

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 12, 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 . . . . .	iv
Chapter 1: Event Selection . . . . .	2
1.1 Introduction . . . . .	2
1.2 Boosted Decision Tree . . . . .	8
1.3 $H \rightarrow \mu\tau_h$ channel . . . . .	10
1.4 $H \rightarrow \mu\tau_e$ channel . . . . .	14
1.5 $H \rightarrow e\tau_h$ channel . . . . .	15
1.6 $H \rightarrow e\tau_\mu$ channel . . . . .	19
Bibliography . . . . .	24

## 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). . . .	3
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. . . . .	4
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. . . . .	5
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. . . . .	7
1.5	Illustration of decision tree. . . . .	8
1.6	Distribution of the input variables to the BDT for the $H \rightarrow \mu\tau_h$ process. 12	
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). . . . .	14
1.8	Distribution of the input variables to the BDT for the $H \rightarrow \mu\tau_e$ process. 16	
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). . . . .	17
1.10	Distribution of the input variables to the BDT for the $H \rightarrow e\tau_h$ process. 18	
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). . . . .	19
1.12	Distribution of the input variables to the BDT for the $H \rightarrow e\tau_\mu$ process. 22	
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). . . . .	23

## TABLES

1.1	Event selection criteria for the kinematic variables for the $H \rightarrow \mu\tau$ channels . . . . .	13
1.2	Event selection criteria for the kinematic variables for the $H \rightarrow e\tau$ channels . . . . .	20

$\tau$

## 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



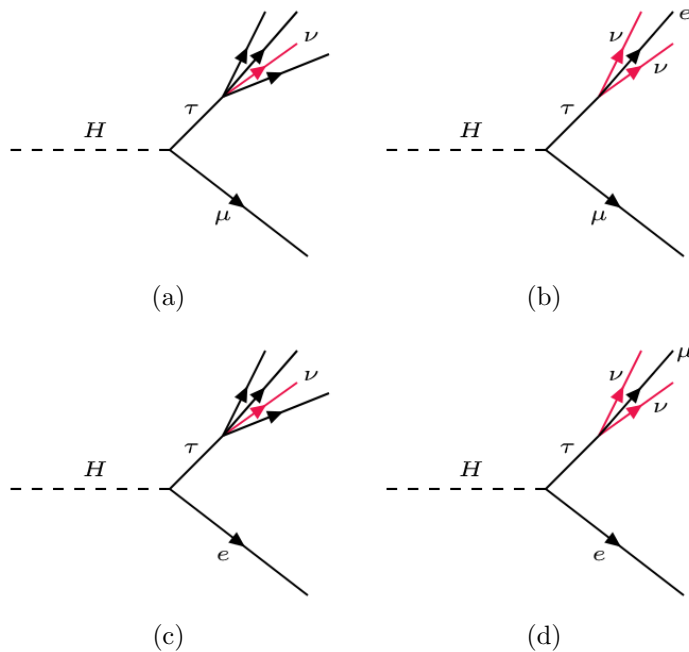


Figure 1.1. Feynman diagrams of lepton-flavor violating Higgs-boson decays. The first row shows diagrams for the Higgs boson coupling to  $\mu\tau$  (a,b). Couplings to  $e\tau$  (c,d) are shown in the second row. Taus can decay leptonically or hadronically. Feynman diagrams are shown for the leptonic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

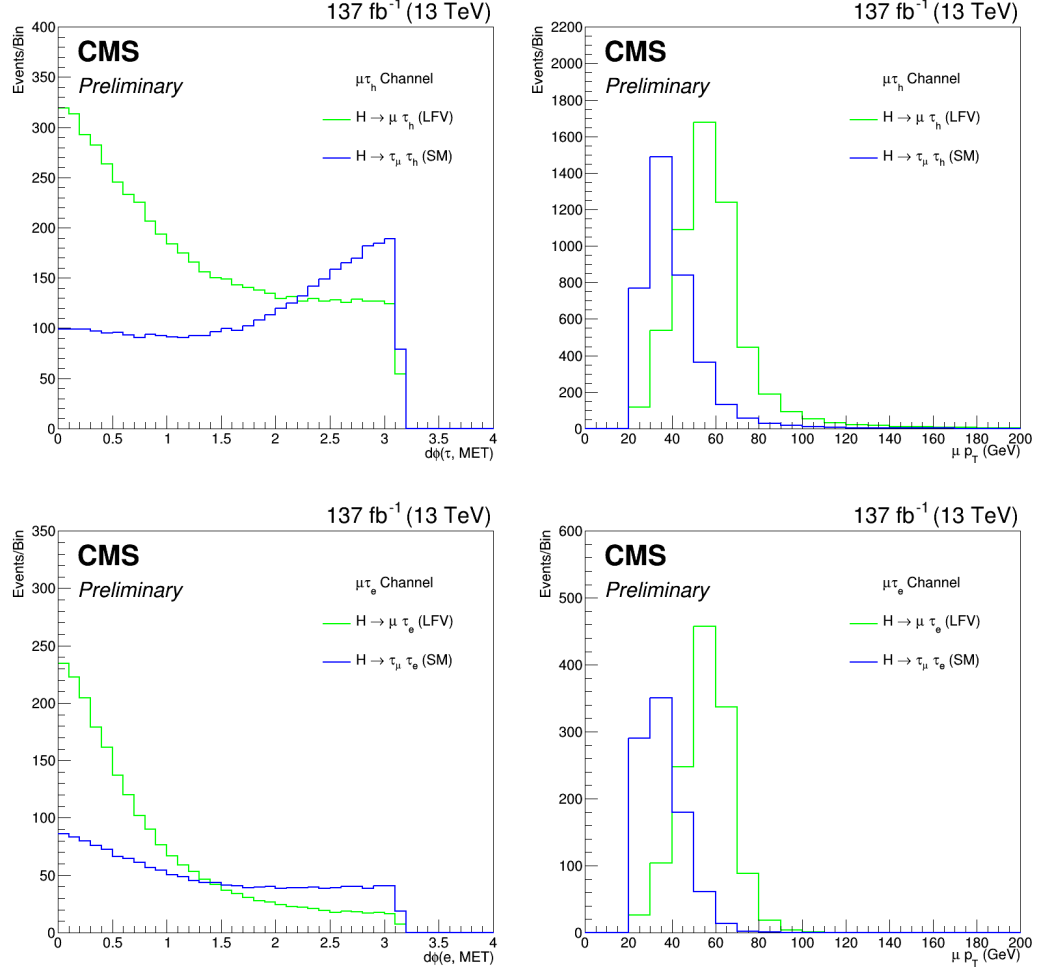


Figure 1.2. Illustration of the differences in  $d\phi(\ell = \tau \text{ or } e, MET)$  and  $p_T^\mu$  spectrums in LFV and SM  $H \rightarrow \tau\tau$  processes.

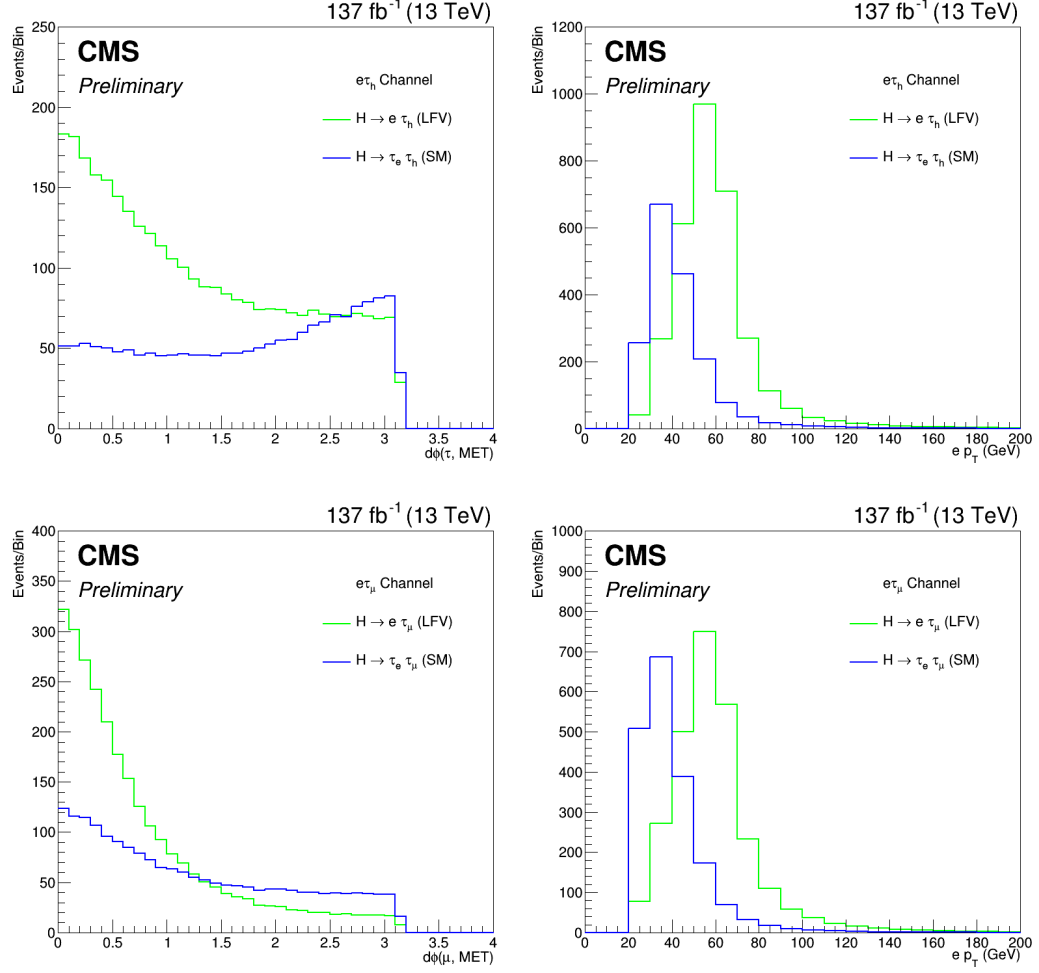


Figure 1.3. Illustration of the differences in  $d\phi(\ell = \tau \text{ or } \mu, MET)$  and  $p_T^e$  spectrums in LFV and SM  $H \rightarrow \tau\tau$  processes.

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

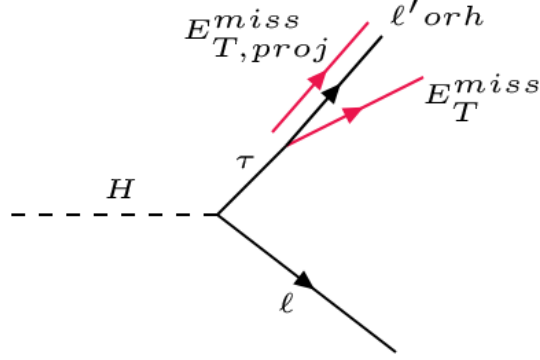


Figure 1.4. Estimation of the neutrino momentum  $\cancel{E}_{T,proj}$  by using the component of the missing transverse energy  $\cancel{E}_T$  which is collinear to the visible decay products of tau in the transverse plane.

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.

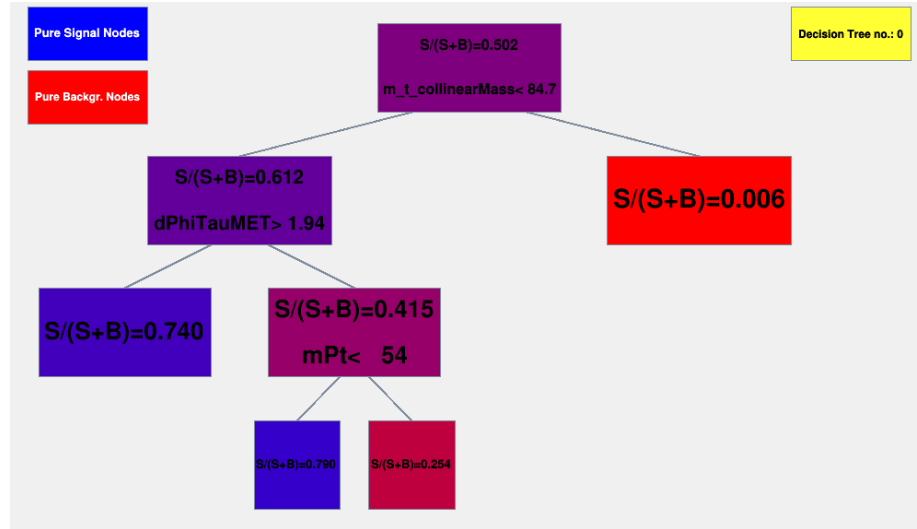


Figure 1.5. Illustration of decision tree.

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.

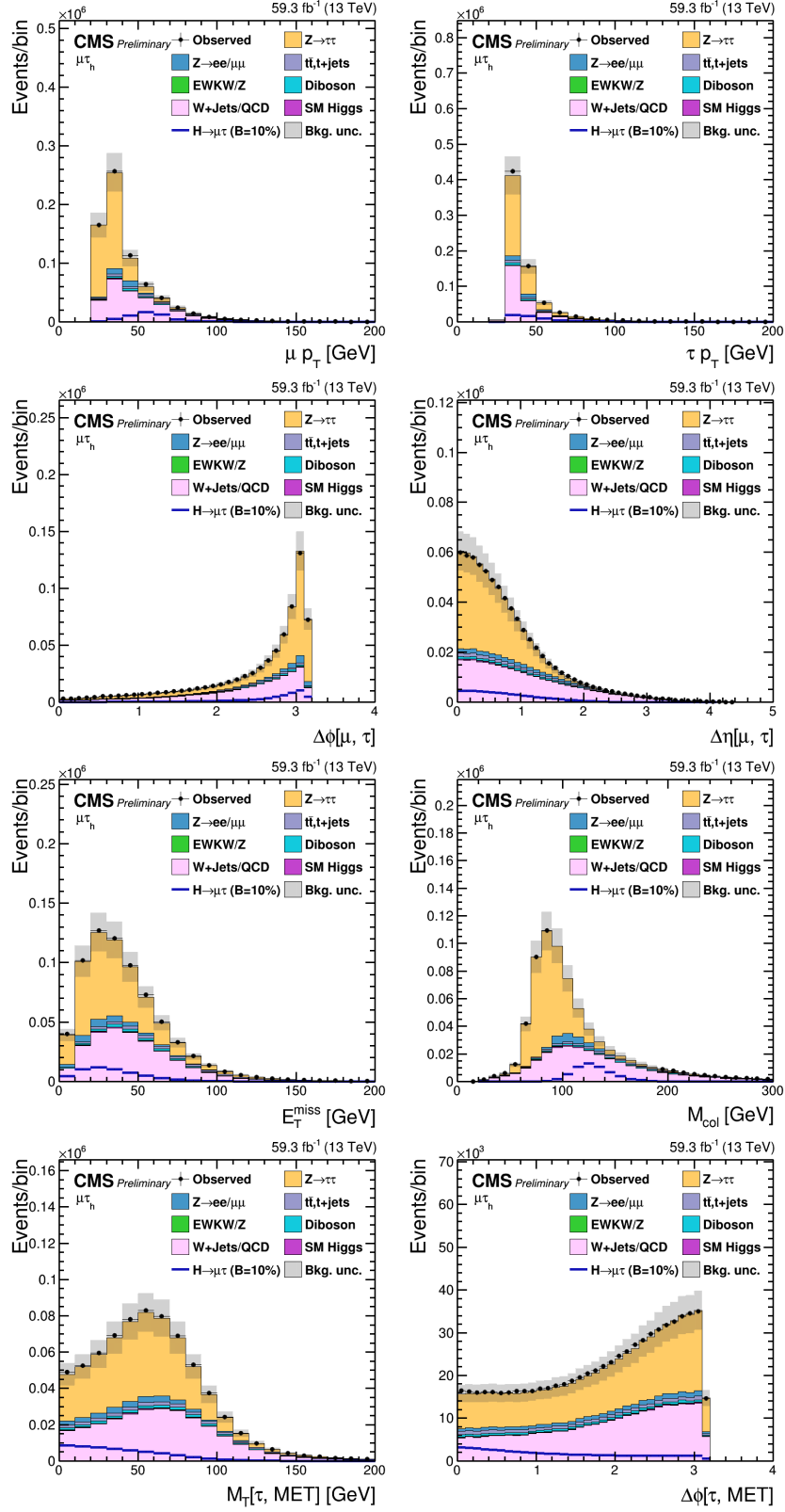


Figure 1.6. Distribution of the input variables to the BDT for the  $H \rightarrow \mu\tau_h$  process.

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
$M_{jj}$	—	—	$< 550$	$\geq 550$	—	—	$< 550$	$\geq 550$
$p_{\text{T}}^{\mu}$			—		$> 30$	$> 26$	$> 26$	$> 26$
$M_{\text{T}}(\mu)$			—		$> 60$	$> 40$	$> 15$	$> 15$
$M_{\text{T}}(\tau_{\text{h}})$	$< 105$	$< 105$	$< 105$	$< 85$			—	
$\Delta\phi(\text{e}, \vec{p}_{\text{T}}^{\text{miss}})$			—		$< 0.7$	$< 0.7$	$< 0.5$	$< 0.3$
$\Delta\phi(\text{e}, \mu)$			—		$> 2.5$	$> 1.0$	—	—

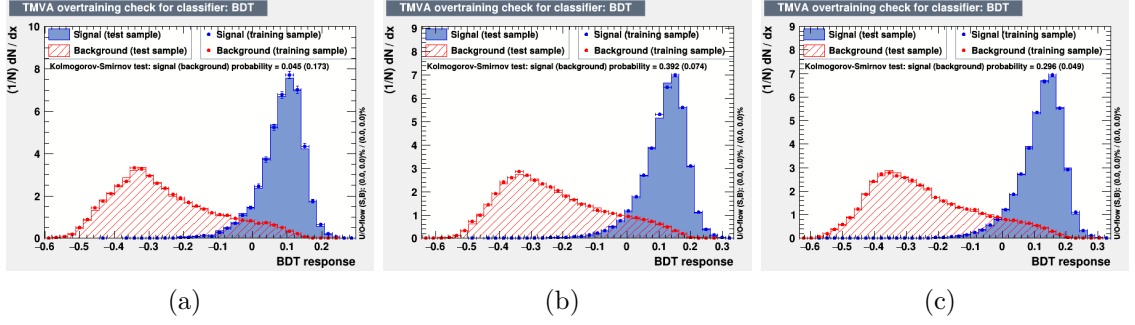


Figure 1.7. Overtraining check as performed in TMVA for the trained BDT in  $H \rightarrow \mu\tau_h$  channel for 2016 (a), 2017 (b), and 2018 (c).

#### 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

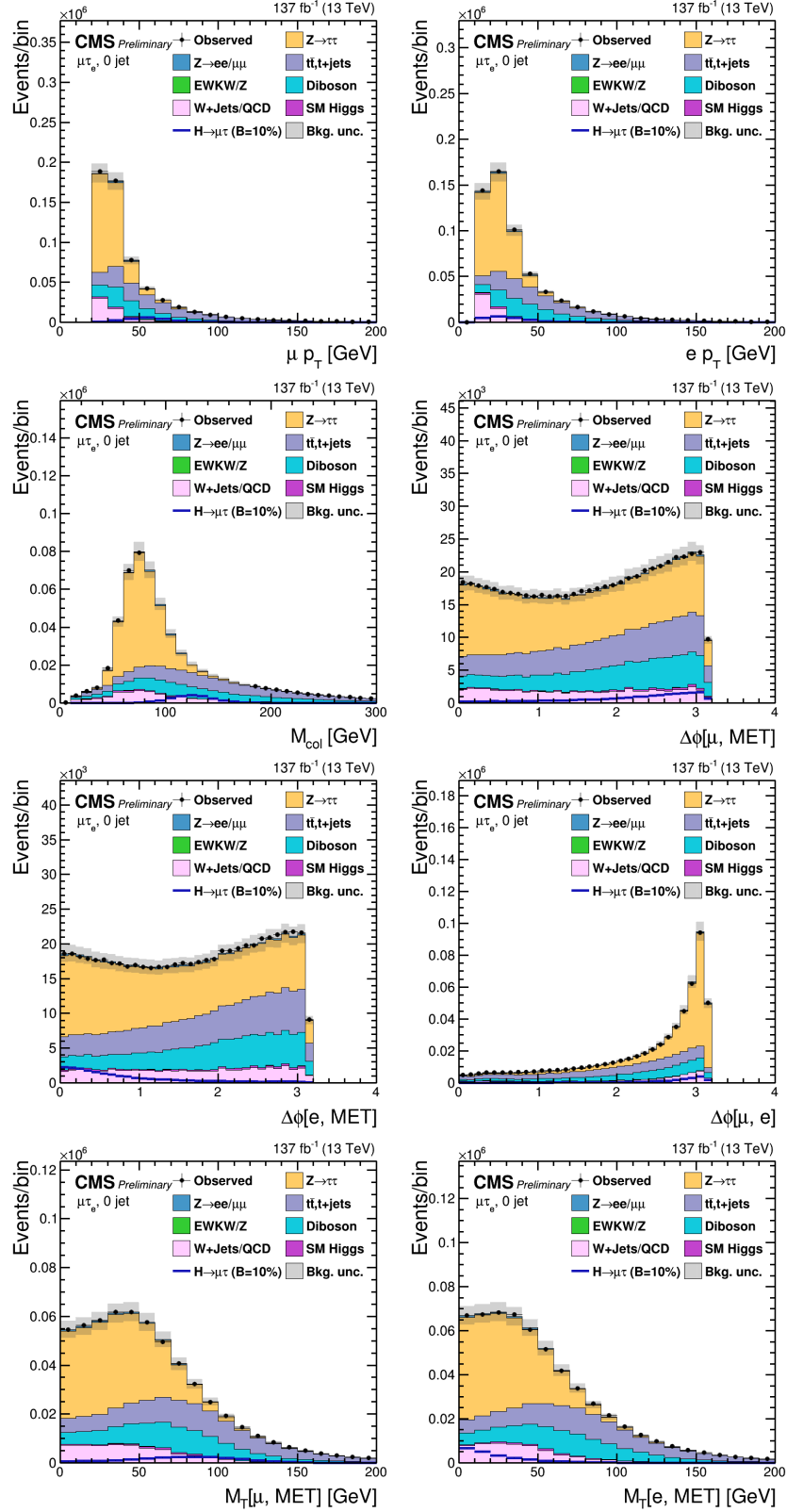


Figure 1.8. Distribution of the input variables to the BDT for the  $H \rightarrow \mu\tau_e$  process.

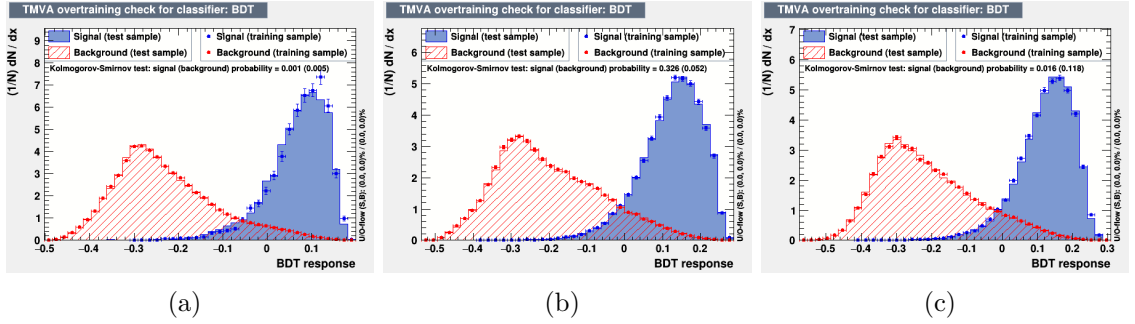


Figure 1.9. Overtraining check as performed in TMVA for the trained BDT in  $H \rightarrow \mu\tau_e$  channel for 2016 (a), 2017 (b), and 2018 (c).

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.

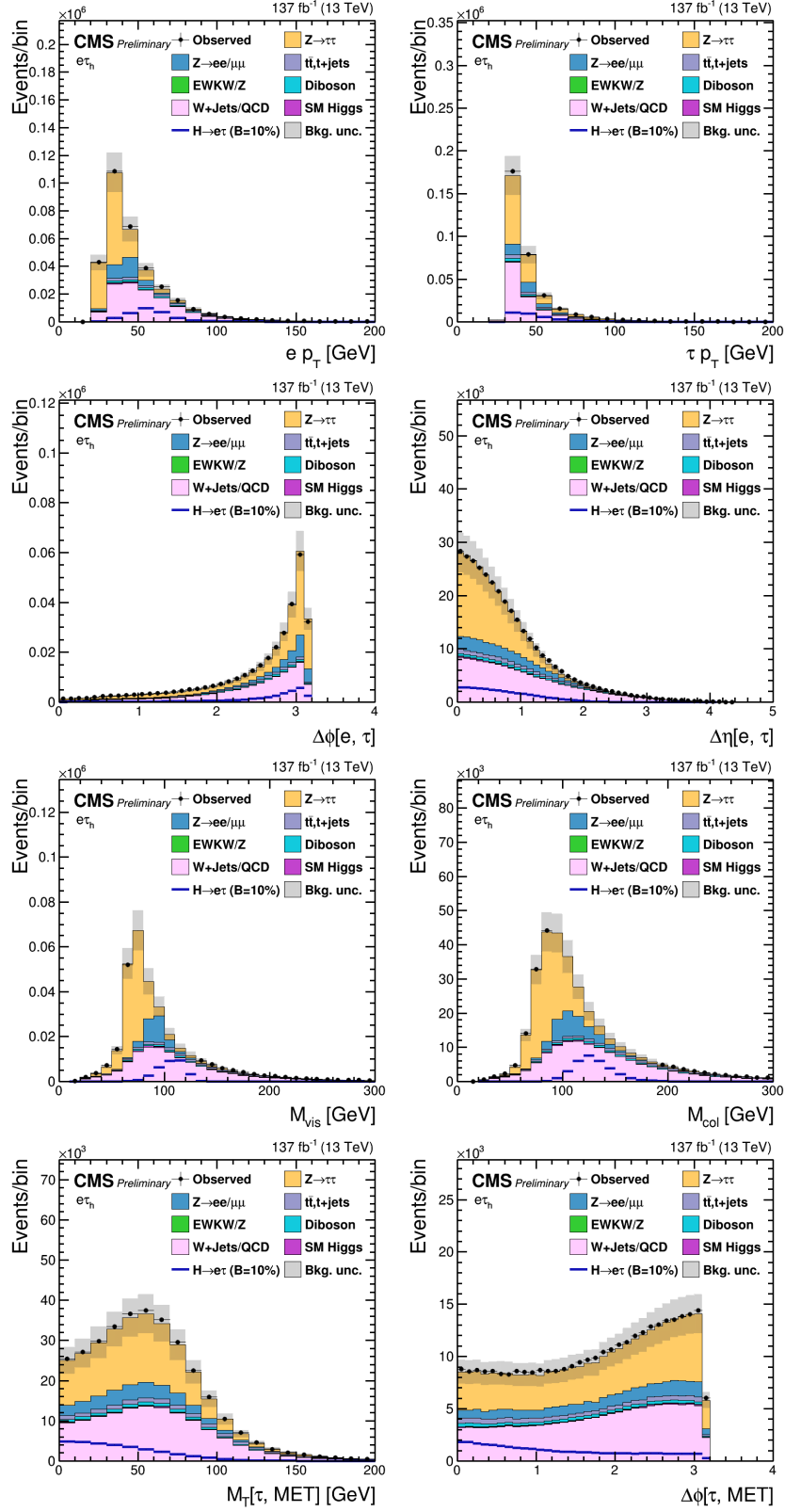


Figure 1.10. Distribution of the input variables to the BDT for the  $H \rightarrow e\tau_h$  process.



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.

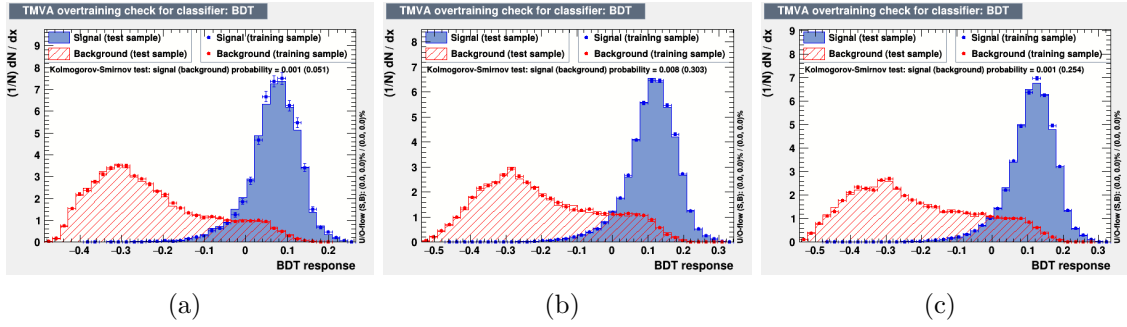


Figure 1.11. Overtraining check as performed in TMVA for the trained BDT in  $H \rightarrow e\tau_h$  channel for 2016 (a), 2017 (b), and 2018 (c).

## 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.

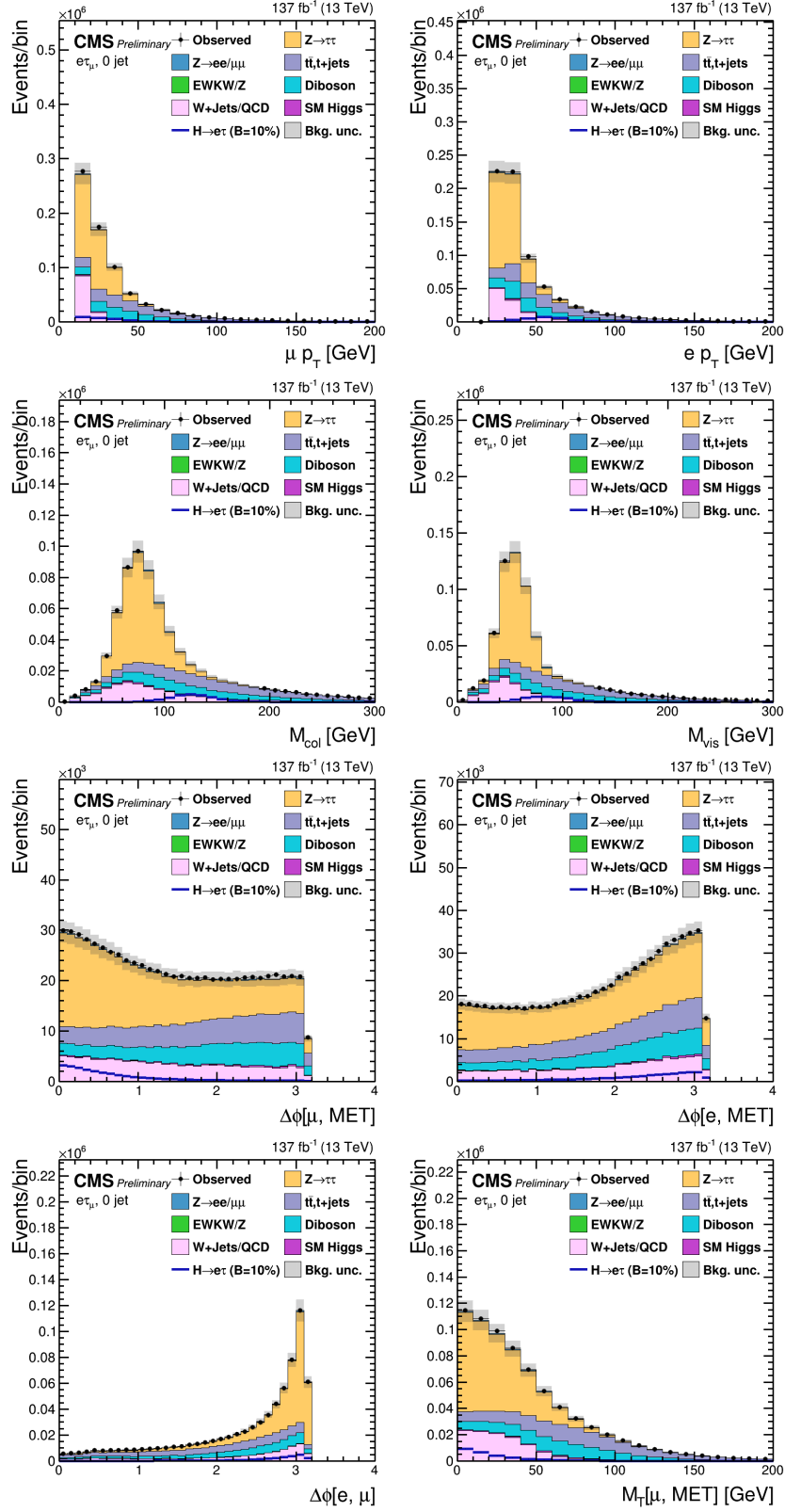


Figure 1.12. Distribution of the input variables to the BDT for the  $H \rightarrow e\tau\mu$  process.

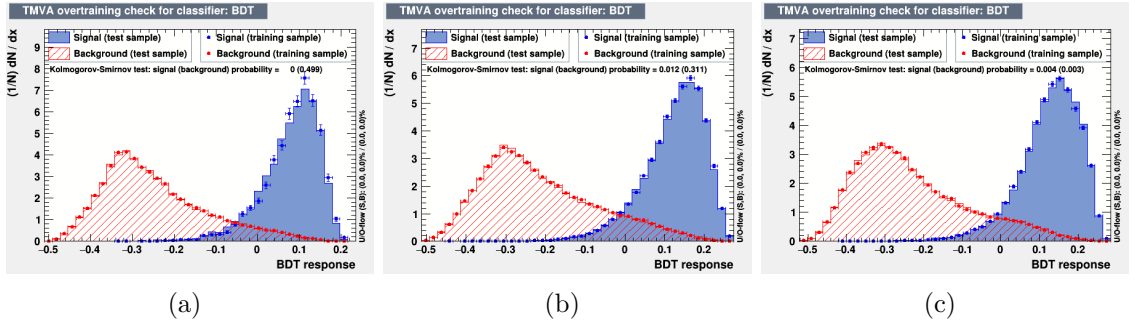


Figure 1.13. Overtraining check as performed in TMVA for the trained BDT in  $H \rightarrow e\tau_\mu$  channel for 2016 (a), 2017 (b), and 2018 (c).

## 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.