SEARCH FOR EXTENDED HIGGS SECTOR AT THE LHC USING MULTILEPTONS AND PHOTONS IN THE FINAL STATE

BY SHRUTI M. PANWALKAR

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Physics and Astronomy

Written under the direction of

Dr. Sunil Somalwar

and approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2014

ABSTRACT OF THE DISSERTATION

Search for Extended Higgs Sector at the LHC using multileptons and photons in the final state

by Shruti M. Panwalkar

Dissertation Director: Dr. Sunil Somalwar

This dissertation describes the search for an extension of the Higgs sector to Two Higgs Doublet Models. This entails searching for decays of the heavy scalar (H) and pseudo-scalar (A) Higgs bosons. These decays include the Standard Model-like Higgs h in the final state and lead to events with isolated leptons and photons. This search is carried out using a 19.5 fb^{-1} data sample at $\sqrt{s} = 8$ TeV collected by the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC). Observed multilepton events with or without diphoton candidates are organized into exclusive search channels based on event kinematics. The search channels are ordered by the amount of expected Standard Model background. Standard Model backgrounds are estimated using data-based techniques and data-validated simulations. The observed events are consistent with the Standard Model expectation. Limits are placed on the cross section of gluon-gluon fusion production of H and A. Exclusions of the Two Higgs Doublet Models (TYPE I and TYPE II) are presented in terms of the relevant parameters of the model.

Acknowledgements

I would like to take this opportunity to express my sincere gratitude for my advisor Sunil Somalwar. He not only gave me an insight into the world of particle physics but also taught me to think like an experimentalist. He has always been concerned about the overall welfare of his students, even outside the physics department.

I would like to thank Richard Gray for guiding me through the twists and turns of a good analysis. He always had immense patience with silliest of my questions and words of encouragement when the road ahead seemed uphill. Matt Walker encouraged us to write better and efficient code. It has been a privilege to work with these brilliant postdocs. Emannuel, Partick, Peter, Sanjay have been the most amazing fellow multileptoners. It has been a pleasure working with them and I have learnt a lot from them in the process. I would like to thank Eva Halkiadakis, Amit Lath, John Paul Chou, Steve Schnetzer and Yuri Gershtein for their valuable advise and always having time to discuss any of my queries, however insignificant they may be. Scott Thomas has been like my second advisor. He gave me the much needed physics intuition, crucial for any kind of analysis. Thanks to Nathaniel and Jared for answering all the theory related questions (Trust me, the list was never ending). David, Kelvin, Kin and Stephen are the brightest undergrads I have met. I wish them all the best for their future endeavors. Thanks to Pieter Jacques and John Doroshenko for making sure the computer farm was running smoothly every day of the year.

I would like to thank Sourabh for introducing me to the world of particle physics. He has been my mentor since my undergrad days and always has the best advice on every topic; be it applying to graduate schools or giving presentations or writing thesis. Thank you for paving an exemplary path that I could trace right from school to seeking my graduate advisor!

Dedication

For

Aai, Baba, Kt

& the pure joy of learning

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction to High Energy Physics

What is the 'recipe' for making all the things around us? Are there any basic 'ingredients' that make up everything in the universe? Questions like these have been asked by curious minds since time immemorial. The study of basic building blocks of matter and the forces which act between them is a fundamental area of research, known as particle physics. It is also known as 'High Energy Physics' (HEP) for the following reason. The structures under scrutiny are extraordinarily small, well below 10^{-15} m. In order to perform experiments at this scale, probes with corresponding high spatial resolution are needed. Visible light, with a wavelength $\lambda \sim 500$ nm, is absolutely inadequate. A microstructure may only be resolved by a probe, if its wavelength is small compared to the size of the structure. Thus a wavelength (λ) well below 10^{-15} m, and hence extremely high energy, is required in elementary particle physics.

At present, our knowledge and understanding of the field of particle physics can be satisfactorily explained by the 'Standard Model' (SM). The SM is a wonderful collection of theories about all the known elementary particles and their interactions.

Though SM successfully describes particle physics phenomena probed by experiments, it is still unsatisfactory in several ways. One of these is the incomplete understanding of the mechanism for electroweak symmetry breaking (EWSB). Fulfilling this goal not only includes the characterization of the recently discovered SM-like Higgs boson [1], but also warrants a search for additional Higgs states which could provide a window into the underlying physics of EWSB [2].

This dissertation is the search for an extension of the SM Higgs sector to such additional Higgs states. In chapter 2, I shall describe the SM and the need to go beyond its current scope. In chapter 3, I will discuss the apparatus for the search, that

is the Large Hadron Collider (LHC) and the Compact Muon Solenoid (CMS) detector. In the subsequent chapters, I will describe the analysis in detail. These chapters are based on the background estimations and allied material that our Rutgers high energy group has developed over the past few years, which has been documented in earlier internal CMS notes and publications([3], [4], [5] and [6]). Finally, results for the search are presented in chapter 7.

Chapter 2

The Standard Model And Beyond

2.1 The Standard Model

The basis of the Standard Model (SM) of particle physics was developed in the first half of the 1970s out of the work of Glashow,Weinberg, Salam and confirmed by numerous experimental probes. It describes three out of the four known fundamental interactions (viz. strong interaction, weak interaction, electromagnetic interaction) and the elementary particles that take part in these interactions. These particles make up all matter in the universe. The SM is a gauge theory of the electroweak (electromagnetism+weak interaction) and strong interactions with the gauge group $SU(2)\times U(1)\times SU(3)$.

2.1.1 Brief History of the Standard Model

The SM was not put together in one attempt. Since the 1800's scientists have been trying to get to the crux of this 'matter'. In 1897 J.J Thompson discovered the electron and came up with the 'plum pudding' model. Thereafter Rutherford discovered the positive nucleus in 1911 and Chadwick discovered neutrons in 1932 .Thus was born the planetary model of an atom. For the next couple of decades it was believed that the proton, electron and neutron are the most fundamental particles of matter. Later in 1964 Murray Gell-Mann and George Zweig tentatively put forth the idea of quarks (Sec.2.1.2). They suggested that mesons and baryons are composites of three quarks or antiquarks, viz. up, down and strange (u, d, s), with spin half and electric charges $2/3, -1/3, -1/3$, respectively. It later turned out that this theory is not completely accurate. In 1967 Steven Weinberg and Abdus Salam separately proposed a theory that unifies electromagnetic and weak interactions into the electroweak interaction. In 1983 the $W^{\pm}$ and Z^0 bosons required by the electroweak theory, were observed by two

experiments at CERN using techniques developed by Carlo Rubbia and Simon Van der Meer to collide protons and antiprotons.

Components of the SM are described in greater detail in the following sections. The SM particles are split into *fermions* (particles that form matter) and *bosons* (particles that describe interaction).

2.1.2   Fermions

Fermions are spin half particles that obey the Fermi-Dirac statistics. They follow Pauli's exclusion principle. They can be further categorized into quarks and leptons.

Quarks: In addition to weak and electromagnetic interaction, the quarks also interact via strong interaction. A quantum field theory of strong interaction, known as Quantum Chromodynamics (QCD), was formulated in 1973 . There are six types of quarks, characterized by flavors: up, down, charm, strange, top and bottom (refer to table tab:leptonprop). Up and down quarks have the lowest masses of all quarks. The heavier quarks rapidly change into up and down quarks through a process of particle decay: the transformation from a higher mass state to a lower mass state. For this reason, up and down quarks are generally stable and the most common in the universe, whereas charm, strange, top and bottom quarks can only be produced in high energy collisions (such as those involving cosmic rays and in particle accelerators). Quarks (or antiquarks) carry a baryon quantum number of B = 1/3 (or B = −1/3), which is approximately conserved. A quark- antiquark pair can form a meson (B = 0) and three quarks can form a baryon (B =±1). Other combinations of quarks have not yet been observed. For example, protons, made of two up and one down quark, and neutrons, made of one up and two down quarks, are baryons. As quarks are spin$-\frac{1}{2}$ particles, they can occur in a left- and right-handed form according to the orientation of their spin. In 1965, O.W. Greenberg, M.Y. Han and Yoichiro Nambu introduced the quark property of color charge. It is a very peculiar characteristic of quarks. Color charge may be *red blue* or *green*. The presence of equal number of quarks of each color makes a particle colorless. All observed hadrons (particles made of quarks and gluons) are color neutral.

Leptons: These are fermions which can interact via weak and electromagnetic interaction only. They can not 'feel' the strong interaction. The first lepton to be identified was the electron. Then in 1930, Wolfgang Pauli postulated the electron neutrino to preserve conservation of energy, conservation of momentum, and conservation of angular momentum in beta decay. The muon was discovered by Carl D. Anderson in 1936. There are 6 types of leptons,characterized by flavors: electron (e^-) and electron neutrino (ν_e); muon (μ^-) and muon neutrino (ν_μ); tau (τ^-) and tau neutrino (ν_τ). Leptons (or antileptons) carry a lepton quantum number of L = 1 (or L = −1), which is approximately conserved . The difference of baryon and lepton quantum number B−L is absolutely conserved.

The elementary fermions can be ordered in three generations, where the first generation is made of up and down quark as well as the electron and its neutrino. The second generation consists of charm and strange quark as well as muon and its neutrino. The third generation is comprises of bottom quark, top quark, tau and its neutrino. Experiments carried out in SLAC and CERN strongly suggest that there are three and only three generations of fundamental particles. This is inferred by showing that the Z^0-boson lifetime is consistent only with the existence of exactly three very light (or massless) neutrinos.

2.1.3 Gauge Bosons

Gauge bosons are integer spin particles that act as carriers of the fundamental forces of nature. They obey the Bose-Einstein statistics. The gauge bosons included in the SM are the photon(γ) which mediates the electromagnetic interaction, the (W^+, W^-, Z^0) bosons which mediate the weak force, and the gluon g, which mediates the strong color force. Due to color confinement, isolated gluons do not occur at low energies. The hypothetical graviton, G, which might be mediating gravity is not included in the SM.

Name	Symbol	Baryon Number B	Lepton Number L	Charge	Mass (MeV)
Leptons	l				
electron	e	0	1	-1	0.511
electron neutrino	ν_e	0	1	0	$< 2 \times 10^{-6}$
muon	μ	0	1	-1	106
muon neutrino	ν_μ	0	1	0	$< 2 \times 10^{-6}$
tau	τ	0	1	-1	1777
tau neutrino	ν_τ	0	1	0	$< 2 \times 10^{-6}$
Quarks	q				
up	u	1/3	0	$+2/3$	1.5 to 3.3
down	d	1/3	0	$-1/3$	3.5 to 6
charm	c	1/3	0	$+2/3$	1270
strange	s	1/3	0	$-1/3$	104
top	t	1/3	0	$+2/3$	1.71×10^5
bottom	b	1/3	0	$-1/3$	4.2×10^3

Table 2.1: Properties of standard model fermions

Name	Symbol	Charge	Mass (GeV)
Photon	γ	0	0
	$W^\pm$	± 1	80.2
	Z	0	91.2
gluon	g	0	0
Higgs	h	0	125

Table 2.2: Properties of the standard model bosons

2.2   Shortcomings in the SM

Even though the SM is a successful theory with a backing from experimental results it still remains incomplete. It falls short of explaining the matter-antimatter asymmetry in the universe. As yet, we do not know the nature of dark matter. SM says that the neutrinos should be massless. However the neutrino oscillation experiments in the past decade have shown the neutrinos to have masses. Gravitational interaction has still not found a place in the SM framework. Until very recently SM could not explain why particles have mass. The discovery of Higgs boson now offers a window into the mechanism of electroweak symmetry breaking, there by making SM particles massive.

2.2.1 Local gauge invariance

Discussion herein follows from Ref [7] , [8] and [9]. SM is built upon the expectation of local gauge invariance. This is crucial to keep any theory 'renormalizable' and hence be able to compute physical effects in the presence of divergences that may be introduced by perturbative expansions of QED (and QCD) amplitudes [7]. For example, consider the lagrangian-

$$L \; = \; i\bar{\psi}\gamma^\mu \partial_\mu \psi \; - \; m\bar{\psi}\psi \tag{2.1}$$

Under local gauge invariance we have-

$$\psi(x) \; \rightarrow \; e^{i\alpha(x)}\psi(x) \tag{2.2}$$

However the lagrangian does not remain invariant under such local gauge (i.e. phase) transformation. In order to keep it invariant, we replace ∂_μ by a modified derivative D_u such that-

$$D_\mu \; \equiv \; \partial_\mu \; - \; ieA_\mu \tag{2.3}$$

Here e is the electric charge. Also, we have introduced a new field A_μ that transforms as-

$$A_\mu \; \rightarrow \; A_\mu \; + \; \frac{1}{e}\partial_\mu \alpha \tag{2.4}$$

The invariance of lagrangian will only be preserved if introduction of A_μ does not introduce an additional mass term (such as $\frac{1}{2}m^2 A_\mu A^\mu$). That means A_μ needs to be massless and can be physically understood as the photon field. Thus we have achieved a interacting field theory of QED. An analogous argument can be presented for QCD by replacing $U(1)$ gauge group by $SU(3)$ since gluons (carries of strong force) are also massless. Yet, these arguments break down for interactions mediated by massive gauge bosons like W and Z. This happens because the mass term for gauge fields destroys

gauge invariance of the lagrangian. Ignoring the local gauge symmtery will make the theory unrenormalizable. Hence the mass term needs to be introduced in a way that preserves local gauge invariance. This is called Higgs mechanism.

2.2.2 Higgs Mechanism

As seen in the previous section, to have massive gauge bosons and preserve local gauge symmetry of $SU(2)$ mass term need to be generated by 'spontaneous symmetry break-ing. It is called 'spontaneous' as no external agency is responsible for this symmetry breaking [10]. This is done by introduction of a scalar field ϕ[1]. This field needs to be $SU(2)$ doublet in order for its coupling with fermions to be gauge invariant. Its coupling to bosons are ensured to be gauge invariant by the same trick of modifying the derivative to covariant derivative.

This doublet is taken as the Higgs doublet and has the form-

$$\phi \; = \; \begin{pmatrix} \phi_\alpha \\ \phi_\beta \end{pmatrix}, \; where \; \begin{aligned} \phi^+ &\equiv (\phi_1 \; + \; i\phi_2)/\sqrt{(}2), \\ \phi^0 &\equiv (\phi_3 \; + \; i\phi_4)/\sqrt{(}2), \end{aligned} \tag{2.5}$$

Higgs potential can be considered to have this general SU(2)-invariant form-

$$V(\phi) \; = \; \mu^2 \phi^\dagger \phi \; + \; \lambda(\phi^\dagger \phi)^2 \tag{2.6}$$

For $\mu^2 \; < 0$, $\lambda \; > 0$, the Higgs potential has a minimum value at $\frac{\mu^2}{2\lambda}$

'This manifold of points at which $V(\phi)$ is minimized is invariant under $SU(2)$ trans-formations' [7]. We need expand $\phi(x)$ about some minimum and to do that let us choose-

$\phi_1 \; = \; \phi_2 \; = \; \phi_4 \; = \; 0$ and $\phi_3 = \sqrt{\frac{-\mu^2}{\lambda}} \; \equiv \; v$

Substituting ϕ_0 in the lagrangian gives rise to three massive gauge fields and one massive scalar h. This choice of expectation value to be that of the neutral field ϕ_3, provides for a massless photon. The equation 2.5 then gives-

_____

[1]Complete derivation can be found in [8]

$$\langle \phi \rangle \;\equiv\; \left\langle 0 \left| \begin{pmatrix} \phi_\alpha \\ \phi_\beta \end{pmatrix} \right| 0 \right\rangle \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.7}$$

As $\tau_i \langle 0 \rangle \neq 0$ and $Y \langle 0 \rangle \neq 0$, the $SU(2)$ and $U(1)_Y$ symmetry is spontaneously broken.

Electroweak symmetry breaking is the breaking $SU(2) \times U(1)$ symmetry and here single Higgs doublet was used to understand the mechanism for EWSB.

2.2.3   Extension of Higgs Sector

The mechanism of EWSB might be simple enough to be explained by single Higgs doublet or it may need to be extended to second Higgs doublet [11]. 'The phenomenology of 2HDMs is rich as five physical Higgs sector particles survive EWSB: two neutral CP-even scalars, h, H; one neutral CP-odd pseudoscalar, A; and two charged scalars, H^+ and H^-. All of these states could have masses at or below the TeV scale, which is a regime that is accessible to the LHC' [2].The two Higgs doublets are denoted by Φ_1 and Φ_2. As a norm, Φ_2 stands for the Higgs doublet and couples to Q$\bar{u}$. Four discrete types of models satisfy the Glashow-Weinberg condition. Out of these four, we will consider TYPE I and TYPE II 2HDM scenarios. The coupling of the two Higgs doublets to $SU(2)_L$ singlet fermions is given Table 2.3-

	TYPE I 2HDM	TYPE II 2HDM
u	Φ_2	Φ_2
d	Φ_2	Φ_1
e	Φ_2	Φ_1

Table 2.3: Couplings of the Higgs boson to $SU(2)_L$ singlet fermions in TYPE I and TYPE II 2HDM. This table is taken from phenomenology paper by Nathaniel Craig et al. [12]

Detailed list of couplings of particles of the extended sector to SM fermions and massive gauge bosons in Type I and II 2HDMs is given in table 2.4.

The production cross section for $gg \rightarrow H \rightarrow hh$ for TYPE I and TYPE II 2HDMs vary as a function of α and β as shown in figure 2.1 and figure 2.2. Similarly figure

1	2HDM I	2HDM II
hVV	$sin(\beta - \alpha)$	$sin(\beta - \alpha)$
hQu	$cos\alpha/sin\beta$	$cos\alpha/sin\beta$
hQd	$cos\alpha/sin\beta$	$-sin\alpha/cos\beta$
hLe	$cos\alpha/sin\beta$	$-sin\alpha/cos\beta$
HVV	$cos(\beta - \alpha)$	$cos(\beta - \alpha)$
HQu	$sin\alpha/sin\beta$	$sin\alpha/sin\beta$
HQd	$sin\alpha/sin\beta$	$cos\alpha/cos\beta$
HLe	$sin\alpha/sin\beta$	$cos\alpha/cos\beta$
AVV	0	0
AQu	$cot\beta$	$cot\beta$
AQd	$-cot\beta$	$tan\beta$
ALe	$-cot\beta$	$tan\beta$

Table 2.4: Couplings of all five particles of the sector to SM fermions and massive gauge bosons are fixed in terms of α and β. This table is taken from phenomenology paper by Craig et. al. [12]
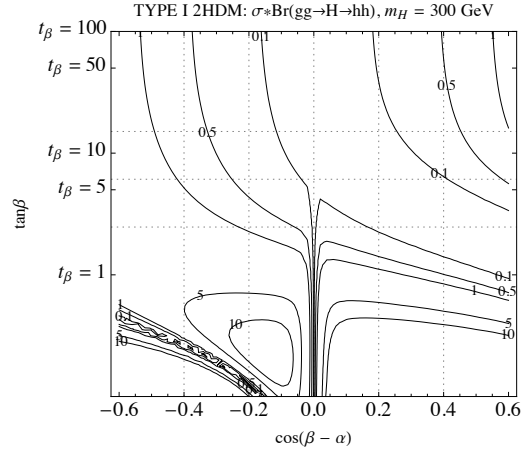
Figure 2.1: Contour plot for Type I 2HDM where contour lines give σ*BR for $H \rightarrow hh$. Parameters α and β give Heavy higgs's couplings to SM fermions and massive gauge bosons. Parameter λ is set to 0 GeV. This figure is similar to one from theory paper by Nathaniel Craig et al. [13], the only difference is plotting of tanβ, instead of β, on the vertical axis.
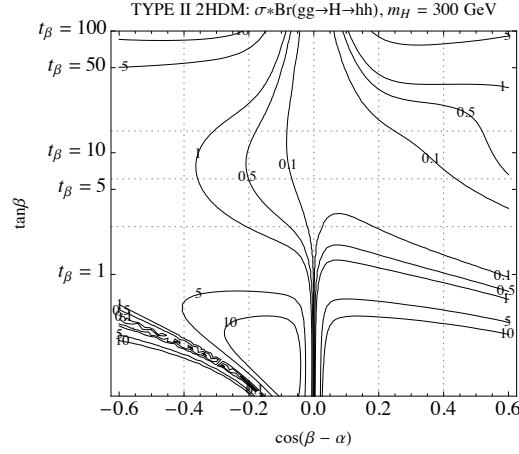
Figure 2.2: Contour plot for Type I 2HDM where contour lines give σ*BR for $H \to hh$. Parameters α and β give Heavy higgs's couplings to SM fermions and massive gauge bosons. Parameter λ is set to 0 GeV. This figure is similar to one from theory paper by Nathaniel Craig et al. [13], the only difference is plotting of $\tan\beta$, instead of β, on the vertical axis.
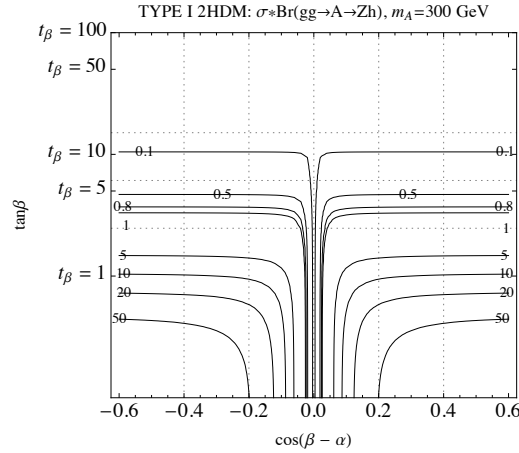
Figure 2.3: Contour plot for Type I 2HDM where contour lines give σ*BR for $H \to hh$. Parameters α and β give Heavy higgs's couplings to SM fermions and massive gauge bosons. Parameter λ is set to 0 GeV. This figure is similar to one from theory paper by Nathaniel Craig et al. [13], the only difference is plotting of $\tan\beta$, instead of β, on the vertical axis.
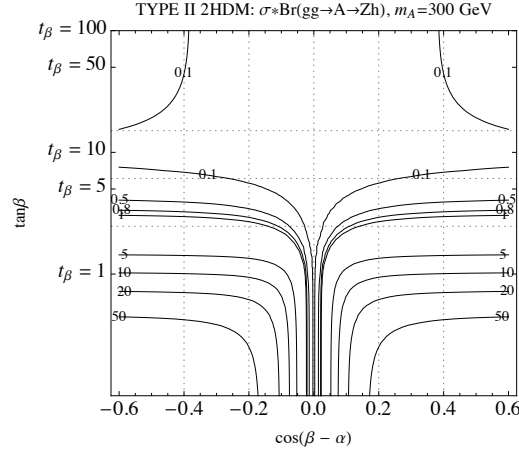
Figure 2.4: Contour plot for Type I 2HDM where contour lines give σ*BR for $H \rightarrow hh$. Parameters α and β give Heavy higss's couplings to SM fermions and massive gauge bosons. Parameter λ is set to 0 GeV. This figure is similar to one from theory paper by Nathaniel Craig et al. [13], the only difference is plotting of tanβ, instead of β, on the vertical axis.

2.3 and figure 2.4 show the variation of production cross section for $gg \rightarrow A \rightarrow Zh$ as a function of α and β. The branching ratio of h to WW, ZZ , $\tau\tau$ and $\gamma\gamma$ may deviate from SM values as we move in the parameter space of α and β. This deviation is shown in figure 2.5 and figure 2.6 for TYPE I and TYPE II 2HDM scenario respectively.

If the mass of the Heavy Higgs H were close to twice the mass of the SM Higgs, then apart from SM channels, it decays to two SM-like Higgs (h) [13]. For the particle A, the decay $A \rightarrow Zh$ is the primary decay mode when the mass of the A lies between two times the SM Higgs mass and two times the top mass. We consider the gluon-gluon fusion production of H followed by a decay to two SM-like Higgs particles. Here we assume h to have a nominal mass of 126 GeV and branching ratios to WW, ZZ, $\tau\tau$, bb and $\gamma\gamma$ channels appropriate to its mass. We will see in chapter 5 that there are seven distinct possibilities for the two SM-like Higgs to decay and provide multilepton final states: $WWWW$, $WWZZ$, $WW\tau\tau$, $ZZZZ$, $ZZ\tau\tau$, $ZZbb$, and $\tau\tau\tau\tau$. When one of the Higgs decays to $\gamma\gamma$, the final states $\gamma\gamma WW$, $\gamma\gamma ZZ$ and $\gamma\gamma\tau\tau$ are also useful in the search. We look for particle A decaying to Z boson and SM-like Higgs ($A \rightarrow Zh$). SM Higgs can then decay to WW, ZZ, $\tau\tau$ or $\gamma\gamma$. For the $A \rightarrow Zh$ search, the final
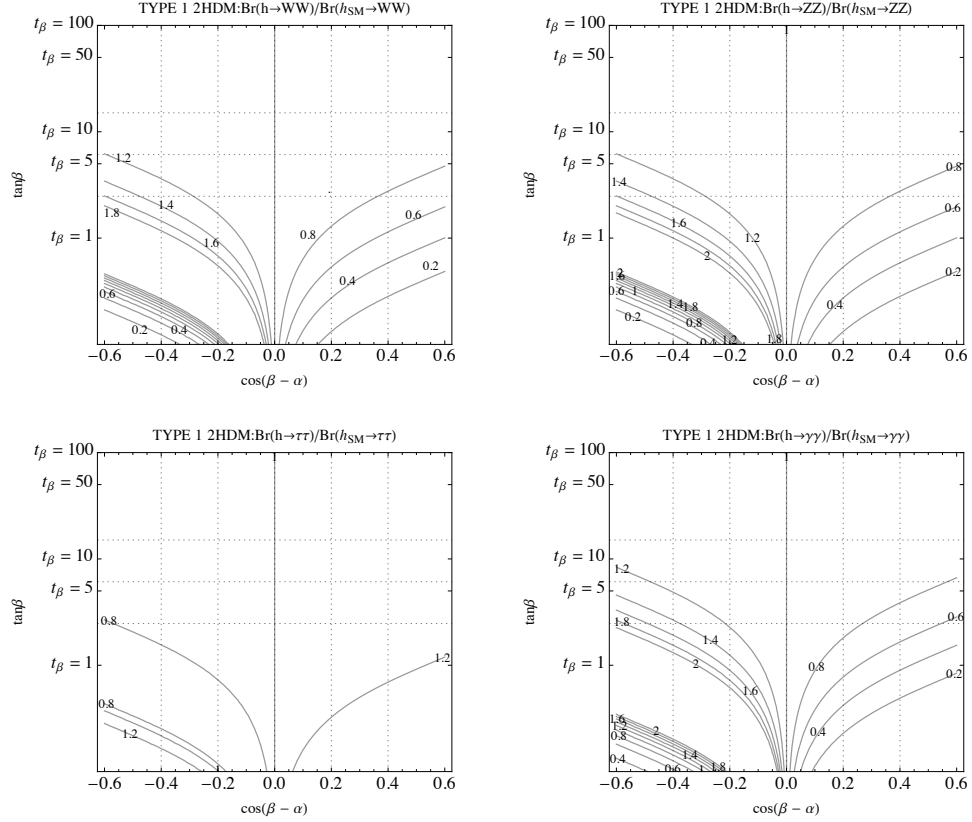
Figure 2.5: Contour plot for Type I 2HDM where contour lines give σ*BR for $H \to hh$. Parameters α and β give Heavy Higgs's couplings to SM fermions and massive gauge bosons. Parameter λ is set to 0 GeV. This figure is similar to one from theory paper by Nathaniel Craig et al. [13], the only difference is plotting of $\tan\beta$, instead of β, on the vertical axis.
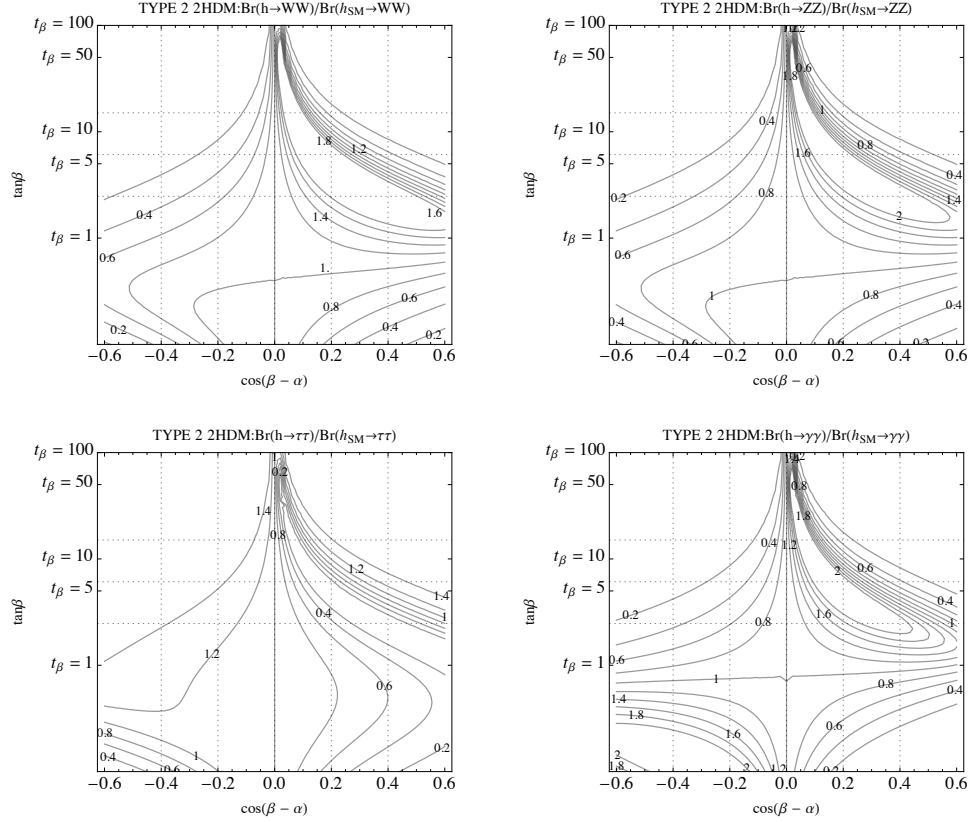
Figure 2.6: Contour plot for Type I 2HDM where contour lines give σ*BR for $H \to hh$. Parameters α and β give Heavy Higgs's couplings to SM fermions and massive gauge bosons. Parameter λ is set to 0 GeV. This figure is similar to one from theory paper by Nathaniel Craig et al. [13], the only difference is plotting of $\tan\beta$, instead of β, on the vertical axis.

states from Z boson and h decays that are useful when the Z decays leptonically are $llWW$,$llZZ$,$ll\tau\tau$ and $ll\gamma\gamma$ (Table 5.3). If the Z does not decay leptonically, then only the $h \rightarrow ZZ$ decay mode of the SM-like Higgs is useful.

2.2.4   Signal Generation

For this analysis all signal simulations were made using MADGRAPH 4.4.5 [14] and then decayed and showered through PYTHIA 6. We produce $H \rightarrow hh$ for H mass ranging from 260 GeV to 360 GeV. This mass range is chosen because Heavy Higgs primarily decays to two SM Higgs when its mass lies between two times the SM Higgs mass and two times the top mass. Each h can decay to WW, ZZ, $\tau\tau$, bb and $\gamma\gamma$ ($h \rightarrow WW$, $h \rightarrow ZZ$, $h \rightarrow \tau\tau$,$h \rightarrow bb$,$h \rightarrow \gamma\gamma$), for example, as shown in figure 2.8. We generate events separately for each of these decay modes for each mass point of H. At the end all the final states are taken together along with weights due to SM-like Higgs branching ratios.

For $A \rightarrow Zh$ search, with similar signal generation method, we produce grid for masses of A between 260 GeV to 360 GeV. As mention before within this mass range, the decay $A \rightarrow Zh$ is the primary decay mode. Hence the particular choice of mass range. Particle A decays to a Z boson and SM Higgs boson. Z can decay leptonically or otherwise and h can decay to WW,ZZ,$\tau\tau$ or $\gamma\gamma$.
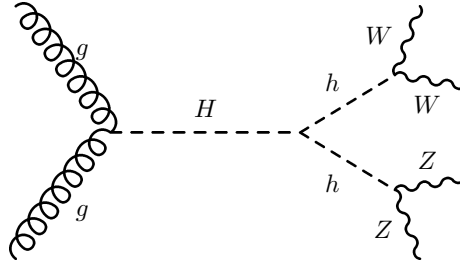
Figure 2.7: Heavy Higgs decays to 2 SM Higgs.One of the SM Higgs decays to WW, other to ZZ.

Next to next leading order (NNLO) results compiled by the Higgs cross section working group [15] for SM Higgs cross section of a given mass were used for gluon-gluon
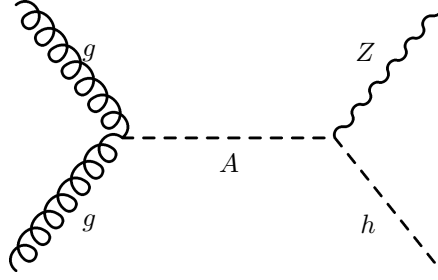
Figure 2.8: Particle Adecays to Z boson and SM Higgs.

fusion production cross sections of Heavy Higgs and particle A. For 2HDM scenario, SusHi [16] program was used to calculate the 2HDM cross sections and branching ratios for SM-like Higgs were calculated using 2HDMC [17]. These branching ratios were compared to the ones provided by the LHC Higgs cross-section group and found them to be the same. For example, for a particular point in 2HDM parameter space, say $cos(\beta - \alpha)$= 0.5 and $tan\beta$=2, the SM-like Higgs branching ratios to WW,ZZ, $\tau\tau$, $\gamma\gamma$ are hWW: 0.49,0.062,0.0032,0.0038 respectively.

Chapter 3

Detector

The LHC is the largest and highest energy accelerator in the world, located at CERN (*Conseil Europé n pour la Recherche Nucléaire*), along the Swiss-France border. This is the same location as that of the Large Electron-Positron (LEP) Collider. A significant number of modifications were made to the LEP structure before it was transformed into the proton-proton (pp) collider that we know today. The two main detectors that analyze the data from the pp collisions are Compact Muon Solenoid (CMS) and ATLAS (A Toroidal LHC Apparatus). The work presented here has been carried out at the CMS detector located at Cessy, France.

3.1 The Accelerator Complex

The Accelerator complex consists of three main components- 1.) The linac 2.) The proton boosters and 3.) The main accelerator ring. The layout of these components can be seen in figure 3.1

3.1.1 Obtaining and boosting protons

The Linac2 duoplasmatron is the source of protons for the LHC [18]. An electric field applied to the duoplasmatron metal cylinder breaks down the hydrogen gas to yield protons. These protons are then fed to the Proton Synchrotron Booster (PSB). Charged particles moving along a curved path emit electromagnetic radiation known as synchrotron radiation. Hence the name 'synchroton' booster. Protons are maintained along a curved path using several quadrupoles, bending and steering magnets. The PSB boosts this beam of protons to 1.4 GeV using varying electromagnetic fields in radio frequency (RF) cavities. The RF cavities are metallic chambers that contain
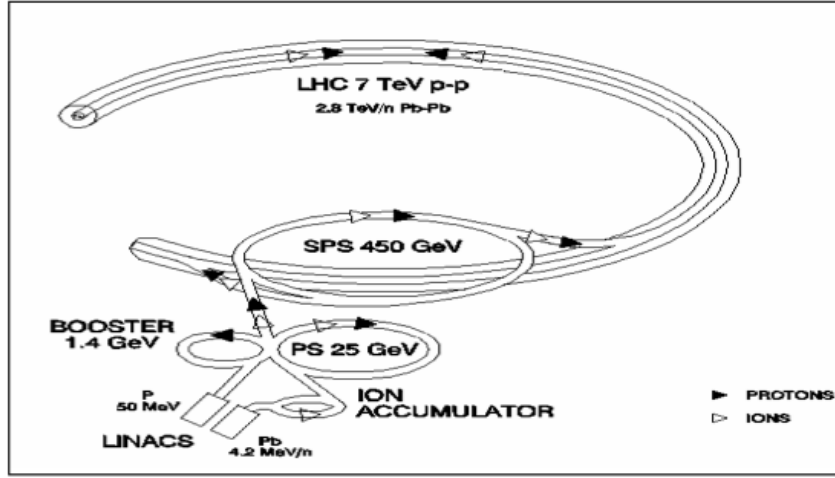
Figure 3.1: Accelerator complex: The protons are boosted as they move through the linac, proton boosters and are then fed into the main LHC ring. This figure has been taken from [18]

an electromagnetic field. They can be structured along the beam pipe of a particle accelerator. Every time protons (or any charged particles) travel through one of these cavities, they get accelerated due to the force and direction of the electromagnetic field in the cavities. A field oscillating at a fixed frequency is applied to the RF cavities. This causes the protons with slightly different energies arriving earlier or later to be accelerated or decelerated so that they stay close to the energy of the ideal particle. This gives rise to a beam with 'bunches' of protons instead of a continuous beam of protons.

The PSB sends the proton beam to the Proton Synchrotron (PS) in six bunches, delivered in two batches. Here the beam is further split and accelerated (using RF cavities with higher harmonics) to 25 GeV and is ready to be injected in the main injector, every 3.6 seconds in batches of 72 bunches spaced by 25 nanoseconds (ns).

3.1.2 Proton Synchrotron

The beam coming from PS is accepted by the Super Proton Synchrotron (SPS) [18]. The SPS acts as the main injector for the LHC ring. The SPS is the final link in the injector chain for the LHC. For a standard LHC filling cycle, input of 25 GeV proton

beam is taken from PS in two to four batches. Once the last batch is inserted, ramping up from 25 GeV to 450 GeV begins. Prior to extraction, the beam is rephased with respect to the LHC in order to inject it into the LHC at the required azimuthal position. This beam consists of 1.7×10^{11} protons per bunch. The number of bunches and bunch spacing is maintained at the values obtained from the PS input. Any beam tails are scraped out. Next input from the PS is accepted only after the 450 GeV beam from the previous cycle has been fully injected into the LHC main ring.

3.1.3   The LHC main ring

The 27 km LHC ring is designed to collide either proton-proton or heavy ion (like lead) beams. Two beams of 450 GeV each, coming from the SPS are further ramped up to an energy of 4TeV each and made to go around in the ring in clockwise and anti-clockwise directions respectively. There are nearly 2808 bunches per beam with nearly 1.15×10^{11} protons per bunch. Interactions can occur at four points through out the ring. Circular path of the beams is maintained by a magnetic field of 8 Tesla provided by nearly $\tilde{9}300$ magnets. This includes a combination of dipoles, quadrupoles, and correcting magnets. All the accelerator design statistics are taken from LHC design report volumes [19].

The number collisions per second, at the LHC, is proportional to luminosity L given by-

$$L \;=\; \frac{N_b^2 n_b f_{rev} \gamma_r}{a\pi\epsilon_n\beta^*}F \tag{3.1}$$

where N_b is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency, γ_r is the relativistic gamma factor, ϵ_n is the normalized transverse beam emittance, β^* is the beta function at the collision point and F is the geometric luminosity reduction factor due to the crossing angle at the impact point [19].

The data from proton-proton collisions is collected by two experiments: CMS and ATLAS. LHC ring is also used for heavy ion (Pb) collisions. Data from these collisions is collected by ALICE and LHCb detectors.
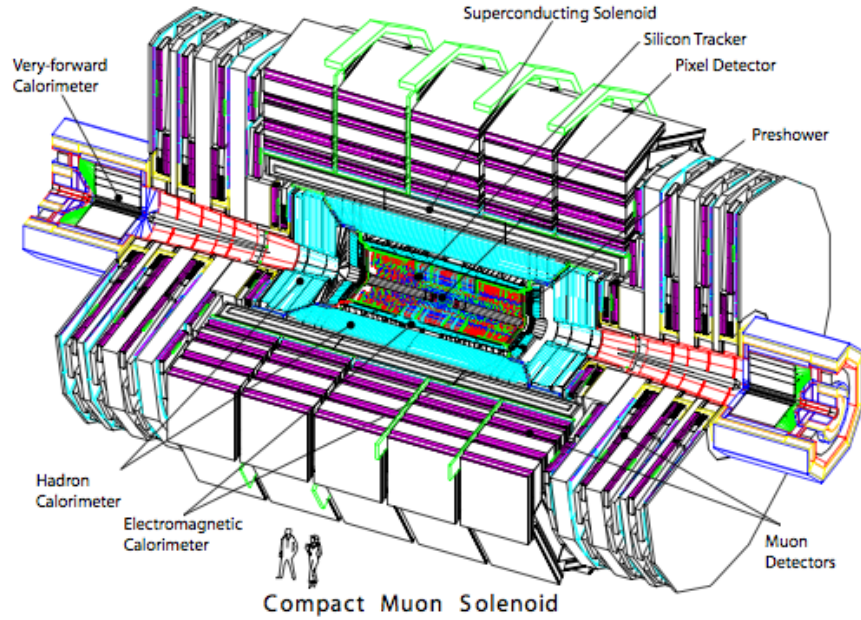
Figure 3.2: A lateral view of the CMS detector taken from [20]. It shows the various components of the detector like the tracking system, the calorimetric system and the muon detecting system.

3.2 The CMS Detector

CMS is a multipurpose detector with barrel plus end cap design and is built around a huge solenoid magnet. The magnetic field is confined by a steel yoke that forms the bulk of the 12,500 ton heavy detector. More than 2000 scientists from 155 institutions from 37 countries collaborate at CMS.

The overall layout of the CMS detector is shown in figure 3.2. Moving radially outwards from the beam is the tracker, the electromagnetic calorimeter, the hadronic calorimeter and finally the muon detectors.

The CMS technical design report offers a concise explanation of the CMS coordinate system- 'The coordinate system adopted by CMS has the origin centered at the nominal collision point inside the experiment, the y-axis pointing vertically upward, and the x-axis pointing radially inward toward the center of the LHC. Thus, the z-axis points along the beam direction towards the Jura mountains from LHC Point 5. The azimuthal angle ϕ is measured from the x-axis in the x-y plane. The polar angle θ is measured from

the z-axis. Pseudorapidity is defined as $\eta = ln(tan(\theta/2))$. Thus, the momentum and energy measured transverse to the beam direction, denoted by p_T and E_T, respectively, are computed from the x and y components. The imbalance of energy measured in the transverse plane is denoted by E_T^{miss}' [20].

3.2.1 The Magnet

The magnitude of momentum is inversely proportional to the curvature of its path. If the curvature of the path is known, we can know the momentum of the particle (equation 3.2).

$$p = qBR \tag{3.2}$$

where p is the momentum of the particle, q is the charge of the particle, B is strength of the magnetic field and R is the radius of curvature of the path of the particle.

A very strong magnet is needed to sufficiently bend the very high momentum particles like the muons. A super conducting solenoid (3 m long and 5.9 m inner diameter) sits between the hadronic calorimeter and muon chambers to provide a strong magnetic field of 4 Telsa. It is the largest magnet of its type ever constructed and allows the tracker and the calorimeters to be placed inside the coil resulting in the 'compact' nature. The enormous magnet also provides most of the structural support to the experiment [20].

3.2.2 The Tracker

The CMS tracker records the paths taken by charged particles by finding their positions at a number of key points. The tracker can reconstruct the paths of high-energy muons, electrons and hadrons as well as record tracks coming from the decay of very short-lived particles such as b quarks. The tracker needs to record particle paths accurately, yet be lightweight so as to disturb the particle as little as possible. Each measurement is accurate to 10 μm. It is also the inner most layer of the detector and hence receives the highest volume of particles. The construction materials were therefore carefully chosen

to resist radiation. The final design consists of a tracker made of silicon pixels and microstrips. The pixels are located at the very core of the detector to deal with high intensity particles. As particles travel through the tracker the pixels and microstrips produce tiny electric signals that are amplified and detected [20].

Silicon Pixels: Each layer is spilt into segments like tiny kitchen tiles. Each of these is a tiny silicon sensor (100 μm by 150 μm) about two hairs widths giving an occupancy of about 10^4 per pixel per LHC crossing. When a charged particle passes through, it provides enough energy for electrons to be ejected from the silicon atoms, creating electron-hole pairs. Each pixel uses an electric current to collect these charges on the surface as a small electric signal. An electronic silicon chip, is attached to each tile using an almost microscopic spot of solder which amplifies the signal. Trajectory of a particle is deduced from the pixels that have been hit.

Silicon Detectors are used in the intermediate region ($20 < r < 55$ cm). Minimum cell size of 10 cm by 80 μm, leads to an occupancy of roughly 23%/LHC crossing. The silicon detectors work on principle similar to the pixels: as a charged particle crosses the material it knocks electron from atoms. These electrons move due to the applied electric field, giving a very small pulse of current lasting a few nanoseconds.

3.2.3   The Electromagnetic Calorimeter (ECAL)

When a photon interacts with material it can give rise to e^+e^- pairs. These electrons can then intern loose energy via bremsstrauhlung radiation. Multiple such processes can lead to a 'shower' of electrons and photons in the material. An electron can initiate such a shower too. Calorimeter is a device that absorbs the full kinetic energy of a particle, and provides a signal that is proportional to the deposited energy. In order to build up a picture of events occurring in the LHC, CMS needs to find the energies of emerging particles. This needs to be done in a high magnetic field, at high levels of radiation and only 25 nanoseconds between collisions. This calls for very particular detector materials. The electromagnetic calorimeter at CMS is made of lead tungstate ($PbWO_4$) crystals with 61,200 of them mounted in the central barrel region and 7,324 crystals in each of the 2 endcaps. The primary reason to choose lead tungstate is

its very short radiation (X_0 =0.89 cm) and Moliere (2.2 cm) length [20]. Radiation length is the mean free length over which an electron looses all but $\frac{1}{e}$ of its energy. So, smaller the radiation length, greater is the chance for electrons and photons to dissipate their entire energy in the calorimeter. Each crystal is highly transparent and *scintillates* when electrons and photons pass through it. This means it produces light in proportion to energy of the particle. These high-density crystals produce light in fast, short, well-defined photon bursts that allow for a precise, fast and fairly compact detector. This means the calorimeter system can be very precise and very compact, fitting within the magnet coil. Lead tungstate is also relatively easy to produce from readily available raw materials. It has its disadvantages such as the yield of light depends strongly on temperature, a problem given how much heat is released by the close by electronics. To solve the cooling problem a special system maintains the temperature of 100 tons of crystal to within 0.1^oC. Another drawback is that the yield of light is low and in order to measure the scintillation light (created in the crystals from the energy of passing particles), it must be captured by the photodetectors, converted to an electrical signal and then amplified. The stronger and now digitized electrical signals are then whisked away through optic fibers. The photodetectors also need to be radiation-hard and operate within a strong magnetic field. Hence the avalanche photodiodes (APD) are used for the crystal barrel, and vacuum phototriodes (VPTs) are used for the endcaps. They are glued onto the back of each of the crystals to detect the scintillation light and convert it to an amplified electrical signal and sent for analysis.

The electromagnetic calorimeter, made up of a barrel section and two 'endcaps', forms a layer between the tracker and the hadron calorimeter. For extra spatial precision, the ECAL also contains preshower detectors that sit in front of the endcaps. These allow CMS to distinguish between single high-energy photons (often signs of exciting physics) and the less interesting close pairs of low-energy photons.

3.2.4   The Hadron Calorimeter (HCAL)

It measures the energy of hadrons, that is particles made of quarks and gluons (for example protons, neutrons, pions and kaons). Measuring these particles is important

as they can tell us if new particles such as the Higgs boson or supersymmetric particles have made an appearance. As these particles decay they may produce new particles that do not leave a record of their presence in any part of the CMS detector. To spot these, the HCAL must be 'hermetic' [20]. This means that, to the extent possible, it should capture every particle emerging from the collisions. This way if we see particles shoot out from one side of the detector, but not the other, with an imbalance in the momentum and energy, we can deduce that we are producing 'invisible' particles. Layers of the HCAL were built in a staggered fashion so that there are no gaps in direct lines. The HCAL is a sampling calorimeter, meaning it finds a particles position, energy and arrival time using alternating layers of absorber and fluorescent scintillator materials that produce a rapid light pulse when the particle passes through.

3.2.5   The Muon Chambers

Muons are expected to be produced in the decay of a number of potential new particles; for instance, one of the clearest signatures of the Higgs Boson is its decay into four muons. Hence measuring their momenta to a very high precision is one of the primary aims of the CMS detectors. Unlike electrons and positrons, muons are heavy ($M_\mu/M_e \approx$200) and can penetrate several meters of iron without interacting. Even CMS's calorimeters cannot stop them. Therefore, chambers to detect muons are placed at the very edge of the experiment where they are the only particles likely to register a signal.

The muon detecting system has three types of gaseous detectors- 1) The Drift Tubes (DTs) 2) Cathode Strip Chambers (CSCs) 3) Resistive Plate Chambers (RPCs). Figure 3.3 shows the layout of components of the muon detector [20]. DTs are placed in the barrel region ($|\eta| < 1.2$), where the neutron induced background is small. Cathode Strip Chambers are used in the endcap region which faces a large flux of muons. The magnetic field in this region is large and so is the neutron induced background. In addition to DTs and CSCs, RPCs are used in both barrel and endcap region. These coarsely spaced detectors provide a good time resolution and unambiguously identify correct bunch crossing. A particle is detected by fitting a curve to the hits among the
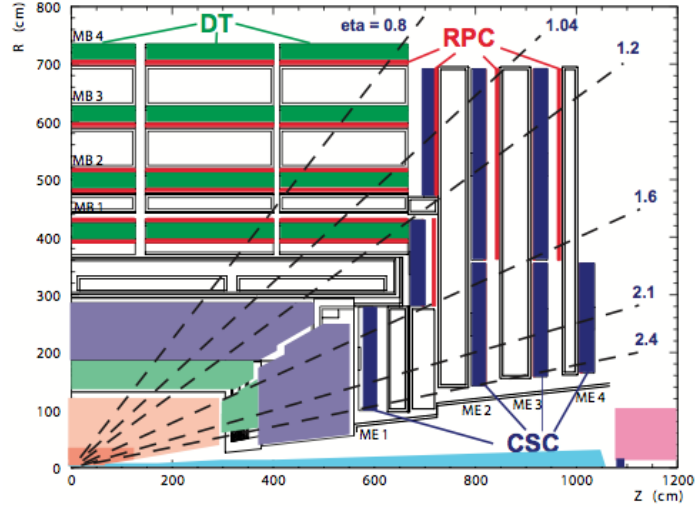
Figure 3.3: Layout of CMS muon chambers taken from [20]. It shows the DTs, CSCs and RPCs.

four muon stations, which sit outside the magnet coil and are interleaved with iron return yoke plates. By tracking its position through the multiple layers of each station, combined with tracker measurements the detectors precisely trace a particle's path. This gives a measurement of its momentum. Hence, the CMS magnet is very powerful so as to bend the paths of very high-energy muons and calculate their momenta.

3.2.6   Luminosity Measurement

Luminosity is the measure of amount of data produced during proton-proton collisions. At CMS, it is measured using signals from forward hadronic calorimeter (HF). HF covers a range of pseudorapidity $3< |\eta| < 5$. Luminosity is measured every few seconds with a statistical accuracy that is better than 1%. This is done in two ways, viz. tower occupancy measurement and total transverse energy measurement. Luminosity can be written in terms of mean number of interactions per bunch crossing as-

$$L \; = \; \frac{\mu.f}{\sigma_{pp}} \tag{3.3}$$

The total uncertainty on luminosity measurement at CMS is 2.6% [21].

Chapter 4

Lepton Identification and Triggers

4.1 Object identification

Armed with the knowledge of the detector, we can now discuss the selection of objects registered by the detector. In this section I will describe the various selection criteria applied to leptons, photons, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ used in the analysis. The selection is designed to maximize the number of good reconstructed objects while minimizing background coming from fakes.

4.1.1 Muons

Muons are reconstructed three times. Once as 'tracker muon'. As the name suggests it is reconstructed based only on the tracker information. Muon reconstructed based only on the information from the muon chambers is called 'stand-alone muon'. Finally information from both, the tracker and the muon chambers is put together to get a combined fit for the muon and such a muon is known as 'Global muon'. Other criteria the the muons must satisfy in order to be selected for the analysis are given below.

Particle Flow muon: Only particles identified as a muons by the particle-flow event reconstruction [22] are used. The particle flow algorithm uses measured quantities from the tracker, calorimeter, and muon system. The matching candidate tracks must satisfy quality requirements and spatially match with the energy deposits in the ECAL and the tracks in the muon detectors.

Impact Parameter : Muons should originate from within half a centimeter of the primary vertex in z direction and that the impact parameter d_{xy} between the track and the event vertex in the plane transverse to the beam axis needs to be small: $d_{xy} \leq 0.02$

cm. This helps reduce background contribution from muons coming from jets and other pile-up[1] vertices.

Global muon fit: The global muon reconstructed by extending the muon trajectories to include hits in the silicon tracker [23]. That is, beginning with a reconstructed standalone muon, its trajectory is extrapolated from the innermost muon station to the outer tracker surface to match the hits in the tracker. Taking this information into account, a track is fit to the muon and its quality is judged by the fit of the track. This is given in terms of χ^2 of the fit per degree of freedom. This helps to suppress background from muons from decays in flight.

Number of hits in muon chambers: The global-muon track fit needs to have at least hit from one of the muon chambers to ensure that the muon did not come from decays in flight or hadronic punch-throughs.

Number of hits in tracker : A minimum of 5 hits are required in the tracker along with the hits in the muon chambers to guarantee a good p_T measurement

Relative Isolation: Muons can also arise from semileptonic decays of heavy quarks. These muons are then a part of the jet that the heavy quark gave rise to. Typically such muons will be surrounded by other particles from the jet. To ensure that we do not selected such muons, checking the isolation of the muon is very useful. That is, we check the ratio of the energy of the muon to all the extraneous energy in the calorimeter with a cone of ΔR defined by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$. Muon is selected only if this ratio is less than 15% of the muon's energy. That is, the ratio of the sum of p_T of other particle flow candidates in the cone around the lepton to the p_T of the lepton. This requirement also helps get rid of stuff coming from vertices other than the primary vertex (i.e. pileup). The average pileup at CMS in 2012 was 19 [24].

All the requirements for muon selection [25] for this analysis are listed in 4.1. Details about muon reconstruction at CMS can be found in [26] .

---

[1]Several proton-proton interactions can happen in a single event. These are called 'pileup' interactions

Cut	Value
$|\eta|$	< 2.4
Global muon	1
PF muon	1
$\chi^2/$d.o.f.	< 10
$|dz|$ from vertex	$< 0.5cm$
$|d0|$ from vertex	$< 0.02cm$
Number of valid pixel hits	> 0
Number of tracker LayersWM*	>5
Number of valid hits in muon chamber	>0
Relative isolation within $\Delta R < 0.3$, with beta corrections for PU	< 0.15

Table 4.1: Selection criteria for muon

4.1.2 Electrons

When electrons pass through electromagnetic calorimeter they radiate bremsstrahlung photons which in turn can produce electron-positron pairs. This cascading continues until the electron has deposited all its energy in the ECAL, giving rise to a 'shower' in the ECAL. Electrons are reconstructed by matching track in tracker and to the energy deposits made by this electron showering in the calorimeter. A number of quality cuts are applied on the track and shower shape of the electron while selecting it for the analysis.

H/E : Electrons deposit almost all of its energy in the ECAL and a very small fraction of its energy in the HCAL. So the fraction of energy deposited by the electron in HCAL to ECAL is very close to zero. Yet we still have a non-zero cut value to account for high p_{T} electrons that leak into the HCAL. This cut also helps distinguish electrons from pions as the later have a much larger HCAL energy deposit fraction.

$\Delta_{\eta In}$, $\Delta_{\phi In}$: These two cuts are used to match the electron track to the energy deposit in the ECAL in η and ϕ respectively. These cuts ensure that the track is match to an electromagnetic object and not charged pions.

$1/E$ $-1/p$: Here 'E' is the energy of the electron and 'p' is the momentum. Unlike for hadronic tracks with π^0s or photons ,these to quantities are nearly equal to each other. Hence keeping the value of this cut at 5% will remove such hadronic tracks from the selection.

Relative Isolation: It is defined in the same way as for muons.

The exact selection [25] criteria are listed in table 4.2. Additional details about muon reconstruction at CMS can be found in [27].

Cut	For Barrel	For endcap
$|\eta|$	<2.4	<2.4
$\Delta_{\eta In}$	<0.007	< 0.009
$\Delta_{\phi In}$	<0.15	< 0.10
$\sigma_{i\eta i\eta}$	<0.01	< 0.03
H/E	<0.12	< 0.10
d0(vtx)	<0.02	< 0.02
dZ(vtx)	<0.1	< 0.2
$|1/E - 1/p|$	<0.05	< 0.05
Relative PF isolation	<0.15	< 0.15
conversion rejection cut	0	0
Number of expected inner hits	< 2	< 2

Table 4.2: Selection criteria for barrel and endcap electrons

See table 4.2. Furthermore,

- Veto transition region: reject electrons with η in range 1.4442–1.566.

- Electron should not be within ΔR of 0.1 of selected muon

- Corrections to the electron isolation, due to pileup events is taken into account.

4.1.3 Taus

'Taus can decay either leptonically (τ_ℓ) to electrons or muons, or hadronically (τ_h). The hadronic decays yield either a single charged track (one-prong) or three charged tracks (three-prong) with or without additional electromagnetic energy from neutral pion decays. The τ_h are reconstructed using the HPS algorithm which reconstructs the various hadronic decay modes and rejects candidates that appear to be poorly reconstructed electrons and muons. We require the visible p_T of the τ to be greater than 20 GeV and $|\eta| < 2.3$'[25]. The selection criteria for τ are given below.

- HPS Taus

- $p_T > 20\,\mathrm{GeV}$, $|\eta| < 2.3$

- ByDecayModeFinding

- AgainstElectronMVA

- AgainstMuonTight

- ByLooseCombinedIsolationDBSumPtCorr

- $\Delta R > 0.1$ from selected leptons

ByDecayModeFinding is a discriminant calculated by the HPS algothrim [28]. It takes boolean values and its value will be one when the algorithm is able to reconstruct a valid hadronic decay of the τ. AgainstElectronMVA rejects tau candidates that are already selected as electron candidates by the PF algorithm. AgainstMuonTight is a discriminator for rejecting fake τ from muons. The ByLooseCombinedIsolationDB-SumPtCorr discriminator calculates sum p_T of charged and neutral candidates ($p_T > 0.5$ GeV) with isolation cone of 0.5. Its value is one when the isolation is less than 2 GeV.

4.1.4 Photons

Photons need to have $p_T \geq 20$ GeV and $|\eta| < 2.5$. The cone size for all isolation sums is $\Delta R = 0.3$. The selection criteria [25] for photons are given below.

Cut	For Barrel	For endcap
Conversion safe electron veto	1	1
Single tower H/E	< 0.05	< 0.05
Sigmaietaieta	< 0.012	< 0.034
Rho corrected PF charged hadron isolation	< 2.6	< 2.3
Rho corrected PF neutral hadron isolation	< 3.5 + 0.04*photon pt	< 2.9 + 0.04* photon pt
Rho corrected PF photon isolation	< 1.3 + 0.005* photon pt	-

Table 4.3: Selection criteria for barrel and endcap photons

4.1.5 Jets

Jets are reconstructed using particles identified via particle flow with $|\eta| \leq 2.5$. The selection criteria [25] for jets are given below.

- pfjets, apply L1FastL2L3 corrections to MC, L1FastL2L3residual corrections to data

- $p_T > 30\,\mathrm{GeV}$, $|\eta| < 2.5$

- Neutral hadron fraction of total jet energy < 0.99

- Neutral EM fraction of total jet energy < 0.99

- Number of constituents in jet > 1

- If $|\eta| < 2.4$ additional cuts are-

    - Charged hadron fraction >0

    - Number of tracks > 0

    - Charged EM fraction < 0.99

- Apply energy corrections

4.1.6   B-tagging

For multilepton channels we additionally classify events according to presence or absence of b-tagged jets. An event is considered to contain b-jets if at least one jet passes the b-tagger which uses the CMS Combined Secondary Vertex algorithm [29]. The tagger has a tagging efficiency of 70% and a misidentification rate of 13% for the medium working point. The systematics uncertainties [30] related to the b-tagging/mis-tagging scale factors is estimated by varying the corresponding p_T dependent scale factors by $\pm 1\sigma$.

MET

- pfmet without Type I corrections

- Use official list of filters

    - CSC tight beam halo filter

    - HBHE noise filter with isolated noise rejection

- Primary vertex filter

- ECAL dead cell trigger primitive (TP) filter

- Tracking failure filter

- Bad EE Supercrystal filter

4.2 Lepton Efficiency and Scale Factors

When selection criteria are applied to leptons, all of the reconstructed lepton candidates do not pass the selection. Hence lepton selection efficiency must be taken into account while calculating acceptance of leptons for the analysis. The 'Tag and probe' method [5] is employed to calculate efficiencies of leptons. 'Tag' is a well reconstructed object passing the selection criteria for the analysis and 'probe' is an object that satisfies loose selection criteria that are a subset of the tag selection. Lepton efficiency is then calculated as the fraction of these probes that pass the strict criteria used for the analysis selection. Events where a Z boson decays to e^+e^- and Z boson decaying to $\mu^+\mu^-$ are used to measure the efficiency of the identification and isolation requirements for electrons and muons respectively. We require the tag muon to have the same selection cuts as for isolated muons used in our analysis, except that we raise the p_t requirement to $p_t > 20$ GeV. For a probe muon, we require only that it be a global muon with $p_t > 5$ GeV and $|\eta| < 2.1$. For a probe electron, we loosen the cut values on $\sigma_{I\eta I\eta}$, $\Delta\phi$, $\Delta\eta$, and H/E to the analysis selection given in the previous section. It is ensured that these cuts are sufficiently loose by looking at their distributions for candidate Z's.

The invariant mass distributions [5] of di-muons and di-electrons for these probe p_t bins are shown in figure 4.1 and figure 4.2, respectively. The Monte Carlo has been normalized to have the same number of events under the Z peak, 80 to 100 GeV/c^2, as the data. For each bin in p_t, we find the number of events on the Z peak before and after applying cuts. The mass distribution between 55 $GeV to 125 GeV$ is fit to sixth order polynomial to remove background under the Z peak. Background is removed using the fit function.

Lepton identification and isolation efficiencies are calculated separately. Lepton
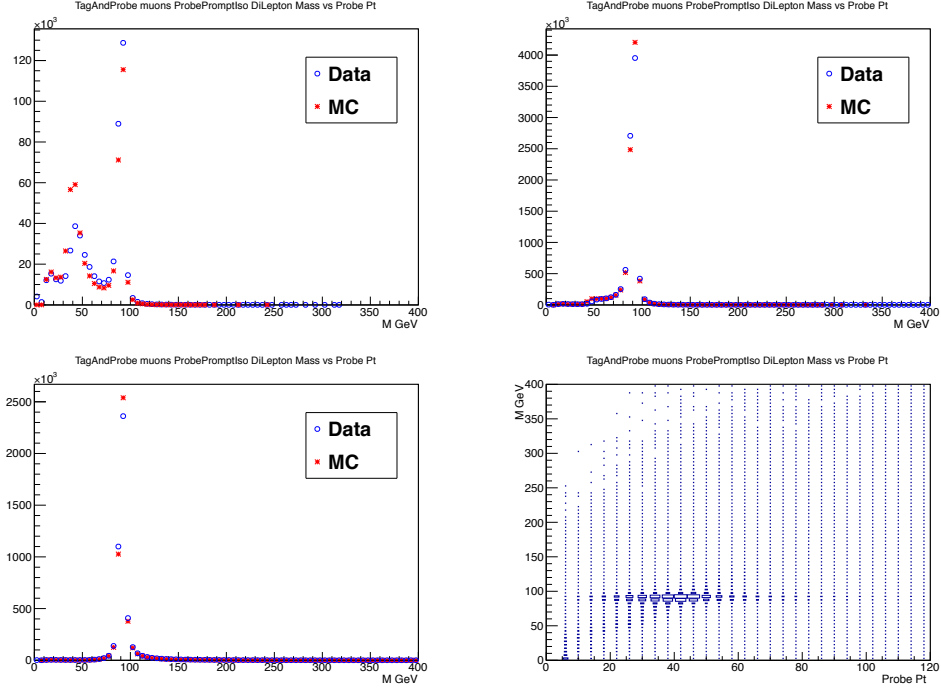
Figure 4.1: Di-muon invariant mass of tag muon and probe muon. Shown is mass for probe p_t from 12-24 GeV (top left), 24-48 GeV (top right), and > 48 GeV (bottom left). The mass versus probe p_t is shown bottom right.
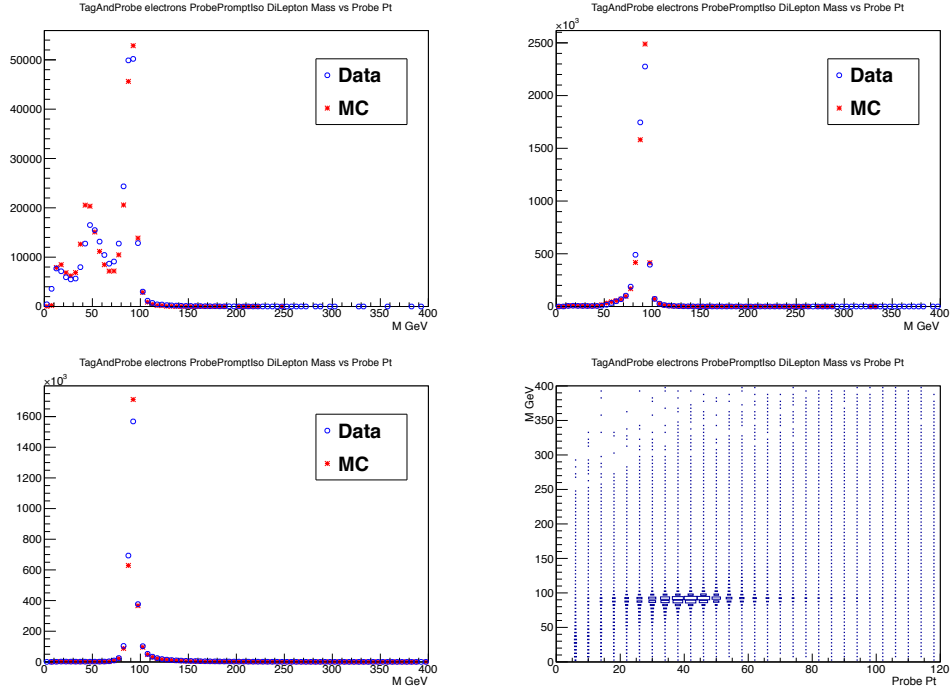
Figure 4.2: Di-electron invariant mass of tag electron and probe electron. Shown is the mass for probe p_t from 12-24 GeV (top left), 24-48 GeV (top right), and > 48 GeV (bottom left). The mass versus probe p_t is shown bottom right.

identification efficiency is determined to be the probe selection efficiency after applying identification requirements to the probes. Isolation efficiency is calculated by applying the isolation requirement to probes that pass the identification selection. The probes that satisfy identification requirements are called 'good probes'. The lepton isolation efficiency is then defined as-

$$\text{Iso Efficiency} = \frac{\text{Number of good probes that satisfy isolation requirement}}{\text{Total number of good probes}} \quad (4.1)$$

Figures 4.3 and 4.4 show the identification efficiency and corresponding data to Monte Carlo ratio as a function of probe p_t for muons and electrons, respectively. Muon identification efficiency is modeled to within a percent by the MC. For lower p_T values ($p_t < 25$ GeV/c) muon and electron isolation efficiencies measured in data are smaller than those measured in MC. Figures 4.5 and 4.6 show the isolation efficiency and corresponding data to Monte Carlo ratio as a function of probe p_t for muons and electrons, respectively.

The difference in data and MC efficiencies is taken into account by multiplying efficiencies in MC by 'scale factors'. We obtain these scale factors by fitting the ratio of data and MC isolation efficiencies to the equation below which was proposed in the same sign dilepton analysis for lepton efficiencies [31].

$$\text{Efficiency}(p_T) = \epsilon_\infty \times \text{Erf}\left(\frac{p_t - C}{\sigma}\right) + \epsilon_C \times \left(1 - \text{Erf}\left(\frac{p_t - C}{\sigma}\right)\right) \quad (4.2)$$

- ϵ_∞ = Value in plateau region at high momenta
- $C = p_T$ cut for leptons
- ϵ_C = value at $p_T = C$
- σ = Describes the rate of change in value as p_T drops.

This equation is also used for calculating electron identification efficiency scale factors. The efficiencies and scale factors depend upon event kinematics. For example isolation efficiency is inversely proportional to pileup. It also goes down with increase

in jet activity. It is important to make sure that efficiencies and scale factors appropriate to the event kinematics are used. To ensure this, efficiencies and scale factors are calculated as a function of p_T in bins of η ($|\eta| < 1.5$ (barrel) and $1.5 < |\eta| < 2.1$ (endcap)), number of jets and number of pileup vertices. Plots showing these various binning can be found in [5].The change in scale factors with the number of jets, number of vertices etc. is taken as systematic uncertainty on their estimation.

The fit parameters for isolation efficiency along with statistical and systematic uncertainty are as follows-

- $\sigma_\mu = 11.6361 \pm 0.3416(stat) \pm 2.3697(syst_{BE}) \pm 1.8662(syst_{jet}) \pm 1.7979(syst_{vert})$

- $(\epsilon\_inf)_\mu = 0.9985 \pm 0(stat) \pm 0.002(syst_{BE}) \pm 0.0009(syst_{jet}) \pm 0.0002(syst_{vert})$

- $(\epsilon\_const)_\mu = 0.9324 \pm 0.0039(stat) \pm 0.0371(syst_{BE}) \pm 0.1041(syst_{jet}) \pm 0.0166(syst_{vert})$

- $\sigma_e = 16.4017 \pm 0.5597(stat) \pm 0.5075(syst_{BE}) \pm 1.9723(syst_{jet}) \pm 2.839(syst_{vert})$

- $(\epsilon\_inf)_e = 0.9982 \pm 0.0001(stat) \pm 0.001(syst_{BE}) \pm 0.0004(syst_{jet}) \pm 0.0001(syst_{vert})$

- $(\epsilon\_const)_e = 0.9316 \pm 0.0052(stat) \pm 0.0054(syst_{BE}) \pm 0.015(syst_{jet}) \pm 0.005(sys_{vert})$
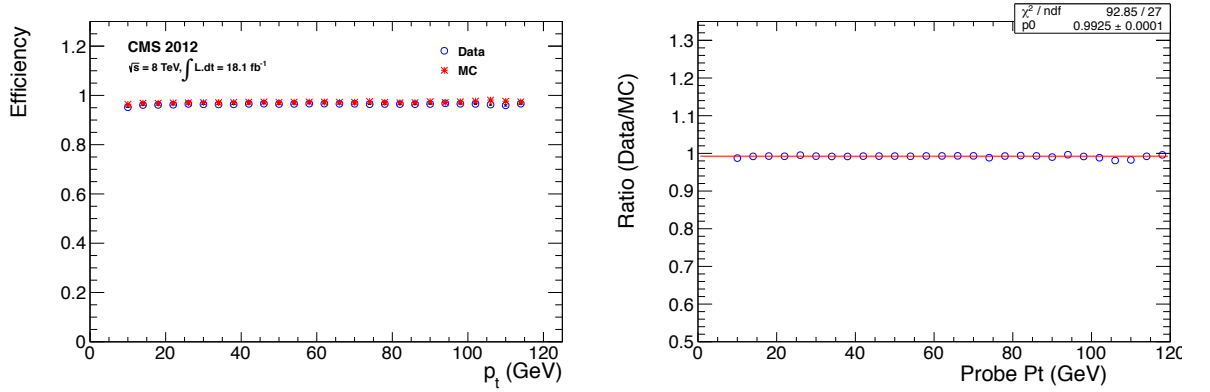
Figure 4.3: Muon identification efficiency as a function of probe p_t (left) and ratio of data and MC (right).

4.3 Trigger Efficiency

We also need to calculate trigger firing efficiency for reasons similar to the need for lepton efficiency calculation. Like any other electronic component, the triggers have a
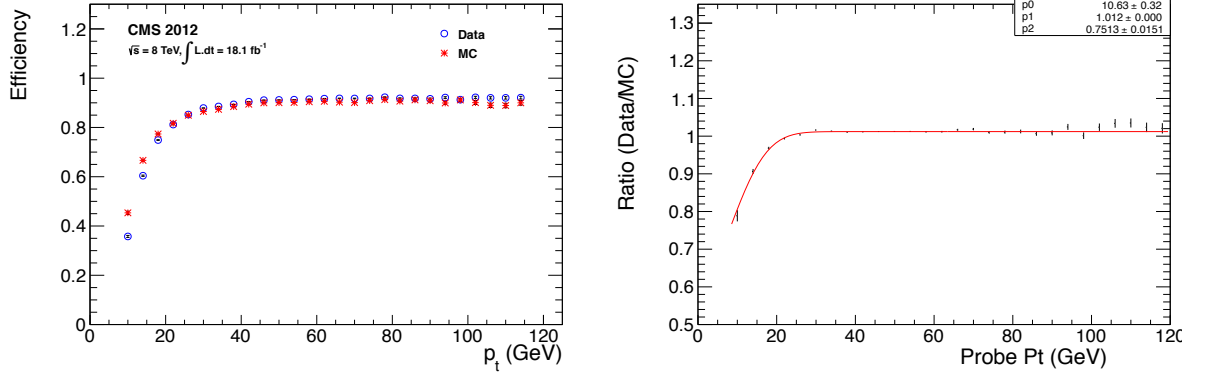
Figure 4.4: Electron identification efficiency as a function of probe p_t (left) and ratio of data and MC (right).
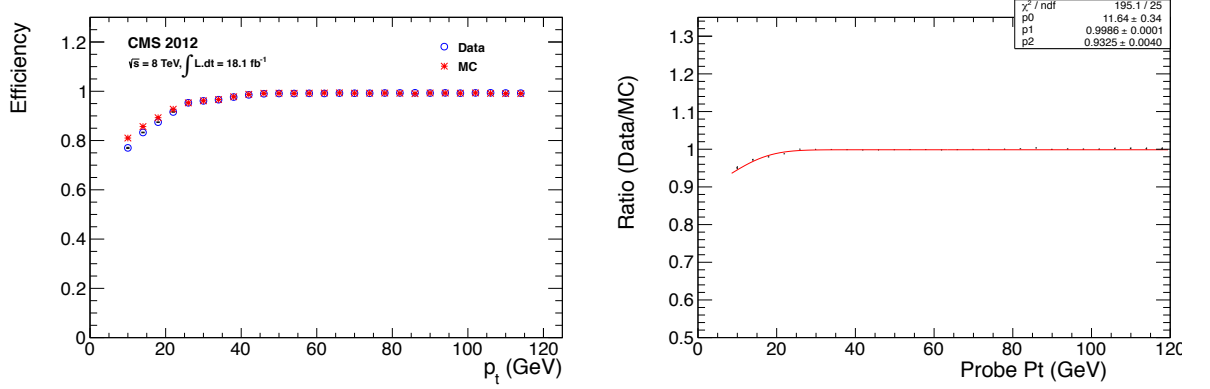
Figure 4.5: Muon isolation efficiency as a function of probe p_t (left) and ratio of data and MC (right).
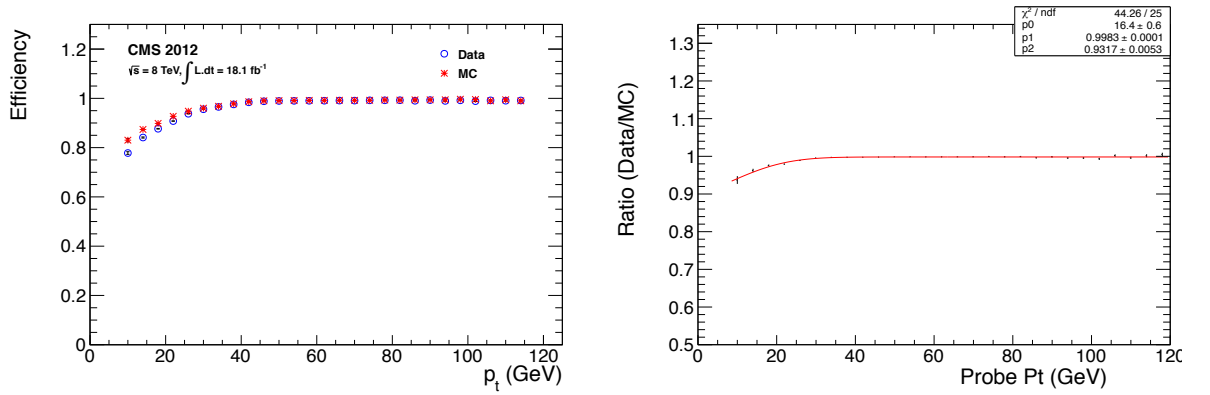
Figure 4.6: Electron isolation efficiency as a function of probe p_t (left) and ratio of data and MC (right).

firing curve. Every single event satisfying the criteria for firing the trigger need not actually fire it. Hence we need to calculate the number of times that a certain type of event actually fires the corresponding trigger. For example double electron trigger efficiency is defined as-

$$\text{Trigger Efficiency } = \frac{\begin{array}{c}\text{Number of events with two reconstructed electrons}\\ \text{\& the double electron trigger fired}\end{array}}{\text{Total number of events with two reconstructed electrons}} \quad (4.3)$$

HT triggered dataset is used for trigger efficiency calculations. As explained in an earlier internal CMS note- 'If we assume that the efficiencies of trigger i, ϵ_i, and trigger j, ϵ_j, are uncorrelated, then the efficiency for an event to satisfy both triggers , ϵ_{ij}, is given by $\epsilon_{ij} = \epsilon_i \times \epsilon_j$. We estimate the trigger efficiency by taking the ratio of the number of events that satisfied both the HT and lepton trigger to the number of events that satisfied the HT trigger. This ratio will be equal to $\epsilon_{ij}/\epsilon_i = \epsilon_j$ if the HT and lepton triggers are uncorrelated' [5].

Events with single(two) isolated lepton(s) that pass selection criteria and fire at least one HT trigger are selected. This forms the denominator. Numerator is a subset of these events with an additional requirement of the event firing a single lepton(dilepton) trigger to calculate single lepton(dilepton) trigger efficiency. For example, double muon trigger efficiency is defined as the ratio of number of events with two good isolated muons and fired at least one of the dimuon analysis trigger to the number of events that have two good isolated muons. HT or E_T^miss triggers used as 'tag' triggers, themselves have a turn-on curve. To remove the effect of the turn-on curve of the tag trigger we need to calculate the dilepton trigger efficiencies in the active region (i.e. region where trigger is completely turned on) of the former trigger. Hence we apply additional conditions like $E_\text{T}^\text{miss} > 180$ or $HT > 550$ or $HT > 300$ and $E_\text{T}^\text{miss} > 70$. This helps remove trigger correlations or biases between HT and lepton triggers, especially for the electron triggers.

Since we use a large number of dilepton triggers we calculate the efficiency for logical "OR" of all the triggers in the same category (e.g. double electron triggers). For 2012

data isolated electron trigger efficiency was 89.9%, and the muon trigger efficiency was 86.8%. Efficiency was 95% dielectron triggers, 90% for dimuon triggers and 93%± for electron-muon cross triggers. Double photon efficiency is measured to be 95%. Efficiencies are calculated using different HT triggers as tag triggers to ensure that the dilepton trigger efficiencies are independent of tag triggers used. Figure 4.7 shows efficiency for dimuon and dielectron triggers as a function of the trailing lepton p_T. The dimuon trigger efficiency is more or less flat as a function of p_T. As this not the case with dielectron and electron-muon cross triggers, p_T dependent trigger efficiencies are used in these cases. The double electron trigger efficiency for trailing lepton pt < 20 GeV is 82% and that for electron-muon cross triggers is 86 % .
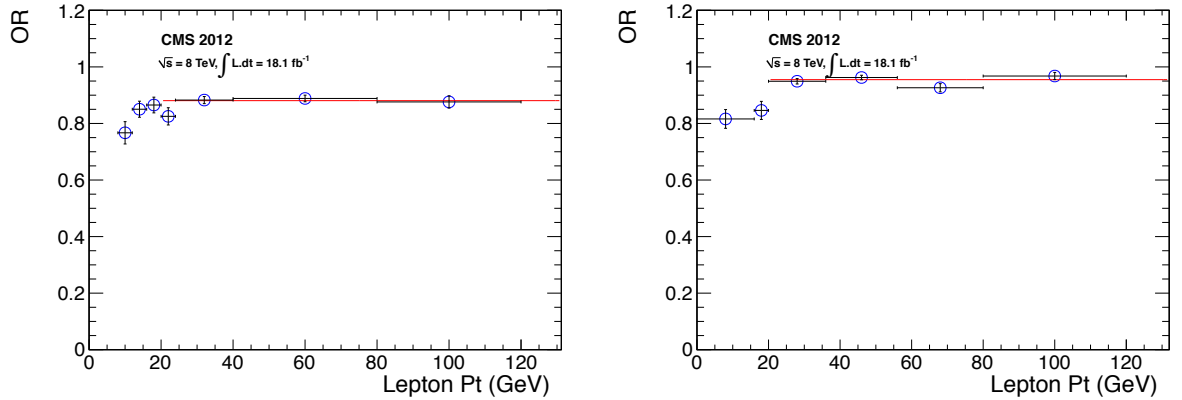
Figure 4.7: , Dimuon "OR" Efficiency (left) and Dielectron "OR" efficiency (right) by the method described in this section.

The data used for this search came from double-lepton (double-electron, double-muon, muon-electron) and double-photon triggers. The uncertainty in the correction to the simulation translates into a systematic uncertainty in the irreducible backgrounds and signal efficiencies. We also use single-lepton triggers for tag and probe studies and for the one-lepton $t\bar{t}$ control region.

Chapter 5

Analysis Overview

Any experimental High Energy (HEX) analysis has three essential components - observed events, background and signal. Observed events are taken from the collision data. Estimating SM backgrounds, that is contributions from known physics processes, is the most crucial task for an experimentalist. New physics scenarios for which the search is being carried out, is termed as the signal. Signal can be decided a priori. That is, one can start the analysis to look for a specific theory. For example a search could be carried out for detection of dark matter, supersymmetry etc. Such type of analyses are called targeted analyses. Another way of analizing is to carry out an open search. As the name suggests, here one does not set out with any particular scenario for new physics in mind. For such a search we calculate SM backgrounds for a given type of events (interactions) and then evaluate data. Then a model can be built to explain discrepancy in background and data, if any.

Elementary steps for any HEX analysis are as follows-

- Decide the search signal and understand what final states the signal will show up in.

- Select the relevant triggers and dataset.

- Choose preliminary event selection.

- Design search regions based on the signal.

- Estimate SM background and signal yield along with uncertainty in the estimation. There are two types of uncertainties- statistical and systematic. Statistical uncertainty, as the name suggests arises due to the number of events available

for the estimate and goes down as you go to higher statistics. Systematic uncertainty is due to some inherent property of the experimental apparatus and/or the method of background estimation. It may be reduced by better precision of the device etc. but cannot be removed entirely.

- Examine the data and check the consistency/discrepancy between number of events observed in data and prediction from SM (i.e. the background).

- If observation is consistent with SM expectation then use the information to put constraints on signal scenario. But if there is some discrepancy in the two, calculate the size of discrepancy and if the observation is more than 5σ away from the background prediction- "Hurray !!!! " we just might have discovered something (NOTE: Of course, a serious of stringent double-triple checks must be done to make sure the excess is real.).

This dissertation is a search for two particles Heavy Higgs and particle A in the extended Higgs sector. This is essentially a counting experiment and the search channels are decided based on the decays of the two particles. We will discuss this in detail in the upcoming sections.

5.1 Triggers

The choice of triggers depends upon the types of objects needed for the search. For example, as we want events with leptons and photons, we use data triggered on leptons and photons. We use un-prescaled [1] lepton and photon triggers with lowest possible threshold on the p_T of these objects. Since we need at least two leptons or two photons in each event, data triggered by double lepton or double photon triggers, is used. Just to give an example one of the dilepton trigger that we use is-

HLT_Mu17_Mu8 : This trigger is fired when an event has at least two muons, one with $p_\mathrm{T} > 17$ GeV and other with $p_\mathrm{T} > 8$ GeV.

---

[1]Lower the threshold on the p_T of the triggering objects more the number of signal events entering the analysis. But as the p_T thresholds go down frequency of such triggers firing increases and due to the limitations on our capacity to store this data prescale factors must be applied to the triggers. Eg. A prescale factor of 10 means the trigger will fire at every 10th event its actually meant to fire.

A complete list of triggers can be found in the Appendix C. Care is taken to ensure that there is no double counting of events coming from two different datasets.

5.2   Event Selection

Datasets used for the analysis are decided based on the above mentioned triggers. For example a dataset collected when at least one dimuon trigger fires is called dimuon dataset. We make sure that for each event coming from a particular dataset, at least one of the relevant trigger was fired and p_T of the leptons or photons in the event is above the threshold requirements of the trigger. Several proton-proton interactions can happen in a single event. These are called 'pileup' interactions. The average number of proton-proton interactions per LHC bunch crossings is 19 [24]. From this pileup we select a reconstructed vertex that has highest Σp_T^2 of tracks associated with it. Additionally this vertex needs to be within 24 cm from the center of the detector in the z direction and within 2 cm in a direction transverse from this beam line [32]. Several quality cuts that are used to identify leptons and photons required for the analysis, are mentioned in chapter 3. Classification of the events into search channels is discussed in the next section.

5.3   Search Channels based on final states

Here onwards, we are going to have two broad categories of channels- multilepton channels and diphotons plus leptons channels. Events with three or more leptons will fall in the multilepton category. Events with two or less number of leptons will be considered only if accompanied by two photons will be part of diphoton + lepton channels. This section has been taken from [25] where I first presented this information.

5.3.1   Multilepton channels

As mentioned earlier, candidate events in this category must have at least three leptons where at most one of them is an hadronic τ. The thresholds on the transverse momenta of the leptons are chosen such that the triggers used are maximally efficient in that p_T

regime. The leading muon (electron) is required to have $p_T > 20$ GeV and the next to leading muon (electron) is required to have $p_T > 10$ GeV.

We classify multilepton events into search channels on the basis of the number of leptons, lepton flavor and relative charges as well as charge and flavor combinations and other kinematic quantities like E_T^{miss}. Since the hadronic τ decays bring in additional background, search channels including these are kept separate from pure electron and muon channels.

Each event is classified in terms of the maximum number of opposite-sign and same-flavor (OSSF) dilepton pairs that can be made by using each lepton only once. For example, both $\mu^+\mu^-\mu^-$ and $\mu^+\mu^-e^-$ are OSSF1, $\mu^+\mu^+e^-$ is OSSF0 and $\mu^+\mu^-e^+e^-$ is OSSF2. Similarly, events are classified in terms of the presence of an opposite-sign and opposite-flavor (OSOF) dilepton pair.

The level of SM background varies considerably across the channels. Channels hadronic tau decays or containing OSSF ($\ell^\pm\ell^\mp$) pairs suffer from larger backgrounds than do channels without OSSF pairs. Hence all these charge combinations are considered as different channels.

5.3.2 Diphotons plus leptons channels

Candidate events in this category must have either one or two leptons and exactly two photons that pass all the identification and isolation requirements. The thresholds on the transverse momenta of leptons and photons depend upon the dataset the event in consideration comes from. For example, event with two leptons and two photons can come from either dilepton or diphoton triggered dataset. For former case, leading lepton should have $p_T > 20$ GeV and next to leading lepton should have $p_T > 10$ GeV while the two photons need to have a minimum p_T of 20 GeV each. For the latter case, leading photon must be $p_T > 40$ GeV and next to leading photon should have $p_T > 25$ GeV while the two leptons require a minimum p_T of 10 GeV. These thresholds are chosen such that triggers for the particular dataset are maximally efficient. In the case of diphoton plus one lepton event, all the data is taken from the diphoton triggered dataset.

5.4 Channel selection based on search signal

Now we have the capacity to search from the large pool of search channels mentioned in the previous section. We need to now finalize the search channels based on the type of signal we are looking for. Multilepton channels selected for the final stage depend upon the type of signal to be probed while all channels with photons are considered for signal search.

5.4.1 Search for $H \to hh$

For example, for the $H \to hh$ search, based on h branching ratios, sensitivity will come from combination of W bosons and taus in final states. Based on this, the multilepton channels that are expected to contribute the most are -

- Channels without OSSF pair (greatly reduces Drell-Yan backround).

- Channels with OSSF pair but the invariant mass of the pair is off-Z (i.e. invariant mass of the OSSF pair < 75 GeV or > 105GeV.).

- Channels with a same sign pair as they have very low SM background.

Table 5.1 gives the various decay modes of h. Various combinations of these modes then populate our search channels as shown in table 5.2.

	$h \to WW^*$	$h \to ZZ^*$	$h \to \tau\tau$	$h \to bb$	$h \to \gamma\gamma$
$h \to WW^*$	✓	✓	✓	X	✓
$h \to ZZ^*$	-	✓	✓	✓	✓
$h \to \tau\tau$	-	-	✓	X	✓
$h \to bb$	-	-	-	X	X
$h \to \gamma\gamma$	-	-	-	-	X

Table 5.1: This table shows the various decay modes of h. The combination of these decays that are considered for the analysis are marked with "✓" and those not considered for the analysis are marked with a "X".

Final states from h decays	Search Channels h decays populate
WW^*WW^*	
$\tau\tau\tau\tau$	Three or four leptons (upto one τ_h), OSSF pair off-Z
$WW^*\tau\tau$	or no OSSF pair, in bins of $E_{\text{T}}^{\text{miss}}$ and b-tag
$ZZ^*\tau\tau$	
ZZ^*ZZ^*	
ZZ^*bb	
$\gamma\gamma WW^*$	
$\gamma\gamma ZZ^*$	2photons ($M_{\gamma\gamma}$ within higgs bin)+ 1 or more lepton(upto 1 τ_h),in bins of $E_{\text{T}}^{\text{miss}}$
$\gamma\gamma\tau\tau$	

Table 5.2: This table shows the various combinations to h decay modes and the search channels they populate.

	$h \to WW^*$	$h \to ZZ^*$	$h \to \tau\tau$	$h \to \gamma\gamma$
$Z \to ll$	✓	✓	✓	✓
$Z \to qq$	X	✓	X	X
$Z \to \nu\nu$	X	✓	X	X

Table 5.3: This table show the various decay modes of h and Z boson. The combination of these decays considered for the analysis are marked with "✓" and those not considered for the analysis are marked with a "X".

h & Z boson decays	Search Channels h & Z boson decays populate
$Z(\to ll)WW^*$	
$Z(\to ll)ZZ^*$	
$Z(\to ll)\tau\tau$	1 on-shell Z$(\to ee, \mu\mu)$ + 1 or more lepton(upto 1 τ_h) in bins of $E_{\text{T}}^{\text{miss}}$ & b-tag
$qqZ(\to ll)Z^*$	
$\nu\nu Z(\to ll)Z^*$	
$\gamma\gamma ll$	2 photons($M_{\gamma\gamma}$ within higgs bin)+1 or more lepton(upto 1 τ_h) , in bins of $E_{\text{T}}^{\text{miss}}$

Table 5.4: This table shows the various combinations to h and Z boson decay modes and the search channels they populate.

Search for $A \rightarrow Zh$

For this search, there is a Z boson in the signal itself. So channels with an on-Z OSSF plus a little MET from higgs decays will greatly reduce the Drell-Yan background. Thus these channels are expected to be most sensitive to signal. Hence for this search we only use channels with alteast one OSSF pair that makes a Z mass. We also consider channels with no OSSF to consider non-resonant decays of Z boson.

Table 5.3 gives the various decay modes of h and Z boson. Various combinations of these modes then populate our search channels as shown in table 5.4.

Estimation of SM background for these search channels will be discussed in detail in the next chapter.

Chapter 6

Background Estimation

Any SM physics process that mimics the signal that you are looking for, contributes to the 'background' for the analysis. Estimating standard model backgrounds is a crucial part of any high energy physics analysis. A solid understanding of the background sources is important to understand the events in data, accounted for by SM processes. Any number of extra events can then potentially be signs for new physics. For example, if we are hunting for events with three leptons, then events with associated production of W and Z boson form the background for this signal. W boson can decay to a lepton and a neutrino (W$\to l + \nu$) and Z boson can decay to two leptons (Z $\to l^+ l^-$), thus giving you three leptons. Z+jets type of events can also contribute to background as the third lepton can be a 'fake lepton'. An object is called a fake either when it is misidentified or when it comes from an unwanted source. For this analysis, a lepton is deemed as fake if it is a misidentified lepton or if the lepton comes from a jet.

6.1 Background Reduction

Before delving into the complex background estimation procedure, we try to minimize our labor by reducing the background using relatively simple selection criteria. Following measures are taken to reduce the background [25]-

- Fakes from non-prompt leptons: We are interested in leptons that come from hard interactions and are associated with the primary vertex. Heavy flavor quark decays within jets can produce leptons. Such leptons coming from jets are surrounded by other particles belonging to the jet and tend to be father away from the primary vertex. This type of background can be suppressed by placing isolation and vertex requirement criteria on the leptons.

- E_T^{miss} requirement: As mentioned earlier, SM processes like Z+jet can also lead to multilepton events. This background can be reduced significantly by additionally requiring some missing transverse energy (E_T^{miss}). The magnitude of the vectorial sum of the momenta of all PF candidates is termed as E_T^{miss} [33].

- Low mass resonances: Cutting out events where invariant mass of OSSF leptons in less that 12 GeV helps remove backgrounds from low mass resonances like J/ψ(1S) and Υ .

- Binning in exclusive channels: We bin them according to lepton flavors, number of OSSF pairs, invariant mass of these pairs, number of b-jets etc. This helps to concentrate SM backgrounds in small number of channels that we can use as control channels and thus other channels get an improved signal to background ratio. This is in a way a 'divide an conquer' technique.

- Higgs requirement: Jets can easily fake photons. This background for the diphoton channels can be curbed by requiring the invariant mass of the photons to be within the Higgs mass window (120 GeV-130 GeV).

We consider a number background sources for this analysis: Fake leptons from jets, diboson+jets, Z+jets, asymmetric photon conversion and fake photons. We use both Monte-Carlo (MC) simulations and data-driven estimates for background estimation. Additional details about these methods are given in Ref. [5]. These methods were initially developed for Ref. [34].

6.2   Background estimation using MC

MC simulation based estimation is 'relatively' simple. We need to generate samples of particular type of events that act as background for the analysis. These samples are usually generated centrally by CMS. For example we use samples for VV+jets (V stands for vector boson), $t\bar{t}$ production and rare three-body decays. Care needs to be taken to create these samples with contributions from appropriate physics processes with all the relativistic effects taken into account. A good understanding of the detector is also

essential to get reliable estimates from MC. Also a large number of events (at least a factor of 10 more than what SM will create) need to be generated in order to have sufficient statistics in the tails of the distributions. Despite all these concerns, we do need MC for estimating 'irreducible' backgrounds. As the name suggests, no amount of cuts and selection criteria and can eliminate such backgrounds. Once we have reliable samples, the way ahead is pretty straight forward. We apply various selection criteria to choose a particular region in phase space almost completely dominated by that particular background. This region is called the 'control region'. We validate MC by comparing it with data in these control regions. Based on the level of agreement between the two we apply appropriate scale factors to the simulation and then use it for background estimation.

6.2.1   Diboson+jets

WZ+jets and ZZ+jets are examples of diboson+jet events. As seen earlier WZ+jets events can mimic three lepton signal and can also mimic four lepton signal with a fake from the jets. Similarly ZZ+jets can produce four lepton events. They can also look like three lepton events if we fail to measure one of the leptons.

We define the control region for WZ MC sample as events with exactly three leptons, one OSSF pair and the invariant mass of the pair is within the Z-mass window (75 GeV-105 GeV). Invariant mass of the remaining third lepton and E_T^{miss} is the transverse mass of W (M_T). Figure 6.2 shows the E_T^{miss} distribution of events in this control region when M_T is between 60 GeV and 120 GeV and M_T distribution of events with E_T^{miss} between 50 GeV and 100 GeV. It can be seen that the data agrees well with the MC simulation.

To validate ZZ MC, we select events with four leptons with two OSSF pairs and at least one of them has to be on-Z and $E_T^{miss} < 50$ GeV. Figure 6.3 shows the invariant mass distribution of four leptons in this control region and MC is in good agreement with data.

We have two control regions for validating $t\bar{t}$ MC sample- single lepton control region and dilepton control region. The single lepton control region requires one isolated muon with $p_t > 30$ GeV, three or more jets with $p_T > 40$ GeV, at least one b-tagged jet, and
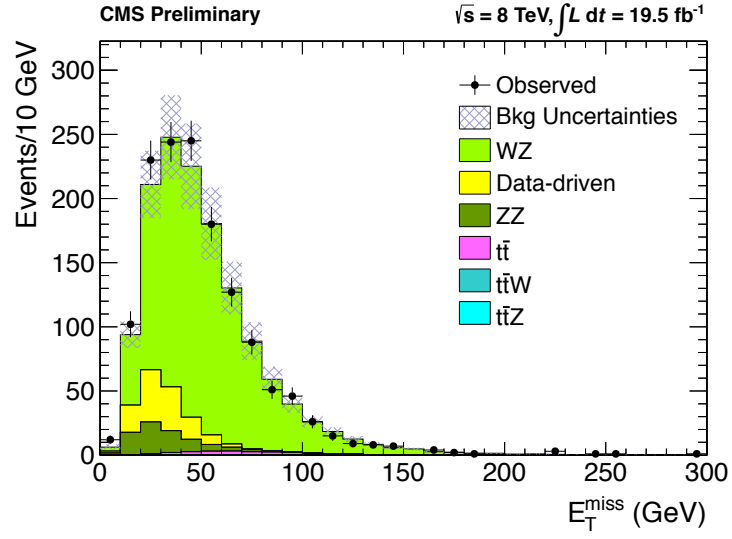
Figure 6.1: $E_\mathrm{T}^\mathrm{miss}$ distribution for WZ validation
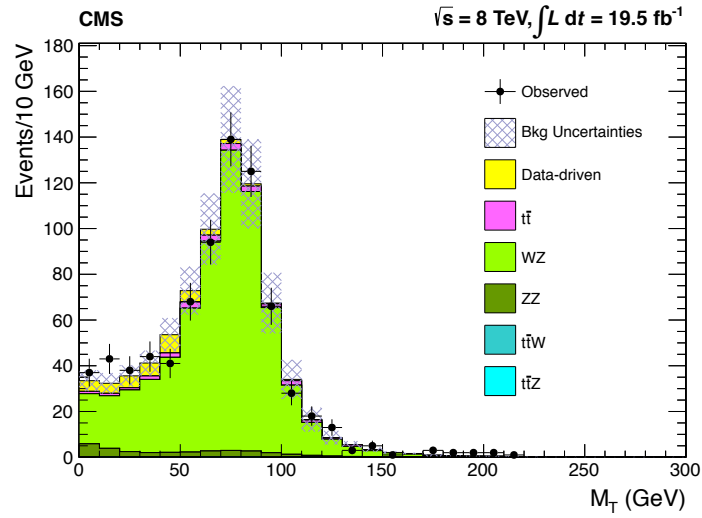
Figure 6.2: M_T distribution for WZ validation
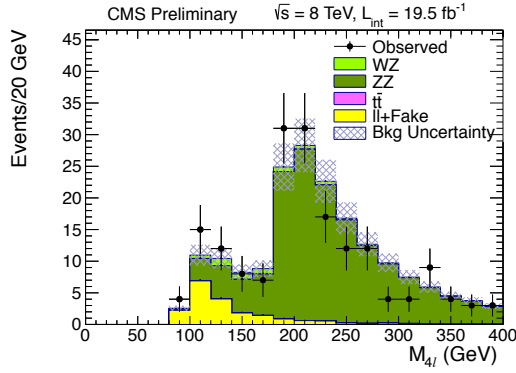
Figure 6.3: Invariant mass distribution of four-leptons with at least one OSSF pair for events with E_T^{miss} between 0-50 GeV. This is for ZZ validation.

sum of H_T, E_T^{miss} and p_t of selected leptons (S_T) > 300 GeV [25]. The $t\bar{t}$ events can fake trilepton events with an additional fake lepton. The single lepton control region is useful to understand the ratio of fakes in $t\bar{t}$ to that in data. Figure 6.4 shows the relative isolation distribution of the non-prompt muon (not the one from control region requirement). Depending upon the level of agreement in the isolation bin, we apply 50% scale factor for fakes coming from $t\bar{t}$ MC.

The dilepton control region, which requires oppositely charged isolated muon and isolated electron [25] is used to determine the overall normalization scale factor. The good agreement between data ad MC in the S_T in figure 6.5 makes the normalization factor almost negligible.

In addition to overall normalization factors, other scale factors are applied for lepton efficiency corrections, trigger efficiency reweighing, pileup corrections, jet energy scale corrections [35] etc.

6.3 Data Driven Background Estimation

MC simulations may at times fail to simulate peculiar events due to some inherent cuts (as can be seen in the case of asymmetric photon conversions). It may mis-model the tails of certain distributions like in the case of Z+jets. Such and more reasons call for
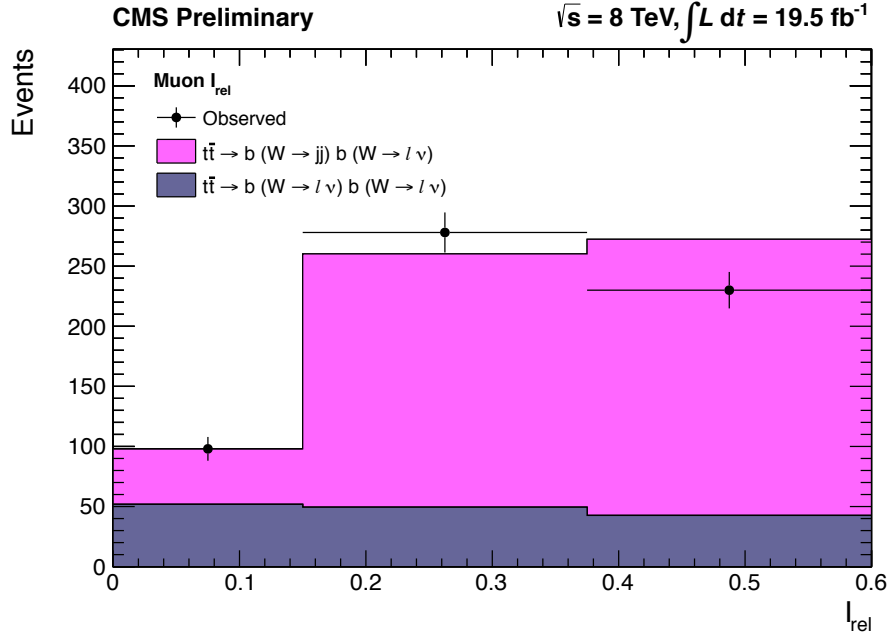
Figure 6.4: Relative isolation distribution of the non-prompt muon in the $t\bar{t}$ single lepton control region.
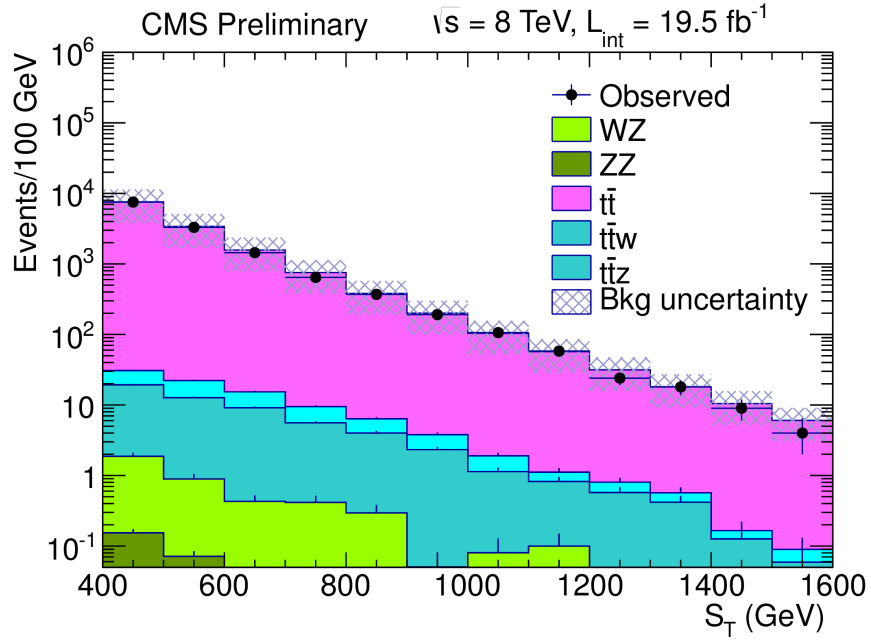
Figure 6.5: The S_T distribution for by $t\bar{t}$ control region.
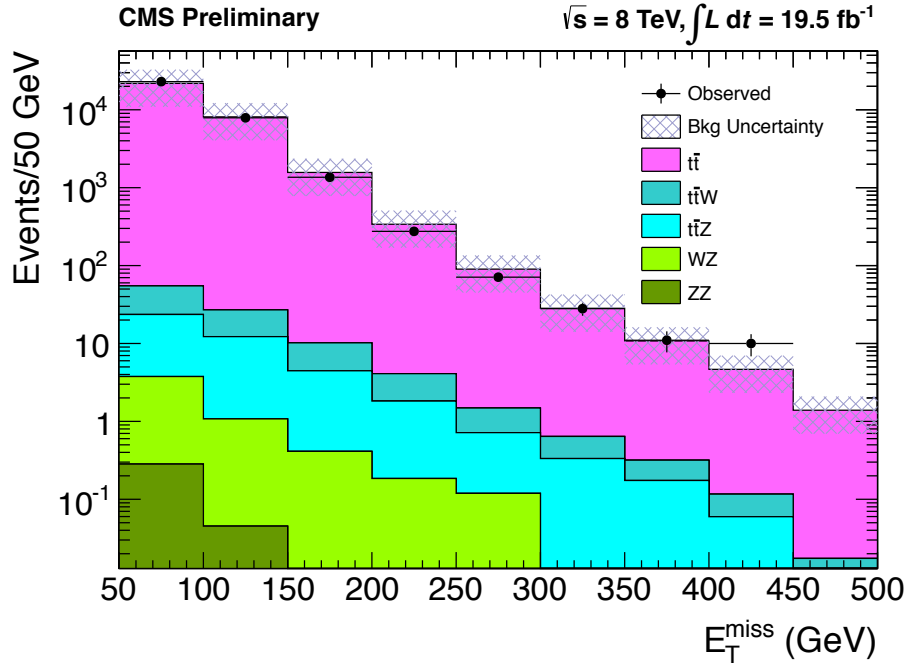
Figure 6.6: MET distribution for $t\bar{t}$ dilepton control region.
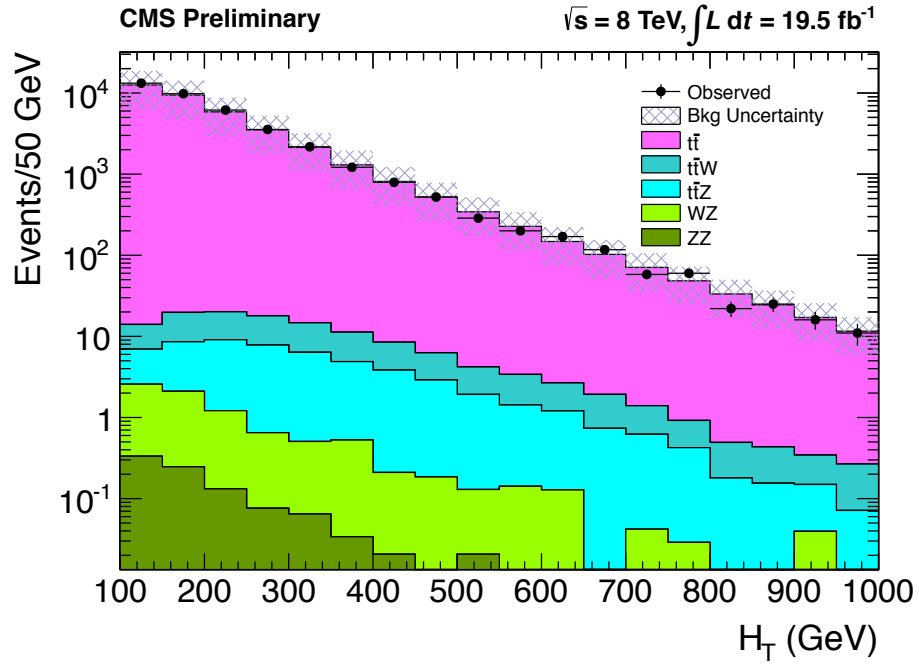
Figure 6.7: HT distribution for $t\bar{t}$ dilepton control region.

a data driven (DD) background estimates. The principle behind a DD method is to find another object (proxy) that resembles your object of interest (fake) in the detector. Usually the proxy is chosen such that it has similar dependence on event kinematics as the fake and occurs in abundance as compared to the fake object. This helps factor out several additional sources of systematic errors that otherwise creep into the estimate and reduce statistical error. There after ratio of production rates of the proxy object to fakes is calculated. This is tested in a control set and applied to 'seed' events in the signal region. By seed events we mean events closely resembling the signal minus a fake object. For example, for three lepton background estimation, dilepton+one proxy object will be the seed event. Methods based on this principle will be discussed in the following sections.

6.3.1 Fake light leptons from jets

MC simulation may not be relied on to simulate the rate for jets to produce light fake leptons as it depends on several factors. Factors such as jet shape, jet spectra, jet flavor etc. can influence this rate. Hence there is a need for DD estimate.

The DD procedure is used to predict the probability of jets to produce light lepton candidates that appear to be prompt and isolated. While establishing this method we used QCD jet triggered data to study the lepton candidates from jets. Later in 2012 the dilepton data had sufficient statistics to make robust predictions. This method is tested in controlled data regions before applying it to signal regions. In doing so we eliminate the need for simulations. Isolated tracks are used as proxy objects for fake leptons. We determine the ratio of rate for jets to produce isolated lepton candidates to the rate for jets to produce isolated tracks. Equation 6.1 is used to predict the three lepton (2l+fake) background from 2l+isolated track seed events. In the equation f_μ and f_e are conversion factors.

$$N_\mu^{\text{Iso}} = f_\mu \times N_{\text{Track}}^{\text{Iso}}, \quad N_e^{\text{Iso}} = f_e \times N_{\text{Track}}^{\text{Iso}} \tag{6.1}$$

The conversion factors depend on the relative composition of light and heavy flavored jets in the sample. This composition may change while going from control region to signal region, causing the background estimate to be erroneous. Combined Fakeable Object (CFO) method is used to meticulously understand the change in this composition and to assess any systematic changes that have occurred. Additional details about the CFO method have been described in Ref. [36]. Traditional 'fake rate' methods rely only on the proxy object, whereas CFO method uses multiple objects to understand the variation in conversion factors across different data sets. In this case isolated track acts as a proxy object and other objects like non-isolated leptons, non-isolated prompt and non-prompt tracks are used to test for changes in conversion factors. The level of confidence in determining how close the factors measured in the control region are to factors measured in signal region, determine the systematic uncertainties. Equation 6.2 expresses the conversion factors in terms of objects that can be measured in both signal and control region [25]. N_μ and N_T are measured directly from the sample where conversion factors are applied. Parameterization of ratio of isolation efficiencies is taken from control sample.

$$f_{\mu/e} = \frac{N_{\mu/e}}{N_T} \times \frac{\epsilon_{\mu/e}^{\text{Iso}}}{\epsilon_T^{\text{Iso}}} \qquad (6.2)$$

$N_{\mu/e}$: Number of non-isolated muons/electrons candidates in the dataset.

N_T : Number of non-isolated tracks.

$\epsilon_{\mu/e}^{\text{Iso}}/\epsilon_T^{\text{Iso}}$: Ratio of muon/electron and track isolation efficiencies.

Jets with decays of heavy flavored quarks are more likely to produce fake prompt lepton candidates than light flavored jets. Hence the jet flavor composition affects the conversion factors. To address this we define another variable called R_{dxy} defined as the ratio of the number of tracks with an impact parameter $|d_{xy}| > 0.02$ cm to the number of tracks with an impact parameter $|d_{xy}| < 0.02$ cm. Tracks coming from heavy flavored jets have high impact parameter and hence a sample with majority of heavy flavored jets will have high R_{dxy} value where as a sample dominated by light flavored jets will

have low R$_{dxy}$ value.

We now need to parameterize efficiency ratio in terms of R$_{dxy}$ for different data samples with varying amount of contribution from light and heavy flavored jets. Values of R$_{dxy}$ and efficiency ratio in two extreme scenarios (w.r.t. to jet flavor composition) are calculated. The two extremes are Z+jet like events and $t\bar{t}$ events. Events with two good leptons that make a Z and an additional lepton are selected to create Z+jet enriched data samples. Efficiency ratio and Rdxy value is calculated after subtracting out contribution from events with real third lepton (e.g. WZ events). High purity $t\bar{t}$ sample is obtained by selecting events with oppositely charged prompt and isolated electron and prompt and isolated muon. Efficiency ratio and R$_{dxy}$ is calculated in this region. Continuous relationship between efficiency ratio and R$_{dxy}$ value is established using analytical relationship between the two, resulting in figure 6.8. Detailed derivation of this relationship can be found in the Appendix D.
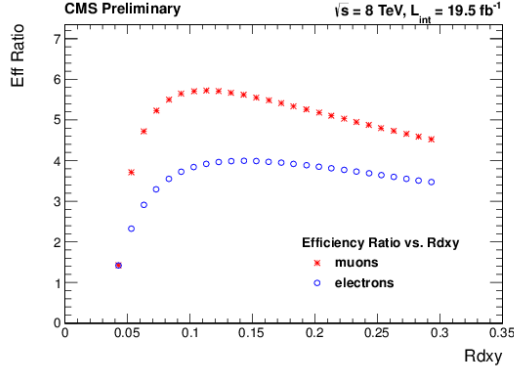
Figure 6.8: Relationship between R_{dxy} and the efficiency ratio for electrons and muons.

Events with two isolated leptons are used as seed events to estimate background with two leptons plus fakes. In signal region number of isolated tracks, non-isolated tracks, non-isolated leptons are measured and contribution from backgrounds determined using MC are subtracted out and R$_{dxy}$ is measured. Corresponding efficiency ratio is determined from figure 6.8. and conversion factor is calculated using equation 6.2. Thirty percent systematic uncertainty is placed on the estimate to account for factors like p_{T} weighted fake rate, assuming correlation between isolation efficiencies of

leptons and tracks coming from jets and how well the slope of R_{dxy} is known.

Events with two isolated leptons plus an isolated track are scaled by the conversion factor to get three lepton background estimate. The four lepton background estimate is obtained by selecting events with two isolated leptons and two isolated tracks and applying the conversion factor twice.

6.3.2   Jets faking taus

It is tricky to determine background to hadronically decaying τ leptons based on isolation, as these taus appear like jets in the detector. To estimate this background in a dataset the isolation sideband method [5] is used. Ratio of number of τ candidates in the side band region to number of τ candidates that pass the isolation cut is called conversion factor f_t. This conversion factor is sensitive to factors like jet flavor, jet spectra etc.

Consider figure 6.9. It shows isolation distributions of τ candidates-one for soft jet spectra in red and other for hard jet spectra in blue. Isolation region is marked in green and sideband region is marked in pink. It can be seen that the above mentioned ratio heavily depends on the shape of the spectra. It will be high for soft jet spectra whereas it will low for hard jet spectra. This could lead to different conversion factors in control and signal region. To account for this another variable called f_{sb} is used. It is defined as ratio of τ candidates in sideband region to total number of non-isolated τ candidates (i.e. sideband+other in fig 6.9.).

To determine the relationship between f_t and f_{sb} we divide dilepton data with $E_{\mathrm{T}}^{\mathrm{miss}}$ < 100 GeV and H_T < 200 GeV in bins of sum p_{T} of tracks ($\sum p_T$) associated with the primary vertex. Figure 6.10 shows isolation distribution of τ candidates for such different jet activity bins [37]. Values of f_t and f_{sb} are determined in each of these bins and plotted as shown in figure 6.11.

These events are separately binned according to leading jet p_{T} and similar plots are made again. The difference in the two plots with different binning gives rise to 30% systematic on the background estimation. To estimate background of the type $N\ell$+fake τ ($N\ell$ is number of light leptons), f_{sb} is measured in signal region and appropriate f_t
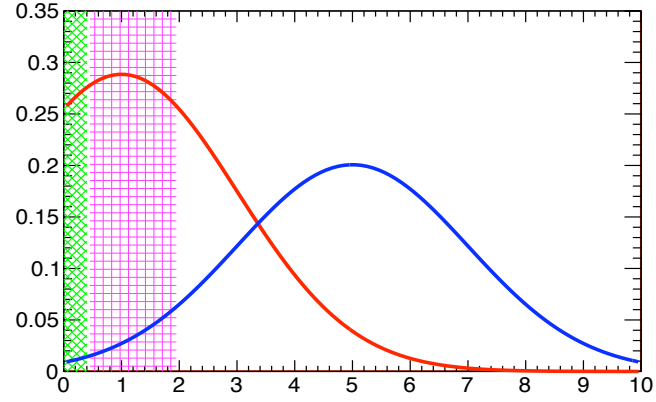
Figure 6.9: Cartoon isolation distribution showing various regions used in estimation of fake τ background.
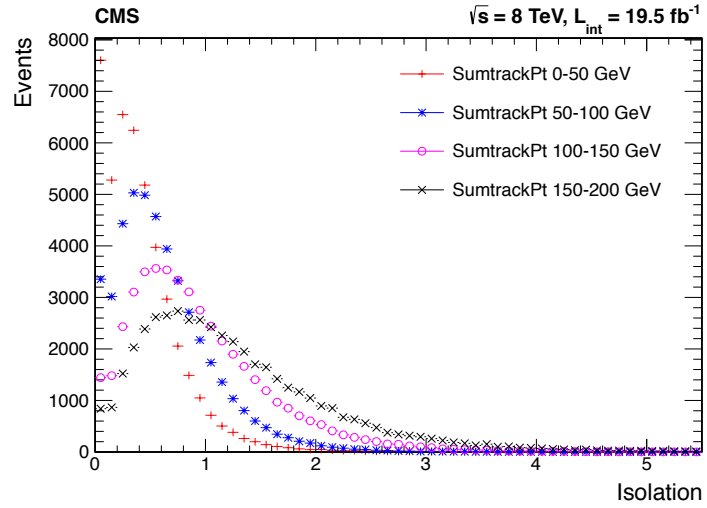
Figure 6.10: Tau fakeable object isolation distribution plot for different jet activities bins.
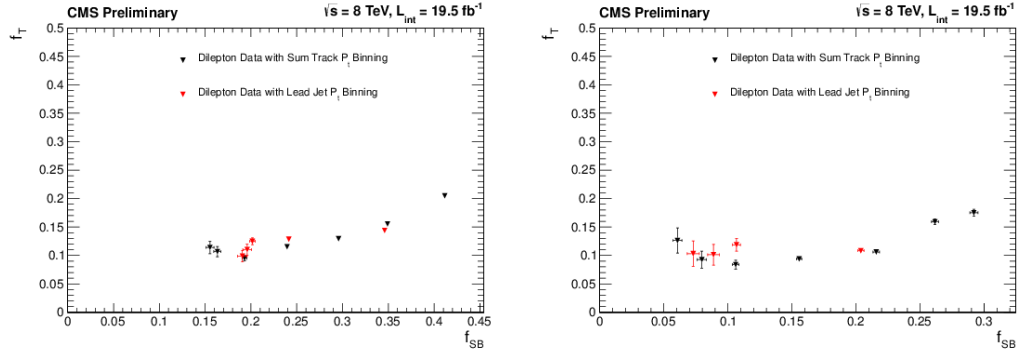
Figure 6.11: f_t versus f_{SB} for p_T of τ candidates between 20 GeV and 40 GeV (L) and p_T between 40 GeV and 60 GeV (R).

is calculated. Then the number of τ candidates in the sideband are scaled by f_t to get background prediction.

6.3.3   Asymmetric photon conversion

Internal and external conversions of photons to $\ell^+\ell^-$ pairs can give rise to backgrounds for multilepton channels. An external conversion occurs when the photon interacts with the detector material or radiates in the external magnetic field. Such type of conversions predominantly result in electron-positron pairs. In such events, ratio of getting e^+e^- pair as opposed to $\mu^+\mu^-$ is of the order of 10^4-10^5. This type of background is very well taken care of by the electron selection cuts. When the photon is virtual, it converts internally and can produce muons as often as electrons. Despite simulating such a background, MC cannot predict this type of background accurately due to an inherent cut on the lepton momentum. When a photon converts asymmetrical, that is one of the leptons takes away all the photons momentum, it is easy to loose the soft lepton due to this cut in MC simulation.

Consider a Z boson decaying to two leptons and one of the lepton radiates. If this radiated photon converts asymmetrically to two leptons and we loose the soft lepton, then the remaining lepton along with the previous two leptons (from Z) become background for three lepton channels. This background is missed when we look for

candidate Z events as the initial leptons do not build a Z (because one of them has radiated). To account for this type of background, we assume that the rate for virtual photons to undergo asymmetric conversion is the same as that for on-shell photons. We define conversion factor as the ratio of probability of photon to give good lepton candidate to the photon to be on-shell and pass our selection criteria. This conversion factor is measured in low $E_{\mathrm{T}}^{\mathrm{miss}}$, low H_T region as our experience from 2010 and 2011 shows that it is devoid of new physics. The conversion factor can be written as-

$$Conversion factor(C) = \frac{\ell^+\ell^-\ell^\pm \ events \ that \ make \ a \ Z}{\ell^+\ell^-\gamma \ events \ that \ are \ on-Z} \tag{6.3}$$

Figure 6.12 shows invariant mass of three muons where no two OSSF muons make a Z and figure 6.13 shows similar plot for 3-body $\ell^+\ell^- e^\pm$. The conversion factor for muon is measure to be $0.7\% \pm 0.1\%$ and that for electrons is $2.1\% \pm 0.3\%$.
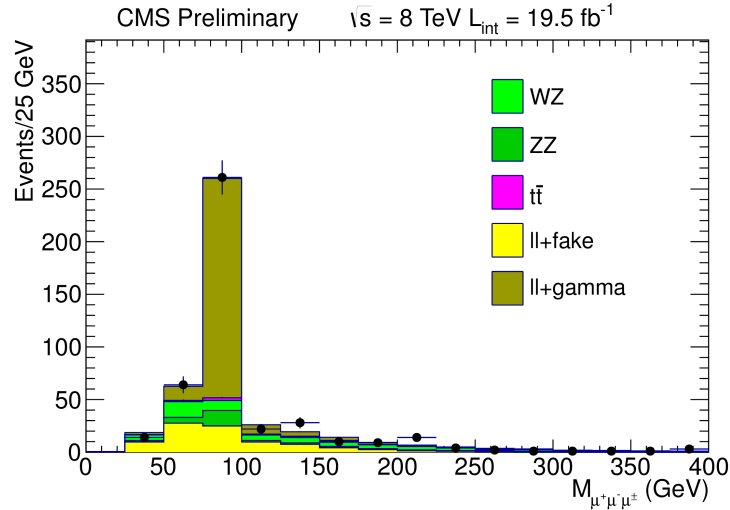
Figure 6.12: $M(\mu^+\mu^-\mu^\pm)$ where both $M(\mu^+\mu^-)$ are either < 75 GeV or > 105 GeV.

6.3.4 Backgrounds for diphoton channels

We use channels with photons when the SM Higgs from Heavy Higgs (one of the SM Higgs) or A decays to two photons. So essentially when we plot the invariant mass of
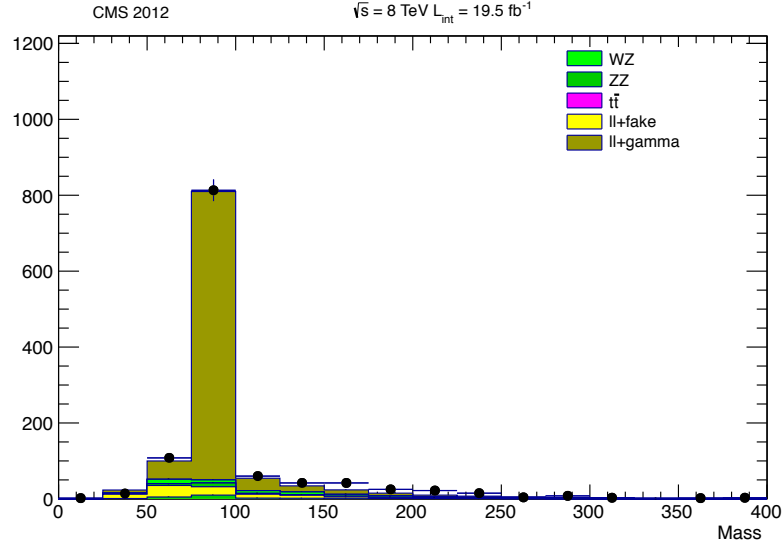
Figure 6.13: $M(\ell^+\ell^- e^\pm)$ where $M(\ell^+\ell^-)$ is either < 75 GeV or > 105 GeV. Note that most external conversions to electrons have already been removed in the electron identification requirements.

photons in these channels, the invariant mass corresponding to Higgs range (120 GeV-130 GeV) is the signal region, everything else is background. We then use the sidebands around the Higgs mass window to fit for background shapes and use the fit function to estimate background contribution in the signal region (i.e. Higgs mass window).

6.3.5   Two taus and two photons (Tau2Gam2)

Channels with up to two hadronic taus and two photons have enough statistics to do fit for background shape in the sideband region. We use fit function of the type $A \times e^{-x}$ for the fits and remove the Higgs bin (120 GeV-130 GeV) from the fitting procedure. Figure 6.14 shows background fits in different E_T^{miss} regions. We use the same function in all the E_T^{miss} bins and will use the deviation in the fits in the various E_T^{miss} bins as systematic. We use a conservative 50% systematic uncertainty.
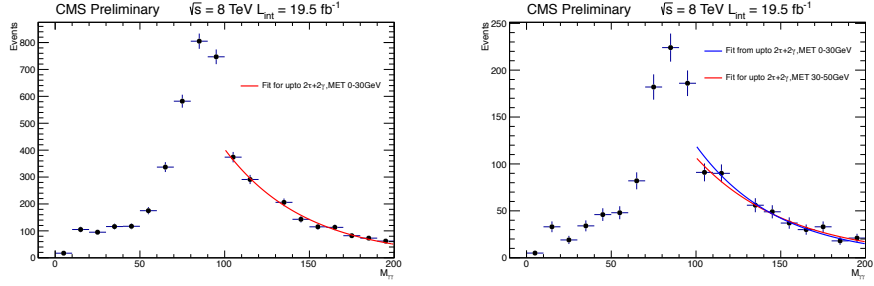
Figure 6.14: Invariant distribution of photons in Tau2Gam2 channels for E_T^{miss} less than 30 GeV (L) and E_T^{miss} between 30-50 GeV (R) . Right hand side plot shows the fit function taken from less than 30 GeV E_T^{miss} (blue curve) applies well to 30-50 GeV E_T^{miss} bin.Original fit for 30-50 E_T^{miss} region is given in red curve.
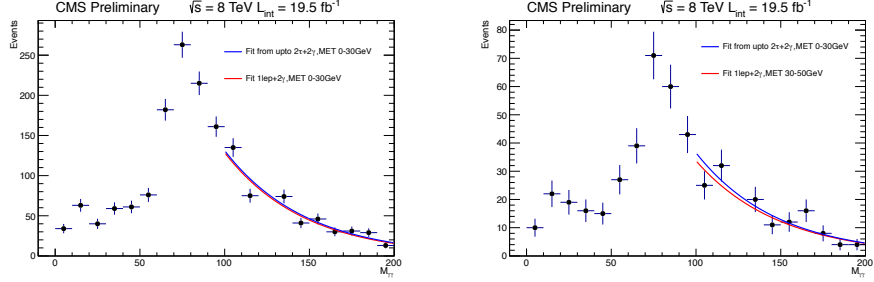
Figure 6.15: Invariant mass distribution of photons in L1Gam2 channel ,0-30 GeV E_T^{miss} (L) and 30-50 GeV E_T^{miss} bin (R). The blue fit is the one taken from Tau2Gam2 channels and red fit corresponds to independent fit for that particular bin.

6.3.6 One lepton and two photons (L1Gam2)

The invariant mass distribution of the photons in these channels is very similar to the distribution from theTau2Gam2 channels. Figure 6.15 shows the invariant mass distribution for L1Gam2 channel with E_T^{miss} between 0-30 GeV and E_T^{miss} between 30-50 GeV. The blue fit is taken from Tau2Gam2 E_T^{miss} 0-30 GeV channels and the red fit corresponds to an independent fit for that particular channel. The good agreement between the red and the blue fit allows us to use fits from Tau2Gam2 channels. The fit is normalized for each signal channel based on the ratio of number events in the sideband region of the channel of interest to the number of side band events from the original fit. A conservative systematic of 50% is placed on the background estimate.

6.3.7 One lepton, one tau and two photons (L1Tau1Gam2) and Two leptons and two photon channels (L2Gam2)

Channels of this type are statistically limited to do background fits independently. In this region too, we use the same fit function that we use for Tau2Gam2 $E_{\mathrm{T}}^{\mathrm{miss}}$ 0-30 GeV channel along with appropriate normalization. We apply a conservative 50% systematic to account for the fact that we cannot fit these regions for backgrounds individually.

6.3.8 Two leptons and two photon channels

Like the rest of the diphoton channels the invariant mass distribution of photons is fit with the function from Tau2Gam2 $E_{\mathrm{T}}^{\mathrm{miss}}$ 0-30 GeV channel along

This section for background estimation of diphoton channels has be take from CMS internal note [4], where it was first presented.

6.4 Systematic Uncertainties

This section describes the different sources of systematic uncertainties that were considered for this analysis. The systematic uncertainties are listed in table 6.1.

Luminosity systematic: As the name suggests, this systematic arises due to the uncertainty in the the measurement of luminosity. It is taken into account while calculating signal yields and background predictions from MC simulations. It does not affect the DD background estimates. It is taken into consideration during statistical tests and is not a part of quoted background uncertainties.

Lepton Efficiency: The deviation in lepton efficiencies (identification and isolation) calculated in MC and data is taken into account by scaling MC by appropriate scale factors (see chapter 4). Systematic uncertainty is determined based on their variation of parameters used to fit for the ratio between data and MC w.r.t. to barrel and endcap probe, number of jets and number of pileup vertices.

Fakes: Uncertainty in the measurement of fake rate is taken as a systematic for estimation of backgrounds coming from fakes.

Source of Uncertainty	Uncertainty
Luminosity	4.4%
PDF	14%
$E_\mathrm{T}^\mathrm{miss}$ Res ($E_\mathrm{T}^\mathrm{miss}$): 0–50 GeV, 50–100 GeV, > 100 GeV	(−3%, +4%, +4%)
Jet Energy Scale $W^\pm Z$	0.5% (WZ)
B-Tagging scale factor	0.1% (WZ), 6% ($t\bar{t}$)
Muon ID/Isolation at 30 GeV	0.2%
Electron ID/Isolation at 30 GeV	0.6%
$t\bar{t}$ xsec	10%
$t\bar{t}$ fake rate contribution	50%
WZ xsec	15%
ZZ xsec	15%

Table 6.1: The systematic uncertainties associated with this analysis.

Jet Energy Scale: Jets in the event are fluctuated up and down within the uncertainty of jet p_T. This cause the jet spectra to move around and largest deviation in each bin from the nominal spectra is taken as systematic. This variation is also propagated to calculation of $E_\mathrm{T}^\mathrm{miss}$.

PDF uncertainty: These are taken into account for signal. They have been from official LHC cross section site [38].

Chapter 7

Results

Analysis procedure in general and selection of analysis channels, optimum for our signal scenario, was reviewed in chapter 5. Background estimation methods were discussed in detail along with the sources of systematic error in chapter 6. We are now ready to discuss the results of this analysis. We will first take a look at observed events and background predictions for all the different categories of channels starting with highest number of lepton channels, that is, four leptons and all the way up to one lepton+one τ_h+ diphoton channels. These results are presented without any signal consideration. Results for 2HDM scenarios will be presented in the later sections of this chapter.

7.1 Results

Table 7.1 and table 7.2 show the number of observed events (data) along with the expected SM background in the respective channel. The SM background is the total background from DD background predictions and MC simulations. For easy presentation, we have combined the electron and muon exclusive channels. However channels with and without at least one τ_h are kept separate. Results for dilepton+diphoton channels are presented in table 7.3 and table 7.4 shows the results for channels with one lepton+diphoton. Tables 7.5 and 7.6 give results for channels with up to two τ_h+diphotons and one lepton+one τ_h+diphotons, respectively.

The abbreviations[25] used in the tables can be explained as follows-

- L4/L3/L2/L1 : Total number of leptons

- DY0/DY1/DY2 : Number of opposite sign same flavor (OSSF) pairs

- ZV : Invariant mass of OSSF pair is outside the Z mass window

- SS : Contains same sign pair

The results for some multilepton channels are presented in figures 7.1 through 7.7. The various backgrounds contributing to the channel are stacked on top of each other along with data on the top.
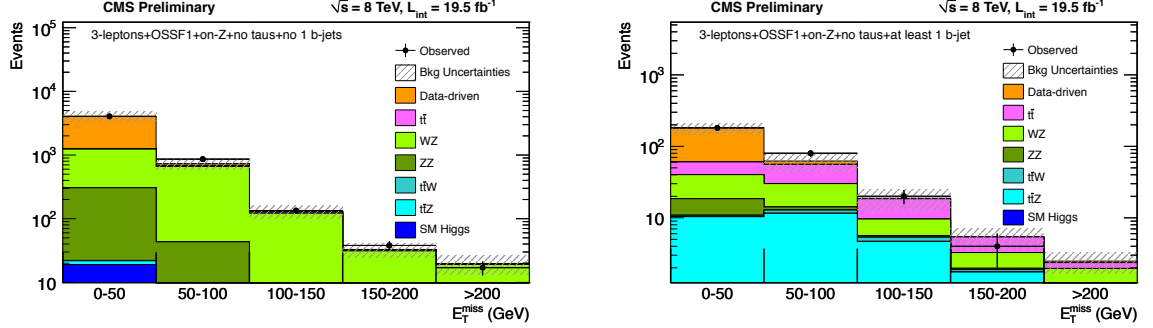
Figure 7.1: E_T^{miss} distribution for 3-lepton + OSSF1 + on-Z + Tau0 + b0 (L) and 3-lepton + OSSF1 + on-Z + Tau0 + b1 (R).
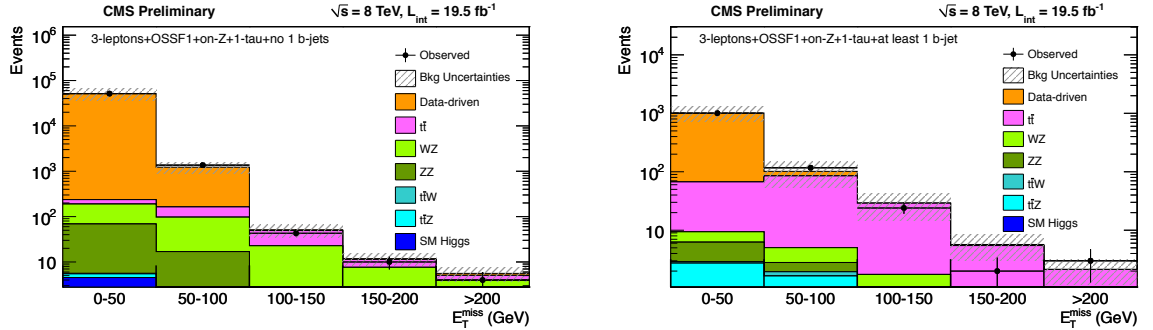
Figure 7.2: E_T^{miss} distribution for 3-lepton + OSSF1 + on-Z + Tau1 + b0 (L) and 3-lepton + OSSF1 + on-Z + Tau1 + b1 (R).

7.2 Statistical method

This section describes method for setting statistical limits for various models with multiple channels. Details about the statistical procedure are taken from Sourabh Dube's thesis [9] and CMS internal note for this analysis [4]. CL_s is a useful technique to calculate combined limits for independent search channels. Each channel is treated as statistically independent bin and then combined exclusion levels are calculated [9]. Consider a test statistic Q that distinguishes between background-like and signal-like

4 Lepton Results

Selection	on- or off-Z	$E_{\mathrm{T}}^{\mathrm{miss}}$	$N_\tau=0$, $N_{b-Jet}=0$ obs	expect	$N_\tau=1$, $N_{b-Jet}=0$ obs	expect	$N_\tau=0$, $N_{b-Jet}\geq1$ obs	expect	$N_\tau=1$, $N_{b-Jet}\geq1$ obs	expect
OSSF0	NA	$(100,\infty)$	0	0.07 ± 0.07	0	0.18 ± 0.09	0	0.05 ± 0.05	0	0.16 ± 0.1
OSSF0	NA	$(50,100)$	0	0.07 ± 0.06	2	0.8 ± 0.35	0	0 ± 0.03	0	0.43 ± 0.22
OSSF0	NA	$(30,50)$	0	0.001 ± 0.02	0	0.47 ± 0.24	0	0 ± 0.02	0	0.11 ± 0.09
OSSF0	NA	$(0,30)$	0	0.007 ± 0.02	1	0.4 ± 0.16	0	0.001 ± 0.02	0	0.02 ± 0.04
OSSF1	off-Z	$(100,\infty)$	0	0.07 ± 0.04	4	1 ± 0.33	0	0.14 ± 0.09	0	0.46 ± 0.2
OSSF1	on-Z	$(100,\infty)$	2	0.6 ± 0.2	2	3.4 ± 0.8	1	0.8 ± 0.41	0	0.6 ± 0.26
OSSF1	off-Z	$(50,100)$	0	0.21 ± 0.09	5	2.6 ± 0.6	0	0.21 ± 0.11	1	0.7 ± 0.32
OSSF1	on-Z	$(50,100)$	2	1.3 ± 0.39	10	12 ± 2.5	2	0.6 ± 0.33	1	0.8 ± 0.3
OSSF1	off-Z	$(30,50)$	1	0.16 ± 0.07	4	2.4 ± 0.5	0	0.06 ± 0.06	0	0.47 ± 0.21
OSSF1	on-Z	$(30,50)$	3	1.2 ± 0.35	11	14 ± 3.1	0	0.22 ± 0.12	0	0.8 ± 0.31
OSSF1	off-Z	$(0,30)$	1	0.38 ± 0.18	11	5.7 ± 1.7	0	0.05 ± 0.04	0	0.5 ± 0.26
OSSF1	on-Z	$(0,30)$	1	2 ± 0.5	32	30 ± 9.2	1	0.19 ± 0.11	3	1.3 ± 0.42
OSSF2	TwoZ	$(100,\infty)$	0	0.02 ± 0.15	–	–	0	0.21 ± 0.13	–	–
OSSF2	OneZ	$(100,\infty)$	1	0.43 ± 0.15	–	–	0	0.5 ± 0.29	–	–
OSSF2	off-Z	$(100,\infty)$	0	0.06 ± 0.03	–	–	0	0.09 ± 0.07	–	–
OSSF2	TwoZ	$(50,100)$	3	2.8 ± 2.1	–	–	0	0.33 ± 0.11	–	–
OSSF2	OneZ	$(50,100)$	1	2 ± 0.7	–	–	1	0.5 ± 0.28	–	–
OSSF2	off-Z	$(50,100)$	2	0.2 ± 0.14	–	–	0	0.12 ± 0.1	–	–
OSSF2	TwoZ	$(30,50)$	19	22 ± 9	–	–	2	0.7 ± 0.24	–	–
OSSF2	OneZ	$(30,50)$	6	6.5 ± 2.4	–	–	0	0.32 ± 0.12	–	–
OSSF2	off-Z	$(30,50)$	3	1.4 ± 0.6	–	–	1	0.15 ± 0.08	–	–
OSSF2	TwoZ	$(0,30)$	118	109 ± 28	–	–	3	2 ± 0.5	–	–
OSSF2	OneZ	$(0,30)$	24	29 ± 7.6	–	–	1	0.6 ± 0.17	–	–
OSSF2	off-Z	$(0,30)$	5	7.8 ± 2.3	–	–	0	0.18 ± 0.06	–	–

Table 7.1: Observed yields for four lepton events from 19.5 fb^{-1} recorded in 2012. The channels are broken down by the number of and mass of any opposite-sign, same-flavor pairs (whether on or off Z), whether there are any b-jets present and the MET. Expected yields are the sum of simulation and data-driven estimates of backgrounds in each channel. The channels are exclusive.

3 Lepton Results

Selection		N(τ)=0, NbJet=0		N(τ)=1, NbJet=0		N(τ)=0, NbJet≥1		N(τ)=1, NbJet≥1	
OSSF0(SS)	(200,∞)	1	1.3 ± 0.6	2	1.4 ± 0.5	0	0.7 ± 0.36	0	0.7 ± 0.5
OSSF0(SS)	(150,200)	2	2.1 ± 0.9	0	3 ± 1.1	1	2.1 ± 1	0	1.5 ± 0.6
OSSF0(SS)	(100,150)	9	10 ± 4.9	4	9.9 ± 3	12	12 ± 5.9	4	6.3 ± 2.8
OSSF0(SS)	(50,100)	34	37 ± 15	54	66 ± 14	32	32 ± 15	24	22 ± 10
OSSF0(SS)	(0,50)	47	46 ± 11	196	221 ± 51	28	24 ± 11	21	31 ± 9.6
OSSF0	(200,∞)	–	–	5	4.8 ± 2.4	–	–	6	5.9 ± 3.1
OSSF0	(150,200)	–	–	12	18 ± 9.1	–	–	21	20 ± 10
OSSF0	(100,150)	–	–	94	96 ± 47	–	–	91	121 ± 61
OSSF0	(50,100)	–	–	351	329 ± 173	–	–	300	322 ± 163
OSSF0	(0to50)	–	–	682	767 ± 207	–	–	230	232 ± 118
OSSF1 below-Z	(200,∞)	2	2.5 ± 0.9	4	2.1 ± 1	1	1.9 ± 0.7	2	2.4 ± 1.2
OSSF1 on-Z	(200,∞)	17	19 ± 6.3	4	5.6 ± 1.9	1	2.4 ± 0.8	3	2.1 ± 0.9
OSSF1 below-Z	(150,200)	7	4.4 ± 1.7	11	9.3 ± 4.6	3	4.7 ± 2.1	7	11 ± 5.9
OSSF1 on-Z	(150,200)	38	32 ± 8.5	10	11 ± 3.6	4	5.4 ± 1.7	2	5.7 ± 2.7
OSSF1 below-Z	(100,150)	21	26 ± 9.9	45	56 ± 27	20	23 ± 11	56	66 ± 33
OSSF1 on-Z	(100,150)	134	129 ± 29	43	51 ± 16	20	18 ± 6	24	28 ± 14
OSSF1 below-Z	(50,100)	157	129 ± 30	383	380 ± 104	58	60 ± 28	166	173 ± 87
OSSF1 on-Z	(50,100)	862	732 ± 141	1363	1227 ± 323	80	62 ± 17	117	101 ± 48
OSSF1 below-Z	(0,50)	543	559 ± 93	10186	9171 ± 2714	40	52 ± 14	257	256 ± 79
OSSF1 on-Z	(0,50)	4041	4061 ± 691	51361	51369 ± 15340	181	181 ± 28	1003	1012 ± 286

Table 7.2: Observed yields for three lepton events from 19.5 fb^{-1} recorded in 2012. The channels are broken down the number of and mass of any opposite-sign, same-flavor pairs (whether on or off Z), whether there are any b-jets present and the MET. Expected yields are the sum of simulation and data-driven estimates of backgrounds in each channel. The channels are exclusive.

2 Lepton and 2 Photon Results				
Selection		$E_{\mathrm{T}}^{\mathrm{miss}}$	obs	expect
OSSF1	off-Z	$(50,\infty)$	0	0.19 ± 0.25
OSSF1	on-Z	$(50,\infty)$	0	0.1 ± 0.17
OSSF1	off-Z	$(30,50)$	1	0.17 ± 0.25
OSSF1	on-Z	$(30,50)$	1	0.33 ± 0.28
OSSF1	off-Z	$(0,30)$	1	1.2 ± 0.74
OSSF1	on-Z	$(0,30)$	0	1.01 ± 0.55
OSSF0	NA	$(0,\infty)$	0	0 ± 0.17

Table 7.3: Observed yields for three lepton events from 19.5 fb^{-1} recorded in 2012. The channels are broken down the number of and mass of any opposite-sign, same-flavor pairs (whether on or off Z), and the MET. Only channels where invariant mass of photons lies in the higgs mass window (120-130 GeV)are considered. Expected yields are the sum of simulation and data-driven estimates of backgrounds in each channel. The channels are exclusive.

1 Lepton and 2 Photon Results		
$E_{\mathrm{T}}^{\mathrm{miss}}$	obs	expect
$(50,\infty)$	9	14.3 ± 7.15
$(30,50)$	31	22.1 ± 11.05
$(0,30)$	74	79.1 ± 39.55

Table 7.4: Observed yields for one lepton and diphoton events from 19.5 fb^{-1} data recorded in 2012. The channels are broken down in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$. There are no hadronic taus in these channels. Only channels where the invariant mass of photons lies in the higgs mass window (120-130 GeV) are considered. Expected yields are data-driven estimates of backgrounds in each channel. The channels are exclusive.

Upto 2 Hadronic Tau and 2 Photon Results		
$E_{\mathrm{T}}^{\mathrm{miss}}$	obs	expect
$(50,\infty)$	16	11.39 ± 5.6
$(30,50)$	73	69.1 ± 34.6
$(0,30)$	235	241.4 ± 120.7

Table 7.5: Observed yields for up to two hadronic taus plus diphoton events from 19.5 fb^{-1} data recorded in 2012. The channels are broken down in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$. Only channels where invariant mass of photons lies in the higgs mass window (120-130 GeV) are considered. Expected yields are data-driven estimates of backgrounds in each channel. The channels are exclusive.

1 Lepton, 1 Hadronic Tau and 2 Photon Results		
$E_{\mathrm{T}}^{\mathrm{miss}}$	obs	expect
$(50,\infty)$	0	0.16 ± 0.66
$(30,50)$	0	0.5 ± 0.57
$(0,30)$	0	0.76 ± 0.6

Table 7.6: Observed yields for one lepton, one hadronic tau plus diphoton events from 19.5 fb^{-1} recorded in 2012. The channels are broken down in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$. Only channels where invariant mass of photons lies in the higgs mass window (120-130 GeV) are considered. Expected yields are data-driven estimates of backgrounds in each channel. The channels are exclusive.
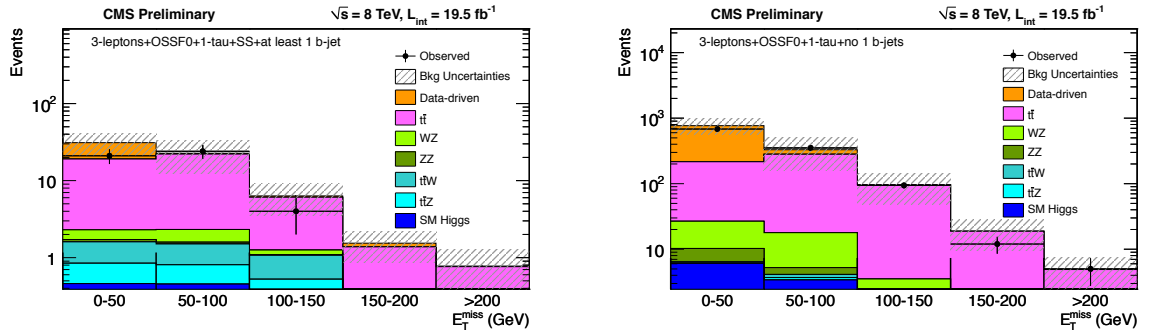
Figure 7.3: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for 3-lepton + OSSF0 + Tau1+ b1 +SS 1 (L) and 3-lepton + OSSF0 + Tau1 + b0 (R).
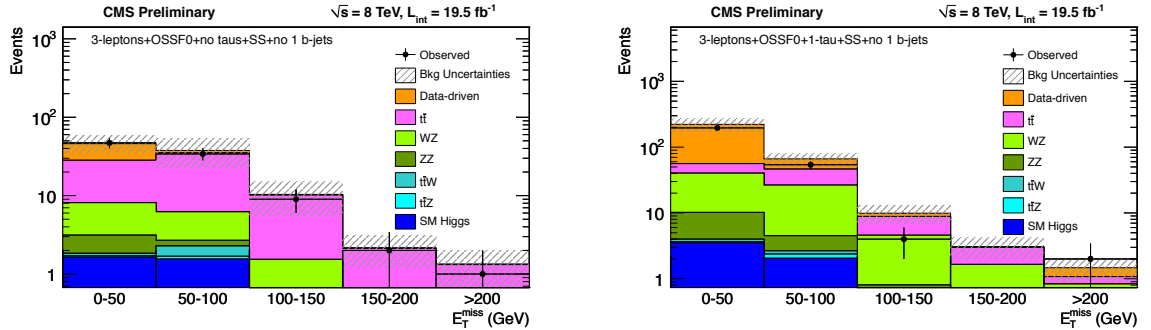
Figure 7.4: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for 3-lepton + OSSF0 + Tau0 + b0 + SS1 (L) and 3-lepton + OSSF0 + Tau1 + b0 + SS1 (R).

Figure 7.5: E_T^{miss} distribution for 4-lepton + OSSF1 + off-Z + Tau0 + b0 (L) and 4-lepton + OSSF1 + off-Z + Tau1 + b0 (R).
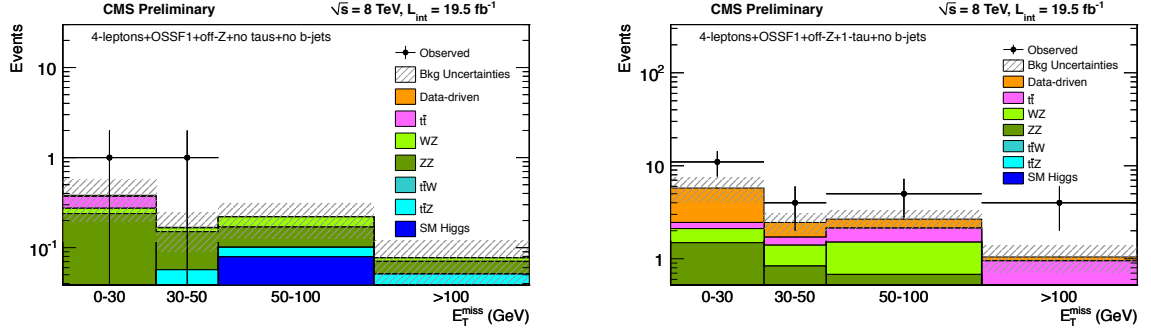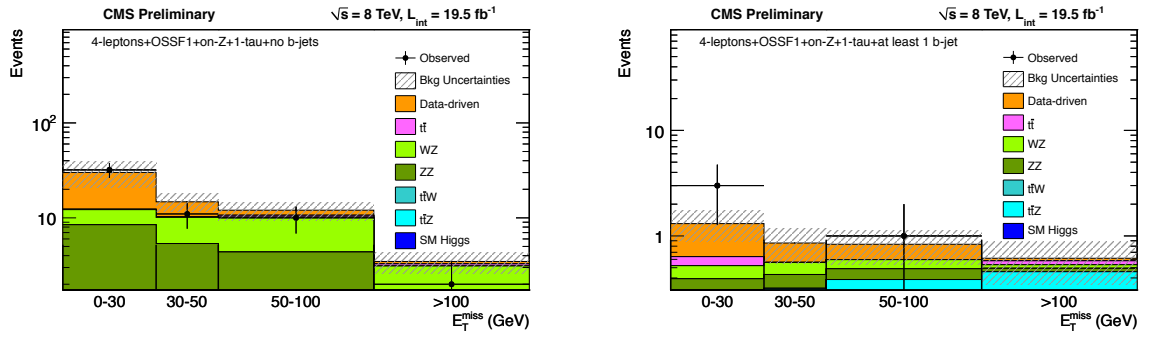

Figure 7.6: E_T^{miss} distribution for 4-lepton + OSSF1 + on-Z + Tau1 + b0 (L) and 4-lepton + OSSF1 + on-Z + Tau1 + b1 (R).
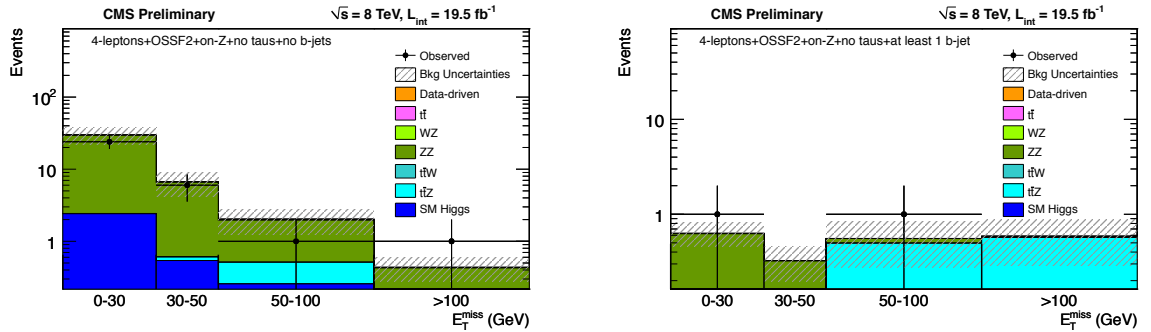

Figure 7.7: E_T^{miss} distribution for 4-lepton + OSSF2 + on-Z + Tau0 + b0 (L) and 4-lepton + OSSF2 + on-Z + Tau0 + b1 (R).

scenario. Given signal 's', background 'b' and observation 'n', the likelihood ratio for 'k' channels is given by-

$$\prod_{i=1}^{k} \frac{e^{-(s_i+b_i)}(s_i+b_i)^{n_i}}{e^{-b_i}b_i^{n_i}} \tag{7.1}$$

Then the probability for Q to be less than or equal to the observed data gives the confidence level (C.L.) for exclusion of signal plus background (s+b) hypothesis [39]. The confidence level is thus given by-

$$CL_{s+b} = P_{s+b}(Q \leq Q_{obs}) = \sum_{Q(n_i') \leq Q(n_i)} \prod_{i=1}^{k} \frac{e^{-(s_i+b_i)}(s_i+b_i)^{n_i}}{e^{-b_i}b_i^{n_i}} \tag{7.2}$$

where $Q(n_i)$ is the test statistic for observed data and $Q(n_i')$ stands for the various possible n_i' that can give test statistic less than or equal to the observed one.

Here we place 95% C.L. upper limits on signal production cross section and other model parameters.

The LandS [40] tool is used compute these 95% C.L LHC-type CL_s limits. We would like to thank the developers of LandS for this helpful tool. LandS takes input in the form of datacards and yield results in the form of observed as well as the expected limit. The datacards contain observations, background estimation and signal expectation of each selected channel along with the statistical and systematic uncertainties on both background and signal. Nuisance parameters are assigned to describe the various uncertainties and proper care is taken to take into account the different correlations between systematic uncertainties across channels in signal and background.

Given the large number of channels studied in the analysis, every channel will not contribute to every particular signal scenario at each point in the signal parameter space. Hence to speed up the limit calculation process, channels with no signal expectation are excluded from the datacard. The channel are then ordered by decreasing r-value (low to high) where the r-value is defined as $\frac{\sigma(95\%excluded)}{\sigma_{theory}}$. Channels that make top 90% of the signal are then selected for the datacard. We have tested this '90% ' cut off to be robust by changing it between 85% to 95% and this resulted in very negligible

change in the limits.

Roughly 30-35 channels get used for each point. Practically only top 15 channels have significant contribution to the limit. Other channels act as safety net so the we do not miss any important channels and the limit calculation process remains stable.

7.3 σ*BR 95% Upper Limit

Since the data agrees reasonably well with the SM background with uncertainties our next step is to probe the sensitivity of the analysis to signal production. We will first present our search sensitivity in the most model independent manner. Using the method for calculating limits in previous section, we place 95% upper limit [41, 42, 43] on the production cross section of Heavy Higgs and particle A. That is, we place limit on the process of gluon-gluon fusion production of Heavy Higgs that decays to two SM-like Higgs ($gg \rightarrow H \rightarrow hh$) and gluon-gluon fusion production of particle A that decays to Z boson and a SM-like Higgs ($gg \rightarrow A \rightarrow Zh$).

As I have described in Ref. [25], figure 7.8 shows 95% upper limit on σ * Br for $H \rightarrow hh$ search along with 1-sigma and 2-sigma bands on the expected contour using only the multilepton channels. Figure 7.9 shows 95% C.L. upper limits on σ * Br for the same signal using the multilepton and the diphoton channels. It can be seen that the photon channels add up to half a picobarn (pb) to the sensitivity and a couple of pb in the observed limit. They also serve as an important cross check in case we see any excess in multilepton channels.

The discrepancy between expected and observed 95% C.L. limits in Figure 7.8 and Figure 7.9 is driven by the $3(e/\mu)+\tau_h$ off-Z channels with no b-tags. The three relevant entries in table 7.2 refer to an observed (expected) yield of 11 (5.7 ± 1.7), 4 (2.4 ± 0.5), 5 (2.6 ± 0.6) events for the three different E_T^{miss} bins. The probability of a statistical fluctuation in a single measurement with the expected yield of 10 ± 1.9 events to result in 20 or more observed events is only about 1.5%. However, once trial factors are incorporated to account for this search looking simultaneously at 40 individual channels, the probability of encountering as significant a fluctuation in the sum of the three E_T^{miss}

bins as observed in data is about 46%.Since this is the only $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution where the total number of events is this far off from the expected number of events, we are within the expectations for statistical fluctuations [25].

Next we probe the sensitivity for $A \rightarrow Zh$. Figure 7.10 shows 95% C.L. upper limits for $A \rightarrow Zh$ with 1-sigma and 2-sigma bands on expected limit with only multilepton channels. Figure 7.11 shows the same signal probed with multilepton plus diphoton channels. Here too, the diphoton channels help little in expected sensitivity but substantially in the observed limit.

7.4   Exclusion Limits

Previous section gives model independent limits on for search for $H \rightarrow hh$ and $A \rightarrow Zh$. In this section we present these limits on Two Higgs Doublet Models (2HDMs) scenario. Type I and Type II 2HDMs [12] scenarios are considered in this dissertation. For each type the 2HDM parameter space is generated by varying for $\cos(\beta - \alpha)$ and $\tan\beta$. As a reminder, α is the mixing angle between H and h and $\tan\beta$ decides the relative contribution of each Higgs doublet to electroweak symmetry breaking. 95% C.L. are calculated at each point in the parameter space. The cross section and branching ratios for SM-like Higgs change as we move through the parameter space and this has been taken into account while calculating limits. Hence these are model dependent limits.

It can be seen from the previous section that the model independent limits are not very sensitive to the mass of Heavy Higgs or particle A. Hence we choose mass of 300 GeV as candidate mass to show the next model dependent limits.

As previously described in Ref. [25]- figures 7.12 and 7.13 show expected limits for Heavy Higgs of mass 300 GeV for Type I 2HDM and Type II 2HDM respectively along with the associated σ * Br contours(theory) for the respective model. These theory contour plots are similar to one from theory paper by Nathaniel Craig et al. [13], difference being $\tan\beta$ plotted on the y-axis as opposed to β in the orginal paper.

Figures 7.14 and 7.15 are similar plots for the particle A. At each point in the $\tan\beta$ vs $\cos(\beta - \alpha)$ plane , appropriate branching ratio for SM-like Higgs, depending upon

the values of $\tan \beta$ vs $\cos(\beta - \alpha)$, are used. Similar contours, but with the assumption that the SM-like Higgs branching ratios are the same as the SM Higgs at 126 GeV are in the Appendix B. We also look for Heavy Higgs and A signal together, for the case where these two particles are mass degenerate. Figure 7.16 shows exclusion in β and $\cos(\beta - \alpha)$ plane for the combined signal for Type I and Type II 2HDMs. At each point in this plane, appropriate branching ratios for SM-like Higgs are used.

In figures 7.12 to 7.15 the regions below the curve is excluded. In case of a closed curve, the region enclosed by the curve is excluded [25]. When compared to the theory contour plot alongside each exclusion limit, it can be seen that the exclusion shapes closely follow the cross section contours. As $\cos(\beta - \alpha)$ approaches zero, we reach the 'alignment limit' , where the branching ratio for Heavy Higgs decaying to two SM-like Higgs and A decaying to Z and SM-like Higgs, vanishes [13]. Hence, as we approach $\cos(\beta - \alpha) = 0$, the exclusion power rapidly declines, creating a valley shaped region around this point.

7.5   Conclusion

In this dissertation, I have presented the search for an extension of Higgs sector in $H \rightarrow hh$ and $A \rightarrow Zh$ channels using multilepton and diphotons final states with 19.5 fb^{-1} data recorded in 2012. The SM backgrounds were calculated using MC simulations and data-driven techniques. Overall the observed data agrees well with SM expectations with some local deviations. These null results are used to place 95% C.L. upper limits on production cross sections of particles in the extended Higgs sector. A 95% C.L. upper limit at 7pb is placed on the production cross section gluon-gluon fusion of Heavy Higgs decaying to two SM-like Higgs bosons. A similar limit is placed on the gluon-gluon fusion of particle A decaying to a Z and SM-like Higgs, at 2pb. Results are interpreted in the context of Type I and Type II Two Higgs Doublet Models. Exclusion regions are presented in tanβ vs $\cos(\beta - \alpha)$ plane.
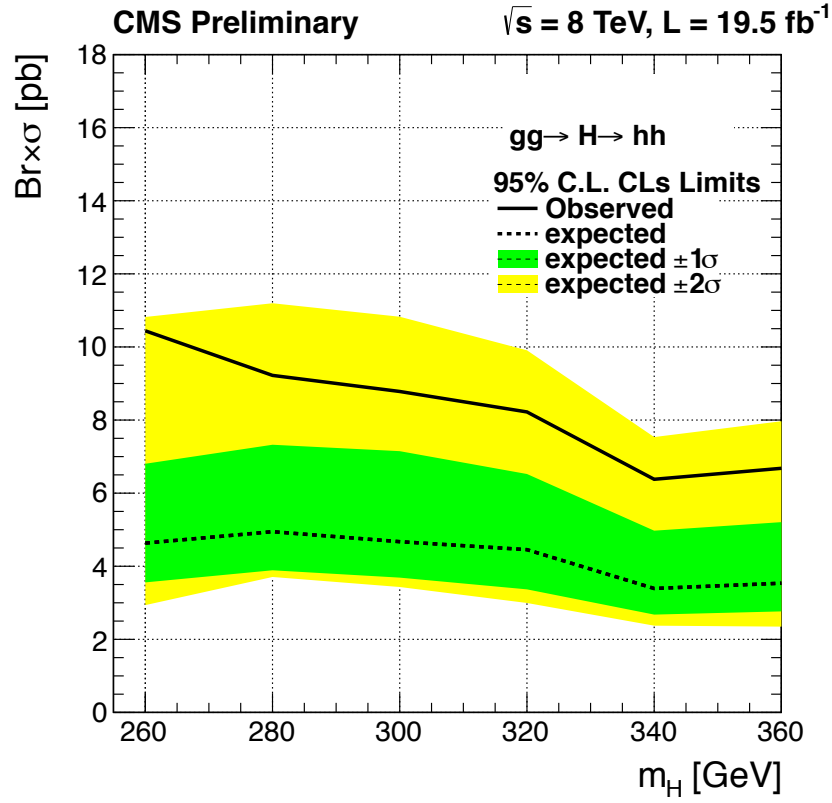
Figure 7.8: Observed and expected limits with 1 and 2- σ bands for $H \to hh$ in terms of $\sigma * BR$. These limits are based only on multilepton channels. Branching ratios for h are assumed to have Standard Model values. No contribution from gg→A→Zh is considered in this limit.

Figure 7.9: Observed and expected limits with 1 and 2- σ bands for $H \to hh$ in terms of $\sigma * BR$. These limits are based on multilepton and diphoton channels. Branching ratios for h are assumed to have Standard Model values. No contribution from gg→A→Zh is considered in this limit.
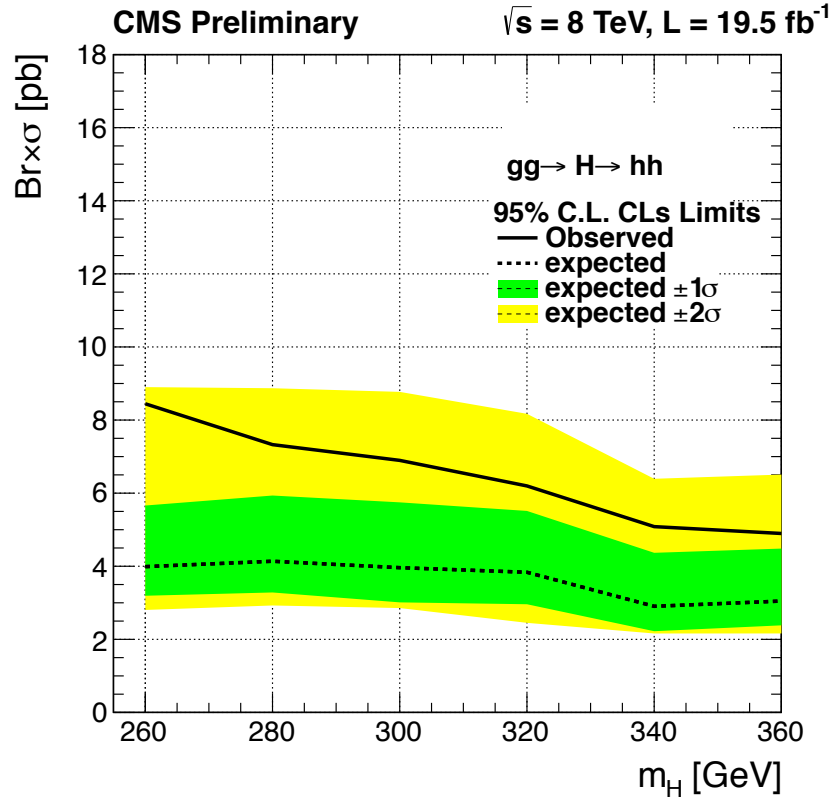
Figure 7.10: Observed and expected limits with 1 and 2- σ bands for $A \rightarrow Zh$ in terms of $\sigma * BR$. These limits are based only on multilepton channels . Branching ratios for h are assumed to have Standard Model values. No contribution from gg→H→hh is considered in this limit.
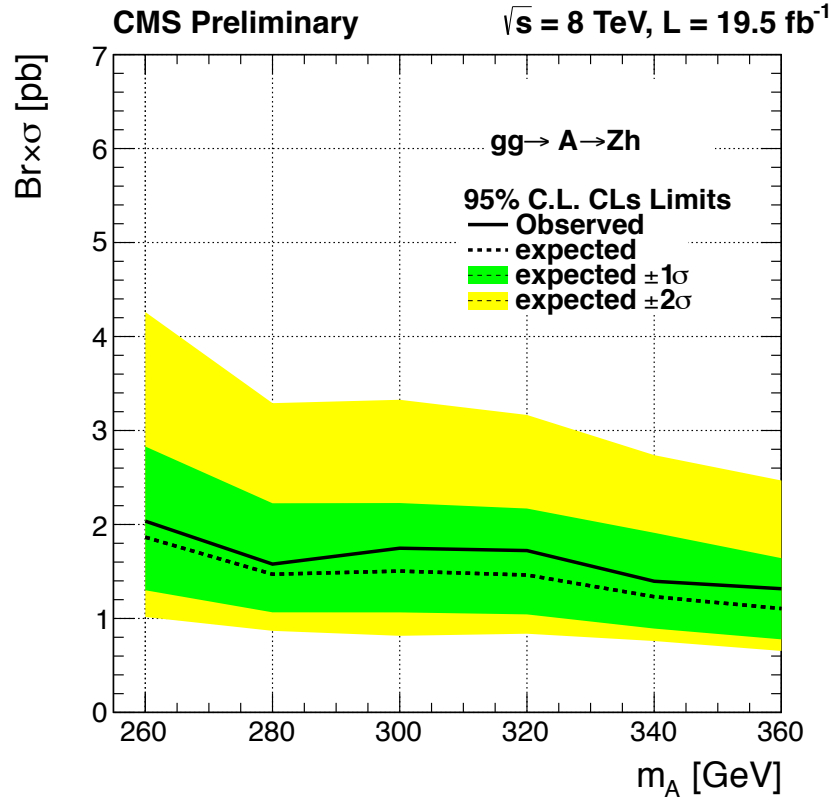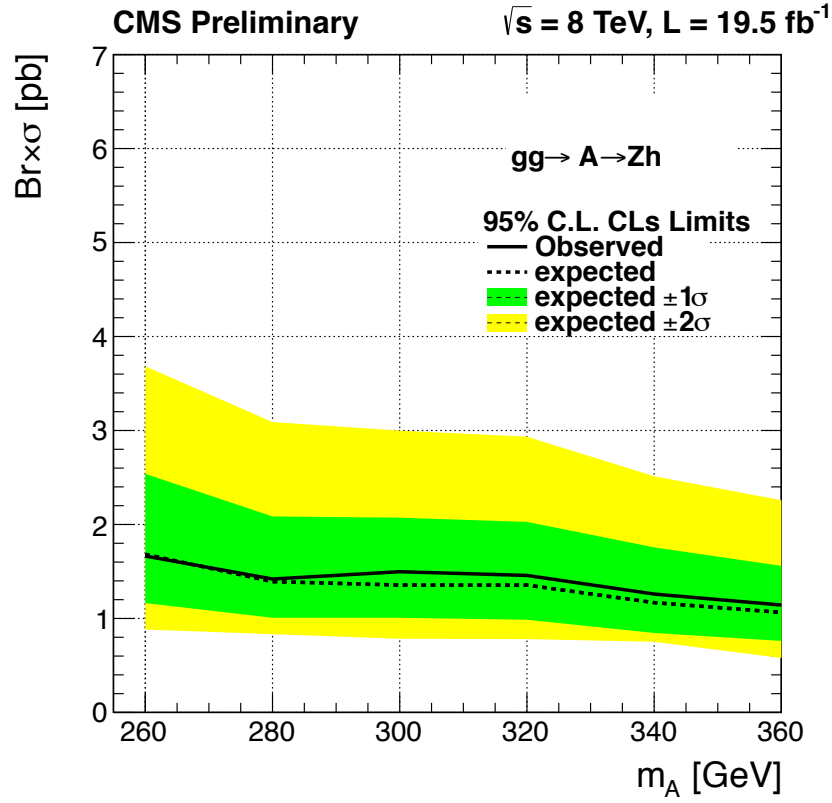
Figure 7.11: Observed and expected limits with 1 and 2- σ bands for $A \rightarrow Zh$ interms of $\sigma * BR$. These limits are based on multilepton and diphoton channels. Branching ratios for h are assumed to have Standard Model values. No contribution from gg→H→hh is considered in this limit.

Figure 7.12: Left: Observed and expected limits on Heavy higgs of mass 300 GeV in Type I 2HDM. The parameters α and β determine the cross section for H production, the $\mathrm{Br}(H \to hh)$ and the $\mathrm{Br}(h \to WW, ZZ, \tau\tau, \gamma\gamma)$. Right: The $\sigma * \mathrm{Br}(H{\to}hh)$ contours for TYPE I 2HDM. This figure is similar to one from theory paper by Nathaniel Craig et al. [13], the only difference is plotting of $\tan\beta$, instead of β, on the vertical axis. The regions below the observed limit lines and within the loop by marked by observed limit are excluded.
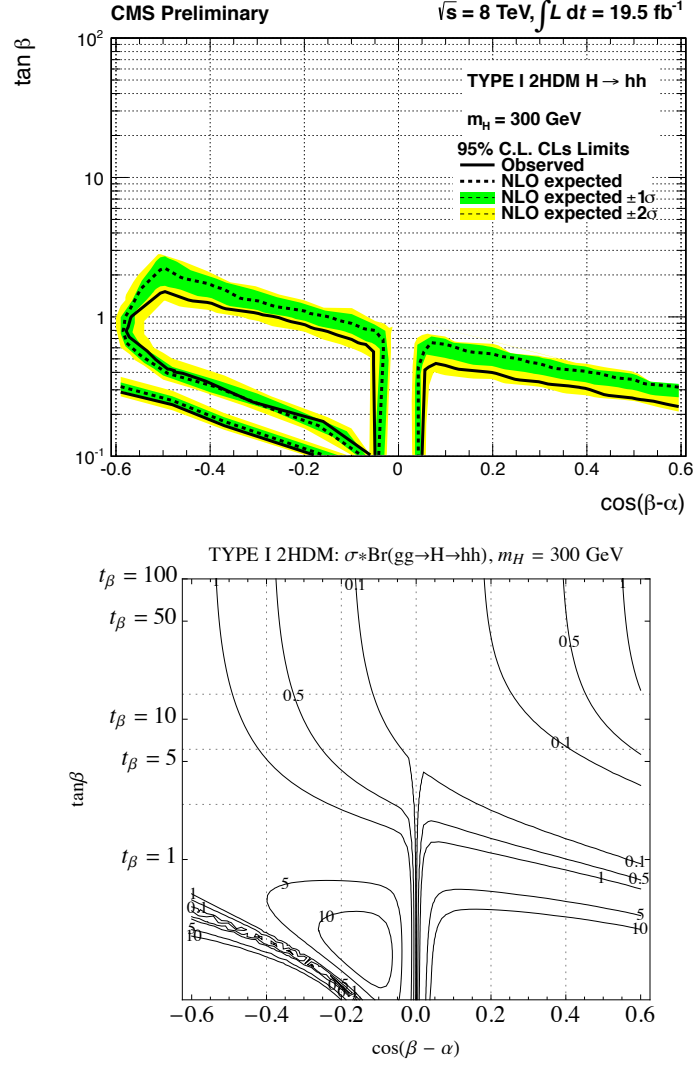
Figure 7.13: Left: Observed and expected limits on Heavy higgs of mass 300 GeV in Type II 2HDMs. The parameters α and β determine the cross section for H production, the $\text{Br}(H \to hh)$ and the $\text{Br}(h \to WW, ZZ, \tau\tau, \gamma\gamma)$. Right: The $\sigma * \text{Br}(\text{H}\to\text{hh})$ contours for TYPE II 2HDM. This figure is similar to one from theory paper by Nathaniel Craig et al. [13], the only difference is plotting of $\tan\beta$, instead of β, on the vertical axis. The regions below the observed limit lines and within the loop by marked by observed limit are excluded.
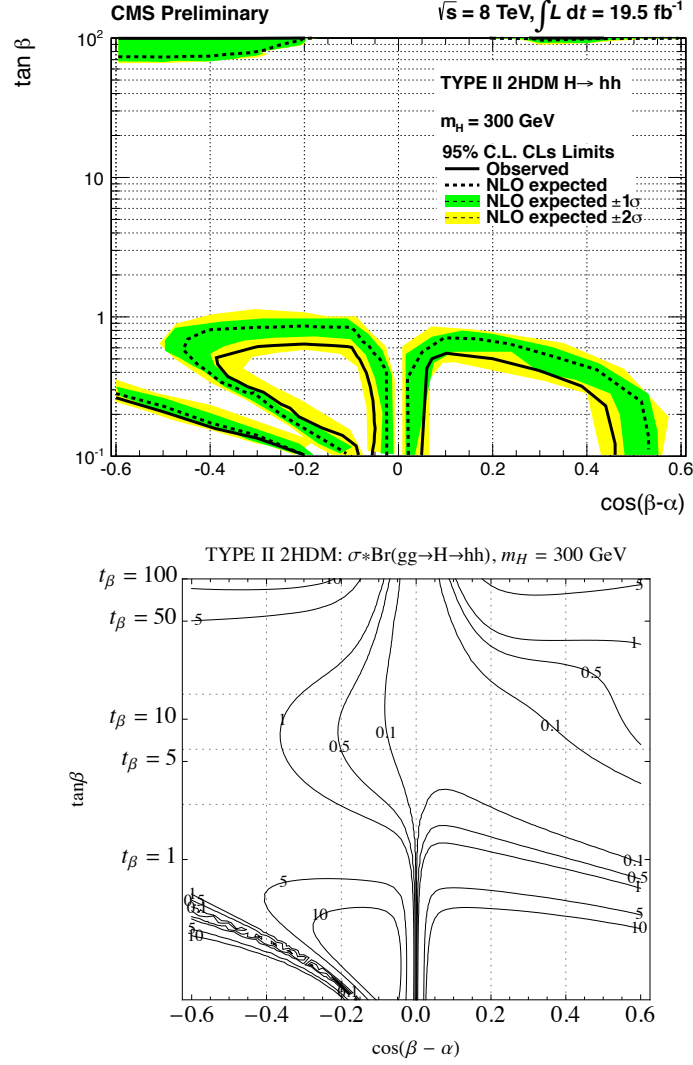
Figure 7.14: Left: Observed and expected limits on A of mass 300 GeV in Type I 2HDMs .The parameters α and β determine the cross section for H production, the $\mathrm{Br}(A \to Zh)$ and the $\mathrm{Br}(h \to WW, ZZ, \tau\tau, \gamma\gamma)$. Right: The $\sigma * \mathrm{Br}(A{\to}Zh)$ contours for TYPE I 2HDM. This figure is similar to one from theory paper by Nathaniel Craig et al. [13], the only difference is plotting of $\tan\beta$, instead of β, on the vertical axis. The region below the observed limit line is excluded.
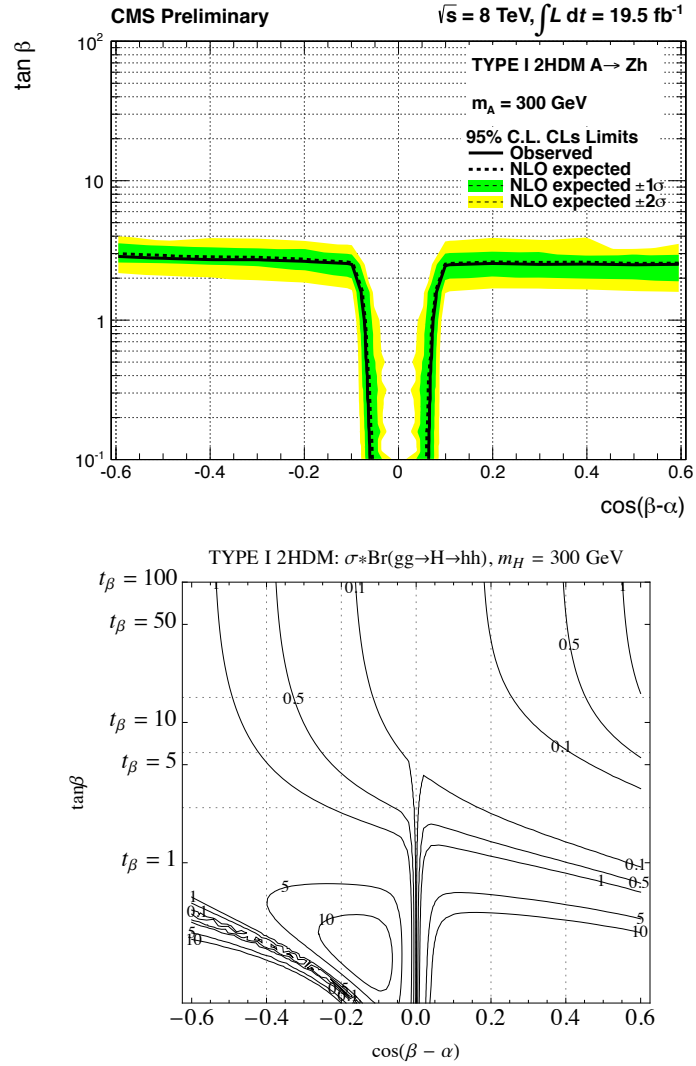
Figure 7.15: Left: Observed and expected limits on A of mass 300 GeV in Type II 2HDMs. The parameters α and β determine the cross section for H production, the $\mathrm{Br}(A \to Zh)$ and the $\mathrm{Br}(h \to WW, ZZ, \tau\tau, \gamma\gamma)$. Right: The $\sigma * \mathrm{Br}(A{\to}Zh)$ contours for TYPE II 2HDM. This figure is similar to one from theory paper by Nathaniel Craig et al. [13], the only difference is plotting of $\tan\beta$, instead of β, on the vertical axis. The region below the observed limit line is excluded.
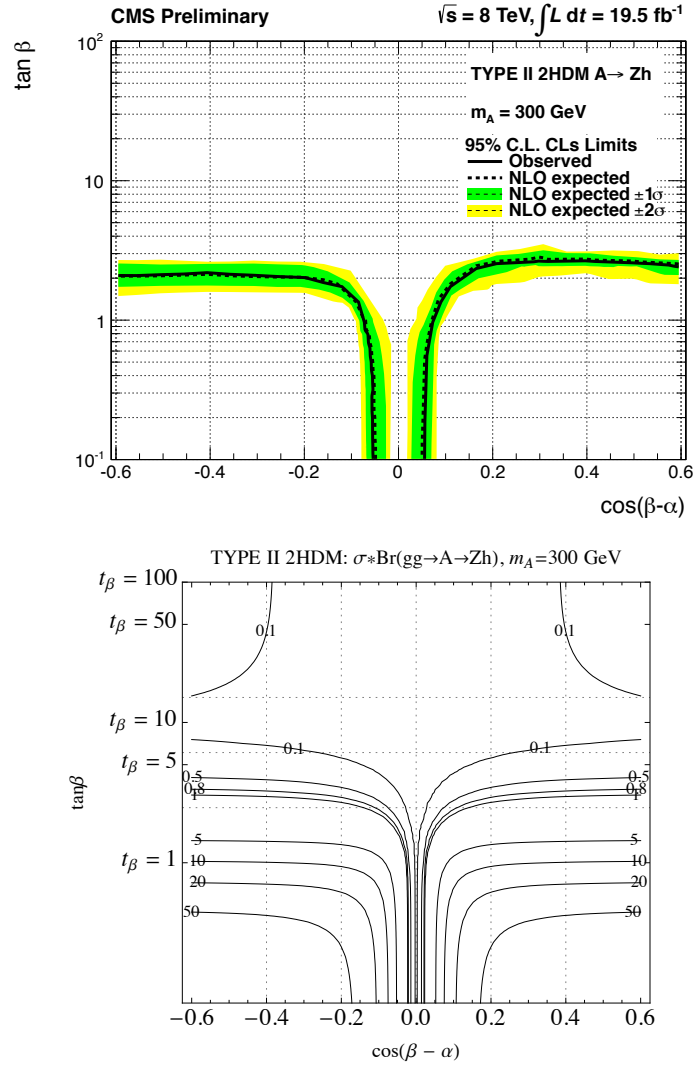
Figure 7.16: Left: Observed and expected limits with 1 and 2- σ bands on combined signal for Heavy Higgs and A inType I 2HDMs ($m_H = m_A = 300 GeV$). The parameters α and β determine the cross section for H and A production, the Br($H \to hh$) and B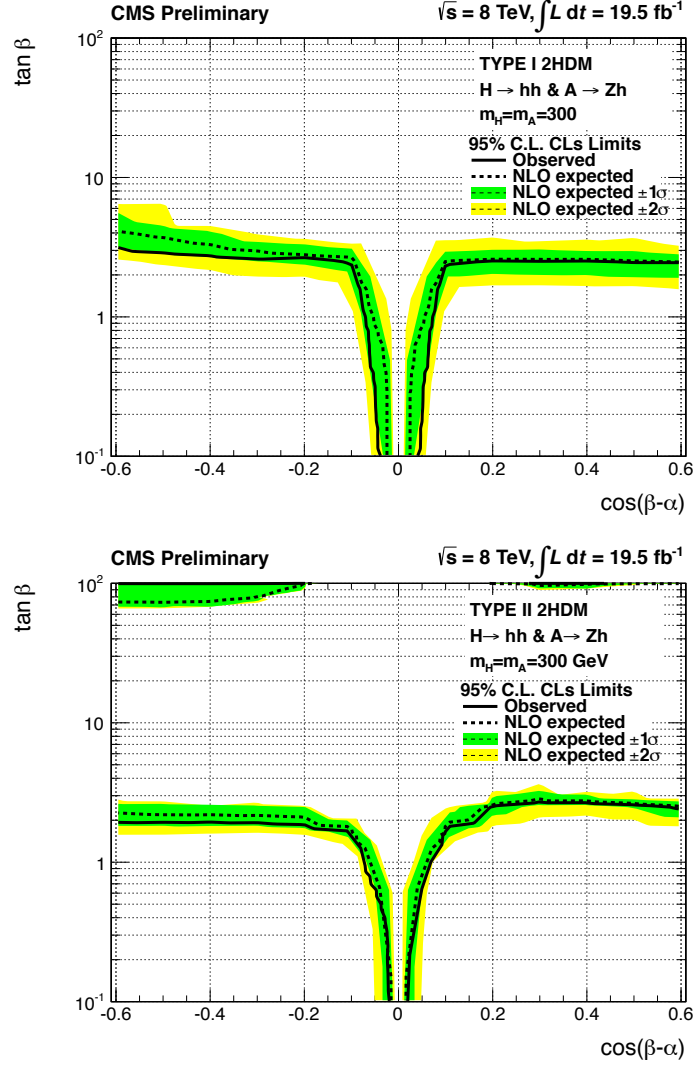r($A \to Zh$) and the Br($h \to WW, ZZ, \tau\tau, \gamma\gamma$). Right: Observed and expected limits with 1 and 2- σ bands on combined signal for Heavy Higgs and A inType II 2HDMs ($m_H = m_A = 300 GeV$). The region below the observed limit line is excluded.

Appendix A

Additional Results Plots

A.1   **Signal:** $HeavyHiggs \rightarrow hh$

The following plots show observations and background estimations for various three lepton and four lepton bins along with signal for Heavy Higgs of mass 300 GeV.

Figure A.1: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for 4-lepton + OSSF1 + off-Z + Tau0 + b0 (L) and 4-lepton + OSSF1 + off-Z + Tau1 + b0 (R).

Figure A.2: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for 4-lepton + OSSF2 + off-Z + Tau0 + b0 (L) and 4-lepton + OSSF2 + off-Z + Tau0 + b0 (R).

Figure A.3: E_T^{miss} distribution for 4-lepton + OSSF0 + Tau0 + b0 (L) and 4-lepton + OSSF0 + Tau1 + b0 (R).

Figure A.4: E_T^{miss} distribution for 4-lepton + OSSF1 + off-Z + Tau0 + b1 and 4-lepton + OSSF1 + off-Z + Tau1 + b1 (R).
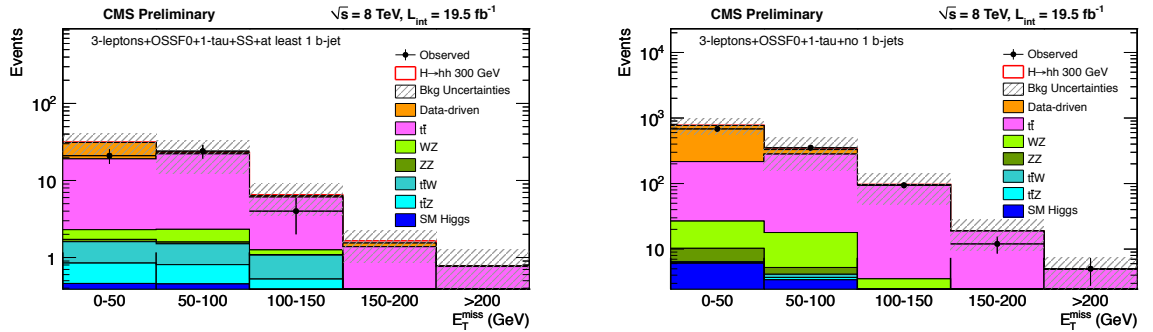
Figure A.5: E_T^{miss} distribution for 3-lepton + OSSF0 + Tau1+ b1 +SS 1 (L) and 3-lepton + OSSF0 + Tau1 + b0 (R).
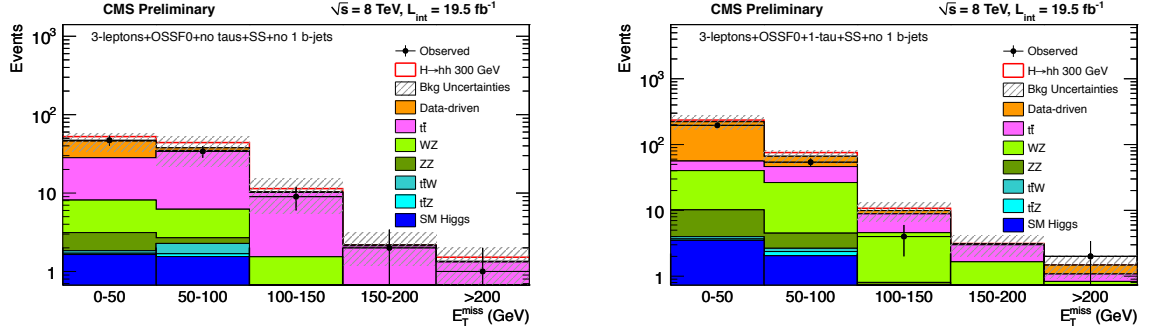
Figure A.6: $E_\mathrm{T}^\mathrm{miss}$ distribution for 3-lepton + OSSF0 + Tau0 + b0 + SS1 (L) and 3-lepton + OSSF0 + Tau1 + b0 + SS1 (R).
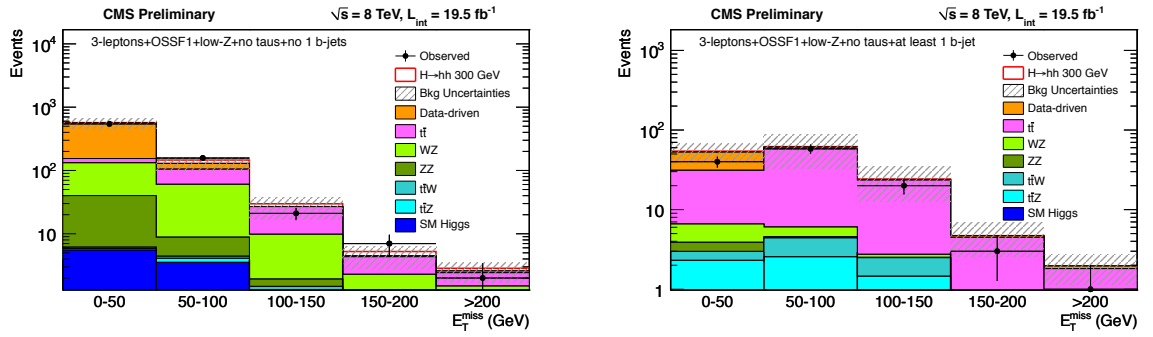
Figure A.7: $E_\mathrm{T}^\mathrm{miss}$ distribution for 3-lepton + OSSF1 + below-Z + Tau0 + b0 (L) and 3-lepton + OSSF2 + off-Z + Tau0 + b1 (R).
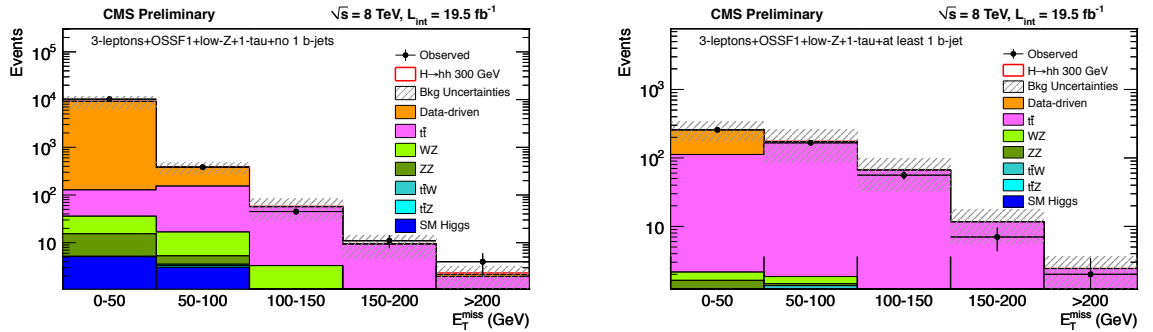
Figure A.8: $E_\mathrm{T}^\mathrm{miss}$ distribution for 3-lepton + OSSF1 + below-Z + Tau1 + b0 (L) and 3-lepton + OSSF1 + off-Z + Tau1 + b1 (R).

A.2   Signal: $A \to Zh$

The following plots show observations and background estimations for various three lepton and four lepton bins with signal for A of mass 300 GeV.
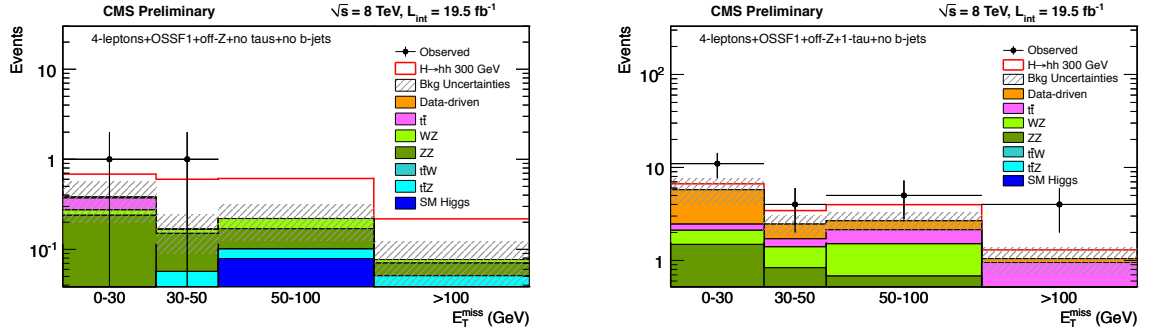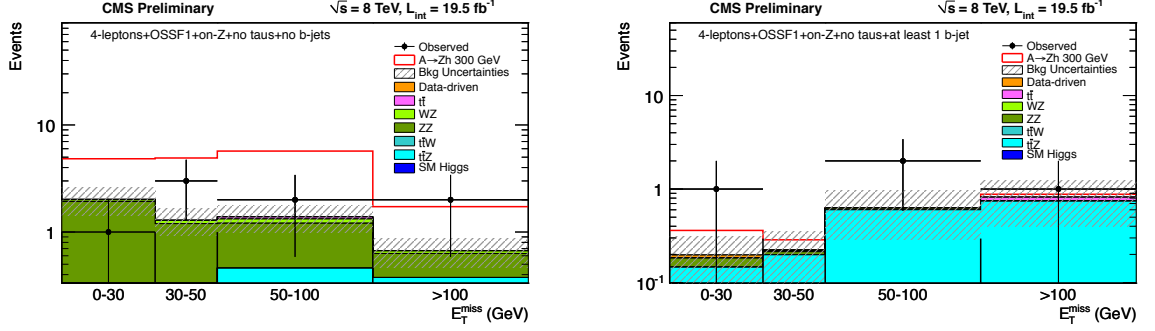

Figure A.9: E_T^{miss} distribution for 4-lepton + OSSF1 + on-Z + Tau0 + b0 (L) and 4-lepton + OSSF1 + on-Z + Tau0 + b1 (R).
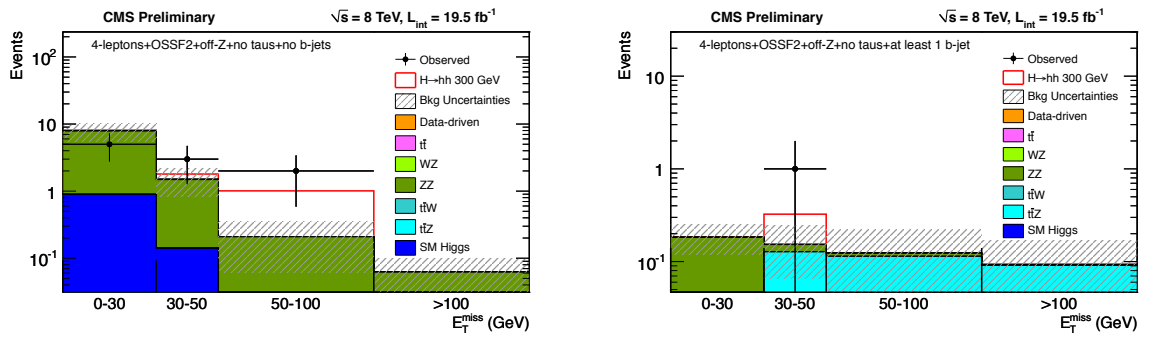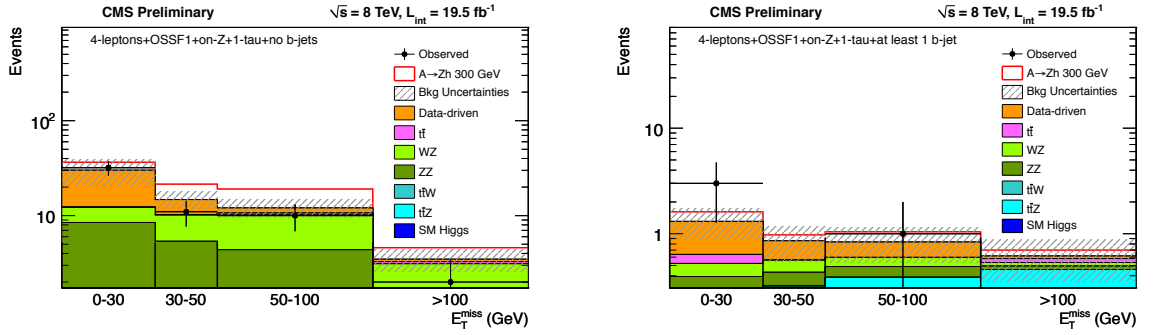

Figure A.10: E_T^{miss} distribution for 4-lepton + OSSF1 + on-Z + Tau1 + b0 (L) and 4-lepton + OSSF1 + on-Z + Tau1 + b1 (R).
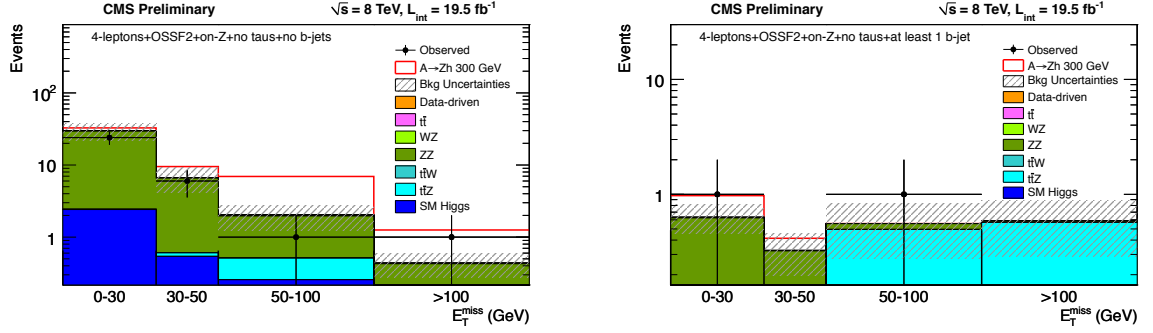
Figure A.11: E_T^{miss} distribution for 4-lepton + OSSF2 + on-Z + Tau0 + b0 (L) and 4-lepton + OSSF2 + on-Z + Tau0 + b1 (R).
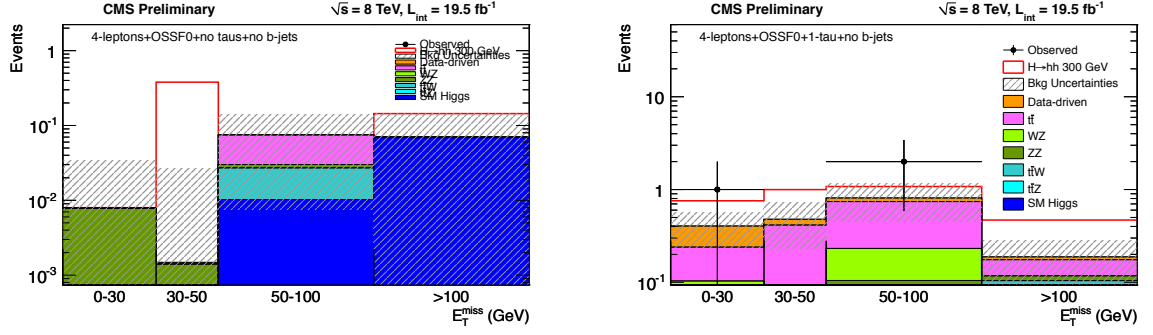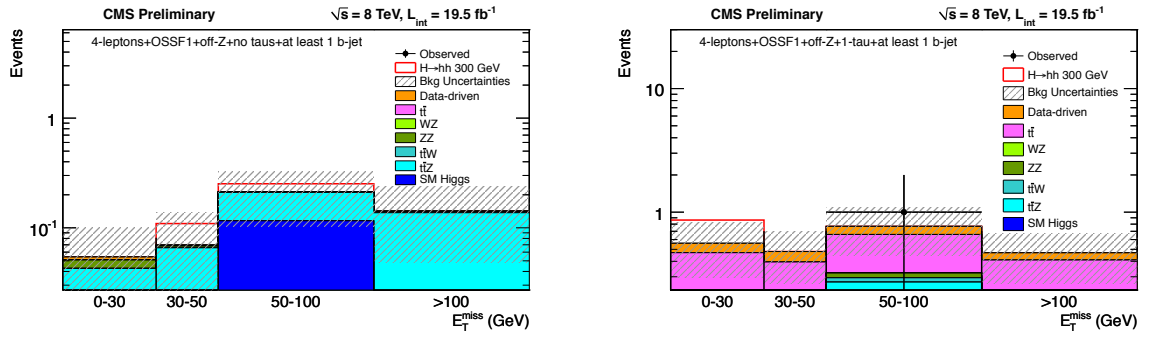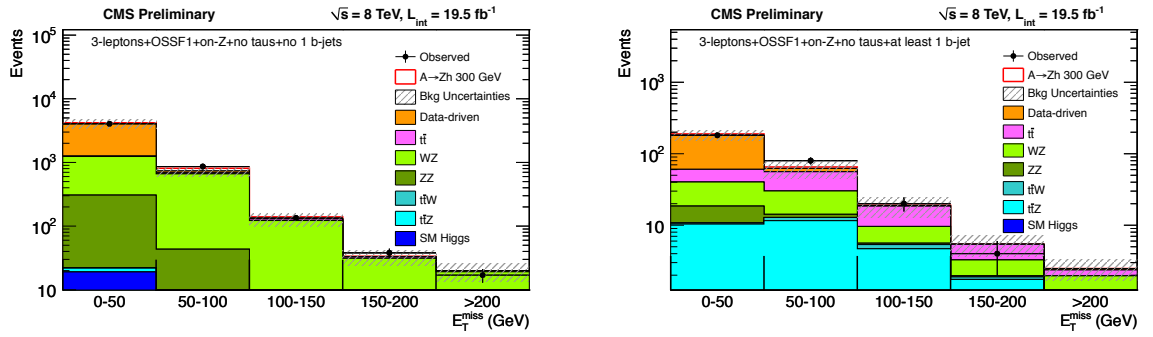

Figure A.12: E_T^{miss} distribution for 3-lepton + OSSF1 + on-Z + Tau0 + b0 (L) and 3-lepton + OSSF1 + on-Z + Tau0 + b1 (R).
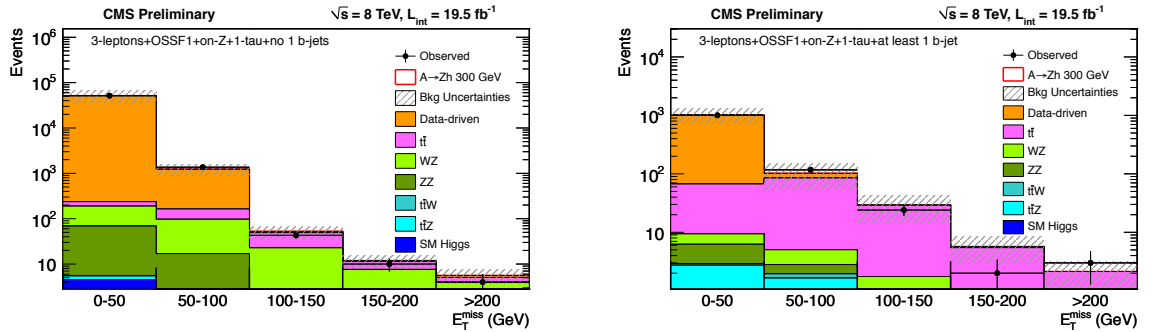
Figure A.13: E_T^{miss} distribution for 3-lepton + OSSF1 + on-Z + Tau1 + b0 (L) and 3-lepton + OSSF1 + on-Z + Tau1 + b1 (R).

Appendix B

Results using SM Higgs branching ratios

B.1 σ*BR 95% upper limit with SM higgs (126 GeV) branching ratios for SM-like Higgs in 2HDMs

Plots are presented for exclusion of 2HDM parameter space in chapter 7. Branching ratios for SM-like Higgs (to various final states) used in these plots are taken from 2HDM scenario and the vary for each value of α and β.
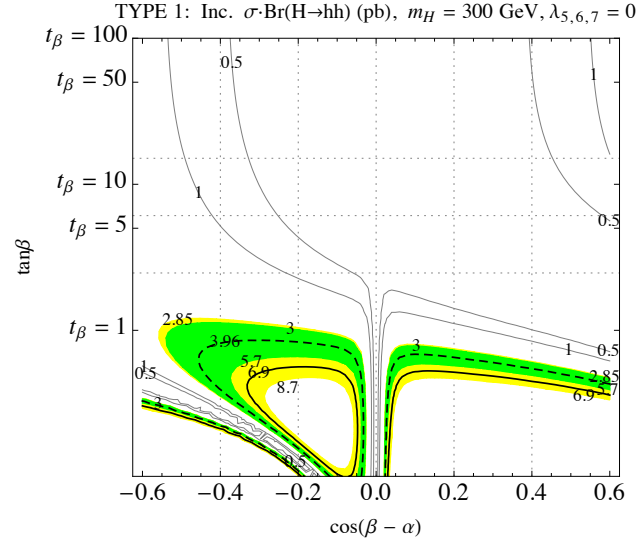
Figure B.1: Expected limit on Heavy higgs of mass 300 GeV in Type I 2HDMs. The different contour lines give σ*BR for $H \to hh$. Parameters α and β give Heavy Higgs's couplings to SM fermions and massive gauge bosons.

What if this SM-like Higgs in the 2HDM regime is the recently discovered SM Higgs? Figure B.1 and figure B.2 show regions of 2HDM parameter space that can be excluded if SM Higgs branching ratios are used for the light neutral Higgs for gg→H→hh case.
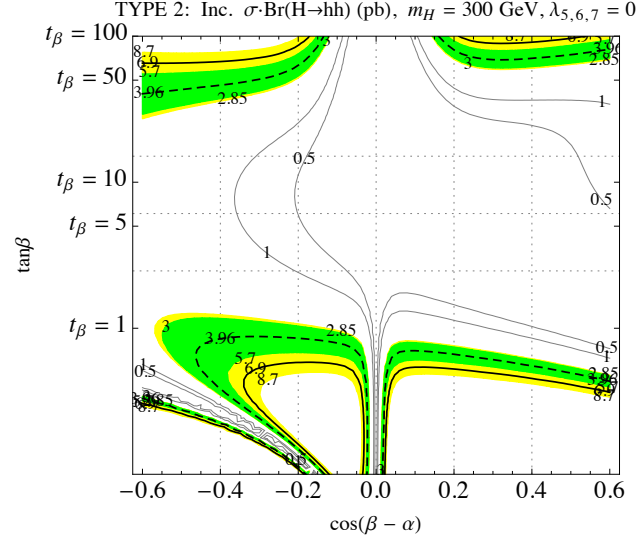
Figure B.2: Expected limit on Heavy higgs of mass 300 GeV in Type II 2HDMs . The different contour lines give σ*BR for $H \to hh$. Parameters α and β give Heavy Higgs's couplings to SM fermions and massive gauge bosons.
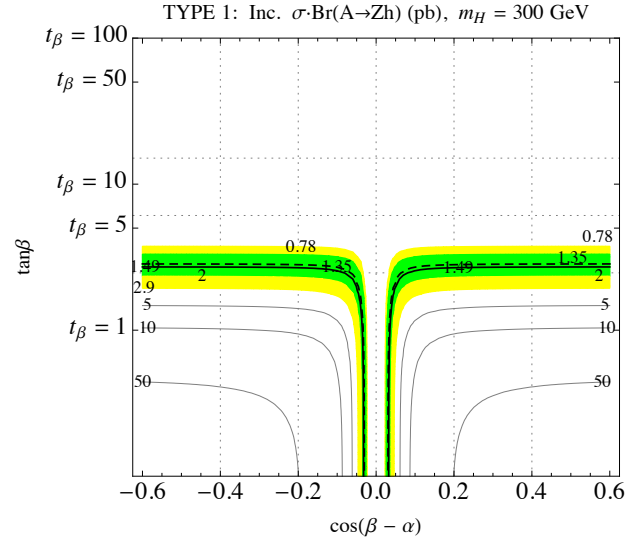
Figure B.3: Expected limit on A of mass 300 GeV in Type I 2HDMs. The different contour lines give σ*BR for $A \to Zh$. Parameters α and β give A's couplings to SM fermions and massive gauge bosons.

Figure B.3 and figure B.4 show similar exclusions for the gg→A→Zh case. It can be seen that the plots made with 2HDM branching ratios give limits comparable to the ones made using SM Higgs branching ratios. This just happens to be the case as the exclusions lie more or less along the region where the two branching ratios are very similar.
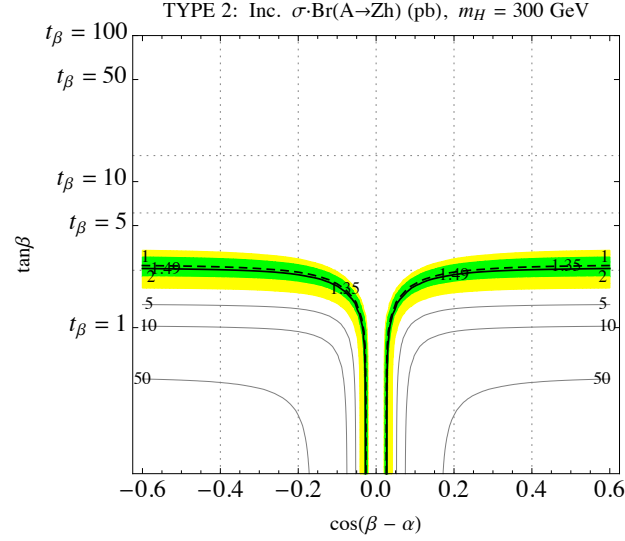
Figure B.4: Expected limit on A of mass 300 GeV in Type II 2HDMs. The different contour lines give σ*BR for $A \to Zh$. Parameters α and β give A's couplings to SM fermions and massive gauge bosons.

Appendix C

Triggers

C.1   Trigger List

The following is a list of all of the unprescaled triggers [5] used for the analysis. The 'v*' for different versions of the same trigger. Minor changes/updates to triggers happen throughout the period of data taking. This gives rise to different versions of the trigger. To make sure that no event is missed in the course of the analysis we use the logical OR of all the different triggers and their versions. Trigger efficiency is calculated for this logical OR. These efficiencies are monitored through out the course of data taking in case there are some major deviations which then will affect the trigger scale factors for MC.

DoubleMuon

- HLT_Mu17_Mu8_v*

Double Electron

- HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

Muon-Electron

- HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

- HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

DoublePhoton

- HLT_Photon26_R9Id85_OR_CaloId10_Iso50_Photon18_R9Id85_OR_CaloId10_Iso50_Mass60_v*

- HLT_Photon26_R9Id85_OR_CaloId10_Iso50_Photon18_R9Id85_OR_CaloId10_Iso50_Mass70_v*

- HLT_Photon26_CaloId10_Iso50_Photon18_R9Id85_Iso50_Mass60_v*

- HLT_Photon26_R9Id85_Photon18_CaloId10_Iso50_Mass60_v*

- HLT_Photon26_R9Id85_Photon18_R9Id85_Mass60_v*

- HLT_Photon26_R9Id85_Photon18_CaloId10_Iso50_Mass60_v*

- HLT_Photon36_CaloId10_Iso50_Photon22_CaloId10_Iso50_v*

- HLT_Photon36_CaloId10_Iso50_Photon22_R9Id85_v*

- HLT_Photon26_R9Id85_OR_CaloId10_Iso50_Photon22_R9Id85_OR_CaloId10_Iso50_v*

- HLT_Photon36_R9Id85_Photon22_CaloId10_Iso50_v*

- HLT_Photon36_R9Id85_Photon22_R9Id85_v*

Single Electron

- HLT_Ele8_CaloIdT_TrkIdVL_v*

- HLT_Ele8_CaloIdL_CaloIsoVL_v*

- HLT_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

- HLT_Ele17_CaloIdL_CaloIsoVL_v*

- HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

- HLT_Ele22_CaloIdL_CaloIsoVL_v*

- HLT_Ele27_CaloIdL_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

- HLT_Ele30_CaloIdVT_TrkIdT_v*

- HLT_Ele32_CaloIdL_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

- HLT_Ele65_CaloIdVT_TrkIdT_v*

- HLT_Ele80_CaloIdVT_TrkIdT_v*

- HLT_Ele100_CaloIdVT_TrkIdT_v*

- HLT_Ele90_CaloIdVT_GsfTrkIdT_v*

Single Muon

- HLT_Mu5_v*

- HLT_Mu8_v*

- HLT_Mu12_v*

- HLT_Mu17_v*

- HLT_Mu15_eta2p1_v*

- HLT_Mu24_eta2p1_v*

- HLT_Mu30_eta2p1_v*

- HLT_Mu40_eta2p1_v*

- HLT_Mu50_eta2p1_v*

- HLT_IsoMu20_eta2p1_v*

- HLT_IsoMu24_eta2p1_v*

- HLT_IsoMu30_eta2p1_v*

- HLT_IsoMu34_eta2p1_v*

- HLT_IsoMu40_eta2p1_v*

- HLT_IsoMu20_eta2p1_v*

- HLT_IsoMu24_eta2p1_v*

- HLT_IsoMu30_eta2p1_v*

- HLT_IsoMu34_eta2p1_v*

- HLT_IsoMu40_eta2p1_v*

- HLT_Mu24_v*

- HLT_Mu30_v*

- HLT_Mu40_v*

- HLT_IsoMu24_v*

- HLT_IsoMu30_v*

- HLT_Mu40_v*

Appendix D

Rdxy Derivation

D.1 Rdxy derivation

Here is the derivation for analytical relationship between R_{dxy} and ϵ_{ratio} [5]. Definitions:

$$N_{\text{NP}} \equiv \text{The number of non-prompt tracks } (d_{xy} > 0.02 \text{ cm}). \tag{D.1}$$

$$N_{\text{P}} \equiv \text{The number of prompt tracks } (d_{xy} < 0.02\text{cm}). \tag{D.2}$$

$$R_{dxy} \equiv \frac{N_{\text{NP}}}{N_{\text{P}}} \tag{D.3}$$

$$N_{\text{non}-\text{iso}}^{\text{T}} \equiv \text{The number of non-isolated tracks.} \tag{D.4}$$

$$N_{\text{iso}}^{\text{T}} \equiv \text{The number of isolated tracks.} \tag{D.5}$$

$$N_{\text{non}-\text{iso}}^{\ell} \equiv \text{The number of non-isolated leptons.} \tag{D.6}$$

$$N_{\text{iso}}^{\ell} \equiv \text{The number of isolated leptons.} \tag{D.7}$$

$$\epsilon_{\ell}^{\text{Iso}} \equiv \frac{N_{\text{iso}}^{\ell}}{N_{\text{non}-\text{iso}}^{\ell}} \tag{D.8}$$

$$\epsilon_{\text{T}}^{\text{Iso}} \equiv \frac{N_{\text{iso}}^{\text{T}}}{N_{\text{non}-\text{iso}}^{\text{T}}} \tag{D.9}$$

R_{dxy} as a function of β, which goes from 0 to 1, a is the 0-b-jet sample and b is the

maximum b-jet sample:

$$R_{dxy}(\beta) = \frac{(1-\beta) * \frac{N_{\mathrm{NP}}^a}{N_{\mathrm{NP}}^a + N_{\mathrm{P}}^a} + \beta * \frac{N_{\mathrm{NP}}^b}{N_{\mathrm{NP}}^b + N_{\mathrm{P}}^b}}{(1-\beta) * \frac{N_{\mathrm{P}}^a}{N_{\mathrm{NP}}^a + N_{\mathrm{P}}^a} + \beta * \frac{N_{\mathrm{P}}^b}{N_{\mathrm{NP}}^b + N_{\mathrm{P}}^b}} \tag{D.10}$$

$$\beta = \frac{1}{1 + \frac{\frac{R_{dxy} N_{\mathrm{P}}^b - N_{\mathrm{NP}}^b}{N_{\mathrm{P}}^b + N_{\mathrm{NP}}^b}}{\frac{N_{\mathrm{NP}}^a - R_{dxy} N_{\mathrm{P}}^a}{N_{\mathrm{P}}^a + N_{\mathrm{NP}}^a}}} \tag{D.11}$$

$$\beta = \frac{1}{1 + \frac{R_{dxy} - R_{dxy}^b}{R_{dxy}^a - R_{dxy}} \frac{1 + R_{dxy}^a}{1 + R_{dxy}^b}} \tag{D.12}$$

The efficiency ratio is the ratio of the lepton efficiency to the track efficiency:

$$\epsilon_{\mathrm{ratio}} \equiv \frac{\epsilon_{\ell}^{\mathrm{Iso}}}{\epsilon_{\mathrm{T}}^{\mathrm{Iso}}} \tag{D.13}$$

$\epsilon_{\mathrm{ratio}}$ as a function of β:

$$\epsilon_{\ell}^{\mathrm{Iso}}(\beta) = \frac{(1-\beta) * \frac{N_{\mathrm{iso}}^{\ell,\mathrm{a}}}{N_{\mathrm{iso}}^{\ell,\mathrm{a}} + N_{\mathrm{non-iso}}^{\ell,\mathrm{a}}} + \beta * \frac{N_{\mathrm{iso}}^{\ell,\mathrm{b}}}{N_{\mathrm{iso}}^{\ell,\mathrm{b}} + N_{\mathrm{non-iso}}^{\ell,\mathrm{b}}}}{(1-\beta) * \frac{N_{\mathrm{non-iso}}^{\ell,\mathrm{a}}}{N_{\mathrm{iso}}^{\ell,\mathrm{a}} + N_{\mathrm{non-iso}}^{\ell,\mathrm{a}}} + \beta * \frac{N_{\mathrm{non-iso}}^{\ell,\mathrm{b}}}{N_{\mathrm{iso}}^{\ell,\mathrm{b}} + N_{\mathrm{non-iso}}^{\ell,\mathrm{b}}}} \tag{D.14}$$

$$\epsilon_{\ell}^{\mathrm{Iso}}(\beta) = \frac{\epsilon_{\ell}^{\mathrm{Iso,a}} \epsilon_{\ell}^{\mathrm{Iso,b}} + \epsilon_{\ell}^{\mathrm{Iso,a}} + \beta * (\epsilon_{\ell}^{\mathrm{Iso,b}} - \epsilon_{\ell}^{\mathrm{Iso,a}})}{1 + \epsilon_{\ell}^{\mathrm{Iso,b}} + \beta * (\epsilon_{\ell}^{\mathrm{Iso,b}} - \epsilon_{\ell}^{\mathrm{Iso,a}})} \tag{D.15}$$

$$\epsilon_{\mathrm{T}}^{\mathrm{Iso}}(\beta) = \frac{\epsilon_{\mathrm{T}}^{\mathrm{Iso,a}} \epsilon_{\mathrm{T}}^{\mathrm{Iso,b}} + \epsilon_{\mathrm{T}}^{\mathrm{Iso,a}} + \beta * (\epsilon_{\mathrm{T}}^{\mathrm{Iso,b}} - \epsilon_{\mathrm{T}}^{\mathrm{Iso,a}})}{1 + \epsilon_{\mathrm{T}}^{\mathrm{Iso,b}} + \beta * (\epsilon_{\mathrm{T}}^{\mathrm{Iso,b}} - \epsilon_{\mathrm{T}}^{\mathrm{Iso,a}})} \tag{D.16}$$

$$\epsilon_{\mathrm{ratio}}(\beta) = \frac{\frac{\frac{1}{\beta} * (\epsilon_{\ell}^{\mathrm{Iso,a}} \epsilon_{\ell}^{\mathrm{Iso,b}} + \epsilon_{\ell}^{\mathrm{Iso,a}}) + (\epsilon_{\ell}^{\mathrm{Iso,b}} - \epsilon_{\ell}^{\mathrm{Iso,a}})}{\frac{1}{\beta} * (1 + \epsilon_{\ell}^{\mathrm{Iso,b}}) + (\epsilon_{\ell}^{\mathrm{Iso,b}} - \epsilon_{\ell}^{\mathrm{Iso,a}})}}{\frac{\frac{1}{\beta} * (\epsilon_{\mathrm{T}}^{\mathrm{Iso,a}} \epsilon_{\mathrm{T}}^{\mathrm{Iso,b}} + \epsilon_{\mathrm{T}}^{\mathrm{Iso,a}}) + (\epsilon_{\mathrm{T}}^{\mathrm{Iso,b}} - \epsilon_{\mathrm{T}}^{\mathrm{Iso,a}})}{\frac{1}{\beta} * (1 + \epsilon_{\mathrm{T}}^{\mathrm{Iso,b}}) + (\epsilon_{\mathrm{T}}^{\mathrm{Iso,b}} - \epsilon_{\mathrm{T}}^{\mathrm{Iso,a}})}} \tag{D.17}$$

Substitue for β:

$$\epsilon_{\mathrm{ratio}}(R_{dxy}) = \frac{\frac{(1 + \frac{R_{dxy} - R_{dxy}^b}{R_{dxy}^a - R_{dxy}} \frac{1 + R_{dxy}^a}{1 + R_{dxy}^b}) * (\epsilon_{\ell}^{\mathrm{Iso,a}} \epsilon_{\ell}^{\mathrm{Iso,b}} + \epsilon_{\ell}^{\mathrm{Iso,a}}) + (\epsilon_{\ell}^{\mathrm{Iso,b}} - \epsilon_{\ell}^{\mathrm{Iso,a}})}{(1 + \frac{R_{dxy} - R_{dxy}^b}{R_{dxy}^a - R_{dxy}} \frac{1 + R_{dxy}^a}{1 + R_{dxy}^b}) * (1 + \epsilon_{\ell}^{\mathrm{Iso,b}}) + (\epsilon_{\ell}^{\mathrm{Iso,b}} - \epsilon_{\ell}^{\mathrm{Iso,a}})}}{\frac{(1 + \frac{R_{dxy} - R_{dxy}^b}{R_{dxy}^a - R_{dxy}} \frac{1 + R_{dxy}^a}{1 + R_{dxy}^b}) * (\epsilon_{\mathrm{T}}^{\mathrm{Iso,a}} \epsilon_{\mathrm{T}}^{\mathrm{Iso,b}} + \epsilon_{\mathrm{T}}^{\mathrm{Iso,a}}) + (\epsilon_{\mathrm{T}}^{\mathrm{Iso,b}} - \epsilon_{\mathrm{T}}^{\mathrm{Iso,a}})}{(1 + \frac{R_{dxy} - R_{dxy}^b}{R_{dxy}^a - R_{dxy}} \frac{1 + R_{dxy}^a}{1 + R_{dxy}^b}) * (1 + \epsilon_{\mathrm{T}}^{\mathrm{Iso,b}}) + (\epsilon_{\mathrm{T}}^{\mathrm{Iso,b}} - \epsilon_{\mathrm{T}}^{\mathrm{Iso,a}})}} \tag{D.18}$$

References

[1] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012, 1207.7235.

[2] Nathaniel Craig, Jared A. Evans, Richard Gray, Can Kilic, Michael Park, et al. Multi-Lepton Signals of Multiple Higgs Bosons. *JHEP*, 1302:033, 2013, 1210.0559.

[3] CMS Collaboration. Search for anomalous production of multilepton events in pp collisions at $\sqrt{s}$=7 TeV. *JHEP*, 2012, 2012.

[4] S.R. Arora et. al. Search for extended Higgs sectors in the H hh and A Zh channels in $\sqrt(s) = 8$ TeV pp Collisions with multileptons and photons in final states. *CMS Analysis Note*, CMS-AN-2013/065, 2013.

[5] S. Arora, M. El Bahri, K. Chan, J.P. Chou, C. Contreras-Campana, E. Contreras-Campana, R.C. Gray, D. Kolchmeyer, A. Lath, K. Mei, S. Panwalkar, S. Schnetzer, S. Somalwar, S. Thomas, P. Thomassen, M. Walker, E. Williams, and P. Zywicki. Background and Efficiency Determination Methods for Multilepton Analyses. *CMS Analysis Note*, CMS-AN-2012/257, 2012.

[6] S.R. Arora et. al. MET Resolution Dependence on Pileup. *CMS Analysis Note*, CMS-AN-2012/342, 2012.

[7] Halzen Francis and Martin Alan. *Quarks & Leptons: An introductory course in modern particle physics*. John Wiley & Sons, New York, USA, 1984.

[8] Collins Peter D B, Martin Alan Douglas, and Squires Euan J. *Particle physics and cosmology*. Wiley, New York, NY, 1989.

[9] Sourabh Dube. *Search for supersymmetry at the Tevatron using the trilepton signature*. Ph.d., Rutgers University, 2008.

[10] D. Griffiths. *Introduction to Elementary Particles*. John Wiley & Sons, New York, USA, 1987.

[11] G.C. Branco, P.M. Ferreira, L. Lavoura, M.N. Rebelo, Marc Sher, et al. Theory and phenomenology of two-Higgs-doublet models. *Phys.Rept.*, 516:1–102, 2012, 1106.0034.

[12] Nathaniel Craig and Scott Thomas. Exclusive Signals of an Extended Higgs Sector. *JHEP*, 1211:083, 2012, 1207.4835.

[13] Nathaniel Craig, Jamison Galloway, and Scott Thomas. Searching for Signs of the Second Higgs Doublet. 2013, 1305.2424.

[14] Fabio Maltoni and Tim Stelzer. MadEvent: Automatic event generation with MadGraph. *JHEP*, 0302:027, 2003, hep-ph/0208156.

[15] S Heinemeyer et al. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. 2013, 1307.1347.

[16] Robert V. Harlander, Stefan Liebler, and Hendrik Mantler. SusHi: A program for the calculation of Higgs production in gluon fusion and bottom-quark annihilation in the Standard Model and the MSSM. *Computer Physics Communications*, 184:1605–1617, 2013, 1212.3249.

[17] David Eriksson, Johan Rathsman, and Oscar Stal. 2HDMC: Two-Higgs-Doublet Model Calculator Physics and Manual. *Comput.Phys.Commun.*, 181:189–205, 2010, 0902.0851.

[18] Michael (ed.) Benedikt, Paul (ed.) Collier, V (ed.) Mertens, John (ed.) Poole, and Karlheinz (ed.) Schindl. Lhc design report. Technical Report 2004/003, 2004.

[19] (Ed.) Bruning et al. LHC design report. Vol. I: The LHC main ring. Technical report, 2004. CERN-2004-003-V-1.

[20] CMS Collaboration. Cms physics : Technical design report volume 1: Detector performance and software. Technical Report 2006/001, 2006.

[21] CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update. Technical Report CMS-PAS-LUM-13-001, CERN, Geneva, 2013.

[22] CMS Collaboration. Commissioning of the particle-flow reconstruction in minimum-bias and jet events from pp collisions at 7 TeV. *CMS Physics Analysis Summary*, CMS-PAS-PFT-10-002, 2010.

[23] CMS. Muon selections https://twiki.cern.ch/twiki/bin/view/cmspublic/swguidemuonid.

[24] Serguei Chatrchyan et al. Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 1306:081, 2013, 1303.4571.

[25] CMS Collaboration. 2hdm scenario, h to hh and a to zh. *CMS Physics Analysis Summary*, CMS-PAS-HIG-13-025, 2013.

[26] CMS Collaboration. Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV. *CMS Physics Analysis Summary*, CMS-PAS-MUO-10-002, 2010.

[27] CMS Collaboration. Electron reconstruction and identification at $\sqrt{s} = 7$ TeV. *CMS Physics Analysis Summary*, CMS-PAS-EGM-10-004, 2010.

[28] CMS Collaboration. Performance of τ-lepton reconstruction and identification in cms. *Journal of Instrumentation*, 7, 2012.

[29] Serguei Chatrchyan et al. Identification of b-quark jets with the CMS experiment. *JINST*, 8:P04013, 2013, 1211.4462.

[30] CMS. B-tagging https://twiki.cern.ch/twiki/bin/viewauth/cms/btagpog.

[31] D. Barge et. al. Search for New Physics with Same Sign Dileptons using 2011 dataset of CMS. *CMS Analysis Note*, CMS-AN-2011/258, 2011.

[32] Serguei Chatrchyan et al. Measurement of the $t\bar{t}$ production cross section and the top quark mass in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 1107:049, 2011, 1105.5661.

[33] CMS Collaboration. Missing transverse energy performance in minimum-bias and jet events from proton-proton collisions at $\sqrt{s} = 7$ TeV. *CMS Physics Analysis Summary*, CMS-PAS-JME-10-004, 2010.

[34] Serguei Chatrchyan et al. Search for Physics Beyond the Standard Model Using Multilepton Signatures in pp Collisions at sqrt(s)=7 TeV. 2011, 1106.0933.

[35] Sanjay R. Arora. *Search for pair-produced heavy fourth-generation bottom-like quarks decaying to bZ and tW in 8,TeV proton-proton collisions with multilepton final states.* Ph.d., Rutgers University, 2013.

[36] R. Gray, S. Somalwar, E. Halkiadakis, A. J. Richards, S. Panwalkar, M. Klimek, and D. Hidas. Combined Fakeable Object (CFO) Method to Measure Lepton Fake Rates Using Multiple Data Objects. *CMS Analysis Note*, CMS-AN-2010/159, 2010.

[37] S.R. Arora et. al. A Search for Anomalous Production of Events with three or more leptons using 19.5 fb^1 of $\sqrt{}(s)$ =8 TeV CMS data. *CMS Analysis Note*, CMS-AN-2012/343, 2012.

[38] LHC. cross sections https://twiki.cern.ch/twiki/bin/view/lhcphysics/cernyellowreportpageat8tev.

[39] A L Read. Modified frequentist analysis of search results (the CL_s method). (CERN-OPEN-2000-205), 2000.

[40] Mingshui Chen and Andrey Korytov. Lands: software package for calculations of exclsion limits and signifcance. *CMS Analysis Note.*

[41] Thomas Junk. Confidence Level Computation for Combining Searches with Small Statistics. *Nucl. Instrum. Meth.*, A434:435–443, 1999, hep-ex/9902006.

[42] Alexander L. Read. Modified frequentist analysis of search results (The CL(s) method). Prepared for Workshop on Confidence Limits, Geneva, Switzerland, 17-18 Jan 2000.

[43] Alexander L. Read. Presentation of search results: The CL(s) technique. *J. Phys.*, G28:2693–2704, 2002.