

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 December 30, 2020 for

Prasanna Kumar Siddireddy

This L^AT_EX 2 _{\mathcal{E}} 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 _{\mathcal{E}} 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 _{\mathcal{E}} 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

December 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 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 corresponding to an integrated luminosity of 137 fb^{-1} was used to perform the search. The observed (expected) upper limits on the branching fraction of Higgs boson decaying into a muon and a tau lepton, $\mathcal{B}(H \rightarrow \mu\tau)$, were set at 0.15 (0.15) % at 95% CL and the branching fraction of Higgs boson decaying into an electron and a tau lepton, $\mathcal{B}(H \rightarrow e\tau)$, were set at 0.22 (0.16) % at 95% CL. These are the most stringent limits set on these processes to date.

CONTENTS

Figures	iv
Tables	x
Chapter 1: Introduction	1
Chapter 2: Theoretical Motivation	6
2.1 Theoretical building blocks	8
2.2 The Standard Model	11
2.2.1 Electroweak symmetry breaking and the Higgs boson	14
2.3 Lepton flavor violating decays of the Higgs boson	15
2.4 Constraints from low-energy measurements	22
2.5 Lepton flavor violating decays of the Higgs boson at the LHC	26
Chapter 3: LHC and the CMS detector	29
3.1 Introduction	29
3.2 The Large Hadron Collider	29
3.3 The Compact Muon Solenoid experiment	31
3.3.1 The coordinate system of CMS	34
3.3.2 Kinematic quantities	35
3.3.3 Detector requirements	36
3.3.4 Magnet	37
3.3.5 Inner tracking system	37
3.3.6 Electromagnetic calorimeter	38
3.3.7 Hadron calorimeter	41
3.3.8 Muon system	42
3.3.9 Trigger and data acquisition system	43
3.3.10 Luminosity measurement	46
3.3.11 CMS Phase-2 upgrade	50
Chapter 4: Monte Carlo Event Generation	58
4.1 Introduction	58
4.2 Monte Carlo simulation	59
4.3 Monte Carlo generators	62
4.4 Detector simulation	64

4.5 Monte Carlo samples	65
Chapter 5: Event Reconstruction	67
5.1 Track and primary vertex reconstruction	68
5.2 Muon reconstruction	71
5.3 Electron reconstruction	72
5.4 Tau reconstruction	76
5.4.1 Hadrons Plus Strips	77
5.4.2 DeepTau	78
5.5 Jet reconstruction	82
5.6 Missing transverse energy	84
5.7 Relative isolation	85
Chapter 6: Event Selection	86
6.1 Introduction	86
6.2 Boosted Decision Tree	92
6.3 $H \rightarrow \mu\tau_h$ channel	94
6.4 $H \rightarrow \mu\tau_e$ channel	95
6.5 $H \rightarrow e\tau_h$ channel	99
6.6 $H \rightarrow e\tau_\mu$ channel	105
Chapter 7: Background Estimation	109
7.1 Introduction	109
7.2 Embedding technique	111
7.3 Misidentified lepton background	113
7.3.1 Fully data-driven approach	113
7.3.2 Semi data-driven approach	120
7.4 Monte Carlo simulations	126
Chapter 8: Statistical Methods and Systematic Uncertainties	128
8.1 Introduction	128
8.2 Statistical methods	128
8.3 Systematic uncertainties	133
Chapter 9: Results	143
Chapter 10: Conclusion	160
Appendix A: SVFit Mass	163
A.1 “Classic” SVfit algorithm	164
Appendix B: Fake Factor method	167
Bibliography	170

FIGURES

2.1	Elementary particles of the SM.	7
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	13
2.3	Diagrams contributing to the flavor violating decay $\tau \rightarrow \mu\gamma$, mediated by a Higgs boson with flavor violating Yukawa couplings.	23
2.4	Expected and observed 95% CL upper limits for each category and their combination of the search performed by the CMS experiment [25].	28
2.5	Constraints on the LFV Yukawa couplings, $ Y_{\mu\tau} - Y_{\tau\mu} $ (left), and $ Y_{e\tau} - Y_{\tau e} $ (right) of the search performed by the CMS experiment.	28
3.1	Cumulative luminosity as a function of time delivered to CMS during stable beams for pp collisions. The luminosity is shown for 2015 (purple), 2016 (orange), 2017 (light blue), and 2018 (dark blue) [37]. .	32
3.2	Schematic view of the CMS detector.	33
3.3	Coordinate system convention of CMS (left) and the relation between pseudorapidity η and polar angle θ (right)	35
3.4	Schematic view of the ECAL.	39
3.5	Schematic view of the muon detectors.	44
3.6	The ECAL electronics online monitoring system in operation.	46
3.7	Schedule of LHC and HL-LHC operation.	50
3.8	APD dark current shown for APDs at $ \eta = 1.45$ operated at 18°C (red curve), at 9°C (blue curve), and in a mixed scenario including a further step to 6°C after the LHC Long Shutdown 4.	53
3.9	Performance estimates from High-Level Synthesis.	57
4.1	MC simulation of an event in pp collisions.	59
5.1	The tracking efficiency for “all-tracks” collection as a function of pseudorapidity and transverse momentum for muons coming from the Z decay.	70
5.2	Efficiency of muon identification as a function of η and p_T , for data (black) and simulation (blue) [75].	73

5.3	Receiver operating curve showing the background versus signal efficiency for the BDT with and without isolation incorporated in training.	77
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.	80
5.5	Comparison of the collinear mass distribution from the previous analysis (left) compared to the current analysis (right). As can be seen from the plots, the Drell-Yan background and misidentified lepton background are significantly reduced in the current analysis. This can be attributed to using the DeepTau identification for hadronic taus.	81
6.1	Feynman diagrams of LFV 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).	87
6.2	Illustration of the differences in $d\phi(\ell, \text{MET})$ and muon transverse momentum spectrums in LFV and SM $H \rightarrow \tau\tau$ processes.	88
6.3	Illustration of the differences in $d\phi(\ell, \text{MET})$ and electron transverse momentum spectrums in LFV and SM $H \rightarrow \tau\tau$ processes.	89
6.4	Estimation of the neutrino momentum $E_{T,\text{proj}}$ by using the component of the missing transverse energy E_T which is collinear to the visible decay products of tau in the transverse plane.	91
6.5	Illustration of decision tree.	92
6.6	Distribution of the input variables to the BDT for the $\mu\tau_h$ process.	96
6.7	Overtraining check as performed in TMVA for the trained BDT in $\mu\tau_h$ channel for 2016 (a), 2017 (b), and 2018 (c).	98
6.8	Distribution of the input variables to the BDT for the $\mu\tau_e$ process.	100
6.9	Overtraining check as performed in TMVA for the trained BDT in $\mu\tau_e$ channel for 2016 (a), 2017 (b), and 2018 (c).	101
6.10	Single electron trigger efficiency as a function of p_T in different η regions.	102
6.11	Distribution of the input variables to the BDT for the $e\tau_h$ process.	103
6.12	Overtraining check as performed in TMVA for the trained BDT in $e\tau_h$ channel for 2016 (a), 2017 (b), and 2018 (c).	105
6.13	Distribution of the input variables to the BDT for the $e\tau_\mu$ process.	107
6.14	Overtraining check as performed in TMVA for the trained BDT in $e\tau_\mu$ channel for 2016 (a), 2017 (b), and 2018 (c).	108
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) $W\gamma^{(*)}$	110

7.2	Schematic of embedding technique	112
7.3	Distributions of M_{col} discriminator in the $Z \rightarrow \tau\tau$ control regions for the (a) $\mu\tau_h$, (b) $\mu\tau_e$, (c) $e\tau_h$, and (d) $e\tau_\mu$ channels.	114
7.4	Distributions of BDT discriminator in the $Z \rightarrow \tau\tau$ control regions for the (a) $\mu\tau_h$, (b) $\mu\tau_e$, (c) $e\tau_h$, and (d) $e\tau_\mu$ channels.	115
7.5	Signal region (green) contrasted with the background enriched regions used for estimating the misidentified background	116
7.6	Fit performed to τ_h misidentification rates for $\mu\tau_h$ (a) and $e\tau_h$ (b) channel as a function of τ_h p_T for the different years. The misidentification rates used are further parametrized based on τ_h decay mode along with the pseudorapidity of τ_h . However, here only the inclusive misidentification rates are shown.	118
7.7	Fit performed to the muon (a) and electron (b) misidentification rates as a function of their p_T for 2016 (left), 2017 (center), and 2018 (right). The hyperbolic tangent function is used for performing the fit.	119
7.8	Distributions of M_{col} discriminator in the same sign (left) and W boson enriched (right) control regions for the $\mu\tau_h$ (top) and $e\tau_h$ (bottom) channels.	121
7.9	Distributions of BDT discriminator in the same sign (left) and W boson enriched (right) control regions for the $\mu\tau_h$ (top) and $e\tau_h$ (bottom) channels.	122
7.10	The 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.	123
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.	124
7.12	Distribution of M_{col} discriminator in the muon anti-isolated control regions for the $\mu\tau_e$ channel.	125
7.13	Distributions of M_{col} (BDT) discriminator in $t\bar{t}$ enriched control region for $\mu\tau_e$ and $e\tau_\mu$ channel.	127
8.1	Test statistic distributions for ensembles of pseudo-data generated for background-only (blue) and signal-plus-background (red) hypotheses [91].	131

8.2	(left) An example of differential distribution of possible limits on μ 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 μ for the background-only hypothesis.	133
8.3	Shape uncertainty estimated for the $H \rightarrow \mu\tau_h$ channel. (a) Distribution of $\Delta\phi(\mu, MET)$ from the W boson enriched control region. (b) The discrepancy seen above is parametrized with a linear function. (c) Distribution of $\Delta\phi(\mu, MET)$ from the signal region.	139
8.4	Shape uncertainty estimated for the $H \rightarrow e\tau_h$ channel. (a) Distribution of $\Delta\phi(e, MET)$ from the W boson enriched control region. (b) The discrepancy seen above is parametrized with a linear function. (c) Distribution of $\Delta\phi(e, MET)$ from the signal region.	140
9.1	BDT discriminator distributions for the observed and estimated background in the $\mu\tau_h$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow \mu\tau) = 20\%$. The $\mu\tau_h$ 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 statistical and systematic uncertainties added in quadrature.	144
9.2	BDT discriminator distributions for the observed and estimated background in the $\mu\tau_e$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow \mu\tau) = 20\%$. The $\mu\tau_e$ 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 statistical and systematic uncertainties added in quadrature.	145
9.3	M_{col} distributions for the observed and estimated background in the $\mu\tau_h$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow \mu\tau) = 10\%$. The $\mu\tau_h$ 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 statistical and systematic uncertainties added in quadrature.	146

9.4	M_{col} distributions for the observed and estimated background in the $\mu\tau_e$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow \mu\tau) = 10\%$. The $\mu\tau_e$ 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 statistical and systematic uncertainties added in quadrature.	147
9.5	BDT discriminator distributions for the observed and estimated background in the $e\tau_h$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow e\tau) = 20\%$. The $e\tau_h$ 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 statistical and systematic uncertainties added in quadrature. . . .	148
9.6	BDT discriminator distributions for the observed and estimated background in the $e\tau_\mu$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow e\tau) = 20\%$. The $e\tau_\mu$ 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 statistical and systematic uncertainties added in quadrature. . . .	149
9.7	M_{col} distributions for the observed and estimated background in the $e\tau_h$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow e\tau) = 10\%$. The $e\tau_h$ 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 statistical and systematic uncertainties added in quadrature.	150
9.8	M_{col} distributions for the observed and estimated background in the $e\tau_\mu$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow e\tau) = 10\%$. The $e\tau_\mu$ 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 statistical and systematic uncertainties added in quadrature.	151
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.	153

9.10 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) [103] and $\tau \rightarrow \mu\gamma$ or $\tau \rightarrow e\gamma$ (lighter green) [20]. The blue diagonal line is the theoretical naturalness limit $ Y_{ij} Y_{ji} \leq m_i m_j / v^2$	154
10.1 Expected limits from the previous analysis projected to the full Run 2 luminosity for (a) $\mu\tau_h$, (b) $e\tau_h$, (c) $\mu\tau_e$, (d) $e\tau_\mu$, (e) $\mu\tau$, and (f) $e\tau$ channels. The red cross represents the median expected limits obtained in the current analysis and are an improvement compared to the projections.	162
A.1 Illustration of the variables θ_{inv} and ϕ_{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.	164
A.2 Collinear mass vs SVFit mass.	166
B.1 Illustration of the fake factor method. For each of the processes (W + jets, QCD, and $t\bar{t}$), a dedicated fake factor is measured in a process enriched determination region. The measured fake factor is corrected and then applied on an event by event basis inside the application region. The application of the fake factor inside the application region also involves the fractions f_i , representing the probabilities of an event in the application region originating from process i	168

TABLES

2.1	Constraints on LFV Yukawa couplings from low-energy measurements [20].	25
3.1	Integrated luminosity considered for physics analysis at the CMS experiment during Run 2	31
6.1	Event selection criteria for the kinematic variables for the $H \rightarrow \mu\tau$ channels	97
6.2	Event selection criteria for the kinematic variables for the $H \rightarrow e\tau$ channels	104
8.1	Systematic uncertainties in the expected event yields for the $H \rightarrow \mu\tau$ channels.	135
8.2	Systematic uncertainties in the expected event yields for the $H \rightarrow e\tau$ channels.	136
9.1	Observed and expected upper limits at 95% CL and best fit branching fractions for each individual category, and combined, in the $H \rightarrow \mu\tau$ process from BDT fit analysis.	155
9.2	Observed and expected upper limits at 95% CL and best fit branching fractions for each individual category, and combined, in the $H \rightarrow e\tau$ process from BDT fit analysis.	156
9.3	Observed and expected upper limits at 95% CL and best fit branching fractions for each individual category, and combined, in the $H \rightarrow \mu\tau$ process from M_{col} fit analysis.	157
9.4	Observed and expected upper limits at 95% CL and best fit branching fractions for each individual category, and combined, in the $H \rightarrow e\tau$ process from M_{col} fit analysis.	158
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.	159

CHAPTER 1

INTRODUCTION

Our universe's nature is explored in particle physics by understanding elementary particles' properties and the fundamental forces. 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 the Standard Model (SM) by introducing a new scalar field associated with a new particle, the Higgs boson. 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 and dark energy, which are inferred from astrophysical/cosmological observations and our universe's accelerated expansion [1]. Also, the CP-violating effects of the SM cannot explain the large difference between matter and anti-matter in the universe.

The flavor is a term used to describe the different kinds of leptons (electron, muon, or tau). A lepton flavor violation (LFV) is a transition that does not conserve the lepton flavor. The neutrino oscillations have been experimentally observed at several experiments (Super-Kamiokande collaboration [2] and Sudbury Neutrino Observatory collaboration [3]), suggesting that LFV occurs in the neutral lepton sector (electron neutrino, muon neutrino, or tau neutrino) and that the neutrinos mix and have a mass. However, no charged LFV has been observed to date. Yukawa interactions between the scalar field and the fermion fields give rise to the mass of the fermions. In the SM, the mass matrix and the Yukawa interaction matrix are

diagonalized simultaneously by choosing a particular eigenstate basis as they have the same coefficient in the Lagrangian. This corresponds to the SM Higgs boson not having LFV decays.

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 Cabibbo-Kobayashi-Maskawa (CKM) matrix is the observed CP-violation source, a 3×3 unitary matrix that can be parameterized by three mixing angles and one complex CP-violating phase [4]. The CKM matrix describes the quarks' mixing because the mass eigenstates are not equal to the weak eigenstates.

A baryon is a type of composite subatomic particle that contains an odd number of valence quarks. A baryon number is a quantum number equal to the number of baryons in a system of subatomic particles minus the number of antibaryons. A lepton number is a quantum number representing the difference between the number of leptons and the number of antileptons in an elementary particle reaction. 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. The mixing of the neutrinos is described by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [4]. 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 (~ 173 GeV for the top quark) [4].

One question that needs to be addressed is any relation between quark mixing and neutrino mixing. Yukawa interaction is an interaction between a scalar field (or pseudoscalar field) and a Dirac field. The Yukawa interactions within the SM cannot be constrained by gauge invariance. To address the open questions, we need new physics models. Supersymmetry is a possible extension of the SM that could resolve the mass hierarchy problems within gauge theory by guaranteeing that quadratic

divergences of all orders will cancel out in perturbation theory. Grand unification theories 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 [5]. 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 LFV Yukawa couplings [5].

The Large Hadron Collider (LHC) [6] 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 $\sqrt{s} = 14 \text{ TeV}$ 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 [7, 8]. 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 the final states with different lepton flavors to avoid the large Drell-Yan background. The Drell-Yan process occurs when a quark of one hadron and an antiquark of another hadron annihilate, creating a virtual photon or Z boson, decaying into a pair of oppositely charged leptons. 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 (pp) 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 fb^{-1} , 41.5 fb^{-1} , and 59.3 fb^{-1} , respectively. Thus, the total integrated luminosity analyzed in this search is 137 fb^{-1} , which corresponds to a four times larger dataset than the one used in the previous searches [9, 10].

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 Monte Carlo (MC) simulations. This gives rise to a better description of the event kinematics and an improvement regarding 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 improvement 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. The experimental setup of LHC and the CMS experiment is discussed in chapter 3. In chapter 4, we 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 discuss the systematic uncertainties in chapter 8 and the corresponding statistical analysis to obtain the results discussed in chapter 9, and finally concluding in chapter 10. Some studies for the future LFV analysis have been performed, and they are detailed in the appendix A and appendix B.

CHAPTER 2

THEORETICAL MOTIVATION

The SM is a set of mathematical formulae and measurements describing elementary particles and their interactions. 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: electron (e), muon (μ), and tau (τ). Each lepton has its corresponding neutral partner: electron neutrino (ν_e), muon neutrino (ν_μ), and tau neutrino (ν_τ). The six flavors of leptons and quarks can be arranged into three generations.

Every fermion has a corresponding anti-particle with the same quantum numbers except for the electric charge. Elementary particles of the SM can be seen in Figure 2.1. The electromagnetic, weak, and strong interactions are part of the SM and can be described as quantum fields. Their interactions are mediated by the gauge bosons, which are spin-one particles. Quantum electrodynamics (QED) is the relativistic quantum field theory (QFT) of the electromagnetic force. The electroweak theory is the unified theory of weak interaction and electromagnetism. The theory of strong interaction is called quantum chromodynamics (QCD).

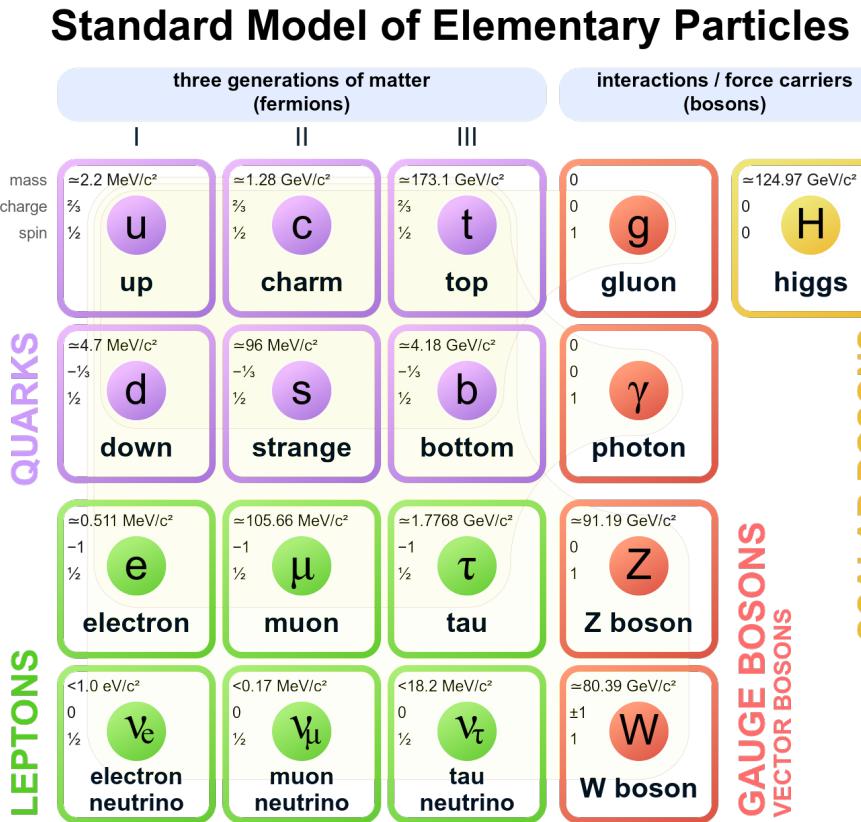


Figure 2.1. Elementary particles of the SM.

2.1 Theoretical building blocks

The SM is a renormalizable QFT and has a $SU(3) \times SU(2) \times U(1)$ symmetry structure. The group theory aims to understand the general, abstract details of structures. The Lagrangian describes the dynamics of a system. A type of QFT is a gauge theory, where gauge fields require symmetry under local transformations. In the following paragraphs, a brief overview of groups and Lagrangian is given, and their application in a QFT is described.

A group is a set, G , together with an operation \star that combines any two elements a and b to form another element, denoted $a \star b$. The set and operation, (G, \star) , must satisfy four requirements to qualify as a group: closure, associativity, the existence of an identity element, and an inverse element. A group is Abelian if the group's elements commute under the group operation ($a \star b = b \star a$). If the elements do not commute, the group is called non-Abelian. The number of elements in the group determines the order n of the group. A group can be represented in terms of $n \times n$ matrices.

Lie groups are a special type of groups that are parameterized by one or more continuous variables. The transformation of elements in Lie groups can be described as matrices M . These matrices are unitary if the Hermitian conjugate is also the inverse of the transformation, $M^\dagger = (M^*)^T = M^{-1}$. A unitary group is a set of such $n \times n$ matrices that preserve the norm under transformations and correspondingly the probability amplitude. If the unitary matrix's determinant is $\det |M| = 1$, then the volume is preserved under such transformations. A set of such matrices form the special unitary group $SU(n)$.

The 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 minus the potential energy $L = T(q, \dot{q}) - V(q)$. In a field theory, the Lagrangian is replaced by the Lagrangian den-

sity \mathcal{L} , which is a function of the fields $\phi(x^\mu)$, which are functions of spacetime and their derivatives $\partial_\mu\phi$. The action S of the system is defined as $S = \int d^4x \mathcal{L}$.

In quantum mechanics, space is treated as an operator, and time is treated as a parameter. In a relativistic theory, space and time have to be treated on an equal footing. QFT interprets the fields as operators parameterized by the spacetime coordinates. It is impossible to measure both the field and its time rate of change simultaneously to infinite precision at a given spacetime point. The vacuum is the state with the lowest possible energy level. When a field operator acts on the vacuum, it produces a state with some energy.

The Lorentz group is the group of all Lorentz transformations of Minkowski spacetime. The Lorentz group contains two copies of the SU(2) group, where the SU(2) group represents spin. The spin of the particle j characterizes the different representations. The Lorentz group can be represented by (j, j') for each of the two SU(2) subgroups. There are three irreducible representations of the Lorentz group:

- The $(0, 0)$ scalar representation.
- The $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ left-handed/right-handed spinor representations.
- The $(\frac{1}{2}, \frac{1}{2})$ vector representation.

Spin is a rotation in the spinor space of SU(2). The spin-0 fields are called scalar fields. Complex scalar fields have the form $\phi = \phi_1 + i\phi_2$, with two real degrees of freedom and hence the fields ϕ and ϕ^\dagger are treated independently.

The spin-1/2 fields are called spinor fields Ψ . The left-handed Ψ_L and right-handed Ψ_R spinors transform the same under rotations but differently under boosts. 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):

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = -2\eta^{\mu\nu} \mathbb{1}$$

with Minkowski metric $\eta^{\mu\nu}$.

A particular type of Lorentz transformations called Parity transformations Λ_P switches the coordinate frame's handedness. Under parity transformations, left-handed spinors are transformed into right-handed spinors. To be Lorentz invariant, the adjoint field of a left-handed spinor field should be the same as the right-handed spinor field $\bar{\Psi}_L = \Psi_R$ and vice versa. The field can be projected onto the left and right-handed components using the projection operators: $P_+ \Psi = \Psi_R$ and $P_- \Psi = \Psi_L$. The charge conjugation is another important transformation: $\Psi \rightarrow C\bar{\Psi}^T = -i\gamma^2\Psi^*$, which swaps the charge of the field.

The spin-1 fields are called vector fields A_μ . Under Lorentz transformations the four components of the vector fields transform as a spacetime vector. The Lorentz invariant terms are:

$$A_\mu A^\mu, \quad (\partial_\mu A_\nu)(\partial^\mu A^\nu), \quad (\partial_\mu A^\mu)(\partial_\nu A^\nu) \quad (2.1)$$

In QFT, the fundamental fermions are represented as spinor fields, while the gauge bosons are represented as vector fields. An interaction term in the combined Lagrangian densities of both fields leads the two fields to interact. Such a term can be introduced using gauge theory.

Gauge theories are theories in which the Lagrangian is invariant under a continuous group of local transformations with a spacetime dependence. The possible transformations are associated with a Lie group. The elements of a group can be expressed in terms of the group generators. The Lagrangian density of a fermion field is defined to satisfy the Dirac equation. To make the Lagrangian density invariant under a local gauge transformation, a new vector field A_μ has to be introduced for each group generator. The vector field can be expressed in terms of the group generators, $A_\mu = A_\mu^a T_a$, and the gauge bosons are the quanta of these fields.

The derivative ∂_μ in the Lagrangian density is replaced by the covariant derivative

D_μ , which is ∂_μ along with a term proportional to a gauge field. The second term of the covariant derivative introduces the interaction terms. The vector field dynamics can be included with a kinetic term proportional to the fields' derivatives. The local gauge symmetry is broken if the Lagrangian has terms proportional to $A_\mu A^\mu$, hence the vector fields are required to be massless. The general principle of local gauge invariance can be used to derive all fundamental interactions of the SM. In the SM, the W and Z bosons acquire their mass through the Higgs mechanism, which combines the principles of gauge theory and spontaneous symmetry breaking.

When a local symmetry is broken, the gauge fields can acquire mass. A global symmetry that is broken results in massless bosons called Goldstone bosons [11, 12]. If we consider a complex scalar field with a Lagrangian that has a local U(1) symmetry and a potential with the vacuum V_{minimum} , then choosing a particular ground state breaks the system's symmetry and is considered “spontaneous” because there are no external means by which this occurs. After a change of basis, the field is expanded around the constant vacuum value, writing the fields in terms of fluctuations around the chosen vacuum. We can choose the vacuum for a local U(1) symmetry so that the vacuum is real and that ϕ is always real; therefore, ϕ can be expanded as $\phi = v + H$, with H being a real scalar field. When a local U(1) symmetry is broken, it results in a real scalar with a mass, and the field A_μ , originating from the local gauge symmetry, acquires mass.

2.2 The Standard Model

The SM 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, L indicates that the interaction only involves left-handed states, and Y stands for the hypercharge. The transformations in the color space are represented by the group $SU(3)_C$ and described by QCD. QCD describes the strong interaction mediated by eight gauge bosons (G_μ^a) called gluons. Quarks

are described as a triplet of spinors that differ in color (red, blue, or green).

Group	Operation	Coupling	Vector field	(2.2)
$U(1)_Y$	Phase	g'	B_μ	
$SU(2)_L$	Weak isospin	g	W_μ^a	
$SU(3)_C$	Color	g_S	G_μ^a	

The electroweak theory is represented by the group $SU(2)_L \times U(1)_Y$. The neutrinos only interact via the weak interaction, which is represented by $SU(2)_L$. As only left-handed fields interact under $SU(2)_L$, only a left-handed neutrino is needed. The left-handed and right-handed fields can interact under the $U(1)_Y$ group, therefore charged leptons have to exist in both a left-handed and a right-handed state. Left-handed doublet fields

$$\psi_L = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix} \quad (2.3)$$

and right-handed singlet fields e_R are introduced. The group $SU(2)_L$ has three associated gauge fields W_μ^a and the group $U(1)_Y$ has one gauge field B_μ . The symmetry groups, couplings, and the associated vector fields are summarized in Table 2.2.

The Higgs field is a complex scalar $SU(2)_L \times U(1)_Y$ doublet field. The Higgs boson is a spin-0 particle and is the field quanta of the Higgs field. The Higgs boson's fundamental vertices with the fermions, the gauge bosons of the weak interaction, and its self-interactions can be seen in Figure 2.2. The Higgs mechanism breaks the electroweak theory into two separate forces: the broken weak theory and the unbroken theory of electromagnetism associated with the $U(1)_{em}$ symmetry of QED. The resulting fields A_μ , W_μ^\pm , and Z_μ^0 are linear combinations of the vector fields of

$SU(2)_L \times U(1)_Y$:

$$\begin{aligned} A_\mu &= \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu \\ W_\mu^\pm &= (W_\mu^1 \mp i W_\mu^2) / \sqrt{2} \\ Z_\mu^0 &= -\cos \theta_W W_\mu^3 + \sin \theta_W B_\mu \end{aligned} \quad (2.4)$$

with the weak mixing angle $\theta_W = \tan^{-1}(g'/g)$, $\sin \theta_W = g'/\sqrt{g^2 + g'^2}$ and $\cos \theta_W = g/\sqrt{g^2 + g'^2}$ written 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. Furthermore, the electromagnetic coupling e of $U(1)_{em}$ is given by $e \equiv g \sin \theta_W$.

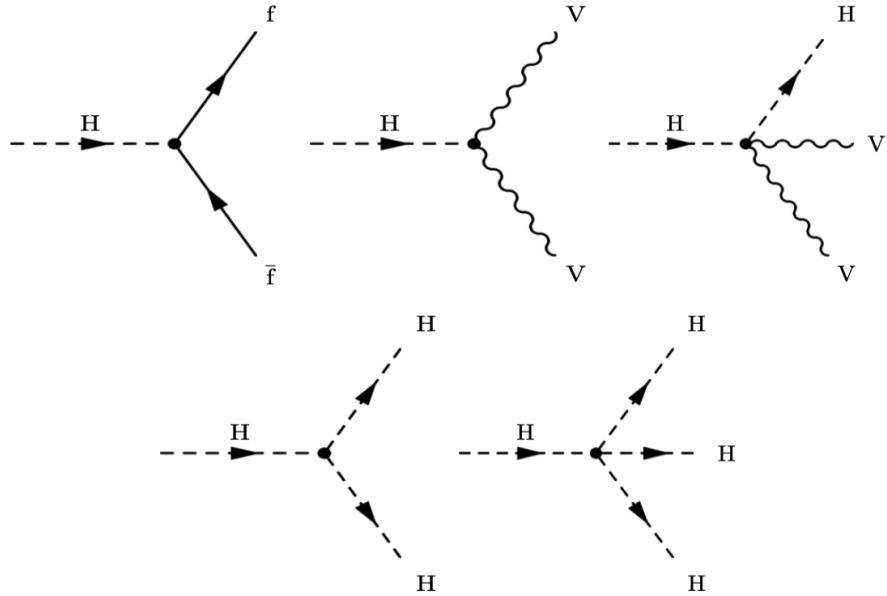


Figure 2.2. Fundamental vertices of the Higgs boson. Fermions are denoted f , anti-fermions are denoted \bar{f} , and gauge bosons of the weak interaction are denoted V .

2.2.1 Electroweak symmetry breaking and the Higgs boson

For the Higgs mechanism, an arbitrary scalar potential

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (2.5)$$

with a self-interacting scalar doublet field Φ is introduced. The term with the coefficient μ^2 is the mass term of the field. The scalar doublet field Φ has four degrees of freedom, which are the CP-even and CP-odd components ϕ^0 and a^0 , and the complex charged component ϕ^+ :

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\phi^+ \\ \phi^0 + ia^0 \end{pmatrix} \quad (2.6)$$

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

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \quad (2.7)$$

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

$$M_W^2 = \frac{g^2 v^2}{4}, \quad M_Z^2 = \frac{(g'^2 + g^2)v^2}{4} \quad (2.8)$$

After electroweak symmetry breaking, there is one remaining degree of freedom of the Higgs field, which is the Higgs boson. The mass of this new scalar particle is

given by $M_H = \sqrt{2\lambda}v$, where λ is the self-coupling parameter. One crucial aspect of the electroweak symmetry breaking is the sign of $\mu^2 = -\lambda v^2$. The expectation value of the Higgs field is fixed by the Fermi coupling G_F : $v = (\sqrt{2}G_F)^{-1/2} \approx 246$ GeV which is measured very precisely in muon decays. In the SM, the fermions acquire mass through the Yukawa interactions between the Higgs field and the fermions. The charge conjugated Higgs doublet $\tilde{\Phi}$ is needed to allow interactions with up-type quarks. The Lagrangian of the Yukawa term is given by:

$$\mathcal{L}_{\text{Yukawa}} = -\hat{\lambda}_{d_{ij}} \bar{q}_{L_i} \Phi d_{R_j} - \hat{\lambda}_{u_{ij}} \bar{q}_{L_i} \tilde{\Phi} d_{R_j} - \hat{\lambda}_{\ell_{ij}} \bar{\ell}_{L_i} \Phi e_{R_j} + \text{h.c.} \quad (2.9)$$

where $\hat{\lambda}$ are 3×3 coupling coefficient matrices for the up and down-type quarks and the charged leptons. After electroweak symmetry breaking, the Higgs field acquires a VEV and the fermion mass eigenstate basis is chosen, so that the Higgs interactions are diagonalised: $\hat{\lambda}_{f_{ij}} \rightarrow \lambda_f \delta_{ij}$. The mass of the fermions is given by $m_{f_i} = \lambda_f v / \sqrt{2}$, with the corresponding Yukawa coupling λ_f . The Higgs boson was discovered by the ATLAS and CMS collaborations in 2012 [13]. Its mass and full width were measured to be $M_H = 125.09 \pm 0.24$ GeV and $\Gamma < 1.7$ GeV.

2.3 Lepton flavor violating decays of the Higgs boson

The SM is considered to be an effective theory up to a certain scale Λ . The theoretical description of the SM does not incorporate the flavor structure. There are three main puzzles in the SM related to flavor [5]:

- The reason for three generations of quarks and leptons.
- The principle underlying the formation of the Yukawa matrices describing the SM Yukawa interactions.
- The peculiar pattern of fermion masses and mixing.

The Yukawa interactions with the Higgs field gives charged leptons their mass.

In the SM, LFV decays of the Higgs boson are forbidden because the Yukawa interaction matrix can be diagonalized in the mass basis. The SM extensions like multi-Higgs doublet models can introduce LFV Yukawa couplings along with additional CP-violation. These models can have tree-level Higgs mediated flavor changing neutral currents leading to LFV Yukawa couplings. Also, models like composite Higgs models [14] and models with extra dimensions can introduce LFV Yukawa couplings.

A direct mass term of the charged leptons is not gauge invariant, but interactions of the charged leptons to the scalar Higgs field can be introduced with the following term in the Lagrangian:

$$\mathcal{L}_{\text{SM}} = -\lambda_{ij} L_L^i \Phi \ell_R^j + \text{h.c.} \quad (2.10)$$

where with $\bar{L}_L^i = (\bar{\nu}_\ell^i, \bar{\ell}_L^i)$, are the $SU(2)_L$ doublets, and $\bar{\ell}_R^i$ are 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 VEV and the physical Higgs boson leading to the following Lagrangian:

$$\begin{aligned} \mathcal{L}_{\text{SM}} &= -\frac{\lambda_{ij}}{\sqrt{2}}(v + H)\bar{\ell}_L^i \ell_R^j + \text{h.c.} \\ &= -\frac{\lambda_{ij}}{\sqrt{2}}v \bar{\ell}_L^i \ell_R^j - \frac{\lambda_{ij}}{\sqrt{2}}\ell_L^i \ell_R^j H + \text{h.c.} \end{aligned} \quad (2.11)$$

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 :

$$m = \begin{pmatrix} m_e & 0 & 0 \\ 0 & m_\mu & 0 \\ 0 & 0 & m_\tau \end{pmatrix} = \frac{1}{\sqrt{2}}V_L \lambda V_R^\dagger v = \frac{\lambda^m}{\sqrt{2}}v \quad (2.12)$$

where $\lambda^m = V_L \lambda V_R^\dagger$ is written in terms of the mass basis. The Yukawa interaction

matrix can also be written in terms of the mass basis:

$$Y = V_L \frac{\lambda}{\sqrt{2}} V_R^\dagger = \frac{\lambda^m}{\sqrt{2}} = \frac{m}{v} = \begin{pmatrix} Y_{ee} & 0 & 0 \\ 0 & Y_{\mu\mu} & 0 \\ 0 & 0 & Y_{\tau\tau} \end{pmatrix} \quad (2.13)$$

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

Dimension-6 operators: Assuming that the Higgs field is the only field that causes electroweak symmetry breaking and that the particle spectrum is only composed of the SM particles up to an energy scale $\Lambda \gg 200$ GeV, we can integrate out the additional heavy fields, which gives us an effective field theory. The SM Lagrangian, which has dimension four, can be extended with higher-dimension operators. An effective Lagrangian with additional dimension-6 operators O_n^{ij} can be written as:

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \sum_{nij} \frac{\alpha_n^{ij}}{\Lambda^2} O_n^{ij} \quad (2.14)$$

where i, j denote flavor indices, n runs over the number of independent operators, Λ is the scale of new physics, and α_n^{ij} are coefficients [15].

The effects of new physics at the electroweak scale can be described in a model-independent way using this effective Lagrangian. The effective Lagrangian can also contain a dimension-5 operator, which can generate neutrino masses, but it cannot generate LFV interactions for the Higgs boson. However, LFV interactions of the Higgs boson can be generated using the dimension-6 operator $O_{L\phi}^{ij} = (\Phi^\dagger \Phi)(\bar{L}_i \ell_{Rj} \Phi)$ which is a Yukawa-type operator. The dimension-6 operator also contributes to the fermion mass matrices. After including these contributions in the fermion mass matrices, which are then diagonalized and the interaction term of the Lagrangian is written in terms of the mass basis. The Lagrangian has the following additional term

after including the Yukawa-type operator:

$$\Delta\mathcal{L} = -\frac{\lambda'_{ij}}{\Lambda^2}(\Phi^\dagger\Phi)(\bar{L}_L^i\ell_R^j\Phi) + h.c. \quad (2.15)$$

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

Electroweak symmetry breaking leads to the following Lagrangian terms:

$$\begin{aligned} \Delta\mathcal{L} &= -\frac{\lambda'_{ij}}{\Lambda^2}\left(\frac{v+H}{\sqrt{2}}\right)^2\bar{\ell}_L^i\ell_R^j\left(\frac{v+H}{\sqrt{2}}\right) \\ &= -\frac{\lambda'_{ij}}{\Lambda^2}\frac{1}{(\sqrt{2})^3}(v^3 + 3v^2H + 3vH^2 + H^3)\bar{\ell}_L^i\ell_R^j \\ &= -\frac{1}{\sqrt{2}}\frac{\lambda'_{ij}}{2\Lambda^2}(v^3\bar{\ell}_L^i\ell_R^j + 3v^2\bar{\ell}_L^i\ell_R^jH + 3v\bar{\ell}_L^i\ell_R^jH^2 + \bar{\ell}_L^i\ell_R^jH^3) \end{aligned} \quad (2.16)$$

where the term proportional to $\bar{\ell}_L^i\ell_R^j$ leads to an additional mass term of the charged leptons and the term proportional to $\bar{\ell}_L^i\ell_R^jH$ adds a further Yukawa-interaction term. Thus, the mass term of the effective Lagrangian is the sum of the SM mass term and the additional contribution:

$$\begin{aligned} \mathcal{L}_{\text{mass}} &= \mathcal{L}_{\text{mass,SM}} + \Delta\mathcal{L}_{\text{mass}} \\ &= -\left(\frac{\lambda_{ij}}{\sqrt{2}}v + \frac{1}{\sqrt{2}}\frac{\lambda'_{ij}}{2\Lambda^2}v^3\right)\bar{\ell}_L^i\ell_R^j \\ &= -\frac{1}{\sqrt{2}}\left(\lambda_{ij} + \frac{v^2}{2\Lambda^2}\lambda'_{ij}\right)v\bar{\ell}_L^i\ell_R^j \end{aligned} \quad (2.17)$$

The term of the effective Lagrangian which describes the Yukawa interactions is given by:

$$\begin{aligned} \mathcal{L}_Y &= \mathcal{L}_{Y,\text{SM}} + \Delta\mathcal{L}_Y \\ &= -\left(\frac{\lambda_{ij}}{\sqrt{2}} + \frac{1}{\sqrt{2}}\frac{\lambda'_{ij}}{2\Lambda^2}3v^2\right)\bar{\ell}_L^i\ell_R^jH \\ &= -\frac{1}{\sqrt{2}}\left(\lambda_{ij} + 3\frac{v^2}{2\Lambda^2}\lambda'_{ij}\right)\bar{\ell}_L^i\ell_R^jH \end{aligned} \quad (2.18)$$

The mass matrix of the charged leptons can be diagonalized as for the SM using

other matrices V_L, V_R :

$$\sqrt{2}m = V_L[\lambda + \frac{v^2}{2\Lambda^2}\lambda']V_R^\dagger v \quad (2.19)$$

The Yukawa interaction matrix can then be written in terms of the mass basis:

$$\begin{aligned} \sqrt{2}Y &= V_L[\lambda + 3\frac{v^2}{2\Lambda^2}\lambda']V_R^\dagger \\ &= \frac{\sqrt{2}m}{v} + V_L 2\frac{v^2}{2\Lambda^2}\lambda' V_R^\dagger \\ &= \frac{\sqrt{2}m}{v} + \frac{v^2}{\Lambda^2}V_L\lambda' V_R^\dagger \\ &= \frac{\sqrt{2}m}{v} + \frac{v^2}{\Lambda^2}\hat{\lambda} \end{aligned} \quad (2.20)$$

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} with possible LFV Yukawa interactions:

$$Y_{ij} = \frac{m_i}{v}\delta_{ij} + \frac{v^2}{\sqrt{2}\Lambda^2}\hat{\lambda}_{ij} \quad (2.21)$$

Type-III 2HDM: The two Higgs doublet model (2HDM) is one of the most studied extensions of the SM. The structure of Yukawa couplings can characterize them. In Type-I, Type-II, and Type-X 2HDM, each fermion type is coupled to only one scalar doublet. Thus there is no flavor violating Yukawa couplings of the neutral scalar boson in this case. In Type-III 2HDM, however, both of the scalar doublets couple to all the fermions. As a result, there is a possibility of flavor violation in the neutral scalars Yukawa couplings [16].

In the Higgs basis, the VEV resides only in Φ_1 and the fields Φ_1 and Φ_2 can be

expanded as

$$\Phi_1 = \begin{pmatrix} G^+ \\ \frac{1}{\sqrt{2}}(v + \phi_1 + iG^0) \end{pmatrix}, \quad \Phi_2 = \begin{pmatrix} H^+ \\ \frac{1}{\sqrt{2}}(\phi_2 + iA) \end{pmatrix} \quad (2.22)$$

where v is the VEV, G^\pm and G^0 are the would-be Goldstone bosons, H^\pm is the charged Higgs, A is the neutral CP-odd Higgs, ϕ_1 and ϕ_2 are the neutral CP-even Higgs. The fields ϕ_1 and ϕ_2 are not mass eigenstates. They are related to the mass eigenstates h and H by

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} h \\ H \end{pmatrix} \quad (2.23)$$

The Yukawa sector in the Type-III 2HDM is given by

$$\mathcal{L}_{\text{Yukawa}} = -\frac{\sqrt{2}m_\ell^i}{v}\delta^{ij}\bar{L}_L^i\ell_R^j\Phi_1 - \sqrt{2}Y_\ell^{ij}\bar{L}_L^i\ell_R^j\Phi_2 \quad (2.24)$$

where m_ℓ are lepton masses, Y_f are the Yukawa coupling matrices, and the indices i, j, k run over lepton families. After electroweak symmetry breaking, the Yukawa couplings in the physical basis read

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} &\supset -y_{\ell,h}^{ij}\bar{\ell}_L^i\ell_R^j h - y_{\ell,H}^{ij}\bar{\ell}_L^i\ell_R^j H - y_{\ell,A}^{ij}\bar{\ell}_L^i\ell_R^j A + \text{h.c.} \\ y_{f,h}^{ij} &= \frac{m_f^i}{v}\delta^{ij}\cos\alpha - Y_f^{ij}\sin\alpha \\ y_{f,H}^{ij} &= \frac{m_f^i}{v}\delta^{ij}\sin\alpha + Y_f^{ij}\cos\alpha \end{aligned} \quad (2.25)$$

LFV Yukawa couplings: The inclusion of LFV Yukawa interactions were discussed by introducing a dimension-6 operator to the Lagrangian and in theories with more than one Higgs doublet, where the scalar fields can mix. There are several possible additional terms to the Lagrangian due to new physics, which prevent the simultaneous diagonalization of mass matrix and the Yukawa matrix. 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:

$$\mathcal{L}_{\text{Yukawa}} = -m_i \bar{\ell}_L^i \ell_R^i - Y_{ij} \bar{\ell}_L^i \ell_R^j H + \text{h.c.} \quad (2.26)$$

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

$$Y = \begin{pmatrix} Y_{ee} & Y_{e\mu} & Y_{e\tau} \\ Y_{\mu e} & Y_{\mu\mu} & Y_{\mu\tau} \\ Y_{\tau e} & Y_{\tau\mu} & Y_{\tau\tau} \end{pmatrix} \quad (2.27)$$

which can have LFV Yukawa couplings Y_{ij} . LFV Yukawa couplings allow LFV decays of the Higgs boson.

The Higgs boson couples to all particles according to their mass in the SM. The LFV Yukawa couplings might be related closely to the fermion mass matrices in the presence of new physics, reflecting the observed fermion mass hierarchy [17]. The LFV couplings then have a hierarchical structure given by $\Delta_{ij}\sqrt{m_i m_j}$, where i, j are the generation indices, and Δ_{ij} is related to the mass mixing. The mass corrections to the lepton masses coming from mass mixing are assumed to be small in the natural assumption. The SM Yukawa couplings are taken for the flavor diagonal couplings. The LFV branching fractions can be expressed in terms of the SM ones:

$$\mathcal{B}_{\text{nat}}(H \rightarrow \ell_i \ell_j) = \mathcal{B}_{\text{SM}}(H \rightarrow \ell_i \ell_i) \cdot \frac{m_{\ell_j}}{m_{\ell_i}} \quad (2.28)$$

with the total decay width assumed to be the SM one, $\mathcal{B}_{\text{nat}}(H \rightarrow \ell_i \ell_j)$ being the natural branching fraction and $\mathcal{B}_{\text{SM}}(H \rightarrow \ell_i \ell_i)$ being the branching fraction for the SM Higgs boson decay to two leptons. The LFV branching fractions are expected to be of similar size or smaller than $\mathcal{B}_{\text{nat}}(H \rightarrow \ell_i \ell_j)$. A naturalness assumption referred

to as theoretical naturalness limit can also be derived for the Yukawa couplings:

$$|Y_{ji} Y_{ij}| \lesssim \frac{m_i m_j}{v^2} \quad (2.29)$$

2.4 Constraints from low-energy measurements

Constraints on the Higgs boson's LFV decays have been derived using low-energy measurements and with certain assumptions on the value of flavor diagonal Yukawa couplings. The constraints from several low-energy measurements on LFV Yukawa couplings are discussed in this section. These measurements and the assumptions used to derive constraints on the LFV Yukawa couplings are presented.

Constraints from LFV decays $\ell_i \rightarrow \ell_j \gamma$: LFV Yukawa couplings also contribute to LFV decays of the type $\ell_i \rightarrow \ell_j \gamma$, where i, j are flavor indices with $i \neq j$. One-loop and two-loop diagrams of this process are shown in Figure 2.3. Constraints of the type $\sqrt{|Y_{ij}|^2 + |Y_{ji}|^2}$ can be derived assuming SM values for $Y_{\tau\tau}$, $Y_{\mu\mu}$, and Y_{tt} . 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.

Constraints from LFV decays $\ell_i \rightarrow 3\ell_j$: LFV Yukawa couplings also contribute to LFV decays of the type $\ell_i \rightarrow 3\ell_j$. Such decays have been searched for and constraints of the type $\sqrt{|Y_{ij}|^2 + |Y_{ji}|^2}$ can be derived assuming SM values for $Y_{\tau\tau}$, $Y_{\mu\mu}$, and Y_{tt} . The Z boson contribution to the LFV decay $\tau \rightarrow \mu$ are neglected because their contributions are negligible [18].

Constraints from muonium-antimuonium oscillations: The muonium (M) bound state $\mu^+ e^-$ can oscillate to the antimuonium (\bar{M}) bound state $e^+ \mu^-$. An upper limit on the conversion probability $P(M \rightarrow \bar{M}) < 8.3 \times 10^{-11}$ [19] was derived by the muonium-antimuonium conversion spectrometer experiment. 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 $|Y_{\mu e} + Y_{e\mu}^*|$ [20].

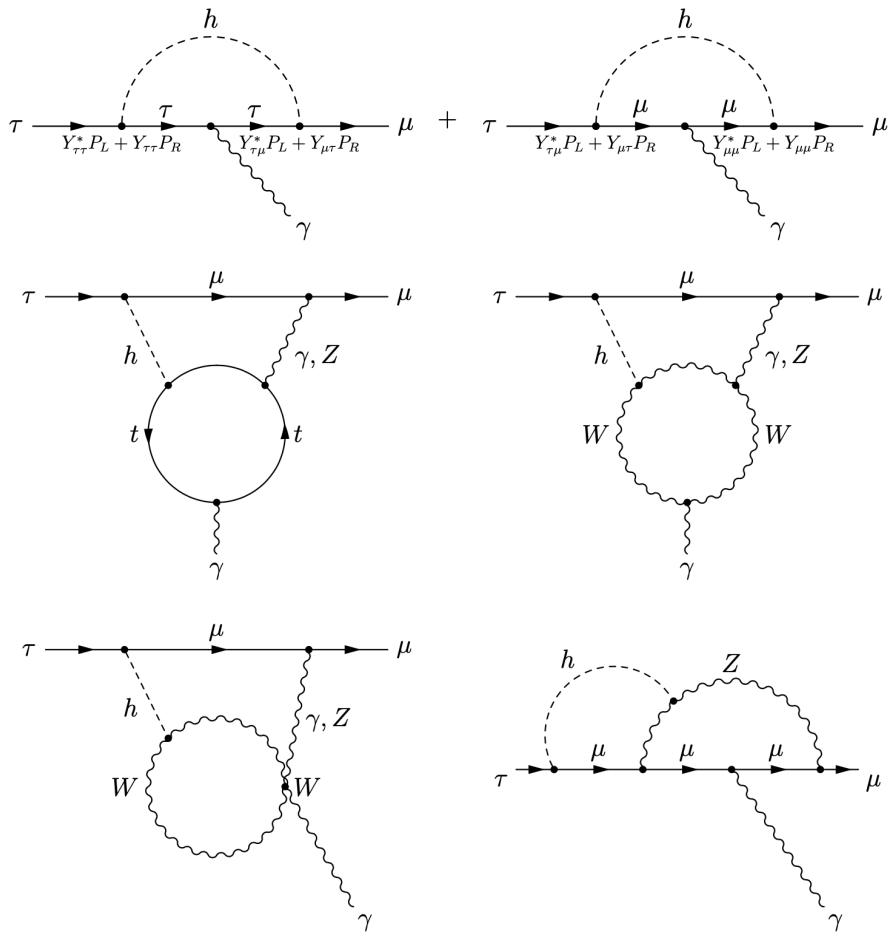


Figure 2.3. Diagrams contributing to the flavor violating decay $\tau \rightarrow \mu\gamma$, mediated by a Higgs boson with flavor violating Yukawa couplings.

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_μ/m_τ or m_τ/m_H , the LFV contribution to $g_\mu - 2$ due to the one-loop diagram has the form

$$a_\mu \equiv \frac{g_\mu - 2}{2} \propto \text{Re}(Y_{\mu\tau} Y_{\tau\mu}) \quad (2.30)$$

with the discrepancy between measurement and SM prediction having the size

$$\Delta a_\mu \equiv a_\mu^{\text{exp}} - a_\mu^{\text{SM}} = (2.87 \pm 0.63 \pm 0.49) \times 10^{-9} \quad (2.31)$$

where a_μ^{exp} and a_μ^{SM} are the measured and predicted value, respectively. The one-loop diagram can contribute to the muon's electric dipole moment if the LFV Yukawa couplings are complex. If the terms suppressed by m_μ/m_τ or m_τ/m_H are neglected, the dependence on the LFV Yukawa couplings of the electric dipole moment d_μ is

$$d_\mu \propto -\text{Im}(Y_{\mu\tau} Y_{\tau\mu}) \quad (2.32)$$

Constraints from $\mu \rightarrow e$ conversion in nuclei: LFV Yukawa couplings can also contribute to $\mu \rightarrow e$ conversion in nuclei via a tree-level exchange of the Higgs boson and one-loop diagrams with the Higgs boson and a photon exchange. The two-loop diagrams' contribution is larger than the one-loop diagrams because they are only suppressed by the weak gauge coupling or Y_{tt} and were taken into account for deriving the constraints on the LFV Yukawa couplings.

Constraints on the LFV Yukawa couplings from low-energy measurements are summarised in Table 2.1. The LFV decays of type $\ell_i \rightarrow \ell_j \gamma$ give the strongest constraints. 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 calculated to be $\lesssim \mathcal{O}(10^{-8})$. Constraints on the branching fractions

were derived under the assumption that only one of them contributes in addition to the SM Higgs boson's total width.

TABLE 2.1

CONSTRAINTS ON LFV YUKAWA COUPLINGS FROM
LOW-ENERGY MEASUREMENTS [20].

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

2.5 Lepton flavor violating decays of the Higgs boson at the LHC

In the SM, the Higgs boson couples 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 $\mathcal{B}_{\text{nat}}(H \rightarrow \ell_i \ell_j)$ using the natural assumption. Furthermore, LFV branching fractions $\mathcal{B}_{\text{nat}}(H \rightarrow \ell_i \ell_j)$ 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%, between decays to taus and muons. LFV decays of the Higgs boson to $e\tau$ would have a smaller natural branching fraction in the order of $0.001\% = 10^{-5}$.

Low-energy measurements 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. 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.

The CMS experiment published the first direct search for $H \rightarrow \mu\tau$ [21], followed by searches for $H \rightarrow e\tau$ and $H \rightarrow e\mu$ decays [22], using pp collision data corresponding to an integrated luminosity of 19.7 fb^{-1} at a center-of-mass energy of 8 TeV. A small excess of data with respect to the SM background-only hypothesis at $M_H = 125 \text{ GeV}$ was observed in the $H \rightarrow \mu\tau$ channel, with a significance of 2.4 standard deviations (σ), and the best fit for the branching fraction was found to be $\mathcal{B}(H \rightarrow \mu\tau) = (0.84^{+0.39}_{-0.37})\%$. A constraint was set on the observed (expected) branching fraction $\mathcal{B}(H \rightarrow \mu\tau) < 1.51\%(0.75\%)$ at 95% CL. This search has improved the limit on the branching fraction $\mathcal{B}(H \rightarrow \mu\tau)$ from the indirect limit by one order of magnitude.

No excess of events over the estimated background was observed in the $H \rightarrow e\tau$ or

$H \rightarrow e\mu$ channels, and observed (expected) upper limits on the branching fractions $\mathcal{B}(H \rightarrow e\tau) < 0.69\%(0.75\%)$ and $\mathcal{B}(H \rightarrow e\mu) < 0.035\%(0.048\%)$ at 95% CL were set. The ATLAS Collaboration reported searches for $H \rightarrow e\tau$ and $H \rightarrow \mu\tau$ using pp collision data at a center-of-mass energy of 8 TeV, finding no significant excess of events over the background expectation, and set observed (expected) limits of $\mathcal{B}(H \rightarrow \mu\tau) < 1.43\%(1.01\%)$ and $\mathcal{B}(H \rightarrow e\tau) < 1.04\%(1.21\%)$ at 95% CL [23, 24].

The CMS experiment placed an upper limit on the observed (expected) branching fraction $\mathcal{B}(H \rightarrow \mu\tau) < 0.25\%(0.25\%)$ and $\mathcal{B}(H \rightarrow e\tau) < 0.61\%(0.37\%)$ at 95% CL using the 2016 dataset corresponding to an integrated luminosity of 35.9 fb^{-1} [9]. The ATLAS experiment published the search results using the 2016 dataset corresponding to an integrated luminosity of 36.1 fb^{-1} and placed an upper limit of 0.28% (0.37%) and 0.47% (0.34%) on the $\mathcal{B}(H \rightarrow \mu\tau)$ and $\mathcal{B}(H \rightarrow e\tau)$ with a 95% CL, respectively [10]. Figure 2.4 shows the expected and observed 95% CL upper limits for each category and their combination of the search performed by the CMS experiment. Figure 2.5 shows the corresponding Yukawa couplings.

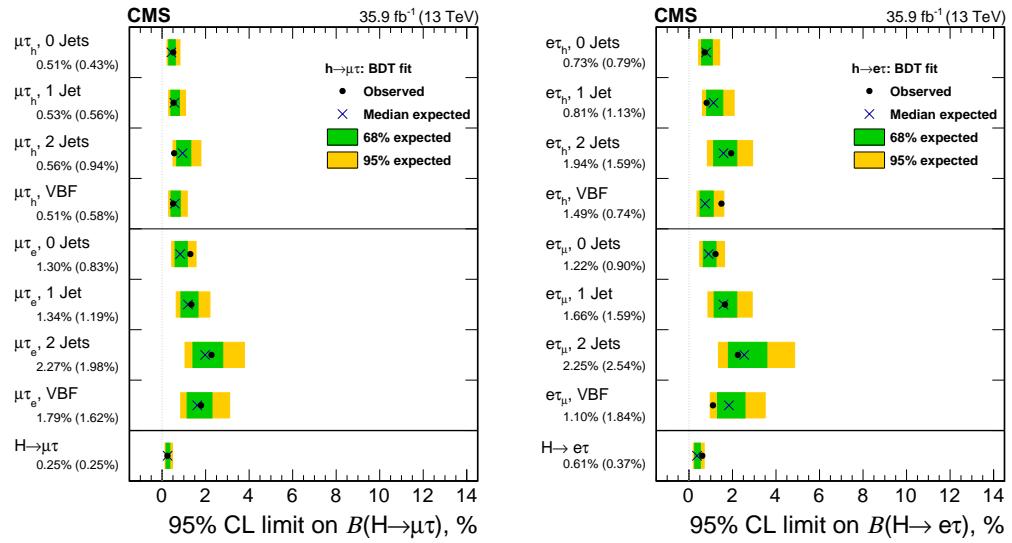


Figure 2.4. Expected and observed 95% CL upper limits for each category and their combination of the search performed by the CMS experiment [25].

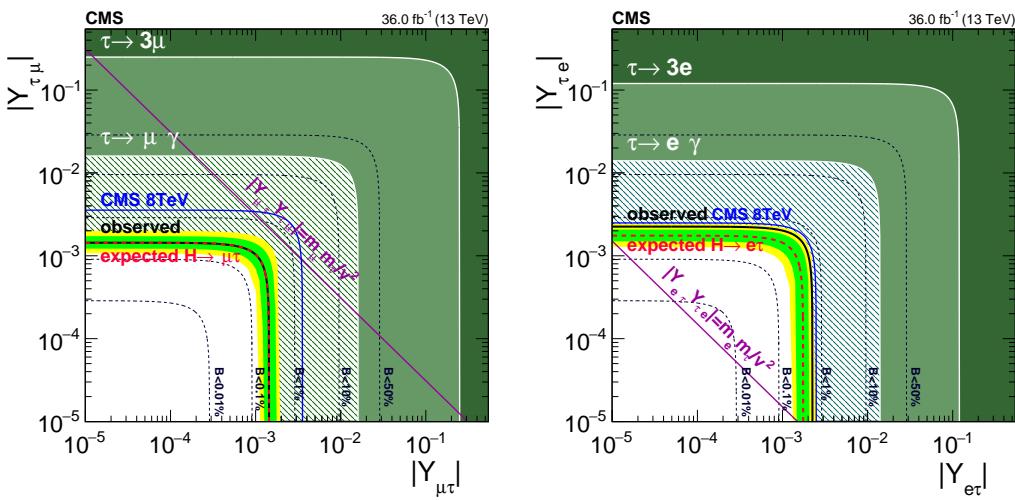


Figure 2.5. Constraints on the LFV Yukawa couplings, $|Y_{\mu\tau}| - |Y_{\tau\mu}|$ (left), and $|Y_{e\tau}| - |Y_{\tau e}|$ (right) of the search performed by the CMS experiment.

CHAPTER 3

LHC AND THE CMS DETECTOR

3.1 Introduction

The prime goals of the LHC and the CMS [26] experiment are exploring the physics at the TeV scale and studying the mechanism of electroweak symmetry breaking, including studying the SM Higgs boson's properties along with searches for new particles predicted by beyond the SM theories. The physics program's other main interests include studying 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 quarks and gluons; therefore, two colliding partons' initial energy is not known. In contrast, the collision 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 $\sqrt{s} = 14 \text{ TeV}$ and reach a luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. Lead ions can be accelerated up to an energy of 2.8 TeV per nucleon and reach a luminosity of $10^{27} \text{ cm}^{-2}\text{s}^{-1}$. The 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 [27] 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 [28], where they are accelerated to 50 MeV. They are further accelerated in the Proton Synchrotron Booster [29] to 1.4 GeV. They are then injected in the Proton Synchrotron [30], where they are accelerated to 25 GeV. A bunch train is produced within the Proton Synchrotron before extraction. The protons are then accelerated in the Super Proton Synchrotron [31] to 450 GeV before injecting in the LHC.

The four interaction points are equipped with particle detectors: the CMS experiment, the “A Toroidal LHC ApparatuS” ATLAS experiment [32], the “A Large Ion Collider Experiment” ALICE experiment [33], and the “LHC beauty” LHCb experiment [34]. Two further smaller experiments, the “Total Cross Section, Elastic Scattering and Diffraction Dissociation” TOTEM [35] and the “LHC forward” LHCf [36] experiments, are located near the CMS and the ATLAS interaction points, respectively.

The ATLAS and CMS experiments have both multi-purpose detectors installed. The detectors were built to detect particles from pp or heavy-ion collisions. One of the main tasks currently is to study the production and decay of the discovered Higgs boson [13], 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 2 data-taking, a bunch spacing of 25 ns was used, and pp 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 pp collision data-taking period. CMS does not record the whole delivered data, and only part of the recorded data that is considered good is used for physics analysis.

TABLE 3.1

INTEGRATED LUMINOSITY CONSIDERED FOR PHYSICS
ANALYSIS AT THE CMS EXPERIMENT DURING RUN 2

Year	Integrated luminosity
2016	35.9 fb^{-1}
2017	41.5 fb^{-1}
2018	59.3 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 at the heart of CMS. A considerable bending power for the momentum measurement of charged particles within a compact design is achieved using a high magnetic field of 3.8 T. The internal part of the solenoid 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

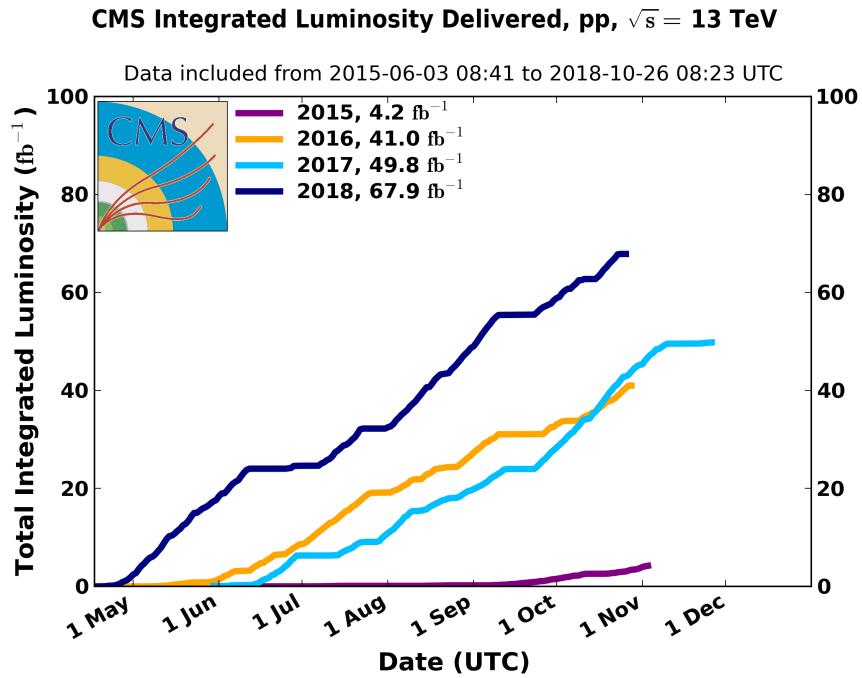


Figure 3.1. Cumulative luminosity as a function of time delivered to CMS during stable beams for pp collisions. The luminosity is shown for 2015 (purple), 2016 (orange), 2017 (light blue), and 2018 (dark blue) [37].

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. The trajectories are precisely measured with the high granular silicon strip tracker, which can deal with high charged particle multiplicities.

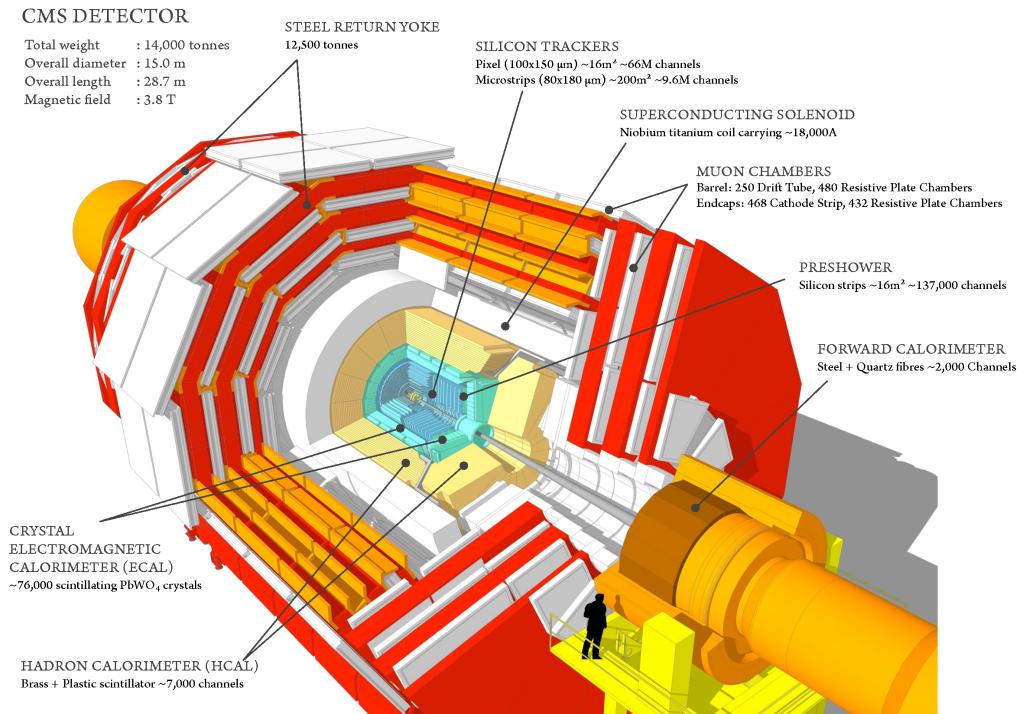


Figure 3.2. Schematic view of the CMS detector.

The energy of the particles is measured in calorimeters. The calorimeter material initiates electromagnetic 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 λ_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 electromagnetic calorimeter (ECAL) made of lead tungstate (PbWO_4) crystals 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 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

The CMS experiment uses a right-handed coordinate system, with the origin at the nominal collision point. The x-axis points towards the LHC center, while the y-axis points upward towards the surface (perpendicular to the LHC plane), and the z-axis points along the anti-clockwise beam direction. Two angles are defined, where the azimuthal angle ϕ is measured from the x-axis in the x-y plane, and the polar angle θ 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 ϕ and η are denoted $\Delta\phi$ and $\Delta\eta$. These distance measures are used to define the spatial separation between physics objects by ΔR with

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

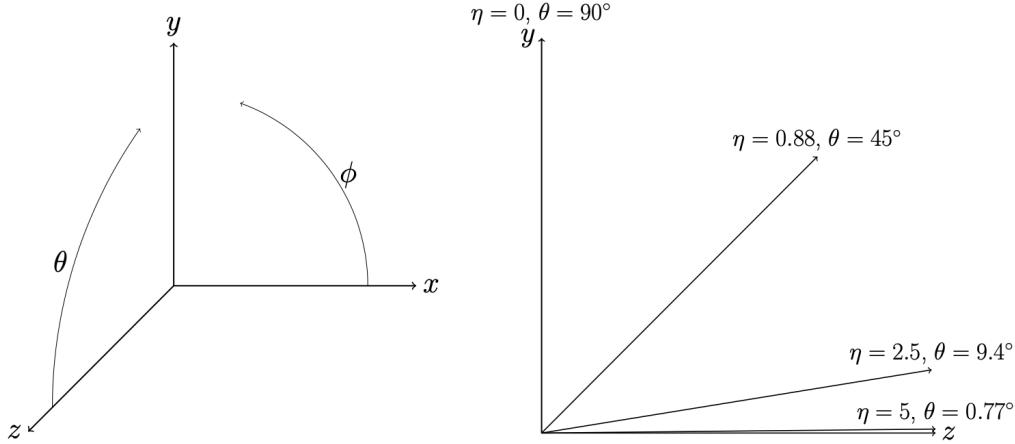


Figure 3.3. Coordinate system convention of CMS (left) and the relation between pseudorapidity η and polar angle θ (right)

3.3.2 Kinematic quantities

At LHC, the protons carry half of the collision energy. The hard interactions occur between the colliding protons' partons, and each parton carries a fraction 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

$$\begin{aligned} p_1 &= x_1 \cdot \frac{\sqrt{s}}{2} (1, 0, 0, 1) \\ p_2 &= x_2 \cdot \frac{\sqrt{s}}{2} (1, 0, 0, -1) \end{aligned} \tag{3.3}$$

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

$$\hat{s} \equiv M^2 = (p_1 + p_2)^2 = \frac{s}{4} \cdot [(x_1 + x_2)^2 - (x_1 - x_2)^2] = x_1 x_2 s \tag{3.4}$$

where $\sqrt{\hat{s}}$ denotes the center-of-mass energy of the parton-parton collision. The hard collision products have a zero total momentum 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. The transverse momentum is denoted by p_T and the transverse energy is denoted by E_T with $E_T = E \cdot \sin\theta$. The weakly interacting particles, like 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

Isolated leptons in the final states leave clean signatures in the CMS detector. The muon has to be well-identified, and its momentum resolution needs to be good over a wide range of momenta. The charge of the muon has to be determined unambiguously. In addition, electrons and charged hadrons need to be reconstructed and identified with good efficiency and momentum resolution. The momenta of charged particles are measured using the curvature of their trajectory. Hence, considerable bending power is needed to measure precisely the charge of particles with large momentum. Furthermore, a good electromagnetic energy resolution is needed over a wide geometrical coverage and correct localization of the primary interaction.

The CMS detector can detect the hadrons produced from the hadronization of the quarks. Tau leptons that decay hadronically need to be identified with good efficiency for the Higgs boson physics. A good measurement of the impact parameter of charged particle tracks and good position measurement of the secondary vertices can improve the triggering and identification efficiency of hadronically decaying taus. This requires having the pixel detector 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 measuring the hadrons energy with good resolution and for the estimation of \cancel{E}_T .

The collisions are happening at a rate of 40 MHz 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 ~ 1 kHz. The short time between two bunch crossings, 25 ns, has a major implication on the read-out and trigger system design. Multiple pp interactions (pileup) happen in one bunch crossing. The interactions' products overlap and can be wrongly linked, and 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 [38]. The conductor is made of high purity and is aluminum stabilized with a four-layer winding and is composed of five coils and has to withstand outward pressure. The superconducting wire is cooled using an indirect cooling by thermosyphon.

3.3.5 Inner tracking system

The inner tracking system [39] 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 cylindrical shape with a length of $5.8\ \text{m}$ and a diameter of $2.6\ \text{m}$. The pixel detector consists of three layers at radii of $4, 7$, and $11\ \text{cm}$ 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 $23\text{--}34\ \mu\text{m}$ in the r -direction and $23\ \mu\text{m}$ in z , while it is $35\text{--}52\ \mu\text{m}$ 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 endcap made of 9 disks and the tracker inner disks made of three small disks, which fill the gap between the TIB and the tracker endcap.

3.3.6 Electromagnetic calorimeter

The ECAL is a hermetic, homogeneous calorimeter with a coverage of $|\eta| < 3$, and it measures the energy of electrons and photons. It has an excellent energy resolution and high granularity and can separate close clusters very well [40, 41]. ECAL uses PbWO_4 crystals, which have a short radiation length of $X_0 = 0.89\text{cm}$. These crystals are radiation hard and emit blue-green scintillation light with a maximum at $420\ \text{nm}$. The crystals allow a compact calorimeter inside the solenoid that is fast, has a fine granularity and is radiation-resistant. The decay time for the scintillation is short and in the same order as the LHC bunch crossing time, and about 80% of the

light is emitted within 25 ns. However, the light output depends on the temperature. Therefore, to preserve energy resolution, a cooling system is needed. Besides, as the light yield is low, this requires photodetectors with intrinsic gain that can also operate in a high magnetic field. Figure 3.4 illustrates the layout of the ECAL.

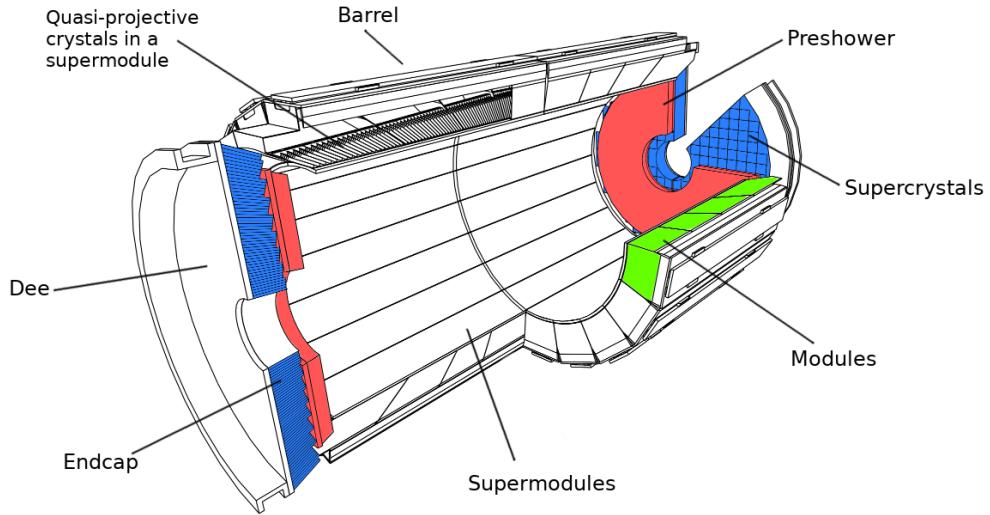


Figure 3.4. Schematic view of the ECAL.

The calorimeter is divided into a ECAL barrel (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 are installed in front of each endcap and cover the region $1.653 < |\eta| < 2.6$. In the EB, the scintillation light is detected with silicon avalanche photodiodes (APD). These photodiodes respond to temperature changes and 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 $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$. In the EE, the scintillation light is detected with vac-

uum phototriodes. A 5×5 array of crystals is termed supercrystals. The crystals and the supercrystals are arranged in a rectangular x-y grid. The preshower system is installed in front of EE to identify and reject π^0 mesons, which dominantly decay into two photons. It has a higher granularity than the EE and improves the position determination of electrons and photons. It is a sampling calorimeter consisting of two layers of a lead absorber, which initiate the electromagnetic showers from incoming electrons or photons. The lead radiators are followed by silicon strip sensors measuring the energy deposits and the transverse shower profiles. The silicon strip sensors are oriented orthogonally and are placed after a radiation length of $1 X_0$ and $2 X_0$ of the lead absorber.

The ECAL calibration and performance is estimated from the supercluster energy deposited by the electrons and photons. The calibration is performed, which corrects the detector's response to the energy deposited to match expectations. The calibration is applied to the superclusters to take into account the η and ϕ -dependent geometry and material effects, as well as the fact that electrons and photons shower slightly differently. The calibrations were tuned and validated using electrons' energy from W boson decays, the reconstructed mass from η decays to two photons, and the energy resolution measured with $Z \rightarrow ee$ events. I implemented an event skimmer using the updated electron identification and isolation selections to select events with $Z \rightarrow ee$ resulting from the pp collisions. The primary motivation for implementing the event skimmer is to obtain datasets enriched with $Z \rightarrow ee$ events that can be further used for performing calibration in ECAL, and because trigger selections are not applied in the event skimmer, it can be used for trigger efficiency studies. One primary difference from the old event skimmer is using the effective area method to calculate electrons' isolation instead of the delta beta method. The description of these methods is given in section 5.7. The event skimmer's acceptance is tested on the 2017 data, and the obtained event rate is within expectations. It is currently

integrated into the CMS software framework.

3.3.7 Hadron calorimeter

The HCAL is a sampling calorimeter, and it measures the energy of hadrons. The HCAL system is located surrounding the ECAL system, and most of it is located inside the solenoid. The HCAL design has an important requirement to minimize the non-Gaussian tails in the energy resolution and to provide a good hermeticity for the determination of E_T . Another design requirement is to maximize the interaction length of the material within the solenoid. This corresponds to minimizing the space allocated to the active material. It is divided into a HCAL barrel (HB) detector, covering the region $|\eta| < 1.4$ and two HCAL endcap (HE) detectors, covering $1.3 < |\eta| < 3.0$. The absorber material is brass. Brass is chosen because it has a relatively short interaction length, is easy to machine, and is non-magnetic. The active material is a plastic scintillator. In order to provide structural support, stainless steel is used in the innermost and outermost layers. The 3.7 mm thick scintillator plates are sandwiched between the absorber plates. Wavelength-shifting fibers are embedded in the scintillator tiles, which channel the light to photodetectors that can operate in high axial magnetic fields called hybrid photodiodes. In the HB, the segmentation is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087 \approx 5^\circ \times 5^\circ$, which is termed tower. In the HE, the ϕ segmentation is $5^\circ - 10^\circ$ and the η segmentation is 0.087–0.35 depending on η .

The energy from the hadron showers that leak through the calorimeters' rear are samples with additional layers of scintillators are installed outside the coil within the return yoke using the iron as an absorber. This helps in improving the central shower containment and the E_T resolution of the calorimeter. This sampling calorimeter is referred to as hadron outer detector. It is located along with the muon barrel system and closely follows its segmentation. It is divided into five sections termed rings along η . The hadron outer detector follows the HB geometry in η and ϕ and covers the

region $|\eta| < 1.26$. Two further detectors, which cover the region $2.9 < |\eta| < 5.0$, are specialized in measuring energetic forward hadronic showers and ensuring full geometric coverage for the transverse energy measurement. The HCAL forward (HF) sampling calorimeters use steel as absorber material and quartz fibers as the active material. Cherenkov light is emitted by the shower particles in the quartz fibers and is channeled to photomultipliers. The hadron showers' neutral components are preferentially sampled in the HF, leading to narrower and shorter hadronic showers, and is ideally suited for the forward region.

3.3.8 Muon system

The muon system is installed in the magnet return yokes of the CMS detector. The muon system's main task is to identify muons with good efficiency and reconstruct its momentum with good resolution and assign the correct charge even for high- p_T muons. It is also tasked with triggering events with muons. It is divided into a barrel detector covering $|\eta| < 1.2$ and two endcap detectors, 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 ($1.2 < |\eta| < 2.4$). The muon system has three different types of gaseous detectors in different radiation environments. As the muon rate, the neutron-induced background rate, and the residual magnetic field are low in the barrel region, the drift tube (DT) chambers are used. In the two endcaps, the muon rate, the neutron-induced background, and the magnetic field are high, and hence the 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 and are thus used to identify the correct bunch crossing.

Figure 3.5 gives an overview of the muon system’s layout. The barrel is divided into four stations arranged in cylinders interleaved with the iron yoke. It is also divided into five wheels along the beam direction following the five wheels of the return yokes. There are 12 layers divided into 3 Super Layers (SL) made out of four DTs layers in each chamber. The two SL measure the $r\text{--}\phi$ coordinate, while a third SL sandwich between them measures the z coordinate. In the last muon station, there are only two SL to measure the $r\text{--}\phi$ coordinate. There are one or two associated RPCs in each DT chamber. The DTs have a single point resolution of $\approx 200\mu\text{m}$. The CSCs and RPCs are arranged in four disks in each endcap. They are divided into three concentric rings in the innermost station or two concentric rings in the other stations. Up to six space coordinates (r, ϕ, z) are measured in each CSC, 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 ϕ is ~ 10 mrad. The DTs or CSCs and the RPCs provide two independent and complementary information sources.

3.3.9 Trigger and data acquisition system

The online event selection process, trigger, reduces the rate to about 1 kHz. 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 $3.2\ \mu\text{s}$, during which the data is held in pipelined memory buffers. However, the time needed for Level-1 trigger calculations is only $1\ \mu\text{s}$. 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_T or p_T thresholds are used for the decision making. Reduced granularity and resolution data are used to form the

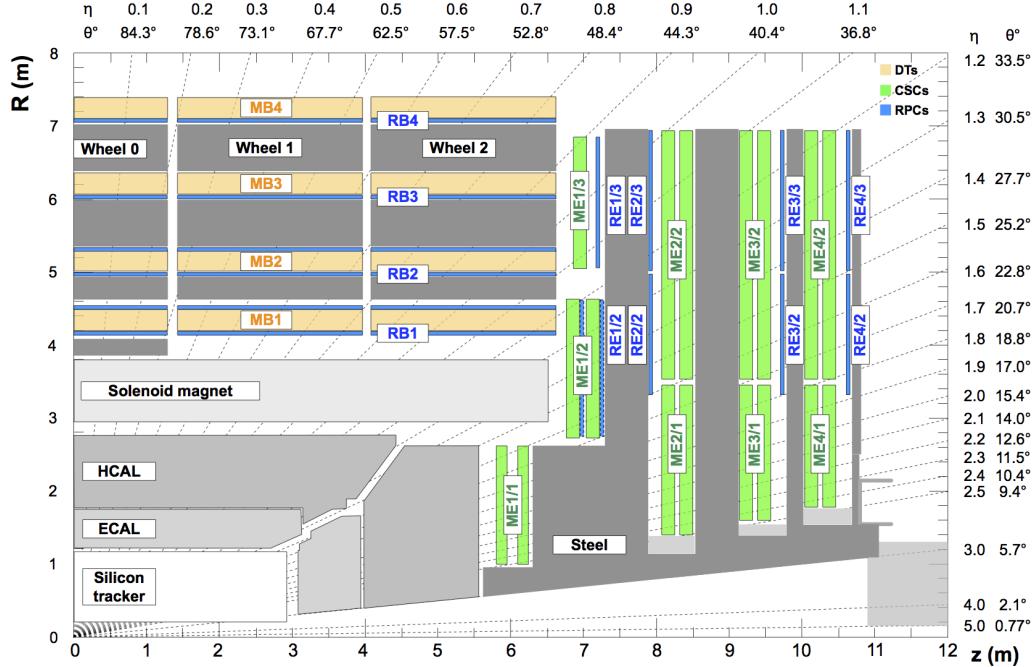


Figure 3.5. Schematic view of the muon detectors.

trigger objects. Sums of E_T and \cancel{E}_T are also employed. If a decision is made to keep the event, the data is transferred to the front-end read-out 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 100 kHz to 1 kHz. 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.

The data acquisition path of the ECAL sub-detector will be described in the following paragraphs. The process starts when the global trigger distributes the Level-1 decision. The data concentrator card is the main ECAL read-out unit and is responsible for collecting the crystal data. It also aggregates the trigger primitive information and performs an extensive integrity check of the data received. About 100 kB per event have been allocated for ECAL data, but the full event payload, if all channels are read out, exceeds this target by a factor of nearly 20. The selective read-out processor can perform the reduction of the data volume. The suppression applied to a channel takes account of energy deposits in the vicinity to achieve the best resolution for large energy deposits [42].

The calorimeter regions are classified into three groups by comparing the transverse energy deposited in each trigger tower with two configurable thresholds. Upon receiving the Level-1 decision, the information is transmitted to the selective read-out processor, determining the read-out mode that the data concentrator card has to apply to each trigger tower. There are two read-out modes: full read-out and zero suppression. With full read-out, the crystal time samples are saved without any suppression. With zero suppression, the signal's amplitude for each crystal is calculated by applying the finite impulse response filter over the digitalized time samples and compared with a configurable threshold: if the amplitude is above the suppression level, the crystal data is saved; otherwise, it is discarded.

I upgraded the selective read-out processor monitor from the legacy monitoring system and the error database to quickly diagnose re-appearing errors for the ECAL. The online monitoring system checks the status of various online electronic components of ECAL. The ECAL experts regularly use the new monitoring system to identify errors so that there is minimal data collection loss [43]. Figure 3.6 shows the online monitoring system in operation during data-taking.



Figure 3.6. The ECAL electronics online monitoring system in operation.

3.3.10 Luminosity measurement

Luminosity is defined as the ratio of the event rate to the cross-section of a given process. It is the effective area quantifying the likelihood of a scattering event. The cross-section is measured in units of area, thus the luminosity is given in units of events per time per area, $(b.s)^{-1} = 10^{24} \text{ cm}^{-2}\text{s}^{-1}$. The integrated luminosity is the integral over the instantaneous luminosity $L = \int \mathcal{L}(t)dt$. The number of expected events for a given process is given by the product of the integrated luminosity and the production cross-section, $N_{\text{exp}} = L \cdot \sigma_{\text{exp}}$.

The performance of LHC is monitored in real-time using luminosity measurements and they provide an overall normalization for physics analyses [44]. The cross-section of a given process can be estimated using a reference process, $\sigma_{\text{exp}} = \frac{N_{\text{exp}}}{N_{\text{ref}}} \cdot \sigma_{\text{ref}}$, where N_{ref} and σ_{ref} are the number of reference events and the reference process's cross-section. The above equation assumes that the same integrated luminosity has been used for both processes. The modern colliders use bunched beams [4]. If two bunches with n_1 and n_2 particles collide head-on with frequency f , then the instantaneous luminosity is given by:

$$\mathcal{L} = f \cdot \frac{n_1 n_2}{\sqrt{\epsilon_x \beta_x^* \epsilon_y \beta_y^*}} \quad (3.5)$$

where x, y are the beams transverse coordinates, $\beta_{x,y}^*$ are the interaction points amplitude functions, where the beam optics produce a narrow focus. Emittance is a measure of the beam width defined as $\epsilon_x \equiv \frac{\sigma_x^2}{\beta_x}$, with the root mean square of the transverse beam sizes in the horizontal or vertical direction being σ_x and σ_y , respectively. The instantaneous luminosity can be increased by colliding bunches with a high population of low emittance at high frequency at locations where the beam optics provide low values of the amplitude functions. As the beam parameters change, the instantaneous luminosity has to be measured correspondingly.

The event rate of a reference process can estimate the instantaneous luminosity if the cross-section is known. The visible cross-section is given by: $\sigma_{\text{vis}} = \sigma(E) \cdot A(t, \mu, \dots)$, where the cross-section depends on the collision energy and the detector acceptance depends on time, the mean number of interactions per bunch crossing, and other parameters. There are independent detectors or parts of the detector for measuring the instantaneous luminosity called luminometers. Once the visible cross-section has been determined for a luminometer, the luminosity can be estimated using $\mathcal{L} = \frac{\dot{N}}{\sigma_{\text{vis}}}$, where \dot{N} is the event rate for this luminometer.

The luminometer's visible cross-section should not be time-dependent nor depend on experimental conditions. Whenever the beam parameters change, the visible cross-section has to be determined again. The same detector configuration that is used during calibration has to be used during data-taking. Thus, there are two important parts of the luminosity measurement: the event rate and the luminometer's calibration. These two parts are described in detail in the following paragraphs.

Luminometer: The monitoring and measurement of luminosity at CMS uses five detectors based on rate measurements [45–47]. The luminometers used are the CMS silicon pixel detector, the DTs in the muon system's barrel, the HF calorimeter, the fast beam conditions monitor (BCM1f), and the pixel luminosity telescope (PLT). The PLT, BCM1f, and HF luminometers have an independent, fast read-out system.

The luminometer of the pixel detector and the DTs use the standard CMS trigger and data acquisition system because they have very low occupancy and very good stability over time.

The HF is most sensitive to the hadronic showers' electromagnetic component and has a high rate of acquisition. There are two methods that have been studied to estimate the luminosity. The first method, which is referred to as zero-counting, counts the hits above the single physical towers' threshold and averages each tower's result. In the second method, the linear relationship between the total transverse energy deposit in the HF and the number of interactions and the luminosity is exploited. These two methods require that the mean value of interactions is proportional to the luminosity.

Many pixels in the CMS detector are employed for the pixel cluster counting (PCC) method. Two different tracks from the same bunch crossing have an exceedingly small probability of hitting a given pixel. The fraction of occupied pixels is less than per mille for $\mu = 25$. Thus, the number of hit pixel clusters is a linear function of the number of interactions per crossing, and the number of hit pixel clusters is a good measure of the instantaneous luminosity, which is given by:

$$\mathcal{L} = \frac{\langle N_{\text{cluster}} \rangle \cdot f}{\sigma_{\text{vis}}^{\text{PCC}}}, \quad (3.6)$$

where $\langle N_{\text{cluster}} \rangle$ is the mean number of hit pixel clusters, f is the LHC orbit frequency, and $\sigma_{\text{vis}}^{\text{PCC}}$ the visible cross-section of the PCC method.

In the offline luminosity measurement, the pixel modules that have not been fully operational during calibration are omitted. The precision offline luminosity measurement is based on the PCC method because it has a very small dependence on pileup and other experimental conditions. The smaller statistical uncertainty of the HF measurements makes it useful for cross-checks or studies of the luminosity measurements' systematic uncertainties.

Van der Meer (VdM) scans: The luminosity per colliding bunch pair can be measured from machine parameters using VdM scans. The counting rate proportional to the rate of the beam-beam interaction is measured by the counter system. The luminometer rate is measured as a function of the beam-beam separation after one of the two beams is displaced, resulting in a maximum at zero separation. It is used under the assumption that the two bunch densities factorize in x and y. A dedicated LHC machine set up is required for performing these scans. The two beams are scanned through one another in the transverse plane of the detector. The luminosity rate and the luminosity must be determined at the same time to measure the visible cross-section:

$$\sigma_{\text{vis}} = 2\pi \Sigma_x \Sigma_y \langle n \rangle_0 \quad (3.7)$$

where Σ_x and Σ_y are the effective beam widths and $\langle n \rangle_0 = \frac{1}{2} (R_x + R_y)$ with normalization rates R_x and R_y , which are the fitted scan curves amplitudes. The beam shapes are studied using the beam imaging scans. The transverse and longitudinal interaction point centroids are also determined using the VdM scan.

Luminosity integration: The instantaneous luminosity can be measured after the VdM scan using the luminometer's visible cross-section. The integrated luminosity is obtained by summing the luminosities of the luminosity section (LS). The LS is a convenient minimal time interval to consider for the estimation of luminosity corresponding to t_{LS} (~ 23 s). The average number of clusters per event is computed for each LS, and the luminosity for the LS is derived and multiplied by t_{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%.

3.3.11 CMS Phase-2 upgrade

During the LHC Phase-2, the High Luminosity LHC (HL-LHC) program's main objective is to deliver to the experiments a much larger dataset than during the Phase-1 for new physics searches, Higgs boson coupling measurements, and precision tests of the standard model. After 2026, the aim is to operate the upgraded accelerator at an instantaneous luminosity as high as $7.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. This corresponds to a factor of 3 increase with respect to the Phase-1 peak luminosity. This will allow the delivery of up to 4500 fb^{-1} over about 12 years of operations. A major experimental requirement for the upgrades will be to distinguish the primary interactions among the hundreds of softer collisions that will pileup (~ 200) in each beam crossing. Therefore, the new systems will need high resolution to separate the trajectories and energy deposits of particles produced in these different collisions and then associate them with their correct origin. In Figure 3.7 we can see the schedule for HL-LHC operation.

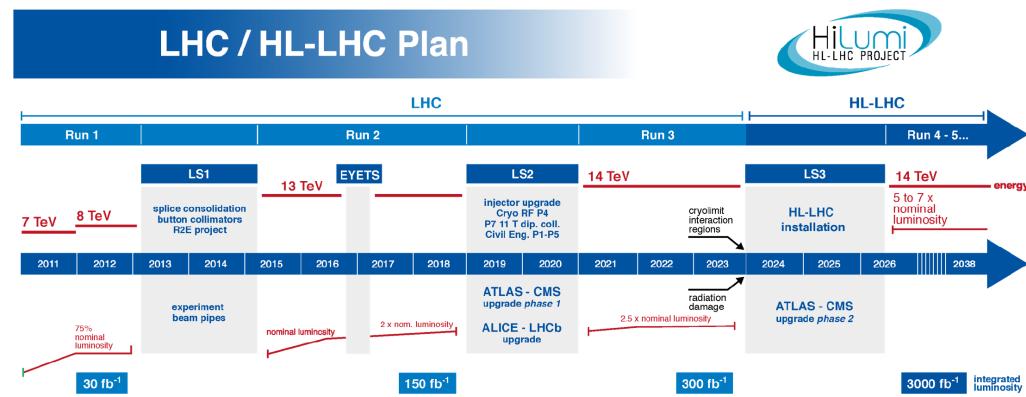


Figure 3.7. Schedule of LHC and HL-LHC operation.

In order to improve the event reconstruction and enhance its resilience against pileup, the CMS experiment foresees the use of per-particle time information at the level of 30 ps precision. Combining the time information for electrons and photons with that of charged hadrons widens the physics reach for precise measurements in the Higgs sector and searches of new physics beyond the standard model from simulation studies [48]. As a part of the CMS Phase-2 upgrade, several projects have been proposed, including a new CMS silicon tracker, a new endcap calorimeter, and a minimum ionizing particle timing detector.

The ECAL subdetector needs to be upgraded to maintain its current performance up to the luminosity goals of the LHC Phase-2. As a part of these upgrades, the front-end and off-detector electronics of the EB need to be upgraded along with the complete replacement of the EE [49]. The primary technical motivation for the upgrade of EB is the Phase-2 trigger requirements to have a latency of 12.5 μs , compared to the current 4 μs , and a Level-1 trigger rate of about 750 kHz, compared to the current 100 kHz. In order to meet these requirements, the EB front-end cards and off-detector electronics need to be replaced.

The effect of radiation on lead tungstate crystals has been measured, and it is observed that the scintillation process remains unchanged. However, the formation of scattering color centers affects crystal transparency. The main concern for operating the calorimeter at the HL-LHC is the permanent loss of light transmission due to hadron irradiation. The studies to parametrize and predict the evolution of the crystals under the aging conditions expected at the HL-LHC have been performed at test beam campaigns over the past years [?]. By the end of the LHC Phase-2, the energy resolution's constant term is expected to increase to approximately 1.5% at large η in the EB. Thus, the CMS Phase-2 upgrade will retain the EB crystals and the APDs because the constraints measured are compatible with the physics requirements.

The APDs dark current will increase significantly due to the bulk damage of the silicon caused by hadron irradiation. In order to circumvent this and reduce the APD dark current by a factor of 2, the EB operating temperature will be lowered from the nominal Phase-1 18°C to 9°C. With this lowered temperature, it is expected that by the end of the LHC Phase-2, the dark current will reach about 50 μ A (250 MeV energy-equivalent noise) for APDs at $|\eta| = 0$ and about 100 μ A at $|\eta| = 1.45$. Figure 3.8 shows the evolution of the dark current for the two operating temperatures, along with a combined scenario where a further lowering of the temperature to 6°C is actuated after about 1600 fb^{-1} . The cooling system currently in place will provide sufficient power to stabilize the temperature at 9°C. Only minor interventions to the distribution pipes are required to provide chilled water from the surface to the underground cavern and are currently being undertaken during the long LHC shutdown. A more complex modification to the coolant supply system is required for a further temperature reduction step and is still under evaluation.

The CMS Phase-2 upgrade will upgrade the very front-end and front-end cards. The upgrade will maintain the mechanical structure of the cooling system and the motherboards. The motherboards are placed below the cooling structure and distribute the low voltage to the electronics and the bias voltage to the photodetectors. They also channel the APD signals to the very front-end cards. A redesigning has been done for the very front-end card to keep low noise and have a shorter pulse shaping and higher sampling rate. A trans-impedance amplifier (TIA) is selected to accomplish this. The TIA generates a voltage image of the photocurrent in the output of the APD, and only the total bandwidth of the system limits the shaping time. The dynamic range of signals up to 2 TeV is maintained using two gains different by a factor of 10. A lossless data compression algorithm is used to reduce the needed bandwidth for transmitting the TIA signals digitized at 160 MHz by 12-bit ADCs to the front-end cards.

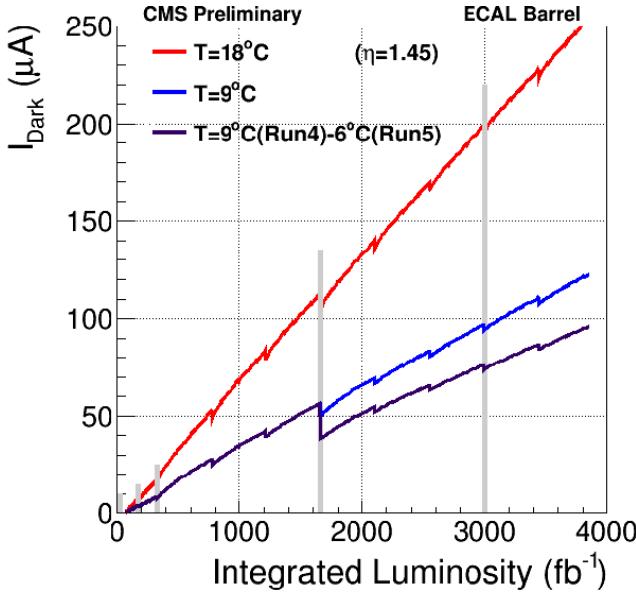


Figure 3.8. APD dark current shown for APDs at $|\eta| = 1.45$ operated at 18°C (red curve), at 9°C (blue curve), and in a mixed scenario including a further step to 6°C after the LHC Long Shutdown 4.

The front-end card will use new technologies like Low Power Gigabit Transceiver and Versatile Link plus to send single crystal data sampled at 160 MHz (compared to the current 40 MHz) to the back-end electronics system. This removes the current requirement of having on-detector data selection and buffering. The data transmitted to the back-end electronics will be processed by the trigger primitive generation algorithm with a crystal level granularity (compared to the current trigger tower level granularity) for usage at the Level-1 trigger. This will increase by a factor of 25 the Level-1 trigger granularity. This also has added benefits in the topological rejection of signals due to direct ionization in the APDs (spikes). Powerful commercially available FPGAs and high-speed optical links will be used in the off-detector electronics. As the electronics are being upgraded, extensive algorithm development is ongoing to test novel algorithms for signal processing, trigger, and data acquisition.

Firmware Development: I took part in the CMS Phase-2 upgrade algorithm

development and implemented the trigger primitive generation algorithm for ECAL’s proposed Level-1 trigger. ECAL’s Level-1 trigger proposal is to perform the trigger primitive generation algorithm in the off-detector electronics and use the crystal-level instead of the tower-level information to calculate the trigger primitives. There are three main stages in which the raw data is processed: linearizer, amplitude filter, and peak finder. During Phase-1 trigger primitive generation, the amplified and digitized signal from 5 crystals of a strip is first linearized. The linearization takes into account the gain and corrects the digitized signal by the calibration coefficients. The amplitude of the strip pulse is measured using the amplitude filter. The filter is based on linear weighted sums, and the weights take into account the expected shape of the signal and subtract the residual pedestal dynamically. This output is then filtered by a peak finder that keeps only the maximum as a measure of the transverse energy contained in the strip.

For Phase-2 trigger primitive generation, the three steps involved are kept the same, except for implementing the algorithm on crystal level digitized signal instead of implementing it on the digitized signal from 5 crystals of a strip. The input to the linearization step is the digitized signal along with the linearizer coefficient. The linearizer coefficient is 24 bits in length. The first 12 bits are the pedestal, subtracted from the digitized signal to get the corrected signal. The next 8 bits are the multiplicative factor applied to the corrected signal to calibrate it. The remaining 4 bits is the bit shift applied to the calibrated signal to get the linearized signal. The amplitude of the linearized signal is estimated using the amplitude filter, which is implemented using a shift register. A shift register is a type of digital circuit using a cascade of flip flops where the output of one flip-flop is connected to the next’s input. A 4 stage shift register is used for implementing the amplitude filter, which takes as input the linearized signal along with five weights. The amplitude filter’s output is given as input to the peak finder, which is implemented using a 2 stage shift register

that finds the peak by comparing the current amplitude output to the previous and next output.

The main challenge for implementing the algorithm lies in the latency constraints. High-Level Synthesis is a tool that converts programs written in high-level programming languages such as C++ into hardware description languages such as VHDL. The algorithm has been implemented using High-Level Synthesis, and the synthesized code is tested on custom hardware (Virtex-7 field-programmable gate array (FPGA)). A clock cycle of ~ 8.33 ns or 120 MHz is used for testing the initial implementation of the algorithm. The initiation interval corresponds to the time interval after which the algorithm can read the next input. As the collisions are happening every 25 ns, the initiation interval is required to be 25 ns as well. From the synthesis report in Figure 3.9, the initiation interval obtained after the algorithm has been synthesized is three clock cycles (~ 25 ns). Latency corresponds to the time interval from processing the digitized signal to obtaining the peak finder's output. The latency obtained after the algorithm has been synthesized is six clock cycles (~ 50 ns). The overall resource utilization in the hardware is minimal. In Figure 3.9, BRAM corresponds to block random access memory, which is not used, DSP corresponds to digital signal processing, which is used for the multiplication steps by the algorithm, FF corresponds to flip-flops, and LUT corresponds to look-up tables which are used for storage. The DSP usage is 9%, FF usage is 14%, while LUT usage is 19%. This suggests that other algorithms, like multi-fit, spike rejection, etc., can easily be implemented on the hardware as many resources are remaining.

The synthesized VHDL code is converted into a bit file. The bit file contains all of the information necessary to properly configure the device created by our design's implementation tools. The bit file is loaded onto the Virtex-7 FPGA and has been successfully tested. The hardware's output matches the simulation output, which suggests that the algorithm is working as expected. The algorithm implementation

has shown the feasibility of using High-Level Synthesis to do the CMS Phase-2 upgrade's firmware development. High-Level Synthesis allows physicists like me to get involved in firmware development using object-oriented languages like C++ without knowing the nitty-gritty details of using hardware description languages like VHDL, which remains in the realm of engineers.

Performance Estimates**Timing (ns)****Summary**

Clock	Target	Estimated	Uncertainty
ap_clk	8.33	7.29	1.04

Latency (clock cycles)**Summary**

Latency	Interval			
min	max	min	max	Type
6	6	3	3	function

Detail**Instance****Loop****Utilization Estimates****Summary**

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	-	-	-
Expression	-	-	0	1504
FIFO	-	-	-	-
Instance	-	352	64592	62480
Memory	16	-	0	0
Multiplexer	-	-	-	21987
Register	-	-	63430	-
Total	16	352	128022	85971
Available	2940	3600	866400	433200
Utilization (%)	~0	9	14	19

Figure 3.9. Performance estimates from High-Level Synthesis.

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 and the decay processes in an event can be described by perturbative QFT. However, perturbative QCD (pQCD) cannot describe the 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. 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 behavior in simulation as in collision data. Event signature beyond SM particles can be generated to compare its signature to one of the generated background processes.

General-purpose Monte Carlo (GPMC) generators, like PYTHIA [50], provide fully exclusive simulations of high energy collisions. However, there are also event generators that are specialized in 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 the initial state's composition and substructure. An overview of different steps in MC generation for pp collision events can be seen in Figure 4.1.

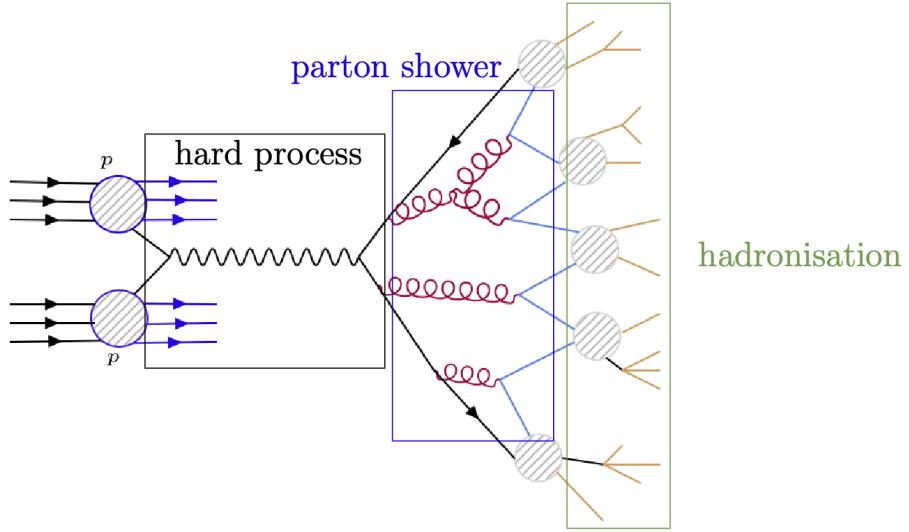


Figure 4.1. MC simulation of an event in $p\bar{p}$ collisions.

4.2 Monte Carlo simulation

The primary hard interaction process and the decay of short-lived particles happen at short distance scales. The QCD and QED radiation at a time scale much below $\frac{1}{\Lambda}$, where Λ 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 factorized into parton distribution functions (PDFs). Therefore, the cross-section remains accurate up to higher-order corrections if interpreted as an inclusive cross-section. If this is not the case, then the QCD singularities can lead to a non-convergence of the fixed order expansion.

Matrix element generator: Matrix element generators generate the exact ma-

trix 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 the 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.

Parton shower algorithm: The parton shower algorithm is used for computing the cross-section for a generic hard process. 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 standard [51]. The kinematics of the basic process is first generated, followed by a sequence of independent shower splittings. The cross-section for the given final state is calculated by assigning a probability to each splitting vertex. Collinear emissions and soft gluon emissions 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.

Matching: QCD color confinement restricts quarks and gluons from existing as isolated particles. The hadronization 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 evolution. A matching has to be done if different generators have been used to generate matrix elements and parton showers or extra partons generated by the hard process generator.

Hadronization models: The hadronization scale Q_{had} is by construction equal

to the infrared cut-off where the parton shower ends. Colored partons are transformed into a set of colorless hadrons by GPMCs. This happens at scales with low momentum transfers and at long distances, where non-perturbative effects become important. GPMCs use models that rely on the color flow information between partons as a starting point for hadronization.

Soft hadron-hadron physics modeling : Underlying event is the additional activity beyond the basic process and its associated initial state and final state radiation. The dominant part is coming from additional color 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. Most MPI are soft, and they influence the color flow and the event's total scattered energy. 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.

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 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 constrains them and is referred to as generator tuning. MC generators are not tuned beyond the constraints in theoretical and experimental precision to avoid overfitting. The final state of the particles and their spectra are influenced by event modeling and generator tuning. Events generated with different generators or tunes can differ and might not describe the collision data in the entire phase space.

4.3 Monte Carlo generators

PYTHIA has been developed for multi-particle production in pp collisions and simulation of jets. PYTHIA can generate hard subprocess, initial and final state parton showers, hadronization, decays, and the underlying event. Many hard processes have been implemented for generating the matrix elements for final state and phase space calculation. PYTHIA can optimally generate $2 \rightarrow 1$ and $2 \rightarrow 2$ processes. Resonance decays with the resonance masses above the b quark system are implemented. Their branching fractions and partial width can be dynamically calculated as a function of their mass. If the spin information is available for resonance decays, it leads to angular correlations of the resonance decay products; otherwise, the resonance decays isotropically. GPMC generators like PYTHIA can simulate the full process. However, there are specialized generators that deal with a certain aspect of the event simulation.

MadGraph: MadGraph generates the matrix element with leading order (LO) accuracy [52]. MadGraph 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 that are renormalizable or from an effective field theory written in a Lagrangian can be generated. The full amplitude is split into gauge invariant sub-amplitudes. The matrix element contains the full spin correlation and Breit-Wigner effects but is not valid far from the mass peak.

POWHEG: POWHEG is a framework for implementing NLO matrix element calculations [53]. It includes NLO virtual corrections and radiation of an extra parton in the matrix element. It needs the LO matrix elements and the finite part of the virtual corrections as input from which it finds all the singular regions. The singular regions are characterized by a 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.

aMC@NLO: aMC@NLO implements all aspects of NLO computation and its matching with parton showers [54, 55]. NLO calculations can be achieved by combining one-loop matrix elements and tree-level matrix elements. Tree level computations are performed using MadGraph, and one-loop amplitudes are evaluated with MadLoop [56]. The matched samples which differ by their final state multiplicity can be merged using the FxFx merging scheme.

MLM matching: MLM matching scheme is a matching algorithm [57, 58] 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. The jet closest in (η, ϕ) to the hardest parton 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: FxFx merging scheme is an NLO merging procedure [59]. 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 the real emission contribution, infrared subtraction terms, and the one-loop contributions to processes with $(J - 1)$ partons. Events are reweighted, and a certain amount of events might carry negative weights.

4.4 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 analyzed. Geant4 [60] is used for detector simulation. It is a toolkit for simulating particles' passage 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 shapes 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 digitization. The detector's geometrical structure is divided into logical and physical volumes. Logical volumes contain the information of the material and the sensitive detector behavior. 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 ionization described by energy loss and range tables, bremsstrahlung, pair production of electron-positrons from photons, photoelectric 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, and 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 digitizations generated by simulation are added. Trajectories of simulated particles can be added optionally for the recording of “simulation truth.”

4.5 Monte Carlo samples

MC simulated event samples are used to model signal and background contributions to all the analysis regions with several event generators. In all cases, parton showering, hadronization, and underlying event properties are modeled using PYTHIA version 8.212. The PYTHIA parameters affecting the description of the underlying event are set to the CUETP8M1 tune in 2016 [61], except for the $t\bar{t}$ sample where the CP5 tune is used, which is also the tune in 2017 and 2018 [62]. The NNPDF3.0 PDF set is used for all 2016 samples, and the NNPDF3.1 PDF set is used for the 2017 and 2018 samples [63].

Simulation of interactions between particles and the CMS detector is based on Geant4. The same reconstruction algorithms used for data are applied to simulated samples as well. The Higgs bosons are produced in pp collisions predominantly by gluon gluon fusion (ggF) [64], but also by vector boson fusion (VBF) [65], and in association with a vector boson (W/Z) [66]. The ggF, VBF, and associated production Higgs boson samples are generated with POWHEG generator in the implementation described in Ref. [67, 68]. We only consider Higgs boson produced in ggF and VBF production mechanisms and do not use associated production Higgs boson samples for the signal.

Embedded samples are data samples with well-identified $Z \rightarrow \mu\mu$ events from which muons are removed, and simulated tau leptons are embedded with the same kinematics as the replaced muons. These samples are employed for the data-driven estimation of the $Z \rightarrow \tau\tau$ and some $t\bar{t}$ /Diboson/Single Top background. The MadGraph generator is used to simulate the $Z \rightarrow ee/\mu\mu + \text{jets}$ process along with the $W + \text{jets}$ background process. They are simulated at LO with MLM jet matching and merging schemes [69].

Diboson production is simulated at NLO using aMC@NLO generator with FxFx jet matching and merging scheme. Top quark pair and single top quark production

simulated samples are generated using POWHEG. Due to the high instantaneous luminosities attained during data taking, events have multiple pp interactions per bunch crossing. The effect is taken into account in simulated samples by generating concurrent minimum bias events. All simulated samples are weighted to match the pileup distribution observed in the data.

CHAPTER 5

EVENT RECONSTRUCTION

The particle flow (PF) event reconstruction aims to reconstruct and identify each particle in an event with an optimized combination of information from the various CMS detector elements [70]. In this process, identifying 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 η - ϕ 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 electrons are defined as PF electrons if the tracker's extrapolated tracks have a corresponding energy deposit in the ECAL. The sequence now proceeds with identifying photons and hadrons. In the sequence, tracks with momentum uncertainty more than the calorimeters' resolution 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.

5.1 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 [71]. 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 tracks' global coordinate system during track reconstruction. Kalman filter (KF) [72] based algorithm is used to reconstruct tracks and is called the combinatorial track finder. Tracks with high p_T and those 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, 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.

KF 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, 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 candidate. The added hits' information is combined with the original track candidates' extrapolated trajectory to update the new candidates' trajectory parameters. Figure 5.1 illustrates the reconstruction efficiency of tracks in the case of isolated muons.

The hits from the last step are refitted using a KF 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 (PV), and how good a fit it yields.

The interaction vertices in the pp 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 closest approach to the beam spot center. 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 [73]. The PV has the most significant sum of squared transverse momenta of tracks originating from it.

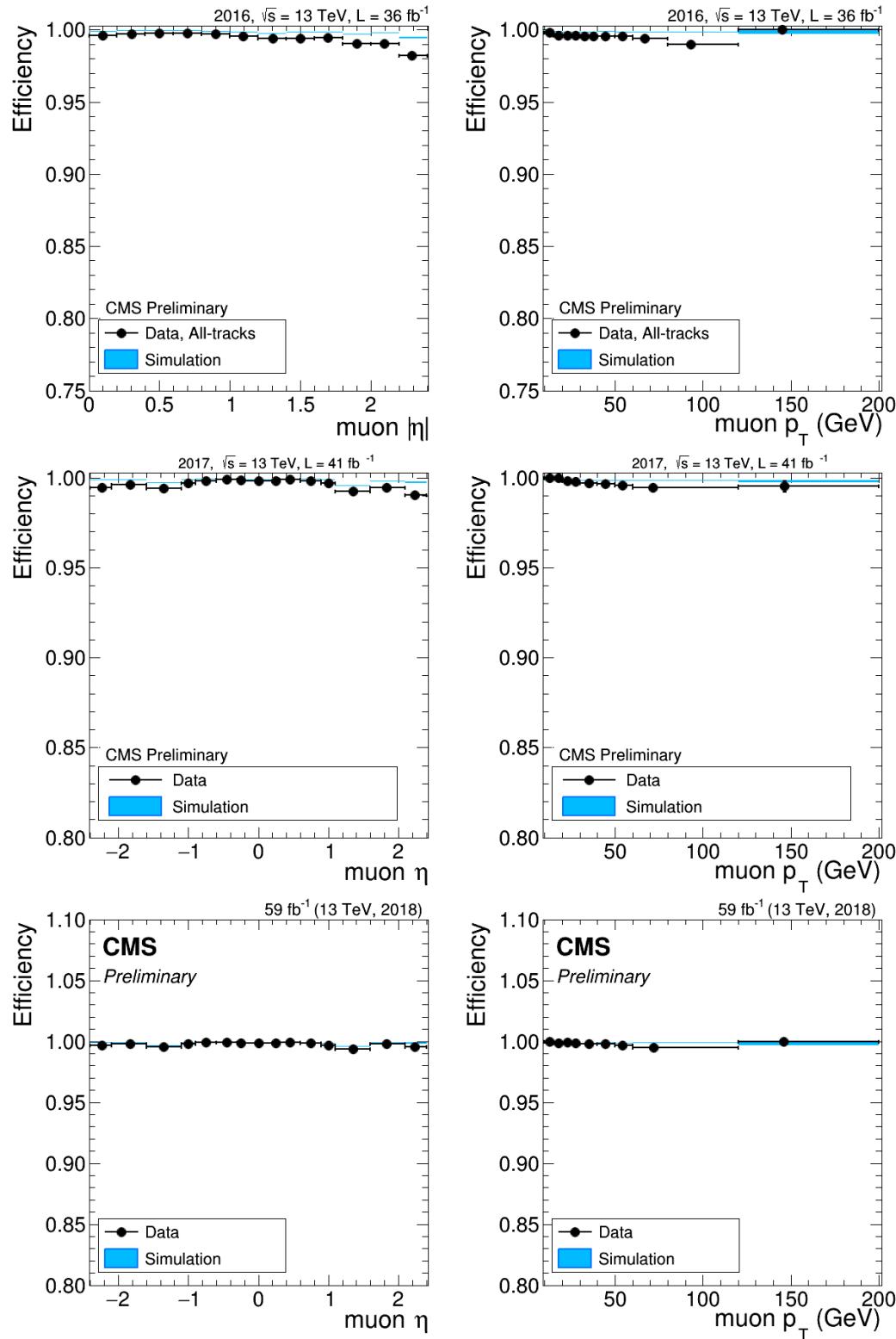


Figure 5.1: The tracking efficiency for “all-tracks” collection as a function of pseudorapidity and transverse momentum for muons coming from the Z decay.

5.2 Muon reconstruction

Muons are reconstructed using the hits in the muon system, and tracks from the tracker [74]. 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. KF is used for reconstructing the hits from the muon system. These tracks are called standalone-muon tracks. Tracker tracks with $p_T > 0.5 \text{ GeV}$ 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 KF 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.
- χ^2/ndof of the global muon track fit < 10 .
- At least one muon chamber hit included in the global muon track fit.
- There should be 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 $|d_{xy}| < 2$ mm with respect to the PV.
- The longitudinal distance of the tracker track with respect to the PV is $|dz| < 5$ 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 η and p_T to adjust the simulation efficiency to match the efficiency in data.

5.3 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 these bremsstrahlung photon showers' energy to measure the electron's energy. The bremsstrahlung photons' energy spreads primarily in the ϕ direction, and the spread in the η 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 the ϕ direction but are narrow in the η direction. A seed crystal contains the most significant energy deposited in the considered region above a 1 GeV threshold. A 5×1 array of crystals are added in $\eta \times \phi$ around the seed crystals in both directions of ϕ if the energy contained in the array is above the 0.1 GeV threshold. Contiguous arrays are merged into clusters. An electron supercluster is formed from all such strip

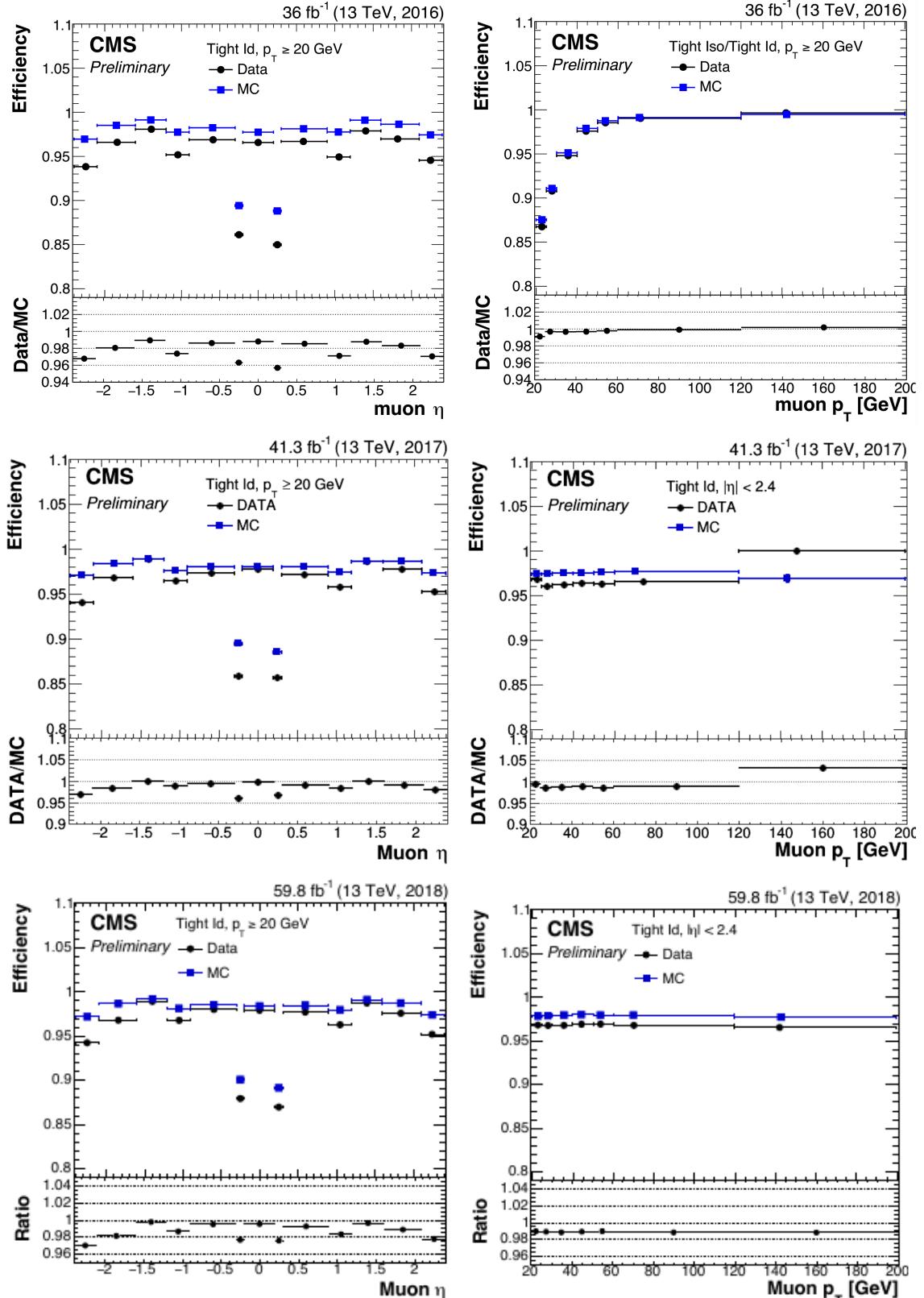


Figure 5.2: Efficiency of muon identification as a function of η and p_T , for data (black) and simulation (blue) [75].

clusters, which have at least one seed strip with energy above the 0.35 GeV threshold. A different clustering algorithm is used in the ECAL endcap due to the crystals' different geometrical arrangements. This algorithm is called the 5×5 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 5×5 crystals around the seed crystal. The added clusters need to have energy over 1 GeV and be within ± 0.7 and ± 0.3 respectively in η and ϕ 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 ϕ 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 from the tracker and the ECAL. The first step in electron track reconstruction is seeding. The reconstructed superclusters' position and energy can 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 standard 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. The 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 KF 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 performs 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 η - ϕ 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 electron's charge is estimated using the GSF track curvature, ECAL supercluster's relative position in ϕ to that of the first hit in the GSF track, and 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 η and ϕ 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 1×5 and a 5×5 grid around the supercluster seed, respectively.
- Shape variable $R9 = \frac{E_{3\times 3}}{E_{SC}}$, with $E_{3\times 3}$ the energy in a 3×3 grid of cells around the supercluster seed and E_{SC} the raw energy of the supercluster.
- The number of valid hits in the track fit, the χ^2 of the track fit, and the χ^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 η 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.

The BDT was trained on a Z/γ MC sample generated with MadGraph5, in 3 η bins for electrons with $p_T > 10 \text{ GeV}$. There are two versions of the MVA available for electron identification. The first one includes the electron isolation in training and another which does not include this, and their performance is illustrated in Figure 5.3. This analysis uses the version without the isolation included in the training. An additional selection of electron isolation is used. This is done so that in the $e\tau_h$ channel, the misidentified lepton background can be estimated using the isolation based sideband regions. Instead of using fixed cuts on the MVA score, the cuts are alternatively varied exponentially with p_T to achieve a more constant efficiency of 80%. 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 $|dxy| < 0.045 \text{ cm}$ and $|dz| < 0.2 \text{ cm}$ to ensure the electron is associated with the PV.

5.4 Tau reconstruction

Hadrons plus strips (HPS) algorithm is used for reconstructing the hadronic decays of tau [76]. “DeepTau” based identification is used for discrimination of hadronic decays of tau against jets, electrons, and muons. A detailed description of the HPS algorithm, followed by the DeepTau Identification algorithm, is given in the following subsections.

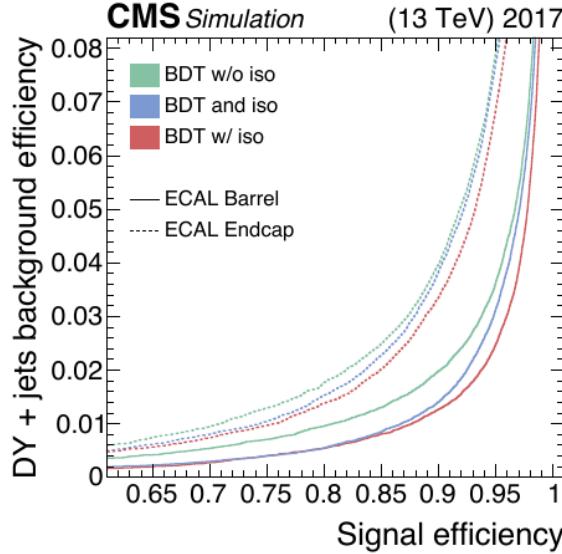


Figure 5.3: Receiver operating curve showing the background versus signal efficiency for the BDT with and without isolation incorporated in training.

5.4.1 Hadrons Plus Strips

The HPS 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 π^0 candidates leave in the ECAL, photon and electron constituents of the jet that seeds the hadronic tau (τ_h) reconstruction are clustered into strips. The e or γ (not yet included in a strip) with the highest p_T is used to build a new strip. The η and ϕ of this candidate determine the initial position of the strip. The next highest p_T e or γ 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

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

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

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 π^0 candidate. Hadronic taus are reconstructed by combining charged particles and strips into different signatures compatible with a specific decay mode 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 considered for reconstructing taus are:

- One charged particle, no strips.
- 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 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 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.

5.4.2 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 τ_h candidates with low-level information from the inner tracker, calorimeters, and muon sub-detectors using PF candidates reconstructed within the τ_h signal and isolation cones. DeepTau also takes advantage of using the updated decay mode definitions.

A tau lepton decaying into an electron is referred using τ_e , while if it decays into a muon it is referred using τ_μ , and a jet that is misidentified as a tau lepton is referred with τ_j . A balanced mix of τ_e , τ_μ , τ_h , and τ_j candidates coming from $t\bar{t}$, $W + \text{jets}$, and $Z + \text{jets}$ MC simulation is used to perform the training. The τ_h has a loose preselection: $p_T \in [20, 1000] \text{ GeV}$, $|\eta| < 2.3$, 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 (ρ). For each candidate reconstructed within the tau signal or isolation cones, the 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 11×11 (21×21) cells with a cell size of 0.02×0.02 (0.05×0.05) for the signal (isolation) cone. In cases where more than one object of the given type belongs to the same cell, only the object with the highest p_T is considered input. Within each cell, the input variables are split into three blocks: e-gamma, muon, hadrons.

The low-level inputs' organization 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 reconstructed tau candidate's probabilities to be τ_e , τ_μ , τ_j ,

or a genuine τ_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.

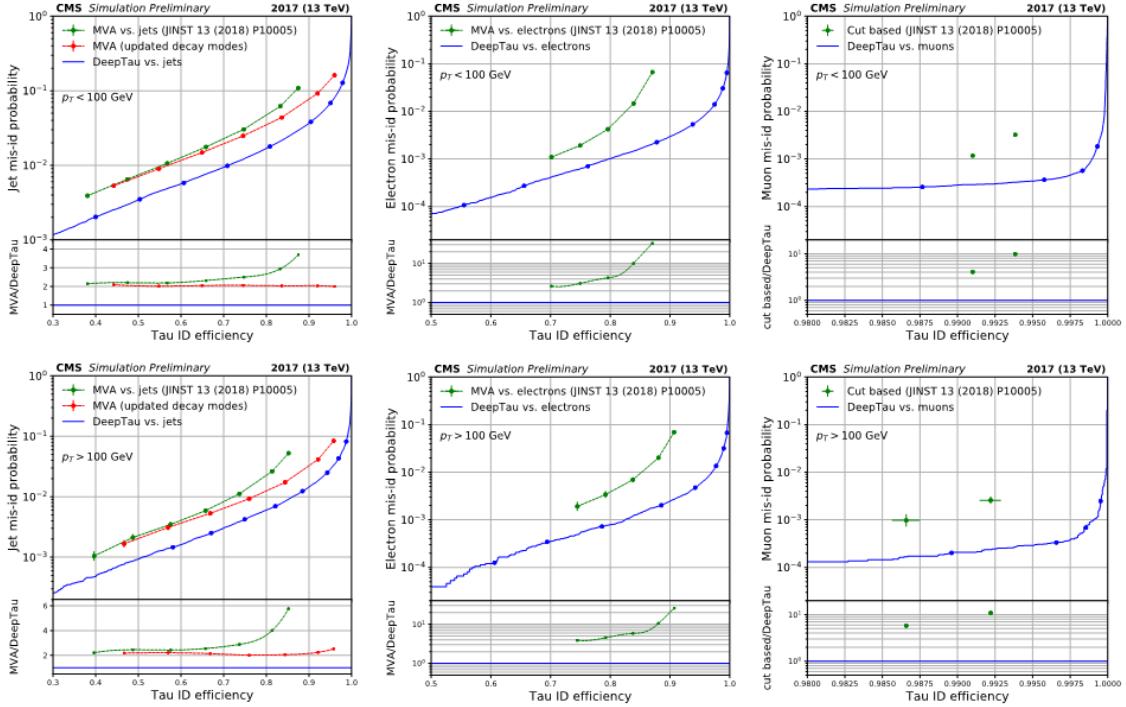


Figure 5.4. Performance of tau discrimination against quark and gluon induced jets (left), electrons(middle), and muons(right) for DeepTau and the previously available discriminators.

A WP defined using a p_T dependent threshold on the output of the DNN is used to distinguish τ_h from jets. This WP has a τ_h identification efficiency of about 70% with a misidentification probability of $\approx 1\%$. This helps to provide a good tau identification efficiency along with high background rejection. The DNN can reject electrons and muons misidentified as τ_h candidates using dedicated criteria based

on the consistency between the tracker's measurements, the calorimeters, and the muon detectors. In the $\mu\tau_h$ ($e\tau_h$) channel, the WP that is used has an efficiency of about 97.5% (87.5%) with a misidentification probability of $\approx 1 - 2\%$ ($0.2 - 0.3\%$) to discriminate τ_h against electrons, and has an efficiency of about 99.6% (99.8%) with a misidentification probability of $\approx 0.04\%$ (0.06%) to discriminate τ_h against muons. DeepTau identification for hadronic taus helps in reducing the Drell-Yan and misidentified lepton background significantly, and the comparison with the previous analysis can be seen in Figure 5.5.

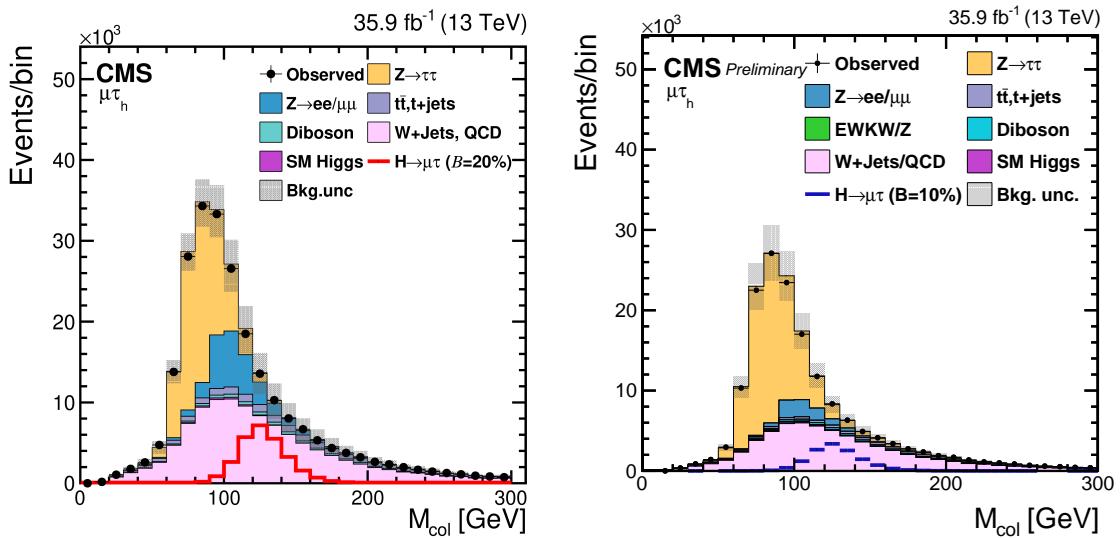


Figure 5.5. Comparison of the collinear mass distribution from the previous analysis (left) compared to the current analysis (right). As can be seen from the plots, the Drell-Yan background and misidentified lepton background are significantly reduced in the current analysis. This can be attributed to using the DeepTau identification for hadronic taus.

5.5 Jet reconstruction

Quarks and gluons hadronize due to color confinement producing a fine spray of particles called jets [77]. Jets are reconstructed using the anti- k_T clustering algorithm [78]. 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:

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

$$d_{iB} = k_{ti}^{2p}$$

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. The d_{ij} represents the distance between 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 which are 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 ensure a uniform response in η [79, 80]. 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 η , and the number of pileup vertices (N_{PV}) averaged over ϕ . The correction is then given by $1 - \frac{\langle \text{Offset}(N_{PV}, \eta) \rangle}{p_T^{\text{RAW}}}$, 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 (ρ) for an event is expected to be insensitive to hard particles. This ρA_i is a good approximation of pileup contribution to the i -th jet. However, this approach has a drawback because it doesn't consider the fact that the detector response is η dependent. Thus, the hybrid jet area method combines these two methods to calculate a jet-by-jet correction depending on η and N_{PV} .

Jet momentum is determined as the vectorial sum of all particle momenta in the jet. It is found from simulation to be, on average, within 5 to 10% of the true momentum over the whole p_T spectrum within detector acceptance [81]. Data collected in the ECAL endcaps were affected by noise during the 2017 run. This effect is mitigated by discarding events with jets having $p_T < 50 \text{ GeV}$ and $2.65 < |\eta| < 3.139$ in 2017 dataset. Hadronic jets that contain b quarks are tagged using a DNN called

the DeepCSV algorithm, and we use a WP with efficiency $\approx 70\%$ [82].

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 η and p_T . A QCD dijet sample is used to correct the dependence on η . Using jets that are approximately back-to-back in the azimuthal direction but at different η regions of the detector, the difference in response between these two η regions can be ascertained and corrected. Using the same method of measuring residual response in the transverse direction in $\gamma + \text{jets}$ or $Z + \text{jets}$ events, the absolute jet energy scale as a function of p_T can be made uniform [80]. Additional selection criteria are applied to each jet to remove jets potentially dominated by instrumental effects or reconstruction failures. When combining information from the entire detector, the jet energy resolution amounts typically to 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV. Any jet within $\Delta R = 0.5$ of the identified leptons is removed.

5.6 Missing transverse energy

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^{miss} is computed as the negative vector sum of all the PF candidates' transverse momenta in an event $\vec{p}_T^{\text{miss}} = -\sum \vec{p}_T$, and its magnitude is denoted as p_T^{miss} [83]. The \vec{p}_T^{miss} is modified to account for corrections to the reconstructed jets' energy scale in the event. Anomalous high- p_T^{miss} events can be due to various reconstruction failures, detector malfunctions, or non-collision backgrounds. Such events are rejected by event filters designed to identify more than 85–90% of the spurious high- p_T^{miss} events with a mistagging rate less than 0.1% [83]. In addition to the event filtering algorithms, the jet identification selection imposed, which requires the neutral hadron energy fraction

of a jet to be less than 0.9, rejects more than 99% of the noisy jets, independent of jet p_T , with a negligible mistag rate.

Corrections to the \vec{p}_T^{miss} are applied to reduce the mismodeling of the simulated Z, W, and Higgs boson samples. The corrections are applied to the simulated events based on 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^{miss} obtained from the simulation by a few GeV. The \vec{p}_T^{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^{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 these objects' precise reconstruction and calibration.

5.7 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 $p_T^{\text{PV charged}}$, p_T^{neutral} , and p_T^γ 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 [84, 85] for the electron or from the sum of transverse momenta of charged hadrons not originating from the PV 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 muon or electron along with an oppositely charged isolated tau lepton (τ_μ , τ_e , or τ_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.

This analysis's final states are similar to the $H \rightarrow \tau\tau$ decay allowed by the SM and since been observed [86]. However, there are some significant kinematic differences. The LFV $\mu\tau_h$ and $\mu\tau_e$ ($e\tau_h$ and $e\tau_\mu$) decays consist of a muon (an electron) that comes directly from the Higgs boson 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 electric charge, and missing transverse momentum from the tau decay. Also, there are fewer neutrinos in LFV decays, coming from the decay of a single tau. The decay products of this highly boosted tau are closely aligned, leading to a narrow separation between the tau's visible decay products and the \vec{p}_T^{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 ($e\tau_\mu$, $e\tau_h$, $\mu\tau_e$, $\mu\tau_h$), a set of loose selection (preselection) for the respective signature is first defined. The jets in the event are required to have a

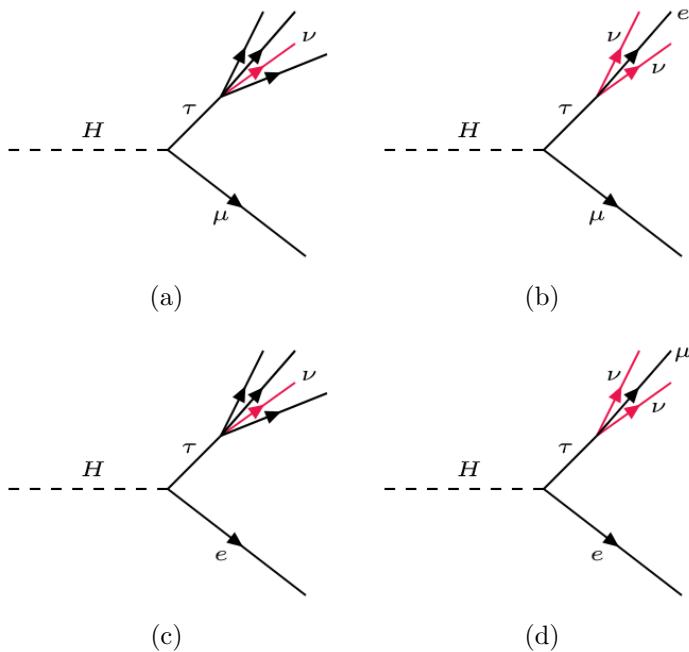


Figure 6.1. Feynman diagrams of LFV Higgs boson decays. The first row shows diagrams for the Higgs boson coupling to $\mu\tau$ (a,b). Couplings to $e\tau$ (c,d) are shown in the second row. Taus can decay leptוניתly or hadronically. Feynman diagrams are shown for the leptonic decay of taus (b,d) and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

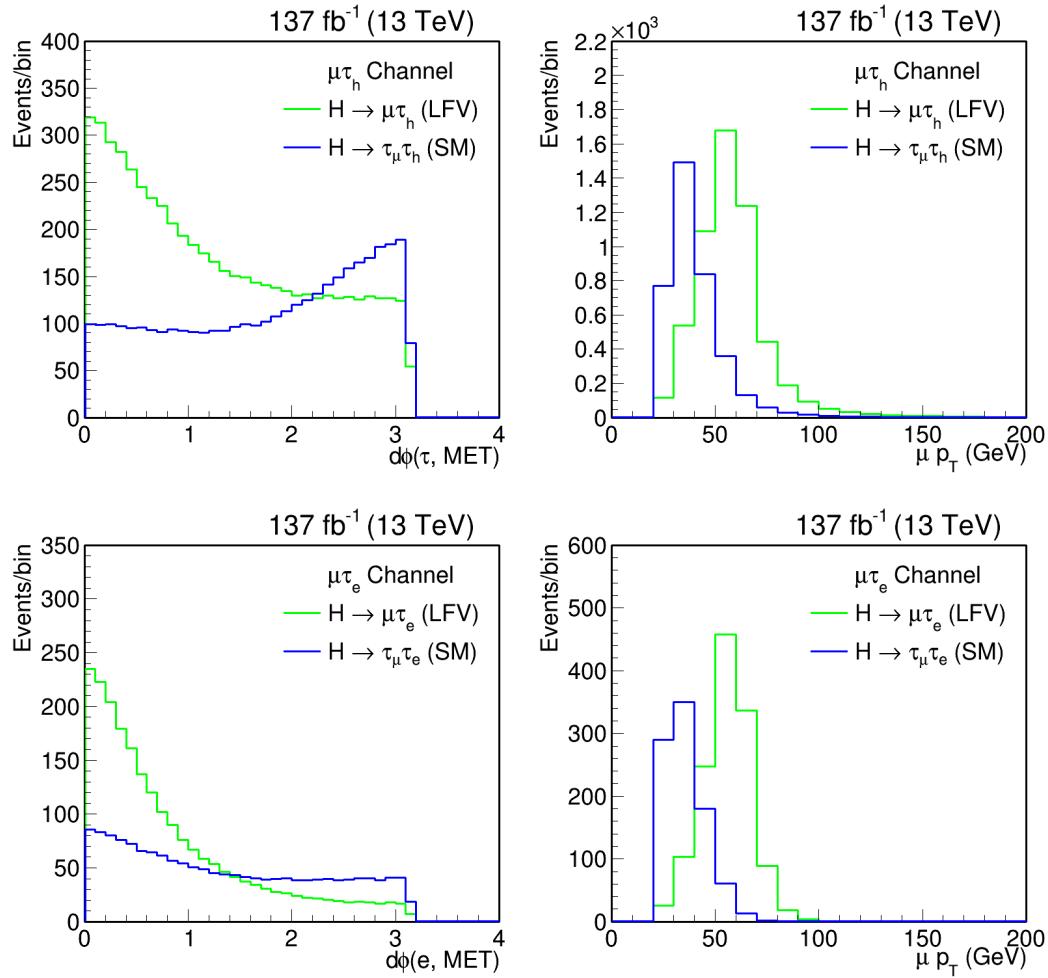


Figure 6.2. Illustration of the differences in $d\phi(\ell, \text{MET})$ and muon transverse momentum spectrums in LFV and SM $H \rightarrow \tau\tau$ processes.

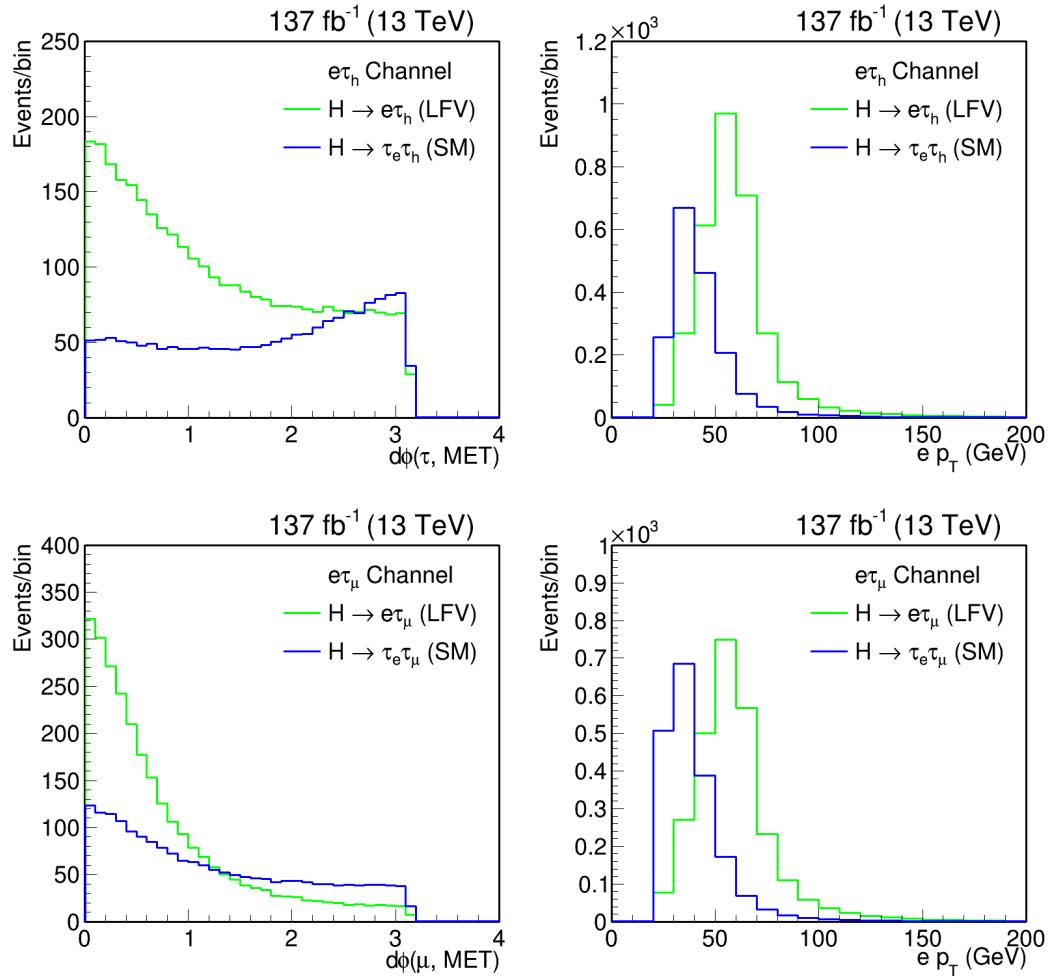


Figure 6.3. Illustration of the differences in $d\phi(\ell, \text{MET})$ and electron transverse momentum spectrums in LFV and SM $H \rightarrow \tau\tau$ processes.

$p_T > 30 \text{ GeV}$ 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 different Higgs boson production mechanisms.

The 0-jet category enhances the 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 two jets' invariant mass (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 BDT discriminator is trained using simulated events, using the TMVA tool of the ROOT analysis package [87]. After applying preselection, a binned likelihood is used to fit a BDT discriminator's distribution for the signal and the background contributions. We call this the BDT fit method. The collinear mass (M_{col}) and the transverse mass ($M_T(\ell)$) 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_{col} provides an estimate of M_H using the Higgs boson candidate's observed decay products. 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 [88]. The momentum of the neutrino coming from tau decay can be approximated to have the same direction as the visible decay products of the τ ($\vec{\tau}^{\text{vis}}$). Figure 6.4 shows the corresponding Feynman diagram for the E_T projected in the direction of the visible decay products of tau.

The component of the \vec{p}_T^{miss} in the direction of the visible tau lepton decay products, is used to estimate the neutrino momentum's transverse component ($p_T^{\nu,\text{est}}$). The collinear mass can then be derived from the visible mass of the $\mu-\tau$ or $e-\tau$ system

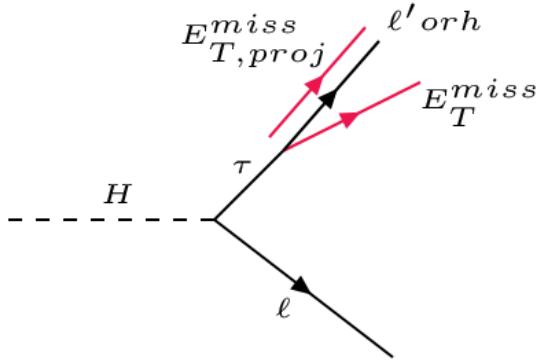


Figure 6.4. Estimation of the neutrino momentum $E_{T,proj}^{miss}$ by using the component of the missing transverse energy E_T^{miss} which is collinear to the visible decay products of tau in the transverse plane.

(M_{vis}) as $M_{\text{col}} = M_{\text{vis}} / \sqrt{x_\tau^{\text{vis}}}$, where x_τ^{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_{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_{col} as a discriminant for a binned likelihood fit. Henceforth, we call this the M_{col} fit method. The selection criterion described was developed without looking at the region's data where the signal is expected [89]. This approach is standard in particle physics analysis and eliminates the experimenter's bias.

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 tree structure's starting point 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 best separation of the signal and the background is used for each split.

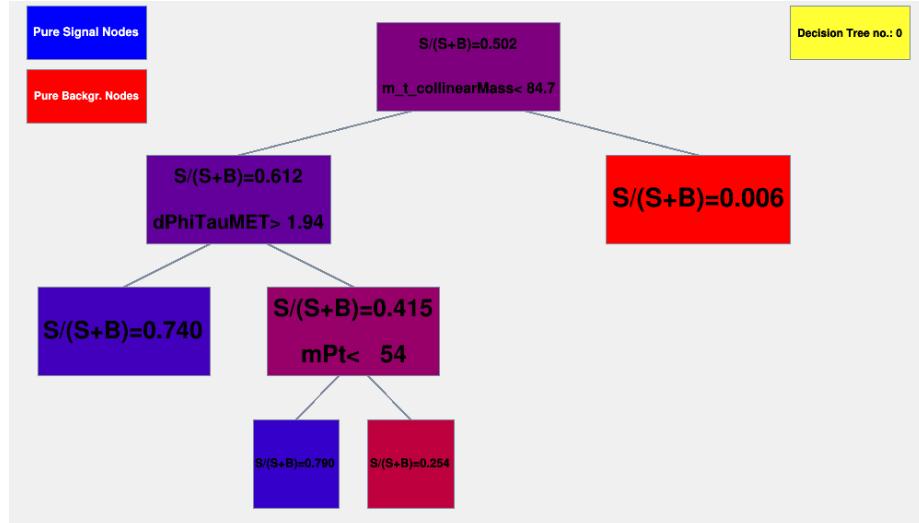


Figure 6.5. Illustration of decision tree.

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

Gini index increase 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 that ends up in a leaf node with most 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 re-weighted 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 α_m which is calculated from the training error and β is the learning rate.

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

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

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, each event's final score 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 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. Hence, 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 muon and an isolated τ_h candidates of opposite electric charge and separated by $\Delta R > 0.5$. The muon candidate is required to have $p_T > 26$ GeV, $|\eta| < 2.1$ and isolation $I_{rel} < 0.15$. The τ_h

candidate is required to have $p_T > 30 \text{ GeV}$ and $|\eta| < 2.3$. Events containing additional electrons, muons, or τ_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 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^μ , $p_T^{\tau_h}$, M_{col} , \vec{p}_T^{miss} , $M_T(\tau, \vec{p}_T^{\text{miss}})$, $\Delta\eta(\mu, \tau_h)$, $\Delta\phi(\mu, \tau_h)$, and $\Delta\phi(\tau_h, \vec{p}_T^{\text{miss}})$. The input variables' distribution to the BDT can be seen in Figure 6.6.

The selection on \vec{p}_T^{miss} is motivated by neutrinos in the tau lepton decays. The neutrino is expected to be collinear with τ_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 $\mu\tau_h$ channel, are shown in chapter 9.

In the M_{col} fit method, additional selection criteria require $M_T(\tau, \vec{p}_T^{\text{miss}}) < 105 \text{ GeV}$ in the 0-, 1-, and 2-jet ggF categories and $M_T(\tau, \vec{p}_T^{\text{miss}}) < 85 \text{ GeV}$ in the 2-jet VBF category. The M_{col} distributions of simulated signal, data, and backgrounds for each category in $\mu\tau_h$ channel, are shown in chapter 9.

6.4 $H \rightarrow \mu\tau_e$ 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

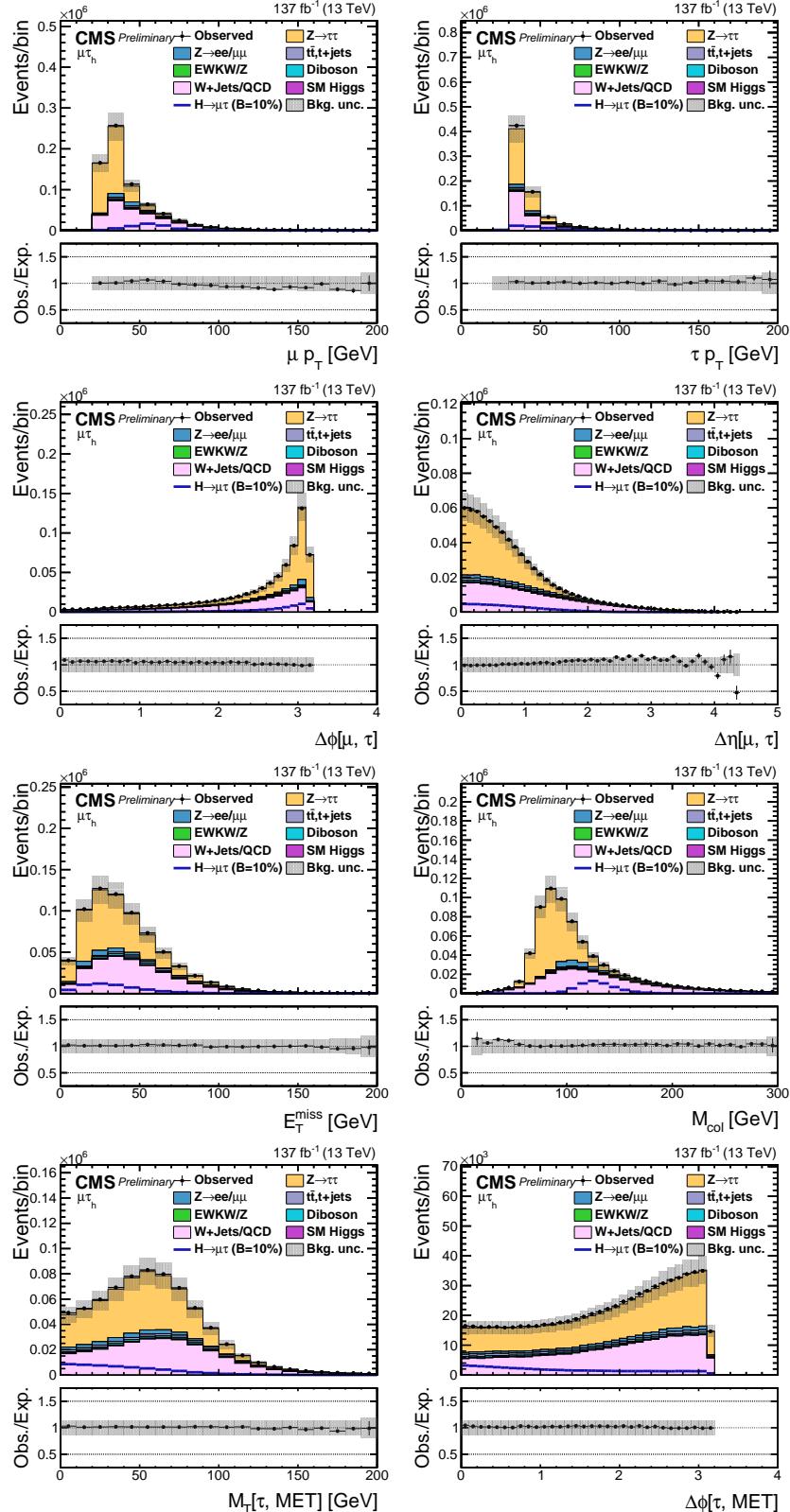


Figure 6.6. Distribution of the input variables to the BDT for the $\mu\tau_h$ process.

TABLE 6.1
 EVENT SELECTION CRITERIA FOR THE KINEMATIC VARIABLES
 FOR THE $H \rightarrow \mu\tau$ CHANNELS

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

M_{col} fit selection								
	0-jet		1-jet		2-jet			
	ggF	VBF	ggF	VBF	ggF	VBF		
M_{jj}	—	—	< 550	≥ 550	—	—	< 550	≥ 550
p_T^μ	—	—	—	—	> 30	> 26	> 26	> 26
$M_T(\mu)$	—	—	—	—	> 60	> 40	> 15	> 15
$M_T(\tau)$	< 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	—	—

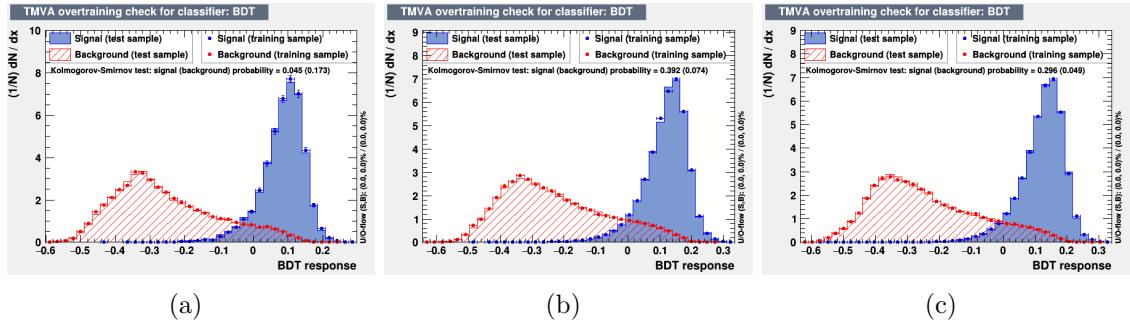


Figure 6.7. Overtraining check as performed in TMVA for the trained BDT in $\mu\tau_h$ channel for 2016 (a), 2017 (b), and 2018 (c).

impact parameter to the PV. 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 $\mu\tau_e$ channel requires an isolated muon and isolated electron candidates of opposite electric 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 $I_{\text{rel}} < 0.15$. The electron candidate is required to have $p_T > 13 \text{ GeV}$, $|\eta| < 2.5$ and isolation $I_{\text{rel}} < 0.1$. 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, or τ_h candidates are vetoed. Events with at least one b jet tagged by the DeepCSV algorithm are rejected to suppress the $t\bar{t}$ background.

Similar to the $\mu\tau_h$ channel, a BDT is trained after applying preselection criteria. The signal training is done similarly, 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. The $Z \rightarrow \ell\ell$ process contributes dominantly for the 0- and 1-jet categories. The QCD multijet back-

ground has the third largest contribution, so we use the lepton anti-isolated control region in data with the same electric charge for both the leptons as additional background for training. The input variables to the BDT are: p_T^μ , p_T^e , M_{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 input variables' distribution to the BDT can be seen in Figure 6.8. The BDT discriminator distributions of simulated signal, data, and backgrounds for each category in $\mu\tau_e$ channel, are shown in chapter 9.

In the M_{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 the azimuthal separation between the electron and \vec{p}_T^{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 $\mu\tau_h$ and $\mu\tau_e$ channels in all categories are summarized in Table 6.1. The M_{col} distributions of simulated signal, data, and backgrounds for each category in $\mu\tau_e$ channel, are shown in chapter 9.

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 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 single electron trigger efficiency used in the $e\tau_h$ channel can be seen in Figure 6.10.

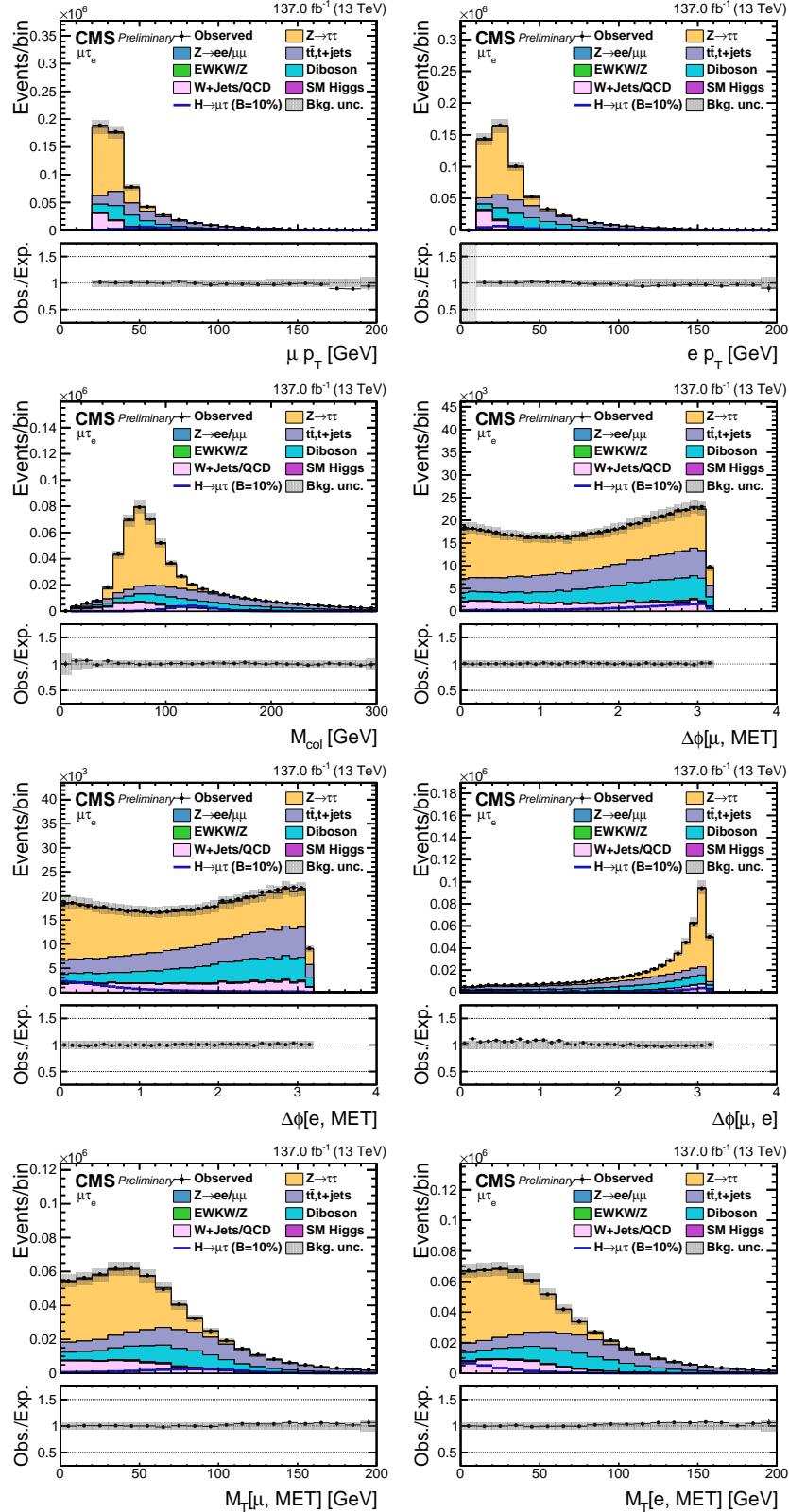


Figure 6.8. Distribution of the input variables to the BDT for the $\mu\tau_e$ process.

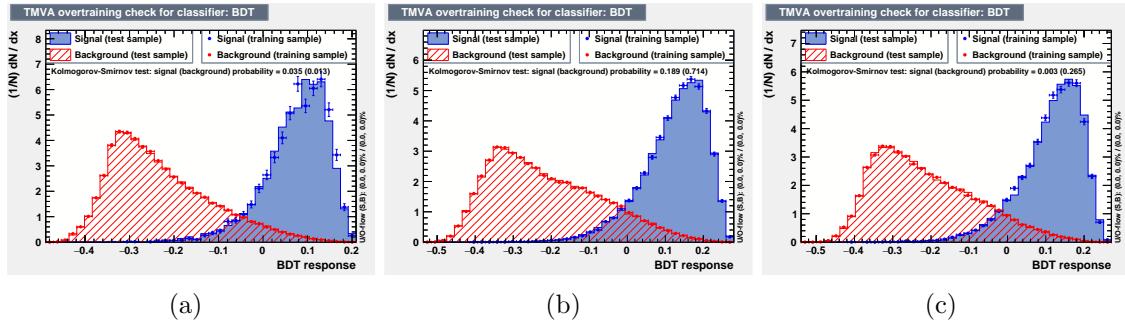


Figure 6.9. Overtraining check as performed in TMVA for the trained BDT in $\mu\tau_e$ channel for 2016 (a), 2017 (b), and 2018 (c).

The preselection in this channel requires an isolated electron and an isolated τ_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 $I_{\text{rel}} < 0.15$. The τ_h candidate is required to have $p_T > 30 \text{ GeV}$ and $|\eta| < 2.3$. Events containing additional electrons, muons, or τ_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 same training samples, as used in the $\mu\tau_h$ channel, are considered. The list of input variables to the BDT training stays the same, except for the addition of the visible mass, M_{vis} variable, and removing \vec{p}_T^{miss} . The M_{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 input variables' distribution to the BDT can be seen in Figure 6.11. The BDT discriminator distributions of simulated signal, data, and backgrounds for each category in $e\tau_h$ channel, are shown in chapter 9.

In the M_{col} fit method, additional selection criteria require $M_T(\tau, \vec{p}_T^{\text{miss}}) < 60 \text{ GeV}$ in all the categories. The M_{col} distributions of simulated signal, data, and backgrounds for each category in $e\tau_h$ channel, are shown in chapter 9.

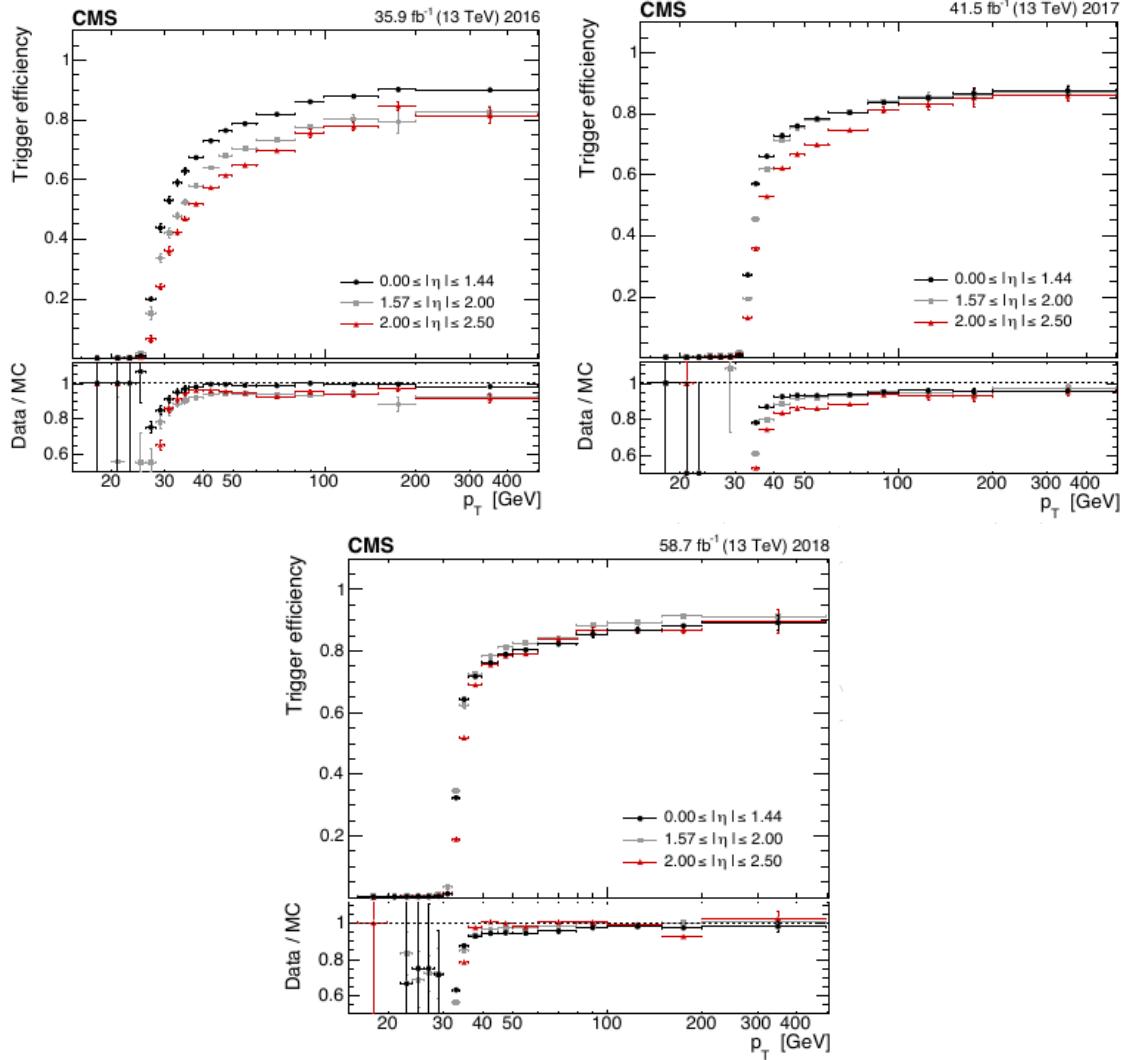


Figure 6.10: Single electron trigger efficiency as a function of p_T in different η regions.

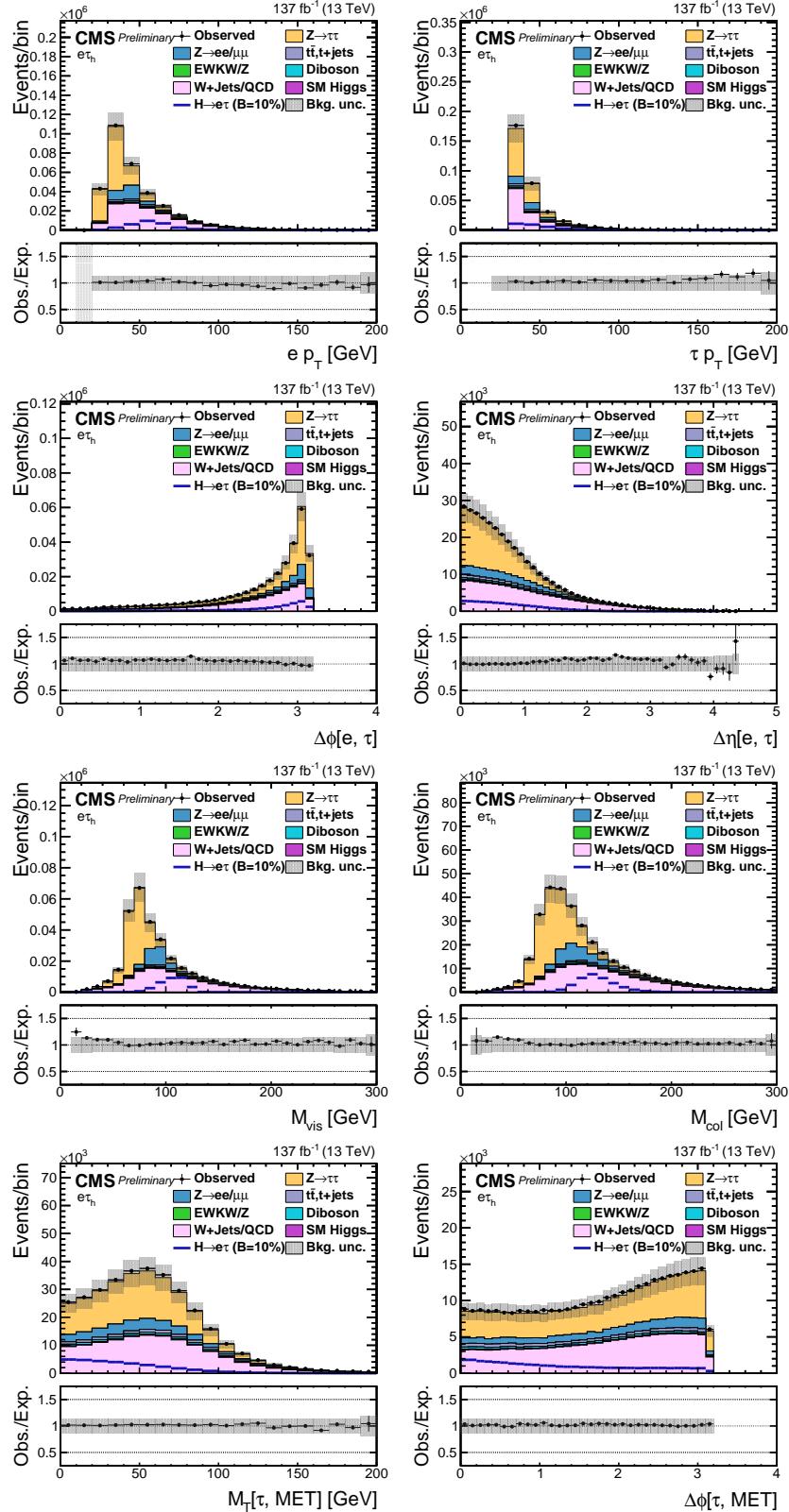


Figure 6.11. Distribution of the input variables to the BDT for the $e\tau_h$ process.

TABLE 6.2

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

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

M_{col} fit selection

	0-jet		1-jet		2-jet				0-jet		1-jet		2-jet	
	ggF	VBF	ggF	VBF	ggF	VBF	ggF	VBF	ggF	VBF	ggF	VBF	ggF	VBF
M _{jj}	—	—	< 550	≥ 550	—	—	< 550	≥ 550	—	—	< 550	≥ 550	—	—
p_T^e	—	—	—	—	> 30	> 26	> 26	> 26	> 30	> 26	> 26	> 26	—	—
M _T (e)	—	—	—	—	> 60	> 40	> 15	> 15	> 60	> 40	> 15	> 15	—	—
M _T (τ)	< 60	< 60	< 60	< 60	< 60	—	—	—	—	—	—	—	—	—
$\Delta\phi(\mu, \vec{p}_T^{\text{miss}})$	—	—	—	—	< 0.7	< 0.7	< 0.5	< 0.3	< 0.7	< 0.7	< 0.5	< 0.3	—	—
$\Delta\phi(e, \mu)$	—	—	—	—	> 2.5	> 1.0	—	—	> 2.5	> 1.0	—	—	—	—

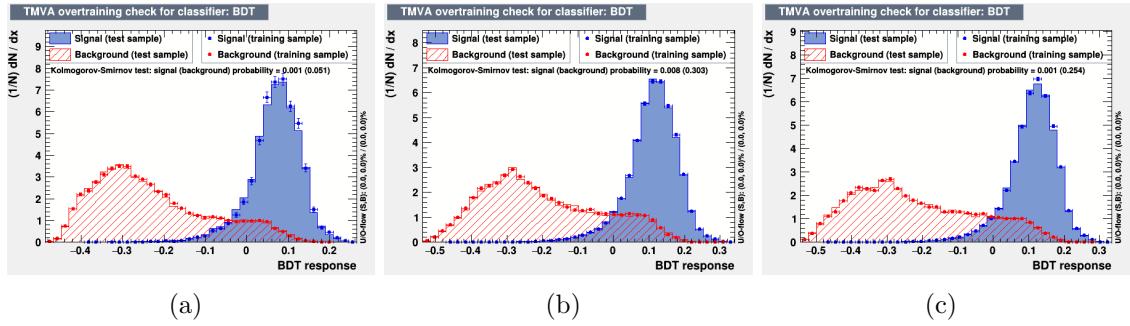


Figure 6.12. Overtraining check as performed in TMVA for the trained BDT in $e\tau_h$ channel for 2016 (a), 2017 (b), and 2018 (c).

6.6 H → e τ_μ channel

The events are required to pass the cross-trigger with p_T thresholds on the electron and the muon. The p_T threshold on the electron is 23 GeV, and on the muon is 8 GeV. The cross-trigger also places a constraint on the longitudinal impact parameter of the two leptons to the PV. 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 $\Delta R < 0.5$.

This channel’s preselection criteria require an isolated electron and isolated muon candidates of opposite charge and separated by $\Delta R > 0.4$. The electron candidate is required to have $p_T > 24 \text{ GeV}$, $|\eta| < 2.4$ and isolation $I_{\text{rel}} < 0.1$. The muon candidate is required to have $p_T > 10 \text{ GeV}$, $|\eta| < 2.5$ and isolation $I_{\text{rel}} < 0.15$. 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, or τ_h candidates are vetoed. Events with at least one b jet tagged by the DeepCSV algorithm are rejected to suppress the $t\bar{t}$ background. The selections for both $e\tau_h$ and $e\tau_\mu$ 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 $\mu\tau_e$ channel are considered. The list of input variables to the BDT training also stays the same, except for the addition of the visible mass, M_{vis} variable, and removal of $M_T(e, \vec{p}_T^{\text{miss}})$. The input variables' distribution to the BDT can be seen in Figure 6.13. The BDT discriminator distributions of simulated signal, data, and backgrounds for each category in $e\tau_\mu$ channel, are shown in chapter 9.

In the M_{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 the azimuthal separation between the muon and \vec{p}_T^{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 $e\tau_h$ and $e\tau_\mu$ channels in all categories are also summarized in Table 6.2. The M_{col} distributions of simulated signal, data, and backgrounds for each category in $e\tau_\mu$ channel, are shown in chapter 9.

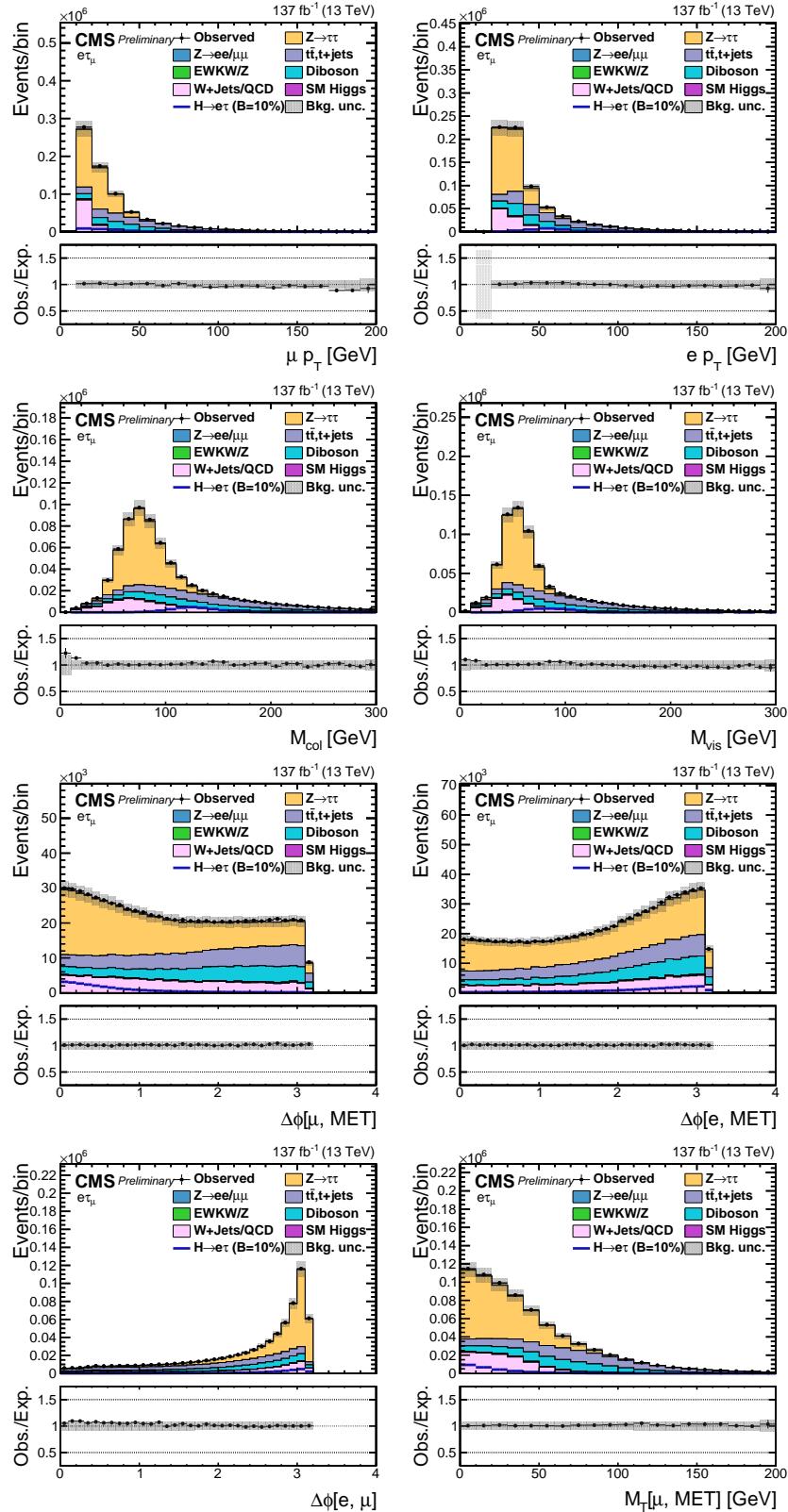


Figure 6.13. Distribution of the input variables to the BDT for the $e\tau_\mu$ process.

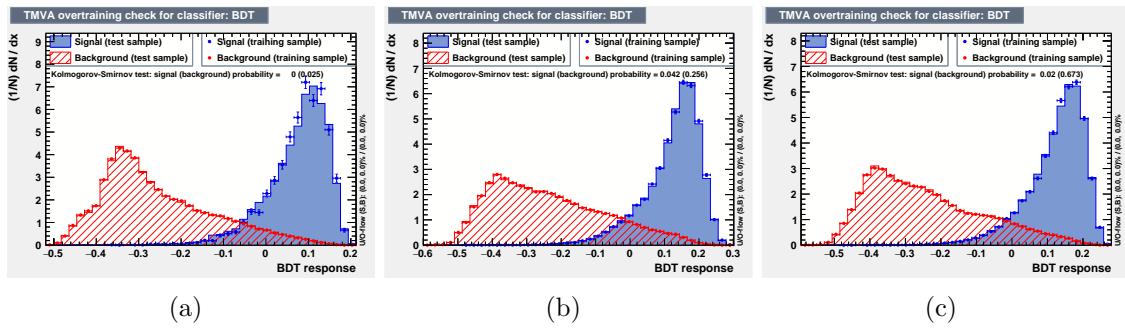


Figure 6.14. Overtraining check as performed in TMVA for the trained BDT in $e\tau_\mu$ channel for 2016 (a), 2017 (b), and 2018 (c).

CHAPTER 7

BACKGROUND ESTIMATION

7.1 Introduction

The dominant background contribution comes from the $Z \rightarrow \tau\tau$ process, in which the muon or electron arises from a tau decay. The other dominant background 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 background contribution.

Other background 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 , 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.

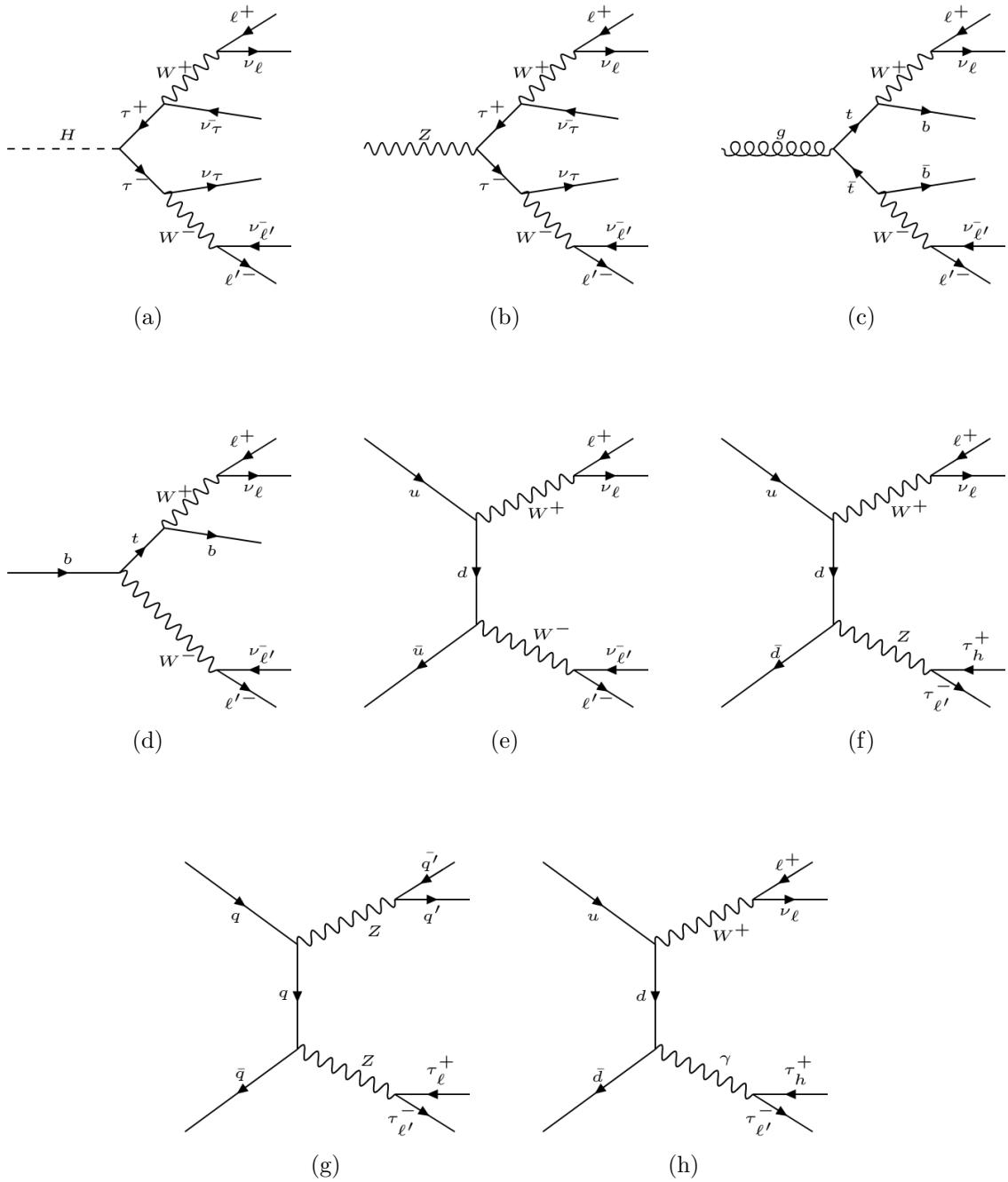


Figure 7.1. Feynman diagrams of background processes to LFV Higgs boson decays: (a) $H \rightarrow \tau\tau$, (b) $Z \rightarrow \tau\tau$, (c) $t\bar{t}$, (d) Single Top, (e) WW , (f) WZ , (g) ZZ , and (h) $W\gamma^{(*)}$.

7.2 Embedding technique

The $Z \rightarrow \tau\tau$ background is estimated from data using the embedding technique [90]. The embedding technique allows for an estimation of the genuine $\tau\tau$ SM 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 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^{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 $\mu\tau_h$ 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_ζ is the difference of the projections of p_T^ℓ plus MET, and the p_T^ℓ on the axis bisecting the two leptons. In $e\tau_h$ channel, the same requirements are placed with the muon variables

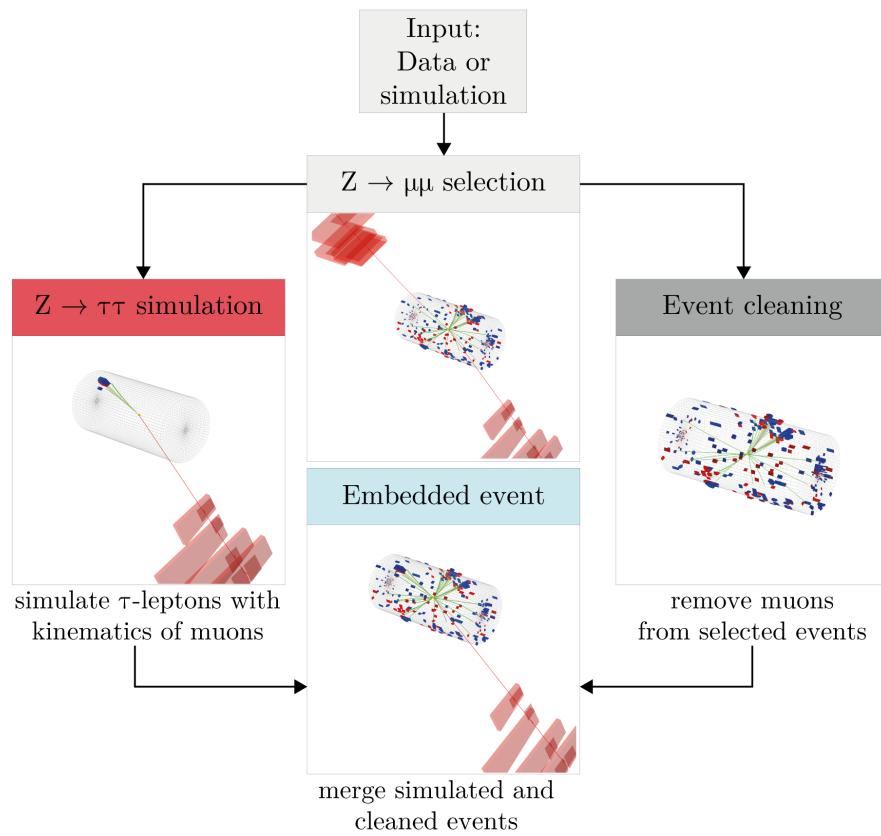


Figure 7.2. Schematic of embedding technique

replaced by corresponding electron variables.

In $\mu\tau_e$ 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 $e\tau_\mu$ 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

The 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 semi data-driven approach results are 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 in the signal region is estimated using misidentification rates measured in a $Z + \text{jets}$ control region and applied to a background enriched sample from collision data. The misidentification rates are evaluated using events with a Z boson candidate and at least one jet that can be misidentified as a lepton. The signal region contrasted with the background enriched regions used for determining the misidentified background can be seen in Figure 7.5.

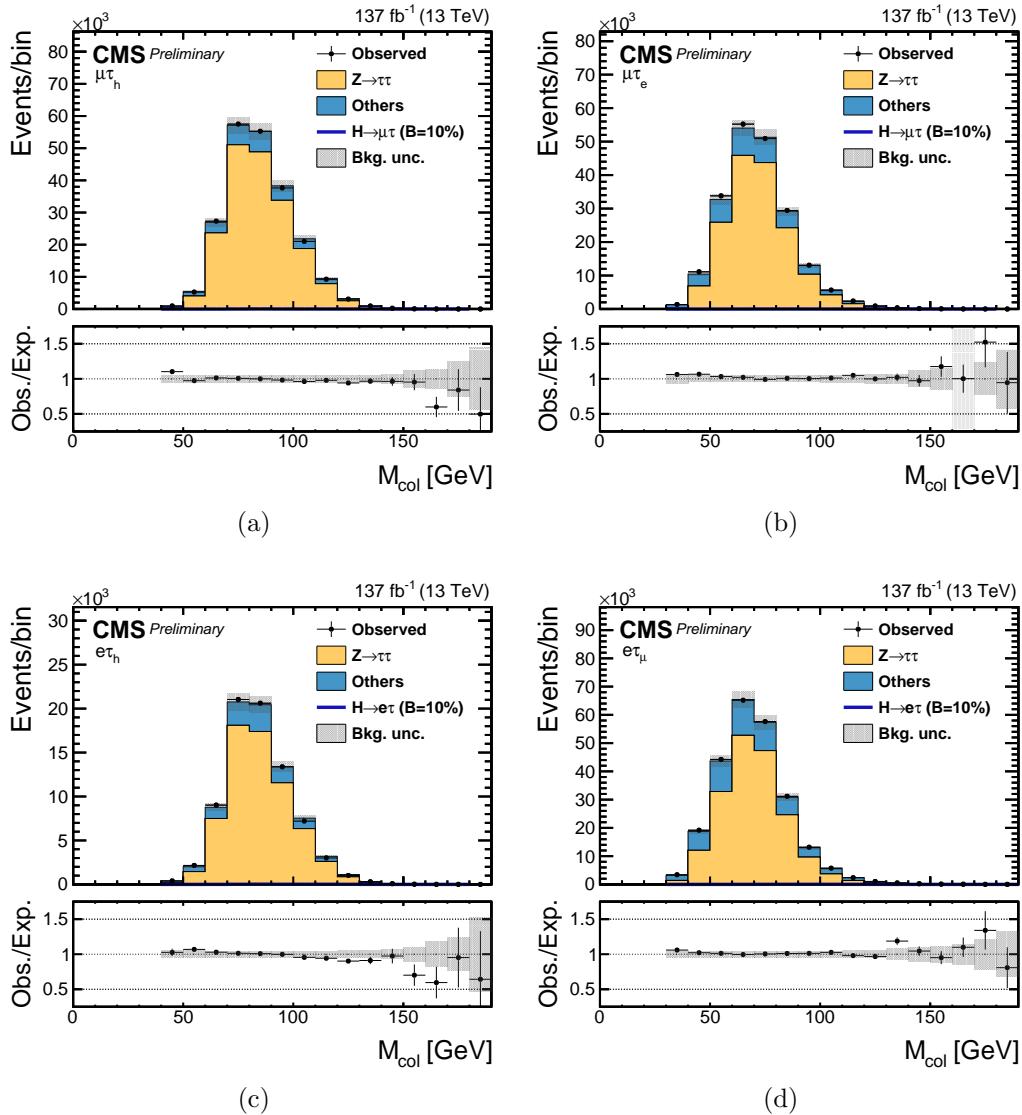


Figure 7.3. Distributions of M_{col} discriminator in the $Z \rightarrow \tau\tau$ control regions for the (a) $\mu\tau_h$, (b) $\mu\tau_e$, (c) $e\tau_h$, and (d) $e\tau_\mu$ channels.

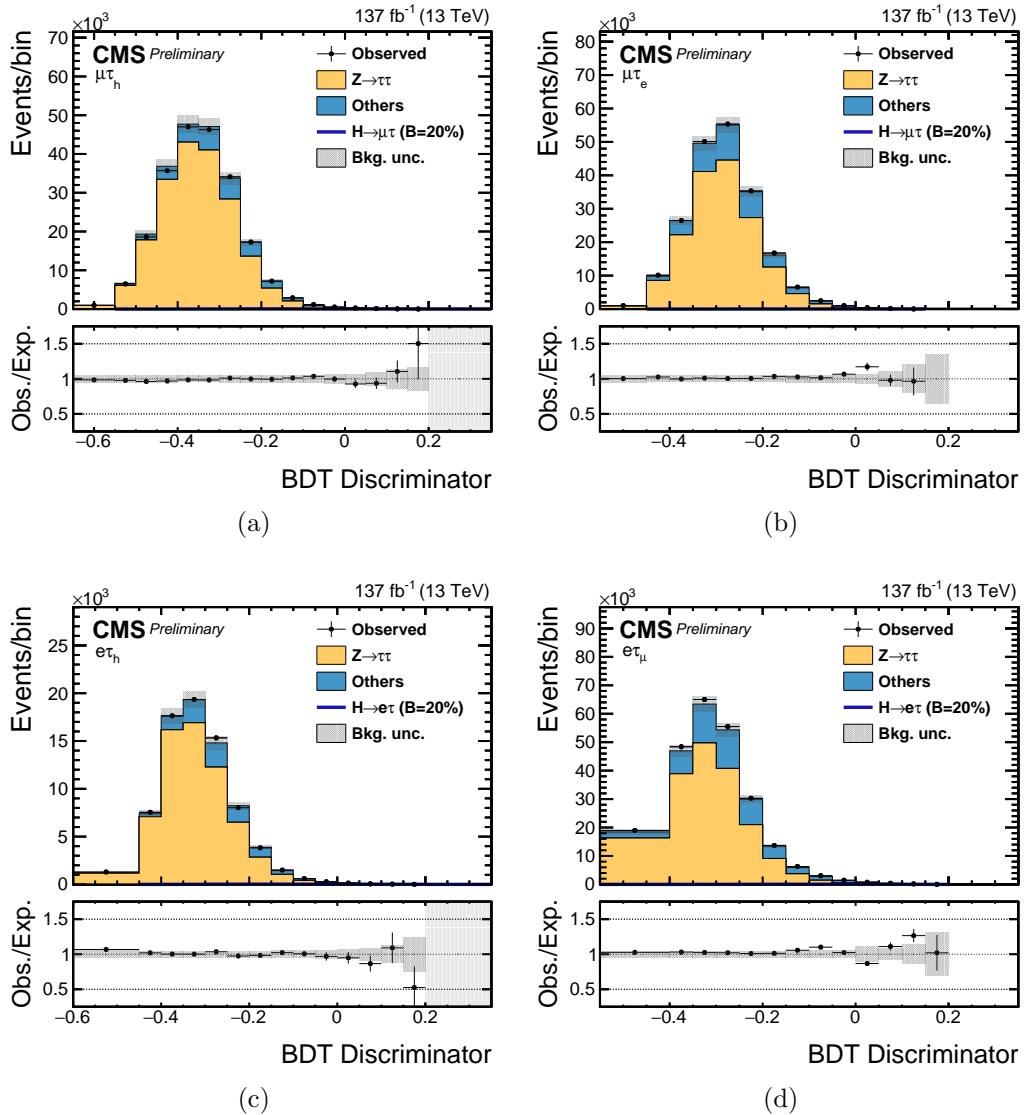


Figure 7.4. Distributions of BDT discriminator in the $Z \rightarrow \tau\tau$ control regions for the (a) $\mu\tau_h$, (b) $\mu\tau_e$, (c) $e\tau_h$, and (d) $e\tau_\mu$ channels.

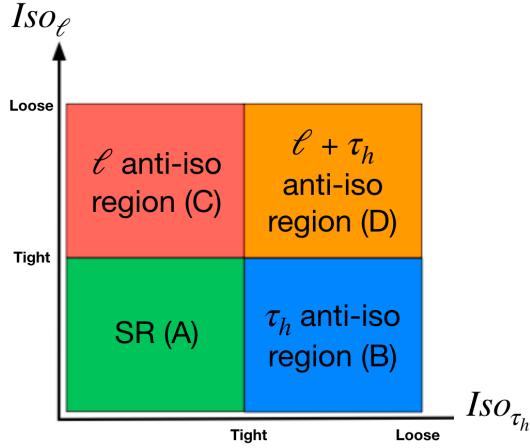


Figure 7.5. Signal region (green) contrasted with the background enriched regions used for estimating the misidentified background

The probabilities with which jets are misidentified as an electron, muon, or τ_h are labeled as f_e , f_μ , and f_{τ_h} , respectively. The Z boson candidate is formed using two muons with $p_T > 26 \text{ GeV}$ and $|\eta| < 2.4$ and $I_{\text{rel}} < 0.15$ for measuring the jet $\rightarrow \tau_h, \mu, e$ misidentification rate. The muons are required to be oppositely charged and have their invariant mass 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 SR. A “signal-like” sample is defined if the jet passes the tight lepton isolation, otherwise a “background-like” sample is defined if it only passes the looser lepton isolation. These two samples are used to estimate f_e , f_μ , and f_{τ_h} 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 τ .

As can be seen in tables 6.1 and 6.2, the isolation for the electron and the muon is required to be $I_{\text{rel}} < 0.15$ and the τ_h is discriminated against jets with a WP that has an identification efficiency of about 70% for “signal-like” sample. For the “background-like” sample, lepton isolation is required to be $0.15 < I_{\text{rel}} < 0.25$ for the muon, $0.15 < I_{\text{rel}} < 0.5$ for the electron, and the τ_h is discriminated against jets with a WP that has an identification efficiency of about $\sim 80\%$ and not pass the “signal-like” WP. After the “signal-like” and “background-like” samples are defined the misidentification rates are computed as a function of the lepton p_T . The τ_h misidentification rate shows a p_T dependence that varies with the τ_h decay mode and $|\eta|$ and are thus evaluated as a function of p_T^τ for the different decay modes and two $|\eta|$ regions ($|\eta| < 1.5$ or $|\eta| > 1.5$).

In the $e\tau_h$ channel, the τ_h misidentification rate is evaluated using events with a Z boson candidate that decays into two electrons with $p_T > 27 \text{ GeV}$ and $|\eta| < 2.5$ and $I_{\text{rel}} < 0.15$. The electrons must be oppositely charged and have their invariant mass between 70 and 110 GeV. The reason for using $Z \rightarrow ee$ events for evaluating the τ_h misidentification rate in $e\tau_h$ channel is that the DNN WPs used for discriminating τ_h against electrons and muons are different in this channel compared to the $\mu\tau_h$ channel. The misidentification rates evaluated using this control region are compatible with the misidentification rates measured in $Z \rightarrow \mu\mu$ events. The electron, muon, and τ_h misidentification rates can be seen in Figure 7.6 and 7.7.

Each event in the background enriched sample 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 , η and decay mode for the τ_h 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 muon (electron) and a misidentified

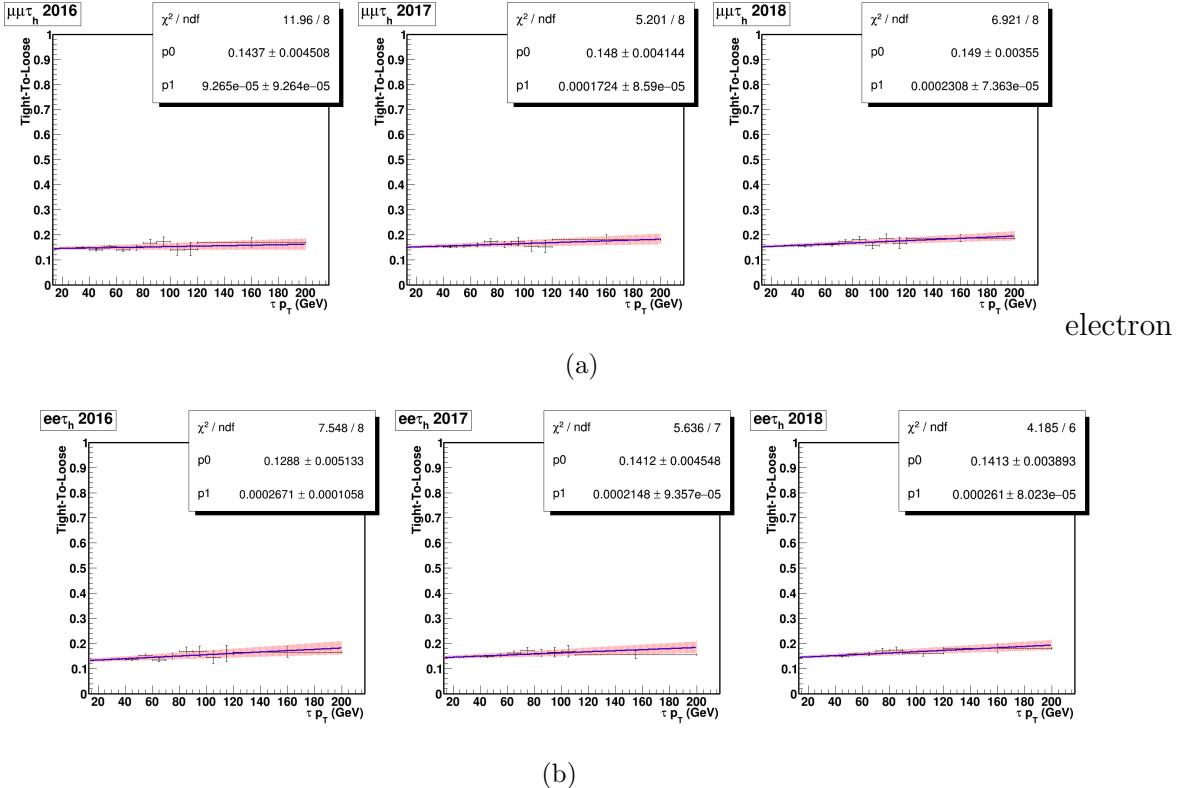


Figure 7.6. Fit performed to τ_h misidentification rates for $\mu\tau_h$ (a) and $e\tau_h$ (b) channel as a function of τ_h p_T for the different years. The misidentification rates used are further parametrized based on τ_h decay mode along with the pseudorapidity of τ_h . However, here only the inclusive misidentification rates are shown.

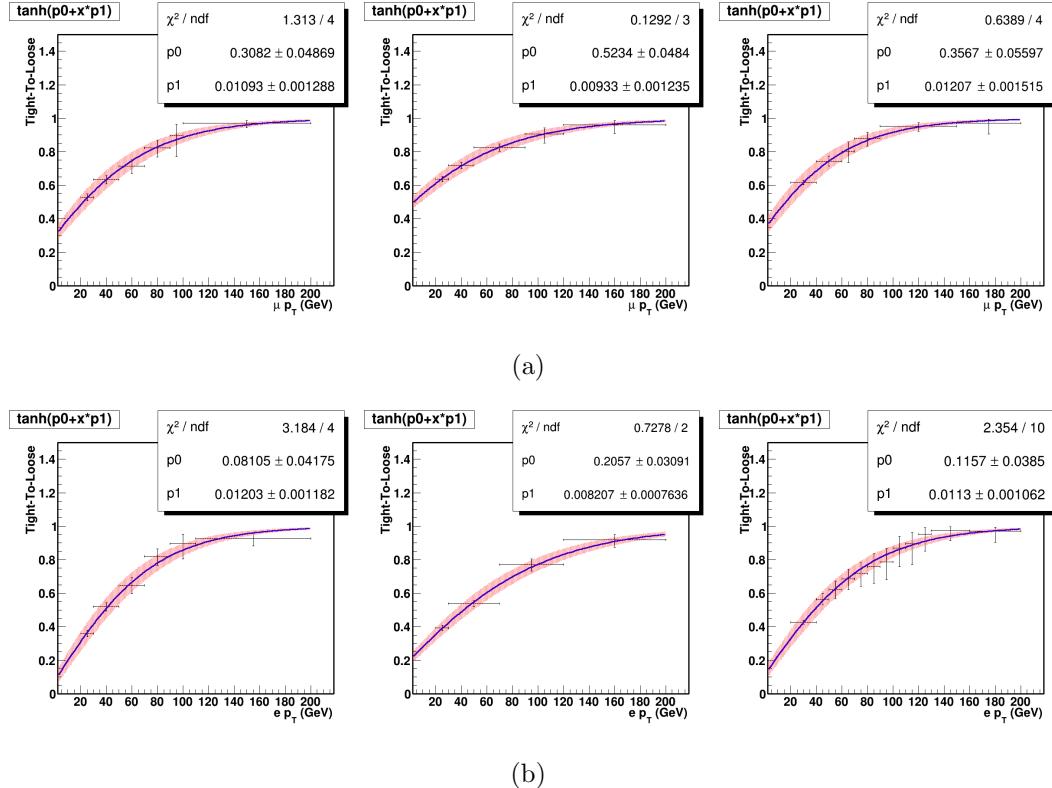


Figure 7.7. Fit performed to the muon (a) and electron (b) misidentification rates as a function of their p_T for 2016 (left), 2017 (center), and 2018 (right). The hyperbolic tangent function is used for performing the fit.

τ_h are subtracted in the $\mu\tau_h$ ($e\tau_h$) channel using a weight, $f_\tau f_\ell / [(1 - f_\tau)(1 - f_\ell)]$, where $\ell = \mu$ or e .

The background estimation is validated in a control region with the same electric charge for both the leptons (same-sign), enhancing the misidentified lepton background. The misidentification rate f_i is applied to events passing preselection and by inverting the lepton pair's charge requirement in both the “background-like” and the “signal-like” regions. The background estimation is also validated in a W boson enriched control region. This control region is obtained by applying the preselection and $M_T(\ell, \vec{p}_T^{\text{miss}}) > 60 \text{ GeV}$ ($\ell = e$ or μ) and $M_T(\tau, \vec{p}_T^{\text{miss}}) > 80 \text{ GeV}$. The same strategy is applied in the $e\tau_h$ 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 $\mu\tau_h$ and $e\tau_h$ channels.

7.3.2 Semi data-driven approach

In the $e\tau_\mu$ and $\mu\tau_e$ channels, QCD multijet background is estimated from the data using events with an electron and a muon with the same electric charge. Contributions from other processes are estimated from simulation and subtracted from data. Extrapolation factors from the control region with the same electric charge for both the leptons to the signal region with an opposite electric charge for both the leptons are measured in data as a function of the jet multiplicity and the ΔR separation between the electron and the muon. The extrapolation factors measured can be seen in Figure 7.10.

The extrapolation factors are 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 ΔR dependency and is parameterized with a linear function. The extrapolation factors are higher for events with low ΔR separation between the electron and the muon, decreasing as the ΔR separation increases. The extrapolation

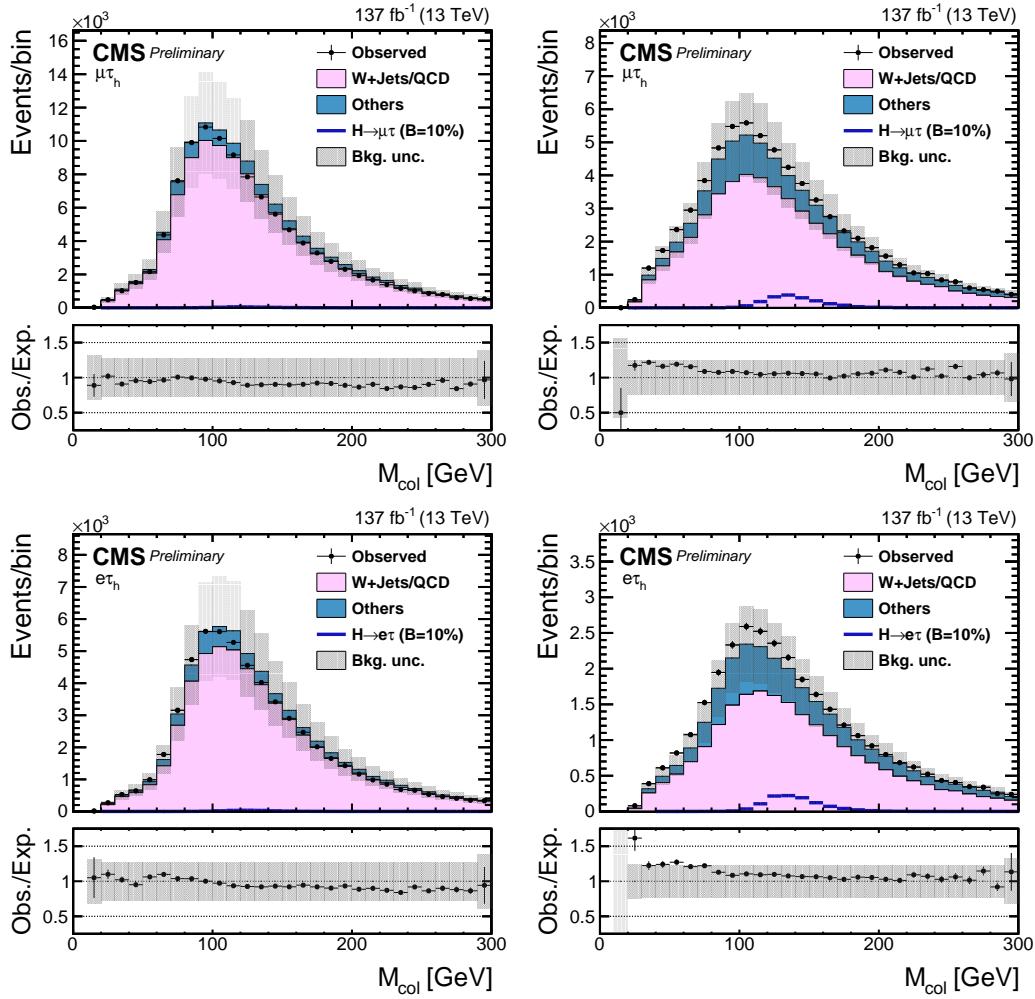


Figure 7.8. Distributions of M_{col} discriminator in the same sign (left) and W boson enriched (right) control regions for the $\mu\tau_h$ (top) and $e\tau_h$ (bottom) channels.

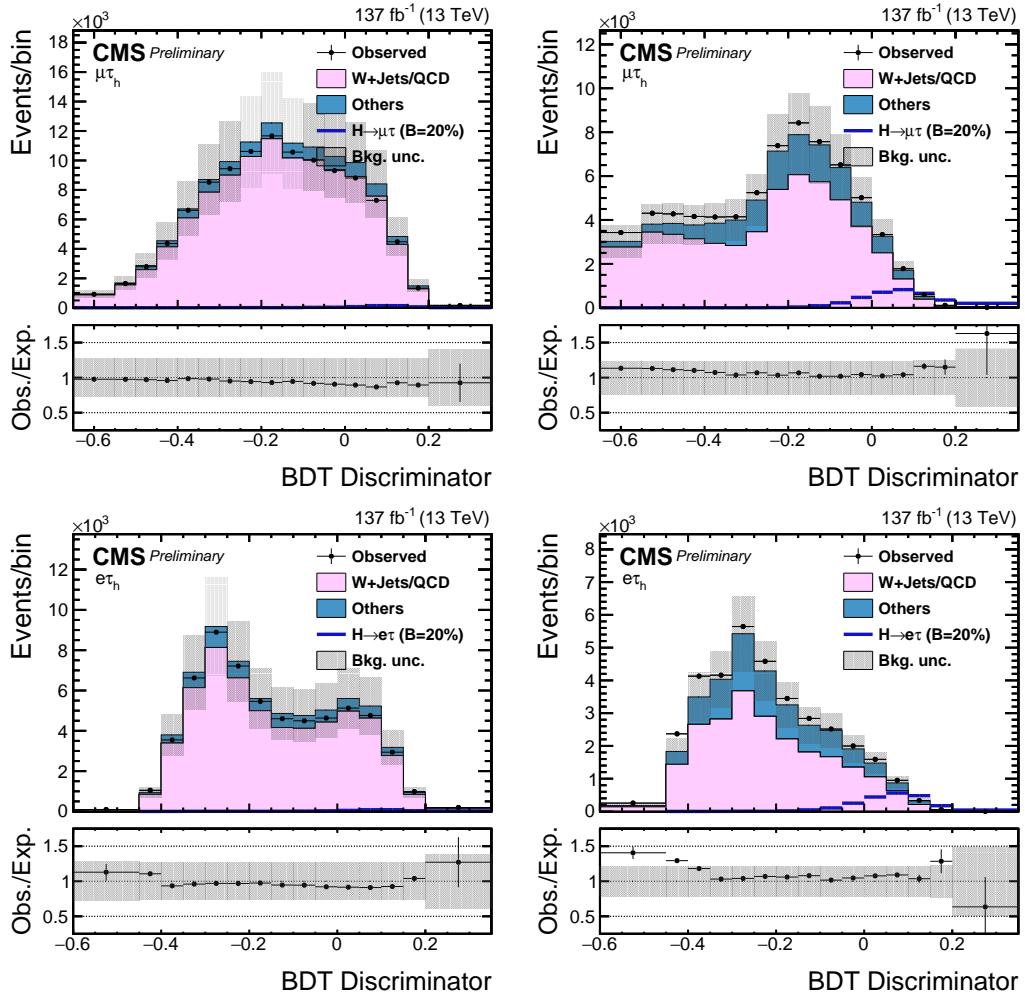


Figure 7.9. Distributions of BDT discriminator in the same sign (left) and W boson enriched (right) control regions for the $\mu\tau_h$ (top) and $e\tau_h$ (bottom) channels.

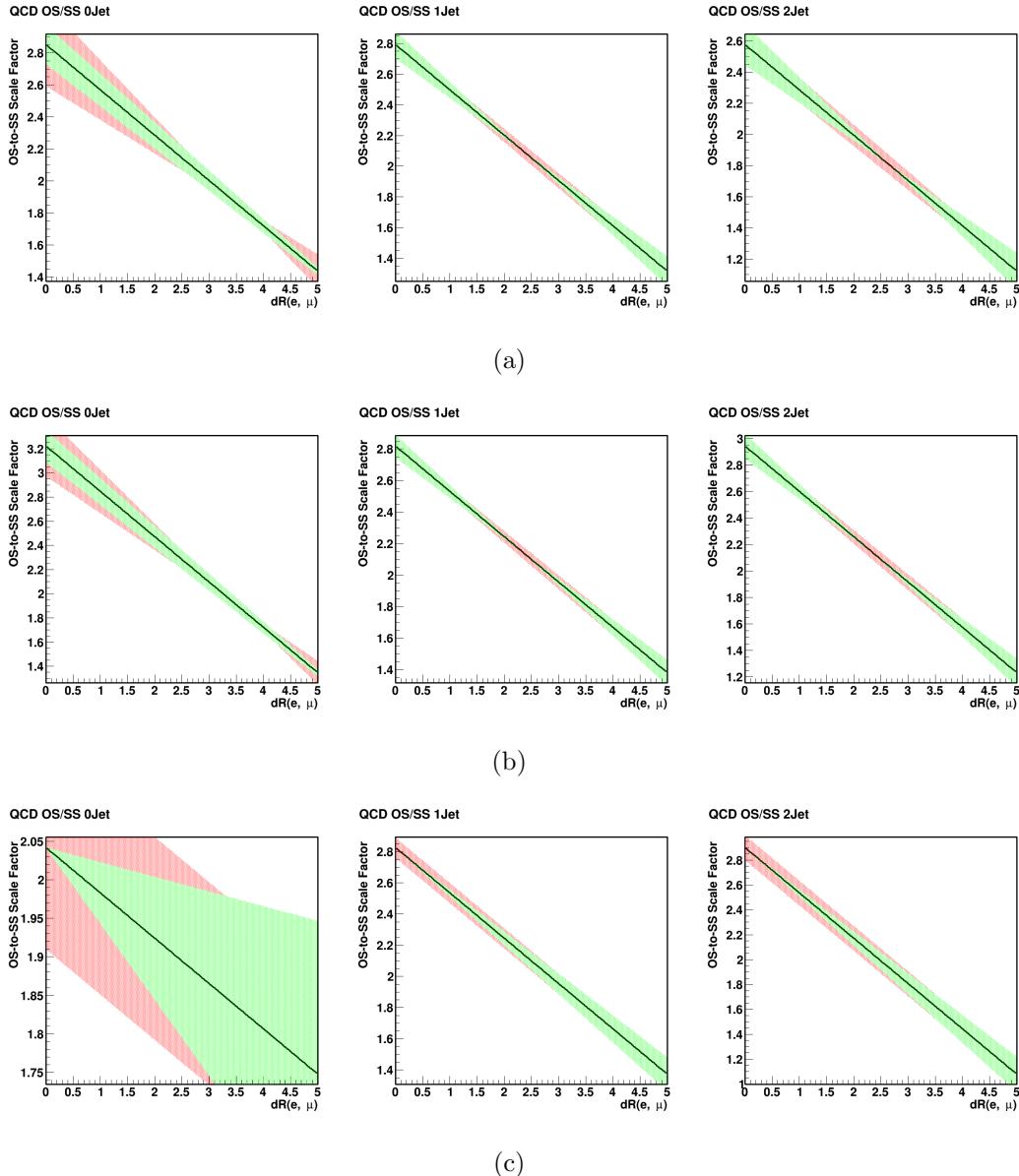


Figure 7.10. The extrapolation factors in events with 0 jets (left), 1 jet (center), and 2 jets (right) for 2016 (a), 2017 (b), 2018 (c). The line is the best fit, and the shaded region corresponds to the shape uncertainties.

factors also depend 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 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.

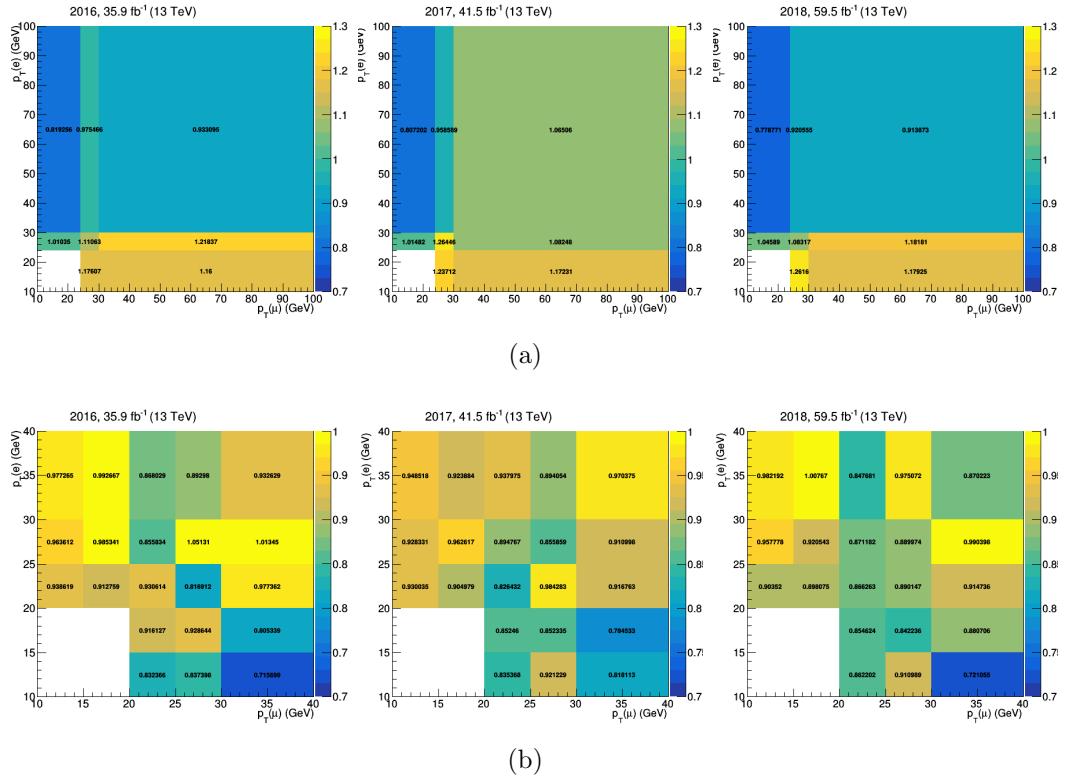


Figure 7.11. (a) Corrections of the QCD OS/SS extrapolation factors determined in the region with an anti-isolated muon as a function of the p_T of the electron and the muon, using data collected in 2016, 2017, and 2018. (b) Correction of the QCD OS/SS extrapolation factors to account for the mismodeling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

As the extrapolation factors are measured in a control region where the muon is anti-isolated, an additional correction is applied to cover a potential mismodeling. This correction is calculated by measuring the 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 for the potential mismodeling induced by anti-isolating the muon. Figure 7.12 shows the comparison of data with estimated background in the muon anti-isolated control regions for the $\mu\tau_e$ channel.

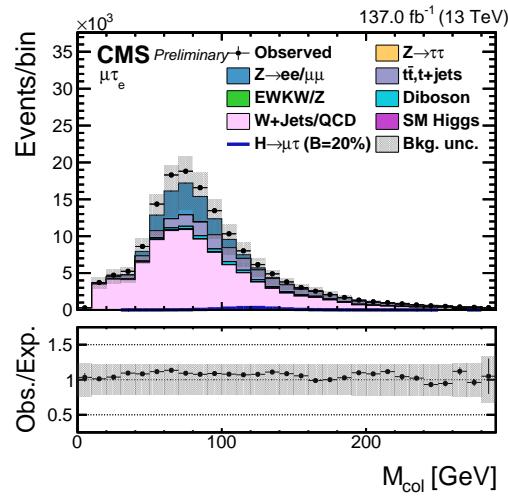


Figure 7.12. Distribution of M_{col} discriminator in the muon anti-isolated control regions for the $\mu\tau_e$ channel.

7.4 Monte Carlo simulations

All the other backgrounds are estimated using MC simulation. In the leptonic channels, the $t\bar{t}$ process has a dominant contribution. This background is validated in a dedicated control region defined by requiring 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 $\mu\tau_e$ and $e\tau_\mu$ channels.

The SM 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.

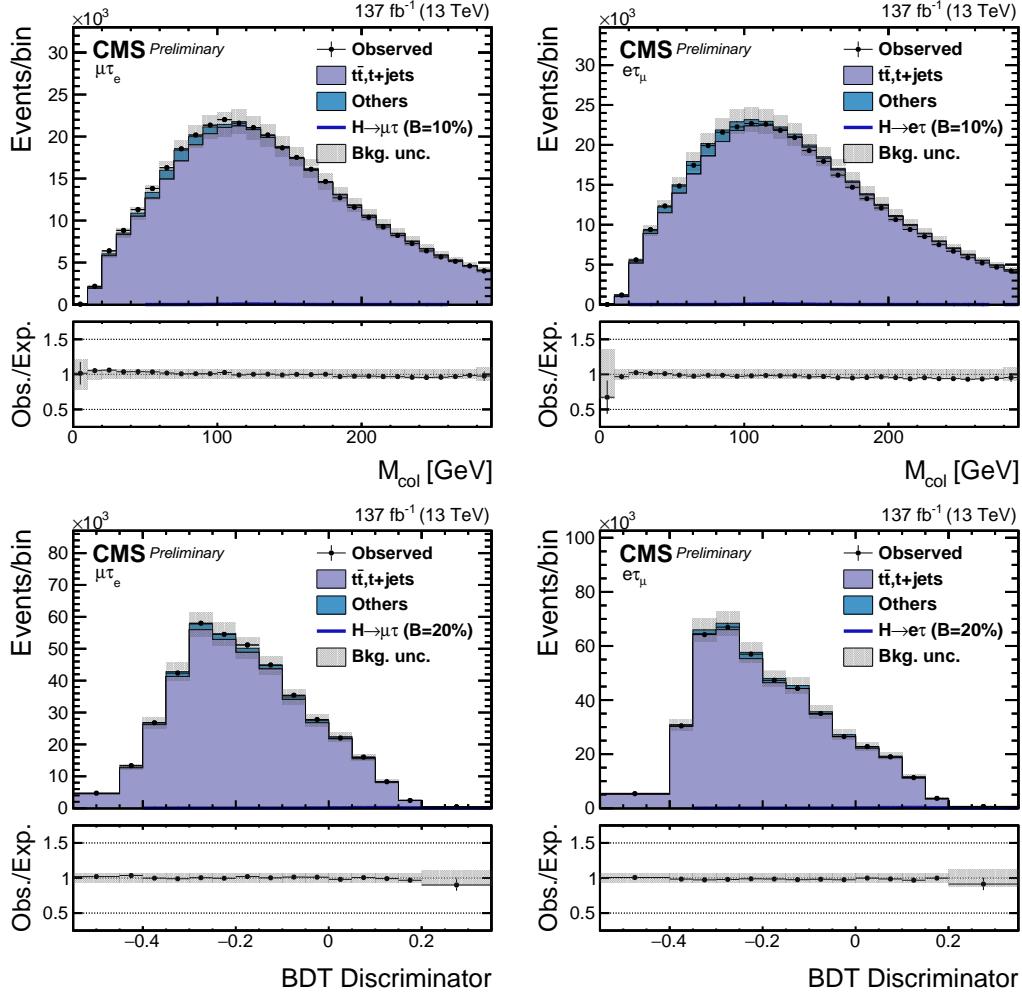


Figure 7.13. Distributions of M_{col} (BDT) discriminator in $t\bar{t}$ enriched control region for $\mu\tau_e$ and $e\tau_\mu$ channel.

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 parame-

ter of interest. In addition to it, the signal and background model contain 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 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 μ is the signal strength modifier, which changes the signal production cross-sections of all the production mechanisms by the same scale of μ . 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 the 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 [91]. 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 component will be ρ , and for the uncorrelated part, it will be $\sqrt{1 - \rho^2}$, where ρ is the

magnitude of the correlation. For example, a 50% correlation will have a correlated component with a magnitude of 0.5 and an uncorrelated part with a magnitude of 0.866.

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 θ . The default value reflects our degree of belief on the real value of θ . The probability distribution function $\rho(\theta|\tilde{\theta})$ 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 θ . The likelihood of the measurement can be constrained by using the probability distribution function of $\tilde{\theta}$. After incorporating the nuisance parameters, the likelihood function now becomes:

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

If no excess is observed then upper exclusion limits can be set on the branching fraction using the CL_s method [92–94]. 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}_μ 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 θ , while $\hat{\mu}$ and $\hat{\theta}$ refer to the global maximum likelihood estimators for μ and θ . 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}_μ^{obs} , is calculated for the signal strength μ , 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 MC pseudo-datasets. These pseudo-datasets are used to construct probability distribution functions 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.

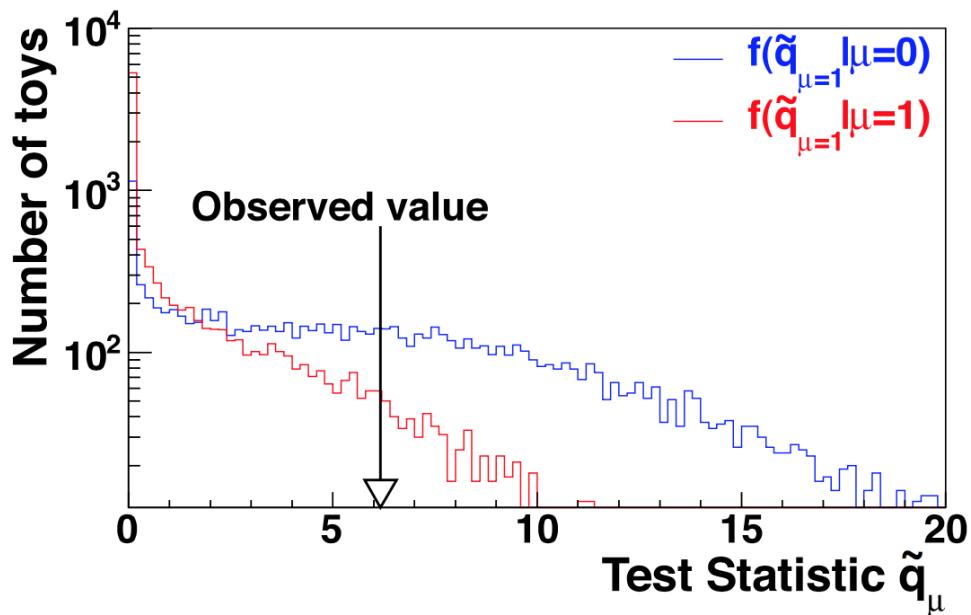


Figure 8.1: Test statistic distributions for ensembles of pseudo-data generated for background-only (blue) and signal-plus-background (red) hypotheses [91].

CL_{s+b} and CL_b correspond to the probabilities of the observations under both hypotheses. CL_{s+b} measures the incompatibility of data under the signal-plus-background hypothesis. CL_b measures the incompatibility of data under the background hypothesis. These quantities alone are not adequate for hypothesis testing. For example, 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 data's incompatibility with the background-only hypothesis alone doesn't tell us that it is compatible with the signal. The ratio of the two quantities referred to as CL_s helps deal with both the situations described above well. The 95% CL is then arrived at by iterating over μ until we have $\text{CL}_s = 0.05$. The μ thus obtained is denoted as $\mu^{95\%\text{CL}}$ and is said to be excluded at 95% CL.

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

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

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

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 CL_s and $\mu^{95\%\text{CL}}$ is calculated for each of them. A distribution is built as a function of the $\mu^{95\%\text{CL}}$ calculated for each 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.

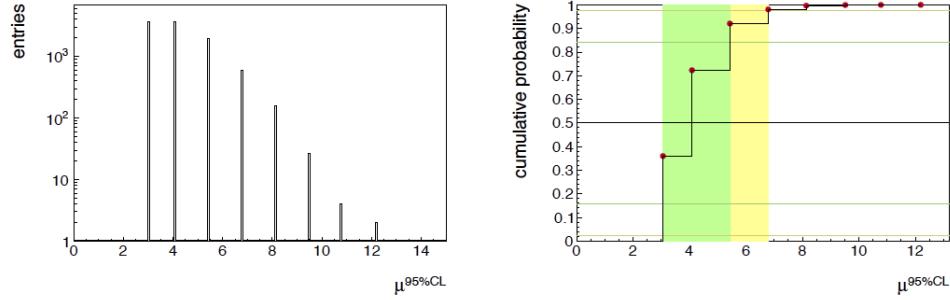


Figure 8.2: (left) An example of differential distribution of possible limits on μ for the background-only hypothesis. (right) The cumulative probability distribution of the plot on the left with 2.5%, 16%, 50%, 84%, and 97.5% quantiles defines the median expected limit as well as the 68% and 95% bands for the expected value of μ 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 probability distribution functions and corresponds to a multiplicative factor in the signal or background yields. These probability distribution functions are characterized by the parameter κ 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 distribution's scale differently in each bin. Such uncertainties are called shape uncertainties [95] and are modeled using a linear interpolation method [96]. Two other distributions obtained by varying the nuisance parameter by ± 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. They are summarized in Tables 8.1 and 8.2.

The uncertainties to reconstruct a τ_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 τ_h candidates correspond to 40% and between 10-70%, respectively, for different bins of p_T , η , and τ_h decay modes. The uncertainty due to τ_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 τ_h energy scale and amounts 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	$\mu\tau_h$	$\mu\tau_e$
Muon Identification/Isolation	2%	2%
Electron Identification/Isolation	—	2%
Trigger	2%	2%
τ_h Identification	p_T dep. (2–3%)	—
$\mu \rightarrow \tau_h$ Identification	10–70%	—
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		
τ_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%
μ 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	$e\tau_h$	$e\tau_\mu$
Muon Identification/Isolation	—	2%
Electron Identification/Isolation	2%	2%
Trigger	2% for SingleElectron, 5% for Cross-Trigger	2%
τ_h Identification	p_T dep. (2–15%)	—
$e \rightarrow \tau_h$ Identification	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%
τ_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%
μ 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 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 PF 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 η are used to correct simulation. The uncertainties in measuring these scale factors are treated as systematic uncertainty.

The uncertainty to track leptons (e, μ), 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% [41, 97–99]. The uncertainty in the measurement of the muon momentum scale is in the range 0.4 – 2.7% for different η , 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 dependent on transverse momentum of τ_h , where a couple of uncertainties are ascribed per fit function. Discriminators with different signal to background efficiency are used to single out τ_h against electrons and muons, which entails an additional 3% uncertainty for the $e\tau_h$ channel. The normalization uncertainties in the data-driven estimation of misidentified lepton backgrounds (jet $\rightarrow \tau_h, \mu, e$), is taken from the orthogonal control region as described in chapter 7.

An additional shape uncertainty is estimated for the misidentified lepton background in the W boson enriched control region. Hence, the signal region is defined

orthogonally to the definition of the W boson enriched control region in hadronic channels. The variables $\Delta\phi(\ell, MET)$ are chosen because a clear discrepancy can be seen for them in the W boson enriched control region. The discrepancy is obtained by subtracting the other background's contribution from data and dividing it by the estimated misidentified lepton background. It is parameterized as a function of $\Delta\phi(\mu, \vec{p}_T^{\text{miss}})$ for $\mu\tau_h$ channel and as a function of $\Delta\phi(e, \vec{p}_T^{\text{miss}})$ for $e\tau_h$ channel. This is not applied as a correction to the signal region because no such discrepancy is observed in the signal region, and hence only a shape uncertainty is considered. The discrepancy in the W boson enriched control region, parametrization for shape uncertainty, and the distribution observed in the signal region is shown in Figure 8.3 for $H \rightarrow \mu\tau_h$ channel and Figure 8.4 for $H \rightarrow e\tau_h$ 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 extrapolation factors as a function of the spatial separation between electron and muon, and lepton p_T are taken into account. The uncertainties from extrapolation factors 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 of 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 PDFs, and the strong coupling constant (α_S). These uncertainties affect the normalization of the signal shape and are taken from Ref. [100]. The variations of the QCD scales results in a 3.9%, 0.5%, 0.9%, and 0.8% uncertainty for the ggF, VBF, ZH, WH cross-sections, respectively, while variations of PDF+ α_S results in 3.2%, 2.1%, 1.3%, 1.9% uncertainty. The acceptance effects are also taken into account while varying

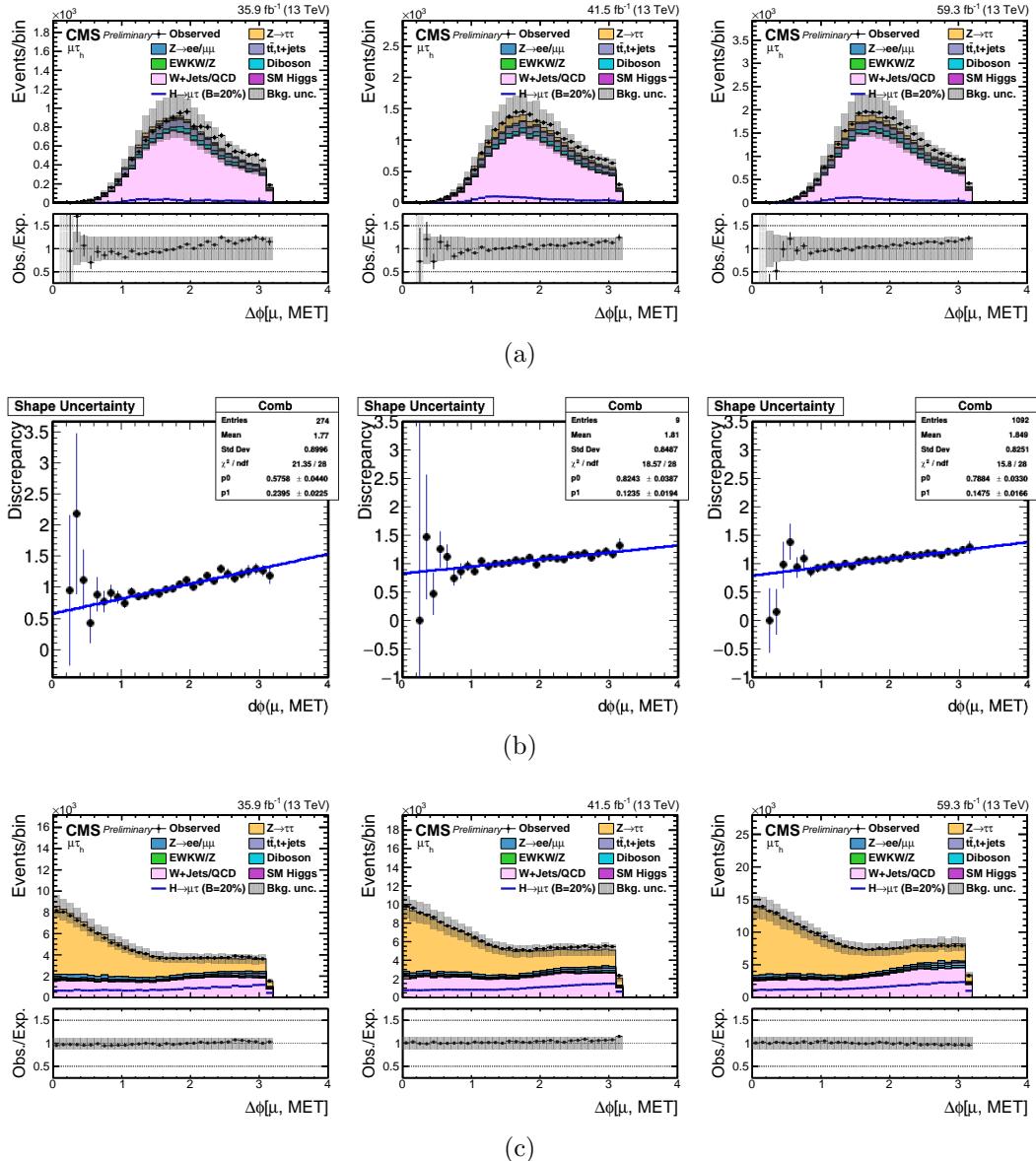


Figure 8.3. Shape uncertainty estimated for the $H \rightarrow \mu\tau_h$ channel. (a) Distribution of $\Delta\phi(\mu, \text{MET})$ from the W boson enriched control region. (b) The discrepancy seen above is parametrized with a linear function. (c) Distribution of $\Delta\phi(\mu, \text{MET})$ from the signal region.

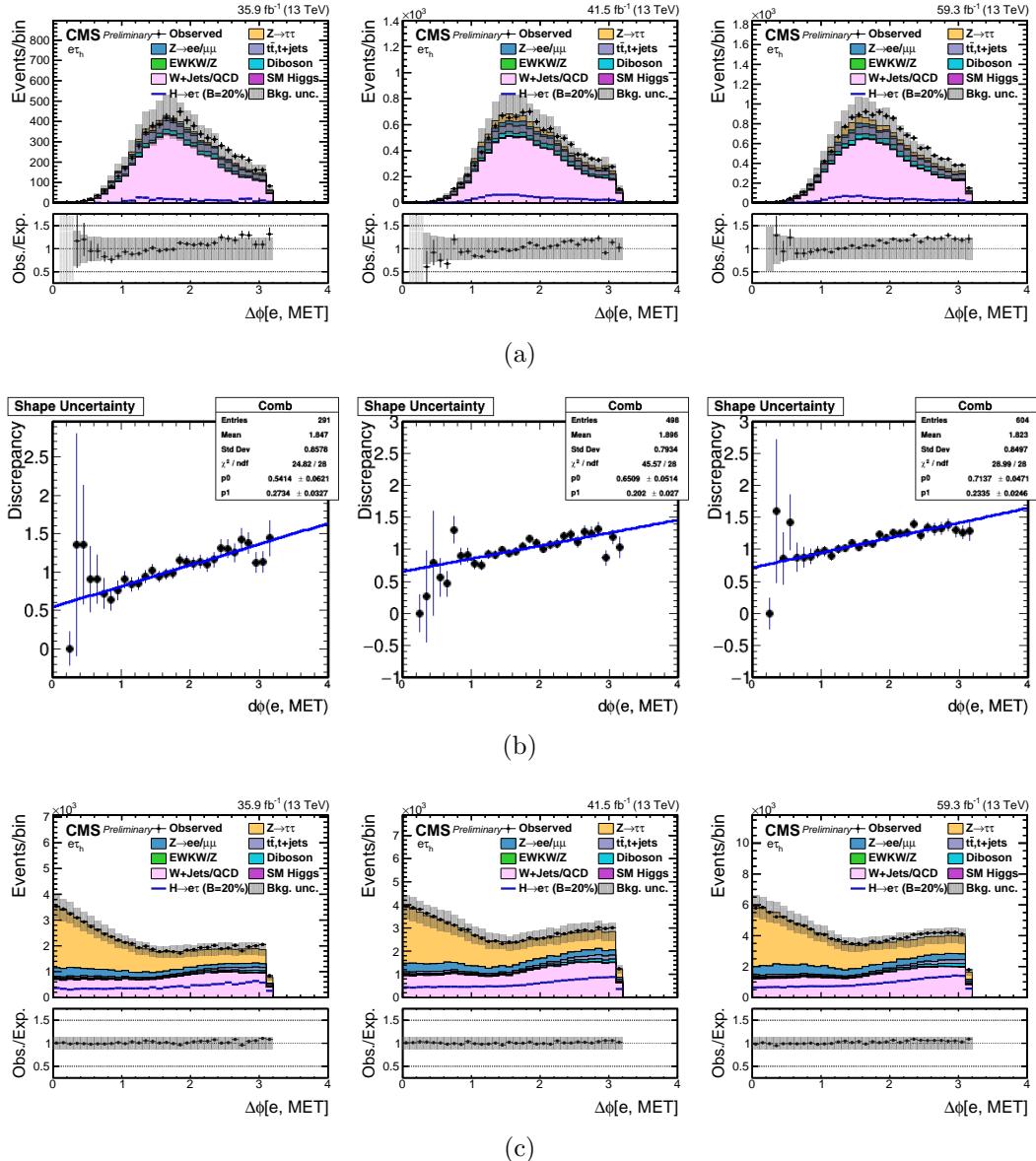


Figure 8.4. Shape uncertainty estimated for the $H \rightarrow e\tau_h$ channel. (a) Distribution of $\Delta\phi(e, \text{MET})$ from the W boson enriched control region. (b) The discrepancy seen above is parametrized with a linear function. (c) Distribution of $\Delta\phi(e, \text{MET})$ from the signal region.

the renormalization and factorization scales and the PDF choice and PDF+ α_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 α_S . In addition, the latest CMS measurement of the $H \rightarrow \tau\tau$ signal strength $\mu = 0.85^{+0.12}_{-0.11}$ has been used. 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 α_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 [95, 101] approach is used, assigning 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 to minimize the number of parameters needed in the maximum-likelihood fit. They are treated uncorrelated between bins, processes, and categories.

The integrated luminosities of 2016, 2017, and 2018 data-taking periods are individually known with uncertainties in the 2.3–2.5% range [45–47], while the total Run 2 (2016–2018) integrated luminosity has an uncertainty of 1.8%, the improvement in precision reflecting the (uncorrelated) time evolution of some systematic effects. The uncertainty 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 estimated 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 Level-1

trigger. A gradual shift in the timing of the inputs of the ECAL Level-1 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 ≈ 50 GeV (≈ 100 GeV), in the region $2.5 < |\eta| < 3.0$ the efficiency loss is $\approx 10\text{-}20\%$, depending on p_T , η , 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 preselection. The cross-check results are obtained from a fit to the M_{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_{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% (0.15%) for $H \rightarrow \mu\tau$ and 0.22% (0.16%) 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_{col} distribution after applying additional selections [9] yields a sensitivity on the Higgs branching fraction of 0.38% (0.26%) for $H \rightarrow \mu\tau$ and 0.32% (0.28%) for $H \rightarrow e\tau$.

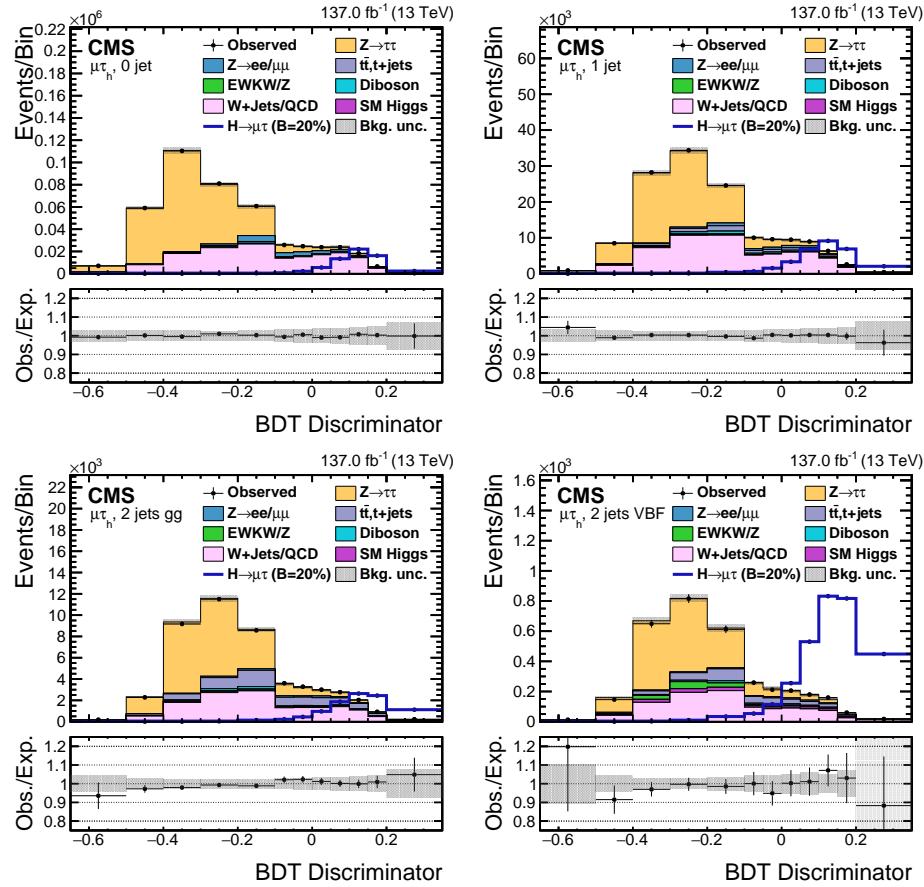


Figure 9.1. BDT discriminator distributions for the observed and estimated background in the $\mu\tau_h$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow \mu\tau) = 20\%$. The $\mu\tau_h$ channel categories 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 statistical and systematic uncertainties added in quadrature.

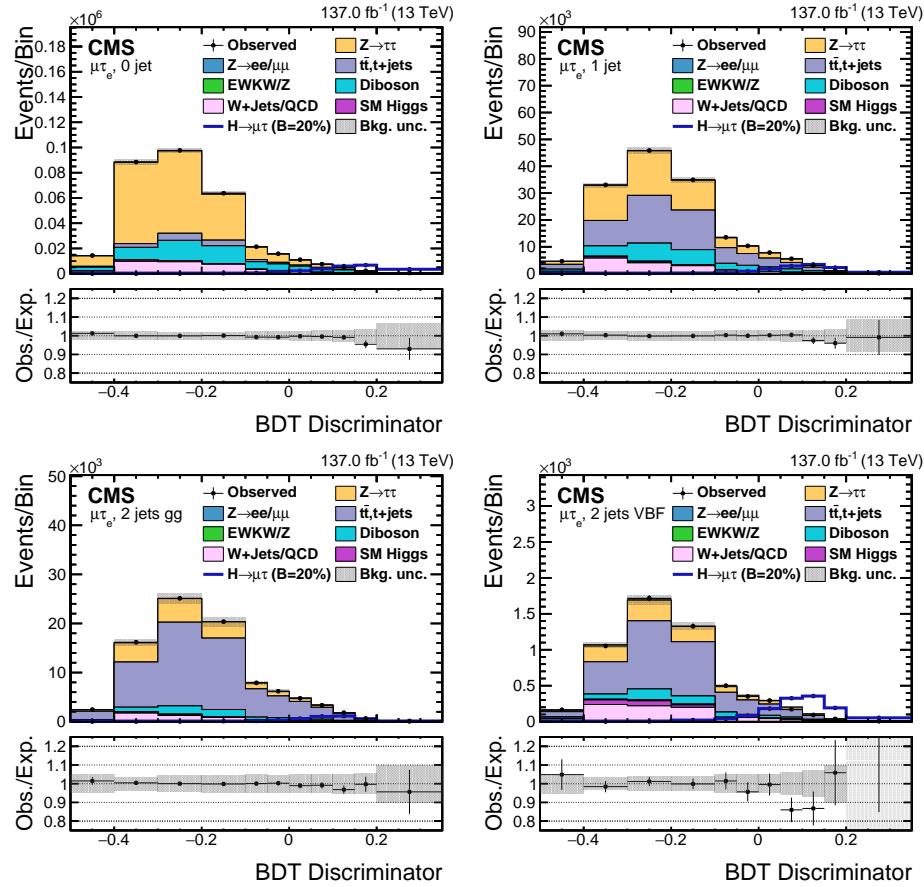


Figure 9.2. BDT discriminator distributions for the observed and estimated background in the $\mu\tau_e$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow \mu\tau) = 20\%$. The $\mu\tau_e$ channel categories 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 statistical and systematic uncertainties added in quadrature.

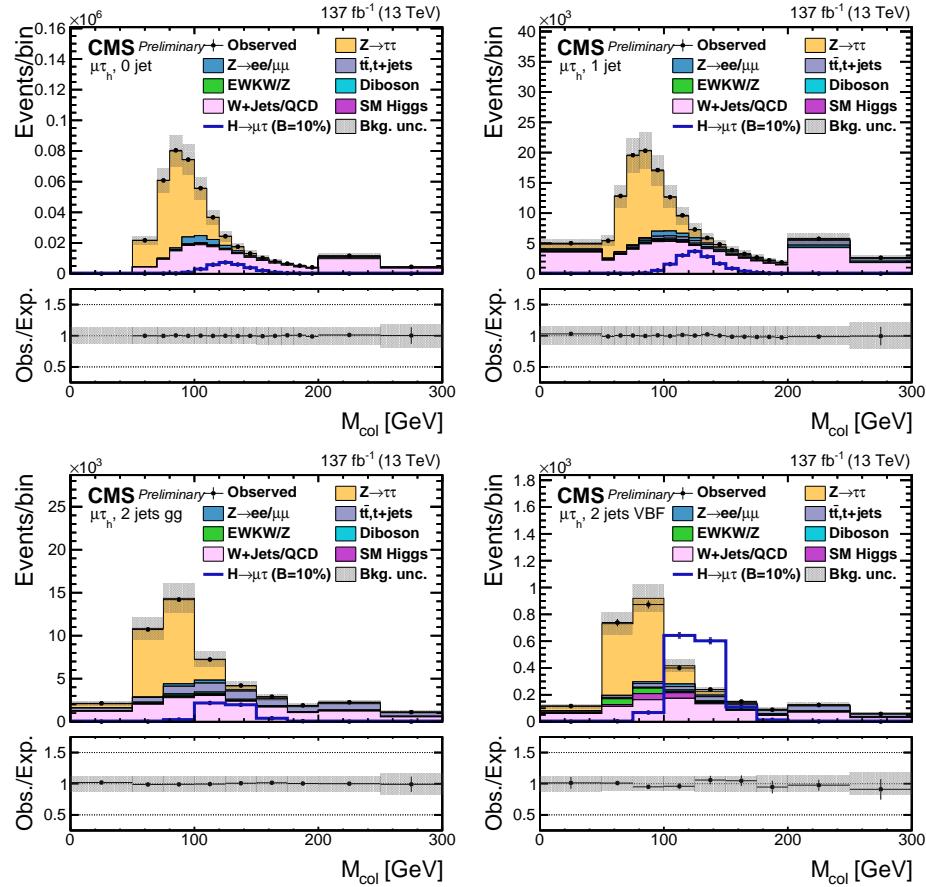


Figure 9.3. M_{col} distributions for the observed and estimated background in the $\mu\tau_h$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow \mu\tau) = 10\%$. The $\mu\tau_h$ channel categories are 0 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 statistical and systematic uncertainties added in quadrature.

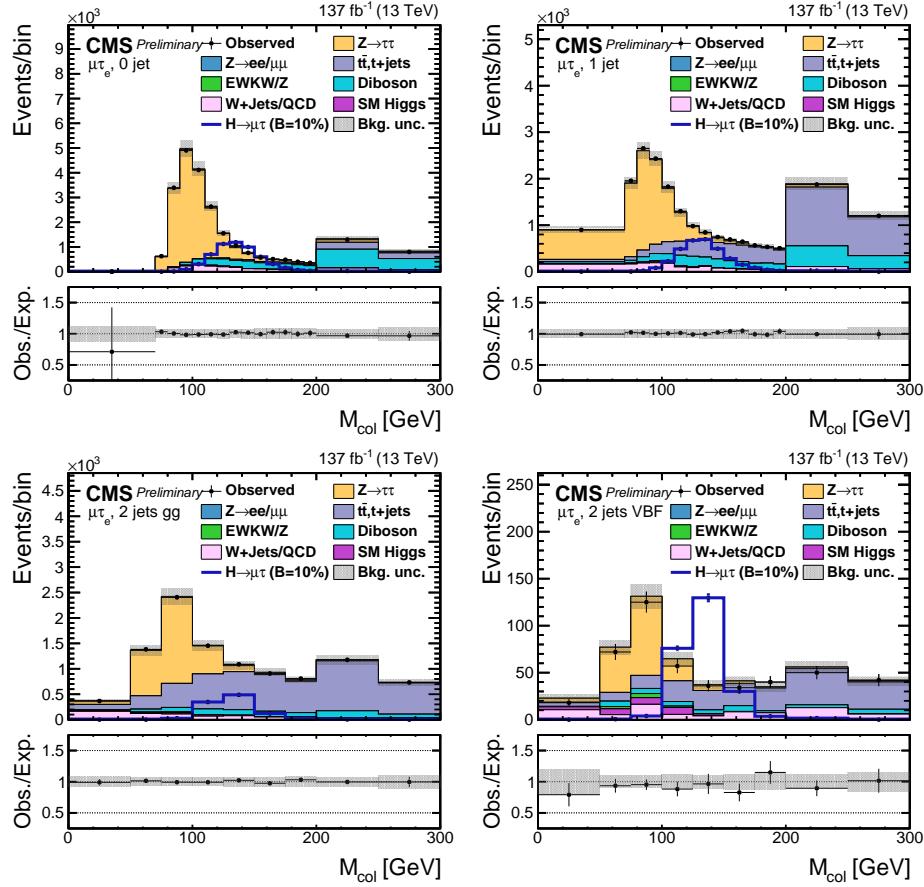


Figure 9.4. M_{col} distributions for the observed and estimated background in the $\mu\tau_e$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow \mu\tau) = 10\%$. The $\mu\tau_e$ channel categories are 0 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 statistical and systematic uncertainties added in quadrature.

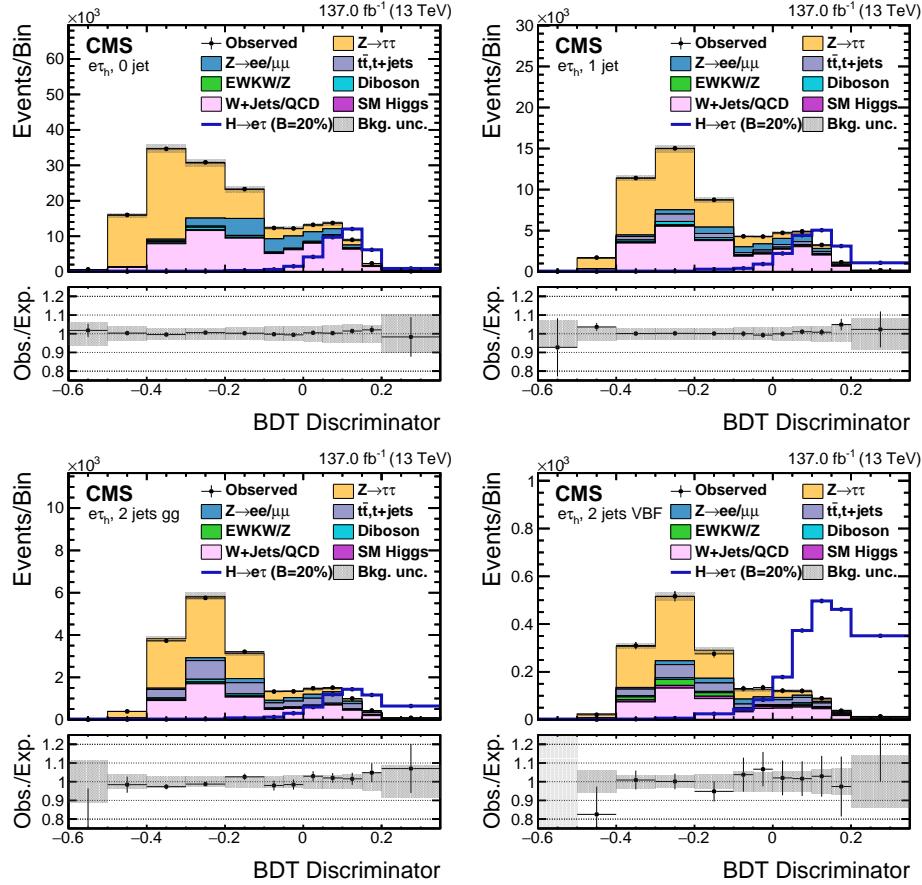


Figure 9.5. BDT discriminator distributions for the observed and estimated background in the $e\tau_h$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow e\tau) = 20\%$. The $e\tau_h$ channel categories are 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 statistical and systematic uncertainties added in quadrature.

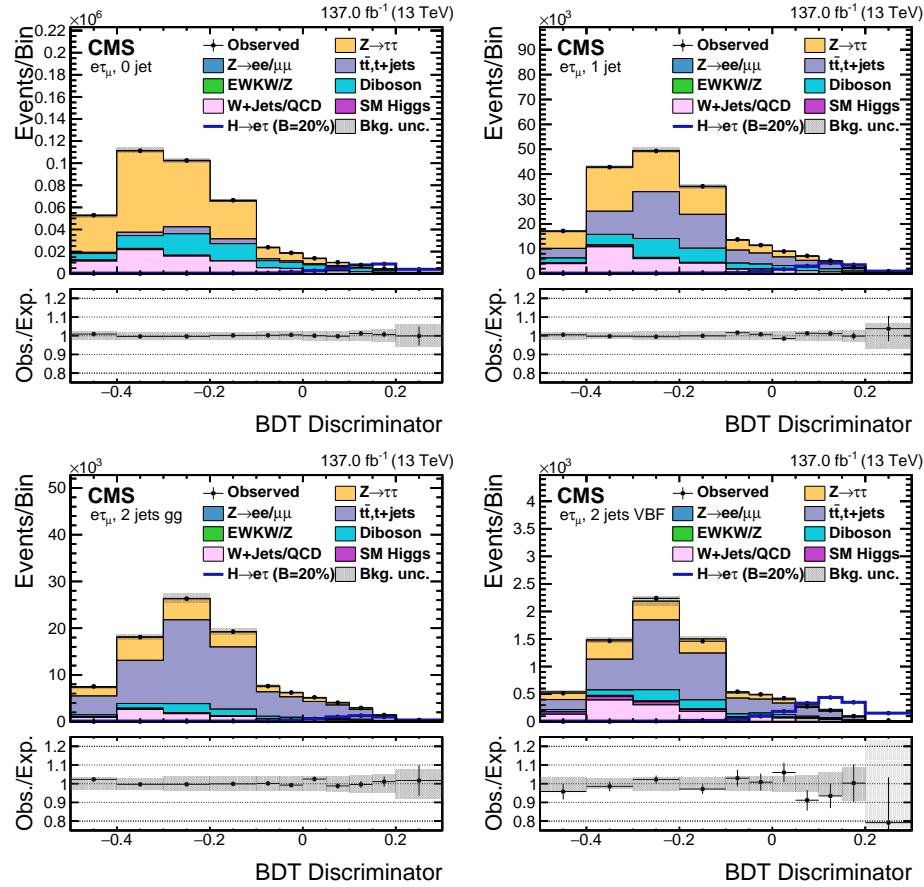


Figure 9.6. BDT discriminator distributions for the observed and estimated background in the $e\tau_\mu$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow e\tau) = 20\%$. The $e\tau_\mu$ channel categories 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 statistical and systematic uncertainties added in quadrature.

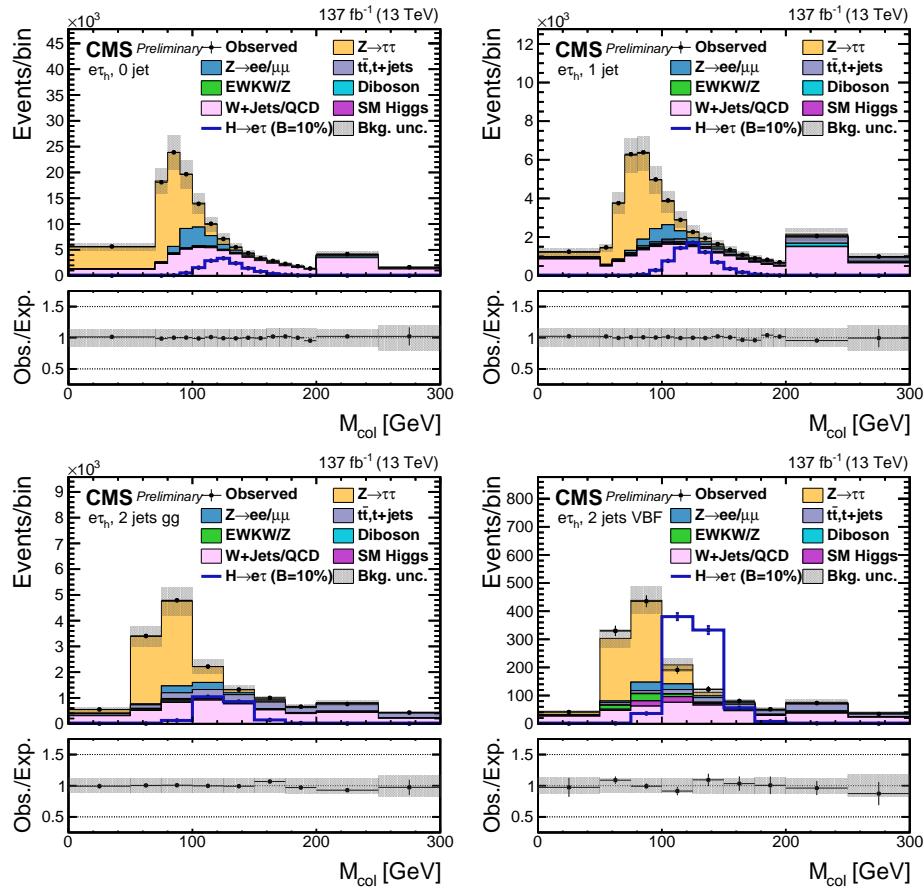


Figure 9.7. M_{col} distributions for the observed and estimated background in the $e\tau_h$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow e\tau) = 10\%$. The $e\tau_h$ channel categories are 0 jets (top 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 statistical and systematic uncertainties added in quadrature.

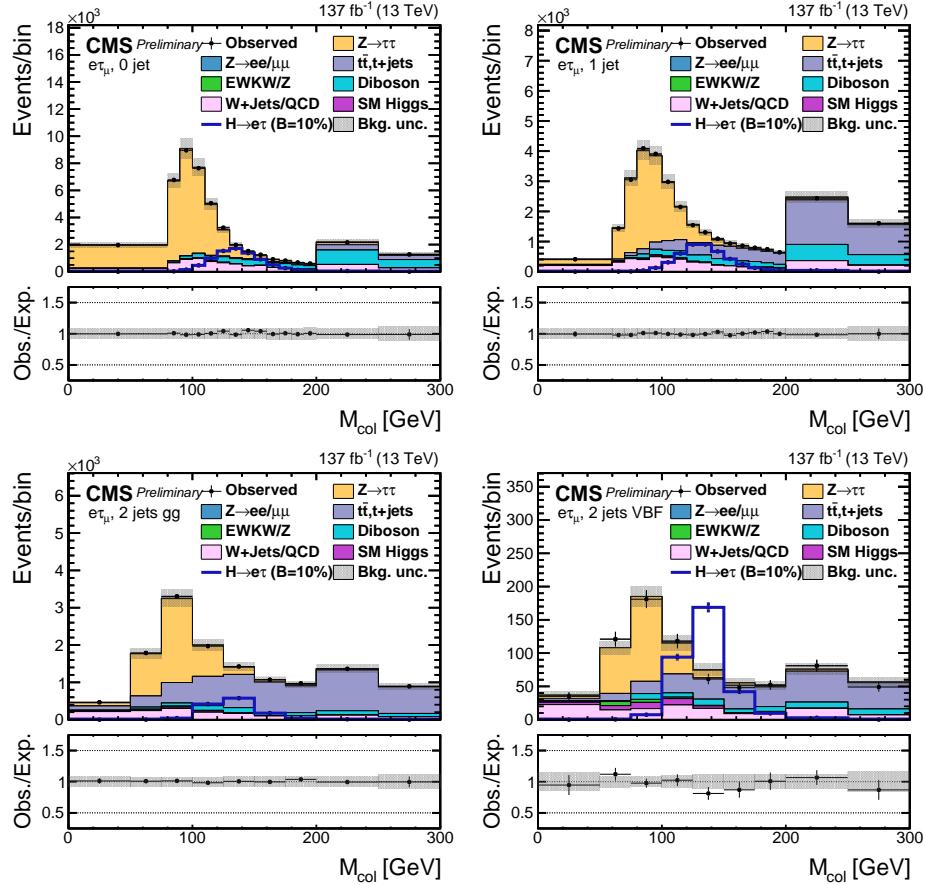


Figure 9.8. M_{col} distributions for the observed and estimated background in the $e\tau_\mu$ process. The background is normalized to the best fit values from the signal plus background fit. Signal corresponds to $\mathcal{B}(H \rightarrow e\tau) = 10\%$. The $e\tau_\mu$ channel categories are 0 jets (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 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 BDT fit analysis is more sensitive than the M_{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 [20]. 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 ℓ^α, ℓ^β are the leptons (e, μ, τ) of different flavors ($\alpha \neq \beta$).

The decay width of LFV Higgs boson decays is

$$\Gamma(H \rightarrow \ell_i \ell_j) = \frac{M_H}{8\pi} (|Y_{ji}|^2 + |Y_{ij}|^2) \quad (9.1)$$

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

$$\mathcal{B}(H \rightarrow \ell_i \ell_j) = \frac{\Gamma(H \rightarrow \ell_i \ell_j)}{\Gamma(H \rightarrow \ell_i \ell_j) + \Gamma_{\text{SM}}} \quad (9.2)$$

where Γ_{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 Γ_{SM} for estimating the branching fractions.

The SM Higgs boson decay width is assumed to be $\Gamma_{\text{SM}} = 4.1$ MeV [102] 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 9.10.

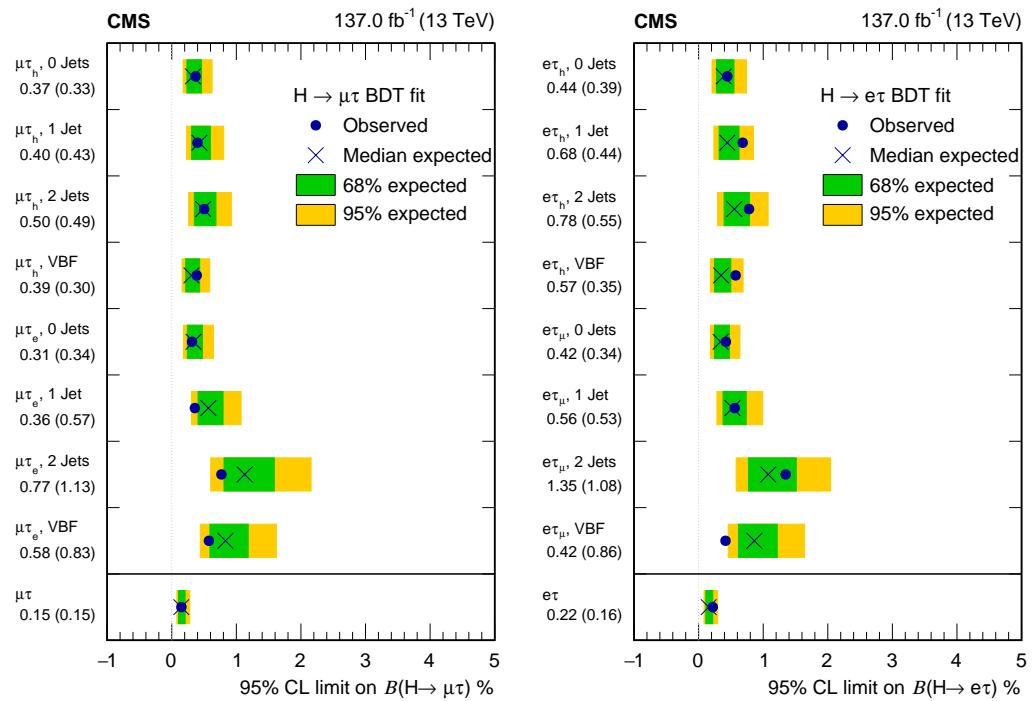


Figure 9.9. Observed (expected) 95% CL upper limits on the $\mathcal{B}(H \rightarrow \mu\tau)$ (left) and $\mathcal{B}(H \rightarrow e\tau)$ (right) for each individual category and combined from the BDT fit analysis.

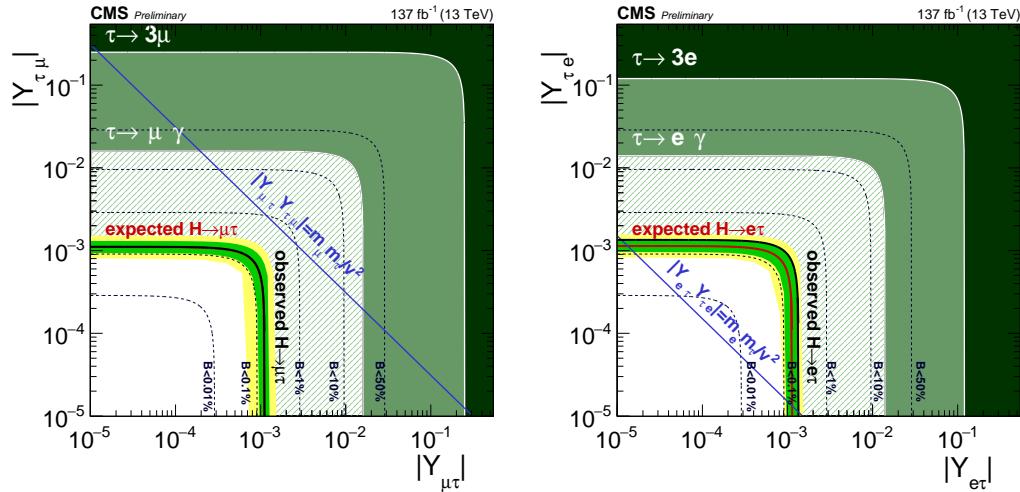


Figure 9.10. Constraints on the LFV Yukawa couplings, $|Y_{\mu\tau}| - |Y_{\tau\mu}|$ (left), and $|Y_{e\tau}| - |Y_{\tau e}|$ (right). The expected (red line) and observed (black solid line) limits are derived from the results shown in Figure 9.9. The flavor-diagonal Yukawa couplings are approximated 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) [103] and $\tau \rightarrow \mu\gamma$ or $\tau \rightarrow e\gamma$ (lighter green) [20]. 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
CATEGORY, AND COMBINED, IN THE $H \rightarrow \mu\tau$ PROCESS FROM
BDT FIT ANALYSIS.

Expected limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	< 0.34	< 0.57	< 1.13	< 0.83	< 0.27
$\mu\tau_h$	< 0.33	< 0.43	< 0.49	< 0.30	< 0.18
$\mu\tau$			< 0.15		
Observed limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	< 0.31	< 0.36	< 0.77	< 0.58	< 0.19
$\mu\tau_h$	< 0.37	< 0.40	< 0.50	< 0.39	< 0.24
$\mu\tau$			< 0.15		
Best fit branching fractions (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	-0.03 ± 0.17	-0.40 ± 0.28	-0.66 ± 0.56	-0.41 ± 0.39	-0.14 ± 0.13
$\mu\tau_h$	0.05 ± 0.17	-0.05 ± 0.22	0.02 ± 0.25	0.10 ± 0.16	0.07 ± 0.09
$\mu\tau$			0.00 ± 0.07		

TABLE 9.2

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

Expected limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$e\tau_\mu$	< 0.34	< 0.53	< 1.08	< 0.86	< 0.26
$e\tau_h$	< 0.39	< 0.44	< 0.55	< 0.35	< 0.20
$e\tau$			< 0.16		
Observed limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$e\tau_\mu$	< 0.42	< 0.56	< 1.35	< 0.42	< 0.22
$e\tau_h$	< 0.44	< 0.68	< 0.78	< 0.57	< 0.37
$e\tau$			< 0.22		
Best fit branching fractions (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$e\tau_\mu$	0.11 ± 0.17	0.04 ± 0.27	0.35 ± 0.55	-1.04 ± 0.44	-0.07 ± 0.13
$e\tau_h$	0.07 ± 0.20	0.29 ± 0.23	0.27 ± 0.29	0.27 ± 0.17	0.20 ± 0.10
$e\tau$			0.08 ± 0.08		

TABLE 9.3

OBSERVED AND EXPECTED UPPER LIMITS AT 95% CL AND BEST
FIT BRANCHING FRACTIONS FOR EACH INDIVIDUAL
CATEGORY, AND COMBINED, IN THE $H \rightarrow \mu\tau$ PROCESS FROM
 M_{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 ± 0.29	0.08 ± 0.42	0.56 ± 0.85	-0.35 ± 0.55	0.23 ± 0.21
$\mu\tau_h$	-0.42 ± 0.32	0.62 ± 0.27	-0.02 ± 0.48	0.29 ± 0.33	0.16 ± 0.17
$\mu\tau$			0.15 ± 0.13		

TABLE 9.4

OBSERVED AND EXPECTED UPPER LIMITS AT 95% CL AND BEST
FIT BRANCHING FRACTIONS FOR EACH INDIVIDUAL
CATEGORY, AND COMBINED, IN THE $H \rightarrow e\tau$ PROCESS FROM M_{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 ± 0.29	0.04 ± 0.40	-0.66 ± 0.84	-1.64 ± 0.68	0.26 ± 0.23
$e\tau_h$	-0.45 ± 0.39	0.04 ± 0.34	0.14 ± 0.53	0.17 ± 0.37	-0.08 ± 0.20
$e\tau$			0.05 ± 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 (%)
$H \rightarrow \mu\tau$	≤ 0.15 (0.15)	0.00 ± 0.07
$H \rightarrow e\tau$	≤ 0.22 (0.16)	0.08 ± 0.08
Yukawa coupling constraints		
$H \rightarrow \mu\tau$		≤ 1.11 (1.10) $\times 10^{-3}$
$H \rightarrow e\tau$		≤ 1.35 (1.14) $\times 10^{-3}$

CHAPTER 10

CONCLUSION

In this dissertation, the search for the LFV decays of the SM Higgs boson 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 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 of High Energy Physics. The observed (median expected) upper limits on $\mathcal{B}(H \rightarrow \mu\tau)$ is 0.15 (0.15) % at 95% CL and on $\mathcal{B}(H \rightarrow e\tau)$ is 0.22 (0.16) % 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 $\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 1.11 \times 10^{-3}$ and $\sqrt{|Y_{e\tau}|^2 + |Y_{\tau e}|^2} < 1.35 \times 10^{-3}$. 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 an improvement over the results published in the previous analysis. The previous analysis's expected limits are projected to the full Run 2 luminosity, which is accomplished by scaling the statistical and systematic uncertainties as inversely proportional to the square root of the luminosity. As can be seen in Figure 10.1, the median expected limits from the current analysis are an improvement

compared to the projections from the previous analysis. Several changes are performed compared to the previous search, and the various details have been explained in this dissertation in detail.

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 previous 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 fb^{-1} in Run 3, which 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 A (“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. In appendix B, an improvement to the fake factor method is presented, and it has the potential to improve the modeling of the misidentified lepton background in hadronic channels and give a better handle over the systematic uncertainties that have a high impact on the median expected limits. Any future iteration of this search is highly recommended to use this mass variable for the signal mass hypothesis and the improved fake factor method to improve the analysis’s sensitivity.

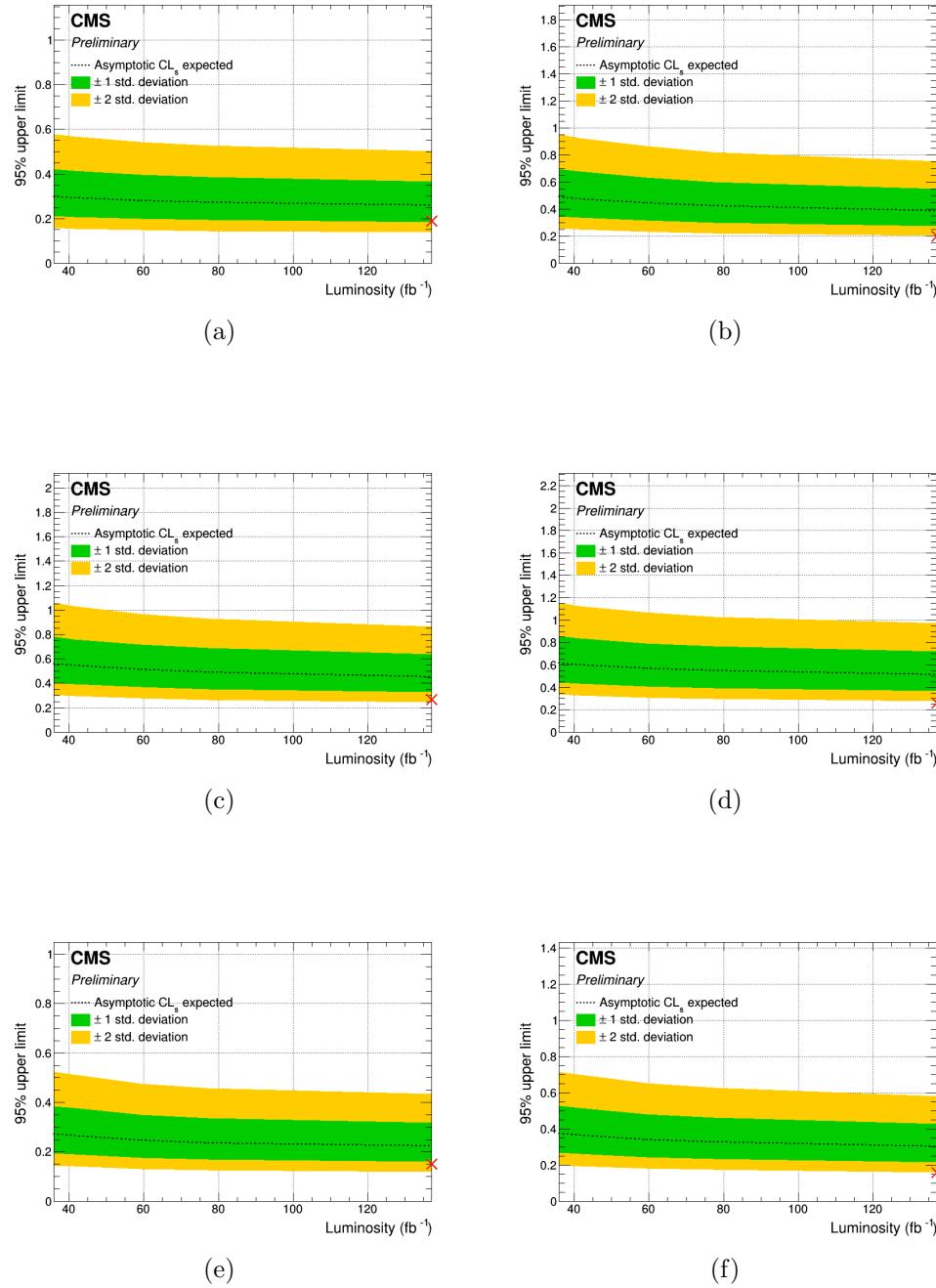


Figure 10.1. Expected limits from the previous analysis projected to the full Run 2 luminosity for (a) $\mu\tau_h$, (b) $e\tau_h$, (c) $\mu\tau_e$, (d) $e\tau_\mu$, (e) $\mu\tau$, and (f) $e\tau$ channels. The red cross represents the median expected limits obtained in the current analysis and are an improvement compared to the projections.

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 (Sensitivity Volume Fit) algorithm can help improve the signal resolution compared to the collinear mass resolution [104]. 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 an improved version of the SVFit algorithm. The SVfit algorithm [105] 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 SM $H \rightarrow \tau\tau$ analysis’s sensitivity for measuring the signal rate by $\approx 40\%$ [106]. 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 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 , η , and ϕ 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. The “Classic” SVFit algorithm was modified to 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, it was modified so that the matrix element of the probability density function defined can reconstruct the Higgs mass $M_{\ell\tau}$. The phase space of the tau decay products along with the angles θ_{inv} and ϕ_{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 ϕ_{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.

$$\begin{aligned} \mathcal{L} (\mathbf{p}^{vis}, p_x^{rec}, p_y^{rec} | M_H) &= \frac{32\pi^4}{s} \int dM_H d\Phi_n |\text{BW}_\tau|^2 |\mathcal{M}_{\tau \rightarrow \dots}(\tilde{\mathbf{p}})|^2 \\ &\quad W(\mathbf{p}^{vis} | \hat{\mathbf{p}}^{vis}) W_{rec}(p_x^{rec}, p_y^{rec} | \hat{p}_x^{rec}, \hat{p}_y^{rec}) \mathcal{F}(\mathbf{p}) \end{aligned} \quad (\text{A.1})$$

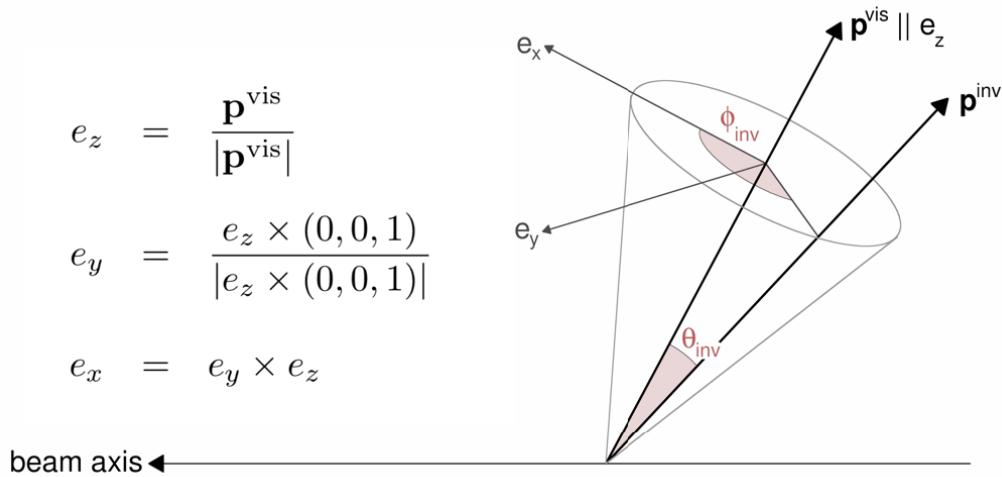


Figure A.1: Illustration of the variables θ_{inv} and ϕ_{inv} that specify the orientation of the \mathbf{p}_{inv} vector relative to the momentum vector \mathbf{p}_{vis} of the visible tau decay products.

The function $\mathcal{F}(\mathbf{p})$ in the integrand may be an arbitrary function of the momenta of the prompt and tau leptons. The integral is evaluated numerically, using a custom implementation of the Markov chain MC integration method with the Metropolis-Hastings algorithm [107]. 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 collinear mass and SVFit mass distributions. The distributions are fit with a double Gaussian, and the goodness of fit as tested from χ^2/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 ~ 129 GeV due to its larger tail, while SVFit mass has a mean of ~ 117 GeV due 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 the study was from the computing point of view. It takes an order of magnitude

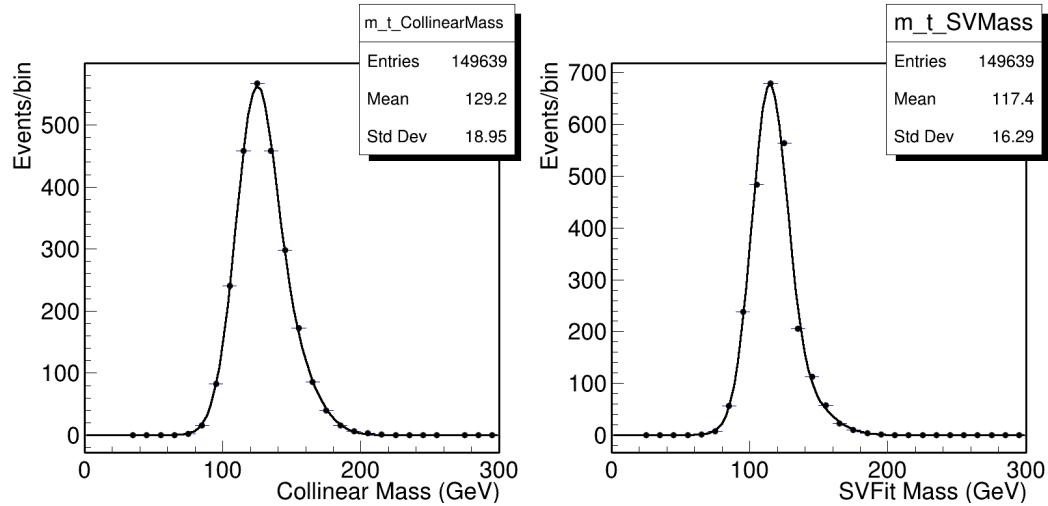


Figure A.2: Collinear mass vs SVFit mass.

longer time to compute the SVFit mass than the collinear mass, which has to be investigated and improved in any future analysis.

APPENDIX B

FAKE FACTOR METHOD

The misidentified lepton background is dominant in hadronic channels, and it is estimated using the fake factor method discussed in chapter 7. The systematic uncertainties associated with estimating the misidentified lepton background are noted to have a high impact on the median expected limits. To recap, we are using a normalization uncertainty of 30% for the misidentified lepton background along with a 10% normalization uncertainty uncorrelated across the categories of each channel. Similarly, the shape uncertainties are estimated in a W boson enriched control region, and these uncertainties have a high impact and are quite constrained by the data. To improve the analysis's sensitivity, we need a better handle on the misidentified lepton background and its associated uncertainties. An improvement to the fake factor method is presented here.

Since the jet flavor composition of different background processes is different, leading to different probabilities of these jets faking τ_h , the fake factor is measured separately for each of the dominant processes. The dominant background processes are W + jets, QCD multijet, and $t\bar{t}$.

First, determination regions are defined, each enriched in one of the above backgrounds. From each determination region, one extracts a dedicated fake factor, which is parametrized by its main dependencies to reduce statistical fluctuation. Further, so-called closure corrections in important variables not used in the fake factor's parametrization are calculated. Contaminations in the different determination regions other than the process of interest are subtracted from data by using MC

simulation and embedding samples. The fake factor is applied on an event by event basis as a weighted sum in the so-called application region. The application region is equivalent to the “background-like” region and signal region is equivalent to the “signal-like” region defined in chapter 7. The weights applied to each fake factor contribution are called fractions and represent the probability of the event being of a particular background process. Fractions are determined from MC simulation and embedding samples. Figure B.1 shows a graphical summary of the fake factor method.

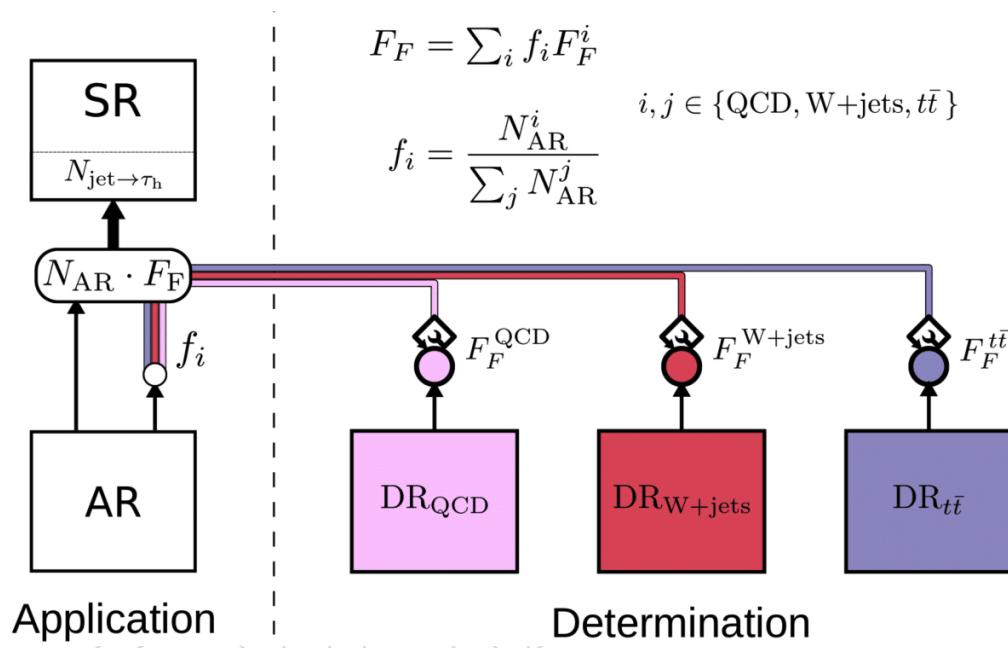


Figure B.1. Illustration of the fake factor method. For each of the processes ($\text{W} + \text{jets}$, QCD , and $t\bar{t}$), a dedicated fake factor is measured in a process enriched determination region. The measured fake factor is corrected and then applied on an event by event basis inside the application region. The application of the fake factor inside the application region also involves the fractions f_i , representing the probabilities of an event in the application region originating from process i .

The basic assumption of the fake factor method is the so-called universality. The fake factor measured in a determination region is the same as if measured in the signal region or any other phase space region. Universality is not strictly given because by measuring the fake factor in a determination region, one introduces biases—for example, the jet-flavor composition changes in the determination region with respect to the signal region. Therefore, the fake factor method also implements process-dependent bias corrections to restore the fake factor universality. In summary, the fake factor is measured in dedicated determination regions and parametrized in variables reflecting its most dominant dependencies. Afterward, corrections are derived for improving the modeling of the jet- τ_h fake contribution in variables not used for the parametrization. A last step bias correction is calculated, accounting for possible differences between the determination region where the fake factor is measured and the application region where the fake factor is applied.

A future search using the full Run 3 data can benefit from using this method for calculating the misidentified lepton background in hadronic channels. The corrections used for improving the modeling and the bias corrections can give a better description of the misidentified lepton background in hadronic channels. This method has been successfully incorporated in the SM $H \rightarrow \tau\tau$ analyses [108].

BIBLIOGRAPHY

1. P.A.R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594:A13, 2016. doi: 10.1051/0004-6361/201525830.
2. Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, 1998. doi: 10.1103/PhysRevLett.81.1562.
3. Q.R. Ahmad et al. Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, 2002. doi: 10.1103/PhysRevLett.89.011301.
4. M. Tanabashi et al. Review of Particle Physics. *Phys. Rev. D*, 98(3):030001, 2018. doi: 10.1103/PhysRevD.98.030001.
5. 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.
6. LHC Machine. *JINST*, 3:S08001, 2008. doi: 10.1088/1748-0221/3/08/S08001.
7. 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.
8. 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.
9. 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.
10. Georges Aad et al. Searches for lepton-flavour-violating decays of the Higgs boson in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector. *Phys. Lett. B*, 800:135069, 2020. doi: 10.1016/j.physletb.2019.135069.
11. J. Goldstone. Field Theories with Superconductor Solutions. *Nuovo Cim.*, 19:154–164, 1961. doi: 10.1007/BF02812722.
12. Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken Symmetries. *Phys. Rev.*, 127:965–970, 1962. doi: 10.1103/PhysRev.127.965.

13. 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.
14. 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.
15. 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.
16. Reinard Primulando and Patipan Uttayarat. Probing Lepton Flavor Violation at the 13 TeV LHC. *JHEP*, 05:055, 2017. doi: 10.1007/JHEP05(2017)055.
17. 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.
18. 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.
19. 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.
20. Roni Harnik, Joachim Kopp, and Jure Zupan. Flavor Violating Higgs Decays. *JHEP*, 03:026, 2013. doi: 10.1007/JHEP03(2013)026.
21. 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.
22. Vardan Khachatryan et al. Search for lepton flavour violating decays of the Higgs boson to $e\tau$ and $e\mu$ in proton–proton collisions at $\sqrt{s} = 8$ TeV. *Phys. Lett. B*, 763:472–500, 2016. doi: 10.1016/j.physletb.2016.09.062.
23. Georges Aad et al. Search for lepton-flavour-violating decays of the Higgs and Z bosons with the ATLAS detector. *Eur. Phys. J. C*, 77(2):70, 2017. doi: 10.1140/epjc/s10052-017-4624-0.
24. Georges Aad et al. Search for lepton-flavour-violating $H \rightarrow \mu\tau$ decays of the Higgs boson with the ATLAS detector. *JHEP*, 11:211, 2015. doi: 10.1007/JHEP11(2015)211.
25. 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.

26. S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.
27. The Large Electron-Positron Collider. Jul 2012. URL <https://cds.cern.ch/record/1997351>.
28. Linear accelerator 2. Sep 2012. URL <https://cds.cern.ch/record/1997427>.
29. The Proton Synchrotron Booster. Jul 2012. URL <https://cds.cern.ch/record/1997372>.
30. The Proton Synchrotron. Jan 2012. URL <https://cds.cern.ch/record/1997189>.
31. The Super Proton Synchrotron. Jan 2012. URL <https://cds.cern.ch/record/1997188>.
32. 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.
33. K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008. doi: 10.1088/1748-0221/3/08/S08002.
34. Jr. Alves, A. Augusto et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. doi: 10.1088/1748-0221/3/08/S08005.
35. 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.
36. 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.
37. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
38. 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.
39. 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.
40. Electron reconstruction and identification at $\sqrt{s} = 7$ TeV. 2010.
41. 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.
42. N. Almeida et al. Data filtering in the readout of the CMS electromagnetic calorimeter. *JINST*, 3:P02011, 2008. doi: 10.1088/1748-0221/3/02/P02011.

43. Prasanna Kumar Siddireddy. The CMS ECAL Trigger and DAQ system: electronics auto-recovery and monitoring. 6 2018.
44. G. L. Bayatian et al. CMS Physics: Technical Design Report Volume 1: Detector Performance and Software. 2006.
45. CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV. 5 2019.
46. CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV. 6 2018.
47. CMS Luminosity Measurements for the 2016 Data Taking Period. 3 2017.
48. Joel N. Butler and Tommaso Tabarelli de Fatis. A MIP Timing Detector for the CMS Phase-2 Upgrade. 2019.
49. Technical Proposal for the Phase-II Upgrade of the CMS Detector. 6 2015.
50. 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.
51. 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.
52. 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.
53. 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.
54. 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.
55. 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.
56. 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.

57. 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.
58. 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.
59. Rikkert Frederix and Stefano Frixione. Merging meets matching in MC@NLO. *JHEP*, 12:061, 2012. doi: 10.1007/JHEP12(2012)061.
60. S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250–303, 2003. doi: 10.1016/S0168-9002(03)01368-8.
61. Vardan Khachatryan et al. Event generator tunes obtained from underlying event and multiparton scattering measurements. *Eur. Phys. J. C*, 76(3):155, 2016. doi: 10.1140/epjc/s10052-016-3988-x.
62. Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements. 8 2018.
63. Richard D. Ball et al. Parton distributions from high-precision collider data. *Eur. Phys. J. C*, 77(10):663, 2017. doi: 10.1140/epjc/s10052-017-5199-5.
64. H.M. Georgi, S.L. Glashow, M.E. Machacek, and Dimitri V. Nanopoulos. Higgs Bosons from Two Gluon Annihilation in Proton Proton Collisions. *Phys. Rev. Lett.*, 40:692, 1978. doi: 10.1103/PhysRevLett.40.692.
65. Robert N. Cahn, Stephen D. Ellis, Ronald Kleiss, and W.James Stirling. Transverse Momentum Signatures for Heavy Higgs Bosons. *Phys. Rev. D*, 35:1626, 1987. doi: 10.1103/PhysRevD.35.1626.
66. S.L. Glashow, Dimitri V. Nanopoulos, and A. Yildiz. Associated Production of Higgs Bosons and Z Particles. *Phys. Rev. D*, 18:1724–1727, 1978. doi: 10.1103/PhysRevD.18.1724.
67. G. Heinrich, S.P. Jones, M. Kerner, G. Luisoni, and E. Vryonidou. NLO predictions for Higgs boson pair production with full top quark mass dependence matched to parton showers. *JHEP*, 08:088, 2017. doi: 10.1007/JHEP08(2017)088.
68. G. Buchalla, M. Capozi, A. Celis, G. Heinrich, and L. Scyboz. Higgs boson pair production in non-linear Effective Field Theory with full m_t -dependence at NLO QCD. *JHEP*, 09:057, 2018. doi: 10.1007/JHEP09(2018)057.
69. Johan Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J. C*, 53: 473–500, 2008. doi: 10.1140/epjc/s10052-007-0490-5.

70. 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.
71. 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.
72. 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.
73. R. Fruhwirth, W. Waltenberger, and P. Vanlaer. Adaptive vertex fitting. *J. Phys. G*, 34:N343, 2007. doi: 10.1088/0954-3899/34/12/N01.
74. 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.
75. CMS Muon Physics Object Group. Muon efficiency scale factors. Website. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsMUO>.
76. A.M. Sirunyan et al. Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV. *JINST*, 13(10):P10005, 2018. doi: 10.1088/1748-0221/13/10/P10005.
77. 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.
78. Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J. C*, 72:1896, 2012. doi: 10.1140/epjc/s10052-012-1896-2.
79. 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.
80. 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.
81. Jet algorithms performance in 13 TeV data. 3 2017.
82. A.M. Sirunyan et al. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *JINST*, 13(05):P05011, 2018. doi: 10.1088/1748-0221/13/05/P05011.
83. 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.

84. 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.
85. 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.
86. Albert M Sirunyan et al. Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector. *Phys. Lett. B*, 779:283–316, 2018. doi: 10.1016/j.physletb.2018.02.004.
87. Andreas Hocker et al. TMVA - Toolkit for Multivariate Data Analysis. 3 2007.
88. 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.
89. Aaron Roodman. Blind analysis in particle physics. *eConf*, C030908:TUIT001, 2003.
90. 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.
91. Procedure for the LHC Higgs boson search combination in summer 2011. 8 2011.
92. 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.
93. Alexander L. Read. Modified frequentist analysis of search results (The CL(s) method). In *Workshop on Confidence Limits*, pages 81–101, 8 2000.
94. 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.
95. J.S. Conway. Incorporating Nuisance Parameters in Likelihoods for Multi-source Spectra. In *PHYSTAT 2011*, pages 115–120, 2011. doi: 10.5170/CERN-2011-006.115.
96. Alexander L. Read. Linear interpolation of histograms. *Nucl. Instrum. Meth. A*, 425:357–360, 1999. doi: 10.1016/S0168-9002(98)01347-3.
97. 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.
98. Vardan Khachatryan et al. Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS. *JINST*, 11(01):P01019, 2016. doi: 10.1088/1748-0221/11/01/P01019.

99. Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2. 7 2016.
100. 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.
101. 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.
102. 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.
103. 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.
104. Lorenzo Bianchini, Betty Calpas, John Conway, Andrew Fowlie, Luca Marzola, Christian Veelken, and Lucia Perrini. Reconstruction of the Higgs mass in events with Higgs bosons decaying into a pair of τ leptons using matrix element techniques. *Nucl. Instrum. Meth. A*, 862:54–84, 2017. doi: 10.1016/j.nima.2017.05.001.
105. Lorenzo Bianchini, John Conway, Evan Klose Friis, and Christian Veelken. Reconstruction of the Higgs mass in $H \rightarrow \tau\tau$ Events by Dynamical Likelihood techniques. *J. Phys. Conf. Ser.*, 513:022035, 2014. doi: 10.1088/1742-6596/513/2/022035.
106. Serguei Chatrchyan et al. Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons. *JHEP*, 05:104, 2014. doi: 10.1007/JHEP05(2014)104.
107. W.K. Hastings. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 57:97–109, 1970. doi: 10.1093/biomet/57.1.97.
108. Measurement of Higgs boson production and decay to the $\tau\tau$ final state. 3 2019.