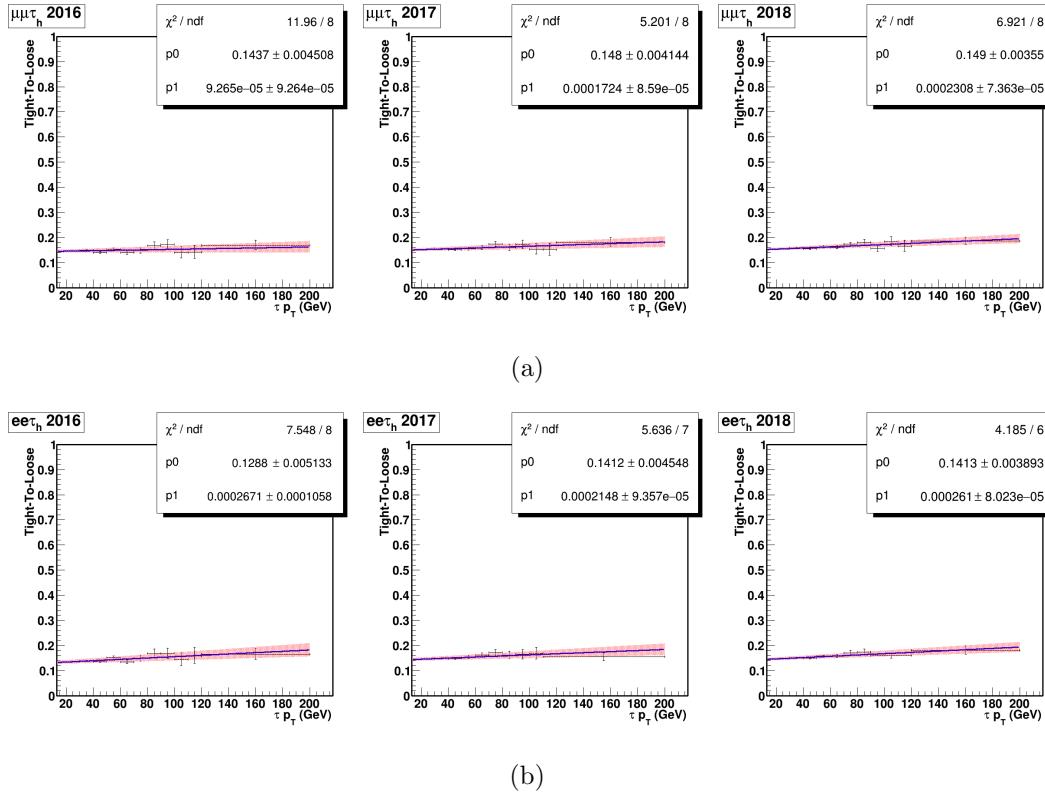
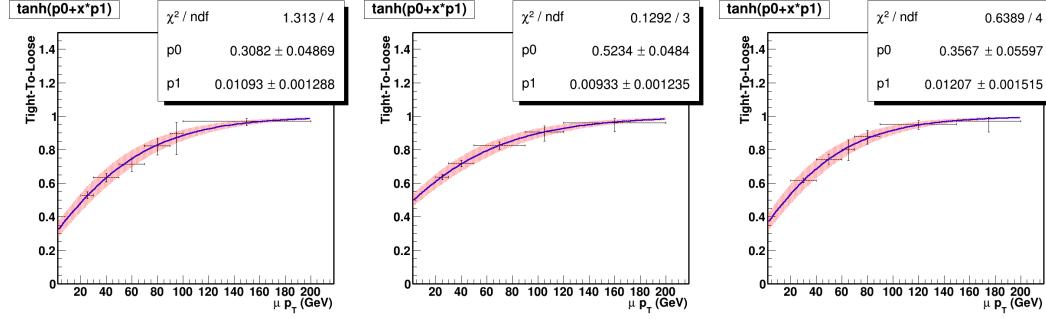
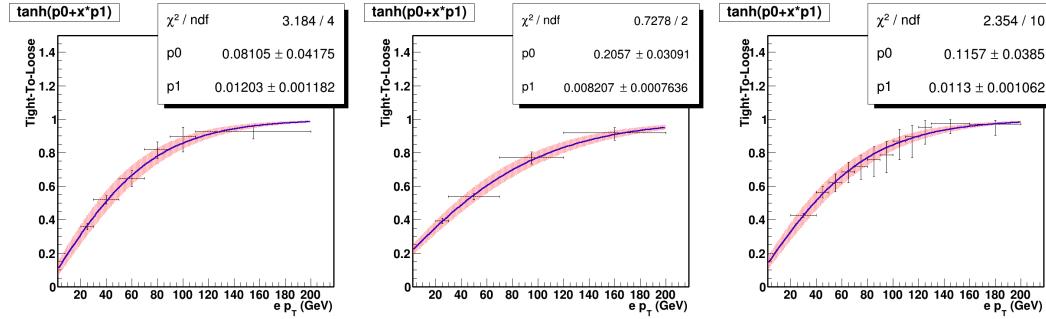


Figure 7.6. Fit performed to  $\tau_h$  misidentification rates for  $H \rightarrow \mu\tau_h$  (a) and  $H \rightarrow e\tau_h$  (b) channel as a function of  $\tau_h p_T$  for the different years. The misidentification rates used are further parametrized based on  $\tau_h$  Decay Mode along with the pseudorapidity of  $\tau_h$ . However, here only the inclusive misidentification rates are shown. The misidentification rates are labeled as “tight-to-loose” to clarify that they are calculated as a ratio of the number of events passing the tight WP to the loose WP of DNN discrimination against jets.



(a)



(b)

Figure 7.7. Fit performed to the  $\mu$  (a) and  $e$  (b) misidentification rates as a function of their  $p_T$  for 2016 (Left), 2017 (Center), and 2018 (Right). The misidentification rates are labeled as “tight-to-loose” to clarify that they are calculated as a ratio of the number of events passing the tight isolation to the loose isolation. The hyperbolic tangent function is used for performing the fit.

Each event in the control region defined using the collision data with the same selection as the signal region, but loosening the isolation requirements on one of the leptons is then weighted by a factor  $f_i/(1 - f_i)$  depending on the lepton  $p_T$  for electrons and muons or  $p_T$ ,  $\eta$  and decay mode for the  $\tau$  lepton candidates. Both background yields and shape distributions are thus estimated. Events with the possibility of double-counting due to two misidentified leptons are subtracted using a weight. For example, events with a misidentified  $\mu(e)$  and a misidentified  $\tau_h$  are subtracted in the  $H \rightarrow \mu\tau_h$  ( $H \rightarrow e\tau_h$ ) channel using a weight,  $f_\tau f_\ell / [(1 - f_\tau)(1 - f_\ell)]$ , where  $\ell = \mu$  or  $e$ .

The estimation of the background is validated in events where the two leptons have the same electric charge. The misidentification rate  $f_i$  is applied to events passing preselection and by inverting the lepton pair's charge requirement. The same-sign selection enhances the misidentified lepton background. The background estimation is also validated in a  $W$  boson enriched control sample. This control region is obtained by applying the signal-like requirement and  $M_T(\ell, \vec{p}_T^{\text{miss}}) > 60 \text{ GeV}$  ( $\ell = e$  or  $\mu$ ) and  $M_T(\tau_h, \vec{p}_T^{\text{miss}}) > 80 \text{ GeV}$ . The same strategy is applied in the  $H \rightarrow e\tau_h$  channel, which results in a similar agreement. Figures 7.8 and 7.9 show the comparison of data with background estimates in the same-sign and  $W$  boson enriched control regions for the  $H \rightarrow \mu\tau_h$  and  $H \rightarrow e\tau_h$  channels.

### 7.3.2 Semi data-driven approach

In the  $H \rightarrow e\tau_\mu$  and  $H \rightarrow \mu\tau_e$  channels, QCD multijet background is estimated from data using events with an electron and a muon with the same electric charge. These events are selected by applying the preselection except for requiring both the leptons to have the same electric charge, and we call this the same sign (SS) control region. Contributions from other processes are estimated from simulation and subtracted from data in this SS control region. Extrapolation factors from the SS control

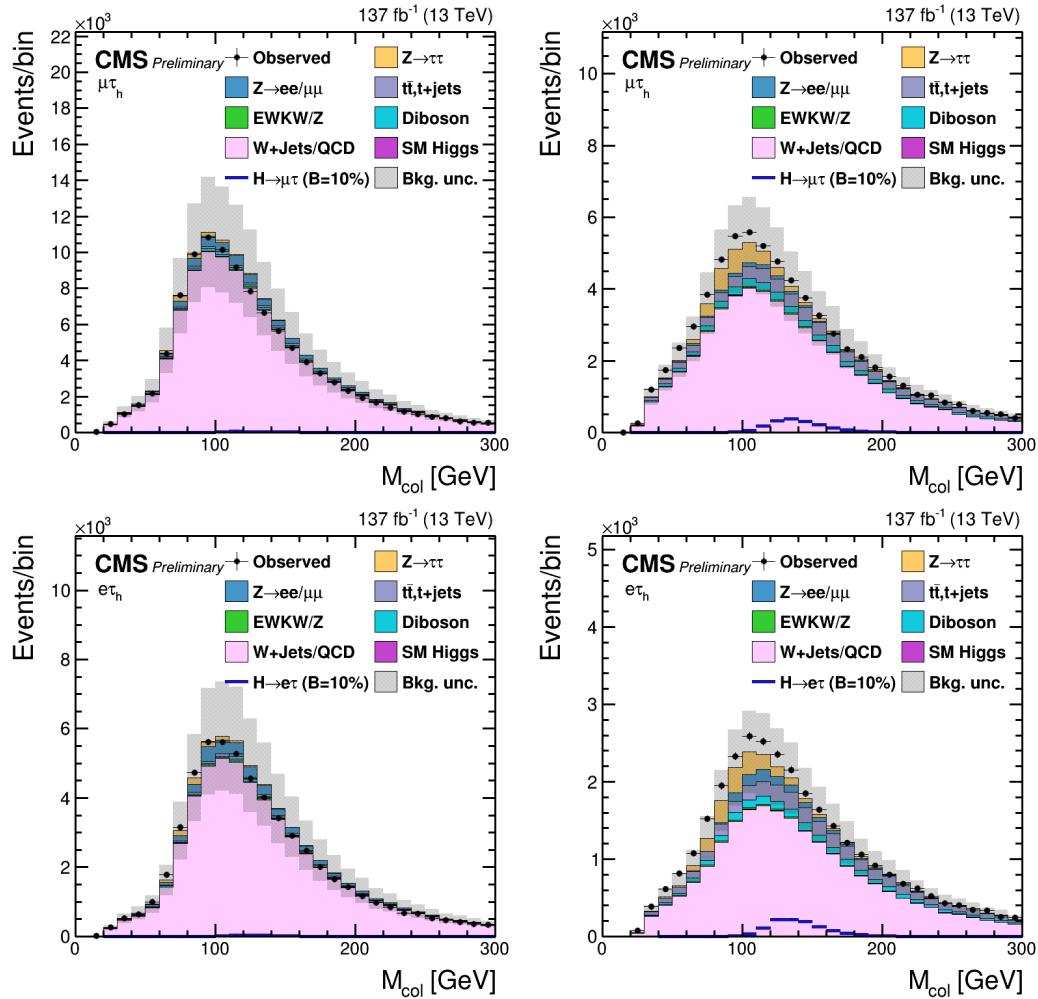


Figure 7.8. Distributions of  $M_{\text{col}}$  discriminator in the same-sign (Left) and W boson enriched (Right) control regions for the  $H \rightarrow \mu\tau_h$  (top) and  $H \rightarrow e\tau_h$  (bottom) channels.

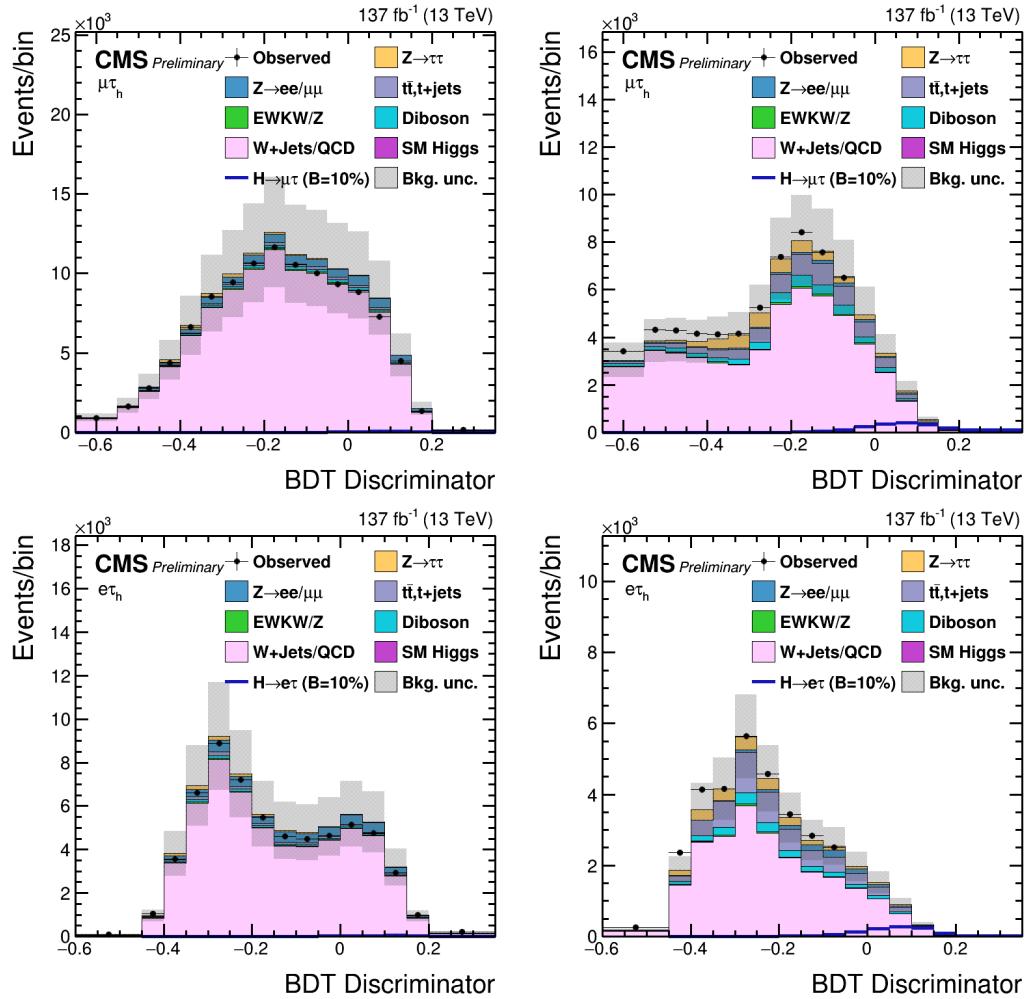


Figure 7.9. Distributions of BDT discriminator in the same-sign (Left) and  $W$  boson enriched (Right) control regions for the  $H \rightarrow \mu\tau_h$  (top) and  $H \rightarrow e\tau_h$  (bottom) channels.

region to the opposite sign (OS) signal region are measured in data as a function of the jet multiplicity and the  $\Delta R$  separation between the electron and the muon. QCD OS/SS extrapolation factors that are measured can be seen in Figure 7.10.

The OS/SS extrapolation factor is estimated using events with an anti-isolated muon and an isolated electron. The contribution from  $b\bar{b}$  events to the QCD multijet background gives rise to the  $\Delta R$  dependency and is parameterized with a linear function. The OS/SS extrapolation factor is higher for events with low  $\Delta R$  separation between the electron and the muon, decreasing as the  $\Delta R$  separation increases. The OS/SS extrapolation factor also depends on the electron and muon  $p_T$ . This  $p_T$  dependence comes from the leptons arising from the semi-leptonic c quark decay. These leptons tend to be softer in  $p_T$  and less isolated resulting in a reduction in the number of such events passing the  $p_T$  and isolation requirements. Corrections of the QCD OS/SS extrapolation factors dependent on lepton  $p_T$  along with the correction to account for the mismodeling introduced by anti-isolating the muon to measure them can be seen in Figure 7.11.

As the OS/SS extrapolation factor is measured in a control region where the muon is anti-isolated, an additional correction is applied to cover for a potential mismodeling. This correction is calculated by measuring the OS/SS extrapolation factors in two different control regions. The first control region has events where the muon is isolated, and the electron is anti-isolated. The second control region has events where both the electron and the muon are anti-isolated. The ratio of the extrapolation factors measured in these two control regions is taken as the correction for accounting the potential mismodeling induced by anti-isolating the muon. Figure 7.12 shows the comparison of data with background estimates in the muon anti-isolated control regions for the  $H \rightarrow \mu\tau_e$  channel.

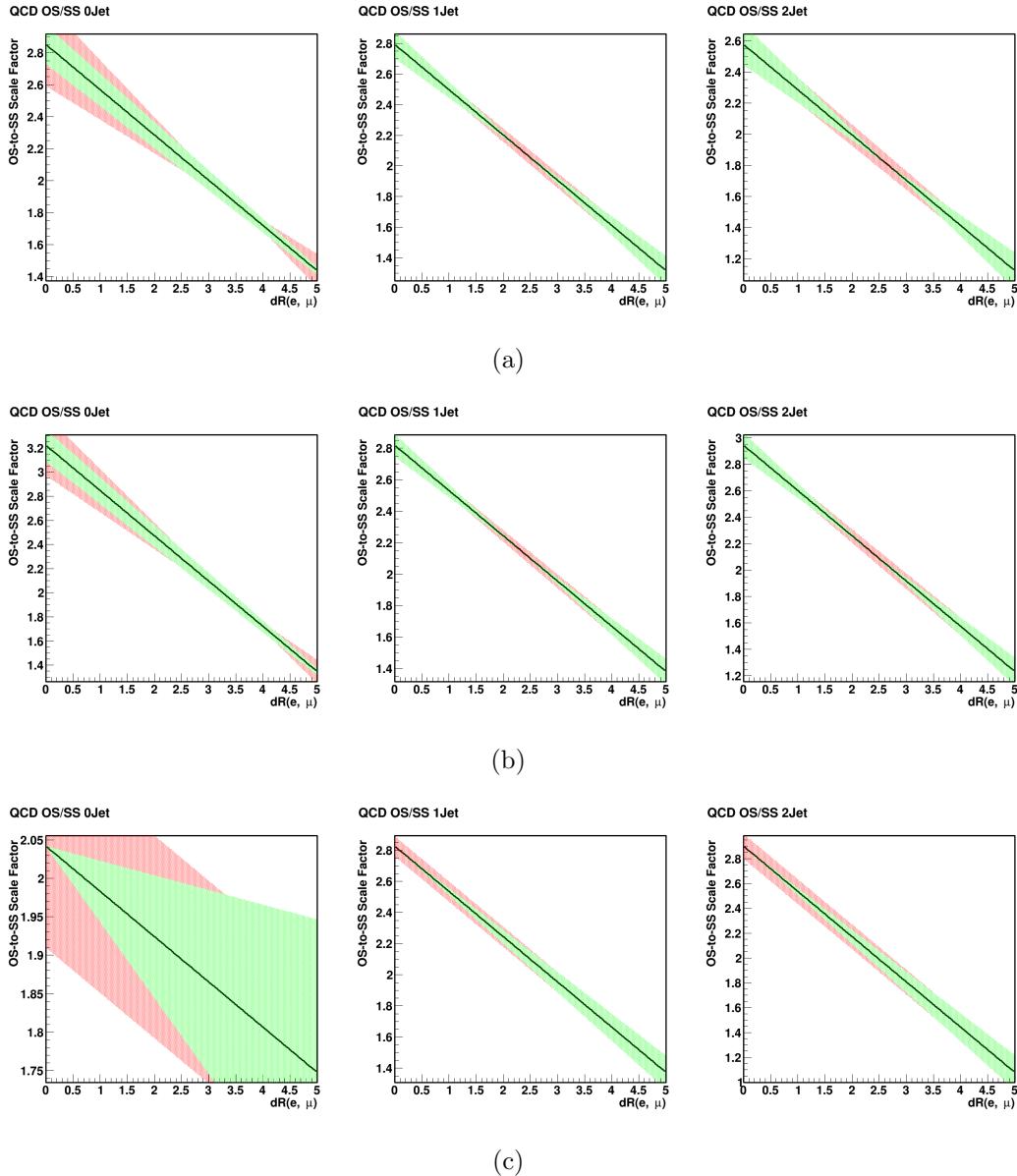


Figure 7.10. QCD OS/SS extrapolation factors in events with 0 Jets (Left), 1 Jet (Center), and 2 Jets (Right) for 2016 (a), 2017 (b), 2018 (c). The line is the best fit, and the shaded region corresponds to the shape uncertainties.

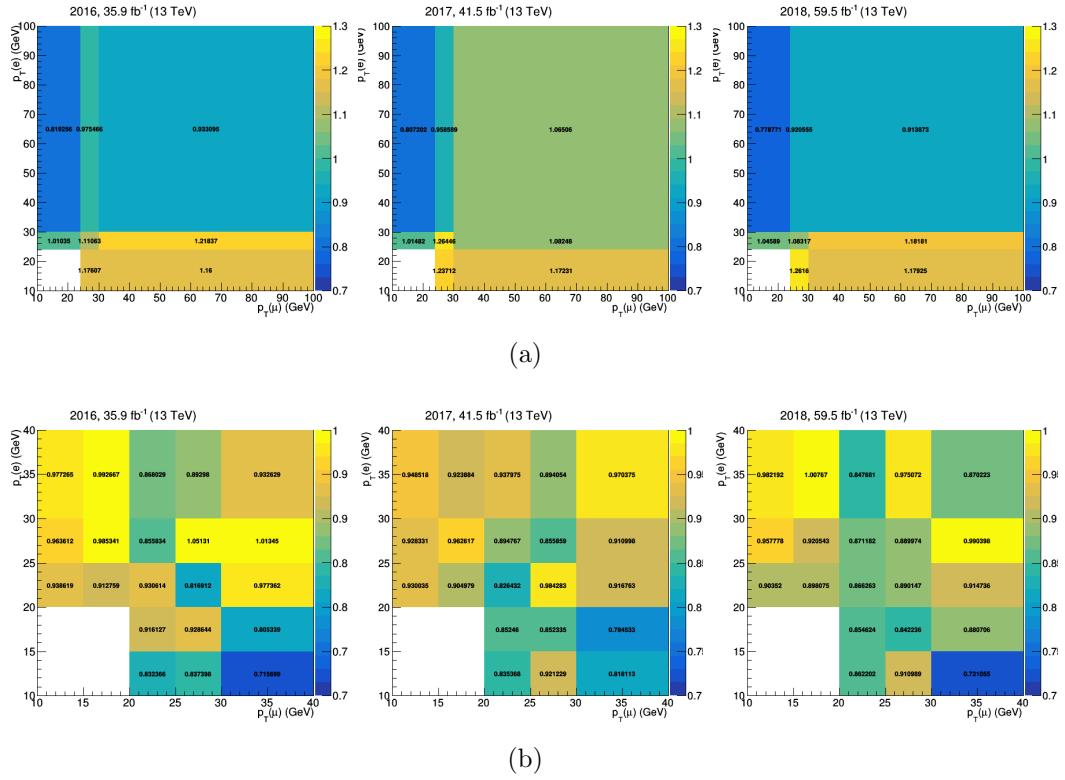


Figure 7.11. (a) Corrections of the QCD OS/SS extrapolation factors determined in the region with an anti-isolated muon as a function of the  $p_T$  of the electron and the muon, using data collected in 2016, 2017, and 2018. (b) Correction of the QCD OS/SS extrapolation factors to account for the mismodeling introduced by anti-isolating the muon to measure the SFs, using data collected in 2016, 2017, and 2018.

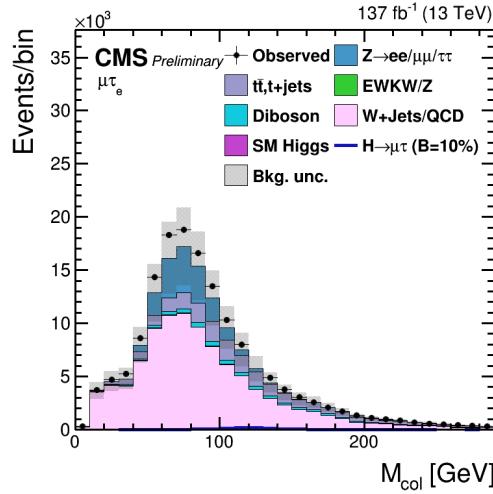


Figure 7.12. Distribution of  $M_{\text{col}}$  discriminator in the muon anti-isolated control regions for the  $H \rightarrow \mu\tau_e$  channel.

## 7.4 MC Simulation

All the other backgrounds are estimated using MC simulation. In the leptonic channels, the  $t\bar{t}$  process has a dominant contribution, and this background is validated in a dedicated control region defined by requiring the presence of at least one b jet tagged by the DeepCSV algorithm in the event. Figure 7.13 shows the background validation in this control region for the  $H \rightarrow \mu\tau_e$  and  $H \rightarrow e\tau_\mu$  channels.

The standard model, Higgs boson production, forms a small but non-negligible background. The contributions come mainly from  $H \rightarrow \tau\tau$  and  $H \rightarrow WW$  decays. The contribution from  $H \rightarrow WW$  peaks at lower values than the signal in the distribution of the BDT discriminator due to the presence of additional neutrinos in the decay. The background is estimated from simulations using selection based on the BDT discriminator and kinematic variables, as described in Chapter 6.

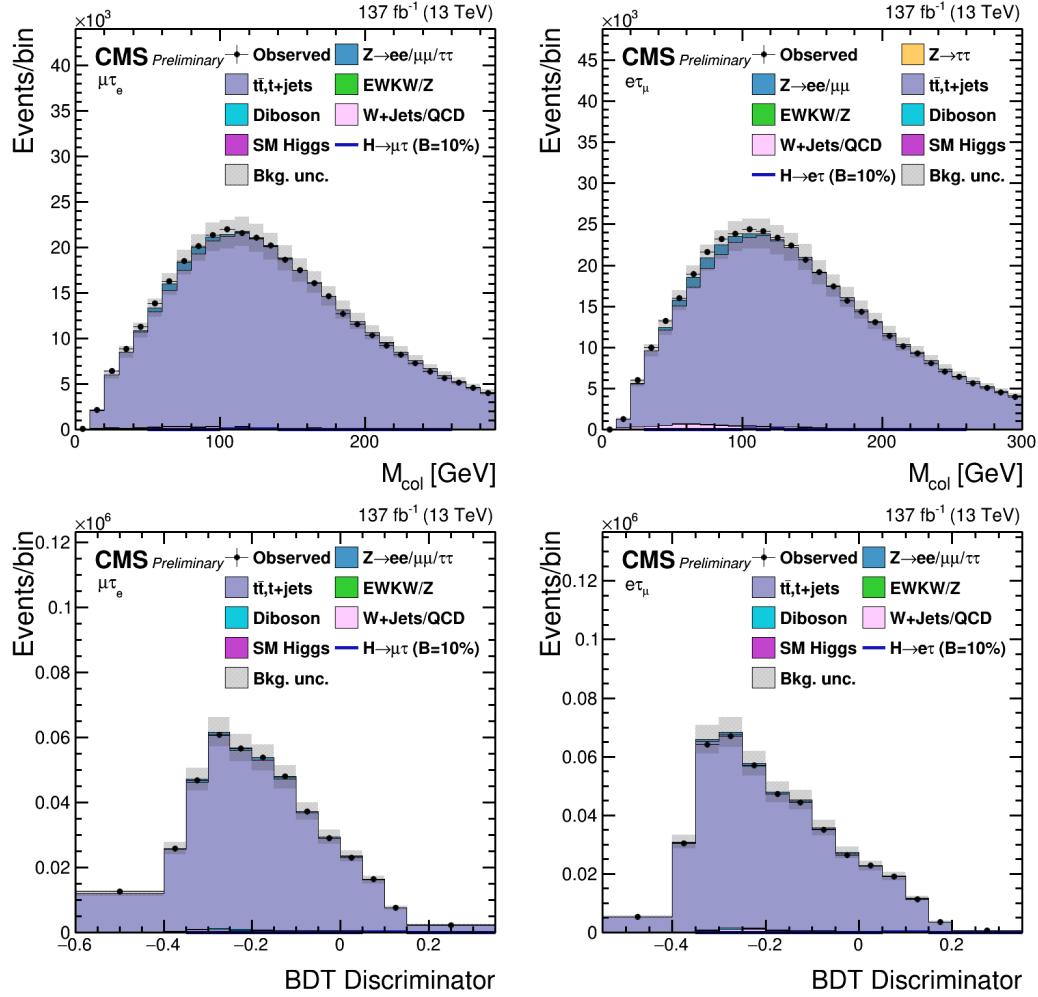


Figure 7.13. Distributions of  $M_{\text{col}}$  (BDT) discriminator in  $t\bar{t}$  enriched control region for  $H \rightarrow \mu\tau_e$  and  $H \rightarrow e\tau_\mu$  channel.

## CHAPTER 8

### STATISTICAL METHODS AND SYSTEMATIC UNCERTAINTIES

#### 8.1 Introduction

In the search for LFV Higgs decays, a discovery will be the observation of events with the Higgs Boson decaying either into  $\mu\tau$  or  $e\tau$ . As the LFV is forbidden in SM, the SM can be taken as the background model, and discovery can be claimed if the observation is not compatible with this background model. The uncertainties resulting from theoretical, experimental, and statistical sources can give rise to an excess in the observation even when there is no signal at all. When an excess is observed, a p-value is computed. The p-value represents the probability that an observed excess is due to statistical fluctuations. The p-value has to be very low to indicate that the excess observed is due to a signal's presence and not merely a statistical fluctuation. However, if no excess is observed, upper exclusion limits can be set on the branching fraction. This chapter describes the statistical methods used for extracting the signal strength, followed by the systematic uncertainties involved in the analysis.

#### 8.2 Statistical methods

The results on the branching fraction of the LFV Higgs-boson decays to  $\mu\tau$ , and  $e\tau$  are estimated using a profile likelihood method. The SM Higgs boson production cross-sections are used for the signal model, while the branching fraction of the Higgs boson to  $\mu\tau$  and  $e\tau$  remain free parameters. The branching fraction is the

parameter of interest. In addition to it, the signal and background model contain in general nuisance parameters whose values must be derived from collision data. The profile likelihood method is implemented, assuming the asymptotic approximation. Distributions of the BDT discriminator and the collinear mass for signal and various background processes are fitted to collision data. Systematic uncertainties are represented as nuisance parameters, and they can affect the normalization or the shape of the distribution.

Poisson distribution can be used to model the expected number of events and the observed events for the situation at hand. The expected number of events is  $\mu \cdot s + b$ , where  $s$  is the expected signal event yields, and  $b$  is the expected background event yields. The parameter  $\mu$  is the signal strength modifier, which changes the signal production cross-sections of all the production mechanisms by exactly the same scale of  $\mu$ . The likelihood function measures the goodness of fit of a statistical model to a sample of data for given values of the unknown parameters. For our situation, the likelihood function  $\mathcal{L}(data|\mu)$  is then given by:

$$\mathcal{L}(data|\mu) = \prod_{i=1}^{bins} \frac{(\mu \cdot s_i + b_i)^{n_i}}{n_i!} e^{-\mu \cdot s_i - b_i} \quad (8.1)$$

where  $n_i$  is the number of events observed in the bin  $i$  of the distribution, and  $s_i$  and  $b_i$  are expected number of signal and background events in that bin, respectively.

The nuisance parameters which represent the systematic uncertainties are embedded into the likelihood function. The uncertainties considered are taken to be 100%-correlated or uncorrelated, thus ensuring that the likelihood function has a clean factorized form [77]. However, certain uncertainties have partial correlations across the years. There are also partial correlations among uncertainties between the embedded and MC samples. A partially correlated uncertainty is separated into 100%-correlated or uncorrelated components. The magnitude of the correlated com-

ponent will be  $\rho$ , and for the uncorrelated part, it will be  $\sqrt{1 - \rho^2}$ , where  $\rho$  is the magnitude of the correlation. For example, a 50% correlation will have a correlated component with 0.5 as magnitude and an uncorrelated part with 0.866 as magnitude.

The expected signal and background yields are effected by nuisance parameters and we parametrize them as  $s(\theta)$  and  $b(\theta)$ . There is a default value  $\tilde{\theta}$  for each component of  $\theta$ . The default value reflects our degree of belief on the real value of  $\theta$ . The probability distribution function (pdf)  $\rho(\theta|\tilde{\theta})$  can then be interpreted as a posterior distribution from measurements of  $\tilde{\theta}$ . Using Bayes' theorem:

$$\rho(\theta|\tilde{\theta}) = \rho(\tilde{\theta}|\theta) \cdot \pi_\theta(\theta), \quad (8.2)$$

The priors  $\pi_\theta(\theta)$  are taken as flat distributions representing no prior knowledge of  $\theta$ . The likelihood of the measurement can be constrained by using the pdf of  $\tilde{\theta}$ . After incorporating the nuisance parameters, the likelihood function now becomes:

$$\mathcal{L}(\text{data}|\mu, \theta) = \prod_{i=1}^{bins} \frac{(\mu \cdot s_i(\theta) + b_i(\theta))^{n_i}}{n_i!} e^{-\mu \cdot s_i(\theta) - b_i(\theta)} \cdot \rho(\tilde{\theta}|\theta) \quad (8.3)$$

If no excess is observed then upper exclusion limits can be set on the branching fraction using the CL<sub>s</sub> method [78–80]. According to Neyman-Pearson lemma, the likelihood ratio is the most powerful discriminator. The likelihood ratio is used generally in the searches at the LHC for hypothesis testing, and it uses profiling of nuisances. The test statistic is denoted by  $\tilde{q}_\mu$  and is given by:

$$\tilde{q}_\mu = -2 \ln \frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})}, \text{ with } 0 \leq \hat{\mu} \leq \mu \quad (8.4)$$

where  $\hat{\theta}_\mu$  refers to the conditional maximum likelihood estimators of  $\theta$ , while  $\hat{\mu}$  and  $\hat{\theta}$  refer to the global maximum likelihood estimators for  $\mu$  and  $\theta$ . The constraint  $0 \leq \hat{\mu}$  ensures that the signal rate cannot be negative. In contrast, the upper constraint

$\hat{\mu} \leq \mu$  is imposed to guarantee that upward fluctuations of data such that  $\hat{\mu} > \mu$  are not considered as evidence against the signal hypothesis.

The observed value of the test statistic,  $\tilde{q}_\mu^{obs}$ , is calculated for the signal strength  $\mu$ , using equation 8.4. Maximum likelihood estimators for the nuisance parameters, under the background-only and signal-plus-background hypotheses, are calculated and are denoted by  $\hat{\theta}_0^{obs}$  and  $\hat{\theta}_\mu^{obs}$  respectively. They are used to generate toy Monte Carlo pseudo-datasets. These pseudo datasets are used to construct pdfs of test statistics  $f(\tilde{q}_\mu|0, \hat{\theta}_0^{obs})$  and  $f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{obs})$  by treating them as if they were real data. An illustration of these distributions can be seen in Figure 8.1.

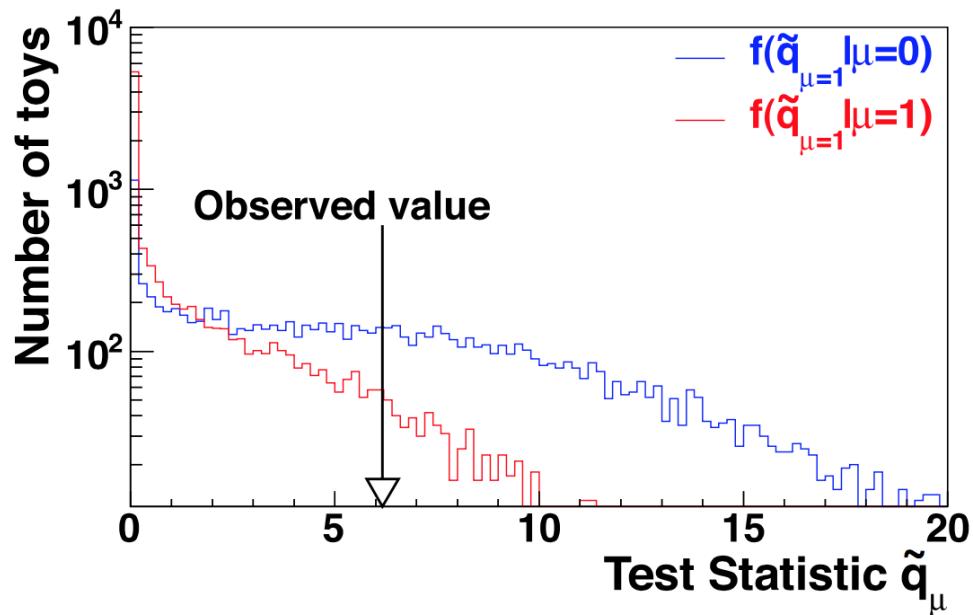


Figure 8.1: Test statistic distributions for ensembles of pseudo-data generated for background-only (blue) and signal-plus-background (red) hypotheses [77].

$CL_{s+b}$  and  $CL_b$  correspond to the probabilities of the observations under both hy-

potheses.  $\text{CL}_{s+b}$  measures the incompatibility of data under the signal-plus-background hypothesis.  $\text{CL}_b$  measures the incompatibility of data under the background hypothesis. These quantities alone are not adequate for hypothesis testing. For example, in situations when the signal is so small that both hypotheses are compatible with the observation and a downward fluctuation of the background can lead to an inference of signal. Also, the incompatibility of the data with the background-only hypothesis alone doesn't tell us that it is indeed compatible with the signal. The ratio of the two quantities referred to as  $\text{CL}_s$  helps deal with both the situations described above well. The 95% CL is then arrived at by iterating over  $\mu$  until we have  $\text{CL}_s = 0.05$ . The  $\mu$  thus obtained is denoted as  $\mu^{95\%CL}$  and is said to be excluded at 95% CL.

$$p_\mu = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs} | \text{signal-plus-background}) = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu | \mu, \hat{\theta}_\mu^{obs}) d\tilde{q}_\mu \quad (8.5)$$

$$1 - p_b = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs} | \text{background-only}) = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu | 0, \hat{\theta}_0^{obs}) d\tilde{q}_\mu \quad (8.6)$$

$$\text{CL}_s = \frac{p_\mu}{1 - p_b} \quad (8.7)$$

The discussion until now pertains to calculating the observed limits when the data is unblinded. However, when the analysis is performed in a blinded manner, we first calculate the expected limits, which are upper exclusion limits calculated using toy datasets of background-only expectation. An extensive set of pseudo-data is generated using the background-only hypothesis, and  $\text{CL}_s$  and  $\mu^{95\%CL}$  is calculated for each of them. A distribution is built as a function of the  $\mu^{95\%CL}$  calculated for each of these pseudo-data. We then calculate the median expected limit from the 50% quantile of the cumulative distribution function. The  $\pm 1\sigma$  and  $\pm 2\sigma$  bands are calculated similarly by integrating the distribution until the appropriate quantiles are reached. The expected limits can be used to maximize the sensitivity of the search.

A more stringent median limit corresponds to a more sensitive search.

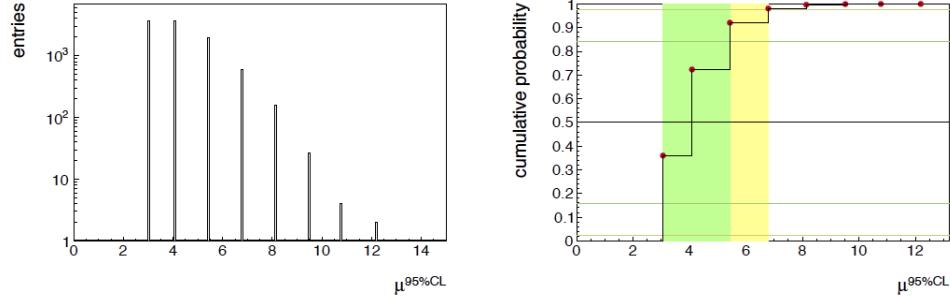


Figure 8.2: (Left) An example of differential distribution of possible limits on  $\mu$  for the background-only hypothesis. (Right) The cumulative probability distribution of the plot on the left with 2.5%, 16%, 50%, 84%, and 97.5% quantiles defines the median expected limit as well as the 68% and 95% bands for the expected value of  $\mu$  for the background-only hypothesis.

### 8.3 Systematic Uncertainties

Several sources of experimental and theoretical systematic uncertainties are taken into account as input to the maximum likelihood fit. These uncertainties affect both the normalization and the shape of the distributions of the different processes. The normalization effect is described by log-normal pdfs and corresponds to a multiplicative factor in the signal or background yields. These pdfs are characterized by the parameter  $\kappa$  and are well-suited for positively valued observables. The log-normal distribution looks like:

$$\rho(\theta|\tilde{\theta}) = \frac{1}{\sqrt{2\pi} \ln(\kappa)} \exp\left(-\frac{\ln(\theta/\tilde{\theta})^2}{2(\ln \kappa)^2}\right) \frac{1}{\theta} \quad (8.8)$$

The distributions' shape is altered when the systematic uncertainties affect the scale of the distribution differently in each bin. Such uncertainties are called shape uncertainties [81] and are modeled using a linear interpolation method [82]. Two other distributions obtained by varying the nuisance parameter by  $\pm 1$  standard deviation are used for implementing the shape uncertainties. A parameter is added to the likelihood that smoothly interpolates between these two other distributions.

As the analysis is categorized into different final states, partial and complete correlations between the uncertainties in different categories are treated appropriately and are summarized in Tables 8.1 and 8.2.

The uncertainties to reconstruct a  $\tau_h$  and estimation of its identification efficiency for different  $p_T$  ranges are measured using a tag-and-probe method and found to be in the range of 2-3%. The uncertainties for different ranges of  $p_T$  are treated in an uncorrelated way. These uncertainties are also considered for the embedded  $\tau\tau$  background, where they are treated 50% correlated with the simulation uncertainties. For the embedded samples, triggering on muons before being replaced by  $\tau$  leptons have an uncertainty of about 4%. This uncertainty is treated as uncorrelated between the three years due to different triggering criteria. Uncertainty due to tracking measurement is also considered in correlated way between  $h^\pm$  and  $h^\pm h^\mp h^\pm$  decay modes and uncorrelated way between  $h^\pm \pi^0$  and  $h^\pm h^\mp h^\pm \pi^0$  decay modes.

Uncertainties due to electrons and muons misidentified as  $\tau_h$  candidates correspond to 40% and between 10-70%, respectively, for different bins of  $p_T$ ,  $\eta$ , and  $\tau_h$  decay modes. The uncertainty due to  $\tau_h$  energy scale uncertainty is treated as uncorrelated for different decay modes and 50% correlated between embedded and simulated backgrounds and ranges from 0.7 to 1.2%. The uncertainty due to the electron (muon) energy scale for misidentified leptons is independent of the  $\tau_h$  energy scale and amount to 7% (1%). The effect of lepton energy resolution is found to be negligible for the study under consideration.

TABLE 8.1

SYSTEMATIC UNCERTAINTIES IN THE EXPECTED EVENT  
YIELDS FOR THE  $H \rightarrow \mu\tau$  CHANNELS.

Systematic Uncertainty	$H \rightarrow \mu\tau_h$	$H \rightarrow \mu\tau_e$
Muon ID/Isolation	2%	2%
Electron ID/Isolation	—	2%
Trigger	2%	2%
$\tau_h$ Identification	$p_T$ dep. (2–3%)	—
$\mu \rightarrow \tau_h$ ID	10–70%	—
$e \rightarrow \tau_h$ ID	—	—
b-tagging veto	6.5%	6.5%
Prefiring	(0.2–1.3%)	(0.2–1.3%)
Jet energy scale	3–20%	3–20%
and resolution		
$\tau_h$ energy scale	0.7–1.2%	—
$e \rightarrow \tau_h$ energy scale	1–7%	—
$\mu \rightarrow \tau_h$ energy scale	1%	—
e energy scale	—	1–2.5%
$\mu$ energy scale	0.4–2.7%	0.4–2.7%
Unclustered energy scale	$\pm 1\sigma$	$\pm 1\sigma$

TABLE 8.2

SYSTEMATIC UNCERTAINTIES IN THE EXPECTED EVENT  
YIELDS FOR THE  $H \rightarrow e\tau$  CHANNELS.

Systematic Uncertainty	$H \rightarrow e\tau_h$	$H \rightarrow e\tau_\mu$
Muon ID/Isolation	—	2%
Electron ID/Isolation	2%	2%
Trigger	2% for SingleElectron, 5.4% for Cross-Trigger	2%
$\tau_h$ Identification	$p_T$ dep. (2–3%)	—
$\mu \rightarrow \tau_h$ ID	—	—
$e \rightarrow \tau_h$ ID	40%	—
b-tagging veto	6.5%	6.5%
Prefiring	(0.2–1.3%)	(0.2–1.3%)
Jet energy scale and resolution	3–20%	3–20%
$\tau_h$ energy scale	0.7–1.2%	—
$e \rightarrow \tau_h$ energy scale	1–7%	—
$\mu \rightarrow \tau_h$ energy scale	1%	—
e energy scale	1–2.5%	1–2.5%
$\mu$ energy scale	—	0.4–2.7%
Unclustered energy scale	$\pm 1\sigma$	$\pm 1\sigma$

The jet energy scale (JES) is affected by several sources, and uncertainty due to them is evaluated as a function of  $p_T$  and pseudorapidity. JES's effect is propagated to the BDT discriminator by varying each source of uncertainty by one standard deviation and disseminating them to the fit template for each process and are of the order of 3–20%. The jet energy resolution uncertainties are also taken into account, and they mostly impact the  $M_{jj}$ -defined categories. In the particle flow candidate list, jets with  $p_T < 10 \text{ GeV}$  are not considered. They fall under unclustered energy. The unclustered energy scale is considered independently for charged particles, neutral hadrons, photons, and very forward particles, affecting both shape and yield, and are treated as uncorrelated. The efficiency to classify a jet as b-tagged is different in data, and simulation and scale factors depend on jet  $p_T$  and  $\eta$  are used to correct simulation. The uncertainties in measuring these scale factors are treated as a systematic uncertainty.

The uncertainty to track leptons ( $e, \mu$ ), reconstruct them, and identify and isolate is measured using the tag-and-probe method in data in  $Z \rightarrow ee$  and  $Z \rightarrow \mu\mu$  events and sums up to about 2% [34, 83–85]. The uncertainty in the measurement of the muon momentum scale is in the range 0.4–2.7% for different  $\eta_\mu$ , while for the electron momentum scale, it is less than 1%. The selection of events using electron and muon based triggers results in additional 2% uncertainty in the yield of simulated processes.

The misidentification rates in the  $e\tau_h$  and  $\mu\tau_h$  final states are parameterized using a linear function on  $\tau_h p_T$ . There are a couple of uncertainties that are ascribed per fit function. The normalization uncertainties in the data-driven estimations of misidentified-lepton backgrounds (jet  $\rightarrow \tau_h, \mu, e$ ), is taken from the orthogonal control region as described in Chapter 7. Discriminators with different signal to background efficiency are used to single out  $\tau_h$  against electrons and muons, which entails an additional 3% uncertainty for the  $e\tau_h$  channel. The misidentified lepton background in the  $e\tau_\mu$  and  $\mu\tau_e$  final states is affected by different shape uncertainties. The statistical

uncertainties arising from both fits of the OS/SS extrapolation factor as a function of the spatial separation between  $e$  and  $\mu$  and lepton  $p_T$  are taken into account. The uncertainties from OS/SS extrapolation factor are estimated in the anti-isolated muon control region, which results in an additional uncertainty due to the inversion of muon isolation, with a combined effect of about 20% on the normalization. The predominant source of uncertainties in the simulated background processes,  $Z \rightarrow ee$ ,  $Z \rightarrow \mu\mu$ ,  $Z \rightarrow \tau\tau$ ,  $WW$ ,  $ZZ$ ,  $W\gamma$ ,  $t\bar{t}$ , and single top production is the measurement of the cross-section for these processes.

The theoretical uncertainties affecting the measurement of Higgs boson production cross-section are the factorization and the renormalization scales, choice of the parton distribution functions (PDFs), and the strong coupling constant ( $\alpha_S$ ). These uncertainties affect the normalization of the signal shape and are taken from Ref. [17]. The variations of the QCD scales results in a 3.9%, 0.5%, 0.9%, and 0.8% uncertainty for the ggH, VBFH, ZH, WH cross-sections, respectively, while variations of PDF+ $\alpha_s$  results in 3.2%, 2.1%, 1.3%, 1.9% uncertainty. The acceptance effects are also taken into account while varying the renormalization and factorization scales along with the PDF choice and  $\alpha_S$ .

The uncertainty on the  $H \rightarrow \tau\tau$  branching-fraction includes a 1.7% uncertainty due to missing higher-order corrections, a 0.99% parametric uncertainty on the quark masses, and a 0.62% parametric uncertainty on  $\alpha_S$ . The uncertainty on the  $H \rightarrow WW$  branching fraction includes a 0.99% uncertainty due to missing higher-order corrections, a 0.99% parametric uncertainty on the quark masses, and a 0.66% parametric uncertainty on  $\alpha_S$ .

The bin-by-bin uncertainties account for the statistical uncertainties in each bin of the template distributions of every process. The Barlow-Beeston Lite [81, 86] approach is used, which assigns a single nuisance parameter to scale the sum of the process yields in each bin, constrained by the total uncertainty, instead of requiring

separate parameters, one per process. This is useful in order to minimize the number of parameters needed in the maximum-likelihood fit. They are treated uncorrelated between bins, processes, and categories.

The uncertainty of  $2.3 - 2.5\%$  on the integrated luminosity affects all processes with the normalization taken directly from the simulation. Shape uncertainties related to the pileup have been considered by varying the weights applied to simulation. The weight variation is obtained by a 5% change of the total inelastic cross-section used to estimate the number of pileup events in data. Other minimum bias event modeling and simulation uncertainties are expected to be much smaller than those on the rate and are therefore neglected.

During the 2016 and 2017 data-taking, pre-firing has impacted the ECAL L1 trigger. A gradual shift in the timing of the inputs of the ECAL L1 trigger in the region at  $|\eta| > 2.0$  caused a specific trigger inefficiency. For events containing an electron (a jet) with  $p_T$  larger than  $\approx 50 \text{ GeV}$  ( $\approx 100 \text{ GeV}$ ), in the region  $2.5 < |\eta| < 3.0$  the efficiency loss is  $\approx 10\text{-}20\%$ , depending on  $p_T$ ,  $\eta$ , and time. Correction factors were computed from data and applied to the acceptance evaluated by simulation. Uncertainty due to this correction factor is of the order of 1%.

## CHAPTER 9

### RESULTS

After applying the selection criteria, the expected limits are calculated from a final discrimination variable using the profile likelihood method with asymptotic approximation. Two different fits are considered, with differences in the discrimination variable employed and changes to the selection. The primary results are obtained from a fit to the BDT discriminator after the loose selection. The cross-check results are obtained from a fit to the  $M_{\text{col}}$  distribution after applying cut-based selection criteria. The fits are performed per channel and category and then combined to set 95% CL upper limits on the branching fraction of LFV Higgs decay in the  $H \rightarrow \mu\tau$  and  $H \rightarrow e\tau$  channels,  $\mathcal{B}(H \rightarrow \mu\tau)$ , and  $\mathcal{B}(H \rightarrow e\tau)$ , respectively. The BDT discriminator distributions of signal and background for each category for  $H \rightarrow \mu\tau$  and  $H \rightarrow e\tau$  can be seen in Figures 9.1, 9.2, 9.5, and 9.6 respectively. The  $M_{\text{col}}$  discriminator distributions of signal and background for each category for  $H \rightarrow \mu\tau$  and  $H \rightarrow e\tau$  can be seen in Figures 9.3, 9.4, 9.7, and 9.8 respectively.

The BDT analysis yields a sensitivity on the Higgs branching fraction of 0.15% (0.16%) for  $H \rightarrow \mu\tau$  and 0.29% (0.19%) for  $H \rightarrow e\tau$ . The observed and median expected 95% CL upper limits and the best fit branching fractions, for  $\mathcal{B}(H \rightarrow \mu\tau)$  and  $\mathcal{B}(H \rightarrow e\tau)$ , assuming  $m_H = 125$  GeV, are reported in Tables 9.1 and 9.2. The limits are also summarized graphically in Figure 9.9 and in Table 9.5.

A cross-check analysis in which a maximum-likelihood fit is performed on the  $M_{\text{col}}$  distribution after applying additional selections [87] yields a sensitivity on the Higgs branching fraction of 0.38% (0.26%) for  $H \rightarrow \mu\tau$  and 0.32% (0.28%) for  $H \rightarrow e\tau$ .

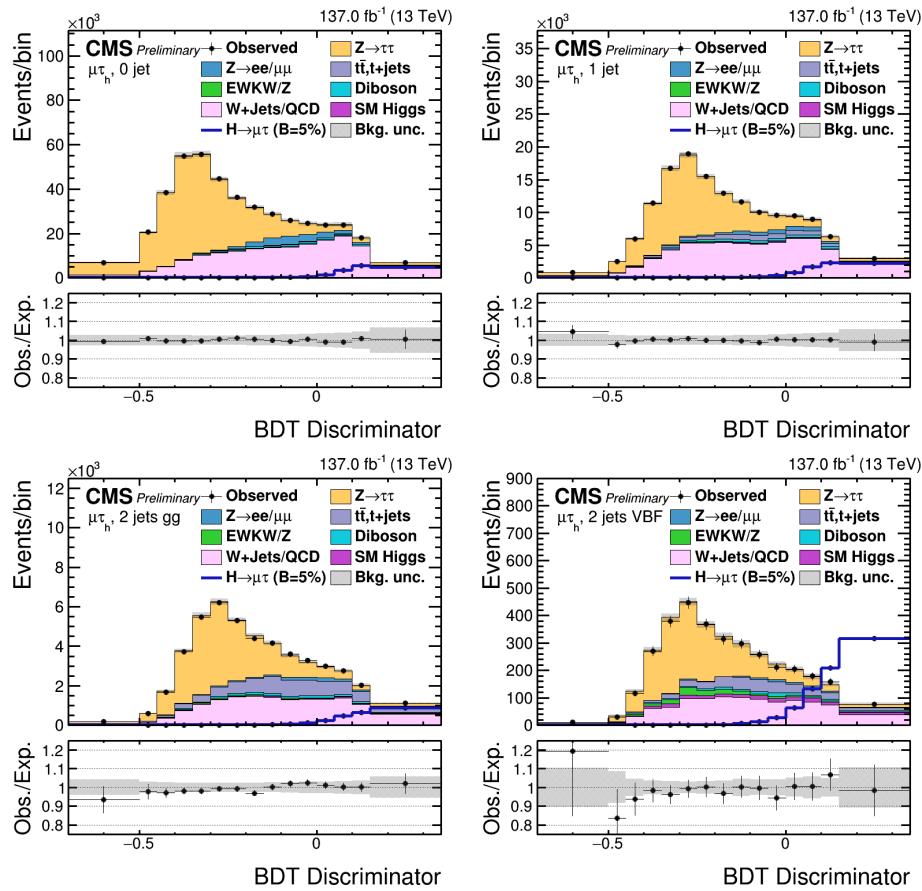


Figure 9.1. BDT discriminator distributions for the observed and estimated background in the  $H \rightarrow \mu\tau_h$  process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to  $\mathcal{B}(H \rightarrow \mu\tau) = 5\%$ .  $H \rightarrow \mu\tau_h$  channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.

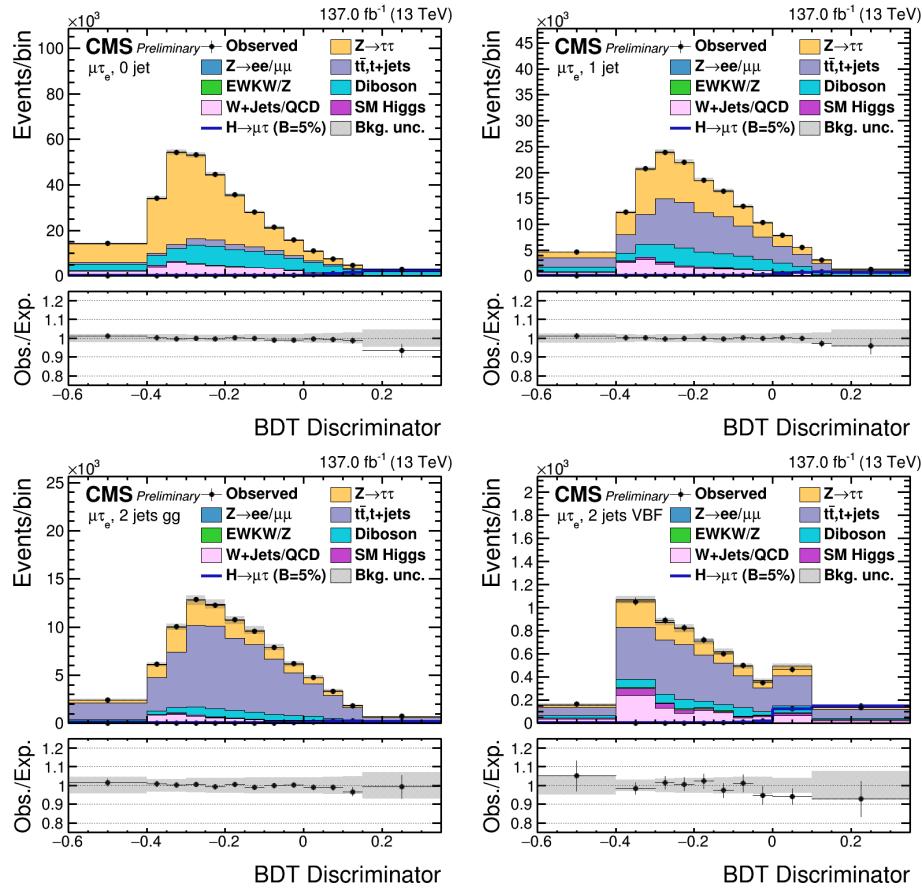


Figure 9.2. BDT discriminator distributions for the observed and estimated background in the  $H \rightarrow \mu\tau_e$  process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to  $\mathcal{B}(H \rightarrow \mu\tau) = 5\%$ .  $H \rightarrow \mu\tau_e$  channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.

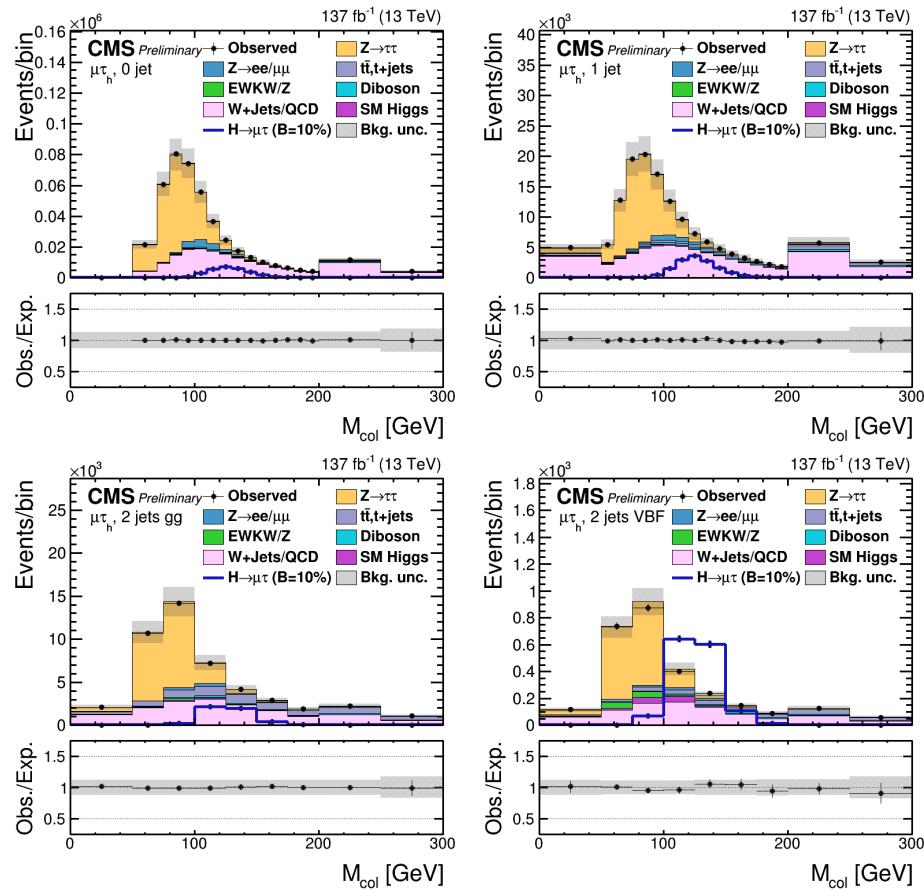


Figure 9.3.  $M_{\text{col}}$  distributions for the observed and estimated background in the  $H \rightarrow \mu\tau_h$  process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to  $\mathcal{B}(H \rightarrow \mu\tau) = 10\%$ .  $H \rightarrow \mu\tau_h$  channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.

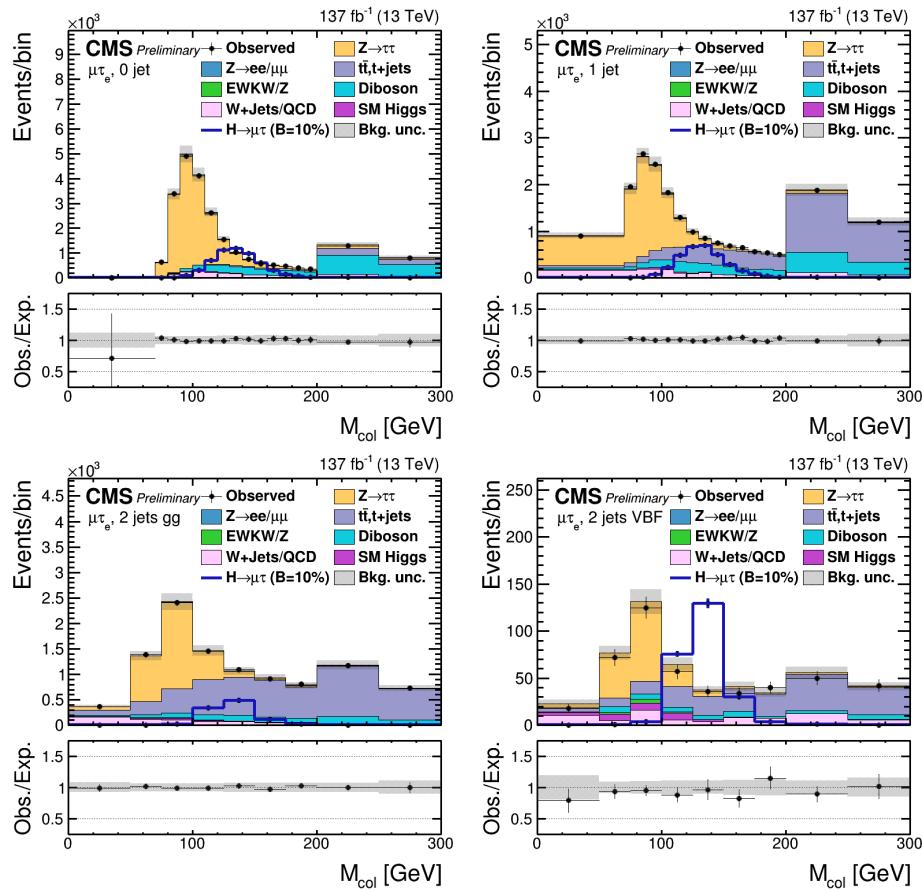


Figure 9.4.  $M_{\text{col}}$  distributions for the observed and estimated background in the  $H \rightarrow \mu\tau_e$  process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to  $\mathcal{B}(H \rightarrow \mu\tau) = 10\%$ .  $H \rightarrow \mu\tau_e$  channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.

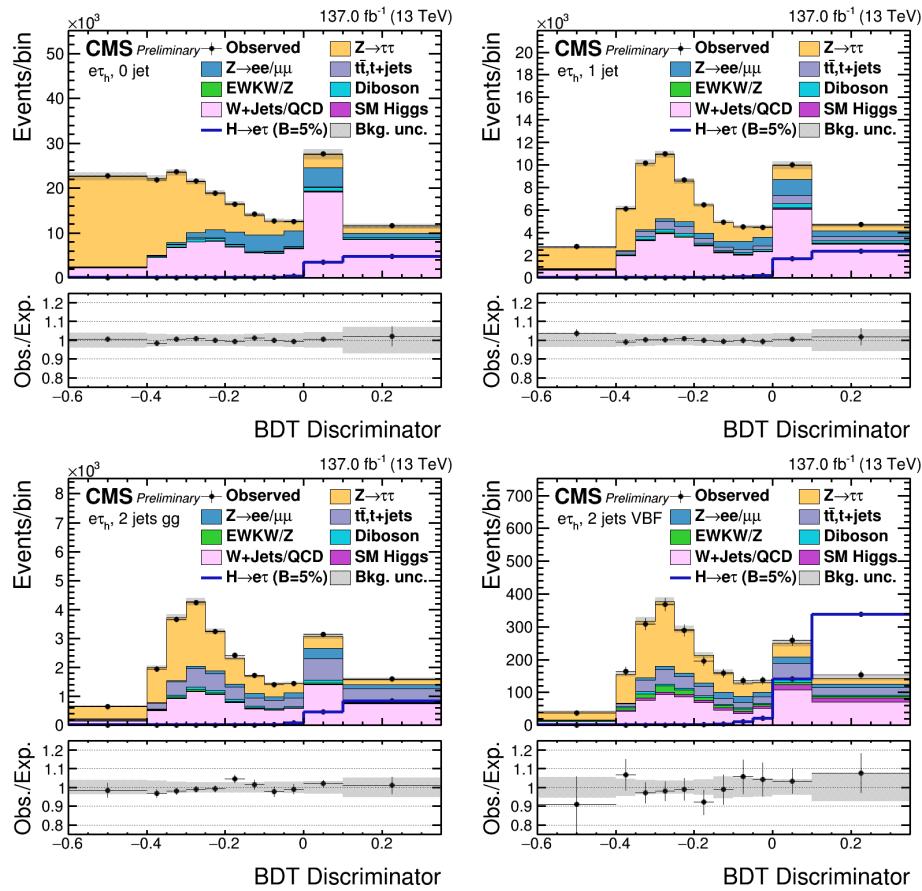


Figure 9.5. BDT discriminator distributions for the observed and estimated background in the  $H \rightarrow e\tau_h$  process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to  $\mathcal{B}(H \rightarrow e\tau) = 5\%$ .  $H \rightarrow e\tau_h$  channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.

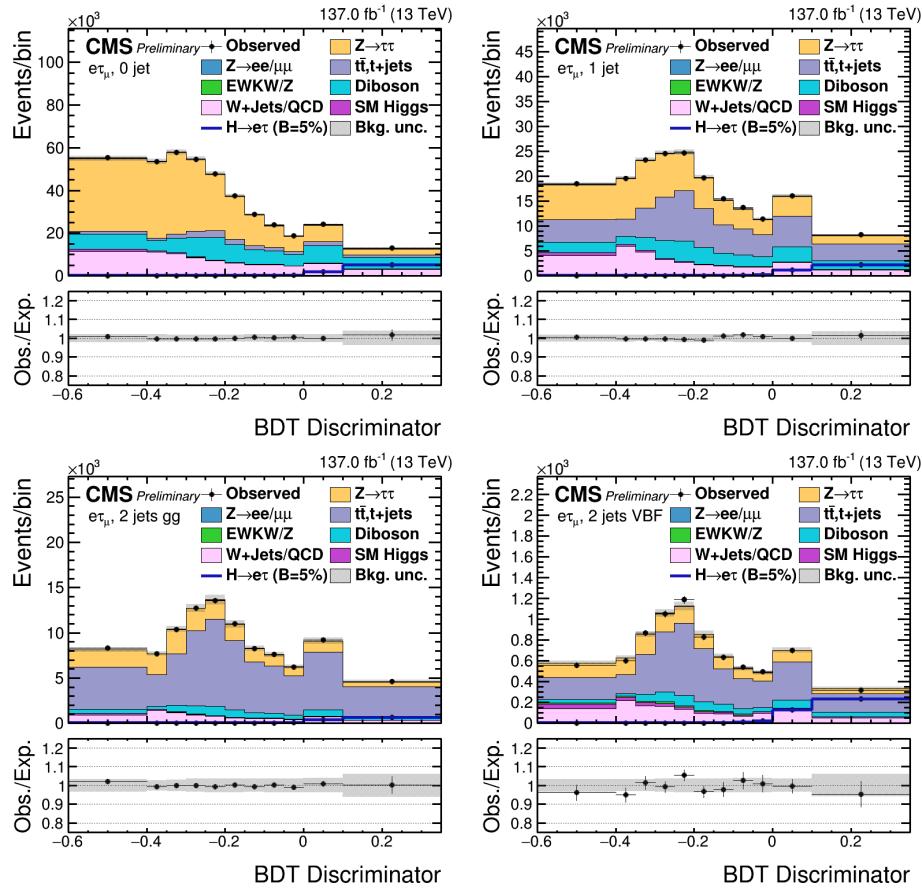


Figure 9.6. BDT discriminator distributions for the observed and estimated background in the  $H \rightarrow e\tau_\mu$  process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to  $\mathcal{B}(H \rightarrow e\tau) = 5\%$ .  $H \rightarrow e\tau_\mu$  channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.

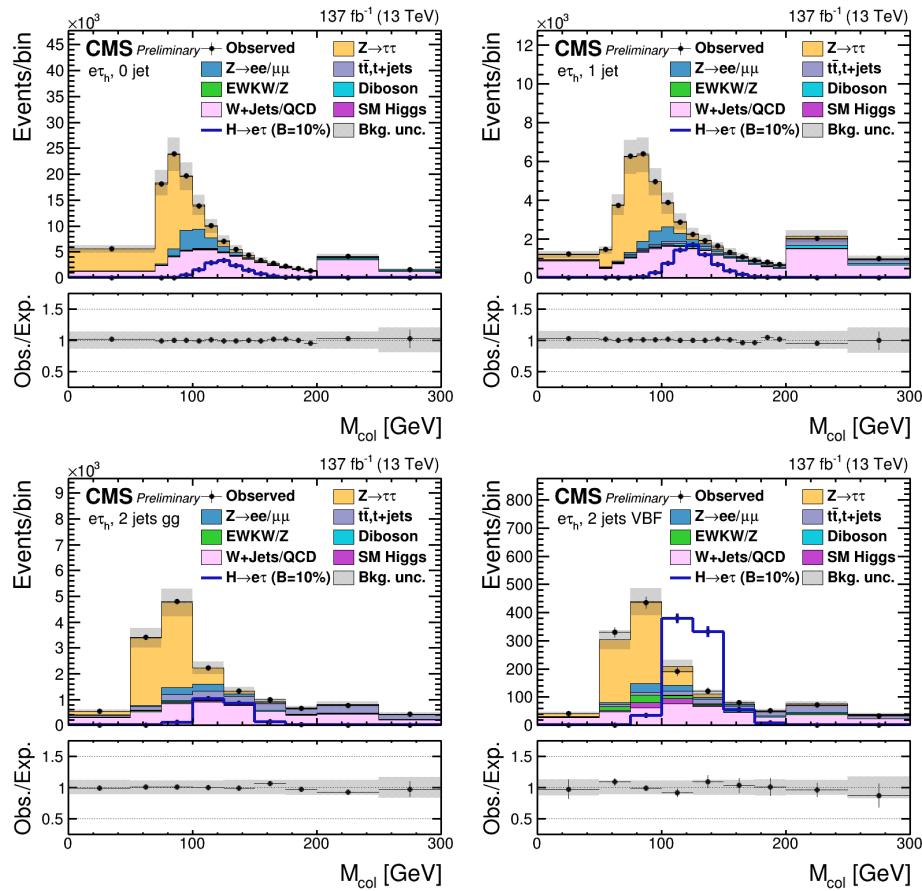


Figure 9.7.  $M_{\text{col}}$  distributions for the observed and estimated background in the  $H \rightarrow e\tau_h$  process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to  $\mathcal{B}(H \rightarrow e\tau) = 10\%$ .  $H \rightarrow e\tau_h$  channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.

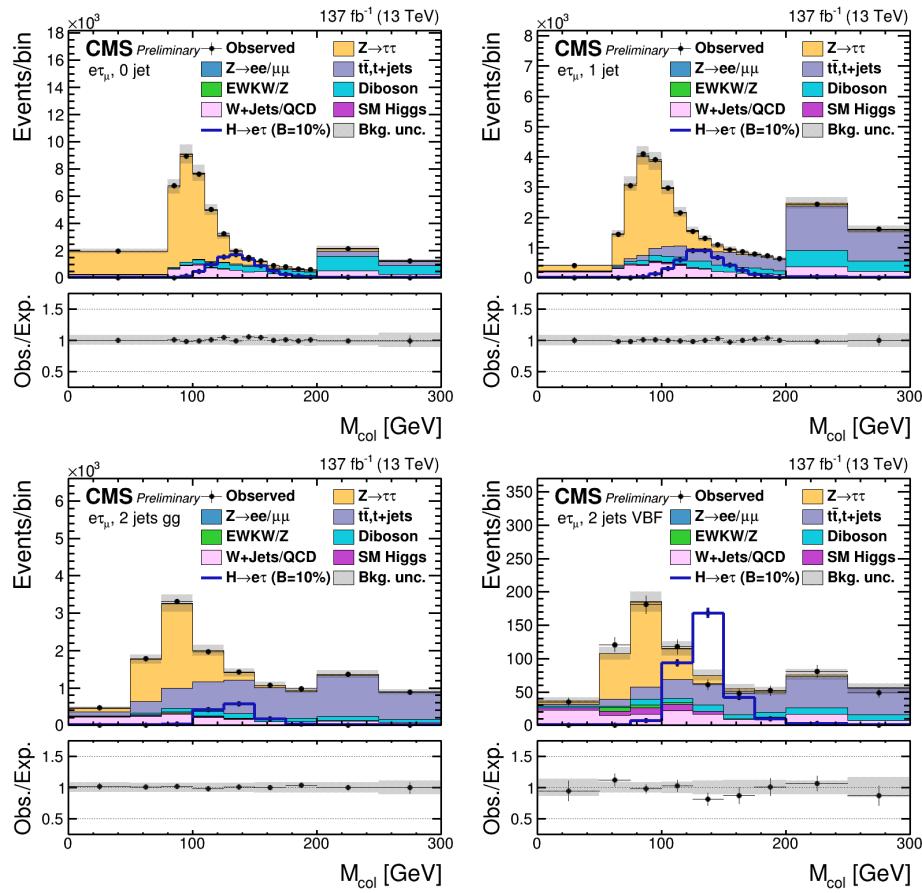


Figure 9.8.  $M_{\text{col}}$  distributions for the observed and estimated background in the  $H \rightarrow e\tau_\mu$  process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to  $\mathcal{B}(H \rightarrow e\tau) = 10\%$ .  $H \rightarrow e\tau_\mu$  channel categories are 0 jets (top left), 1 jet (top right), 2 jets gg (bottom left), and 2 jets VBF (bottom right). The bottom panel in each plot shows the fractional difference between the observed and estimated background. The uncertainty band shows the post fit statistical and systematic uncertainties added in quadrature.

The observed and median expected 95% CL upper limits and the best fit branching fractions, for  $\mathcal{B}(H \rightarrow \mu\tau)$  and  $\mathcal{B}(H \rightarrow e\tau)$ , assuming  $m_H = 125$  GeV, are reported in Tables 9.3 and 9.4. The limits are also summarized graphically in Figure 9.10. The BDT fit analysis is more sensitive than the  $M_{\text{col}}$  fit analysis, and systematic uncertainties dominate results for both cases.

The upper limits on  $\mathcal{B}(H \rightarrow \mu\tau)$  and  $\mathcal{B}(H \rightarrow e\tau)$  are subsequently used to put constraints on LFV Yukawa couplings [6]. The LFV decays  $e\tau$  and  $\mu\tau$  arise at tree level from the assumed flavor violating Yukawa interactions,  $Y_{\ell^\alpha \ell^\beta}$ , where  $\ell^\alpha, \ell^\beta$  are the leptons ( $e, \mu, \tau$ ) of different flavors ( $\alpha \neq \beta$ ). The decay width  $\Gamma(H \rightarrow \ell^\alpha \ell^\beta)$  in terms of the Yukawa couplings is given by:

$$\Gamma(H \rightarrow \ell^\alpha \ell^\beta) = \frac{m_H}{8\pi} (|Y_{\ell^\alpha \ell^\beta}|^2 + |Y_{\ell^\beta \ell^\alpha}|^2),$$

and the branching fractions is given by:

$$\mathcal{B}(H \rightarrow \ell^\alpha \ell^\beta) = \frac{\Gamma(H \rightarrow \ell^\alpha \ell^\beta)}{\Gamma(H \rightarrow \ell^\alpha \ell^\beta) + \Gamma_{\text{SM}}}.$$

The SM H decay width is assumed to be  $\Gamma_{\text{SM}} = 4.1$  MeV [88] for  $m_H = 125$  GeV. The 95% CL upper limit on the Yukawa couplings derived from the expression for the branching fraction above is shown in Table 9.5. The limits on the Yukawa couplings derived from the BDT fit analysis results are shown in Figure 9.11.

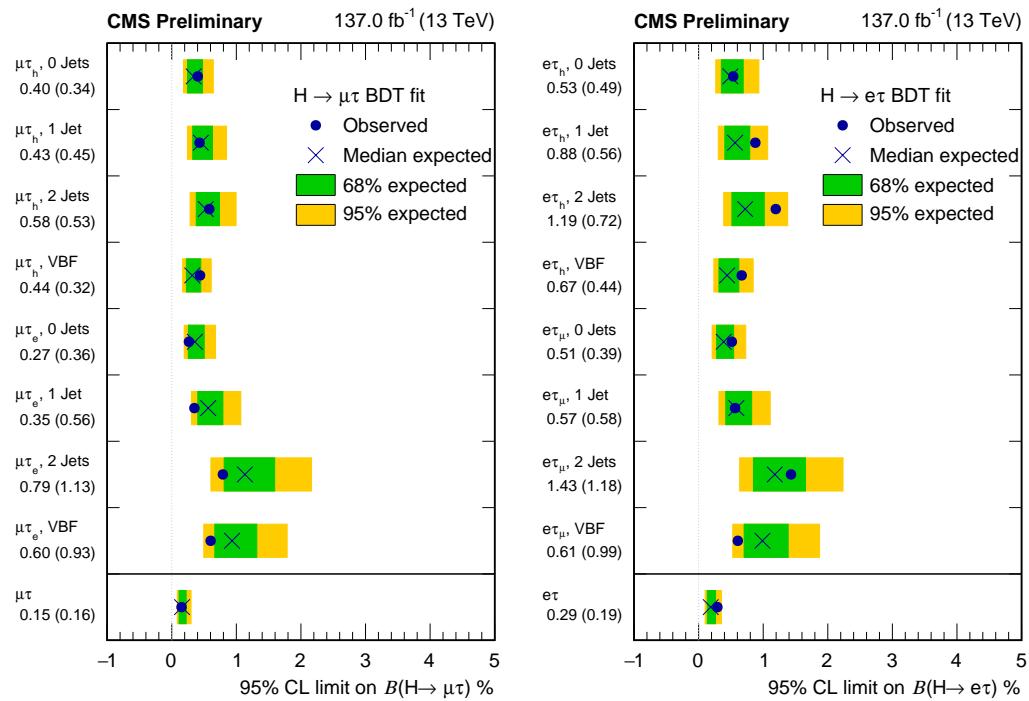


Figure 9.9. Observed (expected) 95% CL upper limits on the  $\mathcal{B}(H \rightarrow \mu\tau)$  (left) and  $\mathcal{B}(H \rightarrow e\tau)$  (right) for each individual category and combined from the BDT fit analysis.

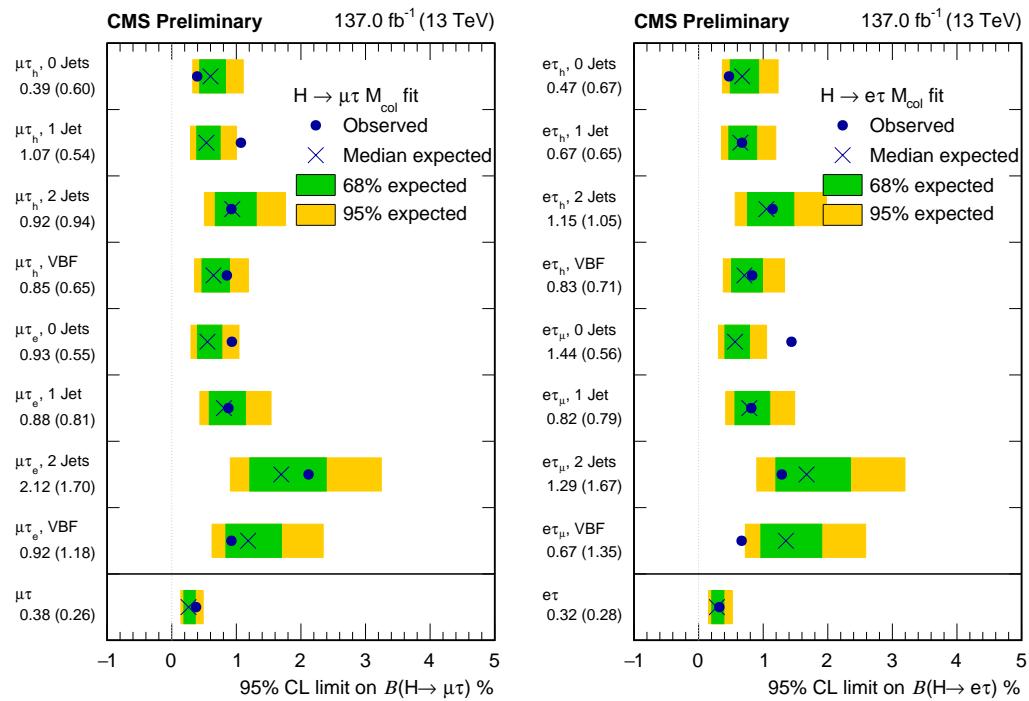


Figure 9.10. Observed (expected) 95% CL upper limits on the  $\mathcal{B}(H \rightarrow \mu\tau)$  (left) and  $\mathcal{B}(H \rightarrow e\tau)$  (right) for each individual category and combined from the  $M_{\text{col}}$  fit analysis.

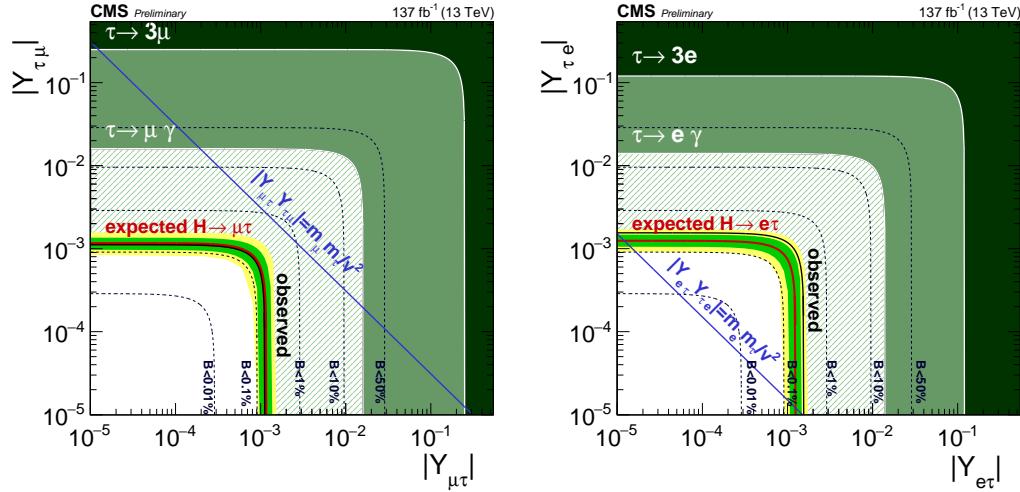


Figure 9.11. Constraints on the LFV Yukawa couplings,  $|Y_{\mu\tau}| - |Y_{\tau\mu}|$  (left), and  $|Y_{e\tau}| - |Y_{\tau e}|$  (right). The expected (red line) and observed (black solid line) limits are derived from the results shown in Figure 9.9. The flavor-diagonal Yukawa couplings are approximated by their SM values. The green hashed region is derived by the CMS direct search presented in this paper. The green (yellow) band indicates the range that is expected to contain 68% (95%) of all observed limit variations from the expected limit. The shaded regions are derived constraints from null searches for  $\tau \rightarrow 3\mu$  or  $\tau \rightarrow 3e$  (dark green) [89] and  $\tau \rightarrow \mu\gamma$  or  $\tau \rightarrow e\gamma$  (lighter green) [6]. The blue diagonal line is the theoretical naturalness limit  $|Y_{ij}Y_{ji}| \leq m_i m_j / v^2$ .

TABLE 9.1

OBSERVED AND EXPECTED UPPER LIMITS AT 95% CL AND BEST  
FIT BRANCHING FRACTIONS FOR EACH INDIVIDUAL JET  
CATEGORY, AND COMBINED, IN THE  $H \rightarrow \mu\tau$  PROCESS.

	Expected limits (%)				
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	< 0.36	< 0.56	< 1.13	< 0.93	< 0.28
$\mu\tau_h$	< 0.34	< 0.45	< 0.53	< 0.32	< 0.19
$\mu\tau$			< 0.16		
	Observed limits (%)				
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	< 0.27	< 0.35	< 0.79	< 0.60	< 0.17
$\mu\tau_h$	< 0.40	< 0.43	< 0.58	< 0.44	< 0.27
$\mu\tau$			< 0.15		
	Best fit branching fractions (%)				
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	-0.16 ± 0.18	-0.41 ± 0.28	-0.61 ± 0.57	-0.58 ± 0.44	-0.22 ± 0.14
$\mu\tau_h$	0.09 ± 0.17	-0.03 ± 0.23	0.08 ± 0.27	0.15 ± 0.16	0.10 ± 0.10
$\mu\tau$			-0.01 ± 0.08		

TABLE 9.2

OBSERVED AND EXPECTED UPPER LIMITS AT 95% CL AND BEST  
FIT BRANCHING FRACTIONS FOR EACH INDIVIDUAL JET  
CATEGORY, AND COMBINED, IN THE  $H \rightarrow e\tau$  PROCESS.

	Expected limits (%)				
	0-jet	1-jet	2-jets	VBF	Combined
$e\tau_\mu$	< 0.39	< 0.58	< 1.18	< 0.99	< 0.29
$e\tau_h$	< 0.49	< 0.56	< 0.72	< 0.44	< 0.25
$e\tau$			< 0.19		
	Observed limits (%)				
	0-jet	1-jet	2-jets	VBF	Combined
$e\tau_\mu$	< 0.51	< 0.57	< 1.43	< 0.61	< 0.31
$e\tau_h$	< 0.53	< 0.88	< 1.19	< 0.67	< 0.45
$e\tau$			< 0.29		
	Best fit branching fractions (%)				
	0-jet	1-jet	2-jets	VBF	Combined
$e\tau_\mu$	$0.17 \pm 0.20$	$-0.03 \pm 0.29$	$0.33 \pm 0.60$	$-0.73 \pm 0.58$	$0.02 \pm 0.15$
$e\tau_h$	$0.06 \pm 0.25$	$0.38 \pm 0.29$	$0.57 \pm 0.37$	$0.28 \pm 0.12$	$0.23 \pm 0.13$
$e\tau$			$0.13 \pm 0.10$		

TABLE 9.3

OBSERVED AND EXPECTED UPPER LIMITS AT 95% CL AND BEST  
FIT BRANCHING FRACTIONS FOR EACH INDIVIDUAL JET  
CATEGORY, AND COMBINED, IN THE  $H \rightarrow \mu\tau$  PROCESS FROM  
 $M_{\text{col}}$  FIT ANALYSIS.

Expected limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	< 0.55	< 0.81	< 1.70	< 1.18	< 0.42
$\mu\tau_h$	< 0.60	< 0.54	< 0.94	< 0.65	< 0.34
$\mu\tau$			< 0.26		
Observed limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	< 0.93	< 0.88	< 2.12	< 0.92	< 0.60
$\mu\tau_h$	< 0.39	< 1.07	< 0.92	< 0.85	< 0.46
$\mu\tau$			< 0.38		
Best fit branching fractions (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	$0.45 \pm 0.29$	$0.08 \pm 0.42$	$0.56 \pm 0.85$	$-0.35 \pm 0.55$	$0.23 \pm 0.21$
$\mu\tau_h$	$-0.42 \pm 0.32$	$0.62 \pm 0.27$	$-0.02 \pm 0.48$	$0.29 \pm 0.33$	$0.16 \pm 0.17$
$\mu\tau$			$0.15 \pm 0.13$		

TABLE 9.4

OBSERVED AND EXPECTED UPPER LIMITS AT 95% CL AND BEST  
FIT BRANCHING FRACTIONS FOR EACH INDIVIDUAL JET  
CATEGORY, AND COMBINED, IN THE  $H \rightarrow e\tau$  PROCESS FROM  $M_{\text{col}}$   
FIT ANALYSIS.

Expected limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$e\tau_\mu$	< 0.56	< 0.79	< 1.67	< 1.35	< 0.44
$e\tau_h$	< 0.67	< 0.65	< 1.05	< 0.71	< 0.38
$e\tau$			< 0.28		
Observed limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$e\tau_\mu$	< 1.44	< 0.82	< 1.29	< 0.67	< 0.65
$e\tau_h$	< 0.47	< 0.67	< 1.15	< 0.83	< 0.34
$e\tau$			< 0.32		
Best fit branching fractions (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$e\tau_\mu$	$0.96 \pm 0.29$	$0.04 \pm 0.40$	$-0.66 \pm 0.84$	$-1.64 \pm 0.68$	$0.26 \pm 0.23$
$e\tau_h$	$-0.45 \pm 0.39$	$0.04 \pm 0.34$	$0.14 \pm 0.53$	$0.17 \pm 0.37$	$-0.08 \pm 0.20$
$e\tau$			$0.05 \pm 0.14$		

TABLE 9.5

SUMMARY OF OBSERVED AND EXPECTED UPPER LIMITS AT 95% CL, BEST FIT BRANCHING FRACTIONS AND CORRESPONDING CONSTRAINTS ON YUKAWA COUPLINGS FOR  $H \rightarrow \mu\tau$  AND  $H \rightarrow e\tau$  PROCESSES.

	Observed (Expected) upper limits (%)	Best fit branching fractions (%)
$H \rightarrow \mu\tau$	$\leq 0.15$ (0.16)	$-0.01 \pm 0.08$
$H \rightarrow e\tau$	$\leq 0.29$ (0.19)	$0.13 \pm 0.10$
Yukawa coupling constraints		
$H \rightarrow \mu\tau$	$\leq 1.12$ (1.15) $\times 10^{-3}$	
$H \rightarrow e\tau$	$\leq 1.55$ (1.24) $\times 10^{-3}$	

## CHAPTER 10

### CONCLUSION

In this dissertation, the search for SM Higgs boson decaying into LFV decays into a muon and tau or an electron and tau has been presented. The search is performed with the full Run 2 data collected at the CMS experiment in 2016, 2017, and 2018 at a center-of-mass energy of 13 TeV corresponding to an integrated luminosity of  $137\text{fb}^{-1}$ . The search found no evidence of the LFV decays of the Higgs boson, and corresponding exclusion upper limits have been placed on the branching fraction of Higgs boson into the  $\mu\tau$  and  $e\tau$  final states.

The analysis has been completed and is currently going through the CMS collaboration's approval process and is expected to be published in the Journal for High Energy Physics. The observed (median expected) upper limits on  $\mathcal{B}(H \rightarrow \mu\tau)$  is 0.15 (0.16) % at 95% CL and on  $\mathcal{B}(H \rightarrow e\tau)$  is 0.29 (0.19) % at 95% CL. The limits on the branching fraction have been correspondingly translated into limits on the off-diagonal Yukawa coupling and are set to be  $\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 1.12 \times 10^{-3}$  and  $\sqrt{|Y_{e\tau}|^2 + |Y_{\tau e}|^2} < 1.55 \times 10^{-3}$ . These are the most stringent limits set in the LFV Higgs decays to date and constitute a significant improvement from the previous results.

The results are a factor of two improvement over the results published with just the 2016 dataset. Several improvements are performed compared to the previous search, and the various details have been explained in this dissertation to a good extent. In summary, using the DeepTau identification for the hadronically decaying tau leptons has significantly improved the sensitivity of the search in the hadronic

channels. In the leptonic channels, the QCD background estimation technique has been updated.  $Z \rightarrow \tau\tau$  background is estimated using the embedding technique for all channels and helps provide a good event description and better control over the systematic uncertainties. The misidentified lepton background in the hadronic channels has been modified. New shape systematics that weren't considered in the 2016 analysis are introduced to account for any residual discrepancies between the data and estimated backgrounds. All these improvements have compounded to set the most stringent limits on these branching fractions to date.

The LHC is expected to start taking data for Run 3 starting from the end of 2021, and currently, the expectation is that the center-of-mass energy will remain the same as for Run 2, i.e., 13 TeV. We anticipate collecting approximately  $150\text{fb}^{-1}$  in Run 3, and this will double the available dataset for Physics analysis. For the next iteration of this search, a new signal mass variable is proposed and is presented in the appendix (“Classic” SVFit mass). From an initial study, this variable has been shown to improve the signal’s mass resolution by approximately 20%, which is a significant improvement. Any future iteration of this search is highly recommended to use this mass variable for the signal mass hypothesis to improve the analysis’s sensitivity.

## APPENDIX A

### SVFIT MASS

A study has been performed to improve the next iteration of the search for LFV decays of the Higgs boson. In this study, we have observed that using the “Classic” SVFit algorithm can help improve the signal resolution compared to the collinear mass resolution. The “Classic” SVFit algorithm has been developed to reconstruct the mass of a Higgs boson decaying into two tau leptons  $m_{\tau\tau}$  and is a more improved version of the SVFit algorithm. The SVfit algorithm [23] has been used to reconstruct the Higgs boson mass in the SM  $H \rightarrow \tau\tau$  analysis and searches for further Higgs bosons predicted by models beyond the SM performed by the CMS collaboration during LHC Run 1. Compared to alternative mass variables, the SVfit algorithm’s usage has improved the sensitivity of the SM  $H \rightarrow \tau\tau$  analysis for measuring the signal rate by  $\approx 40\%$  [13]. The improvement in sensitivity corresponds to a gain by about a factor of two in the integrated luminosity of the analyzed dataset.

The “Classic” SVFit (Sensitivity Volume Fit) algorithm uses a likelihood function of arbitrary normalization. The algorithm allows for the reconstruction of not only the mass  $m_{\tau\tau}$  of the  $\tau$  lepton pair but any kinematic function of the two  $\tau$  leptons, including the  $p_T$ ,  $\eta$ , and  $\phi$  and transverse mass of the  $\tau$  lepton pair. A further improvement concerns the algorithm’s extension to account for the experimental resolution on the reconstruction of hadrons produced in the  $\tau$  decays. I had to modify the “Classic” SVFit algorithm developed for the reconstruction of  $m_{\tau\tau}$  to the case where we reconstruct the mass of the Higgs boson decaying into LFV decay modes  $m_{\ell\tau}$ . A brief description of the algorithm is given in the following paragraphs.

### A.1 “Classic” SVfit algorithm

As only one of the lepton from an LFV Higgs decay is a tau, I had to modify the matrix element of the probability density function defined to reconstruct the Higgs mass  $m_{\ell\tau}$ . The phase space of the tau decay products along with the angles  $\theta_{inv}$  and  $\phi_{inv}$  are illustrated in Figure A.1. The  $\mathbf{p}_{inv}$  vector is located on the surface of a cone, the axis of which is given by the  $\mathbf{p}_{vis}$  vector. The variable  $\phi_{inv}$  represents the angle of rotation, in a counter-clockwise direction, around the cone’s axis. The value  $\phi_{inv} = 0$  is chosen to correspond to the case that the  $\mathbf{p}_{inv}$  vector is within the plane spanned by the  $\mathbf{p}_{vis}$  vector and the beam direction.

$$\begin{aligned}\mathcal{L}(\mathbf{p}^{vis}, p_x^{rec}, p_y^{rec} | m_H) &= \frac{32\pi^4}{s} \int dm_H d\Phi_n |\text{BW}_\tau|^2 |\mathcal{M}_{\tau \rightarrow \dots}(\tilde{\mathbf{p}})|^2 \\ W(\mathbf{p}^{vis} | \hat{\mathbf{p}}^{vis}) W_{rec}(p_x^{rec}, p_y^{rec} | \hat{p}_x^{rec}, \hat{p}_y^{rec}) \mathcal{F}(\mathbf{p})\end{aligned}$$

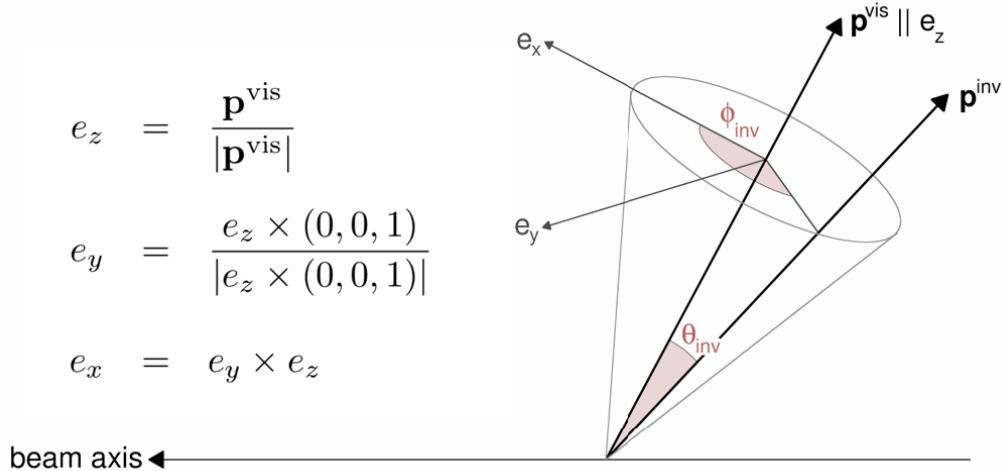


Figure A.1: Illustration of the variables  $\theta_{inv}$  and  $\phi_{inv}$  that specify the orientation of the  $\mathbf{p}_{inv}$  vector relative to the momentum vector  $\mathbf{p}_{vis}$  of the visible  $\tau$  decay products.

The function  $\mathcal{F}(\mathbf{p})$  in the integrand may be an arbitrary function of the momenta  $\mathbf{p}$  of the prompt and  $\tau$  leptons. The integral is evaluated numerically, using a custom implementation of the Markov chain Monte Carlo integration method with the Metropolis-Hastings algorithm [60]. The actual value  $\mathcal{L}(y)$  of the integral is irrelevant. The reconstruction of the mass  $m_{\ell\tau}$  of the prompt and  $\tau$  lepton pair is based on choosing

$$\mathcal{F}(\mathbf{p}) \equiv \left( \hat{E}_\ell + \hat{E}_\tau \right)^2 - \left( (\hat{p}_x^\ell + \hat{p}_x^\tau)^2 + (\hat{p}_y^\ell + \hat{p}_y^\tau)^2 + (\hat{p}_z^\ell + \hat{p}_z^\tau)^2 \right),$$

recording the values of  $\mathcal{F}(\mathbf{p})$  for each evaluation of the integrand by the Markov chain and taking the median of the series of  $\mathcal{F}(\mathbf{p})$  values as the best estimate  $m_{\ell\tau}$  for the mass of the prompt and  $\tau$  lepton pair in a given event. The total number of evaluations of the integrand referred to as Markov chain “states”, amounts to 100000 per event. The first 10000 evaluations of the integrand are used as a “burn-in” period and are excluded from the median’s computation.

Figure A.2 shows the distribution of the collinear mass and the SVFit mass distributions. The distributions are fit with a double Gaussian, and the goodness of fit as tested from  $\chi^2/\text{ndof}$  is close to one, suggesting that it is a good fit. The full width at half maximum is used for inferring the signal resolution, and it has a value of 38.88 GeV for collinear mass, while for SVFit mass, the value is 30.90 GeV. This corresponds to a 20% improvement in signal resolution. The collinear mass and SVFit mass are peaking close to the Higgs mass of 125 GeV. The collinear mass has a higher mean of  $\sim 129$  GeV due to its larger tail, while SVFit mass has a mean of  $\sim 117$  GeV to its smaller tail distribution.

This study has shown that mass resolution is significantly improved by using the SVFit mass instead of the collinear mass, and it can give much more sensitive results for the  $H \rightarrow \mu\tau$  and  $H \rightarrow e\tau$  searches. A future search using the full Run 3 data can benefit from this improved sensitivity, and the SVFit can easily be extended to be used for Heavy Higgs boson searches. The only disadvantage that was noted during

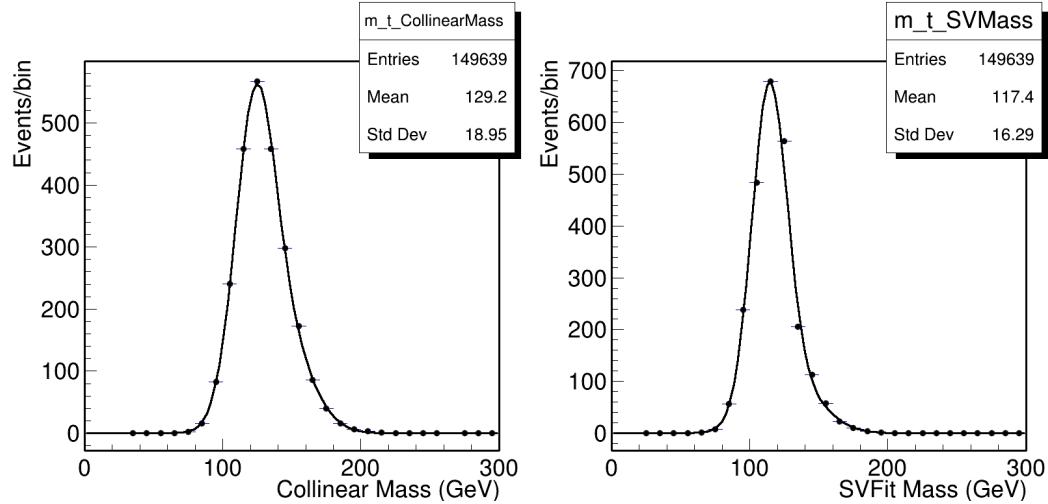


Figure A.2: Collinear mass vs SVFit mass.

the study was from the computing point of view. It takes an order of magnitude longer time to compute the SVFit mass than the collinear mass, which has to be investigated and improved in any future analysis.

## BIBLIOGRAPHY

1. M. Tanabashi et al. Review of Particle Physics. *Phys. Rev. D*, 98(3):030001, 2018. doi: 10.1103/PhysRevD.98.030001.
2. M. Raidal et al. Flavor physics of leptons and dipole moments. *Eur. Phys. J. C*, 57:13–182, 2008. doi: 10.1140/epjc/s10052-008-0715-2.
3. LHC Machine. *JINST*, 3:S08001, 2008. doi: 10.1088/1748-0221/3/08/S08001.
4. Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.
5. Serguei Chatrchyan et al. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett. B*, 716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.
6. Roni Harnik, Joachim Kopp, and Jure Zupan. Flavor Violating Higgs Decays. *JHEP*, 03:026, 2013. doi: 10.1007/JHEP03(2013)026.
7. J. Goldstone. Field Theories with Superconductor Solutions. *Nuovo Cim.*, 19:154–164, 1961. doi: 10.1007/BF02812722.
8. Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken Symmetries. *Phys. Rev.*, 127:965–970, 1962. doi: 10.1103/PhysRev.127.965.
9. Serguei Chatrchyan et al. Observation of a New Boson with Mass Near 125 GeV in  $pp$  Collisions at  $\sqrt{s} = 7$  and 8 TeV. *JHEP*, 06:081, 2013. doi: 10.1007/JHEP06(2013)081.
10. P.A.R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594:A13, 2016. doi: 10.1051/0004-6361/201525830.
11. Gianluca Blankenburg, John Ellis, and Gino Isidori. Flavour-Changing Decays of a 125 GeV Higgs-like Particle. *Phys. Lett. B*, 712:386–390, 2012. doi: 10.1016/j.physletb.2012.05.007.
12. Vardan Khachatryan et al. Search for Lepton-Flavour-Violating Decays of the Higgs Boson. *Phys. Lett. B*, 749:337–362, 2015. doi: 10.1016/j.physletb.2015.07.053.

13. J.Lorenzo Diaz-Cruz and J.J. Toscano. Lepton flavor violating decays of Higgs bosons beyond the standard model. *Phys. Rev. D*, 62:116005, 2000. doi: 10.1103/PhysRevD.62.116005.
14. T.P. Cheng and Marc Sher. Mass Matrix Ansatz and Flavor Nonconservation in Models with Multiple Higgs Doublets. *Phys. Rev. D*, 35:3484, 1987. doi: 10.1103/PhysRevD.35.3484.
15. Toru Goto, Ryuichiro Kitano, and Shingo Mori. Lepton flavor violating  $Z$ -boson couplings from nonstandard Higgs interactions. *Phys. Rev. D*, 92:075021, 2015. doi: 10.1103/PhysRevD.92.075021.
16. L. Willmann et al. New bounds from searching for muonium to anti-muonium conversion. *Phys. Rev. Lett.*, 82:49–52, 1999. doi: 10.1103/PhysRevLett.82.49.
17. D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2/2017, 10 2016. doi: 10.23731/CYRM-2017-002.
18. Albert M Sirunyan et al. Search for lepton flavour violating decays of a neutral heavy Higgs boson to  $\mu\tau$  and  $e\tau$  in proton-proton collisions at  $\sqrt{s} = 13$  TeV. *JHEP*, 03:103, 2020. doi: 10.1007/JHEP03(2020)103.
19. S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.
20. The Large Electron-Positron Collider. Jul 2012. URL <https://cds.cern.ch/record/1997351>.
21. Linear accelerator 2. Sep 2012. URL <https://cds.cern.ch/record/1997427>.
22. The Proton Synchrotron Booster. Jul 2012. URL <https://cds.cern.ch/record/1997372>.
23. The Proton Synchrotron. Jan 2012. URL <https://cds.cern.ch/record/1997189>.
24. The Super Proton Synchrotron. Jan 2012. URL <https://cds.cern.ch/record/1997188>.
25. G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. doi: 10.1088/1748-0221/3/08/S08003.
26. K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008. doi: 10.1088/1748-0221/3/08/S08002.
27. Jr. Alves, A.Augusto et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. doi: 10.1088/1748-0221/3/08/S08005.
28. G. Anelli et al. The TOTEM experiment at the CERN Large Hadron Collider. *JINST*, 3:S08007, 2008. doi: 10.1088/1748-0221/3/08/S08007.

29. O. Adriani et al. The LHCf detector at the CERN Large Hadron Collider. *JINST*, 3:S08006, 2008. doi: 10.1088/1748-0221/3/08/S08006.
30. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
31. CMS ready for winding up. (BUL-NA-2003-150. 49/2003. 49/2003):1, Dec 2003. URL <https://cds.cern.ch/record/686754>. Cover article.
32. Vardan Khachatryan et al. CMS Tracking Performance Results from Early LHC Operation. *Eur. Phys. J. C*, 70:1165–1192, 2010. doi: 10.1140/epjc/s10052-010-1491-3.
33. Electron reconstruction and identification at  $\sqrt{s} = 7$  TeV. 2010.
34. Vardan Khachatryan et al. Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at  $\sqrt{s} = 8$  TeV. *JINST*, 10(06):P06005, 2015. doi: 10.1088/1748-0221/10/06/P06005.
35. G. L. Bayatian et al. CMS Physics: Technical Design Report Volume 1: Detector Performance and Software. 2006.
36. CMS luminosity measurement for the 2018 data-taking period at  $\sqrt{s} = 13$  TeV. 5 2019.
37. CMS luminosity measurement for the 2017 data-taking period at  $\sqrt{s} = 13$  TeV. 6 2018.
38. CMS Luminosity Measurements for the 2016 Data Taking Period. 3 2017.
39. CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update. 10 2013.
40. Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015. doi: 10.1016/j.cpc.2015.01.024.
41. Johan Alwall et al. A Standard format for Les Houches event files. *Comput. Phys. Commun.*, 176:300–304, 2007. doi: 10.1016/j.cpc.2006.11.010.
42. John C. Collins. Sudakov form-factors. *Adv. Ser. Direct. High Energy Phys.*, 5: 573–614, 1989. doi: 10.1142/9789814503266\_0006.
43. Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys. B*, 126:298–318, 1977. doi: 10.1016/0550-3213(77)90384-4.
44. Stefan Hoeche, Frank Krauss, Nils Lavesson, Leif Lonnblad, Michelangelo Mangano, Andreas Schalicke, and Steffen Schumann. Matching parton showers and matrix elements. In *HERA and the LHC: A Workshop on the Implications of*

- HERA for LHC Physics: CERN - DESY Workshop 2004/2005 (Midterm Meeting, CERN, 11-13 October 2004; Final Meeting, DESY, 17-21 January 2005),* pages 288–289, 2005. doi: 10.5170/CERN-2005-014.288.
45. C. Peterson, D. Schlatter, I. Schmitt, and Peter M. Zerwas. Scaling Violations in Inclusive e+ e- Annihilation Spectra. *Phys. Rev. D*, 27:105, 1983. doi: 10.1103/PhysRevD.27.105.
  46. J.D. Bjorken. Properties of Hadron Distributions in Reactions Containing Very Heavy Quarks. *Phys. Rev. D*, 17:171–173, 1978. doi: 10.1103/PhysRevD.17.171.
  47. Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5 : Going Beyond. *JHEP*, 06:128, 2011. doi: 10.1007/JHEP06(2011)128.
  48. Neil D. Christensen and Claude Duhr. FeynRules - Feynman rules made easy. *Comput. Phys. Commun.*, 180:1614–1641, 2009. doi: 10.1016/j.cpc.2009.02.018.
  49. Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010. doi: 10.1007/JHEP06(2010)043.
  50. S. Frixione, Z. Kunszt, and A. Signer. Three jet cross-sections to next-to-leading order. *Nucl. Phys. B*, 467:399–442, 1996. doi: 10.1016/0550-3213(96)00110-1.
  51. S. Frixione. A General approach to jet cross-sections in QCD. *Nucl. Phys. B*, 507:295–314, 1997. doi: 10.1016/S0550-3213(97)00574-9.
  52. Rikkert Frederix, Stefano Frixione, Valentin Hirschi, Fabio Maltoni, Roberto Pittau, and Paolo Torrielli. Four-lepton production at hadron colliders: aMC@NLO predictions with theoretical uncertainties. *JHEP*, 02:099, 2012. doi: 10.1007/JHEP02(2012)099.
  53. J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014. doi: 10.1007/JHEP07(2014)079.
  54. Valentin Hirschi, Rikkert Frederix, Stefano Frixione, Maria Vittoria Garzelli, Fabio Maltoni, and Roberto Pittau. Automation of one-loop QCD corrections. *JHEP*, 05:044, 2011. doi: 10.1007/JHEP05(2011)044.
  55. Michelangelo L. Mangano, Mauro Moretti, and Roberto Pittau. Multijet matrix elements and shower evolution in hadronic collisions:  $Wb\bar{b} + n$  jets as a case study. *Nucl. Phys. B*, 632:343–362, 2002. doi: 10.1016/S0550-3213(02)00249-3.
  56. Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003. doi: 10.1088/1126-6708/2003/07/001.

57. Rikkert Frederix and Stefano Frixione. Merging meets matching in MC@NLO. *JHEP*, 12:061, 2012. doi: 10.1007/JHEP12(2012)061.
58. S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506: 250–303, 2003. doi: 10.1016/S0168-9002(03)01368-8.
59. A.M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017. doi: 10.1088/1748-0221/12/10/P10003.
60. Serguei Chatrchyan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9(10):P10009, 2014. doi: 10.1088/1748-0221/9/10/P10009.
61. R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth. A*, 262:444–450, 1987. doi: 10.1016/0168-9002(87)90887-4.
62. R. Fruhwirth, W. Waltenberger, and P. Vanlaer. Adaptive vertex fitting. *J. Phys. G*, 34:N343, 2007. doi: 10.1088/0954-3899/34/12/N01.
63. A.M. Sirunyan et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV. *JINST*, 13(06):P06015, 2018. doi: 10.1088/1748-0221/13/06/P06015.
64. CMS Muon Physics Object Group. Muon efficiency scale factors. Website. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsMUO>.
65. A.M. Sirunyan et al. Performance of reconstruction and identification of  $\tau$  leptons decaying to hadrons and  $\nu_\tau$  in pp collisions at  $\sqrt{s} = 13$  TeV. *JINST*, 13(10):P10005, 2018. doi: 10.1088/1748-0221/13/10/P10005.
66. Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- $k_t$  jet clustering algorithm. *JHEP*, 04:063, 2008. doi: 10.1088/1126-6708/2008/04/063.
67. Serguei Chatrchyan et al. Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS. *JINST*, 6:P11002, 2011. doi: 10.1088/1748-0221/6/11/P11002.
68. Vardan Khachatryan et al. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST*, 12(02):P02014, 2017. doi: 10.1088/1748-0221/12/02/P02014.
69. Albert M Sirunyan et al. Performance of missing transverse momentum reconstruction in proton-proton collisions at  $\sqrt{s} = 13$  TeV using the CMS detector. *JINST*, 14(07):P07004, 2019. doi: 10.1088/1748-0221/14/07/P07004.
70. Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Catchment Area of Jets. *JHEP*, 04:005, 2008. doi: 10.1088/1126-6708/2008/04/005.

71. Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys. Lett. B*, 659:119–126, 2008. doi: 10.1016/j.physletb.2007.09.077.
72. Albert M Sirunyan et al. Observation of the Higgs boson decay to a pair of  $\tau$  leptons with the CMS detector. *Phys. Lett. B*, 779:283–316, 2018. doi: 10.1016/j.physletb.2018.02.004.
73. Andreas Hocker et al. TMVA - Toolkit for Multivariate Data Analysis. 3 2007.
74. R.Keith Ellis, I. Hinchliffe, M. Soldate, and J.J. van der Bij. Higgs Decay to tau+ tau-: A Possible Signature of Intermediate Mass Higgs Bosons at the SSC. *Nucl. Phys. B*, 297:221–243, 1988. doi: 10.1016/0550-3213(88)90019-3.
75. Aaron Roodman. Blind analysis in particle physics. *eConf*, C030908:TUIT001, 2003.
76. Albert M Sirunyan et al. An embedding technique to determine  $\tau\tau$  backgrounds in proton-proton collision data. *JINST*, 14(06):P06032, 2019. doi: 10.1088/1748-0221/14/06/P06032.
77. Procedure for the LHC Higgs boson search combination in summer 2011. 8 2011.
78. Alexander L. Read. Presentation of search results: The CL(s) technique. *J. Phys. G*, 28:2693–2704, 2002. doi: 10.1088/0954-3899/28/10/313.
79. Alexander L. Read. Modified frequentist analysis of search results (The CL(s) method). In *Workshop on Confidence Limits*, pages 81–101, 8 2000.
80. Thomas Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A*, 434:435–443, 1999. doi: 10.1016/S0168-9002(99)00498-2.
81. J.S. Conway. Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra. In *PHYSTAT 2011*, pages 115–120, 2011. doi: 10.5170/CERN-2011-006. 115.
82. Alexander L. Read. Linear interpolation of histograms. *Nucl. Instrum. Meth. A*, 425:357–360, 1999. doi: 10.1016/S0168-9002(98)01347-3.
83. Serguei Chatrchyan et al. Performance of CMS Muon Reconstruction in  $pp$  Collision Events at  $\sqrt{s} = 7$  TeV. *JINST*, 7:P10002, 2012. doi: 10.1088/1748-0221/7/10/P10002.
84. Vardan Khachatryan et al. Reconstruction and identification of  $\tau$  lepton decays to hadrons and  $\nu_\tau$  at CMS. *JINST*, 11(01):P01019, 2016. doi: 10.1088/1748-0221/11/01/P01019.
85. Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2. 7 2016.

86. Roger J. Barlow and Christine Beeston. Fitting using finite Monte Carlo samples. *Comput. Phys. Commun.*, 77:219–228, 1993. doi: 10.1016/0010-4655(93)90005-W.
87. Albert M Sirunyan et al. Search for lepton flavour violating decays of the Higgs boson to  $\mu\tau$  and  $e\tau$  in proton-proton collisions at  $\sqrt{s} = 13$  TeV. *JHEP*, 06:001, 2018. doi: 10.1007/JHEP06(2018)001.
88. A. Denner, S. Heinemeyer, I. Puljak, D. Rebuzzi, and M. Spira. Standard Model Higgs-Boson Branching Ratios with Uncertainties. *Eur. Phys. J. C*, 71:1753, 2011. doi: 10.1140/epjc/s10052-011-1753-8.
89. K. Hayasaka et al. Search for Lepton Flavor Violating Tau Decays into Three Leptons with 719 Million Produced Tau+Tau- Pairs. *Phys. Lett. B*, 687:139–143, 2010. doi: 10.1016/j.physletb.2010.03.037.