

This Dissertation  
entitled  
SEARCH FOR GAUGE-MEDIATED SUPERSYMMETRY BREAKING  
typeset with `NDdiss2 $\epsilon$`  v3.2013 (2013/04/16) on March 20, 2018 for  
Allison Reinsvold Hall

This  $\text{\LaTeX} 2_{\epsilon}$  classfile conforms to the University of Notre Dame style guidelines as of Fall 2012. However it is still possible to generate a non-conformant document if the instructions in the class file documentation are not followed!

Be sure to refer to the published Graduate School guidelines at <http://graduateschool.nd.edu> as well. Those guidelines override everything mentioned about formatting in the documentation for this `NDdiss2 $\epsilon$`  class file.

It is YOUR responsibility to ensure that the Chapter titles and Table caption titles are put in CAPS LETTERS. This classfile does *NOT* do that!

*This page can be disabled by specifying the “noinfo” option to the class invocation. (i.e., `\documentclass[... ,noinfo]{nddiss2e}` )*

**This page is *NOT* part of the dissertation/thesis. It should be disabled before making final, formal submission, but should be included in the version submitted for format check.**

`NDdiss2 $\epsilon$`  documentation can be found at these locations:

<http://www.gsu.nd.edu>  
<http://graduateschool.nd.edu>

# SEARCH FOR GAUGE-MEDIATED SUPERSYMMETRY BREAKING

A Dissertation

Submitted to the Graduate School  
of the University of Notre Dame  
in Partial Fulfillment of the Requirements  
for the Degree of

Doctor of Philosophy

by

Allison Reinsvold Hall

---

Michael Hildreth, Director

Graduate Program in Physics

Notre Dame, Indiana

March 2018

This is the abstract.

# CONTENTS

FIGURES . . . . .	iv
TABLES . . . . .	vi
CHAPTER 1: INTRODUCTION . . . . .	1
CHAPTER 2: THE STANDARD MODEL AND SUPERSYMMETRY . . . . .	2
2.1 The standard model of particle physics . . . . .	2
2.2 Supersymmetry . . . . .	3
2.3 Gauge-mediated supersymmetry breaking . . . . .	3
CHAPTER 3: The Large Hadron Collider . . . . .	4
3.1 Introduction . . . . .	4
3.2 Injector Chain . . . . .	4
3.3 Vacuum and magnet system . . . . .	4
3.4 Operating parameters . . . . .	4
CHAPTER 4: CMS DETECTOR . . . . .	5
4.1 Coordinate system . . . . .	5
4.2 Superconducting solenoid . . . . .	6
4.3 Tracker . . . . .	8
4.4 Electromagnetic Calorimeter (ECAL) . . . . .	10
4.5 Hadron calorimeter (HCAL) . . . . .	10
4.6 Muon system . . . . .	12
CHAPTER 5: Event Reconstruction in the CMS Detector . . . . .	14
CHAPTER 6: TRIGGER SYSTEM . . . . .	15
6.1 Overview of the CMS trigger . . . . .	15
6.2 Analysis triggers . . . . .	19
6.3 Trigger efficiency . . . . .	22
CHAPTER 7: Sample Selection . . . . .	31
7.1 Object Selection . . . . .	31

CHAPTER 8: DATA ANALYSIS AND BACKGROUND ESTIMATION METH-	
ODS . . . . .	35
8.1 Overview . . . . .	35
8.2 QCD background . . . . .	36
8.3 Electroweak background . . . . .	41
8.4 Irreducible background . . . . .	44
CHAPTER 9: Results and Interpretations . . . . .	46
9.1 Prediction versus observation . . . . .	46
9.2 Simplified Models . . . . .	46
9.3 Confidence level calculations . . . . .	49
9.4 Limits and Interpretations . . . . .	50
CHAPTER 10: Conclusions . . . . .	53
BIBLIOGRAPHY . . . . .	54

## FIGURES

4.1	Schematic of the CMS detector. . . . .	6
4.2	Calculated magnetic field $ \vec{B} $ for a longitudinal slice of the CMS detector when operated at a central magnetic flux density of 3.8 T. In the right-hand portion of the figure, each magnetic field line represents a magnetic flux step of 6 Wb. Reprinted from Reference [11]. . . . .	7
4.3	Schematic of the CMS tracker system. Detector modules are represented by single lines, and back-to-back modules are represented by double lines. The strip tracker is further divided into the Tracker Inner/Outer Barrel (TIB/TOB), Tracker Inner Disk (TID) and the Tracker Endcap (TEC). Reprinted from Reference [6]. . . . .	9
4.4	Longitudinal slice of the CMS HCAL, including the barrel HCAL (HB), outer HCAL (HO), endcap HCAL (HE), and forward HCAL (HF). The dashed lines represent lines of constant pseudorapidity. Reprinted from Reference [6]. . . . .	11
6.1	Schematic showing the overall structure of the calorimeter half of the L1 trigger. Reprinted from Reference [7]. . . . .	25
6.2	Schematic showing the overall structure of the muon half of the L1 trigger. Reprinted from Reference [7]. . . . .	26
6.3	Efficiency of the L1 seed and the leading leg of the primary analysis trigger with respect to photon $p_T$ . For $p_T > 40$ GeV, the trigger is 98.2% efficient. . . . .	27
6.4	Efficiency of the trailing leg of the primary analysis trigger with respect to photon $p_T$ . For $p_T > 40$ GeV, the trigger is 99.8% efficient. . . . .	28
6.5	Efficiency of the primary analysis trigger with respect to the invariant mass of the diphoton system. For $m_{\gamma\gamma} > 100$ GeV, the trigger is 99.4% efficient. . . . .	29
6.6	Efficiency of the trailing leg of the control trigger with respect to photon $p_T$ . For $p_T > 40$ GeV, the trigger is 90.4% efficient. The drop in efficiency with respect to the primary analysis trigger is caused by requiring both electromagnetic objects to be matched to a pixel seed. . . . .	30
8.1	The di-EM $p_T$ vector, shown in light blue, is the vector sum of the $p_T$ of the two photons in the event, shown in blue. The magnitude of the di-EM $p_T$ vector is used to model the hadronic recoil, shown in red. . . . .	37

8.2	Di-EM $p_T$ distributions of the $ff$ control sample (black) and the $\gamma\gamma$ candidate sample (red). The ratio plot on the bottom shows the $\gamma\gamma/ff$ ratios that are used to reweight the $ff E_T^{\text{miss}}$ distribution. . . . .	38
8.3	Di-EM $p_T$ distributions of the $ff$ control sample (black) and the $\gamma\gamma$ candidate sample (red). The ratio plot on the bottom shows the $\gamma\gamma/ff$ ratios that are used to reweight the $ff E_T^{\text{miss}}$ distribution. . . . .	39
8.4	Di-EM $p_T$ distributions of the $ff$ control sample (black) and the $\gamma\gamma$ candidate sample (red). The ratio plot on the bottom shows the $\gamma\gamma/ff$ ratios that are used to reweight the $ff E_T^{\text{miss}}$ distribution. . . . .	40
8.5	Di-EM $p_T$ distributions of the $ff$ control sample (black) and the $\gamma\gamma$ candidate sample (red). The ratio plot on the bottom shows the $\gamma\gamma/ff$ ratios that are used to reweight the $ff E_T^{\text{miss}}$ distribution. . . . .	41
8.6	The ratio of $\gamma\gamma$ to $ff$ events for $E_T^{\text{miss}} < 100$ GeV. The ratio has been fit to a constant function, which can then be used to extrapolate to the $E_T^{\text{miss}} > 100$ GeV signal region. This provides an alternative estimation method for the QCD background. . . . .	42
8.7	Data versus Monte Carlo comparison of the $e\gamma$ control sample $E_T^{\text{miss}}$ distribution. To determine the relative contributions of the $\gamma + \text{jet}$ and $W\gamma$ processes, the data distribution was fit using the $\gamma + \text{jet}$ and $W\gamma$ shapes as templates. The data are shown in black, and the total MC prediction is shown in blue. The $\gamma + \text{jet}$ MC (red) was scaled to 80% of the observed events in data. and the $W\gamma$ MC (green) was scaled to 20% of the data distribution. . . . .	44
9.1	$E_T^{\text{miss}}$ distributions for the full background estimation and the observed data. The black points represent the observed $E_T^{\text{miss}}$ distribution. The QCD background is shown in red, the EWK background is shown in blue, and the $Z\gamma\gamma$ background is shown in green. The combined uncertainty on the background estimation is shown in yellow. . . . .	48
9.2	Diagrams showing the production of signal events in the collision of two protons with four momenta $P_1$ and $P_2$ . In gluino $\tilde{g}$ pair production in the T5gg simplified model (left), the gluino decays to an antiquark $\bar{q}$ , quark $q$ , and neutralino $\tilde{\chi}_1^0$ . In squark $\tilde{q}$ pair production in the T6gg simplified model (right), the squark decays to a quark and a neutralino. In both cases, the neutralino subsequently decays to a photon $\gamma$ and a gravitino $\tilde{G}$ . In the diagram on the right, we do not distinguish between squarks and antisquarks. . . . .	49
9.3	Acceptance $\times$ efficiency for the T5gg simplified model. . . . .	50
9.4	Acceptance $\times$ efficiency for the T6gg simplified model. . . . .	51
9.5	Acceptance $\times$ efficiency for the T5gg simplified model. . . . .	52
9.6	Acceptance $\times$ efficiency for the T6gg simplified model. . . . .	52

## TABLES

6.1	HLT TRIGGER PATHS . . . . .	20
6.2	PRIMARY TRIGGER REQUIREMENTS . . . . .	21
7.1	ELECTROMAGNETIC OBJECT DEFINITIONS . . . . .	32
7.2	EFFECTIVE AREAS FOR ISOLATION CORRECTIONS . . . . .	33
8.1	COMPARISON BETWEEN REWEIGHTING METHOD AND RA- TIO METHOD FOR THE QCD BACKGROUND ESTIMATE . . .	43
9.1	EXPECTED AND OBSERVED EVENTS IN THE SIGNAL REGION	47



## CHAPTER 1

### INTRODUCTION

The Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) is the largest machine in the world and an incredible feat of engineering and science. This machine is used to study elementary particles and their interactions. At the LHC, proton-proton collisions occur at center-of-mass energies up to 13 TeV.

The Standard Model of particle physics describes fundamental particles and their interactions to an incredible level of precision. This is described in more detail in Chapter XX.

The CMS detector and various subsystems will be described in Chapter 4. Finally, Chapter XX will include a discussion of the results of this analysis and suggestions for future work.

Introduction

Chapter 1: Theory Motivation Standard Model Why we need BSM SUSY GMSB

Chapter 2: LHC About CERN How the accelerator works Chapter 3: CMS Detector

Chapter 4: Object Reconstruction

Chapter 5: Trigger

Chapter 6: Event Selection Including object cleaning

## CHAPTER 2

### THE STANDARD MODEL AND SUPERSYMMETRY

#### 2.1 The standard model of particle physics

The standard model (SM) of particle physics is a non-Abelian gauge theory that seeks to describe and quantify everything that is known about elementary particles and their allowed interactions. The current form of the standard model has passed every experimental test to an incredible level of precision. The first piece of the standard model was developed in 1961 (XX references! XX) with the unification of the electromagnetic and weak interactions. The Higgs mechanism was later incorporated into the standard model in 1967 (XX cite XX), and the standard model took on the form we know today with the inclusion of the strong force and quantum chromodynamics (QCD) in the 1970's (XX cite XX).

The standard model is a Lorentz-invariant quantum field theory. The symmetry group of the standard model is

$$SU(3)_c \oplus SU(2)_L \oplus SU(2)_\tau \tag{2.1}$$

where

its roots in the unification of the electromagnetic and weak interactions in the 1950's (XX or some decade XX references! XX). The Higgs mechanism was later incorporated into the standard model in the 1960s, and

## 2.2 Supersymmetry

Motivations: Higgs mass, contributions Hierarchy problem

## 2.3 Gauge-mediated supersymmetry breaking

## CHAPTER 3

### The Large Hadron Collider

#### 3.1 Introduction

The Large Hadron Collider is a particle accelerator located at CERN, the European Organization for Nuclear Research.

#### 3.2 Injector Chain

#### 3.3 Vacuum and magnet system

#### 3.4 Operating parameters

## CHAPTER 4

### CMS DETECTOR

The Compact Muon Solenoid (CMS) detector is a multi-purpose detector designed to accurately measure the energy and momentum of all particles produced in proton-proton or heavy ion collisions. Figure 8.7 shows a schematic of the detector as a whole. Each subdetector will be described in more detail in the following sections. Moving radially outward from the interaction point, the sub-detectors are the silicon pixel and strip tracker (Section 4.3), the electromagnetic calorimeter (Section 4.4), the hadron calorimeter (Section 4.5), and the muon system (Section 4.6). For a full description of the CMS detector, see [6]. The full length of the CMS detector is 21.6 m, the diameter of the detector is 14.6 m, and the total weight is 12500 t.

#### 4.1 Coordinate system

The origin of the CMS detector coordinate system is located at the nominal collision point. The  $z$ -axis is oriented along the beam direction, with the positive  $z$ -axis pointing in the counter-clockwise direction when viewing the LHC from above. The  $y$ -axis points vertically upward, and the  $x$ -axis points toward the center of the LHC. The  $xy$ -plane is referred to as the transverse plane.

Due to the nature of particle collisions, however, Cartesian coordinates are often not the most convenient. Because protons are not elementary particles, it is actually the individual quarks or gluons within the proton that interact during the collision. This means that the collision will not be at rest in the lab frame, but will have some

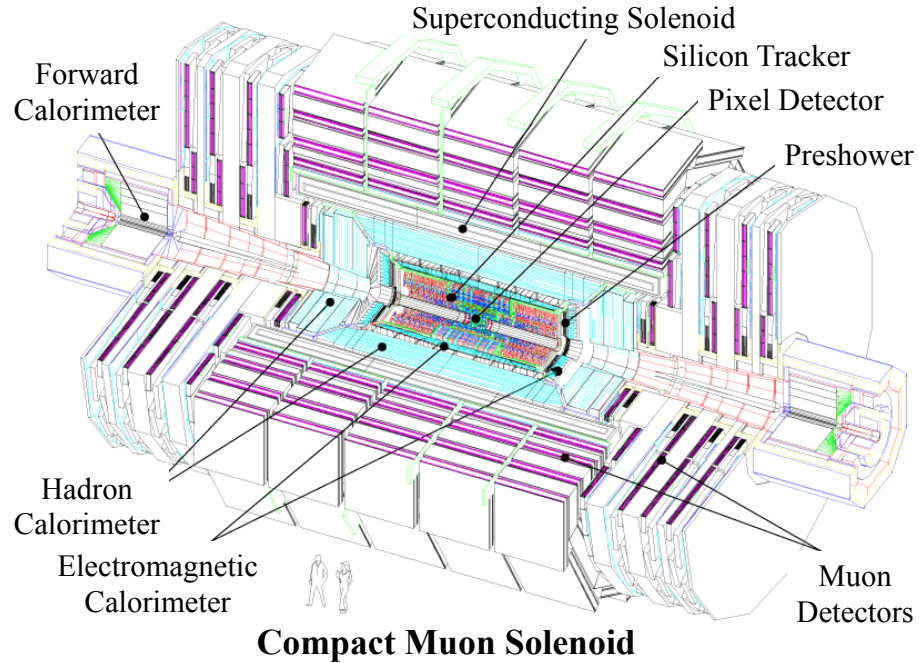


Figure 4.1. Schematic of the CMS detector.

non-zero velocity along the  $z$ -axis. To deal with this, it is beneficial to use coordinates that are invariant under boosts in the  $z$  direction. CMS follows the particle physics convention of describing the position of a particle in terms of its transverse momentum, azimuthal angle, and pseudorapidity. The transverse momentum  $p_T$  is defined as the magnitude of the momentum in the  $xy$  plane. The azimuthal angle  $\phi$  is defined in the transverse plane, with  $\phi = 0$  corresponding to the positive  $x$ -axis. Finally, the pseudorapidity is defined as  $\eta = -\ln \tan(\theta/2)$ , where the polar angle  $\theta$  is measured from the  $z$ -axis.

## 4.2 Superconducting solenoid

One of the most important components of the CMS detector is the superconducting solenoid that provides the bending power necessary to precisely measure the momentum of all charged particles produced in the collision. The solenoid is a 4-layer

niobium-titanium coil embedded in aluminum and aluminum alloy. The magnet is located between the calorimeters and the muon system. It is 12.5 m long and has an inner diameter of 6 m. It is capable of producing magnetic fields up to 4 T, although the magnet is generally operated at 3.8 T to prolong its lifetime. At full current, the magnet has a stored energy of 2.6 GJ. A 12,000 ton steel yoke made up of 5 wheels in the barrel and 3 endcap disks serves to return the magnetic flux. The solenoid is suspended in a vacuum cryostat and cooled to 4.5K with liquid helium. A detailed description of the CMS magnet can be found in [8]. Figure 4.2 shows the calculated magnetic field in a longitudinal slice of the CMS detector.

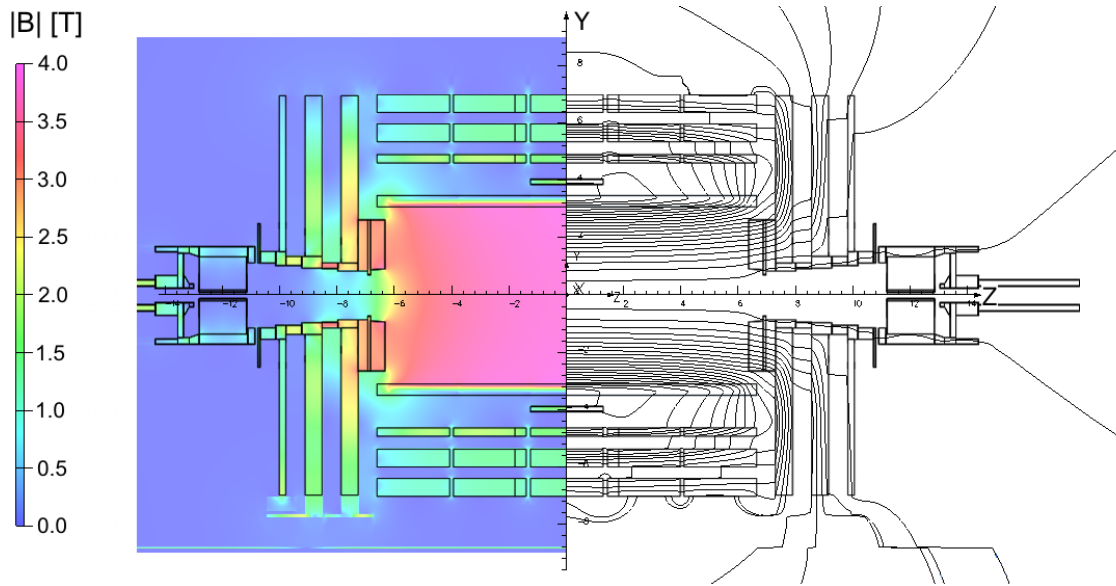


Figure 4.2: Calculated magnetic field  $|\vec{B}|$  for a longitudinal slice of the CMS detector when operated at a central magnetic flux density of 3.8 T. In the right-hand portion of the figure, each magnetic field line represents a magnetic flux step of 6 Wb. Reprinted from Reference [11].

### 4.3 Tracker

The innermost sub-detector is the silicon tracker [9, 10]. The tracker must provide enough information to accurately reconstruct the trajectories of charged particles to a high level of precision. This is accomplished through the use of silicon semiconductors, which rely on the properties of p-n junctions to detect charged particles. A p-n junction is created by bringing a p-type semiconductor into contact with an n-type semiconductor. For p- and n-type semiconductors, the crystal lattice of the semiconductor is doped to either donate or accept extra electrons, respectively. In a p-n junction, the extra electrons in the n-type semiconductor migrate and combine with the electron holes in the p-type semiconductor. This creates a depletion region in the center of the sensor. Charged particles moving through the depletion region will create electron-hole pairs that move towards either end of the junction. The resulting current is proportional to the energy deposited in the detector.

The overall dimensions of the tracker are 5.6 m long and 1.2 m in radius. The full tracking system is cylindrical in shape and is comprised of a barrel and two endcaps, each of which is split into layers of silicon pixel detectors and layers of silicon micro-strip detectors. In total, there are 48 million  $150 \times 100 \mu\text{m}$  pixels and 9.6 million strips that are between 80 and  $180 \mu\text{m}$  wide. Figure 4.3 shows a schematic drawing of the layout of the tracker subsystems. For charged hadrons with transverse momentum  $p_T$  less than 20 GeV, the  $p_T$  can be measured with a resolution of 1%.

#### 4.3.1 Pixel detector

Due to the high occupancy of the tracker within 10 cm of the beam pipe, pixel detectors rather than strip detectors are used as the innermost layers of the tracker. Three layers in the barrel and two disks of pixel detectors in the endcap give three high-precision points for every charged particle moving away from the interaction point. The small pixel size of  $150 \times 100 \mu\text{m}$  is critical for accurate secondary vertex



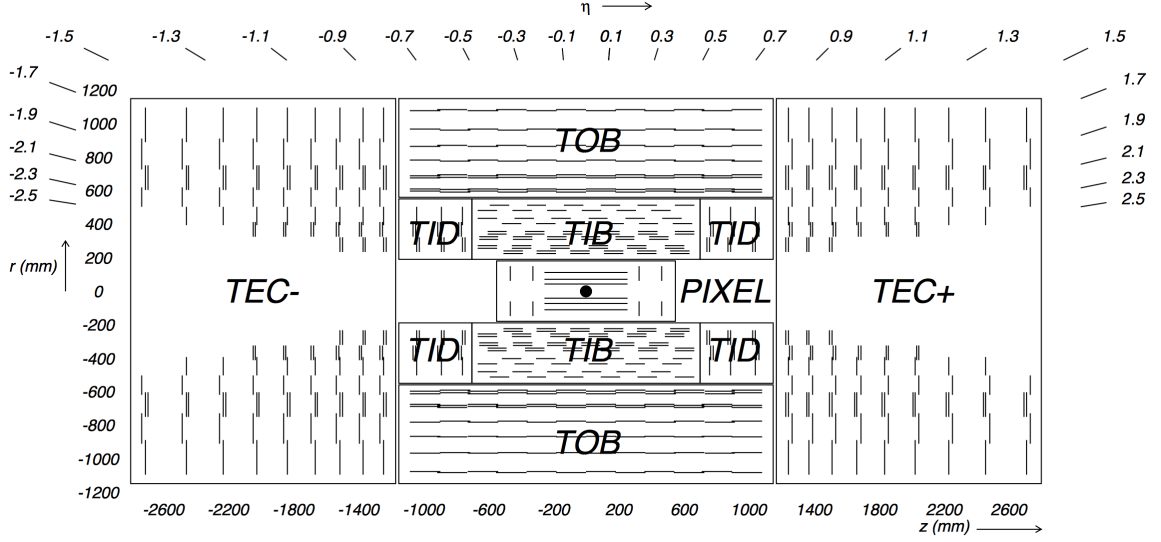


Figure 4.3: Schematic of the CMS tracker system. Detector modules are represented by single lines, and back-to-back modules are represented by double lines. The strip tracker is further divided into the Tracker Inner/Outer Barrel (TIB/TOB), Tracker Inner Disk (TID) and the Tracker Endcap (TEC). Reprinted from Reference [6].

reconstruction, for forming seed tracks used for high level triggering (see Chapter 6), and for the reconstruction of charged particles in the event (see Chapter 5). The three barrel layers are 53 cm long and are located at mean radii of 4.4, 7.3 and 10.2 cm. The endcap modules extend radially from 6 to 15 cm and are placed on each side at  $z = \pm 34.5$  and  $z = \pm 46.5$  cm. The endcaps extend the coverage of the sub-detector to  $|\eta| < 2.5$ .

#### 4.3.2 Strip detectors

Moving radially beyond the pixel detectors are the silicon strip trackers. These operate on the same basic principles as the pixel detectors, but each silicon strip is  $10 \text{ cm} \times 80 \text{ } \mu\text{m}$ , giving a precise position measurement along one dimension only. In the barrel, the strips run parallel to the  $z$ -axis with a pitch, or spacing between strips, between 80 and  $183 \text{ } \mu\text{m}$ . Strips in the endcaps are aligned radially with a pitch between 100 and  $184 \text{ } \mu\text{m}$ . In total there are 10 layers of strip sensors in the

barrel and 12 layers in the endcaps.

The tracker has a spatial resolution of 25-50  $\mu\text{m}$  perpendicular to the strip direction. In order to improve the precision of the detector in the direction parallel to the strips, several of the layers are arranged in pairs. By aligning the second layer 100 mrad off from the strips in the first layer, a spatial resolution of 230 to 530  $\mu\text{m}$  can be achieved. These back-to-back modules are represented by double lines in Figure 4.3.

#### 4.4 Electromagnetic Calorimeter (ECAL)

Very important. Finds photons.

#### 4.5 Hadron calorimeter (HCAL)

The Hadron Calorimeter (HCAL) is the next sub-detector after the ECAL. It is a brass sampling calorimeter, with alternating layers of plastic scintillator and brass absorbers. Heavy particles such as hadrons interact with the brass layers, and the scintillation light from the nuclear showers is collected with wavelength-shifting optical fibers embedded in the plastic scintillator. The light is then guided to pixelated hybrid photodiodes and electronics that amplify the signal. Figure 4.4 shows a longitudinal slice of the full HCAL system.

There are 17 scintillator layers in the barrel of the HCAL (HB), which extends radially from 1.77 m to 1.95 m and to a pseudorapidity  $|\eta| = 1.3$ . The first and last layers of absorber are made of steel rather than brass for structural reasons. The scintillation layers are divided into tiles with area  $\Delta\phi \times \Delta\eta = 0.87 \times 0.87$ . The tiles are further organized into towers that each get a single readout channel. The outer HCAL (HO) is located outside the solenoid and serves to increase the thickness of the calorimeter.

The HCAL endcap (HE) covers the pseudorapidity range  $1.3 < |\eta| < 3$ . There are 17 layers of plastic scintillator and 18 layers of absorber in HE. For  $|\eta| < 1.6$ , the

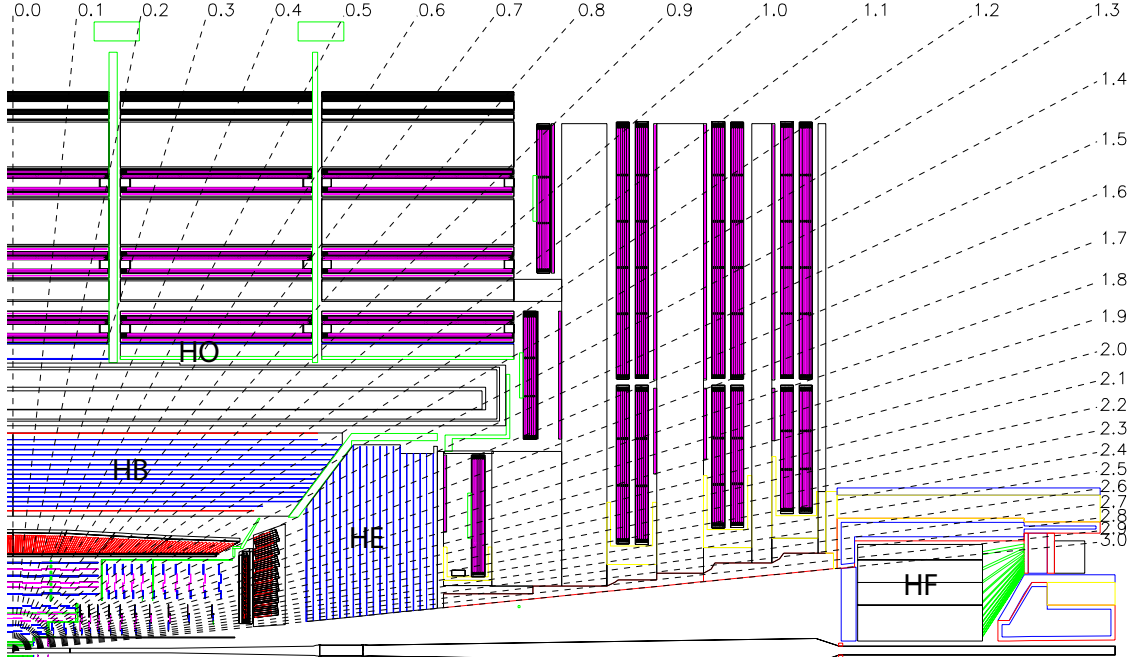


Figure 4.4: Longitudinal slice of the CMS HCAL, including the barrel HCAL (HB), outer HCAL (HO), endcap HCAL (HE), and forward HCAL (HF). The dashed lines represent lines of constant pseudorapidity. Reprinted from Reference [6].

tile size is the same as those in the barrel, and for  $|\eta| > 1.6$  the tile size is reduced to  $\Delta\phi \times \Delta\eta = 0.17 \times 0.17$ . Depending on the pseudorapidity, there are two or three readout channels per tower.

The last part of the HCAL system is the forward HCAL (HF) that covers the pseudorapidity range from  $3 < |\eta| < 5.2$ . Charged particles are detected via Cherenkov radiation in quartz fibers planted in the steel absorber. The quartz fibers were chosen to be able to withstand the high radiation levels of particles emitted close to the beam pipe. All three segments of the HCAL are calibrated using radioactive sources  $\text{Cs}^{136}$  or  $\text{Co}^{60}$  mounted on the tip of a moving wire. The radioactive sources produce photons at known energies, allowing for an absolute calibration of each scintillator.

Because most of the energy from the nuclear showers is deposited in the absorbers rather than the scintillation material, the HCAL naturally has a lower energy resolution than the ECAL. In addition, nuclear showers might start before particles

even reach the HCAL, or charged particles might deposit energy in the ECAL via bremsstrahlung. The mean distance travelled by a hadronic particle before undergoing an inelastic nuclear interaction is known as the nuclear interaction length. The ECAL has a nuclear interaction length of 1.1, and the depth of the HCAL corresponds to 5.82 interaction lengths at  $|\eta| = 0$  and 10.6 interaction lengths at  $|\eta| = 1.3$ .

## 4.6 Muon system

Aside from weakly-interacting particles such as neutrinos, muons are the only particles that make it past the HCAL. Muons do not interact via the strong force and are too heavy to be stopped by electromagnetic interactions alone. For this reason, the muon detector system is the outermost layer of CMS. Three different types of detectors are used in the muon system: drift tube (DT) chambers, cathode strip chambers (CSC), and resistive plate chambers (RPC). The muon system is located at a radial distance  $r > 3.5$  m and is embedded in the return yoke for the magnetic flux.

A drift tube consists of a conducting wire held at high positive voltage in the center of a gas cell. When a charged particle passes through the gas, the gas is ionized. The freed electrons are drawn to the positively charged wire, and in turn cause more ionization as they are accelerated. The time between the passage of the initial particle and the resulting electron avalanche can be used to reconstruct the position of the interaction perpendicular to the wire.

Drift tubes are used in the barrel region of the detector and extend to  $|\eta| < 1.2$ . The CMS drift tubes are made from aluminum plates and contain a mixture of 85% Ar and 15% CO<sub>2</sub>. A gold-plated stainless steel anode wire is located at the center of each tube. The wires are held at 3.6 kV and have a thickness of  $50\mu\text{m}$ . Each drift tube has dimensions  $1.3\text{ cm} \times 4.2\text{ cm} \times 2.4\text{ m}$ . Four layers of drift tubes make up a superlayer, and a group of 2 or 3 superlayers makes up one muon chamber.

The endcap system uses cathode strip chambers (CSC) because of their improved performance in high flux areas and non-uniform magnetic fields. Positively-charged wires are aligned perpendicular to negatively-charged wires in a gas, which gets ionized by passing muons. Electrons produce a signal in the anode wires and ions produce a signal in the cathode wires. This provides two position coordinates for each muon. The spatial resolution in the azimuthal direction is  $150\text{ }\mu\text{m}$ . The CSCs extend the pseudorapidity coverage of the muon system to  $|\eta| < 2.4$

Finally, resistive plate chambers are used in both the barrel and the endcap for more precise timing measurements. This is important for triggering purposes and to make sure each muon is assigned to the correct bunch crossing. A high voltage is applied to two large plates with a gaseous layer between them. The plates themselves have a high resistance, and the cascade caused by the passage of a muon through the gas is detected via readout strips outside the chamber. The CMS RPCs have a timing resolution of  $1\text{ ns}$ .

## CHAPTER 5

### Event Reconstruction in the CMS Detector

5.0.1 Particle Flow Algorithm

5.0.2 Jet reconstruction

5.0.3 Pileup subtraction

## CHAPTER 6

### TRIGGER SYSTEM

#### 6.1 Overview of the CMS trigger

Many more collisions occur within the LHC than can be stored and analyzed. At the four interaction points of the LHC, bunches of protons collide every 25 ns. For every bunch crossing, there are on average 25 “soft-scatter” collisions (pileup). This corresponds to an overall event rate of 1 GHz. Given that each event produces approximately 1 MB of data, it is impossible to store every event or even the majority of events. This has been especially true in recent years, with the LHC continually breaking new records for instantaneous luminosity.

To solve this problem, the CMS detector includes a sophisticated trigger system to determine which events get written to tape and eventually analyzed. The CMS trigger system is divided into two steps: the hardware-based Level 1 (L1) trigger, and the software-based High Level Trigger (HLT). The L1 trigger reduces the rate from 1 GHz to approximately 100 kHz, and the HLT makes the final decisions necessary to reduce the rate to a few hundred Hz, the maximum amount that can be written and stored. The L1 trigger and the HLT are both highly configurable, which allows CMS to define and alter the thresholds as needed to suit different running conditions. Both steps of the CMS trigger are described in more detail below.

##### 6.1.1 Level 1 Trigger

The L1 trigger performs rough calculations of event parameters using field-programmable gate arrays located in the detector cavern. The L1 trigger only has  $3.8 \mu\text{s}$  to make

a decision about each event in order to reduce the output event rate to 100 kHz. For inputs, the L1 trigger receives data from the ECAL, HCAL, and muon detector subsystems. The L1 trigger is naturally divided into two paths. The calorimeter trigger, shown schematically in Figure 6.1, takes data from the ECAL and HCAL subsystems and builds photon, electron, and jet candidates, as well as overall event quantities such as missing transverse momentum and total hadronic activity. The second half of the L1 trigger is the muon trigger, and its overall structure is shown in Figure 6.2. The output from the muon and calorimeter triggers goes into the Global Trigger (GT) processor, which makes the final decision regarding whether or not to accept the event.

#### 6.1.1.1 Calorimeter Trigger

As shown in Figure 6.1, the raw inputs for the calorimeter trigger are trigger primitives (TP) from the ECAL and the HCAL. The trigger primitives from the HCAL are divided between the HCAL Barrel and Endcap (HB and HE, respectively), and the forward HCAL (HF). Trigger primitives consist of coarse information about the energy deposits in the calorimeter for every bunch crossing. The ECAL is divided into trigger towers, groups of crystals corresponding to a region of approximately  $0.087 \times 0.087$  in  $\eta$  and  $\phi$ , and a TP is generated by summing the energy from each crystal in the tower. In the HCAL, TPs are generated for each trigger tower by summing the transverse energy in two consecutive time slices.

Layer-1 of the calorimeter trigger collects the TPs from the ECAL and HCAL. Its primary role is to distribute the data to one of nine Layer-2 nodes. Each Layer-2 node receives the full set of TPs for a particular bunch crossing, and identifies photon, electron, jet, and  $\tau$ -tagged jet candidates using several dynamic clustering and local maxima finding algorithms. The outputs from the Layer-2 nodes are fed into a demultiplexing (demux) node, which ranks the candidates by transverse momentum



and sends the data to the Global Trigger.

#### 6.1.1.2 Muon Trigger

Information from all three of the muon detector subsystems—the cathode strip chamber (CSC) detectors, resistive plate chamber (RPC) detectors, and drift tubes (DT)—are combined in the L1 muon trigger to reconstruct muon candidates and their momenta. The muon trigger has three track-finding subsystems that reconstruct muons for different  $|\eta|$  regions: the barrel track finder ( $|\eta| < 0.83$ ), the endcap track finder ( $|\eta| > 1.24$ ), and the overlap track finder ( $0.83 < |\eta| < 1.24$ ).

Trigger primitives from the DT and RPC detectors are first combined in the TwinMUX system into “super-primitives”. This step serves to increase the precision of the position and timing of muon hits by taking advantage of the redundancy between the two subsystems. The barrel track finder uses the super-primitives as inputs for its track finding algorithm, processing twelve  $30^\circ$  wedges in parallel.

The endcap track finder receives information from the CSC detectors, and the overlap track finder uses information from all three subsystems. Both of these track finders use large look-up tables to convert specific hit patterns into  $p_T$  assignments for the muon candidates.

The outputs from all three track finders are sent to the Global Muon Trigger, which ranks the muons by quality and transverse momentum and sends the top 8 muon candidates to the Global Trigger for a final decision on the event.

#### 6.1.2 Global Trigger

The reconstructed objects and quantities built by the calorimeter trigger and the muon trigger—photon, electron, jet, and muon candidates, missing energy and total hadronic activity—get sent to the Global Trigger (GT) for final processing. The GT can implement up to 512 different trigger algorithms using these objects. For example,

one of the L1 algorithms (or “seeds”) used in this analysis is DoubleEG\_22\_12. This path requires two electron or photon candidates with leading and trailing  $p_T$  greater than 22 and 12 GeV, respectively.

A full set of L1 algorithms is known as the L1 menu. The menu can be adjusted as needed to meet the requirements of the CMS physics program. In particular, the “prescale” of each L1 seed can be adjusted to take full advantage of varying LHC running conditions and instantaneous luminosities. A prescale is an integer value  $N$  used to reduce the rate of a trigger path by only applying the trigger to 1 out of every  $N$  events. For the 2016 data-taking period, the lowest unprescaled DoubleEG seed was DoubleEG\_22\_12, but the menu also included several prescaled algorithms with lower  $p_T$  thresholds.

### 6.1.3 High Level Trigger

The High Level Trigger (HLT) uses a large farm of commercially available PCs to further reduce the rate to a few hundred Hz. The HLT receives fully-built event data from the CMS data acquisition system (DAQ) and processes only those bunch crossings that have passed the L1 trigger. The 13,000 CPU cores used in the HLT run the CMS software framework referred to as CMSSW, the same software framework that is used in the offline analysis.

Approximately 400 HLT “paths” are used to select events of interest. Each HLT path is a single set of criteria for accepting an event if it satisfies a particular physics signature. The full list of HLT paths is referred to as the HLT menu. The HLT menu is often updated to reflect improvements in the software, updates to the calibrations, changes to the beam condition, or revisions to the physics signatures being sought.

The HLT has a limited amount of time it can spend making a decision on a single bunch crossing. For this reason, each HLT path is “seeded” by one or more L1 algorithms. When an event gets passed to the HLT from the DAQ, the HLT

only processes those paths that are seeded by L1 bits that fired. For example, the diphoton HLT path used in this analysis is seeded by a combination of single and double  $e/\gamma$  (EG) L1 algorithms. If none of those particular L1 seeds fired, then the HLT path will not get run on that event.

In addition, significant time is saved at the HLT by running the steps (filters) of each HLT path in order from least CPU intensive to most CPU intensive. If the event fails any of the filters in an HLT path, then the processing is immediately aborted and the remaining filters do not get run. In the diphoton HLT path described below, the most CPU intensive filter is the invariant mass calculation, and therefore it runs only if the event has already satisfied the rest of the path's requirements.

Based on its general physics signature, each HLT path is assigned to a primary data set (PD). For this analysis, signal events are included in the DoubleEG PD. The SingleElectron and SinglePhoton data sets are also used for object identification studies.

## 6.2 Analysis triggers

The HLT paths used in this analysis are listed in Table 6.1. These triggers were developed for the  $H \rightarrow \gamma\gamma$  search, but also serve our analysis well. Two triggers are used: a primary trigger that requires the diphoton invariant mass to be greater than 90 GeV, and a control trigger that was designed to collect  $Z \rightarrow ee$  events.

### 6.2.1 Trigger Requirements

The requirements to pass the various parts of the trigger are listed in Table 6.2. Because we only use photons with  $|\eta| < 1.4442$  in this analysis (see Chapter 7 for full event selection requirements), we list only the trigger requirements in the barrel.

TABLE 6.1

## HLT TRIGGER PATHS

<b>Primary Trigger</b>
HLT_Diphoton30_18_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass90_v*
<b>Control Sample Trigger</b>
HLT_Diphoton30_18_R9Id_OR_IsoCaloId_AND_HE_R9Id_ DoublePixelSeedMatch_Mass70_v*

List of triggers used to accumulate the events in the 35.88 fb<sup>-1</sup> data sample.

Each of the variables used in the trigger are defined below:

- $E_T$ : The transverse energy  $E_T$  of a photon is defined as the magnitude of the projection of the photon momentum on the plane perpendicular to the beam axis.
- $R_9$ : The variable  $R_9$  is a measure of the overall transverse spread of the shower. It is the ratio of the energy deposited in the ECAL inside a 3x3 crystal matrix centered on the most energetic crystal in the supercluster to the supercluster raw energy.
- $\sigma_{i\eta i\eta}$ : The shower width  $\sigma_{i\eta i\eta}$  is the log-fractional energy-weighted spread within the 5x5 crystal matrix centered on the crystal with the largest energy deposit in the supercluster. The symbol  $i\eta$  indicates that the variable is obtained by measuring position by counting crystals.
- ECAL isolation: The ECAL isolation is the sum of all energy deposits in the ECAL within a cone of  $\Delta R < 0.3$  centered on the photon.
- Track isolation: The track isolation is the sum of the energies of tracks in the tracker within a cone of  $\Delta R < 0.3$  centered on the photon.
- H/E: The ratio between the energy deposited in the HCAL tower closest to the supercluster position and the energy deposited to that supercluster in the ECAL is referred to as H/E.

All photons are required to pass the H/E and loose  $R_9$  cuts in \_HE\_R9Id\_, and

TABLE 6.2

## PRIMARY TRIGGER REQUIREMENTS

Name	Cuts
Diphoton30_18_	Leading photon $p_T > 30$ GeV Sub-leading photon $p_T > 18$ GeV
R9Id_	$R_9 > 0.85$
IsoCaloId_	$\sigma_{i\eta i\eta} < 0.015$ ECAL isolation $< (6 + 0.012 \times \text{Photon } E_T)$ Track isolation $< (6 + 0.002 \times \text{Photon } E_T)$
HE_R9Id_	$H/E < 0.1$ $R_9 > 0.5$
Mass90_	$m_{\gamma\gamma} > 90$ GeV

Definition of cuts used in the primary analysis trigger, HLT\_Diphoton30\_18\_R9Id\_OR\_IsoCaloId\_AND\_HE\_R9Id\_Mass90\_v\*

either the tighter  $R_9$  cuts in \_R9Id\_ or the isolation and shape cuts in \_IsoCaloId\_. The leading leg of the filter requires the photon candidate to be matched to an L1 seed. It can be matched to one of several SingleEG and DoubleEG L1 filters, but the largest contribution comes from the lowest unscaled triggers: namely, SingleEG40 and DoubleEG\_22\_15. Both photons must satisfy the trailing filter, which is unseeded. In addition to the cuts listed above, the invariant mass of the diphoton system is required to be greater than 90 GeV.

The control trigger shares all of the same requirements as the primary trigger, with two exceptions: the invariant mass of the two electromagnetic objects is required to be greater than 55 GeV rather than 90 GeV, and both electromagnetic objects are

required to be matched to a pixel seed. A pixel seed is defined as at least two hits in the pixel detectors that are consistent with the location of the energy deposit in the ECAL.

### 6.3 Trigger efficiency

An important input to the analysis is the overall trigger efficiency. Due to the similarity of the ECAL response to electrons and photons, the trigger efficiency can be calculated from  $Z \rightarrow ee$  events in data using the tag-and-probe method. In this method, two electron candidates are required. One serves as the “tag” and is required to pass loose photon identification criteria. The second electron candidate serves as the “probe” and has to satisfy the same selection criteria as our offline photon identification (see Section 7.1). In order to ensure a high purity of electromagnetic objects, the invariant mass of the di-electron system must be between 75 and 105 GeV. For this study, tag-and-probe events were collected by requiring that the tag pass a single electron control trigger, `HLT_Ele27_WPTight_Gsf`.

The efficiency  $\epsilon$  of the HLT path or trigger filter that is being studied is given by the following equation, where  $N_{total}$  is the total number of tag-and-probe pairs passing all requirements, and  $N_{pass}$  is the number of tag-and-probe pairs in which the probe passes the trigger filter.

$$\epsilon_{trig} = N_{pass}/N_{total} \tag{6.1}$$

#### 6.3.1 Efficiency of primary trigger

Because our analysis trigger is seeded by the OR of multiple SingleEG and DoubleEG L1 seeds, the names of which changed over the course of the 2016 data-taking period, calculating the L1 efficiency on its own proved to be tricky. Instead, it was simpler to calculate the total efficiency with which photon candidates pass both the

L1 seed and the leading leg of the HLT path. This efficiency as a function of photon  $p_T$  is shown in Figure 6.3. The efficiency was fit to an error function to calculate the overall efficiency at the plateau. For photon  $p_T > 40$  GeV, the leading filter is 98.2% efficient.

Tag-and-probe objects for the trailing leg efficiency must pass the same set of requirements as those used in the leading leg efficiency calculation, with the additional requirement that the tag must pass the leading filter. This requirement arises from the way HLT modules are structured. Filters are processed sequentially, and if an event fails one filter, the subsequent filters are skipped. Figure 6.4 shows the efficiency of the trailing filter as a function of photon  $p_T$ . For  $p_T > 40$  GeV, the trigger is 99.8% efficient.

Finally, we calculated the efficiency of the trigger with respect to the diphoton invariant mass. For this calculation, we required two photons passing our analysis selection criteria, two photons satisfying the trailing leg of the trigger, and one photon passing the leading leg of the trigger. The efficiency was given by the number of diphoton events passing the full HLT path over the total number of diphoton events passing our requirements. The efficiency of the trigger as a function of invariant mass is shown in Figure 6.5. For  $m_{\gamma\gamma} > 100$  GeV, the trigger is 99.4% efficient.

The efficiency of the trigger as a whole is the product of all three efficiencies. Two factors of the trailing leg efficiency are needed since both photons are required to pass that leg:

$$\epsilon_{trig} = \epsilon_{lead} \times \epsilon_{trail}^2 \times \epsilon_{mass} = 97.2\% \quad (6.2)$$

### 6.3.2 Efficiency of double electron trigger

In the secondary trigger listed in Table 6.1, the pixel seed requirement is applied only to the trailing leg of the trigger. This additional requirement results in a signifi-

cantly lower overall efficiency for this leg. As shown in Figure 6.6, the trigger is only 90.4% efficient for  $p_T > 40$  GeV. Since all other cuts are the same between the two triggers, however, the leading leg efficiency is the same as that shown in Figure 6.3. This results in an overall trigger efficiency of 79.8% for the control trigger.



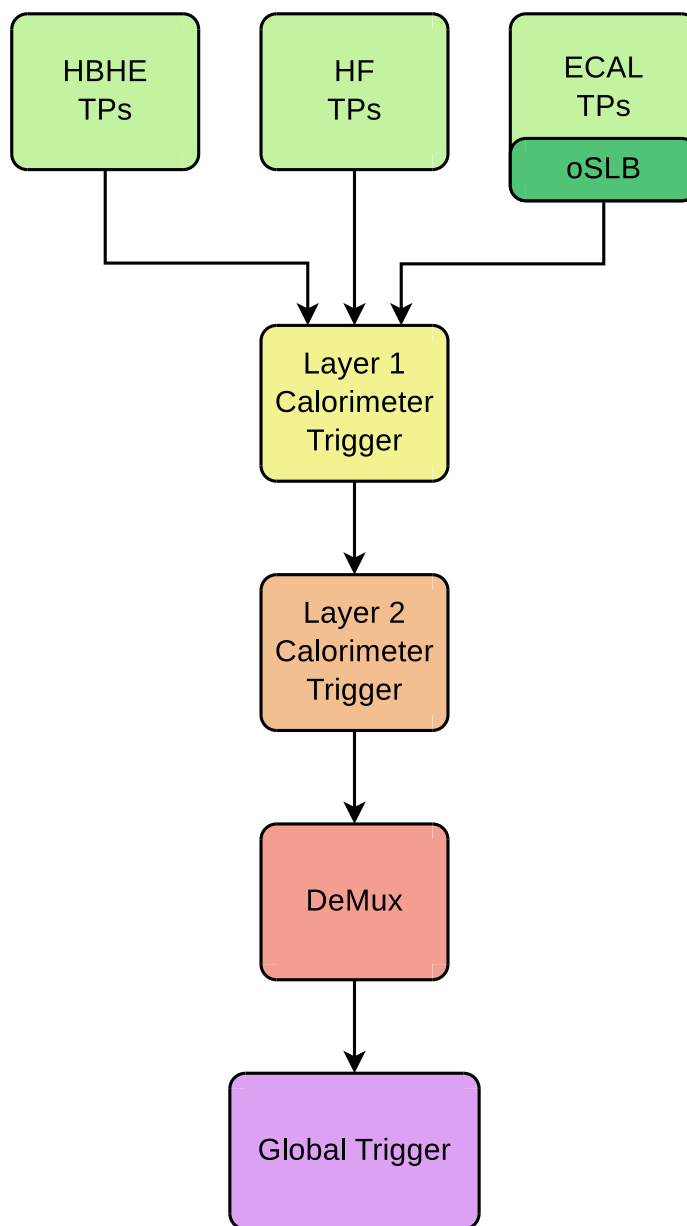


Figure 6.1: Schematic showing the overall structure of the calorimeter half of the L1 trigger. Reprinted from Reference [7].

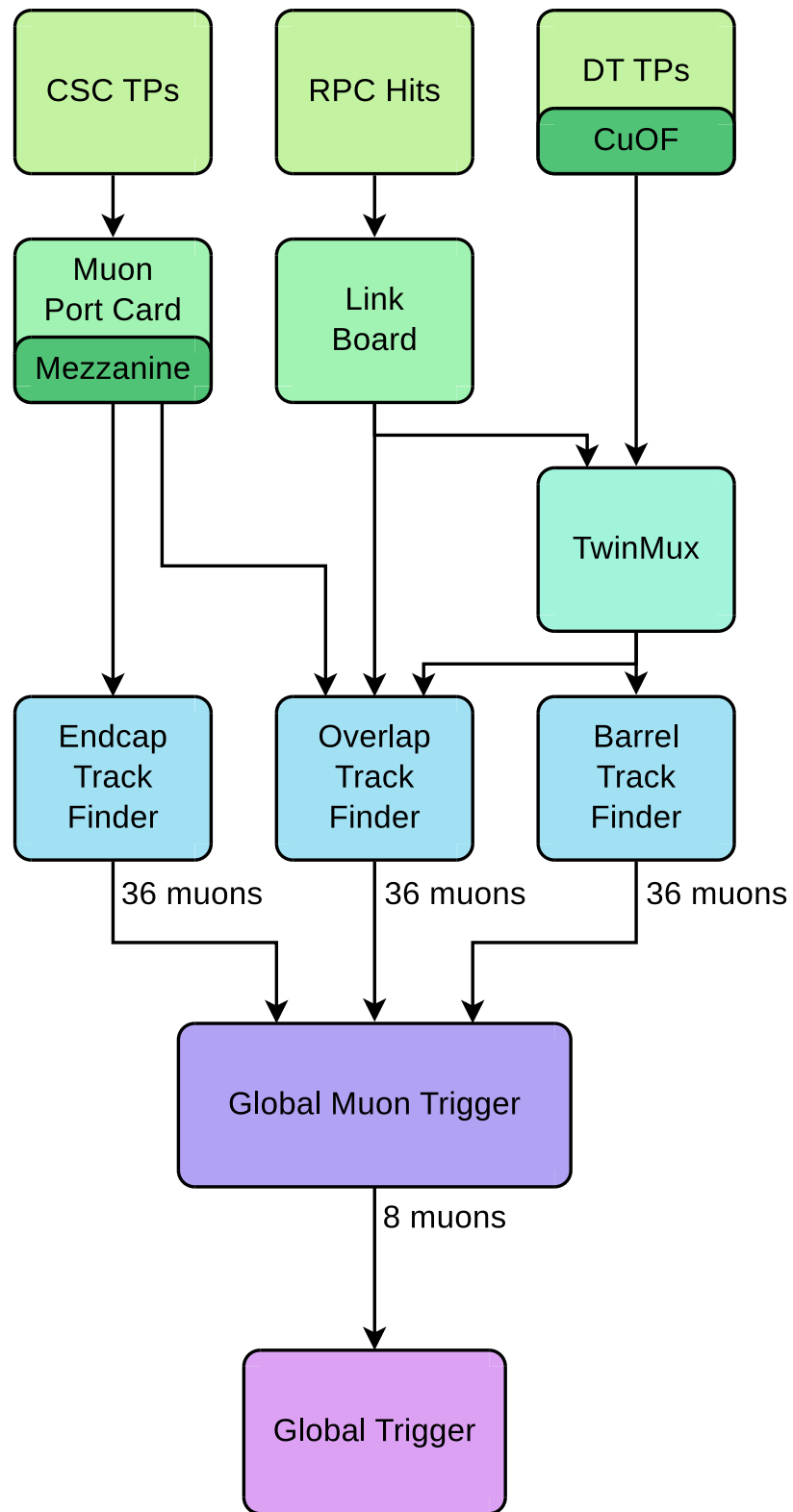


Figure 6.2: Schematic showing the overall structure of the muon half of the L1 trigger. Reprinted from Reference [7].

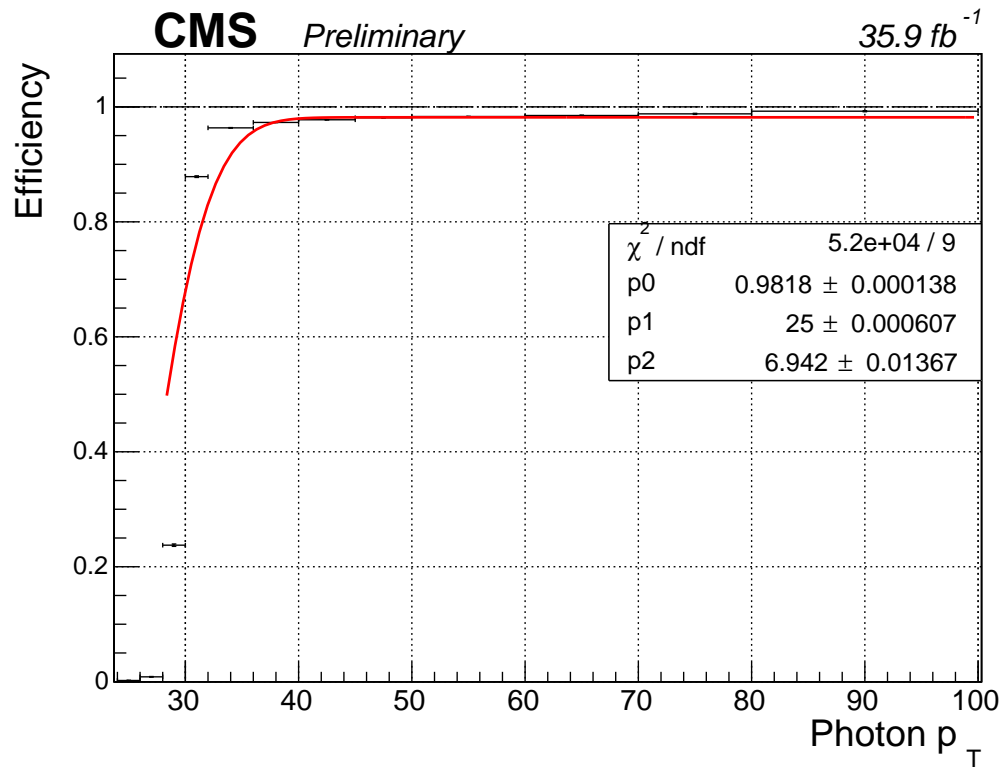


Figure 6.3: Efficiency of the L1 seed and the leading leg of the primary analysis trigger with respect to photon  $p_T$ . For  $p_T > 40$  GeV, the trigger is 98.2% efficient.

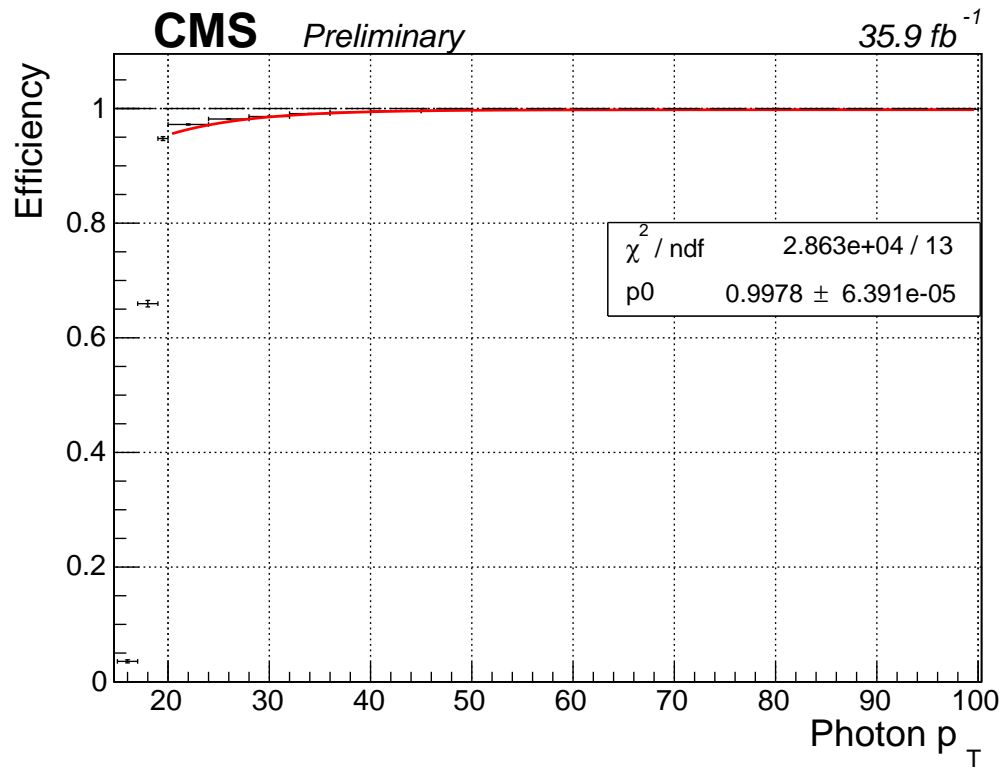


Figure 6.4: Efficiency of the trailing leg of the primary analysis trigger with respect to photon  $p_T$ . For  $p_T > 40$  GeV, the trigger is 99.8% efficient.

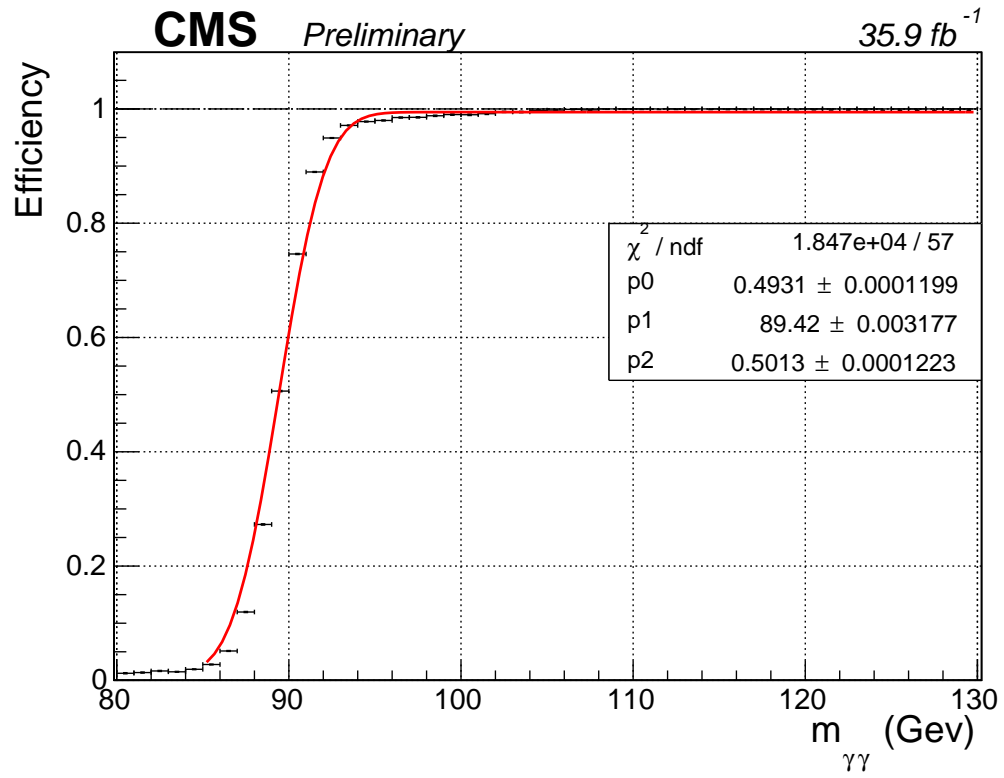


Figure 6.5: Efficiency of the primary analysis trigger with respect to the invariant mass of the diphoton system. For  $m_{\gamma\gamma} > 100$  GeV, the trigger is 99.4% efficient.

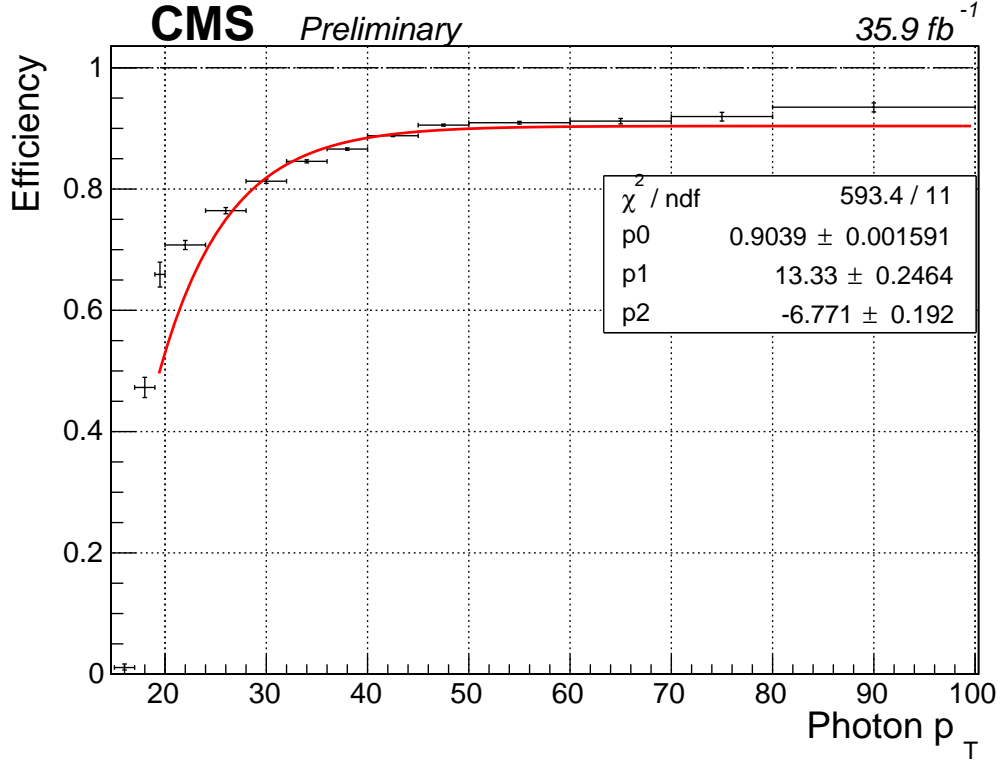


Figure 6.6: Efficiency of the trailing leg of the control trigger with respect to photon  $p_T$ . For  $p_T > 40$  GeV, the trigger is 90.4% efficient. The drop in efficiency with respect to the primary analysis trigger is caused by requiring both electromagnetic objects to be matched to a pixel seed.

## CHAPTER 7

### Sample Selection

#### 7.1 Object Selection

In addition to photons and electrons, a third, orthogonal object definition referred to as “fake” photons is also used. Fakes are primarily electromagnetically-rich jets that have been misidentified as photons. As will be described in Chapter 8, a control sample made up of events with two fakes is used to estimate the QCD background.

##### 7.1.1 Photon identification

##### 7.1.2 Electron identification

##### 7.1.3 Fake identification

Electromagnetically-rich jets that have been misidentified as photons make up the majority of the objects that satisfy our “fake” definition. These objects are taken from a photon ID sideband. Fakes must pass all of the photon identification criteria described in Section 7.1.1, except they are required to fail either the  $\sigma_{i\eta i\eta}$  or the charged hadron isolation requirement. This is described in detail in Table 7.2 (XX better if column widths the same XX)

STOLEN The uncorrected isolation variables are PF objects Pt sums in a cone 0.3 around the photon and also subjected to the photon footprint removal using the association map of the PF objects. The charge hadron Isolation is calculated using the PF objects that their tracks are associated with the primary vertex only, thus removing pile-up charge hadrons. STOLEN

TABLE 7.1

## ELECTROMAGNETIC OBJECT DEFINITIONS

ID Requirement	Photons	Electrons	Fakes
Pixel seed veto	Applied	Reversed	Applied
$\sigma_{i\eta i\eta}$	$< 0.01022$		$0.01022 < \sigma_{i\eta i\eta} < 0.015$
Charged hadron isolation	$< 0.441$		$0.441 < iso < 15$
Photon isolation	$< 2.571 + 0.0047 p_T$		
Neutral hadron isolation	$< 2.2725 + 0.0148 p_T + 0.000017 p_T^2$		
R9	$> 0.5$		
H/E	$< 0.0396$		

Definitions of photon, electron, and fakes used to define the signal and control samples for this analysis. “Fakes” refer to jets that have been misidentified as photons. The definitions of each of the variables used in the object ID’s can be found in Section ??.

## 7.1.4 Object Cleaning

In order to avoid double counting particles, a set of object cleaning rules are applied. First, because muons are reconstructed with a higher purity than any other particle, any electromagnetic object (photon, electron, or fake) that is within  $\Delta R < 0.3$  of a muon candidate is removed. Second, any photon that overlaps within  $\Delta R < 0.3$  of an electron is removed. Finally, if a fake overlaps with an electron or photon candidate within  $\Delta R < 0.4$ , the fake candidate is removed. The larger  $\Delta R$  separation for fakes is due to the fact that fakes are primarily jets. As described in Section XX, jets are reconstructed using the anti- $p_T$  algorithm with a distance parameter of 0.4.



TABLE 7.2

## EFFECTIVE AREAS FOR ISOLATION CORRECTIONS

$ \eta $ Range	Photon Iso	Neutral Hadron Iso	Charged Hadron Iso
$ \eta  < 1.0$	0.120	0.0597	0.0360
$1.0 <  \eta  < 1.479$	0.1107	0.0807	0.0377

Effective areas used in the definition of photon, charged hadron, and neutral hadron isolation values.

## 7.1.5 Lepton veto

In addition to the cuts described above, any event that contains additional muons or electrons is vetoed. No additional leptons are present in the SUSY signals of interest, so applying a lepton veto will not hurt our signal sensitivity. More importantly, by vetoing on the presence of additional leptons, our analysis becomes orthogonal to other CMS searches for gauge-mediated supersymmetry breaking with photons in the final state. As described in Section XX, this is particularly important for the combination paper.

## 7.1.6 Signal region and control samples

After the individual objects are defined and identified, each event is sorted into one of four exclusive categories based on the electromagnetic objects with the highest  $p_T$  in the event. Events with two photons comprise the signal diphoton sample, referred to as the  $\gamma\gamma$  sample. Events with two electrons, one electron and one photon, or two fakes are categorized as  $ee$ ,  $e\gamma$ , or  $ff$  events, respectively.

The  $\gamma\gamma$ ,  $ff$ , and  $e\gamma$  samples are required to pass the primary trigger described in Section 6.2. In order to ensure that the events pass the trigger with a high efficiency, the invariant mass of the two electromagnetic objects is required to be greater than

105 GeV. The  $ee$  sample, on the other hand, is collected using the control trigger listed in Table 6.1 and is required to satisfy  $75 < m_{ee} < 105$  GeV. Chapter 8 will explain in detail how each of these samples is used in this analysis.

## CHAPTER 8

### DATA ANALYSIS AND BACKGROUND ESTIMATION METHODS

#### 8.1 Overview

There are several Standard Model processes that can mimic our signal events. The largest background contribution comes from quantum chromodynamics (QCD) processes. These are primarily multi-jet events, where electromagnetically-rich jets are misidentified as photons, but can also include processes with true photons either from associated photon production or initial-state radiation. In both cases, there is no inherent  $E_T^{\text{miss}}$  in the event. Instead, the measured  $E_T^{\text{miss}}$  is actually the result of mismeasured hadronic activity. As described in Section 8.2, this background is estimated in an entirely data-driven way using a control region derived from a sideband of our photon ID.

The second-largest background is the electroweak (EWK) background. This background is comprised of  $W\gamma$  or  $W+\text{jet}$  events where  $W \rightarrow e\nu$ . In this case, there is inherent  $E_T^{\text{miss}}$  from the neutrino, and these events can mimic our signal topology if the electron is misidentified as a photon. By measuring the misidentification rate in data, we can use an  $e\gamma$  control sample to estimate the contribution from the EWK background. The EWK background estimation method is described in detail in Section 8.3.

Finally, there is an irreducible background from  $Z\gamma\gamma \rightarrow \nu\nu\gamma\gamma$  events. This background is modeled via simulation and is described in Section 8.4.

## 8.2 QCD background

Due to the large QCD cross section, the most significant background for this analysis comes from QCD events without true  $E_T^{\text{miss}}$  and without two real photons. The observed  $E_T^{\text{miss}}$  is the result of mismeasured hadronic activity, and in most cases the “photons” are misidentified jets with a large electromagnetic component.

To estimate the contribution from the QCD background in our signal region, we use the “fake” object selection that was described in Section 7.1. The fake identification criteria is orthogonal to the nominal photon identification, and therefore provides a sideband that can be used as a control region. The  $E_T^{\text{miss}}$  tail of the QCD background is modeled using a “fake-fake” (“ $ff$ ”) control sample made up of events with two fakes that pass the additional criteria outlined in Section 7.1.6.

### 8.2.1 Di-EM $p_T$ reweighting

Because the  $E_T^{\text{miss}}$  in the QCD background and the  $ff$  control sample arises from poorly measured hadronic activity, it is important that the amount of hadronic activity in the control sample matches that of the  $\gamma\gamma$  events we are trying to model.

To account for potential differences between the samples, we define a variable referred to as the “di-EM  $p_T$ ” of an event. Di-EM  $p_T$  is defined as the magnitude of the vector sum of the transverse momentum of the two electromagnetic objects (photons, electrons or fakes):

$$\vec{p}_T^{\text{di-EM}} = \vec{p}_{T1} + \vec{p}_{T2} \quad (8.1)$$

As illustrated in Figure 8.1, the di-EM  $p_T$  variable is used as a measure of the total hadronic recoil. Because CMS measures the energies of electromagnetic objects with greater precision than the energy of jets, this variable is a more accurate representation of the hadronic recoil than simply adding up the transverse momentum of

the jets themselves.

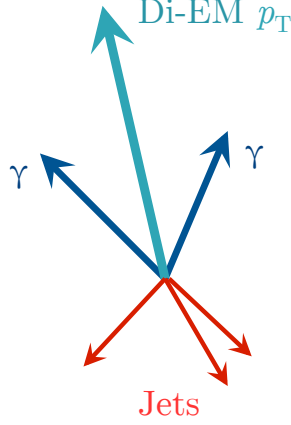


Figure 8.1: The di-EM  $p_T$  vector, shown in light blue, is the vector sum of the  $p_T$  of the two photons in the event, shown in blue. The magnitude of the di-EM  $p_T$  vector is used to model the hadronic recoil, shown in red.

The di-EM  $p_T$  distributions for the  $\gamma\gamma$  candidate sample and the  $ff$  control sample are shown in Figure 8.2. The  $ff$  events are the reweighted using the  $\gamma\gamma/ff$  ratios displayed in the ratio plot of Figure 8.2. The  $ff$   $E_T^{\text{miss}}$  distribution is normalized to the  $E_T^{\text{miss}} < 50$  GeV region of the  $\gamma\gamma$  sample, where signal contamination is minimal.

The unweighted  $ff$  and  $\gamma\gamma$   $E_T^{\text{miss}}$  distributions are shown in Figure 8.3. A comparison between the unweighted and di-EM  $p_T$  reweighted  $ff$   $E_T^{\text{miss}}$  distributions is shown in Figure 8.4. Finally, Figure 8.5 compares the reweighted  $ff$   $E_T^{\text{miss}}$  distribution to the candidate  $\gamma\gamma$   $E_T^{\text{miss}}$  distribution.

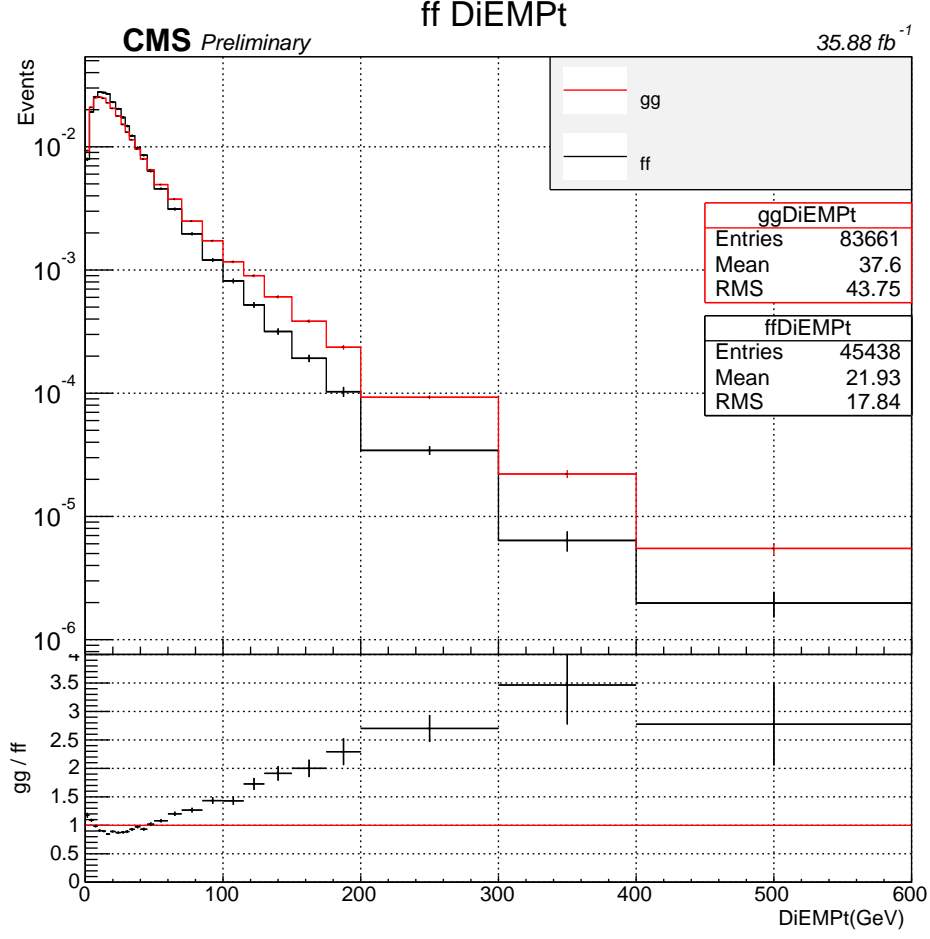


Figure 8.2: Di-EM  $p_T$  distributions of the  $ff$  control sample (black) and the  $\gamma\gamma$  candidate sample (red). The ratio plot on the bottom shows the  $\gamma\gamma/ff$  ratios that are used to reweight the  $ff$   $E_T^{\text{miss}}$  distribution.

### 8.2.2 Cross check on QCD background

In order to set a systematic uncertainty on the overall  $E_T^{\text{miss}}$  shape predicted using the di-EM  $p_T$  reweighting method, we sought an alternate way to estimate the QCD background. This cross check relies on the assumption that the ratio of  $\gamma\gamma$  events to  $ff$  events should not depend sensitively on  $E_T^{\text{miss}}$ . If this assumption is true, then we should be able to extrapolate from the low- $E_T^{\text{miss}}$  control region to the high- $E_T^{\text{miss}}$  signal region.

Figure 8.6 shows the ratio of  $\gamma\gamma/ff$  as a function of  $E_T^{\text{miss}}$  in the  $E_T^{\text{miss}} < 100$  GeV

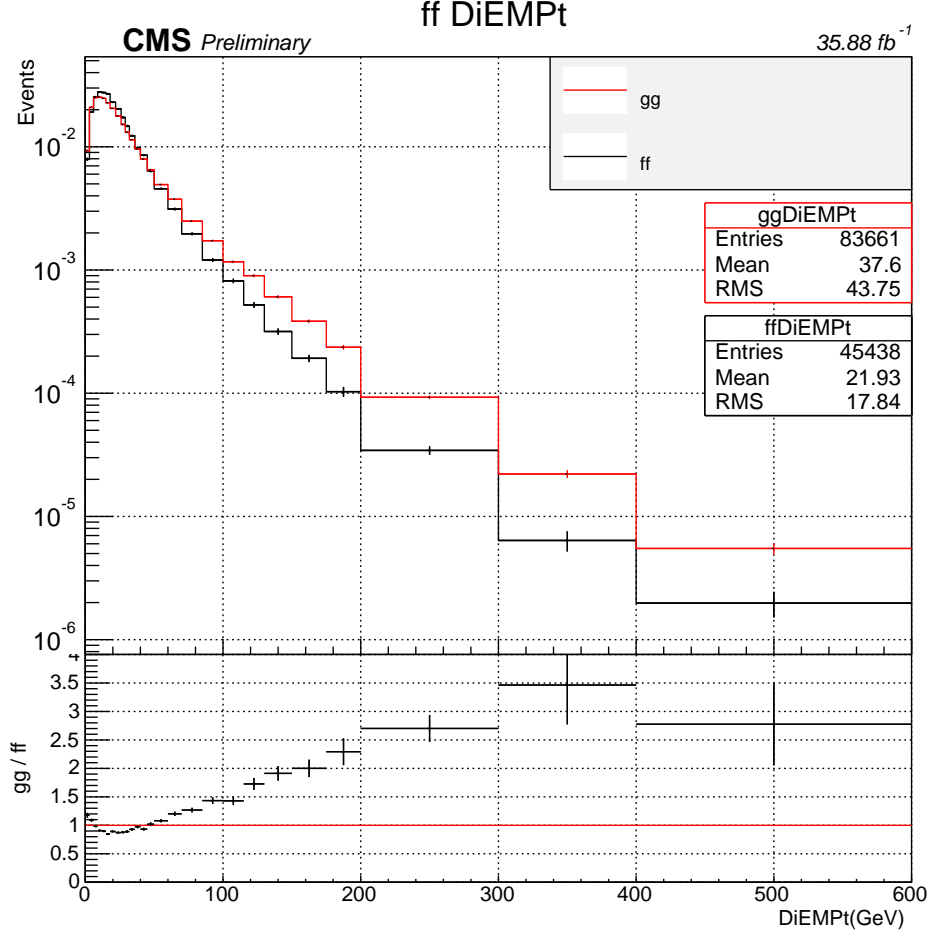


Figure 8.3: Di-EM  $p_T$  distributions of the  $ff$  control sample (black) and the  $\gamma\gamma$  candidate sample (red). The ratio plot on the bottom shows the  $\gamma\gamma/ff$  ratios that are used to reweight the  $ff$   $E_T^{\text{miss}}$  distribution.

control region. The ratio has been fit to a constant function  $f(E_T^{\text{miss}})$ . Using this function, the expected number of  $\gamma\gamma$  events in bin  $i$  of the signal region is given by the following equation:

$$N_{\gamma\gamma}^i = f(E_T^{\text{miss}}) \times N_{ff}^i \quad (8.2)$$

Table 8.1 compares the expected QCD background contribution as predicted with the cross check method to that predicted with the di-EM  $p_T$  reweighting method. For the cross check method, the uncertainty arising from the choice of fitting function is

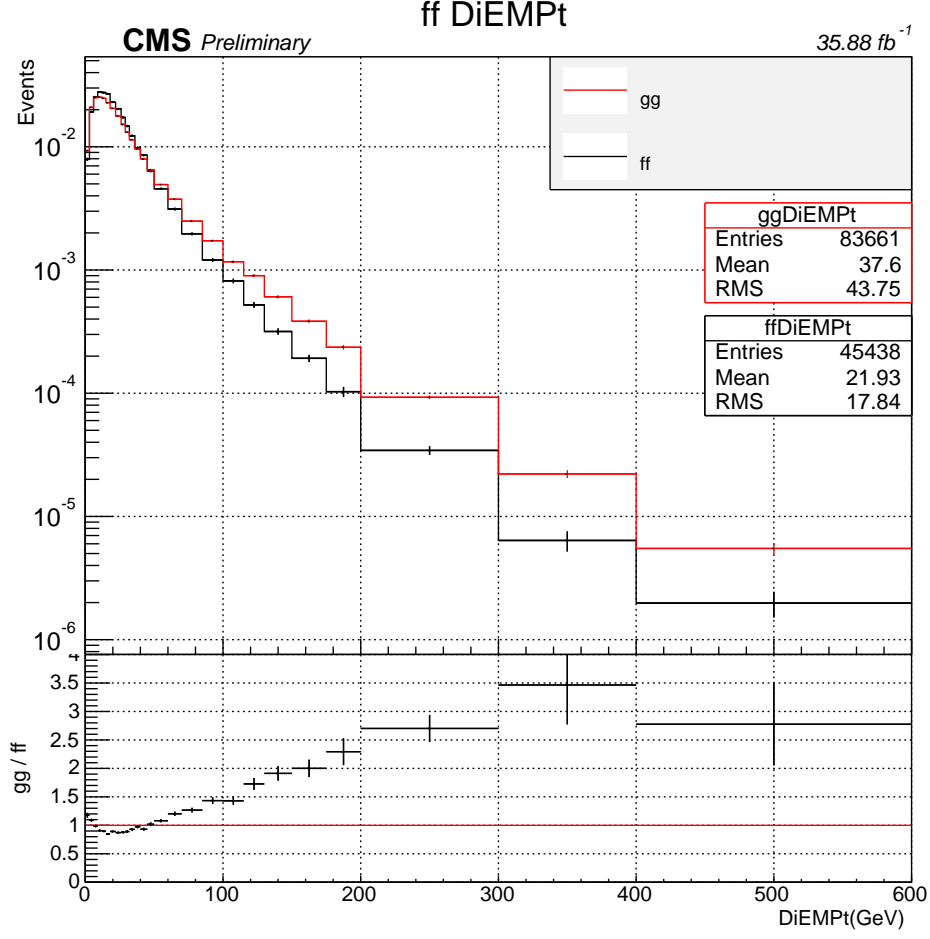


Figure 8.4: Di-EM  $p_T$  distributions of the  $ff$  control sample (black) and the  $\gamma\gamma$  candidate sample (red). The ratio plot on the bottom shows the  $\gamma\gamma/ff$  ratios that are used to reweight the  $ff$   $E_T^{\text{miss}}$  distribution.

estimated by taking the difference between the results when fitting to a constant to the results when fitting to a linear function. In addition, the cross check uncertainties listed in Table 8.1 include the statistical uncertainty from the limited control sample statistics and the  $1\sigma$  uncertainties from the fit. For the di-EM  $p_T$  reweighting method, the uncertainties include the statistical uncertainty and the uncertainty from the di-EM  $p_T$  reweighting procedure (this systematic is described in detail in Section 8.2.3).

The two methods give overlapping predictions within uncertainty, and therefore



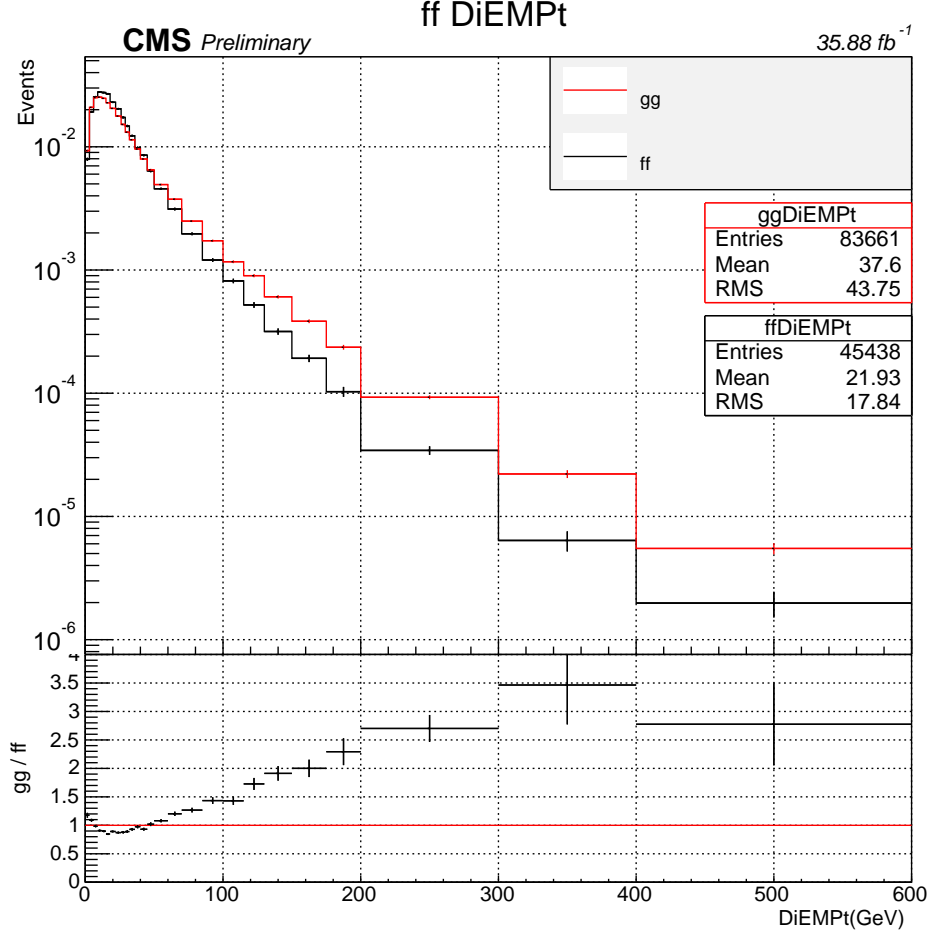


Figure 8.5: Di-EM  $p_T$  distributions of the  $ff$  control sample (black) and the  $\gamma\gamma$  candidate sample (red). The ratio plot on the bottom shows the  $\gamma\gamma/ff$  ratios that are used to reweight the  $ff$   $E_T^{\text{miss}}$  distribution.

this cross check serves to validate our di-EM  $p_T$  reweighting background estimation method. The difference between the two methods is taken as a systematic uncertainty on the overall  $E_T^{\text{miss}}$  shape.

### 8.2.3 Systematic uncertainties on the QCD background

## 8.3 Electroweak background

The subdominant background for this search is comprised of  $W\gamma$  and  $W+\text{jet}$  events where  $W \rightarrow e\nu$  and the electron is misidentified as a photon. This background

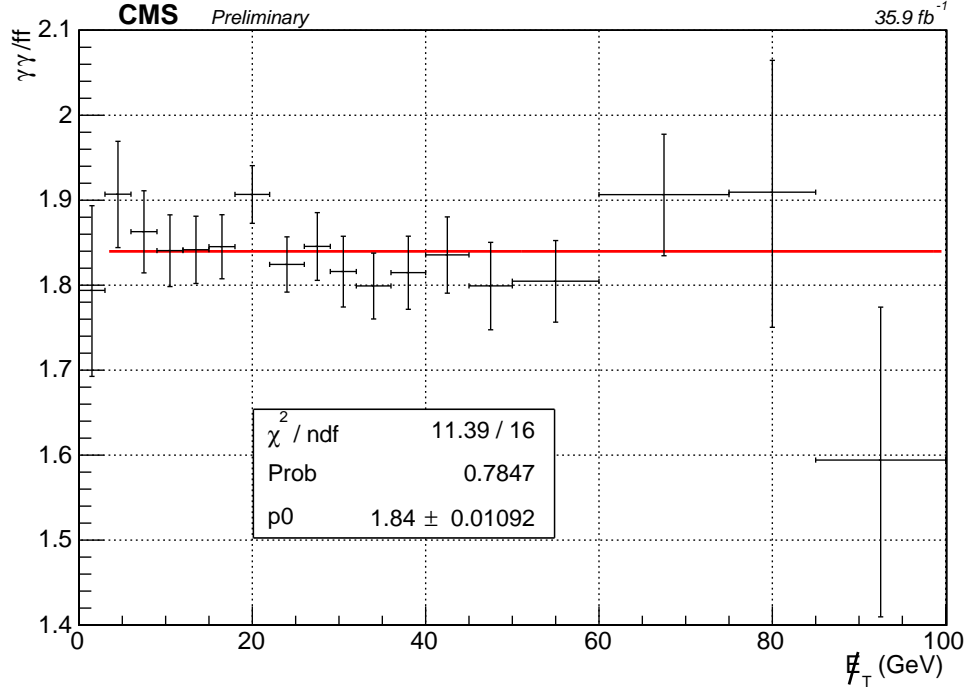


Figure 8.6: The ratio of  $\gamma\gamma$  to  $ff$  events for  $E_T^{\text{miss}} < 100$  GeV. The ratio has been fit to a constant function, which can then be used to extrapolate to the  $E_T^{\text{miss}} > 100$  GeV signal region. This provides an alternative estimation method for the QCD background.

is referred to as the electroweak (EWK) background. Unlike the QCD background, there is inherent  $E_T^{\text{miss}}$  in these events from the escaping neutrino.

To estimate this background, we first calculate the rate at which electrons are misidentified as photons. This is done by comparing the invariant mass peak in a double electron sample with the invariant mass peak in a sample of events with one electron and one photon. The calculation of this "fake rate" is described in detail in Section 8.3.1. To get the final expected contribution from the EWK background in our signal region, the fake rate is used to calculate a transfer factor. By applying the transfer factor to an  $e\gamma$  control sample, we are able to estimate how many of our candidate  $\gamma\gamma$  events are actually events with one photon and one electron that has been misidentified as a photon.

TABLE 8.1

COMPARISON BETWEEN REWEIGHTING METHOD AND RATIO  
METHOD FOR THE QCD BACKGROUND ESTIMATE

$E_T^{\text{miss}}$ bin (Gev)	Reweighted ff estimate	Ratio method estimate
100 – 115	$69.23^{+15.18}_{-12.68}$	$55.19^{+12.03}_{-10.33}$
115 – 130	$30.89^{+11.76}_{-8.82}$	$22.08^{+8.39}_{-6.59}$
130 – 150	$25.98^{+11.95}_{-8.61}$	$16.55^{+7.56}_{-5.69}$
150 – 185	$20.49^{+10.12}_{-7.11}$	$14.72^{+7.26}_{-5.45}$
185 – 250	$8.74^{+11.65}_{-5.89}$	$3.68^{+4.85}_{-2.47}$
> 250	$5.13^{+11.86}_{-4.43}$	$1.84^{+4.23}_{-1.60}$

Comparison of the background estimate using the di-EM  $p_T$  reweighting method and the  $\gamma\gamma/ff$  ratio method. The uncertainties on the di-EM  $p_T$  reweighting method include the statistical uncertainties and the reweighting uncertainty. The uncertainties on the ratio method include the statistical uncertainties, the uncertainties in the fit parameter, and the uncertainty from the choice of fit function.

### 8.3.1 Fake rate calculation

### 8.3.2 Calculating EWK estimate

### 8.3.3 Composition of $e\gamma$ sample

The  $e\gamma$  control sample is primarily made up of  $\gamma + \text{jets}$  or  $W\gamma$  events. A data versus Monte Carlo comparison of this control sample is shown in Figure ???. The data distribution was fit using the simulated  $\gamma + \text{jet}$  and  $W\gamma$  shapes as templates. From the fit, it was determined that 80% of the  $e\gamma$  events observed in data are from  $\gamma + \text{jets}$  processes, and the remaining 20% are  $W\gamma$  events. In the  $E_T^{\text{miss}} > 100$  GeV signal region, however,  $W\gamma$  events dominate. For  $100 < E_T^{\text{miss}} < 130$  GeV, 12.1% of the MC prediction is made up of  $\gamma$  plus jet processes. For  $E_T^{\text{miss}} > 130$  GeV, however,  $\gamma$  plus jet events contribute only 2.6%.

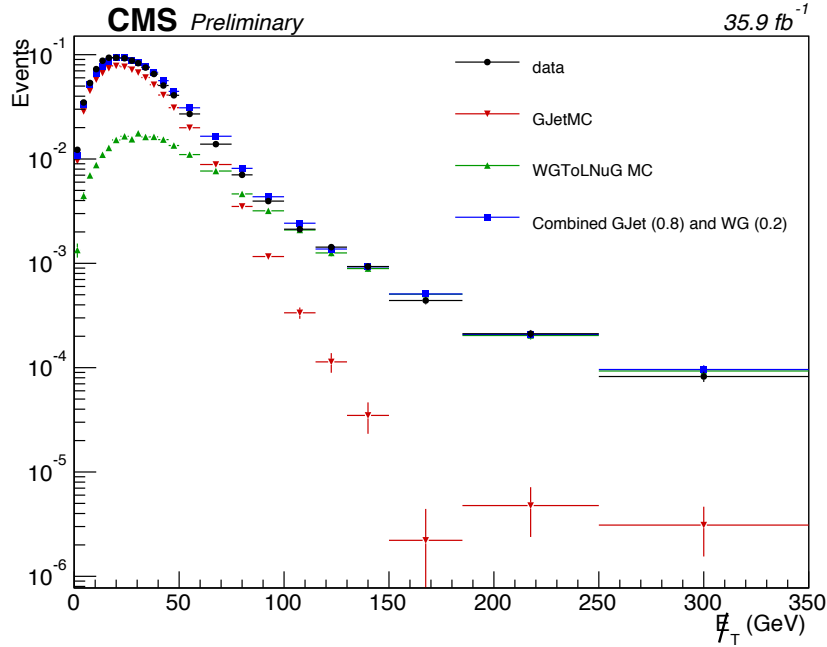


Figure 8.7: Data versus Monte Carlo comparison of the  $e\gamma$  control sample  $E_T^{\text{miss}}$  distribution. To determine the relative contributions of the  $\gamma + \text{jet}$  and  $W\gamma$  processes, the data distribution was fit using the  $\gamma + \text{jet}$  and  $W\gamma$  shapes as templates. The data are shown in black, and the total MC prediction is shown in blue. The  $\gamma + \text{jet}$  MC (red) was scaled to 80% of the observed events in data, and the  $W\gamma$  MC (green) was scaled to 20% of the data distribution.

#### 8.3.4 EWK results

Table XX shows the final EWK estimate in the six signal region bins. There are two sources of uncertainty on the EWK estimate: the statistical uncertainty from the limited number of  $e\gamma$  events observed in data, and the 30% systematic uncertainty from the calculation of the fake rate. Both of these are displayed in Table XX.

#### 8.4 Irreducible background

In addition to the QCD and EWK backgrounds, there is a small irreducible background from  $Z\gamma\gamma \rightarrow \nu\nu\gamma\gamma$  events. Because there is inherent  $E_T^{\text{miss}}$  from this process, it is not included in the QCD background, and because it has two true photons, it

is not included in the EWK background either. We model this background using MC simulation, and assign a 50% uncertainty to the estimate to cover any potential mismodeling.

XX List datasets XX

#### 8.4.1 Cross checking normalization

To cross check the  $Z\gamma\gamma$  normalization obtained from data, we compare the observed number of  $Z\gamma\gamma \rightarrow ee\gamma\gamma$  events seen in data to that predicted via MC. These processes are different only in the decay mode of the  $Z$  boson, and therefore should give an accurate indication of how well the process is simulated in MC. The results are shown in Table XX.

## CHAPTER 9

### Results and Interpretations

#### 9.1 Prediction versus observation

After deriving estimates for all Standard Model backgrounds, we compare the number of observed events with the expected number of background events. The  $E_T^{\text{miss}}$  distributions for the full background prediction and the unblinded data are shown in Figure 9.1. Three benchmark signal models are also shown, corresponding to the T5gg simplified model (described in detail in Section 9.2) with gluino mass equal to 1700 GeV.

Table 9.1 shows the expected and observed numbers of events for each bin in the signal region. Within the uncertainties, no excess is observed with respect to the Standard Model prediction.

#### 9.2 Simplified Models

Two simplified models are used in the interpretation of the results. The T5gg simplified model assumes gluino ( $\tilde{g}$ ) pair production and the T6gg model assumes squark ( $\tilde{q}$ ) pair production. Example decay chains for both models are shown in Figure 9.2.

In both models, the lightest supersymmetric particle (LSP) is the gravitino,  $\tilde{G}$ , which is taken to be nearly massless. The next-to-lightest supersymmetric particle (NLSP) is the neutralino,  $\tilde{\chi}_1^0$ . The models assume 100% branching fraction for  $\tilde{\chi}_1^0 \rightarrow \tilde{G}\gamma$  and  $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_1^0$  and  $\tilde{q} \rightarrow q\tilde{\chi}_1^0$ .

TABLE 9.1

## EXPECTED AND OBSERVED EVENTS IN THE SIGNAL REGION

$E_T^{\text{miss}}$ (GeV)	Exp. QCD	Exp. EWK	$Z\gamma\gamma$ events	Total exp.	Observed
100 – 115	$69.23^{+20.67}_{-18.91}$	$8.17 \pm 2.5$	$1.30 \pm 0.65$	$78.69^{+22.58}_{-20.98}$	65
115 – 130	$30.89^{+14.69}_{-12.46}$	$5.50 \pm 1.70$	$1.14 \pm 0.57$	$37.53^{+16.05}_{-14.04}$	27
130 – 150	$25.98^{+15.21}_{-12.76}$	$4.78 \pm 1.48$	$1.12 \pm 0.56$	$31.88^{+16.44}_{-14.20}$	17
150 – 185	$20.49^{+11.65}_{-9.16}$	$3.95 \pm 1.24$	$1.32 \pm 0.66$	$25.76^{+12.80}_{-10.59}$	13
185 – 250	$8.74^{+12.70}_{-7.765}$	$3.52 \pm 1.11$	$1.28 \pm 0.64$	$13.55^{+13.05}_{-8.31}$	8
$\geq 250$	$5.13^{+12.31}_{-5.514}$	$2.11 \pm 0.69$	$1.14 \pm 0.57$	$8.38^{+12.48}_{-5.88}$	10

In order to study the expected SUSY signal distributions, two signal Monte Carlo scans were produced. The T5Wg scan was produced in bins of gluino mass and neutralino mass, and the T6Wg scan was produced in bins of squark mass and neutralino mass. The leading-order event generator MADGRAPH5\_aMC@NLO [2] is used to simulate the signal samples, which were generated with either two gluinos or two squarks and up to two additional partons in the matrix element calculation. The parton showering, hadronization, multiple-parton interactions, and the underlying event were described by the PYTHIA 8 [15] program with the CUETP8M1 generator tune. The detector response is simulated using CMS fast simulation [1].

A total of 40,000 events were produced for each bin, except for bins with gluino or squark masses above 2.0 TeV, where only 20,000 events were produced per bin. For gluino masses from 1,400 to 2,500 GeV, events were generated in bins of 50 GeV. In the T6Wg scan, the squark masses ranged from 1,400 GeV to 2,050 GeV in bins of 50 GeV. The neutralino masses ranged from 10 GeV up to the mass of the gluino or squark and were binned in 100 GeV segments. Finer binning was used in the compressed region where  $M_{\tilde{\chi}_1^0}$  is within 300 GeV of  $M_{\tilde{g}}$  or  $M_{\tilde{q}}$ , and in the region with

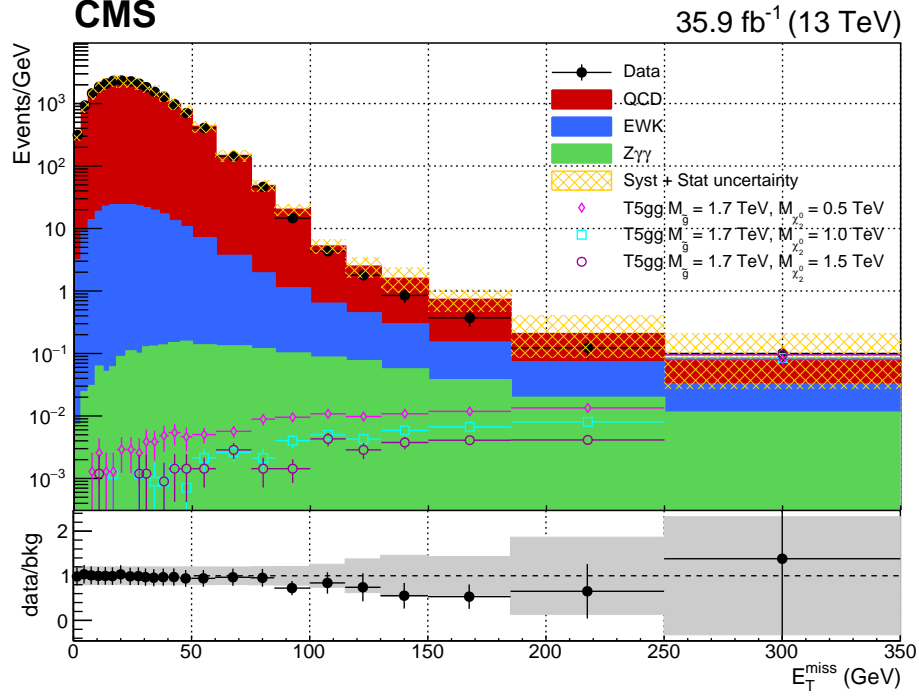


Figure 9.1:  $E_T^{\text{miss}}$  distributions for the full background estimation and the observed data. The black points represent the observed  $E_T^{\text{miss}}$  distribution. The QCD background is shown in red, the EWK background is shown in blue, and the  $Z\gamma\gamma$  background is shown in green. The combined uncertainty on the background estimation is shown in yellow.

low  $M_{\tilde{\chi}_1^0}$ . These mass ranges were selected to overlap and expand upon the mass ranges excluded by previous searches [? ? ].

The parton distribution functions are obtained from NNPDF3.0 [3] The cross sections are calculated at NLO+NLL accuracy [13, 4], with all the unconsidered sparticles assumed to be heavy and decoupled. The uncertainties on the cross sections are calculated as described in Ref. [5].

### 9.2.1 Signal acceptance and efficiency

The generated SUSY signal events undergo the same event selection described in Chapter 7. For every mass point, the number of events expected in each of the



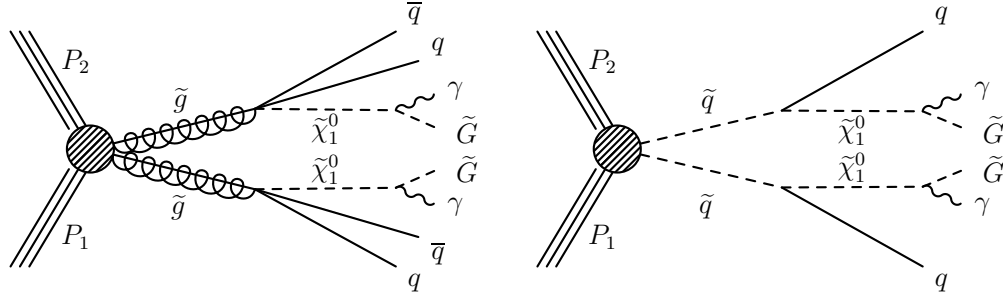


Figure 9.2: Diagrams showing the production of signal events in the collision of two protons with four momenta  $P_1$  and  $P_2$ . In gluino  $\tilde{g}$  pair production in the T5gg simplified model (left), the gluino decays to an antiquark  $\bar{q}$ , quark  $q$ , and neutralino  $\tilde{\chi}_1^0$ . In squark  $\tilde{q}$  pair production in the T6gg simplified model (right), the squark decays to a quark and a neutralino. In both cases, the neutralino subsequently decays to a photon  $\gamma$  and a gravitino  $\tilde{G}$ . In the diagram on the right, we do not distinguish between squarks and antisquarks.

six signal region bins is calculated. In addition, the overall acceptance  $\times$  efficiency ( $A \times \epsilon$ ), defined as the number of events passing the full  $\gamma\gamma$  selection divided by the total number of generated events, is calculated across the full 2D mass plane. Figures 9.3 and 9.4 show the  $A \times \epsilon$  distributions for the T5gg and T6gg simplified model frameworks.

At low neutralino mass, the  $A \times \epsilon$  decreases because the photons are softer and more likely to fail the  $p_T > 40$  GeV cut. There is also more jet energy at low neutralino masses, which increases the chance that photons will overlap with a jet and fail the isolation requirements.

### 9.3 Confidence level calculations

The statistical uncertainties are modeled with gamma distributions, where  $x$  is the expected event yield,  $\alpha$  is the original number of events in the control sample, and  $\beta$  is the transfer factor,  $\beta = x/\alpha$ .

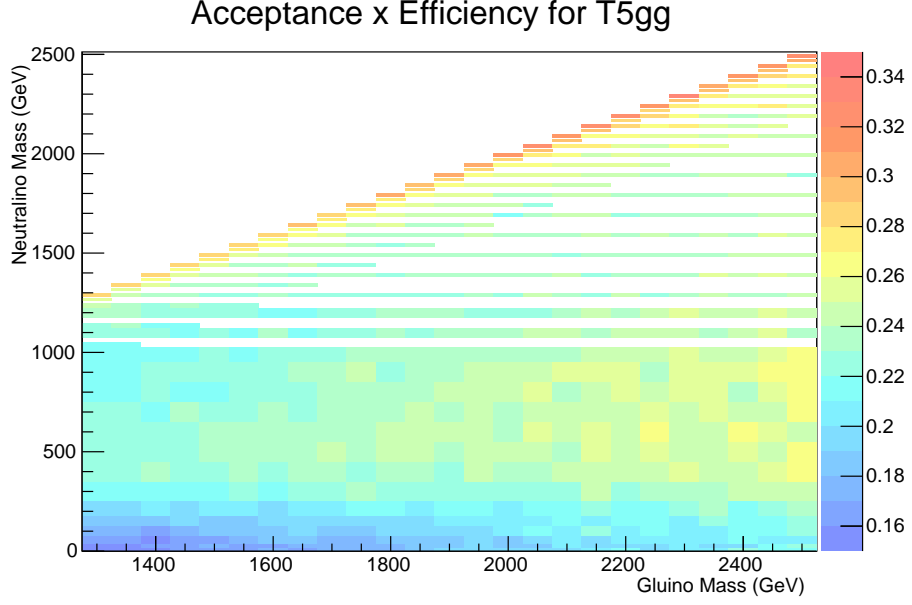


Figure 9.3: Acceptance  $\times$  efficiency for the T5gg simplified model.

$$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad (9.1)$$

The statistical uncertainties include those on the observed number of events in the signal region in data, uncertainties from the limited number of events in the control regions used for the data-driven background estimation methods, and the MC statistical uncertainties in the SUSY signal MC and the MC used to model the  $Z\gamma\gamma \rightarrow ee\gamma\gamma$  background.

All of the systematic uncertainties described in Chapter 8 are treated as nuisance parameters with a log-normal probability distribution. The peak of the probability distribution is set to the nominal background estimate for that bin, and the size of the systematic uncertainty sets the width of the distribution.

#### 9.4 Limits and Interpretations

The  $\text{CL}_s$  method [12, 14] is applied to our analysis.

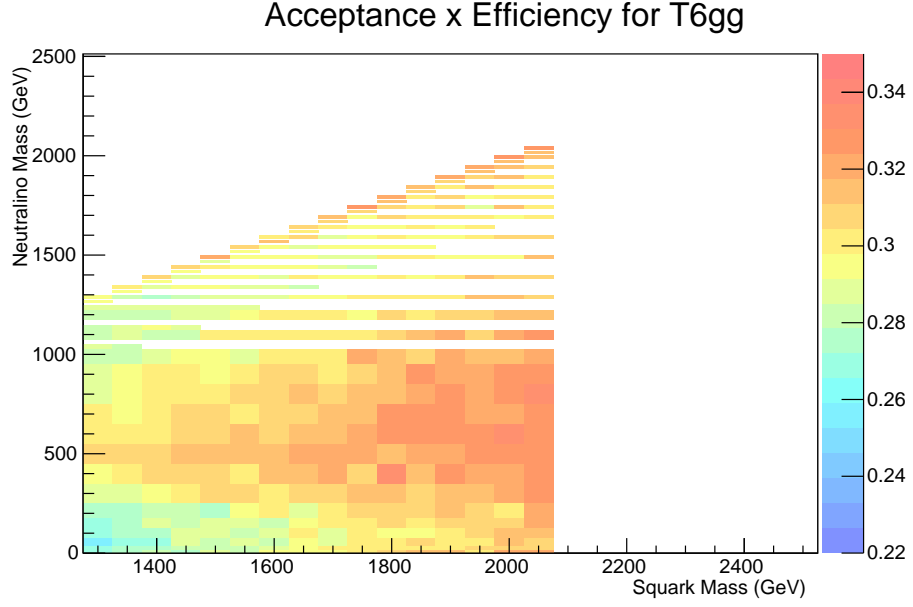


Figure 9.4: Acceptance  $\times$  efficiency for the T6gg simplified model.

In both the T5gg and T6gg simplified models, the cross section is independent of the mass of the neutralino and depends only on the mass of the gluino or squark, respectively. Therefore, it is not surprising that the exclusion contours shown in Figures 9.5 and 9.6 do not depend sensitively on the mass of the neutralino.

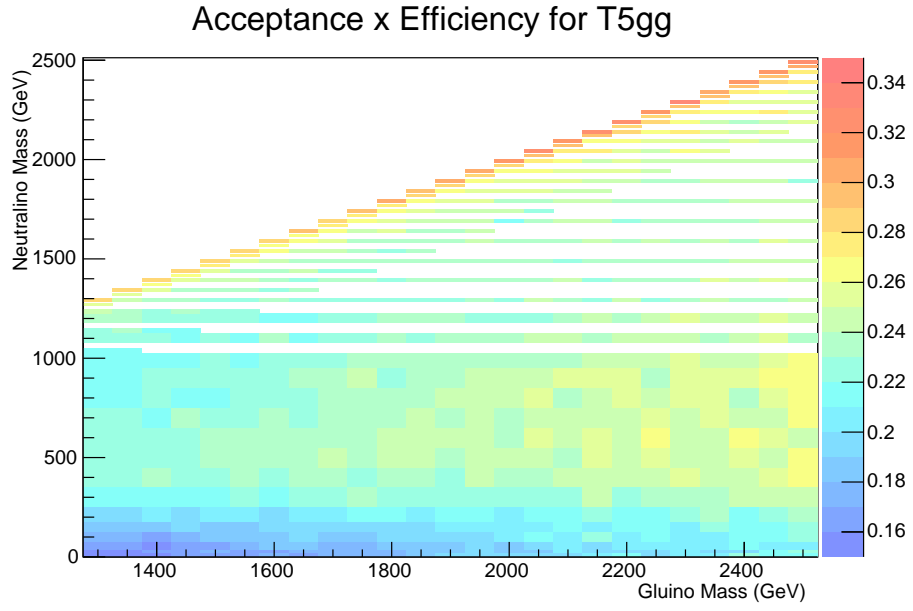


Figure 9.5: Acceptance  $\times$  efficiency for the T5gg simplified model.

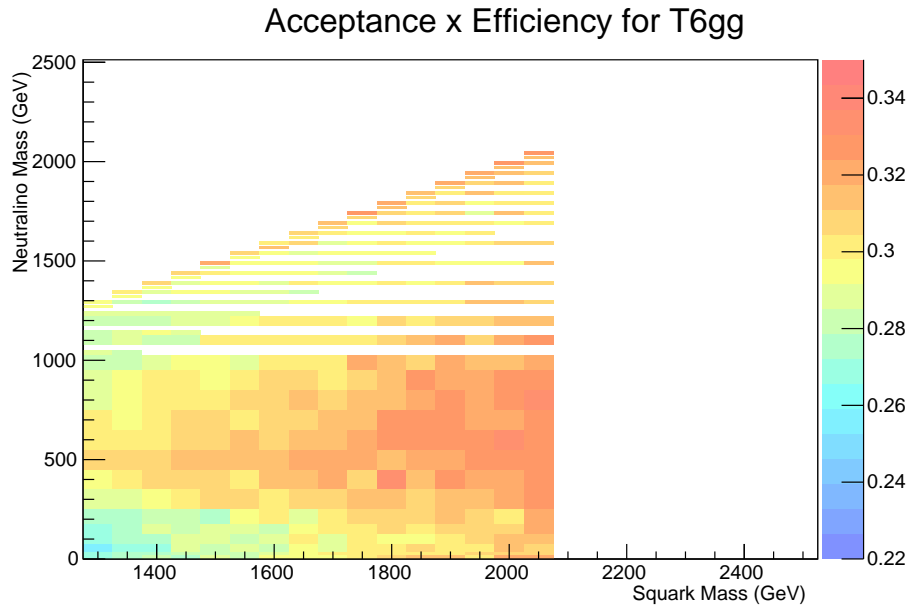


Figure 9.6: Acceptance  $\times$  efficiency for the T6gg simplified model.

## CHAPTER 10

### Conclusions

In this dissertation, a search for gauge-mediated supersymmetry breaking in events with two photons and missing transverse momentum was described. The analysis was performed with data taken with the CMS detector in 2016, but suggestions to improve the background estimation methods for the analysis of the 2017 data were also explored. Results were interpreted in the context of two simplified models, one assuming squark pair production and the other assuming gluino pair production. No evidence for gauge-mediated supersymmetry breaking (GMSB) was observed, and limits were placed on the masses of supersymmetric particles. Work is also ongoing to combine this analysis with other analyses targeting photon final states, in order to set limits on more comprehensive GMSB models.

## BIBLIOGRAPHY

1. S. Abdullin, P. Azzi, F. Beaudette, P. Janot, and A. Perrotta. The fast simulation of the CMS detector at LHC. *J. Phys. Conf. Ser.*, 331:032049, 2011. doi: 10.1088/1742-6596/331/3/032049.
2. J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014. doi: 10.1007/JHEP07(2014)079.
3. R. D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015. doi: 10.1007/JHEP04(2015)040.
4. W. Beenakker, S. Brensing, M. Krämer, A. Kulesza, E. Laenen, and I. Niessen. Soft-gluon resummation for squark and gluino hadroproduction. *JHEP*, 12:041, 2009. doi: 10.1088/1126-6708/2009/12/041.
5. C. Borschensky, M. Krämer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn, and X. Portell. Squark and gluino production cross sections in  $pp$  collisions at  $\sqrt{s} = 13, 14, 33$  and 100 TeV. *Eur. Phys. J. C*, 74:3174, 2014. doi: 10.1140/epjc/s10052-014-3174-y.
6. S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.
7. CMS. Twiki. URL <https://twiki.cern.ch/twiki/bin/view/CMS/L1TriggerPhaseIDiagrams>.
8. CMS Collaboration. The CMS magnet project: Technical Design Report. CMS Technical proposal CERN-LHCC-97-010, CMS-TDR-1, CERN, 1997. URL <https://cds.cern.ch/record/331056>.
9. CMS Collaboration. The CMS tracker system project: Technical Design Report. CMS Technical proposal CERN-LHCC-98-006, CMS-TDR-5, CERN, 1997. URL <https://cds.cern.ch/record/368412>.
10. CMS Collaboration. The CMS tracker: addendum to the Technical Design Report. CMS Technical proposal CERN-LHCC-2000-016, CMS-TDR-5-add-1, CERN, 2000. URL <https://cds.cern.ch/record/490194>.

11. C. Collaboration. Precise mapping of the magnetic field in the cms barrel yoke using cosmic rays. *Journal of Instrumentation*, 5(03):T03021, 2010. doi: 10.1088/1748-0221/5/03/T03021. URL <http://stacks.iop.org/1748-0221/5/i=03/a=T03021>.
12. T. Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A*, 434:435, 1999. doi: 10.1016/S0168-9002(99)00498-2.
13. A. Kulesza and L. Motyka. Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC. *Phys. Rev. D*, 80:095004, 2009. doi: 10.1103/PhysRevD.80.095004.
14. A. L. Read. Presentation of search results: the  $CL_s$  technique. *J. Phys. G*, 28:2693, 2002. doi: 10.1088/0954-3899/28/10/313.
15. T. Sjöstrand, S. Mrenna, and P. Z. Skands. A brief introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852, 2008. doi: 10.1016/j.cpc.2008.01.036.

*This document was prepared & typeset with pdfL<sup>A</sup>T<sub>E</sub>X, and formatted with  
NDdiss2<sub>ε</sub> classfile (v3.2013[2013/04/16]) provided by Sameer Vijay and updated  
by Megan Patnott.*