SEARCH FOR LEPTON FLAVOR VIOLATING DECAYS OF STANDARD MODEL HIGGS AND NEUTRAL HEAVY HIGGS BOSONS TO A MUON AND AN ELECTRONICALLY DECAYING 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

Nabarun Dev

Colin Philip Jessop, Director

Graduate Program in Physics

Notre Dame, Indiana

November 2019

SEARCH FOR LEPTON FLAVOR VIOLATING DECAYS OF STANDARD MODEL HIGGS AND NEUTRAL HEAVY HIGGS BOSONS TO A MUON AND AN ELECTRONICALLY DECAYING TAU LEPTON

Abstract

by

Nabarun Dev

This dissertation presents searches for Lepton Flavor Violating decay of the Standard Model Higgs Boson (h) as well as neutral Heavy Higgs Bosons (H) into a muon and an electronically decaying tau lepton. Data collected by the CMS detector in 2016, in proton-proton collisions at the LHC, at a center-of-mass energy of 13 TeV were used to perform both searches. No evidence of signal was found in either search, and stringent upper exclusion limits were set on both processes. Observed (expected) upper limits on the branching fraction of h decaying into a muon and a tau lepton, $\mathcal{B}(\mathrm{h} \rightarrow \mu\tau)$, was set at 0.25 (0.25) % at 95% CL. These are the most stringent limits set on this process till date. Observed (expected) upper limits on product of H cross-section and branching ratio to $\mu\tau$, $\sigma(\mathrm{gg} \rightarrow \mathrm{H}) \times \mathcal{B}(\mathrm{H} \rightarrow \mu\tau)$ range from 51.9 (57.4) pb to 1.6 (2.1) pb, at 95% CL, for H masses in the range between 200 and 900 GeV. This search is the first direct search to set limits on this decay [1, 2].

DEDICATED TO

My parents,

Prashanta and Sumitra,

who taught me everything I needed to know.

CONTENTS

FIGURES

TABLES

ACKNOWLEDGMENTS

I would like to begin by thanking my PhD advisor and mentor, Prof. Colin Jessop. Colin guided, encouraged and supported me throughout all my endeavors during my PhD, both during the time spent at Notre Dame and doing research abroad at CERN in Geneva. A lot of what I have learnt and grown to be, in terms of scientific method, collaborating with fellow researchers and communicating what I have learnt to an audience were due to Colin's guidance.

I would also like to thank my committee members Profs. Mike Hildreth, Randal Ruchti and Dinshaw Balsara for taking the time to review my research over the last few years, and providing feedback and encouragement. I would also like to thank Notre Dame and CMS faculty members, Profs. Kevin Lanon, Chris Neu, Nancy Marinelli, Sasha Ledovsky and others who advised me with my research from time to time. I am also thankful to Dr. Alexandre Zabi and Dr. David Petyt for their guidance and encouragement.

Special thanks are in order for Dr. Silvia Taroni and Dr. Maria Cepeda Hermida who worked hard to teach me the ropes of the trade. This includes everything an experimental particle physics student needs to learn from ground up, from programming to big data processing to preparing talks and posters to drafting physics papers.

I would also like to thank the Notre Dame Physics Department for nurturing and supporting my growth as a scientist, and the University of Notre Dame for being my home and family for the last several years. I would especially like to thank Susan, Shari and Shelley in the physics office for helping me smoothly through many tasks over the years. I would like to thank my friends and colleagues Fanbo Meng, Prasanna Siddireddy, Michael Planer, Charlie Mueller, Robin Luo, Matthias Wolf, Anna Woodard, Carlos Alvarado, Fatemeh Elahi, Nathan Kellams for helping me learn and making research fun. I would like to thank my classmates James Kelly, Anthony Ruth, Allison Reinsvold, Elizabeth Louden, Rodolfo Capedevilla, Laura Collins, Ek Adhikari, Matthew Hall, Alison Deatsch and fellow chemistry PhD student Patrick Louden for being my family at Notre Dame.

I would like to thank my parents, family and friends for being there for me over the years, believing in me and standing by me during hard times. I would specially like to thank my father for having encouraged me to do what I wanted during my formative years, and not to be influenced by the crowd, even if the chance for success was small.

It wouldn't make sense for me to thank my wife, Namrata, because I consider everything I do as much hers as mine.

Lastly, I would like to thank all the Yodas, Kenobis, scientists, technicians and engineers, not only associated with CMS and LHC, but every other experimentalist, theorist and Ewok whose collective work over millennia make little things like this thesis possible.

SYMBOLS

c	speed of light
h	Standard Model Higgs
$h125$	Standard Model Higgs
H	Heavy Higgs
m	mass
e	elementary charge
E	energy
$\mathrm{p_T}$	Transverse Momentum
M_{col}	Collinear Mass
M_{T}	Transverse Mass

CHAPTER 1

INTRODUCTION

> Modern science has been a voyage into the unknown, with a lesson in humility waiting at every stop. Many passengers would rather have stayed home.
>
> *Carl Edward Sagan*

The Standard Model (SM) is the most well-tested and elegant description of nature available today. The discovery of the Higgs Boson in 2012 [3–5] completed the SM. In the SM, elementary particles acquire mass from their interaction with the scalar Higgs field, the quantum of which is the Higgs Boson (h). This particle which had eluded physicists for years is a cornerstone of the SM, and in a way, was the last predicted missing piece associated with it. It was introduced, in 1964 by Brout, Englert, Higgs, Guralnik, Hagen and Kibble as a consequence of electroweak symmetry breaking, in order to explain how elementary particles could have mass without violating the gauge invariance of the SM [6–9].

It was nearly 50 years before the h was discovered. During this period many important discoveries such as W/Z bosons (1983 at UA1/UA2 collaborations at CERN) and the top quark (1995 CDF/D0 at FermiLab) were made. The excellent performance of the Large Hadron Collider (LHC) at CERN (European Organization of Nuclear Research) in delivering proton-proton collisions, and the excellent work by the CMS (Compact Muon Solenoid) and ATLAS (A Large Toroidal Apparatus) collaborations made possible the discovery of the Higgs Boson in 2012. Although the CMS and the ATLAS are large general purpose detectors aimed at studying a wide range of physics, their design was optimized to study the phenomenon of Electroweak Symmetry Breaking, making the discovery of the Higgs one of their primary aims. They started collecting data in 2010, and the h discovery was made using the data collected from 2010 to 2012. In 2013, Peter Higgs and Franois Englert, two of the physicists associated with the development of the theory, were jointly awarded the Nobel Prize in Physics.

This discovery was a significant step for particle physics, and while it put an end to the decades old search for the elusive h, it opened up a fertile sector for particle physicists to explore and understand. One of the very important tasks is to ascertain if the properties of the discovered h are indeed compatible with theoretical SM expectations. In fact, many studies since 2012 have found properties of the h such as the spin, couplings, and charge-parity (CP) assignment to be consistent with SM [10]. While more precise studies of the properties and couplings of the h is important, it also provides us with a portal to look for new physics Beyond the Standard Model (BSM). The SM, as mentioned above, is a remarkable theory that has stood the test of time. However, it is has its shortcomings and is not a complete theory. For example, the SM does not explain gravity and thus is inadequate as a candidate for an ideal "Theory of Everything". It also doesn't explain why the weak force is 10^{24} times stronger than gravity. This is usually known as the hierarchy problem. To address such shortcomings, many BSM theories have been proposed that modify the SM in such a way that they are consistent with existing observations, but at the same time try to address its imperfections. Many outcomes these theories predict are non-SM and the recently discovered h unlocks a pristine ground to look for these outcomes. In fact, the constraint on the branching fraction to non-SM decay modes of the h, derived from a combined study by CMS and ATLAS is B(non-SM) $<$ 34% at 95% confidence level (CL) [10]. Thus, a significant contribution from exotic (non-SM) decays is allowed in the BSM Higgs sector.

One such interesting process that is forbidden in the SM but occurs in many new physics scenarios is interactions between charged leptons that violate the conservation of Lepton Flavor. In particular, Lepton Flavor Violating (LFV) decays of the h are allowed by these theories, and could be realized in decays of the h, which is neutral, into two charged leptons of different flavor. In this dissertation, we describe a search looking for LFV decay of the h into a muon (μ) and a tau lepton (τ). The tau lepton is short-lived and can further decay hadronically (τ_h), into a electron, or into a muon. Since we can detect electrons better than tau leptons, the electronic channel leaves a cleaner signature in the detector. The same can be said of taus decaying into muons. However, there is a very large background coming from SM processes when both final state candidates are muons, which reduces the sensitivity of this channel. In particular, the search described here looks for this electronic channel signature of a LFV decay of h boson, i.e. $\mathrm{h} \to \mu\tau_e$. Indirect constraints on $\mathrm{h} \to \mu\tau$ exist through interpretations of measurements of processes such as $\tau \to \mu\gamma$ [11]. These constraints set weak limits on such decays allowing significant branching fractions; $Br(\mathrm{h} \to \mu\tau) < O(10\%)$ [12, 13]. A search was performed by CMS for $\mathrm{h} \to \mu\tau$ with proton-proton collision data at center-of-mass energy of 8 TeV, collected during run I (2010-12) of the LHC. This improved the above limits by an order of magnitude to $Br(\mathrm{h} \to \mu\tau) < O(1.51\%)$ at 95% confidence level [14]. However, an excess of events with a significance of $2.4\,\sigma$ was also observed. This warrants us to do this search with a larger amount of data which would either lead us to confirm this excess, or squash it and set much stricter limits on this process. The dataset collected by the CMS detector in 2016 provides us with such an opportunity. It corresponds to proton-proton collision data at a much higher center-of-mass energy of 13 TeV and is almost two times in size of the run I dataset. Besides using this larger dataset, the analysis described in this thesis improves upon previous searches by introducing multivariate techniques.

An interesting common feature of many of the models that allow LFV decays of the h is that they predict the existence of heavy neutral Higgs bosons, H (CP-even) and A (CP-odd). These are also expected to have LFV decays into charged leptons of different flavor [15]. A direct search for these channels would thus provide a complementary probe of these models. In this dissertation, we also describe such a search for heavy neutral Higgs boson (H) decaying in a lepton flavor violating manner into a muon and an electronically decaying tau, .i.e H $\rightarrow \mu\tau_e$. For this search, we probe H mass (m_H) in the range $200 < m_H < 900\,\mathrm{GeV}$, and use analysis techniques similar to the h $\rightarrow \mu\tau_e$ search. The lower bound is kept at(200 GeV) to avoid probing regions close to the 125 GeV SM h. The heavier Higgs we probe, the harder it becomes to describe the background in the region of the signal. The search gets limited by the amount of simulated and collision data we have. Considering these factors, and given that this search is the first ever direct search to look for this process, we made a (somewhat arbitrary) choice of probing masses up to 900 GeV. In future iterations of this search, it may be possible to probe much heavier H masses. This search is the first ever direct search to look for this process. In this entire document, we denote neutral heavy Higgs boson simply by H and SM Higgs boson as h.

The dissertation is devoted to the description of the h $\rightarrow \mu\tau_e$ and H $\rightarrow \mu\tau_e$ searches using the CMS experiment at the LHC. In chapter 2, we describe theoretical background and motivations for these searches. In the next chapter (3), we describe the experimental apparatus used for the search, i.e. the collider (LHC) and the detector (CMS). In the following chapter (4), the procedure for simulation of events and reconstruction of physics objects such as electrons, muons and jets are outlined. Chapter 5 describes the strategies followed to select events with the signal signature, and to increase the percentage of signal-like events in the sample thereby increasing the sensitivity of the searches. In chapter 6, estimation of background processes for both searches is outlined. Chapter 7 provides a description of the statistical methods used for signal extraction and setting of exclusion limits, and also the uncertainties associated with the searches. Finally, chapter 8 lays out the results of both the searches performed.

CHAPTER 2

THEORETICAL BACKGROUND

If I have seen further it is by standing on the shoulders of Giants.

Isaac Newton

2.1 Introduction

In this chapter, we describe the theoretical motivations that drive the searches described in this thesis. We start with a description of the standard model (SM), its particle content and interactions, and the Higgs mechanism. We then talk about the inadequacies of the SM, and the existence of physics beyond the standard model (BSM). We then outline a few BSM models and how they point towards the possible existence of the decays that we search for in this thesis.

2.2 The Standard model

The SM is the result of human endeavors over centuries to understand what we and the world around us are made of, and capture those ideas in beautiful mathematical form. Our understanding of the world around us has refined progressively from the ancient times, when best tools of observation we had were nothing but our own eyes, to the current day when we are able to collide particles that make up matter at unprecedented speeds, and have sophisticated tools like the CMS detector to aid us. From the ancient greeks who pondered over philosophical questions about what the basic elements of nature were, to the discovery of electron in 1898 by J.J.Thompson, to Rutherford's famous gold foil experiment, to the discovery of the neutron by James Chadwick in 1932, each event has been a stepping stone towards our understanding of nature and the formulation of SM [16]. During the course of its formulation and after, the SM has accurately explained phenomena already known and predicted the existence of particles that were discovered later. The last of these particles is the Higgs Boson (h), discovered in 2012 at CERN by the CMS and the ATLAS experiments [3–5]. The SM is a gauge theory, in which three of the four known natural forces (strong, electromagnetic, weak and not gravity) are represented by the SU(3)×SU(2)×U(1) symmetry group. This symmetry group describes under which transformations the SM is invariant. By Noether's theorem each of the above symmetries associated with the SM Lagrangian is associated with a conserved quantity: color charge, weak isospin and electric charge. The following describes the elementary particles of the SM, the interactions among these and finally, the spontaneous symmetry breaking mechanism.

2.2.1 Elementary particles

There are two kinds of elementary particles in the SM. They are characterized by the intrinsic angular momentum that they carry, i.e. by their spin. Fermions, which have half-integer spins, form the building blocks of matter. Bosons, which have integer spins, are the force-carriers or mediators of interactions.

2.2.1.1 Fermions of SM

Fermions we described here are fundamental particles, i.e. they cannot be broken down into further constituents. The space-time evolution of the fermions is described by the Dirac equation and their behavior follows Fermi-Dirac statistics. All fermions are subject to the Pauli exclusion principle. They can be further categorized into two classes depending on their interaction with the strong force. Fermions which do not interact with the strong force are called leptons, and do not carry any color charge. Quarks carry color charge and interact via the strong force. Both leptons and quarks are further classified into three generations. Each lepton generation consists of a lepton and a neutrino while each quark generation consists of a up-type and a down-type quark. These are outlined in detail below.

Leptons comprise of the familiar electron (e), and its heavier cousins – muon (μ) and tau lepton (τ), which carry the same negative electric charge as the electron ($1.6 \times 10^{-19}C$). The heavier leptons τ ($\sim 1.8\,\mathrm{GeV}/c^2$) and μ ($\sim 105.7\,\mathrm{MeV}/c^2$) have short lifetimes of $\sim 2.9 \times 10^{-13}\,\mathrm{s}$ and $\sim 2.2 \times 10^{-6}\,\mathrm{s}$ respectively. They eventually decay into an electron which is the lightest lepton ($\sim 0.5\,\mathrm{MeV}/c^2$) and has infinite lifetime, or lighter hadrons. In the CMS detector, the μ survives long enough to reach the muon systems, and is thus detected as its own distinct signature. The τ on the other hand, owing to its extremely short lifetime, can travel only a very short distance ($\sim< 10\,mm$) before decaying. Thus, only decay products of tau leptons are able to be directly detected by CMS. Each charged lepton is associated with an electrically neutral neutrino. They are called electron neutrino (ν_e), muon neutrino (ν_μ) and tau neutrino (ν_μ). Because neutrinos carry no electric charge, they do not interact via electromagnetic interaction. This means the only way they interact is via the weak interaction. This makes neutrinos very difficult to detect. In particular, they pass through the CMS detector effectively without interacting at all, and their presence and the energy they carry can only be estimated using imbalance in transverse momentum of observed particles (see section 4.4.7). The three generations of leptons are pictorially shown below.

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}$$

Quarks come in two generations: up-type and down-type. The up-type quarks are the up quark (u), charm quark (c) and top quark (t). Their down-type counterparts are down quark (d), strange quark (s) and bottom quark (b). Each up-type quark carries a positive electric charge of 2/3 times the charge of the electron. Each down-type quark carries a negative electric charge of 1/3 times the charge of the electron. Just like the leptons, each progressive generation is heavier with the third generation consisting of the top and bottom quarks being the heaviest. In fact, the top quark was the last of the SM fermions to be discovered in 1995, and is the most massive particle in the SM ($\sim 173\,\mathrm{GeV}/c^2$). As mentioned above, all quarks carry color charge. Color charge is to strong force as electric charge is to electromagnetic force. This allows quarks to interact via the strong force. Due to a phenomenon called color confinement, quarks aggregate together into color singlets (having zero color charge) particles called hadrons. Hadrons are either formed of 3 (anti-) quarks (baryons) or 2 (anti-) quarks (mesons). The proton and neutron are baryons. The proton is made of two up quarks, and one down quark. It has a mass of $\sim 938.3\,\mathrm{MeV}/c^2$ and is stable (infinite lifetime). The neutron is made of one up quark and two down quarks. It has a mass of $\sim 939.5\,\mathrm{MeV}/c^2$ and has a lifetime of $\sim 880\,\mathrm{s}$. The three generations of quarks are pictorially shown below.

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix}$$

Each particle described above has an anti-particle associated with it. Particles (matter) and their anti-particles (anti-matter) are almost identical except they have opposite physical charges (electric charge, lepton number, baryon number). For example, the anti-particle of an electron is the positron which is nearly identical to the electron except for the fact that it has positive electric charge.

2.2.1.2 Bosons of SM

The bosons in SM are carriers or mediators of force. Their behavior follows Bose-Einstein statistics and they are not constrained by the Pauli exclusion principle. The strong interaction, as its name suggests, is the strongest of the fundamental forces (see table 2.1). The eight gluons mediate the strong interaction between particles with color charge. Photons are the mediators of the next strongest fundamental force, the electromagnetic force. Gluons and photons are massless, electrically neutral and have spin 1. Additionally, gluons carry color charge. This is in contrast to photons which are electrically neutral. The W^+,W^- and Z gauge bosons mediate the weak interaction between particles of different flavors. Both bosons have spin 1. However, unlike the photons and the gluons, they are massive. The W boson has a mass of $\sim 80.4\,\mathrm{GeV}/c^2$ and the Z boson has a mass of $\sim 91.2\,\mathrm{GeV}/c^2$. Finally, the Higgs field, the quanta of which is a massive, scalar (spin 0) and electrically neutral Higgs boson, is responsible for giving masses to W, Z bosons and fermions. Table 2.1 shows the relative strength of fundamental forces and their range.

A pictorial summary of all particles in the SM, divided into different classes is shown in Figure 2.1.

2.2.2 Theory of interactions in SM

The SM follows the Lagrangian formalism to describe interaction between the particles. Given the SM is a gauge theory, symmetries of the Lagrangian are central to its understanding [18]. In a gauge theory, the Lagrangian is invariant under certain (groups of) transformations and each such symmetry is associated with a conservation law (Noether's theorem). The underlying symmetry group that the SM Lagrangian is invariant under is SU(3)×SU(2)×U(1), where the group SU(3) corresponds to the strong interaction while the group SU(2)×U(1) corresponds to the electromagnetic and weak (electroweak) interaction. Each group generator is associated with an

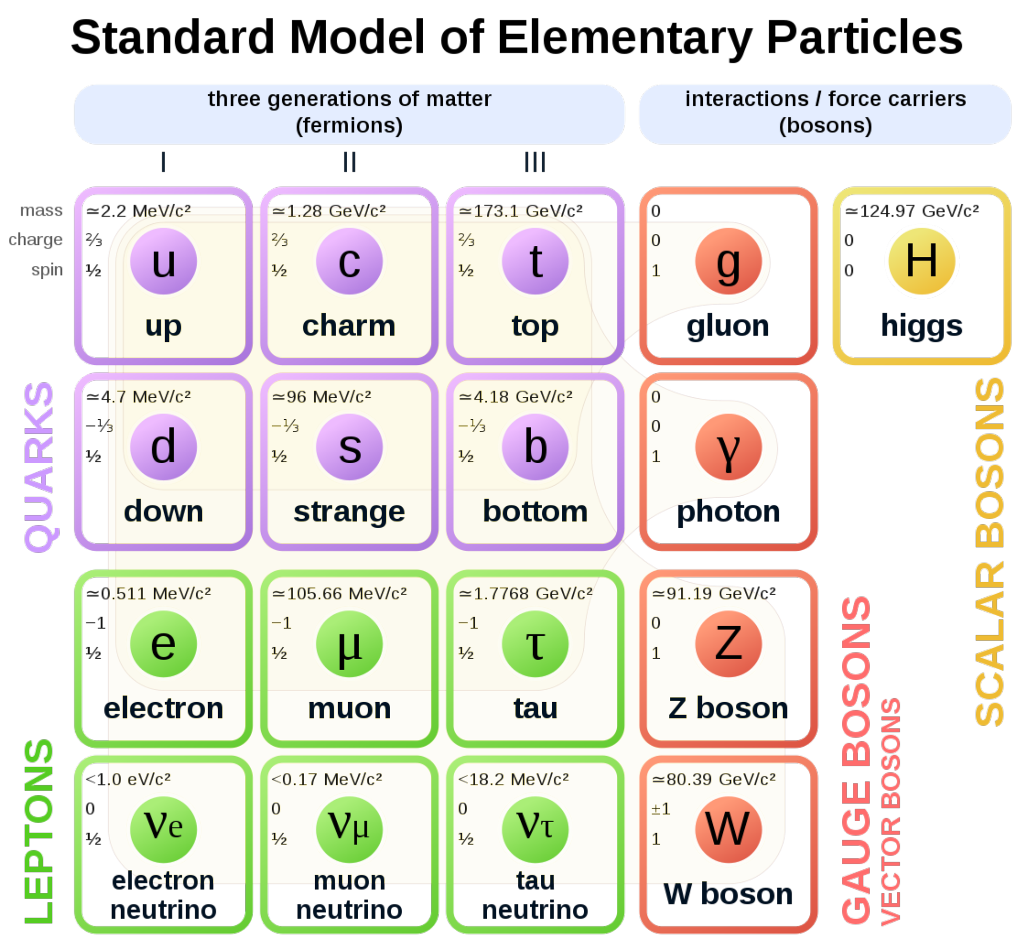


Figure 2.1. A pictorial summary of particles in the SM. The Higgs boson is shown in yellow. Gauge bosons are shown in red. Leptons and quarks are shown in green and violet respectively [17].

TABLE 2.1

RELATIVE STRENGTHS AND RANGES OF ALL FOUR FUNDAMENTAL FORCES, WITH THE STRONG FORCE AS THE BASELINE

Interaction	Relative Strength	Range
Strong	10^{39}	$10^{-15}\,m$
Electromagnetic	10^{36}	∞
Weak	10^{24}	$10^{-18}\,m$
Gravity	1	∞

underlying vector field, the quanta of which are the gauge bosons (gluons, photons, W and Z) described above. We describe the SM interactions briefly below in order of strength.

2.2.2.1 Strong and electroweak interactions

The theory that describes the strong interaction is called Quantum Chromodynamics (QCD). It is a non-abelian gauge field theory based on SU(3) symmetry. Color charge is the quantity conserved under this symmetry. There are three colors: green (g), red (r) and blue (b). Each color has a corresponding anticolor (negative color). As noted earlier, all quarks and gluons have non-zero color charge. Quarks carry a single color, while each of the eight gluons have a color and anticolor charge. The theory being non-abelian, the generator matrices (Gell-Mann matrices) do not commute. The consequence of this is that gluons (unlike photons) can interact with each other.

The theory that was originally formulated to describe the electromagnetic interaction is called Quantum Electrodynamics (QED). It is a gauge field theory based on U(1) symmetry. Electric charge is the quantity conserved under this symmetry and all particles that interact via electromagnetic interaction need to carry electric charge. Unlike the gluons, the photon (because it is electrically neutral) cannot interact with itself. The weak interaction was initially formulated based on the SU(2) symmetry group, with conserved quantity being the weak isospin. The associated gauge bosons are massive and can be electrically neutral (Z) or charged (W). Quarks of (same) different generations can interact with each other via (Z) W bosons. In the 1960s Glashow, Salam and Weinberg combined the theories describing electromagnetic and weak interactions, after realizing that they were different aspects of the same over-arching interaction. This is regarded as electroweak unification, and the electroweak interaction is described by a gauge field theory based on combined SU(2)×U(1) symmetry group. The conserved quantities, weak isospin (T) and electric charge (Q) are related via:

$$Q = T_3 + \frac{Y_W}{2} \tag{2.1}$$

where T_3 is the third component of T and Y_W is a quantum number called the weak hypercharge.

The gauge bosons in this theory are divided into a triplet with two electrically charged and one neutral component (corresponding to Ws and Z), and a singlet with no electric charge (corresponding to the photon). However, in order to maintain gauge invariance of the theory, no mass terms are allowed in the Lagrangian. This would require ALL the gauge bosons (and fermions) to be massless. This is known not to be the case. This broken symmetry (photons being massless and W/Z bosons being massive) is explained by the Higgs mechanism [6–9], described in the next section.

2.2.2.2 The Higgs mechanism

In order to explain how massive gauge bosons come about, the idea of electroweak spontaneous symmetry breaking (EWSB) is introduced. The phenomenon by which EWSB is utilized to give mass to particles is called the Higgs mechanism. Under this mechanism, a new scalar field, ϕ, called the Higgs field and an associated potential, V(ϕ), is introduced. This is represented as doublet and has four degrees of freedom. Three of these four degrees of freedom correspond to the polarizations of the massive W and Z bosons. In order for the Higgs field to interact with W and Z but the not the photon, symmetry has to be broken. The minimum of the potential, i.e. the vacuum state or ground state must be non-zero for this to happen. The parameters of V(ϕ) is so chosen such that it has a Mexican-hat (sombrero) shape, which has infinite degenerate non-zero minima. This non-zero minimum is called the vacuum expectation value (vev), which is measured to be 246 GeV. The direction of symmetry breaking is such that it gives mass to the Z boson but leaves the photon massless. This breaking of symmetry is called spontaneous because there is no particular reason (that we know of) for this direction to have been picked. Nature just happened to spontaneously pick this direction. The Higgs field gives rise to a new massive scalar particle. This particle is the Higgs boson, and corresponds to the fourth remaining degree of freedom of the scalar doublet mentioned above. The fermions acquire mass via Yukawa interaction with the h. The strength of the Yukawa coupling of the h with fermions is proportional to the fermion masses. To summarize, the Higgs mechanism allows the introduction of a mass term for the gauge bosons without breaking the underlying gauge symmetry of the SM Lagrangian. Addition of this field, gives rise to another massive particle, the interactions with which give masses to gauge bosons and fermions. This massive particle is a scalar boson called the Higgs boson, which was discovered in 2012 at CERN by CMS and ATLAS [3–5], almost 50 years after it was first predicted to exist. The most recent measurement of the Higgs mass by

CMS, combining data from both run I and run II, is: $125.35 \pm 0.15\,\mathrm{GeV}$ [19]. Before the LHC, experiments at LEP and Tevatron looked for existence of the h. It was the last missing piece of the SM, and its discovery can be thought to have concluded an era in particle physics and led us into a newer equally exciting era.

2.2.3 Higgs boson production and decays at the LHC

There are several different ways the Higgs boson can be produced at the LHC [20]. The LHC collides protons at high energy, and the production modes of the Higgs boson, in order of cross-section, at the LHC are :

- **Gluon-Gluon Fusion (ggH)**: Since gluons are massless, they do not directly couple to the h. This production mode proceeds via quark loop. The ggH production cross-section at 13 TeV is $\sim 48.37\,pb$ at N3LO [21].
- **Vector Boson Fusion (VBF)**: This production mode has the second largest cross section at the LHC. This mode is characterized by two high-momentum quarks in the final state which hadronize to form jets. The VBF production cross-section at 13 TeV center-of-mass energy is $\sim 3.77\,pb$ at NNLO.
- **Associated Production**: The third largest h production mode at the LHC involves the production of a virtual W^*/Z^* boson that splits into a real boson W/Z boson and a h. The WH production cross-section is $\sim 1.36\,pb$ and the ZH production production cross-section is $\sim 0.87\,pb$, at NNLO level for a center-of-mass energy of 13 TeV at the LHC.
- **ttH Production**: In this production mode, the h is produced along with a pair of top quarks. The production cross-section at 13 TeV center-of-mass energy is $\sim 0.50\,pb$ at NLO.

The Feynman diagrams for h production modes described above are shown in Figure 2.2. The cross-section of each process as a function of center-of-mass energy is shown in Figure 2.3. Figure 2.4 shows the branching fractions of various SM decays of Higgs boson as a function of its mass, illustrating how branching fractions depend on the Higgs mass. It is interesting to note that the 2012 discovery was made combining the channels where the Higgs decays into ZZ*, or into $\gamma\gamma$ (di-photon). Even though the Higgs doesn't directly interact with the massless photon, there are loop order contributions. Although, the cross-section of this channel is small compared to the others, its clean final state signature made it one of the primary channels for the discovery.

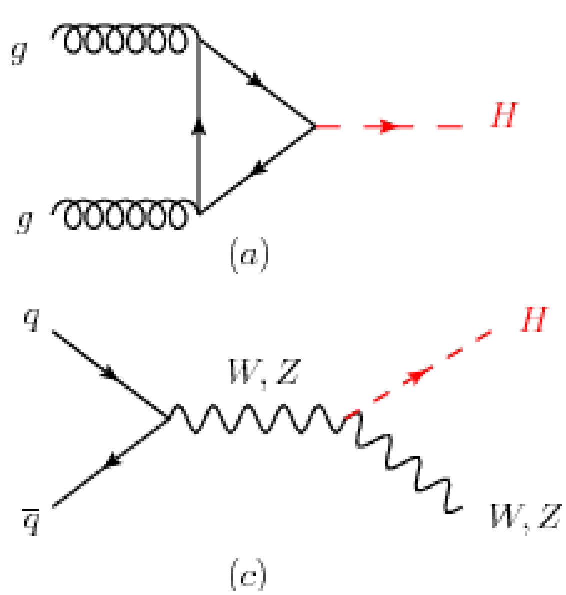


Figure 2.2. Feynman diagrams for Higgs production modes at LHC: (a) gluon-gluon fusion, (b) vector boson fusion, (c) associated production and (d) ttH [22], are shown.

2.3 Inadequacies of the SM

Despite being a faithful description of nature, the SM is not perfect. There are several motivations that suggest the existence of physics beyond the SM. We outline some of these here.

To start, the SM falls short of being an ideal theory of everything, because it

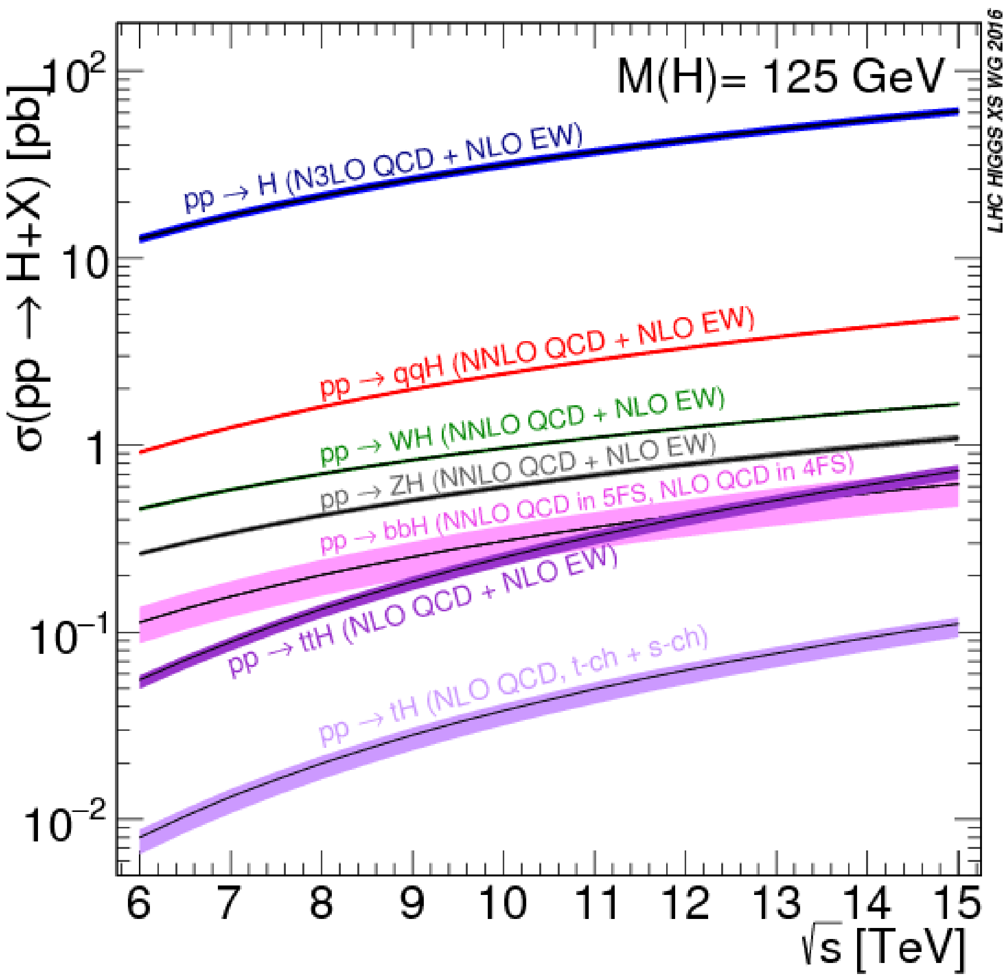


Figure 2.3. The SM Higgs boson production cross-section as a function of the center-of-mass energy in proton-proton collisions at the LHC [22].

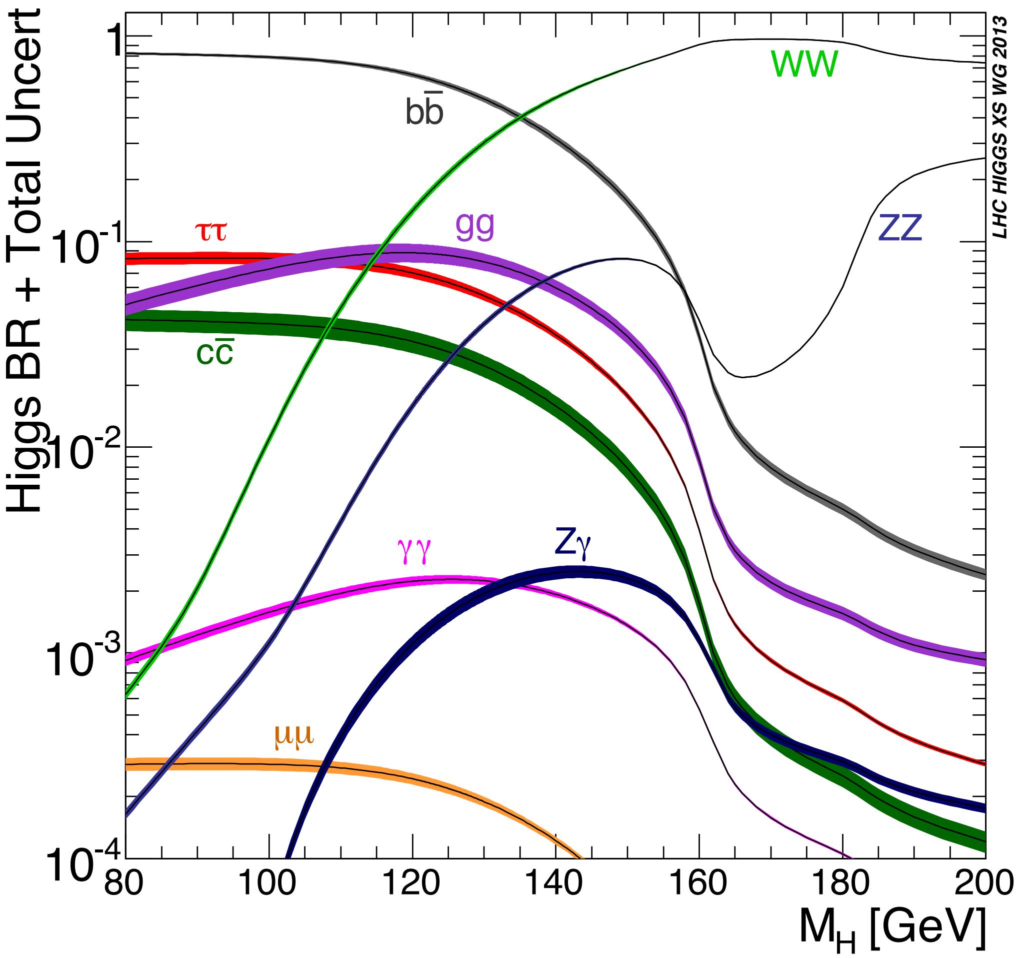


Figure 2.4. Branching fractions to SM decays of the Higgs boson as function of mass [23].

doesn't include gravitation. Including gravitational interaction into the SM has proven to be a difficult challenge. It hasn't yet been possible to incorporate the most successful theory of gravity, General Relativity, and the SM into a single framework. Secondly, neutrinos which are electrically neutral leptons (see section 2.2.1.1) are strictly massless within the SM. However, it has been well-established by experiments that neutrinos oscillate (change flavor). This is only possible if neutrinos have mass. The small but finite masses that neutrinos are now known to have doesn't fit with the SM formulation. Thirdly, cosmological observations point to the existence of a type of matter and energy, the origin of which cannot be explained within the SM. They are referred to as dark matter and dark energy. About 26% of the universe is known to be made of dark matter and 69% is known to be made of dark energy. Thus, particles of the SM form only 5% of the observable universe. Finally, it is believed that matter and anti-matter were produced in almost equal amounts at the Big Bang. However, the universe is made almost entirely of matter. There is no mechanism in the SM that explains how we ended up with a matter dominated universe. Besides the unexplained phenomena outlined above, our understanding of some theoretical features of the SM is inadequate. The SM contains no less than 19 numerical free parameters. The values of these parameters are known but we do not have an understanding of their origins.

To address such shortcomings, many theories have been proposed that modify the SM in such a way that they are consistent with existing observations, but at the same time try to address its imperfections. These theories, called BSM (beyond the standard model) theories predict many outcomes that are otherwise not allowed by the SM. The recently discovered h unlocks a portal to look for these outcomes. As mentioned earlier, the constraint on the branching fraction to non-SM decay modes of the h, derived from a combined study by CMS and ATLAS is B(non-SM) $< 34\%$ at 95% confidence level (CL) [10]. These limits suggest a significant contribution from exotic (non-SM) decays in the BSM Higgs sector. One such interesting process that is forbidden in the SM but occurs in many new physics scenarios is interactions between charged leptons that violate the conservation of Lepton Flavor. In particular, Lepton Flavor Violating (LFV) decays of the h are allowed by these theories, and could be realized in decays of the h, which is neutral, into two charged leptons of different flavor. Looking for LFV decays of charged leptons is also interesting in the light of neutrino oscillations mentioned earlier, which also violate lepton flavor, a phenomenon that remains unexplained by the SM [24].

2.4 BSM models with lepton flavor violation

Like all fermions, charged leptons acquire mass from their interaction with the Higgs. Higgs interacts with these leptons via Yukawa couplings. The Yukawa interaction matrix is diagonal in SM:

$$Y = \begin{pmatrix} Y_{ee} & 0 & 0 \\ 0 & Y_{\mu\mu} & 0 \\ 0 & 0 & Y_{\tau\tau} \end{pmatrix}$$

However, in BSM models, the above doesn't hold true [13] and off-diagonal Yukawa couplings are possible. In a model containing only the SM Higgs as the source of EWSB, an effective field theory approach can be used to introduce off-diagonal couplings [25]. If only SM particles (quarks, leptons, gauge and Higgs bosons) are considered to exist up to a certain energy scale, Λ, additional heavy fields can be integrated out, leading to an effective field theory. Higher dimensional operators of dimension 6 then suffice to introduce LFV couplings. Interestingly, dimension 5 operators introduce neutrino oscillations into the SM, but not LFV in interactions of charged leptons. Dimension 6 operators decouple the values of fermion couplings to the h from the fermion masses. The Yukawa couplings can be then written as:

$$Y_{ij} = \frac{m_{ij}}{v}\delta_{ij} + \frac{v^2}{\sqrt{2}\Lambda^2}\hat{\lambda}_{ij} \tag{2.2}$$

where $\hat{\lambda}_{ij}$s are coefficients associated with dimension 6 operators. In the limit $\Lambda \to \infty$, we recover the SM and off-diagonal couplings are zero. Thus, LFV couplings can be introduced as long as the mass scale is finite. In BSM models with several sources of EWSB, LFV couplings can be introduced without this restriction. Two Higgs doublet models (2HDM) models constitute general models of this class, and allow the violation of lepton flavor [26]. Supersymmetry models [27–32], such as the Minimal Supersymmetric Standard model (MSSM) and the Next-to-Minimal Supersymmetric Standard model (NMSSM), also postulate multiple Higgs bosons, and give rise to LFV interactions. Supersymmetric models introduce a alternate (supersymmetric) boson partner to every SM fermion and an alternate fermion partner for every SM boson. These alternate particles, if discovered, could be suitable candidates for explaining dark matter and dark energy. Other models that allow LFV interactions include [1] composite Higgs models [33, 34] which consider the SM h to be a bound state of other BSM particles, partially composite Higgs models such as Randall-Sundrum models [35–37], and several others [38–47].

2.5 Pre-LHC constraints on LFV couplings

Indirect low-energy measurements from pre-LHC era can be used to constrain the h $\to \mu\tau$ decay. These constraints were derived and summarized in [13]. For example, constraints on $\tau \to \mu\gamma$ transition which proceed via a virtual Higgs boson can be used to constrain h $\to \mu\tau$ decay. Feynman diagrams contributing to this process at one loop level are shown in Figure 2.5. Further constraints come from $\tau \to 3\mu$ decays, and also from anomalous magnetic dipole moments, and are shown in the same figure. The constraints on Yukawa couplings derived from the above measurements can be converted to constraints on $Br(\mathrm{h} \to \mu\tau)$, following the procedure described in Section 8. These constraints set the upper limit on $Br(\mathrm{h} \to \mu\tau) \lesssim 10\%$, thus leaving a lot of room to search for this decay. Similar constraints exist for $\mathrm{h} \to \mathrm{e}\tau$ LFV decay, and are set at $Br(\mathrm{h} \to \mathrm{e}\tau) \lesssim 10\%$. Indeed, CMS searches looking for $\mathrm{h} \to \mathrm{e}\tau$ decay have been performed along with searches for $\mathrm{h} \to \mu\tau$. Finally, it is interesting to note the LFV decay, $\mathrm{h} \to \mu\mathrm{e}$, is very strongly constrained by $\mu \to \mathrm{e}\gamma$ decays, giving an very stringent upper limit $Br(\mathrm{h} \to \mu\mathrm{e}) \lesssim 2 \times 10^{-8}$. Due to such strong existing constraints, this search is not been performed by CMS, but rather $\mathrm{h} \to \mu\tau$ (this thesis) and $\mathrm{h} \to \mathrm{e}\tau$ searches are performed.

2.6 Constraints from previous LHC searches

The first direct search for LFV Higgs decays was published by CMS collaboration in 2015 [14]. This search improved the limits listed above by an order of magnitude to $\mathcal{B}(\mathrm{h} \to \mu\tau) < 1.51\%$ (0.75%) for observed (expected) limits at 95% CL. This was followed by another search (2016) which set observed (expected) upper limits on the branching fractions $\mathcal{B}(\mathrm{h} \to \mathrm{e}\tau) < 0.69\%$ (0.75%) at 95% CL [48]. Both searches were performed with $19.7\,\mathrm{fb}^{-1}$ of pp collision data collected at 8 TeV center-of-mass energy by CMS during Run I of LHC . The limits from these searches are summarized graphically in Figure 2.6. In 2015 and 2017, the ATLAS Collaboration also published results from similar searches performed with data collected by the atlas detector [49, 50]. The observed (expected) limits were set at $\mathcal{B}(\mathrm{h} \to \mu\tau) < 1.43\%$ (1.01%) and $\mathcal{B}(\mathrm{h} \to \mathrm{e}\tau) < 1.04\%$ (1.21%) at 95% CL.

The 2015 CMS search for $\mathrm{h} \to \mu\tau$ saw an excess of events with a significance of $2.4\,\sigma$. Although this excess is not quite enough to claim evidence for this decay, this gives us a strong motivation to perform this search with a larger amount of data which would either lead us to confirm this excess, or squash it and set much stricter limits

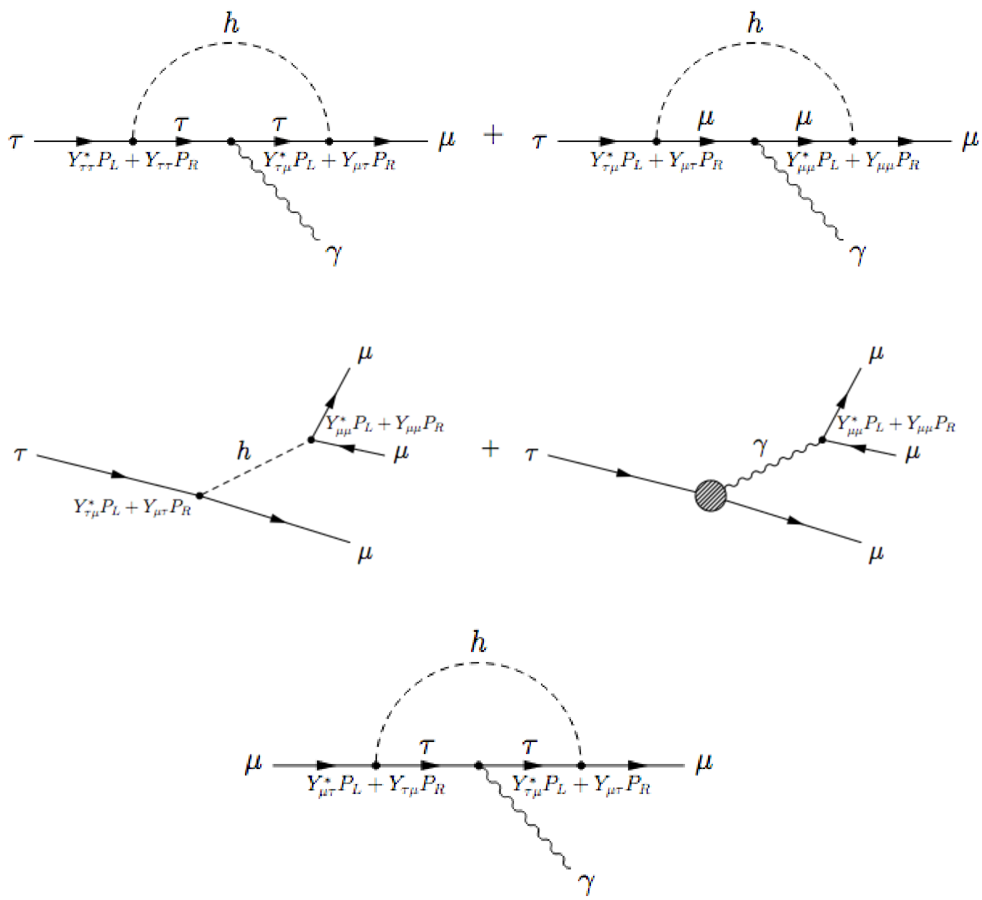
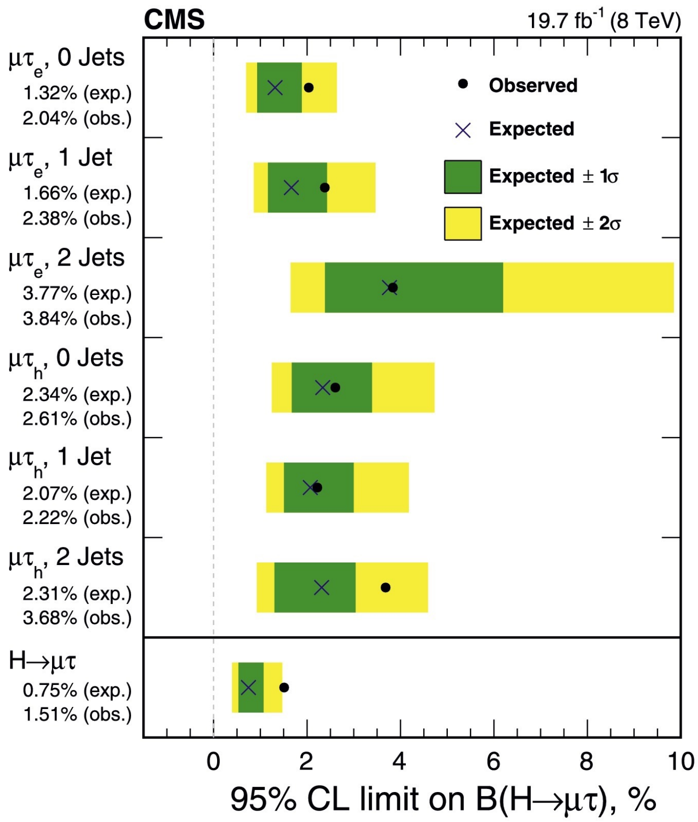


Figure 2.5. Diagrams contributing to flavor violating process $\tau \to \mu\gamma$ (top), $\tau \to 3\mu$ (middle) and anomalous magnetic moment of the muon (bottom) [13].



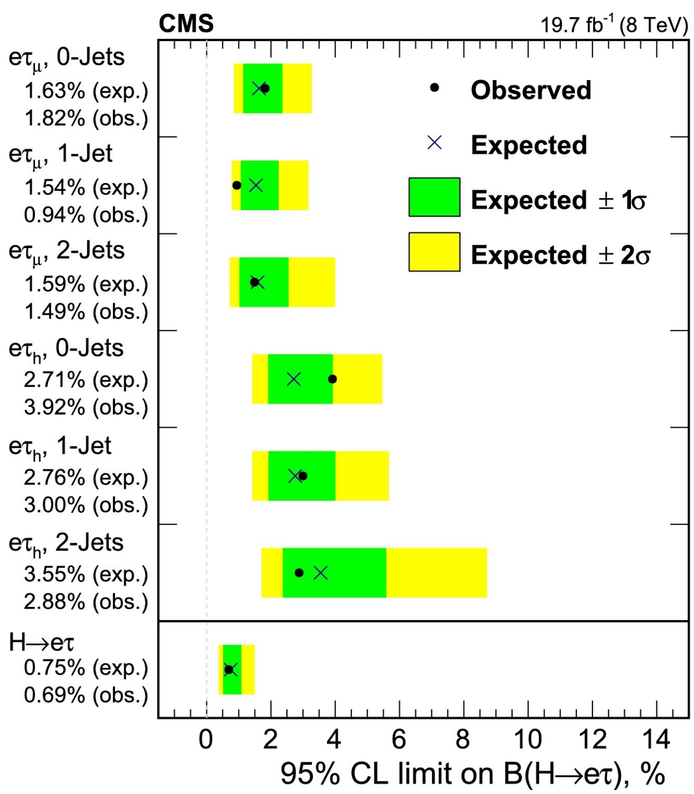


Figure 2.6: Limits from Run I searches performed by CMS for h $\to \mu\tau$ (top) and h $\to e\tau$ (bottom) [14, 48].

on this process. The dataset collected by the CMS detector in 2016 provides us with such an opportunity. It corresponds to proton-proton collision data at a much higher center-of-mass energy of 13 TeV. The number of h bosons produced depends on the cross-section. Since the cross section scales up at higher center-of-mass energies (see Figure 2.3), a much larger number of h bosons would be produced. Also, the 2016 dataset has a size of 36 fb^{-1} which is almost two times in size of the run I dataset. This thesis describes this search specifically in the channel where the τ decays into a electron, i.e. the h $\rightarrow \mu\tau_e$ channel.

2.7 Motivations for H $\rightarrow \mu\tau_e$ search

As mentioned in section 2.4, many of the BSM models, that allow LFV decays of the h, predict the existence additional heavy Higgs bosons. For example, 2HDM predicts the existence of two heavy neutral Higgs bosons, H(CP-even) and A(CP-odd). According to a theoretical study published in 2016 [15], these heavy bosons (henceforth referred to as H) are expected to decay in a Lepton Flavor Violating manner just like their SM counterpart, h . A direct search for H $\rightarrow \mu\tau$ would thus provide a complementary probe of these BSM models that postulate the existence of such heavy neutral H bosons. In fact, the 2015 CMS search for h $\rightarrow \mu\tau$ was reinterpreted as a search for H $\rightarrow \mu\tau$ decay [51], and limits on $\sigma(\mathrm{gg} \rightarrow \mathrm{H}) \times \mathcal{B}(\mathrm{H} \rightarrow \mu\tau)$ were set for H bosons in the mass range of 150 GeV to 300 GeV. We describe here the first direct search to look for H $\rightarrow \mu\tau$ decay, in the channel where the τ decays into an electron, i.e. H $\rightarrow \mu\tau_e$ channel. Only the primary H production mode (gluon fusion) is considered for this search. This search uses the same dataset as the h $\rightarrow \mu\tau$ search, i.e. 36 fb^{-1} of pp collision data at 13 TeV center-of-mass energy collected in 2016, and probes H masses in the range range $200 < m_H < 900$ GeV.

CHAPTER 3

EXPERIMENTAL SETUP

> If it disagrees with experiment, its wrong. In that simple statement is the key to science. It doesnt make any difference how beautiful your guess is, it doesnt matter how smart you are who made the guess, or what his name is . . . If it disagrees with experiment, its wrong. Thats all there is to it.
>
> *Richard Phillips Feynman*

3.1 Introduction

In this chapter, we describe the experimental setup used to collect the data used for the searches. We first describe the collider which smashes protons together. Then we describe the detector used to collect the data from those proton-proton collisions.

3.2 The Large Hadron Collider

The Large Hadron Collider (LHC) [52] 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 foreseeable 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 Synchrotron Booster), PS (Proton Synchrotron) and SPS (Super Proton Synchrotron). This is illustrated in Figure 3.1. 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 pressurized 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 with protons, and about 20 minutes for the proton beam to reach its current peak energy of

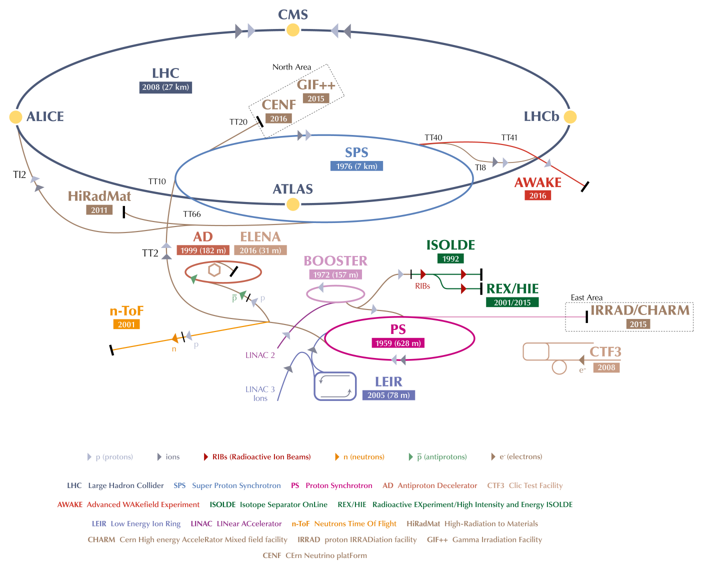


Figure 3.1: Cern Accelerator Complex [53]

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 2021. 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 also expressed in terms of only beam parameters as:

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

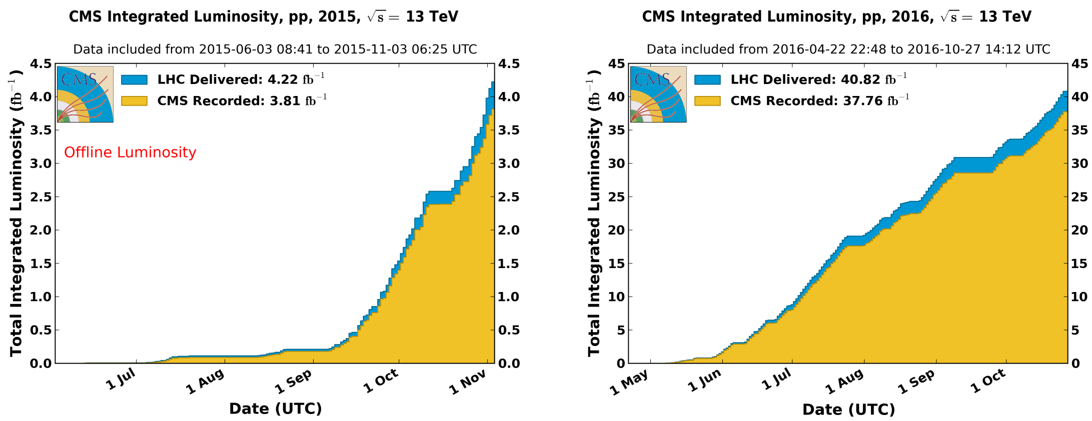


Figure 3.2: Evolution of integrated luminosity in 2015 and 2016 delivered by LHC (blue), and collected by CMS detector (orange) [54].

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 integrated 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 delivered data amounting to a more than 36 fb^{-1}, only in 2016. Figure 3.2 shows the amount of data delivered by the LHC overlaid with the subset collected by the CMS detector in 2015 and 2016. The data analyzed in the searches described in this thesis were collected by the CMS detector (section 3.3) in 2016 during proton-proton collisions delivered by the LHC, and correspond to an integrated luminosity of 35.9 fb^{-1}.

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. In this timeframe, it is anticipated to operate at increasingly higher luminosities over extended periods of time, helping collect unprecedented amounts of data. Figure 3.3 shows an overview of the long term LHC schedule.

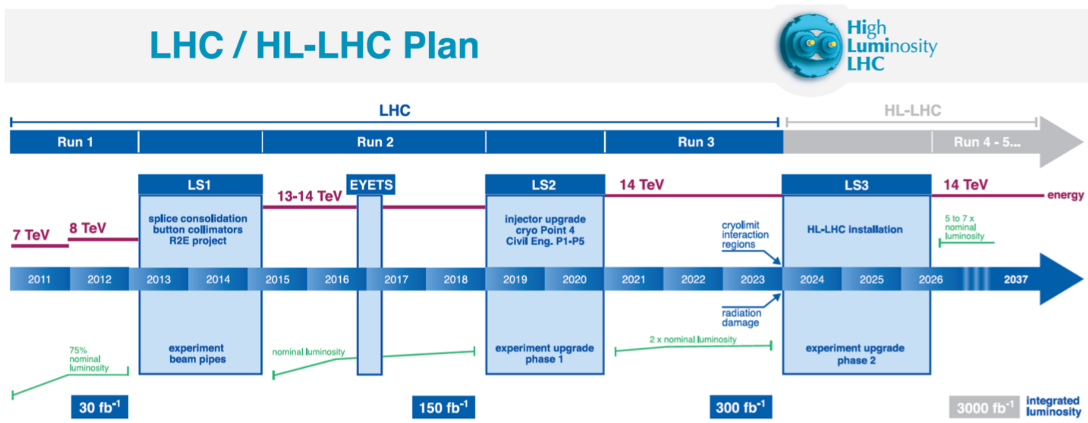


Figure 3.3: Overview of the long term LHC schecdule [55].

3.3 The CMS detector

The Compact Muon Solenoid [56] is a general multipurpose particle physics detector that is placed in one of the four collision points of the LHC. It is 28.7 m long with a diameter of 15.0 m, weighs 14000 tonnes, and is composed of several subdetectors. Its aim is to study a broad array of physics, from making precise measurements of known processes 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 3.8 T. 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,

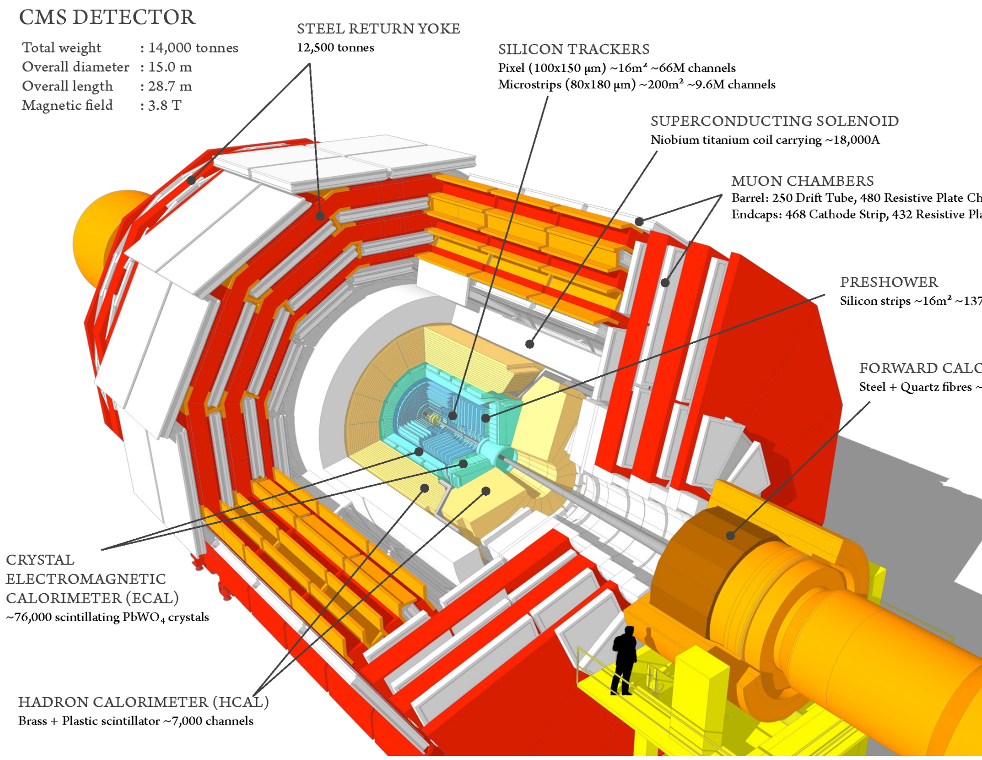


Figure 3.4: Layered View of the CMS detector

and the hadronic calorimeter (HCAL) which measures the energy of heavier particles (mesons, baryon, jets of such 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 3.4 shows a layered view of the detector. Figure 3.5 shows a transverse cross-sectional view along with specific particle interactions. The following sections describe the detector in further detail.

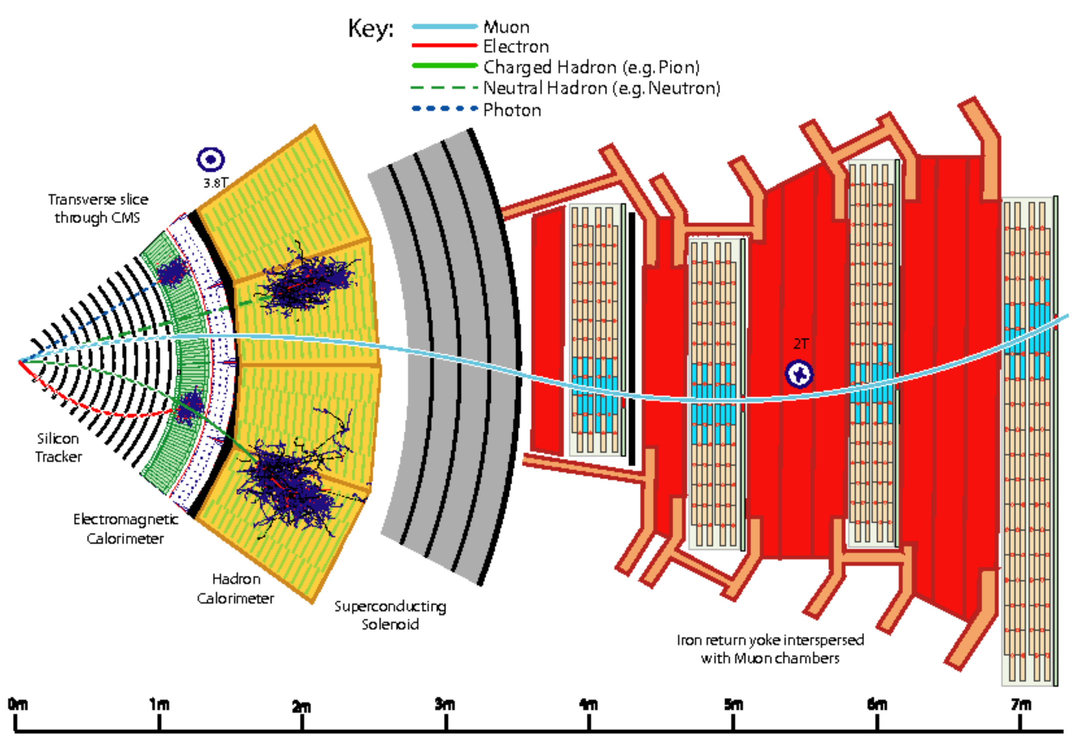


Figure 3.5: Transverse cross-sectional view of the CMS along with specific particle interactions [57].

3.3.1 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 anti-clockwise 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.

3.3.2 Charged particle tracking system

The Charged Particle Tracking system or the tracker measures, efficiently and precisely, the trajectories of charged particles passing through it [56, 58, 59]. It is cylindrical in shape with a length of 5.8 m and diameter of 2.5 m. The momentum of charged particles can be measured using these tracks with high precision. The tracker also helps in reconstruction of secondary vertices of long-lived particles. Given the tracker is innermost the layer of the CMS and the challenging LHC conditions, the tracker experiences high particle flux. Therefore, the material it is built of needs to be radiation hard. This makes silicon a good choice for detector material. The tracker also needs to be highly granular and have fast response in order to be able to efficiently and precisely measure the trajectories. However, it is also necessary to keep the amount of material of the tracker low to limit photon conversion, multiple scattering, bremsstrahlung and nuclear interactions. A compromise between these two aspects led to the current design of the tracker which consists of two subsystems. The inner layer consists of a detector made of 66 million silicon pixels. The pixels are $100\,\mu\text{m}\times 150\,\mu\text{m}$ and arranged in the form of three barrel layers with radii of 4.3 cm, 7.2 cm and 11 cm respectively, and two endcap disks on each side of the barrel at 34.5 cm and 46.5 cm from the interaction point. Figure 3.6 shows illustrations of the CMS pixel system and that of a section showing a pixel element. Outside the region of the pixel detector, the density of tracks is low and the high granularity requirement can be somewhat relaxed. The outer subsystem of the tracker is thus made of strips of silicon. The strip detector consists of 10 barrel layers which are organized into 4 inner layers, called Tracker Inner Barrel (TIB) and 6 outer layers, called Tracker Outer Barrel (TOB). They range in radii from 25 cm to 110 cm. It also consists of 12 endcap disks on each side of the barrel which have radii up to 110 cm, and are up to 280 cm from the interaction point. They are organized on each side into 3 inner endcap disks, called Tracker Inner Disks (TID) and 9 outer disks, called Tracker End Cap (TEC). Figure 3.7 shows the layout of the tracker subsystems.

The combination of the pixel and strip trackers give the CMS excellent track resolution and efficiency. The algorithm used to reconstruct these tracks from hits in the pixels and strip detectors is described in detail in Section 4.4.2. Promptly produced, isolated muons of $p_T > 0.9\,\text{GeV}$ are reconstructed with close to 100% efficiency for $|\eta| < 2.4$ [61]. The transverse momentum resolution for 100 GeV muons in the central region of $|\eta| < 1.4$ is about 2.8%, while the impact parameter resolution is about 10μm and 30μm in the transverse and longitudinal directions.

3.3.3 Electromagnetic calorimeter

The Electromagnetic Calorimeter or ECAL is the next layer of the CMS detector. It fits snugly in between the tracker and the HCAL (next section) and measures the energy of electrons and photons. Besides being a important for many other interest-

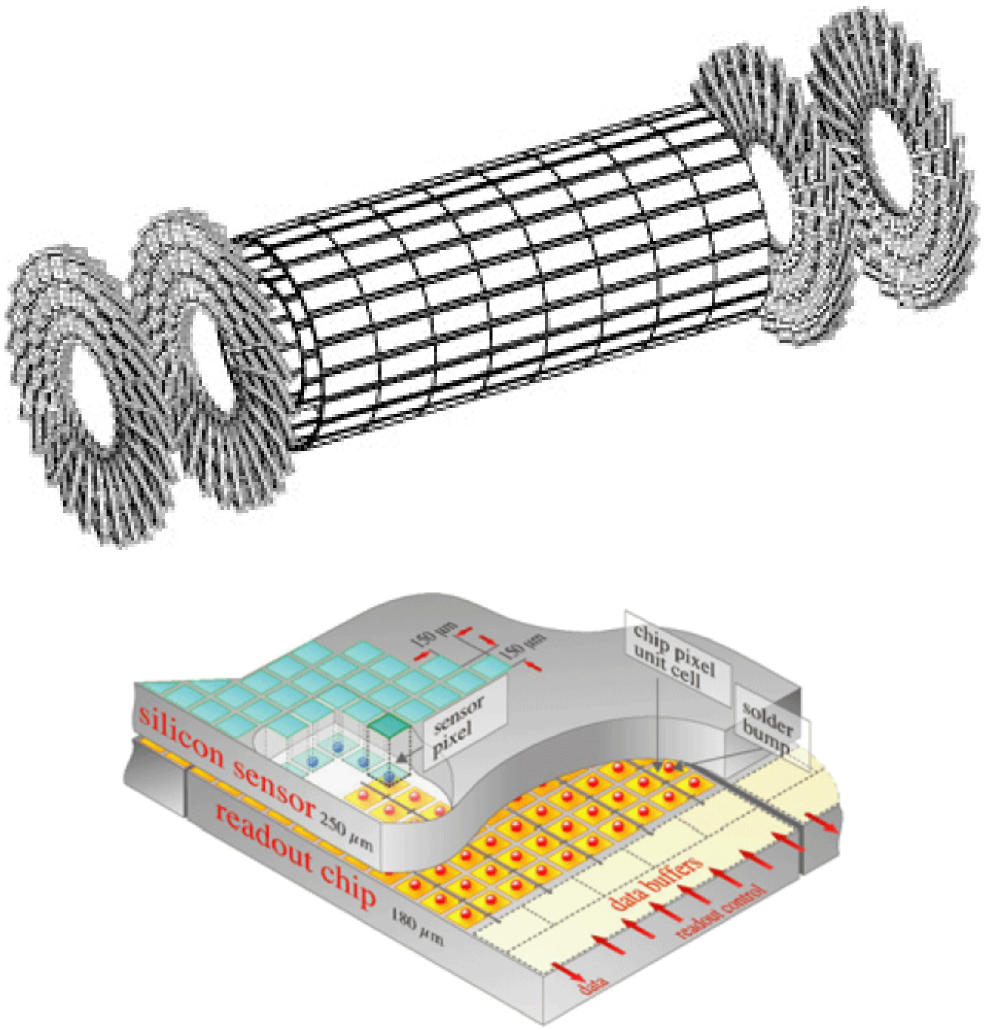


Figure 3.6: Perspective view of the CMS pixel detector (top) [58], and illustration of a section showing individual pixel elements (bottom) [60].

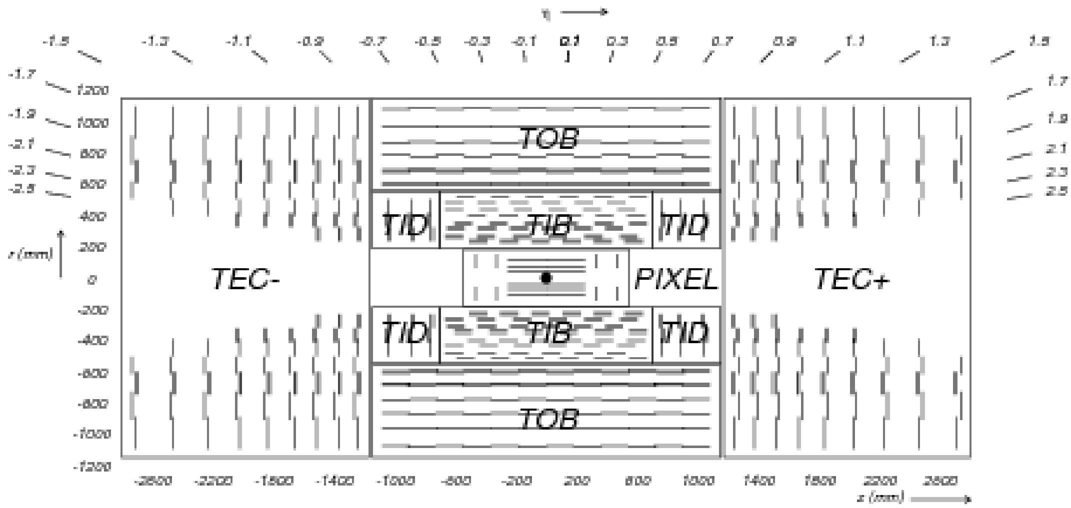


Figure 3.7: The CMS tracker layout [56].

ing analysis, one of the two final state particle in the searches described here is an electron. The ECAL is a homogenous calorimeter; the absorber and scintillator materials are one and the same. In other words, the entire volume of the calorimeter can potentially contribute towards a signal. This does away with the error introduced by sampling (which is needed in case of calorimeters which are not homogenous). The material chosen for the ECAL is Lead Tungstate ($PbWO_4$), which despite being heavier (density=$8.3\,g/cm^3$) than stainless steel (density$<8\,g/cm^3$) is highly transparent in its crystalline form. The $PbWO_4$ crystals (Figure 3.8) used in the ECAL are 23 (22) cm in length and $2.2 \times 2.2\,(2.86 \times 2.86)$ cm in transverse size for the barrel (endcap) [62]. These high density crystals have a short radiation length ($X_\circ = 0.85\,cm$) and a small Moliére radius ($R_M = 2.19\,cm$). The radiation length of a material is the mean distance over which an incident electron loses all but 1/e of its energy, or 7/9 of the mean free path for pair production by a high energy photon. Electromagnetic showers are initiated in the ECAL when a high energy electron or photon enters its volume. An electron emits photons via bremsstrahlung radiation. These photons in

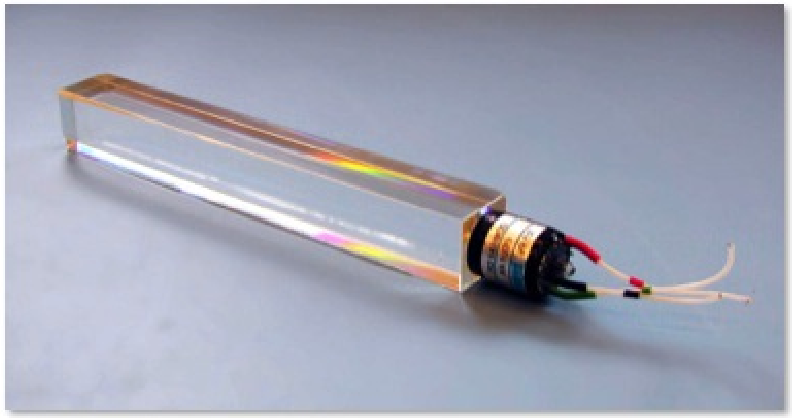


Figure 3.8: A $PbWO_4$ ECAL crystal [63].

turn pair produce electron-positron pairs which can in turn release more photons. This stops when the photon energy becomes lower than what is required for pair production (usually a few MeV). The "depth" of an electromagnetic shower scales logarithmically with energy and linearly with radiation length. Therefore, having a short radiation length is important. Since, the length of $PbWO_4$ crystals (23 cm) used in the ECAL is about 27 times the radiation length (0.85 cm), they are able to contain full showers. Another important quantity, the Moliére radius is the radius of the cylinder that contains 90% of shower in the lateral direction. The small Moliére radius ($2.19\,cm$) of $PbWO_4$ thus helps the ECAL in delivering a high position resolution. Further, these crystals also produce light fast with 80% of light being emitted in 25 ns (LHC bunch crossing time). Because this light yield depends strongly on temperature ($-2\%/^{\circ}C$ at $18^{\circ}C$), it is necessary to maintain the ECAL at a constant temperature ($18 \pm 0.05^{\circ}C$). Also, despite being radiation hard, the crystals still suffer from reduction of transparency due to irradiation. Damage and recovery of the crystals is measured using laser light (Figure 3.9), and this is corrected for periodically in measurements made by the ECAL operations team.

The crystals of the ECAL are organized into a barrel and two endcaps, one on each

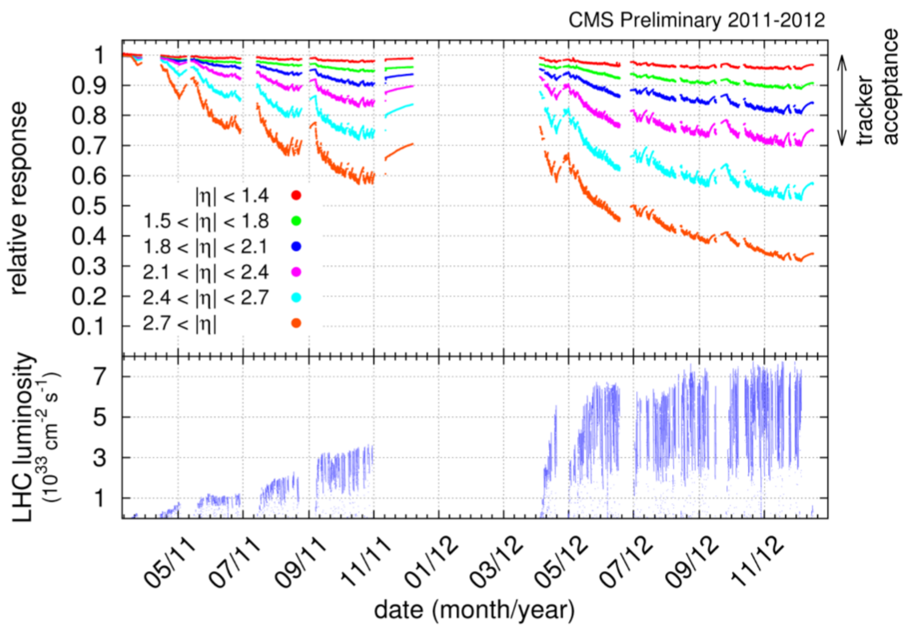


Figure 3.9: Relative response loss of ECAL due to radiation effects, as function of time in different η regions as measured by the laser monitoring system [62].

end. The layout is illustrated in Figure 3.10. The barrel extends up to $|\eta| < 1.479$ and is made of 61200 crystals, which are further organized into 36 supermodules of 1700 crystals each. The endcaps cover the region between $1.65 < |\eta| < 3.0$, and are each made of two Dees with 3,662 crystals in each Dee. The 75,848 crystals together weigh 92 tonnes. The crystals are equipped with photodetectors which convert the light output to amplified electrical signals. In the barrel each crystal is equipped with two $5 \times 5\,\mathrm{mm}^2$ APDs (Avalanche Photo Diodes) which have a gain of 50. In the endcaps, the radiation is too high to use silicon photodiodes. Here, VPTs (Vacuum Phototriodes), with a gain of $\sim$10, are used instead. Each endcap crystal has a VPT glued to the end of it (as seen in Figure 3.8). The energy resolution of the ECAL can be parametrized as the sum of a stochastic term, a noise term and a constant term. The resolution depends on energy, and for the ECAL barrel this is can be approximated by: $(\frac{\sigma}{E})^2 = (\frac{0.028}{\sqrt{E}})^2 + (\frac{0.12}{E}) + 0.003^2$ [64].

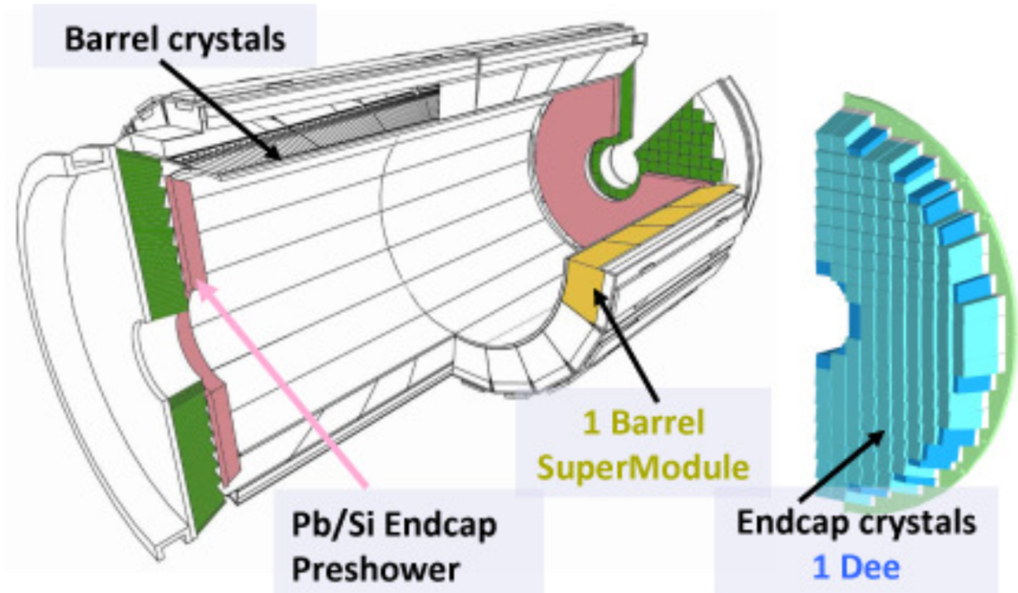


Figure 3.10: CMS ECAL layout [62].

I worked closely with the ECAL during the time spent at CERN. The next subsection briefly describes a critical project related to the operational performance of the ECAL that I worked on.

3.3.3.1 Mitigation of anomalous APD signals in the ECAL barrel

As mentioned in the last section, APDs are used as photodetectors to collect, amplify and convert light to electronic signals in the barrel section of the ECAL. The APDs use silicon as active element and have an active area of $5 \times 5\text{mm}^2$. Each ECAL crystal has a pair of APDs. Figure 3.11 shows an ECAL APD mounted in a module. The APDs suffer from a problem which can be considered both strange and interesting in nature. Particles produced in the collisions can occasionally strike the APDs directly and interact with the material, that cause large deposits through direct ionization of the silicon [65]. These deposits cause anomalous signals to be registered as coming from a real electron or photon that has scintillated the ECAL.

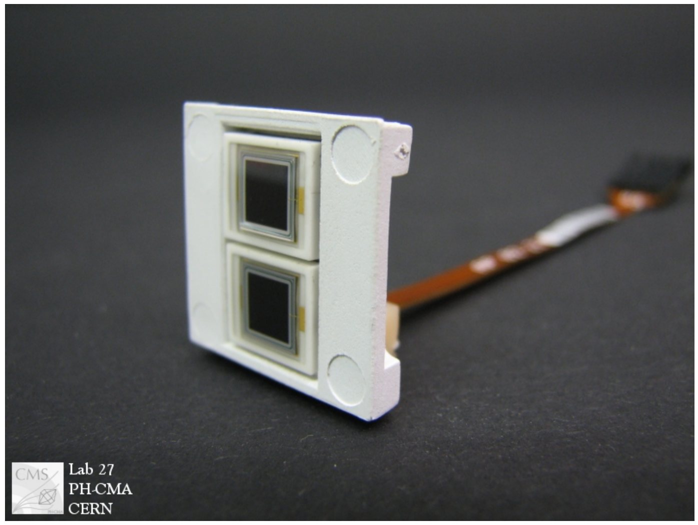


Figure 3.11: Two CMS ECAL Avalanche Photodiodes [66].

These anomalous signals are usually called “spikes”, and they can range from a few GeV to to a few hundred GeV in energy. When the collected data is being processed, these deposits must be eliminated with high accuracy and efficiency for subsequent analysis. This is called offline rejection. An algorithm called the swiss-cross algorithm along with timing information is used to get rid of these deposits with almost cent percent efficiency. However, offline rejection of spikes is not enough. Because spikes are produced at a rate proportional to number of collisions, at a high luminosity output collider like the LHC, spikes cause problems with triggering, as they “eat up” the $\sim 100\,\mathrm{kHz}$ bandwidth available at the first level of triggering (see section 3.3.7 for a description of the two-level CMS trigger system). Thus a majority of spikes need to be eliminated “online”, that is when the data is being collected live. Given that the hardware-based first level of the CMS trigger (L1) has only $3.2\,\mu\mathrm{s}$ to make a decision, spikes need to be identified and rejected fast in the firmware.

Spikes masquerade well as true electromagnetic (EM) energy deposits. However, the defining characteristic of spikes is that they deposit all their energy in a single

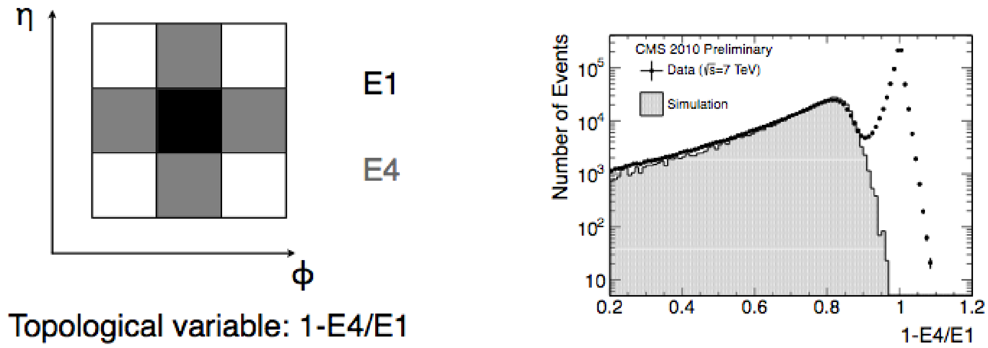


Figure 3.12: Definition of the Swiss-cross variable (left) and its distribution in pp collision data and simulated EM events [65].

ECAL crystal. In contrast, as mentioned in the last section, EM showers spread. The spread is even larger in the direction perpendicular to the magnetic field (ϕ-direction). A well-centered EM shower deposits about $\sim 80\%$ of its energy (E1) in the central crystal. However, the remaining $\sim 20\%$ of its energy is spread in neighboring crystals [65]. Based on the above ideas, a discriminating variable is defined as $1 - \frac{E4}{E1}$, where E4 is the energy deposited in the 4 nearest neighbors of the central crystal. This variable is named the "Swiss-cross" variable. Figure 3.13 (left) shows how the Swiss-cross variable is defined. The distribution of this variable is shown Figure 3.13 (right), for data collected from pp collisions and Monte-Carlo simulated data containing only EM energy deposits. A secondary peak at high Swiss-cross values can be clearly seen in only the pp collision data, and corresponds to spikes. Thus, this variable can be used to discriminate very well against spikes. In fact 99% spikes can be rejected using this variable alone. This can be further improved by using other characteristics such as pulse shape and timing that separate spikes from EM deposits.

Although the Swiss-cross algorithm described above is relatively simple, it still isn't possible to implement it online, and run it on firmware under the 3.2 μs of time that the L1 trigger has to make a decision. Therefore, a much simpler adaptation of the Swiss-cross algorithm was implemented. This algorithm exploits the fact that the

EM showers spread in the ϕ direction due to the magnetic field. The smallest building block of an EM candidate, i.e. an electron or a photon, is a trigger tower (TT). A TT is a block of 5×5 crystals in the $\eta \times \phi$ direction. These TTs are sent downstream by L1 trigger to the next triggering level (HLT, see section 3.3.7). A base block associated with a TT is called a trigger primitive (TP). The aim of this algorithm is to thus reduce the proportion of TPs induced by spikes within a short amount of time available. For each of the 5, 1×5 strips lengthwise across the phi direction in a TT, the number of crystals having energy above a certain threshold is counted. If in any of these strips more than 1 channel is above the threshold (hopefully owing to EM shower spread), a logic flag called the sFGVB (strip Fine Grain Veto Bit) is set to 1. Otherwise, it is set to 0. A TP with sFGVB 0 is considered to have been induced by a spike and is discarded. Not all TPs are subject to the above rule. Only TP with total energy above a certain threshold are eliminated. Otherwise, we would end up discarding many real low energy EM candidates thereby reducing the overall efficiency of detecting such candidates. The sFGVB threshold mentioned earlier is η dependent, and its average value corresponded to 350 MeV in run I while TP-killing threshold was set at 12 GeV. Using this algorithm, the proportion of spike-induced TPs were kept low in run I ($\sim 15\%$).

LHC conditions have become more an more challenging since the start of Run II in 2015. The instantaneous luminosity has risen continuously, and the bunch crossing has been reduced to 25 ns. This means CMS can now collect a lot more data at a faster rate which is great. However, this also means the ECAL APDs get hit with much more spikes compared to RUN I. In data collected in 2015 and early 2016, this effect was observed as expected. This is shown in Figure 3.14. The proportion of spike induced TPs had increased progressively, and the online spike rejection algorithm needed to be tuned to the newer LHC conditions.

There are several things that need to be considered when tuning the above algo-

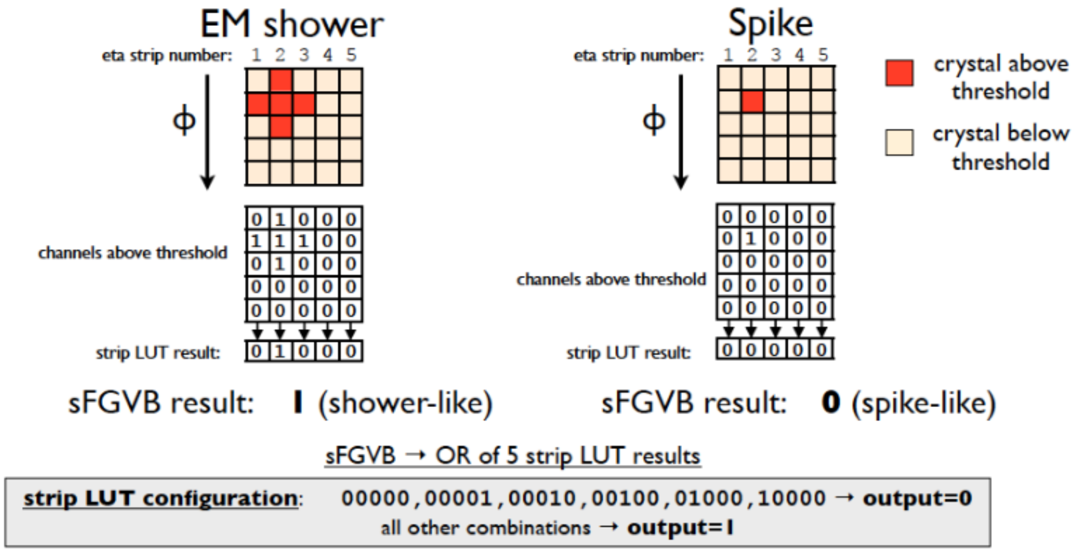


Figure 3.13: Illustration of the Strip Fine Grain Veto Bit [67].

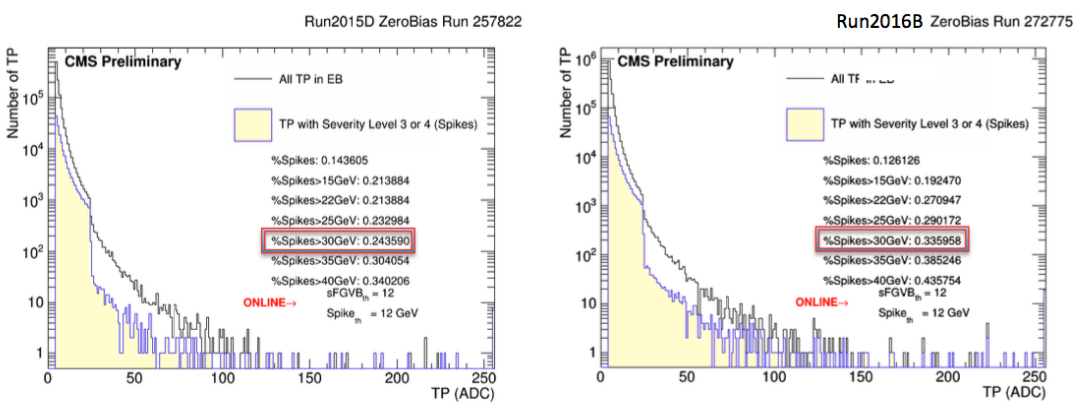


Figure 3.14: Spectrum of energy for all TPs (black line) and TPs induced by spikes (blue line with yellow fill) for pp collision runs from 2015 (left) and 2016 (right). Percentage of spike induced contamination above different energy thresholds are also noted.

rithm. The spike contamination should go down, however this should not come at the cost of reduction in efficiency of detection of real EM candidates. Further, the overarching L1 triggering rate should not increase, and if possible decrease. Therefore, in order to tune the algorithm one needs to adjust both thresholds: sFGVB threshold and the TP-killing threshold. Simply increasing the sFGVB threshold would eliminate more spikes, but it would also end up rejecting some real EM candidates. Thus, in order to keep the efficiency of EM candidate detection effective, we need to increase the TP-killing threshold as well. An array of working points (possible values for the thresholds) were studied. They ranged from 14 to 22 GeV for the TP-killing threshold, and $\sim$ 400 to $\sim$ 900 MeV for the sFGVB threshold. For each of these working points, the residual contamination from spikes was calculated. The efficiency for detection of EM candidates was also calculated. Figure 3.15 shows the purity (1-proportion of spike contaminated TPs), i.e. proportion of real EM candidates, plotted v/s EM candidate efficiency, for various working points. The ideal working point would corresponding to the point closest to the top right corner of the plot. However, we would also need to consider the third factor mentioned above, i.e. the triggering rate. Increasing the TP-killing threshold increases the rate at which all triggers below that threshold fire (or select events). However, each increasing TP-killing threshold is accompanied by a higher sFGVB threshold which brings the rate of firing of all triggers above the TP-killing threshold down. The L1 rate for each of the working points being studied were also calculated. The optimal working point would thus have a higher rate decrease and a lower rate increase thus bringing the overall L1 triggering rate down, and keeping spike contamination low and EM candidate selection efficiency high.

Based on all the aforementioned factors, the new thresholds chosen were 500 MeV for sFGVB and 16 GeV for TP-killing. The tuned spike rejection algorithm was able to reduce the contamination rate down to acceptable levels, and reduce the overall

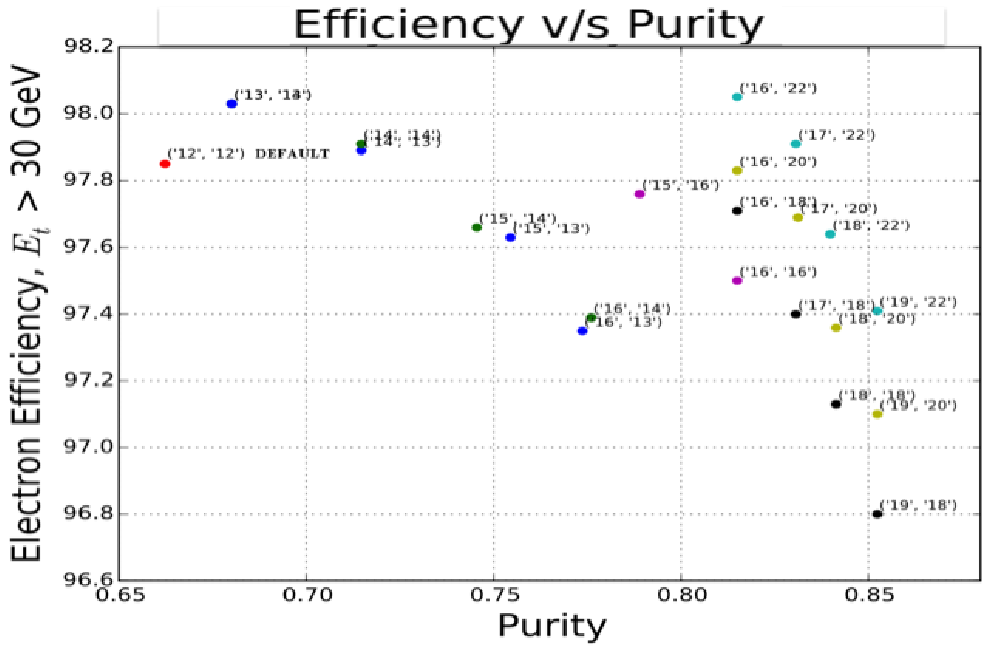


Figure 3.15: Purity (proportion of TPs which are not spike induced) v/s efficiency of selection of EM candidates. The pink point annotated (16,16) was chosen as the new threshold (see text).

L1 trigger rate from EM candidates.

3.3.4 Hadronic calorimeter

The Hadronic Calorimeter (HCAL) is designed to detect and measure the energy of particles that interact via the strong force [56]. It also detects any remnant electron or photon energy not measured in the ECAL. In CMS, this typically means hadrons. In addition, it is also indirectly responsible for measurement of missing transverse energy, $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ (see 4.4.7). The $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ is calculated from imbalance in momentum in the transverse direction, and helps measure the energy of neutrinos, or other such particles, which do not interact with the CMS. In order to help measure $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ accurately, the HCAL needs to be hermetic, i.e. it should effectively capture each and every particle produced in the collision irrespective of the region in which it is produced. Unlike the ECAL which is homogeneous, the HCAL belongs to a different class of calorimeters called sampling calorimeters [68]. In such calorimeters, the material that produces light and measures energy (scintillator) is different from the material that produces the particle shower (absorber). It is chosen to be made of alternating and repeating layers of dense absorber and tiles of plastic scintillator. The dense absorber tiles are made of brass and steel. Hadrons interact with absorber plates to produce numerous secondary particles. These secondary particles interact with successive layers of absorber to give rise to a hadron shower. When the shower passes through the plastic scintillator layers alternating with absorber, it causes the scintillating material to emit light. This light is collected by wave-shifting optical fibers that are embedded within each scintillator tile, and then sent to photosensors. The photosensors used in the HCAL are fast and radiation resistant, and convert the light to electronic signals. The energy deposited can then be measured using these signals. Obviously, because of the absorber layers, the entire energy deposited is not converted into electronic signals. However, we can estimate the total energy deposited in the calorimeter from the energy collected by sampling (and hence the name sampling calorimeter) the shower (using energy deposited) in multiple layers of scintillator.

Like all other CMS subdetectors, the layout of the HCAL is also very modular and divided into four distinct sections. This is illustrated in Figure 3.16. The inner hadronic calorimeter barrel (HB) lies between the outer radius of the ECAL (R=1.77 m) and the inner radius of the magnet (R=2.95 m), and collects particles within the region $|\eta| < 1.3$. Just like the ECAL, it is segmented into towers in $\eta \times \phi$ of size 0.87×0.87. Given the constraint on the size of the HB, an outer hadron calorimeter (HO) is placed just outside the solenoid to complement the HB. The HO is meant as a *tail-catcher*, and increases the effective thickness of the HCAL barrel, helping detect particles that have punched through the HB. The hadron calorimeter endcaps (HE) extend the pseudorapidity coverage to $1.3 < |\eta| < 3.0$. Further, the

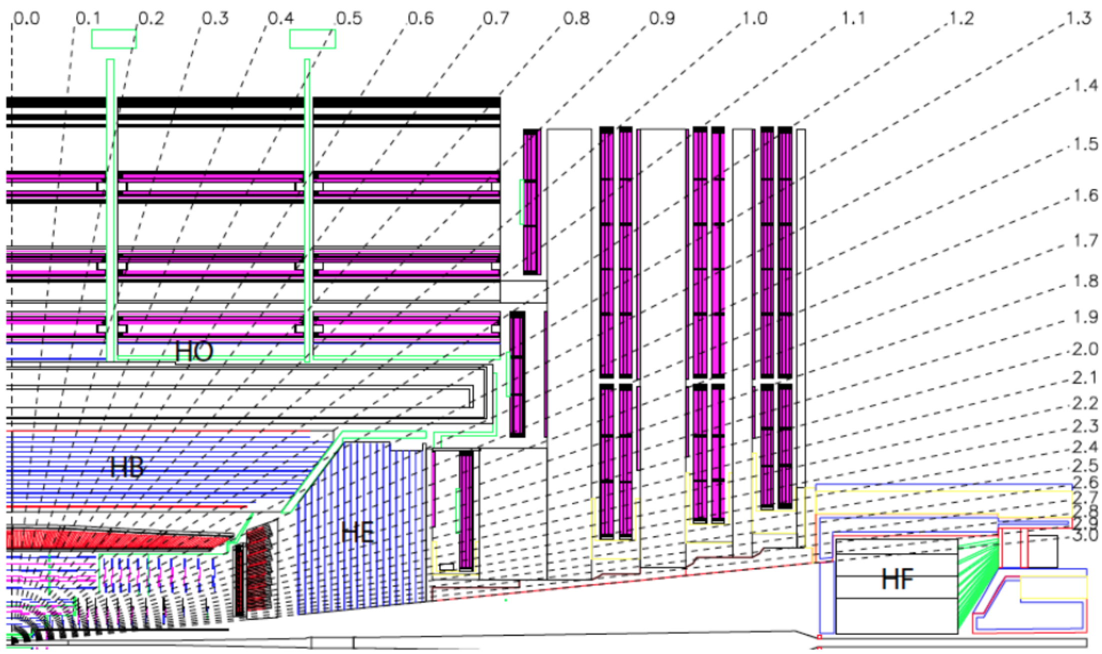


Figure 3.16: The CMS HCAL layout [56].

forward hadron calorimeters (HF) placed at a distance of 11 m from the interaction point cover the high pseudorapidity regions in the range $3.0 < |\eta| < 5.2$. These regions receive very high particle fluxes, and the HF is correspondingly instrumented with Cherenkov light detectors made of radiation-hard quartz fibers inserted in steel absorbers. The combined energy resolution of the CMS HCAL and ECAL for pions is approximately given by: $\frac{\sigma(E)}{E} = \frac{84.7\%}{\sqrt{E}} \oplus 7.4\%$ [69].

3.3.5 Muon system

Detecting muons precisely and efficiently is one of the most important tasks that the CMS detector was built to perform. When the CMS was first designed, one of its primary aims was the discovery of the SM Higgs Boson (h). One of the cleanest (easily separable from background), and so called "golden" decay channels of the h is its decay into 4 muons. Excellent muon detection was therefore considered paramount to CMS design, and hence the name Compact *Muon* Solenoid.

Muons are minimally interacting particles. They are capable of traveling through several meters of iron without interacting and are not stopped by the ECAL, HCAL or the solenoid. The muon system is thus the outermost layer of the CMS, where muons are the only particles that register a signal. This makes their efficient and accurate identification and energy measurement easier than that of other particles. Incidentally, a muon is one of the two final state particles of both the searches described here. We also use muons for triggering (selecting) events (see section 3.3.7) for the search.

The muon system is comprised of three different types of gas detection systems, each serving their own purpose: 250 Drift Tubes (DT), 540 Cathode Strip Chambers (CSC) and 610 Resistive Plate Chambers (RPC). The layout of the CMS muon system is illustrated in Figure 3.17. In each of the three designs, a gas is ionized by the muons passing through the system. The electrons thus knocked off travel in an electric field towards the positive anode, and register electrical signals (hits). These hits can be reconstructed into tracks (similar to the tracker tracks, see 4.4.2) and can be potentially used to determine the momentum of the muon. However, in order to achieve a much higher muon reconstruction accuracy, information from both the tracker and the muon system is used. In particular, the muon system complements the tracker in better identification of muons, and also in improving the energy resolution of muons with high transverse momentum.

Like all other CMS subdetectors, the muon system is organized into a cylindrical barrel and endcaps. In the barrel region ($|\eta| < 1.2$)) the particle rates are low, and magnetic field is uniform [56]. This allows for the use of DT chambers as detection elements. Each drift tube (42mm by 13 mm, shown in Figure 3.18) contains a positively charged stretched wire within a gas volume. The gas (85% Argon (Ar) and 15% Carbon Dioxide (CO_2)) is ionized by muons. Using the positions where the electrons hit on the anode wire and the distance away from the wire (calculated by multiplying

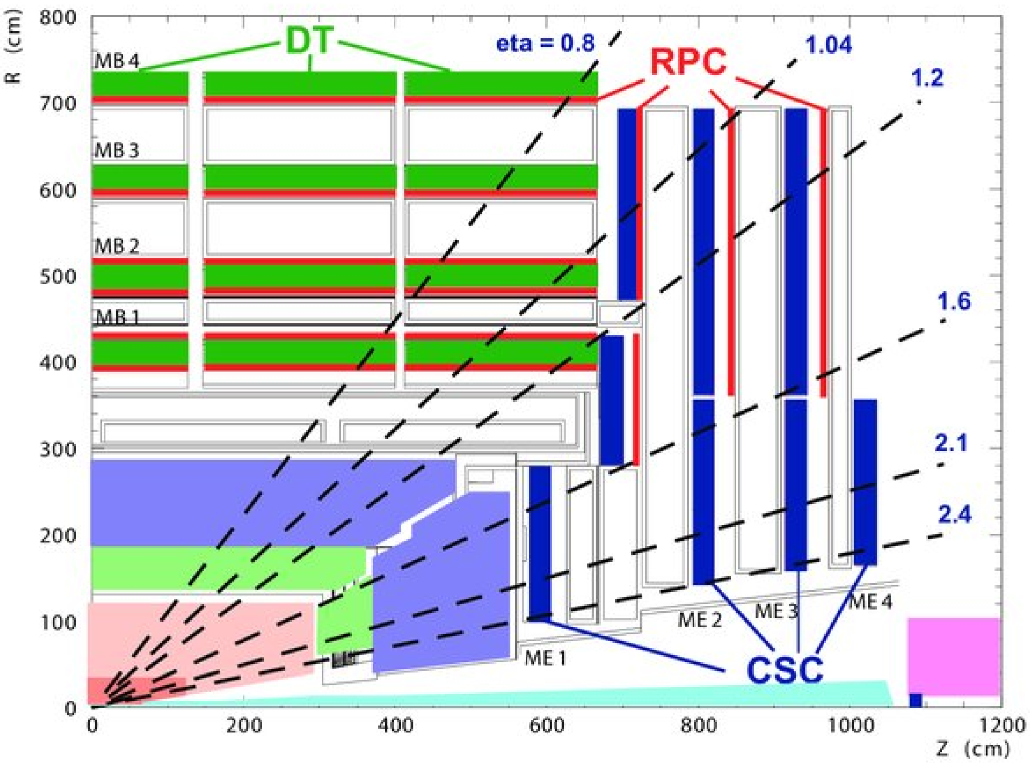


Figure 3.17: The CMS Muon system layout [70].

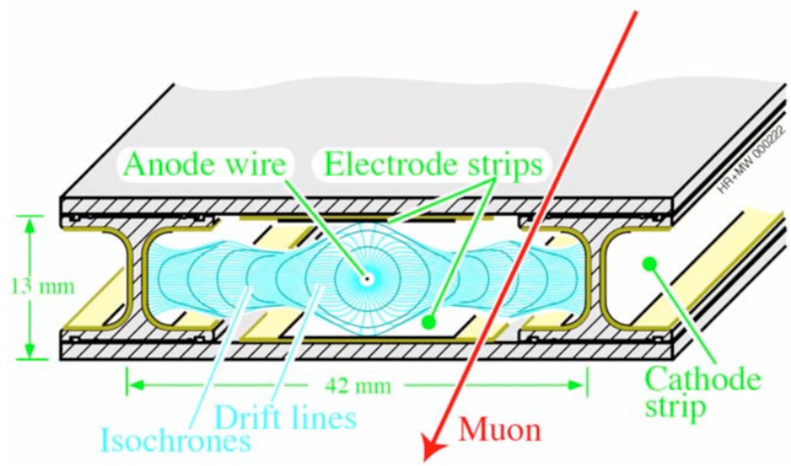


Figure 3.18: A Drift tube schematic [56].

the speed of an electron in the tube by the time taken), the coordinates for a muon's position can be calculated. The DT chambers give a position resolution of 100 μm, and consist of several groups of alternating layers, with each layer containing up to 60 DTs.

In the endcap region, particle intensities are high and the magnetic field is non-uniform. This makes DTs unsuitable for endcaps, and CSCs are used instead. Each CSC chamber is trapezoidal in shape (Figure 3.19) and consists of "intersecting" arrays of positively-charged wires and negatively-charged copper strips. This is housed within a gas volume comprised of 40% Argon (Ar), 50% Carbon Dioxide (CO_2) and 10% Tetrafluoroethane. Electrons and ions, produced when passing muons ionize the gas, travel to the anode wires and cathode strips respectively. Charge pulses are produced in both the strips and the wires (perpendicular to strips), providing two coordinates for each muon hit. Unlike DTs, CSCs can support the high particle flux and cope with the large and non-uniform magnetic field in this region.

The third component of the muon system are the RPCs which are used specifically for the purpose of triggering on muons, i.e. to make very fast decisions whether or not a muon (having a certain amount of energy or more) is present in the event. Because

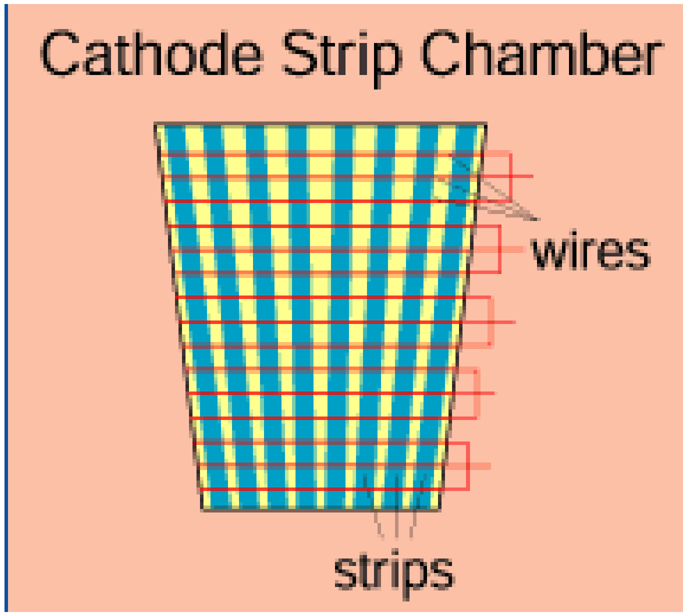


Figure 3.19: A Cathode Strip Chamber schematic [71].

the response of CSCs and DTs are restricted by their relatively longer drift times, RPCs are placed in both the barrel and endcaps for the purpose of triggering. Just like DTs and CSCs, RPCs are also gas based detectors, consisting of positively-charged anode and negatively-charged cathode plates (made of high resistivity plastic) placed parallel to each other (Figure 3.20). The gas that separates these plates (2.3 m in length and ∼2 m wide) is composed of 96.2% Tetrafluoroethane ($C_2H_2F_4$), 3.5% isobutane (C_2H_{10}) and 0.3% Sulfur Hexafluoride (SF_6). The fast timing response (1 ns) of RPCs is achieved by keeping the plates separated by a small consistent distance of about 2 mm. An avalanche of electrons is produced when a muon passes through, travels towards the anode plate and is picked up by external metallic strips as a pattern of hits. This gives a quick measure of muon momentum and helps in making fast trigger decisions. Given the bunch crossing time is 25ns, the 1ns timing of RPCs makes them suitable for quickly identifying a muon associated with a particular bunch crossing.

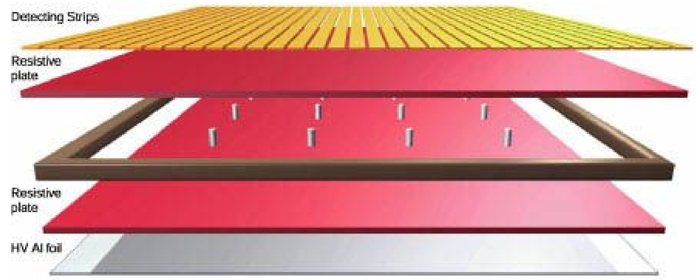


Figure 3.20: A resistive plate chamber schematic [71].

3.3.6 The magnet

The description of the CMS i.e., Compact Muon *Solenoid* is incomplete without a description of its *solenoid* magnet. The CMS magnet is the largest superconducting magnet ever built [72]. It weighs 12,000 tonnes and has an axial magnetic field of 3.8 T. Like any other solenoid, the CMS magnet is also made of coils of wire that produce a magnetic field when electricity passes through them. The niobium-titanium coils are cooled down to a temperature of $-268.5°$ C (a degree higher than outer space) making the magnet superconducting, allowing electricity to flow without resistance and thus producing a very strong magnetic field. The steel return yoke of the magnet (within which the muon detectors are interleaved) is responsible for most of the detector's weight, and also provides structural support to the detector. When the magnet is fully powered up it contains about 2.3 GJ or 638.94 kWh of energy.

Charged particles trajectories bend in a magnetic field due to the Lorentz force. The stronger the magnetic field is the smaller the radius of curvature of a particle's trajectory is. The more energetic a particle is the larger is its radius of curvature. So, if we know the trajectory of the particle (determined precisely in CMS by the tracker), we can determine its momentum. A strong magnetic field thus necessary for precise measurement of momentum of particles with very high energy.

3.3.7 CMS trigger

The current LHC produces pp collisions at about the rate of 40 MHz. Considering a typical event size at CMS is 150 kB, this comes out to be about 5.6 Petabytes per second. Only a small fraction of these events can thus be recorded offline. Even if it were somehow feasible to record this entire volume of data, most of the collisions produced are mundane soft collisions. We are interested only in events that have certain characteristics, such as events having high transverse momentum of the particles associated with it. The trigger system is used to select these events of interest, making its efficient functioning crucial for physics searches performed with the data collected by CMS [73]. A sophisticated two-level trigger system is thus employed, organized in two consecutive stages- the Level-1 (L1) trigger and the High Level Trigger (HLT).

The L1 trigger is the first stage of CMS trigger system. It reduces the rate from 40 MHz to less than 100 kHz and has latency of 3.8 μs. In other words, it has to make a decision in under 3.8 μs. It thus uses only coarse information from the calorimeters and the muon systems, and runs relatively simple algorithms on custom hardware. In particular, it does not use any information from the tracker. The L1 trigger is divided into a L1 Calorimeter Trigger and L1 Muon Trigger; each send information to the Global L1 trigger which makes the decision on whether to keep or discard the event. The information sent by the L1 Calorimeter Trigger and L1 muon trigger include energy sums (e.g. sums of transverse energies deposited on locally grouped sets of crystals, called trigger towers), position, isolation (see section 4.4.8) and quality flags. The global trigger combines this information from the two systems and runs about 300 algorithms in order to make a decision [74]. The hardware used by the L1 trigger is custom made. Reprogrammable circuits like Field Programmable Gate Arrays (FPGA) form the primary component while application-specific integrated circuits (ASICs) are also used in certain specific use cases.

Everything that is deemed fit to record by the L1 trigger is received by the HLT.

The HLT has relatively longer (150 ms) to make a decision and is fully software-based. It uses information from the entire detector (including the tracker) and partially reconstructs the event. It runs over 400 complex selection algorithms and runs on a massive computer farm. The HLT brings down the output event rate to less than 1 kHz. The events that successfully pass the HLT selection process are permanently recorded for analysis.

3.3.7.1 Triggers for h $\rightarrow \mu\tau_e$ and H $\rightarrow \mu\tau_e$ searches

For the analyses, only a subset of pp collision data collected by CMS are used. This is chosen by requiring that those events pass (satisfy) a certain trigger depending on the final state signature of the analysis. In the both the analyses, the final state consists of a high p_T muon that comes directly from the Higgs accompanied by a relatively low p_T electron from the tau lepton decay (see sections 5.2.1 and 5.3.1). In h $\rightarrow \mu\tau_e$ analysis, the data used is composed of events that satisfy a trigger requiring an isolated muon (see section 4.4.8) having a p_T of at least 24 GeV to be present in the event. In H $\rightarrow \mu\tau_e$ analysis, the Higgs Bosons are heavier and p_T requirement is increased to 50 GeV and there is no isolation requirement on the muon. The same trigger requirement is also required to be passed by simulated samples used in the analysis. The efficiency of the trigger in selecting pp collision events is different to that of its efficiency in selecting simulated samples. This is corrected by using scale factors (depending on p_T and η) to match the efficiency in simulation to that of the data. The efficiencies were calculated by the CMS Muon Physics Object Group via tag-and-probe methods [75] using $Z \rightarrow \mu\mu$ events. Briefly, the tag-and-probe method works in the following way. One of the muons (called the tag) is required to pass strict selection criterion, while the other (called the probe) is required to pass more relaxed criterion. Given the invariant mass of the $\mu - \mu$ system is required to within a narrow window of the Z mass, the probe muon is also very likely to be a real muon. The percentage of probe muons that pass the criterion we are testing for (identification, isolation, trigger etc.) gives the efficiency. The efficiency plots are shown in Figure 3.21.

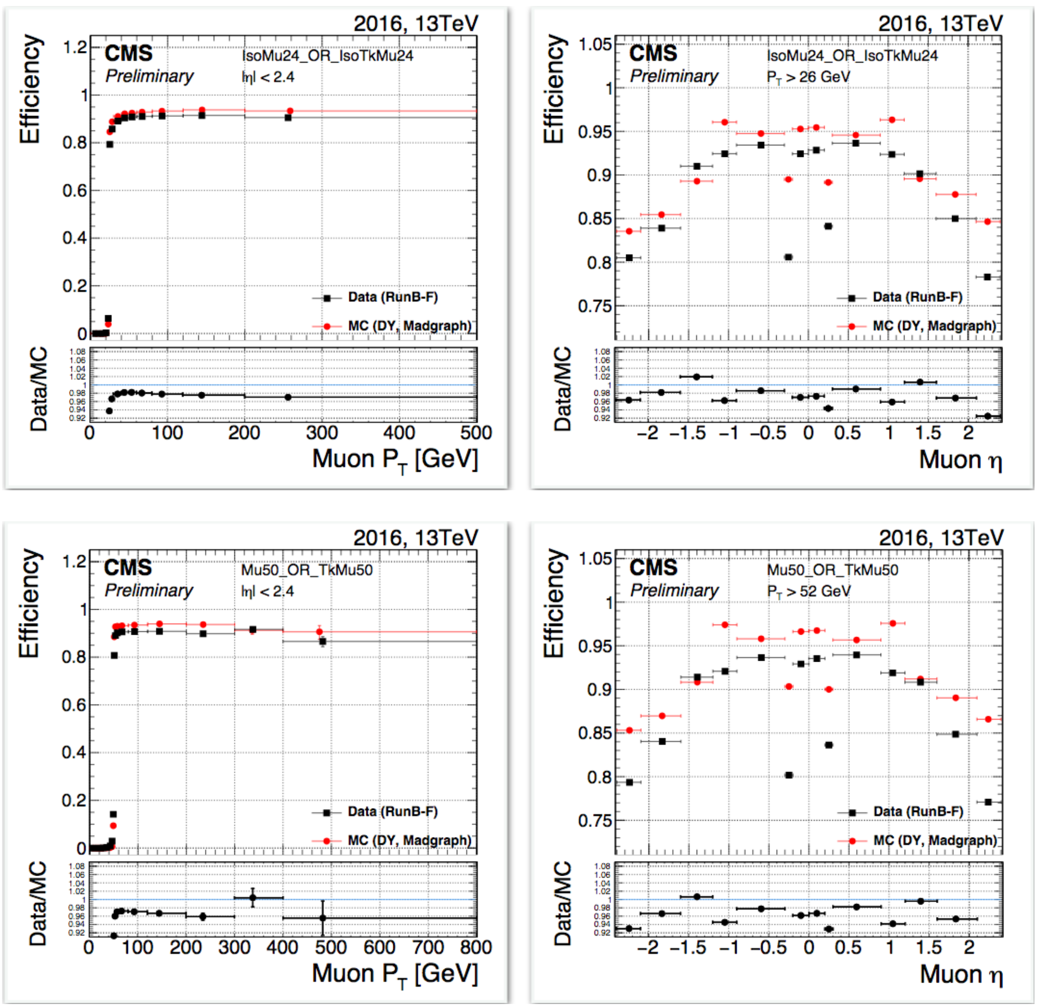


Figure 3.21: Efficiency of trigger used in the $h \to \mu\tau_e$ analysis (top) and $H \to \mu\tau_e$ analysis (bottom), as a function of p_T (left) and η (right), for data (black) and simulation (red) [76].

CHAPTER 4

OBJECT RECONSTRUCTION AND EVENT GENERATION

> Any one who considers arithmetical methods of producing random digits is, of course, in a state of sin. For, as has been pointed out several times, there is no such thing as a random number — there are only methods to produce random numbers, and a strict arithmetic procedure of course is not such a method.
>
> *Jon von Neumann*

4.1 Introduction

This chapter is divided into two parts. In the first part, the procedure for the generation of simulated events is described. This is done in several distinct stages with the output of one stage serving as an input for the next. A suite of software packages, developed mostly by the particle and nuclear physics communities, is used to achieve this. This part concludes by detailing the simulated datasets used in the analyses described in this thesis. In the second part of this chapter, the reconstruction of physics objects is described in detail. It starts with a description of the particle-flow algorithm which is a global event reconstruction scheme for the entire event. This is followed by descriptions of track, muon and electron reconstructions. Reconstruction of jets is described next followed by description of composite objects used in the analysis such as collinear mass and transverse mass.

4.2 Event simulation

A pp collision at the LHC, like any hadronic collision, is more complex than the hard interaction of two participating partons. The proton being a composite object, the colliding partons from the hard interaction are accompanied by other quarks and gluons that interact and rearrange themselves into color singlets due to color confinement. A pp collision thus consists of: the Hard Scattering which represents the part of the collision where two partons in the initial state interact by exchanging high transverse momentum, and the Underlying Event that represent the interaction of the everything else in the collision except the partons in hard scattering. In addition to the implementing the above, i.e. physics of a pp collision that produces a bunch of final state particles, the event simulation also has to include interactions of these particles with the CMS detector. Monte Carlo methods, that use generation of random numbers to simulate sampling from a given probability distribution, are used to model the above event simulations [77].

4.2.1 Monte Carlo method

Monte Carlo (MC) methods (named after a famous casino in the city state of Monaco) are a broad class of computational algorithms that rely on repeated random sampling to obtain numerical results [78]. In particle physics, these methods play a key role in generation of events and are used primarily for : generation of samples from specified probability distributions, and the calculation of integrals. Programs which implement the above method, called MC event generators, use generation of random numbers to make decisions about physics processes. These can range from selection of processes that are generated in the collision, to which decay channel a particle decays in, to making decisions on how the particle interacts with detector material. Usually, each such decision is the result of a draw from a distribution which depends only on the current state the process is in, and not on previous states. The MC generator is provided as input the distributions that represent the physics of the generated particles, their production, their decay modes and their couplings. A MC generator starts by using a pseudo-random number generator that usually outputs a random number between 0 and 1. Although, true random number generation can only be done by physical processes, modern pseudo-random number generators are known to generate numbers with a high degree of randomness. Starting from this distribution, the MC event generator uses one of the various methods such as the inverse-transform method, or the rejection sampling method to convert this uniform distribution into a desired probability distribution, $p(x)$. It is then possible to generate random numbers according to this distribution to simulate physical processes.

4.2.2 CMS simulation pipeline

The MC simulation of events in CMS consists of the following sequential steps. The first step is simulation of the Hard Scattering. As mentioned earlier, this represents the primary hard interaction in a collision where two partons in the initial state interact by exchanging high transverse momentum resulting in a final state with two or more partons. The parton density function (pdf) which parametrizes the distributions of the partons inside each hadron are used to model the momenta of incoming partons. It represents the probability of finding a parton of a certain flavor at a certain longitudinal momentum fraction, when the hadron, that contains it, is probed at a certain scale. The PDF are extracted from fits to the data, mainly from ep collisions, and various PDF sets are available for each parton flavor. Commonly used pdf sets include ones provided by the CTEQ, HERA (H1 and ZEUS) and NNPDF collaborations. The LHAPDF library provides a unified C++ interface to all major PDF sets. The matrix element formulation is used to model the hard scattering process to leading order in perturbative QCD, or to higher orders depending on the generator. The next step is simulation of the parton shower. The hadronization and radiation of quarks and gluons in the initial and final states cannot be feasibly encapsulated in the matrix element computation. Parton shower describes these missing parts. The matrix element calculations are combined with the parton shower by one of the different matching schemes which ensure that there is no double counting of terms present in both the matrix element and the parton shower expansion. The matching schemes that are most often used are MLM [79], CKKW [80] and FxFx [81]. The simulation of the Underlying Event comes next. Underlying event includes everything in the collision that is not associated with the primary hard scattering process. This consists mostly of soft QCD interactions, and implemented using the MC event generators and interfaced with the matrix element simulation. The hadronization of the quarks and gluons is simulated next and it consists of recombination of individual partons into colorless hadrons. Lastly, the decay of short-lived particles is simulated.

An important part of the event generation chain is the simulation of pileup. 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). Event generators add pile-up events to the hard scattering samples by randomly simulating soft inelastic collisions and overlapping them. The distribution of the number of pileup interactions in data is hard to predict. MC event generators usually produce events for a scenario with a higher number of pileup vertices, and with a flat distribution of number of vertices . This is afterwards reweighted to match the observed distribution of pileup interactions in data.

Several MC generators have been developed. Some of these can produce all components of the above simulation pipeline while some calculate only the matrix element and need to be interfaced with other generators for the simulation of remaining parts. Pythia [82] and Herwig [83] can produce the entire chain while Powheg [84–89], aMC@NLO [90] and Madgraph [91] produce up to matrix element stage. Powheg and aMC@NLO can perform next-to-leading order calculations.

Finally, the Geant4 (GEometry ANd Tracking) [92] package is used to simulate the interaction of physical particles after the collision, produced by pipeline described above, with a sophisticated and complex simulation of the detector itself. This simulated detector response is used as input for the same physics reconstruction algorithms (described in the next section), that are used to reconstruct the data, thus enabling a direct comparison of the two. If differences are observed in the behavior of these reconstruction algorithms for MC events in comparison to observed data, the MC events are tuned to the behavior observed in data.

4.3 MC samples used for the analyses

The ggH and VBF Higgs boson samples are generated with POWHEG 2.0 while an extension of POWHEG 2.0 [93] is used for the WH and ZH simulated samples. For the H $\rightarrow \mu\tau_e$ analysis, only the gluon fusion (ggH) production mode has been considered. Samples are generated for a range of H masses from 200 to 900 GeV.

The Z + jets and W + jets processes are simulated using the MG5_aMC@NLO generator at leading order (LO) with the MLM jet matching and merging scheme. The same generator is also used for diboson production which is simulated at next-to-LO (NLO) with the FxFx jet matching and merging scheme. POWHEG 2.0 and 1.0 are used for top quark-antiquark ($t\bar{t}$) and single top quark production, respectively. The POWHEG and MADGRAPH generators are interfaced with PYTHIA 8 for parton showering, fragmentation, and decays.

As mentioned earlier in this chapter, 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 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 isolation and b-jet tagging efficiencies.

4.4 Physics object reconstruction

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

4.4.1 Particle flow

The overarching algorithm used by CMS to produce a unified global (synchronized for all sub-detectors) description of an event is the particle-flow (PF) algorithm [57]. The idea behind the PF algorithm is that if the basic building blocks or elements from the various sub-detectors can be correlated in a well-defined way, then the description of the event and that of each particle in it can be refined by using the global information from the entire detector. The ALEPH experiment at the CERN LEP collider was the first experiment to use such a holistic approach towards event reconstruction. The CMS experiment, owing to its very granular layers of sub-detector, is the first hadron collider experiment to successfully use particle-flow. The first step of the PF algorithm is the linking of the several building-blocks or PF elements that a single particle can give rise to, across different sub-detector layers. The link algorithm tests pairs of neighbors in the $\eta - \phi$ plane and combines (links) them to form PF blocks. Reconstruction and identification algorithms are run according to a predefined sequence in each of these PF blocks. First, muon candidates are reconstructed and identified. If a muon candidate successfully passes PF quality criterion, the PF elements associated with it are removed from the block. Electron reconstruction proceeds next with electron candidates successfully becoming PF electrons if their tracks in the tracker, when extrapolated, have a corresponding energy deposit in the ECAL. The reconstruction procedure of muons, electrons and tracks are discussed in detail in the sections 4.4.3, 4.4.2, 4.4.4. The PF block now consists of photons and hadrons. To reduce fake track identification, tracks with momentum uncertainty larger than the resolution of the calorimeters are removed at this stage. The remaining tracks are then associated with charged hadrons. The remaining calorimeter energy deposits are then associated with photons (ECAL) and hadrons (HCAL). In this manner, PF finally produces a list of all electrons, photons, muons, charged hadrons and neutral hadrons in the event with optimally determined direction, charge and energy.

4.4.2 Track and primary vertex reconstruction

Tracks of charged particles, that traverse the CMS tracker (described in section 3.3.2), are reconstructed [61] using hits from the pixel and strip detectors in the tracker. Hits are reconstructed by clustering signals above specified thresholds in the pixel and strip channels, and then estimating the cluster positions and uncertainties in a local orthogonal system plane of each sensor. During track reconstruction, a translation is made between the local coordinate system of these hits to the global coordinate system of the tracks. The software used to reconstruct tracks by CMS is called the Combinatorial Track Finder (CTF) and is adaptation of the Kalman filter [94]. Tracks are reconstructed using a iterative procedure with the basic idea being, tracks that are easiest to find (e.g., high p_T tracks, and tracks produced near the interaction region) are searched in the initial iterations, with subsequent iterations looking for more difficult sets of tracks (e.g., low p_T tracks , or tracks produced far from the interaction region). Hits unambiguously assigned to the track in the previous iterations are removed for the subsequent ones, thus reducing the combinatorial complexity. Each iteration can be divided into four sequential steps.

The first step is seed generation which provides initial track candidates that define the starting trajectory parameters and associated uncertainties of potential tracks. Charged particles follow helical paths in the quasi-uniform magnetic field of the tracker, requiring a total of five parameters to determine the trajectory. These five parameters are extracted using two or three hits in the inner region of the tracker. The seeds are constructed in the inner part (and then tracks constructed outwards, and not in the opposite manner) because the high granularity of pixel detectors (in contrast to outer strip layers) ensure that low fraction of channels are hit. Also, particles like pions and electrons interact inelastically with tracker material or lose energy due to bremsstrahlung radiation as they traverse through the tracker to its outer regions, making the idea of constructing seeds in the inner region a better choice.

The second step in track generation is track finding which is closely based on the Kalman filter. It extrapolates the seed trajectories along the expected path of a charged particle, beginning with an estimate of the track parameters provided by the trajectory seeds generated in the last step. It then uses the location and uncertainty of detected hits, and estimations of effects such as Coulomb scattering, at successive detector layers, to build track candidates, updating the parameters at each layer. First, using the parameters of the track candidate, evaluated at the current layer, an analytical extrapolation is done that determines which adjacent layers of the detector the trajectory can intersect. This takes into account the current uncertainty in that trajectory just like a Kalman filter. Secondly, a search is performed for silicon modules in these layers that are compatible with the extrapolated trajectory. All compatible modules in each layer are then grouped into mutually exclusive groups, such that no two modules in each group overlap. The collection of all hits from from one such module group forms a group of hits. Finally, new track candidates are formed by adding exactly one of the compatible hits from each group, to each original track candidate. The modules in a given group are mutually exclusive and a contribution of more than one hit from each group is not expected. The trajectory parameters of the new candidates are then updated by combining the information from the added

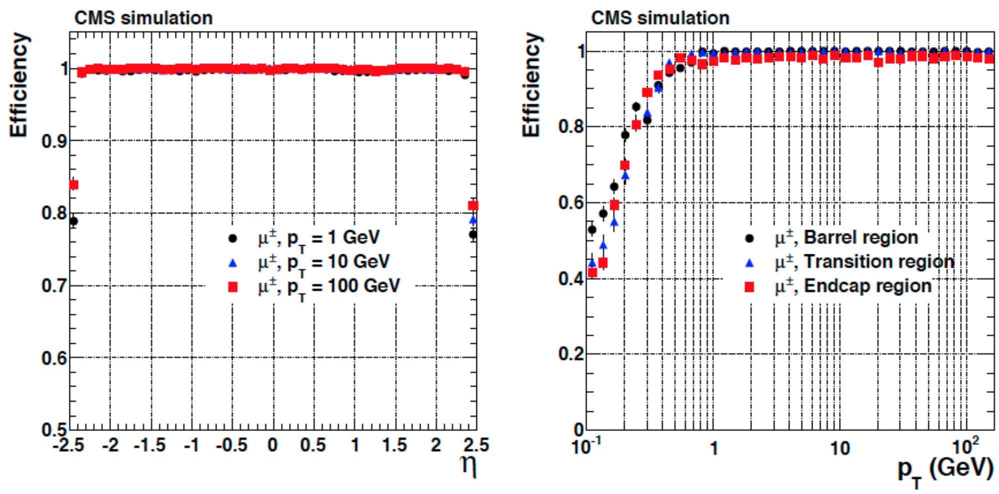


Figure 4.1: Track reconstruction efficencies for single isolated muons as a function of η and $\mathrm{p_T}$ [61].

hits with the extrapolated trajectory of the original track candidates. Figure 4.1 illustrates the reconstruction efficiency of tracks in case of isolated muons.

The third step in track generation track fitting. In this step the collection of hits from the last step are refitted using a Kalman filter and smoother, to provide a best possible estimate of parameters for each track trajectory. The procedure described above, in conditions as challenging as the LHC, can yields several fake tracks that are not associated with any charged particle passing through the tracker. The fourth and final step applies several quality requirements to the set of reconstructed tracks and substantially reduces the fake contribution. The requirements are based on criteria such as the minimum number of layers the track has hits in, how compatible its origin is with a primary vertex, how good a fit it yields etc.

Proton-proton interaction vertices are reconstructed by selecting tracks that are produced promptly in the primary interaction region. The selected tracks are then clustered on the basis of their z-coordinates at their point of closest approach to the centre of the beam spot, which represents a 3-D profile of the region where the LHC beams collide inside the CMS detector. The exact positions of the vertices are then obtained from these clustered candidates, by using a fitting procedure, called the adaptive vertex fitter [95]. The vertex which has the largest sum of squared transverse momenta of tracks originating from it is considered the primary interaction vertex.

4.4.3 Muon reconstruction

Hits in the muon system (described in section 3.3.5) and tracks (muons being charged particles leave tracks in the tracker) from the tracker are used to reconstruct muons [75]. When muons traverse a muon subdetector (such as RPC, CSC or DT) in the muon system, they ionize the gas in the chambers. The electrical signals produced on the wires and strips as a consequence of the ionization are read out by electronics systems that associate these "hits" with well-defined locations in the detector. Various algorithms depending on the subdetector technology are used to reconstruct these hits. Reconstruction of muon tracks using these hits first proceeds independently of track reconstruction in the tracker. These tracks, called *standalone-muon tracks*, are built using these reconstructed hits from the muon system using a Kalman filter. Muon tracks are also built inside-out by propagating tracker tracks (described in previous section) with transverse momentum above 0.5 GeV to the muon system and matching them to (straight-line) segments of hits in DT or CSC. If a match is found, the tracker track qualifies as a *tracker muon track*. Muon tracks are also built outside-in by matching standalone-muon tracks with tracker tracks, and combining information from both using a Kalman filter fit. These are called *global muon tracks*. The global muon reconstruction is especially efficient for muons leaving hits in several muon stations. The *tracker muon* reconstruction is more efficient for low p_T muon candidates but it can cause fake muon tracks due to hadronic particles which *punch-through* to the innermost muon stations. The *global muon* reconstruction has high efficiency for muons penetrating through more than one muon station, and reduces the muon misidentification rate compared to tracker muons. Combining both *tracker* *muon tracks* and *global muon tracks*, the efficiency for reconstructing a muon is as high as 99%. The particle-flow algorithm applies a set of requirements, based on various quality parameters from muon reconstruction as well as information from other sub-detectors, to reconstructed candidates. The PF muon candidates used in the analyses described in this thesis were required to satisfy the following set of criterion to be identified as a muon:

- Must be a global muon or a tracker muon.
- Must have at least one hit in the pixel subdetector of the tracker
- χ^2 of the compatibility between the position of the standalone and trackers tracks < 12
- Transverse impact parameter of the associated tracker track with respect to the primary vertex $d_{xy} < 2mm$
- Longitudinal distance of the (origin of)associated tracker track with respect to the primary vertex $d_z < 5mm$
- constraints on muon segment matching compatibility between tracker and muon system dependent on if it is a global muon

The efficiency of the above selection for muon identification is illustrated using a plot from a study performed by the CMS Muon Physics Object group in Fig. 4.2. As can be seen from the plots, there is a difference in the efficiencies in data and MC simulation. This is corrected using a set a of scale-factors applied as a function η and $\mathrm{p_T}$ to adjust the efficiency in simulation to get it to match the efficiency in data.

The momentum of muons is measured by CMS using one among different possible ways involving the tracker and muon system [96] and then using the PF algorithm to refine this measurement exploiting information from the full event.

4.4.4 Electron reconstruction

Besides muons, electron form the other primary part of the final state of the decay we are searching for in this thesis. Electrons, in the CMS, are reconstructed using

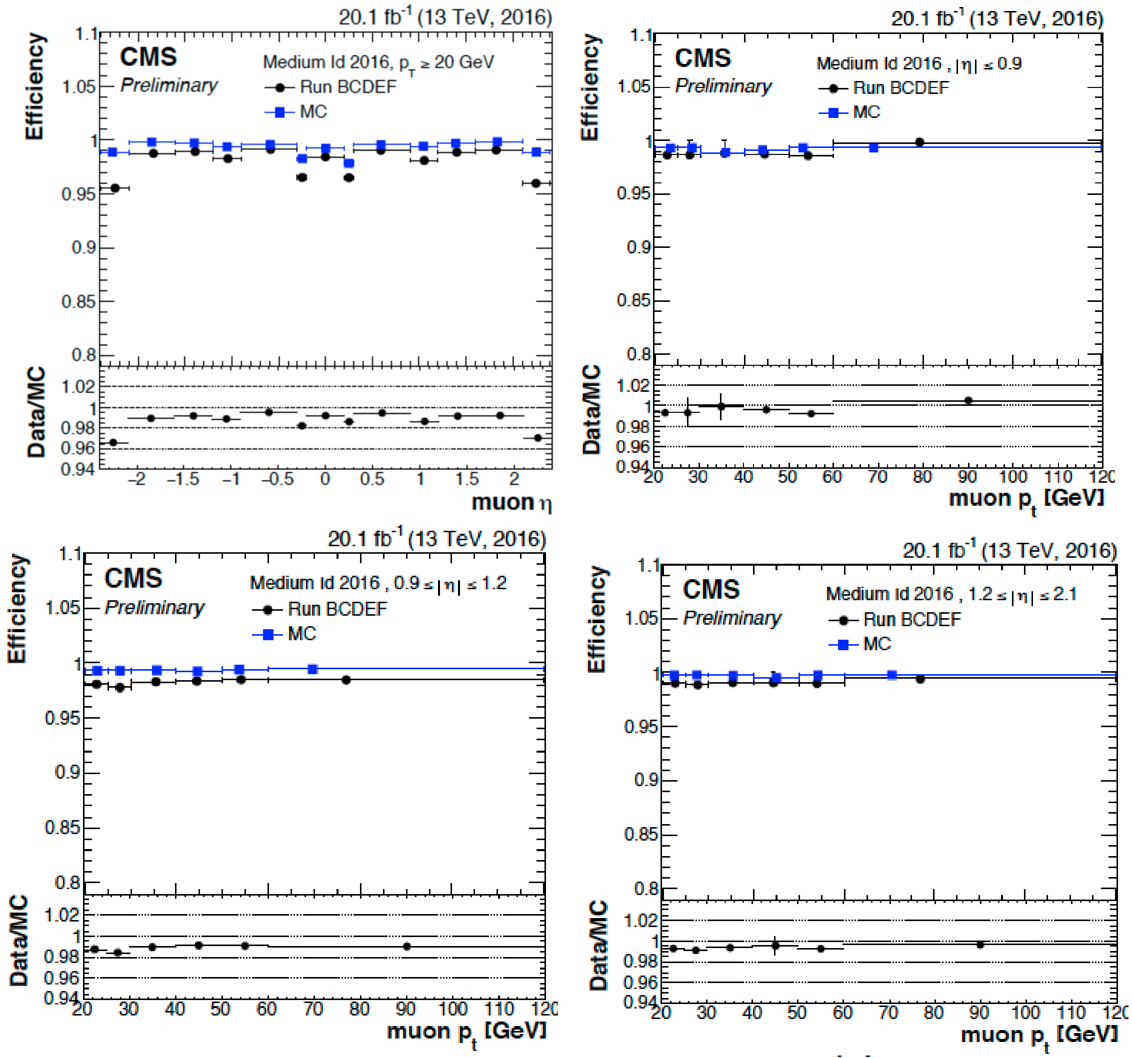


Figure 4.2: Efficiency of muon identification as a function of η and p_T, for data (black) and simulation (blue) [76]

clusters of energy formed in the ECAL (described in section 3.3.3) and associating them with tracks from the tracker [97]. The reconstruction of electrons is made complicated by the fact that they can radiate a significant amount of energy before reaching the ECAL. This happens due to the radiation of bremsstrahlung photons caused by the interaction of electrons with atoms as they pass through the tracker. This loss can range from 33% to as high as 86% depending on η (as a consequence of the fact that the amount of detector material the electron has to cross is η dependent). In order to measure the an electron's energy, clustering algorithms thus need to take into account the energy from these bremsstrahlung photon showers, together with the deposit made in the ECAL by the electron. The energy from these radiated photons spreads primarily in the ϕ direction, owing to bend in electron trajectory in the magnetic field of CMS. The spread in the η direction is relatively small. These facts are used by the clustering algorithms.

The algorithm used to cluster the electron energy deposit in the ECAL barrel is called the *hybrid* algorithm. It exploits the above property of the electron shower shape, and uses the geometry of the ECAL to form clusters that are narrow in η direction but wide in ϕ direction. Starting with a (seed) crystal containing the largest amount of energy deposited in a considered region above a certain threshold (1 GeV), it adds 5x1 arrays of crystals in $\eta \times \phi$ around the seed crystals in both directions of ϕ if the energy contained in the arrays is above another predefined threshold (0.1 GeV). Contiguous arrays are merged into clusters, and finally a electron supercluster is formed from all such strip clusters which have at least one seed strip with energy above another predefined threshold (0.35 GeV). The position of the supercluster is computed as the energy-weighted mean of the cluster positions, whereas its energy is simply taken as the sum of the energy of all its constituent clusters. In the ECAL endcap a different clustering algorithm is used owing to different geometrical arrangement of the crystals. This algorithm called the 5x5 algorithm starts similarly with a seed crystal with maximum energy in a local region, and satisfying the minimum energy requirement of 0.18 GeV. Clusters of 5x5 crystals are progressively grouped around the seed crystal, making a supercluster, if the total cluster energy exceeds 1 GeV and they are within ± 0.7 and ± 0.3 respectively in η and ϕ around the seed crystal. The position and energy of the supercluster is calculated in the same manner as the barrel. The energy from the preshower is also added into the supercluster, using it's most energetic cluster and it's maximum distance in ϕ to other clusters, and extrapolating it to the preshower plane to define the spread in the preshower. The thresholds used in the above algorithms were optimized using simulation and adjusted during data taking periods.

The standard track reconstruction (section 4.4.2) is not efficient for electrons. This is because the standard approach is compromised by the large radiative losses in the tracker leading to a poor estimation of track parameters [61]. Therefore, a dedicated tracking procedure is used for electron candidates that uses information not only from the tracker but also the ECAL. Just like the standard track reconstruction procedure, the first step in electron track reconstruction is seeding. This is done in two ways and the results are then combined. In the first method, superclusters from ECAL are used. As mentioned earlier, owing to strong magnetic field, the bremsstrahlung photons emitted by the electrons deposit energy in the ECAL at η values similar to that of the electron, but at different ϕ leading to a spread. The ECAL supercluster algorithms described above recover this energy. The position and energy of these reconstructed superclusters along with the assumption that the electrons originated close to the center of the beam spot can be used to constrain the trajectory of the electron through the tracker. Hits in the first layers of the trackers compatible with these trajectories are deemed electron seeds. In the second method of seeding, the "opposite" is done. Tracks constructed by the regular tracking algorithm are extrapolated to the ECAL and matched with a supercluster. The seeds corresponding to such matching tracks are retained as electron seeds. The seed collections from these two methods are merged leading to a increase in overall efficiency of the seeding procedure. These seeds are then used to initiate electron track finding and fitting phases. This track finding procedure is similar to that used in standard tracking except for small adjustments. The χ^2 fit thresholds used by the Kalman filter to decide whether a hit is compatible with a trajectory (see section 4.4.2) is weakened to accommodate tracks that deviate from their expected trajectory because of bremsstrahlung. Similar adjustments are made to the penalties assigned to track candidates for passing through a tracker layer without being assigned a hit. The final track fit uses a modified version of the Kalman filter, called the Gaussian Sum Filter (GSF), to account for the fact that the energy loss of an electron traversing the tracker material is non-Gaussian. This makes it unsuitable to use a conventional Kalman filter algorithm which assumes gaussian distribution. The GSF technique deals with this by approximating this non-Gaussian energy-loss distribution as the sum of several Gaussian functions, and is found to perform much better than the regular fitting procedure.

Finally, electron candidates are constructed by associating a electron track (called GSF track) produced by the above procedure with a supercluster in the ECAL. For ECAL-seeded candidates this association is made by a geometrical matching in $\eta - \phi$, while for tracker-seeded candidates a multivariate (MVA) technique that combines information from supercluster and GSF track is used. The electron charge is estimated using a combination of three procedures involving the use of the GSF track curvature, use of ECAL supercluster position and its relative position in ϕ to that of the first hit in the GSF track, and also by using KF tracks that have common hits with the GSF tracks. The combination of a best vote of three methods reduces the charge misidentification probability to 1.5% compared to 10% when using just the GSF track curvature method. Like other variables, the momentum of electrons is also estimated using a combination of tracker and ECAL measurements.

Further, several quality requirements are used on reconstructed electron candidates to identify (real/signal) electrons to suppress fake sources such as photon conversions, jets misidentified as electrons etc. These requirements are based on variables that fall into three broad categories: variables that compare measurements from ECAL and the tracker, variables that come only from ECAL (such as transverse shape of electromagnetic showers, ratio of energy fractions deposited in the HCAL to the ECAL) and purely tracking based variables (such as information from GSF track, difference between the information from GSF and KF-fitted tracks). These variables can be used in two ways: a cut-based method that uses the variables above directly to apply threshold requirements, or a multivariate (MVA) technique that uses all these variables as an input to a Boosted Decision Trees classifier to obtain a combined discriminator variable on which a threshold is applied. The BDT based method has much better performance as is illustrated in Fig 4.3. Two separate BDTs are trained depending on whether electron is required to pass a HLT triggering requirement or it is not. The trigger selection used in the analyses described in this thesis uses trigger based on muons. The BDT based identification criterion for non-triggering electrons is thus used in this analysis. The threshold corresponding to the working point used in this analysis has an efficiency of approximately 80%. The difference in efficiencies of electron identification based on the above criteria in data and MC simulation is corrected using a set a of scaled factors, applied as a function η and p_T. This adjusts the efficiency in simulation to get it to match the efficiency in data.

4.4.5 Hadronic tau leptons

Tau leptons decay hadronically in several ways or decay modes. The primary decay modes consist of: one charged hadron and up to two neutral pions, or three charged hadrons. The algorithm that is used to reconstruct hadronically decaying tau leptons in CMS is called the hadrons-plus-strips (HPS) algorithm [98, 99]. This

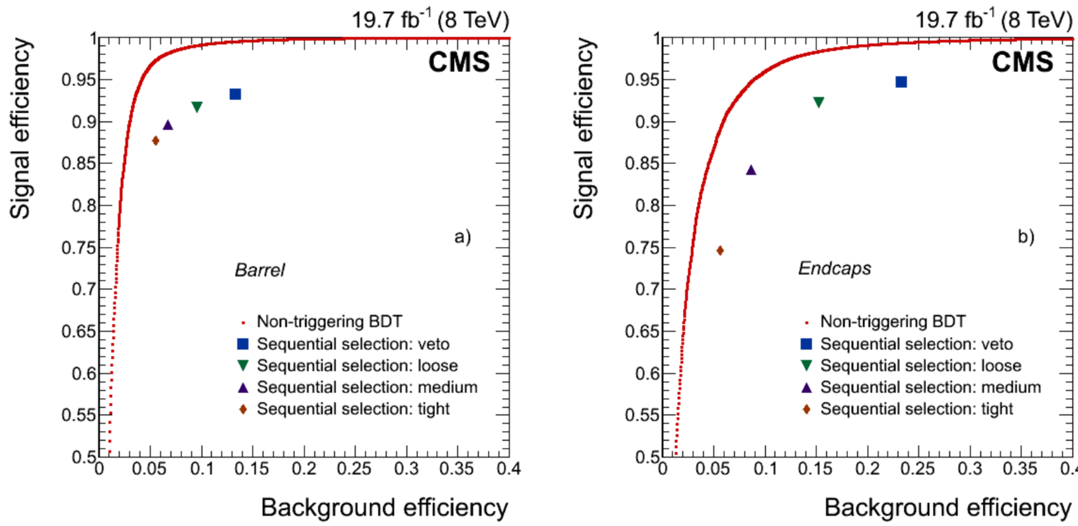


Figure 4.3: Performance of the BDT-based electron identification algorithm (red dots) compared with results from several working points of cut-based selection for electron candidates in the ECAL barrel (left), and endcaps (right) [97].

algorithm proceeds in two steps. In the first step, the topology of the candidate is checked to match the topology of one of the decay modes. The second step consists of a MVA-based discriminator (built using variables such as lifetime information, decay mode etc.), that is used to reject electrons, muons, quarks or gluon jets wrongly identified as hadronic taus. The reconstruction efficiency is also improved, in the case of converted photons from a neutral-pion decay, by considering PF photons and electrons from a strip along the ϕ direction. The final states in the analyses presented here do not contain hadronically decaying taus. Hence, all events which contain hadronically decaying taus are rejected.

4.4.6 Jet reconstruction

Jets are clusters of particles that are experimental signatures of quark and gluons which hadronize (due to color confinement) producing a narrow spray or “jet” of particles [100]. Jets are reconstructed in CMS by clustering PF objects. In order to group together objects into a jet, CMS uses the anti-k_T clustering algorithm.

This belongs to a broader class of clustering algorithms called sequential clustering algorithms which cluster objects into jet in a sequential order following a predefined set of rules. The general form of a sequential clustering algorithm is based on the quantities d_{ij}, which represents the distance between two entities, and d_{iB} which represents the distance of the i-th entity from the beam axis. These distances are defined as:

$$d_{ij} = min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta_{ij}^2}{R^2} \tag{4.1}$$

$$d_{iB} = k_{ti}^{2p} \tag{4.2}$$

where $\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$, k_{ti} is the transverse momentum of the i-th entity and R is the radius parameter which is set as 0.4. The parameter p governs the relative power of energy versus geometrical scales and the particular value of $p = -1$ defines the anti-k_T clustering algorithm. The algorithm first computes distance d_{ij} between all entity pairs present at that stage. If the minimum of those distances (say between entity i and j) is smaller than the minimum distance d_{iB} of any entity from the beam axis, those entities i and j are combined into a single entity. Otherwise, the entity closest to the beam axis is considered a jet and removed from the list of entities to be further clustered. This continues until all entities are clustered. The anti-k_T is dominated by high $\mathrm{p_T}$ particles which it clusters first, subsequently including softer and softer constituents. Soft particles tend to cluster with hard particles before they tend to cluster among themselves. A hard particle that has no hard neighbors within a distance 2R accumulates all the soft particles within a circle of radius R. It tries to produce jets with fairly conical shapes that are centered around the hardest particles of the event, and with boundaries that resilient to the effect of soft radiation.

Jets being complex objects suffer from several effects that cause their energy, reconstructed as described above, to differ from their true values. Multiplicative correction factors are applied to calibrate their $\mathrm{p_T}$ and to ensure a uniform response in η [101, 102]. A total of four multiplicative corrections are applied. Firstly, the energy coming from pileup that has been clustered into the jet needs to be corrected for. A correction is applied based on the *hybrid jet area* method which is a combination of two methods viz., the average offset method and the jet area method. The average offset method uses zero bias events to measure the average amount of energy added to the event due to pileup. The assumption is that averaging over zero bias events makes this measurement insensitive to high $\mathrm{p_T}$ objects and primarily represents soft pileup contributions. The average offset is measured in bins of η and number of pileup vertices (N_{PV}) averaged over ϕ. The correction is then given by $1 - \frac{<Offset(N_{PV},\eta)>}{p_T^{RAW}}$, where p_T^{RAW} is the uncorrected jet $\mathrm{p_T}$. The drawback of this method is that it assumes that every jet contains the same amount of pileup contribution. The jet area method, on the other hand , calculates corrections on a jet-by-jet basis. It calculates an energy density per event by clustering jets using the k_T algorithm (this has a value of parameter p=1 and favors clustering soft jets as opposed to hard ones) and dividing the $\mathrm{p_T}$ by jet area, which is defined as the region in $\eta - \phi$ occupied by soft particles clustered in the jet. The median of this distribution (ρ) for an event is expected to be insensitive to hard particles and thus ρA_j is a good approximation of pileup contribution to the i-th jet. The drawback of this second approach, however is that it doesn't take into account the fact that the detector response is η dependent. The *hybrid jet area* method combines these two methods to calculate a jet-by-jet correction depending on η and N_{PV}. Secondly, a MC calibration factor, which corrects the energy of reconstructed jets to match the generated MC particle jet energy on average, is applied. This factor is based on simulated events. Finally, two other factors are used that each calibrate the energy response of reconstructed jets to be uniform with respect to η and $\mathrm{p_T}$. These are also measured using simulated events. A QCD dijet sample is used to uniformize the dependence in η. Conservation of momentum in the transverse plane tells us that the sum of momentum in the transverse plane should be zero. Using jets that are approximately back-to-back in the azimuthal direction but at different η regions of the detector, the difference in response between these two η regions can be ascertained and corrected/uniformized. Using the same method of measuring residual response in the transverse direction in $\gamma + jets$ or $Z + jets$ events, the absolute jet energy scale as a function of $\mathrm{p_T}$ can be made uniform.

4.4.7 Missing transverse energy: $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$

The CMS detector is unable to detect neutrinos (and other hypothetical particles) that are weakly interacting. However, the momentum balance (or imbalance) in the plane transverse to beam direction can be used to infer their presence. This "missing" transverse momentum vector is referred to as the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ and its magnitude is referred to as $\mathrm{p_T^{miss}}$. It is defined as the negative vector sum of the $\mathrm{p_T}$ of entire list of objects reconstructed in the event by the above reconstruction algorithms and refined by particle-flow (PF objects):

$$\vec{p}_{\mathrm{T}}^{\mathrm{miss}} = -\Sigma \vec{p}_{\mathrm{T}} \tag{4.3}$$

The $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ plays an important role in this analysis as it helps gauge the momentum of the neutrinos from the decaying tau lepton. The $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction is directly dependent on the reconstruction of all the other objects in the event, from jets to muons to electrons. Consequently, it is sensitive to all the effects that influence the precise reconstruction and calibration of these objects. The largest effects come from biases in jet reconstruction and pileup (which are interconnected). Jet energy corrections described in previous section and pileup mitigation techniques discussed earlier can help significantly reduce the bias in $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction [103].

4.4.8 Relative isolation

The isolation of an object is the measure of the absence of other objects in its vicinity. In other words, it is the measure of how "isolated" an object is. It is calculated by summing up the p_T of all objects in a cone with predefined radius $\Delta R = 0.4$ around the lepton. The relative isolation which is obtained by dividing the isolation by the p_T of the lepton is then given by:

$$I^{\ell}_{\mathrm{rel}} = \left(\sum p_{\mathrm{T}}^{\mathrm{charged}} + \max \left[0, \sum p_{\mathrm{T}}^{\mathrm{neutral}} + \sum p_{\mathrm{T}}^{\gamma} - p_{\mathrm{T}}^{\mathrm{PU}} (\ell) \right] \right) / p_{\mathrm{T}}^{\ell}, \tag{4.4}$$

where $p_{\mathrm{T}}^{\mathrm{charged}}$, $p_{\mathrm{T}}^{\mathrm{neutral}}$, and p_{T}^{γ} indicate the p_T of a charged particle, a neutral particle, and a photon within the cone, respectively. The contribution to isolation sum from charged particles in the cone coming from pileup is excluded by requiring the track origins be consistent with the primary vertex associated with the hard interaction. However, the same procedure cannot be adopted for neutral particles in the cone coming from pileup. This contribution, $p_{\mathrm{T}}^{\mathrm{PU}} (\ell)$, is estimated using a jet area method for electrons [104, 105], and simply as half of the scalar p_T sum of charged particles from pileup inside the cone for muons. The definition ensures that the total contribution to the isolation cone sum for neutral particles is always greater than or equal to 0.

The relative isolation is an useful variable in this analysis as prompt objects are usually isolated. More importantly, for an analyses with electrons and muons in the final state such as this, a tightened isolation requirement helps to reject those events where a jet is misidentified as either one of these leptons. Strict isolation requirements are thus used in the event selection in this analysis as described in sections 5.2.2 and 5.3.2.

4.4.9 Collinear mass: M_{col} and transverse mass: M_T

An important variable of interest in this analysis is the collinear mass, M_{col}. As mentioned in later chapters, M_{col} is used as the signal variable in the $\mathrm{H} \to \mu\tau_e$ analysis and in the M_{col} fit method of the $\mathrm{h} \to \mu\tau_e$ analysis. The visible mass of the muon-electron system, M_{vis} is not a very good estimator of the Higgs boson mass. This is because the neutrinos from the tau decay do not interact with the detector and the energy they carry is "lost". M_{col} provides a better estimate of the Higgs mass by approximating the neutrino momenta using the collinear approximation [106]. The primary idea is the following. The mass of the Higgs being much larger than that of the tau lepton causes the tau lepton to become highly boosted. Consequently, the decay products of the tau lepton, i.e. the electron, the tau neutrino and the electron neutrino, are produced in a highly collimated region around the direction of the tau lepton momentum. The momenta of the neutrinos ($\mathrm{p}_{\mathrm{T}}^{\nu,\,\mathrm{est}}$) can thus be approximated from the projection of $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ in the direction of momenta of the visible tau decay product (electron), i.e. $\mathrm{p}_{\mathrm{T}}^{\vec{\mathrm{e}}}$. The visible fraction of the tau lepton momentum is then given as $x_{\tau}^{\mathrm{vis}} = \mathrm{p}_{\mathrm{T}}^{\vec{\tau}^{\mathrm{vis}}}/(\mathrm{p}_{\mathrm{T}}^{\vec{\tau}^{\mathrm{vis}}} + \mathrm{p}_{\mathrm{T}}^{\nu,\,\mathrm{est}})$. Finally, M_{col} is given as $M_{\text{col}} = M_{\text{vis}}/\sqrt{x_{\tau}^{\mathrm{vis}}}$. A very simple illustration in Figure 4.4 shows the superimposition the M_{col} and M_{vis} spectrums for a 300 GeV Higgs boson. Evidently, M_{col} is a better estimator of the mass and also has a sharper peak.

Another variable that is used in this analyses and is useful to discriminate signal from background (see chapter 5) is the transverse mass, $M_{\mathrm{T}}(\ell)$ ($\ell = \mu, \mathrm{e}$). It is defined as follows: $M_{\mathrm{T}}(\ell) = \sqrt{2|\vec{p}_{\mathrm{T}}^{\ell}||\vec{p}_{\mathrm{T}}^{\mathrm{miss}}|(1 - \cos\Delta\phi_{\ell-\mathrm{p}_{\mathrm{T}}^{\mathrm{miss}}})}$, where $\Delta\phi_{\ell-\mathrm{p}_{\mathrm{T}}^{\mathrm{miss}}}$ is the angle between the lepton transverse momentum and $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$.

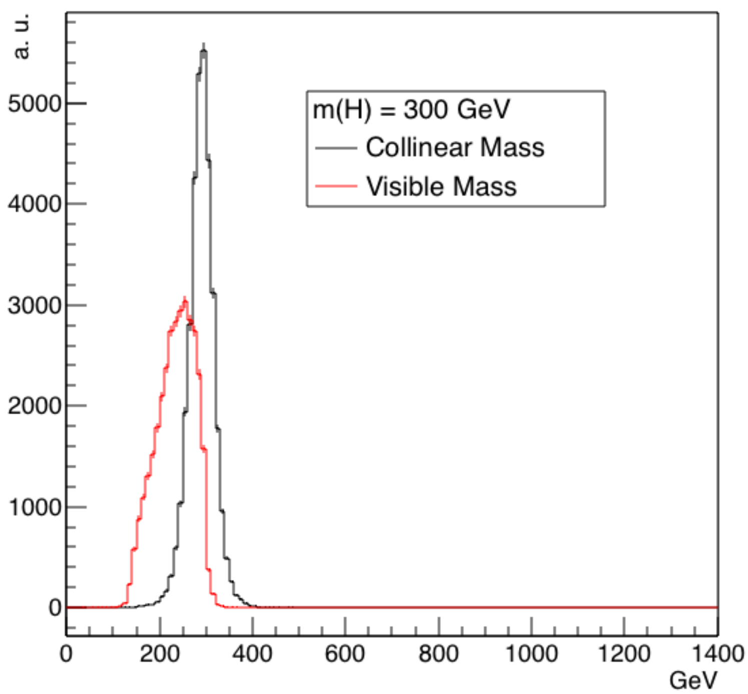


Figure 4.4: M_{col} and M_{vis} distributions for Higgs mass of 300 GeV.

CHAPTER 5

EVENT SELECTION

> It is, indeed an incredible fact that what the human mind, at its deepest and most profound, perceives as beautiful finds its realization in external nature. What is intelligible is also beautiful.
>
> ---
>
> *Subrahmanyan Chandrasekhar*

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 that 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 $\rightarrow \mu\tau_e$ analysis, two methods of selection were developed. The first method developed involves placing requirements on several kinematic variables, and then using the resulting distribution of M_{col} as discriminant for a binned likelihood fit (see section 7.2 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 greater sensitivity, as discussed later in the chapter. However, the M_{col} fit method is also presented as a complementary method and acts like a cross-check for the BDT method. For H $\rightarrow \mu\tau_e$ analysis, only the $M_{\rm col}$ fit method is developed. This is in part due to the difficulties 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 [107] 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. It is important to note here that the signal region is known for the h $\rightarrow \mu\tau_e$ analysis. However, for the H $\rightarrow \mu\tau_e$ analysis, the signal region is presupposed - as there is no current evidence for H or its mass.

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

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

The signature of the h $\rightarrow \mu\tau_e$ 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 $\rightarrow \tau_\mu\tau_e$ decay that is allowed by the SM and since been observed [108], but with significant kinematic differences. In h $\rightarrow \mu\tau_e$ 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 $\rightarrow \mu\tau_e$, 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 $\vec{p}_T^{\rm miss}$ in the azimuthal plane. The same is not true

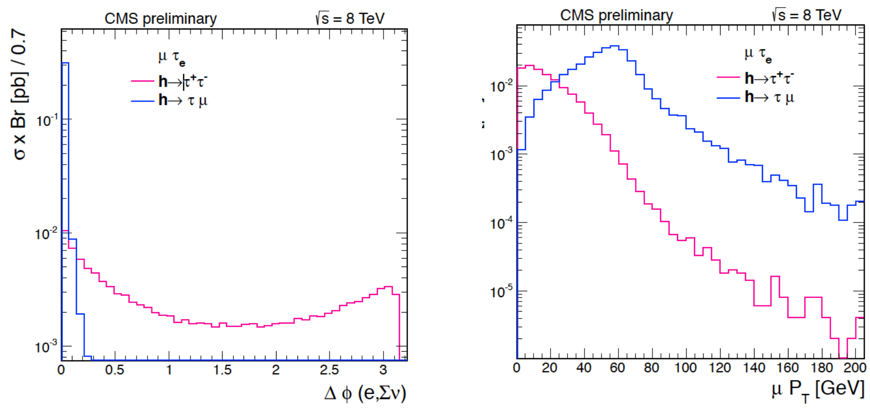


Figure 5.1: Illustration of the differences in $\mathrm{p}_{\mathrm{T}}^{\mu}$ and $\Delta\phi(\mathrm{e}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$ spectrums in $\mathrm{h} \to \mu\tau_e$ and $\mathrm{h} \to \tau_\mu\tau_\mathrm{e}$ processes.

in the $\mathrm{h} \to \tau_\mu\tau_\mathrm{e}$ decays. These differences are illustrated pictorially in Fig. 5.1.

The most dominant backgrounds consists of $\mathrm{Z} \to \tau\tau$ events coming from Drell-Yan production and $t\bar{t}$ production. In $\mathrm{Z} \to \tau\tau$ 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 ($\mathrm{H} \to \tau\tau, \mathrm{WW}$), $\mathrm{W}\gamma^{(*)} + \mathrm{jets}$,single top production, $\mathrm{W} + \mathrm{jets}$ 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.

5.2.2 $\mathrm{h} \to \mu\tau_e$: 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 a 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 4.4.3 and 4.4.4. Isolation criterion, as measured by I_{rel} (described in 4.4.5), are required to have values $I^{\text{e}}_{\text{rel}} < 0.15$ and $I^{\mu}_{\text{rel}} < 0.1$. 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 $p^{\mu}_{\text{T}} > 26\,\text{GeV}$ and $|\eta^{\mu}| < 2.4$.The e is required to have $p^{\text{e}}_{\text{T}} > 10\,\text{GeV}$ and $|\eta^{\text{e}}| < 2.3$. Only events with two or fewer jets are considered. All jets considered must have $p_{\text{T}} > 30\,\text{GeV}$, $|\eta| < 2.4$ and satisfy the loose identification criterion described in section 4.4.6. 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 as $t\bar{t}$ and single top. Also, as described in 4.4.6, any event with one or more jets within $\Delta R < 0.4$ of either lepton candidates 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 5.1. 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 3.3.7. 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 of signal and backgrounds distributions for different variables.

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 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

TABLE 5.1

BASELINE SELECTION CRITERIA FOR h $\rightarrow \mu\tau_e$ ANALYSIS

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

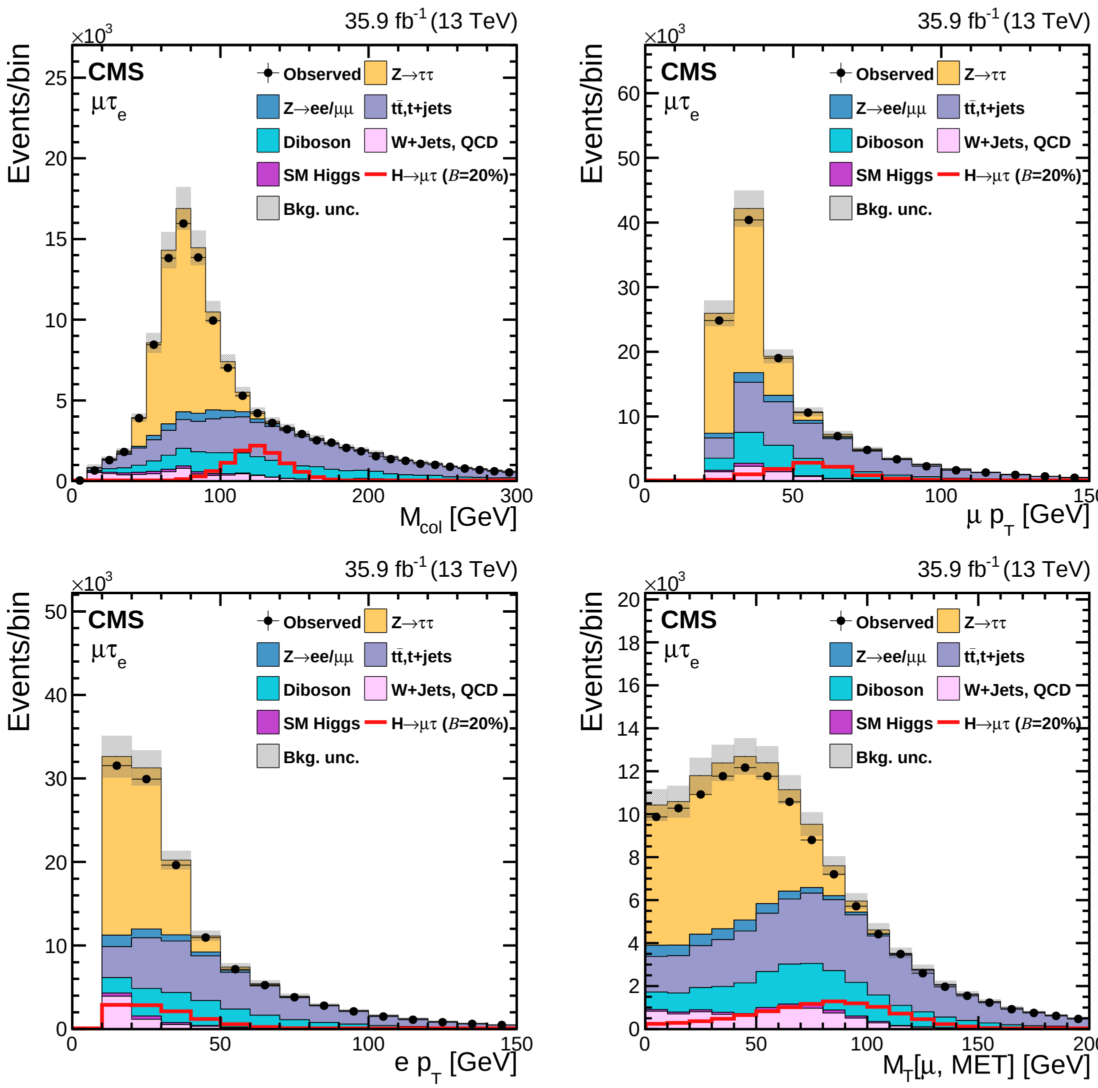


Figure 5.2: Distributions of kinematic variables after baseline selection for $h \rightarrow \mu\tau_e$ analysis (1).

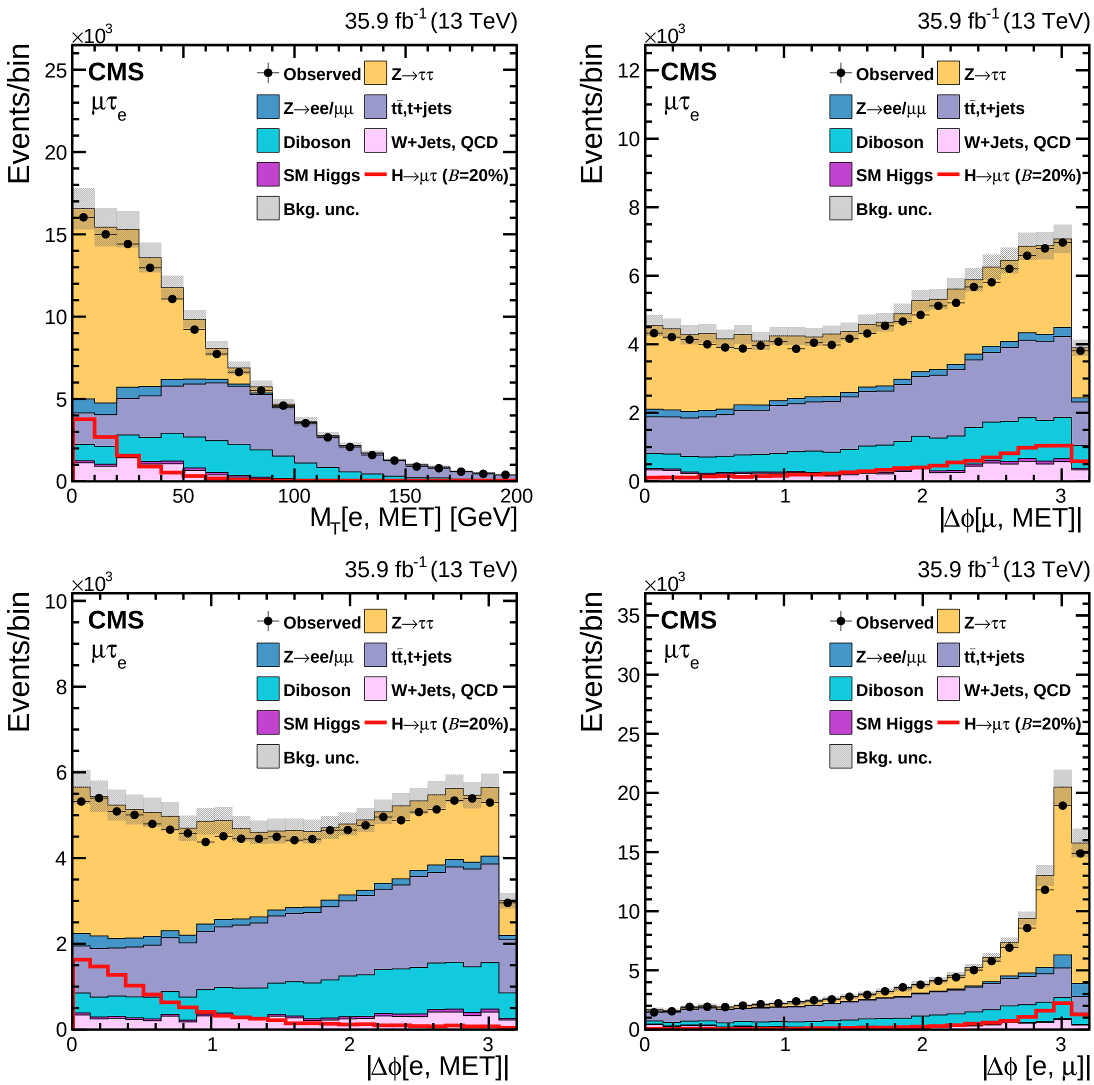


Figure 5.3: Distributions of kinematic variables after baseline selection for $h \to \mu\tau_e$ analysis (2).

than events with no jets. Each of this categories enhance 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 $M_{jj} < 550\,\mathrm{GeV}$. 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 $M_{jj} \geq 550\,\mathrm{GeV}$. The dominant contribution comes from VBF production which is characterized by presence of two jets with high dijet mass.

5.2.3 h $\rightarrow \mu\tau_e$: 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 $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$, between μ and $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$, denoted respectively by $\Delta\phi(\mathrm{e},\mu)$, $\Delta\phi(\mathrm{e},\vec{p}_{\mathrm{T}}^{\mathrm{miss}})$, $\Delta\phi(\mu,\vec{p}_{\mathrm{T}}^{\mathrm{miss}})$, and the transverse mass between μ and $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$, between e and $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$, denoted respectively by $M_T(\mu)$ and $M_T(\mathrm{e})$. The h $\rightarrow \mu\tau_e$ 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 4.4.9) 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 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 paragraph. 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 $\mathrm{p_T}$ of the μ is made marginally stricter by requiring $\mathrm{p}_{\mathrm{T}}^{\mu} > 30\,\mathrm{GeV}$. All other lepton $\mathrm{p_T}$ 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 along-with, that the entire spectrum of distribution of the discriminant variable (that is used in the final max-likelihood fit to extract results) is well-populated, especially 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 extracting the expected limit is given in section 7.2.3. To do the optimization of the kinematic thresholds, we start by requiring the baseline selection. Then for a variable in consideration, e.g.- $\Delta\phi(\mathrm{e}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$, 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 successive rounds. This optimization was done separately for each of the four categories. The final set of thresholds arrived at in this way for the $\mathrm{h} \to \mu\tau_e$ M_{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 $\to \mu\tau_e$ $M_{\rm col}$ FIT ANALYSIS

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

5.2.4 h $\to \mu\tau_e$: BDT method selection

In the BDT method, a boosted decision trees (BDT) classifier is used to discriminate signal events from background events. A decision tree is a 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 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

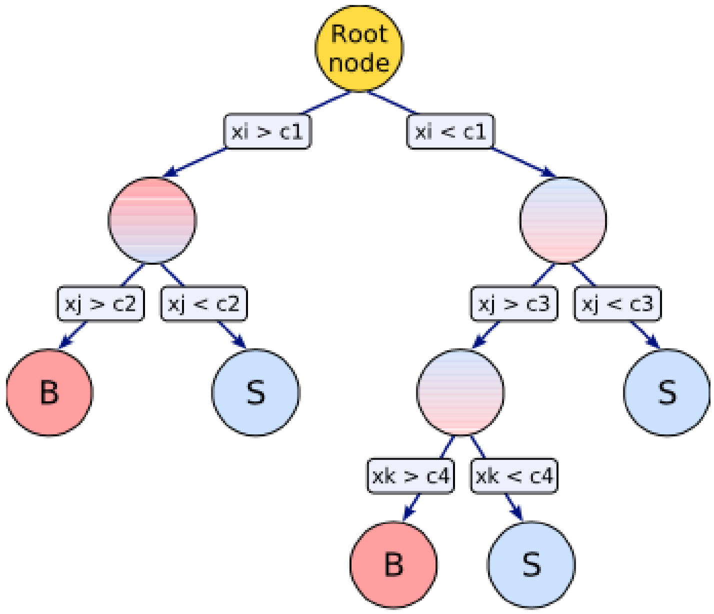


Figure 5.4: Illustration of decision tree [109].

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.

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 backgrounds. A suite of input variables is used in training of 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 $\mathrm{p_T}$: $\mathrm{p_T^{\mu}}$.
- Electron $\mathrm{p_T}$: $\mathrm{p_T^{e}}$.

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

The training was done with a 850 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 boosting) method was used for boosting. A training to testing split of 70:30 split was used. Fig. **??** 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.

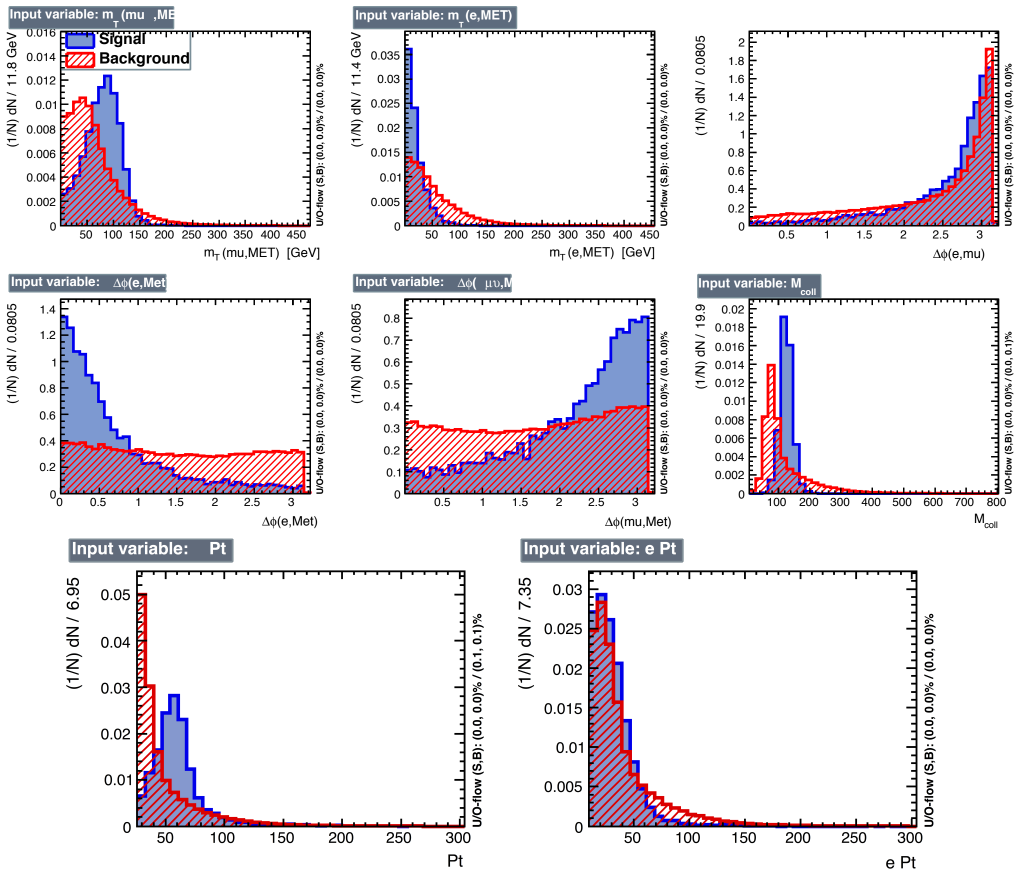


Figure 5.5: Normalized distributions of the input variables for BDT method. The signal (blue) is composed of a weighted mixture of GGF and VBF events, whereas the background (red) is made of $t\bar{t}$ and Drell-Yan events. All events were required to satisfy the baseline selection criteria.

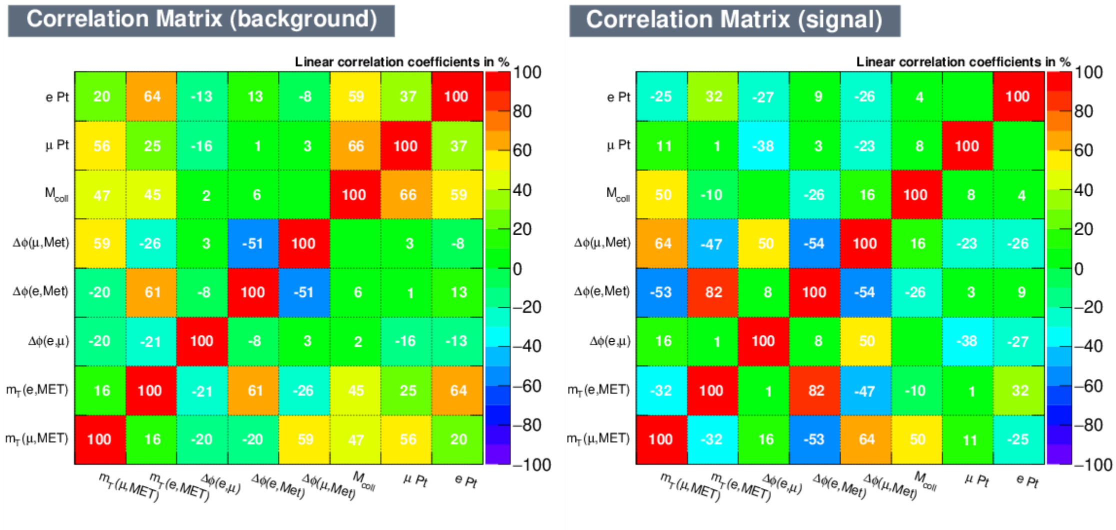


Figure 5.6: Correlations between input variables for signal events (right) and background events (left).

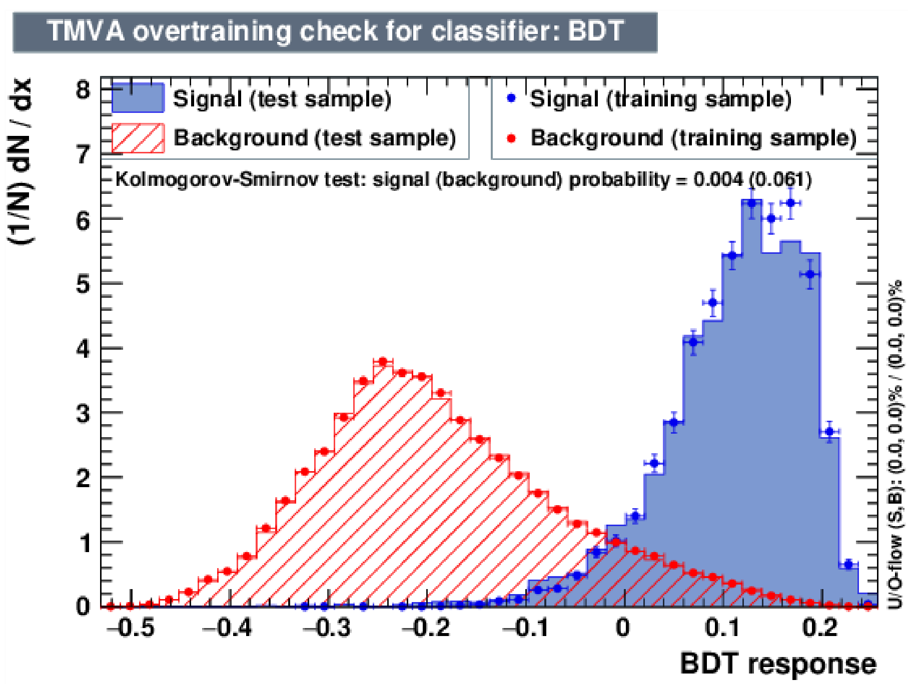


Figure 5.7: Distribution of BDT response for training (dots) and test(fill) distributions for both signal(blue) and background(red) events.

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

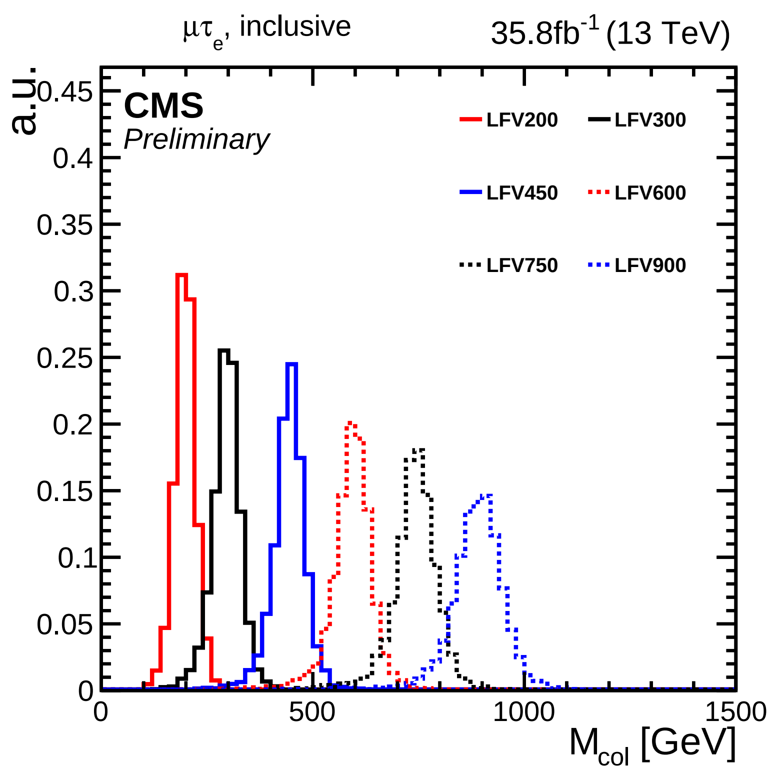


Figure 5.8: Illustration of simulated signal mass peaks for H $\rightarrow \mu\tau_e$ analysis for different H boson masses.

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 form a very small background as the Z $\to \tau\tau$ spectrum peaks at much lower values (around Z boson mass) of collinear mass than the signal events coming from heavy H boson decays. The other backgrounds come from h boson decays ($h \to \tau\tau$, WW), W$\gamma^{(*)}$ + jets ,single top production, W + jets events, $Z \to \ell\ell$ ($\ell =$ e, μ) + jets and QCD multijet backgrounds. These backgrounds are described in more detail, along with their 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 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 isolation criteria have been described in sections 4.4.3, 4.4.4 and 4.4.8. 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 3.3.7. The μ is required to have $p_T^{\mu} > 53$ 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 zero or one jet are considered. Jets must have $p_T > 30$ GeV, $|\eta| < 2.4$ and satisfy the loose identification criterion described in section 4.4.6 to be considered. As only GGF production mode is considered for the H $\to \mu\tau_e$ analysis, events with more than one jet make negligent contribution and are rejected. All other other criteria are same as the h $\to \mu\tau_e$ analysis. The entire set of baseline selection criteria for H $\to \mu\tau_e$ has been summarized in table 5.3.

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

TABLE 5.3

BASELINE SELECTION CRITERIA FOR H $\rightarrow \mu\tau_e$ 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 $p_T > 30\,\text{GeV}$		

- **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).

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

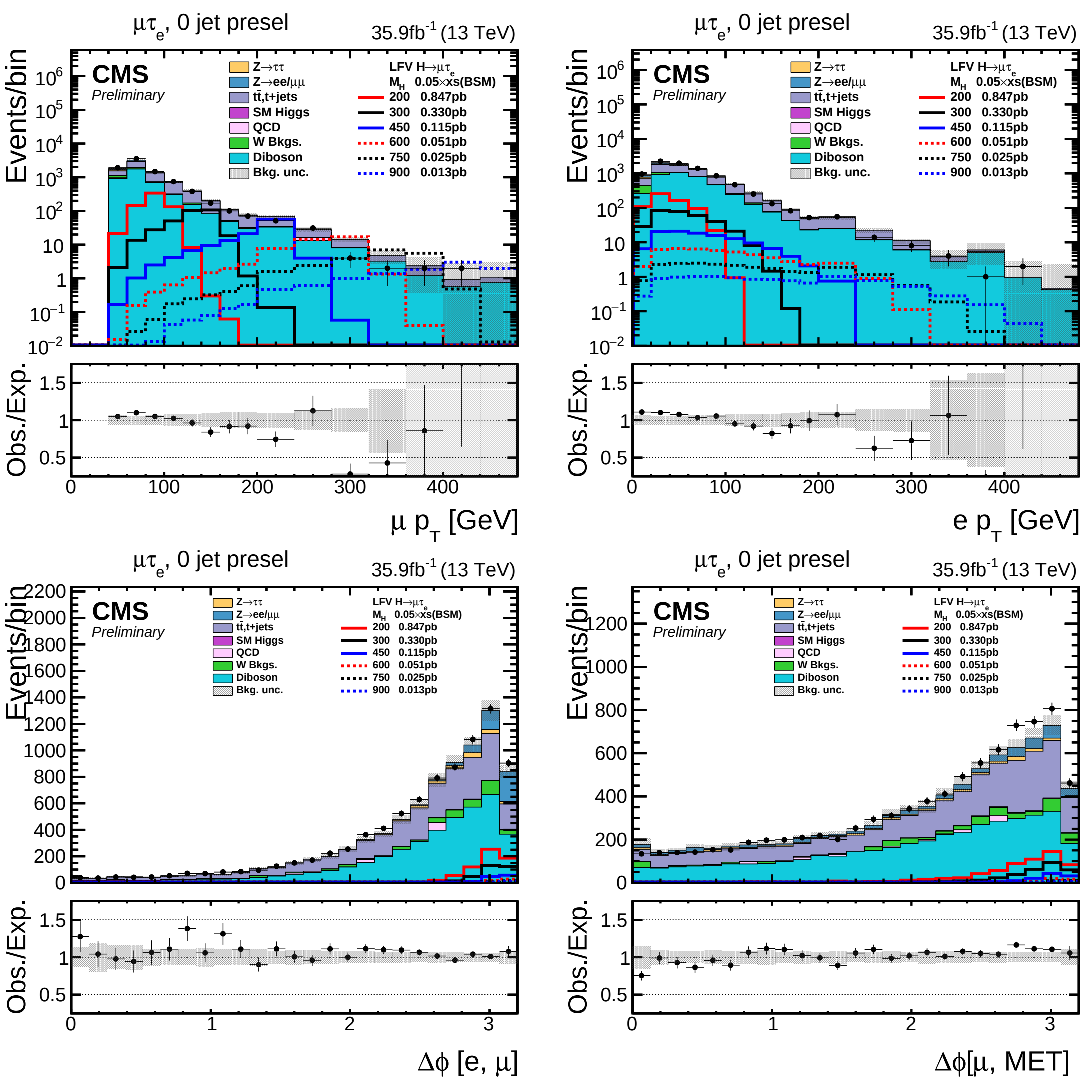


Figure 5.9: Distributions of kinematic variables after baseline selection for 0-jet category of H $\rightarrow \mu\tau_e$ analysis.

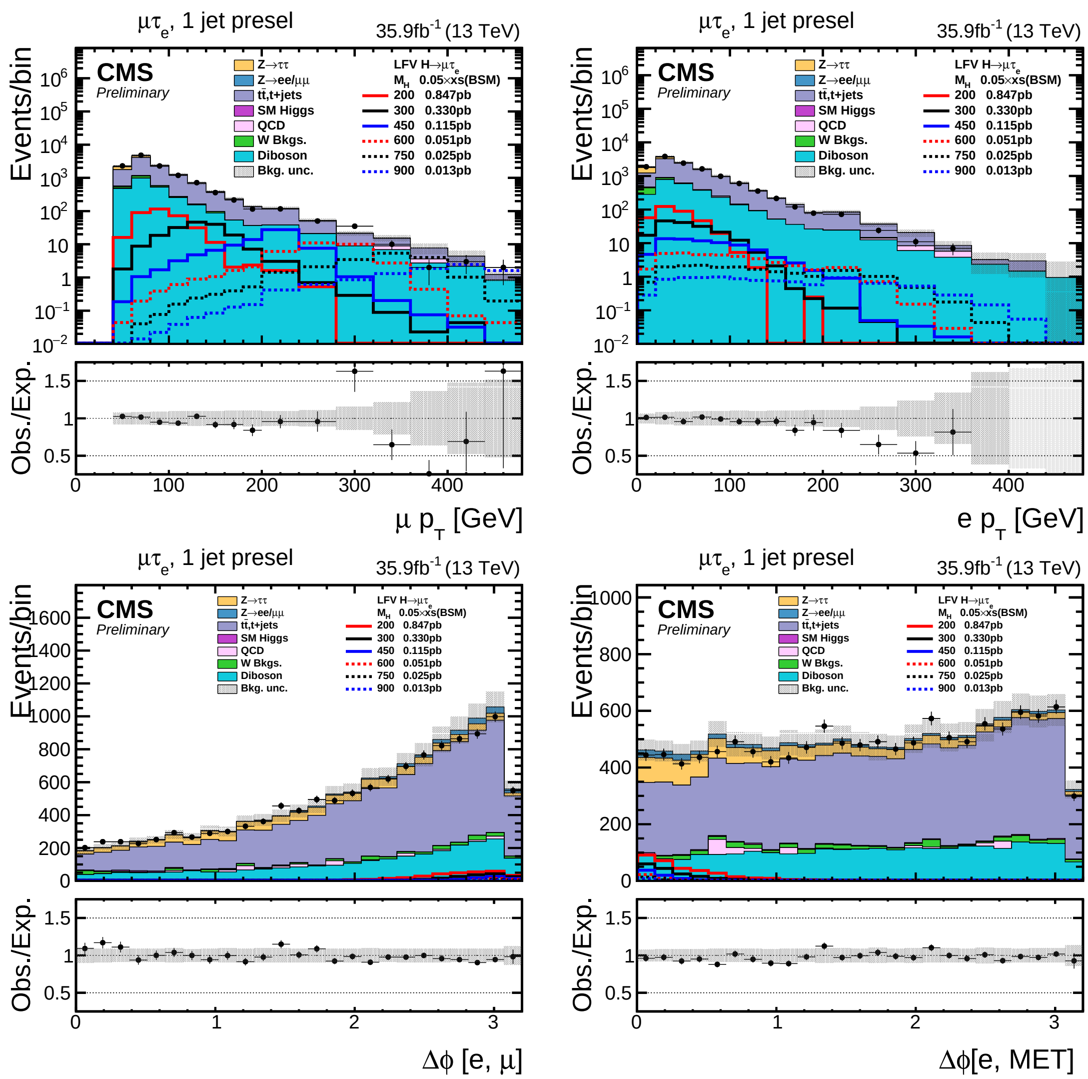
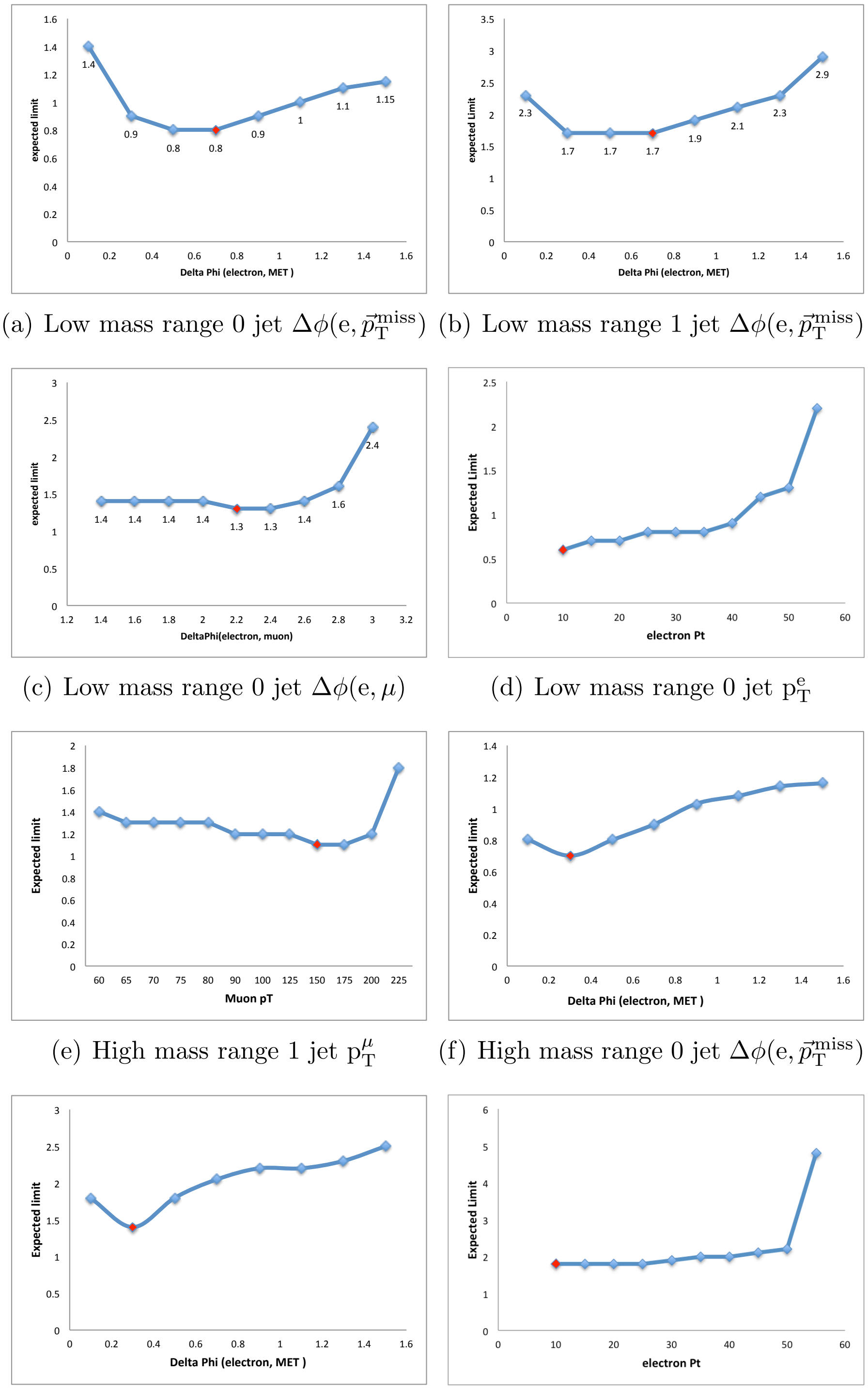


Figure 5.10: Distributions of kinematic variables after baseline selection for 1-jet category of $\mathrm{H} \to \mu\tau_e$ analysis.

5.3.3 $\mathrm{H} \to \mu\tau_e$: M_{col} fit selection

Just like the M_{col} fit method in $\mathrm{h} \to \mu\tau_e$, the selection is performed by placing kinematic cuts on several variables to enhance the signal-to-background ratio. The variables considered are: $\Delta\phi(\mathrm{e}, \mu)$, $\Delta\phi(\mathrm{e}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$, $\Delta\phi(\mu, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$, $M_T(\mu)$ and $M_T(\mathrm{e})$. In addition, the $\mathrm{p_T}$ 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 $\rightarrow \mu\tau_e$, and not in a particular region like the h $\rightarrow \mu\tau_e$ analysis, the potential effect of background mimicking the signal, in particular due to higher $\mathrm{p_T}$ thresholds of the leptons, is not apparent. The motivations for using these variables remain much the same like the h $\rightarrow \mu\tau_e$ 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 $\mathrm{p_T}$ spectrum of the prompt lepton (μ) is hard.

The procedure for optimization of the thresholds of for these variables is exactly the same as described in section 5.2.3. Further to get better sensitivity in the entire mass range from 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 in 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 manner, for both mass ranges and both categories of the H $\rightarrow \mu\tau_e$ M_{col} fit analysis, are listed in Table. 5.4. The M_{col} distributions after requiring these selections is used in a max-likelihood fit to extract results, as discussed in section 7.



(a) Low mass range 0 jet $\Delta\phi(\mathrm{e}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$ (b) Low mass range 1 jet $\Delta\phi(\mathrm{e}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$

(c) Low mass range 0 jet $\Delta\phi(\mathrm{e}, \mu)$ (d) Low mass range 0 jet $\mathrm{p}_{\mathrm{T}}^{\mathrm{e}}$

(e) High mass range 1 jet $\mathrm{p}_{\mathrm{T}}^{\mu}$ (f) High mass range 0 jet $\Delta\phi(\mathrm{e}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$

(g) High mass range 1 jet $\Delta\phi(\mathrm{e}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$ (h) High mass range 1 jet $\mathrm{p}_{\mathrm{T}}^{\mathrm{e}}$

Figure 5.11. Examples of cut optimization for the $\mathrm{H} \rightarrow \mu\tau_e$ analysis.

TABLE 5.4

FINAL SELECTION CRITERIA IN EACH CATEGORY OF THE

$H \rightarrow \mu\tau_e$ ANALYSIS

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

CHAPTER 6

BACKGROUND ESTIMATION AND VALIDATION

Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.

Marie Sklodowska Curie

6.1 Introduction

This chapter describes the techniques used for estimation of the backgrounds in the analyses. Each background is estimated individually. For large backgrounds, the estimation is validated using regions enriched in those backgrounds.

6.2 h125: h $\rightarrow \mu\tau_e$ backgrounds

6.2.1 Z $\rightarrow \tau\tau$ background

The Z $\rightarrow \tau\tau$ background is the dominant background in 0-jet and 1-jet categories of the analysis. It is an irreducible background and arises when one τ coming from the Z boson decay further decays into a μ and the other decays into a e. This background is estimated from simulated monte-carlo events. In Z $\rightarrow \ell\ell$ events from Drell-Yan production including Z $\rightarrow \tau\tau$, the $m_{\ell\ell}$ and $\mathrm{Z}_{\mathrm{p_T}}$ distributions are found to be different in data and simulation. In order to correct for this, a set of reweighting factors is calculated using a dedicated control region enriched in Z $\rightarrow \mu\mu$ events. The set of reweighting factors are applied as a function of generator-level $m_{\ell\ell}$ and $\mathrm{Z}_{\mathrm{p_T}}$ in the signal region of the analysis. A more detailed study of this effect and calculation of the reweighting factors can be found in the following references [108].

To validate this estimation, we look at agreement between observed data and simulation in a region enriched in Z $\rightarrow \tau\tau$ events. This region is constructed by requiring, in addition to the baseline selection, the $\mathrm{p_T}$ of the $\mu < 40$ GeV. The $\mathrm{p_T}$ in Z $\rightarrow \tau\tau$ events is on softer side of the spectrum compared to other backgrounds which are more spread out, as seen in Fig. 5.2 (top right). The $M_T(\mu)$, as seen in Fig. 5.2 (bottom right), is required to be less than 60 GeV following similar reasoning. Further the invariant mass of the e and μ is required to be in between 30 GeV and 70 GeV in order to isolate the Z peak. The distributions of BDT response and M_{col} in this Z $\rightarrow \tau\tau$ enriched region are shown in Fig. 6.1, for the categories where this background is dominant. The plots show good agreement between data and background.

6.2.2 $t\bar{t}$ background

Tops decay into W bosons and a b-quark more than 90% of the time. The W boson can decay leptonically into a μ and e making it a background for the analysis. The b-tagging veto applied at the baseline selection level is able to somewhat suppress this background. However it still forms a large fraction of the background for the analysis. In fact, it is the largest background in both 2-jet categories. It is also large in the 1-jet category. We estimate the $t\bar{t}$ background using simulation. The background estimation is validated in two separate control regions enriched in $t\bar{t}$. The first control region is formed requiring the baseline selection but with a inverted b-tagging veto. In other words, at least 1 b-tagged jet is required to be present in the event. The distributions of BDT response (top) and M_{col} (bottom) in this region are shown in Fig. 6.6 for categories where the $t\bar{t}$ background is large. The second control region is constructed using kinematic selection criteria. In particular, in addition to the baseline selection criteria with the b-tag veto removed, we require M_T(e) (see

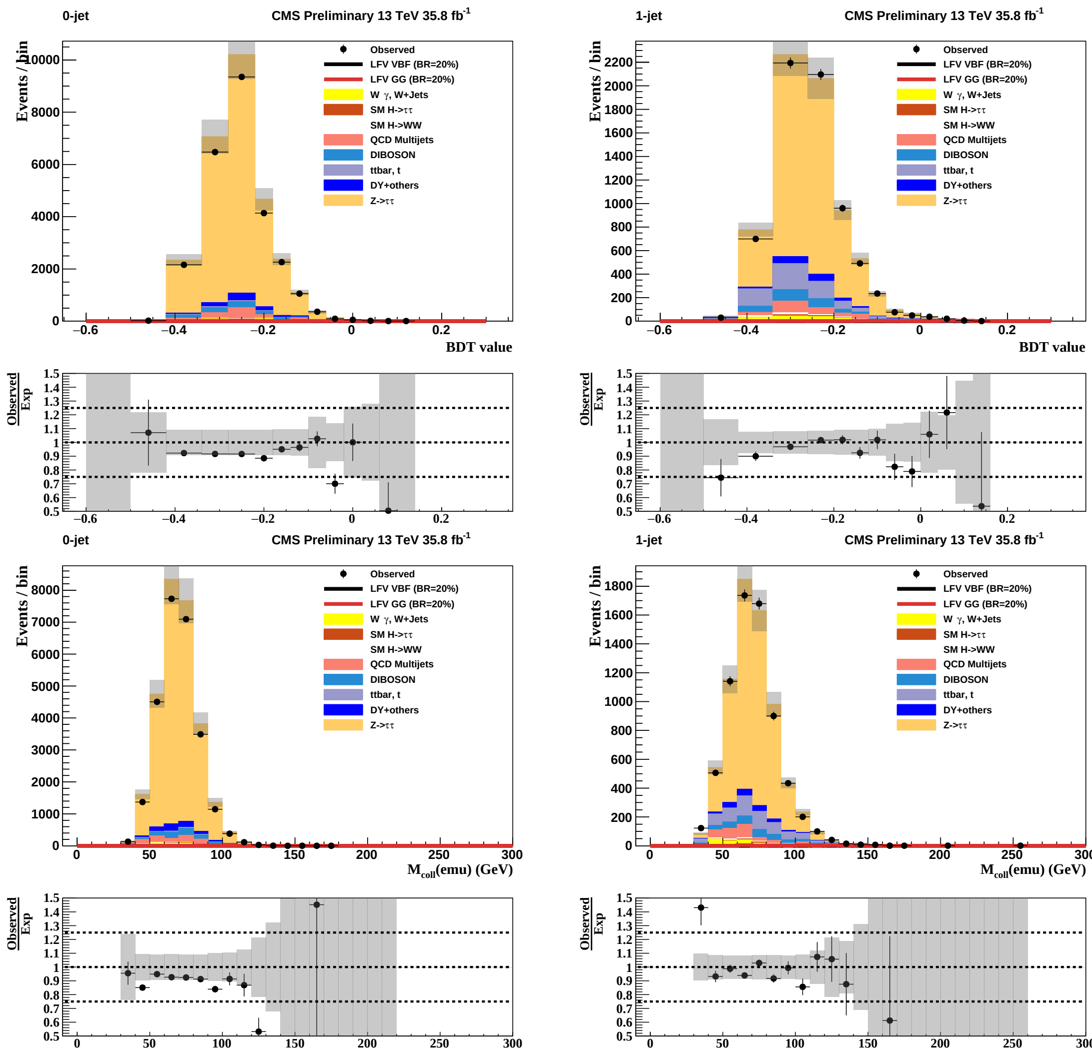


Figure 6.1: Distributions of BDT response (top) an $M_{\rm col}$ (bottom) in $\mathrm{Z} \to \tau\tau$ enriched region for 0-jet (left) and 1-jet (right) categories.

Fig. 5.3 top left) to be greater than 50 GeV. The distributions of BDT response (top) and M_{col} (bottom) in this second control region are shown in Fig. 6.3. Given that the uncertainty bands in these control region plots only contain uncertainties on normalization (and not shape-based uncertainties, as discussed in section 7, and included in the max likelihood fit used to extract results), the data over background estimation ratio is reasonable in these regions. Further, a normalization uncertainty of 10% is applied on the $t\bar{t}$ estimation in the signal region based on these control regions.

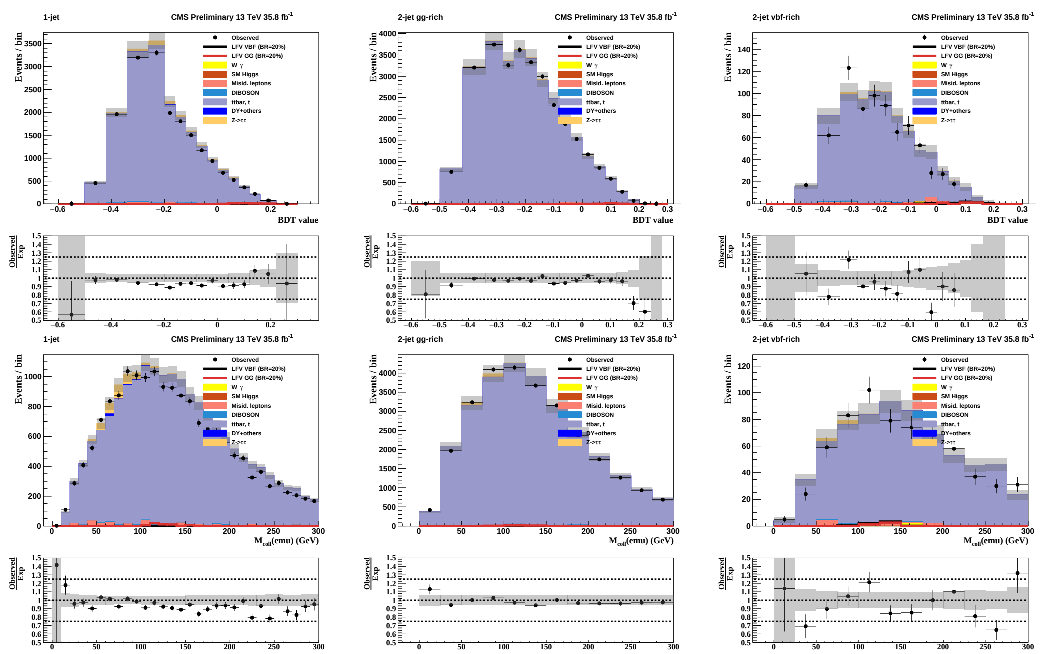


Figure 6.2. Distributions of BDT response (top) an M_{col} (bottom) in the first $t\bar{t}$ enriched region, as described in the text.

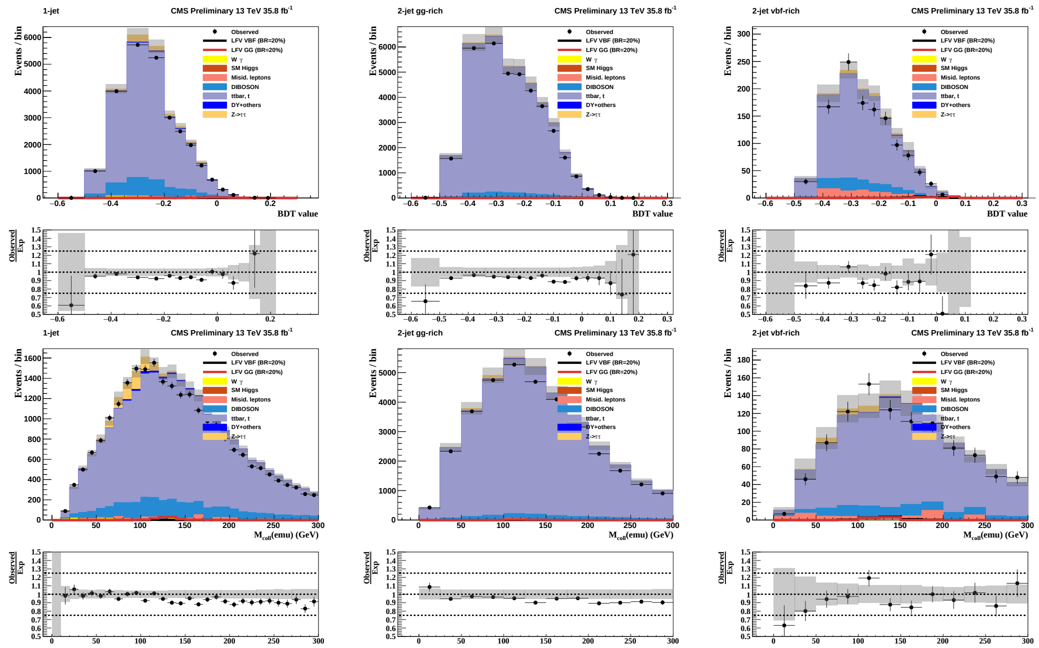


Figure 6.3. Distributions of BDT response (top) an M_{col} (bottom) in the second $t\bar{t}$ enriched region, as described in the text.

6.2.3 Misidentified lepton background

Another source of background which is relatively much smaller than $t\bar{t}$ or $\mathrm{Z} \rightarrow \tau\tau$ arises from jets misidentified as leptons in W+jets or SM events comprised uniquely of jets produced through the strong interaction, referred to as quantum chromodynamics (QCD) multijet events. In W+jets events, one lepton candidate is a real lepton from the W boson decay while the other lepton is a misidentified jet. In QCD events, both leptons in the final state are misidentified jets. The baseline selection criteria requires the leptons to be well identified and isolated. This makes it difficult for a jet to masquerade as a lepton. In case of the μ, this is even more so since it is required to satisfy high $\mathrm{p_T}$ thresholds as well. Consequently, these events form a small part of the background. This is in contrast to a final state where the non-prompt lepton is a hadronically decaying τ instead of an electronically decaying one. This background would be much larger in such a case.

The W + jets background contribution to the misidentified-lepton background is estimated using simulation. The QCD multijet contribution is estimated from collision data events where the leptons have like-sign charge. The expected yield from non-QCD processes in this region is subtracted using simulation. The resulting sample is then rescaled to account for the differences between the composition in the like- and opposite-sign charge regions. The scaling factors are extracted from samples enriched in QCD multijet events, and the procedure is illustrated in Ref. [108]. This background is validated in a control region that is obtained by requiring the baseline selection but inverting the isolation criteria. In other words events with well-isolated μ and e are rejected. The particular isolation thresholds required for this region are:$0.1 < I_{\mathrm{rel}}^{\mathrm{e}} < 1$ or $0.15 < I_{\mathrm{rel}}^{\mu} < 0.25$. The distributions of BDT response and M_{col} in this qcd enriched region are shown in Fig. 6.4. The plots show good agreement between data and background.

6.2.4 Other backgrounds

The other backgrounds in the analysis make relatively much smaller contributions. Electroweak diboson production (WW, WZ and ZZ) contributes a similar number of events as the misidentified lepton background, and is estimated from simulation. WW events make the largest contribution, followed by WZ and ZZ events. This is because WZ and ZZ events have additional leptons in their final state which have to miss detection in order for the event to be a background. SM decays of the h boson also forms a small but non-negligible background. These come particularly from h $\rightarrow \tau\tau$ and h $\rightarrow$ WW decays. Other backgrounds include $Z \rightarrow \ell\ell$ ($\ell = \mathrm{e}, \mu$)+jets, single-top quark production and W$\gamma^{(*)}$ + jets. All of these are estimated using simulation.

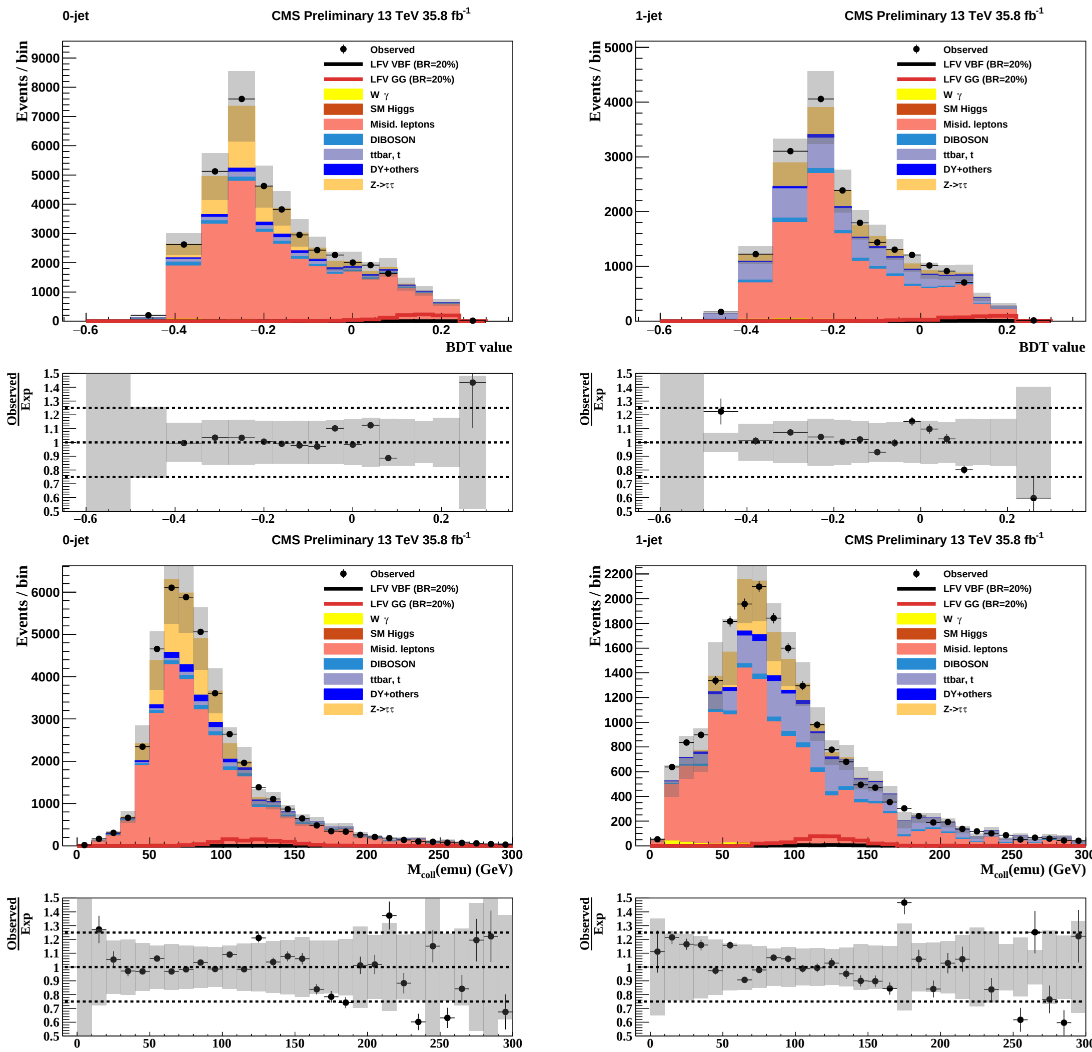


Figure 6.4: Distributions of BDT response (top) an M_{col} (bottom) in QCD enriched region for 0-jet (left) and 1-jet (right) categories.

6.3 Heavy Higgs: H $\rightarrow \mu\tau_e$ backgrounds

The background processes in the H $\rightarrow \mu\tau_e$ analysis are similar to h $\rightarrow \mu\tau_e$ but differ in relative contribution, and are overall much smaller. This is due to the fact that the H $\rightarrow \mu\tau_e$ analyses searches for LFV decay in a higher mass, higher p_T region. In particular, Z $\rightarrow \tau\tau$ background which is the most dominant in h $\rightarrow \mu\tau_e$ is now very small. The Z $\rightarrow \tau\tau$ background peaks around the Z boson mass, and the high p_T cuts in this analysis reject most of these events. The dominant backgrounds in H $\rightarrow \mu\tau_e$ are $t\bar{t}$ production, followed by electroweak diboson production which have a relatively flatter p_T distribution and survive the strict p_T requirements.

$t\bar{t}$ production is the largest background in the H $\rightarrow \mu\tau_e$ analysis. We estimate this background using simulation. A control region enriched in $t\bar{t}$ events is constructed by requiring the baseline selection with the b-tag veto removed, and with the additional requirement that at least 1 b-tagged jet be present. Fig. **??** (left) shows the M_{col} distribution of this sample. To take into account the residual data to background estimation difference, an overall normalization scale factor of 0.886 is extracted from this region, and is applied to the background estimation in the signal region. The same control region above is shown in Fig. **??** (right), after the background has been scaled by the above factor for illustration. Distributions of several other kinematic variables (after the above rescaling) in the $t\bar{t}$ control region are shown in Fig. 6.6. They show reasonable agreement between data and estimated background.

Electroweak diboson production (WW, WZ and ZZ) forms the next largest background in H $\rightarrow \mu\tau_e$ analysis. It is estimated using simulation. All other backgrounds are much smaller. This can be seen from the distributions of kinematic variables after baseline selection, as can be seen from Figs. 5.9 and 5.10. The misidentified lepton background is even smaller here than h $\rightarrow \mu\tau_e$. The higher p_T requirement makes it even less likely for jets to be able to be misidentified as leptons. This background is estimated using the same technique as h $\rightarrow \mu\tau_e$, as described in section 6.2.3. The

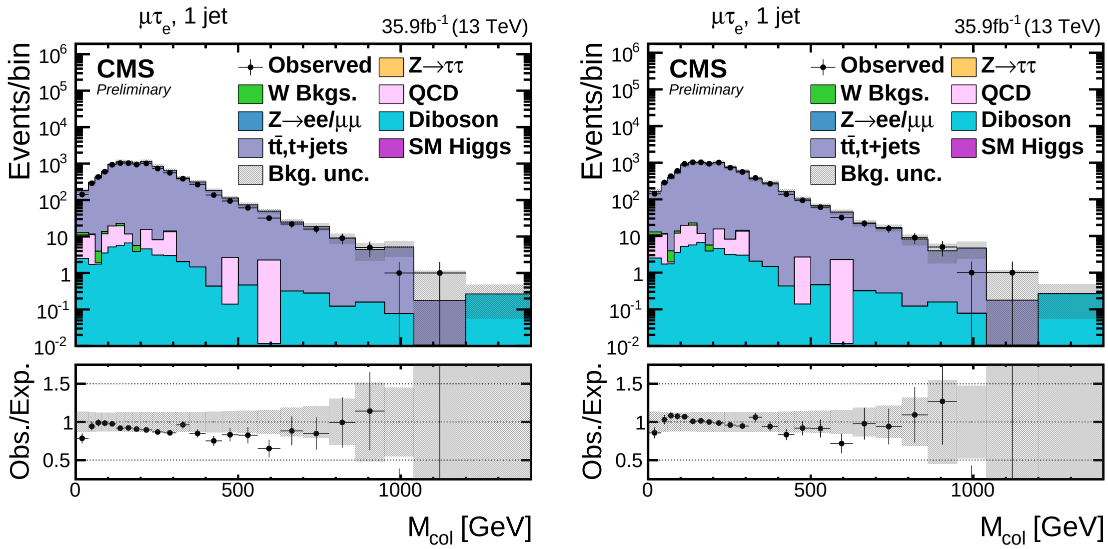


Figure 6.5: $M_{\rm col}$ distribution in $t\bar{t}$ enriched control region as defined in the text before the application of the scale factor (left) and after (right), for the $\mathrm{H} \rightarrow \mu\tau_e$ analysis.

$Z \rightarrow \ell\ell$ ($\ell = \mathrm{e}, \mu$) + jets and $\mathrm{Z} \rightarrow \tau\tau$ backgrounds are estimated from simulation. Other backgrounds include SM h boson decays, $\mathrm{h} \rightarrow \mathrm{WW}$, $\mathrm{h} \rightarrow \tau\tau$, single-top quark production and $\mathrm{W}\gamma^{(*)}$ + jets, and are also estimated using simulation.

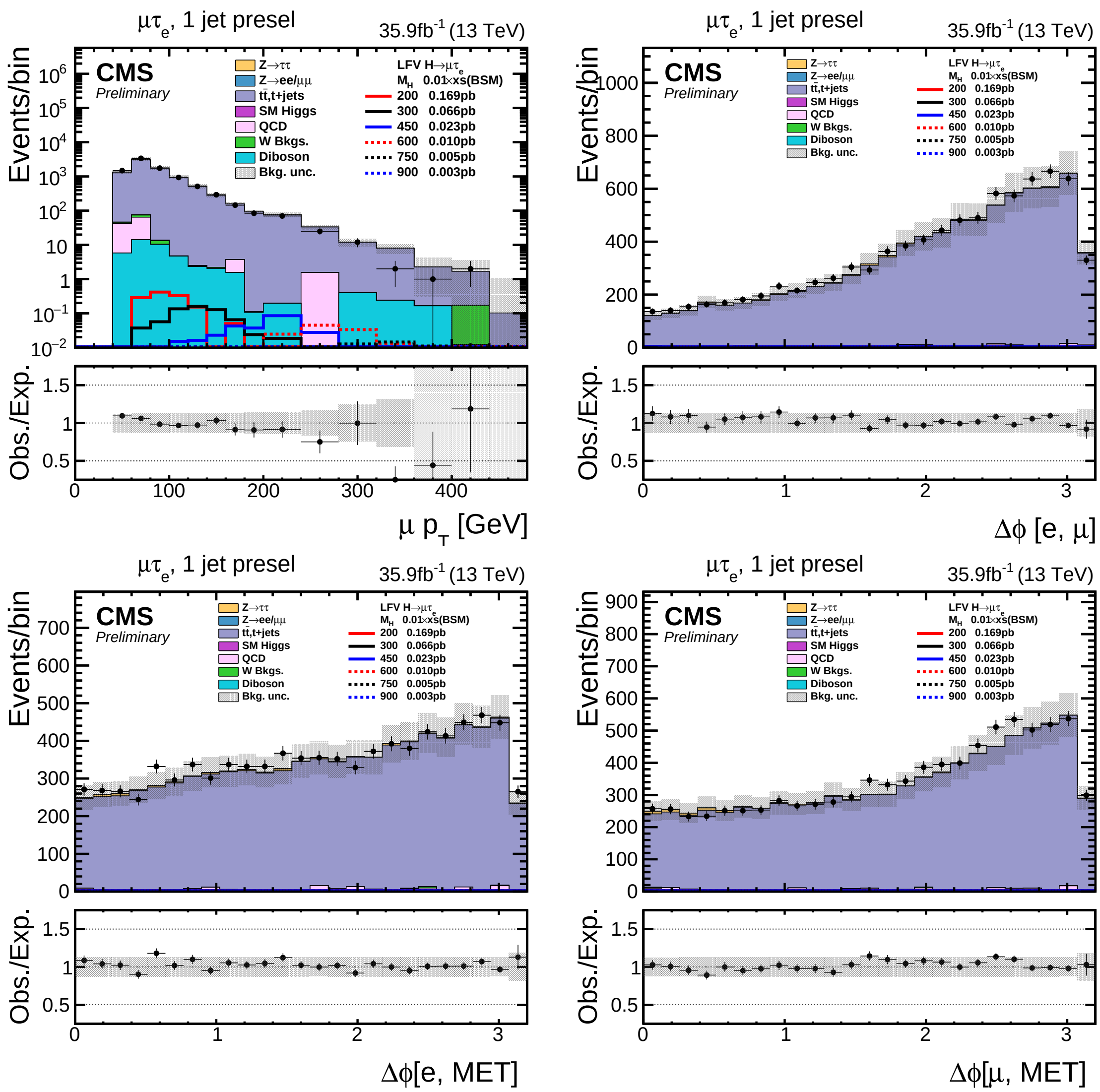


Figure 6.6: Distributions of several kinematic variables in the $t\bar{t}$ enriched control region for $\mathrm{H} \to \mu\tau_e$ analysis.

CHAPTER 7

SIGNAL EXTRACTION AND SYSTEMATIC UNCERTAINTIES

> Statistics is the grammar of science.
>
> *Karl Pearson*

7.1 Introduction

The analysis is, in its essence, a sophisticated counting experiment. The presence of a signal is indicated by an excess of events over the predicted background, in the distribution of a signal variable. For our analyses the signal variables are collinear mass or BDT output, as described in chap 4 and 5. Given that there are several uncertainties, both experimental and theoretical and also due to the innate randomness in the process, it is possible that an excess is observed when there is no signal. So, when an excess is observed, a p-value which represents the probability that the excess is due to statistical fluctuations is computed. A very low p-value is taken to indicate that the excess corresponds to an observed signal and not merely a statistical fluctuation. Conversely, if no excess is observed (upper exclusion) limits are set on the product of branching fraction and production cross-section. A 95% CL (confidence level) is taken as a requirement for ruling out a signal at or above a certain value known i.e. upper exclusion limit. The first part of this chapter describes the statistical methods used, that very closely follow the procedure used for LHC Higgs boson search described in [110].

Several sources of systematic uncertainties need to be considered when making the above measurement. The sources of these uncertainties can be theoretical, experimental or purely statistical in nature. Further, they can affect only the overall scale of the distributions (used to make the measurement), or affect there shape i.e. change the scale differently in each bin of the distribution. All the uncertainties used in the analyses and their sources are described in the second part of this chapter.

7.2 Statistical methods for signal extraction

In the following section, the expected signal event yields are denoted by s, and backgrounds by b. The parameter μ that appears below is the signal strength modifier, which changes the signal production cross-sections of all the production mechanisms by exactly the same scale μ.

7.2.1 Likelihood function

The Poisson distribution is an appropriate model for n, the number of times an event occurs in an interval if the following assumptions are true [111].

- The occurrence of one event does not affect the probability that a second event will occur. That is, events occur independently.
- The rate at which events occur is constant. The rate cannot be higher in some intervals and lower in other intervals. This rate is the average number of events in the interval, λ.
- Two events cannot occur at exactly the same instant; instead, at each very small sub-interval exactly one event either occurs or does not occur.

The poisson probability of distribution is then given by:

$$P(n_{events}) = \frac{e^{-\lambda}\lambda^n}{n!} \tag{7.1}$$

For a counting experiments such as ours, the above conditions approximately hold. The expected number of events is $\mu \cdot s + b$. The likelihood function $\mathcal{L}(data|\mu)$ is then given by:

$$\mathcal{L}(data|\mu) = \prod_{i=1}^{bins} \frac{(\mu \cdot s_i + b_i)^{n_i}}{n_i!} e^{-\mu \cdot s_i - b_i} \tag{7.2}$$

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

7.2.2 Treatment of systematic uncertainties

All systematic uncertainties are handled by introducing them as nuisance parameters. Nuisance parameters are parameters that influence the model but are not of interest in our measurement, e.g., if we are interested in knowing only the mean of a population that is expected to be distributed as a gaussian, the standard deviation becomes a nuisance parameter for the model that we fit. In our experiment, the nuisance parameters are embedded into the likelihood function. In order for the likelihood function to have a clean factorized form [110], all sources of uncertainties considered are taken to be 100%-correlated or uncorrelated. If an uncertainty is partially correlated, it is either separated into 100%-correlated or uncorrelated components, or considered 100%-correlated or uncorrelated, depending on whichever is a more conservative estimate. The full suite of nuisance parameters is represented as θ. These effect the expected signal and background yields which are now represented as $s(\theta)$ and $b(\theta)$. Each component of θ is associated with a default value $\tilde{\theta}$, reflecting our degree of belief on the real value of θ. The pdf (probability distribution function) $\rho(\theta|\tilde{\theta})$ can then be interpreted as a posterior distribution from measurements of $\tilde{\theta}$. Using Bayes' theorem:

$$\rho(\theta|\tilde{\theta}) = \rho(\tilde{\theta}|\theta) \cdot \pi_\theta(\theta), \tag{7.3}$$

where the priors $\pi_\theta(\theta)$ are taken as flat distributions representing no prior knowledge of θ. This reformulation allows us to use the pdf of $\tilde{\theta}$ instead, i.e. $\rho(\tilde{\theta}|\theta)$ to directly constrain the likelihood of the measurement. The likelihood function after the introduction of systematic uncertainties now becomes:

$$\mathcal{L}(data|\mu,\theta) = Poisson(data|\mu \cdot s(\theta) + b(\theta)) \cdot \rho(\tilde{\theta}|\theta) \tag{7.4}$$

Systematic uncertainties that effect only the overall scale of the distributions, correspond to a multiplicative factor in the signal and/or background yields, and are described by log-normal pdfs. Log-normal pdfs are characterized by the width κ, and are well-suited for positively valued observables. The log-normal distribution looks like:

$$\rho(\theta|\tilde{\theta}) = \frac{1}{\sqrt{2\pi}\,\ln(\kappa)} \exp\,(\frac{\ln(\theta/\tilde{\theta})^2}{2(\ln\,\kappa)^2})\frac{1}{\theta} \tag{7.5}$$

Systematic uncertainties that effect the scale of the distribution differently in each bin have the effect of altering its shape along with its scale. Such uncertainties are called shape uncertainties [112], and are modeled using a linear extrapolation method [113]. In practice, two alternate distributions obtained by varying the nuisance by ± 1 standard deviation are used, and a parameter is added to the likelihood that smoothly interpolates between these shapes.

7.2.3 Calculation of exclusion limits

The $\mathrm{CL_s}$ method [114–116] is used to set upper exclusion limits when no excess of data over background is observed. The test statistic used generally for hypothesis testing in searches at the LHC, uses profiling of nuisances as described above, and is based on the likelihood ratio [117], which by the Neyman-Pearson lemma is known as the most powerful discriminator. This is denoted by $\tilde{q_\mu}$, and is given by:

$$\tilde{q_\mu} = -2\ \ln \frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})},\ \text{with}\ 0 \leq \mu \leq \hat{\mu} \tag{7.6}$$

where $\hat{\theta}_\mu$ refers to the conditional maximum likelihood estimators of θ, i.e. the set of nuisances parameters that maximize the likelihood for a given signal strength μ, while $\hat{\mu}$ and $\hat{\theta}$ refer to the global maximum likelihood estimators for μ and θ. The lower constraint on $\hat{\mu}$ i.e., $\hat{\mu} \geq 0$ ensures that the signal rate cannot be negative, while the upper constraint that $\hat{\mu}$, which is the global maximum value, cannot be less than the value of μ under consideration is imposed to guarantee that upward fluctuations of data such that $\hat{\mu} \geq \mu$ are not considered as evidence against the signal hypothesis, i.e., a signal of strength μ.

Now, using equation 7.6, the observed value of the test statistic, $\tilde{q_\mu}^{obs}$, is calculated for the signal strength μ. Also, maximum likelihood estimators for the nuisance parameters, for the background-only($\mu = 0$) and signal-plus-background(current $\mu > 0$ under consideration) hypotheses are calculated. They are denoted by $\hat{\theta}_0^{obs}$ and $\hat{\theta}_\mu^{obs}$ respectively, and are used to generate toy Monte Carlo pseudo-datasets. These pseudo datasets are used to construct pdfs, using equation 7.6, of test statistics $f(\tilde{q_\mu}|0, \hat{\theta}_0^{obs})$ and $f(\tilde{q_\mu}|\mu, \hat{\theta}_\mu^{obs})$ by treating them as they were real data. Example of these distributions are shown in Fig. 7.1.

Having constructed the above pdfs, it is now possible to calculate the probabilities of the observations under both hypotheses. The first quantity that we calculate is:

$$p_\mu = P(\tilde{q_\mu} \geq \tilde{q_\mu}^{obs}|\text{signal-plus-background}) = \int_{\tilde{q_\mu}^{obs}}^{\text{inf}} f(\tilde{q_\mu}|\mu, \hat{\theta}_\mu^{obs}) d\tilde{q_\mu} \tag{7.7}$$

The above quantity corresponds to $\mathrm{CL_{s+b}}$ and measures the incompatibility of data with signal-plus-background hypothesis. This quantity alone is not adequate for hypothesis testing in situations when the signal is so small that both hypotheses

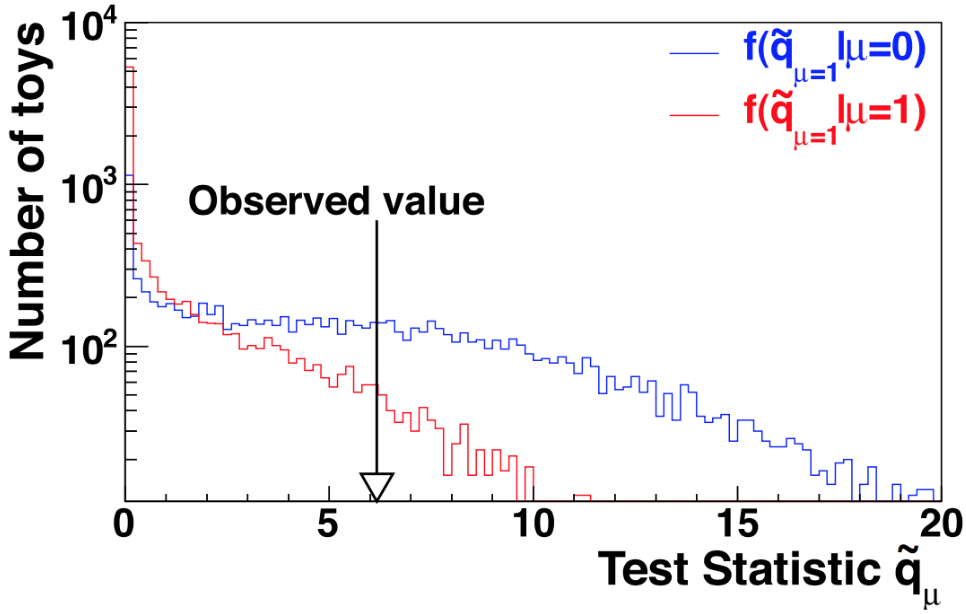


Figure 7.1: Test statistic distributions for ensembles of pseudo-data generated for signal-plus-background (red) and background-only (blue) hypotheses [110].

are compatible with the observation and a downward fluctuation of the background can lead to an inference of signal.

The second quantity we calculate is:

$$1 - p_b = P(\tilde{q}_\mu \geq \tilde{q}_\mu{}^{obs}|\text{background-only}) = \int_{\tilde{q}_\mu{}^{obs}}^{\text{inf}} f(\tilde{q}_\mu|0, \hat{\theta}_0{}^{obs}) d\tilde{q}_\mu \tag{7.8}$$

This quantity corresponds to $\mathrm{CL_b}$ and measures the incompatibility of data with the background. The incompatibility of the data with background-only hypothesis alone doesn't tell us that it is indeed compatible with the signal, and so is not considered a good test of the signal hypothesis.

The ratio of the two quantities referred to as $\mathrm{CL_s}$ [114–116] helps deal with both situations above well, and is given by:

$$\mathrm{CL_s} = \frac{p_\mu}{1 - p_b} \tag{7.9}$$

The 95% CL is then arrived at by iterating over μ until we have $\mathrm{CL_s} = 0.05$. And the amount of signal or above, given by that μ, denoted as $\mu^{95\%CL}$, is said to be excluded at 95% CL.

7.2.4 Median expected limits

Upper exclusion limits calculated using toy datasets of background-only expectation, are called expected limits. A large set of background-only pseudo-data is generated, and $\mathrm{CL_s}$ and $\mu^{95\%CL}$ is calculated for each of them. The median expected limit is calculated by integrating over this distribution until the 50% quantile is reached. The $\pm 1\sigma$ and $\pm 2\sigma$ bands are calculated similarly by integrating the distribution to the appropriate quantiles are reached. The calculation of median expected limits does not involve using the observed data and hence can be calculated when the analyses is blinded to prevent experimenter's bias (as mentioned in Section 5.1). This can be use to maximize the sensitivity of the search, as described in Sections 5.2.3 and 5.3.3. A more stringent (lower) median limit corresponds to a more sensitive search.

7.2.5 Quantifying an excess of events

In case an excess of data over background is observed, it is necessary to make sure beyond a reasonable doubt that the excess is not merely a fluctuation. This is quantified using the background-only p-value, which is the probability for the background to fluctuate and give an excess of events as large or larger than that observed. The same test statistic as equation 7.6 is used with the signal strength set to 0 to correspond to the background-only hypothesis:

$$\tilde{q}_0 = -2 \ln \frac{\mathcal{L}(\text{data}|0, \hat{\theta}_0)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})}, \text{ with } 0 \leq \hat{\mu} \tag{7.10}$$

The constraint on $\hat{\mu}$ being greater than 0 is required so that a deficit of events in observed data is not interpreted in the same manner as we would an excess. In other words a departure from the background hypothesis in the form of deficit of events is not considered in favor of the signal hypothesis. Following the same procedure as calculation of observed limits (as described in 7.2.3) and generating pseudo-data, the distribution $f(\tilde{q_0}|0, \hat{\theta_0}^{obs})$ is constructed. The p-value is then given by:

$$p_0 = P(\tilde{q_0} \geq \tilde{q_0}^{obs}) = \int_{\tilde{q_0}^{obs}}^{\inf} f(\tilde{q_0}|0, \hat{\theta_0}^{obs}) d\tilde{q_0} \tag{7.11}$$

The p-value can be converted to significance $\mathcal{Z}_0$, which is an equivalent way of quantifying an excess and is related to the p-value by the following:

$$p_0 = \int_{\mathcal{Z}_0}^{\inf} \frac{1}{\sqrt{2\pi}} exp(-x^2/2) dx \tag{7.12}$$

Broadly, the significance corresponds to how far into the tail of the distribution (i.e., away from the most probable value), assuming background hypothesis, the test statistic value corresponding to the observed data lies. The farther it is, the less likely it is to have been a fluctuation. The conventional standard in high energy physics to be able to claim observation of a process is a significance of 5σ, which corresponds to a p-value of 2.8×10^{-7}.

7.3 Systematic uncertainties

It is important to consider all relevant sources of uncertainties when performing sophisticated counting experiments such as these. Uncertainties that are introduced as a result of imprecise/inaccurate knowledge of the system or gaps in prior knowledge that is used in the measurement are called systematic uncertainties. They are a different class of uncertainties than those arising purely out randomness in statistical measurements, called statistical uncertainties. The sources of systematic uncertainties range from purely theoretical in nature to purely experimental. They can be categorized in the two following ways:

7.3.1 Normalization uncertainties

The value of these uncertainties are independent of the signal/discriminant variable. To be more precise, these uncertainties are independent of the value of M_{col} or BDT response. Hence, they effect each bin of those distributions in exactly the same manner and thus change only the overall scale of the distribution without altering its shape.

The muons in the analysis are required to pass certain identification, isolation and triggering criteria (see chapter 5). The efficiencies for muon to pass these criteria are measured via tag-and-probe methods [75] using $Z \to \mu\mu$ events, and the scale factors are used to match the efficiency in MC to that in data (see section 3.3.7.1). The efficiencies determined by this process, like any other quantity, are associated with systematic uncertainties. For the muons used in the analyses described here, a combined normalization uncertainty of 2% is associated with muon trigger, identification and isolation. Similar to muons, the efficiencies for electrons used in the analyses have also been measured via tag-and-probe methods [97] using $Z \to$ ee events. The uncertainties in efficiencies of electron identification and isolation criterion are also included as a normalization uncertainty of 2% in the fit. Both the above uncertainties are applied to processes which are derived from MC simulation. As mentioned earlier, a b tagging veto is applied in the analysis in order suppress backgrounds involving top quarks. The efficiency of b tagging procedure is different in MC simulation than data. A scaling procedure is applied to match these efficiencies, and the uncertainties associated with these factors are found to not effect the shape of the M_{col} or BDT distributions. They are thus included in the fit as normalization uncertainties and range across categories from 2-4.5% and 2-2.5% for h $\to \mu\tau_e$ and H $\to \mu\tau_e$ analysis respectively.

Several backgrounds in the analyses are estimated using MC simulations (see chapter 6). These include Z $\rightarrow \tau\tau$, $t\bar{t}$, W+jets, WW, WZ and ZZ, $Z \rightarrow \ell\ell$ ($\ell =$ e, μ)+jets, single-top quark production, W$\gamma^{(*)}$ + jets. The production cross-sections for these backgrounds determine the number of events each background would contribute. These cross-sections are measured experimentally and the uncertainty in those measurements are included in the fit. Given that a change in cross-section changes the overall number of events produced, it has no effect on the shape of distributions. Hence these uncertainties are included as normalization uncertainties. These uncertainties in general arise from: uncertainties on the parton distribution functions and strong coupling constant (called PDF+α_s); variations in renormalization and factorization scales. In the H $\rightarrow \mu\tau_e$ analysis a separate uncertainty is applied for PDF+α_s and renormalization/factorization scales for each of the backgrounds. In h $\rightarrow \mu\tau_e$ analysis, a combined uncertainty for each background is applied to cover both sources. All the above uncertainties are considered 100% correlated among categories. For each background, a 5% uncertainty, uncorrelated among all categories, is also applied to conservatively cover differences across categories. The QCD multijet background is estimated using a data-driven procedure. An uncertainty of 30% associated with this procedure (corresponding to the uncertainty in the extrapolation factor from the same-sign to opposite-sign region) is included in the fit. All uncertainties are summarized in table 7.1 [1, 2]. All uncertainties in the table are treated as correlated between the categories, except those with more values separated by the $\oplus$ symbol. In the case of two values, the first value is the correlated uncertainty and the second value is the uncorrelated uncertainty for each individual category. In the case of three values, the first and second values correspond to the uncertainties arising from factorization and renormalization scales and PDF variations and are correlated between categories, while the third value is the uncorrelated uncertainty for each individual category. Two values separated by the "–" sign represent the range of the uncertainties from the different sources and/or in the different jet categories.

Just like MC backgrounds described above, the signal in both h $\rightarrow$ $\mu\tau_e$ and H $\rightarrow \mu\tau_e$ analysis comes from MC simulation. The signal process and the SM Higgs background are associated with uncertainties in the Higgs boson production cross sections. These come from variations in factorization/renormalization scales, as well as PDF+α_s, and result in changes only in normalization. These uncertainties are summarized for SM Higgs boson and heavier Higgs of different masses in table 7.1. They are taken from Handbook of LHC Higgs cross-sections found in Ref. [21]. The PDF and α_s uncertainties are computed following the recommendation of the PDF4LHC working group. The remaining Gaussian uncertainty accounts for additional intrinsic sources of theory uncertainty described in detail in the reference [2].

The estimation of a particular background, that is derived from simulation, needs to correspond to the number of events (of that background, having a particular cross-section) that would be produced in the amount of proton-proton collision data that we are using for this search. In other words, the background estimations need to be normalized to (brought to the same scale as) the integrated luminosity of the data collected. This integrated luminosity (defined in chapter 3) is a measured quantity, and like all measured quantities, has an uncertainty associated with it. This amounts to 2.5% and, like other normalization uncertainties, only effects only the overall scale of distributions.

7.3.2 Shape uncertainties

In this section, we describe the systematic uncertainties which not only alter the scale but also the shape of the distributions. We start with the description of uncertainties associated with jet energy corrections. As described in Section 4.4.6, the reconstruction of complex objects such as jets need to corrected using several

TABLE 7.1

THE SYSTEMATIC UNCERTAINTIES FOR BOTH ANALYSES (SEE TEXT FOR EXPLANATION OF THE $\oplus$ SYMBOL)

Systematic uncertainty	$\mathrm{h} \to \mu\tau_e$	$\mathrm{H} \to \mu\tau_e$
Muon trigger/ID/isolation	2%	2%
Electron trigger/ID/isolation	2%	2%
b tagging veto	2.0–4.5%	2.0–2.5%
QCD multijet background	30%	30%
$\mathrm{Z} \to \tau\tau + \mathrm{jets}$ background	10%$\oplus$5%	0.1%$\oplus$2%$\oplus$5%
$t\bar{t}$ background	10%$\oplus$5%	10%$\oplus$5%
$W + \mathrm{jets}$ background	10%$\oplus$5%	0.8%$\oplus$3.8%$\oplus$5%
WW, ZZ, WZ background	5%$\oplus$5%	3.5%$\oplus$5%$\oplus$5%
$\mathrm{W}\gamma^{(*)}$ background	10%$\oplus$5%	10%$\oplus$5%
Single top quark background	5%$\oplus$5%	3%$\oplus$5%$\oplus$5%
$Z \to \mu\mu$/ee background	10%$\oplus$5%	0.1%$\oplus$2%$\oplus$5%
Jet energy scale	3–20%	3–20%
μ energy scale	0.2%	0.2%
e energy scale	0.1–0.5%	0.1–0.5%
Unclustered energy scale	$\pm 1\sigma$	$\pm 1\sigma$
pileup	$\pm 1\sigma$	$\pm 1\sigma$
Integrated luminosity	2.5%	2.5%

TABLE 7.2

THEORETICAL UNCERTAINTIES APPLIED TO THE HIGGS BOSON PRODUCTION CROSS SECTIONS FOR THE DIFFERENT MASSES

m_{H} (GeV)	Production mode	Theory, Gaussian (%)	PDF+α_s (%)
125	GGF	± 3.9	± 3.2
125	VBF	± 0.4	± 2.1
200	GGF	± 1.8	± 3.0
300	GGF	± 1.8	± 3.0
450	GGF	± 2.0	± 3.1
600	GGF	± 2.1	± 3.5
750	GGF	± 2.1	± 4.0
900	GGF	± 2.2	± 4.6

correction factors. These factors have uncertainties associated with them and these are included in the fit as shape uncertainties. There are several different sources of these uncertainties and the effect of each is propagated to the fit by including alternate distributions for each process where each source of uncertainty has been moved by one standard deviation on either side. The effect of changing these sources is propagated through the jets in the analysis, and also other affected quantities such as the p_T^{miss}. A total of 27 sources are considered, and these include effects of pileup, composition of jets, η dependence etc. They vary in the range from 3-20% and are considered uncorrelated.

Just like jets, there are shape uncertainties associated with the energy scale of electrons and muons. The effect of electron energy scale uncertainty is treated in a similar manner by propagating the effect of varying the scale to process distributions which are then included in the fit. The uncertainties are a result of the sum in quadrature of the following components: electron selection efficiency, pseudorapidity dependence, and shower-shape related categorization. The resolution systematics result to be negligible and are thus not considered in the fit. The value of this uncertainty ranges from 0.1-0.5%. The muon energy scale uncertainty amounts to 0.2% and is treated in the same manner as above.

Jets with $p_T < 15$ GeV or PF candidates which do not get clustered inside any jets are called unclustered energy. Their scale that effects the p_T^{miss} in particular, and is associated with a shape uncertainty that is treated the same way as others [103]. Four sources of unclustered energy scale uncertainty are considered, and estimated independently for these four particle categories: charged particles, photons, neutral hadrons, and HF (very forward, high $|\eta|$) particles which are not contained in jets. The effect of these sources on the unclustered energy is propagated in the same manner as described above, and they are considered uncorrelated.

A set of weights is applied in order make the distribution of pileup in MC simulation match that of pp collision data. There is an uncertainty associated with this process. This is included by varying the weights by changing by 5%, in each direction, the total inelastic cross section used in the estimation of the pileup events in data [118]. These new set of weights are then applied event-by-event, producing alternate distributions that are included in the fit as shape uncertainties.

A shape uncertainty is also considered to deal with the fact that purely statistical fluctuations can change the shape of the distribution. These uncertainties are called bin-by-bin uncertainties as they account for statistical uncertainty in every bin of the distribution. Alternate distributions are created by varying the contents of each bin up and down, and these distributions are included in the fit as shape uncertainties [119]. Given that considering all bins of all processes will result in a very large number of such nuisances being considered, shape uncertainties for only those bins are considered in which there is more than 10% variation in the up and down shift. All shape uncertainties are summarized in Table 7.1.

CHAPTER 8

RESULTS

> There are two possible outcomes: if the result confirms the hypothesis, then you've made a measurement. If the result is contrary to the hypothesis, then you've made a discovery.
>
> *Enrico Fermi*

In this chapter the results of both the searches are presented. The results for the h $\rightarrow \mu\tau_e$ search are first presented. Results for the H $\rightarrow \mu\tau_e$ search follow.

8.1 h $\rightarrow \mu\tau_e$ results

The resulting distributions of the signal variable (after applying all selection requirements as outlined in 5.2) are fit using a binned maximum likelihood fit. The entire procedure is described in detail in 7.2. All systematic uncertainties are included as nuisance parameters, and the fit is performed simultaneously across all categories. The BDT response distributions of signal and background are shown superimposed for each category in Fig 8.1. The distribution of M_{col} for the M_{col}-fit analysis are also shown in Fig 8.2. We do not observe an excess of signal over expected background. Hence, upper exclusion limits on $\mathcal{B}(\mathrm{h} \rightarrow \mu\tau_e)$ are set, following the procedure described in 7.2.3. In table 8.1, the median expected limits, observed limits and the best fit branching fractions for $\mathcal{B}(\mathrm{h} \rightarrow \mu\tau_e)$ are summarized. As noted earlier in this thesis, the tau lepton coming from the Higgs can also decay hadronically. This channel of the LFV Higgs decay, i.e. $\mathrm{h} \to \mu\tau_{\mathrm{h}}$ is studied in an analyses by different members of the same research team [1]. The limits on $\mathcal{B}(\mathrm{h} \to \mu\tau_{\mathrm{h}})$ from that search are combined with limits on $\mathcal{B}(\mathrm{h} \to \mu\tau_e)$, as calculated above from the search described here. All limits are summarized graphically in Figure 8.3. The combined observed (median expected) upper limits on $\mathcal{B}(\mathrm{h} \to \mu\tau)$ is 0.25 (0.25) % at 95% CL, for the BDT-fit analysis. The combined best fit branching fraction of $\mathcal{B}(\mathrm{h} \to \mu\tau)$ is found to be 0.00 ± 0.12, also for the BDT-fit analysis. Figure 8.4 shows a historical compilation of results from $\mathrm{h} \to \mu\tau_e$ searches until 2017. It is important to note that the $2.4\,\sigma$ excess observed by the earlier CMS search with 8 TeV data [14] has now been excluded by this search.

The constraints on $\mathcal{B}(\mathrm{h} \to \mu\tau)$ can be transformed into constraints on Lepton Flavor Violating Yukawa Couplings $(Y_{\mu\tau}, Y_{\tau\mu})$. These couplings represent the strength of an interaction and are related to the decay width $\Gamma(\mathrm{h} \to \mu\tau)$ in the following way [13]:

$$\Gamma(\mathrm{h} \to \mu\tau) = \frac{m_h}{8\pi}(|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2). \tag{8.1}$$

The decay width is also related to the branching fraction, $\mathcal{B}(\mathrm{h} \to \mu\tau)$ according to the following equation:

$$\mathcal{B}(\mathrm{h} \to \mu\tau) = \frac{\Gamma(\mathrm{h} \to \mu\tau)}{\Gamma(\mathrm{h} \to \mu\tau) + \Gamma_{SM}}. \tag{8.2}$$

where the SM Higgs decay width is assumed to be $\Gamma_{SM} = 4.1$ MeV [120] for $m_{\mathrm{H}} = 125\,\mathrm{GeV}$. Using equations 8.1 and 8.2, we derive the constraints on Yukawa couplings at 95% CL. The limits for the Yukawa couplings are summarized in Table 8.2. Fig. 8.5 pictorially summarizes all existing limits on Yukawa couplings from different direct and indirect searches. It also shows the theoretical “naturalness” limit considering/expecting LFV couplings to be smaller than those of couplings for SM decays

TABLE 8.1

EXPECTED AND OBSERVED UPPER LIMITS AT 95% CL, AND BEST FIT BRANCHING FRACTIONS IN PERCENT FOR EACH INDIVIDUAL JET CATEGORY, AND COMBINED, FOR THE h $\rightarrow \mu\tau_e$ ANALYSIS

Expected limits (%)

	0-jet	1-jet	2-jets ggH	2-jets VBF	Combined
BDT fit analysis	<0.83	<1.19	<1.98	<1.62	<0.59
M_{col} fit analysis	<1.01	<1.47	<3.23	<1.73	<0.75

Observed limits (%)

	0-jet	1-jet	2-jets	VBF	Combined
BDT fit analysis	<1.30	<1.34	<2.27	<1.79	<0.86
M_{col} fit analysis	<1.08	<1.35	<3.33	<1.40	<0.71

Best fit branching fractions (%)

	0-jet	1-jet	2-jets	VBF	Combined
BDT fit analysis	0.61 ± 0.36	0.22 ± 0.46	0.39 ± 0.83	0.10 ± 1.37	0.35 ± 0.26
M_{col} fit analysis	0.13 ± 0.43	−0.22 ± 0.75	0.22 ± 1.39	−1.73 ± 1.05	−0.04 ± 0.33
combined $\mu\tau$ (BDT fit)			0.00 ± 0.12		

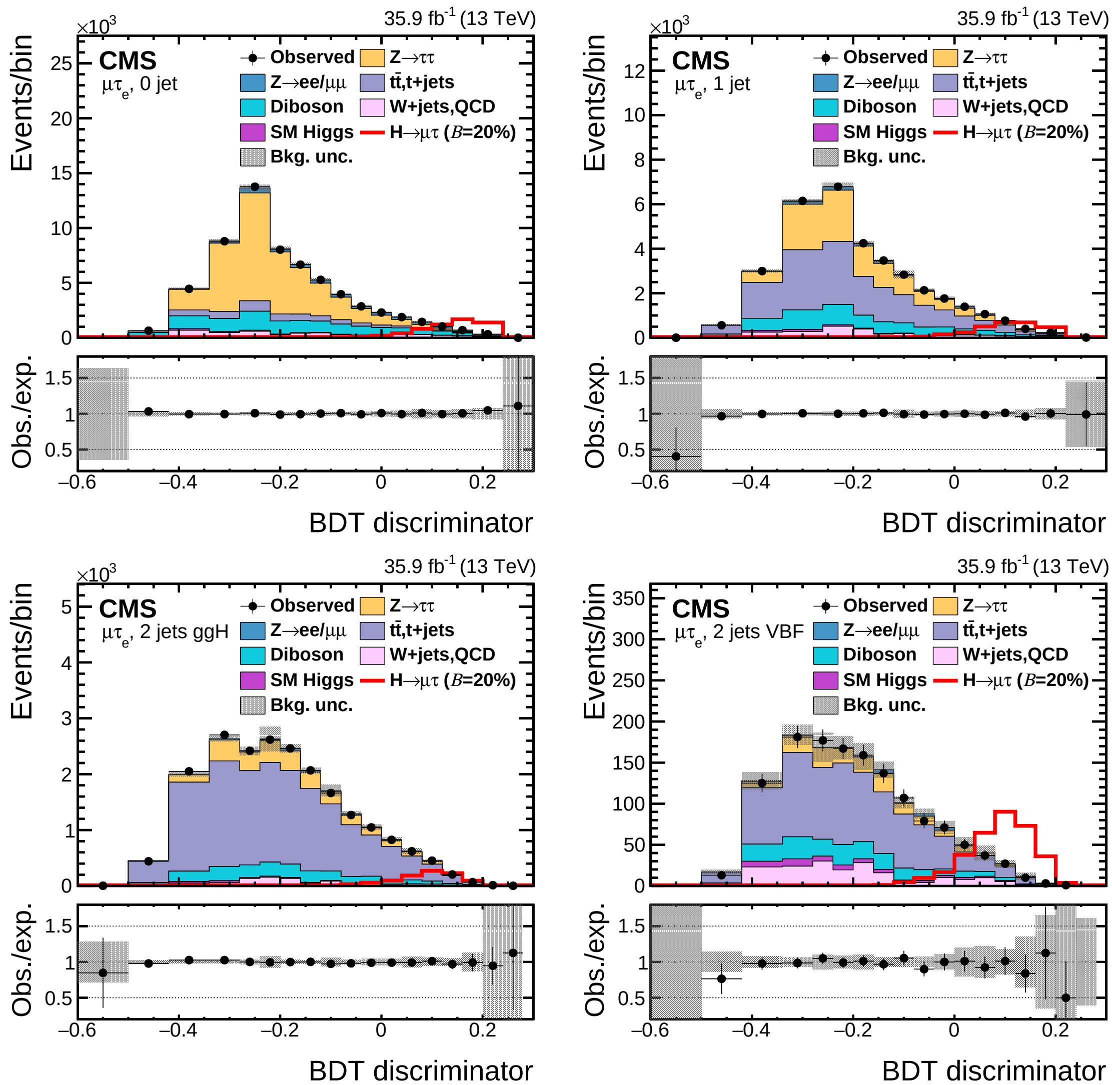


Figure 8.1: Distribution of BDT response in each category comparing signal and background estimations to observed collision data, for h $\rightarrow \mu\tau_e$ analysis. The bottom panel show the ratio of observed data and fitted background in each bin [1].

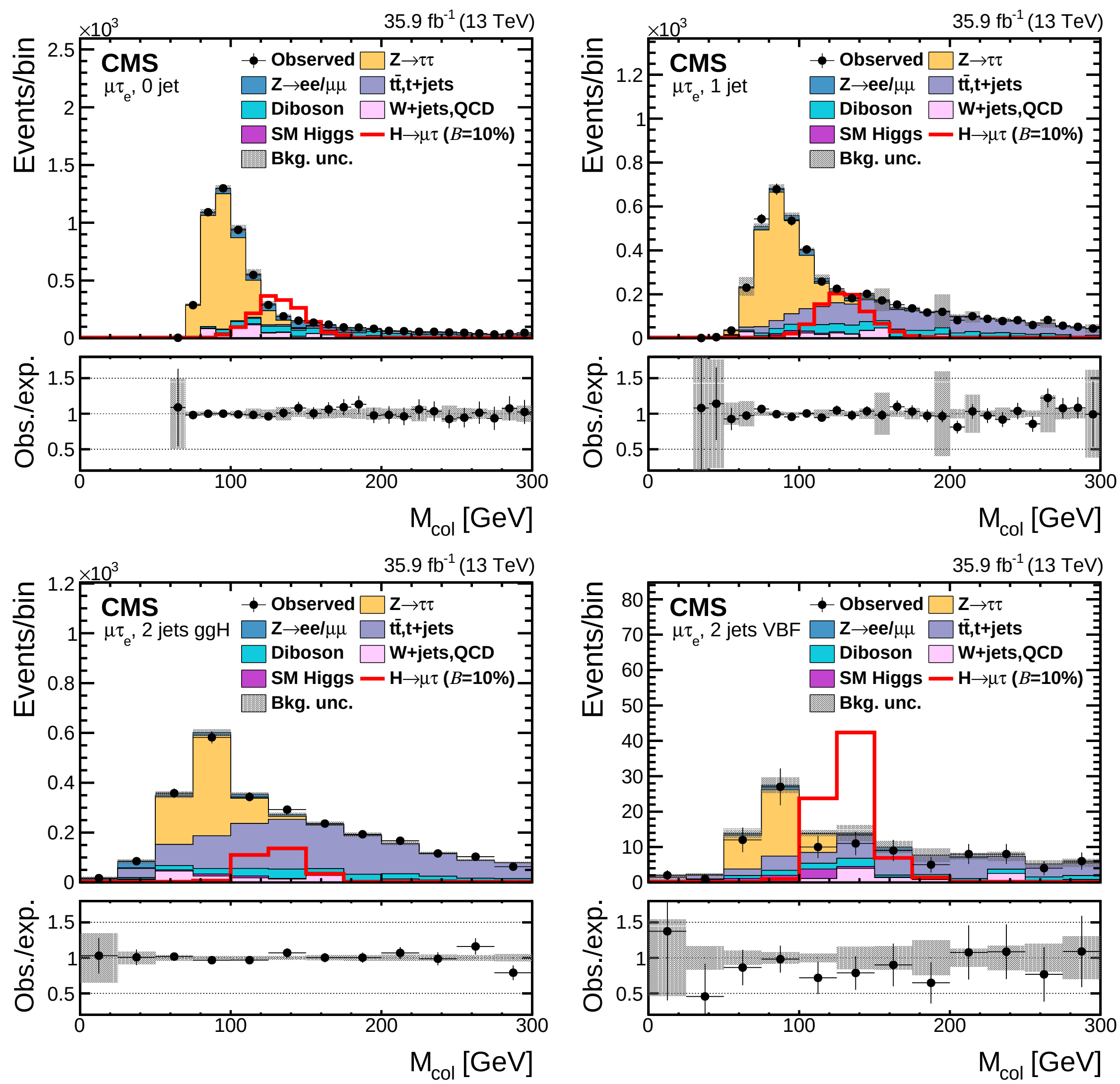


Figure 8.2: Distribution of M_{col} response in each category comparing signal and background estimations to observed collision data, for $\text{h} \to \mu\tau_e$ analysis. The bottom panel show the ratio of observed data and fitted background in each bin [1].

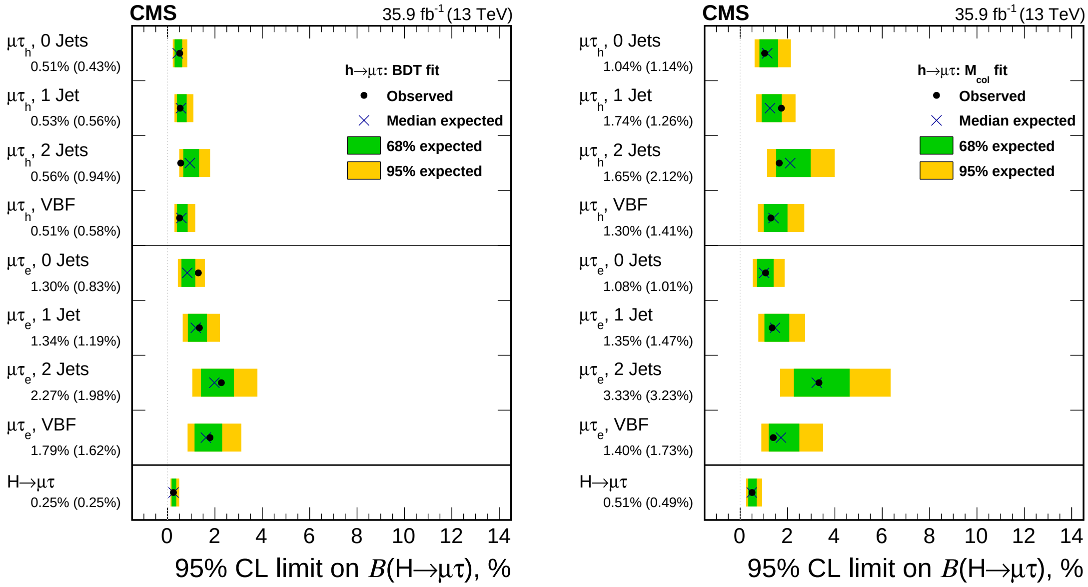


Figure 8.3: Observed and median expected upper exclusion limits for $\mathrm{h} \to \mu\tau_e$, $\mathrm{h} \to \mu\tau_{\mathrm{h}}$ and combined $\mathrm{h} \to \mu\tau$ channels, for the BDT fit (left) and M_{col} fit analysis (right). The $\pm 1\sigma$ and $\pm 2\sigma$ bands for expected limits are also shown in light green and yellow respectively [1].

of the Higgs [13], which can be considered a benchmark for sensitivity of this search. The limits derived from this search are most stringent till date, and surpass the above benchmark.

TABLE 8.2

95% CL OBSERVED UPPER LIMIT ON THE YUKAWA COUPLINGS, FOR THE BDT FIT AND THE M_{col} FIT ANALYSIS

	BDT fit	M_{col} fit
$\sqrt{\lvert Y_{\mu\tau}\rvert^2 + \lvert Y_{\tau\mu}\rvert^2}$	$< 1.43 \times 10^{-3}$	$< 2.05 \times 10^{-3}$

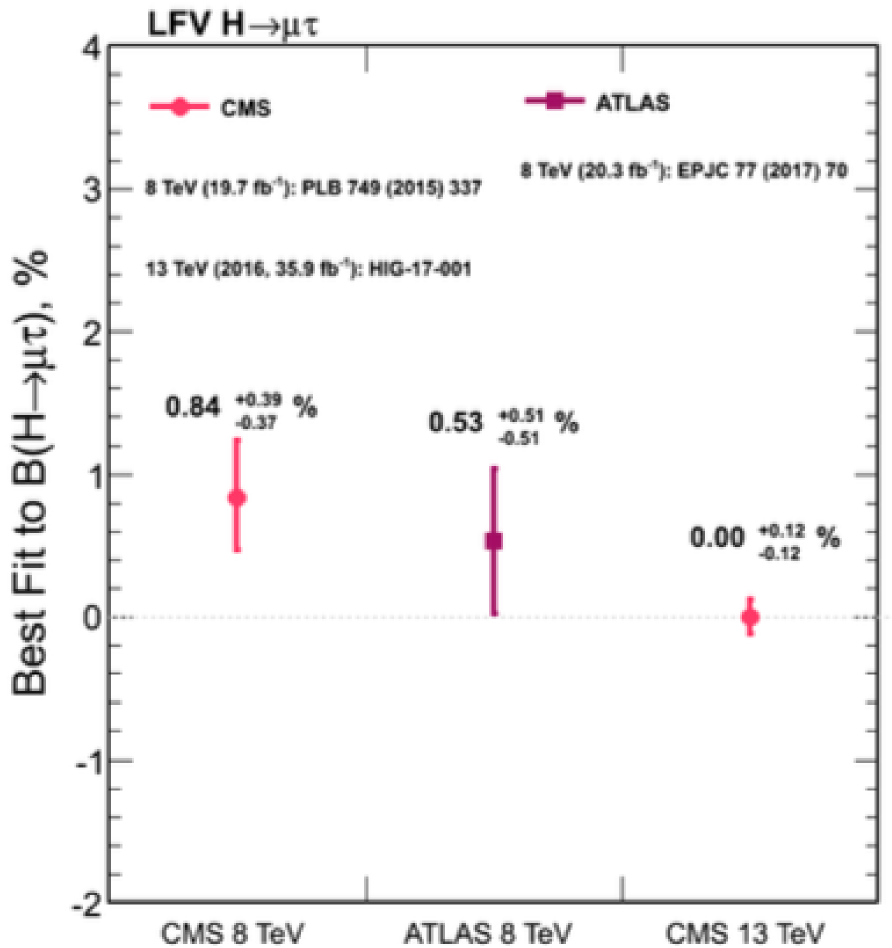


Figure 8.4: Historical compilation of results from direct searches for $h \to \mu\tau$ decay. This search is represented by the rightmost point.

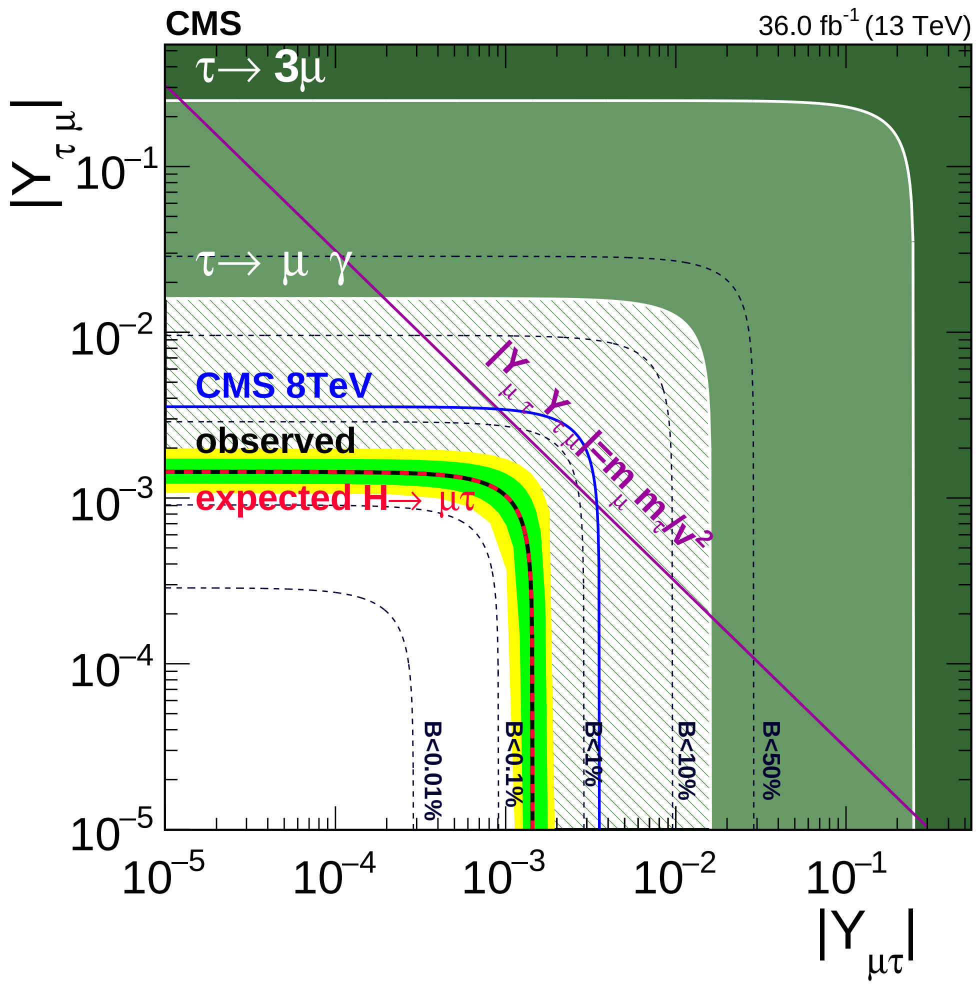


Figure 8.5: Observed (black solid)and median expected (red dashed) upper limits on h $\rightarrow \mu\tau$ Yukawa couplings from this analysis. The light green and yellow bands show the $\pm1\sigma$ and $\pm2\sigma$ spreads of the expected limit. Blue solid line shows the result from the previous CMS search with 8 TeV data [14]. The naturalness limit is shown as a purple straight line [1].

8.2 H $\rightarrow \mu\tau_e$ results

The resulting $M_{\rm col}$ distributions for signal and background estimation (after applying all selection requirements as outlined in 5.3.1), after a binned maximum likelihood fit, are shown superimposed along with the observed data Fig 8.6. All systematic uncertainties are included as nuisance parameters, and the fit is performed simultaneously across all categories. We do not observe an excess over expected background in the entire range. Unlike the $\mathrm{h} \to \mu\tau_e$ analysis described above where the production cross-section of the SM Higgs boson is known, here we are looking for LFV decay of a hypothetical heavy Higgs bosons of different masses. Hence, we set upper exclusion limits on production cross-section times branching fraction, $\sigma(\mathrm{gg} \to \mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mu\tau_e)$. The procedure is the same as used above and described in 7.2.3. The observed and median expected upper limits at 95% CL on $\sigma(\mathrm{gg} \to \mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mu\tau_e)$ are summarized in table 8.3 for different categories and Higgs masses. The limits are also summarized graphically in Figure 8.7. The observed (median expected) limits range from 159.4 (95.6) pb to 2.9 (4.9) pb for heavy Higgs masses in the range between 200 and 900 GeV. This search was combined with LFV heavy Higgs decay search with the tau lepton decaying hadronically, i.e. $\mathrm{H} \to \mu\tau_\mathrm{h}$ to produce constraints $\mathrm{H} \to \mu\tau$. The combined observed (median expected) upper limits on $\sigma(\mathrm{gg} \to \mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mu\tau)$ range from 51.9 (57.4) pb to 1.6 (2.1) pb. These limits are shown graphically in Figure 8.8. This is the first direct search till date to set limits on this decay.

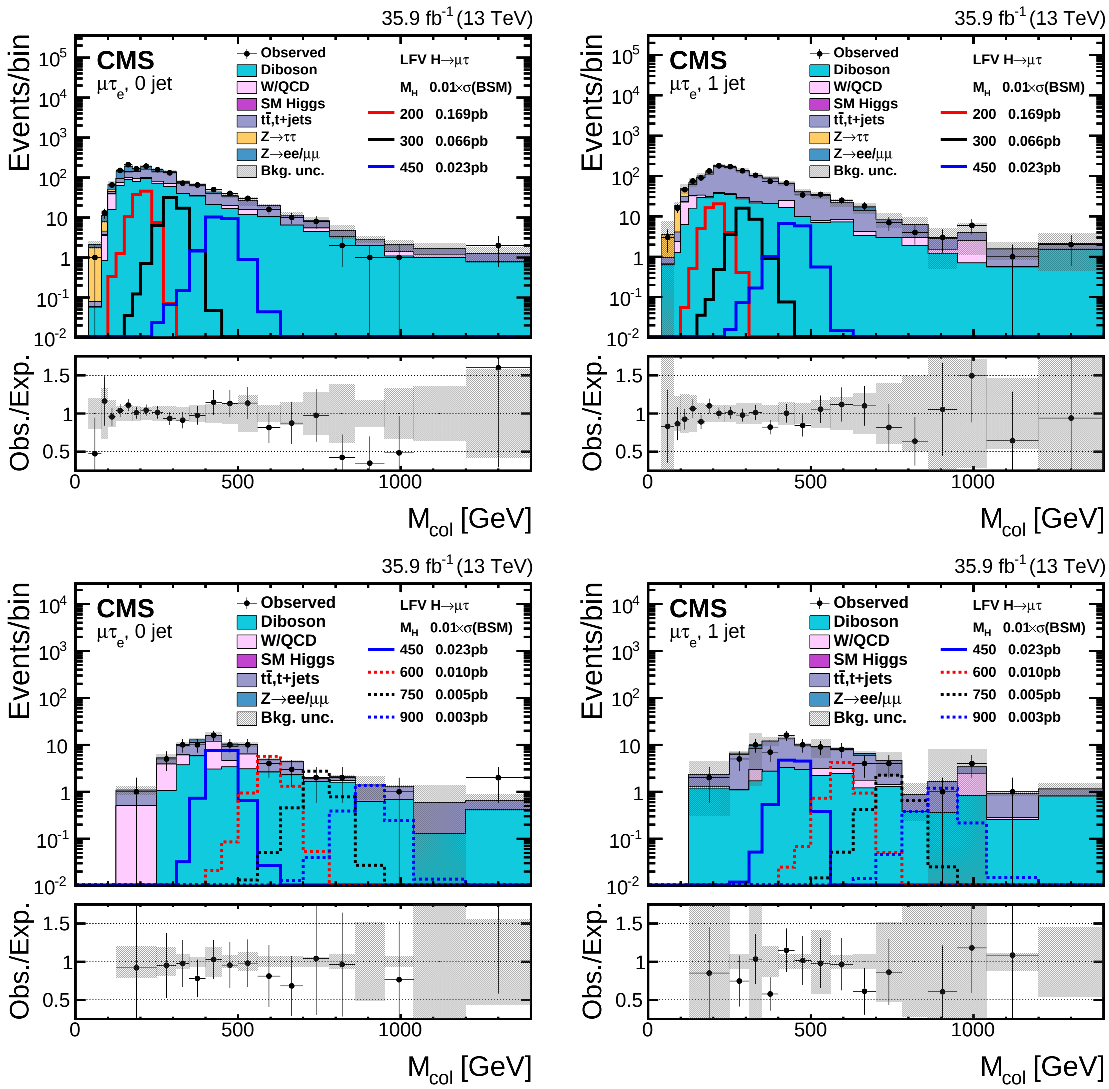


Figure 8.6: Distribution of M_{col} in 0-jet (left) and 1-jet (right) for lowmass (top) and highmass (range), comparing signal and background estimations to observed collision data, for $\text{H} \rightarrow \mu\tau_e$ analysis. The bottom panel show the ratio of observed data and fitted background in each bin [2].

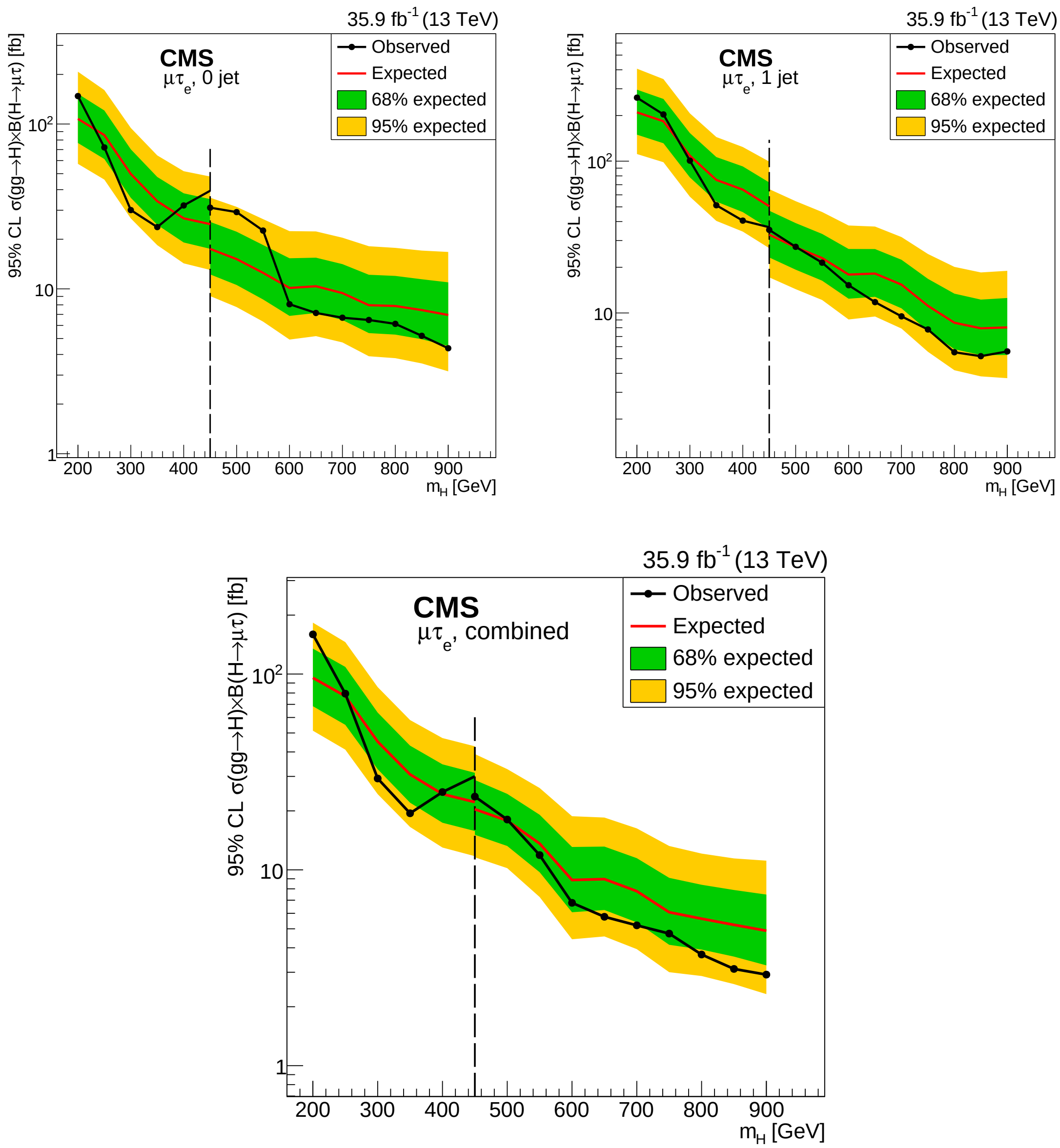


Figure 8.7: Observed and Median expected 95% upper exclusion limits for 0-jet (upper left), 1-jet (upper right) and combined (bottom),for the H $\rightarrow \mu\tau_e$ analysis [2].

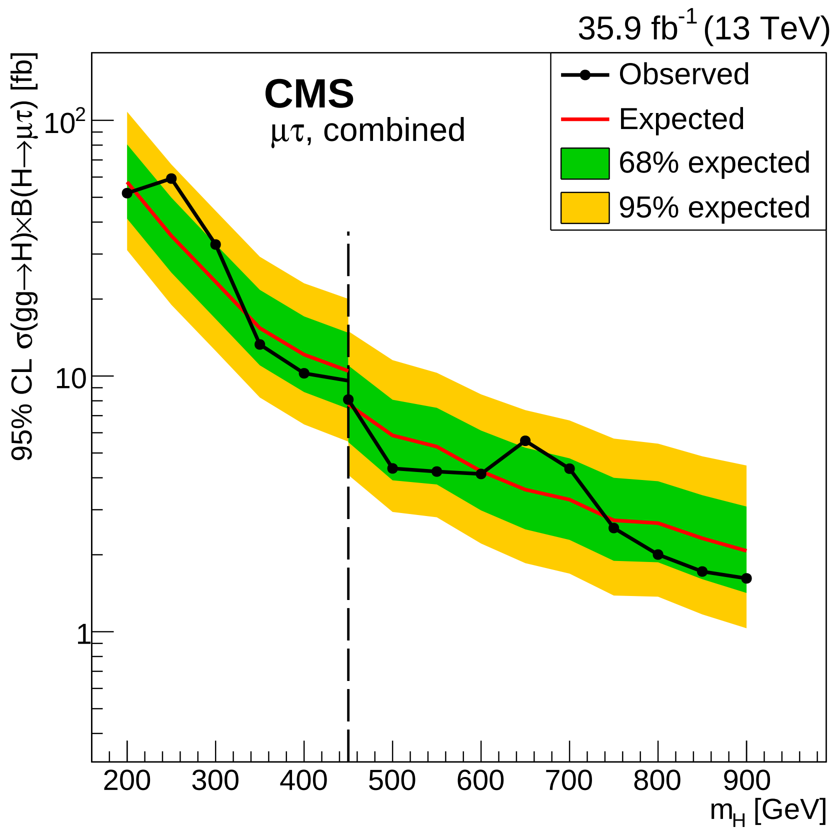


Figure 8.8: Observed and Median expected 95% upper exclusion limits for combined $H \to \mu\tau$ analysis [2].

TABLE 8.3

THE OBSERVED (MEDIAN EXPECTED) 95% CL UPPER LIMITS ON

$\sigma(\mathrm{gg} \rightarrow \mathrm{H}) \times \mathcal{B}(\mathrm{H} \rightarrow \mu\tau_e)$

m_{H} (GeV)	0 jet	1 jet	comb
200	147.8 (107.5)	262.1 (209.8)	159.4 (95.6)
300	30.1 (49.8)	100.8 (108.6)	29.3 (45.2)
450	31.1 (17.5)	35.3 (32.8))	23.7 (20.4)
600	8.1 (10.4)	15.2 (17.9)	6.8 (8.9)
750	6.5 (8.0)	7.8 (18.2)	4.7 (6.1)
900	4.4 (6.9)	5.6 (15.4)	2.9 (4.9)

CHAPTER 9

CONCLUSION

A physicist is an atom's way of knowing about atoms.

George Wald

This dissertation describes two searches for lepton flavor violating decays. Both searches were performed with data collected by the CMS detector in 2016, in proton-proton collisions at the LHC, at a center-of-mass energy of 13 TeV. Neither search found evidence of such BSM decays, and stringent upper exclusion limits have been set.

The search for SM Higgs (h) decaying into a muon and an electronically decaying tau ($\mathrm{h} \rightarrow \mu\tau_e$) is now a published result in a peer-reviewed journal [1]. This search (in combination with $\mathrm{h} \rightarrow \mu\tau_\mathrm{h}$) has set most stringent upper bounds till date on the branching fraction of h decaying to $\mu\tau$. The observed (median expected) upper limits on $\mathcal{B}(\mathrm{h} \rightarrow \mu\tau)$ is 0.25 (0.25) % at 95% CL. Upper limit on off-diagonal $\mu\tau$ Yukawa couplings, derived from the above constraint, is also set to be $\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 1.43 \times 10^{-3}$ at 95% CL. These limits constitute a significant improvement over all previous results.

The search for lepton flavor violating decays of a neutral heavy Higgs boson (H) into a muon and an electronically decaying tau ($\mathrm{h} \rightarrow \mu\tau_e$) is complete, and on its way to being a peer-reviewed publication. This search (in combination with $\mathrm{H} \rightarrow \mu\tau_\mathrm{h}$) has set upper limits on product of H cross-section and branching ratio to $\mu\tau$. These observed (median expected) limits on $\sigma(\mathrm{gg} \to \mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mu\tau)$ range from 51.9 (57.4) pb to 1.6 (2.1) pb for H masses in the range between 200 and 900 GeV. This search is the first direct search to set limits on this decay.

As the LHC's excellent performance in delivering proton-proton collisions continues, and CMS collects more an more data, there is room for above results to improve. The very next step would be to perform these searches with the entire dataset collected in the run II (2015-2018) period. This would amount to more than 3 times the data used in the above searches, and with more innovative techniques to select signal events and reduce background processes, the sensitivity of the searches can be increased.

BIBLIOGRAPHY

1. CMS Collaboration. Search for lepton flavour violating decays of the Higgs boson to $\mu\tau$ and eτ in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 06:001, 2018. doi: 10.1007/JHEP06(2018)001.

2. CMS Collaboration. Search for lepton flavour violating decays of neutral heavy higgs boson to $\mu\tau$ and eτ in proton-proton collisions at $\sqrt{s} = 13$ tev. CMS Physics Analysis Summary CMS-PAS-HIG-18-017, 2019. URL `http://cdsweb.cern.ch/record/2674882`.

3. Georges Aad et al. Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716: 1, 2012. doi: 10.1016/j.physletb.2012.08.020.

4. Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716:30, 2012. doi: 10.1016/j.physletb.2012.08.021.

5. Serguei Chatrchyan et al. Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 06:081, 2013. doi: 10.1007/JHEP06(2013)081.

6. F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321, 1964. doi: 10.1103/PhysRevLett.13.321.

7. Peter W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12:132, 1964. doi: 10.1016/0031-9163(64)91136-9.

8. Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508, 1964. doi: 10.1103/PhysRevLett.13.508.

9. G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585, 1964. doi: 10.1103/PhysRevLett.13.585.

10. ATLAS and CMS Collaborations. Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collisio n data at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 08:045, 2016. doi: 10.1007/JHEP08(2016)045.

11. Toshihiko Ota Shinya Kanemura and Koji Tsumura. Lepton flavor violation in Higgs boson decays under the rare tau decay results. *Phys. Rev. D*, 73:016, 2006. doi: LeptonflavorviolationinHiggsbosondecaysundertheraretaudecayresults.

12. Gianluca Blankenburg, John Ellis, and Gino Isidori. Flavour-changing decays of a 125 GeV Higgs-like particle. *Phys. Lett. B*, 712:386, 2012. doi: 10.1016/j.physletb.2012.05.007.

13. Roni Harnik, Joachim Kopp, and Jure Zupan. Flavor violating higgs decays. *JHEP*, 03:26, 2013. doi: 10.1007/JHEP03(2013)026.

14. Vardan Khachatryan et al. Search for lepton-flavour-violating decays of the Higgs boson. *Phys. Lett. B*, 749:337, 2015. doi: 10.1016/j.physletb.2015.07.053.

15. Marc Sher and Keith Thrasher. Flavor-changing leptonic decays of heavy Higgs bosons. *Phys. Rev. D*, 93:055021, 2016. doi: 10.1103/PhysRevD.93.055021.

16. Paolo Gunnellini. *Study of double parton scattering using four-jet scpenarios in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the CMS experiment at the Large Hadron Collider*. PhD thesis, University of Hamburg, November 2014.

17. Wikipedia foundation. The Standard model particle zoo. Website. `https://en.wikipedia.org/wiki/Standard_Model`.

18. Charles Mueller. *Evidence for a standard model Higgs boson produced in association with a top quark pair and decaying to leptons*. PhD thesis, University of Notre Dame, October 2017.

19. CMS Collaboration. A measurement of the higgs boson mass in the diphoton decay channel. CMS Physics Analysis Summary CMS-PAS-HIG-19-004, 2019. URL `http://cdsweb.cern.ch/record/2691211`.

20. Cécile Caillol. *Scalar boson decays to tau leptons : in the standard model and beyond*. PhD thesis, Université libre de Bruxelles, April 2016.

21. LHC Higgs Cross Section Working Group. Handbook of LHC Higgs cross sections: 4. deciphering the nature of the Higgs sector. *CERN-CYRM-2017-002*, 2016. doi: 10.23731/CYRM-2017-002.

22. M. Carena et al. Status of Higgs boson physics, in Review of Particle Physics. Website. `http://pdg.lbl.gov/2016/reviews/rpp2016-rev-higgs-boson.pdf`.

23. LHC Higgs Cross Section Working Group. Higgs cross sections and decay branching ratios. Website. `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG`.

24. Nathan Kellams. *Search for lepton flavor violating decay of the Higgs boson in the* $\mu\tau_{hadronic}$ *channel*. PhD thesis, University of Notre Dame, September 2016.

25. J. Lorenzo Diaz-Cruz and J. J. Toscano. Lepton flavor violating decays of Higgs bosons beyond the standard model. *Phys. Rev. D*, 62:116005, 2000. doi: 10.1103/PhysRevD.62.116005.

26. James D. Bjorken and Steven Weinberg. Mechanism for nonconservation of muon number. *Phys. Rev. Lett.*, 38:622, 1977. doi: 10.1103/PhysRevLett.38.622.

27. Tao Han and Danny Marfatia. $h \to \mu\tau$ at hadron colliders. *Phys. Rev. Lett.*, 86:1442, 2001. doi: 10.1103/PhysRevLett.86.1442.

28. Ernesto Arganda, Ana M. Curiel, Maria J. Herrero, and David Temes. Lepton flavor violating Higgs boson decays from massive seesaw neutrinos. *Phys. Rev. D*, 71:035011, 2005. doi: 10.1103/PhysRevD.71.035011.

29. Abdesslam Arhrib, Yifan Cheng, and Otto C. W. Kong. Comprehensive analysis on lepton flavor violating Higgs boson to $\mu^{\mp}\tau^{\pm}$ decay in supersymmetry without R parity. *Phys. Rev. D*, 87:015025, 2013. doi: 10.1103/PhysRevD.87.015025.

30. M. Arana-Catania, E. Arganda, and M. J. Herrero. Non-decoupling SUSY in LFV Higgs decays: a window to new physics at the LHC. *JHEP*, 09:160, 2013. doi: 10.1007/JHEP09(2013)160.

31. E. Arganda, M. J. Herrero, R. Morales, and A. Szynkman. Analysis of the $h, H, A \to \tau\mu$ decays induced from SUSY loops within the Mass Insertion Approximation. *JHEP*, 03:055, 2016. doi: 10.1007/JHEP03(2016)055.

32. E. Arganda, M. J. Herrero, X. Marcano, and C. Weiland. Enhancement of the lepton flavor violating Higgs boson decay rates from SUSY loops in the inverse seesaw model. *Phys. Rev. D*, 93:055010, 2016. doi: 10.1103/PhysRevD.93.055010.

33. Kaustubh Agashe and Roberto Contino. Composite Higgs-mediated flavor-changing neutral current. *Phys. Rev. D*, 80:075016, 2009. doi: 10.1103/PhysRevD.80.075016.

34. Aleksandr Azatov, Manuel Toharia, and Lijun Zhu. Higgs mediated flavor changing neutral currents in warped extra dimensions. *Phys. Rev. D*, 80:035016, 2009. doi: 10.1103/PhysRevD.80.035016.

35. Gilad Perez and Lisa Randall. Natural neutrino masses and mixings from warped geometry. *JHEP*, 01:077, 2009. doi: 10.1088/1126-6708/2009/01/077.

36. S. Casagrande, F. Goertz, U. Haisch, M. Neubert, and T. Pfoh. Flavor physics in the Randall-Sundrum model I. theoretical setup and electroweak precision tests. *JHEP*, 10:094, 2008. doi: 10.1088/1126-6708/2008/10/094.

37. Andrzej J. Buras, Bjoern Duling, and Stefania Gori. The impact of Kaluza-Klein fermions on Standard Model fermion couplings in a RS model with custodial protection. *JHEP*, 09:076, 2009. doi: 10.1088/1126-6708/2009/09/076.

38. Monika Blanke, Andrzej J. Buras, Bjoern Duling, Stefania Gori, and Andreas Weiler. $\Delta F = 2$ observables and fine-tuning in a warped extra dimension with custodial protection . *JHEP*, 03:001, 2009. doi: 10.1088/1126-6708/2009/03/001.

39. Gian F. Giudice and Oleg Lebedev. Higgs-dependent Yukawa couplings. *Phys. Lett. B*, 665:79, 2008. doi: 10.1016/j.physletb.2008.05.062.

40. J. A. Aguilar-Saavedra. A minimal set of top-Higgs anomalous couplings. *Nucl. Phys. B*, 821:215, 2009. doi: 10.1016/j.nuclphysb.2009.06.022.

41. Michaela E. Albrecht, Monika Blanke, Andrzej J. Buras, Bjorn Duling, and Katrin Gemmler. Electroweak and flavour structure of a warped extra dimension with custodial protection. *JHEP*, 09:064, 2009. doi: 10.1088/1126-6708/2009/09/064.

42. Andreas Goudelis, Oleg Lebedev, and Jae Hyeon Park. Higgs-induced lepton flavor violation. *Phys. Lett. B*, 707:369, 2012. doi: 10.1016/j.physletb.2011.12.059.

43. David McKeen, Maxim Pospelov, and Adam Ritz. Modified Higgs branching ratios versus CP and lepton flavor violation. *Phys. Rev. D*, 86:113004, 2012. doi: 10.1103/PhysRevD.86.113004.

44. Apostolos Pilaftsis. Lepton flavour nonconservation in H^0 decays. *Phys. Lett. B*, 285:68, 1992. ISSN 0370-2693. doi: 10.1016/0370-2693(92)91301-O.

45. J. G. Körner, A. Pilaftsis, and K. Schilcher. Leptonic CP asymmetries in flavor-changing H^0 decays. *Phys. Rev. D*, 47:1080, 1993. doi: 10.1103/PhysRevD.47.1080.

46. E. Arganda, M. J. Herrero, X. Marcano, and C. Weiland. Imprints of massive inverse seesaw model neutrinos in lepton flavor violating Higgs boson decays. *Phys. Rev. D*, 91:015001, 2015. doi: 10.1103/PhysRevD.91.015001.

47. Hajime Ishimori, Tatsuo Kobayashi, Hiroshi Ohki, Yusuke Shimizu, Hiroshi Okada, and Morimitsu Tanimoto. Non-Abelian discrete symmetries in particle physics. *Prog. Theor. Phys. Suppl.*, 183:1, 2010. doi: 10.1143/PTPS.183.1.

48. CMS Collaboration. Search for lepton flavour violating decays of the Higgs boson to $\mathrm{e}\tau$ and $\mathrm{e}\mu$ in proton-proton collisions at $\sqrt{s} = 8$ TeV. *Phys. Lett. B*, 763:472, 2016. ISSN 0370-2693. doi: 10.1016/j.physletb.2016.09.062.

49. Georges Aad et al. Search for lepton-flavour-violating decays of the Higgs and Z bosons with the ATLAS detector. *Eur. Phys. J. C*, 77:70, 2017. doi: 10.1140/epjc/s10052-017-4624-0.

50. Georges Aad et al. Search for lepton-flavour-violating H $\rightarrow \mu\tau$ decays of the Higgs boson with the ATLAS detector. *JHEP*, 11:211, 2015. doi: 10.1007/JHEP11(2015)211.

51. Malte Buschmann, Joachim Kopp, Jia Liu, and Xiao-Ping Wang. New signatures of flavor violating Higgs couplings. *JHEP*, 06:149, 2016. doi: 10.1007/JHEP06(2016)149.

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

53. CERN. CERN accelerator complex. Website. `http://public-archive.web.cern.ch/public-archive/en/research/AccelComplex-en.html`.

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

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

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

57. A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12:P10003, 2017. doi: 10.1088/1748-0221/12/10/P10003.

58. Danek Kotliski. The CMS pixel detector. *Nucl. Instrum. Meth. A*, 465:46, June 2001. doi: 10.1016/S0168-9002(01)00345-X.

59. V Veszpremi. Operation and performance of the CMS tracker. *JINST*, March 2014.

60. CMS Collaboration. Silicon pixels. Website, 2011. `http://cms.web.cern.ch/news/silicon-pixels`.

61. Vardan Khachatryan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9, 2014. doi: 10.1088/1748-0221/9/10/P10009.

62. Cristina Biino. The CMS Electromagnetic Calorimeter: overview,lessons learned during Run 1 and future projections. *J. Phys.: Conf. Ser.*, 587, June 2015. doi: 10.1088/1742-6596/587/1/012001.

63. CMS Collaboration. Ecal crystal. Website. `https://www3.nd.edu/~cjessop/research/CMS/CMS.html`.

64. P Adzic et al. Energy resolution of the barrel of the CMS Electromagnetic Calorimeter. *JINST*, 2, April 2007.

65. David Petyt and the CMS collaboration. Mitigation of Anomalous APD signals in the CMS Electromagnetic Calorimeter. *J. Phys.: Conf. Ser.*, 404, August 2012.

66. CERN. The construction of the cms ecal: avalanche photodiodes (apds). Website, 2005. `http://cds.cern.ch/record/929105`.

67. Alexandre Zabi. Status of ECAL TPG and preparation for RUN II. *CMS ECAL Readiness Review*, 2015.

68. CMS Collaboration. Silicon pixels. Website, 2011. `http://cms.web.cern.ch/news/hadron-calorimeter`.

69. V Veszpremi. Performance of calorimeters at the LHC. *J. Phys.: Conf. Ser.*, 293, March 2011. doi: 10.1088/1742-6596/293/1/012001.

70. CMS Collaboration. CMS Physics : Technical Design Report Volume 1: Detector Performance and Software, Feb 2006.

71. CMS Collaboration. Muon detectors. Website, 2011. `http://cms.web.cern.ch/news/muon-detectors`.

72. CMS Collaboration. Superconducting magnet. Website, 2011. `http://cms.web.cern.ch/news/superconducting-magnet`.

73. Nabarun Dev et al. The cms level-1 electron and photon trigger: for run ii of lhc. *JINST*, 12, Feb 2017. doi: 10.1088/1748-0221/12/02/C02014.

74. Luca Cadamuro. The cms level-1 trigger system for lhc run ii. *JINST*, 12, Oct 2016. doi: 10.1088/1748-0221/12/03/C03021.

75. A. M. Sirunyan et al. Performance of the CMS muon detector and muon reconstruction with proton proton collisions at $\sqrt{s} = 13$ TeV. *JINST*, 13:P06015, 2018. doi: 10.1088/1748-0221/13/06/p06015.

76. CMS Muon Physics Object Group. Muon efficiency scale-Factors. Website. `https://twiki.cern.ch`.

77. Andy Buckley et al. General-purpose event generators for lhc physics. *Physics Reports*, 504:243, July 2011. doi: 10.1016/j.physrep.2011.03.005.

78. Wikipedia. Monte carlo method. Website, . `https://en.wikipedia.org/wiki/Monte_Carlo_method`.

79. F. Krauss et al. J. Alwall, S. Hche. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *EPJC*, 53, 2008.

80. Leif Lonnblad. Correcting the color dipole cascade model with fixed order matrix elements. *JHEP*, 0205, 2002.

81. Rikkert Frederix and Stefano Frixione. Merging meets matching in MC@NLO. *JHEP*, 12:061, 2012. doi: 10.1007/JHEP12(2012)061.

82. 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.

83. Johannes Bellm et al. Herwig 7.0/Herwig++ 3.0 release notes. *EPJC*, 76, Apr 2016. doi: 10.1140/epjc/s10052-016-4018-8.

84. 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.

85. 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.

86. 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.

87. 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.

88. 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.

89. 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.

90. J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014. doi: 10.1007/JHEP07(2014)079.

91. 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.

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

93. Gionata Luisoni, Paolo Nason, Carlo Oleari, and Francesco Tramontano. $HW^{\pm}/HZ$ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to

GoSam and their merging within MiNLO. *JHEP*, 10:083, 2013. doi: 10.1007/JHEP10(2013)083.

94. Rudolf Frhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth. A*, 262:444, 1987. doi: 10.1016/0168-9002(87)90887-4.

95. Rudolf Frhwirth Wolfgang Waltenberger and Pascal Vanlaer. Adaptive vertex fitting. *J. Phys. G*, 34, 2007.

96. Serguei Chatrchyan et al. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002, 2012. doi: 10.1088/1748-0221/7/10/P10002.

97. Vardan Khachatryan et al. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10: P06005, 2015. doi: 10.1088/1748-0221/10/06/P06005.

98. Vardan Khachatryan et al. Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS. *JINST*, 11:P01019, 2016. doi: 10.1088/1748-0221/11/01/P01019.

99. A. M. Sirunyan et al. Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV. *JINST*, 13:P10005, 2018. doi: 10.1088/1748-0221/13/10/p10005.

100. Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008. doi: 10.1088/1126-6708/2008/04/063.

101. Serguei Chatrchyan et al. Determination of jet energy calibration and transverse momentum resolution in CMS. *JINST*, 6:11002, 2011. doi: 10.1088/1748-0221/6/11/P11002.

102. Vardan Khachatryan et al. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST*, 12:P02014, 2017. doi: 10.1088/1748-0221/12/02/P02014.

103. Albert M Sirunyan et al. Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector. *JINST*, 14:P07004, 2019. doi: 10.1088/1748-0221/14/07/P07004.

104. Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The catchment area of jets. *JHEP*, 04:005, 2008. doi: 10.1088/1126-6708/2008/04/005.

105. Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys. Lett. B*, 659:119, 2008. doi: 10.1016/j.physletb.2007.09.077.

106. 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.

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

108. CMS Collaboration. Observation of the Higgs boson decay to a pair of τ leptons. *Phys. Lett. B*, 779:283, 2018. doi: 10.1016/j.physletb.2018.02.004.

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

110. CMS and ATLAS Collaborations. Procedure for the lhc higgs boson search combination in summer 2011. Technical report, August 2011.

111. Wikipedia. Poisson distribution. Website, . `https://en.wikipedia.org/wiki/Poisson_distribution`.

112. J. S. Conway. Incorporating nuisance parameters in likelihoods for multisource spectra. In *PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva*, January 2011.

113. A. L. Read. Linear interpolation of histograms. *Nucl. Instrum. Meth*, 425, April 1999.

114. A. L. Read. Presentation of search results: The $\mathrm{CL_s}$ technique. *Journal of Physics G*, 28, September 2002.

115. A. L. Read. Modified frequentist analysis of search results (The $\mathrm{CL_s}$ method). In *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000*, August 2000.

116. Thomas Junk. Confidence level computation for combining searches with small statistics. *Nuclear Instruments and Methods A*, 434, September 1999.

117. Eilam Gross Glen Cowan, Kyle Cranmer and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *European Physics Journal C*, 71, February 2011.

118. Albert M Sirunyan et al. Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV. *JHEP*, 07:161, 2018. doi: 10.1007/JHEP07(2018)161.

119. Roger Barlow and Christine Beeston. Fitting using finite monte carlo samples. *Comp. Phys. Comm.*, 77:219, 1993. doi: 10.1016/0010-4655(93)90005-W.

120. A. Denner, S. Heinemeyer, I. Puljak, D. Rebuzzi, and M. Spira. Standard model Higgs-boson branching ratios with uncertainties. *Eur. Phys. J. C*, 71:1753, 2011. doi: 10.1140/epjc/s10052-011-1753-8.

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