This Dissertation

entitled

SEARCH FOR LEPTON FLAVOR VIOLATING DECAYS

OF HIGGS BOSONS

WITH THE CMS EXPERIMENT

typeset with NDdiss2$_\varepsilon$ v3.2017.2 (2017/05/09) on February 16, 2019 for

Nabarun Dev

This LaTeX 2$_\varepsilon$ 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$_\varepsilon$ 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$_\varepsilon$ documentation can be found at these locations:

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

SEARCH FOR LEPTON FLAVOR VIOLATING DECAYS OF HIGGS BOSONS WITH THE CMS EXPERIMENT

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
in
Physics

by

Nabarun Dev

Colin Philip Jessop, Director

Graduate Program in Physics
Notre Dame, Indiana
February 2019

SEARCH FOR LEPTON FLAVOR VIOLATING DECAYS OF HIGGS BOSONS WITH THE CMS EXPERIMENT

Abstract

by

Nabarun Dev

DEDICATED TO

To my family

CONTENTS

FIGURES

TABLES

PREFACE

Long time ago in a galaxy far far away....(preface is optional)

ACKNOWLEDGMENTS

I would like to acknowledge the light side of the force, Master Kenobi and Grand Master Yoda.

SYMBOLS

c	speed of light
m	mass
e	elementary charge
E	energy

CHAPTER 1

INTRODUCTION

The standard model of particle physics is the most complete description of nature available today. The discovery of the Higgs Boson added another feather to the hat of the standard model...

...expand...

CHAPTER 2

THEORETICAL BASES

2.1 The Standard Model

2.2 Physics beyond the standard model

CHAPTER 3

EXPERIMENTAL SETUP

..intoduce...

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [1] is a powerful proton-proton synchrotron. It was built and is operated at the European Center for Nuclear Research (CERN) and is situated about 100 m underground close to Geneva, Switzerland. It has a circumference of 26.7 km and uses a tunnel previously built for LEP (Large Electron Positron Collider). Being a particle-particle collider, it consists of two rings with counterrotating beams which are steered using magnets and accelerated using radiofrequency resonating cavities. These beams are made to intersect at four collision points around the LHC ring, at one of which rests the CMS detector. Besides proton-proton collisions the LHC can also collide heavy ions (lead-lead collisions) or heavy ions with protons (lead-proton collisions). Since starting operation in September 2008 the LHC has been the world's most powerful apparatus and will probably remain so in the forseeable future. The following section describes proton-proton collisions at the LHC as the data used in the subsequent physics analysis corresponds to events from these collisions.

The injector chain that supplies protons to the LHC consists of four CERN accelerators that actually predate the LHC: Linac 2, PSB (Proton Synchroton Booster), PS (Proton Synchotron) and SPS (Super Proton Synchotron). This is illustrated in figure **??**. The proton source is simply a tank of hydrogen gas. The hydrogen atoms

are ionized to yield protons which are then fed in to the Linac 2, a linear accelerator. This accelerates the protons to an energy of about 50 MeV which are then fed into a series of circular accelerators starting with the PSB which accelerates the protons to 1.4 GeV. The PS then accelerates them to 25 GeV, and they are then sent to the SPS which accelerates them to 450 GeV before being finally fed into the LHC beampipe. Inside the LHC the protons are accelerated by sixteen radiofrequency cavities which are made to oscillate at 400 MHz and the proton beam is sorted into discrete packet called 'bunches'. The beam is steered by 1232 Niobium-Titanium superconducting dipole magnets and collimated using quadrupole magnets. This magnet system is kept at a temperature below 2 K, using a pressurised bath of superfluid helium at about 0.13 MPa, and operates at fields above 8T. The LHC has three sophisticated vacuum systems: the insulation vacuum for cryomagnets, the insulation vacuum for helium distribution, and the beam vacuum.

It takes about 4 minutes and 20 seconds to fill up each of the LHC rings wih protons, and about 20 minutes for the proton beam to reach its current peak energy of 6.5 TeV. At this point, each LHC beam contains 2808 bunches with 1.5×10^{11} protons per bunch, colliding at a center of mass energy (COM) of 13 TeV. It is anticipated for the COM energy to increase to 14 TeV in 2018. Looking for physics beyond the standard model by colliding protons at such high energies is one of the primary aims of the LHC.

Another important parameter for a collider like the LHC is the instantaneous luminosity (referred to as just luminosity in the following), $\mathscr{L}$. The number of events (N) generated per second for some processes is given by:

$$\frac{dN}{dt} = \sigma \mathscr{L} \tag{3.1}$$

where σ is the cross-section of the processes. The luminosity of the LHC can be

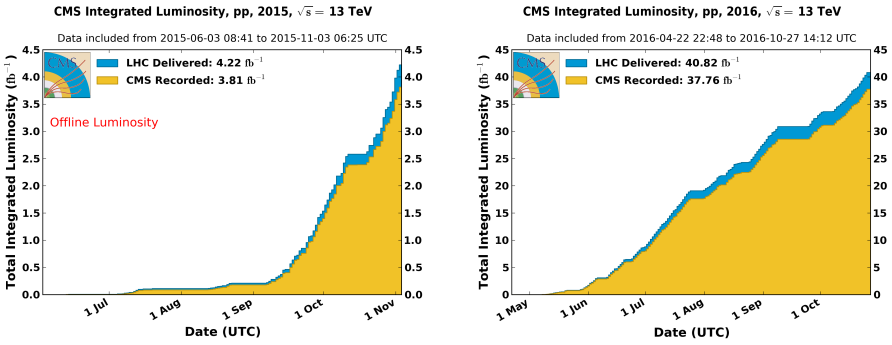


Figure 3.1: Evolution of integrated luminosity in 2015 and 2016 delivered by LHC (blue), and collected by CMS detector (orange) [2].

also expressed in terms of only beam parameters as:

$$\mathscr{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F \tag{3.2}$$

where N_b is number of protons in a bunch, n_b is number of bunches per beam, f_{rev} is the revolution frequency, γ_r the relativistic gamma factor, ϵ_n the transverse beam emittance, β^* the beta function at the collision point, and F is a reduction factor coming from the fact that the beams cross at an angle.

This luminosity intergrated over time represents the total number of events collected per unit cross section and is called the integrated luminosity (L).The LHC has already reached its nominal design luminosity of $10^{34} cm^{-2} s^{-1}$, and it has delivered data amounting to a more than $36 fb^{-1}$, only in 2016. Figure 3.1 shows the amount of data delivered by the LHC overlaid with the subset collected by the CMS detector in 2015 and 2016.

In the longer term, it is planned to keep the LHC running, punctuated with several scheduled stops for upgrades and maintenance, at least until late 2030s. During this period it is anticipated to operate at increasingly higher luminosities helping collect unprecedented amounts of data. Figure 3.2 shows an overview of the long term LHC schedule.

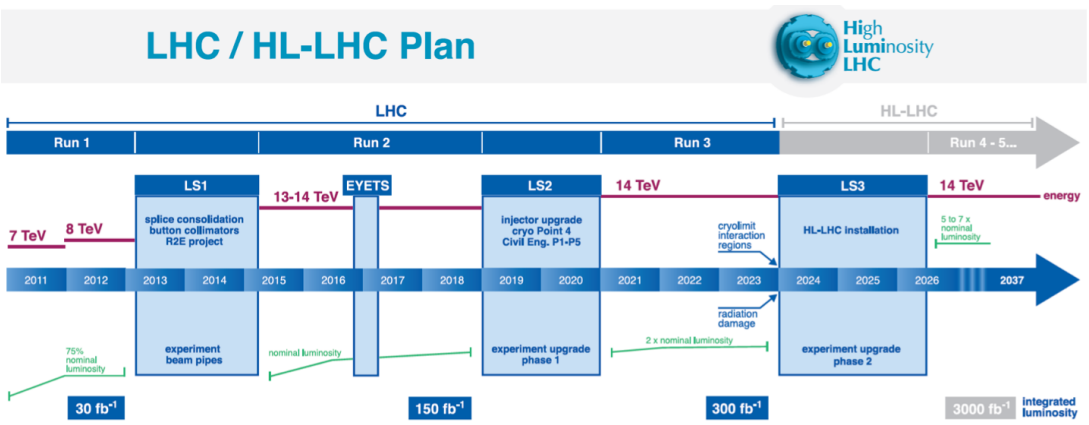


Figure 3.2: Overview of the long term LHC schecdule [3].

3.2 The CMS Detector

The Compact Muon Solenoid [4] is a general multipurpose particle physics detector that is placed in one of the four collision points of the LHC. It is 28.7m long with a diameter of 15.0m, weighs 14000 tons and is composed of several subdetectors. Its aim is to study a broad array of physics, from making precise measurements of known processess to searches for exotic processes predicted by a multitude of BSM theories. In order to be able to pursue its physics aims at the challenging LHC conditions, the CMS experiment needs to meet several requirements which primarily include good muon identification and momentum resolution over a wide range of momenta and angles, good dimuon mass resolution, good charged-particle momentum resolution and reconstruction efficiency, good electromagnetic energy resolution, good diphoton and dielectron mass resolution, good missing-transverse-energy and dijet-mass resolution. The backbone of the CMS is a superconducting solenoid that houses its tracking and calorimetry systems and provides an axial magnetic field of 3.8T. The inner-most layer is the silicon pixel and strip tracker that measures the trajectories of charged particles. Surrounding the tracker are the lead tungstate crystal electromagnetic calorimeter (ECAL) which measures the energy of electrons and photons,

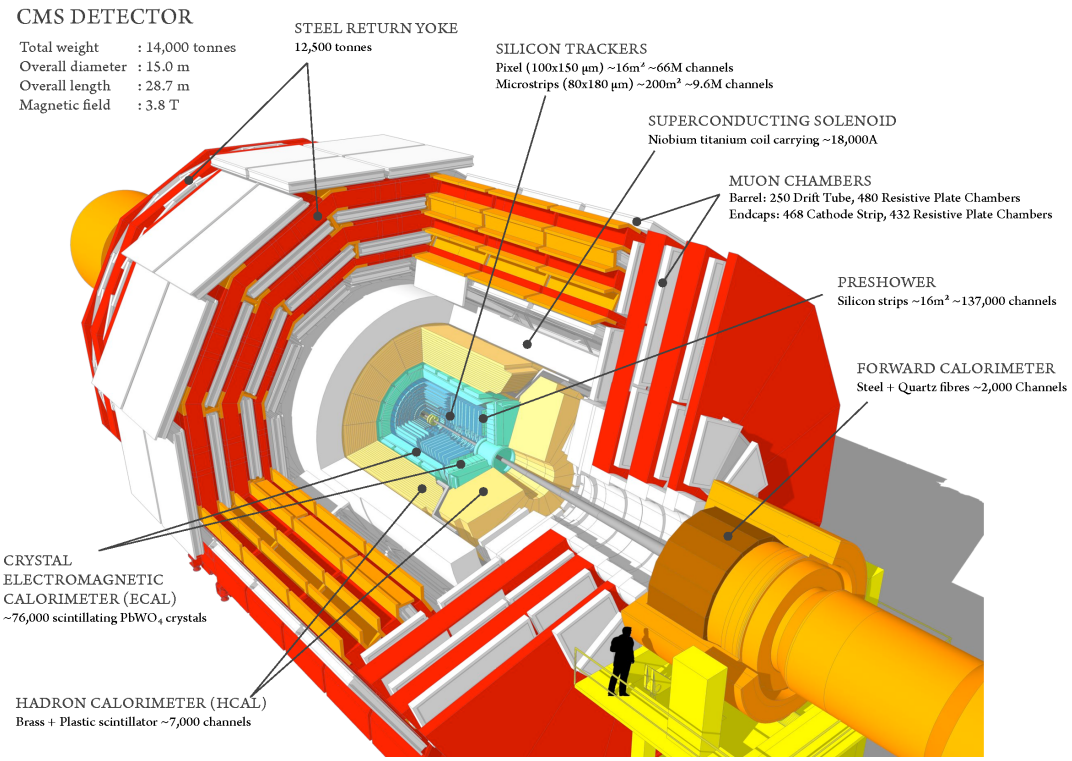


Figure 3.3: Layered View of the CMS detector

and the hadronic calorimeter (HCAL) which measures the energy of heavier particles that pass through the ECAL. The ECAL also contains a preshower detector for extra spatial precision. Outside the solenoid is the muon system which has gas-ionization detectors placed in the steel yoke of the magnet. This is the outermost component of CMS and measures the momenta of muons that traverse through it. A sophisticated two-level trigger system that helps filter out a small fraction of most interesting events among millions produced at the LHC also forms a vital part of the CMS.The powerful solenoid, sophisticated muon system and its compact design (given its complexity) give CMS its name. Figure 3.3 shows a layered view of the detector. The following sections describe it in further detail.

3.2.1 Coordinate Conventions

The CMS detector has adopted a right-handed coordinate system, the origin of which lies at the nominal collision point inside the experiment.The x-axis points radially inward towards the center of the LHC while the y-axis points vertically upwards. This makes the z-axis point along the beam direction. At point 5 of LHC (a village named Cessy in France) where the CMS is, the z axis points toward the Jura Mountains. In cylindrical co-ordinates, the polar angle θ is measured from the z-axis while the azimuthal angle ϕ is measured from the x-axis in the x-y plane. The polar angle is used to define the pseudo-rapidity $\eta = -ln(tan(\frac{\theta}{2}))$ which is a close approximation for rapidity if $E \gg m$. The rapidity is a Lorentz invariant quantity under boosts in the z-direction. Since it is typical of particles that CMS sees to have $E \gg m$, the Lorentz invariance approximately holds for pseudo-rapidity as well.

3.2.2 CMS Trigger

3.2.3 Charged Particle Tracking System

3.2.4 Electromagnetic Calorimeter

3.2.4.1 ECAL trigger project

3.2.4.2 Anomaly detection project for ECAL DQM

3.2.5 Hadronic Calorimeter

3.2.6 Muon System

CHAPTER 4

OBJECT RECONSTRUCTION AND EVENT GENERATION

4.1 Introduction

4.2 Physics Object Reconstruction

4.2.1 Particle Flow

4.2.2 Track Reconstruction

4.2.3 Electron Reconstruction

4.2.4 Muon Reconstruction

4.2.5 Jet Reconstruction

4.2.6 MET, MT and Collinear Mass

4.2.7 Tau Lepton and others

4.3 Datasets

The data analysed in this search was gathered by the CMS detector in 2016 during proton-proton collisions at the LHC, corresponding to an integrated luminosity of $35.9 fb^{-1}$. This data corresponds to a center-of-mass energy of 13 TeV and a spacing of 25ns between bunch crossings in the LHC with an average of about 30 collisions per bunch crossing. The subset of samples used among all collected by CMS are the ones having at least one isolated muon having transverse energy over 24 GeV, as triggered by the CMS high level isolated muon trigger (HLT_IsoMu24 in CMS parlance).

4.4 Monte Carlo Generation

Monte Carlo simulated SM Higgs Boson events produced by gluon fusion (GF) [5], vector boson fusion (VBF) [6] and associated production (production in association with a W or Z boson) [7] mechanisms and decaying into a muon and a tau lepton are used as signal samples for the analysis.POWHEG [8–13] is used to produce these samples. It is also used to simulate $t\bar{t}$ and single-top quark production processes. Several other event generators were used to simulate various other background processes for the analysis. MADGRAPH5.1 [14] is used for Z+Jets, W+Jets and Wγ processes and also for diboson production. All the event generators are interfaced to PYTHIA 8.1 [15] for the showering of partons and hadronization, as well as including a simulation of the underlying event (UE) and multiple interaction (MPI) based on the CUET8PM1 tune [16]. After the generation step, the detector response is simulated using a detailed description of the simulated detector based on the GEANT4 [17] package.

The protons circulate inside the LHC not as a continuous beam but in discrete closely packed bunches. This leads to more than one proton-proton collision per bunch crossing, i.e. pileup both in-time and out-of-time (see chapter 3). As mentioned in the chapter on event simulation (4), additional pileup interactions are also a part of the MC generation pipeline. All simulated samples are reweighted to the pileup distribution observed in data. An event weight is applied based on the number of simulated pileup events and the instantaneous luminosity per bunch-crossing, averaged over the run period.

Several other scale factors (see chapter 4 for details) are used to reweight the events in order to get the MC simulation to match the data closely. These include scale factors based on trigger, lepton identification, lepton isolaton and b-jet tagging efficiencies.

CHAPTER 5

EVENT SELECTION

5.1 Introduction

This chapter describes in detail the event selection criteria for the analyses, and how they were chosen. It starts by introducing the backgrounds th~~at each of selection~~ criterion is trying to reduce in order to get a higher ratio of number of signal events to background events, leading to a better sensitivity for the search. This is followed by the procedure for arriving at the best possible set of selection criterion. For the h $\rightarrow \mu\tau_e$, two methods of selection were developed. The first method developed involves placing requirements on several kinematic variables, and then using the resulting ditribution of M_{col} as discriminant for a binned likelihood fit (see section **??** for description of statistical procedures). We call this method M_{col} fit method. The second method developed involves using a Boosted Decision Trees (BDT) discriminator for classification of signal and background events. The output distribution of the BDT discriminator is then used to perform the fit. ~~We call this method BDT~~ method. The BDT method is found to have ~~greate~~r sensitivity, as discussed later in ~~the~~ chapter. However, the M_{col} fit method is also presented as a cross-check analysis. For H $\rightarrow \mu\tau_e$ analysis, only the M_{col} fit method is developed. This is in part due to the dificulties foreseen in training a BDT with much fewer events available in H $\rightarrow \mu\tau_e$ analysis, and in part since this is the very first time the H $\rightarrow \mu\tau_e$ search is being performed, a simpler analysis was felt to be adequate.

Both analyses were performed blinded [18] in the signal region. All selection criterion and methods described below were developed without the knowledge of the

observed data in the range of variable spectra where the signal is expected to be present. This is considered an optimal way of eliminating the unintended biasing of a result in a particular direction and is a standard methodology in particle physics analyses.

5.2 h125: $h \to \mu\tau_e$ analysis

5.2.1 $h \to \mu\tau_e$: Final state signature and backgrounds

The signature of the $h \to \mu\tau_e$ analysis final state consists of a muon that comes promptly from the Higgs and has a hard p_T spectrum, along with a softer electron of opposite sign charge that comes from the tau lepton, and missing transverse momentum from the tau decay. It is interesting to note that the signature is similar to the $h \to \tau_\mu\tau_e$ decay that is allowed by the SM and since been observed [], but with significant kinematic differences. In $h \to \mu\tau_e$ decay the μ comes directly from the Higgs resulting in its p_T spectrum peaking and spreading out to much higher values. Also there are fewer neutrinos in $h \to \mu\tau_e$, coming from the decay of the single τ. The decay products of this highly boosted tau are closely aligned, leading to a narrow separation between the e and the $\vec{p}_T^{\,miss}$ in the azimuthal plane. The same is not true in the $h \to \tau_\mu\tau_e$ decays. These differences are illustrated pictorially in Fig. 5.1.

The most dominant backgrounds consists of $Z \to \tau\tau$ events coming from Drell-Yan production and $t\bar{t}$ production. In $Z \to \tau\tau$ events, one τ can decay to an e and the other to a μ. This background peaks at lower values of M_{col} than the signal events but there is significant overlap with the signal spectrum. In $t\bar{t}$ production, each of the top quarks can decay into a bottom and a W with the W bosons then decaying to a e and μ. The other backgrounds are smaller and include (in no particular order) electroweak diboson production (WW, WZ and ZZ), h boson decays allowed by the SM ($H \to \tau\tau, WW$), $W\gamma^{(*)}$ + jets ,single top production, W + jets events, $Z \to \ell\ell$

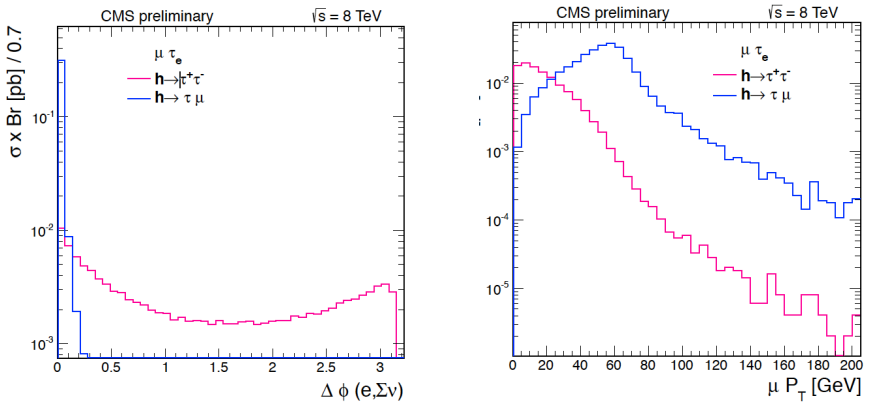


Figure 5.1: Illustration of the differences in p_T^{μ} and $\Delta\phi(e, \vec{p}_T^{miss})$ spectrums in $h \to \mu\tau_e$ and $h \to \tau_\mu\tau_e$ processes.

$(\ell = e, \mu) +$ jets and QCD multijet backgrounds. These backgrounds are described in more detail, along with there estimation and validation techniques in section 6.

5.2.2 $h \to \mu\tau_e$:Baseline selection and categorization

A baseline selection is defined first in order to ensure that we have clean and well-defined events faithful to the final state signature of the signal process. An isolated and well-identified μ is thus required to be present along with an well-identified and isolated e of opposite sign charge. They are required to be separated by $\Delta R > 0.3$. The identification criterion applied for μ and e have been described in sections 4.2.4 and 4.2.3. Isolation criterion, as measured by I_{rel} (described in 4.2.7), are required to have values $I_{rel}^{e} < 0.15$ and $I_{rel}^{\mu} < 0.1$. The p_T of these candidates are required to be above minimal thresholds required by trigger, identification and isolation requirement. Both candidates are also required to be within the fiducial region of the detector. The μ is required to have $p_T^{\mu} > 26\,GeV$ and $|\eta^{\mu}| < 2.4$.The e is required to have $p_T^{e} > 10\,GeV$ and $|\eta^{e}| < 2.3$. Only events with two or fewer jets are considered. All jets considered must have $p_T > 30\,GeV$, $|\eta| < 2.4$ and satisfy the loose identification criterion described in section 4.2.5. Events with one or more jets arising from a

b-quark (b-tagged jets) are vetoed. Cleaning events with b-tagged jets reduce some contribution from backgrounds which give rise to b-quarks such has $t\bar{t}$ and single top. Also, as described in 4.2.5, any event with one or more jets within $\Delta R < 0.4$ of either lepton candidate is also rejected. Further, an event is rejected if it has additional μ or e, or any τ_{had} candidates. All the above baseline selection requirements have been summarized in Table 5.5. All the events were required to pass isolated muon triggers with a p_T threshold of 24 GeV. The trigger selection has been described in detail in section 3.2.2. The distributions of the M_{col} and several other kinematic variables after the baseline selection just described are shown in Figs. 5.2 and 5.3. These distributions act as the starting point for development of stricter kinematic selections looking at the different shapes signal and backgrounds for different variables.

Table 5.1: Baseline selection criteria for $h \to \mu\tau_e$ analysis.

Variable	μ	e
p_T	$> 30\,\text{GeV}$	$> 10\,\text{GeV}$
$\lvert\eta\rvert$	< 2.4	< 2.3
I_{rel}	< 0.15	< 0.1
Cleaning requirements		
$\Delta R(\mu, \text{e}) > 0.3$		
No additional μ, e or τ_{had}		
No b-tagged jets with $p_T > 30\,\text{GeV}$		
No jets with $\Delta R(\mu, jet) < 0.4$ and $p_T > 30\,\text{GeV}$		
No jets with $\Delta R(\text{e}, jet) < 0.4$ and $p_T > 30\,\text{GeV}$		

At this point the events are divided into several buckets, called categories. This is done on the basis of number of jets present in the event. In events with 2 jets the invariant mass of the di-jet system (M_{jj}) is also used for categorization. The topology of events containing different number of such jets can be different. For

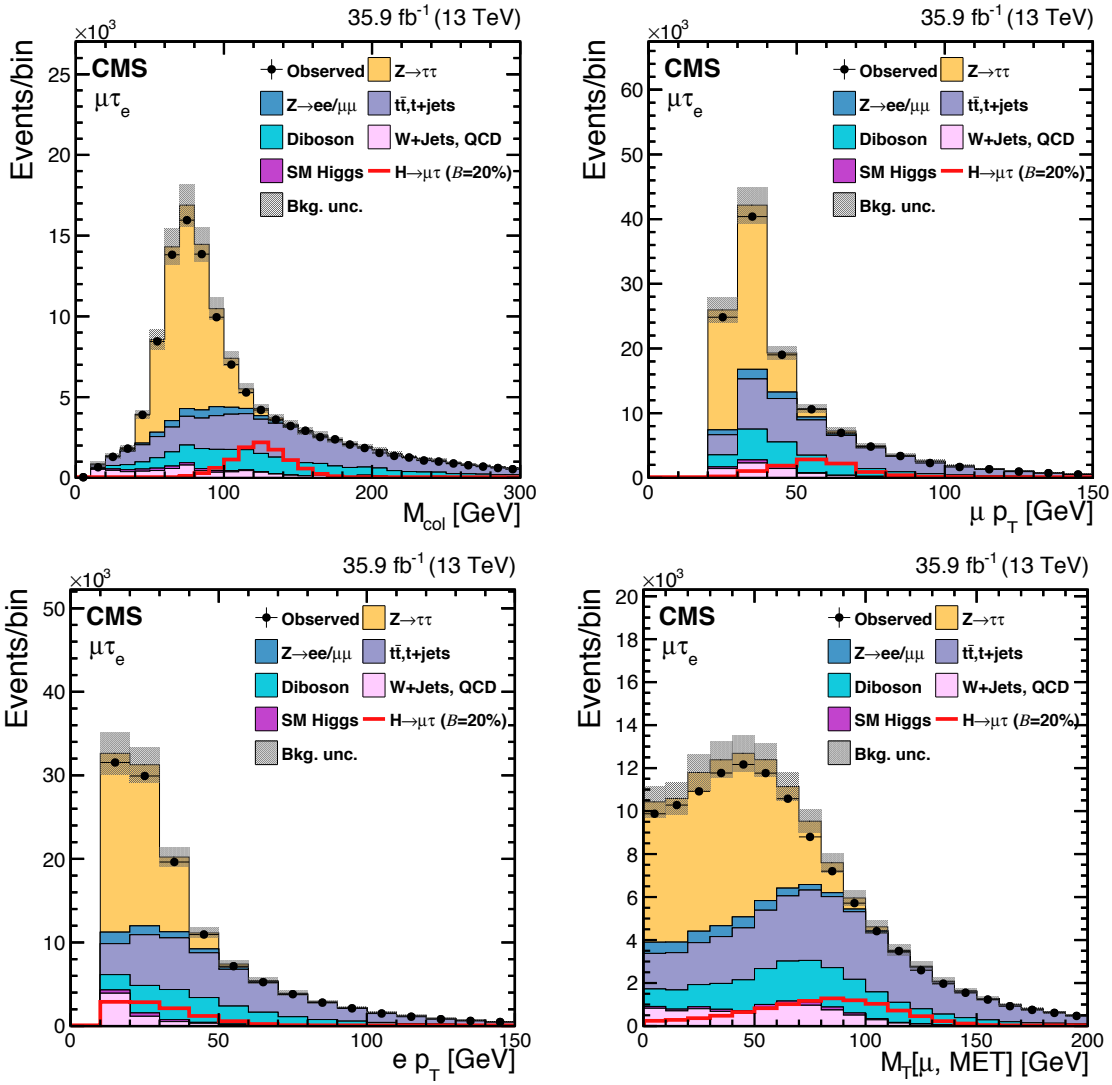


Figure 5.2: Distributions of kinematic variables after baseline selction for h $\rightarrow \mu\tau_e$ analysis (1).

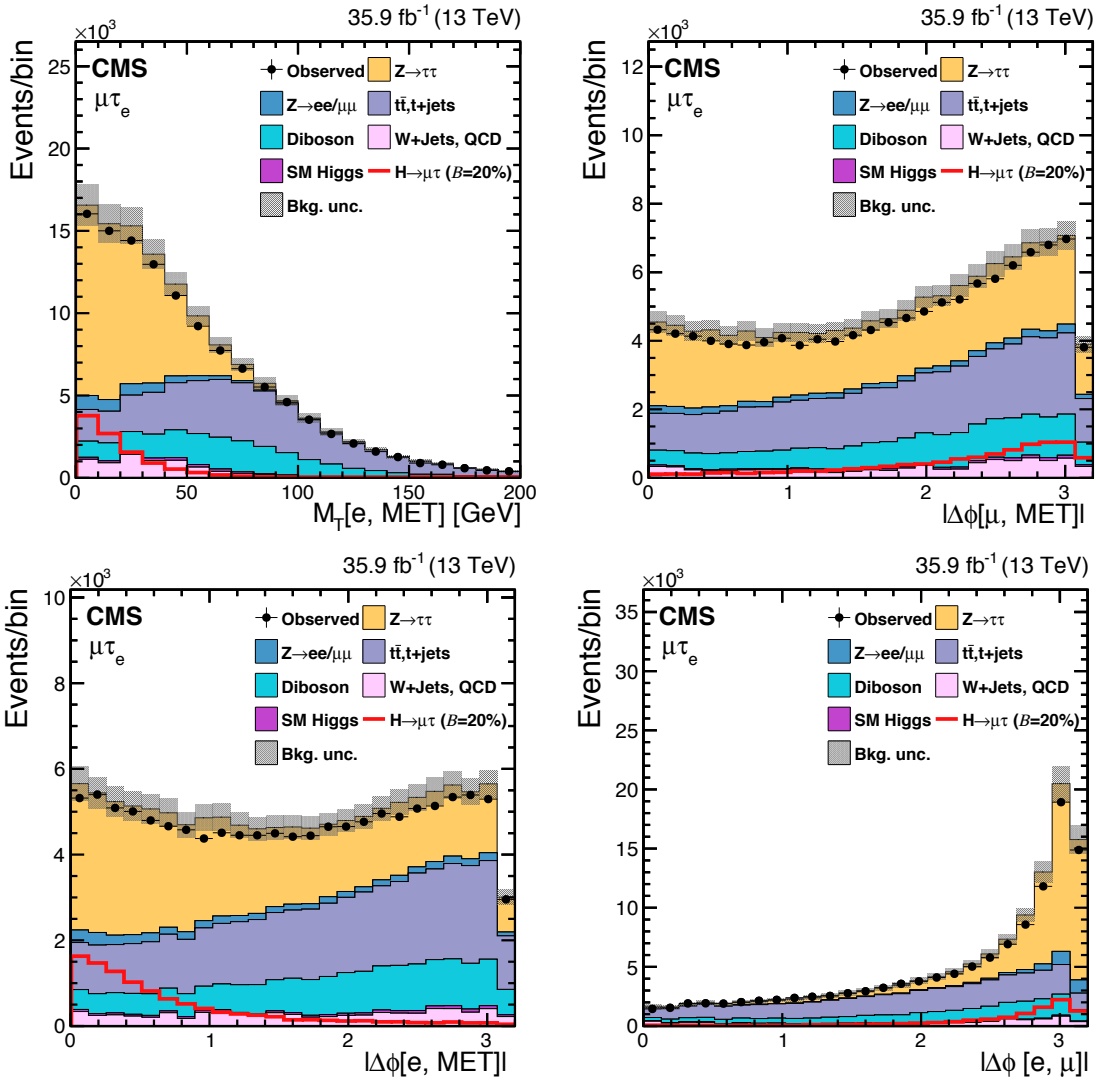


Figure 5.3: Distributions of kinematic variables after baseline selction for h $\rightarrow \mu\tau_e$ analysis (2).

example, in events with one energetic jet the h produced can be boosted resulting in the azimuthal separation of the μ and e (that come from its decay) to be narrower than events with no jets. Each of this categories enchance the contribution of different h boson production mechanisms, and requiring different optimal selection criteria in each category helps increase the sensitivity of the search. The categories in order of decreasing number of signal events are:

- **0-jet category**: These are events that do not have any jet. This category enhances the gluon-gluon fusion (GGF) contribution.

- **1-jet category**: Events that have 1 jet are put in this category. This category enhances the GGF production with initial state radiation (ISR). Some VBF events where one jet has escaped detection can also enter this category.

- **2-jet GGF category**: This category contains events that have 2 jets with the additional requirement that $M_{jj} < 550\,\text{GeV}$. The dominant contribution comes from GGF production in association with two jets.

- **2-jet VBF category**: This category contains events that have 2 jets with the additional requirement that $M_{jj} \geq 550\,\text{GeV}$. The dominant contribution comes from VBF production which is characterized by presence of two jets with high dijet mass.

5.2.3 $\text{h} \to \mu\tau_e$: M_{col} fit selection

In the M_{col} fit method, the selection is performed by placing kinematic cuts on several variables to enhance the signal-to-background ratio. There are several variables considered for this and they include: the azimuthal separation ($\Delta\phi$) between μ and e, between e and $\vec{p}_{\text{T}}^{\text{miss}}$, between μ and $\vec{p}_{\text{T}}^{\text{miss}}$, denoted respectively by $\Delta\phi(\text{e},\mu)$, $\Delta\phi(\text{e},\vec{p}_{\text{T}}^{\text{miss}})$, $\Delta\phi(\mu,\vec{p}_{\text{T}}^{\text{miss}})$, and the transverse mass between μ and $\vec{p}_{\text{T}}^{\text{miss}}$, between

e and $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, devnoted respectively by $M_T(\mu)$ and $M_T(\mathrm{e})$. The $\mathrm{h} \to \mu\tau_e$ decay being a 2-body decay, the μ and e are expected to be well separated in the azimuthal plane.Therefore, selecting events with a $\Delta\phi(\mathrm{e},\mu)$ larger than a threshold can help reject background events while keeping the signal that is peaked at high $\Delta\phi(\mathrm{e},\mu)$ values. This can be seen from Fig 5.3 (bottom right). Both neutrinos in the signal process come from the decay of the same τ. These neutrinos form the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. As mentioned earlier, the τ being much lighter than the h, it is highly boosted and its decay products i.e. e and the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ are expected to be close to each other in the azimuthal direction. Thus $\Delta\phi(\mathrm{e},\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})$ is expected to peak at values close to zero for signal events, as seen in Fig 5.3 (bottom left). Given that all backgrounds have relatively flat shape for this variable throughout the $\Delta\phi$ range, requiring $\Delta\phi(\mathrm{e},\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})$ to be lower than a threshold works as a strong rejection criterion against the backgrounds. Following, a similar line of reasoning, the μ is expected to be well separated from the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ resulting in $\Delta\phi(\mu,\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})$ for signal events to peak at high values, as seen in Fig 5.3 (top right). Further, as the $M_T(\ell)$ (defined in section 4.2.6) contains negative of the cosine of $\Delta\phi(\ell,\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})$ term, it is expected to be peak at values similar to $\Delta\phi(\ell,\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})$. This can be seen from Fig 5.3 (top left) and Fig 5.3 (bottom right) which show signal events for $M_T(\mu)$ and $M_T(\mathrm{e})$ peak at relatively higher and lower values than most backgrounds respectively. In particular, requiring $M_T(\mu)$ to be larger than a threshold can help reject a lot of $\mathrm{Z} \to \tau\tau$ events which is the most dominant background in the 0-jet category. All the above variables have a some amount of correlation with one another (see the correlation matrix shown in Fig. **??**. The optimization procedure used to arrive at the most optimal set of kinematic thresholds for these variables is described in detail in the next paragrhaph. The thresholds on the $\mathrm{p_T}$ of the μ and e have not been made stricter to avoid biasing the selection toward energetic leptons that sculpt the background M_{col} distribution to mimic the signal peak. This effect could potentially reduce the shape discrimination power of

the signal extraction procedure. Only in the 0-jet category category the requirement on p_T of the μ is made marginally stricter by requiring $p_T^{\mu} > 30\,\text{GeV}$. All other lepton p_T requirements are allowed to remain the same as baseline selection and are not included in the optimization procedure.

The aim of the optimization procedure is to maximize the sensitivity of the analysis. In other words, we want to select a set of thresholds which increases a quantity such as the $\frac{S}{\sqrt{S+B}}$ ratio where S and B are the number of estimated signal and background events respectively. It is also necessary to ensure alongwith that the entire spectrum of distribution of the discriminant variable (that is used int the final max-likelihood fit to extract results) is well-populated, escpecially in the region where the signal is expected to appear. A bad fit can potentially degrade the sensitivity of the analysis. Taking both of the above points into consideration the thresholds have been optimized to obtain the most stringent (lowest) possible expected limits. The definition and procedure of extacting the expected limit is given in section **??**). To do the optimization of the kinematic thresholds, we start by requiring the baseline selection. Then for a variable in consideration,e.g.- $\Delta\phi(\text{e}, \vec{p}_T^{\,\text{miss}})$, we look at the expected limit while making the threshold progressively stricter until we reach a point where making the threshold any stricter degrades (increases) the expected limit. We repeat this procedure for all variables and note the stringent expected limit for each (by tightening thresholds of only that variable). This concludes one round of the optimization. For the next round we start by requiring the baseline selection. In addition we require that the variable that achieved the best possible expected limit among all variables in the last round satisfy its corresponding threshold. Lets call this variable variable1. We now repeat the same procedure as the last round for all but variable1. Say the variable that gave us the best possible expected limit this round is variable2. For the start of the following round variable2 is is required to satisfy its corresponding threshold. Then all the other variables (including variables

that were had chosen thresholds in earlier rounds such as variable1 here) are made to go through the same procedure. This is done because the optimum value of threshold for variables chosen earlier might shift as new variables are chosen. This process is continued until the expected limit becomes no further stringent in sucessive rounds. This optimization was done separately for each of the four categories. The final set of thresholds arrived at in this way for the h $\rightarrow \mu\tau_e$ $M_{\rm col}$ fit analysis are listed in Table. 5.2. This method of choosing the optimal set of thresholds is sometimes called the n-1 procedure, and the idea is conceptually similar to forward/backward selection methods used in statistical learning to build optimal models.

TABLE 5.2

FINAL SELECTION CRITERIA FOR h $\rightarrow \mu\tau_e$ $M_{\rm col}$ FIT ANALYSIS.

Category	0-jet	1-jet	2-jet GGF	2-jet
$\mathrm{p}_{\mathrm{T}}^{\mu}$	$> 30\,\mathrm{GeV}$	–	–	–
$M_T(\mu)$	$> 60\,\mathrm{GeV}$	$> 40\,\mathrm{GeV}$	$> 15\,\mathrm{GeV}$	$> 15\,\mathrm{GeV}$
$\Delta\phi(\mathrm{e}, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})$	< 0.7	< 0.5	< 0.3	< 0.3
$\Delta\phi(\mathrm{e}, \mu)$	> 2.5	> 2.0	–	–

5.2.4 h $\rightarrow \mu\tau_e$: BDT method selction

In the BDT method, a boosted decision trees (BDT) classifier is used discriminate signal events from background events. A decision tree is classifier which works by building a tree structure based on binary splits (as shown in Fig. 5.4). Starting from

the root node of the tree (which contains all the events which we want to classify), a sequence of binary splits is made using input variables provided to the classifier. At each split, the variable which provides best purity of split or equivalently, in our case the best separation of signal and background events is used. The same variable can thus be used for splitting several nodes and the splitting is continued until a desired some stopping criterion such as depth of the tree, purity of leaf nodes , minimum number of events in a leaf node etc. is reached. All events end up in one of the leaf nodes. If an event ends up in a leaf node in which signal events form the majority fraction it is classified as a signal event. Otherwise, it is classified as a background event. Boosting is a class of ensemble machine learning techniques which help in enhancing performance of weak classifiers by sequentially building classifiers using reweighted (boosted) versions of the training data and then taking a weighted majority vote of the sequence of classifiers thus produced. Boosting also stabilizes the response of the classifiers with respect to fluctuations in the training data. In other words it helps avoid overfitting to the training data. When the boosting technique is a applied to produce an ensemble of decision trees the resulting ensemble of classifiers is called a Boosted Decision Trees classifier. A detailed overview of how decision trees and boosting works, and the chosen value of parameters used in training the BDTs for this analysis is given in appendix A.

The BDT is trained using events that satisfy the baseline selection criteria. Simulated GGF and VBF events weighted by their cross-section are used as signal events for training. For background, a mixture of $t\bar{t}$ and Drell-Yan events are used, also weighted by their respective cross-sections. The $t\bar{t}$ and Drell-Yan backgrounds are the most dominant backgrounds. The Drell-Yan background is the most dominant background in 0-jet and 1-jet category, while the $t\bar{t}$ background is the most dominant in both 2-jet categories. It also has many kinematic characteristics in common with diboson and single-top bankgrounds. A suite of input variables is used in training of

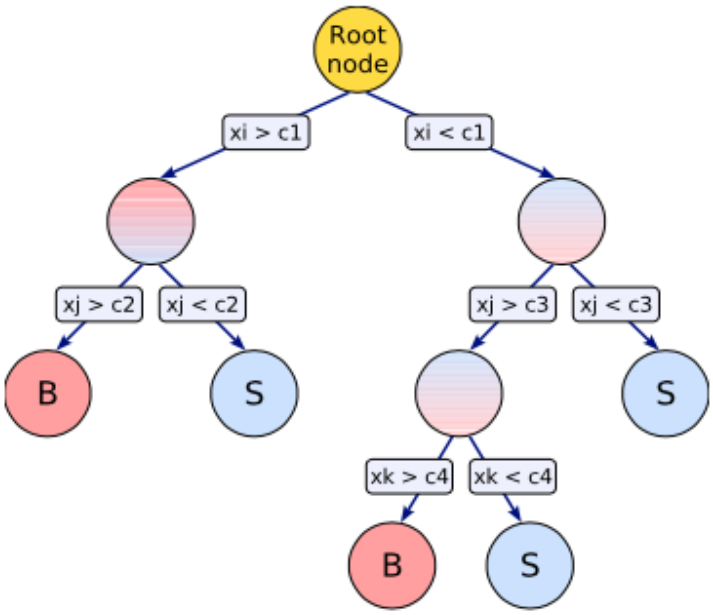


Figure 5.4: Illustration of decision tree. [19]

the BDT. They are as follows:

- Transverse mass between the μ and $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$: $M_T(\mu)$.
- Transverse mass between the e and $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$: $M_T(\mathrm{e})$.
- Azimuthal angle between the e and μ: $\Delta\phi(\mathrm{e},\mu)$.
- Azimuthal angle between the e and $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$: $\Delta\phi(\mathrm{e},\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})$.
- Azimuthal angle between the μ and $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$: $\Delta\phi(\mu,\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})$.
- Collinear mass: M_{col}.
- Muon $\mathrm{p_T}$: $\mathrm{p}_{\mathrm{T}}^{\mu}$.
- Electron $\mathrm{p_T}$: $\mathrm{p}_{\mathrm{T}}^{\mathrm{e}}$.

The distributions of these variables normalized to the total number of events in the input sample to the BDT is shown in Fig. 5.5. The correlations between these variables in signal and background events are shown in Fig. 5.6.

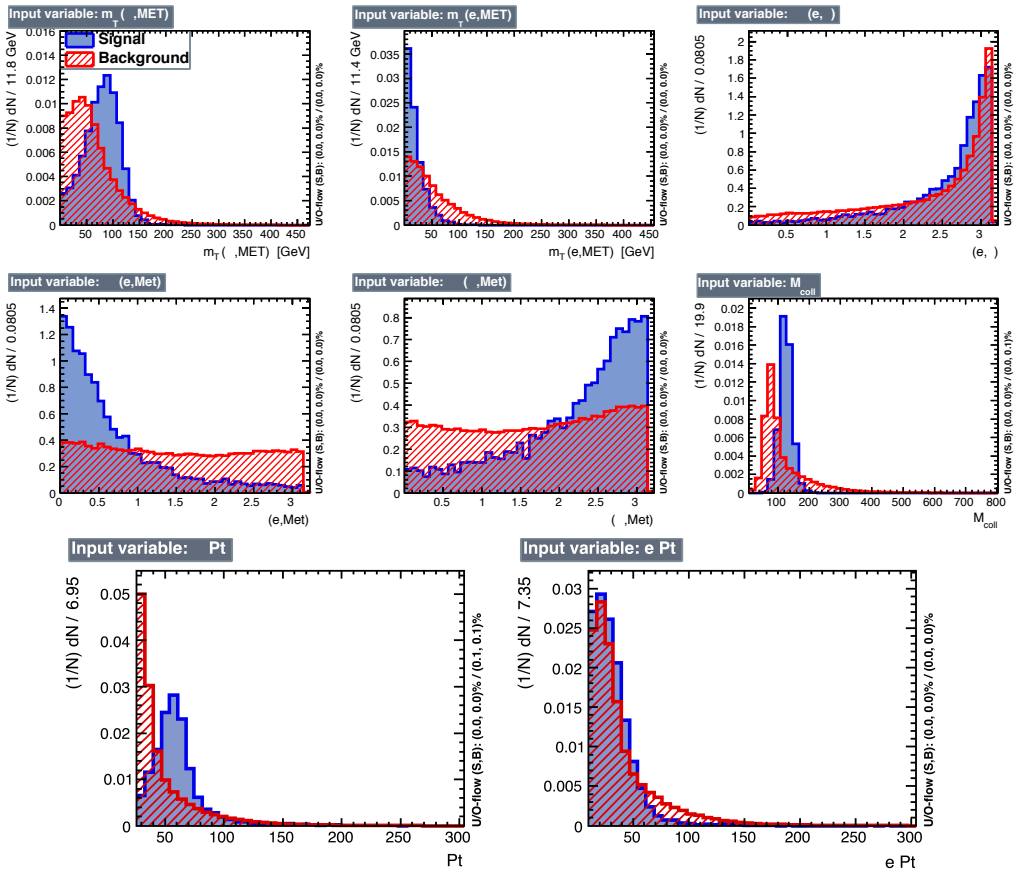


Figure 5.5: Normalized distributions of the input variables for BDT method. The signal (blue) is composed of a weighted mixture of GGF and VBF events, whereas the background (red) is made of $t\bar{t}$ and Drell-Yan events. All events were required to satisfy the baseline selection criteria.

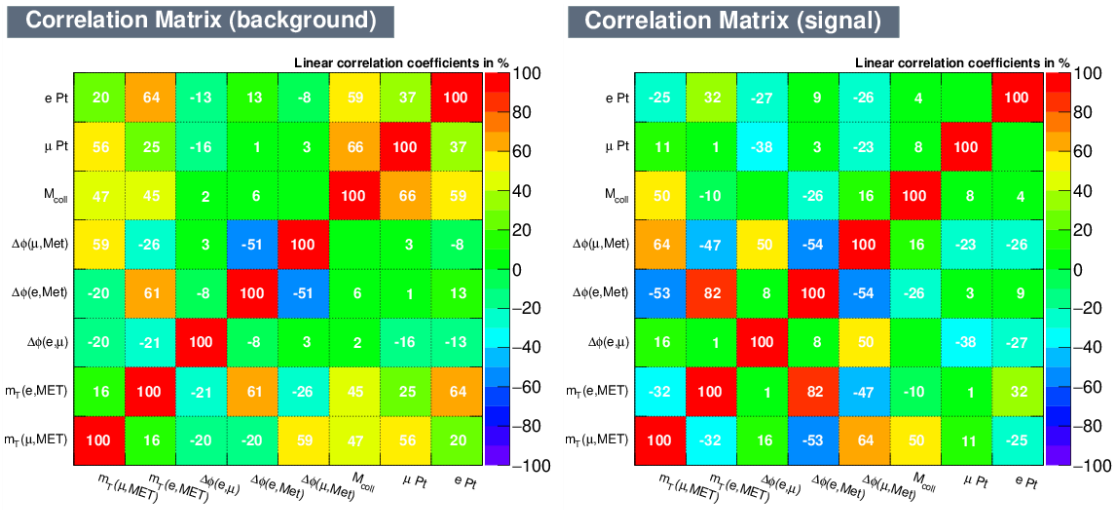


Figure 5.6: Correlations between input variables for signal events (right) and background events (left).

The training was done with a 800 decision tree ensemble, each tree having a maximum depth of 4. The gini-index criterion was used for splitting the data at each node. Further, AdaBoost (adaptive bossting) method was used for boosting (see appendix A for details of these techniques). A training to testing split of 70:30 split was used. Fig. 5.7 shows the distribution of the BDT response for training and testing samples. The training and testing distributions for both signal and background events match well suggesting that there is no overtraining. The distribution of BDT response is used in max-likelihood fit to extract results, as discussed in section 7.2.3.

5.3 Heavy higgs:H $\rightarrow \mu\tau_e$ analysis

5.3.1 H $\rightarrow \mu\tau_e$: Final state signature and backgrounds

The signature of the H $\rightarrow \mu\tau_e$ analysis final state is very similar to that of h $\rightarrow \mu\tau_e$. It also consists of a muon that comes promptly from the Higgs and has a hard p_T spectrum, along with a softer electron that comes from the tau lepton, and missing transverse momentum from the tau decay. The p_T^{μ} spectrum is expected to be harder

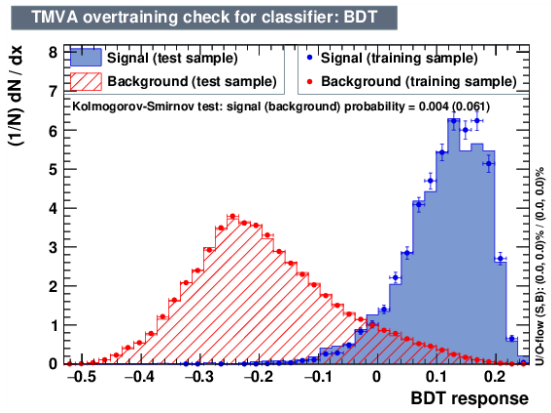


Figure 5.7: Distribution of BDT response for training (dots) and test(fill) distributions for both signal(blue) and background(red) events.

for higher H boson masses. The topologies being similar, the kinematic properties discussed in section 5.2.1 for h $\to \mu\tau_e$ analysis also apply to the H $\to \mu\tau_e$ analysis. The H boson mass peaks for all the simulated samples illustrated in Fig 5.8.

The most dominant backgrounds for H $\to \mu\tau_e$ consists of events from $t\bar{t}$ and electroweak diboson production. Unlike h $\to \mu\tau_e$ analysis, Z $\to \tau\tau$ events from Drell-Yan production a very small background as the Z $\to \tau\tau$ spectrum peaks at much lower values (around Z boson mass) of collinear mass that the signal events coming from heavy H boson decays. The other backgrounds come from h boson decays (H $\to \tau\tau, \mathrm{WW}$), W$\gamma^{(*)}$ + jets ,single top production, W + jets events, $Z \to \ell\ell$ ($\ell = \mathrm{e}, \mu$) + jets and QCD multijet backgrounds. These backgrounds are described in more detail, along with there estimation and validation techniques in section 6.

5.3.2 H $\to \mu\tau_e$: Baseline selection and categorization

The baseline selection for H $\to \mu\tau_e$ is similar to that of h $\to \mu\tau_e$ with the exception of higher $\mathrm{p_T}$ thresholds. Just like h $\to \mu\tau_e$, an isolated and well-identified μ is thus required to be present along with an well-identified and isolated e of opposite sign charge. They are required to be separated by $\Delta R > 0.3$. The identification and

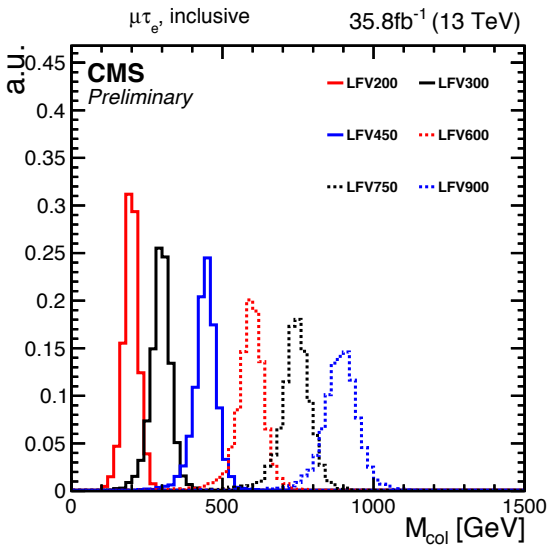


Figure 5.8: Illustration of simulated signal mass peaks for H $\rightarrow \mu\tau_e$ analysis for different H boson masses.

isolation criteria have been described in sections 4.2.4, 4.2.3 and 4.2.7. All events are required to pass a single muon trigger with the threshold of 50 GeV. The trigger selection has been described in detail in section 3.2.2. The μ is required to have $p_T^{\mu} > 53\,\text{GeV}$ and $|\eta^{\mu}| < 2.4$. The e is required to have $p_T^{e} > 10\,\text{GeV}$ and $|\eta^{e}| < 2.3$. Only events with zero or one jet are considered. Jets must have $p_T > 30\,\text{GeV}$, $|\eta| < 2.4$ and satisfy the loose identification criterion described in section 4.2.5 to be considered. As only GGF production mode is considered for the H $\rightarrow \mu\tau_e$ analysis, events with more than one jet make negligent contribution and are rejected. All other other criteria are same as h $\rightarrow \mu\tau_e$ analysis. The entire set of baseline selection criteria for H $\rightarrow \mu\tau_e$ has been summarized in table 5.3.

The events are then divided into categories, with motivations similar to the h $\rightarrow \mu\tau_e$ analysis (see section 5.2.2), on the basis of number of jets present in the event. The two categories for H $\rightarrow \mu\tau_e$ are:

- **0-jet category**: These are events that do not have any jet. This category enhances the gluon-gluon fusion (GGF) contribution.

Table 5.3: Baseline selection criteria for H $\to \mu\tau_e$ analysis.

Variable	μ	e
p_T	$> 53\,\text{GeV}$	$> 10\,\text{GeV}$
$\lvert\eta\rvert$	< 2.4	< 2.3
I_{rel}	< 0.15	< 0.1
Cleaning requirements		
$\Delta R(\mu, \text{e}) > 0.3$		
No additional μ, e or τ_{had}		
No b-tagged jets with $p_T > 30\,\text{GeV}$		
No jets with $\Delta R(\mu, jet) < 0.4$ and $p_T > 30\,\text{GeV}$		
No jets with $\Delta R(\text{e}, jet) < 0.4$ and $p_T > 30\,\text{GeV}$		

- **1-jet category**: Events that have 1 jet are put in this category. This category enhances the GGF production with initial state radiation (ISR).

The distributions of several kinematic variables after the baseline selection and categorization are shown in Figs. 5.9 and 5.10.

5.3.3 H $\to \mu\tau_e$: mcol fit selection

Just like the M_{col} fit method in h $\to \mu\tau_e$, the selection is performed by placing kinematic cuts on several variables to enhance the signal-to-background ratio. The variables considered are: $\Delta\phi(\text{e}, \mu)$, $\Delta\phi(\text{e}, \vec{p}_{\text{T}}^{\,\text{miss}})$, $\Delta\phi(\mu, \vec{p}_{\text{T}}^{\,\text{miss}})$, $M_T(\mu)$ and $M_T(\text{e})$. In addition, the p_T of the μ and e are also considered. Since we are looking for a decay in an extended mass range (200-900 GeV) in H $\to \mu\tau_e$, and not in a particular region like the h $\to \mu\tau_e$ analysis, the potential effect of background being mimic the signal, in particular due to higher p_T thresholds of the leptons is not apparent. The motivations for using these variables remain much the same like the h $\to \mu\tau_e$ analysis owing to similarities in topology. They are motivated by the facts that the only source of MET is the τ, and the τ being lighter than the H, its visible products are closely aligned, and the p_T spectrum of the prompt lepton (μ) is hard.

The procedure for optimization of the thresholds of for these variables is exactly

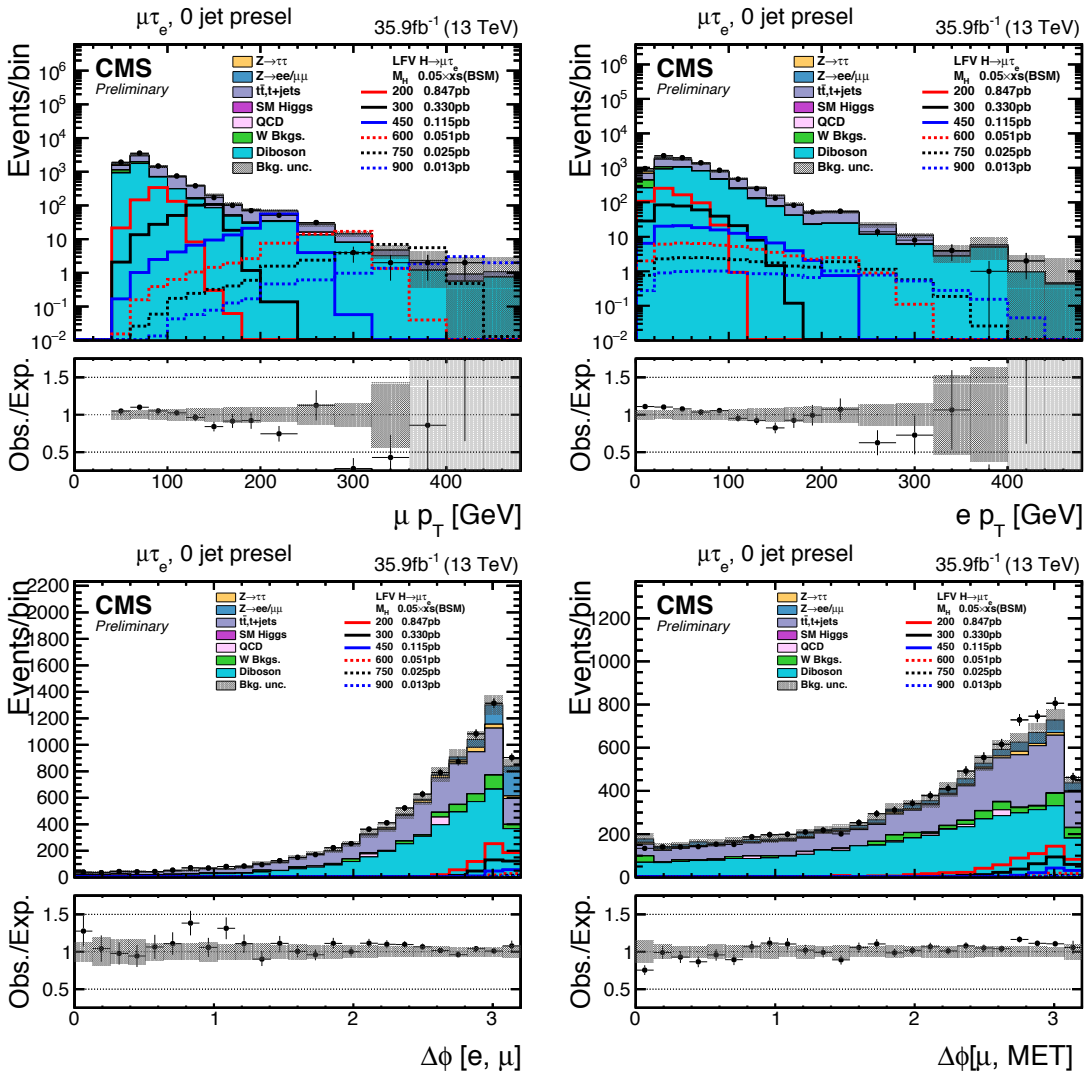


Figure 5.9: Distributions of kinematic variables after baseline selction for 0-jet category of H $\rightarrow \mu\tau_e$ analysis.

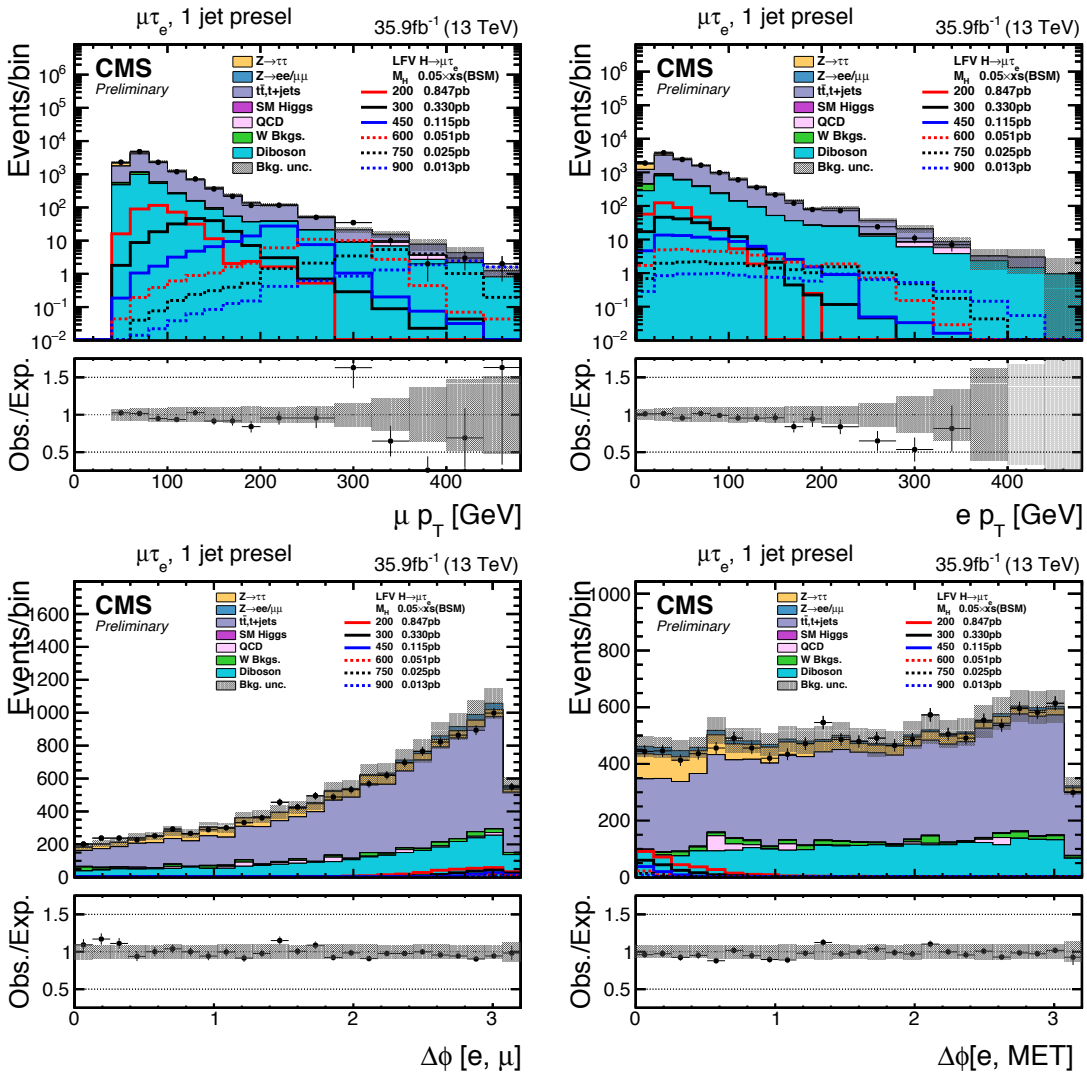


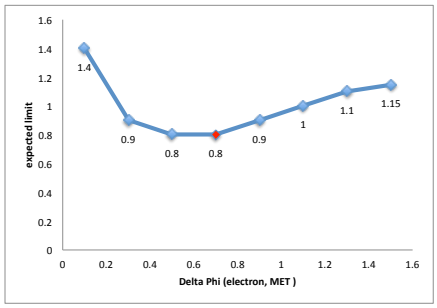
Figure 5.10: Distributions of kinematic variables after baseline selction for 1-jet category of H $\rightarrow \mu\tau_e$ analysis.

the same as described in section 5.2.3. Further to get better sensitivity in the entire mass range fom 200 to 900 GeV, two separate sets of thresholds are optimized, for each category. One set is optimized to provide better sensitivity in the 200-450 GeV mass range. The simulated signal for the H mass of 200 GeV is used when calculating expected limits during the optimization procedure for this mass range. The other set is optimized to provide better sensitivity 450-900 GeV mass range. The simulated signal for H mass of 450 GeV is used when calculating expected limits during the optimization procedure for this mass range. A few illustrations of the optimization procedure are shown in Fig. 5.11. The final set of thresholds arrived at in this way for both mass ranges and both categories for the $H \to \mu\tau_e$ $M_{\rm col}$ fit analysis are listed in Table. 5.4. The $M_{\rm col}$ distributions after requiring these selections is used in a max-likelihood fit to extract results, as discussed in section 7.2.3.

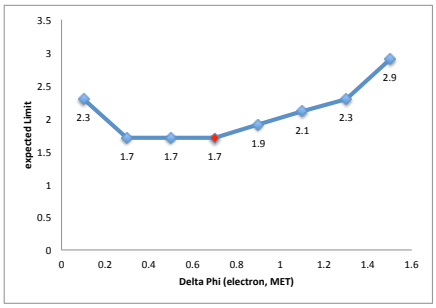
TABLE 5.4

FINAL SELECTION CRITERIA IN EACH CATEGORY OF THE $H \to \mu\tau_e$ ANALYSIS.

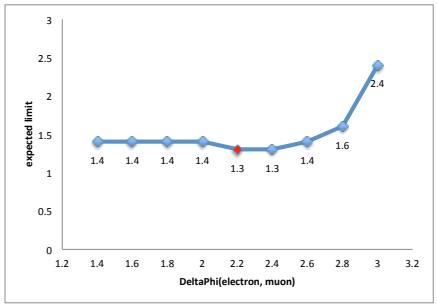
	Low mass range	High mass range
3*0-jet	$p_T^{\mu} > 60$ GeV, $p_T^{e} > 10$ GeV	$p_T^{\mu} > 150$ GeV, $p_T^{e} > 10$ GeV
	$\Delta\phi(e, \vec{p}_T^{\,miss}) < 0.7$	$\Delta\phi(e, \vec{p}_T^{\,miss}) < 0.3$
	$\Delta\phi(e, \mu) > 2.2$	$\Delta\phi(e, \mu) > 2.2$
3*1-jet	$p_T^{\mu} > 60$ GeV, $p_T^{e} > 10$ GeV	$p_T^{\mu} > 150$ GeV, $p_T^{e} > 10$ GeV
	$\Delta\phi(e, \vec{p}_T^{\,miss}) < 0.7$	$\Delta\phi(e, \vec{p}_T^{\,miss}) < 0.3$
	$\Delta\phi(e, \mu) > 2.2$	$\Delta\phi(e, \mu) > 2.2$



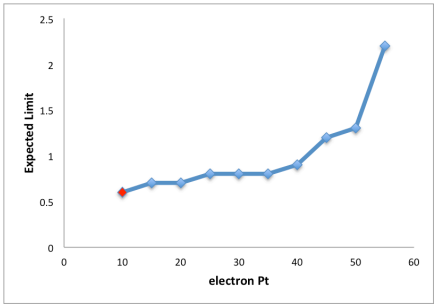
(a) Low mass range 0 jet $\Delta\phi(,\mathrm{p_T^{miss}})$



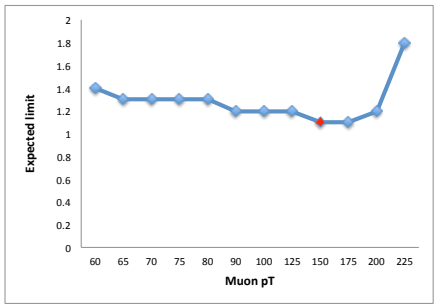
(b) Low mass range 1 jet $\Delta\phi(,\mathrm{p_T^{miss}})$



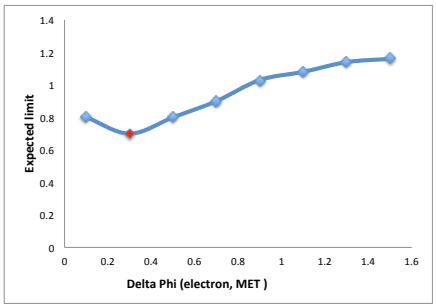
(c) Low mass range 0 jet $\Delta\phi(,\mu)$



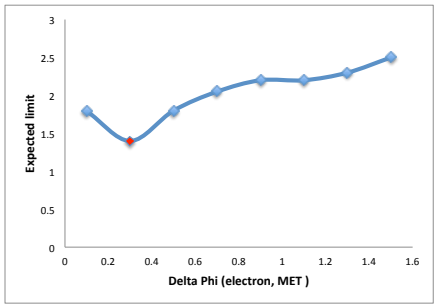
(d) Low mass range 0 jet $\mathrm{p_T^e}$



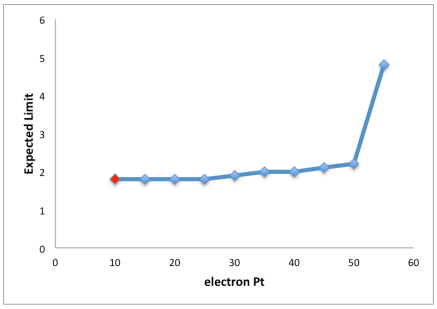
(e) High mass range 1 jet $\mathrm{p_T^\mu}$



(f) High mass range 0 jet $\Delta\phi(,\mathrm{p_T^{miss}})$



(g) High mass range 1 jet $\Delta\phi(,\mathrm{p_T^{miss}})$



(h) High mass range 1 jet $\mathrm{p_T^e}$

Figure 5.11. Examples of cut optimisation for the H $\rightarrow \mu\tau_e$ analysis

CHAPTER 6

BACKGROUND ESTIMATION AND VALIDATION

6.1 H125 Analysis

6.2 Heavy Higgs Analysis

CHAPTER 7

SIGNAL EXTRACTION AND SYSTEMATIC UNCERTAINTIES

7.1 H125 Analysis

7.1.1 Theoretical uncertainties

7.1.2 Experminetal uncertainties

7.1.3 Signal extraction

7.2 Heavy Higgs Analysis

7.2.1 Theoretical uncertainties

7.2.2 Experminetal uncertainties

7.2.3 Signal extraction

CHAPTER 8

INTERPRETATION OF RESULTS

CHAPTER 9

CONCLUSION

APPENDIX A

BOOSTED DECISION TREES

A.1 Introduction

BIBLIOGRAPHY

1. L. Evans and P. Bryant. LHC machine. In *Journal of Instrumentation*, volume 3, August 2008.

2. CMS Collaboration. CMS integrated luminosity - public results. Website. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`.

3. Joel Butler et al. Phase II Upgrade Scope Documentt. *CERN-LHCC-2015-019*, 2015.

4. CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3, August 2008.

5. H. M. Georgi, S. L. Glashow, M. E. Machacek, and Dimitri V. Nanopoulos. Higgs bosons from two gluon annihilation in proton proton collisions. *Phys. Rev. Lett.*, 40:692, 1978. doi: 10.1103/PhysRevLett.40.692.

6. Robert N. Cahn, Stephen D. Ellis, Ronald Kleiss, and W. James Stirling. Transverse-momentum signatures for heavy Higgs bosons. *Phys. Rev. D*, 35: 1626, 1987. doi: 10.1103/PhysRevD.35.1626.

7. S. L. Glashow, Dimitri V. Nanopoulos, and A. Yildiz. Associated production of Higgs bosons and Z particles. *Phys. Rev. D*, 18:1724, 1978. doi: 10.1103/PhysRevD.18.1724.

8. Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004. doi: 10.1088/1126-6708/2004/11/040.

9. Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070, 2007. doi: 10.1088/1126-6708/2007/11/070.

10. Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010. doi: 10.1007/JHEP06(2010)043.

11. Simone Alioli, Keith Hamilton, Paolo Nason, Carlo Oleari, and Emanuele Re. Jet pair production in POWHEG. *JHEP*, 04:081, 2011. doi: 10.1007/JHEP04(2011)081.

12. Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. NLO Higgs boson production via gluon fusion matched with shower in POWHEG. *JHEP*, 04:002, 2009. doi: 10.1088/1126-6708/2009/04/002.

13. E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini. Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM. *JHEP*, 02:088, 2012. doi: 10.1007/JHEP02(2012)088.

14. Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5: going beyond. *JHEP*, 06:128, 2011. doi: 10.1007/JHEP06(2011)128.

15. Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852, 2007. doi: 10.1016/j.cpc.2008.01.036.

16. Vardan Khachatryan et al. Event generator tunes obtained from underlying event and multiparton scattering measurements. *Eur. Phys. J.*, C76(3):155, 2016. doi: 10.1140/epjc/s10052-016-3988-x.

17. S. Agostinelli et al. GEANT4 — a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003. doi: 10.1016/S0168-9002(03)01368-8.

18. Aaron Roodman. Blind Analysis in Particle Physics . 2003. doi: arXiv:physics/0312102v1.

19. A. Hoecker et al. *TMVA - Toolkit for Multivariate Data Analysis*, 2017. `http://tmva.sourceforge.net/`.

This document was prepared & typeset with pdfLaTeX*, and formatted with* NDdiss2$_\varepsilon$ *classfile (v3.2017.2[2017/05/09])*