# Florida State University Libraries

2016

# A Search for Supersymmetry with Two Photons and Missing Transverse Energy at CMS at a Center of Mass Energy of 13 TeV

Arka Santra

FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

A SEARCH FOR SUPERSYMMETRY WITH TWO PHOTONS AND MISSING

TRANSVERSE ENERGY AT CMS AT A CENTER OF MASS ENERGY OF 13 TEV

By

ARKA SANTRA

A Dissertation submitted to the
Department of Physics
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

2016

Arka Santra defended this dissertation on June 29, 2016.
The members of the supervisory committee were:

Andrew Askew

Professor Directing Dissertation

Gary Tyson

University Representative

Todd Adams

Committee Member

Laura Reina

Committee Member

Jorge Piekarewicz

Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the dissertation has been approved in accordance with university requirements.

To my parents, Asit and Sukla Santra, for consistently supporting me.

# ACKNOWLEDGMENTS

I am really thankful to my advisor Andrew Askew for his immense patience, relentless guidance and support during my PhD. Discussions on physics related to this search and other new possibilities really helped me a lot to evolve from a mere student to a critical thinker. His critical comments, be it in talk slides or in grant proposal writing, taught me a great deal on how scientific writings should be done. I also want to thank him for spending countless hours on correcting this dissertation and helping me to polish this document.

I am also grateful to other faculty members of the Department of Physics, Florida State University. The advices and suggestions of Todd Adams were extremely helpful throughout my graduate career. Guidance of Harrison Prosper taught me a lot in the statistical analysis of modern high energy experiments. I also like to thank Jorge Piekarewicz for answering numerous questions related to physics and helping me to overcome emotional issues.

I like to thank my colleagues in the supersymmetry di-photon analysis, specially Mike Hildreth, Marc Weinberg, Iasonas Topsis-Giotis and Allison Reinsvold for making this analysis possible.

The acknowledgement section will be incomplete if I don't mention my friends and officemates. Jordon Adams was really helpful during my early days as a graduate student. Innumerous discussions with Ajeeta Khatiwada and Sam Bein in physics and in other subjects helped me to evolve as a good physicist as well as a good human being. Andrew Ackert lent his helping hands whenever necessary. I am also grateful to Brian Francis who helped me learn many aspects of CMSSW software and Root though numerous online chats. Soham Subhra Ghosh and Eric Coulter, apart from being nice roommates, taught me many physics lessons.

None of this work or my accomplishments would be possible if my family did not support me. My parents have always been there whenever I needed emotional support. They make sure I take good care of my health and get back up after I fall. I am also grateful to Sudeshna Chakraborty for taking her time out from her busy schedule and going through the draft of this document and bearing with me while I was under stress.

I am sure I left out many others who helped me in one or other way throughout this endeavour. I thank them and am forever grateful to them.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF SYMBOLS

1. $\gamma$ - photon

2. $e$ - electron

3. $f$ - fake

4. $\tilde{g}$ - gluino

5. $\tilde{q}$ - squark

6. $\tilde{G}$ - gravitino

7. $\tilde{\chi}_1^0$ - neutralino

8. $E_T^{miss}$ - missing transverse energy

9. $p_T$ - transverse momentum

10. $f_{e \to \gamma}$ - electron to photon mis-identification rate (fake rate)

11. $\eta$ - pseudo-rapidity

# LIST OF ABBREVIATIONS

1. LHC - Large Hadron Collider

2. CMS - Compact Muon Solenoid

3. SM - Standard Model

4. BSM - Beyond Standard Model

5. SUSY - Supersymmetry

6. MSSM - Minimally Supersymmetric Standard Model

7. GMSB - Gauge-Mediated Supersymmetry Breaking

8. ECAL - Electromagnetic Calorimeter

9. EB - ECAL Barrel

10. EE - ECAL Endcap

11. HCAL - Hadronic Calorimeter

12. HB - HCAL Barrel

13. HE - HCAL Endcap

14. HF - Forward Calorimeter

15. HLT - High Level Trigger

# ABSTRACT

The present theoretical framework to describe the universe in particle level, the Standard Model describes only 4% of the matter-energy budget of the universe. There are many theories which attempt to describe the universe beyond the scope of the Standard Model. This dissertation describes a search for supersymmetry, a beyond Standard Model theory, using data collected by the Compact Muon Solenoid detector with integrated luminosity of 2.3 fb$^{-1}$ at a center of mass energy of 13 TeV during 2015. The data were produced in proton-proton collisions at the Large Hadron Collider near Geneva, Switzerland. This search was performed with events having two photons and missing transverse energy in the final state. This final state was motivated by general gauge mediated supersymmetry breaking, one of the theories on how breaking of supersymmetry can be mediated. No significant excess over the expected background was observed. The result was interpreted with simplified model scans and 95% upper limit on production cross sections are provided.

# CHAPTER 1

# INTRODUCTION

Mankind has poised a plethora of questions about the universe to date. Among those questions are what the universe is composed of, and how different components of matter interact with one another. These questions are still in the forefront of science. With the advent of modern scientific techniques, human beings were able to study the building blocks of nature. After decades of empirical study, the Standard Model (SM) of particle physics was constructed and describes the particles in the known universe and their interactions with each other. A brief description of the SM and the particles known to us is given in Chapter 2.

Even though the SM describes all the particles known presently in this universe accurately, by no means is this theory complete. It has limitations in the explanation of certain fundamental quantities and at very high energies. There are many Beyond Standard Model (BSM) theories which attempt to extend and complete the SM. This dissertation presents a search for new physics beyond the SM. This search is motivated by supersymmetry, a BSM theory described briefly in Chapter 2 which addresses many of the limitations of the SM.

This search was performed with data produced in proton-proton collisions at the Large Hadron Collider (LHC) located near Geneva, Switzerland. The center of mass energy of the collisions was 13 TeV and the data were collected during 2015 by the Compact Muon Solenoid (CMS) detector, corresponding to an integrated luminosity of 2.3 fb$^{-1}$. CMS is a multi-purpose hermetic detector which is designed to detect all the SM particles (except neutrinos) and measure potential momentum imbalances with high precision. The search was performed in events with two photons and missing transverse energy.

This document is organized as follows: Chapter 2 describes the SM, supersymmetry and briefly motivates the search for new physics with double photon and missing transverse energy final state. Chapter 3 describes the LHC and different sub-detectors of CMS. The object reconstruction at CMS is described in Chapter 4. The experimental procedures for this analysis are described in Chapters 5 - 7. Chapter 5 gives a brief description of datasets and trigger requirements, Chapter 6 presents the

event selection and photon identification efficiency measurement using the CMS detector, and finally Chapter 7 describes in detail how the backgrounds of this search were estimated using data-driven techniques. Finally Chapter 8 interprets the results obtained after the estimated backgrounds were compared with the data observed. Chapter 9 contains the summary and conclusions.

# CHAPTER 2

# THEORY

This chapter will describe the theoretical motivation and the framework which affects the experimental analysis and its interpretation. This description is extremely biased towards the topic of this dissertation, and by no means is this a complete description of the theoretical framework of the Standard Model and beyond Standard Model theories. First a brief description of the Standard Model will be given, then a beyond Standard Model theory, supersymmetry, will be described.

## 2.1   The Standard Model

The Standard Model of particle physics is very successful in describing all known physical phenomena except gravity to date. It is the best theoretical framework to describe and understand the dynamics of all the particles that have been observed in high energy collisions. The SM represents particles as discrete excitations of quantized fields which are described by Quantum Field Theory (QFT). In this construction, particles are minimal excitations of their own fields. The interactions of these excitations/particles are mediated through other particles which are known as 'force-carriers'.

There are four fundamental forces in the universe: gravity, electromagnetism, strong force and weak force. Gravity interacts with all the particles in the universe. Electromagnetism influences all the particles having electric charge (discussed in Section 2.2.2) and/or magnetic moment. All of the classical world can be described by these two forces. The strong force works within very short distances and is responsible for binding quarks together. The weak force is responsible for nuclear beta decay. The SM describes all the forces except gravity.

A method for combining the electromagnetic and weak forces was discovered by Sheldon Glashow [1]. Later, the Higgs mechanism [2–4] was incorporated into the electroweak interaction [5] to describe the origin of the masses of the particles, technically required for $W^{\pm}$ and $Z^0$ by Steven Weinberg and Abdus Salam. In 1970's, the strong interaction was added and that led to a non-

abelian gauge theory based on the gauge symmetry group:

$$SU(3)_C \times SU(2)_L \times U(1)_Y. \tag{2.1}$$

Here $SU(3)_C$ denotes the quantum chromodynamic (QCD) interaction, $SU(2)_L$ describes the weak interaction and $U(1)_Y$ for hypercharge. The theoretical framework describing these interactions is called the Standard Model. A more detailed description of the SM is given in Appendix A.

## 2.2 Quantum Numbers of the Standard Model

The quantum numbers of particles described by the SM play an important role in deciding particle properties. There are many quantum numbers in the SM like electric charge, weak isospin, baryon number and lepton number. Some of the most relevant quantum numbers for this dissertation are described below.

### 2.2.1 Spin

"Spin" is an intrinsic angular momentum of the elementary particles and composite particles like hadrons and atomic nuclei[1]. Since spin is an intrinsic property of elementary particles, particles of same kind possess the same magnitude of spin. Spin is a vector quantity, having magnitude and direction. But due to quantization, spin magnitude and direction can only be changed in certain discrete steps. The conventional definition of the spin quantum number is n/2 where n is a non-negative integer. Hence spin can have values of 0, 1/2, 1, 3/2, 2.

### 2.2.2 Electric Charge

Electric charge is again an intrinsic property of elementary particles. This is a physical property by virtue of which particles interact with electric and magnetic fields. Electric charges can be of two types, positive and negative. Particles having same sign of charge repel each other while particles with opposite sign of charge attract. An elementary particle with zero charge is called a neutral particle and does not interact with electric or magnetic fields.

---

[1]Quantum particles are by definition a representation of the Poincare group in quantum field theory. There are two casimir operators in the Poincare group and they define mass and spin for a particular representation. One consequence of Poincare invariance is the existence of spin for quantum particles. There is no classical analogue of this statement [6].

### 2.2.3 Color Charge

Color charge is another intrinsic property of particles, such as quarks and gluons, interacting with the strong force. There are three different possible color charges, labeled as 'red', 'blue' and 'green'. Each color has its opposite color. If these three color charges or one color and its anti-color charge are grouped together, the resulting state is colorless. Combinations of quarks only exist in color neutral states, i.e. the overall color of the combined state should be zero. This can be achieved by three quarks having three different colors (forming baryons) or two quarks having one color and its anti-color (forming mesons).

## 2.3 Particle Types of the Standard Model

Depending on spin, particles are of two types, 1) fermions (half-integer spins) and 2) bosons (integer spins). All the matter particles are fermions and the force carriers are bosons. Fermions follow Fermi-Dirac statistics and bosons follow Bose-Einstein statistics [7].

### 2.3.1 Fermions

Fermions include all the matter constituents in the known universe. The SM fermions are quarks and leptons and are listed in Table 2.1. The quarks interact via the strong, electromagnetic and weak forces, whereas electrons, muons and taus interact via the electromagnetic and weak forces. The neutrinos interact only via the weak interaction.

Fermions are further grouped into three 'generations' (as listed in Table 2.1). Between generations, the mass and flavor quantum numbers change, but the interactions and all other quantum numbers remain identical. Each member of a higher generation has greater mass than the corresponding member from the previous generation (except neutrinos). Higher generation fermions eventually decay into the lowest generation (first generation) fermions (again, neutrinos are exceptional in this case) and as a result, ordinary stable matter around us (up and down quarks and electrons) are just the first generation fermions.

### 2.3.2 Bosons

Vector bosons (having spin greater than zero) are the force carriers of the observable universe. Photons couple to quarks and charged leptons and mediate the electromagnetic force, gluons

Table 2.1: Summary of SM particles [8]

| Type | Name | Symbol | Mass | Charge | Spin | Generation |
|---|---|---|---|---|---|---|
| Quarks | Up | u | 2.3 MeV | $+2/3$ | 1/2 | I |
| | Down | d | 4.8 MeV | $-1/3$ | 1/2 | I |
| | Charm | c | 1.28 GeV | $+2/3$ | 1/2 | II |
| | Strange | s | 95 MeV | $-1/3$ | 1/2 | II |
| | Top | t | 173.34 GeV | $+2/3$ | 1/2 | III |
| | Bottom | b | 4.18 GeV | $-1/3$ | 1/2 | III |
| Leptons | Electron | e | 0.511 MeV | $-1$ | 1/2 | I |
| | Muon | $\mu$ | 105.7 MeV | $-1$ | 1/2 | II |
| | Tau | $\tau$ | 1.777 GeV | $-1$ | 1/2 | III |
| | Electron Neutrino[1] | $\nu_e$ | $< 2$ eV | 0 | 1/2 | I |
| | Muon Neutrino[1] | $\nu_\mu$ | $< 0.19$ MeV | 0 | 1/2 | II |
| | Tau Neutrino[1] | $\nu_\tau$ | $< 18.2$ MeV | 0 | 1/2 | III |
| Gauge Bosons | Gluon | g | 0 | 0 | 1 | N/A |
| | Photon | $\gamma$ | 0 | 0 | 1 | N/A |
| | Z Boson | Z | 91.2 GeV | 0 | 1 | N/A |
| | W Boson | $W^\pm$ | 80.4 GeV | $\pm 1$ | 1 | N/A |
| Scalar | Higgs Boson | H | 125.7 GeV | 0 | 0 | N/A |

[1]These are weak eigenstates, not mass eigenstates.

interact with quarks and with themselves and mediate the strong force. $W^\pm$/Z bosons mediate the electroweak interactions. These are all spin-1 vector bosons.

Since the photons are massless, the range of the electromagnetic force is infinite. The $W^\pm$/Z bosons are massive, meaning the weak interaction is very short range (on the order of $10^{-18}$ m). Electroweak symmetry is a symmetry which unifies the electromagnetic and weak forces. But it is a broken symmetry, and that is why photons and $W^\pm$/Z bosons have very different masses. The strong force is different from the electromagnetic and weak forces: here the potential increases with increasing distance (greater than $10^{-18}$ m) and force remains constant. At sufficient distance, the potential between two quarks is so huge that the energy stored in the field is capable of creating two new quarks in a bound state with opposite color charges. This is caused by the fact that gluons, carrying color, interact with themselves. As a consequence of this phenomenon, known as color confinement, isolated single quark states are not possible.
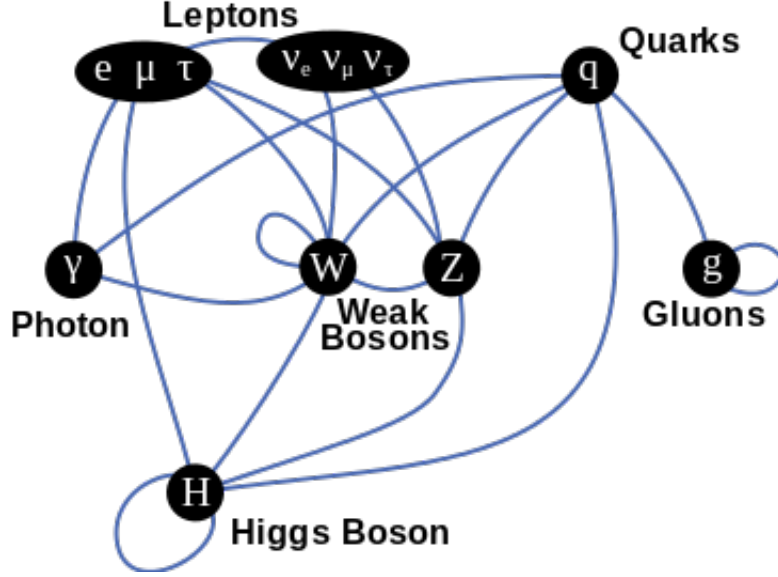
Figure 2.1: The SM particles and interactions are shown in this figure. If two particles are connected by a line, they interact in some way. If a particle is connected to itself, that indicates a self-interaction[9].

Although the photon and Z boson are elementary particles, they are related to fundamental SM fields $W^0$ and $B^0$ with the Weinberg mixing angle $\theta_W$:

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B^0 \\ W^0 \end{pmatrix} \tag{2.2}$$

The Higgs boson, named after physicist Peter Higgs and discovered in 2012 by the CMS and ATLAS experiments at CERN, is another boson with spin 0. So this is a scalar particle. The Higgs boson is the particle of the Higgs field which permeates all space and responsible for the mass of the elementary particles. Unlike the electromagnetic field, the Higgs field takes a non-zero constant vacuum expectation value (VEV) everywhere. This is the reason why after spontaneous symmetry breaking of electroweak symmetry [2], Higgs boson can give masses to all the particles that interact with Higgs boson.

All the SM particles and their interactions are shown in a schematic diagram in Figure 2.1
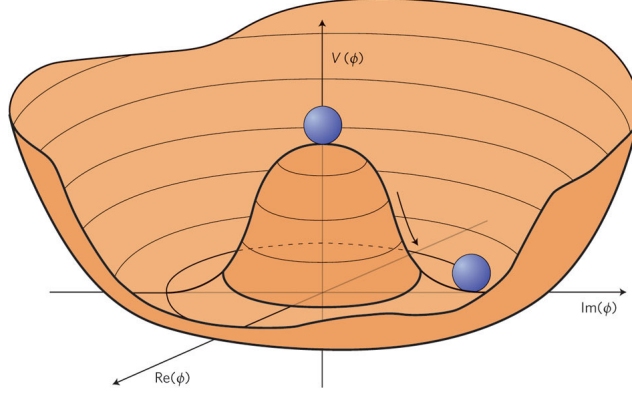
Figure 2.2: Mexican hat potential showing stable and unstable minima for Higgs boson[12].

## 2.4 Need for Beyond Standard Model Theories

The Standard Model does an excellent job in describing the constituents and interactions of visible universe except gravity. It has passed numerous tests with incredible precision. But it is not a complete theory of the universe. There are some open questions which the SM does not address. Questions like why there are three generations of quarks and leptons, or why there is a matter anti-matter asymmetry in the present universe are yet to be answered. The SM also does not give a candidate for dark matter or dark energy [10].

### 2.4.1 Deficiencies of the Standard Model

The SM Higgs field (H) is a complex scalar with the potential:

$$V = \mu^2 |H|^2 + \lambda |H|^4 \tag{2.3}$$

where $\mu^2 < 0$ and $\lambda > 0$. $H$ must have a non-zero VEV at the minimum of the potential $V$ in order to break electroweak symmetry. This results in a potential colloquially known as a "Mexican hat potential" (Figure 2.2). The system moves from an unstable minimum to the stable minimum and in the process it introduces mass terms in the Lagrangian for the Higgs, $Z^0$ and $W^{\pm}$ bosons. The photon remains massless. Here the VEV becomes $\langle H \rangle = \sqrt{-\mu^2/2\lambda}$ [11].

The Higgs boson mass is experimentally measured to be $125.09 \pm 0.21(\text{sys}) \pm 0.11(\text{stat})$ GeV [13]. However, the mass of the Higgs boson receives quantum corrections from every particle it
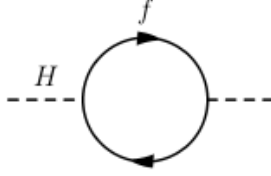
Figure 2.3: Loop diagram showing mass correction to Higgs boson mass due to fermions

couples to. For example, the one-loop correction to Higgs boson mass ($m_H$) coming from a fermion $f$ is depicted in Figure 2.3.

The Higgs field coupling to this fermion will generate a term in the Lagrangian $-\lambda_f H \bar{f} f$ where $f$ is the Dirac fermion and $\lambda_f$ is the coupling strength of the fermion $f$ to the Higgs field. This will give a correction to the Higgs boson mass:

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda_{UV}^2 + ... \tag{2.4}$$

Here $\Lambda_{UV}$ is the ultraviolet momentum cut off used to regulate the loop integral. It has a very high energy value several orders of magnitude higher than the Higgs boson mass. This can be interpreted as the energy scale where new physics enters the theory. Here higher order terms proportional to $\ln(\Lambda_{UV})$ are neglected. The coupling strength $\lambda_f$ is proportional to the mass, meaning the coupling is the strongest in the case of top quark, as it is the heaviest elementary particle discovered to date. Here each SM fermion can contribute to the correction. If one chooses the $\Lambda_{UV}$ to be Planck scale, this correction is huge ( $10^{30}$ GeV$^2$). These corrections will pull the corrected Higgs boson mass to the Planck scale ($\sim 1.2 \times 10^{19}$ GeV) rather than near the electroweak scale ($\sim 246$ GeV). Unless there is an incredible fine-tuning cancellation between the quadratic radiative corrections and the bare mass of Higgs boson, there is no explanation from the SM on why the Higgs boson is so much lighter than the Planck scale. This is known as the "Hierarchy Problem" [14].

It is experimentally observed that only 4.9% of the matter and energy of our universe is composed of SM particles[15], the rest is dark matter ( 26.8%) and dark energy ( 68.3%) [16]. Vera Rubin discovered that as the distance from galactic center increases, the rotational velocity of the galactic constituents do not decrease, as expected from the observable matter constituents, but remained constant [17]. From this study, it could be inferred that there is more matter in
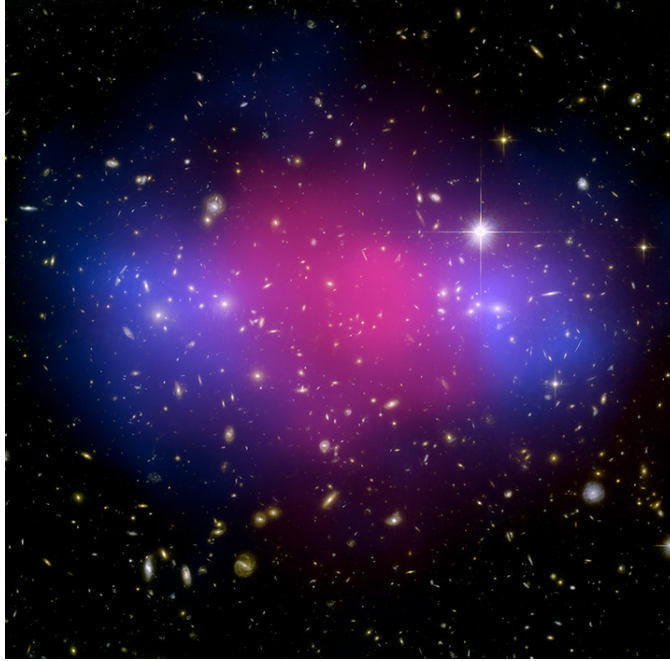
Figure 2.4: Reconstructed image of Bullet Cluster. The red color shows the x-ray emission from ordinary baryonic matter and the blue color represents the mass distribution of the cluster as observed through gravitational lensing. Reprinted from [19].

the universe which is not described by the SM. Additional evidence for dark matter comes from the "Bullet Cluster" which is the remnant of two colliding galaxies (Figure 2.4). The red color in the figure represents the x-ray radiation from ordinary baryonic matter. The blue color represents the gravitational centroid of mass distribution observed through gravitational lensing. This clearly shows that while the observed baryonic mass is near the center of the cluster, the bulk of the matter distribution is away from the center, which is not observable matter [18]. There is no suitable candidate for dark matter in the SM.

### 2.4.2 Motivation for Supersymmetry

Consider the case where the Higgs boson receives a mass correction from a massive scalar $S$. The Feynman diagram is shown in Figure 2.5. Here the correction term will look like (again neglecting the higher order terms):

$$\Delta m_H^2 = \frac{|\lambda_S|}{8\pi^2} \Lambda_{UV}^2 + ...$$

(2.5)
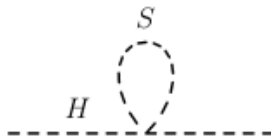
Figure 2.5: Loop diagram showing mass correction to Higgs boson mass due to scalar (S)

where $\lambda_S$ is the coupling strength of the scalar S. Comparing Equations 2.4 and 2.5, one can note the presence of a relative negative sign. So if each of the SM fermions is partnered with a complex scalar such that $|\lambda_S| = |\lambda_f|^2$, the $\Lambda^2_{UV}$ contributions from Equations 2.4 and 2.5 will cancel out exactly. More restrictions on the theory is necessary so that higher order terms are also canceled out neatly [20].

The partnering of SM fermions to complex scalars immediately implies that there is a symmetry between fermions and bosons. This is known as supersymmetry (SUSY). SUSY requires that each SM particle will have a SUSY partner differing by a spin of $\frac{1}{2}$. So SUSY can help to cancel higher order corrections to the Higgs boson mass and thus explains why the Higgs boson mass is so light compared with the Planck scale. Thus SUSY has the potential to solve the hierarchy problem.

As described above, the SM does not provide a suitable dark matter candidate. SUSY can help in this case too. The lightest supersymmetric particle (LSP), if stable and neutral, can be a potential candidate for dark matter.

Another motivation for SUSY is that it provides a prediction for the unification of strong, weak and electromagnetic interaction at a scale of $\sim 10^{16}$ GeV. Though the unification is not necessary theoretically, it is certainly aesthetically pleasing.

## 2.5 Supersymmetry

Supersymmetry (SUSY) is a proposed symmetry of spacetime that relates all the SM particles to partners having their spin shifted by 1/2 [21–23]. In general, SUSY solves a number of problems by introducing many new particles. In this dissertation, the discussion of SUSY will be limited to the Minimal Supersymmetic Standard Model (MSSM) which is the minimal extension

of the Standard Model in the sense that it introduces the minimal set of new particle states and interactions consistent with phenomenology [24].

The supersymmetric transformation turns a fermionic state into a bosonic state and vice versa. The operator $\mathcal{S}$ that generates this transformation must be an anti-commuting spinor and its effect is the following:

$$\mathcal{S}|Boson\rangle = |Fermion\rangle, \quad \mathcal{S}|Fermion\rangle = |Boson\rangle \tag{2.6}$$

Spinors are complex objects, so $\mathcal{S}^{\dagger}$ (the hermitian conjugate of $\mathcal{S}$ ) is also a symmetry generator. Since $\mathcal{S}$ and $\mathcal{S}^{\dagger}$ are fermionic operators, they carry spin angular momentum of spin $1/2$. The $\mathcal{S}$ and $\mathcal{S}^{\dagger}$ operators follow the commutation relations:

$$\left\{\mathcal{S}, \mathcal{S}^{\dagger}\right\} = P^{\mu}, \tag{2.7}$$

$$\{\mathcal{S}, \mathcal{S}\} = \left\{\mathcal{S}^{\dagger}, \mathcal{S}^{\dagger}\right\} = 0, \tag{2.8}$$

$$[P^{\mu}, \mathcal{S}] = \left[P^{\mu}, \mathcal{S}^{\dagger}\right] = 0 \tag{2.9}$$

where $P^{\mu}$ is the four momentum, generator of spacetime translation.

The single-particle states of SUSY are called *supermultiplets*, which are irreducible representation of the SUSY algebra [20]. Each supermultiplet will contain both fermionic and bosonic states. So the SM particles and their supersymmetric partners (known as superpartners of the corresponding SM particles) are grouped into supermultiplets. Each supermultiplet must have equal numbers of fermionic and bosonic degrees of freedom. So there can be two possibilities [20]:

1. chiral supermultiplets where there will be a single Weyl fermion and two real scalars,

2. gauge supermultiplets consisting of a spin-1 vector boson and a spin-1/2 Weyl fermion.

Suppose two states $|\psi\rangle$ and $|\psi'\rangle$ belong to the same supermultiplet. The $|\psi'\rangle$ will be equal to some linear combination of $\mathcal{S}$ and $\mathcal{S}^{\dagger}$ acting on $|\psi\rangle$. Also as the mass squared operator, $P^2$ commutes with both $\mathcal{S}$ and $\mathcal{S}^{\dagger}$ (Equation 2.9), $|\psi\rangle$ and $|\psi'\rangle$ will have the same mass eigenvalue. The $\mathcal{S}$ and $\mathcal{S}^{\dagger}$ commute with the generator of gauge transformation as well and hence $|\psi\rangle$ and $|\psi'\rangle$ will have the same electric charge, weak isospin and color charge. In summary, all the particles belonging to same supermultiplet will have the exact same quantum numbers and masses apart from their spin.

The naming scheme for supersymmetric particles are as follows:

1. For scalars, the names are constructed by prepending an 's' to the name of the corresponding SM partners, e.g. quark→squark.

2. For fermions, the names are made by either adding 'ino' or replacing 'on' with 'ino' at the end of the name of the SM partner particle, e.g. Higgs→Higgsino, gluon→gluino etc.

The SUSY particles are abbreviated by adding a tilde above the SM abbreviations, so squark= $\tilde{q}$, gluino= $\tilde{g}$ etc. The properties of the gauge and chiral supermultiplets are summarized in Tables 2.2 and 2.3 [20].

Table 2.2: Gauge Supermultiplets in the Minimal Supersymmetric Standard Model

| Names | spin-1/2 | spin$-1$ | SU(3)$_C$, SU(2)$_L$, U(1)$_Y$ |
|---|---|---|---|
| gluino, gluon | $\tilde{g}$ | $g$ | $(8,1,0)$ |
| winos, W | $\tilde{W}^{\pm}, \tilde{W}^0$ | $W^{\pm}, W^0$ | $(1,3,0)$ |
| bino, B | $\tilde{B}^0$ | $B^0$ | $(1,1,0)$ |

Table 2.3: Chiral Supermultiplets in the Minimal Supersymmetric Standard Model

| Names | Symbols | spin-0 | spin$-1/2$ | SU(3)$_C$, SU(2)$_L$, U(1)$_Y$ |
|---|---|---|---|---|
| squarks, quarks | $Q$ | $(\tilde{u}_L\ \tilde{d}_L)$ | $(u_L\ d_L)$ | $(3,2,\frac{1}{6})$ |
| (three families) | $\bar{u}$ | $\tilde{u}_R^*$ | $u_R^\dagger$ | $(\bar{3},1,-\frac{2}{3})$ |
| | $\bar{d}$ | $\tilde{d}_R^*$ | $d_R^\dagger$ | $(\bar{3},1,\frac{1}{3})$ |
| sleptons, leptons | $L$ | $(\tilde{\nu}\ \tilde{e}_L)$ | $(\nu\ e_L)$ | $(1,2,-\frac{1}{2})$ |
| (three families) | $\bar{e}$ | $\tilde{e}_R^*$ | $e_R^\dagger$ | $(1,1,1)$ |
| Higgs, higgsinos | $H_u$ | $(H_u^+\ H_u^0)$ | $(\tilde{H}_u^+\ \tilde{H}_u^0)$ | $(1,2,\frac{1}{2})$ |
| | $H_d$ | $(H_d^0\ H_d^-)$ | $(\tilde{H}_d^0\ \tilde{H}_d^-)$ | $(1,2,-\frac{1}{2})$ |

If SUSY were a perfect, unbroken symmetry, all the sparticles mentioned above would have the same masses as their SM superpartners. If this were the case, the sparticles would have been discovered at collider experiments already. This implies that if SUSY exists, it must be a broken symmetry.

But for SUSY to be still a solution to the Hierarchy Problem, the mass difference between sparticles and their SM partners must not be too large. To provide the cancellations to the corrected Higgs boson mass, $|\lambda_S|$ should still be equal to $|\lambda_f|^2$. The combined correction term to the Higgs

boson mass must follow the equation:

$$\Delta m_H^2 = \frac{|\lambda_S| - |\lambda_f|^2}{8\pi^2}\Lambda_{UV}^2 + ...$$ (2.10)

Keeping this constraint in mind, and also for the quadratic dependence on $\Lambda_{UV}^2$ to cancel, "soft" SUSY breaking is introduced. A Lagrangian term is added to the MSSM Lagrangian:

$$\mathcal{L} = \mathcal{L}_{SUSY} + \mathcal{L}_{Soft}$$ (2.11)

where $\mathcal{L}$ is the complete Lagrangian and $\mathcal{L}_{SUSY}$ contains all the terms necessary for gauge and Yukawa couplings of the unbroken MSSM. The newly introduced $\mathcal{L}_{Soft}$ term breaks supersymmetry and thus spontaneous supersymmetry breaking is required to extend the MSSM. Details of SUSY breaking can be found in the literature [20, 25, 26].

The MSSM soft terms in $\mathcal{L}_{Soft}$ arise radiatively or indirectly, not from the tree level renormalizable couplings to the SUSY breaking parameters. SUSY is broken in a "hidden sector" of particles that have either no or very small direct coupling to the "visible sector" chiral supermultiplets of the MSSM. These two sectors share some interactions which mediate the SUSY breaking from the hidden sector to the visible sector. There are currently three proposals about the mediating interactions: one where gravity is the mediator for the breaking of symmetry, one where anomaly in violation of a local superconformal invariance [27] is responsible for mediating the breaking of the symmetry and the third one where the general gauge interactions are responsible for mediation. This analysis is largely motivated by the third method, namely general gauge-mediated SUSY breaking (GMSB) [20].

### 2.5.1 Gauge Mediated SUSY Breaking (GMSB)

In GMSB models, the SM ordinary gauge interactions communicate the breaking of SUSY from the hidden sector to the visible sector, namely the MSSM. New chiral supermultiplets are introduced which couple to the source of SUSY breaking and indirectly to the (s)quarks, (s)leptons and higgs(inos), through the gauge bosons and gauge interactions of $SU(3)_C \times SU(2)_L \times U(1)_Y$. There is also gravitational communication, but this effect is relatively unimportant compared to the effect of gauge interactions [20]. Since the messenger particles have not been observed experimentally, it is thought that they have very high masses.

14

### 2.5.2 R-Parity

R-parity is a multiplicative quantum number which is defined as:

$$P_R \equiv (-1)^{3(B-L)+2s} \tag{2.12}$$

where $B$ is the baryon number, $L$ is the lepton number and $s$ is the spin of the particle. R-parity is designed in this way so that it can suppress $B$ and $L$ violating term in the theory. All the SM particles have $P_R = +1$ while all SUSY particles have $P_R = -1$. If R-parity is conserved, SUSY particles must be pair-produced in collider experiments, as there must be even number of R-parity particles for each vertex [28].

### 2.5.3 Lightest Supersymmetric Particle (LSP)

If R-parity is conserved, the lightest supersymmetric particle coming out of a decay chain must be stable. For any decay chain, the SUSY particles will always decay to an odd number of LSPs (along with SM particles), usually only one. For gauge-mediated SUSY breaking models, the LSP is the gravitino ($\tilde{G}$). The gravitino has a mass on the eV-keV scale. Any gravitino produced at CMS will escape without being detected, creating an imbalance in energy. This can be reconstructed by vectorially adding all transverse momenta in the detector and looking for any imbalance, which will give missing transverse energy (MET or $E_T^{miss}$ , for details see Section 4.7).

### 2.5.4 Final State Considered for the Analysis

The next-to-lightest SUSY particles (NLSP) are constrained to decay only to the LSP for an R-parity conserving theory. For this analysis, a long lived NLSP will not be considered, though they are theoretically possible. A neutralino ($\tilde{\chi}_1^0$) mass eigenstate is an admixture of Higgsino, Bino and Wino particles which is electromagnetically charge neutral. For this analysis, a Bino-like neutralino is required to decay to a photon and gravitino, giving us the photon final state for this analysis.

If SUSY particles are pair produced at CMS, their decay chains will include the neutralino, which will eventually decay to photon (sometimes Z boson) and gravitino. Thus the final state of large $E_T^{miss}$ and two photons will be obtained. The possible representative diagrams producing two photons and $E_T^{miss}$ through strong production are shown in Figure 2.6
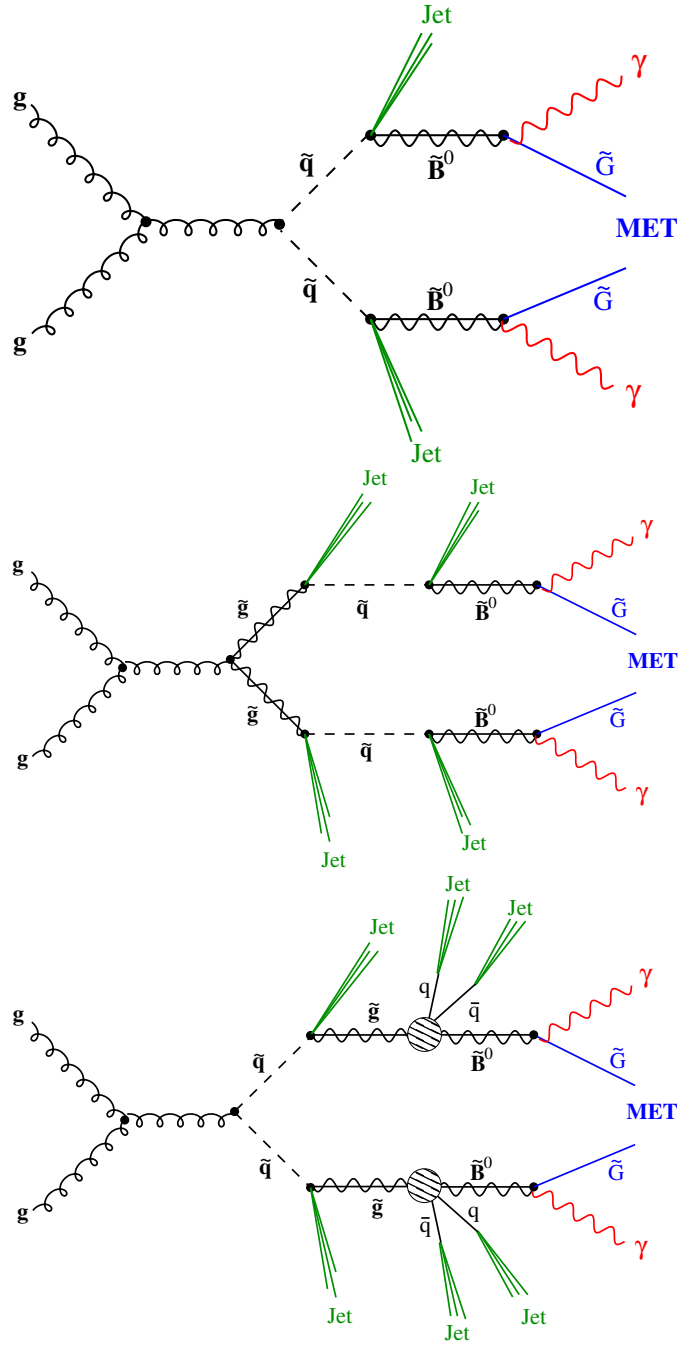
Figure 2.6: Production process showing GMSB SUSY scenario. A bino-like NLSP ($\tilde{B}^0$) decays to a photon and a gravitino with different numbers of jets.

## 2.6 Previous Results from Searching for GMSB SUSY

There were searches for SUSY at LEP experiments at CERN [29]. The GMSB scenario limits from LEP are shown in Figure 2.7. The D0 and CDF experiments also set limits on SUSY scenarios which guided the search strategy for the future [30]. The limit plot from Tevatron searching for GMSB scenario is in Figure 2.8.

In the present LHC era (Run 2), the CMS and ATLAS experiments are searching for SUSY signatures. After Run I, they were able to put limits on SUSY sparticles for different search channels. The combined limit plots from different analyses in gluino mass and LSP mass planes may be found in this summary page [34].

There were searches for SUSY with the two photons and $E_T^{miss}$ final state using CMS with Run I data. The most recent search result is from $\sqrt{s} = 8$ TeV dataset with luminosity of 19.7 fb$^{-1}$. The search produced exclusion contours (Figure 2.9) in the gluino-neutralino mass plane. In this search, gluino masses up to 1.35 TeV were excluded at 95% confidence level [35].
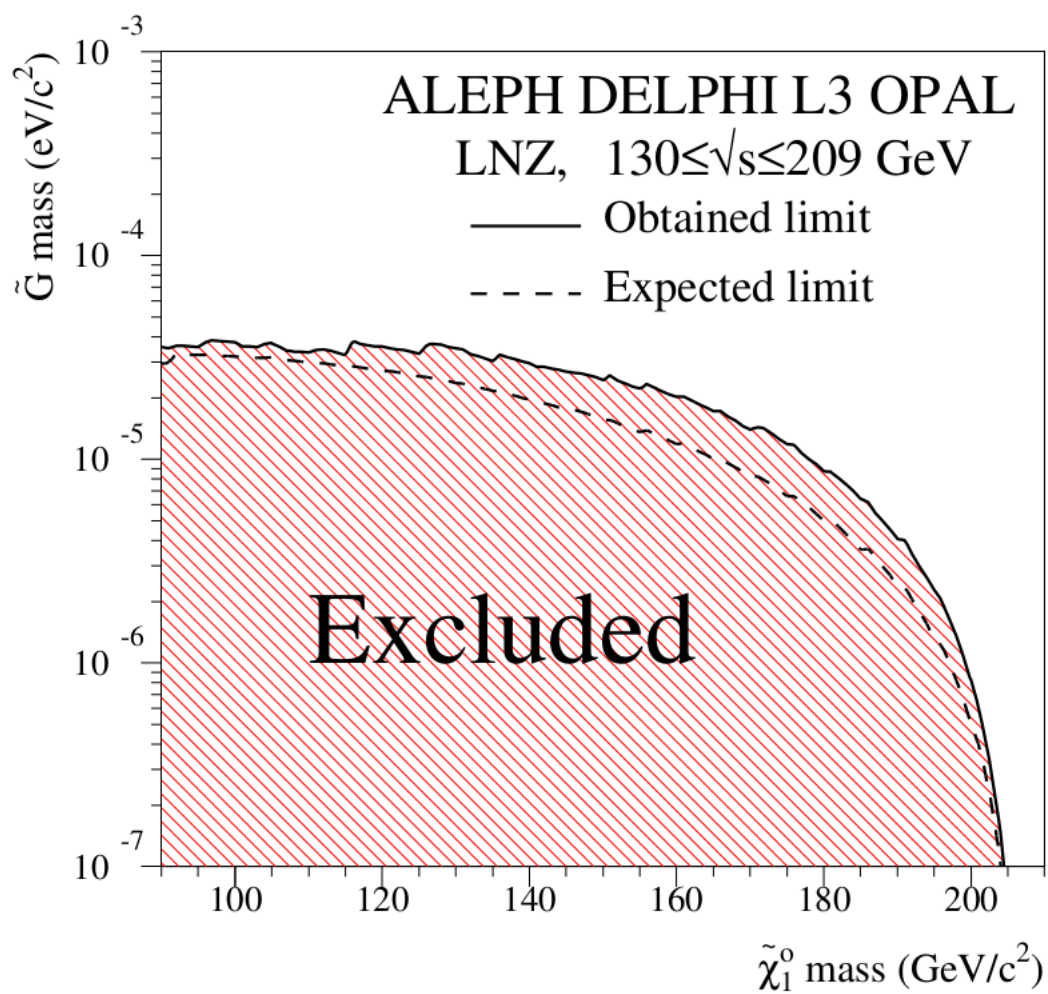
Figure 2.7: The limit on gravitino and neutralino production at LEP [29].
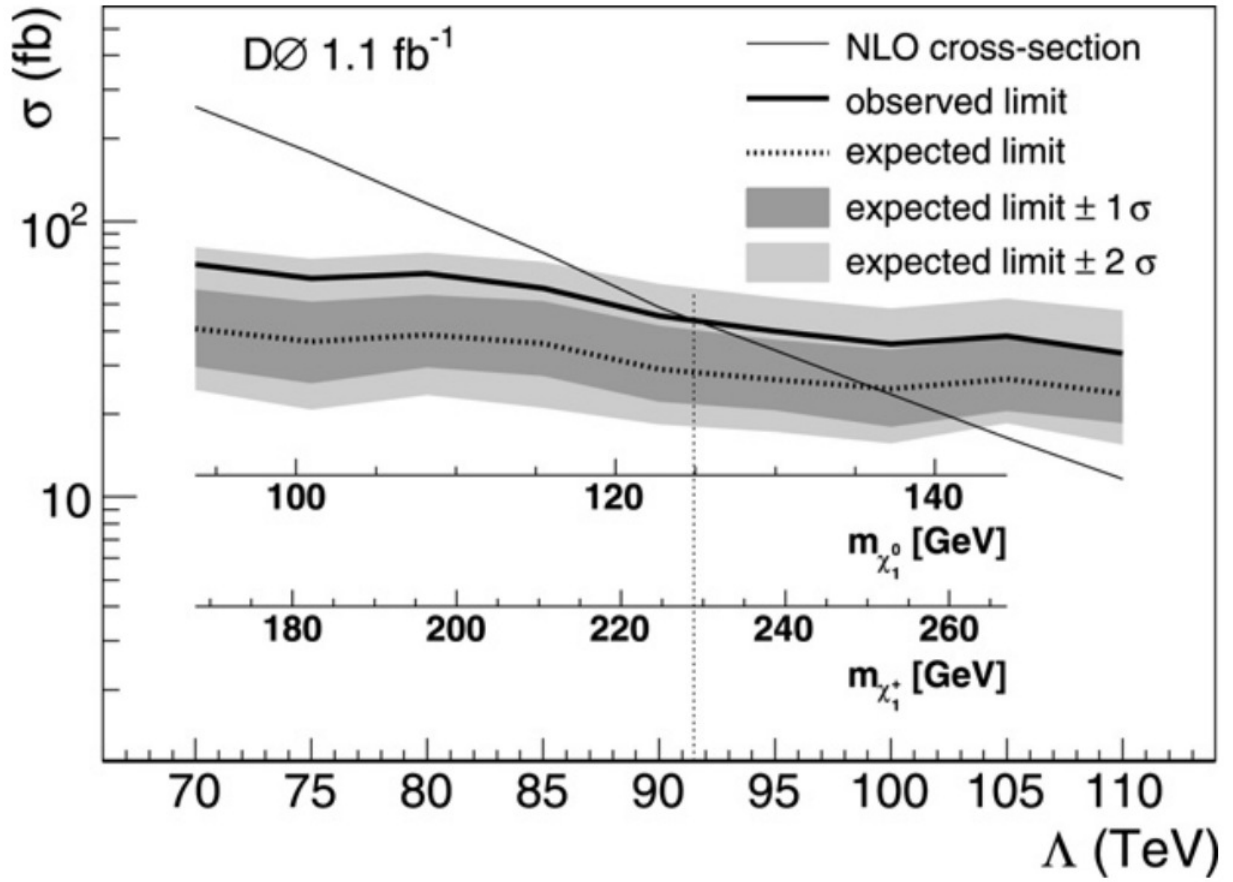
Figure 2.8: The predicted cross-section for the Snowmass Slope model [31, 32] versus Λ, an energy scale that determines the effective scale of SUSY breaking. Reprinted from [33].
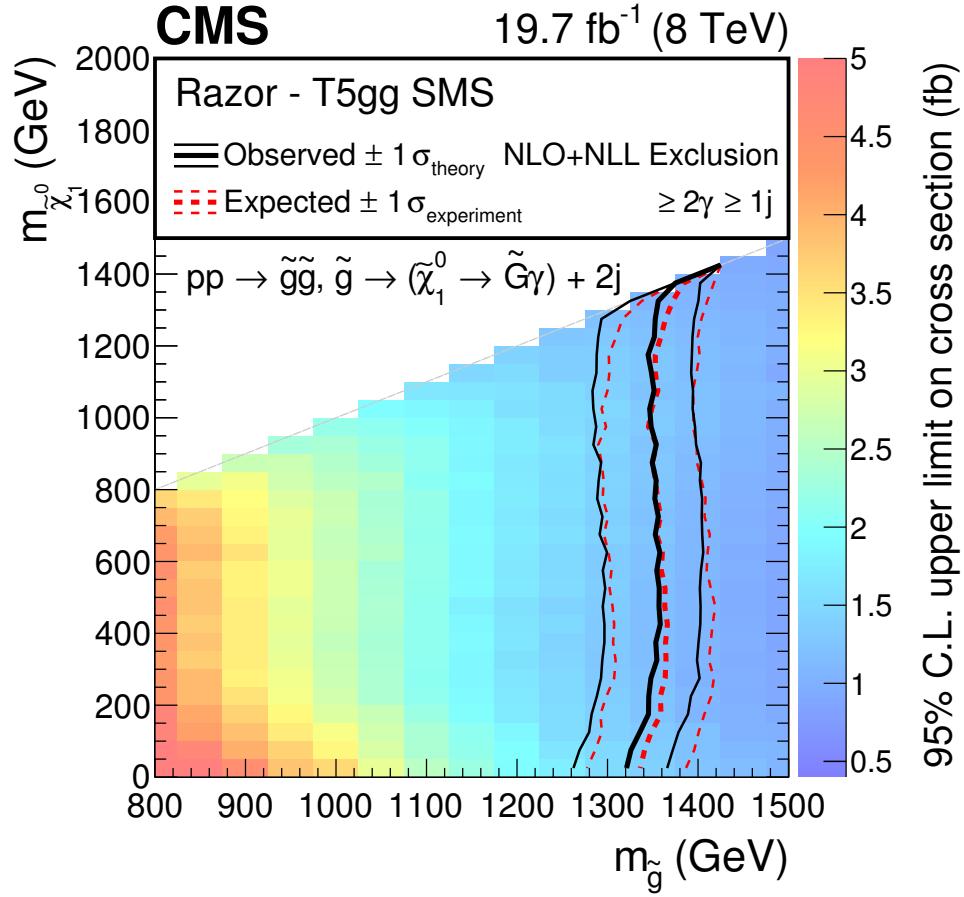
Figure 2.9: Expected and observed limit exclusion contours for T5gg simplified model (discussed in Chapter 8) for gluino-neutralino space. The theoretical uncertainty includes the uncertainty on PDF and NLO renormalization. This is for $\sqrt{s} = 8$ TeV and an integrated luminosity of 19.7 fb$^{-1}$ [35].

# CHAPTER 3

# LARGE HADRON COLLIDER AND CMS DETECTOR

## 3.1   The Large Hadron Collider

The Large Hadron Collider (LHC) is situated just outside of Geneva, Switzerland in a tunnel lying between 45 m and 170 m below the surface of the earth. It has been constructed in the same tunnel where the earlier Large Electron Positron collider was located. The LHC is the largest and most energetic hadron collider to date. The collider ring, split into octants, has eight arcs and eight straight sections (Fig 3.1). The tunnel is approximately 26.7 km in circumference.

### 3.1.1   Performance Goals

The LHC is designed to run at a center of mass energy of $\sqrt{s} = 14$ TeV. The number of events generated per second in collisions is given by:

$$N_{ev} = L\sigma_{ev} \tag{3.1}$$

where $L$ is the luminosity and $\sigma_{ev}$ is the cross section for the process under consideration. The luminosity can be expressed as a function of the number of particles per bunch ($N_b$), the number of bunches per beam ($n_b$), the revolution frequency ($f_r$), the relativistic gamma factor ($\gamma_r$), the normalized transverse beam emittance ($\epsilon_n$) and the beta function at the collision point ($\beta^*$) [36]:

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

In the above expression, $F$ stands for the geometric luminosity reduction factor due to the crossing angle at the interaction point and is given by:

$$F = (1 + (\frac{\theta_c\sigma_z}{2\sigma^*})^2)^{-1/2} \tag{3.3}$$

where $\theta_c$ is the full crossing angle at the interaction point, $\sigma_z$ is the RMS bunch length and $\sigma^*$ is the transverse RMS beam size at the interaction point. Here one assumes $\sigma_z \gg \beta^*$, and equal
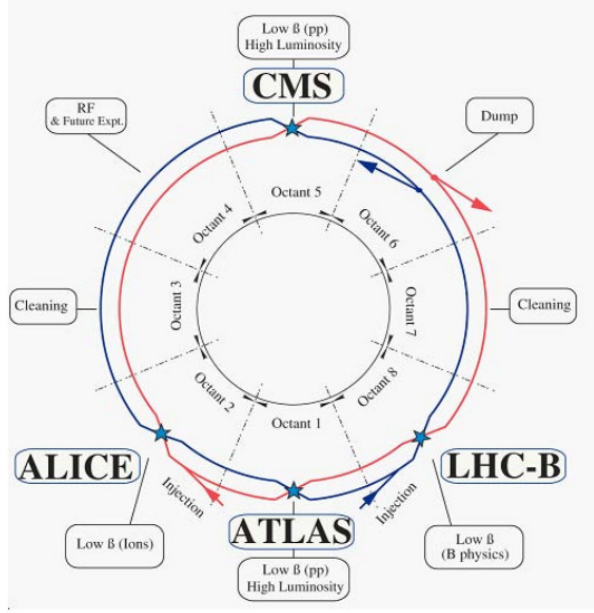
Figure 3.1: Schematic diagram of CERN LHC collider ring split into octants.

beam parameters for both the beams [36]. The assumption behind the expression in Equation 3.2 is that one assumes Gaussian beams. The design instantaneous luminosity at center of mass energy of 14 TeV is $10^{34}$ cm$^{-2}$s$^{-1}$.

The ATLAS and CMS experiments plan for using the peak luminosity of $L = 10^{34}$ cm$^{-2}$s$^{-1}$, whereas LHCb uses a peak luminosity of $L = 10^{32}$ cm$^{-2}$s$^{-1}$ and TOTEM uses $L = 2 \times 10^{29}$ cm$^{-2}$s$^{-1}$ with 156 bunches. The dedicated ion experiment at CERN, ALICE uses peak luminosity of $L = 10^{27}$ cm$^{-2}$s$^{-1}$ for nominal lead-lead operation.

The LHC is designed to be a proton-proton collider where separate magnetic fields and vacuum chambers are used for two colliding beams. The common sections are located only at the interaction regions where the detectors are present [36].

### 3.1.2 Injector Chain

The CERN accelerator complex is not just the collider ring, there are a succession of accelerators with increasingly higher energies. The process starts with stripping hydrogen atoms of their electrons to get the protons for LINAC2, and the protons are brought up to an energy of 0.05 GeV.
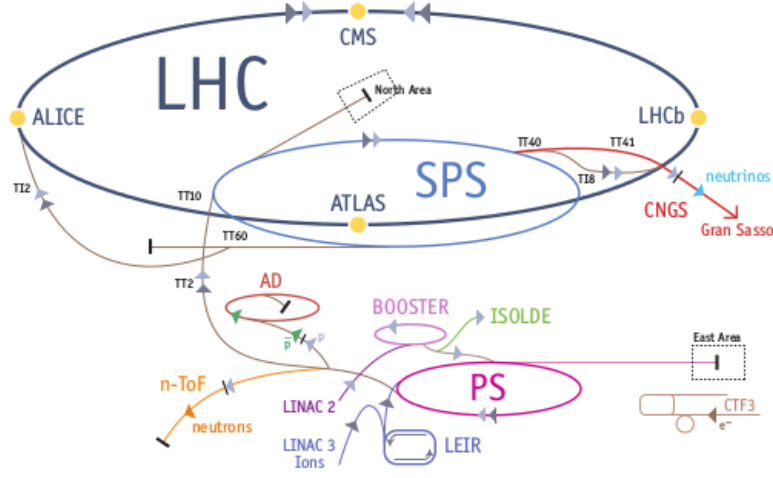
Figure 3.2: Schematic diagram of CERN accelerator complex. The abbreviations used in the diagram are: LHC:Large Hadron Collider, SPS: Super Proton Synchrotron, AD: Antiproton Decelerator, ISOLDE: Isotope Separator OnLine DEvice, PSB: Proton Synchrotron Booster, PS: Proton Synchrotron, LINAC: LINear ACcelerator, LEIR: Low Energy Ion Ring, CNGS: Cern Neutrinos to Gran Sasso.[37].

Then these protons are fed into the Proton Synchrotron Booster which accelerates the protons to an energy of 1.4 GeV. Then they are subsequently injected into Proton Synchrotron (PS). Protons are accelerated to an energy of 25 GeV in the PS and then they are sent to the Super Proton Synchrotron (SPS) which is capable of accelerating protons to 450 GeV. Finally these protons are injected into the main LHC ring in two counter-rotating beams. In the LHC ring, the proton beams are accelerated and the beams are designed to acquire an energy of 7 TeV.

The protons in the LHC are circulated in bunches. Protons can only accelerate when the RF field has the correct orientation. This happens at specified time intervals during an RF cycle. In the PSB, 6 proton bunches are created and injected to the PS where they are divided into 72 bunches. When the protons go into the SPS, between 2 to 4 groups of these 72 bunches are fed. Finally in the LHC, 39 groups of 72 bunches are spaced 25 ns (7 m) apart. That is how proton-proton collisions occur in every 25 ns. The accelerator complex is shown in Fig 3.2.

### 3.1.3 Magnets

There are a large variety of magnets in LHC including dipoles (keeps the particles in nearly circular orbits), quadrupoles (focus the beams), sextupoles, octopoles etc. Each type of magnet optimizes the particle trajectories. There are over 9000 corrector magnets which are embedded in the cold mass of dipoles and quadrupoles[37]. The magnets in LHC have either twin aperture or single aperture. Special magnets like insertion quadrupoles [36] are used to focus the proton beam to the smallest possible size at the collision points which maximizes the probability of collision.

At the LHC, the dipole magnets are superconducting electromagnets and they can provide magnetic field of 8.3 Tesla over their length. They use niobium-titanium (Nb-Ti) cables which are superconducting only below 10 K. The LHC operates at 1.9 K where the dipoles can create 8.3 Tesla of field when $11,850$ A current flows in the dipoles.

### 3.1.4 Vacuum System

The three vacuum systems used by the LHC are:

1. insulation vacuum needed for cryomagnets

2. insulation vacuum required for the helium distribution line

3. vacuum needed for beam.

The insulation vacuum when not at cryogenic temperature can be around 0.1 mbar, but at cryogenic temperature it must be around $10^{-6}$ mbar. The requirement for beam vacuum is much more stricter. For this case, the gas densities required are expressed as normalized to the hydrogen gas density, with proper care taken for different gas ionization cross sections. The equivalent hydrogen gas density should be $10^{15}$ $H_2 m^{-3}$ so that beam lifetime can be up to 100 hours. In the interaction region, the gas density should be $10^{13}$ $H_2 m^{-3}$ and in the room temperature part of the beam vacuum system, the pressure should be $10^{-10}$ to $10^{-11}$ mbar.

The important design parameters for the LHC are tabulated below (Table 3.1).

## 3.2    The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is one of the two multi-purpose detectors at the LHC. It is situated in the village of Cessy, France, around 100 m below the farmland. CMS is built to study

Table 3.1: Important parameters for LHC(Reprinted from [37])

| Parameter | Value |
|---|---|
| Circumference | $26,659m$ |
| Dipole operating temperature | 1.9 K ($-271.3^oC$) |
| Number of magnets | 9593 |
| Number of main dipoles | 1232 |
| Number of main quadrupoles | 392 |
| Number of RF cavities | 8 per beam |
| Nominal energy, protons | 7 TeV |
| Nominal energy, ions | 2.76 TeV/u (Energy per nucleon) |
| Peak magnetic dipole field | 8.33 T |
| Minimum distance between bunches | 7 m |
| Design luminosity | $10^{34}$ cm$^{-2}$ s$^{-1}$ |
| Number of bunches per proton beam | 2808 |
| Number of protons per bunch (at start) | $1.1 \times 10^{11}$ |
| Number of turns per second | $11,245$ |
| Number of collisions per second | 600 million |

a wide range of phenomena and particles, both SM and BSM. It also evaluates the validity of SM at extremely high energies.

### 3.2.1 Overview

The schematic diagram of the CMS detector is shown in Figure 3.3. The most important feature of the CMS detector is the large superconducting solenoid which provides a magnetic field of 3.8 Tesla. The solenoid is 13 m long and 6 m in inner diameter. The strong magnetic field is capable of bending the charged particles produced during the collisions. The four muon stations are integrated in the 1.5 m of iron kept in the return field of the solenoid. This ensures robustness and full geometric coverage. The bore of the solenoid houses the tracker and calorimeters. outside of the solenoid.

### 3.2.2 Coordinate System

The CMS coordinate system is a cartesian coordinate system with the origin located at the center of the CMS detector, which is the nominal beam collision spot. The X-axis points towards the center of the LHC ring and Y-axis is directed vertically upwards. The Z-axis direction is taken along the counter-clockwise beam direction in the LHC. Spherical polar coordinates are also used

CMS DETECTOR

Total weight : 14,000 tonnes
Overall diameter : 15.0 m
Overall length : 28.7 m
Magnetic field : 3.8 T

STEEL RETURN YOKE
12,500 tonnes

SILICON TRACKERS
Pixel (100x150 μm) ~16m² ~66M channels
Microstrips (80x180 μm) ~200m² ~9.6M channels

SUPERCONDUCTING SOLENOID
Niobium titanium coil carrying ~18,000A

MUON CHAMBERS
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
Endcaps: 468 Cathode Strip, 432 Resistive Plate Chambers

PRESHOWER
Silicon strips ~16m² ~137,000 channels

FORWARD CALORIMETER
Steel + Quartz fibres ~2,000 Channels

CRYSTAL
ELECTROMAGNETIC
CALORIMETER (ECAL)
~76,000 scintillating PbWO$_4$ crystals

HADRON CALORIMETER (HCAL)
Brass + Plastic scintillator ~7,000 channels

Figure 3.3: Schematic diagram of CMS detector[38].

26

where the azimuthal angle $\phi$ is measured in X-Y plane from the X-axis. The polar angle $\theta$ is measured from the positive Z-axis. There is another variable called 'pseudorapidity' which is also used in place of polar angle $\theta$. The true rapidity of a particle is given by:

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \tag{3.4}$$

where $E$ and $p_z$ are the energy and momentum in the Z direction respectively. For energetic particles where the mass can be neglected, $p_z$ can be approximated as $E\cos\theta$. Then from Eq 3.4, it can be written as:

$$y \approx \frac{1}{2}\ln\left(\frac{E + E\cos\theta}{E - E\cos\theta}\right) = \frac{1}{2}\ln\left(\frac{1 + \cos\theta}{1 - \cos\theta}\right) = \frac{1}{2}\ln(\cot^2\theta/2) = -\ln(\tan\theta/2) \tag{3.5}$$

Therefore the pseudorapidity is defined as $\eta = -\ln(\tan(\theta/2))$. The $\eta - \phi$ coordinates are adopted because QCD particle production in hadron-hadron collision is approximately constant per unit $\eta$ within the CMS detector. Also the rapidity difference is Lorentz invariant under boosts along the beam direction.

### 3.2.3 Tracker

The CMS tracker is the innermost particle sub-detector and is used to measure the position and trajectory of charged particle to extremely high precision ($\sim 10$ $\mu$m). The tracker allows for the reconstruction of secondary vertices. The tracker volume surrounds the interaction point up to $|\eta| < 2.4$. The length and diameter of the tracker are 5.8 m and 2.5 m respectively [39]. The schematic diagram of the CMS tracker is shown in Figure 3.4. Considering the charged particle flux at various radii at high luminosity, 3 regions can be separated out [40]:

- The particle flux is the highest ($10^7/s$ at radial distance from the interaction point, $r \sim 10$ cm) closest to the interaction vertex. Pixel detectors are placed in this region. The dimensions of a single pixel are roughly $100 \times 150$ $\mu m^2$ which gives an occupancy of about $10^{-4}$ per pixel per LHC crossing at the LHC design luminosity of $10^{34}$ $cm^2$ $s^{-1}$.

- In the region 20 $cm < r < 55$ cm, the particle flux is low enough to use silicon microstrip detectors with a typical size of 10 $cm \times 80$ $\mu m$. This leads to an occupancy of up to 2-3% per strip and per crossing at the LHC design luminosity.

- For the outermost region of the inner tracker, $r > 55$ cm, the particle flux drops sufficiently so that larger-pitch silicon microstrips can be used. The size of silicon microstrips are up to 25 $cm \times 180$ $\mu m$ and the occupancy is $\sim 1\%$.

27
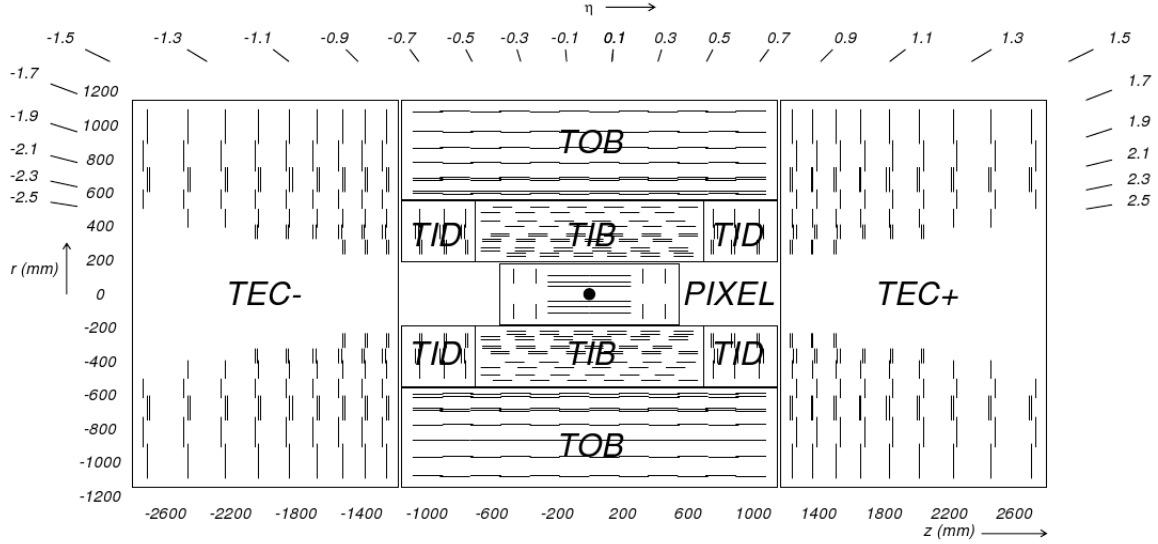
Figure 3.4: Schematic diagram showing the cross section through CMS tracker. Here each line represents a detector module. Each double-line represents back-to-back modules delivering stereo hits.

In the heavy-ion (Pb-Pb) running, the occupancy is expected to be around 1% in the pixel detector. For the outer silicon strip detectors, the occupancy is less than 20%. This permits track reconstruction in the high density environment[40].

In the barrel region, there are 3 layers of hybrid pixel detectors which have radii of 4.4, 7.3 and 10.2 cm respectively. The pixel detector delivers three high precision space points for each charged particle trajectory. The area covered by the pixel detector is $\sim 1\ m^2$ and it has 66 million pixels [40].

The radial region $20\ cm < r < 116\ cm$ is occupied by the silicon strip tracker. The Tracker Inner Barrel and Disks (TIB/TID) extend in radius towards 55 cm. These are composed of 4 layers in the barrel and 3 disks at each end. TIB/TID are capable of delivering up to $4\ r-\phi$ measurements on a trajectory using 320 $\mu m$ thick silicon micro-strip sensors. The strip pitch is 80 $\mu m$ on layers 1 and 2, while for layers 3-4, this value is 120 $\mu m$. This enables a single point resolution of 23 $\mu m$ and 35 $\mu m$ respectively. The TIB/TID is surrounded by Tracker Outer Barrel (TOB). The TOB has outer radius of 116 cm and it consists of 6 barrel layers. Each layer has a microstrip sensor of 500 $\mu m$ thickness. The strip pitches for first 4 layers are 183 $\mu m$ and 122 $\mu m$ for layers 5 and 6. It

28

provides another 6 $r - \phi$ measurements with resolution of 53 $\mu m$ and 35 $\mu m$ respectively.

The TOB coverage is up to $|z| < 118$ cm. Beyond this range, the Tracker EndCaps (TEC) goes from 124 $cm < |z| < 282$ cm and 22.5 $cm < |r| < 113.5$ cm. There are 9 disks in each TEC and each disk carries up to 7 rings of silicon microstrip detectors. The rings have radial strips of 97 $\mu m$ to 184 $\mu m$ average pitch. They provide up to 9 $\phi$ measurements per trajectory.

A second microstrip detector module is added to the first two layers and rings of TIB, TID and TOB as well as rings 1,2 and 5 of the TECs. These are mounted back-to-back with a stereo angle of 100 mrad in order to provide a measurement of the second coordinate (i.e. z in the barrel and r in the disks). The single point resolution of this measurement is 230 $\mu m$ in TIB and 530 $\mu m$ in TOB. The total area of pixel detector is $\approx$ 1 $m^2$ while that of silicon strip detectors is 198 $m^2$.

### 3.2.4 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) of CMS is a hermetic, homogeneous calorimeter. It is made of 61,200 scintillating lead tungstate ($PbWO_4$) crystals mounted in the central barrel part. Each endcap is made of 7,324 crystals. A preshower detector is also present in front of the endcap crystals. The $PbWO_4$ crystals are specifically designed to measure the energy of electromagnetic particles. The primary mechanism by which photons and electrons interact with the crystals are Bremsstrahlung and pair production. When an electron enters the crystal, it will interact with the positively charged nucleus and eventually slow down by emitting photon(s). An electron will continue releasing photons along its way until it has exhausted all its energy. Similarly, when a photon enters the crystal, it will undergo pair-production and create an electron - positron pair. Then this electron and positron will emit photons along their paths as explained earlier. This process is known as showering. The Feynman diagrams showing Bremsstrahlung and pair production are shown in Figure 3.5.

When a high energetic particle enters the crystal, it radiates energy in form of electron, positron or photon. The atoms in the crystals absorb this energy and become excited. The atoms quickly go back to ground state by emitting photons, specifically the blue-green light (with a broad maximum at 420-430 nm wavelength). The crystals are designed to collect this light and deliver it to the photo-sensors mounted (avalanche photodiodes, APDs in the barrel and vacuum phototriodes, VPTs in the endcap) at the back of the crystals. These photo-sensors produce current by the Photo-Electric effect, which is in proportion to the energy deposited.
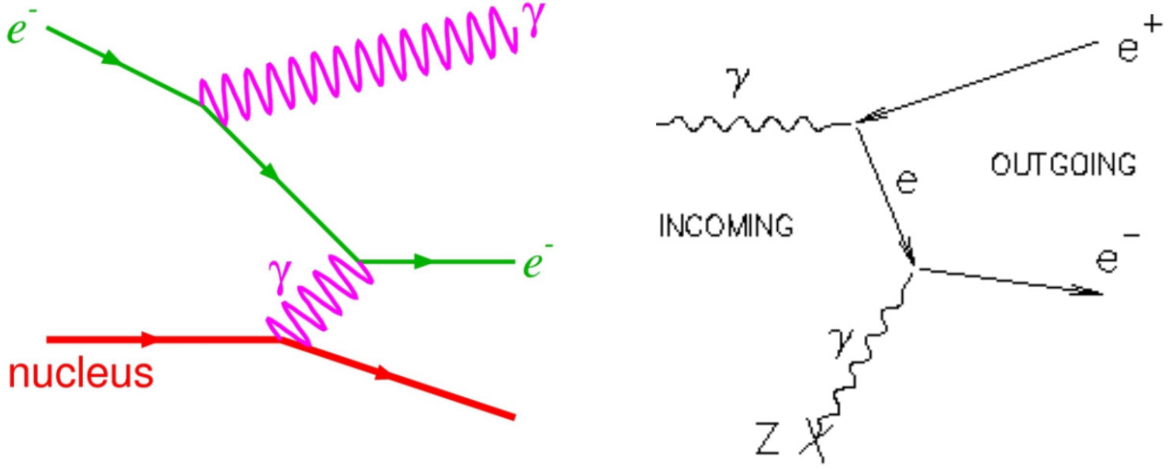
Figure 3.5: Feynman diagrams showing Bremsstrahlung (left) and photon pair producing electron-positron (right).

**Lead tungstate crystals:** Lead tungstate crystals are an appropriate choice for the electromagnetic calorimeter because of their high density (8.3 g/cm$^3$), short radiation length (0.89 cm) and small Moliere radius (defined as the radius of the cylinder which contains 90% of an electromagnetic shower, 2.2 cm). All these things aid in making a granular and compact calorimeter. The scintillation decay time of these crystals is also of the same order of magnitude as the LHC bunch crossing, 80% of the light emitted by the crystals will occur within 25 ns. About 4.5 photoelectrons per MeV are collected by both APDs and VPTs at a temperature of 18℃. Pictures of PbWO$_4$ crystals for barrel and endcap with photo-sensors attached are shown in Figure 3.6 [39].

**The ECAL layout:** The ECAL cross section is shown in Figure 3.7. The granularity of the barrel (EB) part of the ECAL, covering $|\eta| < 1.479$, is 360-fold in $\phi$ and (2×85)-fold in $\eta$ [39]. The crystals are mounted in a quasi-projective geometry, meaning their axes make a small angle (3°) with respect to the vector from the nominal interaction vertex, both in $\eta$ and $\phi$ direction. The crystal cross section in $\eta - \phi$ space is approximately 0.0174×0.0174. This corresponds to 22×22 mm$^2$ at the front face of the crystal and 26×26 mm$^2$ at the rear face, corresponding to a truncated pyramid. The crystal length is 25.8$X_0$ (230 mm), where $X_0$ is the radiation length of PbWO$_4$. The barrel crystal weight is 67.4 t and volume 8.14 m$^3$ [39].

30
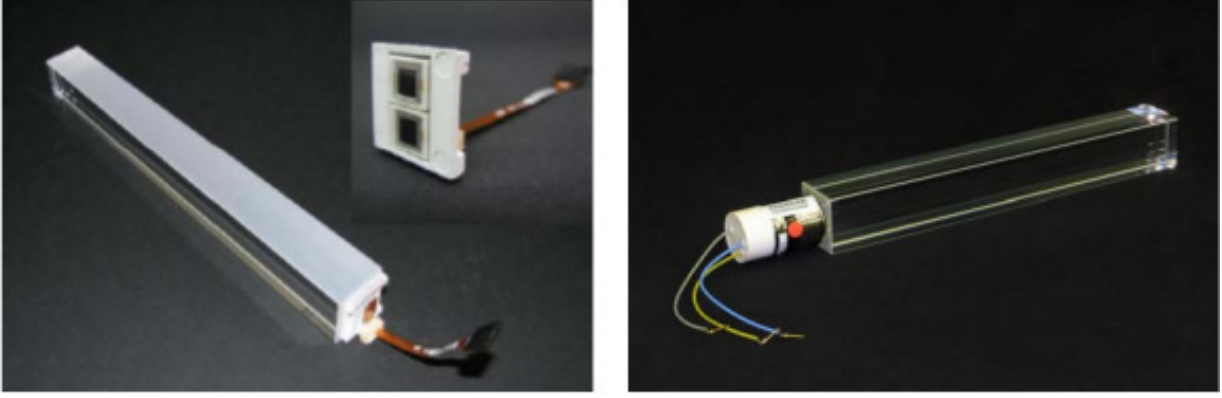
Figure 3.6: PbWO$_4$ crystals with photo-sensors attached. The barrel crystal with one face depolished attached with the APD capsule is shown on the left. The endcap crystal with VPT attached is shown on the right [39].

The endcaps, covering the region $1.479 < \eta < 3.0$ on each side, consist of identically shaped crystals grouped in mechanical units of 5×5 crystals known as supercrystals (SCs). The crystals and SCs are arranged in a Cartesian $X - Y$ grid, where the crystals are pointing at a focus which is 1.3 m beyond the interaction point. This gives an off-pointing angle of $2 - 8°$. The crystal front face cross section is 28.62×28.62 mm$^2$ and rear face cross section 30×30 mm$^2$. They have a length of 220 mm (24.7 X$_0$). The endcap crystal volume and weight are 2.90 m$^3$ and 24 t respectively[39].

**Preshower detector:** The principal goal of the CMS Preshower (ES) detector, covering the region $1.653 < |\eta| < 2.6$, is to provide better position resolution of photons. Often a neutral pion can decay to two photons and in the EE region, the angle between them can be small. In that case, sometimes the two photons are reconstructed as a single energetic photon. The ES detector provides sufficient granularity so that two photons can be resolved lowering the mis-identification rate. The search for the Higgs boson where $H \rightarrow \gamma\gamma$ was the motivating factor for this design.

The ES has two layers: first the lead radiators, which initiate electromagnetic showers from the incoming electrons or photons (showers can also be initiated inside the tracker material), and then the silicon strip sensors placed after each radiator. The silicon strip sensors measure the deposited energy and the transverse shower profiles. The material thickness of the ES at $\eta = 1.653$ before reaching the first sensor is 2X$_0$. The material thickness is 1X$_0$ further before reaching the second sensor plane. Thus 95% of the photons start showering before reaching the second plane. The 2
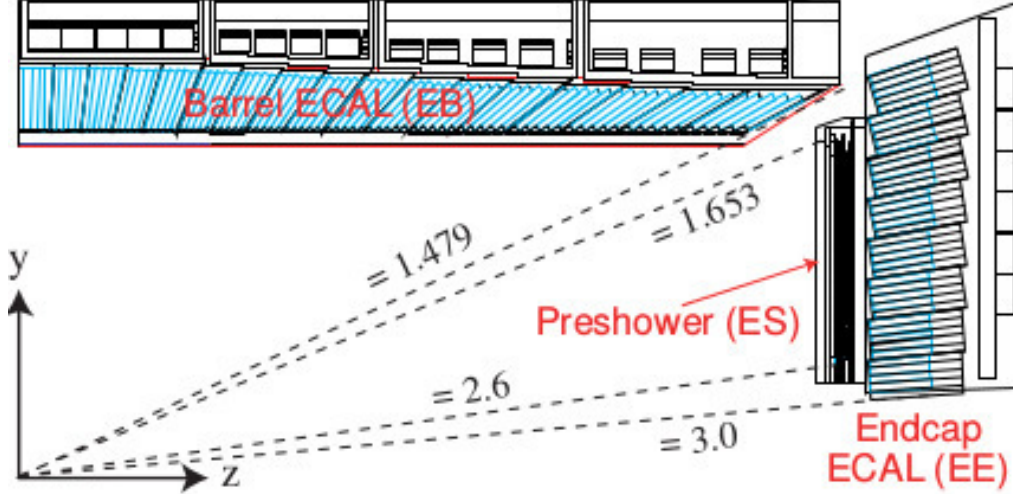
Figure 3.7: Transverse section of ECAL detector showing the layout [40].

sensor planes are oriented orthogonally. Each silicon sensor measures an area of $63 \times 63$ mm$^2$ where $61 \times 61$ mm$^2$ is the active area. This active area is further divided into 32 strips (pitch length is 1.9 mm). The silicon thickness (nominal) is 320 $\mu m$.

**Performance of the ECAL:** The energy resolution of the ECAL is measured by fitting a Gaussian function to the reconstructed energy distributions. This has been parameterized as a function of energy, $E$ [39]:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{3.6}$$

where S is the stochastic term, N is the noise term and C is the constant term. This relationship between energy resolution and energy is pretty generic for calorimeters. The stochastic term includes statistical fluctuations in lateral shower containment and fluctuations in the energy deposited in the preshower absorber (if present) with respect to the energy deposition measured by the preshower silicon detector. The noise term comes from three contributions: electronics noise, digitization noise and pileup fluctuations. Lastly, non-uniformity of the longitudinal light collection, leakage of energy from the back of crystals and intercalibration uncertainties contribute to the constant term. From the fully equipped barrel supermodule which was tested in the CERN H4 beam (2004), a typical energy resolution gave S= 2.8%, N= 0.12 and C= 0.30% where E is in GeV. The actual performance of ECAL is worse than these values, because the test beam did not consider conversion

32

of photons, bremsstrahlung from electrons, material energy loss and pileup which are inevitable in actual experimental environment [41].

**Calibration of ECAL:** The goal of calibration is to measure the energies of photons and electrons as accurately as possible. Many effects like detector non-uniformity and crystal to crystal variation in scintillation light should account for in the ECAL calibration. ECAL calibration is composed of two components: the global component giving rise to absolute energy scale and the intercalibration, which takes care of channel-to-channel differences. Preliminary estimates of the intercalibration coefficients are obtained from laboratory measurements of the light yield of different crystals. Here photodetector and electronic responses are also taken into account. The ultimate intercalibration precision is achieved with physics events and all other sub-detectors in their places (in-situ). Since the energy deposition in the ECAL should be $\phi$-independent, intercalibration coefficients are adjusted within fixed $\eta$ regions to remove $\phi$-dependence. Apart from the effect of detector material, there should be calibrations on electromagnetic objects due to difference in reconstruction algorithms. These calibrations are done with simulated events by comparing the reconstructed and generated parameters of the objects [39].

**Laser monitoring system:** Although radiation resistant, due to the high particle flux and radiation level of LHC, ECAL PbWO$_4$ crystals show a rapid loss of optical transmission from the formation of color centers. This is due to impurities and oxygen vacancies within the crystals. These color centers absorb a fraction of the transmitted light.

A laser monitoring system is used to measure the evolution of the crystal transparency. The laser pulses are injected into the crystals using optical fibers. The response is normalized by the laser pulse magnitude measured by a silicon PN diode. The $R(t) = APD(t)/PN(t)$, i.e. the ratio of the response of the APD to that of the PN diode is used as a measure of the crystal transparency. The injected laser pulse and the scintillating light are different because of the different optical paths and spectra. So the change in response to laser light does not necessarily mean same change in response in the scintillating light. For attenuations less than 10%, the relationship between the changes can be shown to follow a power law [39]:

$$\frac{S(t)}{S(t_0)} = \left[ \frac{R(t)}{R(t_0)} \right]^{\alpha} \tag{3.7}$$

where $S(t)$ represents the response coming from scintillation light and $\alpha$, the characteristics of the crystal, depends on the production method of the crystal. This power law describes the behavior

of all the crystals evaluated at test beams and is expected to be valid in the barrel, both with low and high luminosity at the LHC.

### 3.2.5 Hadronic Calorimeter

The hadronic calorimeter (HCAL) in addition with the ECAL subdetectors form a complete calorimetry system for the measurement of the missing transverse energy and jets (described in the next chapter). The central barrel (HB) and endcap (HE) completely surround the ECAL and are situated fully inside the magnetic field of the solenoid. The barrel, covering the region $|\eta| < 1.3$, and endcaps in the region $1.3 < |\eta| < 3.0$ are joined such that this calorimeter becomes hermetic. There is a forward calorimeter (HF) located 11.2 m from the interaction point, which covers the pseudo-rapidity region of $3.0 < |\eta| < 5.0$. The forward calorimeters are designed to measure the highly energetic forward jets. This increases the hermeticity of the missing transverse energy measurement. There is an array of scintillators located outside the magnet in the outer barrel hadronic calorimeter (HO) which improves the central shower containment in the region $|\eta| < 1.26$. Figure 3.8 shows a longitudinal view of the HCAL.

**HCAL design:** Unlike the homogeneous crystals of the ECAL, the HCAL is a sampling calorimeter, meaning it has layers of different materials: brass and steel plates in the barrel, endcap and forward region to induce hadronic interactions and plastic scintillators which produce scintillating light upon interaction with charged particles. The barrel hadron calorimeter is composed of two half barrels, each of them containing 18 identical wedges each covering 20°in $\phi$. The flat brass alloy (70% Copper and 30% Zinc) absorber plates in the wedge are parallel to the beam axis. The innermost and outermost layers are made of stainless steel which give strength to the HCAL structure. Between the stainless steel and brass absorber plates, 17 active plastic scintillator tiles are interspersed. The longitudinal profile in the barrel going from inner radius of 1.777 m to 2.8765 m is as follows [40]:

- Layer 0: 9 mm scintillator, 61 mm stainless steel.

- Layer 1-8: 3.7 mm scintillator, 50.5 mm brass.

- Layer 9-14: 3.7 mm scintillator, 56.5 mm brass.

- Layer 15 and 16: 3.7 mm scintillator, 75 mm stainless steel, 9 mm scintillator again.
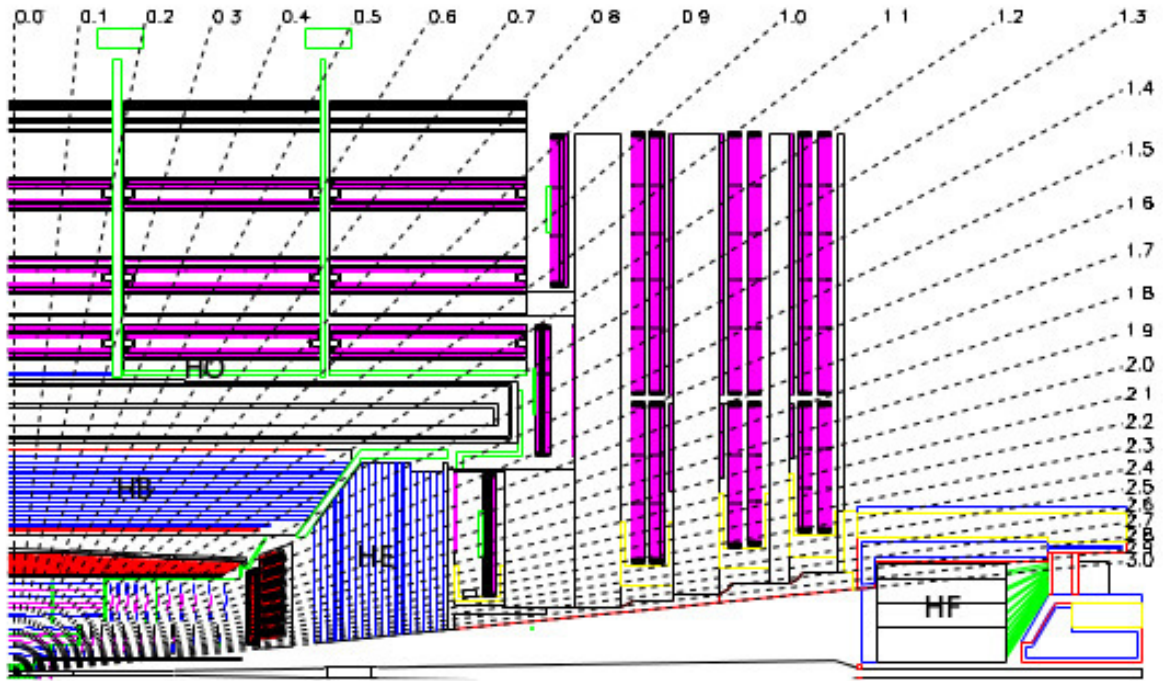
Figure 3.8: Longitudinal view of HCAL. Here the positions of hadron barrel (HB), endcap (HE), outer (HO) and forward calorimeters are shown.

Here layer number refers to the active scintillator layer only. Layer 0 is situated directly behind the ECAL and the scintillator thickness is roughly twice that of other layers. This is to actively gauge the energy loss due to the low energy particle showering in the support material in between ECAL and HCAL.

The individual scintillator tiles have dimensions of 0.087×0.087 in $\eta - \phi$ space. They are each instrumented with a single wavelength shifting fiber (WLS). The WLS fibers are connected to clear fibers and the clear fibers are optically added to corresponding tiles from each of the 17 active layers. In this way they form 32 barrel HCAL "towers" in $\eta$. Towers 15 and 16 are exceptions because they are located at the edge of the HB half-barrel; where multiple optional readouts are present. A pixelated hybrid photodiode (HPD) mounted at the ends of barrel mechanical structure is used to detect the optical signal from the HCAL towers.

The HE also has 18-fold $\phi$-geometry like the barrel calorimeter. But the absorber plates of HE are composed entirely of brass. There are 19 active plastic scintillator layers having thickness of 3.7 mm. The thickness of the absorber plates are 78 mm. In the region $|\eta| > 1.74$, the $\phi$-granularity of the tiles is reduced to 10°. This is to accommodate the bending radius of the WLS fiber readout. The granularity of the calorimeter in the region $|\eta| < 1.6$ is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ and in the region $|\eta| \geq 1.6$ $\Delta\eta \times \Delta\phi \approx 0.17 \times 0.17$.

Steel absorber is used in the HF. Embedded radiation hard quartz fibers provide for fast collection of Cherenkov light. Each HF module is composed of 18 wedges which are kept in a non projective geometry with the quartz fibers running parallel to the beam axis. Long (1.65 m) and short (1.43 m) quartz fibers, placed alternately with a separation of 5 mm, are bundled at the back of the detector and connected to the photo-multiplier tubes.

**HCAL calibration:** During its operation, HCAL is also calibrated regularly using various techniques. The major methods of HCAL calibration are [40]:

- Test beam: When CMS experiment is shut down, a few wedges are taken to the test beam facility and bombarded with particles with known energies. The response is adjusted to match the energy of incident particles.

- Using a radioactive source: The megatile scanner and the wire system contain radioactive Cesium-137 sources. The megatile scanner delivers known amounts of gamma ray radiation and the phototubes read in this energy. Thus the responses of phototubes can be adjusted.

- Using UV laser light: A known amount of light is injected through a single fiber into the megatiles and the corresponding light collected in the photo-sensors are measured. Using this information, the HCAL response is adjusted.

- From data: Reconstructed jets from collision and cosmic muons are also used to calibrate the HCAL [42].

**HCAL performance:** The energy resolution of the HCAL suffers from the presence of very dense detector material of the ECAL. Due to this, hadronic showers can start developing at the ECAL, rather than in the HCAL alone. Therefore, the energy resolution of the HCAL is measured in combination with the ECAL positioned in front of it. The energy resolution can be parameterized by [43] :

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{85\%}{\sqrt{E}}\right)^2 + (7.4\%)^2, \qquad |\eta| < 3.0 \tag{3.8}$$

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{198\%}{\sqrt{E}}\right)^2 + (9.0\%)^2, \qquad 3.0 < |\eta| < 5.0 \tag{3.9}$$

In the above equations, the first terms are the stochastic term and the second terms correspond to the detector non-uniformity and calibration uncertainty.

### 3.2.6 Superconducting Solenoid

The superconducting solenoid is situated radially outward from the hadronic calorimeter. The parameters of the superconducting solenoid is given in Table 3.2.

Table 3.2: Important parameters of CMS superconducting solenoid [40]

| Field | 4T |
|---|---|
| Inner Bore | 5.9 m |
| Length | 12.9 m |
| Number of Turns | 2168 |
| Current | 19.5 kA |
| Stored energy | 2.7 GJ |
| Hoop stress | 64 atm |
| Weight | 12,000t |

The magnetic field generated by the solenoid bends the trajectories of the charged particles. The curved trajectories help the tracker to adequately measure transverse momenta of charged particles. The diameter of the solenoid is 12.9 m which is adequate to fit the entire tracking system

and calorimetry within. This is beneficial because this provides better resolution for measuring energy and transverse momenta of charged particles. The magnetic flux generated inside of the solenoid is returned through a 1.5m thick, 10,000t return yoke. The return yoke is embedded in the muon detection system. The incorporation of return yoke with the muon system concentrates the flux lines within the muon system. This provides a second handle in measuring the momenta of muons.

### 3.2.7   Muon System

At CMS, the muons which are produced centrally are not stopped by inner sub-detectors due to their high penetrative power. Even the muon system does not stop them, rather the muon system is a type of tracking system which pinpoints the positions of muons through 4 concentric layers in the barrel region and 3 layers (and 2 forward layers) in each endcap.

Three types of gaseous particle detectors are used to identify muons and measure their transverse momenta. The detector technology was selected depending upon the radiation environment and the large surface area to be covered. In the barrel region ($\eta < 1.2$), the neutron induced background is small and the muon rate is low. The residual magnetic field in the chambers is also low. So here drift tube (DT) chambers are used. In the both endcaps, the neutron induced background as well as muon rate is very high. Moreover, the magnetic field is also high there. So cathode strip chambers (CSCs) are used to cover up to $\eta < 2.4$. In addition to this, resistive plate chambers (RPCs) are used to cover both the barrel and endcap region. The total numbers of DTs, CSCs and RPCs are listed in Table 3.3.

Table 3.3: Different components of muon system

| Region | Muon chambers | Numbers |
|---|---|---|
| $|\eta| < 1.2$ | drift tubes | 250 |
| $0.9 < |\eta| < 2.4$ | cathode strip chambers | 540 |
| $|\eta| < 1.6$ | resistive plate chambers | 610 |

Each drift tube is a 4 cm wide tube of gas containing a wire held at positive voltage within it. When the charged particles pass through the gas, the gas molecules are ionized and the electrons are attracted towards the positive wire. By measuring the drift time and which wires register the hit, the charged particles's position and trajectories are pin-pointed.

Each CSC consists of wires held at either positive or negative voltages which are orthogonal forming a grid. This grid is confined in a gas tube. Similar to the drift tubes, when charged particles pass through the chambers, they ionize the gas molecules and the electrons are attracted towards the positive wires. This gives two position coordinates for each strip, which will be eventually connected to give a trajectory.

The RPCs are gaseous parallel-plate detectors. They combine adequate spatial resolution with the time resolution comparable to that of scintillators. RPCs are capable of tagging the time of an ionising event in much shorter time than 25 ns, which is the time between two consecutive bunch-crossing (BX) at the LHC. Therefore the muon trigger based on RPCs can identify unambiguously the relevant BXs to which a muon track is associated even in the presence of high rate and high background.

### 3.2.8   Triggers

The Large Hadron Collider is designed such that about one billion proton-proton collisions occur in CMS detector per second when LHC is running at its peak performance with bunch crossing rate of 40 MHz [40]. It is not possible to record each and every collision occurring at this rate. Moreover, these events are largely dominated by QCD processes which are not interesting for the physics program of the CMS detector. The vast majority of these processes are not hard scattering (head-on collision). The probability of finding new physics in hard scattering is higher as the momentum transfer between partons are higher here compared to other soft-scattering. So uninteresting collisions must be rejected.

The trigger system in CMS consists of two levels of trigger architecture: The Level 1 (L1) trigger and the High Level Trigger (HLT).

**Level** 1 **Trigger:** The first level of trigger, L1 trigger operates using the information from the muon chambers and calorimeter. This trigger is a series of electronics associated with the muon and calorimeter systems that are designed to operate very fast and reduce the event rate from 1 GHz to  100 kHz. The L1 trigger reduces the event rate by looking at 'interesting' events, i.e. events which possibly indicate new physics.

Bunch crossings occur at CMS every 25 ns. To cope with this high event rate, the L1 trigger must be extremely fast. Also often the detectors are still filled with particles coming from one bunch crossing when the next bunch crossing takes place. This rate is too high for L1 trigger to
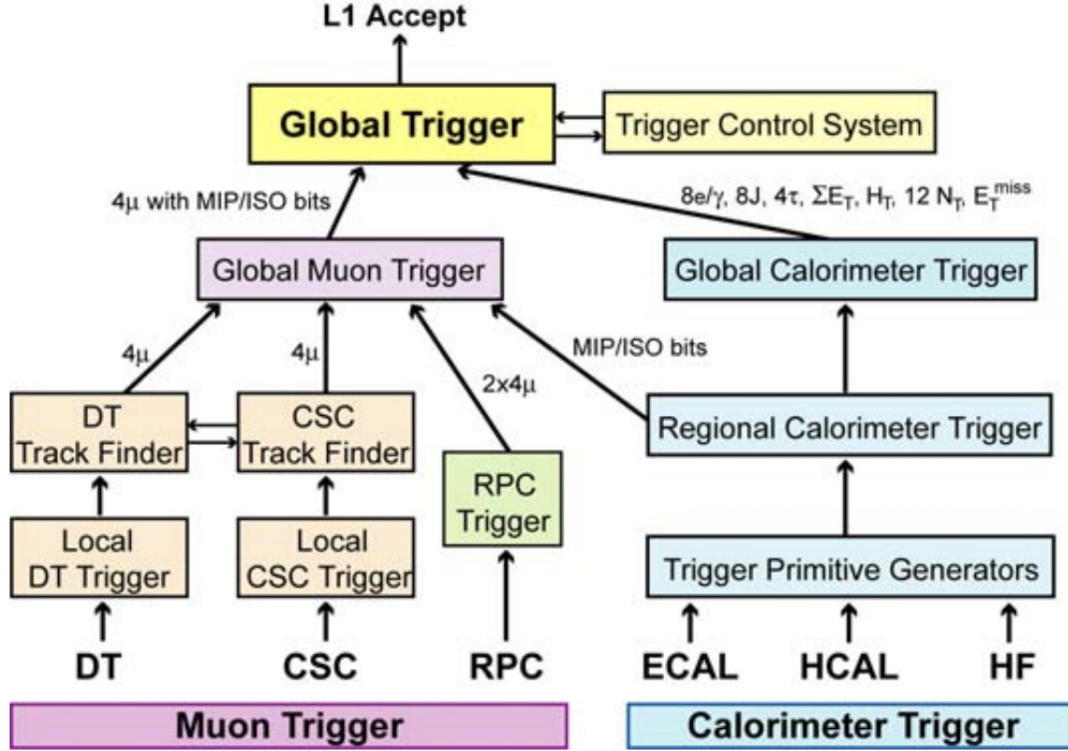
Figure 3.9: The architecture of Level 1 trigger. Reprinted from [39].

access all the information from one bunch crossing. This trigger algorithm processes data in the 'pipeline' using pattern recognition and fast summing techniques, which do not introduce dead time.

The CMS L1 trigger consists of local, regional and global trigger systems associated with the muon systems and calorimeters. The Trigger Primitive Generators (TPG), the Local Triggers, work based on energy deposits in calorimeters and hit patterns and track segments in muon chambers [39].

The TPG is the first step of the Calorimeter Trigger. The calorimeters are divided into trigger towers for triggering purposes. The TPGs measure the total transverse energy deposited in ECAL crystals and HCAL towers and then trigger tower transverse energies are obtained. In the next step, the TPGs are transmitted to the Regional Calorimeter Trigger (RCT) using high-speed serial links. The job of RCTs is to determine photon/electron objects and transverse energy sums from

calorimeters. The next step after RCT is the Global Calorimeter Trigger (GCT) where total transverse energy, missing transverse energy (described in Section 4.7), jet multiplicity etc. are determined. Here $H_T$s, the scalar sum of transverse energy of all jets above some threshold, are also calculated [39].

For the muon trigger, all three detectors in the muon systems (the DTs, the CSCs and the RPCs) are used. The chambers determine the bunch crossing from which muon objects originate. The DT in the barrel provides the track segments of the muon objects which can be treated as local trigger information. Similarly, the CSC in the endcap provides three dimensional track segments. Then DT and CSC link segments to complete tracks and assign physical parameters like track momenta out of the linked tracks. This track finding procedure is done by Regional Muon Trigger. Since RPCs have excellent timing resolution, they provide their own tracks based on regional hit patterns. The Global Muon Trigger (GMT) utilizes the information collected by the three muon systems and achieves an improved momentum resolution and efficiency compared to the individual detectors [39].

The trigger objects from the GCT and GMT finally go to the Global Trigger (GT). This information contains electrons, photons, muons, tau, hadronic jets, total and missing transverse energies. Here the Level 1 Accept decision is made. The whole process takes 144 bunch crossings, i.e. 3.6 $\mu$s [39]. The architecture of Level 1 trigger is shown in Figure 3.9.

**High Level Trigger:** Only events which fire the L1 seeds are passed to the High Level Trigger (HLT). The HLT runs on a large computer farms of fast commercial processors including 13,000 CPU cores. The HLT algorithm must accept a large amount of physics signals, but still the output rate and CPU time should be reasonably small. To achieve this, the HLT algorithms access the data from all the CMS sub-detectors, even from the tracker. As a result, at the HLT level, physics objects are selected and reconstructed in a manner which is as close as possible to the standard offline reconstruction software. The online version differs from the offline version only by different parameter configurations. For example, the tracks from pixel and strip detectors are combined with the energy deposits in the calorimeter to define the electron candidates. Similarly tau candidates found in the calorimeter are combined with the high transverse momenta stubs in the tracker which form the hadronic tau trigger.

The trigger paths, having a modular structure, are sequences of reconstruction and filtering blocks which are arranged according to increasing complexity. Faster algorithms run first and the products from them are filtered. If a filter fails, the rest of the trigger paths are skipped and the event is rejected. The detector read-out and reconstruction are nominally only done around narrow regions of L1 and higher-level objects to reduce the computing time [44].

The output from the HLT has far superior granularity, resolution and purity. The maximum input rate that the HLT can handle is 100 kHz and the output rate is $\sim 400$ Hz.

# CHAPTER 4

# OBJECT RECONSTRUCTION

The highly energetic particles interact with various sub-detectors at CMS and deposit energy observed as hits in the tracking system or energy depositions in the calorimeter. These raw signals are sufficient for triggering, but for analysis, physics object candidates need to be reconstructed from them. At CMS reconstructed objects can be broadly categorized as photons, electrons, neutral hadrons, charged hadrons and muons. Each object is reconstructed differently and may depend on information from several sub-detectors.

## 4.1   Track and Vertex Reconstruction

The reconstruction of vertices and tracks in the CMS detector is very important for the reconstruction of all objects, particularly for charged particles. The presence of a high magnetic field and sophisticated tracking algorithms used at CMS ensure that the tracks and vertices can be reconstructed with great accuracy.

A particle of charge q traveling through magnetic field $\vec{B}$ will feel the force:

$$F = q\vec{v} \times \vec{B} \qquad (4.1)$$

where $\vec{v}$ is the particle's velocity. If the charged particle is moving around the direction of $\vec{B}$ in a circle of radius r, then $\vec{v}\vec{B}$ becomes $v_T B$ where $v_T$ is the velocity component perpendicular to the magnetic field. From the circular motion of the particle, the expression for force, F is given by

$$F = \gamma m v_T^2 / r \qquad (4.2)$$

where $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$. Comparing Equations 4.1 and 4.2, the relativistic transverse momentum $p_T$ of the particle is given by:

$$p_T = \gamma m v_T = qrB \qquad (4.3)$$

. Using Equation 4.3, charged particle transverse momenta can be estimated provided its charge, track curvature, and magnetic field are known.

### 4.1.1 Track Reconstruction

Reconstructing tracks of charged particles requires a good deal of computing power. The tracking algorithm used for reconstruction at CMS is known as the Combinatorial Track Finder (CTF) [45] which is an adaptation of the combinatorial Kalman filter [46]. This is an extension of the Kalman filter which allows pattern recognition and track fitting in the same framework. The collection of reconstructed tracks is produced in a process known as *iterative tracking*, the name coming from multiple passes (iterations) of the CTF track reconstruction sequence. The basic idea behind iterative tracking is that the initial iterations search for tracks which are the easiest to find. These tracks are generally created by particles with relatively large momenta or particles produced near the interaction region. After each iteration, hits associated with tracks are removed and the next iteration continues using hits which are still not associated with tracks. Subsequent iterations search for more difficult classes of tracks e.g. tracks created by low $p_T$ particles or particles which are displaced from the interaction region. These iterations of track finding are utilized to maximize the tracking efficiency and minimize fake rates, where fake tracks are reconstructed tracks leading to false trajectories from an incorrect combination of tracker hits. If charged particles cross seven layers or less of the tracker, the fake rate can be as large as $\sim 80\%$. If more than seven layers are crossed, the fake rate reduces to less than 10%. The fake rate is below 1% for particles crossing ten or more tracking layers [47]. The fake rate for reconstructing tracks having $p_T > 170$ GeV is less than 1% for the barrel region. This number is around 10% for the endcap region. The overall efficiency is highest in the barrel region (95% for $|\eta| < 0.5$). In all other regions, the efficiency never goes below 85% ($|\eta| < 2.2$) [45].

The track reconstruction procedure follows four steps [45]:

- The first step is seed generation which provides the initial track candidates. These candidates are found using very few (2 or 3) hits compatible with the interaction point and above some $p_T$ threshold. In the uniform magnetic field of the tracker, charged particles follow helical trajectories. The parameters needed to define the trajectory can be extracted from 3-D hits (A 3-D hit refers to any hit that can provide a 3-D position measurement). The positions of seeds are very important as misidentification of these seeds can reduce the tracking efficiency significantly.

- The second step is track finding. Once the seeds are identified, the tracks are extrapolated outwards by adding hits from successive detector layers and updating the track parameters at

each layer. The information collected by the CTF algorithm includes location and uncertainty of the detected hits as well as the amount of detector material crossed. This is needed to estimate the effects of multiple Coulomb scattering and energy loss. The track finding also follows a few steps. In the first step all possible hits are associated with the track. In the case where there is more than one hit in a layer, multiple track trajectories are formed. If hits are unambiguously associated with tracks, they are removed from the list of possible collection of hits. The second and third steps repeat the first step, but with looser requirements on the remaining potential tracks. Again, gradually hits are removed once they are associated with the tracks. In the fourth and fifth steps, requirements on vertex position are relaxed which helps to reconstruct particle tracks coming from secondary interaction points within the tracker material or particles decaying inside the tracker material.

- The third step is known as track fitting. The previous step provides a collection of hits and an estimate of track parameters where the full information about the trajectory is available only when all hits are known. This estimate can be biased by various constraints like beam spot constraints applied to the trajectory during track seeding. The trajectory is refit using all found hits with the Kalman filter and this makes the trajectory smoother.

- The final step is track selection. In a typical event containing jets at the LHC, a significant fraction of tracks obtained by track finding procedure will be fake tracks. The number of fake tracks are reduced through quality requirements. Good tracks are selected on the basis of the numbers of layers having hits, how compatible the tracks are with originating from the primary interaction vertex and whether the track fits yield a good normalized $\chi^2$.

There is a variable called *"pixel track seed veto"* which will be used to identify photons in this analysis. This variable will reject any photon having a pixel track seed associated with it. The pixel track seed consists of at least two hits in the pixel detector and is formed in the earliest iterations of tracking. If a photon object is associated with a pixel track seed, then it is considered as an electron for the purpose of this analysis. Details of photon and electron reconstructions are described in Sections 4.3 and 4.4.

### 4.1.2 Vertex Reconstruction

A vertex is the point of interaction and the point of production of particles in the detector. A primary vertex is the position of proton-proton interaction, including the hard-scattering collisions needed for analysis and the pileup collisions. The primary vertex of an event is defined as that vertex for which the total of squared transverse momenta of all the tracks associated with the vertex is the

highest compared to that of all other vertices. The primary goal of primary vertex reconstruction is to measure the location of the proton-proton interaction point from hard scattering. The precise location of vertices are used to distinguish between the primary vertex and vertices from pileup collisions. This in turn helps to reduce the effect of pileup in event reconstruction. Reconstruction of vertices consists of three steps [45]: selection of tracks, clustering the tracks which appear to originate from the same interaction point and fitting the position of each vertex exploiting the associated tracks of the vertex.

Track selection is done by choosing tracks which are consistent with being produced promptly in a hard scattering in the primary interaction region. To choose the good tracks, several quality criteria are required [45]. The selected tracks are then clustered on the basis of their positions in z-axis at the point of closest approach to the center of the beam spot. The beam spot is actually the three dimensional profile of the luminous region, where the LHC beams collide. The track clustering follows a *deterministic annealing* (DA) algorithm [48]. After candidate vertices are identified based on DA clustering of tracks in z, the vertices containing at least two tracks are fit using an *adaptive vertex filter* [49]. This filter computes the best estimates of the vertex parameters which includes its x, y and z positions. It also computes the number of degrees of freedom for the vertex which is an indicator of how well the fit performed.

## 4.2 Particle Flow Algorithm

The Particle Flow (PF) algorithm [50] is a technique for identifying and reconstructing all the individual particles originating from the collisions by combining the information coming from different sub-detectors in CMS. The signals coming from the sub-detectors are combined together optimally to reconstruct basic objects like photons, electrons, charged and neutral hadrons and muons and their properties such as momentum and direction. Particle type is determined by combining the information coming from different sub-detectors. For example, an ECAL energy deposit with a track is identified as electron, where as ECAL energy deposit with no track and no energy deposition in the HCAL is labeled as a photon. Once the basic objects are reconstructed, PF uses those objects to reconstruct jets and missing transverse energy ($E_T^{miss}$).

### 4.2.1 Calorimeter Clustering

The PF clustering algorithm in the calorimeters clusters the energy depositions in each of the calorimeter components separately. The clustering algorithm in the calorimeters helps us to [51]:

(i) detect and measure the energy and direction of neutral particles like neutral hadrons and photons,

(ii) separate neutral particles from energy depositions done by charged hadrons,

(iii) identify electrons and reconstruct their energies taking all accompanying bremsstrahlung photons into account,

(iv) measure the energy of charged hadrons having high $p_T$, low quality tracks.

The clustering is performed for ECAL barrel and endcap, HCAL barrel and endcap, Pre-Shower 1st layer and Pre-Shower 2nd layer separately. The starting point of clustering starts from the energy depositions in the crystals (for ECAL) or towers (for HCAL). Presently no clustering is performed for HF, so here each cell gives one cluster.

A "cluster seed" is identified as the maxima of energy deposition in local calorimeter cells, given that its energy deposition is above a threshold. The energy of seed for ECAL barrel region is taken to be greater than 230 MeV and for endcap region, this is greater than 600 MeV (or $E_T$ greater than 150 MeV). Then the cells around the seed are added to the cluster if their energy depositions are two standard deviations above the calorimeter noise level and they share a common side with the seed or a cell already in the clusters. The threshold energy is 80 MeV in the ECAL barrel and up to 300 MeV in the ECAL endcaps. This way, the cluster seeds form "topological clusters". A topological cluster yields as many "particle flow" clusters as it has seeds [51]. If a cluster contains more than one seed, the cluster is split between constituent seeds and the cluster energy is shared among the seeds in proportion to the distances from the seeds. Lastly, the particle flow clusters are dynamically merged into superclusters for electron and photon reconstruction. The size of the supercluster depends on the $\eta$ and $\phi$ position of the supercluster. Thus it is possible to consider the geometrical variation of the detector.

The clustering in HCAL follows the same procedure as the ECAL method described above. The seed of the HCAL cluster should have energy greater than 1.0 GeV in the barrel region and greater than 1.1 GeV in the endcap region. The adjacent HCAL towers to the seeding towers, if

having energies above some threshold (energy greater than 800 MeV), are added to first seeding towers to create topological clusters.

### 4.2.2  Linking Algorithm

A given particle is expected to give rise to several particle flow elements from different CMS sub-detectors. These particle flow elements include charged-particle tracks, and/or several calorimeter clusters, and/or muon tracks[51]. The information from different sub-detectors must be connected or "linked" properly. The CMS PF algorithm links tracks and clusters by comparing the distance between particle flow elements in $\Delta R$ ($= \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$). The elements linked to another is removed from the list of possible elements, such that double-counting in linking does not occur.

The tracks of charged particles from the inner tracker are extrapolated to the ECAL and HCAL. If a track lies within a cluster boundary, or is tangential to the cluster boundary, the track is matched to the cluster. To collect the energy of all bremsstrahlung photons emitted by charged particles like electrons, tangents to the tracks are extrapolated to the ECAL from the intersection points between the track and each of the tracker layers, because these are the points where potentially bremsstrahlung photons will be emitted by electrons. If the extrapolated tangent from the track is within the boundaries of the cluster, the cluster is linked to the track as a bremsstrahlung photon. The method of extrapolation and linking tracks to clusters work for both ECAL and HCAL clusters. Calorimeter clusters, either an ECAL and an HCAL cluster or an ECAL cluster and a PS cluster, are also linked. When the cluster position in the more granular calorimeter is within the cluster envelope of the less granular calorimeter, a link between them is established [51].

The energy deposition in the HCAL with no track associated with it will be taken as a neutral hadron object, whereas if the energy deposition in the HCAL is linked to a track, that object will be taken as a charged hadron object. Similarly, energy deposition in the ECAL with no track associated with it will be considered as a photon object and energy deposition in the ECAL with track will be considered as an electron object.

The muons are also linked in the similar way where information from the inner tracker and the muon chambers are linked. Tracks are extrapolated from the both systems and the $\chi^2$ of the resulting fit determines if they are matched. Details of the muon reconstruction is described in Section 4.5.

## 4.3 Photon Reconstruction

Photon candidates are reconstructed from the clusters of energy deposited in the ECAL. The clustering algorithm follows the same procedure as described in Section 4.2.1. A fixed supercluster size, which was the output of the stand alone clustering algorithm used at CMS during Run I [52], will associate more energies from pileup in the low $\eta$ region, while missing bremsstrahlung photons in the high $\eta$ region. In both cases, dynamically merged superclusters are better in reconstructing the photon correctly compared to stand alone clustering as their sizes are not fixed. The dynamically merged superclusters are also good for energy containment. These superclusters automatically consider the geometric variation of the detector with $\eta$. A supercluster is promoted to be a photon if the reconstructed transverse energy is greater than 10 GeV.

### 4.3.1 Photon Related Variables

There are several photon related variables used in CMS which help to select good quality photons. Brief descriptions of such variables are given below:

- $R_9$: This is defined as $E_{3\times3}/E_{SC}$ where $E_{3\times3}$ is the energy of $3 \times 3$ crystals surrounding the supercluster seed crystal and $E_{SC}$ is the total energy of the supercluster. This variable is used to discriminate between unconverted (high $R_9$) and converted photons (low $R_9$) [52]. This is also used to discriminate against spikes which are detector noise mimicking the signal of very energetic photons. To reject spikes, $R_9$ is required to be less than 1.

- **Single tower H/E:** This is the ratio of energy deposited (H) in the HCAL tower directly behind the ECAL seed crystal to the total energy of ECAL supercluster.

- $\sigma_{i\eta i\eta}$ : This is lateral extension of electromagnetic shower measured in the $\eta$ direction. This is measured using crystal cells which are in a $5 \times 5$ matrix around the supercluster seed crystal. The mathematical definition of $\sigma_{i\eta i\eta}$ is

$$\sigma_{i\eta i\eta} = \sqrt{\frac{\sum_{5\times5} w_i (\bar{\eta} - \eta_i)^2}{\sum_{5\times5} w_i}} \tag{4.4}$$

where the weighting factor $w_i$ is given by: $w_i = max(0, w_0 + \ln(\frac{E_i}{E_{5\times5}}))$ with the constant $w_0$ set to be 4.7 and $\bar{\eta}$ is the energy weighted mean of the crystals in integer $\eta$ space. This variable is extremely effective in discriminating against hadronic jets as they tend to have larger values for this variable than that of photons.

- **Pixel track seed veto:** This veto, described in Section 4.1.1, rejects any photon which has pixel track seed associated with it.

- **Isolation variables:** Isolation variables are key discriminators against jets. The isolation is a measure of how much additional transverse energy is surrounding a photon object or a supercluster. For a photon within a jet, the surrounding energy deposition around the photon will be very high as opposed to a good isolated photon. Using the isolation variables, photons within jets are discarded and well isolated photons are selected for physics analyses. There are three different types of isolation variables:

    - **PF Charged Hadron iso:** This is the total transverse momentum deposited by the particle flow charged hadrons surrounding a photon supercluster in a hollow cone of $0.02 < \Delta R < 0.3$.

    - **PF Photon iso:** This is the total transverse momenta of particle flow photon objects around the supercluster in $\Delta R < 0.3$ cone, where a strip of $\eta = 0.015$ around the supercluster is excluded.

    - **PF Neutral Hadron iso:** This is the total transverse momenta of particle flow neutral hadrons around the photon supercluster in the cone of $\Delta R < 0.3$.

## 4.4  Electron Reconstruction

The identification and reconstruction of electrons at CMS is a challenging task because electrons can radiate photons and the photons can have discrete depositions of energy in the ECAL. It is difficult to distinguish a photon coming from bremsstrahlung or coming from another source. In addition to this, radiated photons can convert into electron-positron pairs as well. Since electrons can be significantly bent by the magnetic field of the solenoid, the electrons and bremsstrahlung photons can be widely spread along the $\phi$-direction.

The reconstruction of electrons exploits the ECAL granularity optimally and tries to reconstruct electrons almost independently of their isolation or the transverse momenta. The particle flow electron tracks are found by fitting with Gaussian Sum Filter (GSF) [53]. The electron GSF tracks are seeded by the subset of all the tracks compatible with the electron particle hypotheses in the event.

Tracks are extrapolated to the ECAL and the electron energy deposition is found from ECAL. As stated earlier, electrons may radiate photons which are also reconstructed by particle flow clustering algorithm. These photons are collected by a straight line tangent extrapolation to the electron track to the ECAL from each tracker layer. This is because photon is electric charge neutral

50

and can not be bent by the magnetic field. This is the "tracker driven" electron reconstruction algorithm.

There is another method which is the "ECAL driven" seeding approach. Here the electron candidates are seeded from ECAL superclusters. The supercluster energy and position are used to estimate the track position in the inner tracker. This way the pixel based tracker seeds are selected. Those tracks which potentially match the trajectory and momentum of the electron are selected for GSF tracking. The trajectory of an electron candidate from supercluster is extrapolated back along the helical pattern towards the collision vertex.

But many candidate tracks will come from charged hadrons, so additional selection criteria must be applied to reduce the fake rate. These criteria include ratio of energy deposited in HCAL (H) to that deposited in ECAL (E) to be less than certain value ($< 0.15$), and $\Delta R$ matching of PF cluster at ECAL and track. The requirement on $\eta$-direction is less than 0.02 and that in $\phi$-direction is less than 0.15 rad of the PF cluster at ECAL.

### 4.4.1  Electron Momentum Estimation

The electron momentum is estimated using the information from the energy deposition in the ECAL and the tracker. For electrons, the momentum observable is sensitive to bremsstrahlung photons and their subsequent conversion. Electrons are classified according to their bremsstrahlung patterns [54] and this helps to achieve best possible momentum estimation of electrons. The bremsstrahlung fraction of an electron is defined as:

$$f_{brem} = \frac{p_{in} - p_{out}}{p_{in}} \tag{4.5}$$

where $p_{in}$ is the momentum of electron close to the beam spot and $p_{out}$ is the momentum of electron near the ECAL face.

## 4.5  Muon Reconstruction

In the standard CMS reconstruction, muon tracks are first reconstructed, both in the tracker ('tracker-track') and the muon system ('standalone muon-track') independent of each other. Muons are reconstructed using two different approaches [55]:

- Global Muon reconstruction: For each standalone muon track in the muon chamber, a matching track inside the tracker is found by comparing the parameters of the two tracks and propagating the tracks to a common surface. Since this starts with track in muon chamber and goes back to find tracks inside tracker, this approach is basically *'outside-in'*. The Kalman filter technique is used to combine the hits from tracker track and standalone muon track. For $p_T \gtrsim 200$ GeV, the global muon fit improves the momentum resolution compared to the tracker only fit [55] because the tracks are straighter for high $p_T$ muons and so the inner track uncertainty is high.

- Tracker Muon reconstruction: This reconstruction process starts out with tracker tracks and the tracks are extrapolated to the muon system after considering the presence of magnetic field, energy loss of the muon and multiple Coulomb scattering inside the detector material. This process is called *'inside-out'* approach. All tracker tracks with $p_T$ as low as 0.5 GeV (momentum $p > 2.5$ GeV) are considered to be possible muon tracks and they are extrapolated to the muon system. If at least one muon segment (something like short track stub made of DT or CSC hits) matches with the extrapolated track, the track is considered to be a Tracker Muon. Tracker Muon reconstruction is more efficient than the Global Muon reconstruction when the muon has low momentum ($p \lesssim 5$ GeV). This is because Tracker Muon reconstruction requires only a single muon segment whereas the efficiency of the Global Muon reconstruction becomes high for muons penetrating through more than one muon station (typically segments in at least two muon stations are required).

The efficiencies of both tracker-track reconstruction and reconstruction of segments in the muon system are very high. Due to this, a very high fraction of muons ($\sim 99\%$ [55]) produced in pp collisions within the geometrical acceptance of the muon system are reconstructed as either Global Muons or Tracker Muons or both. Muons reconstructed both by the Global Muon approach and the Tracker Muon approach are merged into a single muon object. Muons reconstructed only as standalone muon tracks are more prone to pick up cosmic ray muons than the Global and Tracker Muons and they are not usually used for physics analysis.

### 4.5.1 Particle Flow Muons

The muon reconstruction processes described above are used to make "reco-muons" containing some amount of charged hadrons misidentified as muons. The particle flow algorithm applies particular selection criteria to Global and Tracker muons and rejects the charged hadrons from the sample of muon candidates. Depending on the surrounding environment of the muon, the information from other sub-detectors like energy depositions in calorimeters are used to select the

muons. To select muons within jets with high efficiency and with low misidentification of charged hadrons as muons, the selection criteria is optimized. Three different selections "isolated", "pf-tight" and "pf-loose" are used to identify particle flow muons [56]. A "reco-muon" is isolated if the transverse energy deposition in calorimeters and total $p_T$ of the tracks around the muon in a cone size of $R = 0.3$ is less than 10% of the muon $p_T$. Particle flow techniques are not needed to improve the resolution of those muons as they have very little neighboring activity. After isolated muons are selected, "pf-tight" and "pf-loose" selection criteria are applied to the remaining reco-muons.

The "pf-tight" muon selection requires a minimum number of hits on the muon track. It also requires the compatibility of muon segment and a calorimeter deposition which is defined by a simulation based template [56]. For the "pf-loose" muons, the track momenta are significantly larger than the energy depositions in the calorimeters. Hence they are incompatible with charged hadron tracks. The requirements on the number of hits are relaxed for "pf-loose" muons. Also here the simulation based template compatibility requirement is removed [56].

## 4.6    Jet Reconstruction

Particle flow jets are reconstructed by grouping the clusters of collinear particle flow particles. Partons originating from proton-proton collisions are color connected, and hence they hadronize and create showers of collinear particles, mainly hadrons. In a typical jet, the fraction of charged hadrons, photons and neutral hadrons are approximately 60%, 30% and 10%. At CMS, presently all the PF particles are first reconstructed and clustered as jets and then all other objects are separated from jet collection. There are requirements on jets such as if the jets have tracks, or the ratio of HCAL energy deposit over ECAL energy deposit is small enough to identify the jet as a photon or neutral or charged hadrons.

There are many different jet clustering algorithms employed by CMS, like $k_t$, Cambridge-Aachen (CA) and anti-$k_t$. Each one has its own merits and demerits. The jets used for the analysis in this dissertation use the anti-$k_t$ algorithm with distance parameter $R = 0.4$, commonly known as AK4PF jets.

### 4.6.1    Anti-$k_t$ Algorithm:

To begin with, one introduces the distance $d_{ij}$ between two entities i and j (may be two particle flow objects, or one particle flow object and other one cluster of collinear particles forming

"protojets" or both "protojets") and $d_{iB}$ between ith entity and the beam (B). The inclusive recon-
struction proceeds as follows. The minimum of $d_{ij}$ and $d_{iB}$ is considered and if $min(d_{ij}, d_{iB}) = d_{ij}$,
the two entities i and j are clustered together as jets or "protojets". This process iteratively con-
tinues till one hits the point $min(d_{ij}, d_{iB}) = d_{iB}$. If $min(d_{ij}, d_{iB}) = d_{iB}$, then i will be considered
as a jet and removed from the list of possible entities. The iterations will continue until no entities
are left.

The distance parameters are defined as:

$$d_{ij} = min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta_{ij}^2}{R^2} \tag{4.6}$$

$$d_{iB} = k_{ti}^{2p} \tag{4.7}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$. The $k_{ti}$, $y_i$ and $\phi_i$ refer to transverse momentum, rapidity
and azimuthal angle respectively of ith entity. Here $p$ is an integer to govern the relative power
of the energy versus the geometrical scale $\Delta_{ij}$. This parameter is 1 for the $k_t$ algorithm, 0 for the
Cambridge-Aachen algorithm and $-1$ for the Anti-$k_t$ algorithm. $R$ is the radius parameter: For
AK4PF jets, this value is kept fixed at 0.4. Negative values of $p$ may seem to be pathological, but
these values help the jet algorithm to be infrared and collinear safe [57].

The functionality of the anti-$k_t$ jet algorithm can be understood clearly if one considers an
event with a few well separated hard particles having transverse momenta $k_{t1}$, $k_{t2}$, ... and many
soft particles having low transverse momenta. Consider $d_{1i}$ where 1 is a hard particle and i is a soft
particle. Clearly, here $min(k_{t1}^{-2}, k_{ti}^{-2})\frac{\Delta_{1i}^2}{R^2}$ is dominated by $k_{t1}^{-2}$ and $\Delta_{1i}$ separations. The distance
parameter $d_{ij}$ where i and j are both soft particles will be larger than $d_{1i}$ (given that the separation
parameter remains very similar) and hence soft particles will tend to cluster with hard particles
long before they are clustered among themselves. If a hard particle does not have another hard
particle within a radius of $2R$, then the hard particle will be clustered with all the soft particles
within the radius of $R$ and the resulting jet will be perfectly conical.

If there are two hard particles such that their separation is more than $R$, but less than $2R$, it
is not possible for both the jets to have perfect cones. If transverse momentum of first jet, $k_{t1}$ is
much larger than that of second jet ($k_{t2}$), then jet 1 will be conical and jet 2 will be a truncated
cone. This is because jet 2 will lose few particles which overlap with jet 1. If $k_{t1} = k_{t2}$, neither
of the two jets will be of conical shape, the overlapping part will simply be divided equally by a
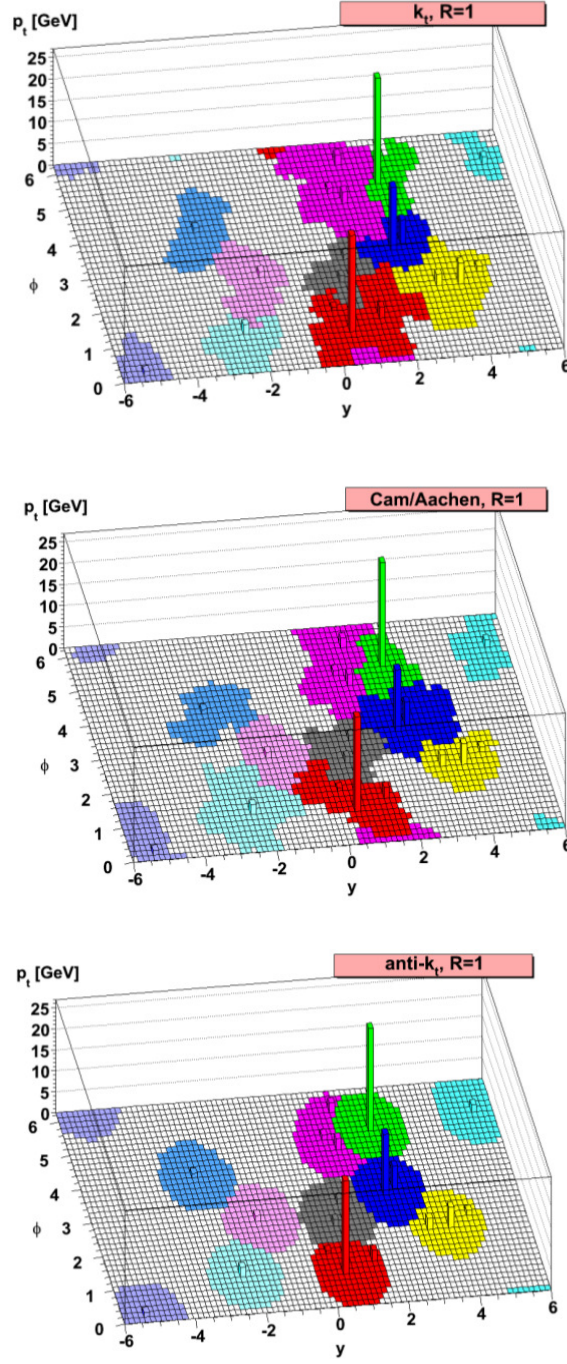
Figure 4.1: A sample parton-level event where different jet algorithms are applied is shown here. Here hard particles are shown with many random soft particles. The shapes of jets with $k_t$ and Cambridge-Aachen algorithm depend on the distribution of soft particles, whereas the shapes have very small dependence on position of soft particles for anti-$k_t$ jet algorithm. Here the jet radius parameter $R$ is 1, but similar qualitative behavior is expected from jet clustering algorithm with radius parameter $R = 0.4$ [57].

straight line between the two. For the case $k_{t1} \sim k_{t2}$, the shape will again be a truncated cone for the two jets, where the separation boundary, b, between them will be given by $\Delta_{1b}/k_{t1} = \Delta_{2b}/k_{t2}$. For the case $\Delta_{12} < R$, particles 1 and 2 will be clustered into a single jet. If $k_{t1} \gg k_{t2}$, then the center of the jet will be around $k_1$. For the case $k_{t1} \sim k_{t2}$, the shape will be more complex one where the shape will be union of cones of radius $< R$ around each hard particle and cone of radius $R$ centered around the final jet.

The key feature of the anti-$k_t$ algorithm is that soft particles do not modify the shape of the jet and this algorithm is resilient with respect to soft radiation, but hard radiation changes the shape of jets. For comparison with different jet algorithm, A sample parton level event with different jet algorithms is shown in Figure 4.1. The figure justifies that shape of jet is not modified by soft radiation in anti-$k_t$ unlike other algorithms. This is ideal for a high pile up scenario.

### 4.6.2 Performance of Particle Flow Jets

The performance of particle flow jets was studied with simulated events. It is already stated that almost 90% of the jet energy is carried by photon and charged hadrons which are measured with high precision by ECAL and tracker respectively. But the rest of the energy is neutral hadrons which are measured mostly by HCAL. The energy resolution of HCAL is poorer than that of ECAL. The energy resolution of jets where jets are reconstructed solely by calorimeters ignoring information from other sub-detectors (so-called "calo-jet") are worse than that of jets which are reconstructed by particle flow algorithm as evident from Figure 4.2. The particle-flow jets are particularly beneficial at the low transverse momentum.

### 4.6.3 Jet Energy Corrections

Jet energy corrections are needed after particle flow particles are clustered into jets to reduce the effect of contributions from electronic noise, non-uniformity of calorimeter response and the contribution from pile up (energy deposition from multiple interactions in the same bunch crossing). Jet energy corrections are multiplicative factors applied to reconstructed jet energies and to attempt to correct reconstructed jet energies as closely as possible to the true values.

These corrections are determined from both data-driven methods and simulation. There are three levels of corrections. The first one is called the Level-1 (L1) correction which corrects the additional energy in jet coming from electronic noise or pileup. This is a $\rho$ correction where $\rho$
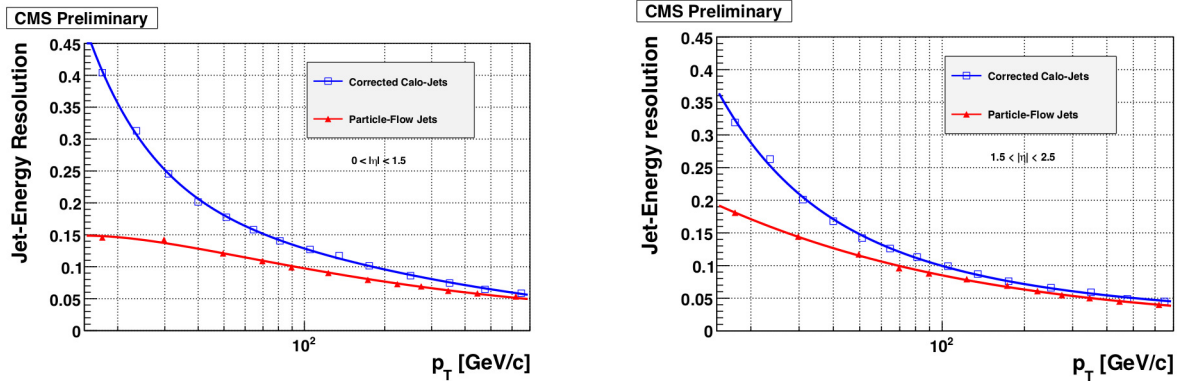
Figure 4.2: Jet resolution of jet-jet events in the barrel (left) and endcap (right). Here resolution of jets reconstructed solely by calorimeters is compared to that of jets reconstructed by the particle flow algorithm. This plot also shows that jet resolution improves when jet $p_T$ increases [51].

represents the median of the distribution of jet energy per unit jet area for a particular event. This correction removes any dataset dependence on luminosity. The jet energy resolution will degrade as the instantaneous luminosity of CMS increases because this will increase pileup. The L1 correction values are found from using noise-only and minimum-bias MC events [58].

After the application of the L1 correction, the next level of correction is the Level-2 (L2) correction which removes the variation in response of the detector as a function of pseudorapidity ($\eta$). This correction makes the jet response flat with respect to $\eta$. Jets in the central region ($|\eta| < 1.3$) of the CMS detector are considered to be the most reliably measured because particles traverse less detector material. On the other hand, the jet energy response degrades as $|\eta|$ increases because the amount of tracker and supporting material becomes more substantial in the higher $|\eta|$ region. Using simulation truth values and data-driven techniques [58], jet energies at arbitrary $\eta$ can be compared with jets in the central region of detector and this way $\eta$ dependent correction factors are obtained.

The Level 3 (L3) correction factors are applied to adjust for detector response and remove the variation coming from jet transverse momentum. This correction factors are obtained from simulation where one can compare L1 and L2 corrected reconstructed jet $p_T$ with the generator level jet $p_T$. Also data driven techniques are used where $\gamma$+jet events are used to derive the L3

correction factors. This sample is used because jet and photon should balance each other in the transverse plane and photon $p_T$ is precisely measured with the ECAL.

The total jet energy correction is given as the product of all the three levels [59]:

$$E_{Corrected} = (E_{L1Corrected} - E_{Offset}) \times C_{Rel}(\eta, p_T^{'}) \times C_{Abs}(p_T^{''}) \tag{4.8}$$

. Here $p_T^{'}$ refers to L1 corrected jet $p_T$. $C_{Rel}(\eta, p_T^{'})$ is the L2 correction factor and $C_{Abs}(p_T^{''})$ is the L3 correction factor. For the L3 correction factor, the jet $p_T$ after both L1 and L2 corrections is used: $p_T^{''} = p_T^{'} C_{Rel}(\eta, p_T^{'})$. The $E_{Corrected}$ typically has values within 0.1 to 10.

## 4.7 Missing Transverse Energy Reconstruction

Unfortunately, not all known particles can be detected at CMS all the time. SM neutrinos are very weakly interacting particles and when they are produced, they leave the detector undetected. Similarly, any particles which are not yet discovered and which are also weakly interacting, will not leave detectable signatures in the detector. These particles must be inferred from the imbalance in total momenta before and after collision.

The colliding partons are highly boosted by the LHC in the Z direction. So the momenta of the partons in the Z direction are very high and not well known at all. But the partons have very small transverse momenta. From momentum conservation in the transverse direction, the transverse momenta of all of the particles should add up (vector addition) to be zero. If there is a significant imbalance in this value, it implies either a particle (or particles) has escaped the detector or there is mis-measurement of visible energy in the event.

The missing transverse energy ($E_T^{miss}$ ) is defined as:

$$\vec{E_T^{miss}} = -\sum_i^n \vec{p_{Ti}} \tag{4.9}$$

where the sum runs over all the visible particles in the event. Due to this, the resolution of $E_T^{miss}$ depends on measurement of transverse momentum of all the visible particles.

### 4.7.1 Correction to $E_T^{miss}$

Equation 4.9 calculates the raw $E_T^{miss}$ of an event. The raw $E_T^{miss}$ is different from the true $E_T^{miss}$ for instrumental reasons. The difference arises because of the non-compensating nature of

the calorimeters, their resolutions and misalignment of the detectors. Corrections for $E_T^{miss}$ are applied to get a better estimate of $E_T^{miss}$ in an event.

**The Type-I correction:** The Type-I correction is the most important $E_T^{miss}$ correction at CMS. Technically, this is just the propagation of the jet energy corrections to $E_T^{miss}$. For the Type-I correction, the raw transverse momenta of particles clustered as jets in the Equation 4.9 are replaced by the transverse momenta of the jets where the jet energy corrections have been applied. The mathematical formulation for $E_T^{miss\ Type-I}$ is given by:

$$\vec{E_T^{miss}}^{Type-I} = -\sum_{i \in jets} \vec{p_{Ti}}^{JEC} - \sum_{i \in other} \vec{p_{Ti}} \tag{4.10}$$

where $\vec{p_{Ti}}^{JEC}$ refers to the momentum of jets where jet energy corrections has been applied and 'other' objects in the sum refer to any object apart from those clustered as jets.

From Equation 4.9, the sum can be broken into particles which are clustered as jets and which are particles other than the jets and rewritten as:

$$\vec{E_T^{miss}}^{raw} = -\sum_{i \in jets} \vec{p_{Ti}} - \sum_{i \in other} \vec{p_{Ti}} \tag{4.11}$$

The first vector sum is on jets where jet energy corrections have not been applied. Therefore the Type-I correction can be written in the following way:

$$\vec{C_T}^{Type-I} = \sum_{i \in jets} \vec{p_{Ti}} - \sum_{i \in jets} \vec{p_{Ti}^{JEC}} \tag{4.12}$$

.

From Equations 4.11 and 4.12, replacing $\sum_{i \in jets} \vec{p_{Ti}}$ one can write:

$$\vec{E_T^{miss}}^{raw} = -\vec{C_T}^{Type-I} - \sum_{i \in jets} \vec{p_{Ti}}^{JEC} - \sum_{i \in other} \vec{p_{Ti}}$$

$$or,\ \vec{E_T^{miss}}^{raw} + \vec{C_T}^{Type-I} = -\sum_{i \in jets} \vec{p_{Ti}}^{JEC} - \sum_{i \in other} \vec{p_{Ti}}$$

$$or,\ \vec{E_T^{miss}}^{Type-I} = \vec{E_T^{miss}}^{raw} + \vec{C_T}^{Type-I} \tag{4.13}$$

where in the last line, Equation 4.10 was used.

**The Type-0 correction:** Broadly, the Type-0 correction is an additive factor which is used to mitigate the effect of pile up interactions. Particles arising from pile up generally have low true $E_T^{miss}$ as very few of them are invisible particles (e.g. neutrinos from Kaon decay). Also the

pileup interactions are symmetric in $\phi$. Therefore, if one could measure all the visible particles precisely and accurately, the $E_T^{miss}$ resolution would not degrade in a high pile up environment. But unfortunately, since our measurement of all the visible particles is not completely efficient always, the $E_T^{miss}$ resolution degrades as the number of pile up interactions increases.

The Type-0 correction attempts to distinguish between particles arising from hard scattering (HS) and particles arising from pileup interactions (PU). This correction removes the additional energy deposited by the pileup by removing the charged and neutral hadrons coming from PU. So to calculate the raw $E_T^{miss}$, particles are classified as originating from hard scattering or from pile up interactions. Hence Equation 4.9 becomes:

$$\vec{E_T^{miss}}^{raw} = -\sum_{i \in HS} \vec{p_{Ti}} - \sum_{i \in PU} \vec{p_{Ti}} \tag{4.14}$$

This equation is theoretically correct, but as one cannot identify all the particles arising from HS and PU, one needs to use some special techniques to effectively remove the second term in Equation 4.14. To find a possible way to do that, the PU particles are classified into neutral particles (neuPU) and charged particles (chPU) and Equation 4.14 becomes:

$$\vec{E_T^{miss}}^{raw} = -\sum_{i \in HS} \vec{p_{Ti}} - \sum_{i \in neuPU} \vec{p_{Ti}} - \sum_{i \in chPU} \vec{p_{Ti}} \tag{4.15}$$

The last sum is taken over the charged particles coming from PU and one can identify such particles from the positions of their vertices. However, the first and second sums in Equation 4.15 cannot be separated. If only the last sum in the above equation is removed, this will make the reconstructed $E_T^{miss}$ more different from its actual value, rather than bringing it close to the actual value.

The neutral pile up particles and charged pile up particles are produced in the same interactions with little to no true $E_T^{miss}$. As a result, the second and last sum in Equation 4.15 are nearly in opposite directions. If the true values are obtained, then these two sums will nearly cancel each other. The Type-0 correction assumes that this cancellation is exact:

$$\sum_{i \in neuPU} \vec{p_{Ti}}^{true} + \sum_{i \in chPU} \vec{p_{Ti}}^{true} = 0 \tag{4.16}$$

This correction also assumes the following:

$$\sum_{i \in chPU} \vec{p_{Ti}}^{true} = \sum_{i \in chPU} \vec{p_{Ti}} \qquad (4.17)$$

i.e. the charged pile up particles can be measured perfectly. The CMS detector has an all silicon tracker kept in a strong magnetic field. The curvature of the tracks can be measured precisely with the CMS tracker (95% in the barrel and more than 85% in the endcap [45]).

The other assumption of Type-0 correction is that the direction of neutral pile up particles are measured perfectly, but their energies are measured systematically off by a factor ($R^0$). Hence one can write:

$$\sum_{i \in neuPU} \vec{p_{Ti}} = R^0 \sum_{i \in neuPU} \vec{p_{Ti}}^{true} \qquad (4.18)$$

The justification for this assumption is that the direction of neutral particles are measured from the energy deposits in calorimeter cells and calorimeters are calibrated to measure high transverse momenta particles. So the energies of low transverse momenta particles are measured systematically off.

So from Equation 4.16, one can write:

$$\sum_{i \in neuPU} \vec{p_{Ti}} = -R^0 \sum_{i \in chPU} \vec{p_{Ti}} \qquad (4.19)$$

The Type-0 correction is addition of this estimate with the estimate coming from charged PU particles. Hence it is written as:

$$\vec{C_T}^{Type-0} = \sum_{i \in neuPU} \vec{p_{Ti}} + \sum_{i \in chPU} \vec{p_{Ti}} = (1 - R^0) \sum_{i \in chPU} \vec{p_{Ti}} \qquad (4.20)$$

This correction factor is added to the $E_T^{miss\ raw}$ and one gets Type-0 corrected $E_T^{miss}$ :

$$\vec{E_T^{miss}}^{Type-0} = \vec{E_T^{miss}}^{raw} + \vec{C_T}^{Type-0} \qquad (4.21)$$

Finally the Type-I and Type-0 corrections can be added so that one gets:

$$\vec{E_T^{miss}}^{Type-0-I} = \vec{E_T^{miss}}^{Type-0} + \vec{E_T^{miss}}^{Type-I} \qquad (4.22)$$

One thing to remember, if both Type-0 and Type-I corrections are applied at the same time, a certain fraction of transverse momenta of the particles are over-corrected. This is because Type-0

corrections are applied on a set of particles produced from pile up interactions and Type-I corrections are applied to particles clustered as jets. These two types of particles are not strictly disjoint. As a result, particles originating from PU interactions and also clustered as jets will receive the corrections from both the Type-0 and Type-I correction. In the current implementation this overcorrection is small and ignored.

**xy-Shift Correction:** The distribution of true $E_T^{miss}$ around the beam axis should be symmetric in $\phi$ direction because of the rotational symmetry, but it is observed that the $E_T^{miss}$ depends on $\phi$ roughly in a sinusoidal curve with period of $2\pi$. This may be caused by the anisotropic detector response, inactive calorimeter cells, misalignment of the detector and the displacement of the beam spot. The amplitude of this distribution varies approximately linearly with the number of pile up interactions.

The amplitude of $\phi$-modulation can be reduced if the origin of the coordinate in the transverse momentum plane is shifted:

$$\vec{p_{Ti}} \rightarrow \vec{p_{Ti}} - \vec{c} \tag{4.23}$$

where $\vec{c}$ is the shift by which origin is offset. Therefore the xy-shift corrected $E_T^{miss}$ becomes:

$$\begin{aligned} \vec{E_T^{miss}}^{xy} &= -\sum_{i \in all} (\vec{p_{Ti}} - \vec{c}) \\ &= -\sum_{i \in all} \vec{p_{Ti}} + \sum_{i \in all} \vec{c} \\ &= \vec{E_T^{miss}}^{raw} + n\vec{c} \end{aligned} \tag{4.24}$$

where n is the number of particles in an event. So the xy-shift correction is $\vec{C_T}^{xy} = n\vec{c}$. Finally the xy-shifted $E_T^{miss}$ is given by:

$$\vec{E_T^{miss}}^{xy} = \vec{E_T^{miss}}^{raw} + \vec{C_T}^{xy}. \tag{4.25}$$

Presently, the xy-shift correction is implemented as a function of number of vertices and the $\vec{C_T}^{xy}$ is rewritten as $\vec{c_A} + n_{vtx}\vec{c_B}$ where $\vec{c_A}$ and $\vec{c_B}$ are constant vectors and $n_{vtx}$ refers to the number of vertices in the event. The overall corrected $E_T^{miss}$ is therefore is the sum of all the corrections:

$$\vec{E_T^{miss}}^{Type-0-I-xy} = \vec{E_T^{miss}}^{Type-0-I} + \vec{C_T}^{xy} \tag{4.26}$$

For this analysis, only Type-I corrected $E_T^{miss}$ was used.

# CHAPTER 5

# DATASETS, SIMULATED SAMPLES AND TRIGGERS

## 5.1 Luminosity

In accelerator physics, instantaneous luminosity is defined as the ratio of number of events detected per unit time to the interaction cross section. Instantaneous luminosity gives an estimate of the mean number of collisions per bunch crossing and the integrated luminosity gives an estimate of the total number of collisions in a given time. This number also gives an estimate of the amount of data taken by the detectors. The LHC conducts a measurement of the integrated luminosity whenever the LHC is operational, irrespective of whether any detector is taking data or not. Since detectors can have downtime, the luminosity accepted by a detector is always less than the luminosity delivered by the LHC. In 2015, the integrated luminosity delivered by the LHC and recorded by the CMS is shown in Fig 5.1[60]. This analysis used the data corresponding to an integrated luminosity of 2.3 fb$^{-1}$.

### 5.1.1 Instantaneous Luminosity Measurement at CMS

At CMS, the instantaneous luminosity is measured online. There are offline methods as well, to cross check the instantaneous luminosity coming from the online method, and provide the most precise determination. [61]. The online method uses signals from the HF, which covers the region $3 < \eta < 5$. Two methods are applied to extract the real time relative instantaneous luminosity with the HF. The first method utilizes the linear relationship between the average transverse energy per HF tower and the luminosity to infer about the mean number of interactions per bunch crossing. The second method is known as "zero-counting" because here the average fraction of empty HF towers is used to infer the mean number of interactions per bunch crossing.

Apart from the HF, CMS uses the Pixel Luminosity Telescope to measure the instantaneous luminosity at the highest energies and highest collision rates foreseen at the LHC. It consists of eight 3-layers telescopes which are based on silicon pixel detectors. These telescopes are placed

around the beam pipe on each end of the CMS. From the hits in the pixel of the sensors, the instantaneous luminosity can be extracted [62].

There are two offline algorithms to monitor the luminosity: one of them uses the energy depositions in the HF and the other is based on tracking and vertex finding. While the offline method takes time (roughly 24 hours), it allows for better background rejection compared to the online method. Offline techniques also provide an independent data-handling path and hence they come up with a completely independent set of systematic uncertainties in the case of the vertex counting method [61].

The online and offline methods need to be calibrated to an absolute luminosity scale so that one can assign a physical significance of the luminosity values. S. Van Der Meer invented a method to calibrate the luminosity measurement, which is known as a Van Der Meer scan. The instantaneous luminosity is given by [61]:

$$L_0 = \frac{N_1 N_2 \nu_{orb} N_b F(0,0)}{\int f_x(\Delta x) d\Delta x \int f_y(\Delta y) d\Delta y} \tag{5.1}$$

where $L_0$ is the peak instantaneous luminosity, $N_i$ is the ith beam bunch intensity, $\nu_{orb}$ is the orbit frequency , $N_b$ is the number of colliding bunches per beam, and $\Delta x$ and $\Delta y$ are the beam separations in horizontal and vertical planes. In order to fit the tails of the distributions at CMS, double Gaussian distributions are used for the functions $f_x$ and $f_y$ having widths $\sigma_{1x}$, $\sigma_{2x}$ and $\sigma_{1y}$ and $\sigma_{2y}$ respectively. In general, the luminosity can be written as:

$$L = L_0 \left( \frac{h_j}{\sqrt{2\pi}\sigma_{1j}} exp \frac{-d^2}{2\sigma_{1j}^2} + \frac{1-h_j}{\sqrt{2\pi}\sigma_{2j}} exp \frac{-d^2}{2\sigma_{2j}^2} \right) \tag{5.2}$$

where d is the beam separation and $h_j$ is the fraction (by area) of the Gaussian having width $\sigma_{1j}$. The uncalibrated interaction rates as a function of d can be fit and from the fit, the absolute luminosity scale can be determined.

## 5.2   Datasets

The entire dataset taken by CMS is further streamed into different primary datasets. This is done to reduce the amount of data any given analyst needs to process, which expedites the analysis procedure. HLT paths which require similar physics objects in the final state are grouped into primary datasets. This way, the datasets are grouped in such a manner that they contain distinct
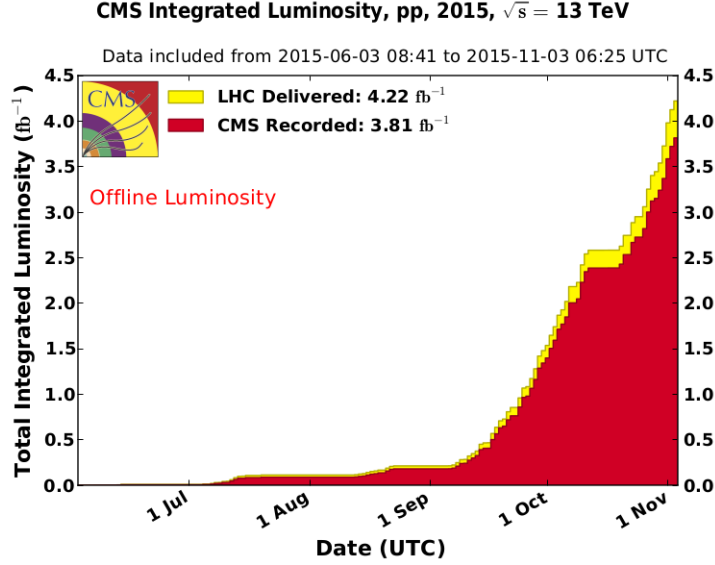
Figure 5.1: Integrated luminosity delivered by the LHC and collected by CMS in 2015 [63].

final states. As an example, if HLT paths applied on the data require two well-measured muons having transverse momenta greater than a pre-determined threshold, the data selected by those HLT criteria are grouped in the "double-muon" datasets. As this analysis required two photons in the final state, the "double EG" (electron or gamma/photon) datasets were used for this analysis. The data were taken during the 2015 run of LHC with 25 ns bunch crossing.

## 5.3 Simulated Samples

In this analysis, the $t\bar{t}$+jets and Z+jet simulated samples were used to check the contamination of real $E_T^{miss}$ in the control samples (mentioned in Section 7.1.3). W+$\gamma$ and $\gamma$+jet samples were utilized to find the components of the electroweak background (details are described in Section 7.2.3). The identification efficiency of photons in simulation was measured using a simulated $Z \rightarrow ee$ sample, as described in Section 6.2.1. Signal simulation samples were used to interpret the result obtained in this analysis (Chapter 8). A brief description of simulations used is given below.

The production of simulated events starts with the Lagrangian coming from the underlying theory of the production process. To generate the events from the Lagrangian of the theory, MadGraph , a leading order parton-level event generator was used [64]. The output of this process

were Les Houches event (LHE) files which contained all the information of the generated events prior to any showering and detector effects [65]. The LHE files were fed into PYTHIA (version 8) [66] which handled the hadronization and parton showering process. The particles are then passed through a simulated version of CMS detector using GEANT [67], known as FULLSIM [68]. For the simulation of signal samples, a simplified simulated version of CMS detector, known as FASTSIM [69], was used. FASTSIM uses a simplified geometry and simpler algorithms compared to FULLSIM for pattern recognition to reduce the computing time.

## 5.4    Trigger

The analysis described in this dissertation required two photons in the final state. The diphoton triggers were used to select the most relevant set of data for this analysis. These diphoton triggers were developed for the study of the Higgs boson decaying to two photons ($H^0 \rightarrow \gamma\gamma$). The primary trigger used for the analysis required the leading photon $p_T$ to be greater than 30 GeV and sub-leading photon $p_T$ to be greater than 18 GeV. All photons were required to pass $R9$ cuts or isolation and shower shape cuts. The invariant mass of the two photons was required to be greater than 95 GeV to reduce the background from Z boson decay ($Z \rightarrow ee$). As a consequence, separate electron triggers had to be used to get electrons in the final state which were needed as one of the control samples, used to determine the background from QCD (details of this event selection are mentioned in Chapter 6). The HLT trigger paths are listed in Table 5.1 and the trigger requirements are listed in Table 5.2.

To find the electron-to-photon mis-identification rate needed for electroweak background modeling (for details, see Chapter 7), a single electron trigger was used to populate the $ee$ and $e\gamma$ invariant mass distributions (mentioned in Section 7.2.1). The single electron trigger is also listed in Table 5.1.

### 5.4.1    Trigger Efficiency

The efficiency for the trigger has two different parts: (i) the hardware based L1 trigger efficiency and (ii) the software based high level triggers. The L1 trigger efficiency was measured by the $H \rightarrow \gamma\gamma$ group using a tag and probe method [71]. For a photon $p_T > 40$ GeV, the L1 seed was found to be 99.3% efficient [72].

Table 5.1: List of Triggers

| Primary Trigger |
|:---:|
| HLT_Diphoton30_18_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass95_v* |
| **Double Electron Triggers** |
| HLT_Diphoton30_18_R9Id_OR_IsoCaloId_AND_HE_R9Id_DoublePixelSeedMatch_Mass70_v* |
| HLT_Diphoton30EB_18EB_R9Id_OR_IsoCaloId_AND_HE_R9Id_DoublePixelVeto_Mass55_v* |
| HLT_Diphoton30PV_18PV_R9Id_AND_IsoCaloId_AND_HE_R9Id_DoublePixelVeto_Mass55_v* |
| **Single Electron Trigger** |
| HLT_Ele27_eta2p1_WPLoose_Gsf |

Table 5.2: Trigger Requirements

| Terms | Meaning |
|:---:|:---:|
| DiphotonXXX_YYY | leading photon $p_T > XXX$ GeV |
| | sub-leading photon $p_T > YYY$ GeV |
| R9Id | $R_9 > 0.85$ |
| IsoCaloId | $\sigma_{i\eta i\eta} < 0.015$ and |
| | ECAL isolation $< (6 + 0.012 \times$ photon $p_T)$ and |
| | Track isolation $< (6 + 0.002 \times$ photon $p_T)$ |
| HE_R9Id | $R_9 < 0.5$ and H/E$< 0.1$ |
| MassXXX | Invariant mass of two leading objects $>$XXX GeV |
| EleXXX | electron $p_T > XXX$ GeV |
| eta2p1 | electron $|\eta| < 2.1$ |
| WPLoose | Selection of loose electrons having efficiency 90% [70] |
| Gsf | tracks fit with Gsf filter [53] |

The primary trigger required the two photons passing the sub-leading filters and one photon passing leading filters. The total efficiency is given by:

$$\epsilon_{tot} = \epsilon_{lead,lead} \times \epsilon_{lead,sub} \times \epsilon_{sub,sub} \qquad (5.3)$$

where $\epsilon_{lead,lead}$ is the efficiency for leading photon to pass the leading filters, $\epsilon_{lead,sub}$ is the efficiency for leading photon to pass the sub-leading filters and finally $\epsilon_{sub,sub}$ is the efficiency of sub-leading photon passing the sub-leading filters. To calculate these numbers, a tag and probe method was applied where the events must have two photons passing the required photon identification criteria, $|\eta| < 1.4442$ and $R_9 > 0.8$. The photons were sorted according to their transverse momenta. The invariant mass of the leading and sub-leading photon was required to be within 75 - 105 GeV. This was to primarily collect electrons from $Z$ boson decay. The sub-leading photon transverse

momentum, used as a tag, was required to be greater than 25 GeV when the efficiencies for leading photon, used as a probe, were calculated. Again for the calculation of efficiencies for sub-leading photon (the probe), the leading photon was used as a tag and its transverse momentum was required to be greater than 40 GeV. Events satisfying all the requirements from the single photon data streams were considered for the trigger efficiency calculation.

The efficiency for the leading photon to pass the sub-leading filter was zero for $p_T < 30$ GeV, because the leading filter of the trigger was applied first. If the event failed this filter, then the event was not further considered. So this means if an event did not have a leading photon with $p_T > 30$ GeV, there was no information on other photons in that event which might have passed the sub-leading filters.

The values required for trigger efficiency calculation are put in Table 5.3. The trigger efficiencies for leading photon and sub-leading photon as a function of photon $p_T$ are shown in Fig 5.2 and Fig 5.3 respectively. To obtain the values shown in Table 5.3, the trigger efficiency plots were fit with an error function and the values were extracted from the plateau of the fits.

Table 5.3: Trigger Values

| Terms | Value |
|---|---|
| $\epsilon_{lead,lead}$ | 99.7% |
| $\epsilon_{lead,sub}$ | 99.5% |
| $\epsilon_{sub,sub}$ | 99.4% |
| $\epsilon_{tot}$ | 98.6% |

The HLT efficiency also depends on the invariant mass of the $\gamma\gamma$ system. This efficiency was calculated with events having two photons in the barrel region of ECAL passing photon identification criteria mentioned in Chapter 4, with $R_9 > 0.8$. The efficiency as a function of invariant mass is shown in Figure 5.4. The efficiency reached a plateau around 105 GeV, so the invariant mass of the double photon sample was required to be more than 105 GeV. The trigger was 98.7% efficient above this point.
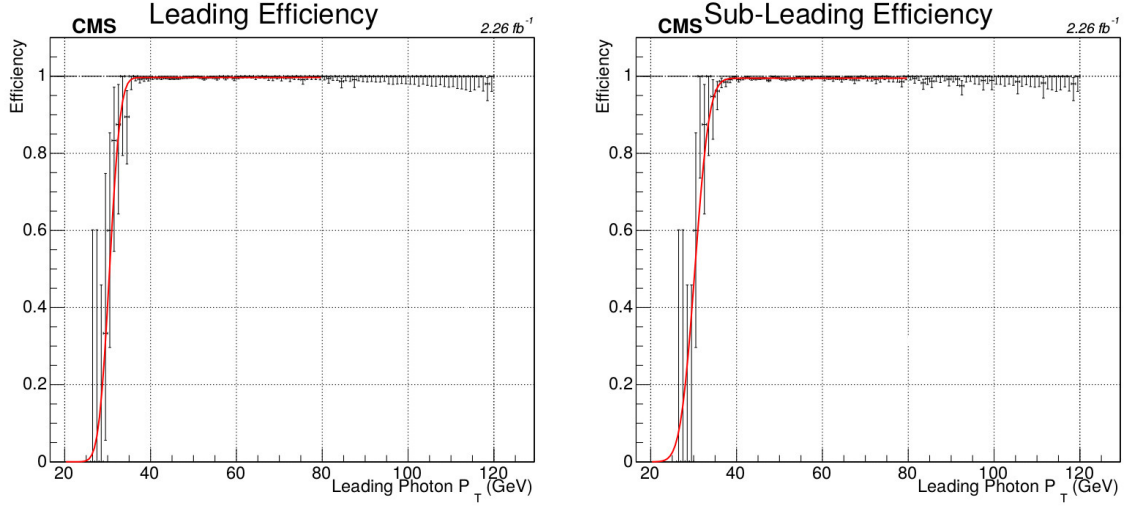
Figure 5.2: Trigger efficiency as a function of leading photon $p_T$ for the primary trigger listed in Table 5.1. On the left, efficiency for the leading photon to pass the leading filters and on the right, the efficiency for the leading photon to pass the sub-leading filters.
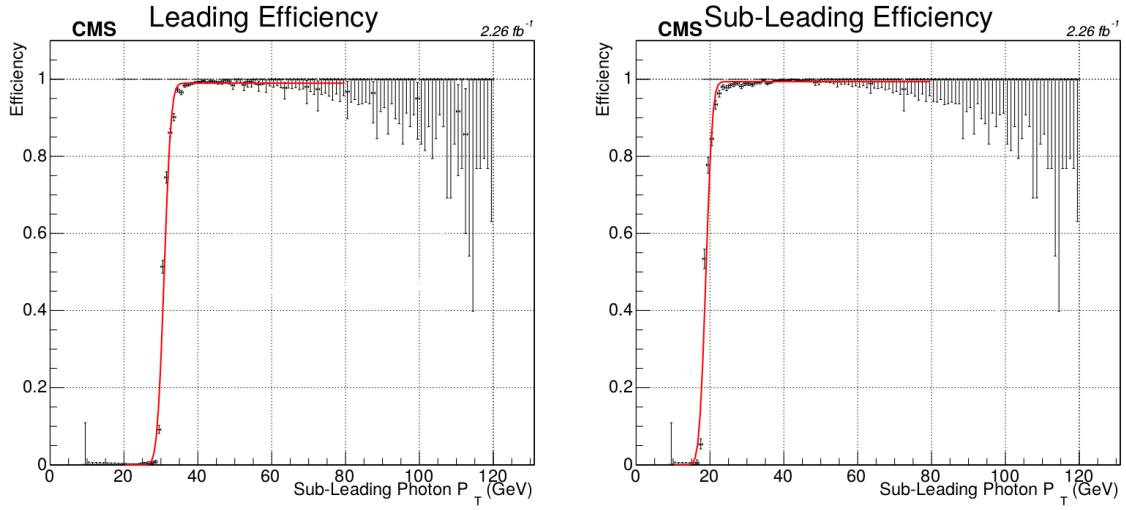


Figure 5.3: Trigger efficiency as a function of the sub-leading photon $p_T$ for the primary trigger listed in Table 5.1. On the left, efficiency for the sub-leading photon to pass the leading filters and on the right, the efficiency for sub-leading photon to pass the sub-leading filters have been plotted.
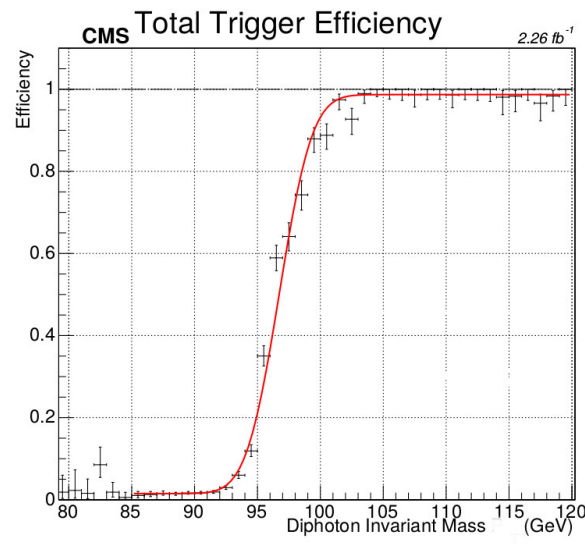
Figure 5.4: Trigger efficiency as a function of double photon invariant mass. The efficiency reaches a plateau at an invariant mass of 105 GeV.

# CHAPTER 6

# EVENT SELECTION

## 6.1 Event Selection

For the purpose of this analysis, several event categories were created for different studies. After the selection of objects, the events were classified as follows:

- photon-photon events ($\gamma\gamma$ , candidate events)

- electron-electron events ($ee$)

- fake-fake events ($ff$)

- electron-photon events ($e\gamma$)

Here fake objects refer to electromagnetically rich jets that can mimic the response of photons. Photon-photon events were the candidate signal events for this analysis. The electron-electron, fake-fake and electron-photon events were control samples used for data driven background estimates. To categorize events, the three highest $p_T$ electromagnetic objects (photon, electron or fake) in the event were put into a list and sorted in descending order of $p_T$. The selection threshold was set at 40 GeV in $p_T$. If from the first two objects, the event could be categorized in one of the above four classes, the event was put in that category. But if it was still not possible to categorize events after looking at the first two objects (for the case of electron-fake or photon-fake objects) the third object was taken into account and this object decided in which category this event fell. For example, in an event where the highest $p_T$ objects were a photon and a fake, and the third object was an electron, the event was categorized as an electron-photon event. A further cleaning was applied depending on whether the two selected objects ($O_1$, $O_2$) were well separated ($\Delta R(O_1, O_2) > 0.3$). For electron-electron events, the invariant mass of the selected objects was required to be between 75 GeV and 105 GeV, as this selects electrons from Z boson decay. For all other events, the invariant mass was required to be greater than 105 GeV. This value was chosen based on where the primary analysis trigger was fully efficient (discussed in Chapter 5). For this analysis, photons, electrons and fake objects were required to be within the barrel region of the ECAL ($|\eta| < 1.4442$). This is because
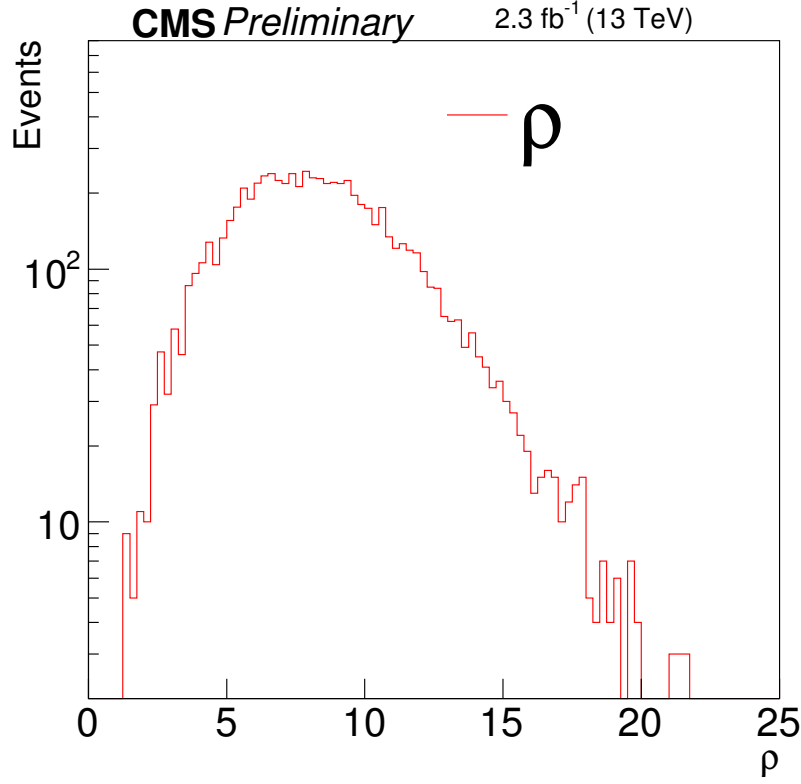
Figure 6.1: Distribution of $\rho$ in data.

in the decay of heavy supersymmetric particles, the final decay products ($\tilde{G}$ and $\gamma$ for this search) will have high transverse energy and low $\eta$.

### 6.1.1 Photon Selection

The photon selection cuts are listed in Table 6.1. The photon isolation is susceptible to pileup, and thus needs a correction. The correction for isolation is applied in the following way:

$$Iso^{Corrected} \;=\; \max(Iso^{Raw} - (\rho \times A_{eff}), 0) \tag{6.1}$$

where $Iso^{Raw}$ are the isolation variables defined in Section 4.3.1, $\rho$ represents the median of the distribution of jet energy per unit jet area for the event and $A_{eff}$ is the effective area. The distribution of $\rho$ in data is shown in Figure 6.1. The concept of effective area is discussed below.

Since $\rho$ has dimensions of energy per unit area, the $\rho$ value must be multiplied with an area to give a sensible correction for the isolation. Here, the effective area is used which is the area of

isolation region weighted by a factor that considers the dependence of the pileup transverse energy density on $\eta$ [52]. The effective area is calculated by taking the slope of a linear fit to the average isolation versus $\rho$. The effective areas for different isolation variables used in this analysis were calculated by EGamma physics object group. The effective areas used for this analysis are given in Table 6.2 [73].

Table 6.1: Photon Selection cuts

| H/E | $< 0.05$ |
|---|---|
| $\sigma_{i\eta i\eta}$ | $> 0.005$ and $< 0.0102$ |
| $R_9$ | $< 1.0$ |
| Pixel track seed veto | True |
| Particle flow neutral hadron isolation | $< 1.06 + 0.014 \times p_T^\gamma + 0.000019 \times (p_T^\gamma)^2$ GeV |
| Particle flow charged hadron isolation | $< 1.37$ GeV |
| Particle flow neutral hadron isolation | $< 0.28 + 0.0053 \times p_T^\gamma$ GeV |

Table 6.2: Effective areas (EA) for the Photon Id, Conditions: bunch crossing 25 ns

| Bin | EA Charged Hadrons | EA Neutral Hadrons | EA Photons |
|---|---|---|---|
| $|\eta| < 1.0$ | 0.0 | $0.0599 \pm 0.001$ | $0.1271 \pm 0.001$ |
| $1.0 < |\eta| < 1.479$ | 0.0 | $0.0819 \pm 0.001$ | $0.1101 \pm 0.003$ |
| $1.479 < |\eta| < 2.0$ | 0.0 | $0.0696 \pm 0.001$ | $0.0756 \pm 0.002$ |
| $2.0 < |\eta| < 2.2$ | 0.0 | $0.0360 \pm 0.001$ | $0.1175 \pm 0.002$ |
| $2.2 < |\eta| < 2.3$ | 0.0 | $0.0360 \pm 0.002$ | $0.1498 \pm 0.00001$ |
| $2.3 < |\eta| < 2.4$ | 0.0 | $0.0462 \pm 0.001$ | $0.1857 \pm 0.005$ |
| $|\eta| > 2.4$ | 0.0 | $0.0656 \pm 0.005$ | $0.2183 \pm 0.003$ |

The tracks of charged hadrons help to distinguish the charged hadrons coming from the primary vertices and the charged hadrons coming from pileup. So the energy correction for the charged hadron isolation is not essential, while for neutral hadrons and photons, the isolation must be corrected. Hence for charged hadrons, the effective areas are zero, and the effective areas for neutral hadrons and photons are non-zero.

Different variables used to identify photons (as mentioned in Table 6.1) are shown in Figures 6.2 - 6.5. To plot the distribution of a particular variable, all other selection criteria for photons apart from the cut on that variable were required. The distributions are shown for leading photon (photon having the highest $p_T$ among all photons in an event) and sub-leading photon (having the second highest $p_T$ among all photons in an event).

74
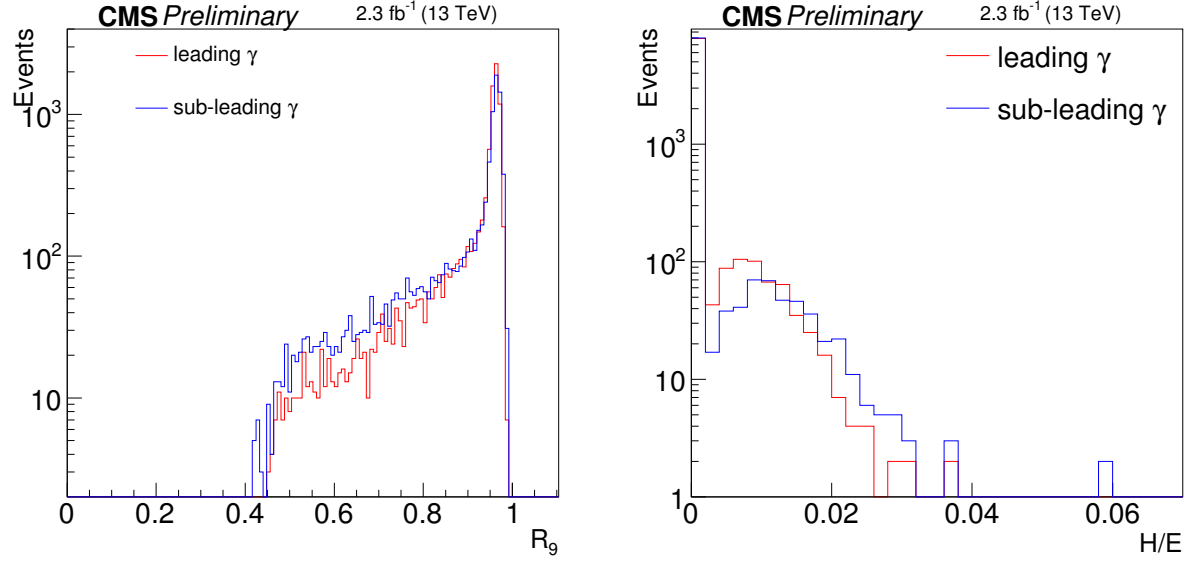
Figure 6.2: Distribution of $R_9$ and H/E in data for leading and sub-leading photon.
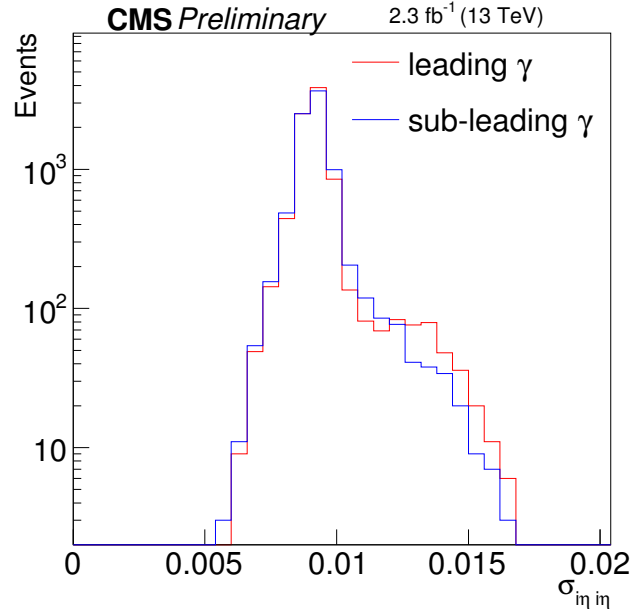


Figure 6.3: Distribution of $\sigma_{i\eta i\eta}$ in data for leading and sub-leading photon.
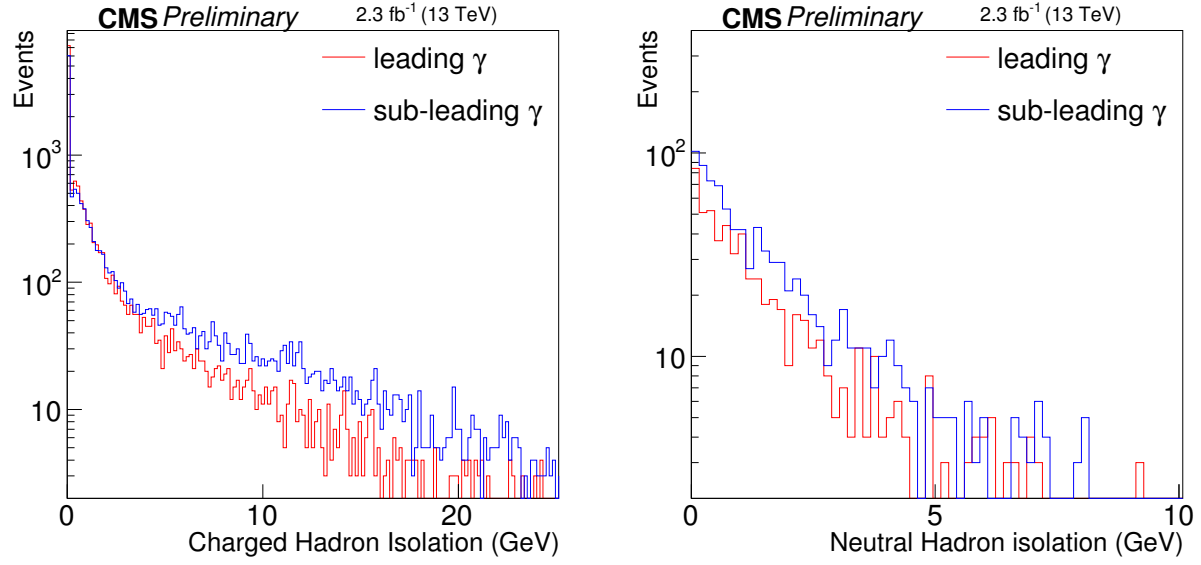
Figure 6.4: Distribution of charged hadron isolation (left) and neutral hadron isolation (right) in data for leading and sub-leading photon. $\rho$ subtraction was performed on the neutral hadron isolation.
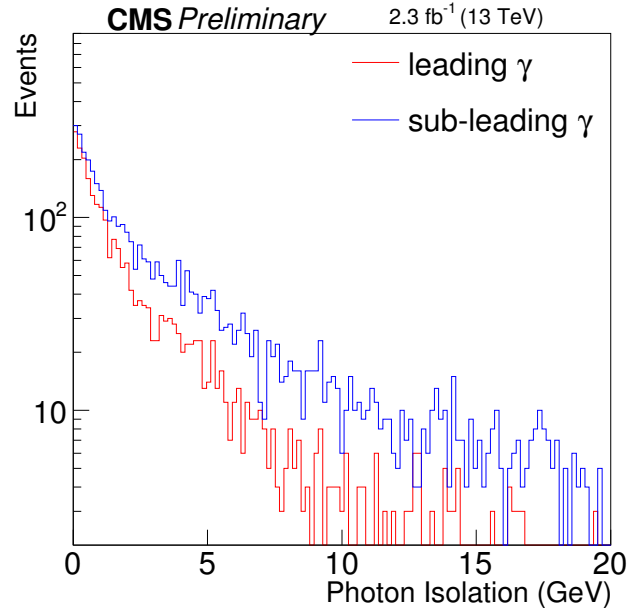


Figure 6.5: Distribution of photon isolation in data for leading and sub-leading photon. $\rho$ subtraction was performed on the photon isolation.

### 6.1.2 Electron Selection

Electron objects were selected from the photon object collection with the pixel track seed requirement on photon reversed.

### 6.1.3 Jet Selection

In this analysis, AK4PF jets (described in Section 4.6) were used when the control sample was reweighted by the jet multiplicity distribution (for details, see 7.1.2). The jet $p_T$ was required to be greater than 30 GeV and the $\eta$ requirement on jets was $|\eta| < 2.4$, where the tracker coverage ends. Jets used here were required to pass PF loose identification criteria. This PF loose identification criteria puts requirements on jet variables such as neutral (charged) hadron fraction which is the fraction of the total jet energy from neutral (charged) hadrons, neutral (charged) EM fraction which is the fraction of the total jet energy coming from neutral (charged) electromagnetic objects such as photons (electrons). There were requirements on the number of jet constituents as well as on the total charge multiplicity of the jet constituents. The requirements on the jet selection is summarized in Table 6.3.

Table 6.3: Loose identification criteria for PF jets

| | |
|---|---|
| Neutral Hadron Fraction | $< 0.99$ |
| Number of Constituents | $> 1$ |
| Neutral EM Fraction | $< 0.99$ |
| Charged Hadron Fraction | $> 0$ |
| Charged Multiplicity | $> 0$ |
| Charged EM Fraction | $< 0.99$ |

### 6.1.4 Muon Selection

Muon candidates used for this analysis were required to pass the medium muon identification criteria (given in Table 6.4, muon was required to pass either 1 or 2) and loose isolation criterion. For loose muon isolation criterion, a combined isolation was defined as follows:

$$\mathrm{Iso}^{combined} = \mathrm{Iso}^{charged} + \max(0, \ \mathrm{Iso}^{neutralHadron} + \mathrm{Iso}^{photonIso} - 0.5 \times \mathrm{Iso}^{PUIso})$$

where $\mathrm{Iso}^{charged}$ is the charged hadron isolation, $\mathrm{Iso}^{neutralHadron}$ is the neutral hadron isolation, $\mathrm{Iso}^{photonIso}$ is the photon isolation and $\mathrm{Iso}^{PUIso}$ is the the isolation energy from pileup. A muon

passed the loose muon isolation criteria if the combined isolation was less than 25% of the muon $p_T$. There were also requirements on $p_T$ ($p_T > 30$ GeV) and $\eta$ ($|\eta| < 1.4442$, to select muons in the barrel region).

Table 6.4: Muon Medium Id. To pass the medium Id, a muon must pass either criteria 1 or criterion 2

| 1. | Global muon | |
| | Normalized global-track $\chi^2$ | normalized $\chi^2 < 3$ |
| | Tracker-Standalone position match | $\chi^2$ LocalPosition $< 12$ |
| | Kick finder | track Kink $< 20$ |
| | Segment compatibility | $> 0.303$ |
| 2. | Tight segment compatibility | $> 0.451$ |

### 6.1.5 Fake Photon Selection

Fake objects originate from electromagnetically rich jets that are reconstructed as photons. In this analysis, the fake objects were used to estimate the Quantum Chromodynamics (QCD) background (for details, see Chapter 7) assuming the QCD background was coming entirely from QCD effects. To identify fake objects, most of the photon selection criteria were kept intact apart from the charged hadron isolation and $\sigma_{i\eta i\eta}$ requirements. The fake selection criteria are shown in Table 6.5. The fake objects passed either the $\sigma_{i\eta i\eta}$ or the charged hadron isolation requirement mentioned in the Table 6.5, but not both. This is because the fake objects were chosen to be a sideband to photon objects.

Table 6.5: Fake selection cuts

| | |
|---|---|
| H/E | $< 0.05$ |
| $\sigma_{i\eta i\eta}$ | $> 0.0102$ and $< 0.015$ |
| $R_9$ | $< 1.0$ |
| Pixel track seed veto | True |
| Particle flow neutral hadron isolation | $< 1.06 + 0.014 \times p_T^\gamma + 0.000019 \times (p_T^\gamma)^2$ GeV |
| Particle flow charged hadron isolation | $> 1.37$ GeV and $< 15.0$ GeV |
| Particle flow neutral hadron isolation | $< 0.28 + 0.0053 \times p_T^\gamma$ GeV |

### 6.1.6 Loose Fake Selection

To improve the statistics of fake objects in the signal region, fake objects with looser selection criteria were selected. The loose fake selection criteria are shown in Table 6.6. Here loose fake could fail either the $\sigma_{i\eta i\eta}$ or the charged hadron isolation requirement or both.

Table 6.6: Loose fake selection cuts

| H/E | < 0.05 |
|---|---|
| $\sigma_{i\eta i\eta}$ | > 0.0102 and < 0.020 |
| $R_9$ | < 1.0 |
| Pixel track seed veto | True |
| Particle flow charged hadron isolation | > 1.37 GeV and < 40.0 GeV |

The loose fake objects were selected such that the $E_T^{miss}$ distribution between fake-fake and loose fake-fake samples did not change significantly. The $E_T^{miss}$ distributions of fake-fake sample and loose fake-fake sample are shown in Figure 6.6.

The QCD background was estimated such that the overall normalization of events came from the fake samples, as they were a sideband to the photons, but the distribution in the signal region of this analysis ($E_T^{miss} > 100$ GeV) came from loose fake samples. Details of how both fake and loose fake samples were used in this analysis, are described in Section 7.1.5.

## 6.2 Photon Identification Efficiency

The efficiency of photon identification was an important factor for this analysis. The photon identification efficiency was measured in both data and simulation. The photon identification efficiency in simulation was required as the result of this search was interpreted using simulated signal samples. The scale factor was defined as the ratio of photon identification efficiency in data to simulation. The efficiencies were measured in a $Z \rightarrow ee$ sample, because for many aspects, the detector response is similar to electrons and photons.

### 6.2.1 Description of the Techniques

The measurement of object identification efficiency utilizes the decay of Z boson which is a well-established resonance. Z bosons, having invariant mass of 91 GeV, decays to two charged leptons $\sim 10\%$ of the time. One of these two reconstructed charged leptons is required to pass
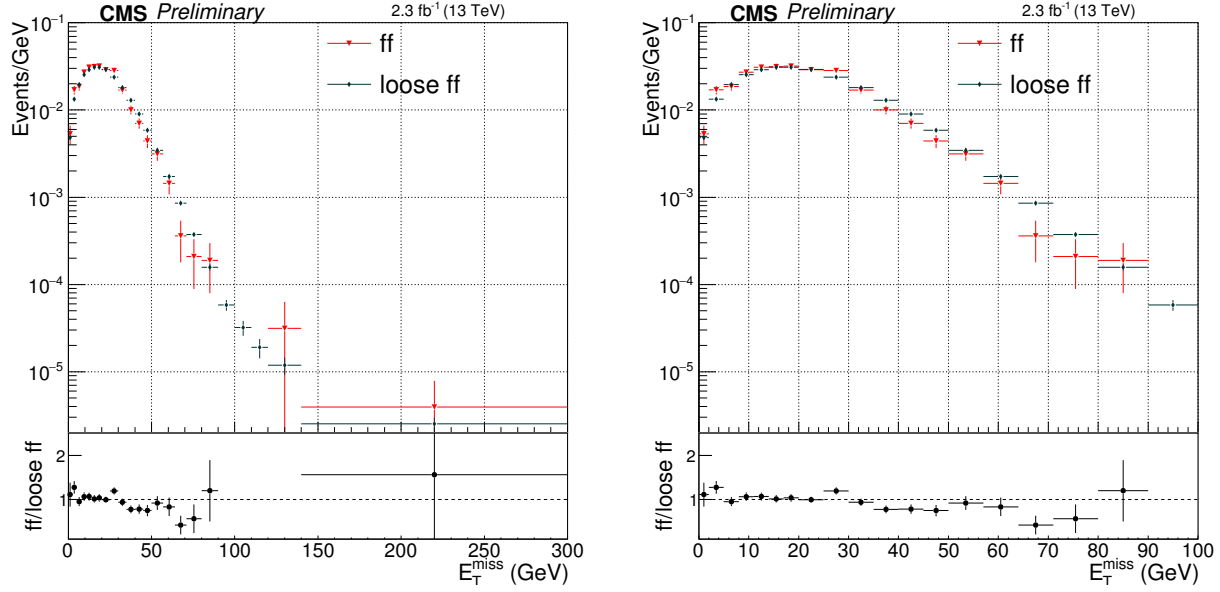
Figure 6.6: The $E_T^{miss}$ distribution of fake-fake(ff) and loose fake-fake sample. From the ratio, it is clear that both the distributions are consistent within their uncertainties. The right hand plot is the zoomed to the lower $E_T^{miss}$ ($< 100$ GeV) region.

tight selection so that the lepton is very well identified. This lepton is known as the tag. The other charged lepton, known as the probe, is required to pass only some of the selection criteria. From the fraction of the total probe leptons falling in the signal region of the Z boson mass peak, the efficiency of lepton identification can be calculated [71].

To obtain the photon identification efficiency for this analysis, events with two electrons having invariant mass between 60 GeV to 120 GeV were selected. One electron (the tag) was required to pass all of the selection criteria of an electron (mentioned in Section 6.1.2), where the other electron (the probe) was just an electron shower reconstructed as photon candidate with no requirement on the presence of a seed track in the pixel detector. The invariant mass of the tag and probe electrons was computed and used to populate the invariant mass histogram. The histogram was fit with signal and background hypotheses and from the number of signal events in the histogram, the identification efficiency was estimated. The details of this measurement are given below.

$Z \rightarrow ee$ **method:** The detector response of photons and electrons are similar, but still they are not identical. There are differences between the electron and photon response which should be corrected in order to use the scale factor for an analysis. The photon identification efficiency is

defined as:

$$\epsilon_\gamma^{data} = \epsilon_\gamma^{MC} \times \frac{\epsilon_e^{data}}{\epsilon_e^{MC}} \qquad (6.2)$$

where $\epsilon_\gamma^{data}$ and $\epsilon_\gamma^{MC}$ are the photon identification efficiencies in data and simulation respectively, $\epsilon_e^{data}$ is the electron identification efficiency in data and $\epsilon_e^{MC}$ is the electron identification efficiency in simulation. The electron identification efficiencies were measured using $Z \rightarrow ee$ tag and probe electrons.

A double electron trigger requiring invariant mass of two electrons greater than 70 GeV (see Table 5.1) was used to select good double electron events from data. A $Z \rightarrow ee$ simulated sample was used for double electron events from simulation. Two samples of tag and probe pairs were formed where in the first sample, the probe electron passed some selection criteria, and in the second one, the probe electron failed. For example, to know the identification efficiency as a function of separation from the nearest jet, the selection criterion for probe electrons was chosen such that $\Delta R(\gamma, jet)$ , the separation of a photon and the nearest jet, would be within some certain range of values. In both samples the invariant mass histograms were fit with a Crystal Ball convolved with a Breit-Wigner, fitting the signal, and an error function multiplied by an exponential for the background. The Crystal Ball function consists of a Gaussian core portion with the tail modeled by a power law, below a certain threshold. The analytical form of this function is found in [74][1].

Typical invariant mass distributions with the fits from data where probe passed or failed a certain selection criterion are shown in Figure 6.7. In this figure, the selection criterion used was $0.35 < \Delta R(\gamma, jet) < 0.85$. This was done to find the photon identification efficiency as a function of $\Delta R(\gamma, jet)$. This helped to understand if the position of the nearest jet affected the identification efficiency of the photon. The simulated signal samples may have many jets in the final state, so it was important to understand if the proximity to jets would degrade the photon identification efficiency in the signal samples.

The fraction of probes which passed the selection was defined to be the selection efficiency. This selection efficiency was measured from both data and simulation and the ratio between them gave the scale factor. The photon identification efficiencies as a function of $\Delta R(\gamma, jet)$ and number

---

[1]Crystal Ball function:

$$f(x; \alpha, n, \bar{x}, \sigma) = N \cdot \begin{cases} exp(-\frac{(x-\bar{x})^2}{2\sigma^2}) & \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\bar{x}}{\sigma})^{-n} & \frac{x-\bar{x}}{\sigma} \leq -\alpha \end{cases}$$

where $A = \left(\frac{n}{|\alpha|}\right)^n \cdot exp\left(-\frac{|\alpha|^2}{2}\right)$, $B = \frac{n}{|\alpha|} - |\alpha|$, $N = \frac{1}{\sigma(C+D)}$, $C = \frac{n}{|\alpha|} \cdot \frac{1}{n-1} \cdot exp(-\frac{|\alpha|^2}{2})$ and $D = \sqrt{\frac{\pi}{2}}\left(1 + erf\left(\frac{|\alpha|}{\sqrt{2}}\right)\right)$.
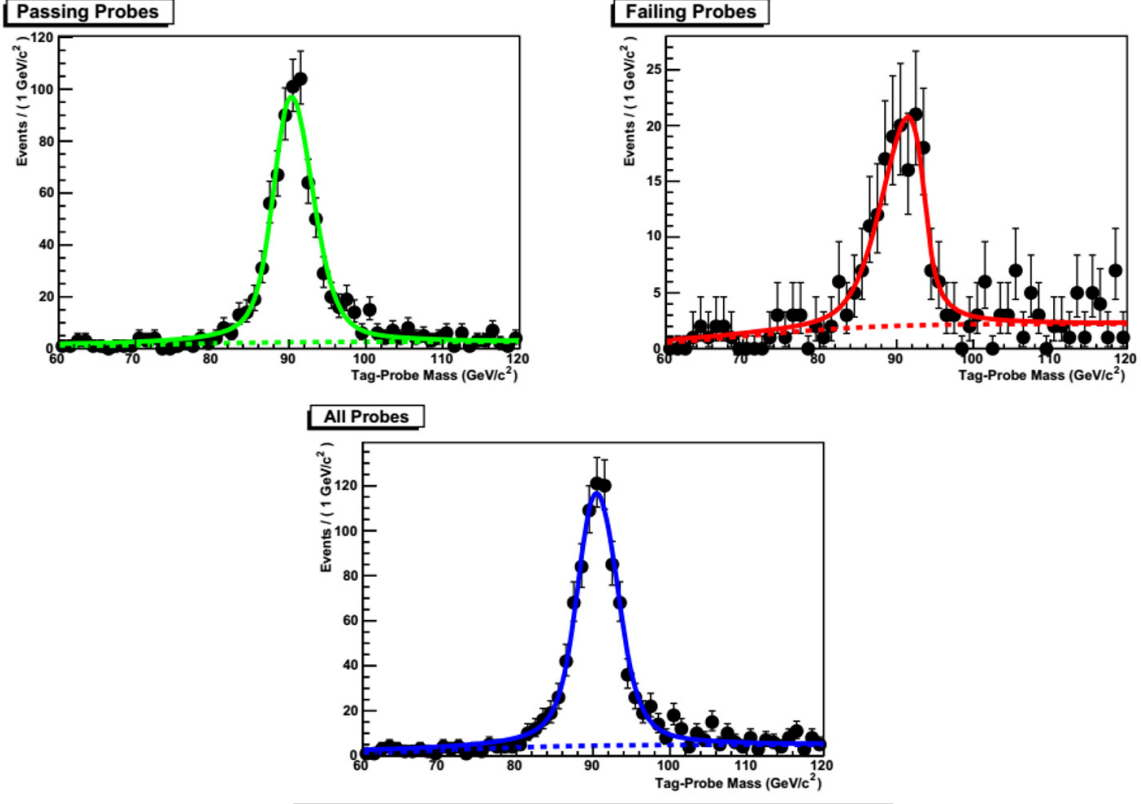
Figure 6.7: Measured $M_{ee}$ distributions in data for probe electrons passing $0.35 < \Delta R(\gamma, jet) < 0.85$, where $\Delta R(\gamma, jet)$ is the separation of a photon and nearest jet, along with the Crystal Ball convolved with Breit-Wigner fits the signal (solid line) and error function multiplied by exponential fits the background (dashed line). The distributions are shown for all probes in blue, and for probes passing (failing) the selection criterion of $\Delta R(\gamma, jet)$ in green (red). The photon identification efficiency as a function of $\Delta R(\gamma, jet)$ needs to be known to understand if the proximity of the nearest jet affects the identification efficiency.

of primary vertices are shown in Figure 6.8. The scale factors are also shown in the plots and it is clear that scale factors did not depend significantly on those variables. The overall scale factor $\frac{\epsilon_e^{data}}{\epsilon_e^{MC}}$ was found to be 0.985±0.011.
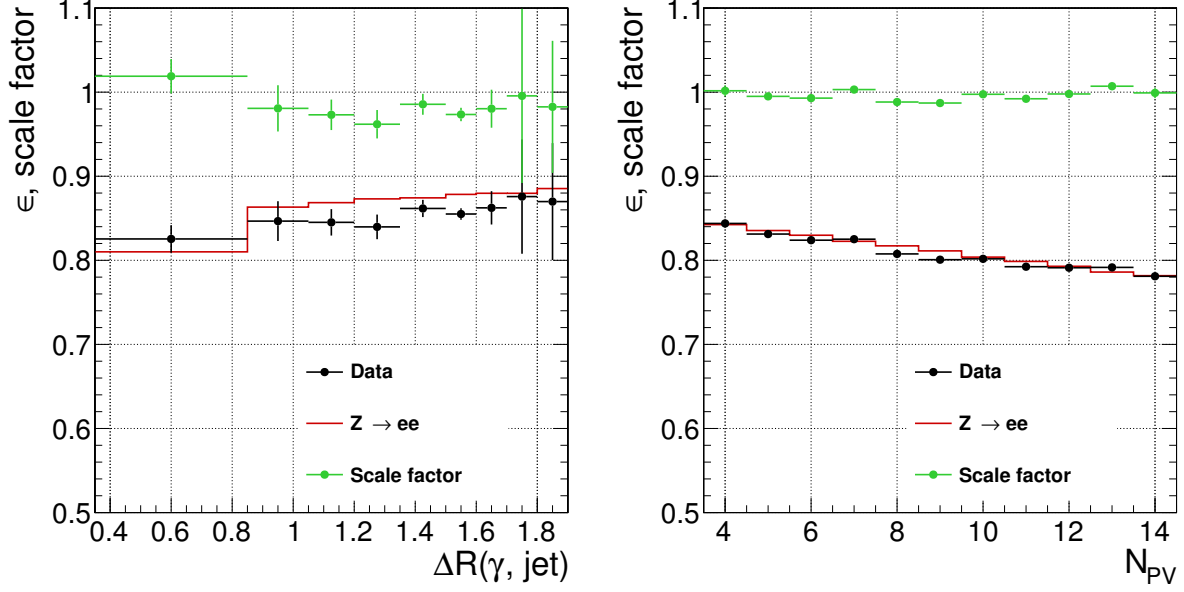


Figure 6.8: The photon identification efficiency and scale factor as a function of separation from nearest jet, $\Delta R(\gamma, jet)$ and number of primary vertices ($N_{PV}$). Errors are statistical only.

The nominal tag-and-probe region did not have any requirement on the $E_T^{miss}$ of the events. But for this analysis, the signal region was defined as $E_T^{miss} > 100$ GeV. The signal simulation samples might also have one or more jets. So it was required to investigate if the selection of electromagnetic objects in this analysis was affected by the presence of multiple jets in the signal region. The $Z \to ee$ sample chosen to calculate the identification efficiency does not have true $E_T^{miss}$, so any $E_T^{miss}$ in the $Z \to ee$ event comes from hadronic activity and jet mis-measurement in the event. $Z \to ee$ events with high value of $E_T^{miss}$ are expected to have one or more jets in the events. So these events were used to investigate if the presence of one or more jets in the signal region affected the identification efficiency of the photons.

A direct comparison of photon identification efficiency in the nominal tag and probe region and $E_T^{miss} > 100$ GeV region is shown in Figure 6.9. It is clear that there was no statistically

significant trend to the ratio of the two efficiencies. This ratio was consistent with a constant factor of $0.99 \pm 0.01$.
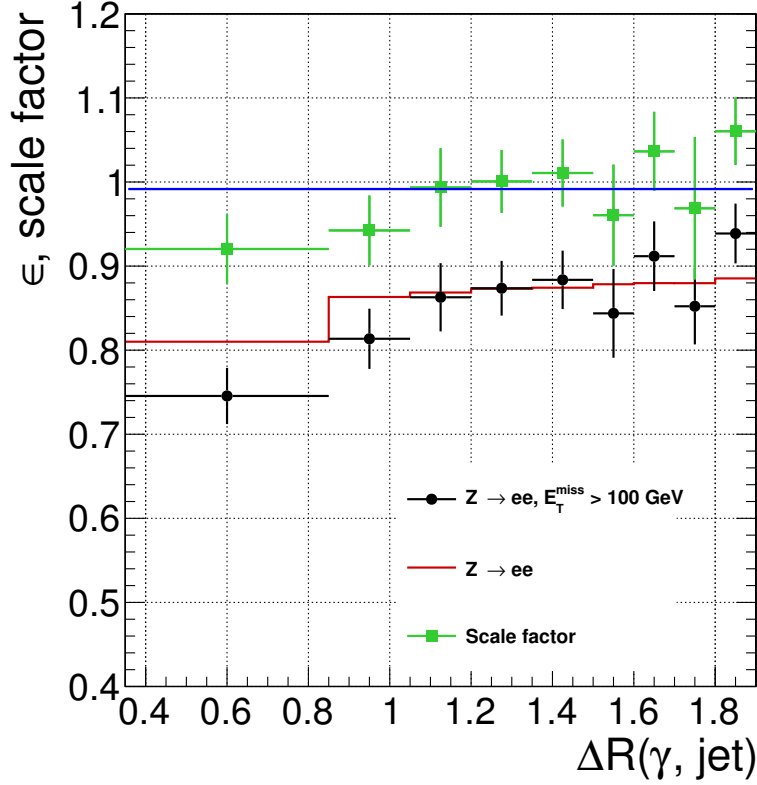


Figure 6.9: Comparison of the simulated efficiencies with respect to $\Delta R(\gamma, jet)$ in the tag-and-probe region and the signal region of the analysis with $E_T^{miss} > 100$ GeV. The blue line is the best constant fit to the scale factor.

**Pixel seed veto efficiency:** The prescription for photon identification is very similar to that for electron identification. It differs only by the presence or absence of a seed track in the pixel detector. A consequence of this is that the efficiency of the pixel seed veto for photons cannot be determined from the tag-and-probe method described above. The pixel seed veto efficiencies were obtained from photons in $Z \to \mu\mu\gamma$ events. In these events, the photon was radiated from one of the muons (final state radiation), so the invariant mass of the two muons and the photon should fall within the Z boson mass window. Here two fully reconstructed muons (the selection criteria for muons are described in Table 6.4) were required to be well separated from the photon ($\Delta R(\mu, \gamma) > 0.5$). The requirement on the three-body invariant mass was $60 < M_{\mu\mu\gamma} < 120$ GeV.

To obtain the pixel seed veto efficiency, first the $\mu\mu\gamma$ invariant mass histogram where the photon passed the pixel seed veto was fit. The signal was fit by Crystal Ball convolved with Breit-Wigner and background was fit with exponential times error function. Next, the $\mu\mu\gamma$ invariant mass where photon passing pixel seed veto was not required was fit with the same functional form. The fraction of these events where the photon passed the pixel seed veto selection criterion was taken as the efficiency, which was measured to be $94.8 \pm 0.3\%$. The data/simulation scale factor was found to be $1.00 \pm 2.5\%$. Typical fits of the $\mu\mu\gamma$ invariant mass distributions where the photon passes pixel seed veto and requires no pixel seed veto are shown in Figure 6.10.
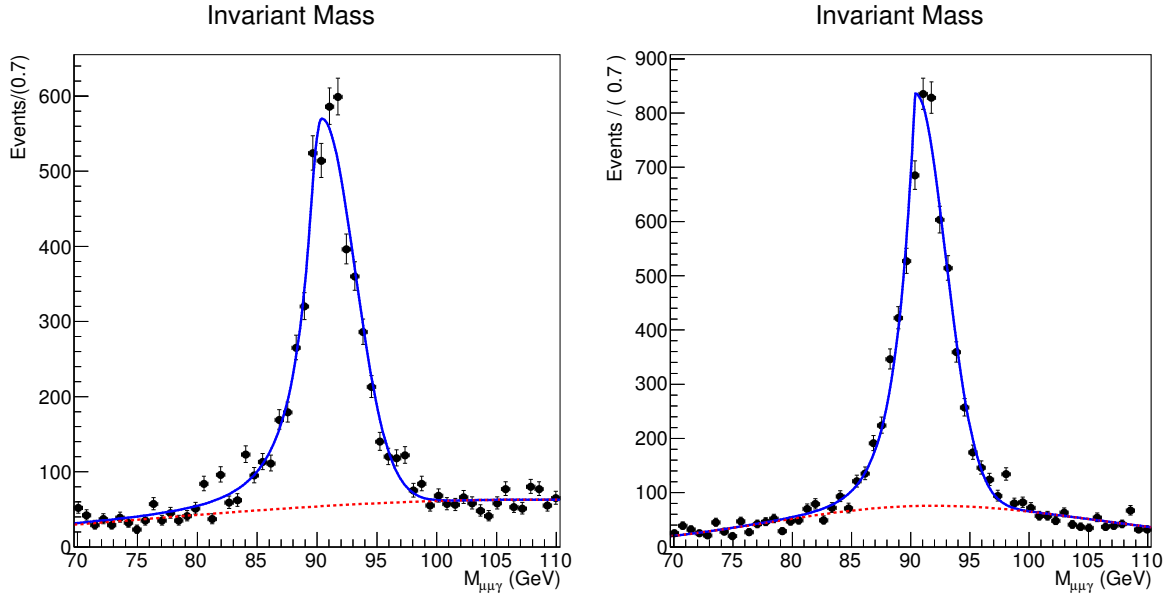


Figure 6.10: The $\mu\mu\gamma$ invariant mass distribution where the signal is fit with Crystal Ball convolved with Breit-Wigner and the background is fit with error times exponential function. Photons passing the pixel seed veto is shown on the left, and for the right plot, no requirement of photon passing pixel seed veto is needed.

# CHAPTER 7

# BACKGROUND ESTIMATION

The final state in this analysis contains two high $p_T$ photons and large missing transverse energy ($E_T^{miss}$ ). Any Standard Model process having this signature is the background for this analysis. The dominant background arises from Quantum Chromodynamic (QCD) processes where two protons collide and produce photons and jets in the final state. An electromagnetically rich jet fragmentation can mimic the response of a photon. This is called 'fake photon'. QCD processes can produce real or fake photons. These processes are labeled as QCD background. They have instrumental $E_T^{miss}$ , which means that the $E_T^{miss}$ arises because of object mis-measurement. A sub-dominant background arises from W-$\gamma$/W-jet events where the W$^{\pm}$ decays to positron or electron and neutrino. This background is labeled as electroweak (EWK) background as it arises from the electroweak interaction. This kind of background has genuine $E_T^{miss}$ as neutrinos are produced in the final state. There are irreducible backgrounds from Z+$\gamma\gamma$ , where Z$\rightarrow \nu\nu$, but in an integrated luminosity of 2.3 fb$^{-1}$, the effect is negligible (less than 1 event overall). Once these backgrounds were modeled, the background $E_T^{miss}$ distinction was compared with the candidate double photon $E_T^{miss}$ distribution. The SUSY signal is expected to manifest as an excess at high $E_T^{miss}$ . The background was modeled while the data was blinded, meaning the data in the signal region (high $E_T^{miss}$ region) were not looked at before the background modeling was performed.

## 7.1   QCD Background

The background to this analysis arising from QCD processes includes di-jet, $\gamma$+jet and diphoton events. It is difficult to accurately simulate the $E_T^{miss}$ of the detector when jets fragment and mimic photons, because for that knowledge of every aspect of jet detection and fragmentation would be required. Also even there are two real photons in the final state, the $E_T^{miss}$ may be mis-measured because of the additional hadronic activity. Because of this, a data driven technique was used to model these backgrounds. Two control samples were chosen - double electron ($ee$) and double fake ($ff$) - to model this background. These samples were chosen because like the QCD background,

they also do not have genuine $E_T^{miss}$ and they each have electromagnetic objects in the final state. The hadronic activity is different in candidate and control samples. This is because the control sample $ee$ comes from Z boson decays, not from any QCD vertex. The additional hadronic activity of the candidate sample has an effect on its $E_T^{miss}$ distribution, so it can be significantly different from that of the control samples. The hadronic recoil of the samples are modeled by the vector sum of two electromagnetic objects in the event. This variable is referred to as the di-EM $p_T$ of the event.

### 7.1.1  Di-EM $p_T$ Reweighting

The di-EM $p_T$ is defined by this equation:

$$p_T^{di-\vec{E}M} = \vec{p_{T1}} + \vec{p_{T2}} \tag{7.1}$$

where $\vec{p_{T1}}$ and $\vec{p_{T2}}$ are the transverse momenta of two electromagnetic objects in the event. Figure 7.1 shows a graphical representation of the di-EM $p_T$. From momentum conservation in the transverse direction, it is clear that di-EM $p_T$ will balance the total transverse momenta of jets in the event having no genuine $E_T^{miss}$ . The resolution of measuring the energy of electromagnetic objects is much better than that of jets. For this reason, the di-EM $p_T$ is used to measure the hadronic recoil in such events.

Any difference in hadronic activity between candidate and control samples must be corrected, because the hadronic activity affects the visible energy resolution of an event. To achieve this, the $E_T^{miss}$ distribution of the $ee$ control sample was reweighted by the ratio of the candidate $\gamma\gamma$ di-EM $p_T$ to that of the $ee$ sample in order to make the hadronic activity of $ee$ control sample similar to that of the candidate sample. The same procedure was followed for the $ff$ control sample as well. The di-EM $p_T$ distribution of the candidate and control samples and the corresponding ratios are shown in Figure 7.2.

The effect of reweighting can be seen from Figure 7.3 where the control sample $E_T^{miss}$ distributions before and after the reweighting are shown. The reweighted $E_T^{miss}$ distributions of the control samples were compared with the $E_T^{miss}$ of $\gamma\gamma$ sample. This comparison is shown in Figure 7.4. From these comparisons, it is clear that the di-EM $p_T$ reweighting of the control samples makes the $E_T^{miss}$ distributions of candidate and control samples very similar. This gave the confidence that di-EM $p_T$ reweighting was indeed able to correct for the hadronic activity in the control samples.
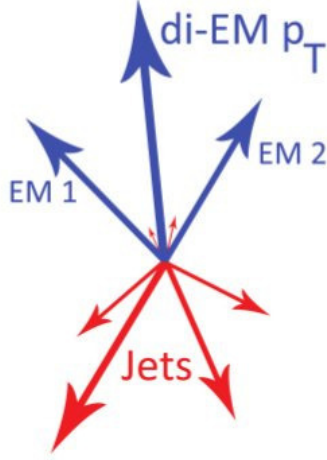
Figure 7.1: Graphical representation of di-EM $p_T$

The error propagation from di-EM $p_T$ reweighting is discussed in Section 7.1.7.

### 7.1.2 Reweighting by Jet Multiplicity

Apart from the difference in hadronic activity, the difference in jet multiplicity between candidate and control samples can also potentially affect the overall $E_T^{miss}$ distribution. For example, the energy resolution of an event where there are 3 jets having a total $p_T$ of 100 GeV will be worse than that of an event where there is only one jet having $p_T$ of 100 GeV. This is because the jet energy resolution improves with increasing energy (for example, see Figure 4.2). The jet multiplicity distributions of candidate and control samples are shown in Figure 7.5. This distributions are similar between candidate $\gamma\gamma$ sample and $ff$ control sample, but they are noticeably different between $\gamma\gamma$ and $ee$ samples. To extract any possible correlation between the jet multiplicity and the di-EM $p_T$ , the jet multiplicity ratio of the $\gamma\gamma$ to the $ee$ sample in bins of di-EM $p_T$ was plotted (Figure 7.6). To investigate the effect of jet multiplicity reweighting on the $ee$ $E_T^{miss}$ distribution, first the di-EM $p_T$ reweighted $E_T^{miss}$ distribution was plotted. Then on the same plot (Figure 7.7), the reweighted $E_T^{miss}$ distribution of the $ee$ sample where the weights were coming from the two dimensional ratio of candidate $\gamma\gamma$ to the $ee$ sample in the di-EM $p_T$ and the jet multiplicity (from
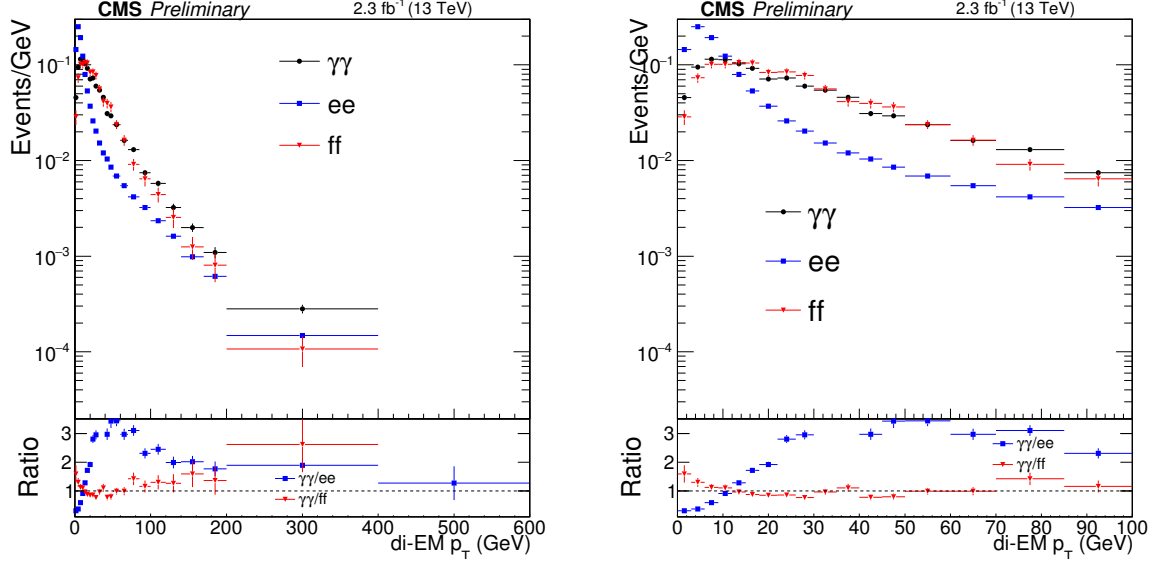
Figure 7.2: Di-EM $p_T$ distributions of $\gamma\gamma$, $ee$ and $ff$ samples. On the right, the same distributions are shown zoomed in to the lower di-EM $p_T$ region.

Figure 7.6) was plotted. From this figure, one can see the effect of jet multiplicity reweighting on $ee$ $E_T^{miss}$ distribution is negligible. So the jet multiplicity reweighting on $ee$ $E_T^{miss}$ distribution was not performed, rather the difference between the two $ee$ estimates (one where $E_T^{miss}$ was reweighted by only di-EM $p_T$ and the other one where $E_T^{miss}$ was reweighted by jet multiplicity in bins of di-EM $p_T$ ) was taken as a source of systematic uncertainty.

### 7.1.3 Subtraction of Other Contributions from the Control Samples

The control samples were chosen such that they do not have any genuine $E_T^{miss}$, but there are other processes containing genuine $E_T^{miss}$ which may contaminate the control samples. Such processes include $t\bar{t}$ events contaminating the $ee$ sample where the top quark decays leptonically, giving rise to neutrinos in the final state. Another possible process which can contaminate the $ff$ control sample is Z+jets, where Z boson decays to two neutrinos, and the two jets pass the fake event selection.

The possible contaminations from these processes were checked using the simulation. The contamination from the Z+jets sample in the $ff$ control sample was negligible (less than 0.1 event).
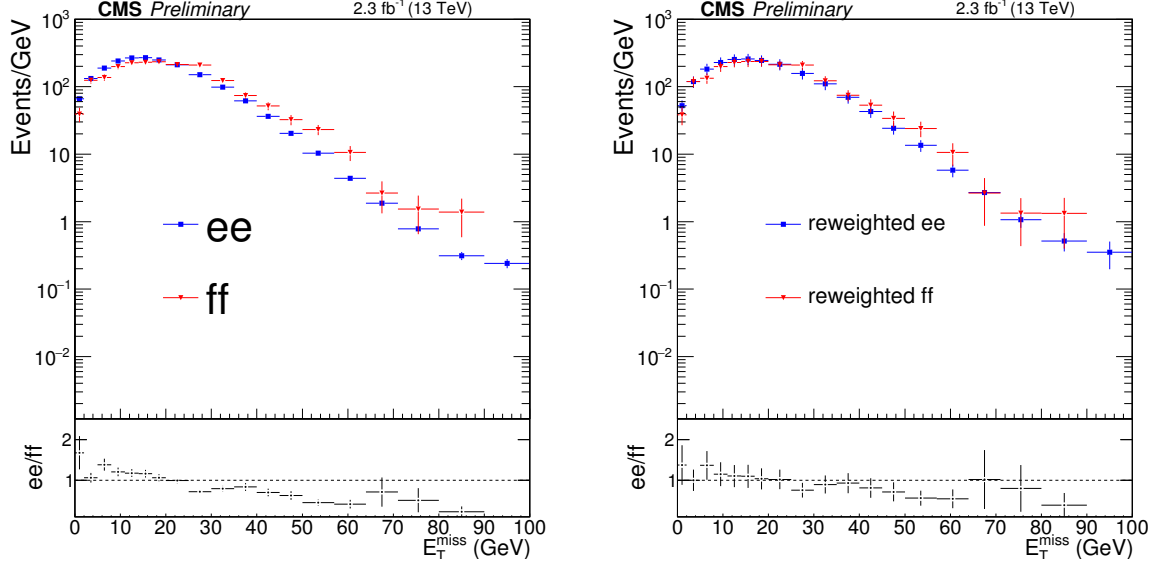
Figure 7.3: $E_T^{miss}$ distributions of the $ee$ and $ff$ control samples. The unweighted distributions are shown on the left and di-EM $p_T$ reweighted distributions are shown on the right. Since the statistics between the $ee$ and $ff$ samples are different, the plots are normalized to the $E_T^{miss} < 50$ GeV region of the candidate $\gamma\gamma$ sample.

But the contamination from $t\bar{t}$ sample in the $ee$ control sample was significant ($17.27\pm0.98$ events in the $E_T^{miss} > 100$ GeV region), so the shape of the $t\bar{t}$ was subtracted from the $ee$ control sample.

### 7.1.4  Di-EM $p_T$ Reweighting Method for $ff$ Control Sample

The $ee$ control sample was used to give an estimate of the background coming from events where both of the objects categorized as photons were actually electromagnetic objects. But the candidate $\gamma\gamma$ sample also includes $\gamma$+jet and di-jet events where one or both of the objects come from jet fragmentation. The $ff$ control sample was used to estimate the QCD background assuming that the QCD background was coming from events where one or two photon-like objects were actually jet fragmentations. The di-EM $p_T$ reweighting method for $ff$ sample followed similar strategy used for $ee$ sample.

Unfortunately, there was an additional complication in using the $ff$ sample: There were only two $ff$ events in the $E_T^{miss} > 100$ GeV region. So it was impossible to get the $E_T^{miss}$ distribution from $ff$ sample in this region. To solve this problem, the fake selection definition was loosened (described in Chapter 6) in order to get more $ff$ statistics at high $E_T^{miss}$ . It was also verified that
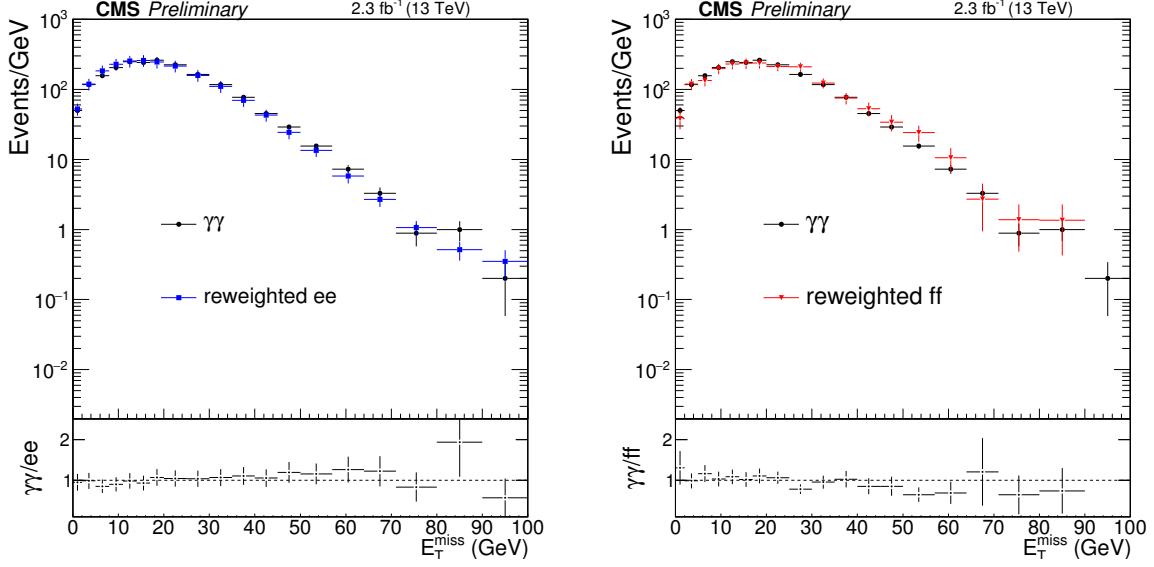
Figure 7.4: $E_T^{miss}$ distributions of the candidate sample, the di-EM $p_T$ reweighted $ee$ sample and the reweighted $ff$ sample. The comparison between the $\gamma\gamma$ sample and the reweighted $ee$ sample is shown on the left and between the $\gamma\gamma$ sample and the reweighted $ff$ sample is shown on the right. Since the statistics of the $ee$ and $ff$ samples are different, the plots here are also normalized to the $E_T^{miss} < 50$ GeV region of the $\gamma\gamma$ sample.

the $E_T^{miss}$ shape of original $ff$ sample and loose $ff$ sample were not significantly different (Figure 6.6).

To get the estimate from the loose $ff$ sample, the di-EM $p_T$ weighted $E_T^{miss}$ distribution was used. It was verified if events having genuine $E_T^{miss}$ contaminated the loose $ff$ sample. The contribution of loose $ff$ events from the Z+jets sample estimated from the simulation was found to be 15.8±0.9 in $E_T^{miss} > 100$ GeV. So the shape of Z+jets was subtracted from the loose $ff$ sample. The di-EM $p_T$ distribution of the candidate and loose $ff$ samples are shown in Figure 7.8. The comparison of reweighted $E_T^{miss}$ distributions of $ff$ and loose $ff$ samples are shown in Figure 7.9. From the ratio plot of this figure, it is clear that the reweighted $E_T^{miss}$ distribution of $ff$ and loose $ff$ samples were consistent within their uncertainties in the control region ($E_T^{miss} < 100$ GeV). Since the control sample $ff$ is not expected to have signal contamination, the distribution of $E_T^{miss}$ should not change significantly between control region and signal region ($E_T^{miss} > 100$ GeV). Hence the loose $ff$ sample was as good as the original $ff$ sample in estimating the QCD background.
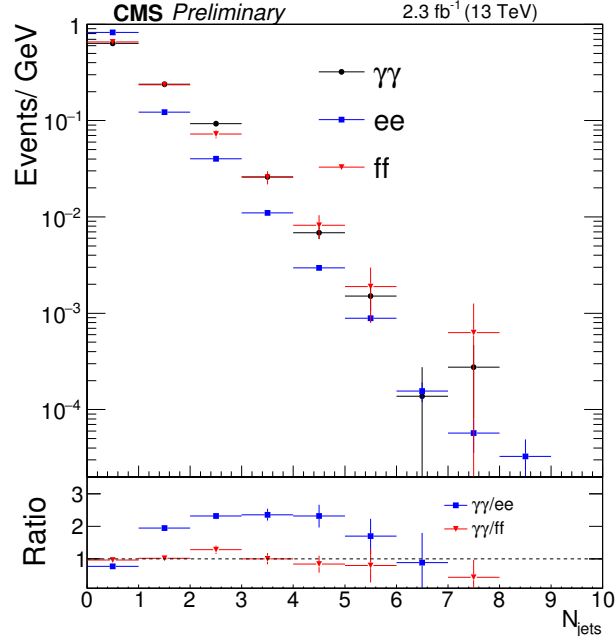
Figure 7.5: Jet multiplicity distributions of the candidate $\gamma\gamma$ , $ee$ and $ff$ samples.
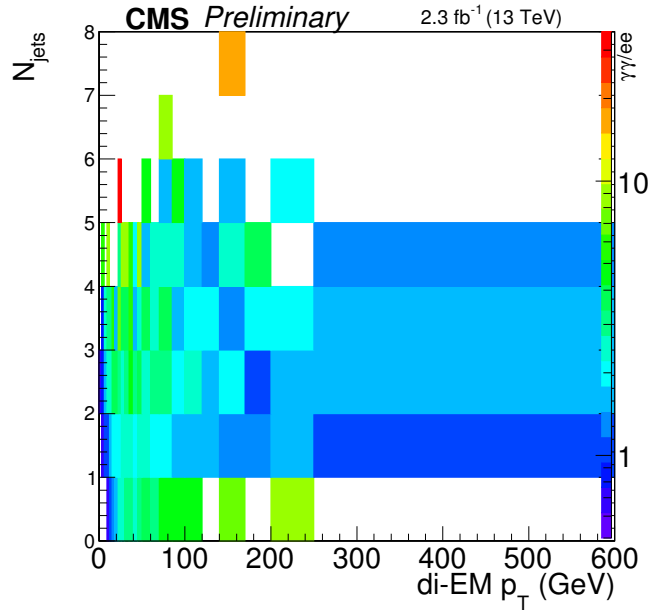


Figure 7.6: Two dimensional ratio of the candidate $\gamma\gamma$ to the $ee$ control sample as a function of the jet multiplicity and the di-EM $p_T$ .
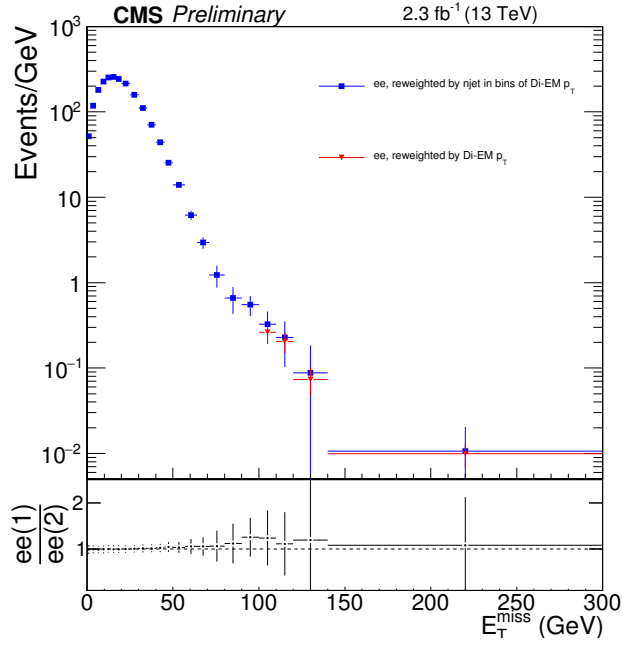
Figure 7.7: $E_T^{miss}$ distributions for the *ee* sample reweighted by di-EM $p_T$ only and $E_T^{miss}$ distribution reweighted by the 2D distribution of di-EM $p_T$ vs jet multiplicity. The ratio plot shows the ratio of blue points to red points and this shows that the distributions are consistent within their uncertainties.
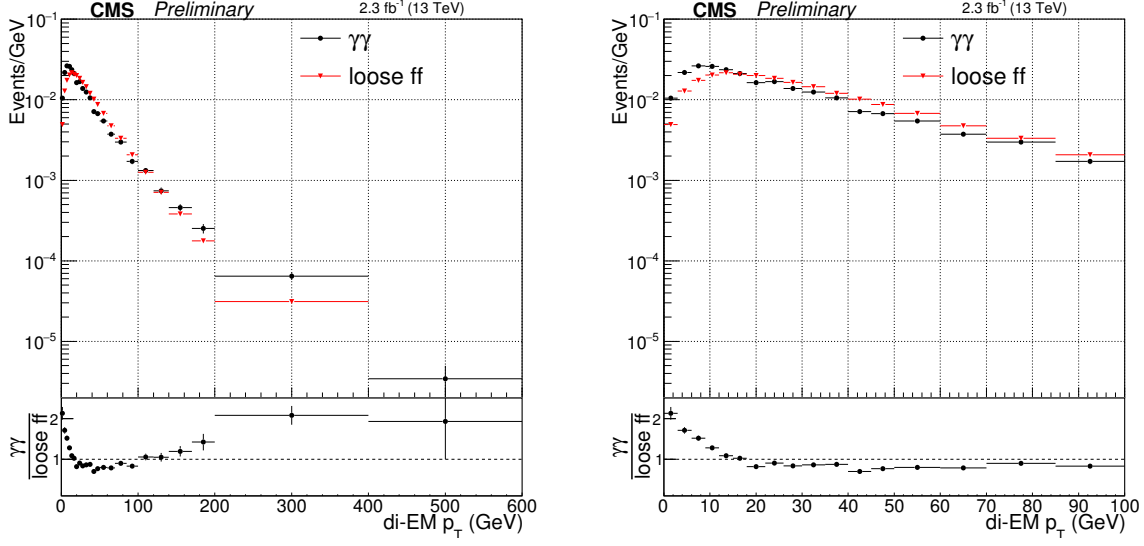
Figure 7.8: Di-EM $p_T$ distributions of candidate $\gamma\gamma$ and loose $ff$ samples. On the right, the same distributions are shown zoomed in to the lower di-EM $p_T$ region.

### 7.1.5   Cross Check on di-EM $p_T$ Reweighting Method

The assumption behind the method of reweighting by the di-EM $p_T$ of the control samples is that the di-EM $p_T$ is an accurate measure of the differences in hadronic recoil between candidate and control samples. This assumption can be validated if an independent method of background estimation with different assumptions is used and then the estimation from di-EM $p_T$ reweighting method is compared with that obtained from this new method. If the two processes give background estimates consistent within their uncertainties, it can be assumed that the di-EM $p_T$ reweighting method gives a credible estimate of the background without genuine $E_T^{miss}$ .

The independent background method relies on the observation that $ff$ sample is just the sideband (in photon identification) to the candidate $\gamma\gamma$ sample. It is not expected that the $E_T^{miss}$ distribution of an event will depend on which electromagnetic object the event has: either photon or fake. So the relative fraction of $\gamma\gamma$ and $ff$ events should not depend sensitively on $E_T^{miss}$ , if there is no signal. The ratio of $\gamma\gamma$ and $ff$ $E_T^{miss}$ distribution can be modeled with a simple function to account for any weak dependence. If the functional form for the $\gamma\gamma$ /$ff$ ratio can be found as a simple function of $E_T^{miss}$ in the low $E_T^{miss}$ ($< 100$ GeV) region, one can extrapolate this function to the signal region ($E_T^{miss} > 100$ GeV) and multiply it by the $ff$ events obtained in
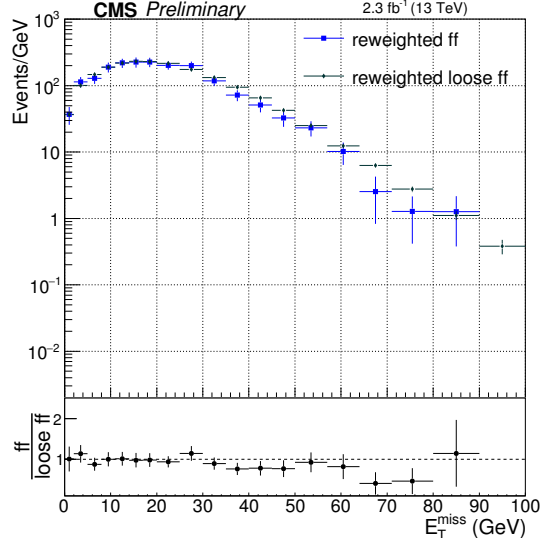
Figure 7.9: $E_T^{miss}$ shapes of the di-EM $p_T$ reweighted $ff$ and loose $ff$ samples are shown. The ratio plot in the bottom pad shows that the $E_T^{miss}$ shapes of these samples are consistent within the uncertainties.

signal region. This will give an estimation of the QCD background in the signal region independent from the di-EM $p_T$ reweighting method. This method is referred to as the $\gamma\gamma$ /$ff$ ratio method of estimating QCD background.

The original fake sample, as it was the sideband to the $\gamma\gamma$ sample, was used to get the functional form of $\gamma\gamma$ /$ff$ as a function of $E_T^{miss}$ and the overall normalization ( i.e. total number of $ff$ events expected ) in the signal region. As there were only two $ff$ events in the signal region, the loose $ff$ sample with increased statistics was used to get the shape of the distribution in the signal region. The $\gamma\gamma$ /$ff$ ratio was fit with a simple exponential function of the form $\exp(p_0\mathrm{x}+p_1)$ where parameters $p_0$ and $p_1$ were the fit parameters. This fit is shown in Figure 7.10.

The background prediction from the di-EM $p_T$ reweighting method and the $\gamma\gamma$ /$ff$ ratio method are compared in Table 7.1. The uncertainties in the ratio method were obtained by varying the parameters a and b by $\pm 1\sigma$ and the uncertainties for di-EM $p_T$ reweighting method is described in Section 7.1.7. The two methods of estimating the QCD background were overlapping within their uncertainties, but the ratio method gave a systematically smaller estimate than that of the di-EM $p_T$ reweighting method. So the di-EM $p_T$ reweighting method which was more conservative

95
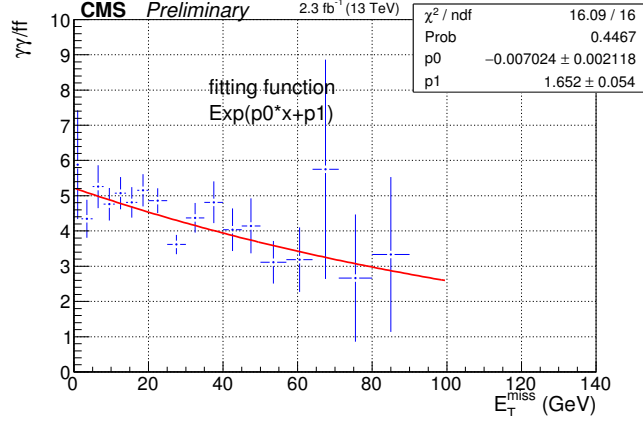
Figure 7.10: Candidate $\gamma\gamma$ over $ff$ $E_T^{miss}$ ratio fit to an exponential function ($\exp(ax+b)$) in the control region ($E_T^{miss} < 100$ GeV).

estimate was chosen to estimate the QCD background. The estimates from the ratio method were only used as a cross-check, they were not used for the final result.

Table 7.1: Estimation of QCD background for $E_T^{miss} > 100$ GeV using the $ff$ control sample

| $E_T^{miss}$ bin (GeV) | Method | Value |
|---|---|---|
| $100 - 110$ | Di-EM $p_T$ reweighting method | $1.97 \pm 0.69$ |
| | $\gamma\gamma$ /$ff$ ratio method | $1.12 \pm 0.59$ |
| $110 - 120$ | Di-EM $p_T$ reweighting method | $1.12 \pm 0.48$ |
| | $\gamma\gamma$ /$ff$ ratio method | $0.61 \pm 0.35$ |
| $120 - 140$ | Di-EM $p_T$ reweighting method | $1.53 \pm 0.76$ |
| | $\gamma\gamma$ /$ff$ ratio method | $0.68 \pm 0.43$ |
| $> 140$ | Di-EM $p_T$ reweighting method | $2.05 \pm 1.32$ |
| | $\gamma\gamma$ /$ff$ ratio method | $0.91 \pm 0.70$ |

### 7.1.6   Final QCD Background Estimation

As the di-EM $p_T$ reweighting method of estimating background using $ee$ and $ff$ samples were consistent within uncertainties, the estimate coming from di-EM $p_T$ reweighted $ee$ sample was used as the central value of prediction for the QCD background. The shape difference between di-EM $p_T$ reweighted $ee$ and $ff$ samples was used as a source of systematic uncertainty on the QCD

background. The procedure to obtain the systematic uncertainty from shape difference of $ee$ and $ff$ sample is described below.
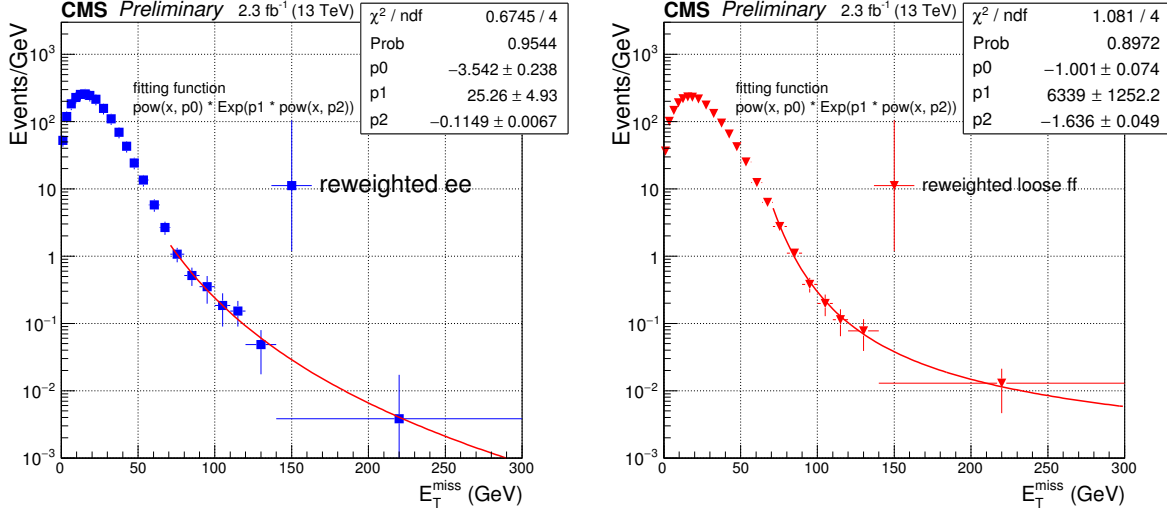


Figure 7.11: Di-EM $p_T$ Reweighted $ee$ and $ff$ $E_T^{miss}$ distributions fit with $x^{p_0} \times \exp(p_1 \cdot x^{p_2})$ (here $x$ is the $E_T^{miss}$ ) from 70 GeV to 300 GeV.

Table 7.2: Shape uncertainty coming from the difference between the $ee$ and $ff$ $E_T^{miss}$ distributions for $E_T^{miss} > 100$ GeV

| $E_T^{miss}$ bin (GeV) | $ee$ prediction | $ff$ prediction | difference | fractional difference |
|---|---|---|---|---|
| $100 - 110$ | 1.87 | 2.21 | 0.34 | 18.18% |
| $110 - 120$ | 1.16 | 1.30 | 0.14 | 12.07% |
| $120 - 140$ | 1.25 | 1.43 | 0.18 | 14.40% |
| $> 140$ | 1.39 | 3.48 | 2.09 | 150.36% |

The di-EM $p_T$ reweighted $ee$ and $ff$ $E_T^{miss}$ distribution were each fit with a function in the range 70-300 GeV of the form $x^{p_0} \times \exp(p_1 \cdot x^{p_2})$, where $p_0$, $p_1$ and $p_2$ were obtained from the fit. The fit results are shown in Figure 7.11. The functions used to fit the di-EM $p_T$ reweighted $ee$ and $ff$ $E_T^{miss}$ were integrated in each signal bin (100-110, 110-120,120-140 and beyond 140 GeV) and their difference in each signal bin gave the shape uncertainty in that bin. The fractional shape uncertainty in each bin was calculated as the ratio of the shape difference uncertainty in each bin to the total number of $ee$ events coming from integrating the $ee$ fit function in that bin. The shape

difference and the fractional shape difference uncertainties are tabulated in Table 7.2. The shape difference uncertainty for the control region ($E_T^{miss} < 100$ GeV) was not used in this analysis. Only for plotting, in this region, a flat extrapolation of fractional shape difference uncertainty of the bin 100-110 GeV (which is 18.18%) was performed to obtain the uncertainty band.

The total QCD background in the signal region with the corresponding uncertainties (both systematic and statistical) are shown in Table 7.3. The uncertainty break down for each bin is shown in Section 7.1.7.

Table 7.3: Estimation of total QCD background for $E_T^{miss} > 100$ GeV

| $E_T^{miss}$ bin (GeV) | Background Prediction |
| --- | --- |
| $100 - 110$ | $1.85 \pm 0.96$ |
| $110 - 120$ | $1.53 \pm 0.63$ |
| $120 - 140$ | $0.97 \pm 0.62$ |
| $> 140$ | $0.61 \pm 2.15$ |

### 7.1.7  Systematic Uncertainties on QCD Background Estimation

The QCD background was estimated using events having no true $E_T^{miss}$ . The central estimate of the QCD background was given by the reweighted $ee$ control sample.

**Systematic uncertainty coming from $ee$ sample:** To get the estimate from the $ee$ sample, the $E_T^{miss}$ distribution of $ee$ sample was di-EM $p_T$ reweighted so that its hadronic activity match with that of the candidate $\gamma\gamma$ sample. This reweighting by the di-EM $p_T$ is a source of systematic uncertainty.

To propagate the statistical uncertainty from the di-EM $p_T$ ratio, a thousand different di-EM $p_T$ ratios of candidate to control ($ee$) sample were generated. In the new ratio plots, the value of each bin was obtained by varying the value of the same bin in the original di-EM $p_T$ ratio using a Gaussian distribution where the spread of the distribution was dictated by the statistical uncertainty of that bin. This way the new di-EM $p_T$ ratios had a new central value but the statistical uncertainty remained the same. The $E_T^{miss}$ distribution of the $ee$ control sample was reweighted by these thousand new di-EM $p_T$ ratios and the uncertainty was then determined from the variation in each $E_T^{miss}$ bin using the thousand $E_T^{miss}$ plot generated by this procedure.

Apart from the di-EM $p_T$ reweighting, there is another source of systematic uncertainty due to jet multiplicity reweighting. But as described in Section 7.1.2, the effect of jet multiplicity is small; so the jet multiplicity reweighting on $ee$ $E_T^{miss}$ was not performed. Instead, the difference between the $ee$ estimate where $E_T^{miss}$ was only di-EM $p_T$ reweighted and $ee$ estimate where the weighting factors on $E_T^{miss}$ were from Figure 7.6 was taken as a systematic uncertainty.

**Systematic uncertainty coming from $ff$ sample:** When the QCD background from $ee$ sample was estimated, implicitly it was assumed that only true electromagnetic objects were contributing to this background. But this is not necessarily true and a large contribution can come from jet fragmentations mimicking the response of photons. In order to account for the difference in sample purity, the estimate from $ff$ sample was also obtained in a similar di-EM $p_T$ reweighting method which was used for the $ee$ sample. The uncertainty coming from di-EM $p_T$ reweighting of $ff$ sample was estimated in exactly the same way it was determined for $ee$ sample. This way two estimates for QCD background were obtained: One assumes that QCD background is composed of pure electromagnetic objects and the other one assumes that QCD background is coming from jets mimicking the response of photon. Their difference (described in Section 7.1.6) gave a systematic uncertainty which accounts for the difference of sample purity of QCD background. The uncertainty values are listed in Table 7.2.

## 7.2   Electroweak Background Estimation

Electroweak background mainly comes from $W\gamma$ events where the W decays leptonically to an electron and a neutrino. If the electron is mis-identified as a photon, events with two photons and $E_T^{miss}$ in the final state are produced. W+jets events can also contribute when the W decays to electron and neutrino and one jet fragments and mimics the response of a photon.

To estimate the electroweak background, the $e\gamma$ control sample was used. Then the electron to photon mis-identification rate ($f_{e\rightarrow\gamma}$ , colloquially referred as fake rate) was estimated. To find $f_{e\rightarrow\gamma}$ , the $Z \rightarrow ee$ invariant mass peak was plotted in both the $ee$ and $e\gamma$ sample. The single electron trigger (Section 5.4) was used to populate both the $ee$ and $e\gamma$ invariant mass distributions. The number of $Z \rightarrow ee$ events in both the invariant mass histograms were found using an extended likelihood fit, where the signal was fit by a Crystal Ball function convolved with a Breit-Wigner and the background was fit by an error function multiplied by an exponential. The fits are shown

Table 7.4: Systematic and Statistical Uncertainties from QCD Background Estimation

| $E_T^{miss}$ bin (GeV) | Systematic Uncertainty | Value |
|---|---|---|
| $100 - 110$ | Di-EM $p_T$ reweighting | 15.11% |
| | Jet multiplicity reweighting | 33.77% |
| | Shape difference between $ee$ and $ff$ | 18.18% |
| | Statistical uncertainty of $ee$ sample | 30.81% |
| $110 - 120$ | Di-EM $p_T$ reweighting | 16.60% |
| | Jet multiplicity reweighting | 14.87% |
| | Shape difference between $ee$ and $ff$ | 12.07% |
| | Statistical uncertainty of $ee$ sample | 33.33% |
| $120 - 140$ | Di-EM $p_T$ reweighting | 33.31% |
| | Jet multiplicity reweighting | 29.39% |
| | Shape difference between $ee$ and $ff$ | 14.40% |
| | Statistical uncertainty of $ee$ sample | 41.75% |
| $> 140$ | Di-EM $p_T$ reweighting | 39.37% |
| | Jet multiplicity reweighting | 20.34% |
| | Shape difference between $ee$ and $ff$ | 150.36% |
| | Statistical uncertainty of $ee$ sample | 70.98% |

in Figure 7.12. From the fits, the number of signal events in $ee$ invariant mass histogram ($N_{ee}$) was 81315±324 (stat.) and in $e\gamma$ invariant mass histogram ($N_{e\gamma}$), it was 3496±66 (stat.).

### 7.2.1 Determination of Fake Rate

The number of observed $Z \to ee$ events in the $ee$ invariant mass spectrum is given by

$$N_{ee} = (1 - f_{e\to\gamma})^2 N_{trueZ} \tag{7.2}$$

where $N_{trueZ}$ is the true number of $Z \to ee$ events. The observed $Z \to ee$ mass peak in the $e\gamma$ mass spectrum is obtained by

$$N_{e\gamma} = 2[f_{e\to\gamma}(1 - f_{e\to\gamma})]N_{trueZ}. \tag{7.3}$$

The factor of 2 comes from the fact that either of the electrons from Z boson decay can be used as the probe electron. So from Equations 7.2 and 7.3, the expression for $f_{e\to\gamma}$ is given by:

$$f_{e\to\gamma} = N_{e\gamma}/(2N_{ee} + N_{e\gamma}) \tag{7.4}$$

To propagate the uncertainty on the fake rate from the chosen shapes of the fits, first the signal was fit with Gaussian while the background was fit with error function multiplied by exponential.
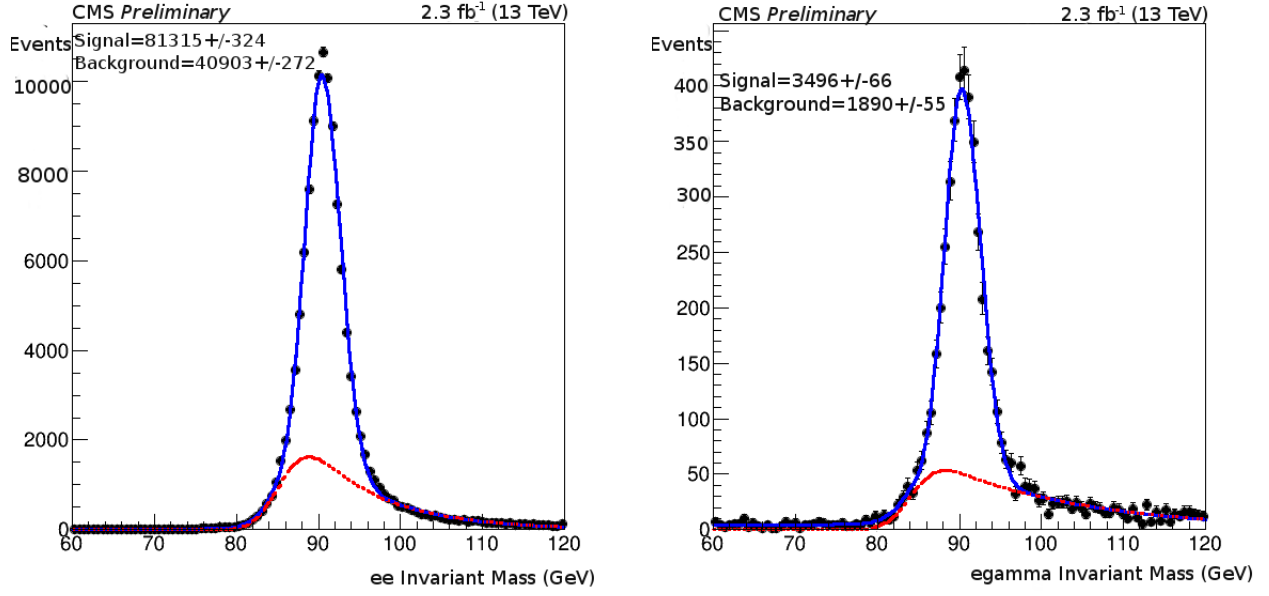
Figure 7.12: $ee$ (left) and $e\gamma$ (right) invariant mass distributions. The signal has been fit with Crystal Ball function convolved with Breit-Wigner and the background has been fit with error function multiplied by exponential.

Again, the signal was fit with Crystal Ball function convolved with Breit-Wigner while this time the background was fit with exponential function only. Changing the signal and background fit functions gave two more estimates of the fake rate. The largest difference from the central value of the fake rate was taken as the uncertainty on the central value of the fake rate. The statistical uncertainty on the fake rate was added in quadrature to give the total uncertainty on fake rate. From the values obtained in Figure 7.12, the fake rate obtained for this analysis was 0.021±0.002.

## 7.2.2 $E_T^{miss}$ Distribution of $e\gamma$ Sample

The number of events in the reconstructed $e\gamma$ mass spectrum is found by

$$N_{e\gamma} = (1 - f_{e\to\gamma})N_{e\gamma,true} \tag{7.5}$$

Here $N_{e\gamma,true}$ is the true number of events with a real photon and real electron. The background $N_{\gamma\gamma}$ which was the fraction of $N_{e\gamma,true}$ that ended up in the candidate $\gamma\gamma$ sample is given by

$$N_{\gamma\gamma} = f_{e\to\gamma}N_{e\gamma,true} = N_{e\gamma}f_{e\to\gamma}/(1 - f_{e\to\gamma}). \tag{7.6}$$

CMS *Preliminary*  2.3 fb$^{-1}$ (13 TeV)

Events/GeV

uncertainty from fake rate

total uncertainty

$E_T^{miss}$ (GeV)
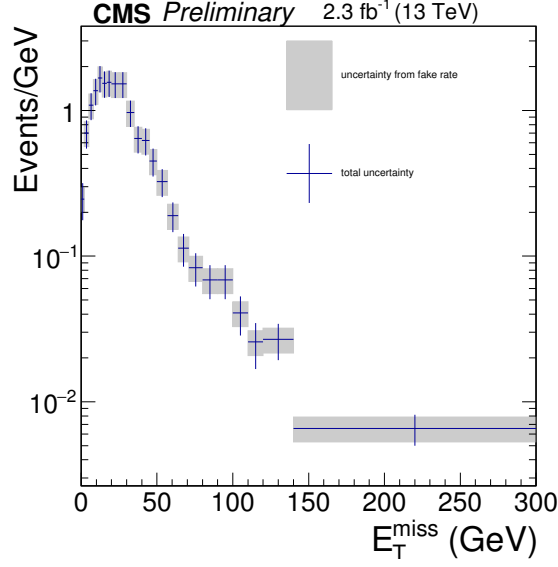
Figure 7.13: $e\gamma$ $E_T^{miss}$ distribution scaled by $f_{e\to\gamma}/(1 - f_{e\to\gamma})$. So this is final electroweak background estimate

The $e\gamma$ $E_T^{miss}$ spectrum was scaled by the factor $f_{e\to\gamma}$ $/(1 - f_{e\to\gamma}$ ) and this gave the estimate of the electroweak background. The systematic uncertainty from the fake rate is described in Section 7.2.4. The final estimation of electroweak background is shown in Figure 7.13. The estimated number of events from the electroweak background in the signal region is given in Table 7.5.

Table 7.5: Estimation of the total EWK background for $E_T^{miss} > 100$ GeV

| $E_T^{miss}$ bin (GeV) | Expected |
|---|---|
| $100 - 110$ | $0.41 \pm 0.12$ |
| $110 - 120$ | $0.26 \pm 0.09$ |
| $120 - 140$ | $0.54 \pm 0.15$ |
| $> 140$ | $1.03 \pm 0.25$ |

### 7.2.3  Components of the Electroweak Background

The entire procedure of estimating electroweak background assumes that there are only $W\gamma$ events in the selected $e\gamma$ sample. This assumption needs to be verified. This was done with the help of simulated samples. The same selection criteria which were used to select the $e\gamma$ sample from data were run on the $\gamma$+jet and $W\gamma$ simulated samples. Then their $E_T^{miss}$ spectra were compared
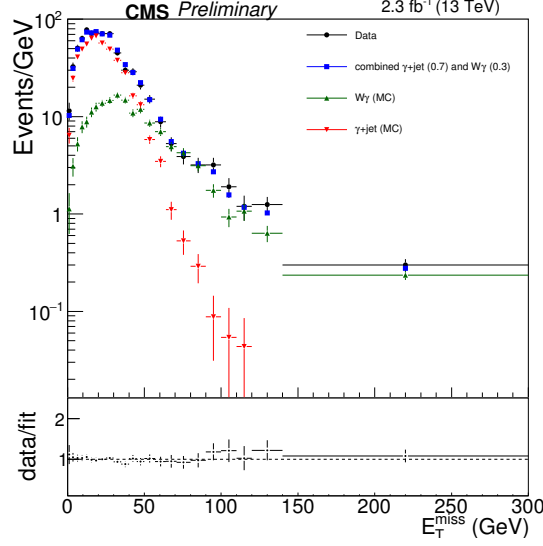
Figure 7.14: The $e\gamma$ $E_T^{miss}$ distribution from data is compared with the $E_T^{miss}$ distributions coming from $\gamma$+jet and $W\gamma$ samples. The data shape can be fit well with combination of $\gamma$+jet and $W\gamma$ $E_T^{miss}$ shape where their respective fractions are 0.7 and 0.3. From the distributions of $W\gamma$ and $\gamma$+jet $E_T^{miss}$, it is clear that the $\gamma$+jet contributes in the bulk, while $W\gamma$ sample contributes in the tail, the signal region.

with that of the $e\gamma$ sample obtained from data (Figure 7.14). From these distributions it is clear that while $\gamma$+jet events populated the bulk of $E_T^{miss}$ in the $e\gamma$ data, the tail was indeed populated by $W\gamma$ events. Since an absolute normalization was used on the $e\gamma$ $E_T^{miss}$ spectrum, the scaled $e\gamma$ $E_T^{miss}$ spectrum indeed gave the correct electroweak background estimate in the tail, which was the signal region.

### 7.2.4 Systematic Uncertainties on EWK Background Estimation

The uncertainty in the EWK background estimate comes from the uncertainty in the extended likelihood fit used to find the fake rate. The systematic uncertainty in each $E_T^{miss}$ bin was found by shifting the fake rate by $\pm 1\sigma$ and scaling the $e\gamma$ $E_T^{miss}$ distribution by the altered fake rate. The difference between the estimates from the two shifted fake rates gave the systematic uncertainty coming from uncertainty on the fake rate. The systematic uncertainty was found to be 19.45% in each $E_T^{miss}$ bin. The statistical uncertainty of electroweak background is shown in Table 7.6.

Table 7.6: Estimation of the statistical uncertainty of EWK background for $E_T^{miss} > 100$ GeV

| $E_T^{miss}$ bin (GeV) | statistical uncertainty |
|:---:|:---:|
| $100 - 110$ | 21.87% |
| $110 - 120$ | 28.63% |
| $120 - 140$ | 19.83% |
| $> 140$ | 14.52% |

# CHAPTER 8

# RESULTS AND INTERPRETATION

## 8.1 Total Estimation of Backgrounds

After the QCD and Electroweak backgrounds were obtained, they were compared with the candidate $\gamma\gamma$ $E_T^{miss}$ distribution. This is shown in Figure 8.1. The total number of expected and observed events in the signal region are summarized in Table 8.1. Here two T5gg signal model points are also shown as a reference. The details about the signal model points are described in Section 8.2.

Table 8.1: Expected, observed and signal events for $E_T^{miss} > 100$ GeV

| $E_T^{miss}$ bin (GeV) | Expected | Observed | T5gg $\tilde{g}$ mass: 1.4 TeV $\tilde{\chi}_1^0$ mass: 0.6 TeV | T5gg $\tilde{g}$ mass: 1.6 TeV $\tilde{\chi}_1^0$ mass: 0.6 TeV |
|---|---|---|---|---|
| $100 - 110$ | $2.26 \pm 0.96$ | 4 | 0.119 | 0.036 |
| $110 - 120$ | $1.79 \pm 0.64$ | 2 | 0.122 | 0.042 |
| $120 - 140$ | $1.51 \pm 0.64$ | 2 | 0.291 | 0.080 |
| $> 140$ | $1.64 \pm 2.16$ | 1 | 12.3 | 4.17 |

From the Table 8.1, one can see that the observed events did not constitute an excess over the predicted events. The result was interpreted using T5gg and T6gg simplified models and the limits on the relevant production cross sections were set (Section 8.4.3).

## 8.2 Theory Modeling

The Standard Model served as a framework for the entire history of hadron colliders, from the discoveries of Z, W$^\pm$ to the Higgs boson. But there are many extensions to the Standard Model and many have qualitatively similar phenomenology. Within the MSSM, each signature which is commonly searched for can be produced in multiple ways. As a consequence, if a signal is observed, it will not be immediately clear which particles are responsible, what their decay modes are and what other species are produced with these particles. That is why the detailed prediction of any
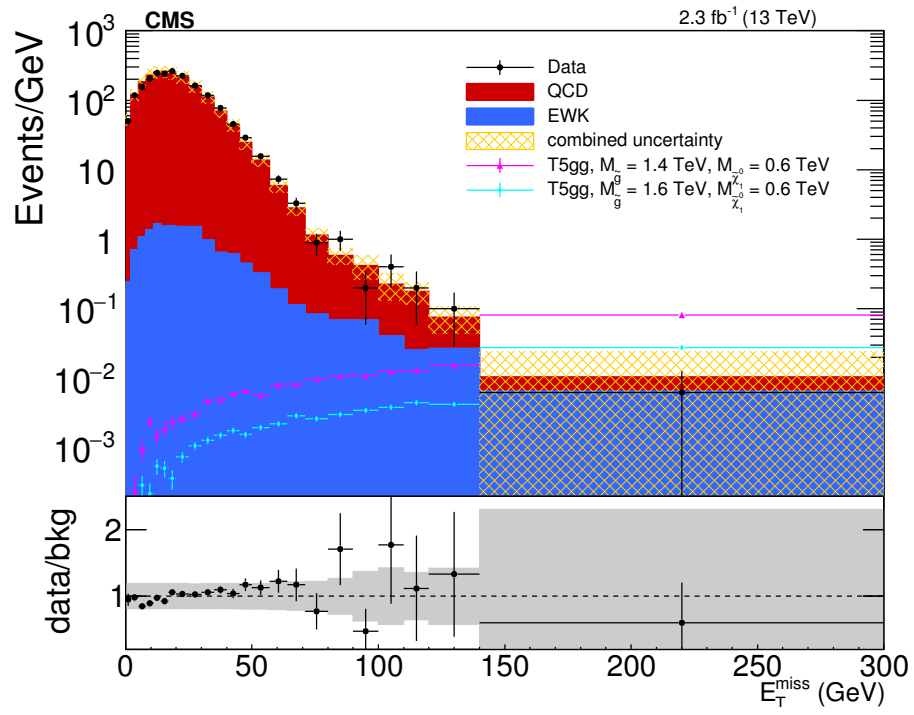
Figure 8.1: Total estimated background compared with the data.
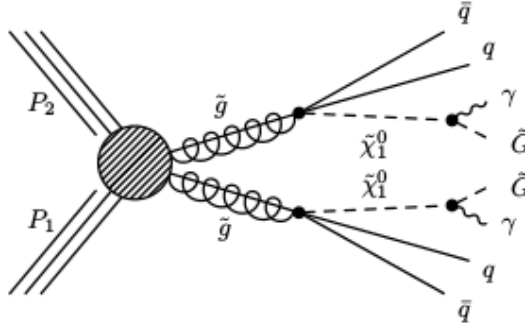
Figure 8.2: Production diagram of T5gg simplified model. Here neutralino ($\tilde{\chi}_1^0$) is the NLSP and gravitino ($\tilde{G}$) is the LSP.

one model or region of parameter space is not used and these properties are characterized in a way that allows comparison to multiple models [75]. Simplified models are presented with a small set of unambiguous parameters. These parameters are based on phenomenology typical to SUSY but devoid of much of the complexity of the full theory. In this analysis, the search was interpreted in the context of GGM SUSY scenarios and simplified model spectra (SMS) scenarios inspired by GGM models. The models considered for interpretation of the result obtained in this analysis are known as T5gg and T6gg.

The T5gg model is based on gluino pair-production where the gluino further decays to $q\tilde{q}\tilde{\chi}_1^0$. The $\tilde{\chi}_1^0$ decays to the gravitino ($\tilde{G}$) and a photon. The branching fraction of all the decays are assumed to be 100%. This way, the final state contains at least two photons, $E_T^{miss}$ and jets. The production process corresponding to T5gg SMS is shown in Figure 8.2.

The T6gg SMS model is based on squark pair production where the squark decays to $q\tilde{\chi}_1^0$ with subsequent decay of $\tilde{\chi}_1^0$ to $\tilde{G}$ and $\gamma$. The production process showing the T6gg SMS is shown in Figure 8.3.

## 8.3 Signal Simulation Production

T5gg signal samples in bins of gluino mass and neutralino mass were produced for this analysis which used the CMS FASTSIM generation [69] and reconstruction process. A total of 150,000 events were produced for each gluino and neutralino mass point. The mass of the gravitino was assumed
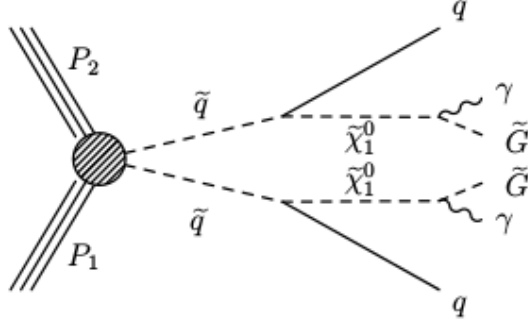
Figure 8.3: Production diagram for T6gg simplified model.

to be 1 GeV. The gluino masses range from 1 TeV to 2 TeV. For gluino masses from 1 TeV to 1.3 TeV, events were generated in steps of 100 GeV. For gluino masses from 1.3 TeV to 2 TeV, events were produced in steps of 50 GeV. The neutralino mass ranges from 100 GeV up to the mass of the gluino in steps of 100 GeV.

For the T6gg signal samples, the squark masses range from 1.2 TeV to 2.0 TeV. The T6gg signal points were generated in 100 GeV mass bins from 1.2 TeV to 1.5 TeV and in 50 GeV mass bins from 1.55 TeV to 2.0 TeV. Following the T5gg model, the neutralino masses range from 100 GeV up to the mass of squark in steps of 100 GeV.

## 8.4   Upper Limits and Exclusions

### 8.4.1   Acceptance times Efficiency for the Generated Signal Samples

The acceptance times efficiency as a function of gluino and neutralino mass for the T5gg model is shown in Figure 8.4. To obtain the acceptance, the double photon candidate event selection criteria were applied to the generated T5gg signal models. Then the acceptance times efficiency for a particular gluino-neutralino mass point was taken as the fraction of the total generated events which had two photons passing the double photon selection criteria. The drop in acceptance times efficiency at low neutralino masses was due to production of low energy photons which failed the photon $p_T > 40$ GeV cut.

The acceptance times efficiency plot as a function of squark and neutralino mass for T6gg model is shown in Figure 8.5.

108

Figure 8.4: Acceptance times efficiency for the T5gg signal samples as a function of gluino and neutralino masses.



Figure 8.5: Acceptance times efficiency for the T6gg signal samples as a function of squark and neutralino masses.

Table 8.2: Summary of systematic uncertainties included in the determination of the expected exclusion contours.

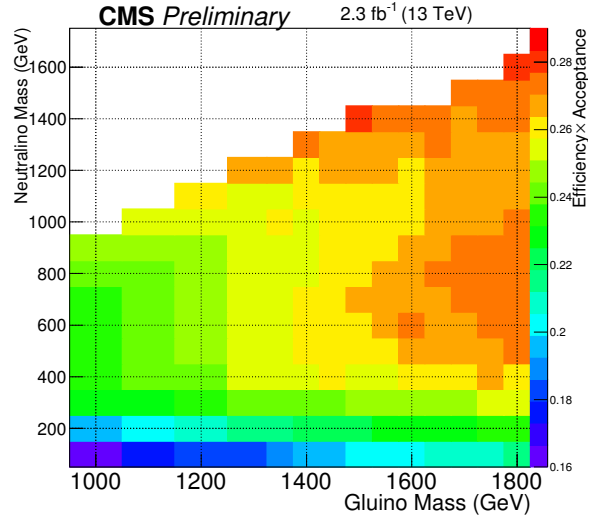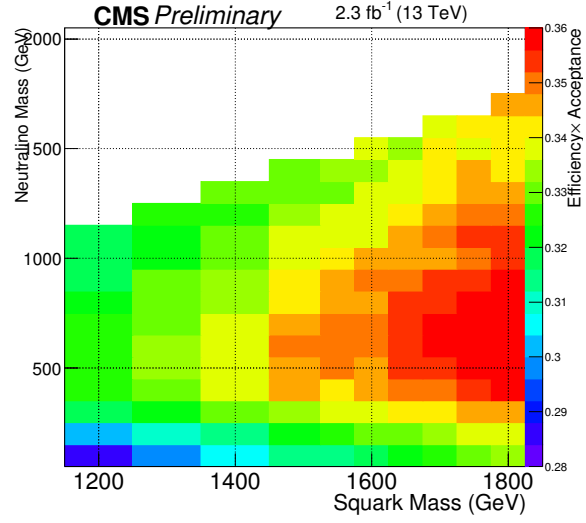| Systematic Uncertainty | [%] |
|---|---|
| Integrated luminosity | 4.6 |
| Photon Data/MC scale factor | 2.4 |
| Jet energy scale | 0 - 23 |
| Finite MC statistics | 0 - 16 |
| PDF error on cross section | 13 - 22 |

## 8.4.2 Sources of Systematic Uncertainties apart from Background Estimation

Apart from the systematic uncertainties arising due to background estimation techniques (discussed in Chapter 7), there were other sources of systematic uncertainties as well. These uncertainties included the uncertainty from the Parton Distribution Functions (PDFs) and the variation of ratios of next-to-leading order cross section to leading order cross section of PDF, commonly known as K factors. The PDF uncertainties were taken from the NNPDF 3.0 variations [76].

Apart from the PDF uncertainty, other sources of systematic uncertainties originate from finite statistics of the simulated samples, the photon Data/MC scale factor (discussed in Section 6.2.1) and the jet energy scale corrections. The uncertainties are listed in Table 8.2. Ranges of uncertainties arose when there were different values for the uncertainty in different signal points.

These additional systematic errors were introduced as nuisance parameters in the calculation of the upper limits of production cross section.

## 8.4.3 95% Confidence Level Limit

This section describes the mathematical method of computing the exclusion limits based on a modified frequentist method, often referred to as $CL_s$ [77, 78]. This supersymmetry search in double photon and $E_T^{miss}$ final state seeks to set limits on the signal production cross section utilizing the signal strength modifier ($\mu$ or R). The signal and background yield predictions are affected by the multiple uncertainties, which are known as nuisance parameters ($\theta$). In this section, the treatment of nuisance parameters in constructing the test statistic and generating pseudo-data [79] for signal (s) and background (b) binned event counts will be discussed. Thus the signal and background predictions are function of nuisance parameters, i.e. $s(\theta)$ and $b(\theta)$.

The systematic error pdfs $\rho(\theta|\bar{\theta})$ where $\bar{\theta}$ is the default value of the nuisance parameter show the degree of belief on the true value of $\theta$. The $\rho(\theta|\bar{\theta})$ are re-interpreted as posteriors coming from some measurements on $\bar{\theta}$:

$$\rho(\theta|\bar{\theta}) \sim p(\bar{\theta}|\theta) \cdot \pi_\theta(\theta) \tag{8.1}$$

where $\pi_\theta(\theta)$ are the hyper-priors of the "measurements". The procedure for obtaining observed limits are the following [79]:

1. At first, the likelihood function $L(data|\mu, \theta)$ is constructed:

$$L(data|\mu, \theta) = Poisson(data|\mu \cdot s(\theta) + b(\theta)) \cdot p(\bar{\theta}|\theta) \tag{8.2}$$

The 'data' represents either the actual experiment observation or pseudo-data coming from montecarlo simulation. $Poisson(data|\mu s + b)$ stands for a product of Poisson probabilities in binned data:

$$\prod_i \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-\mu s_i - b_i}. \tag{8.3}$$

where $n_i$ events are observed in bins $i$.

2. A test statistic $\tilde{q}_\mu$ is formed based on the profile likelihood ratio:

$$\tilde{q}_\mu = -2 \ln \frac{L(data|\mu, \hat{\hat{\theta}}_\mu)}{L(data|\hat{\mu}, \hat{\theta})}, \quad 0 \leq \hat{\mu} \leq \mu \tag{8.4}$$

where $\hat{\hat{\theta}}_\mu$ is the conditional maximum likelihood estimator of $\theta$ for the signal strength parameter $\mu$ and 'data'. The parameters $\hat{\mu}$ and $\hat{\theta}$ are the estimators corresponding to the global maximum of the likelihood. The lower constraint $0 \leq \hat{\mu}$ is because the signal rate is positive while $\hat{\mu} \leq \mu$ is imposed to guarantee a one-sided confidence interval. This means $\hat{\mu} > \mu$ is not considered as evidence against the signal hypothesis. This test statistic is formed to compare the compatibility of the data with the background only and signal+background hypotheses. Here, signal is allowed to be scaled by the factor $\mu$, the signal strength.

3. After that, the observed value of the test statistic $\tilde{q}_\mu^{obs}$ is found for a given signal strength modifier $\mu$.

4. Next, the values of nuisance parameters best describing the experimentally observed data for background only and signal+background hypotheses are found. These are $\hat{\theta}_0$ and $\hat{\theta}_\mu^{obs}$ respectively.

5. The pdfs $f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{obs})$ and $f(\tilde{q}_\mu|0, \hat{\theta}_0^{obs})$ for signal+background hypothesis and background only hypothesis are constructed from toy Monte-Carlo pseudo-data.

111

6. After the $f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{obs})$ and $f(\tilde{q}_\mu|0, \hat{\theta}_0^{obs})$ distributions are found, the p-values associated with actual observation for the signal+background hypothesis and background only hypothesis are calculated:

$$p_\mu = P(\tilde{q}_\mu \geq \tilde{q}_\mu{}^{obs}|signal + background) = \int_{\tilde{q}_\mu{}^{obs}}^{\infty} f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{obs})d\tilde{q}_\mu \qquad (8.5)$$

$$1 - p_b = P(\tilde{q}_\mu \geq \tilde{q}_\mu{}^{obs}|background) = \int_{\tilde{q}_\mu{}^{obs}}^{\infty} f(\tilde{q}_\mu|0, \hat{\theta}_0^{obs})d\tilde{q}_\mu \qquad (8.6)$$

7. The $CL_s(\mu)$ is defined as:

$$CL_s(\mu) = \frac{p_\mu}{1 - p_b} \qquad (8.7)$$

8. If for $\mu = 1$, $CL_s \leq \alpha$, the signal+background hypothesis is excluded with $(1 - \alpha)$ $CL_s$ confidence level (C.L. in short). To obtain the 95% upper limit on signal strength $\mu$ (denoted as $\mu^{95\%CL}$), the value of $\mu$ is adjusted until the value $CL_s = 0.05$ is reached.

The expected median upper limit and also $\pm 1\sigma$ and $\pm 2\sigma$ bands for the background-only hypothesis are obtained by generating a large set of background only simulation samples and $CL_s$ and $\mu^{95\%CL}$ are calculated for each of them [79]. The cumulative probability distributions are used to find the median expected value and its $\pm 1\sigma$ and $\pm 2\sigma$ variations in following way:

1. When the cumulative probability distribution function reaches the quantile of 50%, the point is taken to be the median expected value.

2. The points where the cumulative distribution function reaches 16% and 84% quantiles give $-1\sigma$ and $+1\sigma$ uncertainty on the median expected value respectively.

3. In a similar fashion, the points where the cumulative distribution function reaches 2.5% and 97.5% quantiles, give $-2\sigma$ and $+2\sigma$ uncertainty on the median expected value respectively.

For this analysis, the signal region was $E_T^{miss} > 100$ GeV. In this region, the candidate and background estimates were separated in four bins: 100-110, 110-120, 120-140 and beyond 140. The 95% confidence level upper limit on production cross section for T5gg and T6gg signal models are shown in Figures 8.6 and 8.7. A gluino mass below 1.65 TeV is excluded as seen from Figure 8.6. Squark masses below 1.35 TeV are excluded at 95% confidence level (Figure 8.7).

Figure 8.6: The 95% confidence level upper limits on the production cross section of gluino pair as a function of gluino and neutralino masses. The contours show the observed and median expected exclusions which assumes the NLO + NLL cross sections, with their one standard deviation uncertainties.

Figure 8.7: The 95% confidence level upper limits on the production cross section of squark pair as a function of squark and neutralino masses. Here the squark mass $\sim 1.35$ TeV has been excluded.

# CHAPTER 9

# CONCLUSION

A search has been performed for evidence of new physics under the framework of a gauge-mediated SUSY breaking scenario in the two photon and $E_T^{miss}$ final state. The observed data were found to be in good agreement with the expected background. Confidence level limits on the production cross section in the T5gg and T6gg simplified models were provided. Gluino masses less than 1.65 TeV in the T5gg model are excluded. This is $\sim 300 - 500$ GeV higher than the previous limit on gluino mass (Figure 9.1) using 8 TeV data with 19.6 fb$^{-1}$ integrated luminosity [35]. This analysis was a hadronic search using Razor variables [35], so it required at least one jet in the final state apart from the double photon. For the analysis discussed in this dissertation, there was no such jet requirement, and it was able to extend the limit significantly.

Squark masses less than 1.35 TeV in the T6gg model are excluded. This analysis sets the first limits on squark masses using the T6gg model.

## 9.1 Outlook

This search will be continued with more amount of data taken by CMS in future. With more data, the analysis can be done in separate jet multiplicity bins. Some of the systematic uncertainties of this analysis may be reduced with more statistics. The $E_T^{miss}$ shape difference between $ee$ and $ff$ samples may be understood better than the present analysis and the definition of fake objects may be adjusted to bring down the systematic uncertainty from $ee$ and $ff$ $E_T^{miss}$ shape difference. The electroweak production sector (Figure 9.2) can also be probed in future. Apart from the $E_T^{miss}$ variable, other global variables like $H_T$ (scalar sum of the transverse momenta of all jets in the events), missing $H_T$ (negative vector sum of the transverse momenta of all jets in the events), $E_T^{miss}$ significance [80] and $S_T$ (scalar sum of $H_T$ and transverse momenta of photons, electrons and muons in an event) will also be studied. These variables may play significant roles in discriminating signal from background.
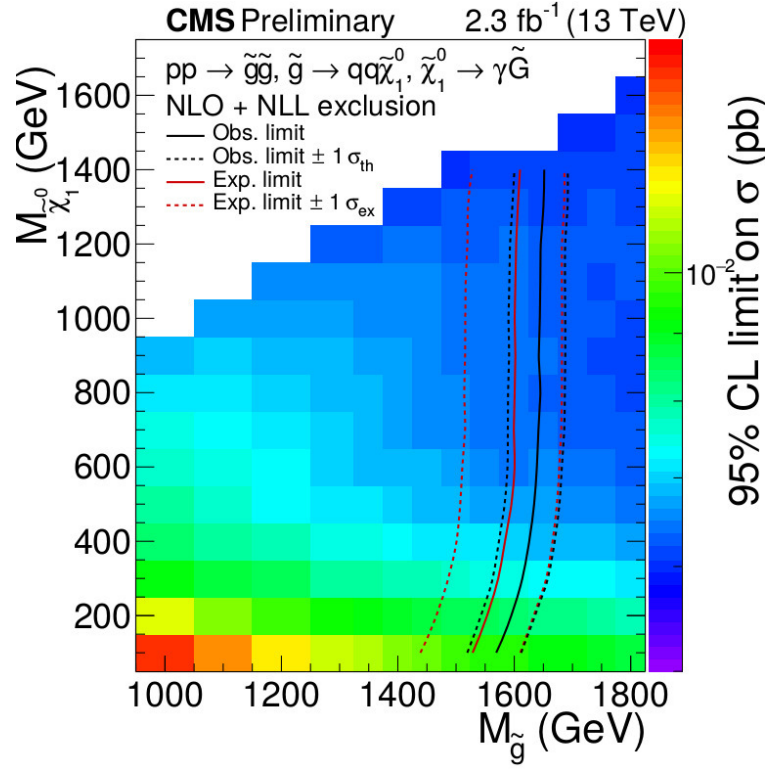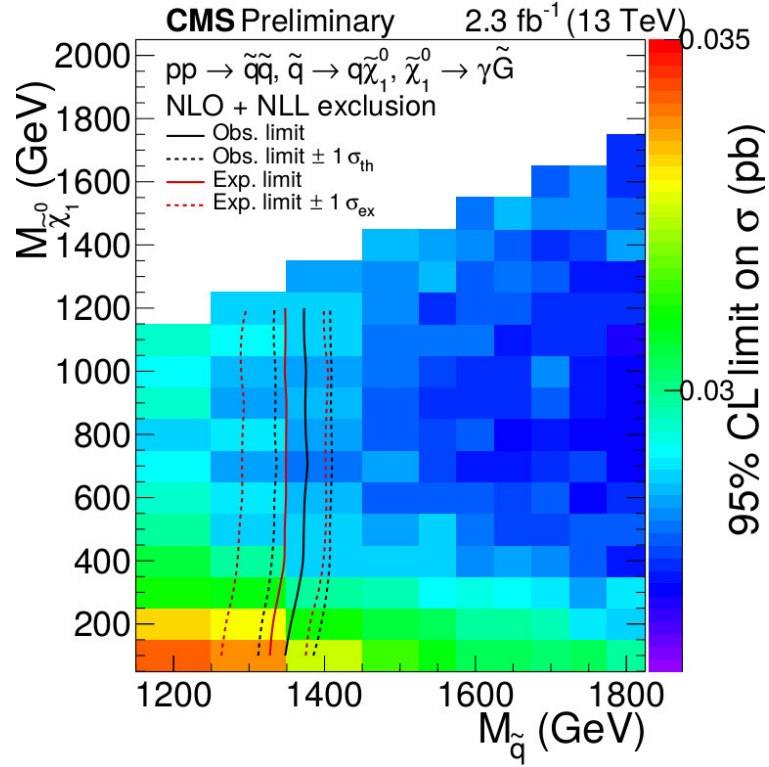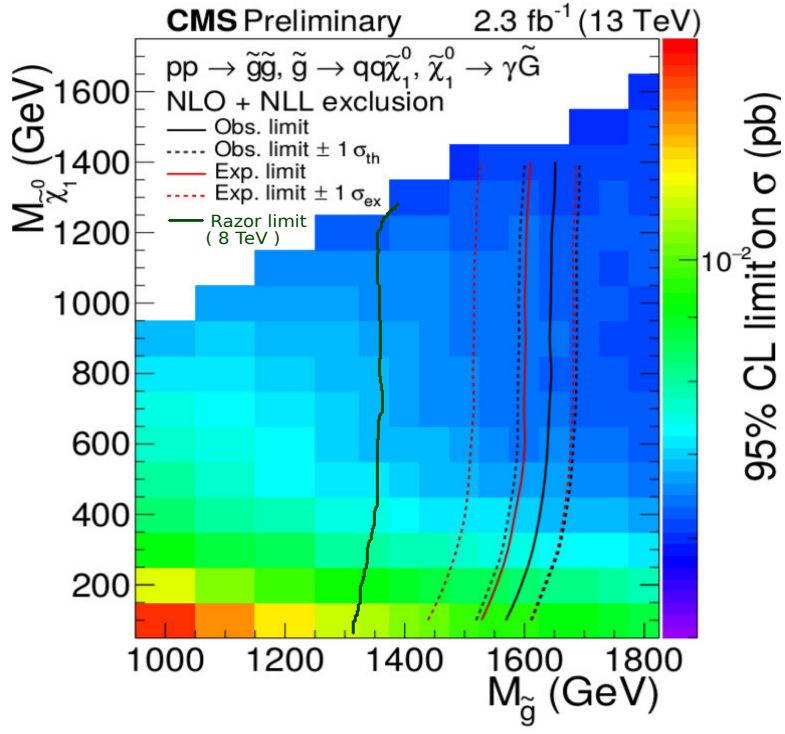
Figure 9.1: The 95% confidence level upper limits on the production cross section of gluino pair as a function of gluino and neutralino masses. The previous limit on the production cross section of gluino pair is also shown in this plot [35].
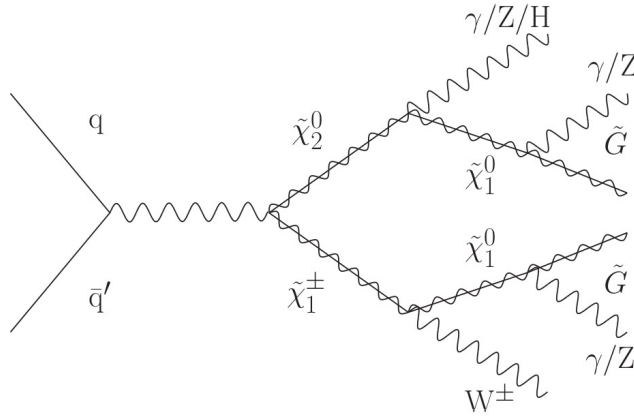


Figure 9.2: Electroweak production diagram of double photon (or Z boson) and $E_T^{miss}$ final state.

# APPENDIX A

# THE STANDARD MODEL

The Standard Model is a quantum field theory based on gauge symmetry group

$$SU(3)_C \times SU(2)_L \times U(1)_Y \tag{A.1}$$

where the gauge group $SU(3)_c$ is for color charge, $SU(2)_L$ is for weak isospin and $U(1)_Y$ is for hypercharge. The SM Lagrangian can be written in the following way [81]:

$$\mathcal{L}_{SM} = \mathcal{L}_{YM} + \mathcal{L}_f + \mathcal{L}_H + \mathcal{L}_{Yu} \tag{A.2}$$

where $\mathcal{L}_{YM}$ denotes the Yang-Mills Lagrangian, $\mathcal{L}_f$ is the fermion Lagrangian, $\mathcal{L}_H$ is the Higgs Lagrangian and $\mathcal{L}_{Yu}$ denotes the Yukawa interactions. The Yang-Mills Lagrangian describes the dynamics of the pure gauge fields. The $\mathcal{L}_{YM}$ can be written as:

$$\mathcal{L}_{YM} = \mathcal{L}_{QCD} + \mathcal{L}_{I_w} + \mathcal{L}_Y = -\frac{1}{4}\sum_{a=1}^{8} G_{\mu\nu}^a G^{a\mu\nu} - \frac{1}{4}\sum_{i=1}^{3} F_{\mu\nu}^i F^{i\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}. \tag{A.3}$$

Here, $G_{\mu\nu}^a$ are the color field strength tensors and are written as

$$G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g_s f^{abc} A_\mu^b A_\nu^c \tag{A.4}$$

where a,b and c run from 1 to 8. In the above equation, the $A_\mu^a$ are the eight color gauge fields (gluons), $g_s$ is the dimensionless strong coupling constant and $f^{abc}$ are the structure constants of $SU(3)$ group. Similarly, the weak isospin field strength tensor $F_{\mu\nu}^i$ is given by

$$F_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g_w \epsilon^{ijk} W_\mu^j W_\nu^k \tag{A.5}$$

where i,j,k run from 1 to 3. The $W_\mu$ are the gauge fields of $SU(2)_L$, $g_w$ is the dimensionless weak isospin coupling constant and $\epsilon^{ijk}$ are the structure constants of $SU(2)_L$ group. The hypercharge field strength tensor is given by

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{A.6}$$

where $B_\mu$ is the gauge field of the weak hypercharge.

The $\mathcal{L}_f$ describes the fermion fields and their interactions with the gauge bosons. The quarks, one of the fermion fields, are triplets under the color gauge group of $SU(3)$ and the leptons, the other fermion fields, have no color. Since the W bosons couple only to the left-handed helicity states of quarks and leptons, the $\mathcal{L}_f$ is built in such a way that right and left-handed components of the fermion fields couple independently to the gauge bosons. The fermion Lagrangian can be written as

$$\mathcal{L}_f = \bar{Q}_L \sigma^\mu \mathcal{D}_\mu Q_L + \bar{u}_R \sigma^\mu \mathcal{D}_\mu u_R + \bar{d}_R \sigma^\mu \mathcal{D}_\mu d_R + \bar{L}_L \sigma^\mu \mathcal{D}_\mu L_L + \bar{e}_R \sigma^\mu \mathcal{D}_\mu e_R + \dots \quad \text{(A.7)}$$

where $Q_L = \begin{pmatrix} u \\ d \end{pmatrix}_L$ is a quark weak doublet, $u_R$ is a up-type quark weak singlet, $L_L = \begin{pmatrix} \nu \\ e \end{pmatrix}_L$ is a lepton weak doublet and $e_R$ is a lepton weak singlet. The dots in the above equation suggests that there are similar terms for the remaining quarks and leptons. Here $\sigma^\mu$ are the Pauli matrices with $\sigma^0 = \mathbf{1}$. The $\mathcal{D}_\mu$ are the covariant derivatives corresponding to each of the fields [81].

A mass term for the fermion fields and for the vector boson fields is not possible because of gauge invariance. To remedy the problem, the last two terms in Equation A.2 are added. The simplest way to preserve the gauge symmetry and generate mass terms for electroweak gauge bosons is known as the Higgs mechanism [2–4]. A separate step is needed to introduce mass terms for the fermions.

The Higgs mechanism starts by adding to the model another complex scalar field, known as the Higgs field $H$. The dynamics of the field $H$ is dictated by $\mathcal{L}_H$ which can be written as

$$\mathcal{L}_H = (\mathcal{D}_\mu H)^\dagger (\mathcal{D}^\mu H) - V(H) \quad \text{(A.8)}$$

where $\mathcal{D}_\mu$ is the covariant derivative of $H$ [81]. The $V(H)$ is the most general renormalizable potential under $SU(2) \times U(1)$ and can be written as

$$V(H) = \mu^2 H^\dagger H + \lambda (H^\dagger H)^2 \quad \text{(A.9)}$$

where $\mu^2$ and $\lambda$ are real arbitrary parameters and $\lambda$ is a dimensionless parameter.

If $\mu^2 < 0$, the field configurations which minimize the potential $V(H)$ will satisfy

$$H_{vac}^\dagger H_{vac} = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2}. \quad \text{(A.10)}$$

So if $\mu^2 < 0$, the Higgs field develops a vacuum expectation value (VEV). This VEV is degenerate over the sphere defined by Equation A.10. If one configuration is chosen $\langle H \rangle = \frac{v}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and the

$\mathcal{L}_H$ is expanded in the vicinity of the chosen minimum by shifting the Higgs field as

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix}, \tag{A.11}$$

the $\mathcal{L}_H$ becomes the Lagrangian of a real scalar field with mass $m_h = 2v^2\lambda$. This is the mass of the Higgs boson. A mass term for the gauge bosons is generated by the first term of Equation A.8 [81].

The $\mathcal{L}_{Yu}$ couples massive fermion fields to Higgs field via Yukawa type interaction. The gauge invariant Yukawa coupling of the Higgs boson to the down quark d is given by

$$-\lambda_d \bar{Q}_L H d_R + \text{hermitian conjugate} \tag{A.12}$$

where $\bar{Q}_L = \begin{pmatrix} \bar{u} & \bar{d} \end{pmatrix}_L$ and $\lambda_d$ is the Yukawa coupling for the down quark. The mass term for down quark will come out to be $\lambda_d v / \sqrt{2}$ if the shift A.11 is applied [81]. Similar terms are added to the $\mathcal{L}_{Yu}$ for each of the massive fermion field.

# BIBLIOGRAPHY

[1] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579 – 588, 1961.

[2] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[3] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[4] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.

[5] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[6] WN Polyzou, W Glöckle, and H Witaa. Spin in relativistic quantum theory. *Few-Body Systems*, 54(11):1667, 2013.

[7] R K Pathria. *Statistical Mechanics*. Butterworth-Heinemann, 2 edition, 2004.

[8] K.A. Olive et al. "Review of Particle Physics". *Chin.Phys.*, C38:090001, 2014.

[9] Whillyard science pages. http://www.whillyard.com/science-pages/interactions.html. Accessed: 2015-07-07.

[10] Dark matter and dark energy. http://home.cern/about/physics/dark-matter. CERN Website.

[11] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, 1 edition, 2013.

[12] Luis Alvarez-Gaume and John Ellis. "Eyes on a prize particle". *Nature Physics*, 2011.

[13] Georges Aad et al. Combined Measurement of the Higgs Boson Mass in $pp$ Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015.

[14] The hierarchy problem and new dimensions at a millimeter. *Physics Letters B*, 429(34):263 – 272, 1998.

[15] Michael S. Turner. "Dark matter, dark energy and fundamental physics". pages 203–217, 1999. FERMILAB-CONF-99-390-A.

[16] P. A. R. Ade et al. Planck 2013 results. I. Overview of products and scientific results. *Astron. Astrophys.*, 571:A1, 2014.

[17] V. C. Rubin, N. Thonnard, and W. K. Ford, Jr. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/. *Astrophys. J.*, 238:471, 1980.

[18] Douglas Clowe, Marusa Bradac, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, Christine Jones, and Dennis Zaritsky. A direct empirical proof of the existence of dark matter. *Astrophys. J.*, 648:L109–L113, 2006.

[19] A clash of clusters provides another clue to dark matter. http://chandra.harvard.edu/photo/2008/macs/. Chandra X-ray Observatory.

[20] Stephen P. Martin. A Supersymmetry primer. *Adv.Ser.Direct.High Energy Phys.*, 21:1–153, 2010.

[21] Martin F. Sohnius. Introducing supersymmetry. *Physics Reports*, 128(2):39 – 204, 1985.

[22] H.E. Haber and G.L. Kane. The search for supersymmetry: Probing physics beyond the standard model. *Physics Reports*, 117(2):75 – 263, 1985.

[23] P. Fayet and S. Ferrara. Supersymmetry. *Physics Reports*, 32(5):249 – 334, 1977.

[24] Howard Baer and Xerxes Tata. *"Weak Scale Supersymmetry, From Superfields to Scattering Events"*. Cambridge University Press, 2012.

[25] Kenneth A. Intriligator and Nathan Seiberg. Lectures on Supersymmetry Breaking. *Class. Quant. Grav.*, 24:S741–S772, 2007. [,125(2007)].

[26] Yael Shadmi. Supersymmetry breaking. In *Particle physics beyond the standard model. Proceedings, Summer School on Theoretical Physics, 84th Session, Les Houches, France, August 1-26, 2005*, pages 147–180, 2006.

[27] Dong-Won Jung and Jae Yong Lee. Anomaly-Mediated Supersymmetry Breaking Demystified. *JHEP*, 03:123, 2009.

[28] Glennys R. Farrar and Pierre Fayet. Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry. *Phys. Lett.*, B76:575–579, 1978.

[29] S. Ask. A Review of the supersymmetry searches at LEP. In *38th Rencontres de Moriond on Electroweak Interactions and Unified Theories Les Arcs, France, March 15-22, 2003*, 2003.

[30] Michel Jaffre. SUSY searches at the Tevatron. *EPJ Web Conf.*, 28:09006, 2012.

[31] B. C. Allanach. Theoretical uncertainties in sparticle mass predictions. *eConf*, C010630:P319, 2001.

[32] S. P. Martin, S. Moretti, Jian-ming Qian, and G. W. Wilson. Direct investigations of super-symmetry: Subgroup summary report. *eConf*, C010630:P346, 2001.

[33] V.M. Abazov et al. Search for supersymmetry in di-photon final states at $\sqrt{s} = 1.96$ tev. *Physics Letters B*, 659(5):856 – 863, 2008.

[34] Summary of comparison plots in simplified models spectra for the 8 TeV dataset. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SUSYSMSSummaryPlots8TeV. CMS Website.

[35] Vardan Khachatryan et al. Search for supersymmetry with photons in pp collisions at $\sqrt{s}$=8TeV. *Phys. Rev.*, D92(7):072006, 2015.

[36] Lyndon Evans and Philip Bryant. "LHC Machine". *Journal of Instrumentation*, 3(08):S08001, 2008.

[37] "CERN brochure-2009-003". http://multimedia-gallery.web.cern.ch/multimedia-gallery/Brochures.aspx. Accessed: 2015-12-27.

[38] "CMS detector design". http://cms.web.cern.ch/news/cms-detector-design. Accessed: 2016-4-27.

[39] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[40] "CMS Physics : Technical Design Report Volume 1: Detector Performance and Software". http://cmsdoc.cern.ch/cms/cpt/tdr/. Accessed: 2016-04-09.

[41] ECAL Detector Performance, 2011 Data. May 2012.

[42] O Kodolova. HCAL calibration in high luminosity environment. page 7 p, 2013.

[43] CMS Collaboration. Performance of the CMS hadron calorimeter with cosmic ray muons and LHC beam data. *Journal of Instrumentation*, 5(03):T03012, 2010.

[44] Daniele Trocino. "The CMS High Level Trigger". *Journal of Physics: Conference Series*, 513(1):012036, 2014.

[45] The CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the cms tracker. *Journal of Instrumentation*, 9(10):P10009, 2014.

[46] Are Strandlie and Rudolf Frühwirth. Track and vertex reconstruction: From classical to adaptive methods. *Rev. Mod. Phys.*, 82:1419–1458, May 2010.

[47] Paolo Azzurri. Track Reconstruction Performance in CMS. *Nucl. Phys. Proc. Suppl.*, 197:275–278, 2009.

[48] K. Rose. Deterministic annealing for clustering, compression, classification, regression, and related optimization problems. *Proceedings of the IEEE*, 86(11):2210–2239, Nov 1998.

[49] R Frhwirth, Wolfgang Waltenberger, and Pascal Vanlaer. Adaptive Vertex Fitting. Technical Report CMS-NOTE-2007-008, CERN, Geneva, Mar 2007.

[50] Florian Beaudette. The CMS Particle Flow Algorithm. In *Proceedings, International Conference on Calorimetry for the High Energy Frontier (CHEF 2013)*, pages 295–304, 2013.

[51] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical Report CMS-PAS-PFT-09-001, CERN, 2009. Geneva, Apr 2009.

[52] Vardan Khachatryan et al. Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10(08):P08010, 2015.

[53] Wolfgang Adam, R. Frhwirth, Are Strandlie, and T. Todor. Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC. 2005.

[54] Vardan Khachatryan et al. Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10(06):P06005, 2015.

[55] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in $pp$ collision events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002, 2012.

[56] Particle-flow commissioning with muons and electrons from J/Psi and W events at 7 TeV. Technical Report CMS-PAS-PFT-10-003, CERN, 2010. Geneva, 2010.

[57] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-$k_t$ jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, 2008.

[58] Jet Energy Corrections determination at 7 TeV. Technical Report CMS-PAS-JME-10-010, CERN, Geneva, 2010.

[59] Jet Performance in pp Collisions at 7 TeV. Technical Report CMS-PAS-JME-10-003, CERN, Geneva, 2010.

[60] "CERN public twiki". https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults. Accessed: 2015-12-28.

[61] CMS Collaboration. Measurement of CMS Luminosity. *CMS-PAS-EWK-10-004*, 2010.

[62] Andreas Kornmayer. The CMS Pixel Luminosity Telescope. Technical Report CMS-CR-2015-121, CERN, Geneva, Jun 2015.

[63] "CMS Luminosity Public Result". https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults. "CMS Twiki Page".

[64] "Generating process using MadGraph". http://madgraph.phys.ucl.ac.be/. Accessed: 2016-06-10.

[65] Johan Alwall et al. A Standard format for Les Houches event files. *Comput. Phys. Commun.*, 176:300–304, 2007.

[66] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.

[67] S. Agostinelli et al. Geant4a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003.

[68] "S Banerjee". "CMS Simulation Software". *Journal of Physics: Conference Series*, 396(2):022003, 2012.

[69] Andrea Giammanco. The fast simulation of the cms experiment. *Journal of Physics: Conference Series*, 513(2):022012, 2014.

[70] "Search for SUSY with multileptons in 13 TeV data". https://cds.cern.ch/record/2140637/files/SUS-16-003-pas.pdf. CMS Physics Analysis Summary.

[71] "CMS Twiki on Tag and Probe". https://twiki.cern.ch/twiki/bin/view/CMSPublic/TagAndProbe. Accessed: 2016-05-25.

[72] CMS Collaboration. First results on Higgs to $\gamma\gamma$ at 13 TeV. 2016.

[73] "Cut Based Photon Identification Twiki for Run 2". https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedPhotonIdentificationRun2. Accessed: 2016-01-09.

[74] Appendix-f charmonium spectroscopy from radiative decays of the j/psi and psi-prime. J. E. Gaiser, Ph.D. Thesis, SLAC-R-255 (1982). Accessed: 2016-5-25.

[75] Alwall, Johan and Schuster, Philip C. and Toro, Natalia. Simplified models for a first characterization of new physics at the LHC. *Phys. Rev. D*, 79:075020, Apr 2009.

[76] Richard D Ball, Valerio Bertone, Stefano Carrazza, Christopher S Deans, Luigi Debbio, Stefano Forte, Alberto Guffanti, Nathan P Hartland, José I Latorre, Juan Rojo, et al. Parton distributions for the LHC Run II. *Journal of High Energy Physics*, 2015(4):1–148, 2015.

[77] Alexander L. Read. Presentation of search results: The CL(s) technique. *J. Phys.*, G28:2693–2704, 2002. [,11(2002)].

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

[79] Procedure for the LHC Higgs boson search combination in Summer 2011. Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug 2011.

[80] Vardan Khachatryan et al. Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV. *JINST*, 10(02):P02006, 2015.

[81] Fernando Febres Cordero. *Next-to-Leading-Order Corrections to Weak Boson Production with a Massive Quark Jet Pair at Hadron Colliders*. PhD thesis, Florida State U., 2007.

# BIOGRAPHICAL SKETCH

## Education

2011-2016  **Ph.D. in Experimental Particle Physics**, Florida State University.
2011-2013  **MS in Physics** Florida State University.
2009-2011  **MSc in Physics** Indian Institute of Technology, Kanpur, India.
2006-2009  **BSc in Physics** Jadavpur University, Jadavpur, Kolkata, India.

## Publications

2016  *Search for supersymmetry in events with photons and missing transverse energy.*
(CMS-PAS-SUS-15-012)
(https://cdsweb.cern.ch/record/2143897/files/SUS-15-012-pas.pdf).
Journal publication in preparation.

2015  *Photon identification efficiency measurement for HGCAL for*
*CMS Phase II Upgrade Scope*
(https://cds.cern.ch/record/2055167/files/LHCC-G-165.pdf) Document.
• Signing all papers on CMS Experiment since December, 2015.

## Research Experience

2012-16  Experimental High Energy Physics
• Searching for supersymmetry at CMS, CERN, Geneva, Switzerland
with two photons and missing transverse energy final state, both with
$\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV data.
• Measurement of the efficiency of HGCAL detector in photon identification
for high pile-up scenario.
• Test beam analysis of Shashlik detector at Fermilab.
• Jet and missing transverse energy performance study
for Shashlik detector.
• Study of detector missing transverse energy performance at CMS
using $\gamma + jet$ sample with $\sqrt{s} = 13$ TeV data.

2010-11  Experimental Electromagnetic Wave Study
• Extra-ordinary transmission of microwave and to enhance microwave
transmission using various structures made on copper plate at Indian
Institute of Technology, Kanpur, India.

2010  Observational High Energy Particle Astro-Physics
• Analyzing the Fermi-Large Area Telescope data at Tata Institute of
Fundamental Research, Mumbai, India.

# Awards/Achievements

2016   Evelyn and John Baugh Research Presentation Scholarship,
Department of Physics, Florida State University
(Award amount $ 500).
• Anna Runyan Graduate Endowment Award based on academic
merit, Department of Physics, Florida State University
(Award amount $ 1000).
• American Physical Society Student Travel Award sponsored by
Division of Particles and Fields (Award amount $ 400).

2014   The Universities Research Association (URA) Visiting Scholars
Award Program at the Fermi National Accelerator Laboratory,
Fall Award cycle (Award amount $ 14, 000).

2010   CSIR-UGC National Eligibility Test (for PhD selection in Indian
institutes and universities) all India rank - 4
(December 2010) among 5, 000 students.
• Academic Excellence Award in the Master of Science (2 Year)
Programme in Physics for the year $2009 - 2010$, Indian Institute of
Technology, Kanpur, India.

2006   KVPY Fellowship (Stream: basic science) from Department of Science
and Technology, Government of India for undergraduate studies
(Award amount $ 2, 500).

2004   Madhyamik Examination (Final school leaving examination in
West Bengal, India in Class X) all West Bengal rank - 10
(among 600, 000 students).

# Talks/Seminars

2016    *"A search for supersymmetry at CMS in events with two photons and missing transverse energy at* $\sqrt{s} = 13$ TeV"
- BARYONS 2016, May 19, Tallahassee, FL, USA.
- PHENO 2016, May 9, Pittsburgh, PA, USA.
- APS April Meeting, April 16, Salt Lake City, UT, USA.

2015    *"A search for supersymmetry at CMS in events with two photons and missing transverse energy at* $\sqrt{s} = 8$ *TeV"*,
- Divisional Seminar, High Energy Nuclear and Particle Physics Division, Saha Institute of Nuclear Physics, July 29, Kolkata, India.
- APS April Meeting, April 13, Baltimore, MD, USA.
- *"Response of jets in Shashlik Detector: Upgrade* 2023 *scenario"*, LPC Physics Forum, Fermi National Accelerator Laboratory, February 26, Batavia, IL, USA.

2014    *"A search for supersymmetry at CMS in events with two photons and missing transverse energy: Modeling the background"*, High Energy Seminar, Florida State University, November 14, Tallahassee, FL, USA.

2011    *"Some basic ideas of String Theory"*, Harish-Chandra Research Institute, May, Allahabad, India.

2010    *"Characteristics of photo-multiplier tube PHOTONIS XP-*2268*B and a study of Fermi-LAT data from month Feb* 2010 *to June* 2010*"*, Visitor Students' Research Program, Tata Institute of Fundamental Research, June, Mumbai, India.

# Poster Presentation

2014    *'Background modeling for events with two photons and missing transverse energy at CMS"*, CMS Week, December 15-19, Miami, FL, USA.

# Teaching Experience

2015    Instructor at
- CMS Data Analysis School held at Fermi National Accelerator Laboratory, January.
- Photon Hands-on Advance Tutorials Session held at Fermi National Accelerator Laboratory, June.

2011-12    Teaching assistant for the undergraduate electromagnetism and mechanics laboratory, Florida State University, Tallahassee, FL, USA.

2012    Grader for the graduate electromagnetism course, Florida State University, Tallahassee, FL, USA.

# Academic Visits/Schools

2015   Visited Division of High Energy Nuclear and Particle Physics,
Saha Institute of Nuclear Physics, July, Kolkata, India.
• Visited LHC Physics Center, Fermi National Accelerator Laboratory,
Batavia, IL, USA, January-June.

2011   Visited Department of Physics, Harish-Chandra Research Institute,
Allahabad, India, May-July, reading project under Prof. Ashoke Sen.

2010   Participant of Visiting Student Research Programme (VSRP) 2010
in the School of Natural Sciences held at Tata Institute of
Fundamental Research, Mumbai, May 17- July 07.
Project under Prof. B.S. Acharya on
*"Characteristics of photo-multiplier tube PHOTONIS XP-2268B and a study
of Fermi-LAT data from month Feb 2010 to June 2010."*

2007   Participant of the summer program at Indian Institute of Science,
Bangalore, India.