Search for di-Higgs production with CMS data at $\sqrt{s}$ = 13 TeV at the LHC

By

Saswati Nandan

PHYS05201304014

Saha Institute of Nuclear Physics

A thesis submitted to the

Board of Studies in Physical Sciences

In partial fulfillment of requirements

for the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE

July, 2019

Homi Bhabha National Institute[1]

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by **Saswati Nandan** entitled **"Search for di-Higgs production with CMS data at √s = 13 TeV at the LHC"** and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Chairman – Prof. Sukalyan Chattopadhyay	8/1/2020
Guide / Convener – Prof. Subir Sarkar	Subir Sarkar 8/1/2020
Co-guide - Prof. Sunanda Banerjee	Sunanda Banerjee 8.1.20
Examiner - Prof. Rohini Godbole	8.1.2020
Member 1- Prof. Satyajit Saha	8/1/2020
Member 2- Prof. Supratik Mukhopadhyay	8/1/2020
Member 3- Prof. Tapan Kumar Nayak	Tapan 8.1.2020

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.

I/We hereby certify that I/we have read this thesis prepared under my/our direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 8/1/2020

Place: Kolkata

Sunanda Banerjee
Signature
Co-guide (if any)

Subir Sarkar
Signature
Guide

[1] This page is to be included only for final submission after successful completion of viva voce.

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Saswati Nandan

Saswati Nandan

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

Saswati Nandan

Saswati Nandan

List of Publications arising from the thesis

Journal

1. "Search for Higgs boson pair production in the $\gamma\gamma bb$ final state in pp collisions at $\sqrt{s}$ = 13 TeV", CMS Collaboration, Phys. Lett. B 788 (2018) 7.

Conferences

1. "di-Higgs search in 4tau final state" at "Double Higgs production at Colliders", 4-8 Sep 2018, Batavia, IL, United States.
2. "di-Higgs search in 4tau final state" at "3rd CMS HH workshop", 4-6 April 2018, LLR, France.

Others

1. CMS Collaboration, "Calibration of the CMS Hadron Calorimeter with collision Data collected in 2016" is submitted to JINST.
2. "Separation of di-Higgs events produced through gluon-gluon and vector boson fusion processes in proton-proton collision at $\sqrt{s}$ = 14 TeV", Saswati Nandan, Sunanda Banerjee, Subir Sarkar, Submitted to Pramana.

Saswati Nandan

Saswati Nandan

DEDICATIONS

To my mother

Acknowledgments

I would like to express my sincere gratitude to my supervisor Prof. Subir Sarkar and co-supervisor Prof. Sunanda Banerjee for giving me the opportunity to take part in such a huge experiment performed at CMS. I am grateful to Prof. Subir Sarkar for allowing me to work independently and taking special classes on C++ and CMS software during my initial phase of PhD. I owe to Prof. Sunanda Banerjee for giving me chance to visit Fermi Lab where I could start a new analysis.

I want to express my special thanks to Edward Laird for suggesting a new analysis. I am indebted to Abdollah Mohamaddi for his valuable guidance to finish the analysis. I thank Prof. Satyaki Bhattacharya for discussing different aspects of statistics. I am grateful to Maxime Gouzevitch for giving me chance to work in their analysis. I would like to take this opportunity to thank my senior colleague Debarati Roy who helped me to understand the work on Hcal Calibration. I will cherish the moments which I spent with my friends Ankita and Pallabi during my CERN visit. I have been fortunate to have friends like Priyanka, Sudipta, Rakhi Di, Maireyee, Vijay who were always supportive.

I would like to take the privilege to thank my school teacher Sunil Bhattacharya for his special care in Science. I am indebted to my Science teacher Jayanta Gangopadhyay who helped during my difficult situations and always motivated to pursue higher studies. I would like to thank Prof. Amitabha Mukhopadhyay for his precious teachings in M.Sc.

I am grateful to other people whom I met during these years and without their incredible support outside the research work, I would not be able to finish my work. I would like to show my earnest gratitude to Lipy Di, Sabyasachi Da, Dipankar Da, Dijen Da, Arindam Da, Anindya Da, who always stand by me whenever I was in need. I am indebted to my uncles who took care of my ailing mother when I was abroad.

One page is not enough to thank everyone, I apologize to those whose names I could not mention but I express my deepest gratitude to everyone for their priceless support, encouragement and standing by me all the time.

Contents

List of Figures

List of Tables

Summary

A search for HH production in the $HH \to \tau\tau\tau\tau$ final state is presented in the thesis in the context of BSM physics where a spin-0 Radion (X) decays to a Higgs boson pair. The HH production cross section in the SM is very small. Nevertheless, different BSM theories consider either anomalous Higgs self-coupling or presence of a resonant particle that decays to HH, that may lead to a significant enhancement of the production cross section.

This analysis is performed using pp collision data recorded by CMS in 2016, which correspond to 35.9 fb^{-1}. The final states of $HH \to \tau\tau\tau\tau$ are selected to achieve the highest S/B value. In this analysis, events where each of the Higgs bosons decays to a pair of tau leptons and subsequently two tau leptons decay leptonically and other two decay hadronically, are considered. Though the Higgs decay branching ratio to the tau lepton pair is small, e.g $\sim 6\%$, the absence of any b-jets in the final state reduces the background contamination. These final states were further divided into opposite-signed (OS) and same-signed (SS) channels based on leptonic decay of two OS or SS taus. This was done keeping in mind that in the SS channels, the background contribution will be smaller compared to the OS channels. Background contributions are estimated from MC simulated events, except for the multijet contribution. The multijet contribution, where jets fake τ_h, has been estimated using a data-driven technique. The contribution of multijet in the signal region is extrapolated from the τ_h anti-isolated region where one or two τ_hs are fake.

The fake rate estimation method is validated in a region which is away from the signal search region and good agreement is seen between the estimated and observed numbers. No excess in data over the expected SM background contribution is observed in the signal region. The 95% CL upper limits on $\sigma(pp \to X) \times BR(HH \to \tau\tau\tau\tau)$ has been calculated where the observed limit is consistent with the expected one. The result has also been compared with the other major search channels, namely, $b\bar{b}b\bar{b}, bb\gamma\gamma, bb\tau\tau, bbWW$. It has been observed that the channel under consideration can be competitive with $bbWW$ at low mass. The sensitivity of the search is limited by statistics. With a lot more data expected after LHC Run-III and eventually during the High-Luminosity LHC running period, improved analysis techniques to separate signal and background and more robust background estimation, the sensitivity of the search is expected to improve significantly.

A calibration procedure for the CMS HCAL and estimation of calibration constants using 2016 data has also been described in the thesis. Reliable estimate of the calibration constants plays an important role in the measurement of energy of hadrons, jets as well as missing transverse energy with high precision. Two different methods, namely iterative and method of moments, have been discussed. In both the methods, either the mean energy or the mean variance is equalized in all the segments along ϕ direction for a particular η segment utilizing the fact that the HCAL detector is symmetric in ϕ direction and also the energy deposition from minimum bias events is ϕ dependent.

A simulation study using helicity angles sensitive to HH production to distinguish vector boson fusion process from gluon-gluon fusion process and use of the helicity angles in CMS search for $HH \to b\bar{b}\gamma\gamma$ has been presented.

Chapter 1

Introduction

Following the discovery of the Higgs boson by both A Toroidal LHC Apparatus (ATLAS) and Compact Muon Solenoid (CMS) experiments in 2012, search for Higgs boson pair (HH) production has become the main focus at the Large Hadron Collider (LHC). The HH production cross section is small in the Standard Model (SM). Different Beyond Standard Model (BSM) Physics consider either anomalous Higgs self-coupling and top Yukawa coupling or presence of a resonant particle decaying to HH, which may lead to enhanced production cross section of HH. In this thesis, search for resonant HH production involving the decay of a spin-0 Radion (X), where each Higgs boson decays to a pair of tau leptons and subsequently two tau leptons decay leptonically and other two decay hadronically, has been studied for the first time in CMS. gluon-gluon fusion (ggF) is the most dominant process of HH production, nevertheless, it is also worthwhile to study the vector boson fusion (VBF) process, which has a cross section 10 times smaller than ggF, but has a better signal-over-background ratio due to the presence of two forward jets.

Chapter 2 describes the Standard Model of Particle Physics, discovery of the Higgs boson and HH production in SM and BSM. The CMS detector and estimation of calibration constants of the CMS Hadron Calorimeter are discussed in **Chapter 3** and **Chapter 4** respectively. Reconstruction of primary vertices, muons, electrons, tau leptons, jets in CMS is discussed in **Chapter 5**. **Chapter 6** presents a study of how ggf and VBF processes can be separated using helicity angles and other variables using a multivariate technique. A detailed description of the analysis of resonant HH search in the BSM scenario is presented in **Chapter 7**. The final chapter presents a summary of the work done.

Chapter 2

The Standard Model of Particle Physics

The Standard Model (SM) of particle physics describes three of the four known fundamental forces, namely, Strong, Electromagnetic and Weak interactions. The SM does not include the Gravitational force. Each fundamental force is mediated by different vector bosons. For example, Strong force is mediated by massless spin one gluons, Electromagnetic force is mediated by massless spin one photons, spin one massive W and Z bosons are mediators of Weak force, whereas quantum gravity is the theoretical framework that attempts to describe the gravitational force where a hypothetical spin-two graviton is considered as the mediator of the force. Some characteristics of all the forces are described in Table 1.1

Table 2.1: Range, relative strength with respect to strong force and mediators of four fundamental forces.

Interaction	Range	Relative strength	Mediators
Strong	10^{-15} m	1	gluons
Electromagnetic	∞	10^{-3}	photons
Weak	10^{-18} m	10^{-14}	$W^{\pm}$, Z
Gravitational	∞	10^{-43}	gravitons

The elementary particles are either bosons or fermions depending on their spin. Particles of half-integer spin are called fermions and follow Fermi-Dirac statistics. On the other hand, particles of integer spin are bosons which follow Bose-Einstein statistics. The Standard Model of particle physics contains 12 elementary fermions, plus their corresponding antiparticles, elementary bosons that mediate the forces as described earlier and the Higgs boson. The 12 fundamental fermions are divided into 3 generations of 4 particles each. Half of the fermions are leptons, three of which have an electric charge of -1, called electron(e^-), muon (μ^-) and tau (τ^-). The other three leptons are neutrinos (ν_e,ν_μ,ν_τ) which are the only elementary fermions with no electric or color charge. The rest of the fermions are quarks : up, down, charm, strange, top and bottom. Each type of quarks appears in three different color quantum numbers. Some of the properties of all the elementary particles are presented in Figure 1.1.

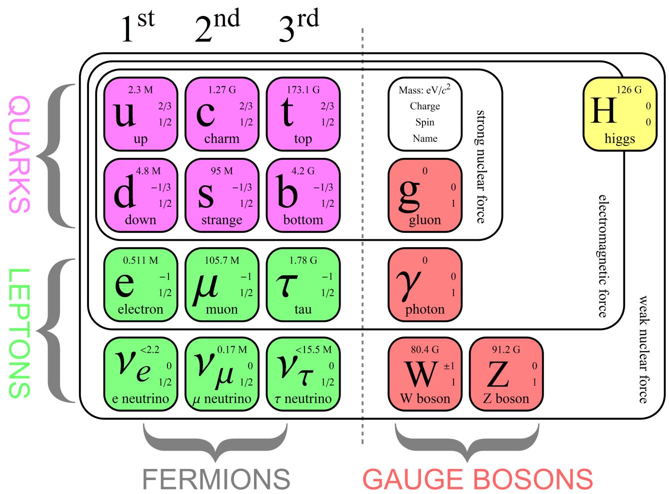

Figure 2.1: Elementary particles included in the Standard Model.

2.1 The Standard Model Lagrangian

The Standard Model is described by Quantum Field Theory (QFT) where the dynamics of every field, which is a function of space and time, is described by a Lagrangian. The Lagrangian should be invariant under Lorentz transformation as well as space-time translation. To keep the Lagrangian of a given field invariant under a local transformation, which is space-time dependent, a gauge transformation is performed where a new field is introduced in the original Lagrangian. The new field is called gauge field and the corresponding theory is known as Gauge theory.

2.1.1 $U(1)$ QED Theory

The Lagrangian of a free Dirac field is described by

$$\mathcal{L} = \overline{\psi}(i \not{\partial} - m)\psi \tag{2.1.1.1}$$

The above Lagrangian is invariant under the global transformation

$$\psi \rightarrow \psi' = e^{-ieQ\theta}\psi \tag{2.1.1.2}$$

where θ is independent of space-time. However, under a local transformation $\psi \rightarrow \psi' = e^{-ieQ\theta(x)}\psi$, where θ is a function of space-time (x), the Lagrangian is no more invariant. To keep the Lagrangian invariant under a local transformation, a covariant derivative is introduced as,

$$D_\mu = \partial_\mu + ieQA_\mu \tag{2.1.1.3}$$

where A_μ transforms in the following way

$$A'_\mu = A_\mu + \partial_\mu \theta(x) \tag{2.1.1.4}$$

A_μ is the gauge field and the quantum of the corresponding field is photon. The full

Lagrangian for the ψ and A_μ field becomes

$$\mathcal{L} = \underbrace{\overline{\psi}(i\not{\partial} - m)\psi}_{\text{free}} - \frac{1}{4}\underbrace{F_{\mu\nu}F^{\mu\nu}}_{\text{kinetic term}} - \underbrace{eQ\overline{\psi}\gamma^\mu\psi A_\mu}_{\text{interaction term}} \qquad (2.1.1.5)$$

where the field tensor, $F_{\mu\nu}$ is defined by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \qquad (2.1.1.6)$$

The field theory with the Lagrangian as mentioned above is called Quantum Electrodynamics (QED) which possesses U(1) gauge symmetry.

2.1.2 $SU(3)$ QCD Theory

As in the previous case, the free Lagrangian for N Dirac fields having the same mass m is given by

$$\mathcal{L} = \Sigma_{\mathrm{k}}\overline{q}_k(i\not{\partial} - m)q_k \qquad (2.1.2.1)$$

For example, in QCD k can represent the 3 color indices of the quarks. Like QED, the above Lagrangian is invariant under global transformation

$$\Psi \rightarrow \Psi' = U\Psi \qquad (2.1.2.2)$$

where, $\Psi = \begin{pmatrix} q_1 \\ q_2 \\ . \\ . \\ . \\ q_{d_R} \end{pmatrix}$ and U, which is space-time independent, is a matrix that represents a group element in the representation R of SU(N) and d_R is the dimensionality of that representation R. To keep the Lagrangian invariant under a local transformation where U is function of space-time, the covariant derivative is defined as follows

$$D_\mu = \partial_\mu + igT_a A_\mu^a \tag{2.1.2.3}$$

where g represents the strength of the interaction and T_a is the generator of the group. These generators are non-abelian and follow the following algebra

$$[\,T_a, T_b] \;= if_{abc}T_c \tag{2.1.2.4}$$

The gauge field A_μ^a transforms in the following way

$$A_\mu^{a\prime} = A_\mu^a + \partial_\mu\theta^a(x) + gf_{abc}\theta^b(x)A_\mu^c \tag{2.1.2.5}$$

The full gauge invariant Lagrangian is given by

$$\mathcal{L} = \underbrace{\overline{q}(i\not{\partial} - m)q}_{\text{free}} - \frac{1}{4}\underbrace{G_{\mu\nu}^a G_a^{\mu\nu}}_{\text{kinetic term}} - \underbrace{g\bar{q}\gamma^\mu T_a q A_\mu^a}_{\text{interaction term}} \tag{2.1.2.6}$$

where $G_{\mu\nu}^a$ transforms as

$$G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - gf_{abc}A_\mu^b A_\nu^c \tag{2.1.2.7}$$

Unlike QED, here $G_{\mu\nu}^a$ contains quadratic terms of gauge fields due to the non-abelian nature of the generators. This means that the gauge bosons, *i.e.* gluons in this case, are self interacting and carry color quantum number unlike the photon. This field theory is known as Yang-Mills Theory, after the names of the two physicists who first described such theories. There are 8 different gluons which are mediators of Strong Interaction.

2.1.3 $SU(2)_L \otimes U(1)_Y$ Electroweak Theory

While only the left handed neutrino (right handed anti-neutrino) is observed experimentally, electrons, which are massive, can be observed both as left handed as well as right handed. The left handed neutrino and the left handed electron can form a doublet under $SU(2)$ whereas the right handed electron behaves as a $U(1)$ singlet, *i.e* the theory is defined by $SU(2)_L \otimes U(1)_Y$ representation where L defines the left chirality and Y is the weak hypercharge, associated with the $U(1)$ part. $SU(2)$ has three generators while $U(1)$ has one generator resulting in a total of 4 gauge bosons which are defined by W_μ^+, W_μ^-, W_μ^3 for the $SU(2)$ group and B_μ for the $U(1)$ group. The gauge invariant Lagrangian is given by

$$\mathcal{L} = (D^\mu \psi)^\dagger (D_\mu \psi) - \frac{1}{4} W^a_{\mu\nu} W_a^{\dagger\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \tag{2.1.3.1}$$

where

$$\begin{aligned} D_\mu &= \partial_\mu + ig\frac{\tau_a}{2} W^a_\mu + ig' Y B_\mu \\ W^a_{\mu\nu} &= \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g f_{abc} W^b_\mu W^c_\nu \end{aligned} \tag{2.1.3.2}$$

where $\tau' s$ are the Pauli matrices : $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. These fields are similar to Yang-Mills fields, *i.e.* the bosons are self interacting. g and g' are the coupling constants for the $SU(2)$ and $U(1)$ groups respectively. As $B_{\mu\nu}$ corresponds to the $U(1)$ group, it is similar to the field tensor introduced in Electrodynamics, *i.e.*

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.1.3.3}$$

The interaction terms for the leptonic field with hypercharge value $Y = -\frac{1}{2}$, is given by

$$\mathcal{L}_{int} = -\frac{1}{2}\begin{pmatrix}\overline{\nu}_L & \overline{e}_L\end{pmatrix}\gamma^\mu \begin{pmatrix} gW^3_\mu - g'B_\mu & g(W^1_\mu - iW^2_\mu) \\ g(W^1_\mu + iW^2_\mu) & -gW^3_\mu - g'B_\mu \end{pmatrix}\begin{pmatrix}\nu_L \\ e_L\end{pmatrix} + \overline{e}_R g'\gamma^\mu B_\mu e_R \quad (2.1.3.4)$$

The $\mathcal{L}$ can be decomposed into charged and neutral currents in the following way:

$$\begin{aligned}\mathcal{L}_{cc} &= -\frac{g}{\sqrt{2}}(\overline{\nu}_L\gamma^\mu W^+_\mu e_L + \overline{e}_L\gamma^\mu W^-_\mu \nu_L) \\ \mathcal{L}_{nc} &= -\frac{1}{2}\overline{\nu}_L\gamma^\mu\nu_L(gW^3_\mu - g'B_\mu) + \frac{1}{2}\overline{e}_L\gamma^\mu e_L(gW^3_\mu + g'B_\mu) + \overline{e}_R\gamma^\mu g'B_\mu e_R\end{aligned} \quad (2.1.3.5)$$

where

$$W^\mp_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \pm iW^2_\mu) \quad (2.1.3.6)$$

Since ${W_\mu}^3$ and B_μ interact with neutral particles, they cannot be the photon field which has non-trivial interactions only with particles with non-zero electromagnetic charge and hence should not have any interactions with the neutrinos. They can be expressed by a linear combination of physical fields Z_μ (Z boson) and A_μ (photon) in following way:

$$\begin{aligned} Z_\mu = &\ \cos\theta_W W^3_\mu - \sin\theta_W B_\mu \\ A_\mu = &\ \sin\theta_W W^3_\mu + \cos\theta_W B_\mu \end{aligned} \quad (2.1.3.7)$$

where θ_W, called Weinberg angle, is defined by the relation

$$\tan\theta_W = \frac{g'}{g} \quad (2.1.3.8)$$

Using the above relation and identities involving the left (L) and right (R) projection operators $i.e$ $L + R = 1$, neutral current can be written as

$$\mathcal{L}_{nc} = -\frac{g}{2\cos\theta_W}\overline{\nu}\gamma^\mu\ L\nu Z_\mu + \frac{g}{2\cos\theta_W}\overline{e}\gamma^\mu(L - 2\sin^2\theta_W)eZ_\mu + g'\cos\theta_W\overline{e}\gamma^\mu e A_\mu \quad (2.1.3.9)$$

The last term is equivalent to the QED interaction between electron and photon if

$$e = g' \cos\theta_W = g \sin\theta_W \tag{2.1.3.10}$$

2.2 Spontaneous Symmetry Breaking

If the Lagrangian of a system is invariant under certain transformation whereas the ground state of the system is not, the associated symmetry is known to be spontaneously broken. In QFT, the spontaneous symmetry breaking (SSB) occurs when there exists a symmetric Lagrangian with degenerate and non-symmetric vacuum states. In case of the Electroweak theory, SSB can be described with the following Lagrangian of a real scalar field ϕ:

$$\mathcal{L} = T - V = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - (\frac{1}{2}\mu^2\phi^2 + \frac{\lambda}{4}\phi^4) \tag{2.2.0.11}$$

which is symmetric under the transformation $\phi \to -\phi$. The value of λ is always positive; otherwise it would not have any ground state. But there are two possibilities in the value of μ^2 as shown in Figure 1.2. If $\mu^2 > 0$, the first term in the potential represents the

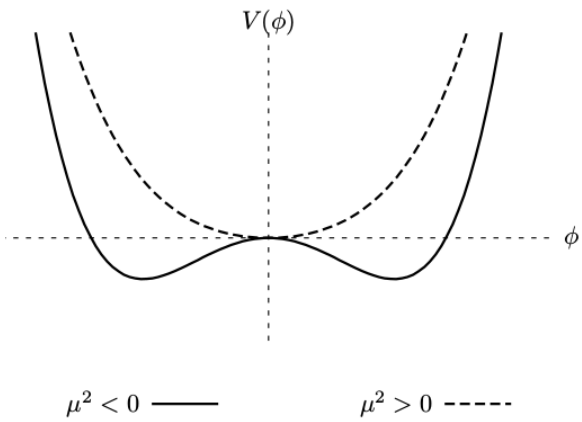

Figure 2.2: Shape of the potential of Eq. (1.2.0.11).

mass term with $m = \sqrt{\mu}$. If $\mu^2 < 0$, the mass term gets a wrong sign. The minimum of potential occurs when

$$\begin{aligned} \frac{\partial V}{\partial \phi} &= 0 \\ \rightarrow \phi(\mu^2 + \lambda\phi^2) &= 0 \end{aligned} \tag{2.2.0.12}$$

The solutions of the above equation are

$$< \phi >= 0, \pm v \tag{2.2.0.13}$$

with

$$v = \sqrt{\frac{-\mu^2}{\lambda}} \tag{2.2.0.14}$$

which gives three real solutions. With the field at $+v$ and small fluctuation of the field around the minimum, the total field becomes

$$\phi(x) = v + H(x) \tag{2.2.0.15}$$

Using this modified field, the Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}(\partial_\mu H)(\partial^\mu H) - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4} H^4 \tag{2.2.0.16}$$

apart from the term containing v^4. This equation shows that the field H has a mass term M_H given by

$$M_H^2 = 2\lambda v^2 \tag{2.2.0.17}$$

The modified Lagrangian is no more symmetric under the transformation $H \rightarrow -H$. This is because as soon as one of the two minima is chosen, the symmetry is lost. This phenomenon is known as spontaneous symmetry breaking as no external force is responsible for breaking the symmetry, the Lagrangian itself leads to a situation where symmetry is not obeyed in physical processes.

2.3 The Higgs Field

The Higgs mechanism is the way the SM generates the mass terms for the massive fundamental particles. A scalar field is introduced in the SM as,

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.3.0.18}$$

Like earlier, the field potential has the form $V = \mu^2 \phi \phi^\dagger + \lambda(\phi^\dagger \phi)^2$ with $\lambda > 0$ and $\mu^2 < 0$. The field can be developed around one of its degenerate minima as

$$\Phi = \frac{1}{\sqrt{2}} exp(i\sigma.\frac{\theta(x)}{v}) \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{2.3.0.19}$$

This gives one massive filed $H(x)$ and three massless fields θ_i known as Goldstone bosons. These massless bosons can be made to disappear by the phase rotation, which is equivalent to gauge transformation, $exp(-i\sigma.\frac{\theta(x)}{v})$ of Φ, leaving only one massive field $H(x)$ and the quantum of this massive field is known as the Higgs (H) boson. This is known as Unitary

Gauge. The $\Phi(x)$ and the corresponding Lagrangian of the system becomes

$$\begin{aligned}\mathcal{L} &= (D_\mu\Phi(x))^\dagger(D^\mu\Phi(x)) - (\mu^2\Phi(x)^\dagger\Phi(x) + \lambda(\Phi^\dagger\Phi)^2) \\ &= \left[\frac{1}{2}\partial_\mu H\partial^\mu H - \frac{1}{2}(2\lambda v^2)H^2\right] + (\frac{1}{4}(gv)^2W^+W^- + \frac{1}{8}v^2(g^2+g'^2)Z^\mu Z_\mu)(1+\frac{H}{v})^2 \\ &- \lambda v^2H^2 - \lambda vH^3 - \frac{\lambda}{4}H^4\end{aligned} \tag{2.3.0.20}$$

The first term represents the dynamics of the Higgs field with mass $M_H = \sqrt{2\lambda v^2}$, where v is the vacuum expectation value (VEV) whose value is 246 GeV. The second term represents the mass of the weak bosons with:

$$\begin{aligned} M_W &= \frac{gv}{2} \\ M_Z &= \frac{v\sqrt{g^2+g'^2}}{2} = \frac{M_W}{\cos\theta_W}\end{aligned} \tag{2.3.0.21}$$

It can be seen that the 3 Goldstone bosons have been absorbed by the 3 gauge fields along their longitudinal polarization. The second term also shows the interaction of Higgs boson with weak bosons like HWW, HZZ, $HHWW$, $HHZZ$. The other terms represent the Higgs self-coupling

$$V(H) = \frac{1}{2}M_H^2H^2 + \lambda_{HHH}vH^3 + \frac{1}{4}\lambda_{HHHH}H^4 \tag{2.3.0.22}$$

with

$$\lambda_{HHH} = \lambda_{HHHH} = \lambda = \frac{M_H^2}{2v^2} \tag{2.3.0.23}$$

The orthogonal combination of Z_μ *i.e* A_μ still remains massless and can be identified with the photon. The process of symmetry breaking can then be summarized into the statement

$$SU(2)_L \times U(1)_Y \rightarrow U(1)_{em} \tag{2.3.0.24}$$

2.3.1 Yukawa interaction

The interaction involving fermions and scalar boson is known as Yukawa interaction. The Lagrangian density of the Yukawa interaction is given by

$$\mathcal{L}_{\mathcal{Y}} = -h(\overline{\psi}_{\ell L}\phi\ell_R + \overline{\ell}_R\phi^\dagger\psi_{\ell L}) \tag{2.3.1.1}$$

where $\psi_{\ell L}$ and ℓ_R are the left and right chiral components of the fermion and h is the strength of the Yukawa interaction. After spontaneous symmetry breaking and with the value of $\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}$, the Lagrangian becomes

$$\begin{aligned} \mathcal{L}_{\mathcal{Y}} &= -\frac{h}{\sqrt{2}}(v(\overline{\ell}_L\ell_R + \overline{\ell}_R\ell_L) + H(\overline{\ell}_R\ell_L + \overline{\ell}_L\ell_R)) \\ &= -(m_\ell\overline{\ell}\ell + \frac{m_\ell}{v}\overline{\ell}\ell H) \end{aligned} \tag{2.3.1.2}$$

The mass of a fermionic field is given by

$$m_\ell = \frac{hv}{\sqrt{2}} \tag{2.3.1.3}$$

In the SM, fermion masses are thus explained as arising from the interaction of fermionic fields with the Higgs field. The strengths of the interactions are directly related to the fermion masses, and are free parameters of the theory. Since there is no right-chiral component of the neutrino, the neutrino is massless in the SM.

2.4 Higgs discovery at the LHC

At the LHC, through proton-proton collision at a center of mass energy of 13 TeV, the Higgs boson can be produced in the following ways :

- **gluon-gluon fusion (ggF)** : The production process indicated in Figure 1.3.a involves heavy quark loop. The cross section of this process is $\sim$43.92 pb [?].
- **Vector Boson Fusion (VBF)** : It is the next dominant production, shown in Figure 1.3.b, with cross section $\sim$ 3.75 pb [?]. In this process, two vector bosons radiated by incoming quarks produce the Higgs boson while the two incoming quarks give rise to two high p_T jets in the forward direction. The presence of two high p_T jets in the forward direction gives an additional handle to separate background from the signal.
- **Associated production** : In this process as shown in Figure 1.3.c, a virtual boson splits into a real boson and the Higgs boson. The cross sections are $\sim$1.38 pb for WH and 0.87 pb [?] for ZH processes respectively.
- $t\bar{t}H$ **production** : In this process as shown in Figure 1.3.d, the Higgs boson is produced in association with a pair of top-antitop quarks. Though the cross section (0.51 pb [?]) of this process is small compared to other processes, nevertheless, it is important to study as the process involves direct coupling of the Higgs boson to the top quark.

Figure 1.3 shows the Feynman diagrams of Higgs boson production.

The cross section of a given process in pp collision, with momentum P and P' of two incoming protons, is given as

$$\sigma(p(P)p'(P')) = \Sigma_{\text{partons i,j}} \int_0^1 dx dx' \hat{\sigma_{i,j}}(xP, x'P')\phi_p^i \phi_{p'}^j \tag{2.4.0.4}$$

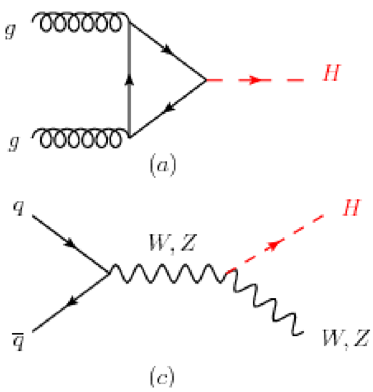

Figure 2.3: Feynman diagrams of Higgs production in (a) ggF, (b) VBF, (c) associated production and (d) $t\bar{t}H$ production.

where $\hat{\sigma_{ij}}$ is the cross section for the scattering of partons i and j to produce the chosen final state and $\phi^i_p, \phi^j_{p'}$ are the probability density for finding partons i and j in the incoming particles p and p' with momentum xP and $x'P'$ respectively, where $0 <= x <= 1$.

The cross sections of different processes as well as the Higgs boson decay branching ratios are shown in Figure 1.4. The maximum branching ratio (BR) of Higgs at mass 125 GeV is in $b\bar{b}$ channel ($\sim$ 58%) followed by the WW^* decay ($\sim$ 21%).

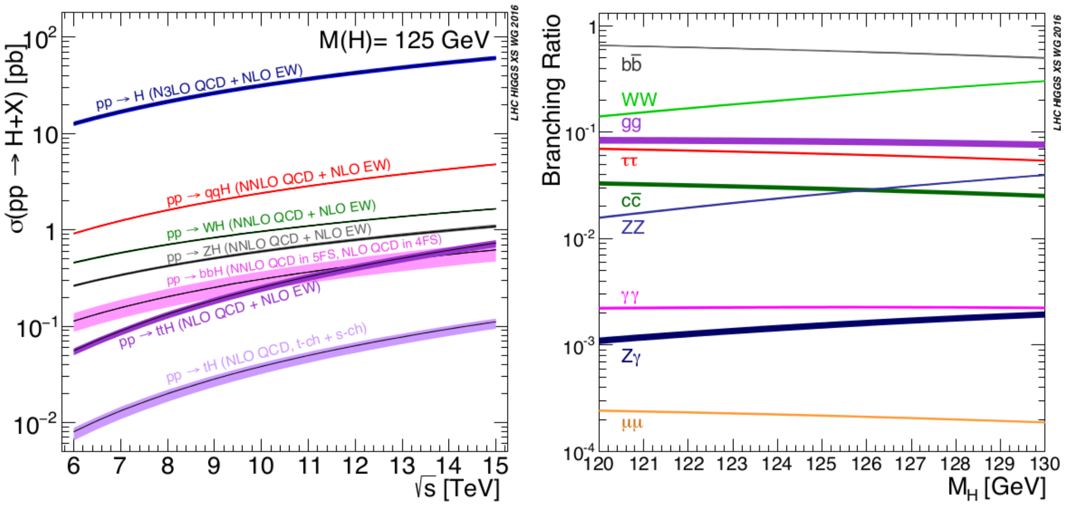

Figure 2.4: Production cross section of the Higgs boson in different channels as a function of $\sqrt{s}$ (left plot) and Branching ratio of the Higgs boson as a function of Higgs boson mass, m_H (right plot) [?].

The Higgs boson production was observed for the first time by both CMS and ATLAS experiments in 2012 using approximately 5 fb^{-1} of 7 TeV and 8 TeV data. An inclusive search was made in $\gamma\gamma$, ZZ^*, WW^*, $b\bar{b}$, $\tau\tau$ channels where an excess of data over SM background is observed in $\gamma\gamma$ [?] and ZZ^* [?] channels. Though the BR of $H \to \gamma\gamma$ is very small, only ∼0.2%, the channel provides a clean final-state topology and also the di-photon mass can be reconstructed with very high precision. Similarly, BR for $H \to ZZ^*$ is only $\sim$ 2.6%, but the presence of four high p_T, isolated leptons in the final state provides a clean signature providing a high signal-to-background (S/B) ratio. The combined measured mass of the Higgs boson is

$$m_H = 125.09 \pm 0.21(stat.) \pm 0.11(syst.)GeV \text{ [?]} \tag{2.4.0.5}$$

2.5 HH production

After discovering the Higgs boson, one of the most important properties to study is the Higgs self-coupling. With the Higgs mass now known with precision, the value of its self-coupling can be computed from Eq. 1.3.0.23 which gives $\lambda = \frac{M_H^2}{2v^2} \approx 0.12$. Experimentally, λ_{HHH} can be probed via di-Higgs production at the LHC. Similarly, with triple Higgs final state, λ_{HHHH} can be studied, although the cross section of this process is small and currently out of reach at the LHC. Experimentally, measuring λ_{HHH} will provide a test of the validity of the SM. Any deviation from the predicted SM value would indicate presence of physics beyond Standard Model (BSM). In the SM, Higgs boson pair can be produced in the following ways:

- **ggF**: The cross section of this dominant HH production at $\sqrt{s} = 13$ TeV is 0.31 pb [?]. In this process the Higgs boson is produced either by trilinear Higgs boson self-coupling or two on shell Higgs bosons are radiated from a heavy quark loop. This process thus involves trilinear Higgs coupling, λ_{HHH} as well as Yukawa coupling y_t as shown in the left and right figure of Figure 1.5 respectively.

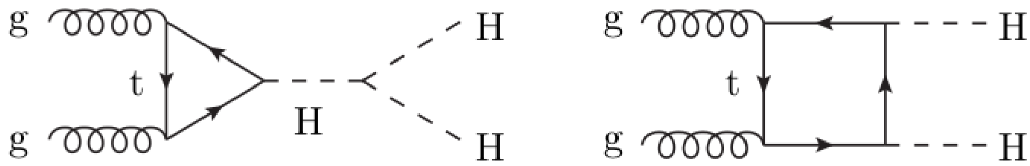

Figure 2.5: HH production in gluon-gluon fusion process.

- **VBF** : In this process shown in the Figure 1.6, two Higgs bosons are produced through vector boson fusion in addition to two high p_T jets in the forward direction. The cross section of this process is 10 times smaller than the ggF process but the presence of two additional jets in the forward direction provides an additional power to separate signal from the background processes.

With the presently available total integrated luminosity and HH identification

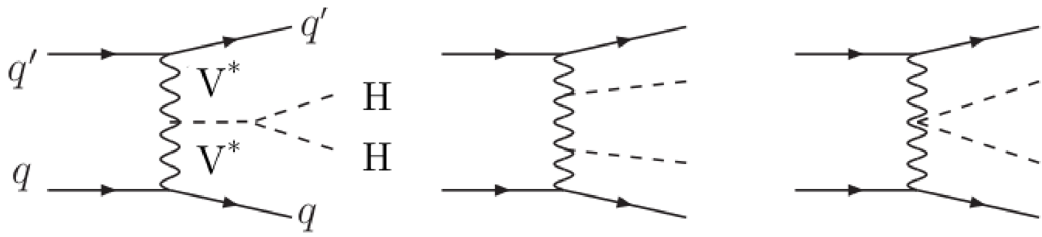

Figure 2.6: HH production in Vector boson fusion process.

efficiency, the absolute number of VBF events that can be tagged at the LHC, is insignificant. At the HL-LHC, however, with an estimated factor of 20 increase in total integrated luminosity, a number of VBF events are expected to be tagged with 3000 fb^{-1} which, aided by improved analysis techniques, will allow us to study VBF HH production modes for which the S/B ratio is better than the ggF process. There is also a small probability where, due to initial state radiation, ggF events may be selected as VBF. The event selection of these two independent processes should be orthogonal to each other to avoid the double counting of the same events in the two independent processes. In this thesis, how the VBF contribution can be distinguished from the contribution of the ggF process is discussed.

2.6 Physics Beyond Standard Model (BSM)

The SM has been very successful in explaining most of the phenomena pertaining to the elementary particle interactions. In spite of this great success, the SM, however, suffers from a number of shortcomings from both theoretical and observational point of views,

- The present version of the SM of particle physics is not considered as the fundamental theory of nature as it does not include the gravitational interaction.

- In the SM, the Higgs boson mass is not protected by any symmetry. When quantum corrections to the Higgs boson mass are taken into account, the correction becomes quadratically divergent [?]. As a consequence, if the SM is assumed to be valid up to the Planck scale $M_{Pl} \sim 10^{19}$ GeV, the Higgs boson mass would also be of order M_{Pl}, which is in clear contradiction with observation. An unnatural fine-tuning at all orders of the perturbation theory is the only way to keep the Higgs boson mass at the observed value.

- In the SM fermion contents, due to the absence of right-handed neutrinos one cannot generate masses for neutrinos. However, data from various neutrino oscillation (solar and atmospheric) experiments , indicate that neutrinos of the SM have tiny masses ($\mathcal{O}$(eV)) and also establish the mixing among the various flavors of neutrinos [?].

- Different cosmological and astrophysical observational data provide substantial evidence of the presence of a weakly interacting neutral massive particle, known as the dark matter, in our universe. So far, there is no direct experimental observation of the dark matter, though various experiments have been dedicated to look for the direct as well as indirect detection of this illusive particle. The SM does not provide any viable dark matter candidate [?].

- The SM cannot predict the observed matter over anti-matter asymmetry of the universe.

The above-mentioned shortcomings of the SM motivate us to look for physics beyond the SM (BSM). The Higgs boson pair production can also be probed in the BSM framework. In BSM, new particles can be produced which subsequently decay to a Higgs boson pair. Alternatively, anomalous Higgs boson coupling may enhance the Higgs boson pair production rate.

2.6.1 Non-resonant HH production in BSM

Non-resonant Higgs boson pair production may proceed through

- **Anomalous coupling**: Though the Higgs trilinear coupling λ_{HHH} is determined by the equation 1.3.0.23, in the BSM, the value of Higgs self-coupling as well as Yukawa coupling can be modified. The modification of these couplings have a direct impact on HH production as well as final state kinematic distributions. Deviation from the SM prediction is quantified in the following way where λ_{HHH}, λ_{HHH}^{SM} are anomalous and SM Higgs trilinear coupling respectively and y_t, y_t^{SM} are respectively anomalous and SM Yukawa coupling.

$$\begin{aligned} k_\lambda &= \frac{\lambda_{HHH}}{\lambda_{HHH}^{SM}} \\ k_t &= \frac{y_t}{y_t^{SM}} \end{aligned} \tag{2.6.1.1}$$

- **Effective Field Theory (EFT)**: In the BSM [?, ?], there may be contact interaction of the Higgs boson with gluons or two Higgs bosons with two gluons or top-antitop pair parametrized by the absolute coupling c_{2g} (Figure 1.7 (c)), c_g (Figure 1.7(d)), c_2 (Figure 1.7(e)) respectively. All these processes that are absent in the SM, can increase the HH cross section. The corresponding Feynman diagrams are shown in Figure 1.7.

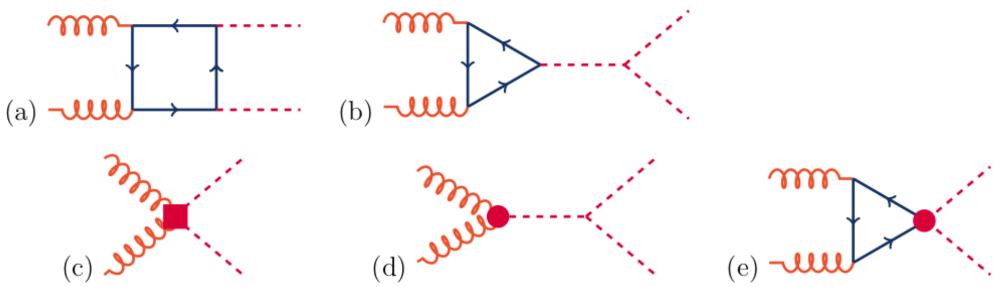

Figure 2.7: HH production in SM processes (a), (b), and in EFT processes (c), (d), (e).

2.6.2 Resonant HH production in BSM

There are various models which deal with resonant production of Higgs boson pair at the LHC.

- **Higgs Singlet, 2HDM, MSSM models** : In the Singlet model [?], one additional Higgs Singlet is introduced to the Higgs doublet

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+H \end{pmatrix}, S = \frac{1}{\sqrt{2}}(v'+h) \tag{2.6.2.1}$$

where v and v' represent the vacuum expectation values for both the fields. After Electroweak Symmetry Breaking, this gives two scalar bosons where one of the scalar bosons (say H) is identified as the Higgs boson discovered at 125 GeV and the other scalar boson (say h) can decay to a SM Higgs boson pair if kinematically possible. In the Minimal Supersymmetric model, two Higgs doublets are present which give five new bosons: two neutral CP-even bosons: h, H (one of which is usually identified as the particle discovered at the LHC); one neutral CP-odd boson (pseudoscalar): A; and two charged bosons: $H^{\pm}$. Models with this characteristics are generally referred as Two-Higgs Doublet Models (2HDM) [?, ?, ?]. These models are also characterized by new phenomena related to the SM-like Higgs boson, either by exotic production or decay. For the first case to happen, at least one of the new resonances in the model, H for example, must be heavier than the SM-like Higgs (h), which then allows decays such as $H \to h + X$, where X can either be a Z, an A or even another h. Alternatively, if A, for example, is lighter than $m_H/2 = 62.5$ GeV, $h \to AA$ becomes possible.

- **Warped extra dimension**: A Warped Extra Dimension model [?] was proposed by Randall and Sundrum (RS) to solve the gauge hierarchy problem in the SM. The model is given in the five-dimensional space-time where one warped extra dimension is compactified on the orbifold. The space-time metric is given by

$$ds^2 = e^{-2ky}\eta_{\mu\nu}dx^\mu dx^\nu - dy^2 \tag{2.6.2.2}$$

where e^{-2ky} is called the warped factor. Two 3-branes are located at $y = 0$ (UV brane) and $y = L$ (IR brane). In the original RS model all the SM fields are confined on the IR brane and the 5-dimensional fundamental scale M_5 at UV brane is scaled down to $M_5 e^{-kL}$ at the IR brane. By taking $kL \approx 35$, the fundamental scale $M_5 = \overline{M_P}$ is scaled down to TeV at IR brane. It is necessary to stabilize L to the above value. Goldberger and Wise showed that a bulk scalar field propagating in the background geometry, Eq. 1.6.2.2 can generate a potential that can stabilize L. The quanta of this scalar field is called Radion.

The detection of Radion at LHC can validate the RS model. The decay channel of $R \to HH$ has a large branching ratio ($\approx 24\%$) [?] for Radion mass greater than 250

GeV and this is almost constant as a function of Radion mass. The cross section of Radion production and it's subsequent decay to HH pair has been studied and will be discussed in this thesis.

Chapter 3

Experimental set-up at CERN

3.1 Large Hadron Collider

The Large Hadron Collider (LHC) [18] is the world's largest and most powerful particle accelerator till date. The LHC consists of a 27 kilometer ring filled with superconducting magnets and radio frequency cavities and is situated about 100 m underground at the French-Swiss border close to Geneva. It was built in the tunnel which was previously used by the Large Electron Positron Collider (LEP). Inside the accelerator two counter rotating beams of high energy proton are made to collide.

Acceleration of proton beam

Protons undergo acceleration through a series of steps before they achieve their final energy and are made to collide at 4 precise locations within the LHC tunnel. The acceleration chain is shown in Figure 3.1. Linac2 is the starting point where protons are produced from hydrogen gas and accelerated up to 50 MeV. Protons are then injected into the Proton Synchrotron Booster (PSB) and accelerated up to 1.4 GeV. In the next step protons are injected into the Proton Synchrotron (PS) where they achieve energy up to 25 GeV. From the PS, protons are injected into the Super Proton Synchrotron (SPS) which accelerates them up to 450 GeV and transfers into the LHC rings where the protons are accelerated to their final energy. The final beam energy has been 3.5 TeV during 2011, 4 TeV during 2012 and 6.5 TeV in 2015 and 2016.

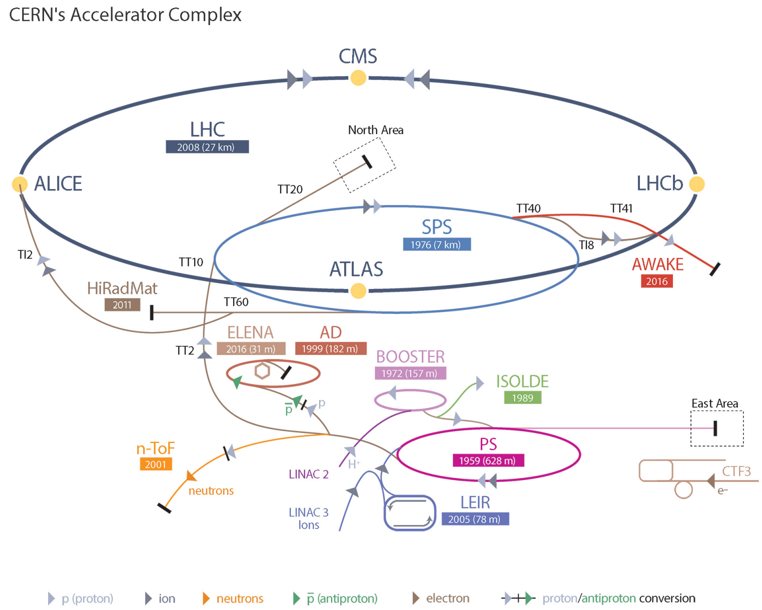

Figure 3.1: Overview of LHC accelerator.

Thousands of magnets of different varieties and sizes are used to direct the proton beams around the accelerator. These include 15 meters long dipole magnets (1232 in number) which bend the beams, and 392 quadrupole magnets of length 5–7 meters each, which focus the beams. Immediately prior to collision, another type of magnet is used to "squeeze" the beams of particles closer together to increase the chances of collisions.

3.2 Detectors at the LHC

The beams inside the LHC tunnel are made to collide at four locations around the accelerator ring, corresponding to the positions of the following four particle detectors,

- Compact Muon Solenoid (CMS) [19], which is built to cover a broad physics programme, from Standard Model (SM) physics to the extra dimensions and particles that could make up dark matter.

- A Toroidal LHC Apparatus (ATLAS) [20], with the same scientific goals as the CMS experiment.

- A Large Ion Collider Experiment (ALICE) [21], a heavy-ion detector, designed to study the physics of strongly interacting matter at extreme energy densities, where a phase of matter called quark-gluon plasma may form.

- Large Hadron Collider beauty (LHCb) [22], which studies the matter-antimatter asymmetry through studies involving b quarks.

One of the important parameters of the LHC is the luminosity $\mathcal{L}$ [23] which is a measure of how many collisions can occur when the beams collide. It depends on beam parameters as,

$$\mathcal{L} = \frac{\gamma f k_B N_p^2}{4\pi\epsilon_n\beta^*}F \tag{3.2.0.1}$$

where γ is the Lorentz factor, f is the revolution frequency, k_B is the number of bunches (a maximum of 2808), N_p is the number of protons per bunch (which could be as high as 1.15×10^{11}), ϵ is the normalized transverse emittance (typically 3.75 μm), β^* is the betatron function at the interaction point (IP) (designed to be 0.55 m) and F is the reduction factor due to the crossing angle. The designed instantaneous luminosity, $\mathcal{L}$, is 10^{34} cm^{-2}s^{-1}.

The total number of particles produced in a particular process is given by

$$N = \sigma \times L \tag{3.2.0.2}$$

where L is the total integrated luminosity, $\int \mathcal{L}dt$ and σ is the cross section of the corresponding process.

3.3 Data recorded by CMS

The first run of the LHC physics started in 2009 at $\sqrt{s} = 1.1$ TeV. The energy was increased to 7 TeV in 2010 and continued through 2011 reaching a maximum instantaneous luminosity of 3.5×10^{33} cm^{-2}s^{-1} with a bunch spacing of 50 ns. During 2012, the center-of-mass energy was increased to 8 TeV, and the machine reached a peak instantaneous luminosity of 7.7×10^{33} cm^{-2}s^{-1}. The total integrated luminosity recorded by the CMS detector during this time period was 19.7 fb^{-1}. During 2013-2014, the LHC experiment went through the first long shut down (LS1) when the accelerator went through several upgrades to cope with its nominal configuration. After LS1, the first collisions started in May 2015 at $\sqrt{s} = 13$ TeV with a bunch spacing of 50 ns and then 25 ns with peak instantaneous luminosity of 5×10^{33} cm^{-2}s^{-1}. The LHC delivered only 2.5 fb^{-1} during 2015. However, the year of 2016 was a record breaking period of the LHC, with peak instantaneous luminosity reaching 1.4×10^{34} cm^{-2}s^{-1}. Figure 3.2 shows the luminosity recorded by CMS for a certain time period. Data collected by CMS in 2016 with total luminosity of 35.9 fb^{-1} have been used in this thesis.

3.4 Overview of Compact Muon Solenoid detector

The Compact Muon Solenoid (CMS) detector [19, 23] is a multi-purpose apparatus operating at point 5 of the LHC. The detector is 21.6 m long with a diameter of 14.6 m and total weight of 12500 t. The name comes from the limited size of the detector, good muon identification capability and the presence of superconducting solenoid. The detector is roughly cylindrical in shape, closed by two endcaps on either side. Different detector components are shown in Figure 3.3.

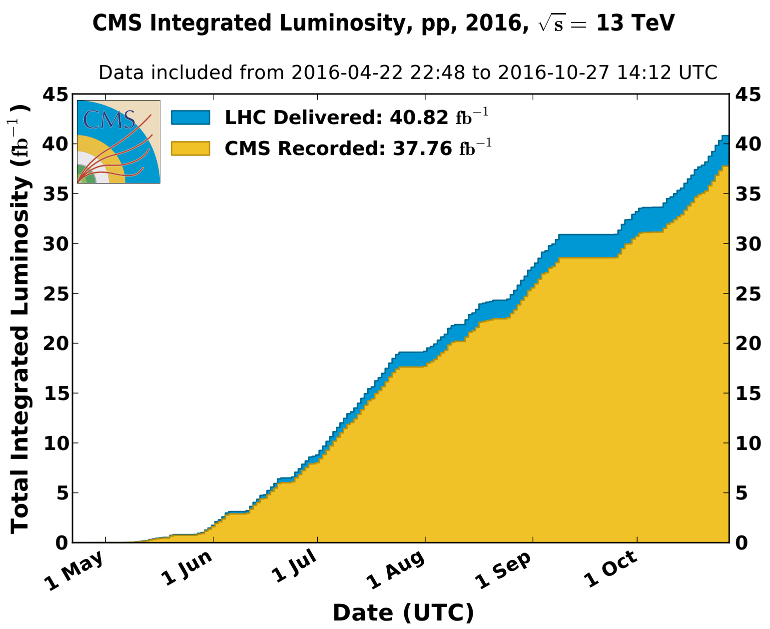

Figure 3.2: Integrated luminosity delivered by the LHC (blue) and recorded by the CMS (yellow) at $\sqrt{s} = 13$ TeV in 2016.

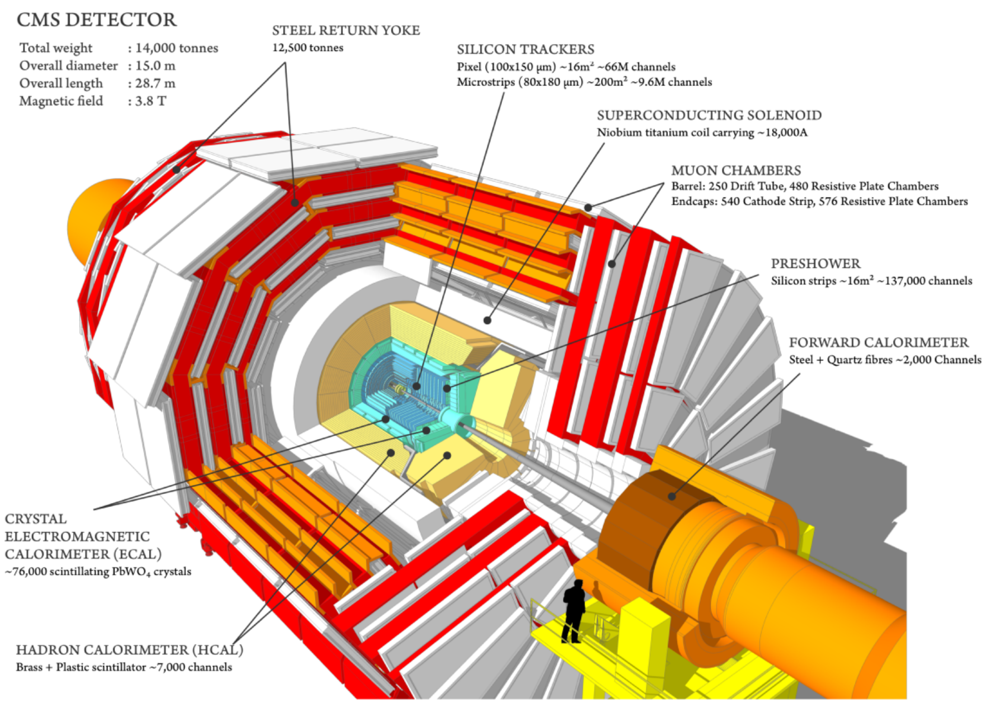

Figure 3.3: Layout of the CMS detector.

The position of a point inside the detector is described by a right-handed co-ordinate system with the origin located at the interaction point. The x axis points towards the center of the LHC, y axis points vertically upward and z axis is given by the direction of anti-clockwise proton beam.

The polar co-ordinate system corresponding to the cylindrical structure of the detector is also used to define the position of a particle inside the detector. The azimuthal angle ϕ is the angle with the positive x axis in the x-y plane and r denotes the radial distance from the origin. The polar angle θ is the angle made by r with the z axis. The angle θ is used to define the pseudorapidity which is given by

$$\eta = -\ln(\tan(\theta/2) \tag{3.4.0.3}$$

In the relativistic limit with $E >> m$, pseudorapidity is equivalent to the rapidity which is given by

$$y = \frac{1}{2}\ln(\frac{E + p_z}{E - p_z}) \tag{3.4.0.4}$$

The projection of momentum in the transverse plane *i.e.* in the x-y plane is called transverse momentum, denoted by p_T. Both η and p_T are invariant under Lorentz boost along the z axis.

The CMS detector has four main sub-detectors. The innermost detector is the Tracker which is surrounded by the Electromagnetic Calorimeter (ECAL) followed by the Hadron Calorimeter (HCAL) and the Muon chambers successively. Between the Hadron calorimeter and the Muon chambers, there is a superconducting coil which provides a magnetic field 3.8 T. A brief description of different layers of the CMS detector and the trigger system is given in the following sections and in the next chapter.

3.4.1 Tracker

The CMS tracker [19] (figure 3.4) is the closest sub-detector surrounding the interaction point. The function of the tracker is to reconstruct the position of primary vertex as well as secondary vertices and the trajectories of all charged particles precisely. To reduce the effect of multiple scattering, nuclear interaction, bremsstrahlung radiation, photon conversion etc. on the measurement of momentum resolution, the material budget around the tracker has been optimized. To reduce thermal runaway effect, the entire tracker is kept at $\leqslant -10°$C.

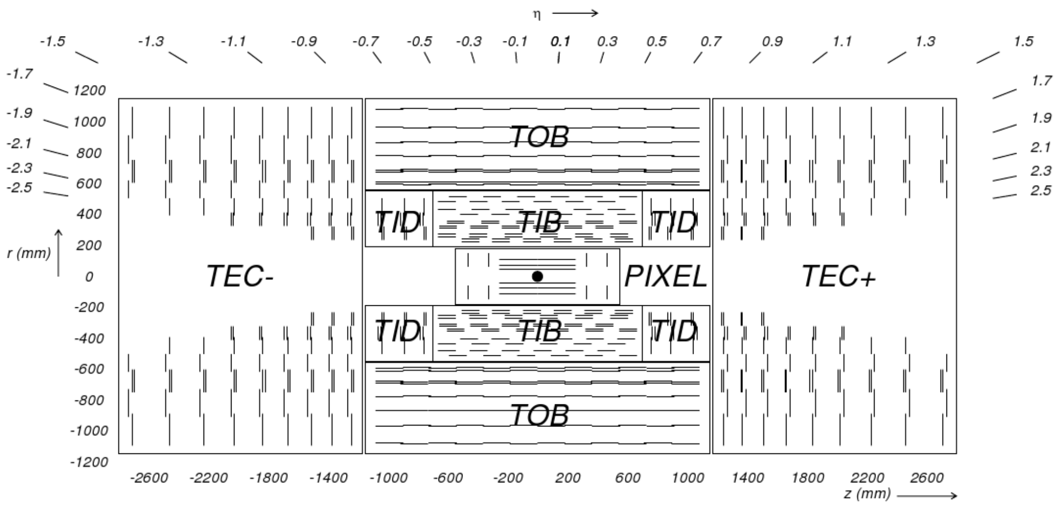

Figure 3.4: Schematic cross section of the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits [19].

The innermost part of the tracker is a pixel system which consists of three barrel layers (BPix) at mean radii of 4.4, 7.3 and 10.2 cm, with two endcap disks (FPix) extending from ≈ 6 to 15 cm in radius. The pixel size is 100×150 μm^2 and the spatial resolution is in the range of 15–20 μm. The pixel detector delivers three high precision space points on each charged particle trajectory. There are 66 million pixels in the pixel detector covering an area of about 1 m^2.

The pixel detector is surrounded by the Silicon Strip Tracker (SST). The SST consists of four parts, tracker inner barrel (TIB), tracker inner disk (TID), tracker outer barrel (TOB) and tracker endcap (TEC). The TIB and TID, consisting of 4 barrel layers and supplemented by 3 disks at each end, deliver up to 4 r–ϕ measurements on a trajectory using 320 μm thick silicon micro-strip sensors and the resolution of single point varies from 23 μm to 35 μm. The TIB/TID is surrounded by TOB with 6 barrel layers of 500 μm thick micro-strip sensors. The TOB provides another 6 r-ϕ measurements with a maximum resolution of 53 μm. TOB is complemented with TEC which is composed of 9 disks providing up to 9 ϕ measurements per trajectory on each side of TOB. A few layers in TIB, TOB and a few rings of TID, TEC, carry double-sided modules where two independent single-sided modules are mounted back-to-back with a stereo angle of 100 mrad to measure the second co-ordinate z in barrel and r on the disks. The resolution of this measurement is 230 μm and 530 μm in TIB and TOB and varies with pitch in TID and TEC. The CMS silicon strip tracker has a total of 9.3 million strips and 198 m^2 of active silicon area. The acceptance of tracker is up to $|\eta| \approx 2.5$.

Figure 3.5 shows the expected resolution of transverse momentum, transverse impact parameter and longitudinal impact parameter of muon with different p_T as a function of η.

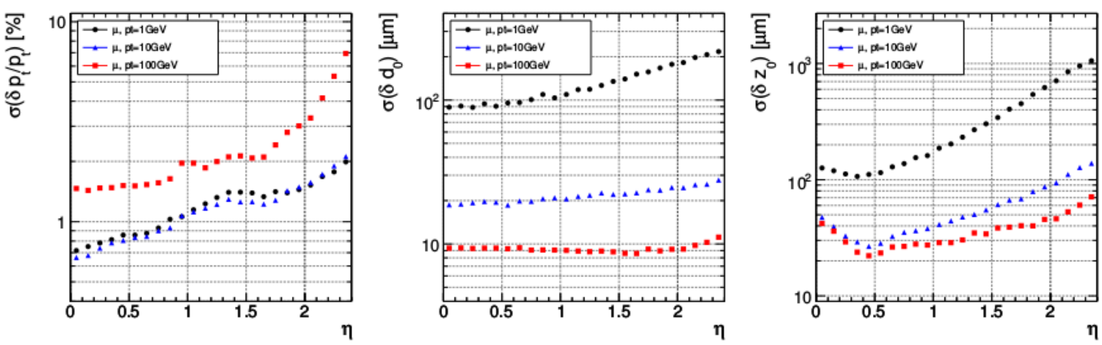

Figure 3.5: Resolution of transverse momentum (left), transverse impact parameter (middle) and longitudinal impact parameter for single muons with p_T of 1, 10 and 100 GeV [19].

3.4.2 Electromagnetic calorimeter (ECAL)

The tracker is surrounded by the ECAL [19, 23] which is a hermetic homogeneous calorimeter made of 61200 lead tungstate ($PbWO_4$) crystals mounted in the central barrel part, complemented with 7324 crystals in each of the two endcaps. The function of the ECAL is to measure the energy and position of incident electrons and photons. An incident electron or photon excites the crystals which emit scintillation light while de-exciting. $PbWO_4$ crystals with high density (8.28 g/cm^3), short radiation length (X_0 = 0.89 cm) and small Moliere radius (R_0 =2.2 cm) provide fine granularity and a compact calorimeter. Around 80% of light is emitted within 25 ns which corresponds to the bunch spacing of the LHC. Figure 3.6 shows a transverse section of the ECAL.

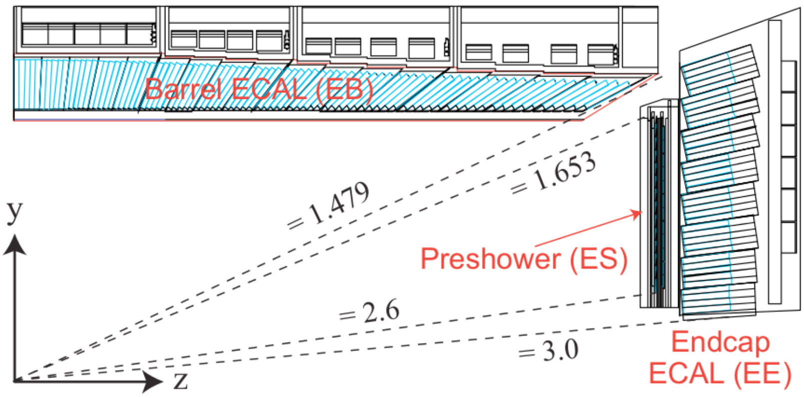

Figure 3.6: Transverse section through the ECAL [23].

The granularity of the barrel part of the ECAL (EB) is 360-fold in ϕ and (2×85)-fold in η, extending up to $|\eta| < 1.479$. The crystals are mounted in quasi-projective geometry with their axes making a small angle (3°) with respect to the vector from the nominal interaction vertex in both ϕ and η directions. Crystals are arranged in 18 super-modules, each covering 20° in ϕ, forming a half barrel. Super-modules are made up of 4 modules containing crystals in a thin walled alveolar structure. The crystal length is 230 mm corresponding to 25.8 X_0. The output of the crystals are read out by avalanche photodiodes (APDs). Each APD, with a gain factor of 50, has an active area of 5×5

mm^2 and a pair of APD mounted on each crystal.

The endcaps (EE), extending within $1.479 < |\eta| < 3.0$, consist of identically shaped crystals grouped in units of 5×5 crystals known as super-crystals. The crystals point at a focus 1300 mm beyond the interaction point, giving off-pointing angles from 2 to 8 degrees. The length of the crystals is 220 mm (24.7 X_0). In EE, the photodetectors are vacuum phototriodes (VPT). One VPT with mean gain of 10.2 at zero magnetic field, is glued to the back of each crystal.

Due to the dependency of number of scintillation photons emitted by crystals and amplification of the APD on temperature $(-3.8 \pm 0.4)\%\ {}^\circ C^{-1}$ the temperature of ECAL should be maintained at a constant value with high precision requiring a cooling system to extract the heat produced by read-out electronics. The nominal temperature of ECAL is stable at 18°C within ± 0.05°C.

A sampling calorimeter known as preshower is installed to identify neutral pions decaying into a pair of photons, to improve the position determination of electrons and photons. It is placed in front of the two endcaps within $1.653 < |\eta| < 2.6$ and consists of two layers: lead radiators to initiate electromagnetic shower from incoming photons/electrons, followed by silicon strip sensors to measure the deposited energy and the transverse shower profiles.

The energy resolution of the ECAL is parametrized as

$$\frac{\sigma}{E}^2 = (\frac{S}{\sqrt{E}})^2 + (\frac{N}{E})^2 + C^2, \tag{3.4.2.1}$$

where the first term is proportional to the secondary particles produced which has Poisson distribution with mean proportional to the energy of incident particle, the second term is proportional to the noise of electronics and digitization and the last term depends on the leakage of shower energy which is again proportional to the incident particle energy. A typical energy resolution for electron beam with momenta between 20–250 GeV/c is found to be [19] :

$$\frac{\sigma}{E}^2 = (\frac{2.8\%}{\sqrt{E}})^2 + (\frac{0.12}{E})^2 + (0.30\%)^2 \tag{3.4.2.2}$$

The energy resolution for a 120 GeV electron is around 0.5%.

3.4.3 Magnet

Large magnetic field is required for a precise measurement of the momentum of charged particles because the particle momenta are measured from the bending of the trajectory within the magnetic field. In CMS, a superconducting solenoid with length of 12.4 m and external diameter of 6.9 m produces 3.8 T magnetic field [24]. The solenoid sits outside the tracker and calorimeter and inside the muon system. The magnet is made of 2168 turns of coils made out of niobium-tin alloy carrying 19.5 kA current. The magnetic field generated by the superconducting coil is returned via yoke which is situated outside the superconducting coil.

3.4.4 Muon chamber

The last component of the CMS detector is the Muon chamber system [25] which sits outside the superconducting coil. Since muons are minimum ionizing particle, they can traverse the other sub-detectors of the CMS and reach the muon chambers. The function of the muon chambers is to trigger identification and momentum measurement of muons. The Muon chamber system consists of three types of detectors. Figure 3.7 shows the layout of the muon chamber system.

In the barrel, there are 250 drift tubes (DT), organized into 12 ϕ-segments in 5 wheels which are situated along z direction and form 4 stations in each wheel at different radii, interspersed between plates of the magnetic flux return yoke. The basic element of the DT system is drift cell with transverse size 42×13 mm^2 and with 50-μm-diameter anode

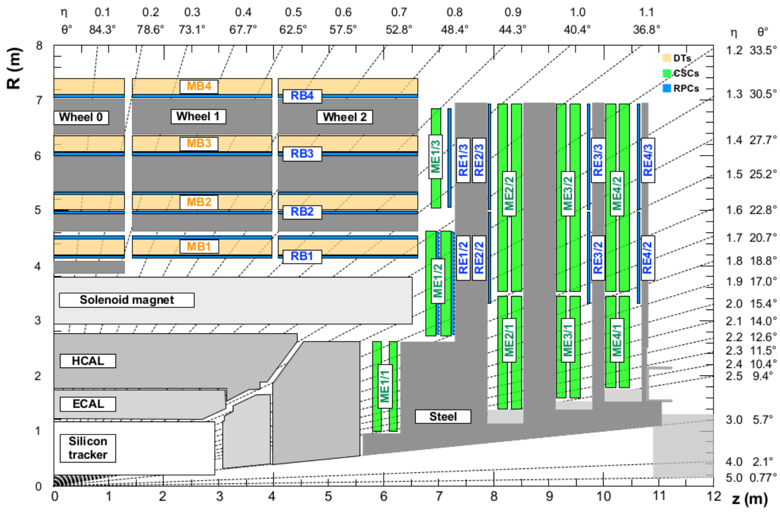

Figure 3.7: Schematic of the layout of the Muon chamber system [25].

wire at the center. The gas mixture of argon and CO_2 inside the cell provides a drift velocity of about 55 μm/ns. Four staggered layers of parallel cells form a super-layer (SL). Each chamber consists of 2 SLs which measure the r–ϕ co-ordinates and an orthogonal SL that measures r–z co-ordinate. A muon traversing the chambers ionizes the gas and the position of the muon is measured from the time needed by the electrons to drift towards the anode wire. The spatial resolution of the DT chamber is 100 μm in the r–ϕ direction and 150 μm along z direction.

To cope with high and non-uniform magnetic field and large background contributions, the endcap cathode strip chambers (CSC), which are multi-wire proportional chambers, are installed within $0.9 < |\eta| < 2.4$. Like the DT, the CSCs also form 4 stations in the endcap which are arranged in 4 disks perpendicular to the beam and in concentric rings, 3 rings in the innermost station and 2 in the others. There are 486 CSC chambers in total in both the endcaps. Each CSC consists of 6 layers, containing 80 cathode strips in each layer. The position resolution of the CSC chambers is around 75–150 μm depending on the position.

In order to assign a hit in the muon chamber to correct the bunch crossing, CMS introduces resistive plate chambers (RPC) which provide fast and independent trigger with excellent time resolution of about 3 ns within $|\eta| < 1.6$. The RPCs are double-gas chambers operated in avalanche mode. Charged particles crossing an RPC ionize the gas in both gas volumes and the avalanches generated by the high electric field induce an image charge, which is picked up by the read-out strips. A total of 6 layers of RPCs are embedded in the barrel muon system, 2 in each of the first 2 stations, and 1 in each of the last 2 stations. The redundancy in the first 2 stations allows the trigger algorithm to work even for low-p_T tracks that may stop before reaching the outer 2 stations. In the endcap region, there is a plane of RPCs in each of the first 3 stations.

3.4.5 Trigger

The time difference between the crossings of two proton bunches is 25 ns which translates to a beam crossing frequency of 40 MHZ. The average event size produced by CMS data acquisition system is about 1 MB which gives a bandwidth of data $\approx$ 40 TB/s which is impossible to store with the current technology available. A two-tier trigger system [26] is introduced to reduce this huge output data while retaining the interesting physics events and at the same time rejecting the background events to a very large extent. The work flow of the first trigger level (L1) is shown in Figure 3.8.

The first layer of the CMS trigger system is called the Level-1 (L1) trigger which is a hardware system based on FPGA and custom ASIC integrated circuits technology, with a fixed latency of 4 μs. The L1 trigger reduces the output rate to 100 kHz. The first step of L1 Calorimeter trigger is to measure the transverse energy in the calorimeters from trigger towers (TT) which is a collection of read-out channels in the ECAL and HCAL separately. The Regional Calorimeter Trigger (RCT) determines the electron/photon, tau, jets, E_T (transverse energy), H_T (the scalar sum of the p_T of all jets with $p_T > 10$ GeV and $|\eta| < 3$) etc. based on the output of TT. All the particles are sorted according to E_T and top four of each type are sent to the Global Calorimeter trigger (GCT) for further

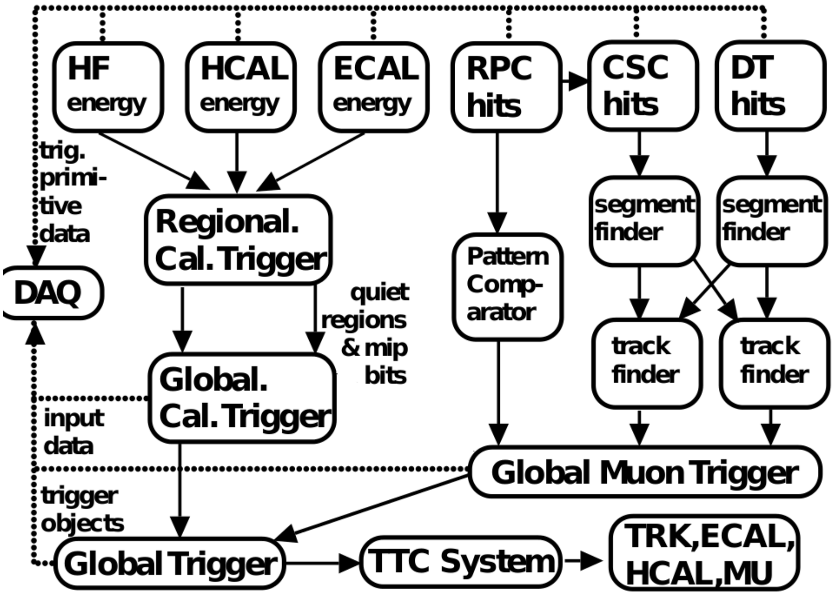

Figure 3.8: Work flow of the L1 trigger system at CMS [26].

processing. The GCT calibrates the clustered jet energies and also total transverse energy supplied by the RCT.

The Muon triggers use information from three muon sub-systems, DT, CSC and RPC. The DT and CSC provide track segments and complete tracks are formed joining the tracks and hits reported by DTs and CSCs. RPCs provide a dedicated detector system with excellent time resolution and produce their own track candidates based on regional hit patterns. The Global Muon Triggers (GMT) combine all regional information from the three sub-detectors.

The output from the GCT and GMT are received by the L1 global trigger system which is the final step of the L1 trigger system and a decision is formed to either select or reject the events based on certain conditions like minimum threshold of E_T or p_T, cut on η and/or cut on the azimuthal angle between two particles etc. Several conditions are combined by simple combinatorial logic to form up to 128 algorithms which are hard-coded in firmware and can be changed by loading an updated firmware version.

Finally, the output from the L1 trigger along-with the complete read-out of all the detector systems is fed into the High Level Trigger (HLT) which reduces the output rate to a few 100 Hz. While the L1 trigger uses information from calorimeter and muon chambers only, the HLT considers the tracker information in addition and the event selection is done in a way which is quite similar to the offline analysis except the calibration constants which is applied during the offline analysis. The object reconstruction is done around the L1 seed to save the object reconstruction time. The events accepted by the HLT are stored on disk for offline processing.

Chapter 4

Hadron Calorimeter

The Hadron Calorimeter (HCAL) of the CMS is the layer of detector outside the ECAL. The HCAL is important for the measurement of hadronic jets and missing transverse energy which are caused due to neutrinos or exotic particles which escape the CMS detector due to their very low interaction cross section with the detector material. Ideally, the HCAL should be compact, hermetic and made of non-magnetic materials for the components located inside the magnet. The HCAL consists of 4 components, HCAL Barrel (HB), HCAL Endcap (HE), outer HCAL (HO), forward calorimeter (HF). Figure 4.1 shows the different parts of the HCAL. A short description of all the sub-detectors of the HCAL is given in the following sub-sections.

4.1 HCAL Barrel (HB)

The HB [27] is a sampling calorimeter, extending up to $|\eta| < 1.39$. The HB is made up of 36 identical azimuthal wedges which form two half-barrels (HB+ and HB−). Each wedge, covering 20° along ϕ direction, is further segmented into four azimuthal sectors, each covering 5° in ϕ. Of the 4 ϕ segmentations within a barrel wedge, the two inner ones

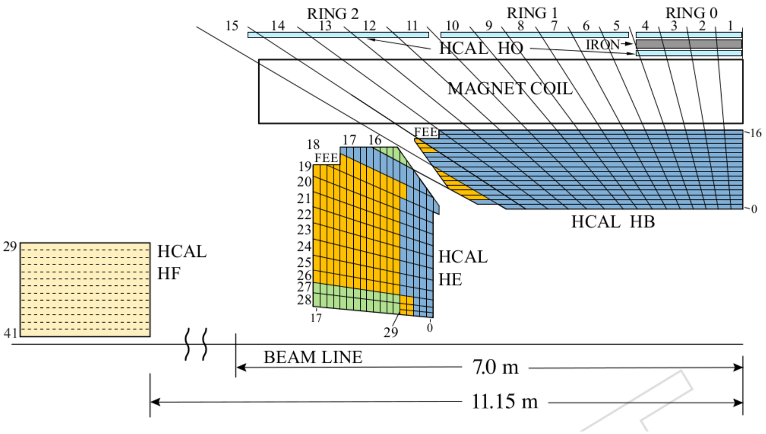

Figure 4.1: A schematic view of the CMS Hadron Calorimeter during 2016 LHC operation, showing the positions of its four major components: HB, HE, HO and HF.

are staggered at larger radius with respect to the outer two to avoid any dead region over the entire azimuthal coverage. The wedges are made of 17 layers of cartridge brass (70% Zn and 30% Cu) and plastic scintillators. The innermost and outermost absorber plates are made of stainless steel for structural strength. The innermost and the outermost scintillators are 10 mm thick while the remaining 15 layers use 3.7 mm thick scintillators. The innermost layer is made out of Bicron BC408 scintillator to have more light output compared to Kuraray SCSN-81 scintillators which are used for the remaining layers. The plastic scintillator is divided into 16 η sectors, resulting in a segmentation $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$. All 17 layers of scintillators corresponding to the same value of $i\eta$ (up to $|i\eta| = 14$) and $i\phi$ are grouped into a single read-out which is called a tower where $i\eta$ and $i\phi$ represent the segment numbers along η and ϕ directions respectively. For $|i\eta|$ beyond 14, there are two towers with depth = 1 and 2. The total absorber thickness at 90° is 5.82 interaction length (λ_I). The effective thickness of HB increases with polar angle (θ) as $1/\sin\theta$, giving in 10.6 λ_I at $|\eta| = 1.3$. The Electromagnetic Calorimeter in front of the HB provides about 1.1 λ_I of material.

4.2 Hadron Endcap (HE)

The HE [28] covers the region $1.3 < |\eta| < 3$ and is made of brass absorbers with sampling layers of plastic scintillators. Each endcap has 18 fold symmetry in ϕ, covering 20° in ϕ. Like HB, HE also has the staggered geometry along ϕ direction to avoid any dead region. The absorber plates are 79-mm thick with 9 mm gaps to accommodate the scintillators. The endcap on either side is divided into 14 parts along η and 18 layers in total are combined into 1, 2 or 3 depth sections. The granularity of the calorimeter is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ (0.17×0.17) for $|\eta| < 1.6$ ($\geqslant 1.6$). The total length of the calorimeter, including the Electromagnetic Calorimeter, is about 10 interaction length, λ_I.

4.3 Hadron Forward (HF)

The HF [29] is located 11.15 m away from the interaction point on either side of the CMS detector and covers $2.9 < |\eta| < 5.2$ and thus significantly improves jet coverage and the missing transverse energy resolution. These are essential in the studies of top quark production, Standard Model Higgs, and all SUSY particle searches. The HF faces a much larger particle flux and on average 760 GeV at the center of mass energy of 14 TeV, is deposited into the two HF calorimeters per proton-proton interaction compared to 100 GeV for the rest of the detectors. Due to this hostile environment, quartz fibers are chosen as the active medium and steel as absorber. The detector utilizes the technique of Cerenkov light emitted by charged particles with energy above the Cerenkov threshold ($E > 190$ keV for electrons) while passing through quartz fibers embedded within the absorbers. Each HF module is constructed from 18 wedges made out of iron with quartz fibers along its length and each wedge contains two ϕ sectors of 10°. There are two types of fibers, with half of the fibers, known as Long fibers, run over the entire length of the absorber (165 cm corresponding to $\approx 10\lambda_I$), while the other half, known as Short fibers, starts at a depth of 22 cm from the front of the detector. These two sets of fibers are read out separately and help to distinguish the electromagnetic shower which are mostly absorbed in the first 22 cm, from the hadronic shower which spread over a much longer length. Size of hadronic shower depends on the energy of the particle and the total length increases logarithmically with energy at high energies. The shower maximum happens between 1.5 to 2 interaction length again depending on the energy of the hadron. In the case of forward calorimeter, the absorber is stainless steel of interaction length 10.4 cm. The total length of the forward calorimeter corresponds to more than 10 interaction length where the entire shower is supposed to be absorbed. The HF segmentation is 0.175×0.175 except for $|\eta|$ above 4.7, where the segmentation is 0.175×0.35 and 0.3×0.35 for the last two η segments.

4.4 HO

In the central pseudorapidity region, the EB and HB do not provide sufficient containment for hadron showers. To ensure adequate sampling depth for $|\eta| < 1.3$, the Hadron Calorimeter is extended outside the solenoid with an additional detector called the HO [30] which uses the solenoid as an additional absorber equal to $1.4/\sin\theta$ interaction lengths. Outside the vacuum tank of the solenoid, the magnetic field is returned through an iron yoke designed in the form of five rings. The HO follows the same geometrical structure of muon chamber in each wheel as mentioned in section 3.4.4. The HO is placed as the first sensitive layer in each of the five rings. All rings have a single HO layer except the ring at $\eta = 0$, where due to minimal absorber depth, two layers of HO scintillators are present on either side of 19 cm thick 'tail catcher' iron. The sizes and positions of the tiles in HO roughly map the layers of HB to make towers of granularity 0.087×0.087 in η and ϕ.

4.5 Signal read-out

Scintillator light inside each scintillator tile tower is collected by an embedded wavelength shifting fiber and the signals are added optically. The colour scheme in Figure 4.1 denotes the longitudinal segmentation of the read-out where all layers with the same coloured tower are summed. The optical signals for the HB, HE are detected by hybrid photo diodes (HPD) with 19 independent pixels; 18 for read-out of fibers and one for monitoring. For the HO, the signal is read out by Silicon Photo Multiplier (SiPM). In the HF, Long and Short fibers are read out separately. The calorimeter is thus functionally subdivided into two longitudinal segments. The photo detectors for HF are eight-stage photo-multipliers (PMT) with a borosilicate glass window, which register the Cerenkov light produced by the charged shower particles in the quartz fibers.

The signals from both types of photodetectors are read by a custom-designed chip, which performs charge integration and encoding (QIE). The QIE is a non-linear multi-range analog to digital converter (ADC) that provides approximately constant fractional precision over a wide dynamic range. For each channel, the measured ADC value is converted into a charge measured in femto Coulomb (fC), using a conversion factor that was determined in the laboratory by charge injection into the QIE.

Figure 4.1 shows a schematic quarter view of the HCAL system in the barrel, endcap and forward regions. Also shown are the locations of some of the Front End Electronics (FEE). The HF FEEs (not shown) are placed around a ring at $|\eta| = 3$ (tower number 29) and HO FEEs are located inside the muon detectors at various locations.

4.6 HCAL Calibration

The purpose of the HCAL calibration is to derive response corrections (RC) and to establish a stable hadronic energy scale. The first step of calibrating the hadron calorimeter with collision data is to equalize the response in ϕ for each η ring. The procedure takes advantage of the azimuthal symmetry of the detector and the corresponding ϕ-symmetric energy deposition from minimum bias (MinBias) events. The barrel and the endcap detectors do not strictly provide a fully ϕ-symmetric layout because of the absorber structure. Figure 4.2 shows expected level of symmetry as a function of ϕ segments from a simulation study with minimum bias events. However, this deviation from strict equality is ignored in the determination of inter-calibration coefficients.

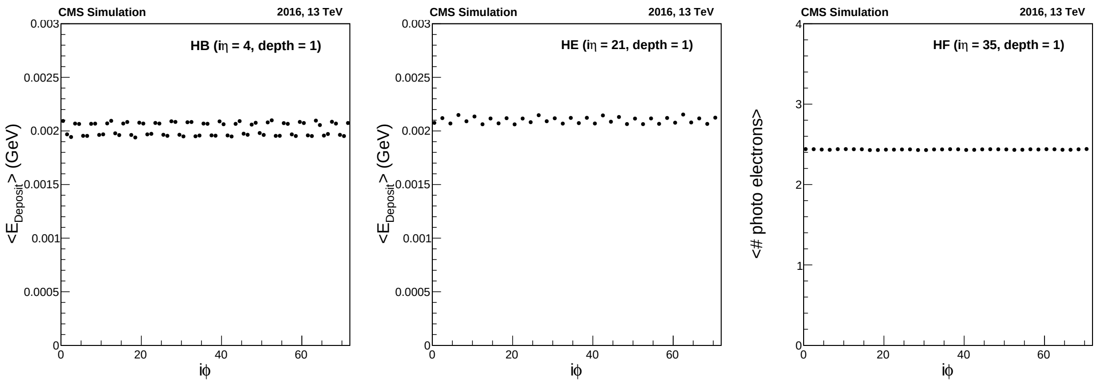

Figure 4.2: Distribution of mean energy deposited as a function of $i\phi$ for 3 η rings. The left plot refers to a HB ring ($i\eta$ = 4, depth = 1), middle one is for a HE ring ($i\eta$ = 21, depth = 1) and the right refers to a HF ring ($i\eta$ = 35, depth = 1).

The inter-calibration is performed by comparing the average energy deposit in a calorimeter cell to the mean of the average energy distributions in the entire $i\eta$-ring. There are two complementary calibration procedures which make use of the azimuthal symmetry.

4.6.1 Iterative Method

Here the inter-calibration is performed by equalizing the mean of read-out energies within some thresholds using physics data streams which are collected using triggers independent of HCAL. The lower threshold is set to the value far from the noise level by 5 sigma (4 GeV at least). The upper threshold is set to 100–150 GeV to avoid accidental high-energy hits. The values of upper and lower energy thresholds vary with detector types. For each channel, the total energy read-out value E_{tot} is estimated from the corresponding energy spectrum dN/dE. A multiplicative correction to the calibration coefficient (scale factor) is calculated as the ratio of E_{tot} to the average energy of $i\eta$-ring at the considered depth for each channel in each iteration. These correction factors are then applied to the energy measurement and the whole process is repeated including the selection of the towers chosen for the energy summation. This process continues till the mean variation in the correction factor falls below a convergence cut off value. The cut on the tower energy in the evaluation of the correction factor applies to the current estimate of the tower energy. So, if the correction factor is modified, the tower energy is changed, and the cut may accept or reject the tower in the new estimation. In fact this is why an iterative method is applied and the correction factor stabilizes after around 10 iterations. Number of iteration is controlled by a convergence criteria.

The statistical uncertainties of the corrections obtained with the 2016 pp-collisions data are of the order of 1% for the HB channels, less than 0.5% for the HF channels and are within a range between 0.1% and 1.0% for the HE channels depending on iη and depth.

4.6.2 Method of Moments

In this method, the first two central moments of the energy distributions are used to obtain the correction coefficients of each channel. The analysis is done using minimum bias events. One event in every 4096 events satisfying level 1 trigger is read out without zero suppression for the entire hadron calorimeter and these events are known as non-zero suppressed (NZS) data. In the Zero Suppression method it is checked if the sum of the energy deposition in two consecutive time slices is above a certain threshold. If this condition is not satisfied, the read-out of that particular channel is suppressed.

First two central moments of the read-out energy distribution are used to obtain the correction factors using the property of azimuthal symmetry of the HCAL detector in the following way:

1. Direct comparison of the mean deposited energy in all the channels after noise subtraction

2. Analysis of the variances of the signal and noise samples

The correction factor for each channel using the first moment or the mean is given by :

$$Corr_{i\eta,i\phi} = \frac{\frac{1}{N_\phi} \times \sum_{N_\phi} < E_{i\eta,j\phi} >}{< E_{i\eta,i\phi} >} \tag{4.6.2.1}$$

where N_ϕ is the number of channels in an iη ring and $E_{i\eta,i\phi}$ is the mean energy deposited in a particular channel.

$$< E_{i\eta,i\phi} >=< E^{signal}_{i\eta,i\phi} > + < E^{noise}_{i\eta,i\phi} > \tag{4.6.2.2}$$

After the noise subtraction and assuming $< E^{noise}_{i\eta,i\phi} >= 0$, we have $< E_{i\eta,i\phi} >=< E^{signal}_{i\eta,i\phi} >$. The uncertainty on the estimation of the correction factors is given by

$$\sqrt{\Delta^2(< E^{signal}_{i\eta,i\phi} >) + \Delta^2(< E^{noise}_{i\eta,i\phi} >)} \tag{4.6.2.3}$$

where Δ^2 denotes the variance. The uncertainty is dominated by the uncertainty on the noise estimation. For a precision of better than 2% in the middle of HB ($|\eta| = 1$), a few tens of million events are required. Therefore, the simplicity and transparency of this approach is offset by the requirement of larger data samples.

The second moment or variance relies on noise removal through subtracting the variance of noise from the variance in the measured energy. The correction factor in this case is given by:

$$Corr_{i\eta,i\phi} = \sqrt{\frac{\frac{1}{N_\phi}\sum_{N_\phi} < \Delta^2 R_{i\eta,j\phi} >}{< \Delta^2 R_{i\eta,i\phi} >}} \tag{4.6.2.4}$$

where

$$\Delta^2 R_{i\eta,i\phi} =< \Delta^2(E^{signal}_{i\eta,i\phi}) + \Delta^2(E^{noise}_{i\eta,i\phi}) > - < \Delta^2(E^{noise}_{i\eta,i\phi}) > \tag{4.6.2.5}$$

The minimum sample for achieving 2% uncertainty on the signal variance is of the order of a few million of events. This method requires substantially smaller samples but it is still sensitive to the noise level in a channel. For noisier channels we need larger statistics. The azimuthal inter-calibration with variances requires knowledge of the electronics noise in a channel. Although azimuthal inter-calibration with mean values does not require the knowledge of pedestals, we still need to investigate them to be sure that the expected value is zero.

4.7 Noise estimation

Due to low luminosity data and large bunch spacing in Run 1, the same data sample can be used for estimating both the signal and the noise. For HB and HE calibration, the first four (0–3) time slices (TS) of width 25 ns are used to estimate the noise, while TS's 4–8, each of width 25 ns are used to estimate the signal. For HF, windows of three TS's are used both for signal and noise. But starting with 25 ns bunch spacing we cannot use the first samples for noise estimation anymore due to the presence of out-of-time pile up (will be described in next chapter) which interferes with the earlier TS's. Local pedestal runs taken at the same period are used for noise estimation. The noise in each channel for HB, HE is measured separately from an independent dataset coming from pedestal runs which are termed as “TestEnablesEcalHcal”. Events from these runs are reconstructed using the standard procedure and the reconstructed energy in each channel gives the noise. Each dataset contains 4–7 million events. There is no data available for the HF, so noise level in HF cannot be examined.

4.8 Correction Factors

Method of moments using variance method is used to calculate the correction factors for HB and HE, as noise in each channel for HB, HE can be measured from local runs. Since there is no measurement of noise for HF, the correction factors for HF is estimated by "Method of moments using means". Figure 4.3 and 4.4, show the distribution of correction factors obtained from 2016 data.

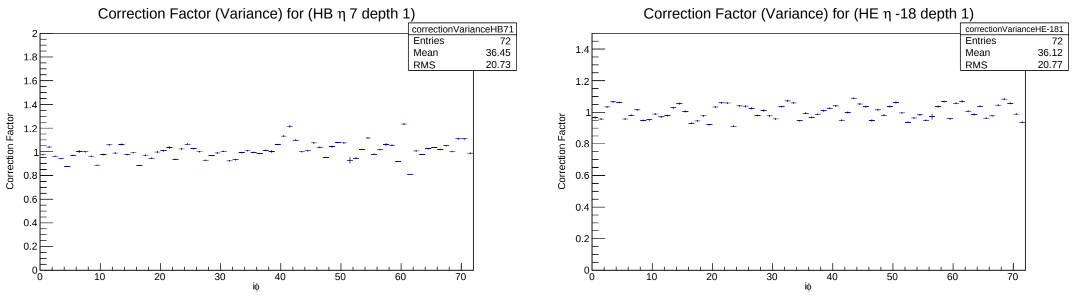

Figure 4.3: Distribution of correction factors obtained from the method of variance for HB ($i\eta$ = 7, depth = 1) (left), and HE ($i\eta$ = −18, depth = 1) (right).

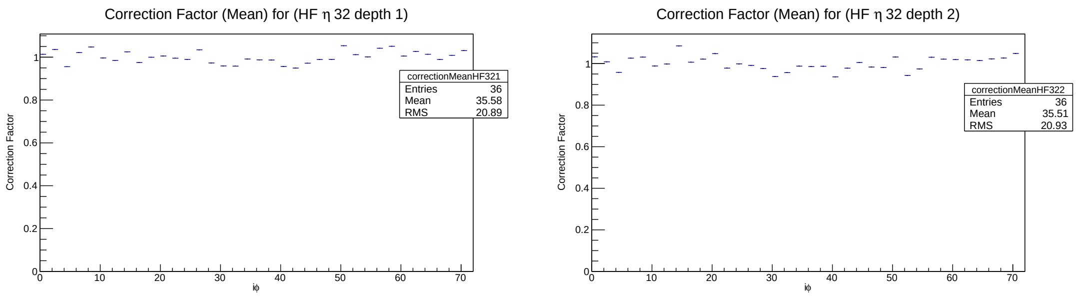

Figure 4.4: Distribution of correction factors obtained from method of mean for HF ($i\eta$ = 32, depth = 1) (left), and HF ($i\eta$ = 32, depth =2) (right).

4.8.1 Combination of the iterative method and method of moments

Final correction factors for the HCAL calibration coefficients are estimated as the error-weighted averages of corrections obtained with the iterative and the method of moments. The arithmetic average of the corrections is used when statistical errors of both methods are below 1% and the weighted average ($w = \frac{1}{\sigma^2}$) is used for the other cases.

4.8.2 Absolute Correction

Once relative correction factors are calculated using iterative and method of moments and applied to the channels, absolute correction factors are used to equalize the η response. Different techniques are used to calculate the absolute correction factors depending on the detector's position. In HB and parts of HE within tracker acceptance, isolated tracks are used to obtain the absolute correction factors. In this method isolated tracks which behave like MIP in the ECAL and deposit energy in the HCAL are selected and calorimeter response is defined as:

$$\frac{E_{HCAL}}{p_{Track} - E_{ECAL}} \tag{4.8.2.1}$$

where E_{HCAL}, E_{ECAL} are the energy deposited in HCAL, ECAL respectively and p_{Track} is the momentum of the track measured in the tracker. The calibration corrections are calculated in an iterative way where the iteration continues until the difference between the calorimeter response at the subsequent steps become three times smaller than the statistical uncertainty.

4.9 Calibration of HF and HO

Calibration of the HF is carried out using $Z \rightarrow ee$ events where one electron is measured in the ECAL and the other electron is reconstructed in the HF. The invariant mass of the electron pair is made consistent with Z boson mass.

In the HO, the relative calibration makes use of muons from the collision data and Cosmic muons that traverse the tiles of the HO. The determination of absolute energy scale makes use of di-jet and γ+jet events. In di-jet events, the energy of two leading jets are balanced by $E_b = 2\frac{(p_{T_{jet1}} - p_{T_{jet2}})}{(p_{T_{jet1}} + p_{T_{jet2}})}$ where $p_{T_{jet1}}$ and $p_{T_{jet2}}$ are the leading and sub-leading jet p_T respectively. In the γ+jet events, absolute response is defined as $R_{abs} = \frac{p_T^{jet}}{p_T^{\gamma}}$ where p_T^{jet} and p_T^{γ} represent jet p_T and γ p_T respectively.

4.10 HCAL Upgrade

Though in Run 1 and the ongoing Run 2 of the LHC, the CMS HCAL detector has run successfully, the radiation damage during this run period has increased the dark current in the scintillator tiles and also the performance of HPD has degraded due to the high operating voltage. During the HL-LHC phase, the instantaneous luminosity and hence the PU will increase by a huge factor. While in 2016, CMS has experienced around 25 PU in an event, PU will increase to around 150–200 in HL-LHC. To cope with this harsh environment, the entire CMS HCAL, especially the HE which is mostly affected by the radiation damage, should undergo some upgrade to be able to perform successfully. The Phase-I upgrade of the CMS, HB and HE detectors aim to update the photo-sensors and read-out electronics instead of replacing the scintillator tiles. While during 2017 (2018), in the HE detector, HPD's in one (all) read-out box (corresponding to a ϕ-sector of 20°) are replaced by the Silicon Photo Multiplier (SiPM), in the long shut down 2 (LS2) in 2019, all the HPD's in the HB will be replaced by SiPM. The advantage of choosing SiPM over the HPD is that it has better photon detection efficiency of around 28-35%, and a very high gain $(2.7 - 3.5 \times 10^5)$ that is two orders of magnitude larger than the HPD gain and works under lower voltages compared to the HPD. The SiPMs are also smaller in size leading to more channels in the same physical space. For 2016 data taking period, the HB has 1 or 2 longitudinal segmentation depending on the iη while HE has a maximum of 3 longitudinal segmentation. In 2017 and 2018, finer longitudinal segmentations are introduced for HE as shown in the Figure 4.5. QIE8 is also replaced by QIE11 in order to match the larger SiPM gain compared to HPDs. QIE11 has a 17-bit dynamic range with 8-bit read-out while QIE8 had a 14-bit dynamic range with 7-bit read-out.

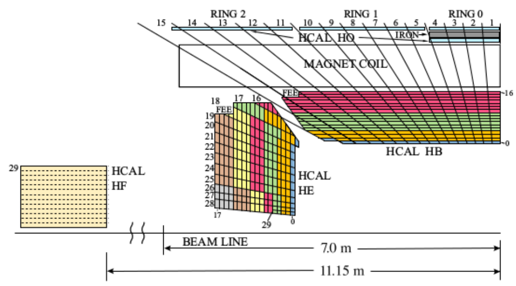

Figure 4.5: The HCAL segmentation in the upgraded system. Layers with same colour are read-out together.

Chapter 5

Particle-Flow

In CMS, the Particle flow (PF) algorithm is used to reconstruct and identify each particle separately using information from all the sub-detectors. For example, to reconstruct an electron, a compatibility is checked between the ECAL energy and momenta of tracks reconstructed in the Tracker. In case of good matching, the particle is identified as an electron, otherwise it is identified as either a neutral hadron or photon. Muons are minimum ionizing particles and can pass through the ECAL and HCAL with negligible energy deposit to reach the muon chambers. Once all the stable particles are identified they are combined to form more complex objects like tau, jet, missing transverse energy etc. A brief description of the PF algorithm is given in the following subsections.

5.1 Tracks and vertices

In CMS, tracks are reconstructed using the Combinatorial Track Finder (CTF) software which is an extension of Kalman filter [31]. The collection of reconstructed tracks is produced in an iterative way. The first iteration looks for prompt tracks with $p_T > 0.8$ GeV and three pixel hits followed by the next iterations with loose selection criteria on p_T

and hits and finally it searches for tracks outside the beam spot. While constructing the tracks, first step is to find out the seeds with three or two 3-D hits near the beam spot. Once the seed is found out, track-finding starts by adding hits through extrapolating the current hit positions in successive layers under uniform magnetic field. This procedure continues until a termination condition is satisfied. In the third step, track candidates are fitted with a Kalman filter to provide an estimate of the track trajectory parameters. In this step, material effect as well as inhomogeneous magnetic field is also considered while fitting the track components. Finally, all the reconstructed tracks are selected through some quality requirements like number of hits, χ^2/dof value, compatibility with the primary vertex etc. to reject fake track candidates. For isolated muons with $1 < p_T < 100$ GeV, the tracking efficiency is $> 99\%$ over the entire η range of Tracker acceptance.

Primary vertex (PV) is formed from the reconstructed tracks. Prompt tracks, selected on the basis of transverse impact parameter significance with respect to the beam line, number of hits in pixel and strip layers, normalized χ^2 value, are clustered with $|\Delta z| <$ 1 cm between the nearest neighbours, where z is the closest approach to the beam line. All the clusters containing at least two tracks are fitted with adaptive vertex [32] fit to estimate the vertex parameters such as the position of vertex. For each track, a weight between 0 and 1 is assigned depending on the compatibility of the concerned track with the vertex. The number of degrees of freedom of the vertex is computed as

$$n_{dof} = 2 \sum_{i=1}^{Tracks} w_i - 3 \tag{5.1.0.1}$$

The number of degrees of freedom of the vertex is used to select the real proton-proton collisions. The PV efficiency is close to 100% with more than two tracks with $p_T > 0.5$ GeV. While the PV is associated with the hard scattering process in the event, there are also other soft vertices (with low p_T tracks) present which are attributed to the underlying events which is not coming from hard scattering and also the effects of events coming from before and after the considered bunch crossing. These vertices are known as pile up (PU) vertices. In 2016 data taking period, there are on average around 25 PU vertices in an event.

5.2 Muon

Muons are reconstructed using the tracks reconstructed in the Tracker and the Muon Chambers [33]. The standalone muons, which are formed using tracks in the Muon Chambers only, have worse momentum resolution and higher admixture of cosmic-ray muons. For each standalone muon track, if there is a corresponding track in the Tracker, a global-muon track is fitted combining hits from the muon chamber and tracker using the Kalman-filter technique and the reconstructed muon is called a Global Muon which has improved momentum resolution for $p_T > 200$ GeV. For all tracks with $p_T > 0.5$ GeV and total momentum $p > 2.5$ GeV, if there is at least one muon segment which is matched to the extrapolated tracks to the Muon Chamber, the corresponding reconstructed muon is called a Tracker Muon which is efficient for $p < 5$ GeV. In 99% cases, reconstructed muons are identified as Global or Tracker Muons and very often as both within the geometrical acceptance of the muon system. To identify the muons, different working points like Loose, Medium and Tight with different reconstruction efficiency are chosen, based on the type of reconstructed muons, normalized χ^2 value of global track, muon segment compatibility, number of hits in the pixel tracker, the impact parameter in longitudinal and transverse plane etc. The reconstruction efficiency of muon with p_T larger than a few GeV is above 95% within $|\eta| < 2.4$ while the mis-identification probability is below 1%.

5.3 Electron

Electrons are reconstructed [34] from energy deposited in the ECAL and tracks in the Tracker. Due to the bremsstrahlung radiation, which depends on material budget in front of the ECAL, electron energy spreads along the ϕ direction with negligible spread in the η direction except for very low p_T (≈ 5 GeV) electrons. The "hybrid" clustering algorithm is used to reconstruct the energy in the barrel whereas "multi-5$\times$5" algorithm, is used in the endcap. In the "hybrid" algorithm, array of 5×1 crystals in $\eta \times \phi$ are added around the most energetic crystal, known as seed crystal with $E_T > 1$ GeV in 17 steps in both directions of ϕ. The contiguous arrays are grouped into clusters, with each distinct cluster having a seed array with energy greater than 0.35 GeV in order to be collected in the final global cluster, called the supercluster (SC). In the "multi 5×5" algorithm, seed crystals are the ones with local maximum energy (> 0.18 GeV) relative to their four direct neighbours. Around these seeds, the energy is collected in clusters of 5×5 crystals that can partly overlap. These clusters with total energy greater than 1 GeV are then grouped into an SC within $|\Delta\eta| = 0.07$ and $|\Delta\phi| = 0.3$ around the seed crystal. The energy-weighted positions of all clusters belonging to an SC are then extrapolated to the planes of the preshower, with the most energetic cluster used as reference point. The maximum distance in ϕ between the clusters and their reference points is used to define the preshower clustering range along ϕ extended by ± 0.15 rad. The range along η is set to 0.15 in both directions. The preshower energies within these ranges around the reference point are then added to the SC energy.

Electron tracks are reconstructed in two steps: seeding and tracking. In seeding, two or three hits in the tracker, from which track can be initiated, are found out in two complementary ways: ECAL-based and tracker-based seeding. In ECAL-based seeding, energy weighted SC positions are used to extrapolate the electron trajectory towards the collision vertex with both positive and negative charge hypotheses. A first compatible hit is then looked for in the innermost (barrel) pixel layer or in the next layer in case no matching is found, within a loose $\Delta\phi$ window and loose Δz interval. The predicted trajectory is then propagated for next compatible pixel hits in the next layers successively. In tracker-based seeding, electron trajectory is reconstructed up to the ECAL, using the KF approach when bremsstrahlung is negligible, with the direction compatible with the position of the ECAL cluster.

The tracking phase is composed of track building and track fitting. The energy loss of electrons in the tracker material does not follow a Gaussian distribution, but a Bethe-Heitler distribution, which has a longer tail. The Gaussian Sum Filter (GSF) algorithm is used to estimate the track parameters from a hit collection obtained with a KF algorithm, by approximating the Bethe-Heitler distribution with a weighted sum of Gaussian distributions. Tracks and superclusters are matched to each other in GSF electron candidates. For an ECAL-seeded electron, the ECAL cluster associated to the track is the one which led to the seed with some geometrical matching between the extrapolated position of the GSF track at the ECAL surface and the SC with $\Delta\eta < 0.02$ and $\Delta\phi < 0.15$. For a tracker-seeded electron, the association is done with a multivariate technique that combines track and supercluster information.

To separate prompt electrons from background sources like photon conversion, jet mis-identification etc., an MVA based Boosted Decision Tree (BDT) discriminator is used. The BDT uses the observables that measure the agreement between ECAL and tracker measurements, compares the energy deposited in ECAL and HCAL and checks the compatibility between GSF and KF tracks. Depending on different electron identification efficiencies (80%, 90%), different threshold values of BDT output have been chosen in CMS.

5.3.1 Isolation variable for Electron and Muon

The isolation variable measures the energy deposited by PF candidates within a cone of fixed radius in η–ϕ plane, which is 0.4 in this case, around the concerned particle. This is defined by the following equation:

$$Iso = \sum p_T^{charged} + \max[0, \sum p_T^{neutralhad} + \sum p_T^{\gamma} - \Delta\beta \times \sum p_T^{PU}] \qquad (5.3.1.1)$$

where,

$\sum p_T^{charged}$ = Energy deposited by charged particles coming from PV, within the isolation cone.

$\sum p_T^{neutralhad}$, p_T^{γ} = Energy deposited by neutral hadrons and $\gamma's$ within the isolation cone.

$\Delta\beta$ = A correction applied to estimate the neutral component coming from PU.

$\sum p_T^{PU}$ = Energy deposited by charged particles coming from PU, within the isolation cone.

The charged component of isolation, coming from PU, can be easily separated from that coming from PV. However, the neutral component cannot be separated as the neutral particles cannot be associated with the PV. To estimate the contribution of the neutral component from PU, a parametrized correction, known as $\Delta\beta$ correction, is used. The value of $\Delta\beta$ is 0.5, which is the ratio of neutral-to-charged pions. The ratio of isolation value to the p_T of concerned particle is known as Relative Isolation.

5.4 Jets

Quarks or gluons cannot be seen as coloured state, rather they manifest themselves as a collimated spray of stable particles, known as jets, through fragmentation and hadronisation. The $anti - k_t$ algorithm, which is infrared and collinear safe, is used to cluster the stable particles into a jet. It proceeds by defining distances d_{ij} between two entities (particles, pseudo-jets) i and j, and distances d_{iB} between an entity i and the beam

B in momentum space:

$$d_{ij} = min(\frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2}) \times \frac{R_{ij}^2}{R} \tag{5.4.0.2}$$

$$d_{iB} = \frac{1}{p_{Ti}^2} \tag{5.4.0.3}$$

where R_{ij} is the distance between two entities i and j, and R whose value is 0.4, is the *radius* parameter of the jet. If the smallest distance is d_{ij}, two entities i, j are merged into a single entity, otherwise i is considered as a stable jet and removed from the list of entities. This process is repeated until all the entities are considered into a stable jet. Once jets are formed, different jet correction factors (JEC) are applied to mitigate the effects of pile up, underlying events, balance the momentum in transverse direction etc. in order to match the reconstructed jet with generator level jet.

Given that the Higgs boson branching ratio to $b\bar{b}$ is $\approx 58\%$, identification of b jets [35] is very important in Higgs searches. For searches of other Higgs decay modes also, it is equally important to identify b jets to exclude the $t\bar{t}$ events which has a large cross section, ~ 730 pb while cross section of HH production is only 0.033 pb. A Combined Secondary Vertex (CSV) discriminator is used in this thesis to identify a b jet. This method utilizes the finite life time ($c\tau \approx 450\mu m$) information of b quark to separate the jets coming from light quarks. A secondary vertex is defined as a vertex sharing less than 65% of its tracks with the primary vertex and separated radially from the primary vertex with a significance of at least 3σ. The secondary vertex with radial distance more than 2.5 cm from the primary vertex or mass associated with the secondary vertex compatible with K^0 or exceeding $6.5 GeV/c^2$ are rejected. The CSV algorithm is still applicable when no secondary vertices are found. The variables used by this method include the number of tracks in the jet, the significance of the track impact parameter in 3D, mass and number of tracks associated to the secondary vertex etc. For different working points like, Loose, Medium, Tight, different CSV values have been fixed. The working points are defined at fixed mis-identification probability of light parton jets. For example, the mis-identification probability of light parton jets for Loose, Medium and Tight working points are 10%, 1% and 0.1% respectively.

5.5 Tau

The Tau lepton with mass 1.777 GeV is the heaviest lepton and has a finite lifetime of $\sim 290 \times 10^{-15}$ s. Tau is the only lepton which can decay into hadrons as well as to leptons. In about two-third of the cases, tau decays hadronically with one or three charged pions or kaons and up to two neutral pions and tau-neutrinos. In the rest of the cases, it decays to either an electron/muon in association with a tau-neutrino and electron/muon neutrino respectively. Table 5.1 shows the branching ratios of tau.

Table 5.1: Approximate branching ratio of different τ decay modes.

Decay mode	Meson resonance	branching ratio %
$\tau^- \to e^- \overline{\nu}_e \nu_\tau$		17.8
$\tau^- \to \mu^- \overline{\nu}_\mu \nu_\tau$		17.4
$\tau^- \to h^- \nu_\tau$		11.5
$\tau^- \to h^- \pi^0 \nu_\tau$	$\rho(770)$	26.0
$\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$	$a_1(1260)$	9.5
$\tau^- \to h^- h^+ h^- \nu_\tau$	$a_1(1260)$	9.8
$\tau^- \to h^- h^+ h^- \pi^0 \nu_\tau$	$a_1(1260)$	4.8
other modes with hadrons		3.2

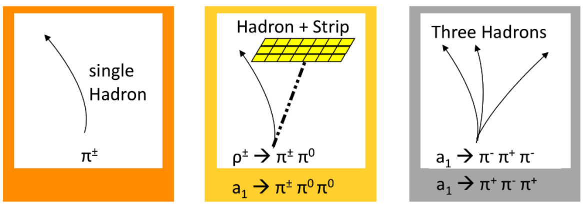

Figure 5.1: Schematic of τ decay mode.

The reconstruction and identification of hadronic tau decay modes (τ_h) is done by the Hadron Plus Strip (HPS) algorithm [36]. The high probability for photons coming from $\pi^0 \to \gamma\gamma$ decays to convert to e^+e^- pairs is accounted for by collecting photon and electron constituents with $p_T > 0.5$ GeV of the jet into dynamic clusters. The clusters are seeded by the most energetic electron and photon and formed by adding the next energetic electrons and photons within a dynamic window defined by:

$$\Delta\eta = f(p_T^{\gamma}) + f(p_T^{strip}) \tag{5.5.0.4}$$

$$\Delta\phi = g(p_T^{\gamma}) + g(p_T^{strip}) \tag{5.5.0.5}$$

The functions f and g are determined from simulation using a single τ gun sample such that 95% electrons and photons coming from the τ decay is confined within the strip. The position of the strip is defined by p_T weighted average of all constituents inside the strip.

$$\eta_{strip} = \frac{1}{p_T^{strip}} \sum p_T^{\gamma} \eta_{\gamma} \tag{5.5.0.6}$$

$$\phi_{strip} = \frac{1}{p_T^{strip}} \sum p_T^{\gamma} \phi_{\gamma} \tag{5.5.0.7}$$

with $p_T^{strip} = \sum p_T^{\gamma}$. This process continues until no more electrons or photons are found to be added into the strip and it starts to cluster new strips. Charged particles with $p_T > 0.5$ GeV and with distance of closest approach between their tracks and τ production vertex, less than 0.4 cm along z direction and < 0.03 cm in the transverse plane, are considered as hadronic tau decay products. Finally, the mass window cut is applied to define the tau decay mode consistent with the decay channels as mentioned in the Table 5.1 in following way :

- $h^{\pm} h^{\mp} h^{\pm}$: combination of three charged particles with mass $0.8 < m_{\tau_h} < 1.5$ GeV.
- $h^{\pm} \pi^0 \pi^0$: combination of single charged particle and two strips, with total mass $0.4 < m_{\tau_h} < 1.2\sqrt{p_T[GeV]/100}$ GeV with upper limit on the mass window constrained to be at least 1.2 GeV and at most 4.0 GeV.

- $h^{\pm}\pi^0$: combination of one charged particle and one strip with mass $0.3 < m_{\tau_h} < 1.3\sqrt{p_T[GeV]/100}$. The upper limit on the mass window is constrained to be at least 1.3 GeV and at most 4.2 GeV.

- $h^{\pm}$: a single charged particle with no strips and mass of a charged pion.

5.5.1 Tau-isolation and Discriminator against jet, electron, muon

Tau-isolation is a good discriminator to separate τ_h from gluon and quark jets as the decay products of τ_h are more collimated along the tau decay axis with less hadronic activity around. Tau-isolation is defined as the scalar sum of p_T of all charged particles and photons with $p_T > 0.5$ GeV within an isolation cone of $\Delta R = 0.4$. The isolation variable is defined as,

$$\begin{aligned} I_{\tau_h} &= \sum p_{T_1}^{ch}(\Delta z < 0.2cm, \Delta R < 0.4) + \\ & max(\sum p_T^{\gamma} - \Delta\beta \sum p_{T_2}^{ch}(\Delta z > 0.2cm, \Delta R < 0.8), 0) \end{aligned} \tag{5.5.1.1}$$

where,

- $\sum p_{T_1}^{ch} =$ Energy deposited within the isolation cone by charged particles coming from the τ_h production vertex with $|\Delta z| < 0.2$ cm and $\Delta xy < 0.03$ cm.

- $\sum p_T^{\gamma} =$ The neutral component of isolation within the isolation cone.

- $\Delta\beta =$ The correction applied to estimate the neutral component, not coming from the τ_h production vertex.

- $\sum p_{T_2}^{ch} =$ Same as $p_{T_1}^{ch}$ but not coming from the τ_h production vertex *i.e* $|\Delta z| > 0.2$ cm and $\Delta xy > 0.03$ cm, within $\Delta R < 0.8$.

As it is not possible to associate the neutral component with the τ_h production vertex as described in section 5.3.1, $\Delta\beta$ correction is introduced and it's value is estimated to be

0.2 in Run2.

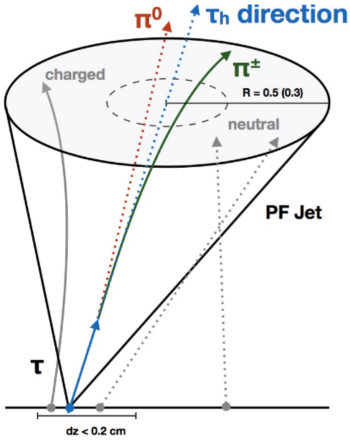

Figure 5.2: Schematic of τ_h identification using the isolation variable.

In Run2, to reduce the jet $\to \tau_h$ mis-identification, another cut on the p_T sum of e/γ which are included in the strips to reconstruct the τ_h but outside the signal cone $R_{sig} = 3.0/p_T$ is applied as follows:

$$p_T^{strip,outer} = \sum p_T^{e/\gamma}(\Delta R > R_{sig}) < 0.10.p_T^{\tau} \tag{5.5.1.2}$$

where maximum(minimum) value of R_{sig} is 0.1(0.05). The signal cone, R_{sig}, is defined around the leading track and the energy within that cone is excluded while calculating the isolation $i.e$ all the decay products of τ_h are assumed to be confined within the signal cone.

A multivariate BDT discriminator is used to reduce the τ_h mis-identification. The inputs to BDT training include tau lifetime information, τ_h isolation value, τ_h decay mode compatible with τ_h decay channels as mentioned in Table 5.1, transverse impact parameter d_0 of the leading track, distance between τ_h production and decay vertices etc. Depending on the τ_h identification efficiency, different working points of MVA are chosen: VLoose, Loose, Tight, VTight. Figure 5.3 shows the efficiency of the τ_h identification and mis-identification probability as a function of τ_h p_T in Run-2 where τ_h isolation value as described above is one of the input variables to BDT. The working points of the MVA isolation discriminator, corresponding to different τ_h identification efficiencies, are defined by cuts on the BDT discriminant. For a given working point, the threshold on the BDT discriminant is adjusted as a function of the transverse momentum of the τ_h candidate to ensure a uniform efficiency over p_T. τ_h passing a given working point is called isolated, otherwise anti-isolated. Similarly, the region containing isolated $\tau_h' s$ only is known as isolated, otherwise anti-isolated.

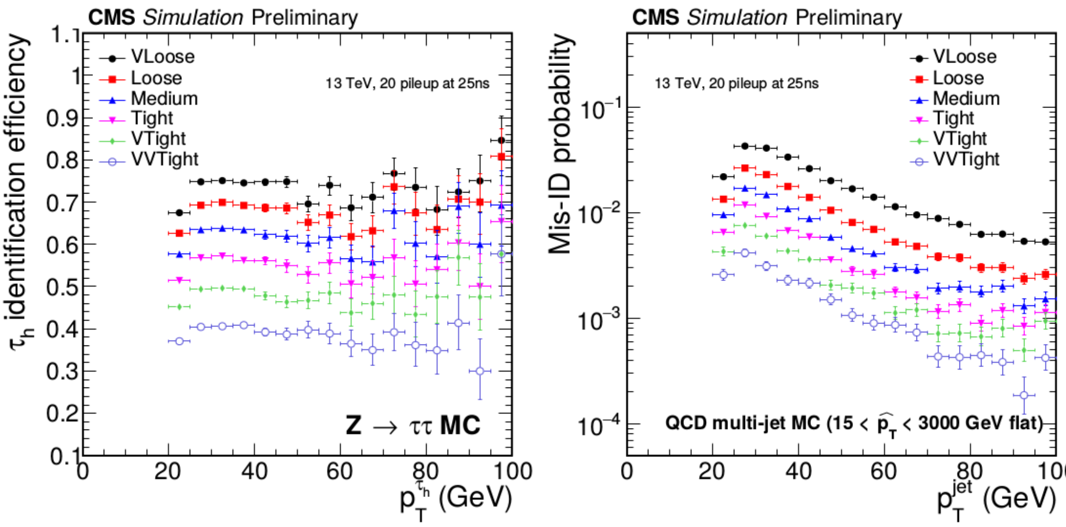

Figure 5.3: Efficiency of τ_h identification estimated with $Z/\gamma^* \rightarrow \tau\tau$ events (left) and mis-identification probability estimated with simulated QCD multi-jet events (right) for the Very Loose, Loose, Medium, Tight, Very Tight, and Very Very Tight working points of the MVA based τ_h isolation algorithm [37].

There is a probability for electron and muon to be identified in $h^{\pm}\pi^0, h^{\pm}\pi^0\pi^0, h^{\pm}$ decay mode. To reduce the e$\rightarrow$ τ_h mis-identification, another set of BDT variables which are sensitive to e/γ shower algorithm, is used. Similarly for $\mu \rightarrow \tau_h$ mis-identification, the variables related to muon chambers are used in BDT. In all cases different working points of BDT have been chosen corresponding to different τ_h identification efficiency.

Chapter 6

Higgs pair production: separation of VBF process from ggF process

6.1 Higgs pair production

In this chapter, a simulation study of Higgs pair (or di-Higgs) production in proton-proton collisions at a center of mass energy of 14 TeV is presented targeting the High Luminosity (HL) LHC [38, 39, 40] data-taking phase. As already discussed, Higgs boson pair production in the SM can proceed in several ways: (a) gluon-gluon fusion (ggF), (b) vector boson Fusion (VBF), and (c) associated production (AP) with a vector boson, W or Z. Figure 6.1 shows the Feynman diagrams for di-Higgs production through ggF and VBF process respectively.

The cross section (σ) of ggF provides the largest contribution to the overall Higgs pair production cross section of $\sigma_{HH} = 33.49^{+4.3}_{-0.6}(\text{scale}) \pm 5.9(\text{theo})$ fb. The next dominant process is VBF whose cross section is about an order of magnitude smaller than that of ggF. At the HL-LHC, the VBF process is expected to become significant because of an order of magnitude increase in the instantaneous luminosity which will result in a yearly

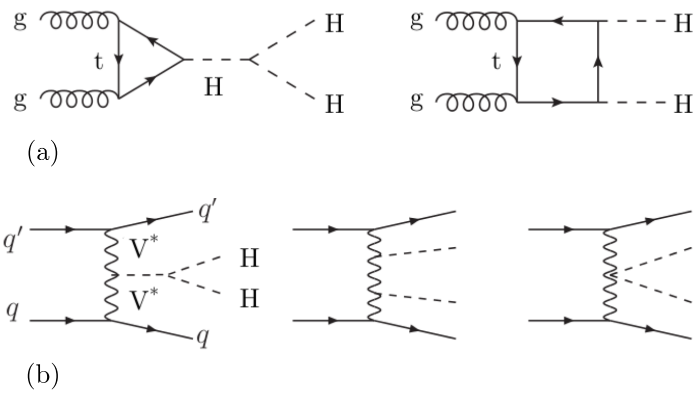

Figure 6.1: Feynman diagrams for the (a) ggF and (b) VBF processes.

integrated luminosity of 300 fb^{-1}.

Presence of two forward jets will make the VBF process distinguishable from other di-Higgs production as well as SM background processes. Selection of Higgs pair events produced in the VBF process with two forward jets and separation of VBF against ggF processes using cut based as well as multivariate approaches are studied in this chapter. Use of helicity angular distributions of the Higgs bosons and their decay products to separate events produced through the two production mechanisms efficiently is also discussed in detail.

For the present analysis, 100000 di-Higgs events at the leading order (LO) are produced in the LHE format [41] using MadGraph5_aMCNLO for both VBF and ggF production processes. As the cross section of the associated production is very small, the process is not considered in this study. We consider VBF as signal and ggF as background in this study.

One of the Higgs bosons is allowed to decay to a pair of bottom (b) quarks while the other to a pair of tau (τ) leptons at the tree level in the event generation step. The generated events are hadronized using Pythia8 [42] followed by detector simulation and event reconstruction with Delphes version 3 [43], using the environment of the CMS detector. Delphes is a modular framework that simulates particle response inside a detector in a parametrized fashion. Delphes is very useful for doing quick analysis but may not be suitable for precise measurements.

The study of the helicity angles has been used for $HH \to b\bar{b}\gamma\gamma$ analysis in CMS using 2016 data, which is also described in this chapter.

6.2 Angular Variables

The helicity angles [44], namely, $\cos\theta^*$, ϕ_1, ϕ, $\cos\theta_1$ and $\cos\theta_2$ can be used to find spin and parity of the particles involved [45]. The corresponding distributions may be studied to disentangle the signal (VBF) from the background (ggF). The first two helicity angles are related to the production mechanism while the other three angles depend on the decay chain. A pictorial description of all these angles is given in Figure 6.2 and the definitions of all the angles are given below:

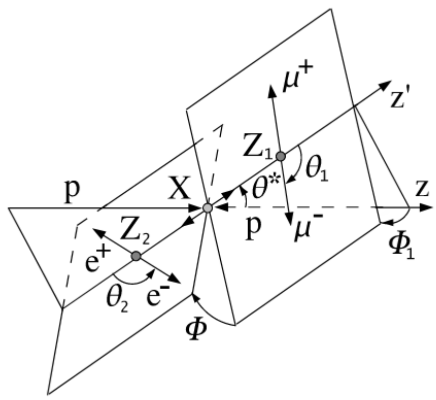

Figure 6.2: Illustration of helicity angles where the resonant particle X decays to Z_1 and Z_2 which subsequently decay to $\mu^+\mu^{-1}$ and e^+e^- pair respectively.

θ^* = Direction of one of the Higgs bosons in the di-Higgs system where the Higgs boson is boosted in the di-Higgs rest frame. Distribution of θ^* is the angular observable that contains information about the spin and parity properties of the resonant particle. For example, the distribution of θ^* would be flat for a spin-0 particle as no spin correlation is involved. However, the angle may be helpful to disentangle signal and background.

$\phi_1 =$ Angle between the decay plane of one of the Higgs bosons and the plane containing one of the incident partons and the Higgs boson under consideration, where all the particles are boosted in the di-Higgs rest frame.

$\phi =$ Angle between the decay planes of the two Higgs bosons measured in the di-Higgs rest frame.

$\theta_i =$ Angle between one of the decay products of one Higgs boson (i) and the opposite direction of the other Higgs boson, where all the particles are boosted in the rest frame of the Higgs boson i.

The angles, ϕ and θ_i, can also be used to infer the spin of resonant particle. For example, the distributions of all these angles would be flat if the decay particles come from a spin-0 particle. The angles can be expressed as

$$\phi_1 = \frac{h_1 \cdot (z' \times n_1)}{|h_1 \cdot (z' \times n_1)|} \cdot \cos^{-1}(z' \cdot n_1) \tag{6.2.0.1}$$

$$\phi = \frac{-h_1 \cdot (n_1 \times n_2)}{|h_1 \cdot (n_1 \times n_2)|} \cdot \cos^{-1}(n_1 \cdot n_2) \tag{6.2.0.2}$$

where

$$z' = \frac{p_1 \times h_1}{|p_1 \times h_1|} \tag{6.2.0.3}$$

p_1 is the momentum of one of the incident partons boosted in the di-Higgs rest frame and h_1 is the momentum of one of the Higgs bosons in di-Higgs rest frame.

$$n_1 = \frac{q_{11} \times q_{12}}{|q_{11} \times q_{12}|}, \tag{6.2.0.4}$$

$$n_2 = \frac{q_{21} \times q_{22}}{|q_{21} \times q_{22}|} \tag{6.2.0.5}$$

where q_{i1}, q_{i2} are the momenta of the two daughters of the Higgs boson i. All the particles are boosted in the di-Higgs rest frame.

$$\cos\theta_{q_1} = \frac{q_1 \cdot (-h_2)}{|q_1| \times |h_2|} \qquad (6.2.0.6)$$

where q_i is the momentum of one of the daughters of the Higgs boson i.

6.3 Selection of the VBF process

Three different methods are explored to select the VBF events, in order to find the optimum one, as described below.

Jet pair with maximum p_T : The first two highest p_T reconstructed jets are matched to quarks that give rise to the VBF jets at the generator level. Each reconstructed jet is matched to a particle which comes from a quark whose status corresponds to the outgoing particles of the hardest sub-process, also requiring that the quark should not come from Higgs decay, which might be case for b quarks. The reconstructed jets are said to be matched if the separation between the jet and particle in the $\eta - \phi$ plane is < 0.5. Figure 6.3 shows the p_T distribution of the quark corresponding to the VBF jet and the matched reconstructed jets.

Jet pair with maximum η difference : Two reconstructed jets with maximum $\Delta\eta$ are matched with quarks corresponding to the VBF jets. The matching requires that the separation in the $\eta - \phi$ plane between the reconstructed jets and particles is within 0.5. The $\Delta\eta$ between the two quarks which give rise to the two VBF jets and maximum $\Delta\eta$ between the two matched reconstructed jets are shown in the Figure 6.4.

Jet pair with maximum invariant mass : Jet pair for which the invariant mass is maximum is matched with the outgoing quarks corresponding to the VBF jets.

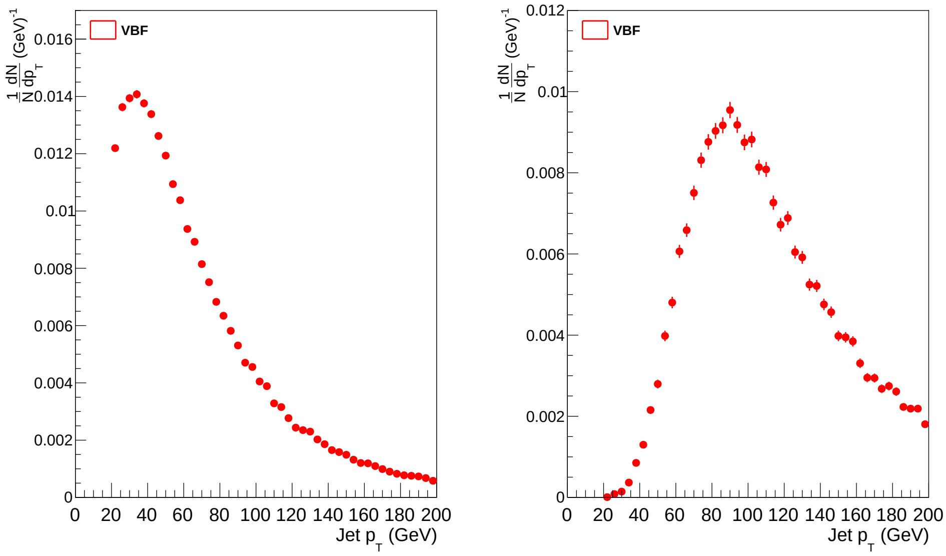

Figure 6.3: (Left) p_T distribution of the quarks giving rise to the two VBF jets. (Right) p_T distribution of the leading and sub-leading reconstructed jets, matched with the parent quarks.

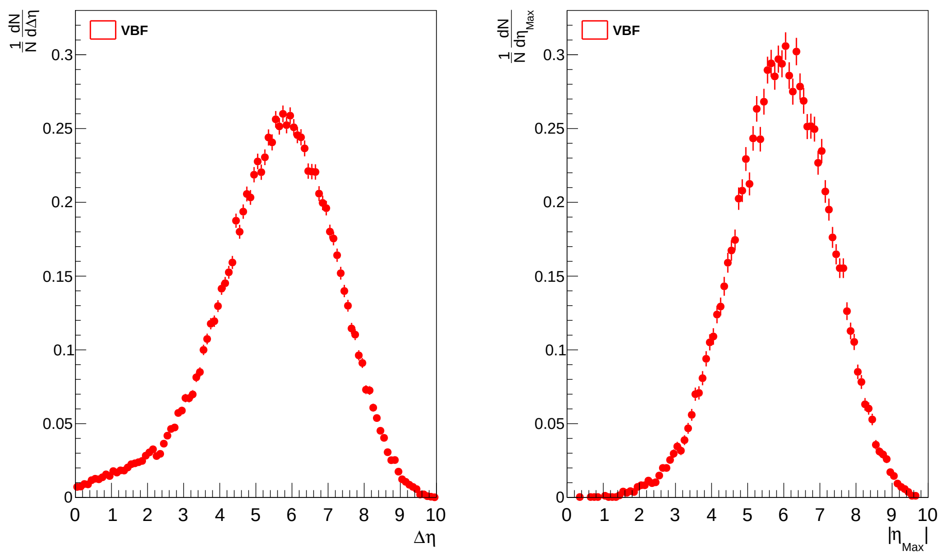

Figure 6.4: (Left) $|\Delta\eta|$ between the VBF quarks. (Right) reconstructed jets with maximum $\Delta\eta$ separation and matched with VBF quarks.

Figure 6.5 shows the invariant mass distribution of the quark pair and reconstructed jet pair matched with the quarks corresponding to the VBF jets.

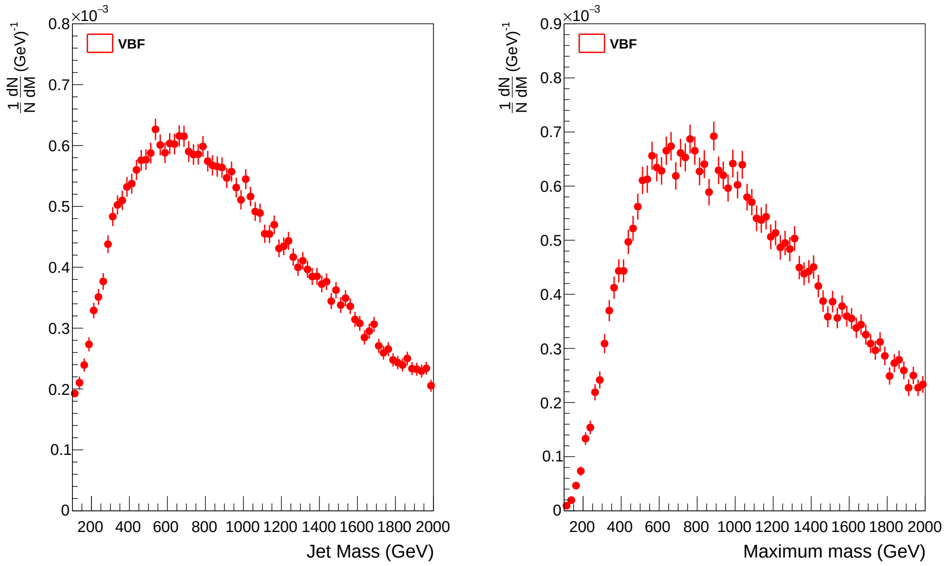

Figure 6.5: (Left) Invariant mass distribution of the two VBF quarks. (Right) Invariant mass distribution of the two jets matched with the VBF quarks and having maximum invariant mass.

The best matching is obtained for the jet pair with maximum invariant mass, in 48% cases followed by the matching by maximum η difference between the two reconstructed jets in 43% cases and finally, in 7% cases, two highest p_T jets can be matched with the VBF jets. Two well separated forward jets with high invariant mass can be used to select di-Higgs events produced in VBF process [47].

6.4 Separation of VBF and ggF events

Events independent of VBF and ggF are selected with two τ leptons and two b jets coming from the Higgs boson decay. All the decay modes of τ are considered. τ is considered to have decayed leptonically if a reconstructed lepton, electron or muon, is matched with the generator level lepton coming from τ decay. Electrons are selected within the acceptance of the CMS Tracker ($|\eta| < 2.5$), while muons are selected requiring $|\eta| < 2.4$, with $p_T >$ 10 GeV in both cases. For a hadronically decaying τ, reconstructed jets are matched with the generator level τ coming from the Higgs boson decay within a distance of 0.5 in the $\eta - \phi$ plane. Similarly, b jets are selected if there is a matching between the reconstructed jets and generator level b quarks which come from Higgs decay. Both b and hadronic τ's are selected with $p_T > 20$ GeV and $|\eta| < 2.3$ and 2.5 for τ and b jets respectively. All the selected particles are required to be separated from each other by a minimum distance of 0.5 in the η–ϕ plane. To select the VBF jets, a jet pair is selected with invariant mass > 300 GeV and $|\Delta\eta| > 3$. The efficiency of VBF and ggF events passing all the cuts are 21% and 7% respectively. These cuts are referred to as pre-selection cuts in Table 6.1. In addition to these cuts, an analysis is done by selecting the jet pair with invariant mass > 300 GeV, 400 GeV and $\Delta\eta > 3$, 3.5, 4 for each mass cut to increase the signal purity due to the presence of two forward jets with high invariant mass. Table 6.1 shows the signal and background efficiencies for the respective cuts. As seen in Table 6.1, a maximum background rejection efficiency (= 1-efficiency) of 47% can be obtained with a signal efficiency of 84% for mass cut > 400 GeV and $|\Delta\eta|$ cut > 4.

Table 6.1: Efficiency (in %) for events produced through the VBF and ggF processes and after passing the pre-selection cuts, for different mass and $|\Delta\eta|$ cuts for the additional jet pair.

Mass cut	300 GeV		
$\|\Delta\eta\|$ cut	3	3.5	4
Efficiency in VBF	100	95	88
Efficiency in ggF	100	83	64

Mass cut	400 GeV		
$\|\Delta\eta\|$ cut	3	3.5	4
Efficiency in VBF	92	89	84
Efficiency in ggF	77	66	53

To improve separation between signal (VBF) and background (ggF) further, events, which pass pre-selection cuts, are fed into a Boosted Decision Tree (BDT). A BDT is a multivariate analysis technique for event classification as signal or background combining several variables and their correlation [48]. The decision tree consists of a set of sequential binary splits of the data where splitting is based on finding the best variables and the corresponding values which can separate signal and background maximally. This procedure continues until some convergence criteria is reached where ending nodes are called leaves and if more than half of the events of a leaf corresponds to signal, the leaf is tagged as signal leaf otherwise as background leaf.

Figures 6.6, 6.7, 6.8 show the distribution of all the angular variables defined in the earlier section for events produced through the VBF and ggF processes after implementing the kinematic and ΔR cuts. From these figures, it can be seen that only the $\cos\theta^*$ distribution has significant power to disentangle the VBF and ggF production processes as it is sensitive to the tensor structure of the production mechanism [45, 46] resulting in distinct distributions for VBF and ggF. The distribution is relatively flat for ggF while it peaks around $\pm$ 1 for VBF. Since Higgs is a spin zero particle, the angular distribution of the decay products should be independent of any directions which is reflected in the distributions of ϕ, $\cos\theta_\tau$ and $\cos\theta_b$.

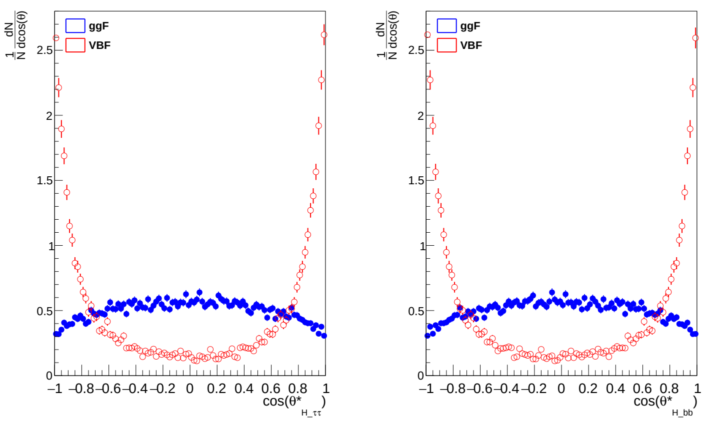

Figure 6.6: Distribution of $\cos\theta^*$ for both the Higgs bosons.

The following variables are used in the BDT,

- $\cos\theta^*$,

- invariant mass of the selected jet pair,

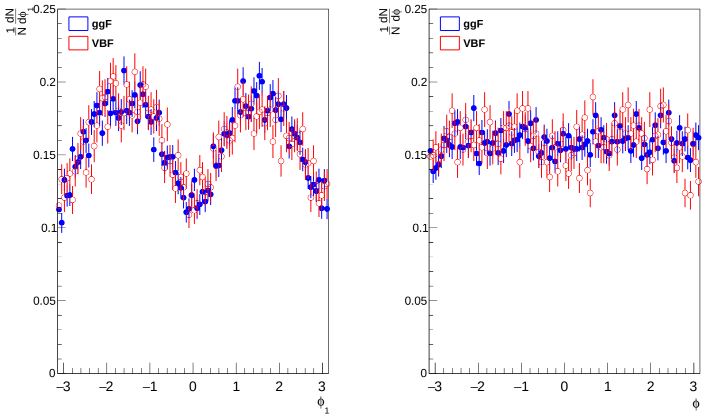

Figure 6.7: Distribution of (Left) ϕ_1, (Right) ϕ.

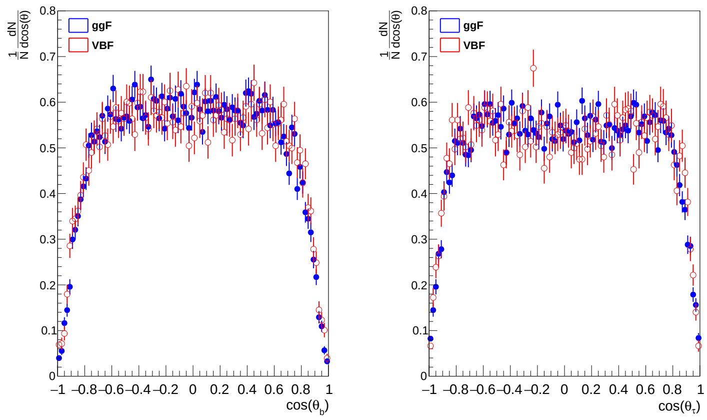

Figure 6.8: Distribution of (Left) $\cos\theta_b$, (Right) $\cos\theta_\tau$.

- $\Delta\eta$ between the selected jet pair.

The BDT is trained with 13581 signal and 6214 background events where VBF and ggF events are treated as signal and background respectively and is tested with an independent set of 3395 and 1553 events for signal and background respectively. Figure 6.9 shows the output response of the BDT classifier and background rejection versus signal efficiency where we can see that for $\sim 84\%$ signal efficiency, background rejection is $\sim 64\%$ which gives about 25% better separation power compared to the cut based analysis presented earlier in Table 6.1.

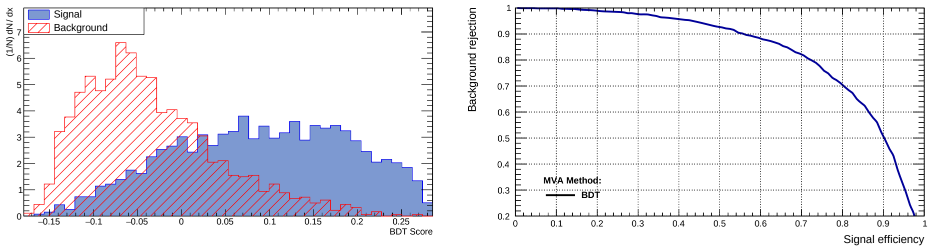

Figure 6.9: (Left) Output of the BDT classifier. (Right) Background rejection versus signal efficiency obtained from BDT training.

6.5 Application of helicity angles in $HH \to \gamma\gamma b\bar{b}$ analysis

The distributions of helicity angles are used in the $HH \to \gamma\gamma b\bar{b}$ analysis to improve signal-background separation using BDT training in 2016 data collected by CMS. For this analysis, events are selected using double-photon triggers, which require two photons with $p_T^{\gamma^1} > 30$GeV and $p_T^{\gamma^2} > 18$GeV for the leading and sub-leading photons, respectively. In addition, a number of pre-selection cuts, as described in Table 6.2, are applied. These cuts are based on shower shape variables,

- R9, which is defined as the ratio between the energy deposited on a 3×3 ECAL crystal matrix around the most energetic crystal in the superclusters and the super-crystal energy,

- isolation variables CHI, which is the energy sum of all charged hadron particle-flow candidate inside a cone of $\Delta R < 0.3$ around the photon axis,
- identification variables, H/E, the ratio between the photon's energy deposit in HCAL and in ECAL
- kinematic variables, namely, E_T and η.

Table 6.2: pre-selection cuts applied on di-photon candidates.

Requirements	Leading Photon	Sub-leading Photon
E_T	30 GeV	20 GeV
$E_T/M(\gamma\gamma)$	$> 1/3$	$> 1/4$
$\|\eta\|$	< 2.5 and outside $1.442 < \|\eta\| < 1.566$	
Shower shape and Isolation	R9 > 0.8 or $CHI < 20$ or $CHI/E_T < 0.3$	
Identification	H/E < 0.08	

Events containing at least one di-photon candidate passing the pre-selection requirements and with di-photon invariant mass ($M_{\gamma\gamma}$), $100 < M_{\gamma\gamma} < 180\,\text{GeV}$ are considered for the analysis. Photons are identified using a multivariate technique where p_T of the electromagnetic shower, its longitudinal leakage into the HCAL and isolation variables are used as the input. The score of the multivariate ID is selected such that efficiency of photon identification is 90% both in barrel and endcap. If more than two photons are found in an event, the photon pair with the highest $p_T^{\gamma\gamma}$ is retained.

b jets are reconstructed using the anti-k_t jet clustering algorithm with radius parameter, R = 0.4. Jets with $p_T > 25$ GeV and $|\eta| < 2.4$ are considered for this analysis. Selected jets should be separated from each other by a distance $\Delta R_{\gamma j} > 0.4$ where $\Delta R_{\gamma j}$ is the separation of selected b jet and γ in $\eta - \phi$ plane. In addition, b tagging discriminator is also used to select the b jets. In case more than two jets are found in an event, the di-jet constructed with the two highest b tagging scores is selected. Finally, an event is accepted if the invariant mass of the jet pair (m_{jj}) satisfies the condition: $70 < m_{jj} < 190$ GeV.

In this analysis, three helicity angles are used,

- The scattering angle, θ^{CS}_{HH} is defined in the Collins-Soper (CS) [49] frame as the angle between the direction of the $H \to \gamma\gamma$ candidate to the CS frame. The CS frame is defined by placing the z axis half way between axes of two beams in the di-Higgs rest frame to minimize the effect of initial state radiation of incoming partons. The θ^* defined in section 6.2 is not defined in CS frame.

- The Higgs decay angles are defined as the angles of the decay products in each Higgs boson's rest frame with respect to the direction of motion of the boson.

The dominant background contribution comes from nγ + *jets* events. The SM single Higgs boson productions are three orders of magnitude smaller than the dominating background process, nγ + *jets* events. The distributions of three helicity angles after the event selection as described above, are shown in Figure 6.10. Only the single Higgs boson production with a significant number of events estimated from MC are shown for clarity of the figure. Resonant signal events are normalized to cross section of 500 fb and the SM-like ggF di-Higgs events (VBF di-Higgs events) signal to 10^4 (10^5) times its cross section. As seen in Figure 6.10,

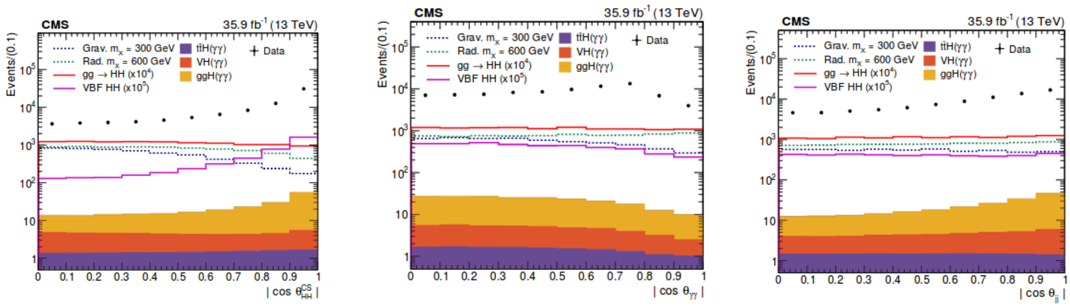

Figure 6.10: Data (dots), dominated by nγ + jets background, compared to different signal hypothesis and three single Higgs boson samples after the selections on photons and jets as described in the text: (left) $|\cos\theta^{CS}_{HH}|$, (middle) $|\cos\theta_{\gamma\gamma}|$ and (right) $|\cos\theta_{jj}|$.

the distribution of $|\cos\theta^{CS}_{HH}|$ is relatively flat for ggF and the spin-0 Radion production. For the spin-2 Graviton, it decreases towards 1 and for VBF and data it rises toward 1. Due to the absence of any spin correlation between the Higgs decay products, the distribution of $H \rightarrow \gamma\gamma$ and $H \rightarrow b\bar{b}$ is expected to be flat, but it decreases to 1 due to the cuts applied on p_T of the decay products.

These angular distributions in addition to the b tagging scores of leading, sub-leading b jets and $p_T(\gamma\gamma)/M(jj\gamma\gamma)$, $p_T(jj)/M(jj\gamma\gamma)$ are used to as input to a BDT where di-Higgs samples are used as signal events and background events are obtained from data by inverting the score of the multivariate ID of one of the two photons. Figure 6.11 shows the BDT output.

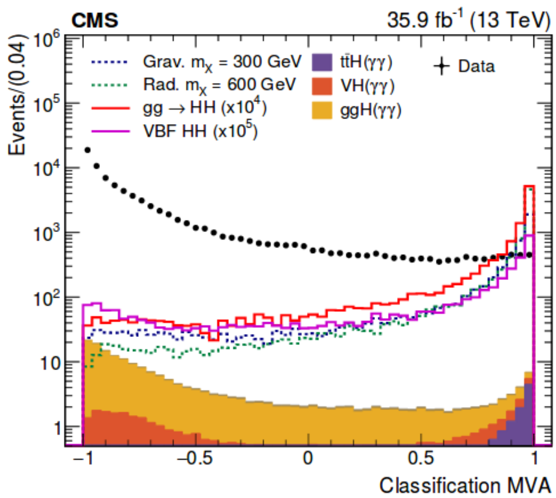

Figure 6.11: Distribution of BDT classifier. Data dominated by nγ+jets background, are compared to different signal hypotheses and three single-Higgs boson samples after the selections on photons and jets as described in the text.

Events from the VBF production are selected less efficiently than those from ggF production as one of the most discriminating variables, $\cos\theta^{CS}_{HH}$ as shown in Figure 6.10 has similar behavior to the n$\gamma + jets$ background and VBF process while it is different for ggF process.

6.6 Conclusion

MC simulation based study of separation of di-Higgs events produced through ggF and VBF processes is presented. It is observed that, VBF events can be selected with high efficiency by taking the jet pair with maximum invariant mass and jet pair with maximum $\Delta\eta$ separation. Five different helicity angles are studied to separate VBF events from ggF events. The angle $\cos\theta^*$ is found to have the highest discriminating power, which, along-with the maximum invariant mass of jet pair and $\Delta\eta$ between the corresponding jets are used in the BDT to maximize the separation power of the analysis. For a signal efficiency of 84%, a background rejection efficiency of 64% is achieved. These variables are unique to the VBF process and can also be used to separate VBF events against SM backgrounds other than ggF events.

The helicity angle distributions can be used to separate the signal and background processes depending on the nature of signal as seen in $b\bar{b}\gamma\gamma$ analysis where ggF di-Higgs events can be separated more efficiently compared to VBF di-Higgs events.

The analysis strategy discussed in this chapter can be put to effective use at the HL-LHC where VBF production will become significant.

Chapter 7

HH search in $\tau\tau\tau\tau$ final state

As discussed in the previous chapter, the discovery of the Higgs boson at the LHC has prompted physicists to take up the challenging task of understanding the Higgs self-potential which can be probed by searching for double Higgs production. Search for HH production has so far been limited to four final states only, namely : $b\bar{b}b\bar{b}$ [50], $bb\gamma\gamma$ [51], $bb\tau\tau$ [52], $bbVV$ [53] (where $V = W$ or Z boson), due to either higher BR or good mass resolution. The $HH \rightarrow \tau\tau\tau\tau$ final state has not been studied previously by CMS and ATLAS. The 4τ final state is affected by a very low BR($\sim$0.4%) which is seen in Figure 7.1. However, the absence of b jets in the final state at the same time involves a relatively smaller background contribution compared to the other final states studied so far. In this chapter, a search for resonant production of HH, where a spin-0 Radion decays to a HH pair, in the 4τ final state has been discussed.

For the time being, only those channels where two τ's decay leptonically (ℓ), where all the combinations of lepton flavours are taken into account, and other two τ's decay hadronically, are considered. The reconstruction of hadronically decaying tau, τ_h, is described in chapter 5, section 5.5, where the charge of τ_h can be measured from decay products. These channels are further divided into same-sign (SS) leptons where the two

SS τ's decay leptonically and opposite-sign (OS) leptons where the two OS τ's decay leptonically, while in both cases the other two τ's decay hadronically, to take advantage of small background contribution in SS lepton channels. These classifications give rise to six independent signal regions

- $\mu^{\pm}\tau_h^{\mp}\mu^{\mp}\tau_h^{\pm}$
- $e^{\pm}\tau_h^{\mp}e^{\mp}\tau_h^{\pm}$
- $\mu^{\pm}(e^{\pm})\tau_h^{\mp}e^{\mp}(\mu^{\mp})\tau_h^{\pm}$
- $\mu^{\pm}\tau_h^{\mp}\mu^{\pm}\tau_h^{\mp}$
- $e^{\pm}\tau_h^{\mp}e^{\pm}\tau_h^{\mp}$
- $\mu^{\pm}(e^{\pm})\tau_h^{\mp}e^{\pm}(\mu^{\pm})\tau_h^{\mp}$

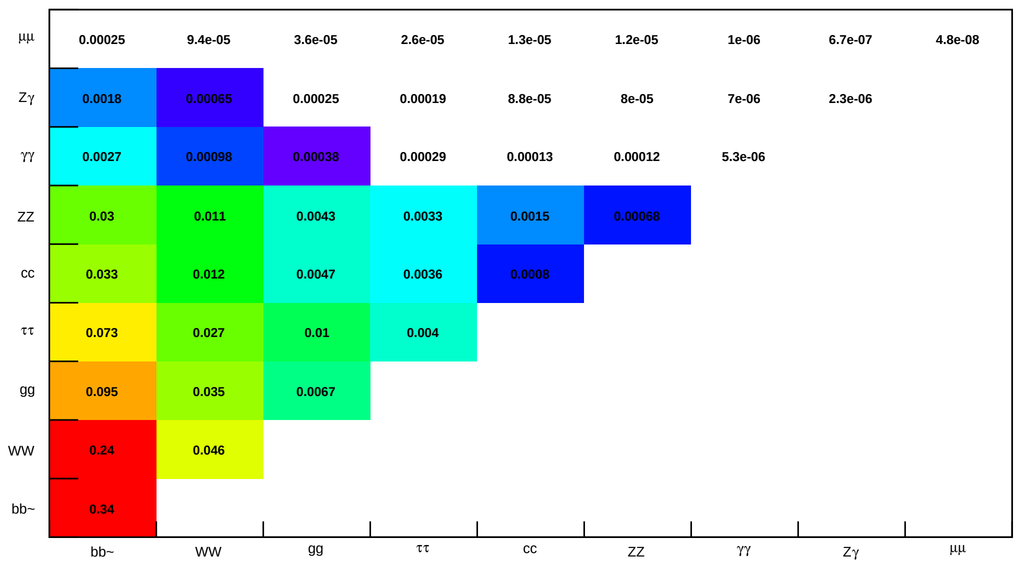

Figure 7.1: HH branching ratio in different final states.

7.1 Data and Monte Carlo Samples

7.1.1 Data Samples

This analysis is based on data corresponding to an integrated luminosity of 35.9 fb^{-1}, collected by the CMS detector in 2016 at a center of mass energy of 13 TeV with 25 ns bunch crossing intervals. The datasets used in this analysis are given in Tables 7.1 and 7.2 along-with the corresponding integrated luminosity. Good luminosity blocks provided by CMS are selected for the analysis.

Table 7.1: Name of datasets used in the analysis.

Dataset name	**Luminosity in fb^{-1}**
Double Muon Dataset	35.839
Double Electron Dataset	30.078
Muon + Electron/Gamma Dataset	30.078

Table 7.2: Name of the Datasets used in the validation of QCD estimation.

Dataset name	**Luminosity in fb^{-1}**
Single Muon Dataset	35.381
Single Electron Dataset	35.381

7.1.2 Monte Carlo Samples

Independent signal samples at different Radion mass points starting from 260 GeV to 900 GeV, are generated at the leading order (LO) using the MadGraph5_aMC@NLO [54] event generator. Both the Higgs bosons are allowed to decay to tau leptons only. The list of all the signal datasets is given in Table 7.3. Cross section for all the signal samples together with the branching ratio for $HH \rightarrow \tau\tau\tau\tau$ is taken to be 1 pb.

The Monte Carlo (MC) samples used to estimate the background contribution are listed in Tables 7.4 along-with the corresponding cross sections in pb. The $W +$n-jets

Table 7.3: CMS datasets used in the analysis for Signal samples of spin-0 resonant production at the leading order (LO) using the MadGraph5_aMC@NLO for different mass points in the range 260-900 GeV.

Dataset name
/GluGluToRadionToHHTo4Tau_M-260_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-270_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-280_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-300_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-320_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-340_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-350_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-400_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-450_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-500_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-550_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-600_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-650_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-700_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-750_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-800_narrow_13TeV-madgraph/
/GluGluToRadionToHHTo4Tau_M-900_narrow_13TeV-madgraph/

and Drell-Yan (DY) + n-jets samples are generated with MadGraph5_aMCNLO 2.3.2, while the Single top and $t\bar{t}$ samples are generated with POWHEG 2.0 [55, 56, 57, 58, 59]. The di-boson samples are generated with either MadGraph5_amcNLO 2.3.2 or POWHEG 2.0. The Pythia8.212 [60] event generator, with the CUETP8M1 tune [61], is used to model the parton shower and hadronization processes, as well as tau decays in all these samples. All the produced particles are simulated through the GEANT4 [62] based simulation of the CMS detector.

7.1.3 Stitching Technique

The DY + jets and $W+$ jets samples are generated with different associated jet multiplicities from 0 to 4 jets at the matrix element level. Due to excessive computing time required to produce the events, the statistics in each exclusive sample is limited. The event statistics can be increased if events in exclusive sample can be combined with the events present in the inclusive sample using the generator level information of jet multiplicities. This process is known as stitching technique which is achieved by providing an appropriate weight to the n-jet sample which is given as follows :

$$weight_{n-jet} = \frac{1}{\frac{N_{exc}}{XS_{exc}} + \frac{N_{inc}}{XS_{inc}}}$$

For the 0-jet sample, the weight applied to the event is :

$$weight_{0-jet} = \frac{1}{\frac{N_{inc}}{XS_{inc}}}$$

where N_{exc} and N_{inc} are the number of events present in exclusive and inclusive n-jets samples respectively while XS_{exc} and XS_{inc} represent the cross section of exclusive and inclusive production of n-jet samples respectively.

Table 7.4: Background MC samples used in the analysis.

Physics Process	**MC Dataset name**	Generator and Parton Shower	Cross section in pb
	Irreducible background		
ZZ→4l	ZZTo4L_13TeV_powheg_pythia8	Powheg+ Pythia8	1.2
Single Higgs	ZHToTauTau_M125_13TeV_powheg_pythia8	Powheg+Pythia8	0.0559
	Reducible background		
DY + n jets	DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	Madgraph+Pythia8	4895
	DY1JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8		1012.5
	DY2JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8		332.8
	DY3JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8		101.8
	DY4JetsToLL_M-50 TuneCUETP8M1_13TeV-madgraphMLM-pythia8		54.8
W + n jets	WJetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	Madgraph+Pythia8	50690
	W1JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8		9644.5
	W2JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8		3144.5
	W3JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8		954.8
	W4JetsToLNu_TuneCUETP8M1_ 13TeV-madgraphMLM-pythia8		485.6
Di-boson	WW_TuneCUETP8M1_13TeV-pythia8	Pythia8	63.21
	WZ_TuneCUETP8M1_13TeV-pythia8		22.82
	ZZ_TuneCUETP8M1_13TeV-pythia8		10.32
$t\bar{t}$	TT_TuneCUETP8M2T4_13TeV-powheg-pythia8	Powheg+Pytia8	730
Single top	ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	Powheg+Pythia8	38.09
	ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1		38.09
	ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1		136.02
	ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1		80.95

7.2 Event Selection

7.2.1 Trigger Requirement

Events are required to satisfy at least one of the trigger paths of HLT defined in Table 7.5, depending on the final state. All the triggers used are without any prescale at the HLT level during the 2016 data taking period. The same trigger requirements are used for both data and MC. All these triggers require either one or two leptons with a p_T cut and loose isolation requirement. For example, HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ requires two muons with leading muon $p_T > 17$ GeV and sub-leading muon $p_T > 8$ GeV with a very loose isolation requirement in the tracker and with a $|dz|$ cut which ensures that muon is close to the primary vertex along the longitudinal direction.

Table 7.5: Trigger paths used in the analysis.

Channel	HLT path name	Dataset
$\mu\mu\tau_h\tau_h$	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v	Double Muon
$\mu\mu\tau_h\tau_h$	HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v	Double Muon
$\mu\mu\tau_h\tau_h$	HLT_IsoMu24_v	Double Muon
$\mu\mu\tau_h\tau_h$	HLT_IsoTkMu24_v	Double Muon
$ee\tau_h\tau_h$	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v	Double Electron
$ee\tau_h\tau_h$	HLT_Ele25_eta2p1_WPTight_Gsf_v	Double Electron
$\mu e\tau_h\tau_h$	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v	MuonEG
$\mu e\tau_h\tau_h$	HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_v	MuonEG
$\mu e\tau_h\tau_h$	HLT_IsoMu24_v	MuonEG
$\mu e\tau_h\tau_h$	HLT_IsoTkMu24_v	MuonEG
$\mu e\tau_h\tau_h$	HLT_Ele25_eta2p1_WPTight_Gsf_v	MuonEG

7.2.2 Event Vertex

The primary vertex of the hard scattering process is selected using the Deterministic Annealing (DA) [63] clustering algorithm. The distance of the vertex from the nominal interaction point is required to be smaller than 24 cm along the direction of the beam and 2 cm in the transverse plane. The number of degrees of freedom of the vertex fit is required to be larger than 4. Out of the selected vertices, the one with the largest Σp_T^2 of the tracks associated to that vertex is chosen as the event vertex corresponding to the hard scattering process. All other vertices are considered to come from additional soft scattering collisions (pile-up) during bunch crossing. Pile-up interactions are simulated by superimposing a number of minimum bias events onto the hard scattering process. In general, pile-up distributions of MC events do not match with that of data and to take that into account, a pile-up re-weighting is applied to each MC event.

7.2.3 Object Selection

Muon selection in $\mu\mu\tau_h\tau_h$ final state

Muons are required to be reconstructed by either the Particle Flow (PF) or the Tracker or the Global muon reconstruction algorithm [64]. The leading and sub-leading muons are selected with the following cuts:

- Leading μ $p_T > 18$ GeV
- Sub-leading μ $p_T > 9$ GeV
- $|\eta| < 2.4$
- μ should pass Physics Object Group (POG) recommended medium muon identification (ID) criteria [65]
- Relative isolation < 0.3

Invariant mass of the muon pair ($M_{\mu\mu}$) is required to be greater than 12 GeV to reject background events with low lepton pair mass. In the case of opposite-sign muons, to reduce the background contribution coming from the DY process, two additional cuts are applied: 70 GeV $> M_{\mu\mu}$ or $M_{\mu\mu} >$ 110 GeV and PF based missing transverse energy (pfMET) $>$ 20 GeV.

Electron selection in ee$\tau_h\tau_h$ final state

Selection criteria for electron candidates demand :

- Leading e p_T > 23 GeV
- Sub-leading e p_T >13 GeV
- $|\eta| < 2.5$
- Relative isolation < 0.25
- All the selected electrons should pass the POG recommended 90% electron MVA ID [66].

To mitigate the charge mis-identification rates, electrons should pass one of the conditions : charge is consistent between the GSF electron track and the pixel track, or the charge is consistent between GSF electron track and Kalman filter track. To reduce the background events with low lepton pair mass, the invariant mass of the electron pair (M_{ee}) should be greater than 12 GeV. For opposite-sign electrons, the DY contribution is reduced by demanding 70 GeV > M_{ee} or M_{ee} > 110 GeV and pfMET > 20 GeV.

Muon and electron selection in μe$\tau_h\tau_h$ final state

To select events in μe$\tau_h\tau_h$ channel the followings cuts are applied :

- p_T of leading lepton > 24 GeV
- p_T of sub-leading lepton > 9 GeV
- Muons should pass medium muon ID with $|\eta| < 2.4$ and relative isolation < 0.3
- Electrons should pass the 90% electron MVA ID with $|\eta| < 2.5$ and relative isolation < 0.25

The Leading lepton may be either a muon or an electron. To reduce charge mis-identification among electrons and muons, the following conditions should be satisfied. For electrons, the charge should either be consistent (a) between the GSF electron track and the pixel track or (b) GSF electron track and Kalman filter track, while for muons, no muon should fail $\sigma_{p_T}/p_T < 0.2$. Charges of the selected leptons could be of opposite sign or of same sign depending on the final sate. Invariant mass of the lepton pair should be > 12 GeV to reduce contribution of background events coming from low lepton pair mass.

τ_h selection for all the final states

Hadronic τ's (τ_h) are identified using the "Hadron Plus Strips" (HPS) algorithm. p_T of the τ_h should be > 20 GeV and $|\eta|$ is required to be < 2.3. The τ_h candidate should pass the loose MVA discriminator against electrons and muons to reduce the contribution of e$\rightarrow \tau_h$ or $\mu \rightarrow \tau_h$ fake. The Loose MVA working point of the MVA based isolation is chosen to increase the statistics.

To select same-sign leptons in the signal region, the charge of the two selected leptons (μ's in $\mu\mu\tau\tau$ channel, e's in ee$\tau\tau$ channel and μ, e in $\mu e\tau\tau$ channel) should be the same while two τ_h's should also have the same-sign but opposite with respect to the lepton charge. For opposite-sign leptons in the signal region, selected leptons should be of opposite charge and the same should hold for τ_h's. Total charge of the selected leptons and τ_h's should be zero and the separation between all the leptons and τ_h is required to be > 0.3, *i.e.* $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.3$ where $\Delta\eta$ and $\Delta\phi$ are the difference in η and ϕ between any two particles.

7.2.4 b jet veto

To reduce contamination from $t\bar{t}$, Single top events, events with well identified b jets are rejected in all the three channels. The rejection criterion is based on jets with $p_T > 20$

GeV, $|\eta| < 2.5$ and BDT distribution score to tag a jet as b jet should be greater than 0.8484. This value corresponds to medium b tagging score for which the mis-identification probability of light-flavour jets is around 1% and b tagging efficiency is around 63%.

7.2.5 Removal of overlap between different final states

To avoid double counting of events among all channels, events are rejected if there is any extra muon with

- $p_T > 9$ GeV
- $|\eta| < 2.4$
- Relative isolation < 0.3
- Medium muon ID

or any extra electron with

- $p_T > 13$ GeV
- $|\eta| < 2.5$
- Relative isolation < 0.25
- 90% electron MVA ID.

7.3 Background Estimation

There are two types of background which can mimic the signal, irreducible background due to $ZZ \to 4\ell$, ZH and reducible process, where one or both $\tau_h^\prime s$ are faked by jets.

7.3.1 $ZZ \to 4\ell$

This is an irreducible background and a major background as both the Z's can decay to tau lepton where two τ's can decay hadronically and the other two τ's can decay leptonically. The contribution of this background is estimated using MC samples.

7.3.2 ZH

A Standard Model Higgs boson produced in association with a Z boson is also an irreducible background. The contribution from this process is also calculated using MC samples.

7.3.3 Drell-Yan + Jets

This is one of the major sources of background where two opposite-sign, same-flavour leptons are present. The contribution of this background is also important where two same-sign electrons are selected due to charge mis-identification. Charge mis-identification rate is larger for electrons than for muons. Two opposite-sign electrons coming from the DY process can be mis-identified as same-sign while two jets can fake two τ_h's to mimic the signal. The contribution of DY background is estimated from MC samples.

7.3.4 $t\bar{t}$, Single top, W + Jets, di-boson

The contribution of $t\bar{t}$ events is important in $\mu e\tau_h\tau_h$, where both the W's can decay leptonically and the two τ_h's are jets faking τ_h's. The contribution of Single top, W + Jets, di-boson are found to be negligible. All background contributions from Single top, $t\bar{t}$, W + Jets, di-boson (ZZ, WZ, WW) are calculated using MC.

7.3.5 QCD Multijets

This is one of the main reducible backgrounds which is estimated using a data driven method that relies on estimation of jet $\rightarrow \tau_h$ fake rate. The fake rates are measured from data as jet $\rightarrow \tau_h$ fake is not well simulated in MC simulated event samples and also due to limited statistics of QCD MC events.

Jet $\rightarrow \tau_h$ Fake Rate Measurement

The probability of a jet faking a τ_h is measured in a region with no genuine τ_h. This is achieved by selecting DY + jets events where Z decays to either a muon or an electron pair. Muons and electrons are selected using the same procedure as in the event selection and requiring that there should be at least two τ_h which satisfy all the τ_h selection conditions except isolation. The total charge of two leptons and two τ_h's should not be zero to remove overlap with signal events where total charge is zero. To select a sample of events dominated by DY + jets events, electron and muon vetoes are applied in both the channels and pfMET is required to be less than 20 GeV. The contribution of all other non-DY processes is calculated from MC and subtracted from data to reject events containing a real τ_h which can come from other non-DY events like WZ, ZZ events. After this selection, the ratio of number of τ_h's which satisfy the loose MVA isolation to the total number of τ_h's is defined as the fake rate. The p_T distribution of the τ_h's in the denominator and numerator for both the channels are shown in Figure 7.2. Around 20% disagreement is observed between data and MC for events in the numerator which is attributed due to mis-modeling of jet $\rightarrow \tau_h$ fake in simulation. The measured fake rate is shown in Figure 7.3 as a function of τ_h p_T.

For the $\mu e\tau_h\tau_h$ channel, the fake rate is obtained by combining the fake rates from $\mu\mu\tau_h\tau_h$ and $ee\tau_h\tau_h$ channels. The measured fake rate is then fitted with the following function in the three channels separately:

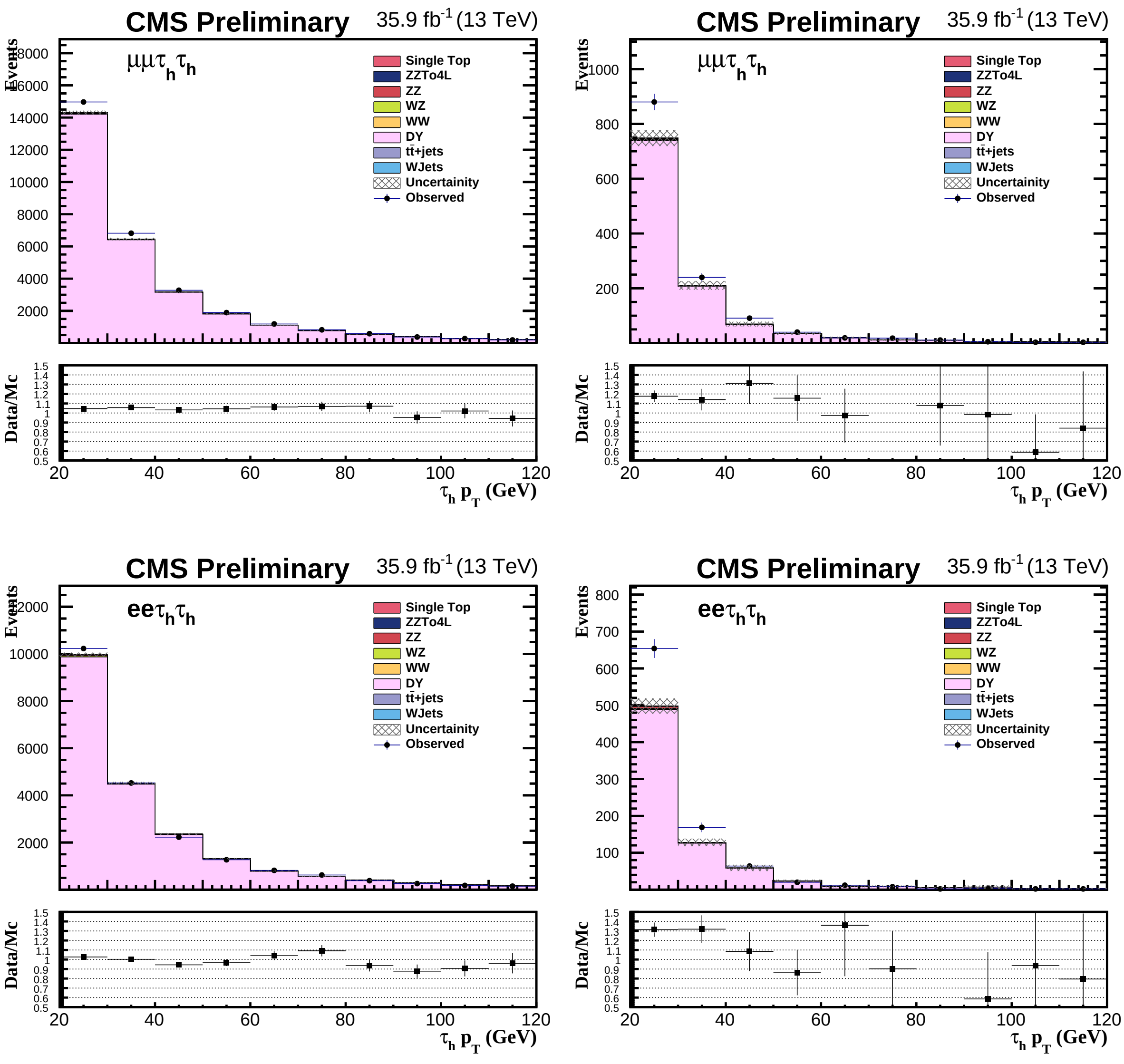

Figure 7.2: p_T distribution of τ_h in the denominator in $\mu\mu\tau_h\tau_h$ (top left), in ee$\tau_h\tau_h$ (bottom left). p_T distribution of τ_h in the numerator in $\mu\mu\tau_h\tau_h$ (top right), in ee$\tau_h\tau_h$ (bottom right).

$$F\left(p_T(\tau_h)\right) \;=\; C_0 e^{C_1 p_T(\tau_h)} + C_2 \tag{7.3.5.1}$$

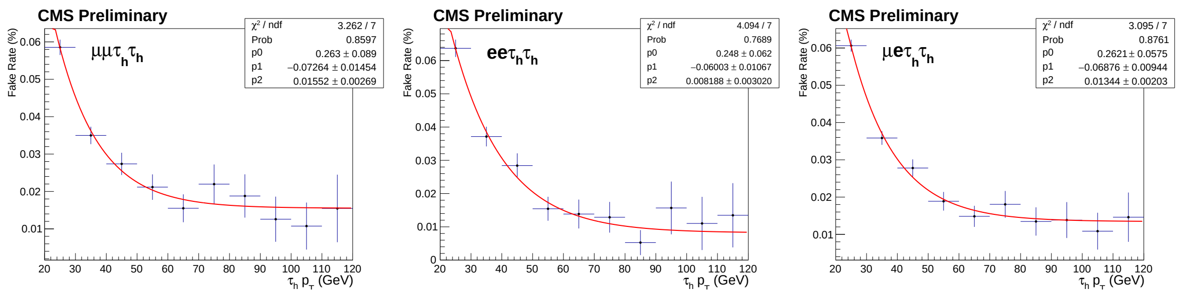

Figure 7.3: Fake rate as a function of τ_h p_T in $\mu\mu\tau_h\tau_h$ (left), in $ee\tau_h\tau_h$ (middle), in $\mu e\tau_h\tau_h$ (right) channels.

Background Estimation using τ_h Fake Rate

To estimate the QCD background where τ_h's are faked by jets, the fit results are applied to the signal region where all the event selection cuts are applied except the cuts for τ_h isolation. The control region is divided into three mutually exclusive regions:

FF: In this region, both τ_h's fail the isolation cut and the following weight is applied to events:

$$w_{FF} \;=\; \frac{F(\tau_1)\,F(\tau_2)}{(1-F(\tau_1))\,(1-F(\tau_2))}$$

FP: In this region, the leading τ_h fails the isolation criterion, but the sub-leading τ_h passes the same. The weight applied to this region is,

$$w_{FP} \;=\; \frac{F(\tau_1)}{(1-F(\tau_1))}$$

PF: This region is the same as the FP region but the role of the leading and sub-leading τ_h is exchanged. The weight applied to this region is as follows:

$$w_{PF} = \frac{F(\tau_2)}{(1 - F(\tau_2))}$$

$F(\tau_i)$ is defined as the fake rate at the p_T of τ_h which fails the isolation criteria. Due to the presence of the FF region in the other two control regions, we need to subtract the contribution of the FF region from these two regions. After subtraction, the total background estimation coming from jet $\rightarrow \tau_h$ fake becomes:

$$\begin{aligned} N_{estimated} &= \frac{N_{FF}F(\tau_1)F(\tau_2)}{(1-F(\tau_1))(1-F(\tau_2))} \\ &+ \frac{(N_{FP} - N_{FF}F(\tau_2))}{(1-F(\tau_2))} \times \frac{F(\tau_1)}{1-F(\tau_1)} + \frac{(N_{PF} - N_{FF}F(\tau_1))}{(1-F(\tau_1))} \times \frac{F(\tau_2)}{1-F(\tau_2)} \\ &= N_{estimated\_in\_FP} + N_{estimated\_in\_PF} - N_{estimated\_in\_FF} \end{aligned} \quad (7.3.5.2)$$

where N_{XY} represents the total number of events in region XY.

The contribution of other non-QCD background in the above three control regions are estimated from MC sample and subtracted from data.

Validation of QCD Measurement

Validation of QCD measurement in data: The measurement of the QCD component in data is validated in the $W+$jets control region for electron and muon channels. To select $W+$jets events, **"SingleMuon"** and **"SingleElectron"** datasets described in Table 7.2 are used. Events are required to pass either the **HLT_IsoTkMu24_v** or the **HLT_IsoMu24_v trigger** for the μ channel and **HLT_Ele25_eta2p1_WPTight_Gsf_v trigger** for the electron channel. Muons are selected with $p_T > 26$ GeV, $|\eta| < 2.4$ and relative isolation < 0.3 with medium muon ID requirement. Electrons are selected with $p_T > 26$ GeV, $|\eta| < 2.5$ and relative isolation < 0.25 with POG recommended 90% electron MVA ID. Muon and electron vetoes are applied in both the channels with $p_T > 9$ GeV for μ and 13 GeV for electron in addition to the other cuts. Tau selection is the same as mentioned in section 7.3.5, except that the selected τ_h's are of the same sign. pfMET and transverse mass (m_T) of the event are required to be larger than 40 GeV and 50 GeV respectively, where m_T is defined as :

$$m_T = \sqrt{2p_t^{\ell} E_T^{miss}(1 - \cos\Delta\phi)} \tag{7.3.5.3}$$

between the lepton ℓ and the pfMET (E_T^{miss}) where ϕ is the azimuthal angle between ℓ and E_T^{miss}. To reject $t\bar{t}$ events, b jet veto is applied. Contribution from other non $W+$jets events, estimated from MC, are subtracted from data. Figure 7.4 shows the τ_h p_T distribution in the denominator and numerator in the $W+$jets control region. The fake rate is fitted with equation 7.3.5.1. Figure 7.5 shows the fake rate as a function of τ_h p_T.

To calculate the QCD contribution, events are selected with one muon (electron) and two τ_h's of opposite-sign in the muon (electron) channel. All the selected τ_h's and muons (electrons) should be separated from each other by a minimum $\Delta\mathrm{R} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.5$. m_T of the selected event is required to be < 40 GeV. This cut is orthogonal to the fake rate measurement region where $m_T > 40$ GeV which is mainly dominated by $W + jets$ region. With this additional cut, event selection goes to a control region which is dominated by DY+jets events where one tau may decay to a muon (an electron) and the other hadronically with the additional τ_h faked by jets. The Figure 7.6 shows that the region with $m_T < 40$ GeV is mainly dominated by DY+jets.

Selected events are categorized into three control regions based on the tau isolation

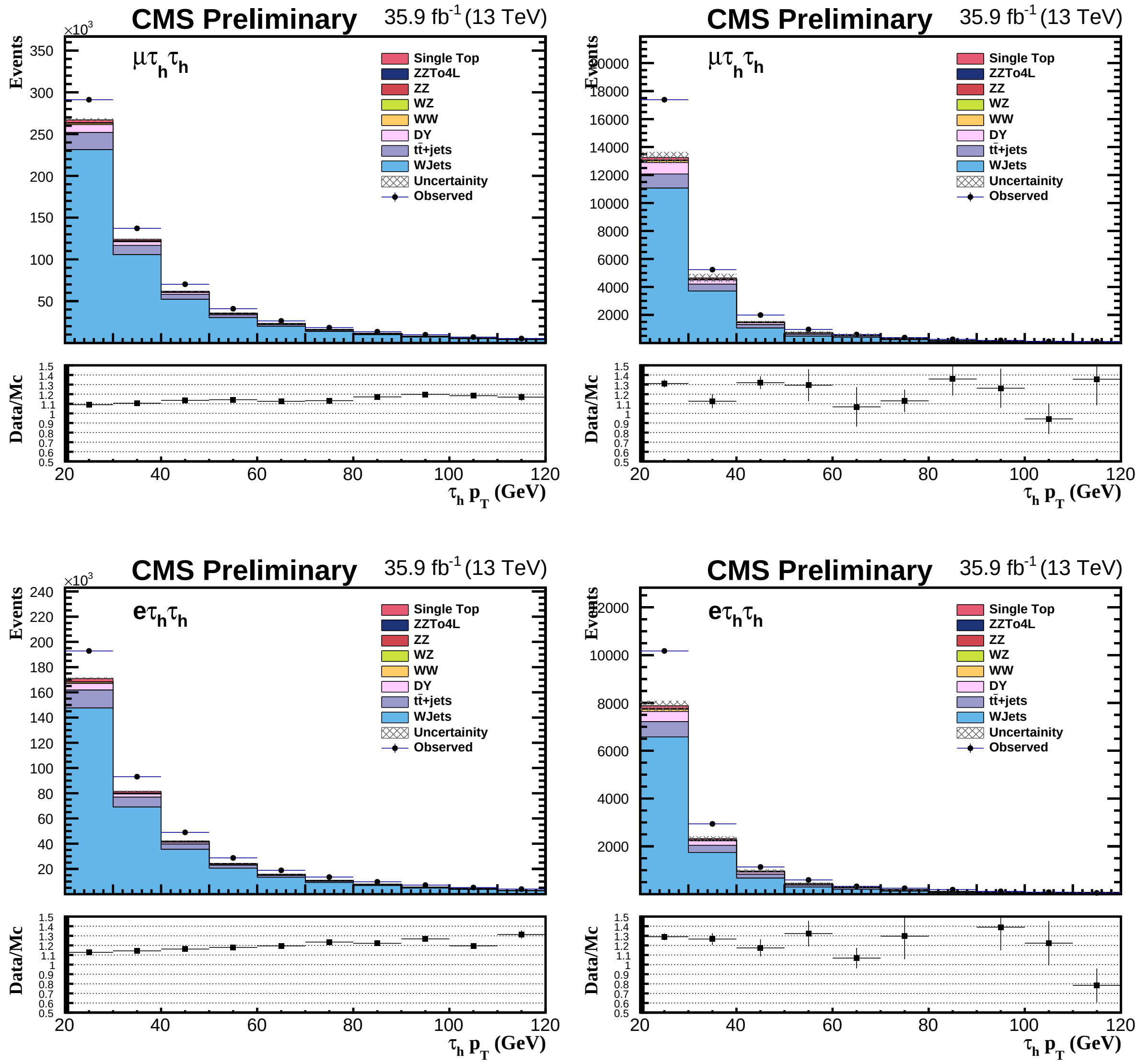

Figure 7.4: p_T distribution of τ_h in the denominator in muon channel (top left), in electron channel (bottom left). p_T distribution of τ_h in the numerator in muon channel (top right), in electron channel (bottom right).

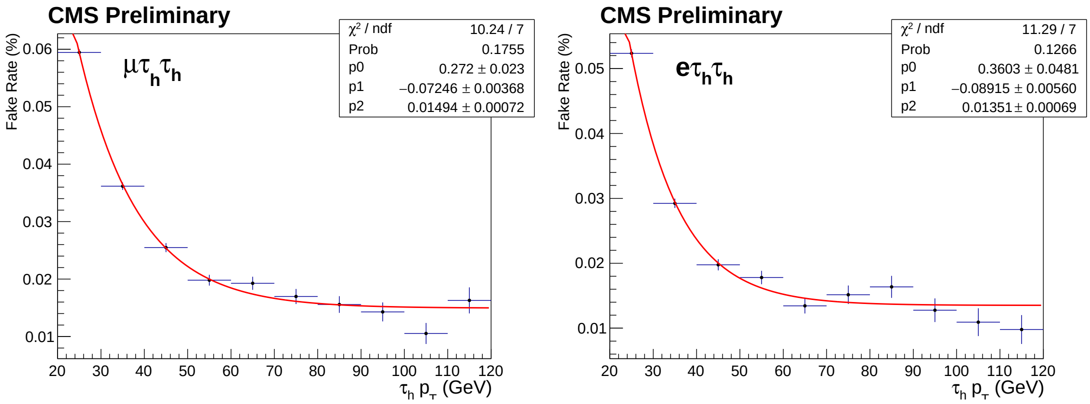

Figure 7.5: Fake rate as a function of τ_h p_T in $W+$jets control region in muon (left) and electron (right) channels.

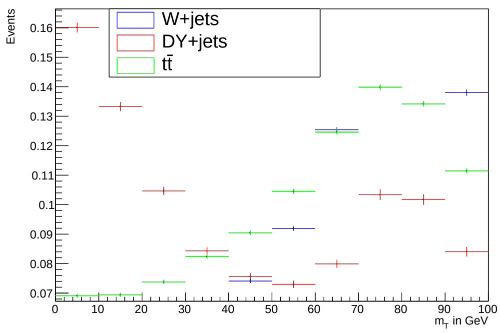

Figure 7.6: m_T distribution for $DY + jets$ (red), $W + jets$ (blue), $t\bar{t}$ (green) events.

as defined in section 7.3.5.

The excess in the data over MC is assumed to be coming from the QCD events as the contribution of other processes are considered from MC. To extrapolate the QCD contribution from the tau anti-isolated to isolated region, fake weight is applied to all the events in these regions as described in section 7.3.5 where value of $F(\tau_i)$ is taken from Figure 7.5. The distribution of invariant mass of $\mu\tau_h$ pair after applying the fake weight in the three anti-isolated regions are shown in Figure 7.7.

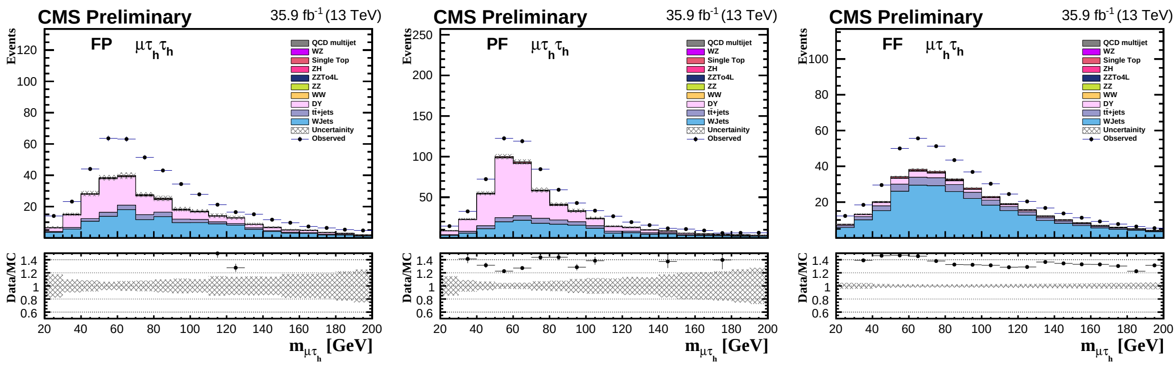

Figure 7.7: Invariant mass distributions of $\mu\tau_h$ pair in the FP (left), PF (middle), FF (right) regions after application of fake weight factors.

Figure 7.8 shows the invariant mass of $\mu\tau_h$ pair in tau isolated region where the contribution due to QCD is estimated from Figure 7.7 using equation 7.3.5.2.

Figure 7.9 shows the invariant mass of $\mu\tau_h\tau_h$ pair in the FP, PF, FF regions. Figure 7.10 shows the invariant mass of $\mu\tau_h\tau_h$ in the tau isolated region. Figures 7.8 and 7.10 show good agreement between the observed and estimated QCD contributions within statistical uncertainty in the muon channel.

The same test is performed for the electron channel as well. Figures 7.11 and 7.12 show the invariant mass of eτ_h pair in the FP, PF, FF and PP regions respectively. Figure 7.13 and 7.14 show the invariant mass of e$\tau_h\tau_h$ in the FP, PF, FF and PP regions. Figures 7.12 and 7.14 show good agreement between the observed and estimated QCD contributions in the electron channel. Figures 7.8 and 7.12 show

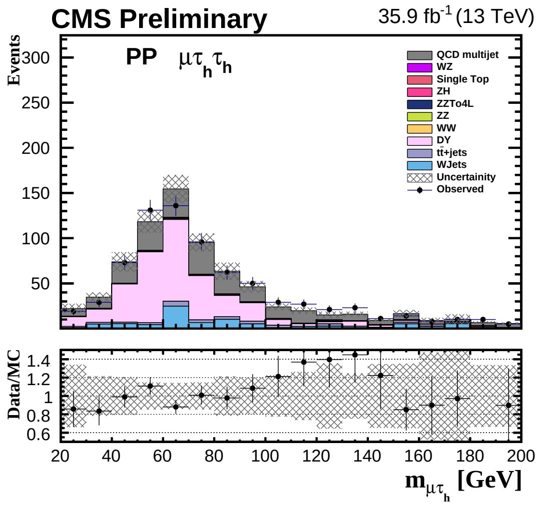

Figure 7.8: Invariant mass of $\mu\tau_h$ pair in the tau isolated region (PP).

that the distributions have a peak around 60 GeV which is due to the Drell-Yan process where one τ_h coming from the Z boson decays leptonically and the other decays hadronically and the invariant mass of $\mu\tau_h$, eτ_h pair peaks around 60 GeV due to the presence of neutrinos in the τ decays.

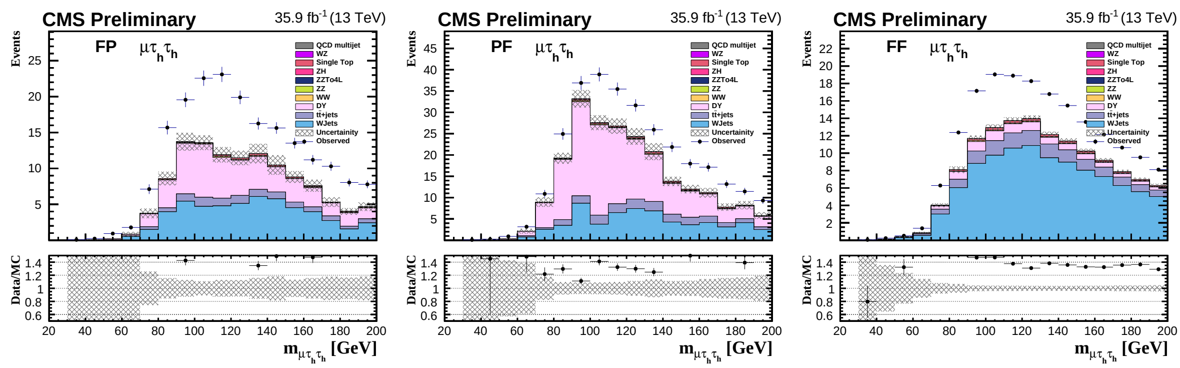

Figure 7.9: Invariant mass distributions of $\mu\tau_h\tau_h$ in the FP (left), PF (middle) and FF (right) regions after application of fake weight factors.

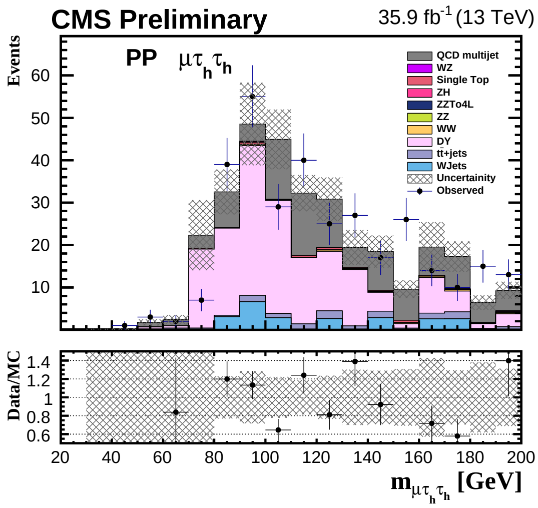

Figure 7.10: Invariant mass distribution of $\mu\tau_h\tau_h$ in the tau isolated region (PP).

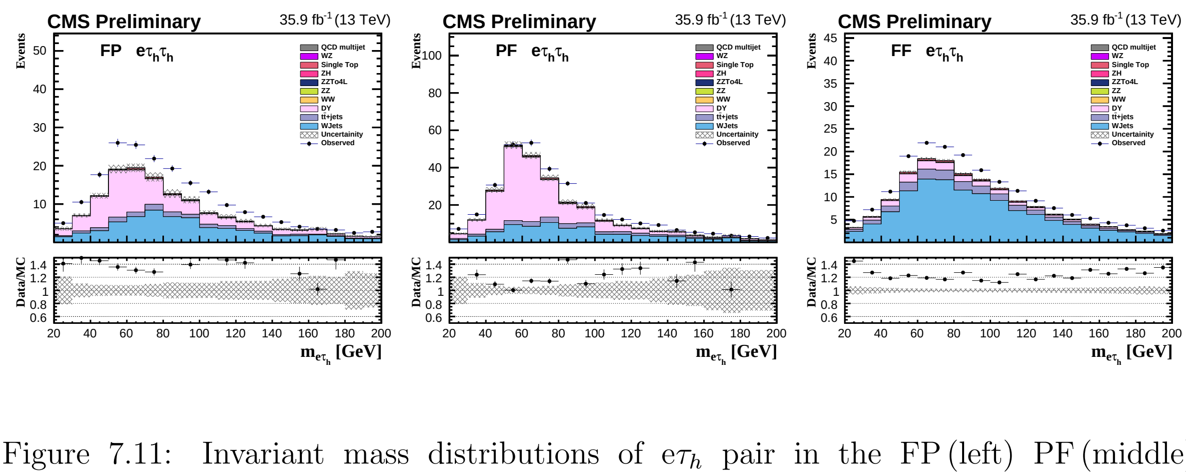

Figure 7.11: Invariant mass distributions of $e\tau_h$ pair in the FP (left) PF (middle), FF (right) regions after application of fake weight factors.

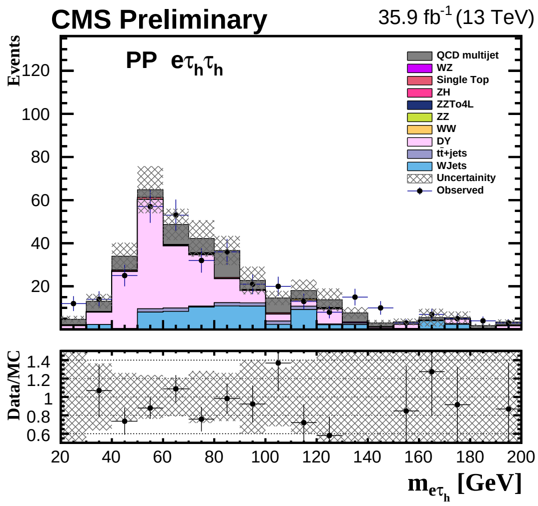

Figure 7.12: Invariant mass distribution of eτ_h pair in the tau isolated region (PP).

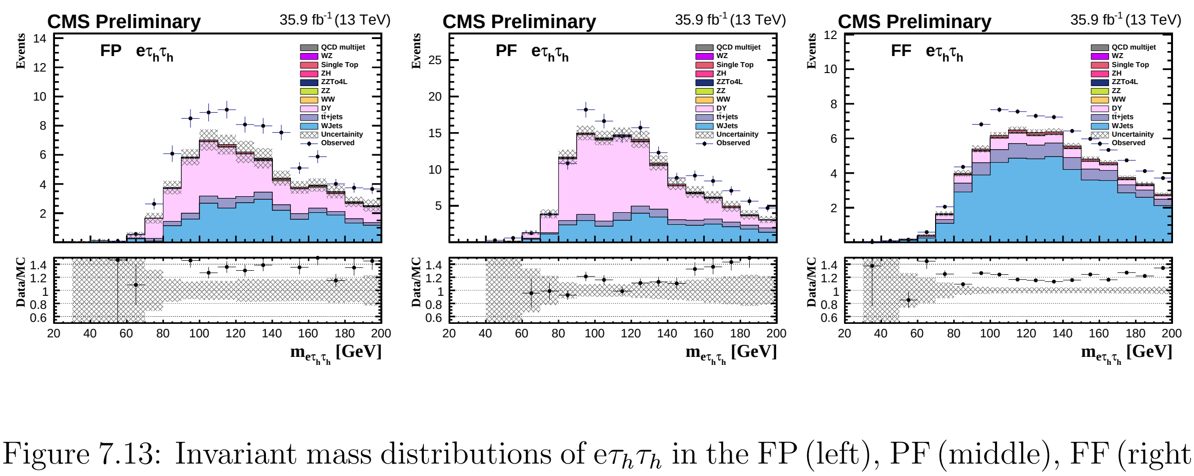

Figure 7.13: Invariant mass distributions of e$\tau_h\tau_h$ in the FP (left), PF (middle), FF (right) regions after application of fake weight factors.

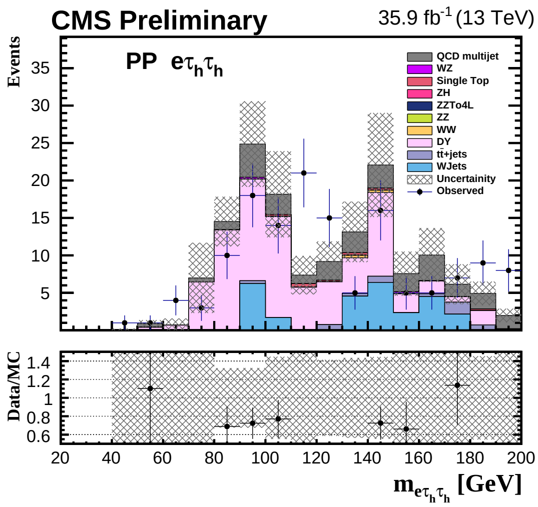

Figure 7.14: Invariant mass distribution of $e\tau_h\tau_h$ in the tau isolated region (PP).

7.4 Control Plot

To check the agreement between data and MC, comparisons are made in regions where the two muons are of opposite sign in the $\mu\mu\tau_h\tau_h$ channel, two electrons are of opposite sign in the $ee\tau_h\tau_h$ and the muon and the electron are of opposite sign in the $\mu e\tau_h\tau_h$ channel and total charge of the two leptons and the two τ_h's is non zero. This region does not overlap with the signal region where total charge of the selected leptons and two τ_h's is zero. To improve the matching between data and MC distributions, different scale factors (SF) like trigger SF [67], muon ID, isolation SF, electron reconstruction efficiency and MVA ID SF [68], τ_h ID SF [70] are applied to MC distributions. SF for each source is basically the ratio of events observed in data to that in MC. Figures 7.15, 7.16, 7.17, 7.18, 7.19, 7.20, 7.21, 7.22 and 7.23 show the kinematic distributions of leptons (muons in $\mu\mu\tau_h\tau_h$, electrons in $ee\tau_h\tau_h$ and muon, electron in $\mu e\tau_h\tau_h$ channels) τ's and invariant mass of the two leptons and two τ_h's. In the $\mu e\tau_h\tau_h$ channel, contribution of the Drell-Yan process, which has a large cross section, is very small due to the requirement of one electron and one muon. This leads to very small statistics in the $\mu e\tau_h\tau_h$ channel compared to the $\mu\mu\tau_h\tau_h$ or the $ee\tau_h\tau_h$ channel. In all these plots, there is a good agreement between data and MC within statistical uncertainty.

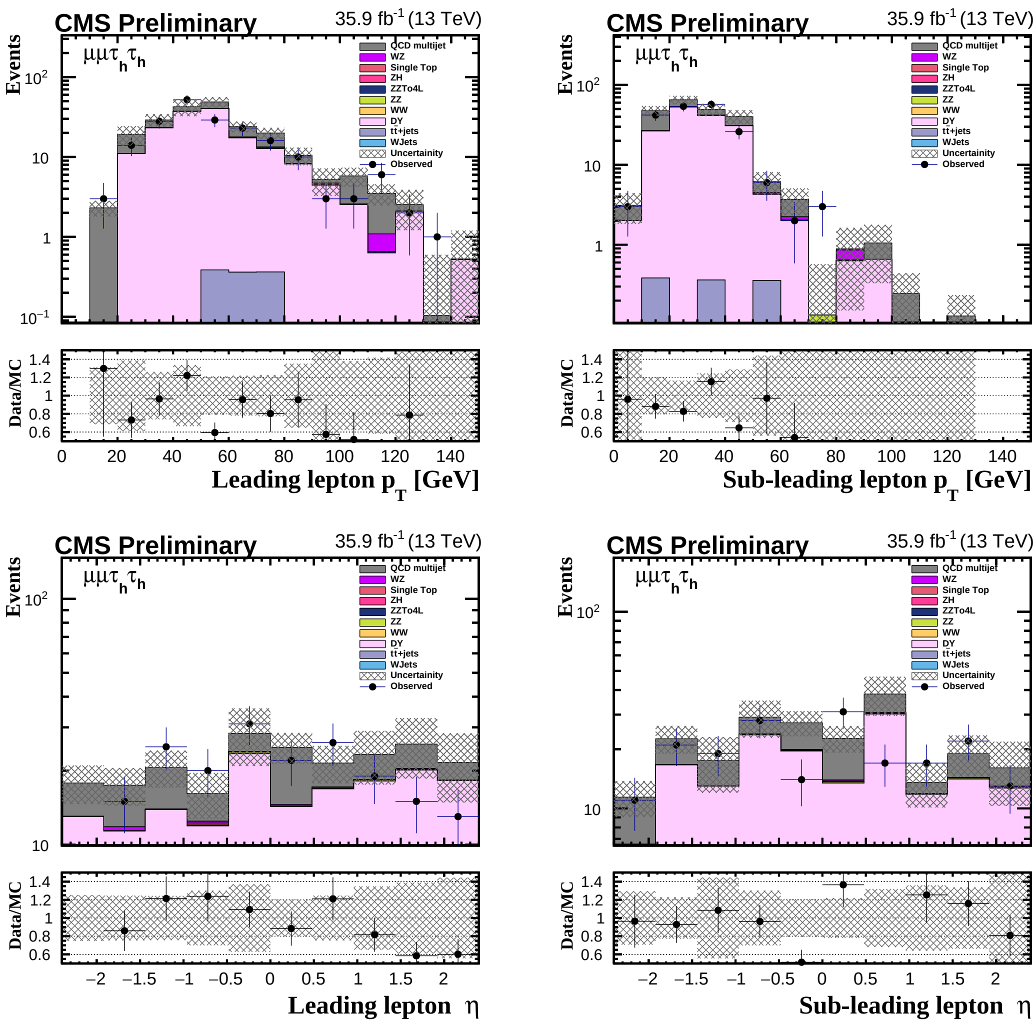

Figure 7.15: Control plots of p_T distribution of the leading muon (top left), sub-leading muon (top right); and of η distribution of the leading muon (bottom left), sub-leading muon (bottom right) in the $\mu\mu\tau_h\tau_h$ channel.

7.5 Signal Estimation

Size of the signal is measured using the invariant mass distribution of $\ell\ell\tau_h\tau_h$. Contribution due to QCD is estimated as mentioned in section 7.3.5 and fake weight is calculated from Figure 7.3. Table 7.6 and 7.7 show the efficiencies of signal in the SS and OS channels for

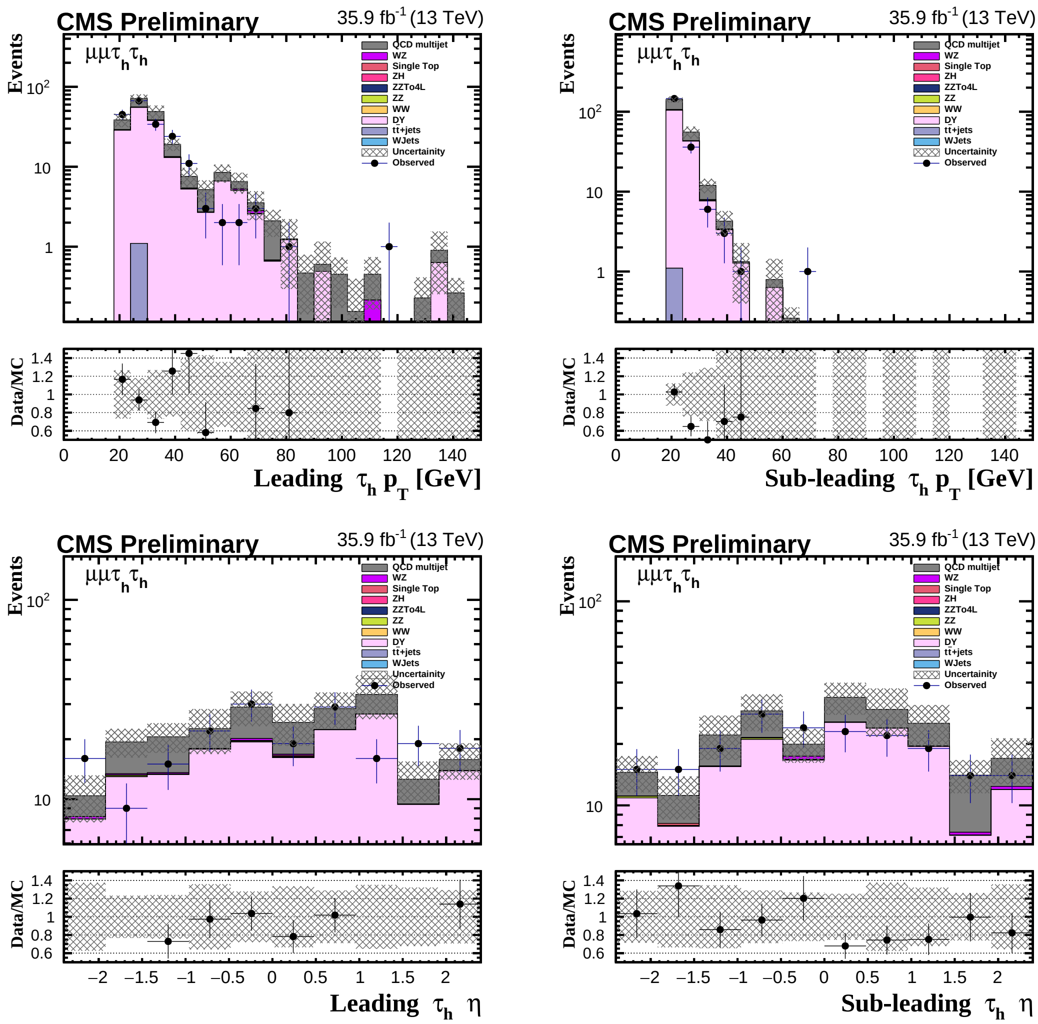

Figure 7.16: Control plots of p_T distribution of the leading τ_h (top left), sub-leading τ_h (top right); and of η distribution of the leading τ_h (bottom left), sub-leading τ_h (bottom right) in the $\mu\mu\tau_h\tau_h$ channel.

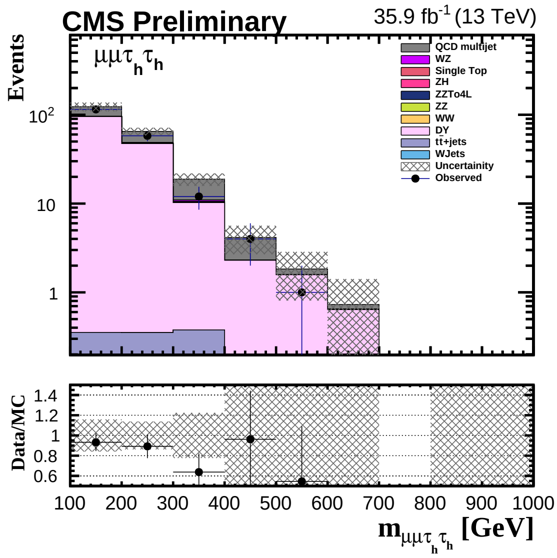

Figure 7.17: Control plot of invariant mass distribution of $\mu\mu\tau_h\tau_h$ for the $\mu\mu\tau_h\tau_h$ channel.

300 and 800 GeV mass Radion respectively. Table 7.8 and 7.9 show the expected event yield in 6 different channels separately for background as well as for signal for Radion mass 300 GeV and 800 GeV.

Table 7.6: Efficiency in % of signal in SS channel for Radion of 300 GeV and 800 GeV mass.

Mass point	$\mu\mu\tau_h\tau_h$	ee$\tau_h\tau_h$	$\mu e\tau_h\tau_h$
300	0.04	0.01	0.03
800	0.23	0.11	0.30

Table 7.7: Efficiency in % of signal in OS channel for Radion of 300 GeV and 800 GeV mass.

Mass point	$\mu\mu\tau_h\tau_h$	ee$\tau_h\tau_h$	$\mu e\tau_h\tau_h$
300	0.06	0.01	0.08
800	0.37	0.23	0.76

Due to low statistics available in the signal region for the same-sign lepton analysis as seen in Table 7.8, it is not possible to perform shape analysis in this case.

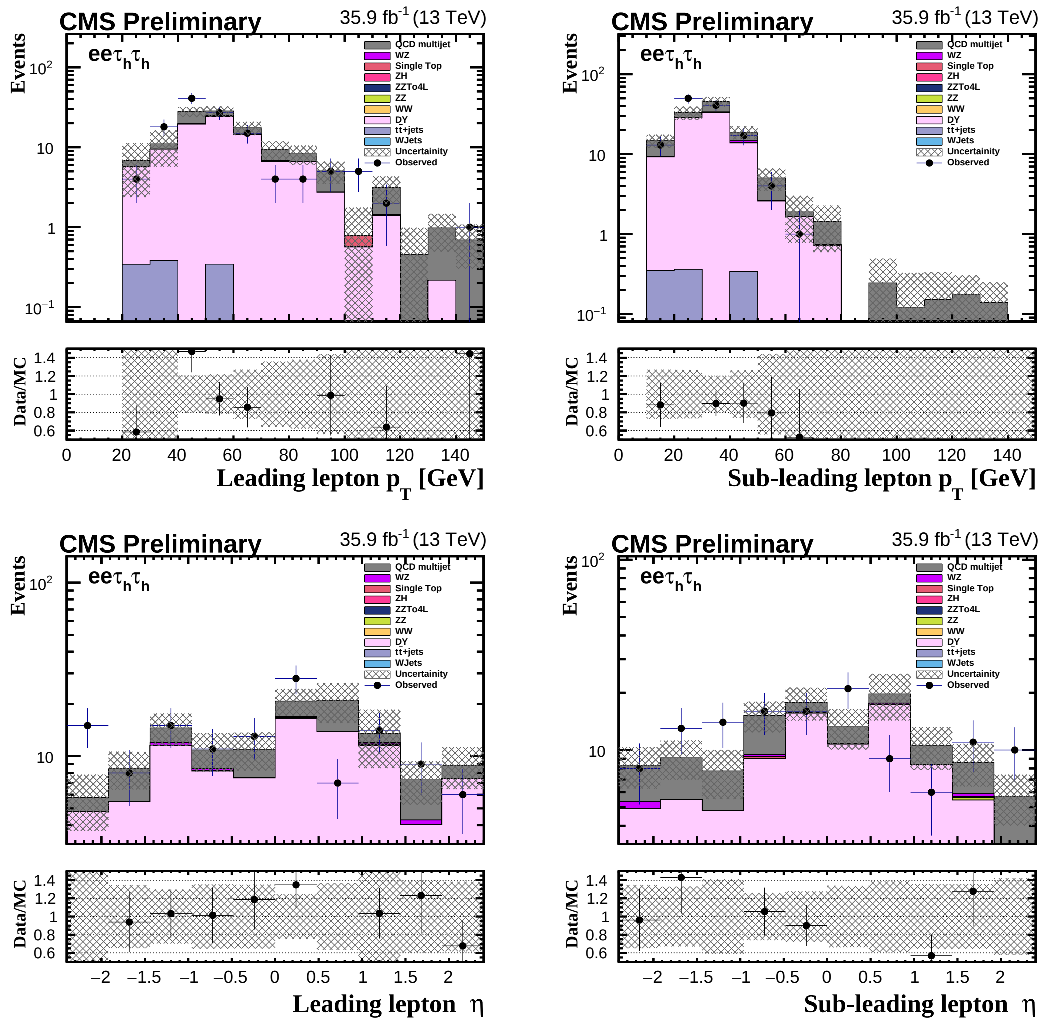

Figure 7.18: Control plots of p_T distribution of the leading electron (top left), sub-leading electron (top right); and of η distribution of the leading electron (bottom left), sub-leading electron (bottom right) in the ee$\tau_h\tau_h$ channel.

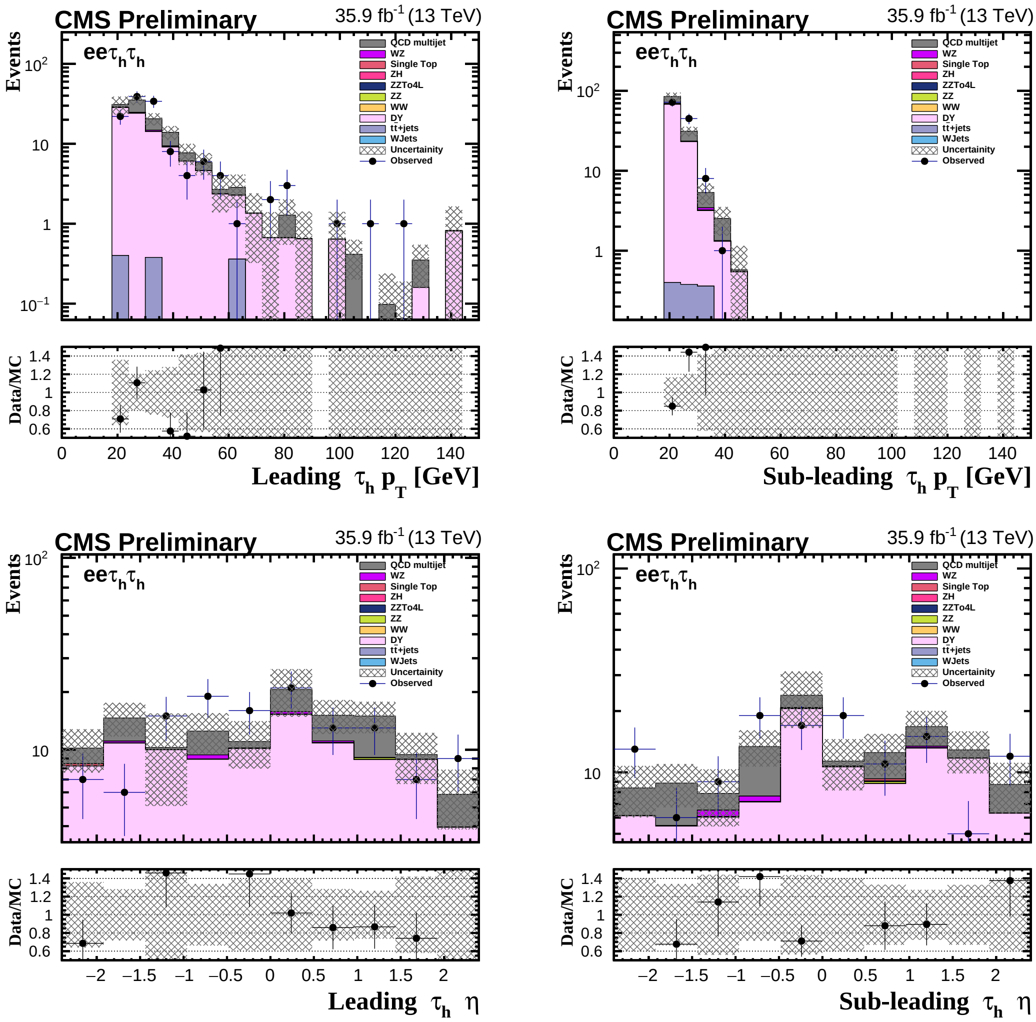

Figure 7.19: Control plots of p_T distribution of the leading τ_h (top left), sub-leading τ_h (top right); and of η distribution of the leading τ_h (bottom left), sub-leading τ_h (bottom right) in the ee$\tau_h\tau_h$ channel.

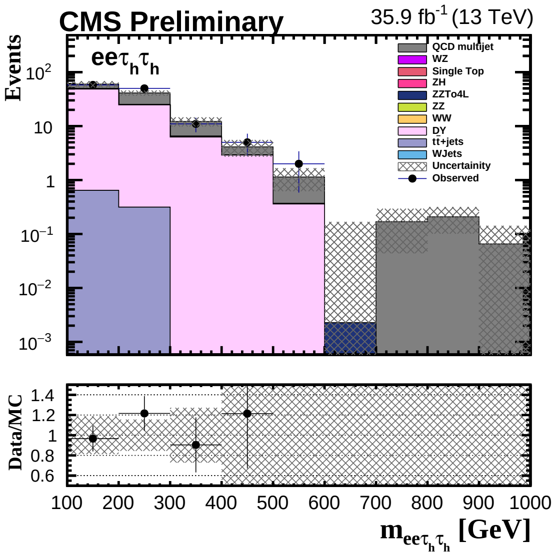

Figure 7.20: Control plot of invariant mass distribution of ee$\tau_h\tau_h$ for the ee$\tau_h\tau_h$ channel.

Table 7.8: Expected event yield in SS lepton final states. All backgrounds are normalized with $\sigma\times$ luminosity where luminosity is 35.9 fb^{-1}. For signal $\sigma \times BR$ is assumed to be 1 pb. The error shows the statistical uncertainty only.

Process	$\mu\mu\tau_h\tau_h$	$ee\tau_h\tau_h$	$\mu e\tau_h\tau_h$
$ZZ\to4\ell$	0.140±0.03	0.105±0.02	0.298 ± 0.04
ZH	0.047±0.01	0.008±0.005	0.082± 0.02
ZZ	0.192±0.19	-	-
$t\bar{t}$	-	0.252±0.25	1.495 ± 0.72
DY+jets	-	0.570±0.57	-
WZ	-	0.186±0.19	-
Single top	-	-	0.600 ± 0.60
Multijet	0.542±0.51	0.323±0.44	0.238 ± 0.60
Total background	0.921±0.54	1.444±0.79	2.713±1.11
300 GeV	24.190±3.64	10.578±1.98	30.944± 3.50
800 GeV	61.713±9.20	29.512±6.17	80.145± 10.28

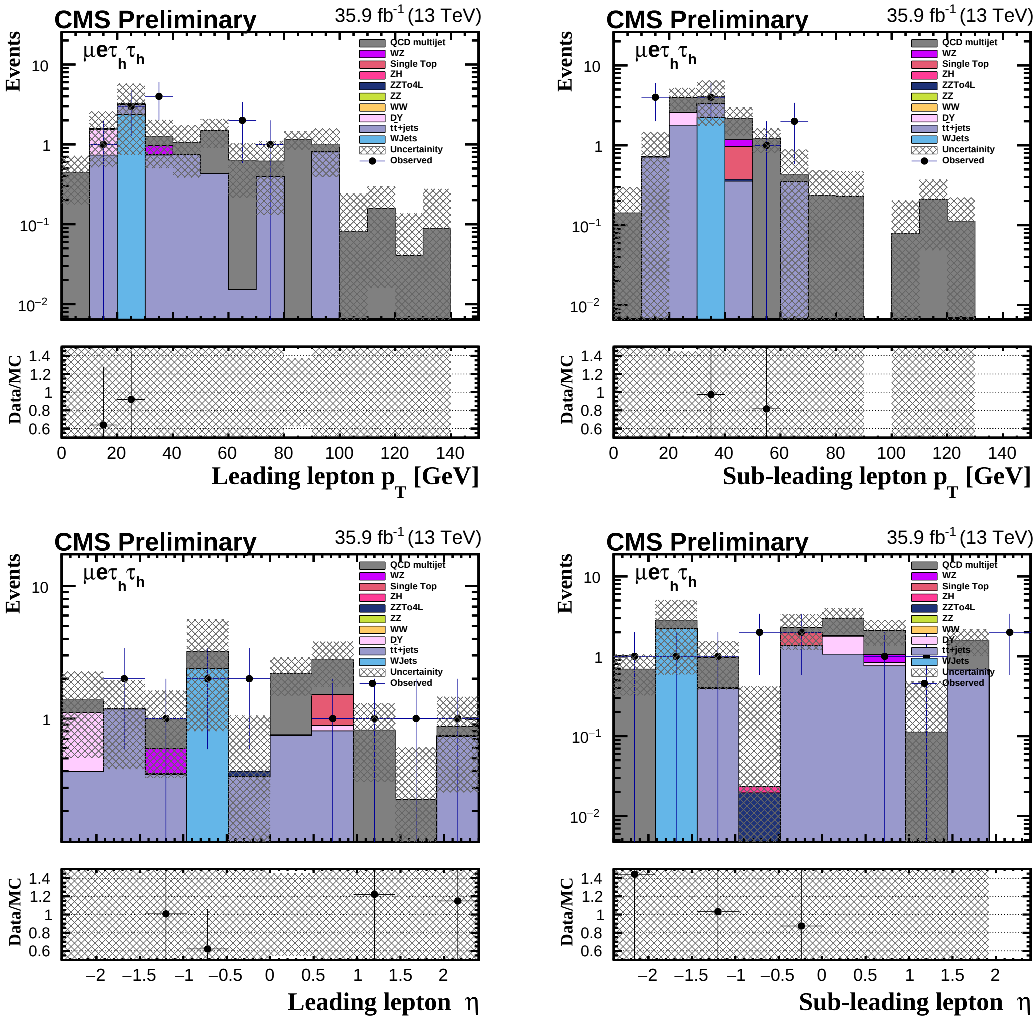

Figure 7.21: Control plots of p_T distribution of the leading lepton (top left), sub-leading lepton (top right); and of η distribution of the leading lepton (bottom left), sub-leading lepton (bottom right) in the $\mu e\tau_h\tau_h$ channel.

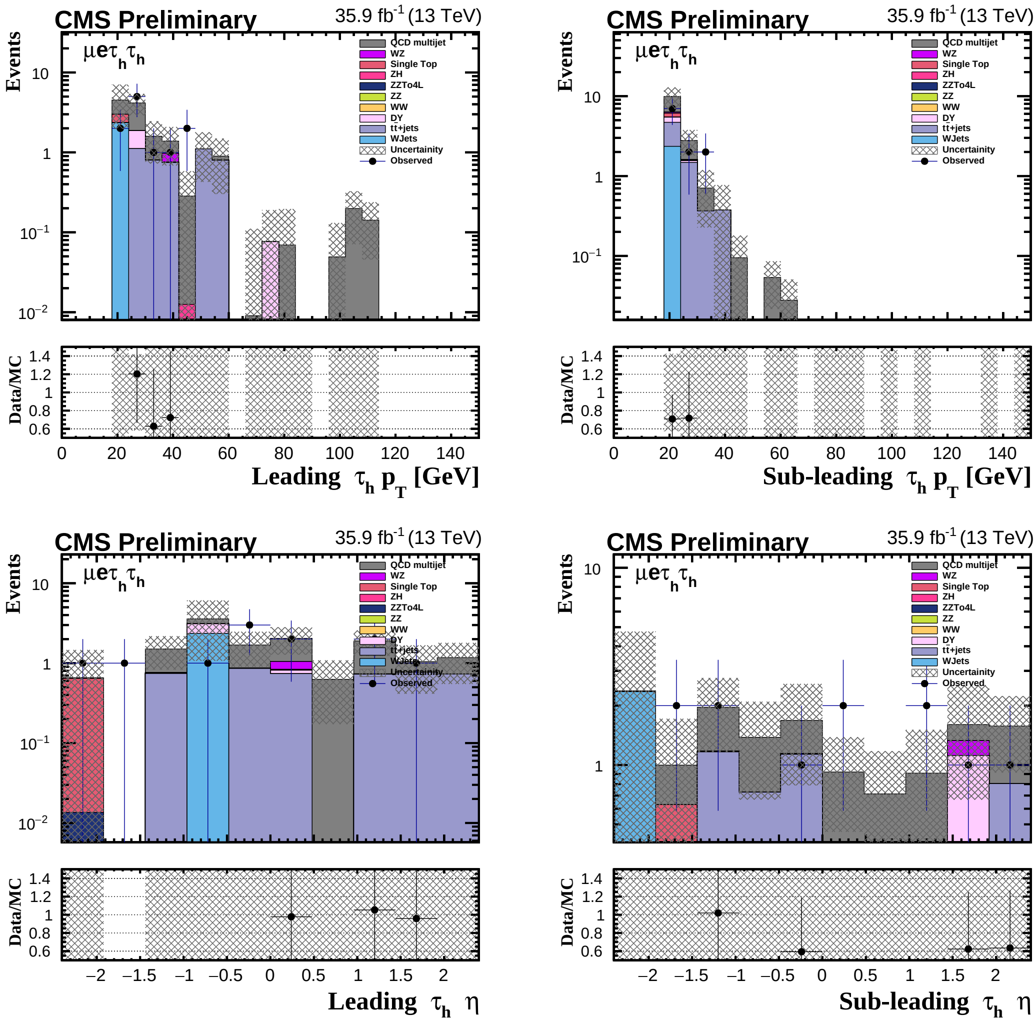

Figure 7.22: Control plots of p_T distribution of the leading τ_h (top left), sub-leading τ_h (top right); and of η distribution of the leading τ_h (bottom left), sub-leading τ_h (bottom right) in the $\mu e \tau_h \tau_h$ channel.

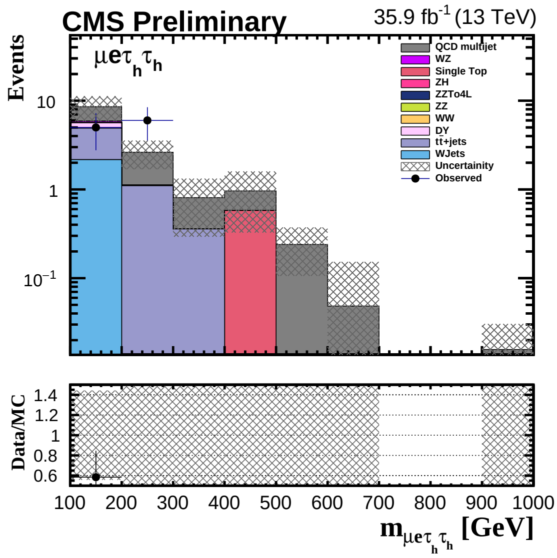

Figure 7.23: Control plot of invariant mass distribution of $\mu e\tau_h\tau_h$ for the $\mu e\tau_h\tau_h$ channel.

Table 7.9: Expected event yield in OS lepton final states. All backgrounds are normalized with $\sigma\times$ luminosity where luminosity is 35.9 fb^{-1}. For signal $\sigma \times BR$ is assumed to be 1 pb. The error represents the statistical uncertainty only.

Process	$\mu\mu\tau_h\tau_h$	$ee\tau_h\tau_h$	$\mu e\tau_h\tau_h$
$ZZ\rightarrow 4\ell$	0.828±0.07	0.604±0.06	0.475 ± 0.05
ZH	0.086±0.02	0.040±0.01	0.086±0.02
ZZ	0.223±0.21	0.652±0.33	0.105±0.10
$t\bar{t}$	2.428± 0.95	0.668±0.47	3.753 ±1.14
DY+jets	10.977±5.11	4.526±1.47	1.238±0.88
WZ	-	0.186±0.19	0.161±0.16
Single top	-	-	0.536±0.54
Multijet	13.576±1.84	4.101±1.41	5.869 ± 1.44
Total background	28.118±5.52	10.777±2.13	12.223±2.12
300 GeV	38.038±4.61	10.473±1.97	69.949± 5.30
800 GeV	100.631±11.85	62.007±8.85	205.833±16.46

A separate mass window is chosen for each mass point of the resonant particle which provides maximum separation between signal and background. This is achieved by maximizing the asymptotic formula [71] given by the following equation where s, b are the expected signal and background yield respectively and yields of background processes are scaled to the total luminosity.

$$\sqrt{\left(2\left(s+b\right)\log\left(1+\frac{s}{b}\right)-s\right)}$$

Table 7.10 shows the mass window for Radion 300 and 800 GeV mass points.

Table 7.10: Mass window after maximizing the asymptotic formula in SS lepton channel for Radion 300 GeV and 800 GeV mass points.

Mass point	$\mu\mu\tau_h\tau_h$	$ee\tau_h\tau_h$	$\mu e\tau_h\tau_h$
300 GeV	100-200	100-300	100-200
800 GeV	300-700	300-600	300-700

After selecting the mass window, all bins within the mass window are considered as a single bin and it is equivalent to the 'cut and count' experiment. Tables 7.11 and 7.12 show the expected event yield in SS region after selecting the mass window for SS channels separately for mass points of 300 GeV and 800 GeV.

Figure 7.24 shows the mass distribution of 4 particles in the signal region with same-sign lepton for 300 GeV and 800 GeV mass points of Radion after applying all the analysis cuts and before selecting the mass window whereas Figures 7.25, 7.26 show the mass distributions after selecting the mass windows for Radions of mass 300 GeV and 800 GeV respectively. Figure 7.27 shows the mass distribution of 4 particles in the signal region with opposite-sign lepton for 300 GeV and 800 GeV Radion mass points with all background contribution after full selection in the signal region.

Table 7.11: Expected event yield in SS lepton final states after fixing the mass window for 300 GeV. All backgrounds are normalized with $\sigma \times$ luminosity where luminosity is 35.9 fb^{-1}. For signal $\sigma \times BR$ is assumed to be 1 pb. The error is statistical uncertainty only.

Process	$\mu\mu\tau_h\tau_h$	$ee\tau_h\tau_h$	$\mu e\tau_h\tau_h$
$ZZ \to 4\ell$	0.096±0.16	0.093±0.14	0.179±0.18
ZH	0.027±0.10	0.003±0.003	0.041±0.11
ZZ	0.192±0.44		
$t\bar{t}$	-	0.252±0.50	1.136±0.79
DY+jets	-	0.570±0.75	-
WZ	-	0.186±0.43	-
Single top	-	-	
Multijet	-	0.151±0.62	-
Total background	0.315±0.48	1.255±1.18	1.356±0.82
300 GeV	20.108±1.824	10.578±1.60	26.380±1.79

Table 7.12: Expected event yield in SS lepton final states after fixing the mass window for 800 GeV. All backgrounds are normalized with $\sigma \times$ luminosity where luminosity is 35.9 fb^{-1}. For signal $\sigma \times BR$ is assumed to be 1 pb. The error represents the statistical uncertainty only.

Process	$\mu\mu\tau_h\tau_h$	$ee\tau_h\tau_h$	$\mu e\tau_h\tau_h$
$ZZ \to 4\ell$	-	0.006±0.006	0.061±0.15
ZH	0.008±0.06	0.005±0.004	0.019±0.11
ZZ	-	-	-
$t\bar{t}$	-	-	-
DY+jets	-	-	-
WZ	-	-	-
Single top	-	-	
Multijet	0.059±0.38	0.049±0.16	0.129±0.49
Total background	0.067±0.38	0.06±0.16	0.209±0.52
800 GeV	49.622±5.65	26.966±5.91	61.304±5.24

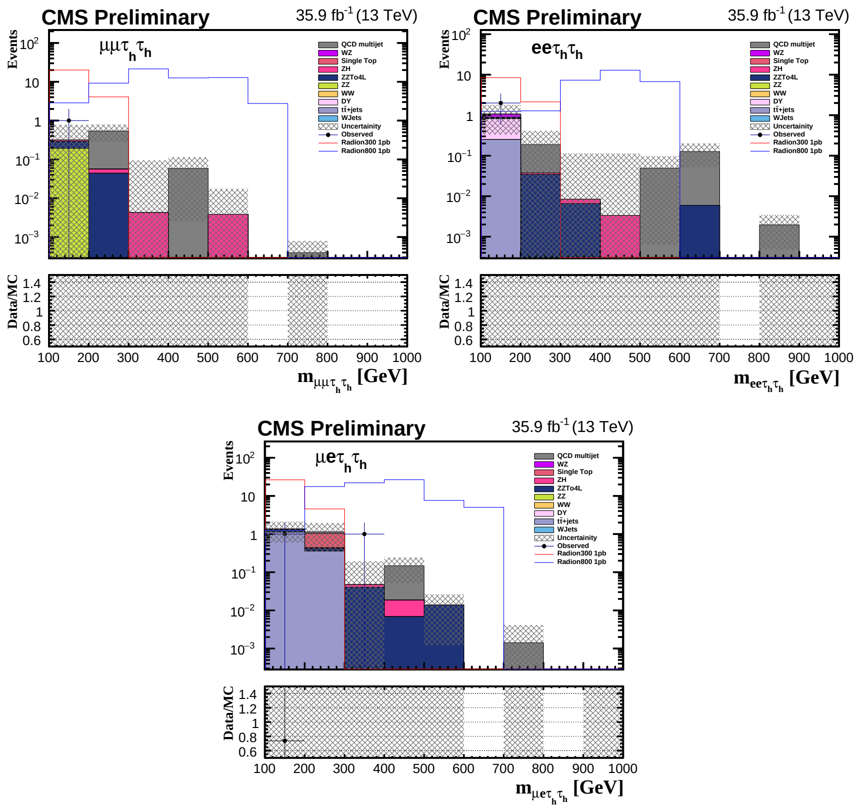

Figure 7.24: Invariant mass of $\mu\mu\tau_h\tau_h$(top left), ee$\tau_h\tau_h$(top right), μe$\tau_h\tau_h$(bottom) in the signal region with same-sign leptons before selecting the mass window. σ x BR of Radions of mass 300 GeV and 800 GeV are assumed to be 1 pb.

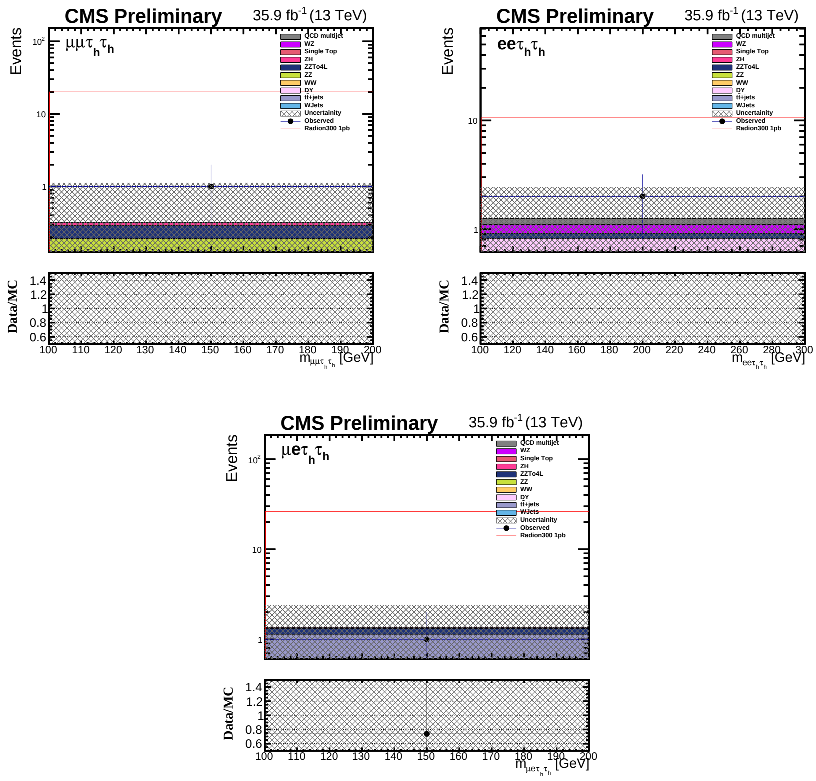

Figure 7.25: Invariant mass of $\mu\mu\tau_h\tau_h$(top left), ee$\tau_h\tau_h$(top right), μe$\tau_h\tau_h$(bottom) in the signal region with same-sign leptons after selecting the mass window. σ x BR of 300 GeV Radion is assumed to be 1 pb.

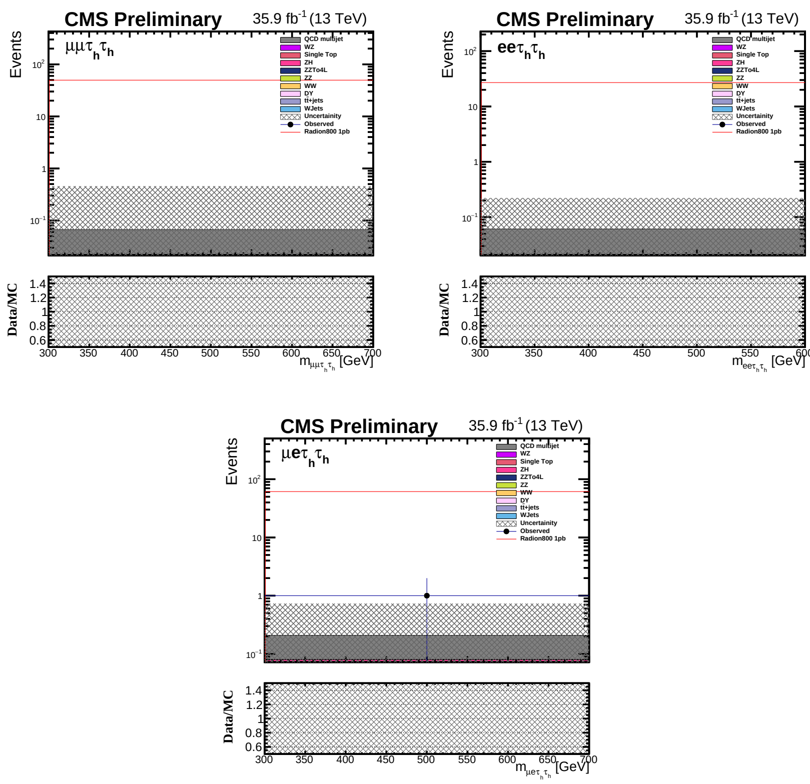

Figure 7.26: Invariant mass of $\mu\mu\tau_h\tau_h$(top left), ee$\tau_h\tau_h$(top right), μe$\tau_h\tau_h$(bottom) in the signal region with same-sign leptons. σ x BR of 800 GeV Radion mass is assumed to be 1 pb.

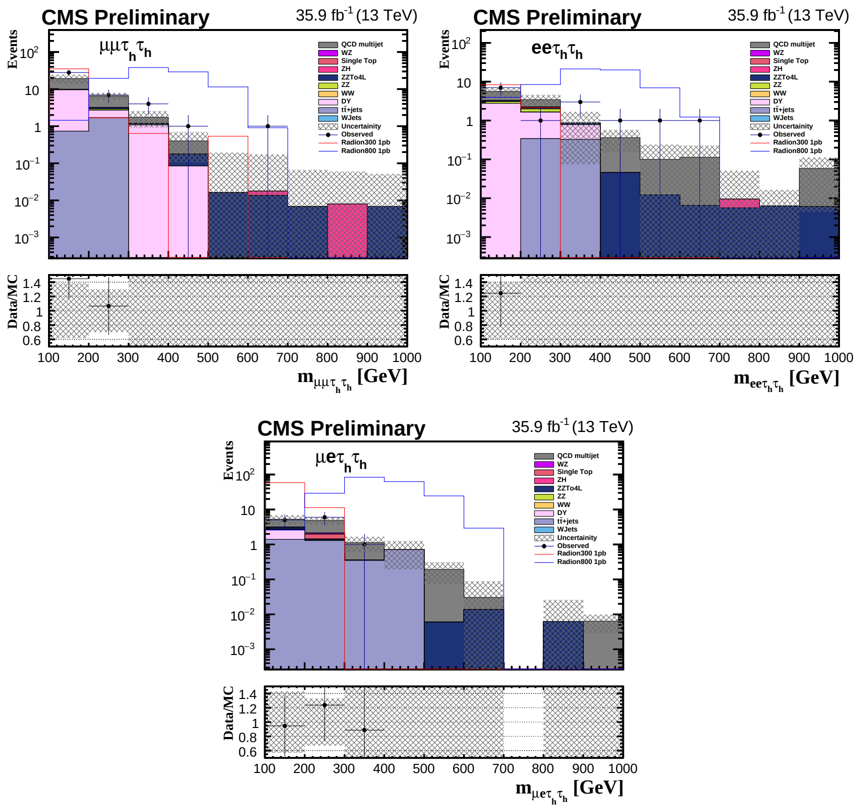

Figure 7.27: Invariant mass of $\mu\mu\tau_h\tau_h$(top left), e$e\tau_h\tau_h$(top right), $\mu e\tau_h\tau_h$(bottom) in the signal region with opposite-sign leptons. σ x BR of Radions of mass 300 GeV and 800 GeV are assumed to be 1 pb.

7.6 Systematic Uncertainties

Sources of systematic uncertainty which can affect the cut and count analysis with same-sign leptons and also the shape analysis with opposite-sign leptons are considered.

7.6.1 Luminosity Uncertainty

An uncertainty of 2.6% [72] on the yield of all the processes except QCD is considered as the uncertainty on luminosity measurements. Since QCD contribution is estimated from data, this uncertainty is not considered for QCD process. This uncertainty is correlated for all the channels.

7.6.2 Uncertainty due to Efficiency of muon, electron and τ_h Identification, Isolation

The uncertainties on muon and electron identification and isolation efficiencies [52] are determined from the uncertainties in the MC to data scale factors. Uncertainties of 2% [73] and 3% [74] are considered for muons and electrons respectively. The uncertainty on the τ_h identification efficiency is measured using $Z/\gamma \to \mu\tau_h$ events and 6% uncertainty is considered as τ_h identification uncertainty.

7.6.3 τ_h Energy Scale

1.2% uncertainty on τ_h energy scale is considered. 4-momentum of the reconstructed τ_h which is matched with the generator level τ decaying hadronically within $\Delta\mathrm{R} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.5$, is changed by $\pm 1.2\%$. The change in the yield for invariant mass of 4 particles in the signal region is passed as nuissance parameter for the cut and count analysis with same-sign leptons. The shape of the invariant mass of 4 particles with up and down variations is used in the case of shape analysis with opposite-sign leptons.

7.6.4 Jet $\rightarrow \tau_h$ Fake Rate

Since jet$\rightarrow \tau_h$ fake rate depends on the quark flavor and in the present analysis it is determined from DY + jets dominated events where softer jets are present, 50% uncertainty is considered for all reducible background processes where one or two τ_h's can be faked by jets. This uncertainty is not considered for QCD process, as it is calculated using a data driven method. This uncertainty is correlated among all the other channels.

7.6.5 Cross Sections

For the signal, a 5% uncertainty due to renormalization and factorization scale variation and 3% due to the uncertainties on parton distribution functions are considered. The cross section uncertainty considered for different background processes are mentioned in Table 7.13.

7.6.6 QCD Background Yields

The QCD background, which is estimated from data in the three different control regions, as explained in Section 7.3.5, affects the number of events in the control regions subject to Poissonian fluctuation [69]. This is modeled as an uncertainty following Gamma distribution and with central value given by $N \times \alpha$, where N is the total number of events in the three different control regions and α is the average weight applied to all the events so that $N\alpha$ is the estimated QCD contribution in signal region. This uncertainty is uncorrelated among the different channels.

7.6.7 Bin-by-bin Uncertainty

To consider the statistical uncertainty, bin-by-bin uncertainty is considered for each process in each bin.

The magnitude of the systematic uncertainties, considered in this analysis, is shown in Table 7.13.

Table 7.13: Sources of systematic uncertainties.

Source of systematic uncertainty	**Magnitude**	**Process**
Luminosity	2.6%	all processes except QCD
Muon ID and Isolation efficiency	2%	all processes except QCD
Electron ID and Isolation efficiency	3%	all processes except QCD
Tau ID and Isolation efficiency	6%	all processes except QCD
Tau energy scale	1.2%	all processes except QCD
PDF	3%	for signal
ZH cross section	4%	ZH
WZ cross section	6%	WZ
Single top cross section	4%	Single top
$t\bar{t}$ cross section	10%	TT
DY cross section	10%	DY
$ZZ \to 4\ell$ cross section	10%	$ZZ \to 4l$

7.7 Results

In the absence of any excess in data over background events as expected from the Standard Model processes, we calculate a 95% CL upper limits on $\sigma\,(\mathrm{pp} \to \mathrm{X}) \times \mathrm{BR}(HH \to \tau\tau\tau\tau)$ as a function of Radion mass. Figure 7.28 shows the observed (expected) limit as a function of Radion mass for $\mu\mu\tau_h\tau_h$, $\mathrm{ee}\tau_h\tau_h$, $\mu\mathrm{e}\tau_h\tau_h$ channels with same-sign leptons. For opposite-sign leptons, limits are shown in Figure 7.29 for $\mu\mu\tau_h\tau_h$, $\mathrm{ee}\tau_h\tau_h$, $\mu\mathrm{e}\tau_h\tau_h$ channels. In all these cases limits are calculated by the Asymptotic method [71] using the combination tool developed by the Higgs combination group. Limits are given in terms of the cross section of $\mathrm{pp} \to \mathrm{X} \to HH \to \tau\tau\tau\tau$. Figure 7.30 shows the observed (expected) combined limit as a function of Radion mass. Figure 7.31 shows the comparison of limit on $\sigma(\mathrm{pp} \to \mathrm{X}) \times \mathrm{BR}(HH)$ in different Higgs boson decay modes.

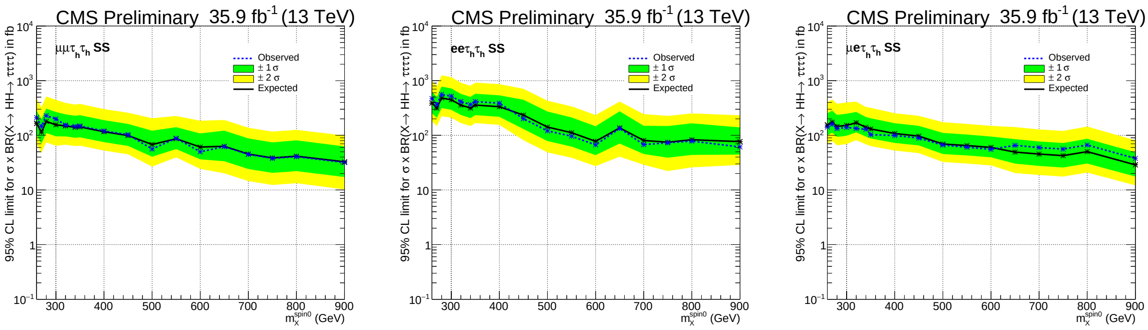

Figure 7.28: Expected 95% CL upper limits on $\sigma(\mathrm{pp}\to\mathrm{X}) \times \mathrm{BR}(HH\to\tau\tau\tau\tau)$ as a function of Radion mass for $\mu\mu\tau_h\tau_h$ (left plot), $\mathrm{ee}\tau_h\tau_h$ (middle plot) and $\mu\mathrm{e}\tau_h\tau_h$ (right plot) with same-sign leptons.

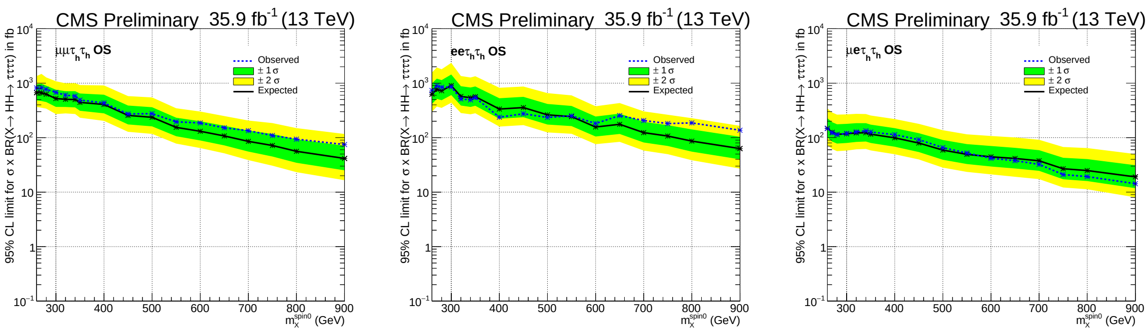

Figure 7.29: Expected 95% CL upper limits on $\sigma(\mathrm{pp}\to\mathrm{X}) \times \mathrm{BR}(HH\to\tau\tau\tau\tau)$ as a function of Radion mass for $\mu\mu\tau_h\tau_h$ (left plot), $\mathrm{ee}\tau_h\tau_h$ (middle plot) and $\mu\mathrm{e}\tau_h\tau_h$ (right plot) with opposite-sign leptons.

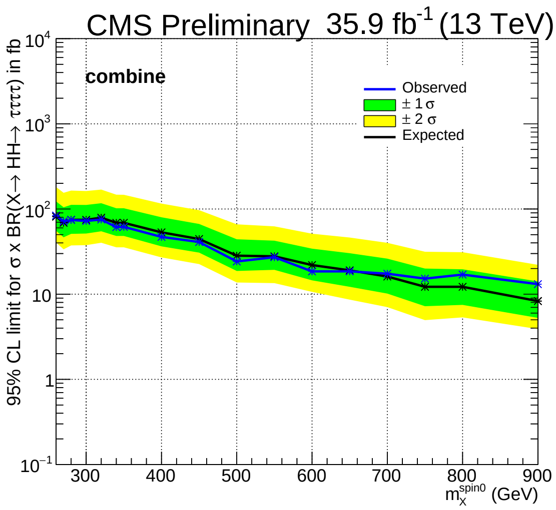

Figure 7.30: Expected 95% CL combined upper limits on $\sigma(\mathrm{pp}\to\mathrm{X}) \times \mathrm{BR}(HH\to\tau\tau\tau\tau)$ as a function of Radion mass obtained from $\mu\mu\tau_h\tau_h$, $\mathrm{ee}\tau_h\tau_h$, $\mu\mathrm{e}\tau_h\tau_h$ channels with opposite-sign and same-sign leptons.

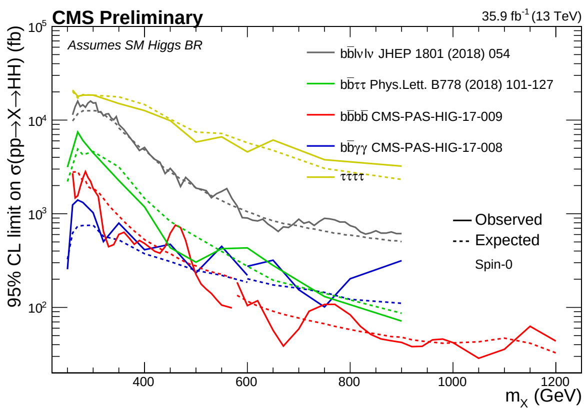

Figure 7.31: Expected 95% CL upper limits on $\sigma(\mathrm{pp}\to\mathrm{X}) \times \mathrm{BR}(\mathrm{X}\to HH)$ as a function of Radion mass for different decay modes of Higgs boson.

7.8 Conclusion and outlook

The results of $HH \to 4\,\tau$ have been compared with other final states in Figure 7.31 where it can be seen that this channel can be competitive in the low mass region. A number of improvements are planned for the future. Only 30% of all the $HH \to 4\,\tau$ final states has been considered in this analysis. Other decay channels of τ, like,

- $HH \to \tau\tau\tau\tau \to \tau_h\tau_h\tau_h\tau_h$
- $HH \to \tau\tau\tau\tau \to \ell\ell\ell\ell$
- $HH \to \tau\tau\tau\tau \to \ell\ell\ell\tau_h$
- $HH \to \tau\tau\tau\tau \to \ell\tau_h\tau_h\tau_h$

must be added to achieve higher sensitivity. Also invariant mass of 4 particles should be reconstructed using the SVFit [75] algorithm where neutrinos from the τ decay are considered properly. This would lead to good mass resolution of signal. Moreover, this analysis is constrained by limited statistics. Improved estimation of background contribution from data, where one or more particles are faked by jets, will reduce statistical uncertainty. In Run3 and beyond, *i.e* during HL-LHC with much higher statistics, there will be other possibilities to improve this analysis so that it can compete with the major final states : $b\bar{b}b\bar{b}$, $bb\gamma\gamma$, $bb\tau\tau$, $bbWW$.

Bibliography

[1] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt1314TeV2014.

[2] CMS Collaboration, "Search for the standard model Higgs boson decaying into two photons in pp collisions at $\sqrt{s}$ = 7 TeV", *Phys.Lett.B* **710** (2012) 403.

[3] CMS Collaboration, "Search for the standard model Higgs boson in the decay channel H to ZZ to 4 leptons in pp collisions at $\sqrt{s}$ = 7 TeV", *Phys.Rev.Lett*. 108, 111804(2012)

[4] http://pdg.lbl.gov/2016/reviews/rpp2016-rev-higgs-boson.pdf.

[5] ATLAS and CMS collaboration, "Combined Measurement of the Higgs boson Mass in pp collisions at $\sqrt{s}$ = 7 and 8 TeV with the ATLAS and CMS Experiments", *Phys.Rev.Lett* **114** (2019) 191803.

[6] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWGHH#Current_recommendations_for_HH_c.

[7] Gautam Bhattacharyya, "Hierarchy problem and BSM physics", $Pramana-J.Phys$ (2017).

[8] https://arxiv.org/pdf/1602.00170.pdf.

[9] R. J. Gaitskell, "Direct detection of dark matter", *AnnualReviewofNuclearandParticleScience*, vol. 54, no. 1, pp. 315-359, 2004.

[10] G. P. A. Azatov, R. Contino and M. Son, "Effective field theory analysis of double Higgs boson production via gluon fusion", *Phys.Rev.* **D92** (2015) 035001.

[11] L. Y. F. Goertz, A. Papefsatathiou and J. Zurita, "Higgs boson pair production in the d = 6 extension of the SM", *JHEP* **04** (2015) 167.

[12] Tania Robens, Tim Stefaniak, "LHC Benchmark Scenarios for the Real Higgs Singlet Extension of the Standard Model", *Eur.Phys.J* **C76** (2016).

[13] G. C. Branco *et al.*, "Theory and phenomenology of two-Higgs-doublet models", *Phys.Rep.* **516** (2012) 1.

[14] "Phenomenology of 2HDM with vector like quarks", *Phys.Rev.D* **97** 095015.

[15] Marcin Badziak,"Interpreting the 750 GeV diphoton excess in minimal extensions of Two-Higgs-Doublet models", *PhysLettB* **759** 2016, 464-470.

[16] L. Randall and R. Sundrum, " A. large mass hierarchy from a small extra dimension", *Phys.Rev.Lett.* **83** (1999) 3370.

[17] Gi-Chol Cho, Yoshiko Ohno, "Production and decay of radion in Randall-Sundrum model at a photon collider", High Energy Physics - Phenomenology (hep-ph).

[18] "The LHC Machine", *JINST* **03** 2008 S08001.

[19] "The CMS experiment at the CERN LHC", *JINST* **3** 2008 S08004.

[20] ATLAS Collaboration, "The ATLAS experiment at the CERN Large Hadron Collider", *JINST*, 3:S08003, 2008.

[21] ALICE Collaboration, "The ALICE experiment at the CERN LHC", *JINST*, 3:S08002, 2008.

[22] LHCb Collaboration, "The LHCb detector at the LHC", *JINST*, 3:S08005, 2008.

[23] "CMS Physics Technical Design Report" Vol. 1, February, 2006.

[24] "The Magnet Project Technical Design Report", CMS TDR, 2 May 1997.

[25] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", $JINST$ **8** (2013) P11002.

[26] "The CMS trigger system", Journal of Instrumentation, **12** 2017.

[27] "Design, performance, and calibration of CMS hadron-barrel calorimeter wedges" $Eur.Phys.J.C55, 1C$. **55** 159-170, (2008).

[28] "Design, Performance, and Calibration of CMS Hadron Endcap Calorimeters", CMS Note (internal) 2008/010.

[29] "Design, performance, and calibration of CMS forward calorimeter wedges" $Eur.Phys.J.C$ **53** , 139-166 (2008).

[30] "Design, performance, and calibration of the CMS hadron-outer calorimeter" $Eur.Phys.J.C$ **57** 2008.

[31] "P. Billoir and S. Qian, Simultaneous pattern recognition and track fitting by the Kalman filtering method,", $Nucl.Instrum.Meth.$, **A 294**, (1990) 219.

[32] "R. Frühwirth, W. Waltenberger and P. Vanlaer, Adaptive Vertex Fitting,"

[33] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", $JINST$ **7** (2012) P10002.

[34] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV", $JINST$ **10** (2015) P06005.

[35] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", $JINST$, **8**, (2013) P04013.

[36] CMS Collaboration, "Reconstruction and identification of tau lepton decays to hadrons and tau neutrino at CMS", *JINST* **11** (2016) P01019.

[37] CMS Collaboration, "Performance of reconstruction and identification of τ leptons in their decays to hadrons and ν_τ in LHC Run-2", CMS PAS TAU-16-002.

[38] https://cds.cern.ch/record/2221747/files/DP2016_064.pdf.

[39] https://cds.cern.ch/record/2243387/files/ATL-PHYS-PUB-2017-001.pdf.

[40] Amit Adhikary, *et al.*, "Revisiting the non-resonant Higgs pair production at the HL-LHC", *JHEP* **1807** (2018), 116.

[41] "A standard format for Les Houches Event Files", J. Alwall *et al*, Computer Physics Communication **176** (2007) 300.

[42] "A Brief Introduction to Pythia 8.1", T. Sjöstrand *et al*, Computer Physics Communication **178** (2008) 852.

[43] "DELPHES 3, A modular framework for fast simulation of a generic collider experiment", J. de Favereau *et al*, JHEP **1402** (2014) 057.

[44] "On the spin and parity of a single-produced resonance at the LHC", Sara Bolognesi *et al*, Physical Review **D86** (2012) 095031.

[45] "On resonance search in dijet events at the LHC", M. V. Chizhov *et al*, https://arxiv.org/abs/1106.4161.

[46] https://arxiv.org/pdf/1301.4965.pdf.

[47] "Higgs pair production in vector-boson fusion at the LHC and beyond", *Eur.Phys.J.C* (2017) 77:48.

[48] "TMVA: Toolkit for Multivariate Data Analysis with ROOT", http://tmva.sourceforge.net/.

[49] J. C. Collins and D. E. Soper, "Angular distribution of dileptons in high-energy hadron collisions", *Phys.Rev.D* **16** (1977) 2219-2225.

[50] CMS Collaboration, "Search for non-resonant Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state at $\sqrt{s} = 13$ TeV", *J.HighEnergyPhys* 2018.

[51] CMS Collaboration, "Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state in pp collisions $\sqrt{s} = 13$ TeV", *PhysicsLettersB* **788** (2018) 7.

[52] CMS Collaboration, "Search for Higgs boson pair production in events with two bottom quarks and two tau leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys.LettB* **778** (2018) 101.

[53] CMS Collaboration, "Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}l\nu l\bar{\nu}$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **01** (2018) 054.

[54] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections and their matching to parton shower simulations",, JHEP 07 (2014) 079.

[55] P.Nason, "A New Method for combining NLO QCD with Shower Monte Carlo Algorithms", *JHEP* **11** (2004) 040.

[56] P. N. S. Frixione and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method" *JHEP* **11** (2007) 070.

[57] C. O. S. Alioli, P. Nason and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043.

[58] S. M. S. Alioli and P. Uwer, "Hadronic top-quark pair-production with one jet and parton showering", *JHEP* **01** (2012) 137.

[59] C. O. S. Alioli, P. Nason and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **09** (2009) 111.

[60] S. M. T. Sjöstrand, and P. Z. Skands, "A brief introductio to pythia 8.1", *Comput.Phys,Commun.* **178** (2008) 852.

[61] C. Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur.Phys.J.C* **76** (2015) 155.

[62] Geant4 User's Guide for Application Developers by Geant4 Collaboration, geant4 **10.2** 2015.

[63] W. Erdmann, "Offline Primary Vertex Reconstruction with Deterministic Annealing Clustering", *CMSInternalNote* 2011/014 (2011).

[64] CMS Collaboration, "Performance of cms muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) P10002.

[65] https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Medium_Muon.

[66] https://twiki.cern.ch/twiki/bin/view/CMS/MultivariateElectronIdentificationRun2#Non_triggering_electron_MVA_deta

[67] https://github.com/CMS-HTT/LeptonEfficiencies.

[68] https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaIDRecipesRun2.

[69] CMS AN-2017/007 – Search for Higgs boson pair production in the bb$\tau\tau$ final state with the CMS detector at the LHC.

[70] https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauIDRecommendation13TeV#Tau_energy_scale.

[71] Asymptotic formulae for likelihood-based tests of new physics, *Eur.Phys.J.C* 71:1554,2011.

[72] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, CERN, 2017.

[73] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018) P06015.

[74] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015) P06005.

[75] Reconstruction of the Higgs mass in $H \to \tau\tau$ Events by Dynamical Likelihood techniques, Lorenzo Bianchini *et al*, J. Phys.: Conf. Ser. **513** (2014) 022035.