

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 `NDdiss2 $\epsilon$`  v3.2017.2 (2017/05/09) on August 19, 2020 for

Prasanna Kumar Siddireddy

This  $\text{\LaTeX} 2_{\epsilon}$  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 $\epsilon$`  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 $\epsilon$`  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

August 2020

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

Abstract

by

Prasanna Kumar Siddireddy

This dissertation presents searches for Lepton Flavor Violating decay of the Standard Model Higgs Boson ( $H$ ) into a muon and a tau lepton or to an electron and a tau lepton. Data collected by the CMS detector in 2016-2018, in proton-proton collisions at the LHC, at a center-of-mass energy of 13 TeV was used to perform the search. Observed (expected) upper limits on the branching fraction of  $H$  decaying into a muon and a tau lepton,  $\mathcal{B}(H \rightarrow \mu\tau)$ , was set at  $XX(0.13)\%$  at 95% CL and branching fraction of  $H$  decaying into an electron and a tau lepton,  $\mathcal{B}(H \rightarrow e\tau)$ , was set at  $XX(0.15)\%$  at 95% CL. These are the most stringent limits set on these processes till date.

## CONTENTS

Figures . . . . .	iii
Tables . . . . .	v
Chapter 1: Event Selection . . . . .	1
1.1 Introduction . . . . .	1
1.2 Boosted Decision Tree . . . . .	7
1.3 $H \rightarrow \mu\tau_h$ channel . . . . .	9
1.4 $H \rightarrow \mu\tau_e$ channel . . . . .	13
1.5 $H \rightarrow e\tau_h$ channel . . . . .	14
1.6 $H \rightarrow e\tau_\mu$ channel . . . . .	18
Chapter 2: Background Estimation . . . . .	23
2.1 Introduction . . . . .	23
2.2 Embedding technique . . . . .	25
2.3 Misidentified lepton background . . . . .	27
2.3.1 Fully data-driven approach . . . . .	27
2.3.2 Semi data-driven approach . . . . .	34
2.4 MC Simulation . . . . .	40
Bibliography . . . . .	42

## FIGURES

1.1	Feynman diagrams of lepton-flavor violating Higgs-boson decays. The first row shows diagrams for the Higgs boson coupling to $\mu\tau$ (a,b). Couplings to $e\tau$ (c,d) are shown in the second row. Taus can decay leptonically or hadronically. Feynman diagrams are shown for the leptonic decay of taus (b,d) and the hadronic decay of taus (a,c).	2
1.2	Illustration of the differences in $d\phi(\ell = \tau \text{ or } e, MET)$ and $p_T^\mu$ spectrums in LFV and SM $H \rightarrow \tau\tau$ processes.	3
1.3	Illustration of the differences in $d\phi(\ell = \tau \text{ or } \mu, MET)$ and $p_T^e$ spectrums in LFV and SM $H \rightarrow \tau\tau$ processes.	4
1.4	Estimation of the neutrino momentum $\cancel{E}_{T,proj}$ by using the component of the missing transverse energy $\cancel{E}_T$ which is collinear to the visible decay products of tau in the transverse plane.	6
1.5	Illustration of decision tree.	7
1.6	Distribution of the input variables to the BDT for the $H \rightarrow \mu\tau_h$ process.	11
1.7	Overtraining check as performed in TMVA for the trained BDT in $H \rightarrow \mu\tau_h$ channel for 2016 (a), 2017 (b), and 2018 (c).	13
1.8	Distribution of the input variables to the BDT for the $H \rightarrow \mu\tau_e$ process.	15
1.9	Overtraining check as performed in TMVA for the trained BDT in $H \rightarrow \mu\tau_e$ channel for 2016 (a), 2017 (b), and 2018 (c).	16
1.10	Distribution of the input variables to the BDT for the $H \rightarrow e\tau_h$ process.	17
1.11	Overtraining check as performed in TMVA for the trained BDT in $H \rightarrow e\tau_h$ channel for 2016 (a), 2017 (b), and 2018 (c).	18
1.12	Distribution of the input variables to the BDT for the $H \rightarrow e\tau_\mu$ process.	21
1.13	Overtraining check as performed in TMVA for the trained BDT in $H \rightarrow e\tau_\mu$ channel for 2016 (a), 2017 (b), and 2018 (c).	22
2.1	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) $W\gamma^{(*)}$ .	24
2.2	Schematic of Embedding Technique	26
2.3	Distributions of $M_{col}$ discriminator in the $Z \rightarrow \tau\tau$ control regions for the (a) $H \rightarrow \mu\tau_h$ , (b) $H \rightarrow \mu\tau_e$ , (c) $H \rightarrow e\tau_h$ , and (d) $H \rightarrow e\tau_\mu$ channels.	28

2.4	Distributions of BDT discriminator in the $Z \rightarrow \tau\tau$ control regions for the (a) $H \rightarrow \mu\tau_h$ , (b) $H \rightarrow \mu\tau_e$ , (c) $H \rightarrow e\tau_h$ , and (d) $H \rightarrow e\tau_\mu$ channels.	29
2.5	Signal region (green) contrasted with the control regions used for estimating the misidentified background . . . . .	30
2.6	Fit performed to $\tau_h$ misidentification rates for $H \rightarrow \mu\tau_h$ (a) and $H \rightarrow e\tau_h$ (b) channel as a function of $\tau_h$ $p_T$ for the different years. The misidentification rates used are further parametrized based on $\tau_h$ Decay Mode along with the pseudorapidity of $\tau_h$ . However, here only the inclusive misidentification rates are shown. The misidentification rates are labeled as “tight-to-loose” to clarify that they are calculated as a ratio of the number of events passing the tight WP to the loose WP of DNN discrimination against jets. . . . .	32
2.7	Fit performed to the $\mu$ (a) and $e$ (b) misidentification rates as a function of their $p_T$ for 2016 (Left), 2017 (Center), and 2018 (Right). The misidentification rates are labeled as “tight-to-loose” to clarify that they are calculated as a ratio of the number of events passing the tight isolation to the loose isolation. The hyperbolic tangent function is used for performing the fit. . . . .	33
2.8	Distributions of $M_{\text{col}}$ discriminator in the same-sign (Left) and W boson enriched (Right) control regions for the $H \rightarrow \mu\tau_h$ (top) and $H \rightarrow e\tau_h$ (bottom) channels. . . . .	35
2.9	Distributions of BDT discriminator in the same-sign (Left) and W boson enriched (Right) control regions for the $H \rightarrow \mu\tau_h$ (top) and $H \rightarrow e\tau_h$ (bottom) channels. . . . .	36
2.10	QCD OS/SS extrapolation factors in events with 0 Jets (Left), 1 Jet (Center), and 2 Jets (Right) for 2016 (a), 2017 (b), 2018 (c). The line is the best fit, and the shaded region corresponds to the shape uncertainties. . . . .	38
2.11	(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. . . . .	39
2.12	Distribution of $M_{\text{col}}$ discriminator in the muon anti-isolated control regions for the $H \rightarrow \mu\tau_e$ channel. . . . .	40
2.13	Distributions of $M_{\text{col}}$ (BDT) discriminator in $t\bar{t}$ enriched control region for $H \rightarrow \mu\tau_e$ and $H \rightarrow e\tau_\mu$ channel. . . . .	41

## TABLES

1.1	Event selection criteria for the kinematic variables for the $H \rightarrow \mu\tau$ channels . . . . .	12
1.2	Event selection criteria for the kinematic variables for the $H \rightarrow e\tau$ channels . . . . .	19

## CHAPTER 1

## EVENT SELECTION

## 1.1 Introduction

This chapter summarizes the event selection criteria for the analysis. The signal topology consists of an isolated lepton,  $\mu$  or  $e$ , along with an oppositely charged isolated tau lepton ( $\tau_\mu$ ,  $\tau_e$ , or  $\tau_h$ ). Jets misidentified as electrons or muons are suppressed by imposing isolation requirements. The events are first categorized into  $\mu\tau$  and  $e\tau$  and then further divided into leptonic and hadronic channels based on tau decay mode. Figure 1.1 shows the corresponding Feynman diagrams for the LFV  $H \rightarrow \mu\tau$  and  $H \rightarrow e\tau$  decays.

The final states of this analysis are similar to the  $H \rightarrow \tau\tau$  decay allowed by the SM and since been observed [1]. However, there are some significant kinematic differences. The LFV  $H \rightarrow \mu\tau_h$  and  $H \rightarrow \mu\tau_e$  ( $H \rightarrow e\tau_h$  and  $H \rightarrow e\tau_\mu$ ) decays consist of a muon (an electron) that comes directly from the Higgs and has a hard  $p_T$  spectrum, along with a hadronically decaying tau or a softer electron (muon) that comes from the tau lepton of opposite sign charge, and missing transverse momentum from the tau decay. Also, there are fewer neutrinos in LFV decays, coming from the decay of the single  $\tau$ . The decay products of this highly boosted tau are closely aligned, leading to a narrow separation between the visible decay products of the tau and the  $\vec{p}_T^{\text{miss}}$  in the azimuthal plane. The same is not true in the  $H \rightarrow \tau\tau$  decays. These differences are illustrated pictorially in Figures 1.2 and 1.3.

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



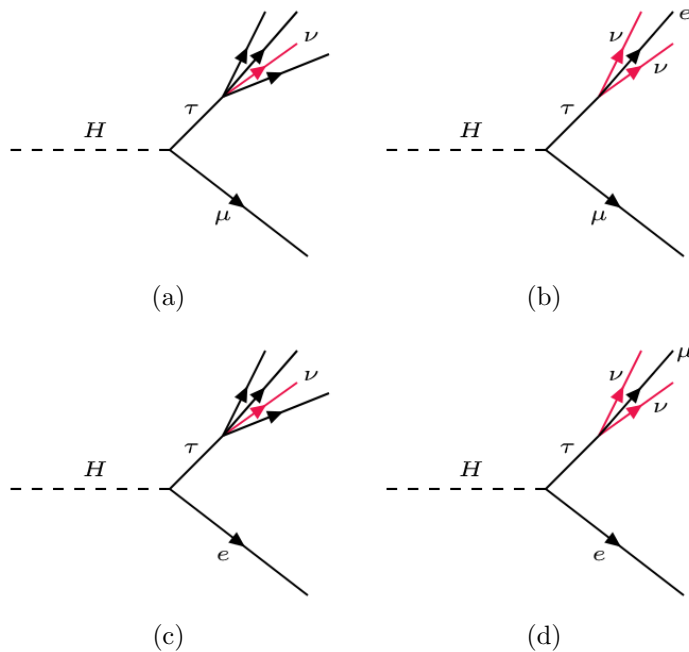


Figure 1.1. Feynman diagrams of lepton-flavor violating Higgs-boson decays. The first row shows diagrams for the Higgs boson coupling to  $\mu\tau$  (a,b). Couplings to  $e\tau$  (c,d) are shown in the second row. Taus can decay leptonically or hadronically. Feynman diagrams are shown for the leptonic decay of taus (b,d) and the hadronic decay of taus (a,c).

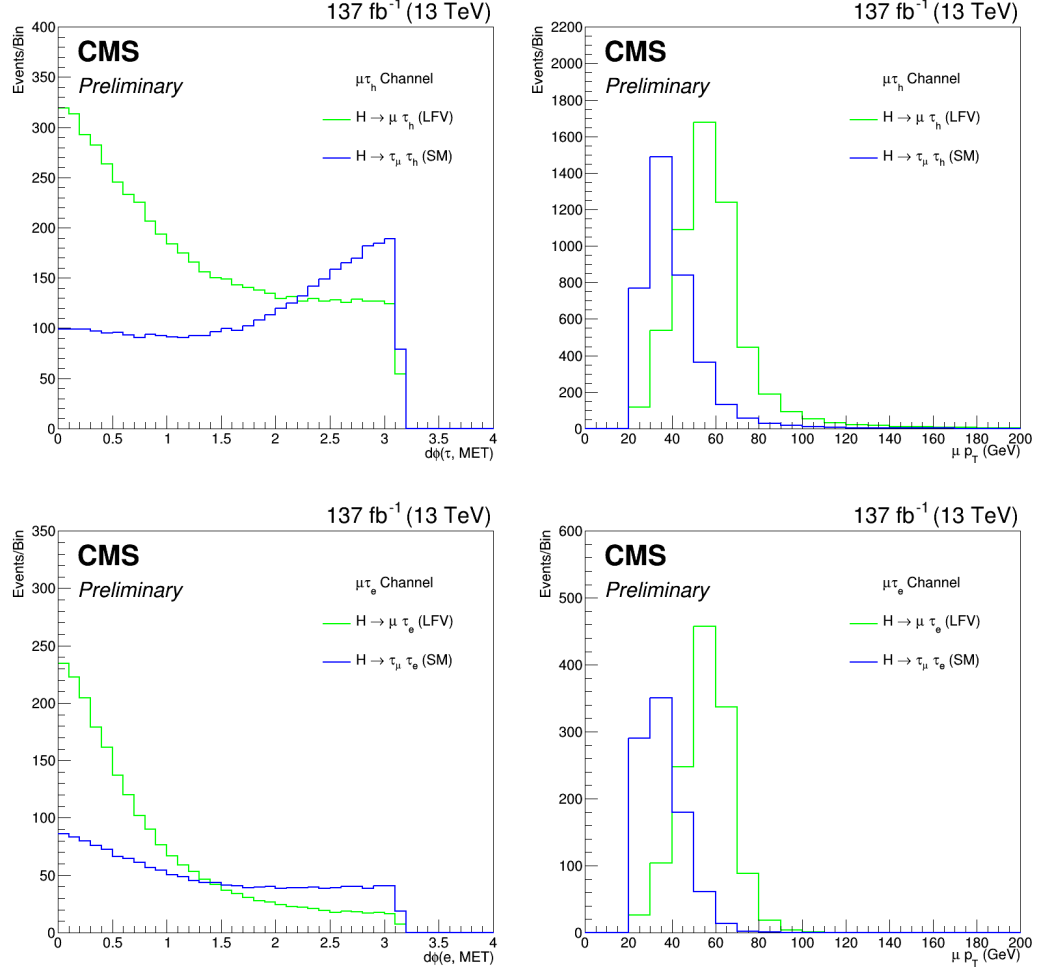


Figure 1.2. Illustration of the differences in  $d\phi(\ell = \tau \text{ or } e, MET)$  and  $p_T^\mu$  spectrums in LFV and SM  $H \rightarrow \tau\tau$  processes.

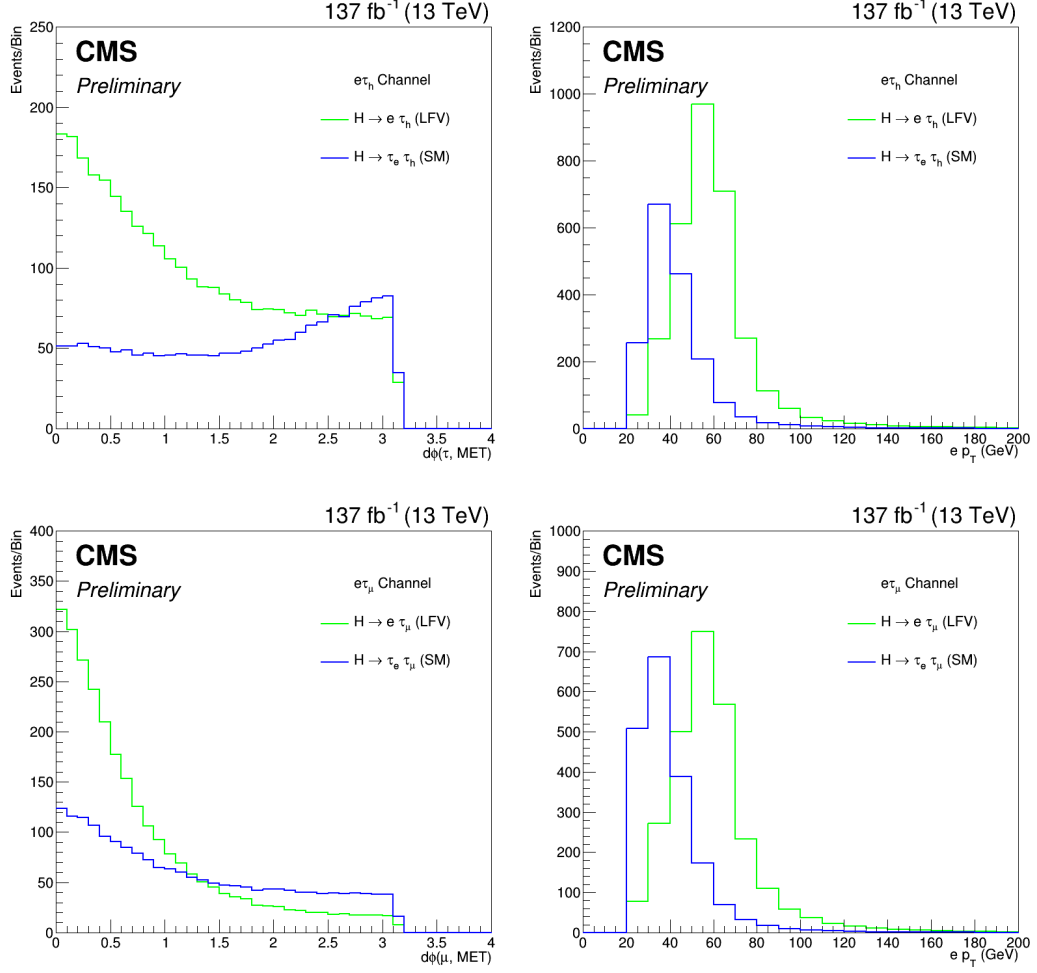


Figure 1.3. Illustration of the differences in  $d\phi(\ell = \tau \text{ or } \mu, MET)$  and  $p_T^e$  spectrums in LFV and SM  $H \rightarrow \tau\tau$  processes.

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

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

To better discriminate between signal and background events, a Boosted Decision Trees (BDT) discriminator is trained using simulated events, using the TMVA tool of the ROOT analysis package [2]. After applying preselection, a binned likelihood is used to fit the distribution of a BDT discriminator for the signal and the background contributions, and we call this the BDT fit method. The collinear mass ( $M_{\text{col}}$ ) and the transverse mass ( $M_T(\ell)$ ) that are used as input variables to the BDT are defined in the following paragraphs. A brief description of the BDT is given in the next section.

The  $M_{\text{col}}$  provides an estimate of  $m_H$  using the observed decay products of the Higgs boson candidate. It is reconstructed using the collinear approximation based on the observation that, since  $m_H \gg m_\tau$ , the  $\tau$  lepton decay products are highly Lorentz boosted in the direction of the  $\tau$  candidate [3]. The momentum of the neutrino coming from  $\tau$  decay can be approximated to have the same direction as the visible decay products of the  $\tau(\bar{\tau}^{\text{vis}})$ . Figure 1.4 shows the corresponding Feynman diagram for the  $\cancel{E}_T$  projected in the direction of the visible decay products of tau.

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

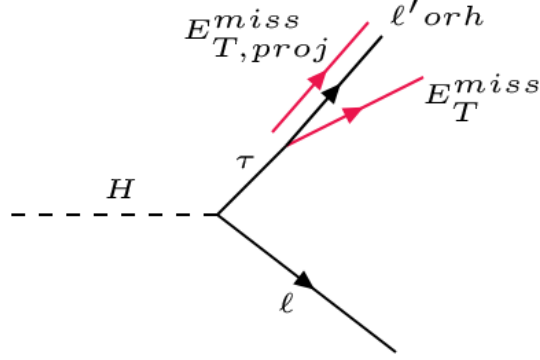


Figure 1.4. Estimation of the neutrino momentum  $\cancel{E}_{T,proj}$  by using the component of the missing transverse energy  $\cancel{E}_T$  which is collinear to the visible decay products of tau in the transverse plane.

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

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

An alternate analysis has been implemented to cross-check the results obtained from the BDT fit method. This approach involves placing requirements on several kinematic variables and then using the resulting distribution of  $M_{\text{col}}$  as a discriminant for a binned likelihood fit. Henceforth, we call this the  $M_{\text{col}}$  fit method. The BDT and  $M_{\text{col}}$  fit methods were performed blinded in the signal region [4]. The selection criterion described was developed without looking at the data in the region where the signal is expected. This approach is standard in particle physics analysis and

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

## 1.2 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 1.5. The starting point of the tree structure is called a root node containing all the events we want to classify. A sequence of binary splits is made using conditions on the input variables provided to the classifier. The variable which ensures the separation of the signal and the background is used for each split.

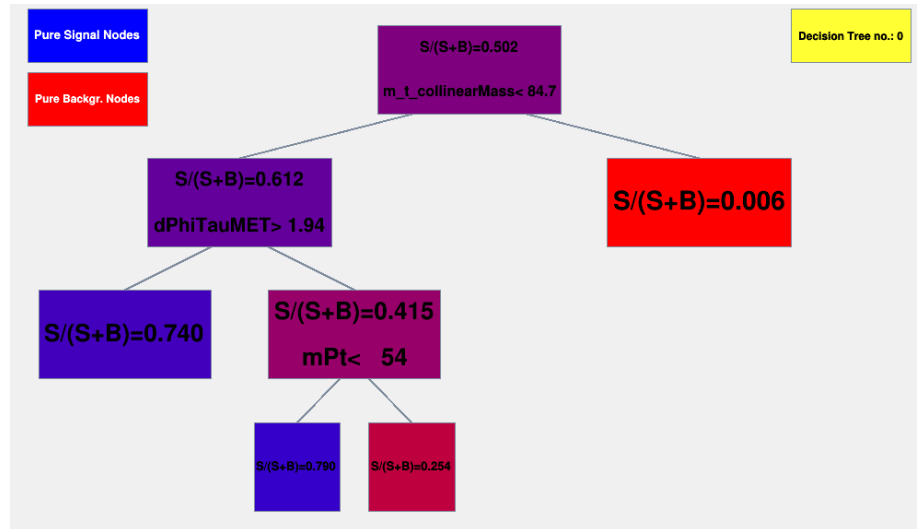


Figure 1.5. Illustration of decision tree.

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

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

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

$$\text{err}_m = \frac{\sum_{i=1}^N w_i I(y_i \neq DT_m(x_i))}{\sum_{i=1}^N w_i} \quad (1.1)$$

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

$$\alpha_m = \beta \times \ln((1 - \text{err}_m) / \text{err}_m) \quad (1.2)$$

$$w_i \rightarrow w_i \times e^{\alpha I(y_i \neq DT_m(x_i))} \quad (1.3)$$

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 reweighting. Event weights in each tree are renormalized to keep the summed weights constant. After the boosting and training processes, the final score of each event is  $DT(x)$ . A high score indicates a signal like event while a low score indicates a background like event.

$$DT(x) = \sum_{m=1}^{N_{tree}} \alpha_m DT_m(x) \quad (1.4)$$

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.

### 1.3 $H \rightarrow \mu\tau_h$ channel

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



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

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

A BDT is trained after applying preselection criteria. The signal training sample considered is a mixture of simulated GGF and VBF events, weighted according to their respective SM production cross-sections. The misidentified lepton background and  $Z \rightarrow \tau\tau$  background are the dominant backgrounds in this channel. The background used for training the BDT is a data sample of misidentified lepton events with the same charge assignment for both leptons along with the Drell-Yan MC sample with signal selections. The input variables to the BDT are  $p_T^\mu, p_T^{\tau_h}, M_{\text{col}}, \vec{p}_T^{\text{miss}}, M_T(\tau_h, \vec{p}_T^{\text{miss}}), \Delta\eta(\mu, \tau_h), \Delta\phi(\mu, \tau_h)$ , and  $\Delta\phi(\tau_h, \vec{p}_T^{\text{miss}})$ . The distribution of the input variables to the BDT can be seen in Figure 1.6.

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

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

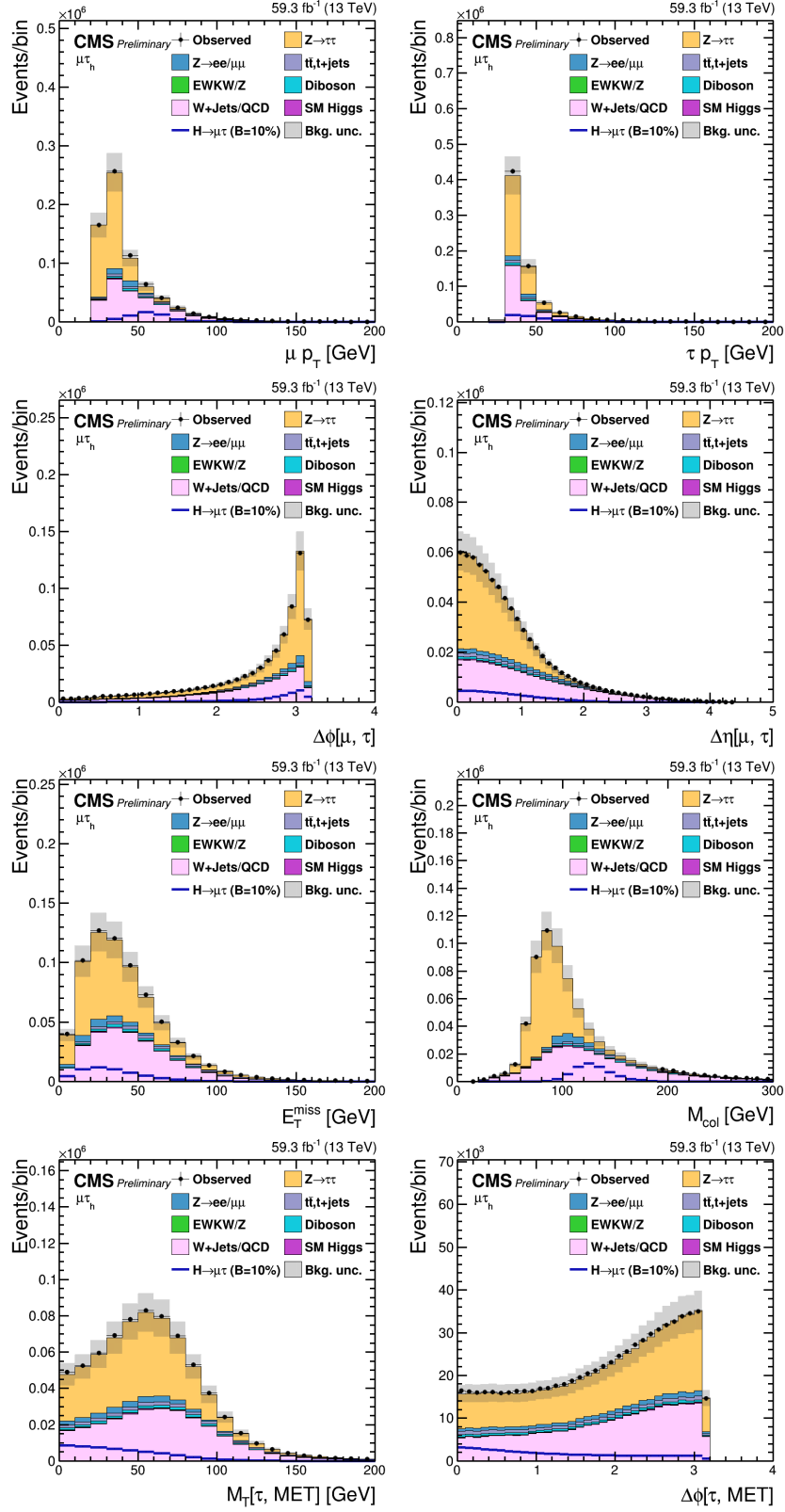


Figure 1.6. Distribution of the input variables to the BDT for the  $H \rightarrow \mu\tau_h$  process.

TABLE 1.1

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

Variable	$\text{H} \rightarrow \mu\tau_{\text{h}}$				$\text{H} \rightarrow \mu\tau_e$			
$p_{\text{T}}^{\text{e}}$	—				$> 13$			
$p_{\text{T}}^{\mu}$	$> 26$				$> 24$			
$p_{\text{T}}^{\tau_{\text{h}}}$	$> 30$				—			
$ \eta ^{\text{e}}$	—				$< 2.5$			
$ \eta ^{\mu}$	$< 2.1$				$< 2.4$			
$ \eta ^{\tau_{\text{h}}}$	$< 2.3$				—			
$I_{\text{rel}}^{\text{e}}$	—				$< 0.1$			
$I_{\text{rel}}^{\mu}$	$< 0.15$				$< 0.15$			
$I_{\text{rel}}^{\tau_{\text{h}}}$	DNN $\tau_{\text{h}}$ ID				—			
Trigger	$\mu(24)$ (all years)				$\text{e}(12)$ and $\mu(23)$ (all years)			
	$M_{\text{col}}$ fit selection							
	0-jet	1-jet	2-jet		0-jet	1-jet	2-jet	
			GGF	VBF			GGF	VBF
$M_{jj}$	—	—	$< 550$	$\geq 550$	—	—	$< 550$	$\geq 550$
$p_{\text{T}}^{\mu}$			—		$> 30$	$> 26$	$> 26$	$> 26$
$M_{\text{T}}(\mu)$			—		$> 60$	$> 40$	$> 15$	$> 15$
$M_{\text{T}}(\tau_{\text{h}})$	$< 105$	$< 105$	$< 105$	$< 85$			—	
$\Delta\phi(\text{e}, \vec{p}_{\text{T}}^{\text{miss}})$			—		$< 0.7$	$< 0.7$	$< 0.5$	$< 0.3$
$\Delta\phi(\text{e}, \mu)$			—		$> 2.5$	$> 1.0$	—	—

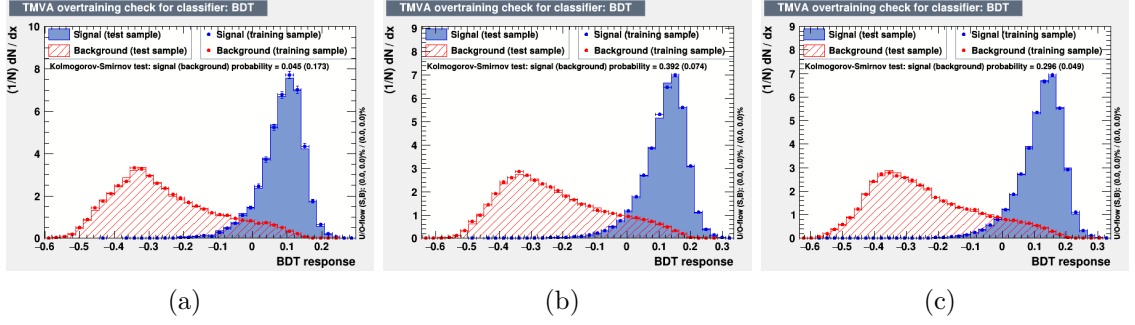


Figure 1.7. Overtraining check as performed in TMVA for the trained BDT in  $H \rightarrow \mu\tau_h$  channel for 2016 (a), 2017 (b), and 2018 (c).

#### 1.4 $H \rightarrow \mu\tau_e$ 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 impact parameter to the primary vertex. However, this constraint is not present in the initial 2016 data samples and 2016 MC samples. In addition to the event passing the trigger, the reconstructed leptons corresponding to the trigger have to match the HLT objects within  $\Delta R < 0.5$ .

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

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

criteria. The signal training is done in a similar way while for background training dominant contributors,  $t\bar{t}$  and  $Z \rightarrow \ell\ell$  ( $\ell = e, \mu, \tau$ ) events are mixed and weighted by their respective production cross-sections. The  $t\bar{t}$  process contributes dominantly for the 2-jet category, with significant contribution to 1-jet category.  $Z \rightarrow \ell\ell$  background processes dominantly contribute the 0- and 1-jet categories. The QCD multijet background has the third-largest contribution, so we use the same sign control region in data as additional background for training. The input variables to the BDT are:  $p_T^\mu, p_T^e, M_{\text{col}}, M_T(\mu, \vec{p}_T^{\text{miss}}), M_T(e, \vec{p}_T^{\text{miss}}), \Delta\phi(e, \mu), \Delta\phi(\mu, \vec{p}_T^{\text{miss}})$ , and  $\Delta\phi(e, \vec{p}_T^{\text{miss}})$ . The distribution of the input variables to the BDT can be seen in Figure 1.8. The BDT discriminator distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow \mu\tau_e$  channel, are shown in results chapter.

In the  $M_{\text{col}}$  fit method, additional selection criteria require a stringent selection on muons,  $p_T^\mu > 30 \text{ GeV}$  for 0-jet category and  $p_T^\mu > 26 \text{ GeV}$  in rest of the categories. The  $M_T(\mu, \vec{p}_T^{\text{miss}})$  is required to be greater than 60, 40, 15 and 15 GeV for 0-, 1-, 2-jet GGF and VBF categories, respectively, while azimuthal separation between the electron and  $\vec{p}_T^{\text{miss}}$  is required to be less than 0.7, 0.7, 0.5 and 0.3 for 0-, 1-, 2-jet GGF and VBF categories, respectively. For the 0- and 1-jet categories  $\Delta\phi(e, \mu) > 2.5$  and 1.0, respectively. The preselections and the selections for  $H \rightarrow \mu\tau_h$  and  $H \rightarrow \mu\tau_e$  channels in all categories are summarized in Table 1.1. The  $M_{\text{col}}$  distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow \mu\tau_e$  channel, are shown in results chapter.

## 1.5 $H \rightarrow e\tau_h$ channel

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

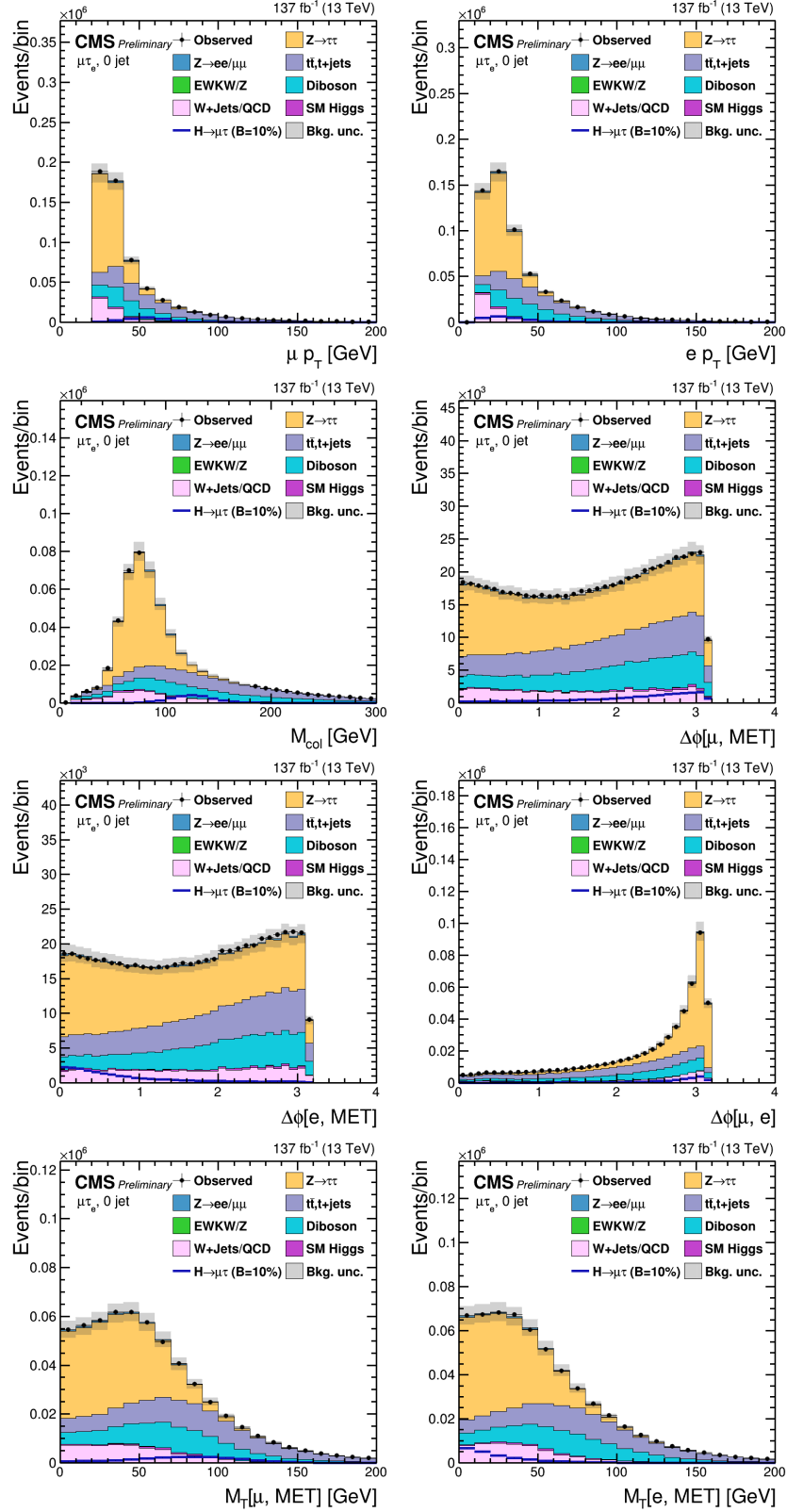


Figure 1.8. Distribution of the input variables to the BDT for the  $H \rightarrow \mu\tau_e$  process.

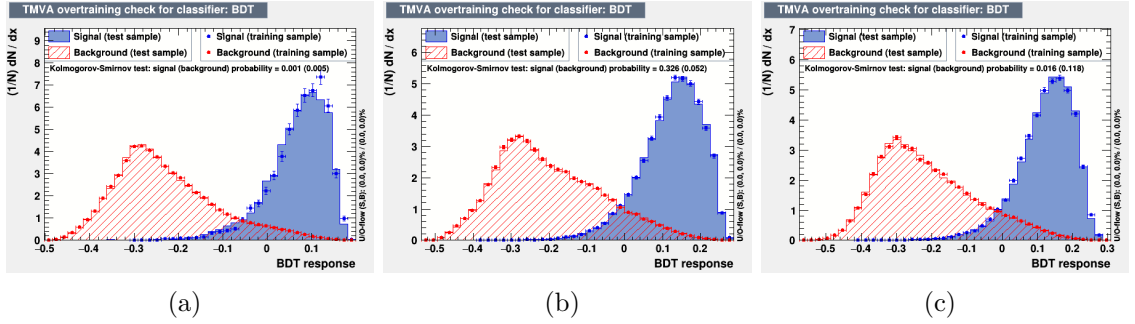


Figure 1.9. Overtraining check as performed in TMVA for the trained BDT in  $H \rightarrow \mu\tau_e$  channel for 2016 (a), 2017 (b), and 2018 (c).

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

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

A BDT is trained after applying preselection criteria. The same training samples, as used in the  $H \rightarrow \mu\tau_h$  channel, are considered. The list of input variables to BDT training stays the same, except for the addition of the visible mass,  $M_{\text{vis}}$  variable, and removal of  $\vec{p}_T^{\text{miss}}$ . The  $M_{\text{vis}}$  variable is more useful as the relative composition of the two channels' backgrounds is different. In particular,  $Z \rightarrow ee+\text{jets}$  background contributes more with respect to  $Z \rightarrow \mu\mu+\text{jets}$  background. The distribution of the input variables to the BDT can be seen in Figure 1.10. The BDT discriminator distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow e\tau_h$  channel, are shown in results chapter.

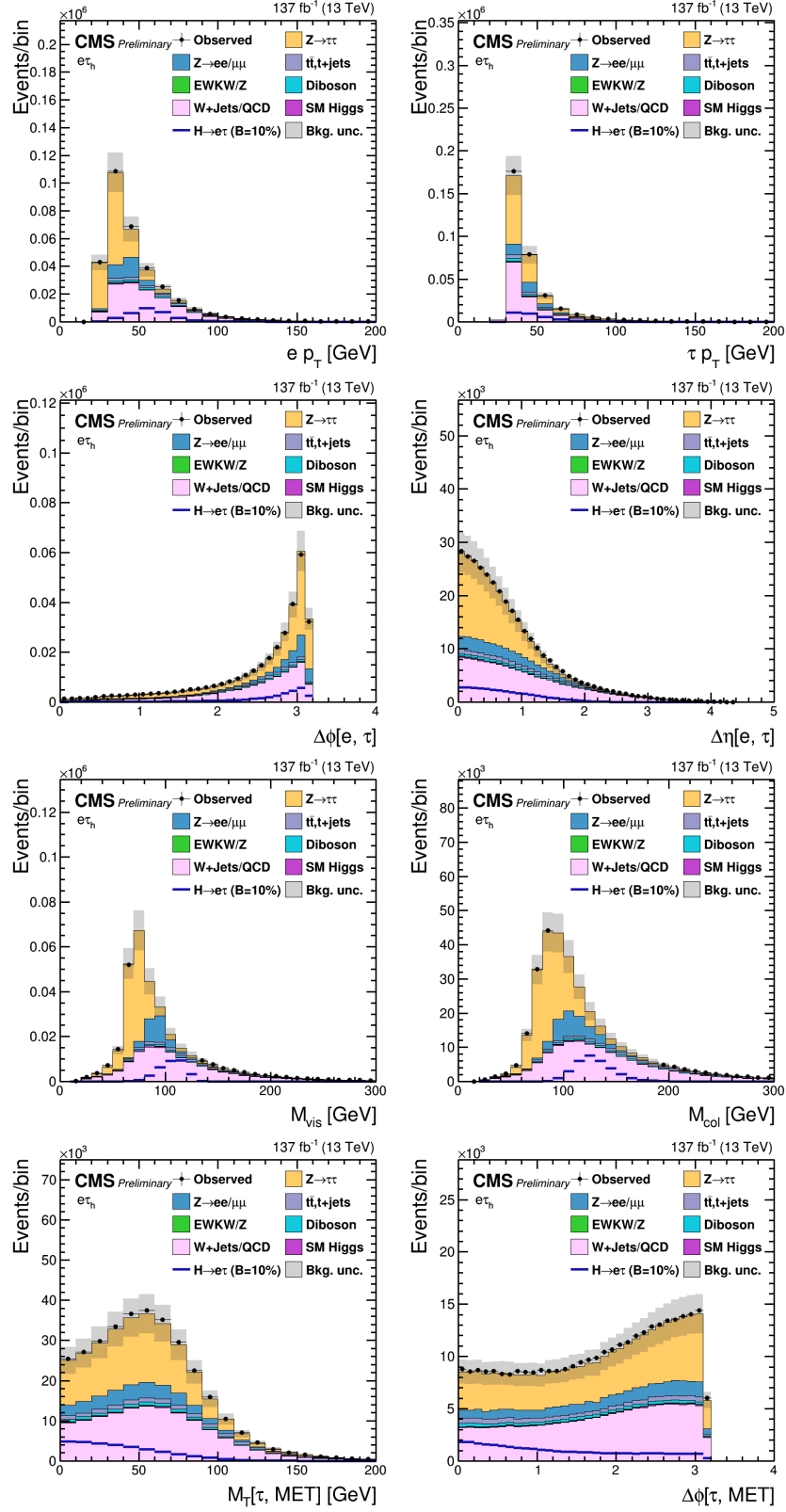


Figure 1.10. Distribution of the input variables to the BDT for the  $H \rightarrow e\tau_h$  process.



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

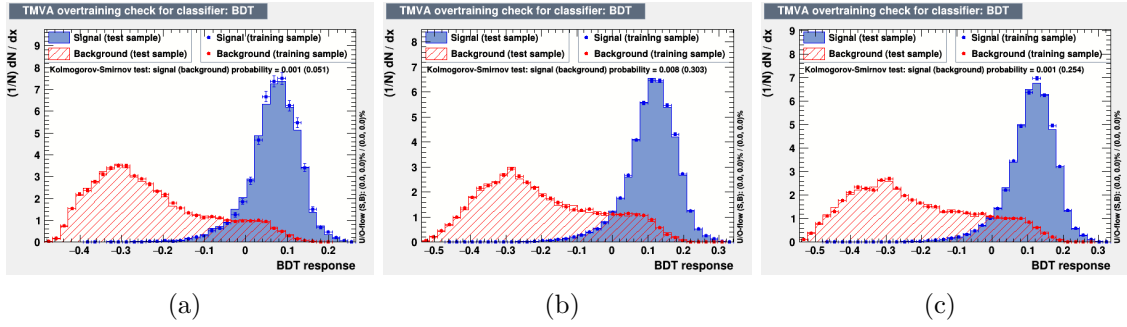


Figure 1.11. Overtraining check as performed in TMVA for the trained BDT in  $H \rightarrow e\tau_h$  channel for 2016 (a), 2017 (b), and 2018 (c).

## 1.6 $H \rightarrow e\tau_\mu$ 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  $23 \text{ GeV}$ , and on the muon is  $8 \text{ GeV}$ . The cross-trigger also places a constraint on the longitudinal impact parameter of the two leptons to the primary vertex. However, this constraint is not present in the initial 2016 data samples and 2016 MC samples. In addition to the event passing the trigger, the reconstructed leptons corresponding to the trigger have to match the HLT objects within  $\Delta R < 0.5$ .

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

TABLE 1.2

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

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

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

candidate is required to have  $p_T^e > 24 \text{ GeV}$ ,  $|\eta| < 2.4$  and isolation  $I_{\text{rel}}^e < 0.1$ . The muon candidate is required to have  $p_T^\mu > 10 \text{ GeV}$ ,  $|\eta| < 2.5$  and isolation  $I_{\text{rel}}^\mu < 0.15$ . The  $p_T$  threshold of the electron and the muon are dictated by the trigger we use for selecting the events of this channel. Events containing additional electrons, muons,  $\tau_h$  candidates or at least one b jet tagged by DeepCSV algorithm are removed. The selections for both  $H \rightarrow e\tau_h$  and  $H \rightarrow e\tau_\mu$  channels in all categories are also summarized in Table 1.2.

The BDT training is done after applying preselection criteria. The same training samples as used in the  $H \rightarrow \mu\tau_e$  channel are considered. The list of input variables to BDT training also stays the same, except for the addition of the visible mass,  $M_{\text{vis}}$  variable, and removal of  $M_T(e, \vec{p}_T^{\text{miss}})$ . The distribution of the input variables to the BDT can be seen in Figure 1.12. The BDT discriminator distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow e\tau_\mu$  channel, are shown in results chapter.

In the  $M_{\text{col}}$  fit method, additional selection criteria require a stringent selection on electrons,  $p_T^e > 30 \text{ GeV}$  for 0-jet category and  $p_T^e > 26 \text{ GeV}$  in rest of the categories. The  $M_T(e, \vec{p}_T^{\text{miss}})$  is required to be greater than 60, 40, 15 and 15 GeV for 0-, 1-, 2-jet GGF and VBF categories, respectively, while azimuthal separation between the muon and  $\vec{p}_T^{\text{miss}}$  is required to be less than 0.7, 0.7, 0.5 and 0.3 for 0-, 1-, 2-jet GGF and VBF categories, respectively. For the 0- and 1-jet categories  $\Delta\phi(e, \mu) > 2.5$  and 1.0, respectively. The selections for both  $H \rightarrow e\tau_h$  and  $H \rightarrow e\tau_\mu$  channels in all categories are also summarized in Table 1.2. The  $M_{\text{col}}$  distributions of simulated signal, data, and backgrounds for each category in  $H \rightarrow e\tau_\mu$  channel, are shown in results chapter.

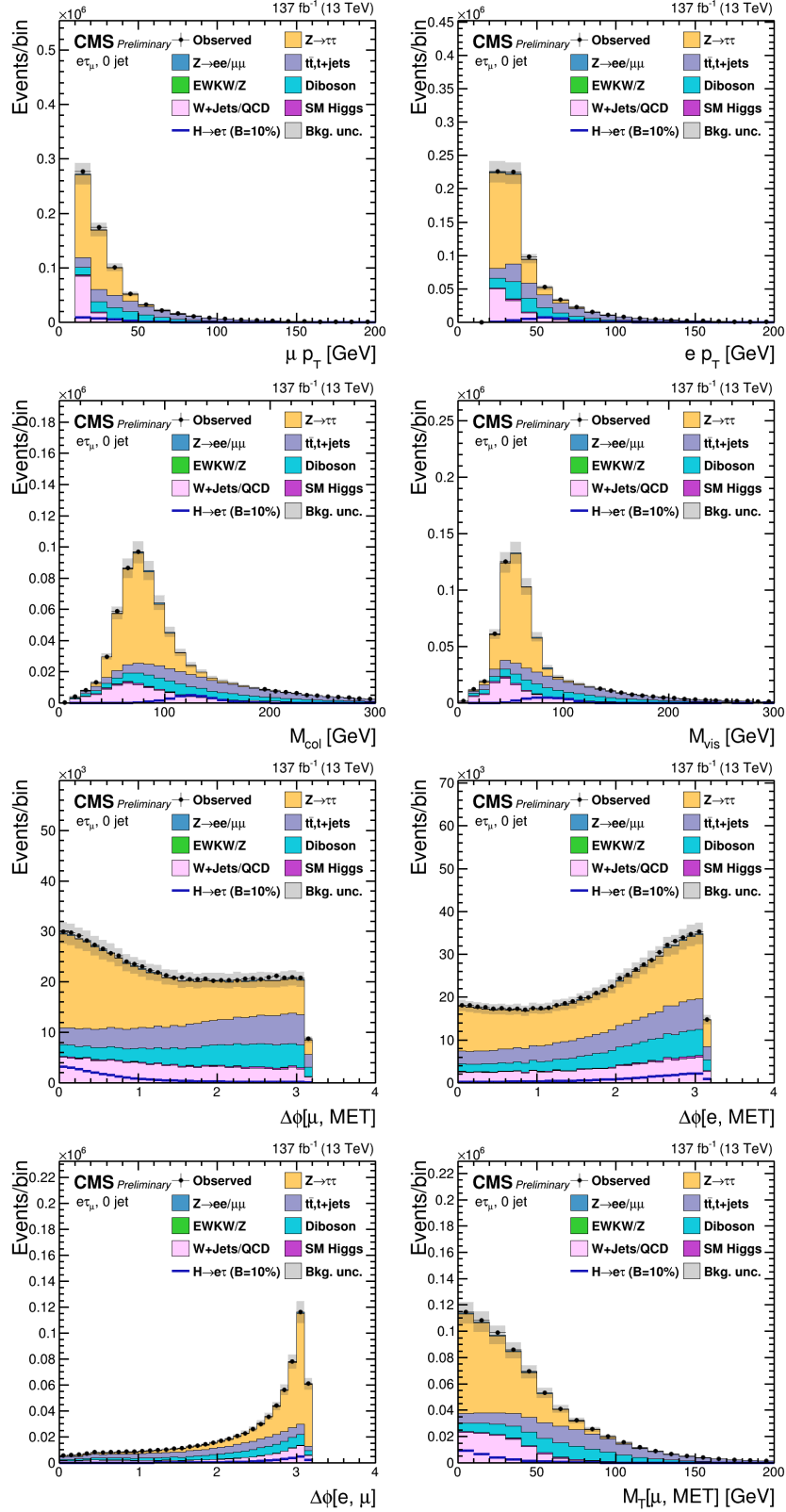


Figure 1.12. Distribution of the input variables to the BDT for the  $H \rightarrow e\tau\mu$  process.

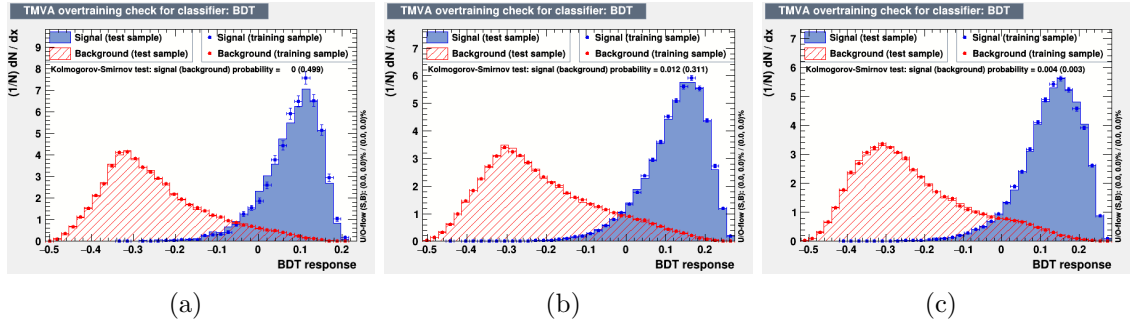


Figure 1.13. Overtraining check as performed in TMVA for the trained BDT in  $H \rightarrow e\tau_\mu$  channel for 2016 (a), 2017 (b), and 2018 (c).

## CHAPTER 2

### BACKGROUND ESTIMATION

#### 2.1 Introduction

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

Other contributions come from the processes in which a lepton pair is produced from the weak decays of quarks and vector bosons. These processes include Higgs boson production ( $H \rightarrow \tau\tau$ ,  $WW$ ),  $WW$ ,  $WZ$ , 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 2.1.

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.

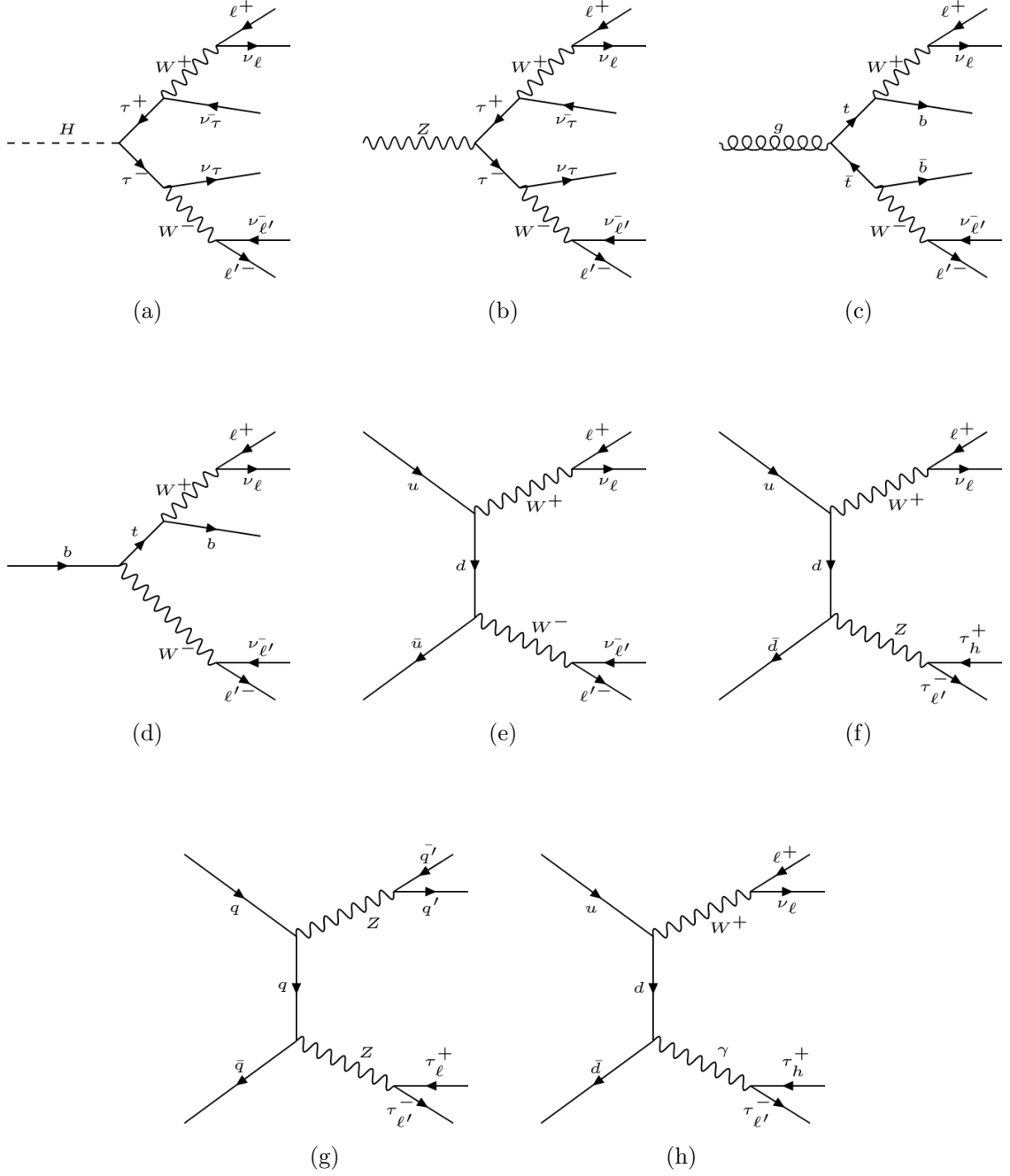


Figure 2.1. Feynman diagrams of background processes to LFV Higgs boson decays: (a)  $H \rightarrow \tau\tau$ , (b)  $Z \rightarrow \tau\tau$ , (c)  $t\bar{t}$ , (d) Single Top, (e)  $WW$ , (f)  $WZ$ , (g)  $ZZ$ , and (h)  $W\gamma^{(*)}$ .

## 2.2 Embedding technique

The  $Z \rightarrow \tau\tau$  background is estimated from data using the embedding technique [5]. The embedding technique allows for an estimation of the genuine  $\tau\tau$  standard model backgrounds from data that minimizes uncertainties arising from a poor event description, with minimal simulation input. Events with a pair of oppositely charged muons are selected in data so that  $Z \rightarrow \mu\mu$  events largely dominate it. These data events are chosen independently of the event selection criteria that are described in Chapter 1.

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

Embedded samples cover all backgrounds with two real  $\tau$  leptons decaying semi-hadronically or leptonically. This includes a small fraction of  $t\bar{t}$ , Diboson, and electroweak W/Z events. The events from the  $t\bar{t}$ , Diboson, and electroweak W/Z MC samples where both tau candidates match genuine taus at the generator level are removed to avoid any double counting. A schematic for the embedding technique can be seen in Figure 2.2.

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



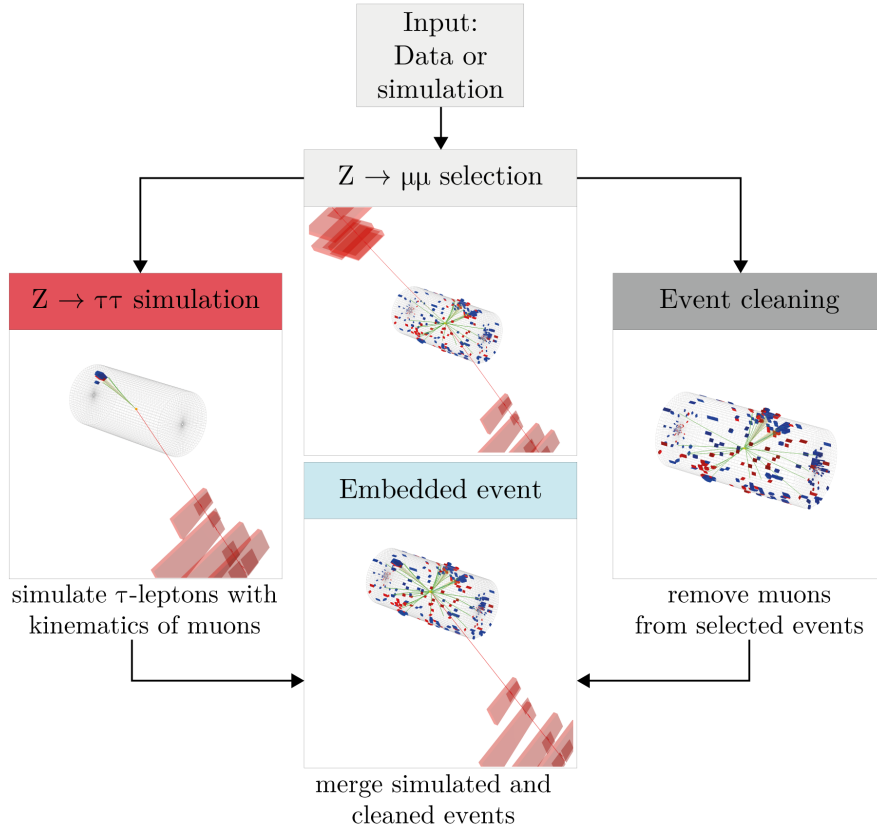


Figure 2.2. Schematic of Embedding Technique

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

In  $H \rightarrow \mu\tau_e$  channel, this region is constructed by requiring, in addition to the preselection,  $M_T(\mu) < 60 \text{ GeV}$ ,  $30 \text{ GeV} < M_{\text{vis}}(\mu, e) < 70 \text{ GeV}$ , and  $p_T^\mu < 40 \text{ GeV}$ . In  $H \rightarrow e\tau_\mu$  channel, the same requirements are placed with the muon variables replaced by corresponding electron variables and vice versa. Figures 2.3 and 2.4 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.

### 2.3 Misidentified lepton background

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

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

#### 2.3.1 Fully data-driven approach

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

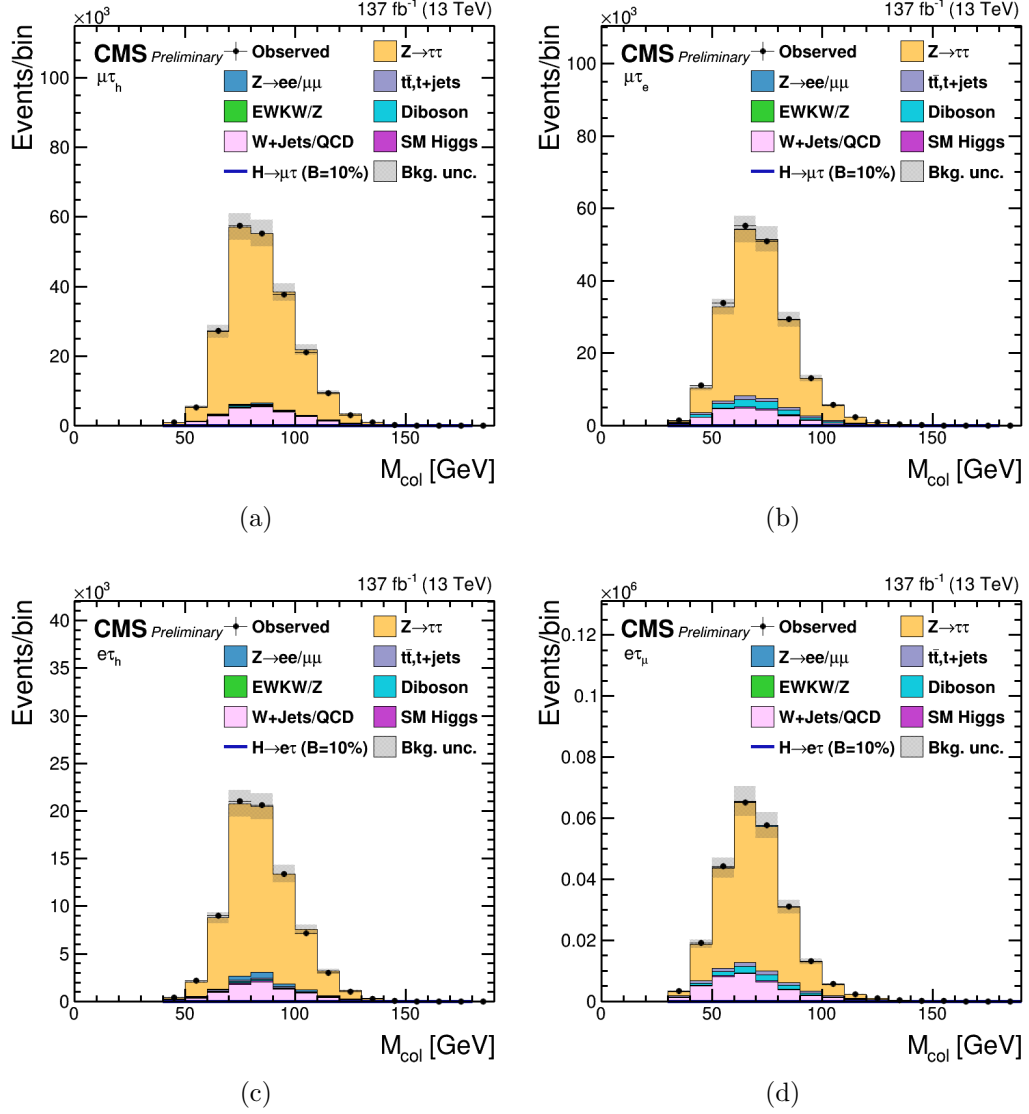


Figure 2.3. Distributions of  $M_{\text{col}}$  discriminator in the  $Z \rightarrow \tau\tau$  control regions for the (a)  $H \rightarrow \mu\tau_h$ , (b)  $H \rightarrow \mu\tau_e$ , (c)  $H \rightarrow e\tau_h$ , and (d)  $H \rightarrow e\tau_\mu$  channels.

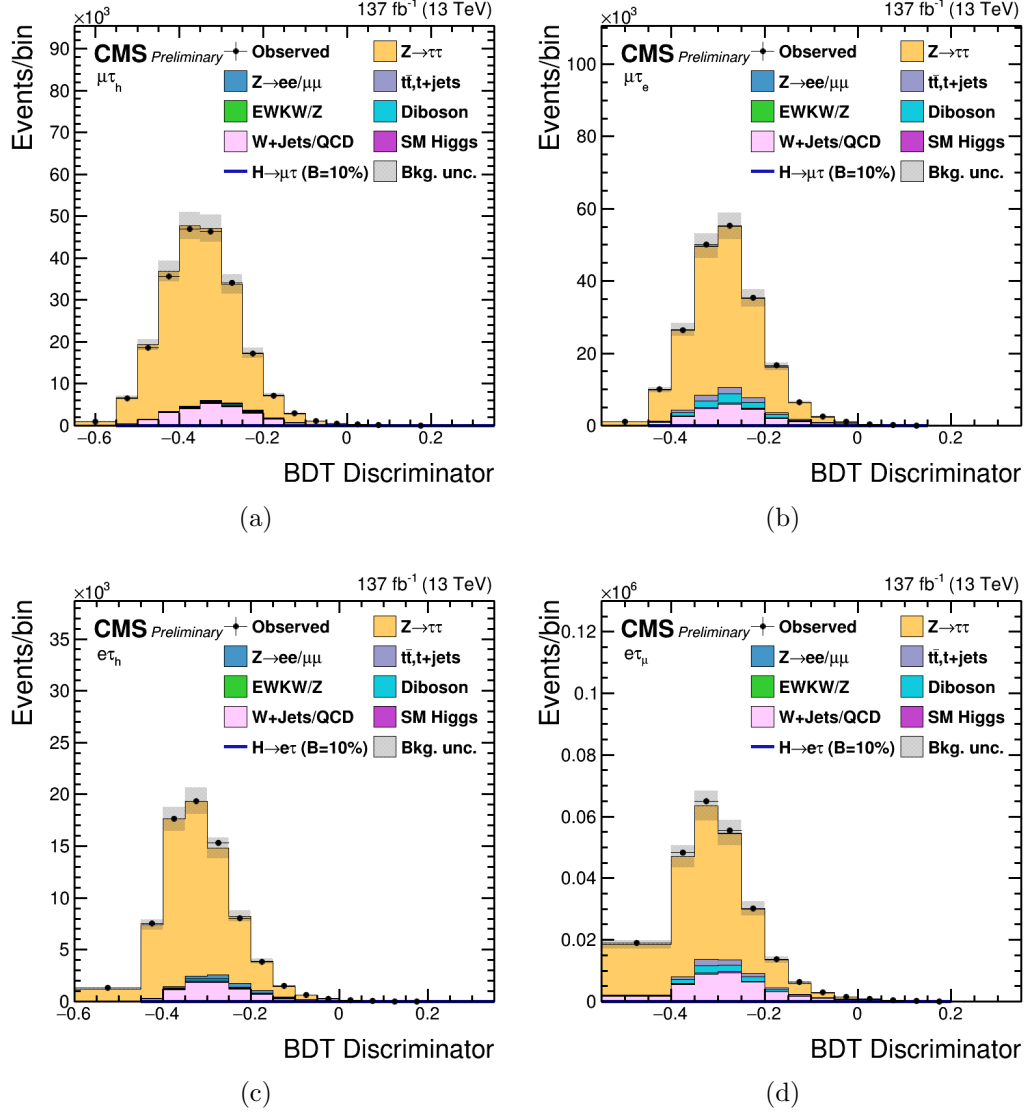


Figure 2.4. Distributions of BDT discriminator in the  $Z \rightarrow \tau\tau$  control regions for the (a)  $H \rightarrow \mu\tau_h$ , (b)  $H \rightarrow \mu\tau_e$ , (c)  $H \rightarrow e\tau_h$ , and (d)  $H \rightarrow e\tau_\mu$  channels.

applied to the control region, to estimate the misidentified background of the signal region. The signal region contrasted with the control regions used for determining the misidentified background can be seen in Figure 2.5.

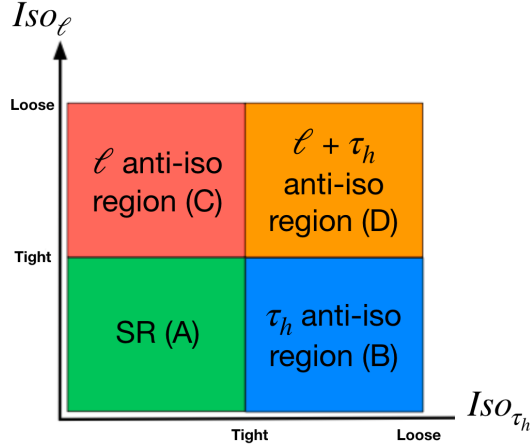


Figure 2.5. Signal region (green) contrasted with the control regions used for estimating the misidentified background

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

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

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

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

where  $N_i(\text{signal-like})$  is the number of events with a third lepton candidate that passes the tight lepton isolation, while  $N_i(\text{background-like})$  is the number of events that pass only the looser lepton isolation and index  $i = e, \mu$ , or  $\tau$ . The lepton selection criteria is summarized in Tables 1.1 and 1.2.

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

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

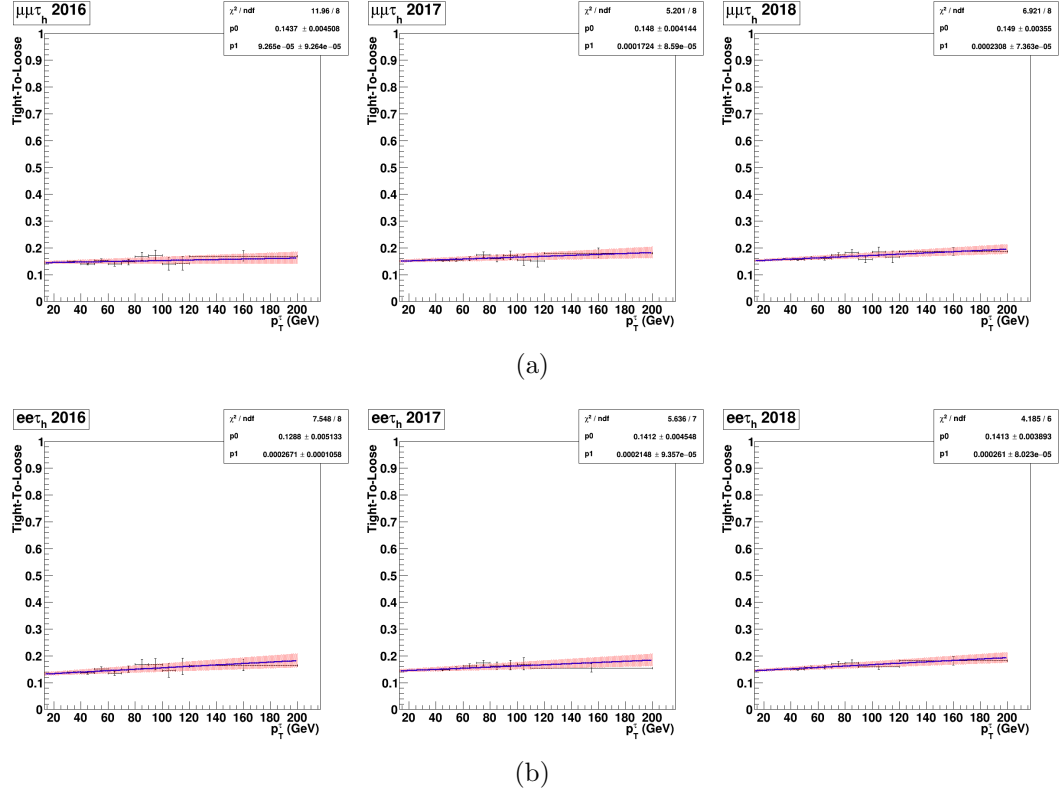


Figure 2.6. Fit performed to  $\tau_h$  misidentification rates for  $H \rightarrow \mu\tau_h$  (a) and  $H \rightarrow e\tau_h$  (b) channel as a function of  $\tau_h$   $p_T$  for the different years. The misidentification rates used are further parametrized based on  $\tau_h$  Decay Mode along with the pseudorapidity of  $\tau_h$ . However, here only the inclusive misidentification rates are shown. The misidentification rates are labeled as “tight-to-loose” to clarify that they are calculated as a ratio of the number of events passing the tight WP to the loose WP of DNN discrimination against jets.

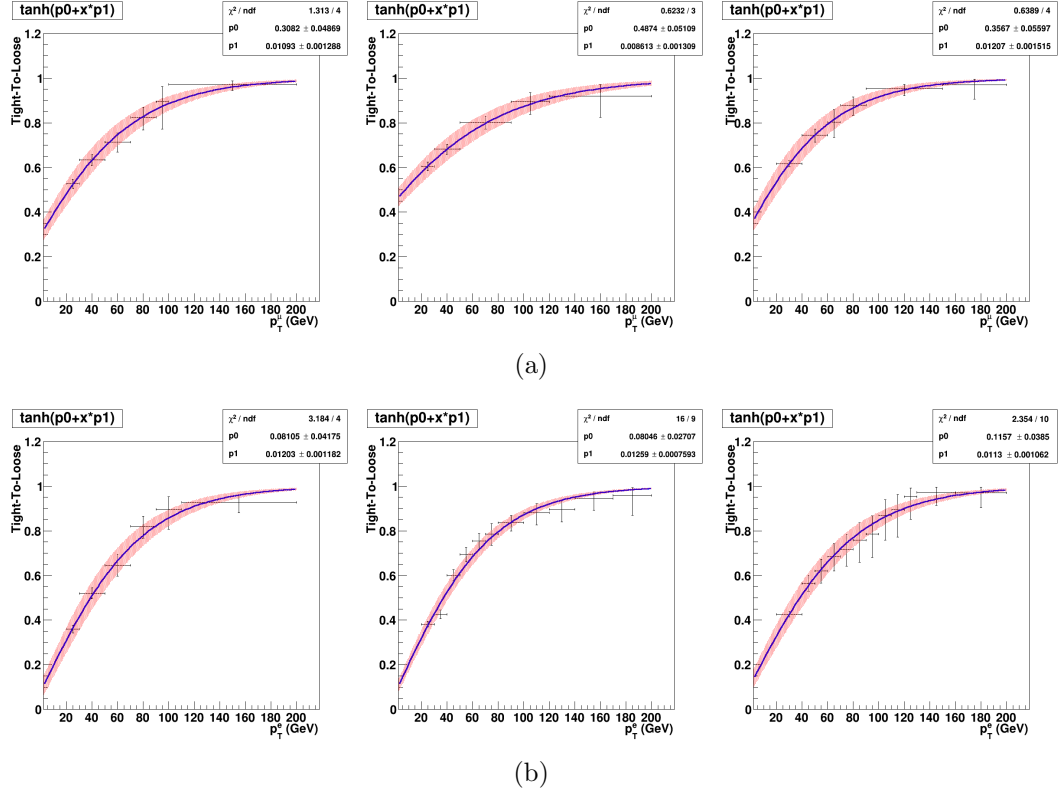


Figure 2.7. Fit performed to the  $\mu$  (a) and  $e$  (b) misidentification rates as a function of their  $p_T$  for 2016 (Left), 2017 (Center), and 2018 (Right). The misidentification rates are labeled as “tight-to-loose” to clarify that they are calculated as a ratio of the number of events passing the tight isolation to the loose isolation. The hyperbolic tangent function is used for performing the fit.



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

The estimation of the background is validated in events where the two leptons have the same electric charge. The misidentification rate  $f_i$  is applied to events passing preselection and by inverting the lepton pair's charge requirement. The same-sign selection enhances the misidentified lepton background. The background estimation is also validated in a  $W$  boson enriched control sample. This control region is obtained by applying the signal-like requirement and  $M_T(\ell, \vec{p}_T^{\text{miss}}) > 60 \text{ GeV}$  ( $\ell = e$  or  $\mu$ ) and  $M_T(\tau_h, \vec{p}_T^{\text{miss}}) > 80 \text{ GeV}$ . The same strategy is applied in the  $H \rightarrow e\tau_h$  channel, which results in a similar agreement. Figures 2.8 and 2.9 show the comparison of data with background estimates in the same-sign and  $W$  boson enriched control regions for the  $H \rightarrow \mu\tau_h$  and  $H \rightarrow e\tau_h$  channels.

### 2.3.2 Semi data-driven approach

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

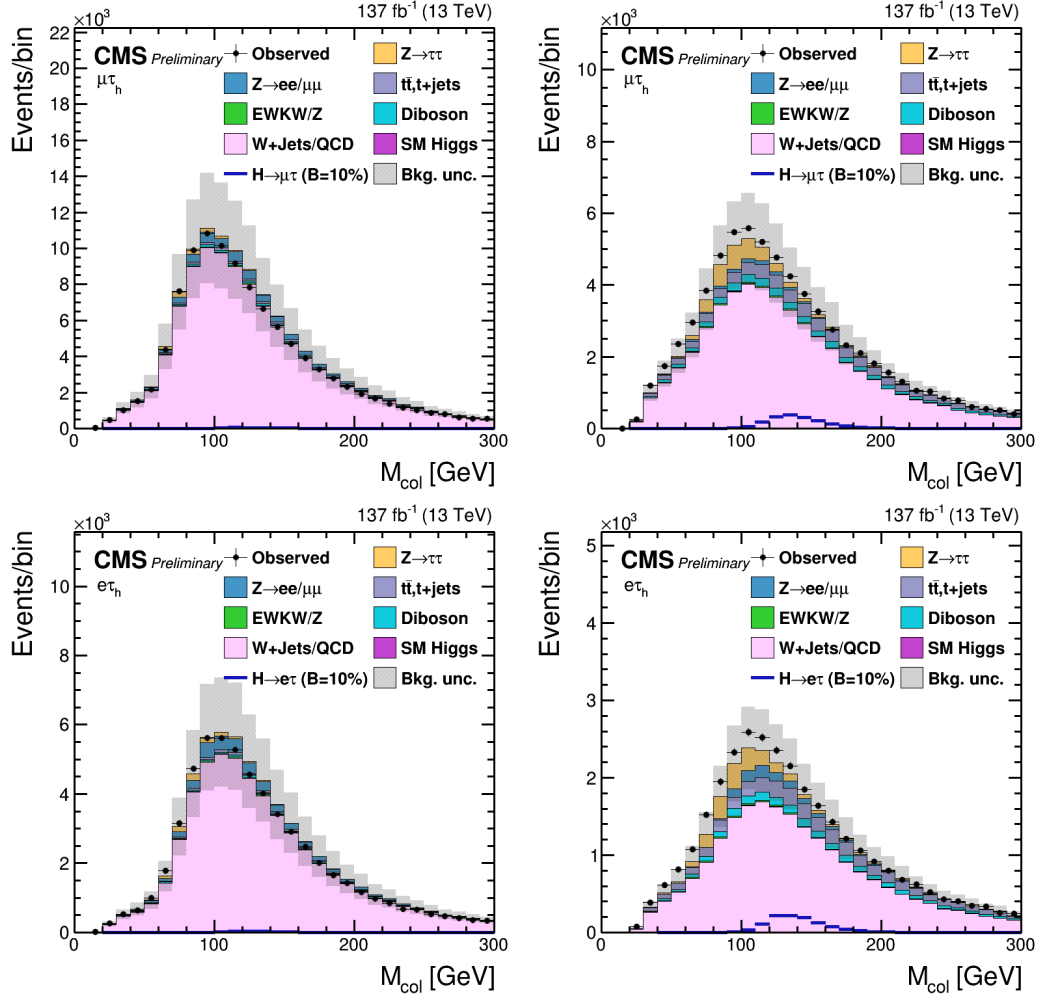


Figure 2.8. Distributions of  $M_{\text{col}}$  discriminator in the same-sign (Left) and W boson enriched (Right) control regions for the  $H \rightarrow \mu\tau_h$  (top) and  $H \rightarrow e\tau_h$  (bottom) channels.

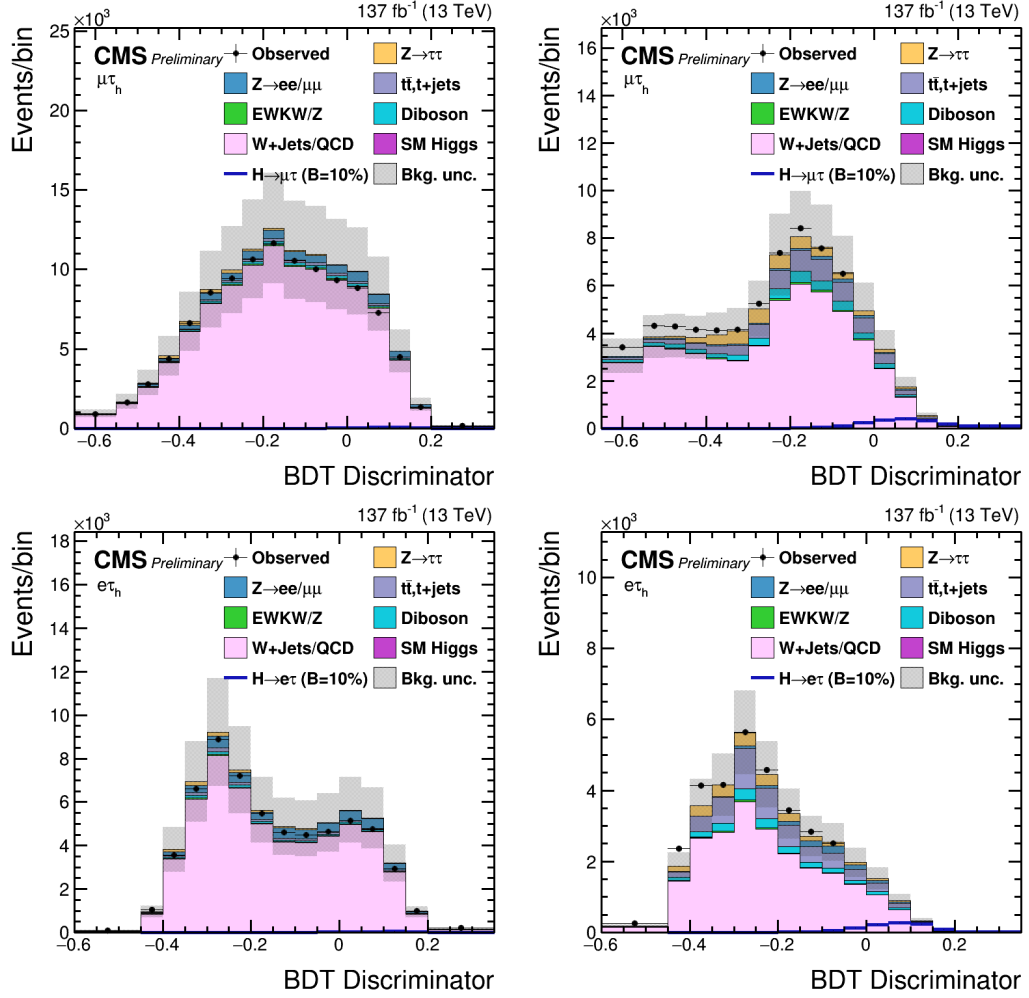


Figure 2.9. Distributions of BDT discriminator in the same-sign (Left) and W boson enriched (Right) control regions for the  $H \rightarrow \mu\tau_h$  (top) and  $H \rightarrow e\tau_h$  (bottom) channels.

region to the opposite sign (OS) signal region are measured in data as a function of the jet multiplicity and the  $\Delta R$  separation between the electron and the muon. QCD OS/SS extrapolation factors that are measured can be seen in Figure 2.10.

The OS/SS extrapolation factor is estimated using events with an anti-isolated muon and an isolated electron. The contribution from  $b\bar{b}$  events to the QCD multijet background gives rise to the  $\Delta R$  dependency and is parameterized with a linear function. The OS/SS extrapolation factor is higher for events with low  $\Delta R$  separation between the electron and the muon, decreasing as the  $\Delta R$  separation increases. The OS/SS extrapolation factor also depends on the electron and muon  $p_T$ . This  $p_T$  dependence comes from the leptons arising from the semi-leptonic  $c$  quark decay. These leptons tend to be softer in  $p_T$  and less isolated resulting in a reduction in the number of such events passing the  $p_T$  and isolation requirements. Corrections of the QCD OS/SS extrapolation factors dependent on lepton  $p_T$  along with the correction to account for the mismodeling introduced by anti-isolating the muon to measure them can be seen in Figure 2.11.

As the OS/SS extrapolation factor is measured in a control region where the muon is anti-isolated, an additional correction is applied to cover for a potential mismodeling. This correction is calculated by measuring the OS/SS extrapolation factors in two different control regions. The first control region has events where the muon is isolated, and the electron is anti-isolated. The second control region has events where both the electron and the muon are anti-isolated. The ratio of the extrapolation factors measured in these two control regions is taken as the correction for accounting the potential mismodeling induced by anti-isolating the muon. Figure 2.12 shows the comparison of data with background estimates in the muon anti-isolated control regions for the  $H \rightarrow \mu\tau_e$  channel.

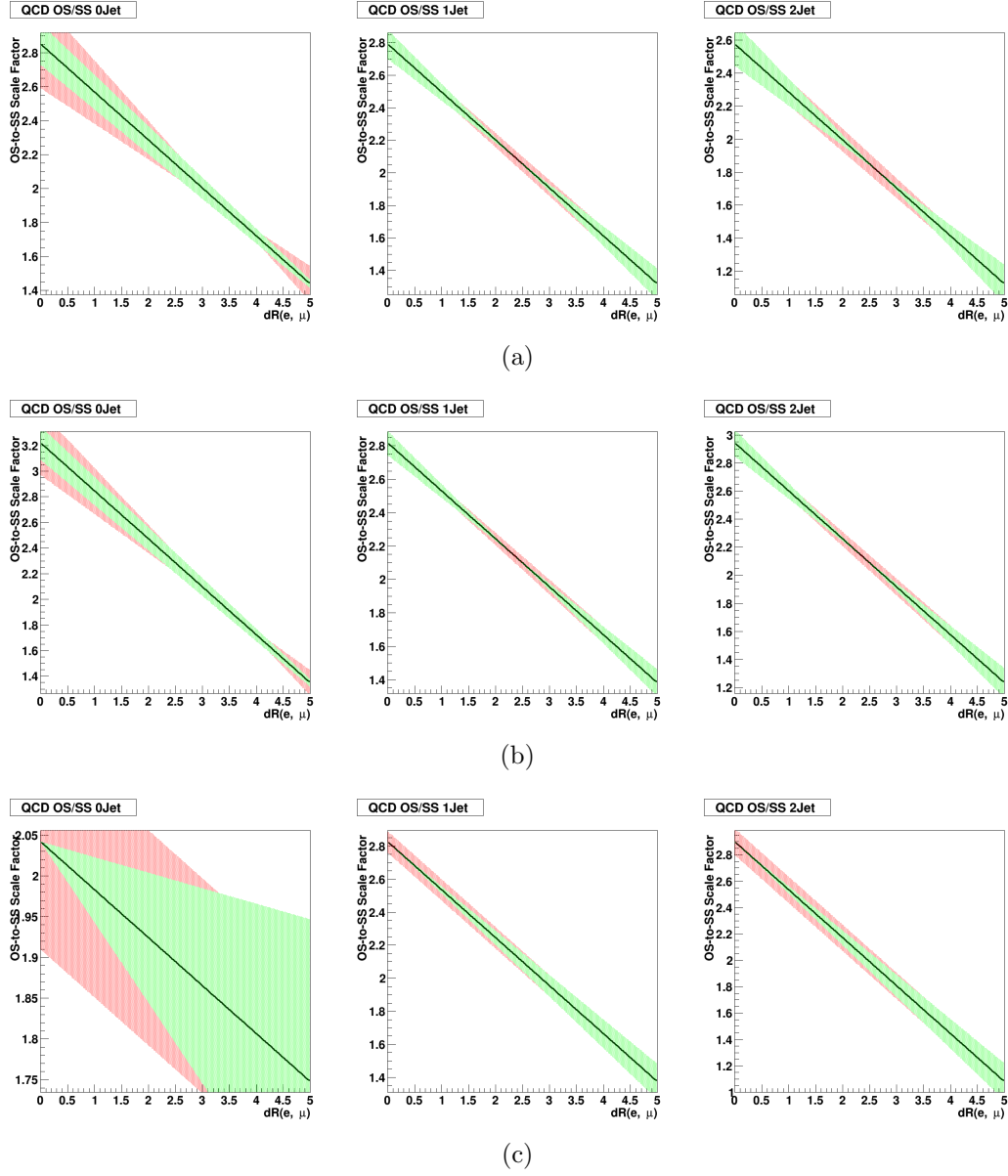


Figure 2.10. QCD OS/SS extrapolation factors in events with 0 Jets (Left), 1 Jet (Center), and 2 Jets (Right) for 2016 (a), 2017 (b), 2018 (c). The line is the best fit, and the shaded region corresponds to the shape uncertainties.

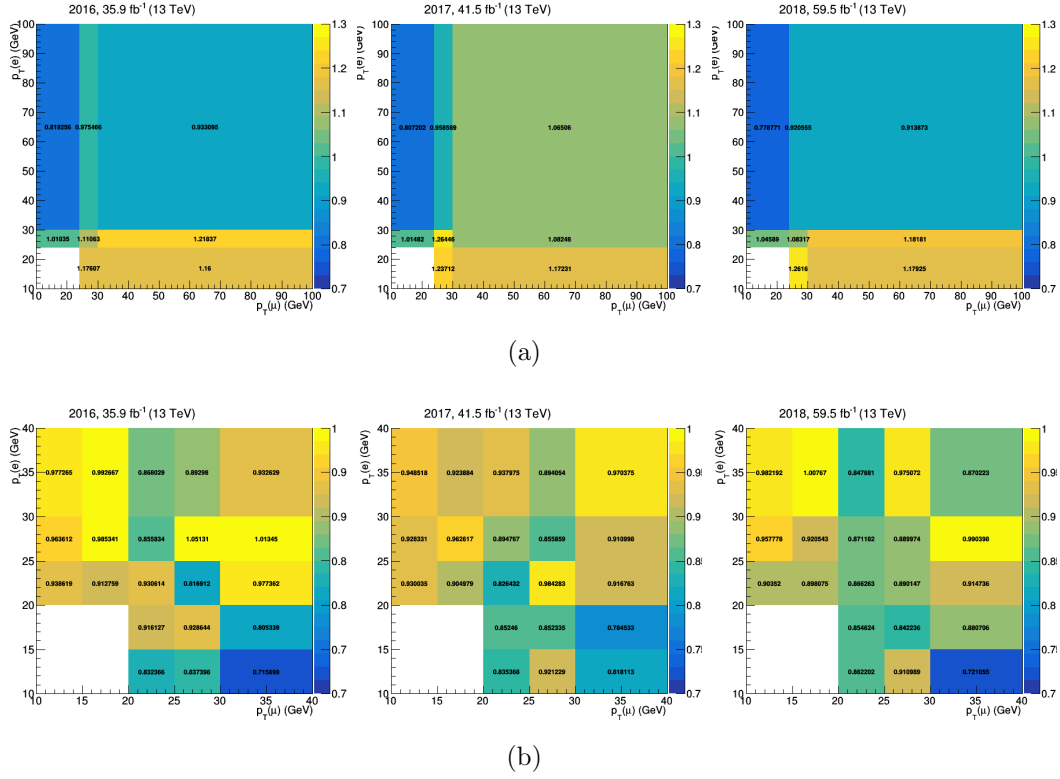


Figure 2.11. (a) Corrections of the QCD OS/SS extrapolation factors determined in the region with an anti-isolated muon as a function of the  $p_T$  of the electron and the muon, using data collected in 2016, 2017, and 2018. (b) Correction of the QCD OS/SS extrapolation factors to account for the mismodeling introduced by anti-isolating the muon to measure the SFs, using data collected in 2016, 2017, and 2018.

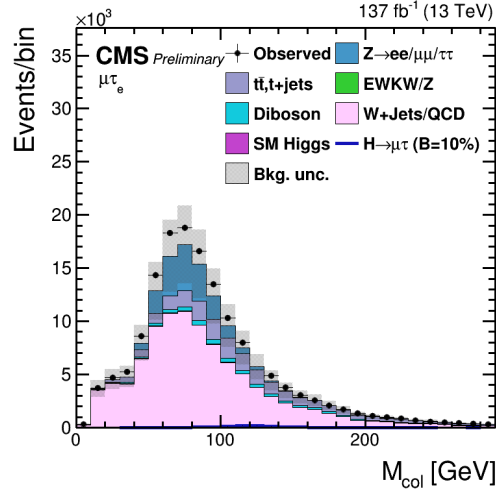


Figure 2.12. Distribution of  $M_{\text{col}}$  discriminator in the muon anti-isolated control regions for the  $H \rightarrow \mu\tau_e$  channel.

## 2.4 MC Simulation

All the other backgrounds are estimated using MC simulation. In the leptonic channels, the  $t\bar{t}$  process has a dominant contribution, and this background is validated in a dedicated control region defined by requiring the presence of at least one b jet tagged by the DeepCSV algorithm in the event. Figure 2.13 shows the background validation in this control region for the  $H \rightarrow \mu\tau_e$  and  $H \rightarrow e\tau_\mu$  channels.

The standard model, Higgs boson production, forms a small but non-negligible background. The contributions come mainly from  $H \rightarrow \tau\tau$  and  $H \rightarrow WW$  decays. The contribution from  $H \rightarrow WW$  peaks at lower values than the signal in the distribution of the BDT discriminator due to the presence of additional neutrinos in the decay. The background is estimated from simulations using selection based on the BDT discriminator and kinematic variables, as described in Chapter 1.

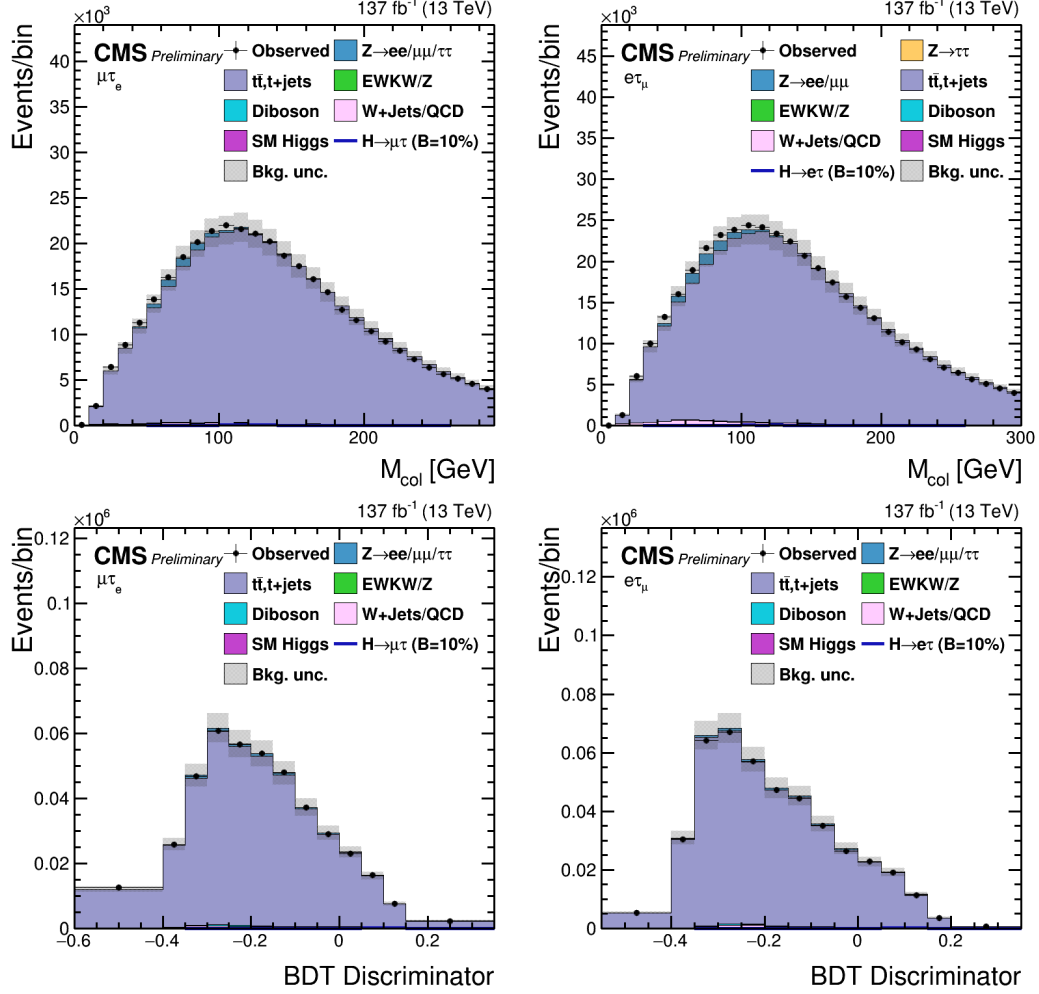


Figure 2.13. Distributions of  $M_{\text{col}}$  (BDT) discriminator in  $t\bar{t}$  enriched control region for  $H \rightarrow \mu\tau_e$  and  $H \rightarrow e\tau_\mu$  channel.



## BIBLIOGRAPHY

1. CMS Collaboration. Observation of the Higgs boson decay to a pair of  $\tau$  leptons. *Phys. Lett. B*, 779:283, 2018. doi: 10.1016/j.physletb.2018.02.004.
2. A. Hoecker et al. *TMVA - Toolkit for Multivariate Data Analysis*, 2017. <http://tmva.sourceforge.net/>.
3. 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 high energy hadron colliders. *Nucl. Phys. B*, 297:221, 1988. doi: 10.1016/0550-3213(88)90019-3.
4. Aaron Roodman. Blind Analysis in Particle Physics . 2003. doi: arXiv:physics/0312102v1.
5. 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.