

This Dissertation
entitled
SEARCH FOR GAUGE-MEDIATED SUPERSYMMETRY BREAKING
typeset with $\text{NDdiss2}_{\mathcal{E}}$ v3.2013 (2013/04/16) on April 2, 2018 for
Allison Reinsvold Hall

This L^AT_EX 2 _{\mathcal{E}} 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 _{\mathcal{E}} 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 _{\mathcal{E}} 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

April 2018

This is the abstract.

CONTENTS

FIGURES	iv
TABLES	viii
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: THE STANDARD MODEL AND SUPERSYMMETRY	2
2.1 The standard model of particle physics	2
2.2 Supersymmetry	2
2.3 Gauge-mediated supersymmetry breaking	3
CHAPTER 3: THE LARGE HADRON COLLIDER	4
3.1 Introduction	4
3.2 Injector chain	5
3.3 LHC tunnel and magnets	5
3.4 Machine luminosity	7
CHAPTER 4: CMS DETECTOR	9
4.1 Coordinate system	9
4.2 Superconducting solenoid	10
4.3 Tracker	12
4.4 Electromagnetic Calorimeter (ECAL)	14
4.5 Hadron calorimeter (HCAL)	17
4.6 Calorimeter performance	19
4.7 Muon system	21
4.8 Instantaneous luminosity measurement	22
CHAPTER 5: TRIGGER SYSTEM	25
5.1 Overview of the CMS trigger	25
5.2 Analysis triggers	29
5.3 Trigger efficiency	32

CHAPTER 6: OBJECT RECONSTRUCTION AND SAMPLE SELECTION	41
6.1 Overview of the CMS reconstruction process	41
6.2 Particle flow algorithm	42
6.3 Photon identification	47
6.4 Jet reconstruction	47
6.5 Missing transverse momentum	51
6.6 Pileup subtraction	52
CHAPTER 7: Sample Selection	55
7.1 MC Samples	55
7.2 Object Selection	56
7.3 Photon ID scale factor	58
CHAPTER 8: DATA ANALYSIS AND BACKGROUND ESTIMATION METHODS	59
8.1 Overview	59
8.2 QCD background	60
8.3 Electroweak background	65
8.4 Irreducible background	70
8.5 Systematic uncertainties	71
CHAPTER 9: Results and Interpretations	77
9.1 Prediction versus observation	77
9.2 Simplified Models	77
9.3 Confidence level calculations	80
9.4 Limits and Interpretations	81
CHAPTER 10: Conclusions	84
BIBLIOGRAPHY	85

FIGURES

<p>3.1 Schematic of the CERN accelerator complex. The LHC is shown in dark grey and includes four interaction points: one each for the CMS, ATLAS, LHCb, and ALICE experiments. Prior to being injected into the LHC, the protons are accelerated through the Linear Accelerator 2 (Linac 2), the Proton Synchrotron Booster, the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS). Reprinted from Reference [8].</p> <p>3.2 Peak instantaneous luminosity per day achieved by the LHC during the 2016 data-taking period. The maximum luminosity was 15.3 Hz/nb, or $1.53 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$.</p> <p>4.1 Schematic of the CMS detector.</p> <p>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 [18].</p> <p>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 [9].</p> <p>4.4 Relative response of the ECAL crystals for the 2011 to 2016 data-taking periods. The average observed change in response is up to 10% in the barrel and 50% for $\eta > 2.5$. The bottom plot shows the instantaneous luminosity delivered during this period. The crystals recovered some, but not all, of their response during 2013-2014 Long Shutdown 1. Reprinted from Reference [10].</p> <p>4.5 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 [9].</p>	<p>6</p> <p>8</p> <p>10</p> <p>11</p> <p>13</p> <p>16</p> <p>18</p>
--	---

4.6	Relative electron energy resolution in bins of pseudorapidity. The energy resolution was calculated from an unbinned likelihood fit to $Z \rightarrow ee$ events. Barrel module boundaries are shown with vertical dashed lines and often correspond to regions where the resolution is somewhat degraded. The gray band at $ \eta = 1.5$ represents the boundary between EB and EE. The data correspond to 2.5 fb^{-1} collected with the CMS detector in 2015. Reprinted from Reference [11].	20
4.7	Integrated luminosity delivered to CMS by the LHC (blue) and the integrated luminosity recorded by CMS (orange) during p-p collisions at $\sqrt{s} = 13 \text{ TeV}$ in 2016.	24
5.1	Schematic showing the overall structure of the calorimeter half of the L1 trigger. Reprinted from Reference [12].	35
5.2	Schematic showing the overall structure of the muon half of the L1 trigger. Reprinted from Reference [12].	36
5.3	Efficiency of the L1 seed and the leading leg of the primary analysis trigger with respect to photon p_T . For $p_T > 40 \text{ GeV}$, the trigger is 98.2% efficient.	37
5.4	Efficiency of the trailing leg of the primary analysis trigger with respect to photon p_T . For $p_T > 40 \text{ GeV}$, the trigger is 99.8% efficient.	38
5.5	Efficiency of the primary analysis trigger with respect to the invariant mass of the diphoton system. For $m_{\gamma\gamma} > 100 \text{ GeV}$, the trigger is 99.4% efficient.	39
5.6	Efficiency of the trailing leg of the control trigger with respect to photon p_T . For $p_T > 40 \text{ GeV}$, 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.	40
6.1	Slice of the transverse plane of the CMS detector showing how different particles interact with the detector subsystems. The most important particle for this analysis is the photon (blue dotted line), which interacts only with the ECAL (shown in green). Reprinted from Reference [8].	42
6.2	Behavior of several jet-clustering algorithms illustrated with a sample parton-level event. CMS uses the anti- k_T algorithm (bottom right) with a distance parameter $R = 0.4$. Reprinted from Reference [7].	48
6.3	Ratio of the average reconstructed jet p_T to the particle-level jet p_T ($p_{T,\text{ptcl}}$) in QCD MC simulation. The ratio is calculated in bins of $p_{T,\text{ptcl}}$ before applying any JEC (top left), after applying the L1 pileup corrections (top right), and after all JEC (bottom). The different curves correspond to different numbers of pileup interactions μ . Reprinted from Reference [22].	50
6.4	Jet energy resolution (JER) in 8 TeV simulation as a function of particle-level jet p_T ($p_{T,\text{ptcl}}$). PF+CHS refers to particle-flow jets with charged hadron subtraction (see Section 6.6). Results are shown for various levels of pileup μ . Reprinted from Reference [22].	51

6.5	Energy composition as a function of the number of pileup (PU) interactions μ for jets in data (markers) and simulation (histogram). The data distributions correspond to 19.7 fb^{-1} collected at $\sqrt{s} = 8 \text{ TeV}$ with the CMS detector in 2012. The contribution from charged hadrons from pileup is shown in dark red. The bottom pane shows the agreement between data and MC for each category. Reprinted from Reference [25].	53
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.	61
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 $\gamma\gamma/f\bar{f}$ ratios that are used to reweight the $f\bar{f} E_T^{\text{miss}}$ distribution.	62
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 $\gamma\gamma/f\bar{f}$ ratios that are used to reweight the $f\bar{f} E_T^{\text{miss}}$ distribution.	63
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 $\gamma\gamma/f\bar{f}$ ratios that are used to reweight the $f\bar{f} E_T^{\text{miss}}$ distribution.	64
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 $\gamma\gamma/f\bar{f}$ ratios that are used to reweight the $f\bar{f} E_T^{\text{miss}}$ distribution.	65
8.6	The ratio of $\gamma\gamma$ to $f\bar{f}$ events for $E_T^{\text{miss}} < 100 \text{ GeV}$. The ratio has been fit to a constant function, which can then be used to extrapolate to the $E_T^{\text{miss}} > 100 \text{ GeV}$ signal region. This provides an alternative estimation method for the QCD background.	66
8.7	Data versus MC comparison of the $e\gamma$ control sample E_T^{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.	68
8.8	Invariant mass distributions for the $e\gamma$ (left) and ee (right) samples. Each plot includes the data (black markers), the signal template (red), and the background template (green). The top plots demonstrate the shape of each distribution with arbitrary normalization, and the bottom plots show the distributions after the fits are performed.	74
8.9	Value of $R_{\gamma/e}$ as a function of various kinematic variables for probes in the barrel: probe p_T , E_T^{miss} , isolation I_\pm , and track multiplicity of the primary vertex (PV). The red band corresponds to a 30% uncertainty on $R_{\gamma/e}$	75

8.10 Value of $R_{\gamma/e}$ as a function of various kinematic variables for probes in the barrel: $\sigma_{in\eta}$ of the probe, vertex multiplicity, jet multiplicity, and probe $ \eta $. The red band corresponds to a 30% uncertainty on $R_{\gamma/e}$.	76
9.1 E_T^{miss} distributions for the full background estimation and the observed data. The black points represent the observed E_T^{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.	79
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 γ and a gravitino \tilde{G} . In the diagram on the right, we do not distinguish between squarks and antisquarks.	80
9.3 Acceptance \times efficiency for the T5gg simplified model.	81
9.4 Acceptance \times efficiency for the T6gg simplified model.	82
9.5 Acceptance \times efficiency for the T5gg simplified model.	83
9.6 Acceptance \times efficiency for the T6gg simplified model.	83

TABLES

5.1	HLT TRIGGER PATHS	30
5.2	PRIMARY TRIGGER REQUIREMENTS	31
7.1	MC SAMPLES FOR SIGNAL AND BACKGROUND	55
7.2	ELECTROMAGNETIC OBJECT DEFINITIONS	57
7.3	EFFECTIVE AREAS FOR ISOLATION CORRECTIONS	58
8.1	COMPARISON BETWEEN REWEIGHTING METHOD AND RATIO METHOD FOR THE QCD BACKGROUND ESTIMATE	67
9.1	EXPECTED AND OBSERVED EVENTS IN THE SIGNAL REGION	78

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

$$SU(3)_c \oplus SU(2)_L \oplus SU(1)_Y \quad (2.1)$$

where $SU(3)_c$ represents the QCD

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 (LHC) is a particle accelerator at CERN, the European Organization for Nuclear Research. It is the world's largest accelerator and hosts several experiments, including the all-purpose Compact Muon Solenoid (CMS) and A Toroidal LHC ApparatuS (ATLAS) experiments, the b -physics experiment LHC beauty (LHCb), and the heavy-ion experiment A Large Ion Collider Experiment (ALICE). The LHC is located 40 to 170 m underground in the countryside outside Geneva, Switzerland on the Switzerland-France border. A full description of the collider can be found in Reference [19].

The LHC consists of two counter-rotating beams of protons that collide at the four interaction points (IP) of the experiments listed above. It can also be used for heavy-ion collisions, in particular ^{208}Pb -p or ^{208}Pb - ^{208}Pb collisions. In Run I of the LHC (2010 - 2012), a maximum center-of-mass energy $\sqrt{s} = 8$ TeV was achieved. After a two-year shutdown, the LHC restarted in 2015 at a center-of-mass energy $\sqrt{s} = 13$ TeV, close to its design value of $\sqrt{s} = 14$ TeV. Run II extends to the end of 2018, when another long shutdown will allow upgrades for the High-Luminosity LHC era to begin.

3.2 Injector chain

A series of smaller accelerators are required to bring protons from rest to the collision energy of 6.5 TeV. The full CERN accelerator complex is shown in Figure 3.1. The journey of a proton in the LHC begins in a small tank of hydrogen. The hydrogen gets ionized and is accelerated to 50 MeV in the Linear Accelerator 2 (Linac 2). Next it enters the Proton Synchrotron Booster and the Proton Synchrotron (PS), which increase the energy to 1.4 GeV and 25 GeV, respectively. The next step is the Super Proton Synchrotron (SPS), which is nearly 7 km in circumference and accelerates the protons to 450 GeV before injecting them into the LHC.

3.3 LHC tunnel and magnets

Due to the prohibitive cost of building a new tunnel, the LHC re-uses the same tunnel as the one previously occupied by the Large Electron-Positron (LEP) Collider. This placed several constraints on the design of the accelerator. Because of the large synchrotron radiation losses that occur when accelerating light particles such as electrons, the tunnel includes eight straight sections where the electrons and positrons could be accelerated. For a hadron collider such as the LHC, however, the losses due to synchrotron radiation are less than 7 keV per rotation. The straight sections of the tunnel mean that the arced sections have a smaller radius of curvature R than if the tunnel was perfectly circular. This directly affects the maximum attainable momentum p of a particle with charge q for a given magnetic field B :

$$B = \frac{p}{qcR} \tag{3.1}$$

To reach the design energy $\sqrt{s} = 14$ TeV, the LHC dipoles must be able to operate at a magnetic field $B = 8.33$ TeV. This requires the magnets to be superconducting. The LHC includes 1232 dipole magnets that are cooled to 1.9 K using liquid helium.

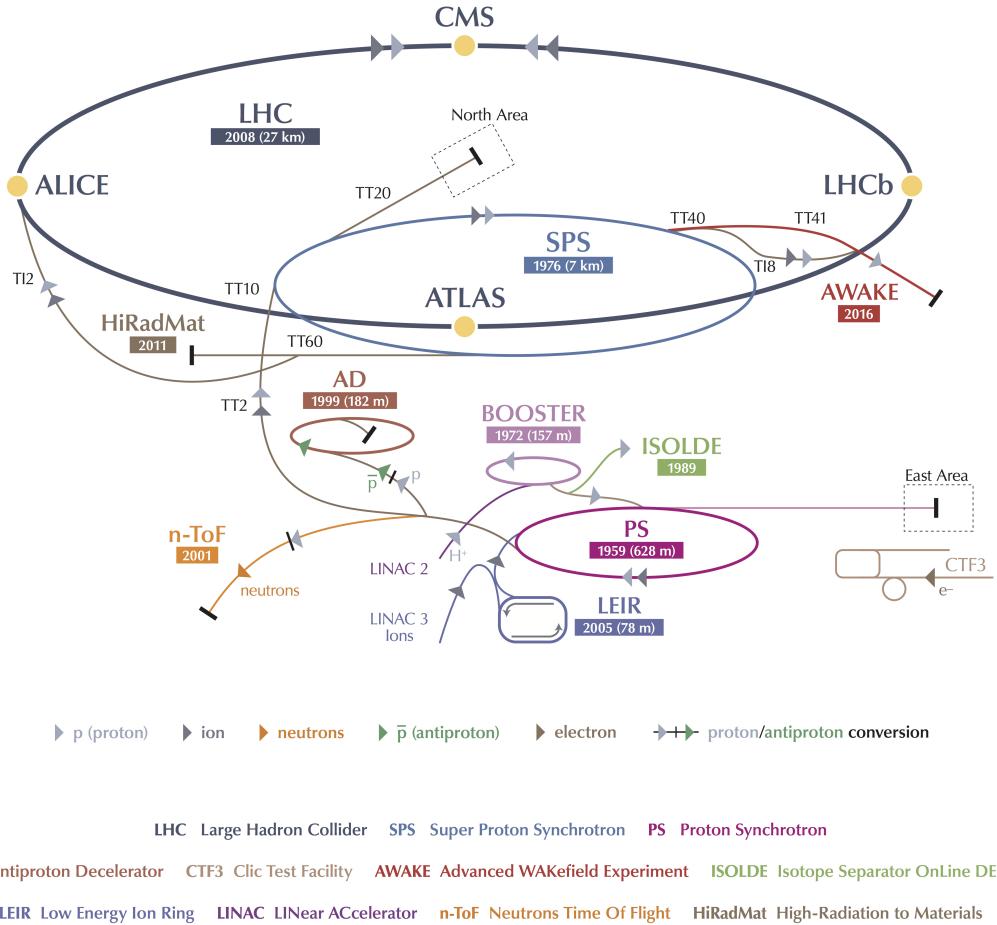


Figure 3.1: Schematic of the CERN accelerator complex. The LHC is shown in dark grey and includes four interaction points: one each for the CMS, ATLAS, LHCb, and ALICE experiments. Prior to being injected into the LHC, the protons are accelerated through the Linear Accelerator 2 (Linac 2), the Proton Synchrotron Booster, the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS). Reprinted from Reference [8].

The solenoid of each magnet is made of stacked niobium-titanium (NbTi) filaments of thickness $6\text{-}7 \mu\text{m}$. The LEP tunnel is too narrow (3.7 m radius) to fit two magnet systems and the associated cryogenics and insulation. Therefore, the LHC magnets use a “twin-bore” design, where both of the beams are embedded in a single magnet system.

The LHC has only one accelerating sector. In the sector, there are eight super-

conducting RF cavities per beam that operate at a frequency of 400 MHz. The RF cavities add 485 keV per rotation. The beams consist of “bunches” of protons with a spacing of 25 ns separating consecutive bunches. Each collision between two bunches is referred to as a bunch crossing.

3.4 Machine luminosity

The number of events N_{event} produced every second for a particular process with cross section σ_{event} is given by

$$N_{\text{event}} = L\sigma_{\text{event}} \quad (3.2)$$

where L is the instantaneous luminosity of the collider.

The instantaneous luminosity depends only on machine parameters:

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta_*} F \quad (3.3)$$

The variables in the above equation include the number of particles per bunch N_b , the number of bunches per beam n_b , the revolution frequency f_{rev} , and the relativistic factor γ_r . For the LHC, n_b is 2808 for the nominal bunch spacing of 25 ns. The number of protons per bunch N_b is limited to 1.15×10^{11} by the nonlinear beam-beam interactions between protons at each IP.

The normalized transverse beam emittance ϵ_n is a measure of the average spread of particles in the beam in position-momentum phase space. In the transverse direction the beam is assumed to have a Gaussian shape with width σ . The beta function is then given by $\beta = \sigma^2/\epsilon$, and β_* refers to the value of the beta function at the interaction point. Finally, F is a geometric factor that depends on the crossing angle at the IP.

The design instantaneous luminosity of the LHC is $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. This goal is only attainable because the LHC is a p-p collider, rather than a p- \bar{p} collider. During the 2016 data-taking period, the LHC met and surpassed its luminosity target, achieving a maximum value of $1.53 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. This is almost double the maximum value of $7.7 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ that was achieved during Run I. The peak luminosity per day in 2016 is shown in Figure 3.2. The measurement of instantaneous luminosity is described in more detail in Section 4.8.

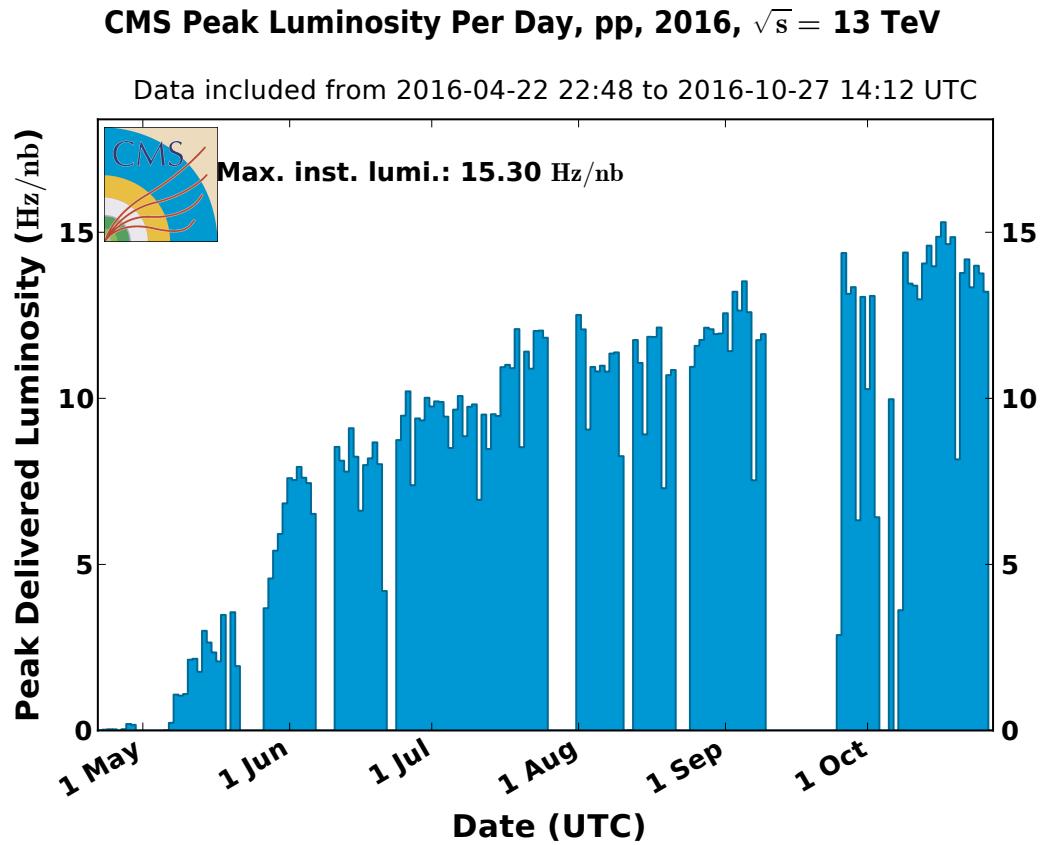


Figure 3.2: Peak instantaneous luminosity per day achieved by the LHC during the 2016 data-taking period. The maximum luminosity was 15.3 Hz/nb , or $1.53 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$.

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 4.1 shows a schematic of the detector as a whole. The CMS detector is 21.6 m long, 14.6 m in diameter, and weighs 12500 t. 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 4.7). For a full description of the CMS detector, see Reference [9].

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 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

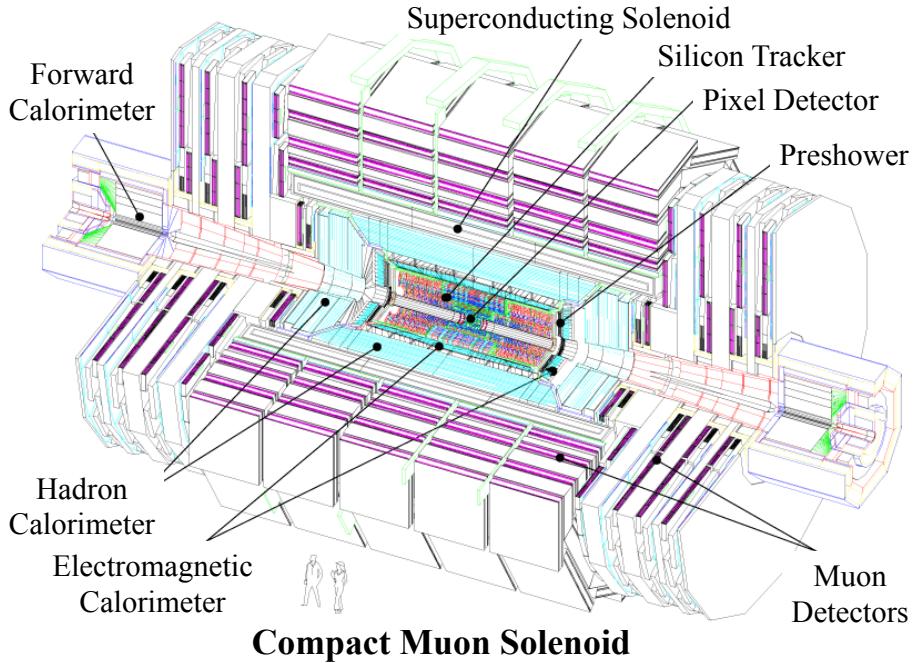


Figure 4.1. Schematic of the CMS detector.

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 ϕ 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 θ 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 4.5 K with liquid helium. A detailed description of the CMS magnet can be found in Reference [15]. Figure 4.2 shows the calculated magnetic field in a longitudinal slice of the CMS detector.

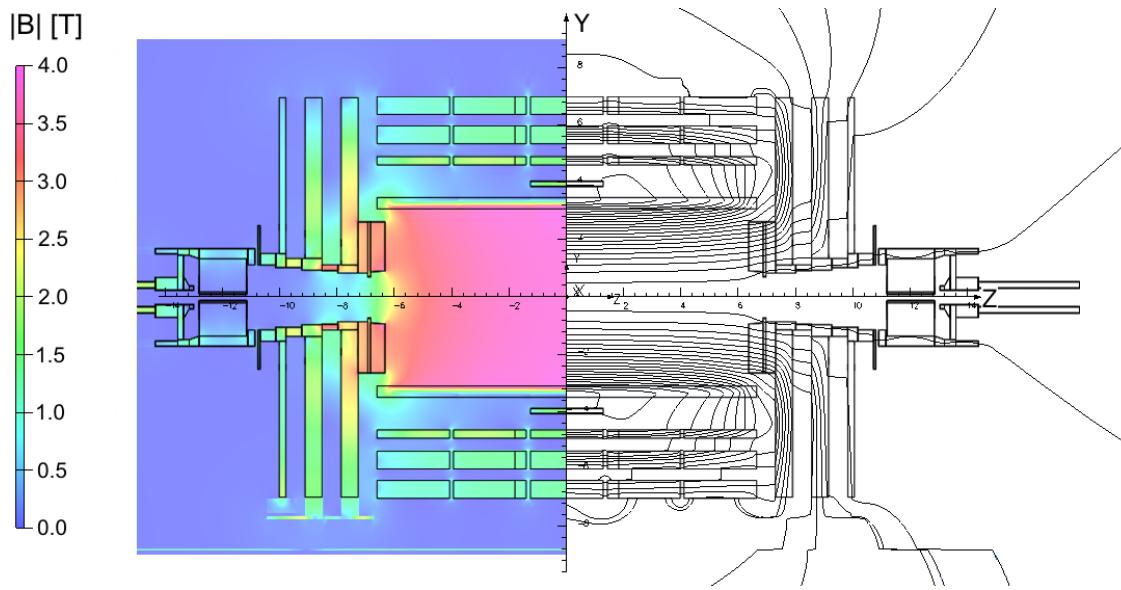


Figure 4.2: Calculated magnetic field $|\vec{B}|$ for a longitudinal slice of the CMS detector when operated at a central magnetic flux density of 3.8 T. In the right-hand portion of the figure, each magnetic field line represents a magnetic flux step of 6 Wb. Reprinted from Reference [18].

4.3 Tracker

The innermost sub-detector is the silicon tracker [16, 17]. 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 crystal. Charged particles passing through the depletion region will create electron-hole pairs that move towards either end of the junction when an external electric field is applied. 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 μm 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 detectors

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

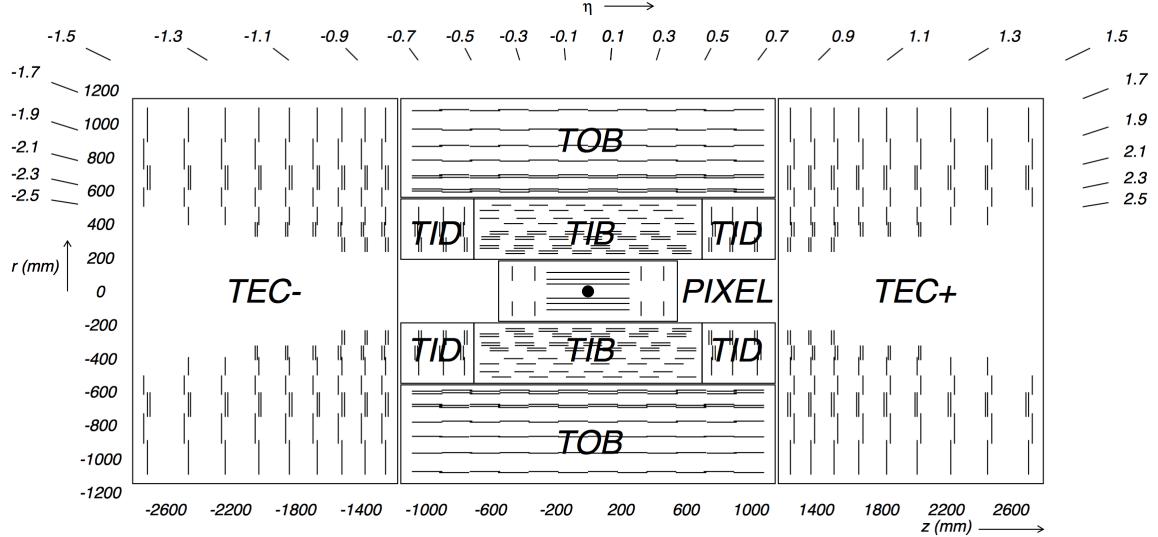


Figure 4.3: Schematic of the CMS tracker system. Detector modules are represented by single lines, and back-to-back modules are represented by double lines. The strip tracker is further divided into the Tracker Inner/Outer Barrel (TIB/TOB), Tracker Inner Disk (TID) and the Tracker Endcap (TEC). Reprinted from Reference [9].

point. The small pixel size of $150 \times 100 \mu\text{m}$ is critical for accurate secondary vertex reconstruction, for forming seed tracks used for high level triggering (see Chapter 5), and for the reconstruction of charged particles in the event (see Chapter 6). 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 $10 \text{ cm} \times 80 \mu\text{m}$, 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 μm . Strips in the endcaps are aligned radially with a

pitch between 100 and 184 μm . 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 μm 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 μm can be achieved. These back-to-back modules are represented by double lines in Figure 4.3.

4.4 Electromagnetic Calorimeter (ECAL)

Beyond the tracker is the electromagnetic calorimeter (ECAL), the most important sub-detector for this analysis [1, 13]. The ECAL measures the energies deposited by photons and electrons when they are stopped by the detector. For energies above 10 MeV, electrons lose energy primarily through the production of photons in bremsstrahlung. Photons lose energy through the production of e^+e^- pairs [20]. Therefore, when an energetic electron or photon is incident on one of the crystals of the ECAL, the result is an electromagnetic cascade (“shower”) of photons and electrons with successively lower energies.

The CMS ECAL is designed to contain the full electromagnetic shower from electrons and photons with initial energies as high as a few TeV. It is a homogeneous calorimeter made with 75,848 scintillating lead tungstate PbWO_4 crystals. The ECAL is divided into a barrel region (EB) covering $|\eta| < 1.479$ and two endcap regions (EE) covering $1.479 < |\eta| < 3.0$. Each region includes a single layer of PbWO_4 crystals. In the barrel, the crystals have a truncated pyramidal shape with a radial depth of 23 cm. The front face of the crystal has dimensions $22 \times 22 \text{ mm}^2$, and the rear face has dimensions $26 \times 26 \text{ mm}^2$. In the endcap, the crystals are 22 cm long and have a cross-section of $28.62 \times 28.62 \text{ mm}^2$ on the front face and $30.0 \times 30.0 \text{ mm}^2$ on the rear face.

Lead tungstate is an inorganic scintillator. The passage of charged particles produce electron-hole pairs in the conduction and valence bands of the material, and light is emitted when the electrons return to the valence band. There are several properties of PbWO_4 that make it an ideal choice for the scintillating material of the CMS ECAL. Its Molière radius, defined as the radius of a cylinder containing 90% of the shower's energy deposition, is only 2.2 cm, allowing for excellent position resolution and separation between showers. The latter is particularly important when trying to distinguish photons from isolated $\pi^0 \rightarrow \gamma\gamma$ decays. The fast response time of lead tungstate is also critical. Approximately 80% of the light is emitted within 15 ns [20], making it possible for each particle to be assigned to the correct bunch crossing.

Another positive feature of PbWO_4 is that it is dense enough (8.3 g/cm^3) that the ECAL can fit into a relatively compact area. The depth of material that is needed is quantified by looking at the radiation length of the material (χ_0). The radiation length is the average distance an electron needs to travel to reduce its energy by a factor $1/e$. The radiation length of PbWO_4 is 0.89 cm, which means $24.7 \chi_0$ fit in the 22 cm radial distance of each EB crystal.

Finally, the last characteristic of PbWO_4 that put it ahead of other inorganic scintillators is its radiation hardness. Nuclear reactions caused by prolonged exposure to severe radiation conditions can cause defects in the crystals of inorganic scintillators. This affects the transparency of the crystals, leading to a degradation in the crystal response. This is shown in Figure 4.4 for the 2011, 2012, 2015, and 2016 data-taking periods. To correct for this effect, each crystal is illuminated with laser light. Part of the light is redirected to a silicon photodiode off the detector for a reference measurement. This correction is important both for long-term damage as illustrated in Figure 4.4, but also for the fast component of the radiation damage that changes the crystal response over the course of a single LHC fill.

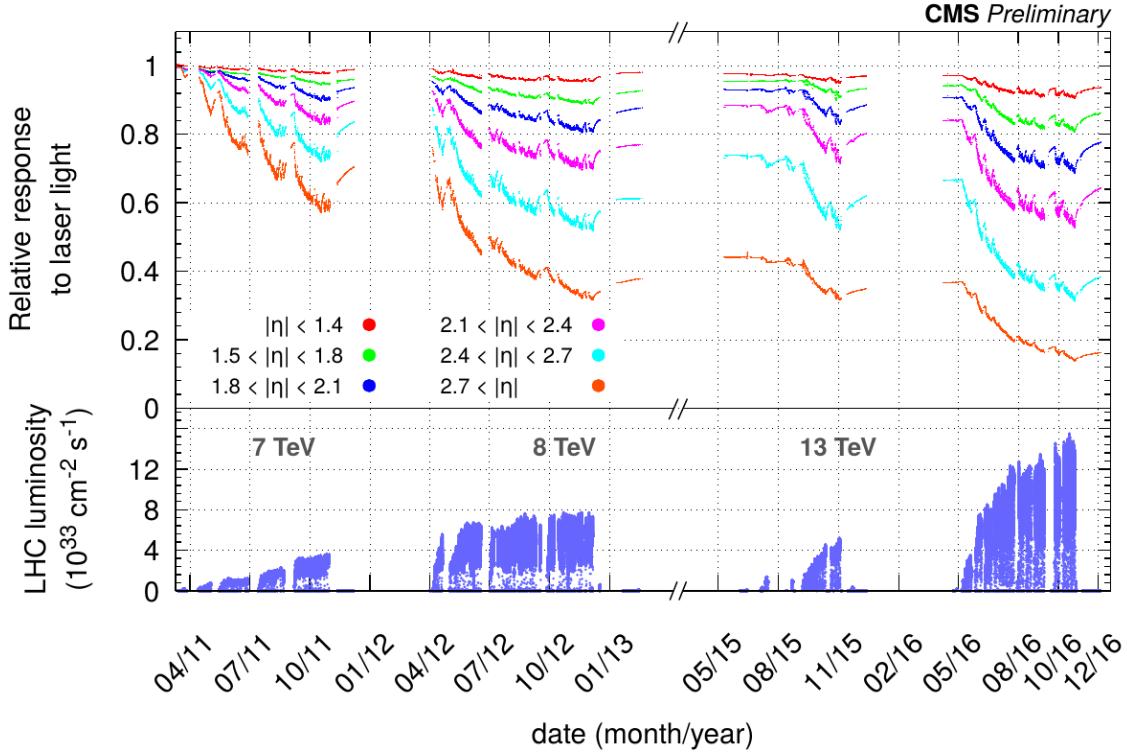


Figure 4.4: Relative response of the ECAL crystals for the 2011 to 2016 data-taking periods. The average observed change in response is up to 10% in the barrel and 50% for $|\eta| > 2.5$. The bottom plot shows the instantaneous luminosity delivered during this period. The crystals recovered some, but not all, of their response during 2013-2014 Long Shutdown 1. Reprinted from Reference [10].

The downside to PbWO_4 is that it has a relatively low light yield of only 2 photoelectrons per MeV of incident energy. This makes the electronics used to collect the signal especially important. The crystal itself acts as an optical waveguide, and the scintillation light is internally reflected until it reaches the photodetectors glued directly onto the rear face of the crystal. Avalanche photodiodes (APD) are used in the barrel, and vacuum phototriodes (VPT) are used in the more radioactive endcap region. Both of these were chosen to be able to operate within the 3.8 T magnetic field, and both increase the gain by a factor of approximately 1000. The signal then goes to a Multi-Gain Preamplifier (MGPA), which dynamically changes the gain based on the energy of the incident particle. Temperature control of the PbWO_4

crystals and the attached APDs is crucial to maintaining an excellent resolution. The crystal response and APD gain change by approximately 2.2% and 2.4% per $^{\circ}\text{C}$, respectively.

The last sections of the ECAL system are two pre-shower detectors located in front of the EE. The pre-shower detector is designed to improve the spatial resolution of the ECAL, so that the two photons from $\pi^0 \rightarrow \gamma\gamma$ decays can be distinguished. Each pre-shower detector consists of alternating strips of lead and silicon detectors. The lead forces the photons to interact, and the silicon detector measures the electrons and positrons produced in the shower with high granularity. The pre-shower detector achieves a spatial resolution of 2 mm, compared to the 3 cm resolution of the EB and EE.

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 4.5 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 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.

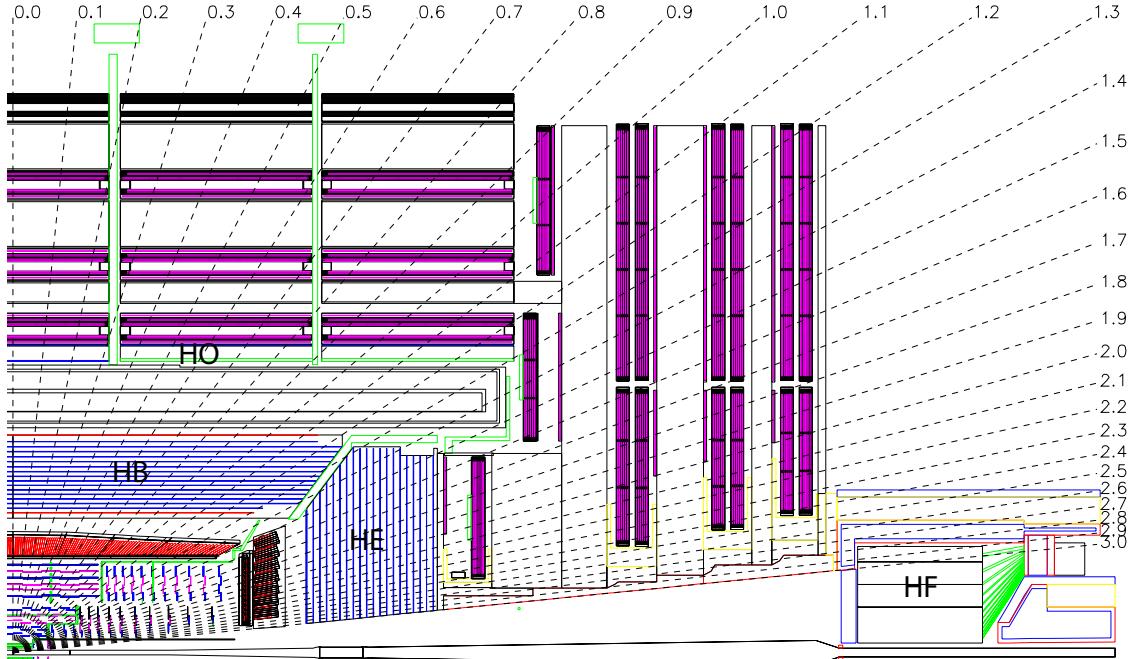


Figure 4.5: Longitudinal slice of the CMS HCAL, including the barrel HCAL (HB), outer HCAL (HO), endcap HCAL (HE), and forward HCAL (HF). The dashed lines represent lines of constant pseudorapidity. Reprinted from Reference [9].

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 the HE. For $|\eta| < 1.6$, the 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 ^{136}Cs or ^{60}Co 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 radial depth of the ECAL corresponds to 1.1 nuclear interaction lengths, and in the HCAL there are 5.82 interaction lengths at $|\eta| = 0$ and 10.6 interaction lengths at $|\eta| = 1.3$.

In practice, the HCAL does not detect individual quarks or gluons or even individual hadrons, but objects known as jets. A jet is a spray of particles in a narrow cone that is produced by the hadronization of quarks and gluons. The reconstruction of jets is described in detail in Chapter 6.

4.6 Calorimeter performance

In general, the energy resolution σ/E of a calorimeter can be modeled with the following equation where \oplus is used to represent summing in quadrature [20]:

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (4.1)$$

The first term is a stochastic term that takes into account random fluctuations in the amount of deposited energy for a particle with incident energy E . Homogeneous detectors like the ECAL have an excellent intrinsic resolution and a very small stochastic term. Sampling calorimeters such as the HCAL, on the other hand, have larger stochastic terms because the number of charged particles that hit the active layers varies from shower to shower.

The b/E term represents electronic noise from the equipment that is used to collect the signal. This term can depend sensitively on the temperature of the elec-

tronics, as already mentioned for the EB APDs. Maximizing the energy yield helps improve this contribution to the energy resolution. The final term is a constant and covers any irregularities in the detector response. This can include variations between ECAL crystals, radiation damage effects, or differences in response due to temperature gradients across the detector.

The resolution of the ECAL as measured using $Z \rightarrow ee$ events in the 2015 CMS dataset is shown in Figure 4.6. For the central region of EB, the electron energy resolution is better than $2\%/\sqrt{E}$.

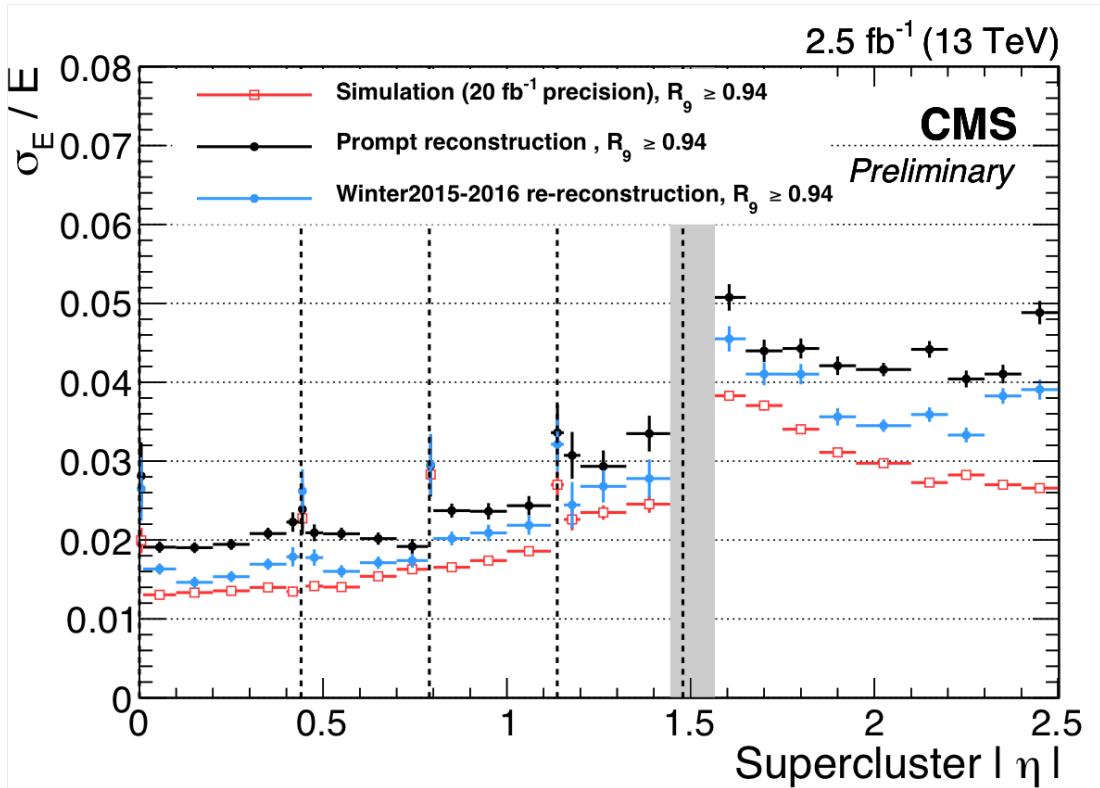


Figure 4.6: Relative electron energy resolution in bins of pseudorapidity. The energy resolution was calculated from an unbinned likelihood fit to $Z \rightarrow ee$ events. Barrel module boundaries are shown with vertical dashed lines and often correspond to regions where the resolution is somewhat degraded. The gray band at $|\eta| = 1.5$ represents the boundary between EB and EE. The data correspond to 2.5 fb^{-1} collected with the CMS detector in 2015. Reprinted from Reference [11].

The combined energy resolution of the ECAL and HCAL can be expressed as the following [25]:

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (4.2)$$

This resolution was measured using a pion test beam. The energy resolution after applying the full event reconstruction will be described in more detail in Chapter 6.

4.7 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% CO₂. 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 μm . Each drift tube has dimensions 1.3 cm \times 4.2 cm \times 2.4 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 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 positive ions produce a signal in the cathode wires. This provides two position coordinates for each muon. The spatial resolution in the azimuthal direction is $150\ \mu\text{m}$. 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 ns.

4.8 Instantaneous luminosity measurement

Five different detectors (luminometers) are used to measure the instantaneous luminosity L delivered by the LHC [14]: the pixel detector, the barrel drift tubes (DT), the forward hadron calorimeter (HF), and two specialized instruments, the Fast Beam Conditions Monitor (BCM1f) and the Pixel Luminosity Telescope (PLT). The last three use a fast readout system that is separate from the rest of the CMS system, and the pixel detector and DT use the standard data acquisition system.

Van der Meer (VdM) scans [27] are used to set an absolute calibration of each detector. Dedicated LHC running conditions allow the two beams to be scanned step-wise through one another in the two transverse directions. By finding the optimum beam position in the horizontal and vertical planes, the size of the beams at the

collision point can be determined. This in turn is used to calculate the visible cross sections σ_{vis} for each detector. The instantaneous luminosity is then given by the following, where R is the measured rate for a given luminometer:

$$L = \frac{R}{\sigma_{vis}} \quad (4.3)$$

The integrated luminosity \mathcal{L} is the integral of all collisions taking place within CMS and represents the total amount of data collected. Figure 4.7 shows the integrated luminosity delivered to CMS by the LHC and the integrated luminosity recorded by CMS during the 2016 data-taking period for 13 TeV p-p collisions. CMS collected 92.5% of the 40.76 fb^{-1} delivered by the LHC. CMS can fail to record events delivered by the LHC if any sub-detector is off or misbehaving. The 2016 “golden JSON”, the full list of runs and events that are certified for use in data analysis, corresponds to 35.92 fb^{-1} . The total uncertainty on the integrated luminosity is 2.5% [14]. The LHC is on track to deliver 100 fb^{-1} over the course of Run II.

CMS Integrated Luminosity, pp, 2016, $\sqrt{s} = 13$ TeV

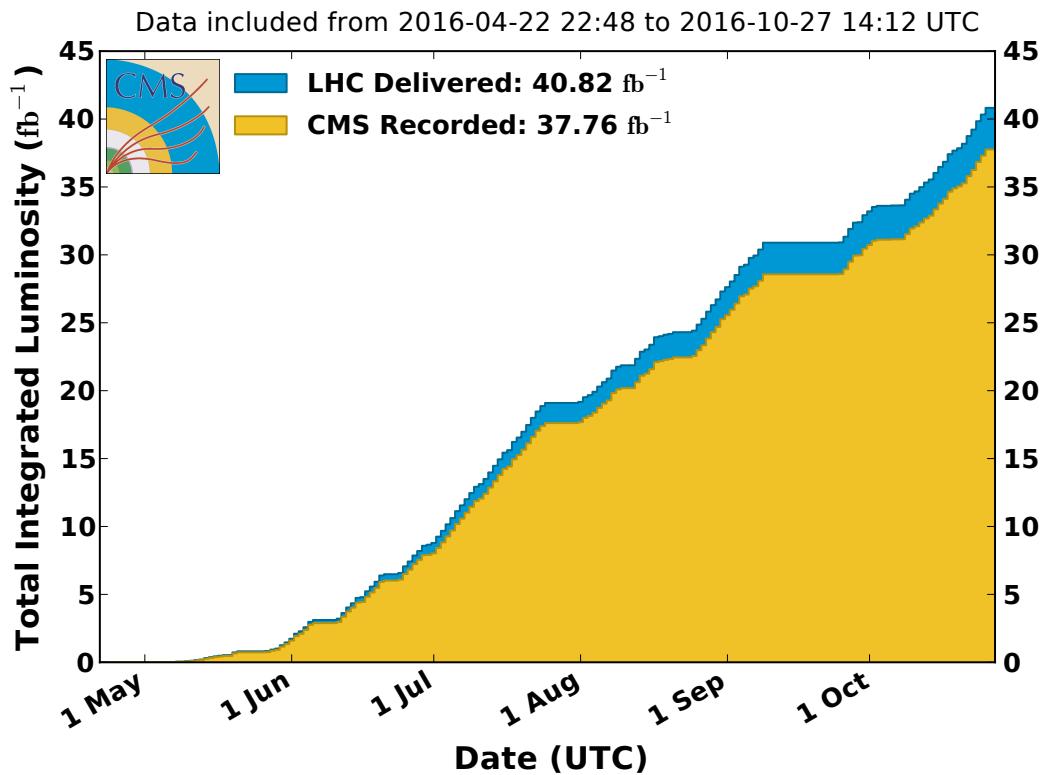


Figure 4.7: Integrated luminosity delivered to CMS by the LHC (blue) and the integrated luminosity recorded by CMS (orange) during p-p collisions at $\sqrt{s} = 13$ TeV in 2016.

CHAPTER 5

TRIGGER SYSTEM

5.1 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.

5.1.1 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 μ 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 5.1, 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 5.2. 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.

5.1.1.1 Calorimeter Trigger

As shown in Figure 5.1, 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 0.087×0.087 in η and ϕ , 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 τ -tagged jet candidates using several dynamic clustering and local maxima finding algorithms. The outputs from the Layer-2 nodes are fed into a de-multiplexing (demux) node, which ranks the candidates by transverse momentum

and sends the data to the Global Trigger.

5.1.1.2 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° 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.

5.1.2 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.

5.1.3 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/γ (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.

5.2 Analysis triggers

The HLT paths used in this analysis are listed in Table 5.1. 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.

5.2.1 Trigger Requirements

The requirements to pass the various parts of the trigger are listed in Table 5.2. 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 5.1

HLT TRIGGER PATHS

Primary Trigger
HLT_Diphoton30_18_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass90_v*
Control Sample Trigger
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^{-1} 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 5.2
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_-	$\sigma_{inj\eta} < 0.015$ 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 unprescaled 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.

5.3 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 7.2). 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 ϵ 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.

$$\epsilon_{trig} = N_{pass}/N_{total} \quad (5.1)$$

5.3.1 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 5.3. 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 5.4 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 5.5. 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:

$$\epsilon_{trig} = \epsilon_{lead} \times \epsilon_{trail}^2 \times \epsilon_{mass} = 97.2\% \quad (5.2)$$

5.3.2 Efficiency of double electron trigger

In the secondary trigger listed in Table 5.1, 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 5.6, 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 5.3. This results in an overall trigger efficiency of 79.8% for the control trigger.

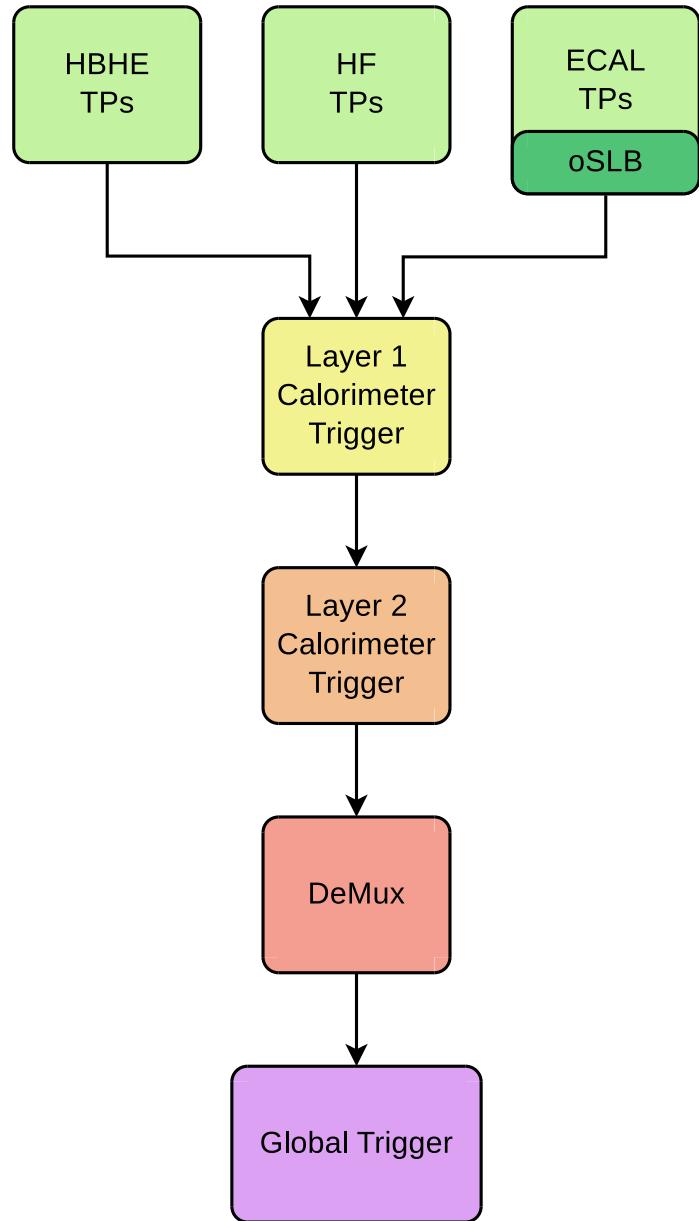


Figure 5.1: Schematic showing the overall structure of the calorimeter half of the L1 trigger. Reprinted from Reference [12].

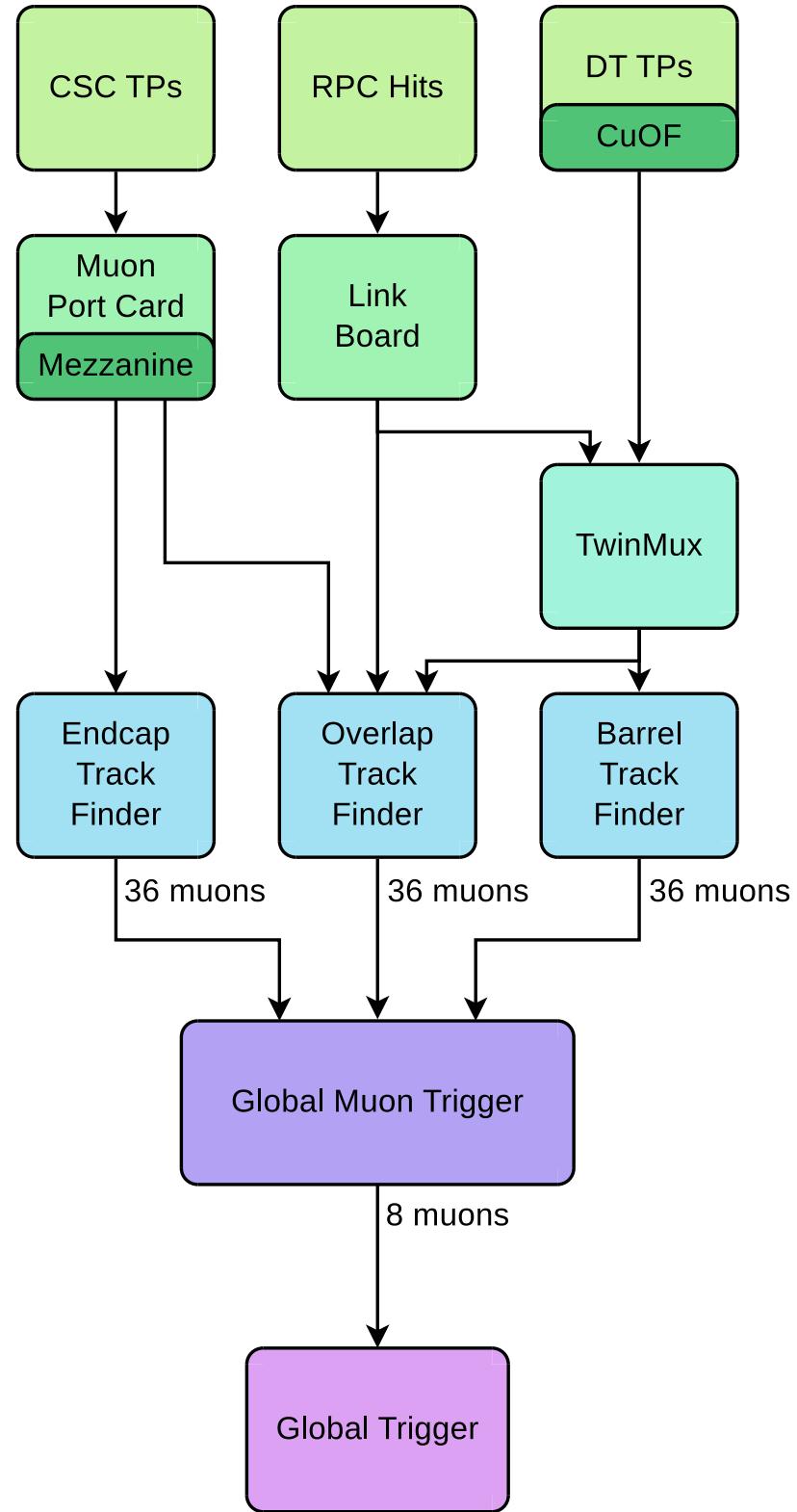


Figure 5.2: Schematic showing the overall structure of the muon half of the L1 trigger.
Reprinted from Reference [12].

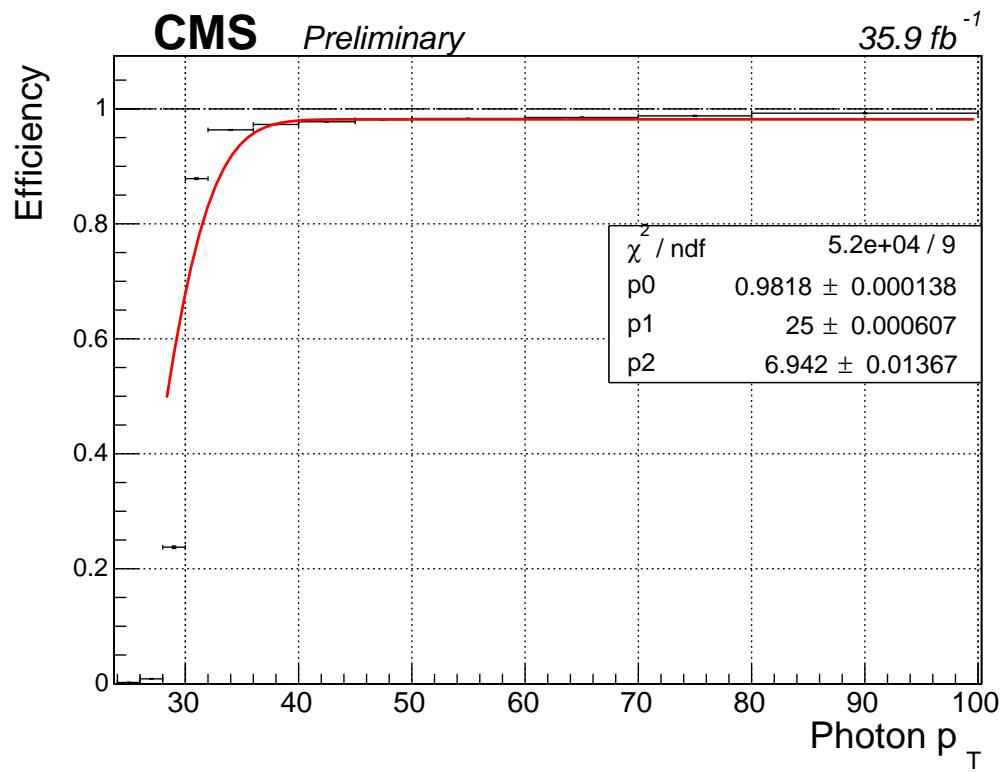


Figure 5.3: Efficiency of the L1 seed and the leading leg of the primary analysis trigger with respect to photon p_T . For $p_T > 40$ GeV, the trigger is 98.2% efficient.

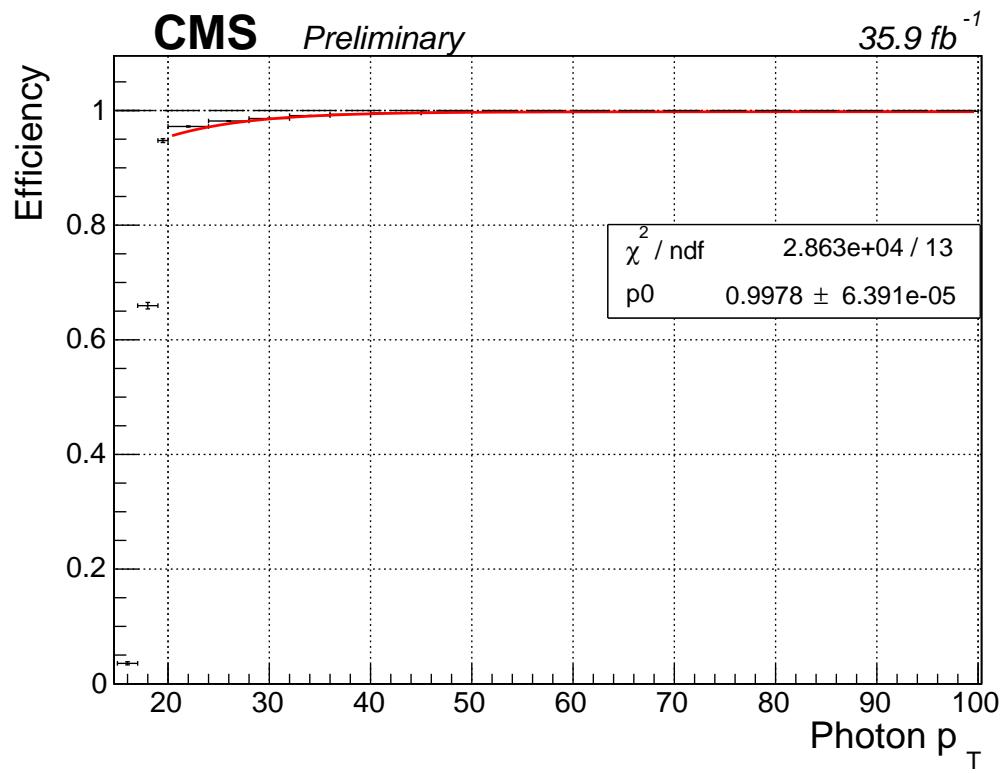


Figure 5.4: Efficiency of the trailing leg of the primary analysis trigger with respect to photon p_T . For $p_T > 40$ GeV, the trigger is 99.8% efficient.

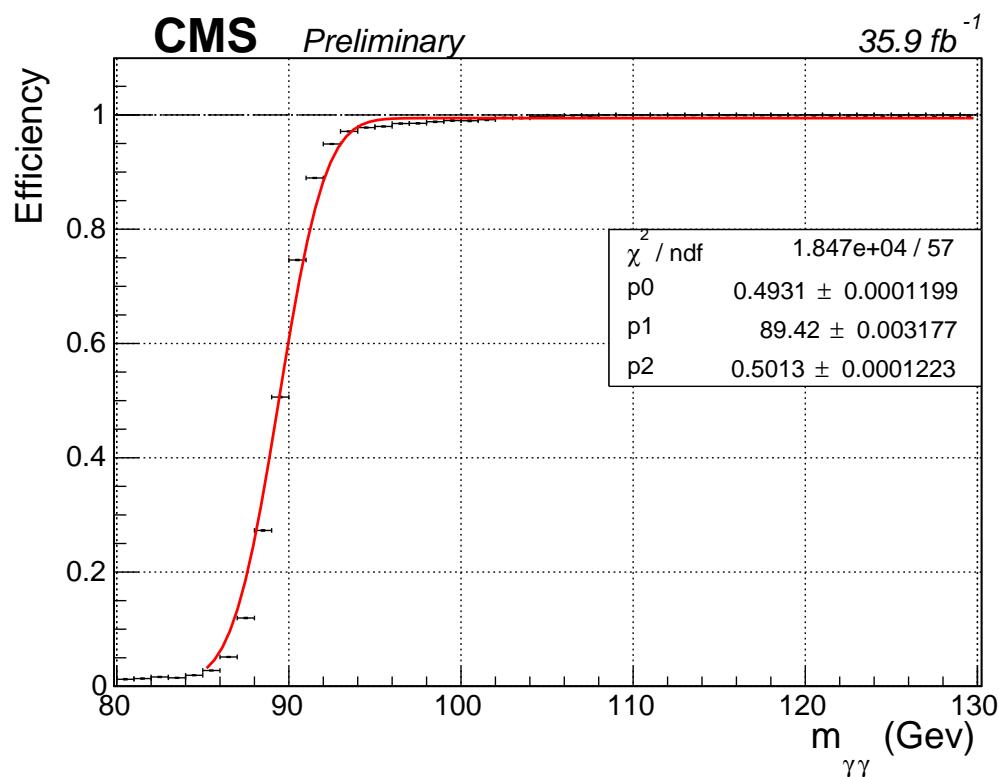


Figure 5.5: Efficiency of the primary analysis trigger with respect to the invariant mass of the diphoton system. For $m_{\gamma\gamma} > 100$ GeV, the trigger is 99.4% efficient.

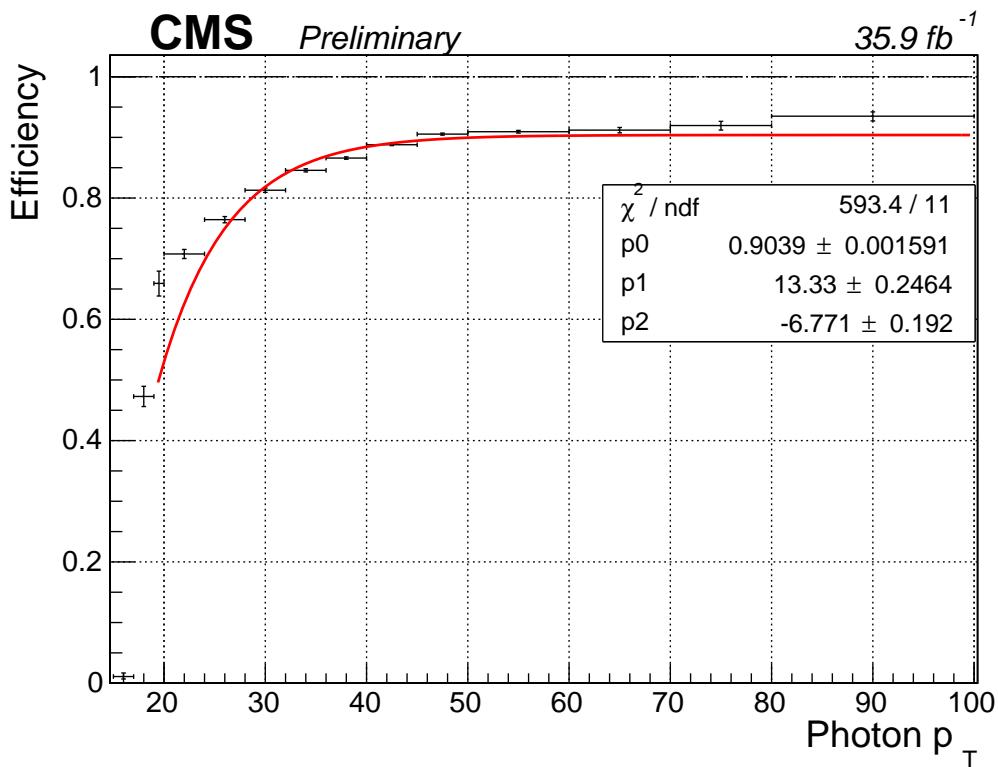


Figure 5.6: Efficiency of the trailing leg of the control trigger with respect to photon p_T . For $p_T > 40$ GeV, the trigger is 90.4% efficient. The drop in efficiency with respect to the primary analysis trigger is caused by requiring both electromagnetic objects to be matched to a pixel seed.

CHAPTER 6

OBJECT RECONSTRUCTION AND SAMPLE SELECTION

6.1 Overview of the CMS reconstruction process

For each event that passes the trigger, the output from all of the sub-detectors described in Chapter 4 is saved by the data acquisition system (DAQ) and eventually recorded to disk and tape. The data format at this stage is referred to as “RAW” data. It includes information about the response of each detector, but the data are still unprocessed, aside from the minimal processing done to determine whether the event passed the trigger. “Reconstruction” is the general term for algorithms that convert the detector response data into lists of object candidates—muons, photons, electrons, jets, etc.—and event quantities such as missing transverse momentum \vec{p}_T^{miss} .

Figure 6.1 shows how each type of particle interacts with the different layers of the CMS detector. Photons, being neutral, do not interact with the tracker, and deposit their energy in the ECAL. Electrons have a very similar response in the ECAL, but unlike photons, they also leave hits in the tracker. Hadrons deposit their energy in the HCAL, and charged hadrons have an associated track in the tracker. Finally, muons interact with the tracker and the muon system.

Of course, the simplified picture in Figure 6.1 does not cover the full complexity of how particles can interact with the detector. Photons can “convert” and produce an electron-positron pair. Hadrons can interact with the material of the tracker or ECAL, and begin to shower before reaching the HCAL. The following sections will describe in more detail how physics objects are reconstructed.

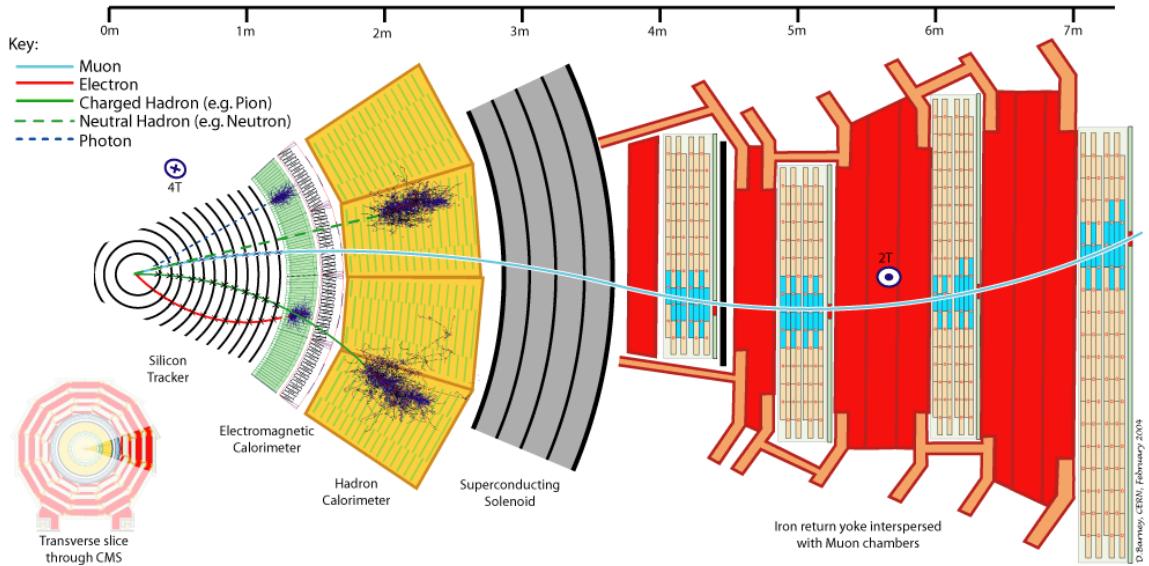


Figure 6.1: Slice of the transverse plane of the CMS detector showing how different particles interact with the detector subsystems. The most important particle for this analysis is the photon (blue dotted line), which interacts only with the ECAL (shown in green). Reprinted from Reference [8].

6.2 Particle flow algorithm

CMS uses a “particle-flow” (PF) reconstruction algorithm [25]. The PF approach seeks to assign each hit in the tracker and each calorimeter energy deposit (known as a cluster) to a final-state particle. This holistic view maximizes our ability to identify particles by using the full set of information provided by the detector. CMS is the first hadron accelerator experiment to use PF for reconstruction. It is made possible through the excellent spatial resolution of the tracker and calorimeters and the large bending power of the magnet. Both of these features are necessary so that signals from nearby particles do not overlap. For example, the ECAL cluster from a charged particle with $p_T = 20$ GeV will be deviated by 5 cm from the ECAL cluster of a neutral particle emitted in the same direction, allowing the two particles to be reconstructed independently [25].

The ingredients to the PF algorithm are reconstructed tracks and calorimeter

clusters. The relevant algorithms and methods for each of these elements will be discussed first, before describing how they are assembled to form the final lists of candidate photons, electrons, muons, and jets.

6.2.1 Track reconstruction

Tracking in CMS takes place with a combinatorial track finder that is applied in ten iterations. The iterations use successively more complex algorithms with looser requirements on the track properties. Each iteration includes three basic steps:

- Seed generation: The first step is to find a track seed consisting of two or three hits compatible with a charge-particle track.
- Track finding: Pattern recognition is used to identify hits in all layers of the tracker that are compatible with the trajectory implied by the track seed.
- Track fitting: A global χ_2 fit is performed to determine the final properties of the track, including origin, direction, and transverse momentum.

Because the charged particles can change direction suddenly through interactions with the tracker material, the track finder uses a Kalman filter method [?]. The goal of each iteration is to find tracks with as high of a purity as possible, at the cost of moderate efficiency. The high purity is accomplished through quality cuts on the seeds and the χ_2 of the track fit. Any hits that are associated with a track are masked for the subsequent iterations. This reduces the combinatorial possibilities, which in turn reduces the likelihood of random hits being built up into a fake tracks.

In addition to particle identification and determination of transverse momentum, tracks are also used to reconstruct vertices in the event. The vertex with the largest value of Σp_T^2 is identified as the primary vertex, associated with the hard-scatter interaction. The other vertices arise from soft-scatter pileup interactions. Thus, the number of reconstructed vertices in an event can be used to approximate the number of pileup interactions. Two of the later iterations of the track finder are applied with

reduced constraints on the origin vertex, in order to reconstruct displaced secondary vertices from the decays of long-lived particles.

Specialized algorithms are used to reconstruct tracks for electrons and muons. Those are described in detail next.

6.2.2 Electron track reconstruction

Electrons have a higher probability of scattering in the tracker than other charged particles, and so electrons are reconstructed using a separate tracking algorithm that is optimized for multiple scatters. In particular, a Gaussian-sum filter (GSF) is used rather than a Kalman filter because it adapts better in the case of sudden and substantial momentum changes. Two approaches are used: an ECAL-driven approach and a tracker-driven approach. In the ECAL-driven method, ECAL clusters are used to infer the hit locations assuming that the cluster came from either a positron or an electron.

The tracker-driven method is better for electrons in jets and for low- p_T electrons, when the ECAL clusters are overlapping with deposits from other particles or are too spread out to reconstruct fully. In this approach, all tracks with p_T greater than 2 GeV are used as seeds for an electron. Each track is propagated to the surface of the ECAL and checked to see if it is consistent with the position of an ECAL cluster. The final step of the algorithm uses a boosted decision tree (BDT) to quantify the goodness of fit. This last step serves to reduce the misidentification rate, which is especially important in the context of the PF algorithm to avoid double counting energy.

6.2.3 Muon track reconstruction

PF muons are classified into three categories:

- All reconstructed tracks with $p_T > 0.5$ GeV and $|\vec{p}| > 2.5$ GeV in the inner

tracker are propagated to the muon system, and classified as **tracker muons** if they match at least one hit in the muon section.

- **Standalone muons** are reconstructed using the muon system only. The process starts with seed tracks in the DT or CSC detectors, and a pattern recognition algorithm identifies DT, CSC, and RPC hits corresponding to the muon trajectory.
- **Global muons** are standalone muons that are consistent with a track in the inner tracker.

Because of the extensive amount of material that muons must traverse before reaching the muon system, the inner tracker gives the best energy resolution for muons up to 200 GeV [25].

6.2.4 Reconstruction of calorimeter clusters

Calorimeter clusters are an important input for the PF algorithm. They are the sole means of detection for photons and neutral hadrons, and they are essential for improving the energy resolution of electrons and high- p_T jets. The reconstruction of a cluster begins with a “seed” cell (ECAL crystal or HCAL scintillating tile) that corresponds to a local energy maxima above a given threshold. Starting from the seed cell, topological clusters are formed by adding cells that share a side or corner with the cluster and have an energy that is at least twice the noise threshold. For the EB, the seed energy must be greater than 230 MeV, and subsequent crystals must have an energy of at least 80 MeV in order to be added to the topological cluster. Finally, a fit is performed to determine what fraction of the energy in the cell is associated with each PF cluster, where the number of PF clusters is equal to the number of seeds incorporated into the topological cluster.

The ECAL clusters are calibrated for photons and electrons using test beam data and radioactive sources, and residual corrections are derived using single photon MC samples. For hadrons, however, the ECAL must be re-calibrated, since the response of the ECAL to hadrons is significantly different than its response to electrons. A

large simulated sample of neutral hadrons is used to derive a simultaneous calibration for the portions of the hadronic shower’s energy deposited in the ECAL and HCAL.

One complication in the reconstruction of ECAL clusters is the presence of signals known as spikes. These occur when a particle directly ionizes an ECAL APD. The result is a signal whose amplitude is 10^5 times higher than the signal from the scintillation light alone. Because these spikes appear in only one or two crystals, they are rejected by considering the ratio of energy deposited in the central crystal to that deposited in the neighboring 4 crystals. In the case where the spike affects two adjacent crystals, the energy deposited in the two crystals is compared to that of their six neighbors. These ratios are referred to as E_4/E_1 and E_6/E_2 , and are required to be less than 5% and 10%, respectively.

6.2.5 Particle identification

A “link” algorithm is used to connect tracks and calorimeter clusters and quantify the likelihood that two elements arose from the same particle. A dedicated procedure is used to link tracks with any ECAL clusters that are consistent with photons from electron bremsstrahlung and to link two tracks compatible with a photon conversion. A group of linked elements is referred to as a PF block.

Particle flow is an iterative procedure. As object candidates are identified within a PF block, their corresponding tracks and clusters are removed from further consideration. This makes PF different from more traditional reconstruction algorithms where clusters or hits could be attributed to more than one particle. The object-identification process is repeated until all tracks and clusters have been assigned to a single PF candidate.

First, muon candidates are identified with two sets of identification criteria: one for isolated muons and one for muons inside or near jets. Next, electron candidates and photon candidates are reconstructed. Finally, any remaining tracks or calorimeter

clusters in the PF block are assigned to a neutral hadron, charged hadron, or photon candidates.

Because photons are the primary object of interest in this analysis, their reconstruction is described in more detail below.

6.3 Photon identification

The reconstruction of photon clusters with a seed as defined in Section 6.2.4. The ratio of the energy in a 3×3 array of crystals to that in a 5×5 region is referred to R9. If R9 is greater than 0.93, than the photon is classified as unconverted, and the energy of the 5×5 region is assigned as the photon super-cluster energy. If the photon has converted (ie, $R9 < 0.93$), then the super-cluster is built starting with the seed crystal and moving outward.

6.4 Jet reconstruction

Jets are reconstructed from PF candidates using the anti- k_T algorithm with a distance parameter $R = 0.4$ [7]. The anti- k_T algorithm is a sequential clustering algorithm where the distance d_{ij} between two particles i and j and the distance d_{iB} between particle i and the beam B are given by the following:

$$d_{ij} = \min(k_{ti}^{2p}, k_{ti}^{2p}) \frac{\Delta_{ij}^2}{R^2} \quad (6.1)$$

$$d_{iB} = k_{ti}^{2p}$$

The value Δ_{ij}^2 is equal to $(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$, k_t is the transverse momentum of the particle, and R is a distance parameter that determines the final radius of the jet.

The clustering algorithm first finds the smallest value of d_{ij} and d_{iB} for all particles in the event. If the minimum distance is d_{ij} , then particles i and j are combined into

a single entity. If the minimum distance is d_{iB} , then particle i is labeled a jet and removed from the list. For the anti- k_T algorithm, the parameter $p = -1$. Setting $p = 0$ corresponds to the inclusive Cambridge/Aachen algorithm, and $p = 1$ is the inclusive k_t algorithm. The result of various jet-clustering algorithms is shown in Figure 6.2. One obvious feature of the anti- k_T algorithm is the production of very circular jets compared to jets identified using other clustering algorithms.

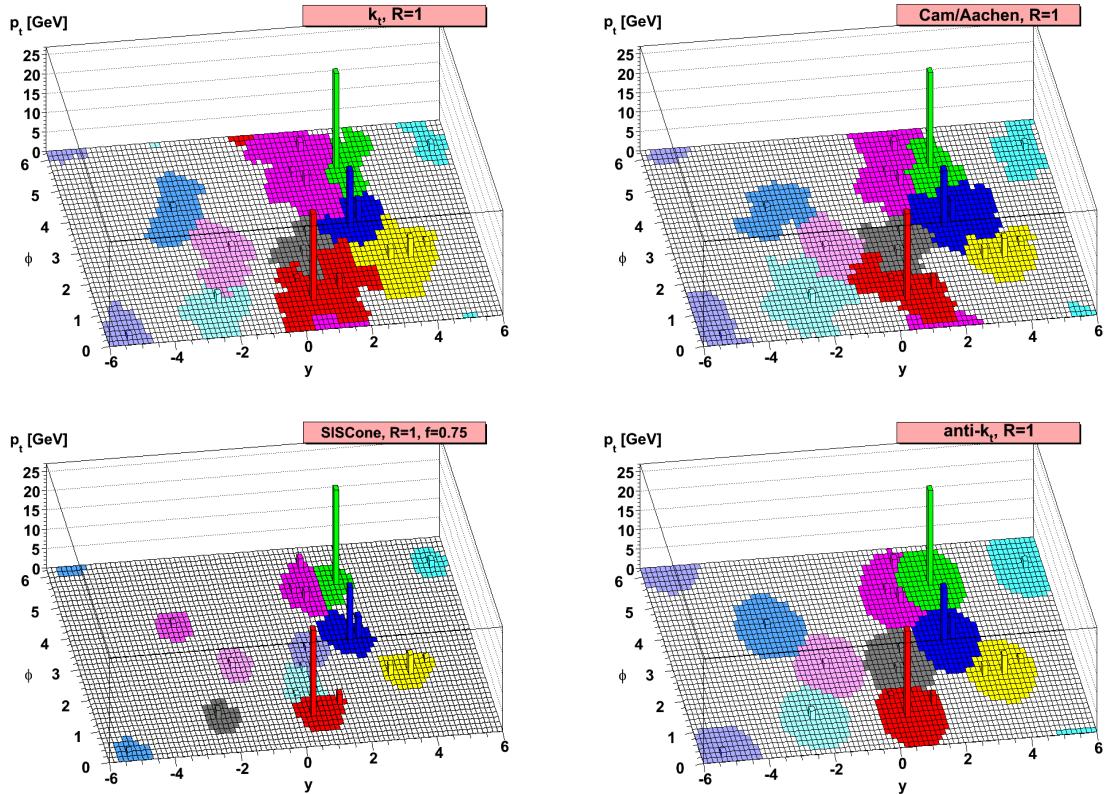


Figure 6.2: Behavior of several jet-clustering algorithms illustrated with a sample parton-level event. CMS uses the anti- k_T algorithm (bottom right) with a distance parameter $R = 0.4$. Reprinted from Reference [7].

One benefit of the anti- k_T algorithm is that it is infrared and collinear (IRC)

safe. Infrared safe means that anti- k_T jets are insensitive to nearby soft particles. In the anti- k_T algorithm, soft particles (those with $k_t \rightarrow 0$) have a large d_{ij} and are therefore clustered last. This leaves hard jets unaffected. Soft particles could possibly be clustered into many soft jets, but those are simple to remove during the analysis. Being infrared safe is particularly important for the high luminosity environment of the LHC, so that the clustering of the hard jets from the primary interaction is not changed by the many soft particles from pileup interactions.

Collinear safe means that very energetic initial quarks will still get reconstructed as a single jet. In the anti- k_T algorithm, collinear particles will have a small value of Δ_{ij}^2 and will be clustered together first. Using an IRC safe algorithm is important for comparing theory and experiment, because if jets are not IRC safe, then their cross sections cannot be calculated using perturbation theory.

6.4.1 Jet energy corrections

A series of jet energy corrections (JEC) are applied to the jets after clustering to improve the calibration and energy resolution [22]. The $L1$ correction is a flat correction designed to remove contributions to the jet energy from pileup (see Section 6.6 for a more thorough description of pileup mitigation techniques). The $L2$ corrections are *eta*-dependent and the $L3$ corrections are p_T -dependent. These are calculated in simulation using generator truth information. In data, the correction factors are derived using dijet events and events where a jet is emitted back-to-back with an photon. In the latter case, the superior energy resolution of the ECAL is exploited to provide a handle on the jet energy. The effects of JEC on jet response in QCD MC is shown in Figure 6.3.

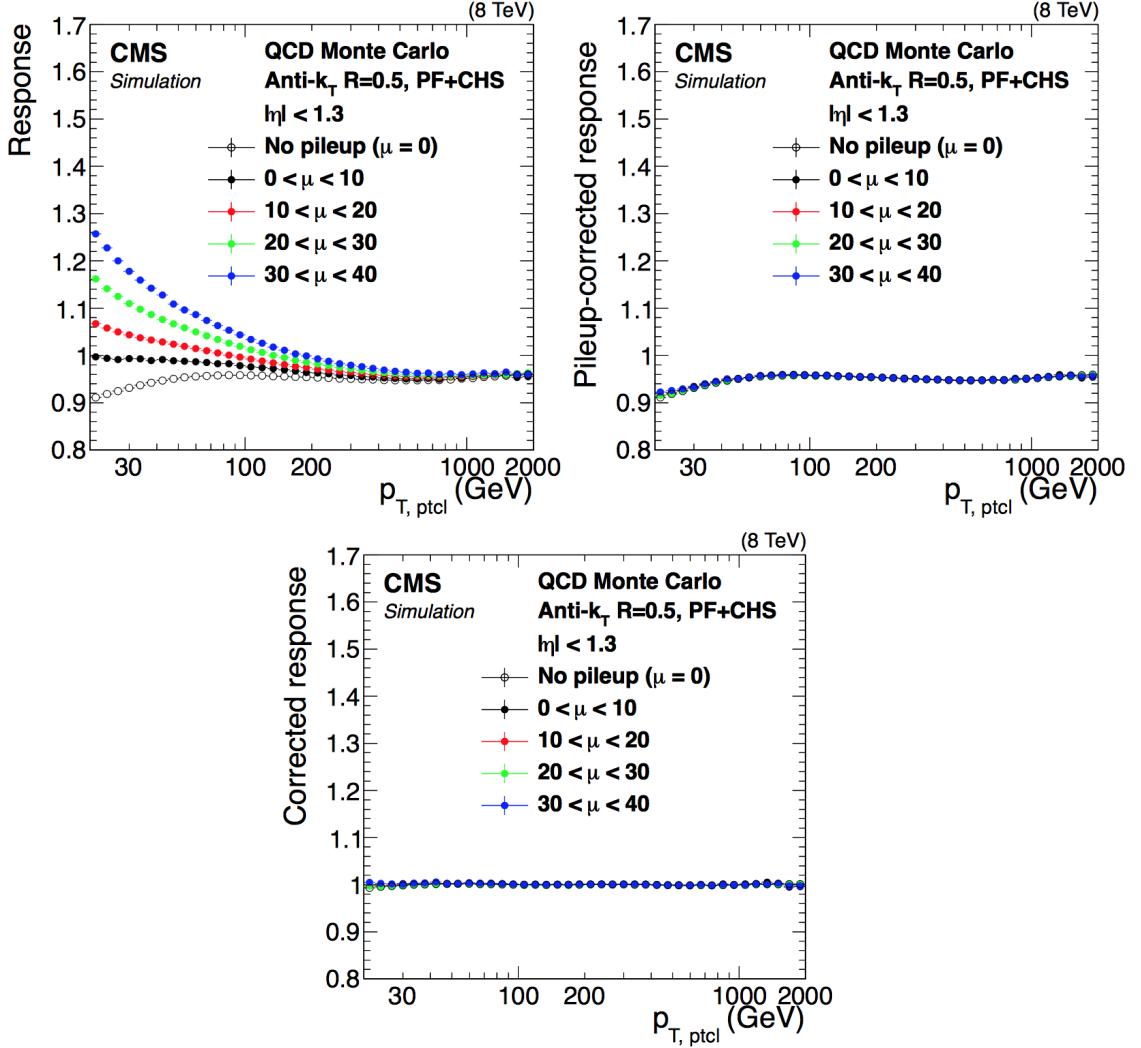


Figure 6.3: Ratio of the average reconstructed jet p_T to the particle-level jet p_T ($p_{T,\text{ptcl}}$) in QCD MC simulation. The ratio is calculated in bins of $p_{T,\text{ptcl}}$ before applying any JEC (top left), after applying the $L1$ pileup corrections (top right), and after all JEC (bottom). The different curves correspond to different numbers of pileup interactions μ . Reprinted from Reference [22].

6.4.2 Jet energy resolution

After the JEC are applied, resolutions of 10% and 5% are achieved for 100 GeV and 1 TeV jets in the barrel, respectively. Figure 6.4 shows the jet energy resolution as a function of jet p_T in 8 TeV simulation.

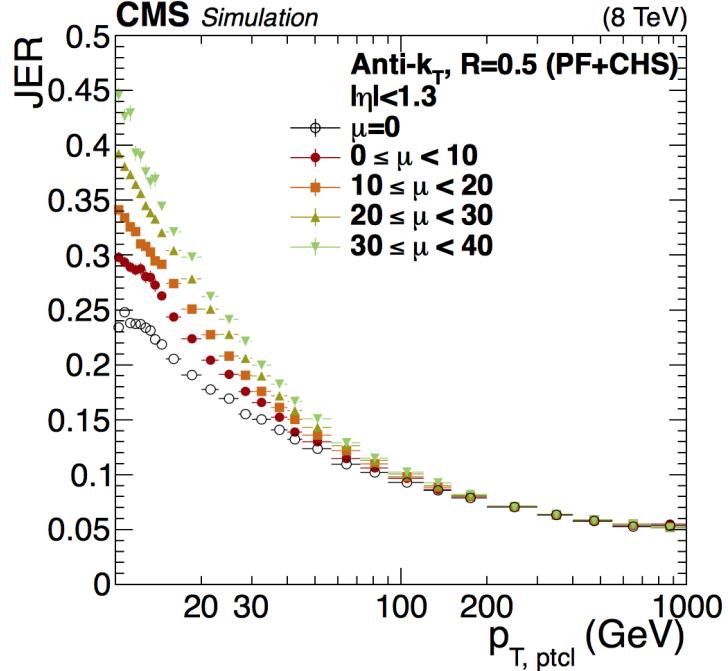


Figure 6.4: Jet energy resolution (JER) in 8 TeV simulation as a function of particle-level jet p_T ($p_{T,\text{ptcl}}$). PF+CHS refers to particle-flow jets with charged hadron subtraction (see Section 6.6). Results are shown for various levels of pileup μ . Reprinted from Reference [22].

6.5 Missing transverse momentum

After the full set of PF objects have been identified and reconstructed, the missing transverse momentum \vec{p}_T^{miss} is given by the following, where the sum is over all Particle Flow objects:

$$\vec{p}_T^{\text{miss}} = \sum \vec{p}_T \quad (6.2)$$

The magnitude of this quantity is referred to as the missing transverse energy E_T^{miss} . For the calculation of E_T^{miss} in this analysis, Type-I corrections are applied, which correspond to propagating the JEC described above to the E_T^{miss} .

There are several situations that can arise where particle misidentification or poor reconstruction leads to an artificially large value of E_T^{miss} . These are dealt with in post-processing by substituting the reconstructed particles with alternative

hypotheses and checking how the E_T^{miss} changes. For example, very energetic charged hadrons can occasionally punch through the HCAL and leave tracks in the muon system. In this case, PF will reconstruct a neutral hadron and a muon, leading to a large value of \vec{p}_T^{miss} pointing in the opposite direction. If replacing the neutral hadron and muon with a charged hadron reduces the E_T^{miss} by 50% or more, then that alternative reconstruction is used instead.

In addition to the post-processing corrections just described, a series of E_T^{miss} filters are applied. The filters are designed to tag events where the E_T^{miss} is poorly reconstructed. At the analysis stage, if an event fails any of the E_T^{miss} filters it is removed from consideration. The following E_T^{miss} filters applied in this analysis:

- Primary vertex filter
- CSC beam halo filter
- HBHE noise filter
- HBHEiso noise filter
- ECAL trigger primitive filter
- ee badSC noise filter
- Bad PF muon filter
- Bad charged hadron filter
- Duplicate muons filter

6.6 Pileup subtraction

The presence of pileup interactions affects many aspects of the reconstruction process, in particular the jet energy resolution, lepton isolation values, and E_T^{miss} . Fortunately, there are several ways in which the PF algorithm can be used to mitigate the effects from pileup. Charged-hadron subtraction (CHS), for example, is the procedure of removing charged hadrons from consideration if they are associated with

a pileup vertex rather than the hard-scatter vertex. Figure 6.5 shows the composition of jets in data and simulation as a function of the number of pileup interactions μ . For an event with $\mu = 30$, 10% of the uncorrected jet energy is due to charged hadrons from pileup interactions.

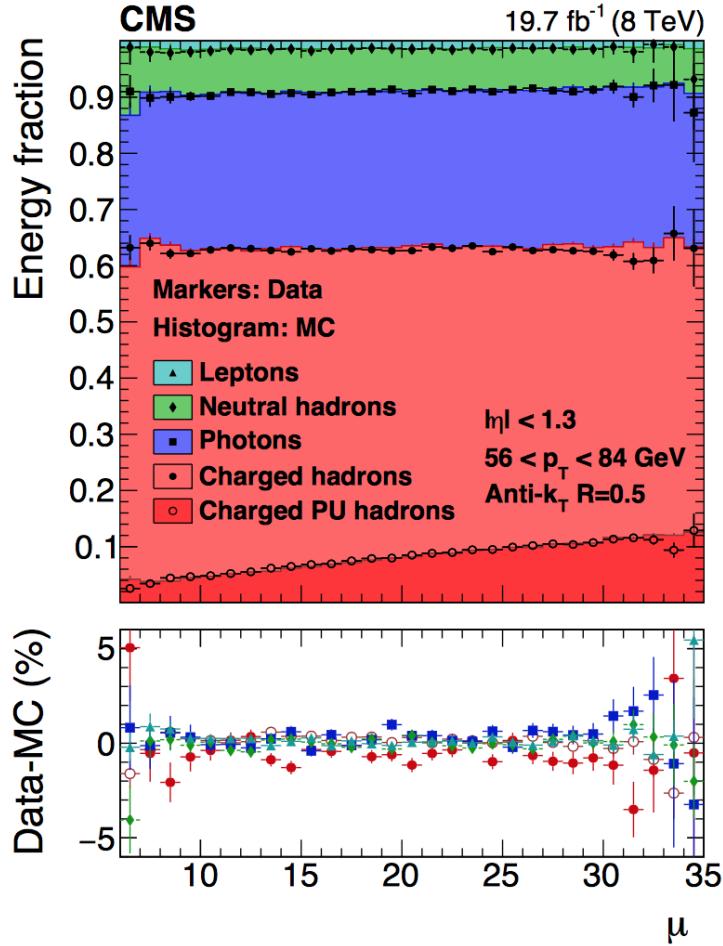


Figure 6.5: Energy composition as a function of the number of pileup (PU) interactions μ for jets in data (markers) and simulation (histogram). The data distributions correspond to 19.7 fb^{-1} collected at $\sqrt{s} = 8 \text{ TeV}$ with the CMS detector in 2012. The contribution from charged hadrons from pileup is shown in dark red. The bottom pane shows the agreement between data and MC for each category. Reprinted from Reference [25].

The pileup contributions from photons and neutral hadrons, however, cannot be subtracted on an object-by-object basis, since there is no good way to identify whether these neutral particles were produced in the primary interaction or in a pileup interaction. These effects can only be subtracted from the event on average. For lepton and photon isolation, this is done by calculating ρ , the p_T density due to pileup interactions [?]. The average contribution to the isolation sum from pileup is then given by ρ times the effective area of the object under consideration. The calculation of the effective areas for photons and electrons is described in Reference [?].

CHAPTER 7

Sample Selection

7.1 MC Samples

The list of all MC samples used for data versus MC comparisons and background studies is shown in Table 7.1.

TABLE 7.1

MC SAMPLES FOR SIGNAL AND BACKGROUND

SUSY Signal Samples
T5Wg
T5Wg ext
T6Wg
T6Wg ext
Background MC Samples
GJet
Zgg
DY

List of MC samples used for signal and background studies.

7.2 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.

7.2.1 Photon identification

7.2.2 Electron identification

7.2.3 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 7.2.1, except they are required to fail either the $\sigma_{inj\eta}$ or the charged hadron isolation requirement. This is described in detail in Table 7.3 (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

7.2.4 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 $\Delta R <$

TABLE 7.2
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 ??.

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 ΔR separation for fakes is due to the fact that fakes are primarily jets. As described in Section XX, jets are reconstructed using the anti- k_T algorithm with a distance parameter of 0.4.

7.2.5 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

TABLE 7.3
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.

the final state. As described in Section XX, this is particularly important for the combination paper.

7.2.6 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 5.2. 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 5.1 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.

7.3 Photon ID scale factor

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^{miss} in the event. Instead, the measured E_T^{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^{miss} from the neutrino, and these events can mimic our signal topography 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^{miss} and without two real photons. The observed E_T^{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 7.2. 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^{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 7.2.6.

8.2.1 Di-EM p_T reweighting

Because the E_T^{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):

$$\vec{p}_T^{\text{di-EM}} = \vec{p}_T^1 + \vec{p}_T^2 \quad (8.1)$$

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.

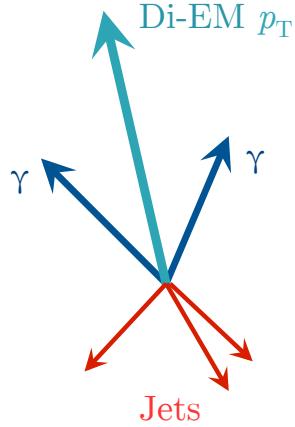


Figure 8.1: The di-EM p_T vector, shown in light blue, is the vector sum of the p_T of the two photons in the event, shown in blue. The magnitude of the di-EM p_T vector is used to model the hadronic recoil, shown in red.

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

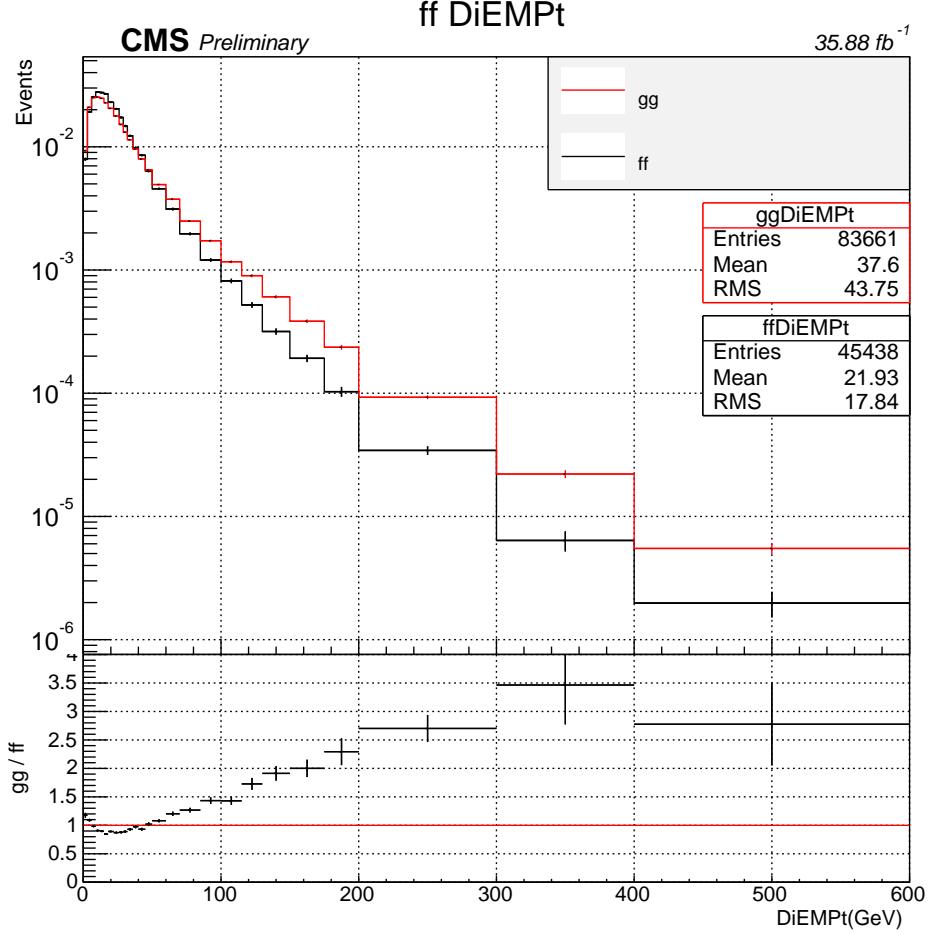


Figure 8.2: Di-EM p_T distributions of the ff control sample (black) and the $\gamma\gamma$ candidate sample (red). The ratio plot on the bottom shows the $\gamma\gamma/ff$ ratios that are used to reweight the ff E_T^{miss} distribution.

8.2.2 Cross check on QCD background

In order to set a systematic uncertainty on the overall E_T^{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^{miss} . If this assumption is true, then we should be able to extrapolate from the low- E_T^{miss} control region to the high- E_T^{miss} signal region.

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

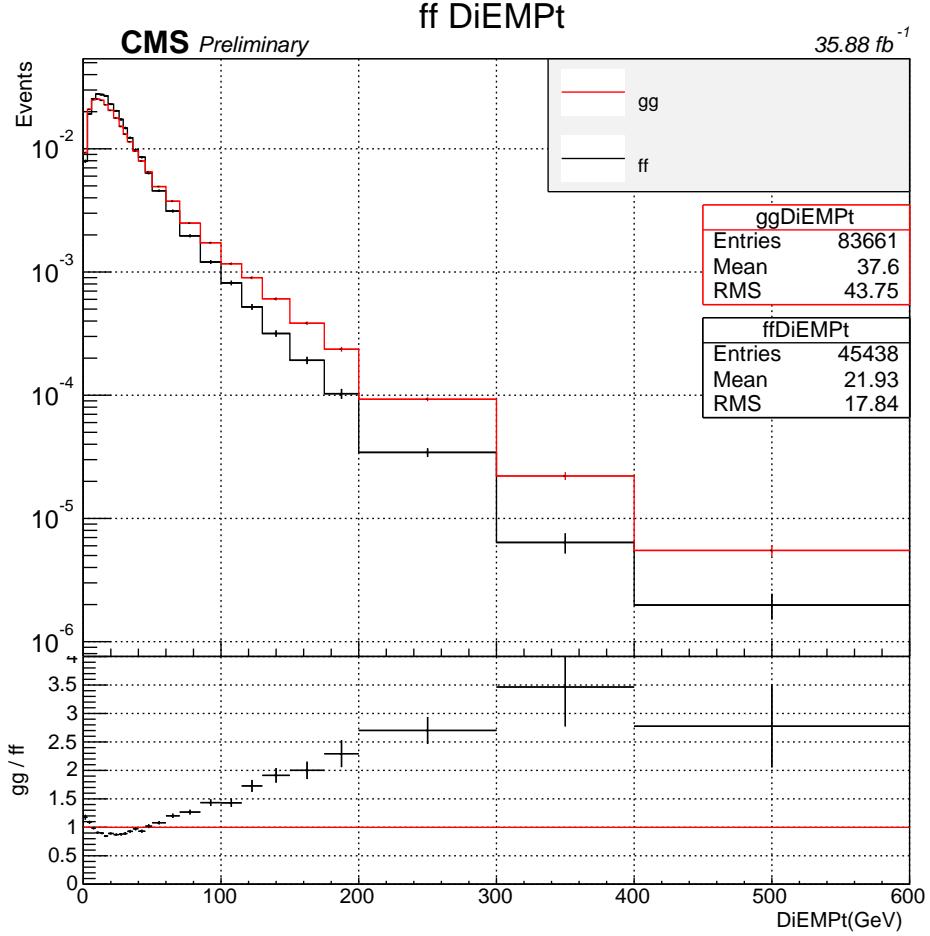


Figure 8.3: Di-EM p_T distributions of the ff control sample (black) and the $\gamma\gamma$ candidate sample (red). The ratio plot on the bottom shows the $\gamma\gamma/ff$ ratios that are used to reweight the ff E_T^{miss} distribution.

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

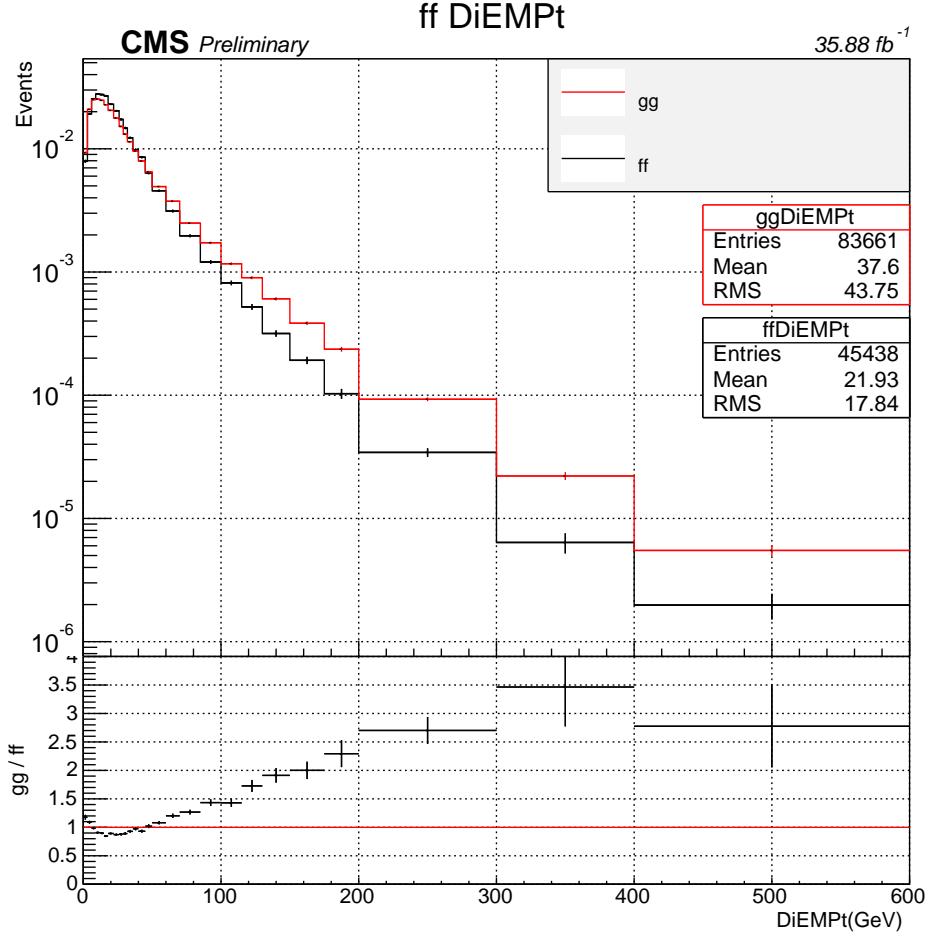


Figure 8.4: Di-EM p_T distributions of the ff control sample (black) and the $\gamma\gamma$ candidate sample (red). The ratio plot on the bottom shows the $\gamma\gamma/ff$ ratios that are used to reweight the ff E_T^{miss} distribution.

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 1σ 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 ??).

The two methods give overlapping predictions within uncertainty, and therefore

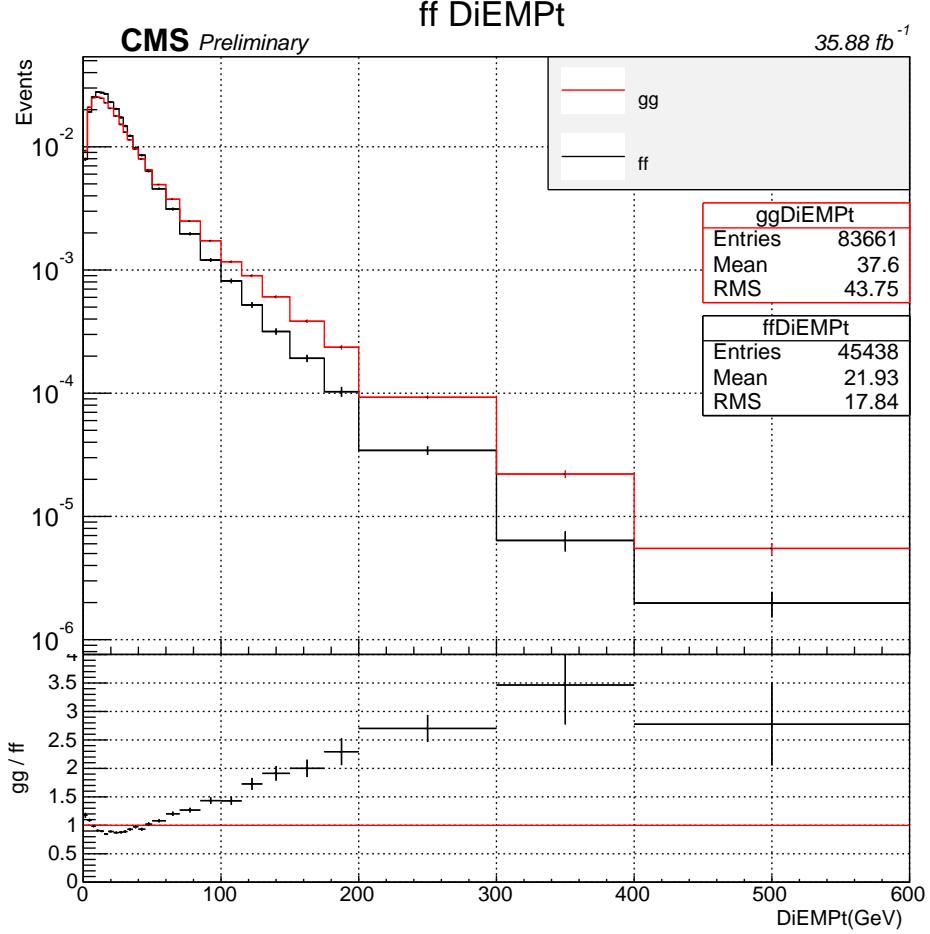


Figure 8.5: Di-EM p_T distributions of the ff control sample (black) and the $\gamma\gamma$ candidate sample (red). The ratio plot on the bottom shows the $\gamma\gamma/ff$ ratios that are used to reweight the ff E_T^{miss} distribution.

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^{miss} shape.

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 is referred to as the electroweak (EWK) background. Unlike the QCD background,

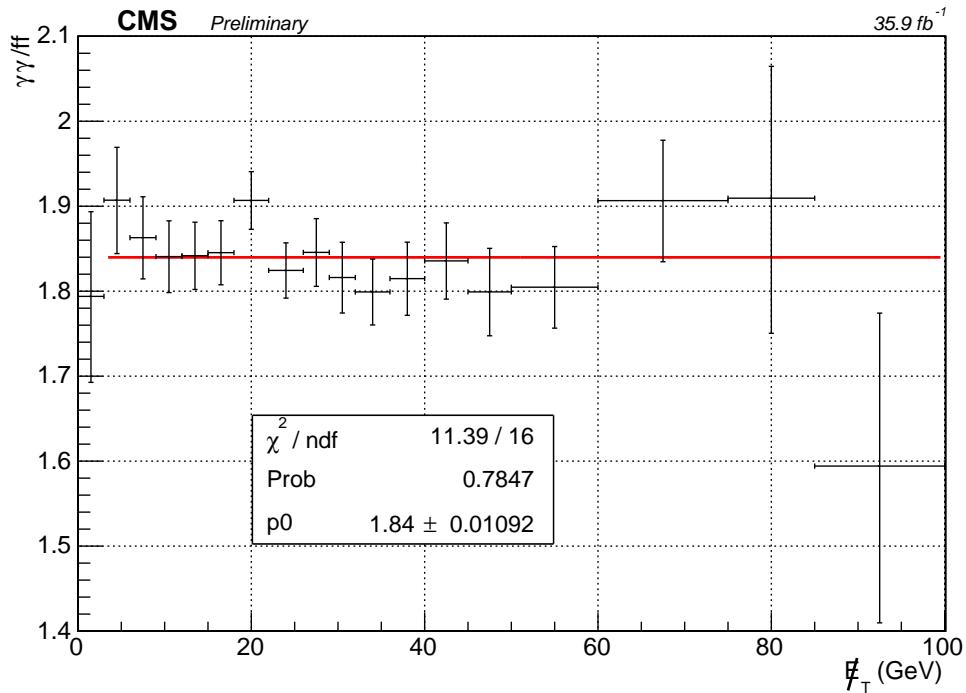


Figure 8.6: The ratio of $\gamma\gamma$ to ff events for $E_T^{\text{miss}} < 100$ GeV. The ratio has been fit to a constant function, which can then be used to extrapolate to the $E_T^{\text{miss}} > 100$ GeV signal region. This provides an alternative estimation method for the QCD background.

there is inherent E_T^{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 (ee) with the invariant mass peak in a sample of events with one electron and one photon ($e\gamma$). The composition of the $e\gamma$ control sample is investigated in Section 8.3.1 and the calculation of this “fake rate” is described in detail in Section 8.3.2. To get the final expected contribution from the EWK background in our signal region, the fake rate is used to calculate a transfer factor (Section 8.3.3). 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 (Section 8.3.4).

TABLE 8.1
 COMPARISON BETWEEN REWEIGHTING METHOD AND RATIO
 METHOD FOR THE QCD BACKGROUND ESTIMATE

E_T^{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 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 MC comparison of this control sample is shown in Figure 8.7. 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 $\gamma + \text{jet}$ processes. For $E_T^{\text{miss}} > 130$ GeV, $\gamma + \text{jet}$ events contribute only 2.6%.

8.3.2 Fake rate calculation

A tag and probe method is used for the fake rate calculation. Events are selected using a single electron trigger. Electrons passing medium ID criteria and satisfying

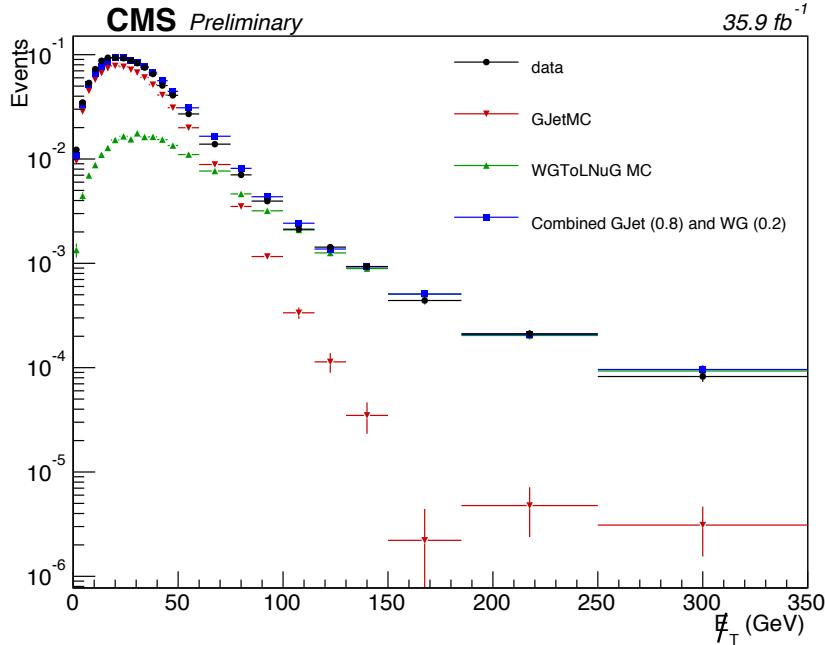


Figure 8.7: Data versus MC comparison of the $e\gamma$ control sample E_T^{miss} distribution. To determine the relative contributions of the $\gamma + \text{jet}$ and $W\gamma$ processes, the data distribution was fit using the $\gamma + \text{jet}$ and $W\gamma$ shapes as templates. The data are shown 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.

$p_T > 30 \text{ GeV}$ and $|\eta| < 2.1$ are used as tags. The tags are required to be matched to the object firing the single electron trigger within $\Delta R < 0.2$. Probes are photon candidates with $p_T > 40 \text{ GeV}$ that pass all of the identification criteria described in Chapter 7 except for the pixel seed veto. The tag and probe objects are required to be separated by $\Delta R > 0.3$ and have an invariant mass between 40 and 140 GeV. If there are multiple tag and probe pairs in a single event, all possible combinations are considered. If the probe has a pixel seed, then it is labeled as an electron and falls into the ee sample. If the probe does not have a pixel seed, however, then it is labeled as a photon and is included in the $e\gamma$ sample.

Because processes other than $Z \rightarrow ee$ decays can be included in this sample, a fit is performed on each invariant mass distribution. The shape of the invariant mass peak

in $Z \rightarrow ee$ MC is used as a signal template, and an electron + muon control region in data is used as the background template. After fitting the ee and $e\gamma$ distributions in data, the integrals of the signal shapes between 80 and 100 GeV are referred to as N_{ee}^Z and $N_{e\gamma}^Z$, respectively. Figure 8.8 shows the shapes of the data distributions as well as the background and signal templates with arbitrary normalization. The results after the fit are shown as well.

The maximum of the background shape at 90 GeV arises from the p_T thresholds of the tag and probe objects. The p_T distributions tend to be sharply falling, so most probes have a p_T around 40 GeV and most tags have a p_T around 30 GeV. Changing the probe p_T threshold changes the position of the peak in the background template.

The ratio $R_{\gamma/e} = N_{e\gamma}^Z/N_{ee}^Z$ is calculated from the fits and found to be $R_{\gamma/e} = 2.63\%$ overall. The fits are performed in bins of different kinematic variables to investigate any potential dependencies of $R_{\gamma/e}$. These are shown in Figures 8.9 and 8.10.

A 30% systematic uncertainty is assigned to the ratio $R_{\gamma/e}$ to cover the observed dependencies. This is shown in the red bands in Figures 8.9 and 8.10. The strongest dependency is with respect to the vertex multiplicity. Recall that vertex multiplicity is a proxy for the number of pileup interactions in the event. For a similar analysis that used this fake rate XX cite Knut XX, $R_{\gamma/e}$ was parameterized as a function of the vertex multiplicity, but no significant difference was found between applying the linear fit shown in Figure 8.10 rather than the constant value of 2.63%.

The distinctive features of $R_{\gamma/e}$ as a function of $|\eta|$ are due to detector effects. Similar features are seen in data quality monitoring plots showing the occupancy of the ECAL crystals as a function of η and ϕ .

8.3.3 Calculating EWK estimate

The number of observed $Z \rightarrow ee$ events in the $e\gamma$ and ee mass spectra can be expressed as a function of $f_{e \rightarrow \gamma}$, the rate at which electrons are misidentified as

photons:

$$\begin{aligned} N_{e\gamma}^Z &= f_{e \rightarrow \gamma} (1 - f_{e \rightarrow \gamma}) N'_Z \\ N_{ee}^Z &= (1 - f_{e \rightarrow \gamma})^2 N'_Z \end{aligned} \tag{8.3}$$

where N'_Z is the actual number of $Z \rightarrow ee$ events.

If we then consider a generic sample of $N'_{e\gamma}$ events with a real photon and a real electron, the number of events $N_{e\gamma}$ that will actually get reconstructed as having one photon and one electron is given by:

$$N_{e\gamma} = (1 - f_{e \rightarrow \gamma}) N'_{e\gamma} \tag{8.4}$$

The background $N_{\gamma\gamma}$, the fraction of $N'_{e\gamma}$ that ends up in the candidate $\gamma\gamma$ sample, can be written as:

$$N_{\gamma\gamma} = f_{e \rightarrow \gamma} N'_{e\gamma} = N_{e\gamma} \frac{f_{e \rightarrow \gamma}}{1 - f_{e \rightarrow \gamma}} \tag{8.5}$$

By looking at Equation 8.3, it can be seen that this last factor is precisely the ratio $R_{\gamma/e} = N_{e\gamma}^Z / N_{ee}^Z = 2.63\%$ that was discussed above. Thus, the EWK background prediction is simply given by $R_{\gamma/e}$ times the number of observed $e\gamma$ events.

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^{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. The predicted background from $Z\gamma\gamma$ events is shown in Table ??.

XX Table XX

8.5 Systematic uncertainties

The systematic uncertainties fall into two main categories: those associated with one of the background estimates, and other uncertainties used in the limit-setting procedure.

8.5.1 Uncertainties on background estimates

The largest uncertainties on the background estimate come from the QCD prediction method. The uncertainties on the QCD estimate are summarized for each E_T^{miss} bin in the signal region in Table ???. All uncertainties are expressed as numbers of events.

One of the uncertainties on the QCD estimate comes from the di-EM p_T reweighting procedure. Recall that to correct for differences in hadronic activity in the events, the E_T^{miss} distribution of the ff sample is reweighted by di-EM p_T , so that its di-EM p_T spectrum matches that of the double photon candidate sample. To propagate the uncertainty from the di-EM p_T reweighting, we generated one thousand different di-EM p_T ratios. In each case, the value of the ratio in each bin is obtained by varying the nominal di-EM p_T ratio using a Gaussian distribution with σ equal to the statistical uncertainty of that bin. We then reweight the E_T^{miss} distributions of the ff control sample with the 1000 generated di-EM p_T ratios. Final uncertainties are then determined from the fluctuations in each E_T^{miss} bin by finding the range that includes 68% of the new E_T^{miss} values.

Another systematic uncertainty on the QCD background estimation comes from the uncertainty in the overall E_T^{miss} shape. To estimate this uncertainty, we use the ratio method described in Section 8.2.2 as a cross-check. By fitting the ratio of $\gamma\gamma$ to ff at low E_T^{miss} to a constant, the expected number of $\gamma\gamma$ events at high E_T^{miss} can be extrapolated from the observed number of ff events. This procedure gives us a second estimates for the QCD background in each bin of E_T^{miss} . The difference between the two estimates is taken as a symmetric systematic uncertainty and is shown in the second-to-last column of Table ??.

As described above, there are two uncertainties on the EWK background: the statistical uncertainty from the limited statistics in the control region and the 30% uncertainty on the $e \rightarrow \gamma$ fake rate. For the $Z\gamma\gamma$ background, the only uncertainty considered is a 50% uncertainty to cover any potential mismodeling of the E_T^{miss} shape or uncertainty on the $Z\gamma\gamma \rightarrow \nu\nu\gamma\gamma$ cross section.

8.5.2 Other sources of systematic uncertainties

In addition to the systematic uncertainties arising from the background estimation techniques described above, there are other systematic uncertainties that affect the final analysis sensitivity. These uncertainties are summarized in Table ???. Ranges of uncertainty arise when there are different values for the uncertainty for different signal points.

The first is the uncertainty in the Parton Distribution Functions (PDF?)s and the variation of cross section ratios (K factors) between leading order PDF? and next-to-leading order PDF? . The PDF uncertainties are taken from the NNPDF30 variations XX cite XX .

Other uncertainties on the signal include finite MC statistics and the photon data/MC scale factor described in Section 7.3. There is also a 2.5% uncertainty on the integrated luminosity of the data sample [14].

Finally, an additional source of uncertainty comes from propagating the jet energy scale uncertainty described in Section 6.4.1 to the E_T^{miss} . The expected signal yields are re-calculated using the $1\ \sigma$ fluctuations of E_T^{miss} if the jet energy scale is varied by its uncertainty. The difference between the yields with and without the jet energy scale changes is taken as a systematic uncertainty.

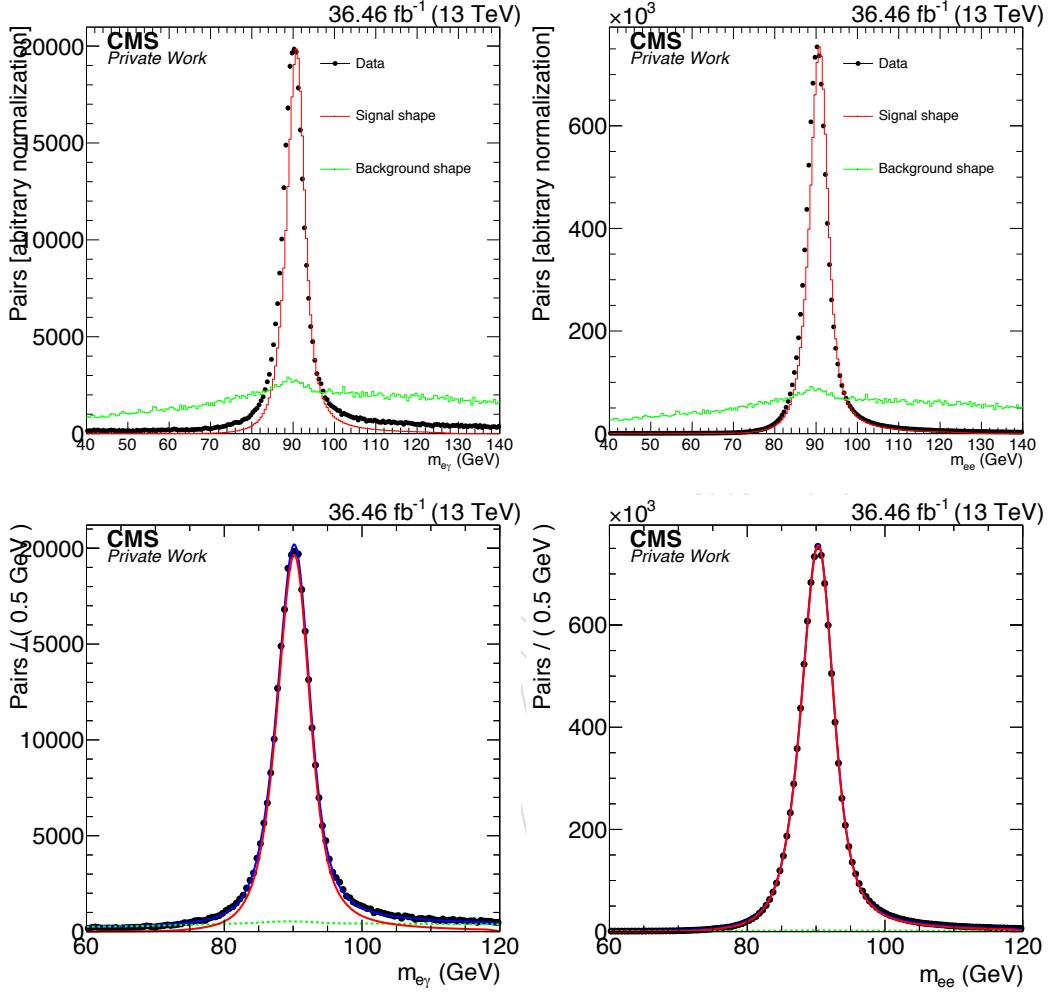


Figure 8.8: Invariant mass distributions for the $e\gamma$ (left) and ee (right) samples. Each plot includes the data (black markers), the signal template (red), and the background template (green). The top plots demonstrate the shape of each distribution with arbitrary normalization, and the bottom plots show the distributions after the fits are performed.

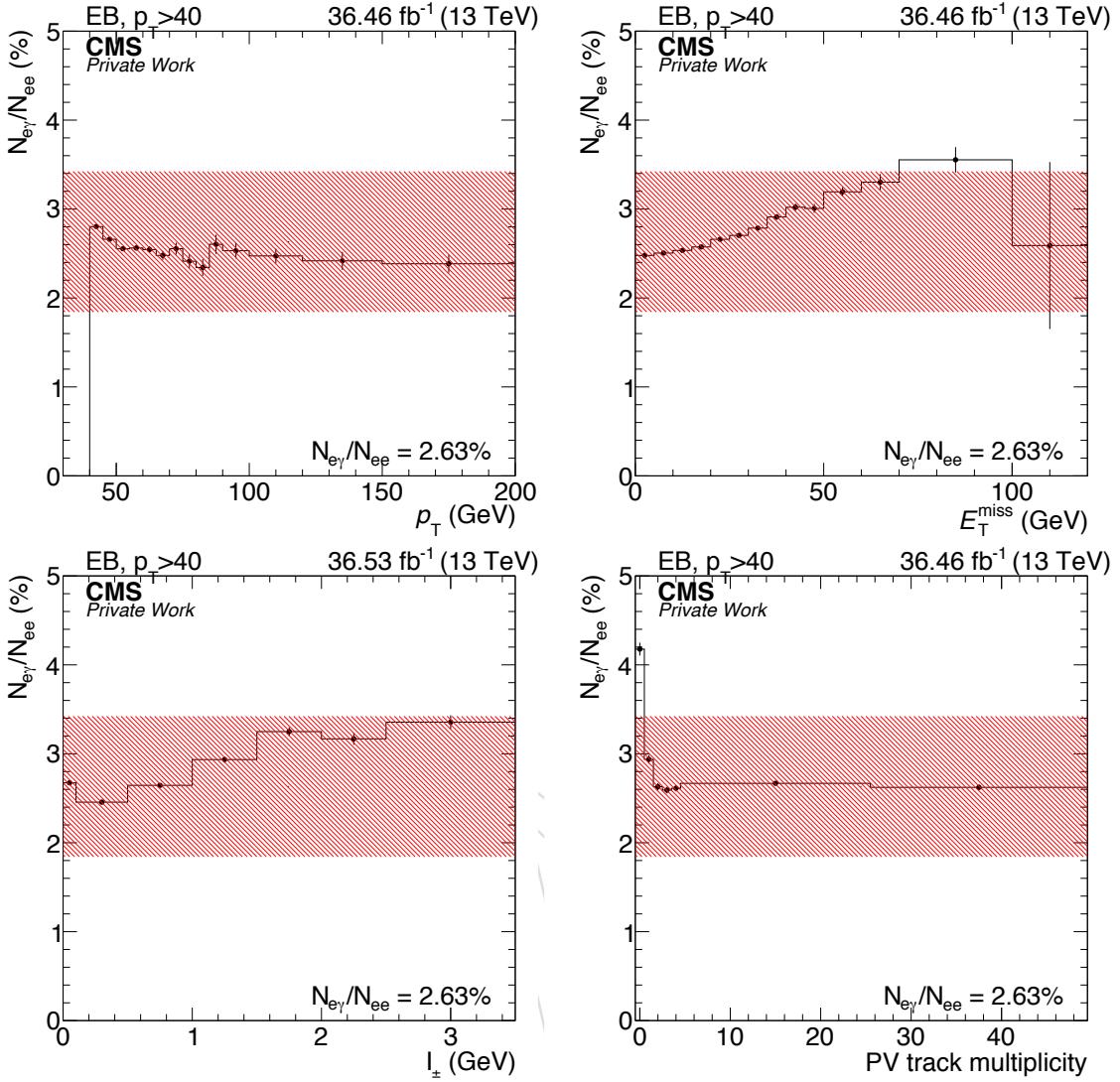


Figure 8.9: Value of $R_{\gamma/e}$ as a function of various kinematic variables for probes in the barrel: probe p_T , E_T^{miss} , isolation I_{\pm} , and track multiplicity of the primary vertex (PV). The red band corresponds to a 30% uncertainty on $R_{\gamma/e}$.

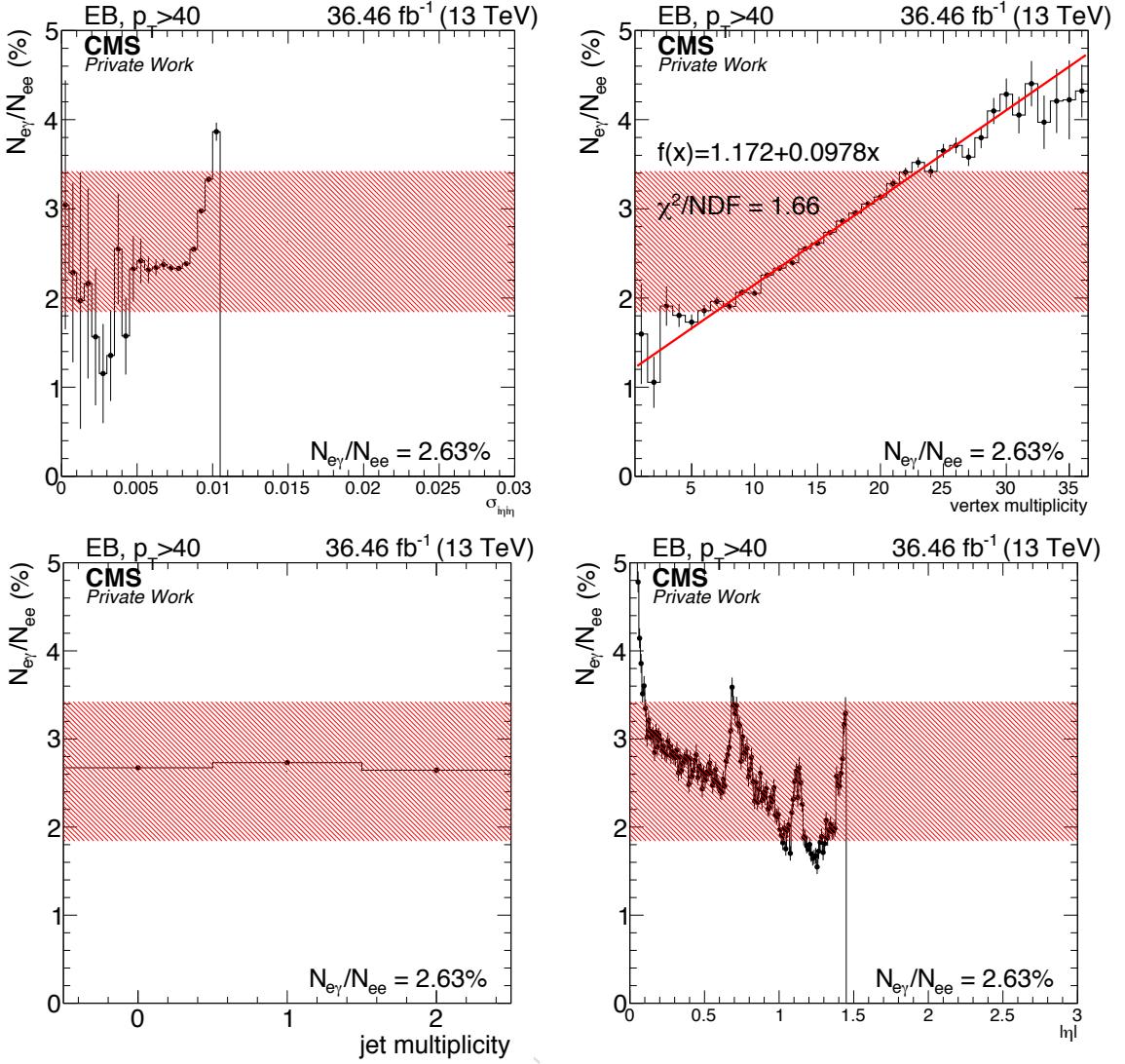


Figure 8.10: Value of $R_{\gamma/e}$ as a function of various kinematic variables for probes in the barrel: $\sigma_{inj\eta}$ of the probe, vertex multiplicity, jet multiplicity, and probe $|\eta|$. The red band corresponds to a 30% uncertainty on $R_{\gamma/e}$.

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^{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^{miss} (GeV)	Exp. QCD	Exp. EWK	Z $\gamma\gamma$ events	Total exp.	Observed
100 – 115	69.23 ^{+20.67} _{-18.91}	8.17 ± 2.5	1.30±0.65	78.69 ^{+22.58} _{-20.98}	65
115 – 130	30.89 ^{+14.69} _{-12.46}	5.50 ± 1.70	1.14±0.57	37.53 ^{+16.05} _{-14.04}	27
130 – 150	25.98 ^{+15.21} _{-12.76}	4.78 ± 1.48	1.12±0.56	31.88 ^{+16.44} _{-14.20}	17
150 – 185	20.49 ^{+11.65} _{-9.16}	3.95 ± 1.24	1.32±0.66	25.76 ^{+12.80} _{-10.59}	13
185 – 250	8.74 ^{+12.70} _{-7.765}	3.52 ± 1.11	1.28±0.64	13.55 ^{+13.05} _{-8.31}	8
≥ 250	5.13 ^{+12.31} _{-5.514}	2.11 ± 0.69	1.14±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` [3] 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` [26] program with the CUETP8M1 generator tune. The detector response is simulated using CMS fast simulation [2].

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

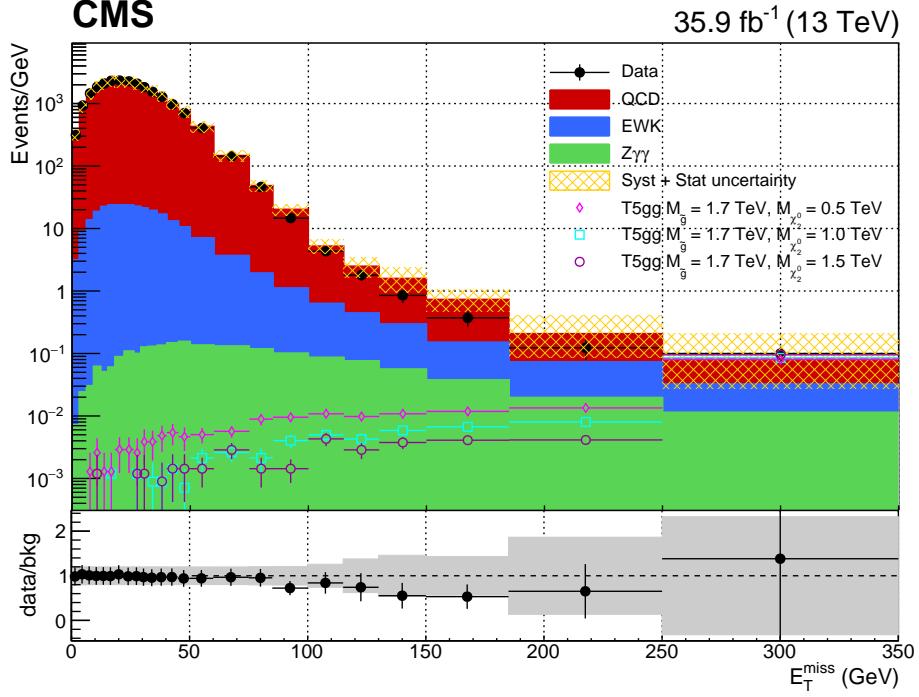


Figure 9.1: E_T^{miss} distributions for the full background estimation and the observed data. The black points represent the observed E_T^{miss} distribution. The QCD background is shown in red, the EWK background is shown in blue, and the $Z\gamma\gamma$ background is shown in green. The combined uncertainty 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 [4]. The cross sections are calculated at NLO+NLL accuracy [23, 5], with all the unconsidered sparticles assumed to be heavy and decoupled. The uncertainties on the cross sections are calculated as described in Ref. [6].

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

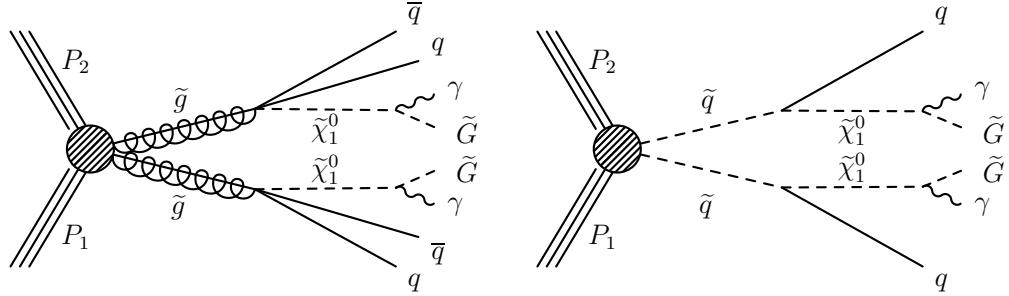


Figure 9.2: Diagrams showing the production of signal events in the collision of two protons with four momenta P_1 and P_2 . In gluino \tilde{g} pair production in the T5gg simplified model (left), the gluino decays to an antiquark \bar{q} , quark q , and neutralino $\tilde{\chi}_1^0$. In squark \tilde{q} 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 γ 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, α is the original number of events in the control sample, and β is the transfer factor, $\beta = x/\alpha$.

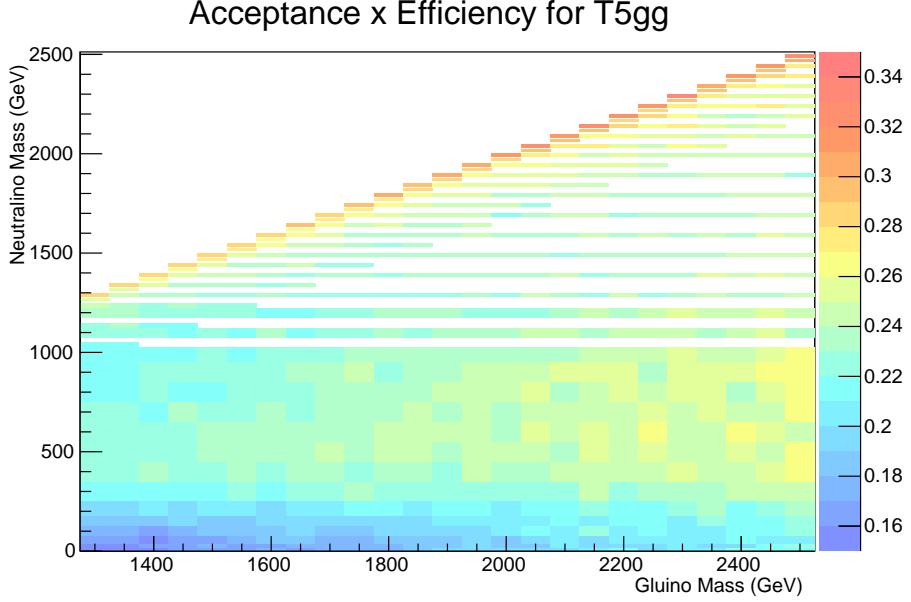


Figure 9.3: Acceptance \times efficiency for the T5gg simplified model.

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

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 CL_s method [21, 24] is applied to our analysis.

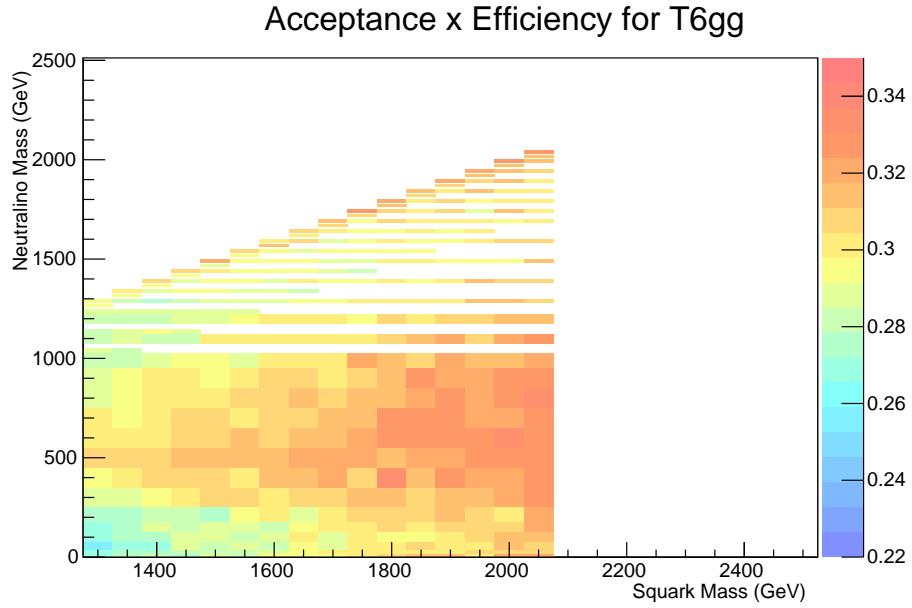


Figure 9.4: Acceptance \times efficiency for the T6gg simplified model.

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.

Acceptance \times Efficiency for T5gg

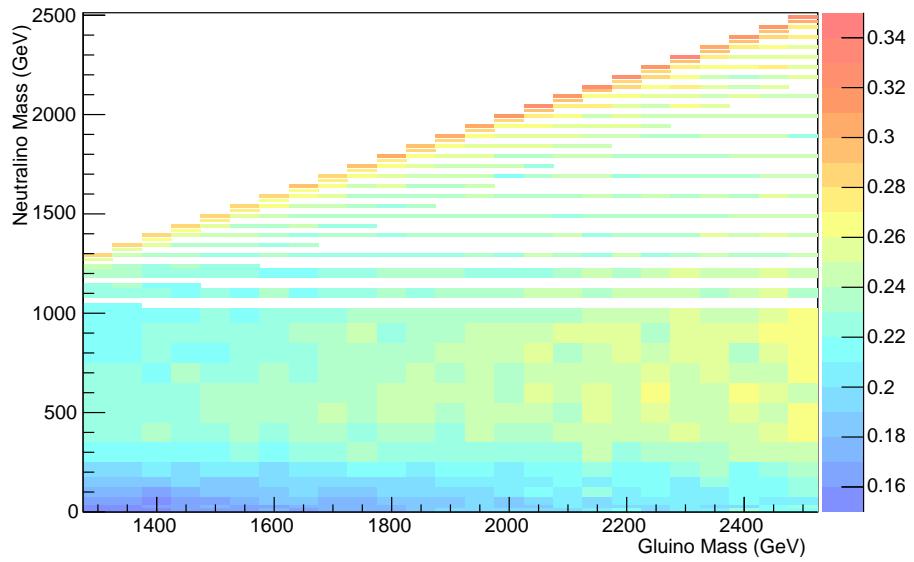


Figure 9.5: Acceptance \times efficiency for the T5gg simplified model.

Acceptance \times Efficiency for T6gg

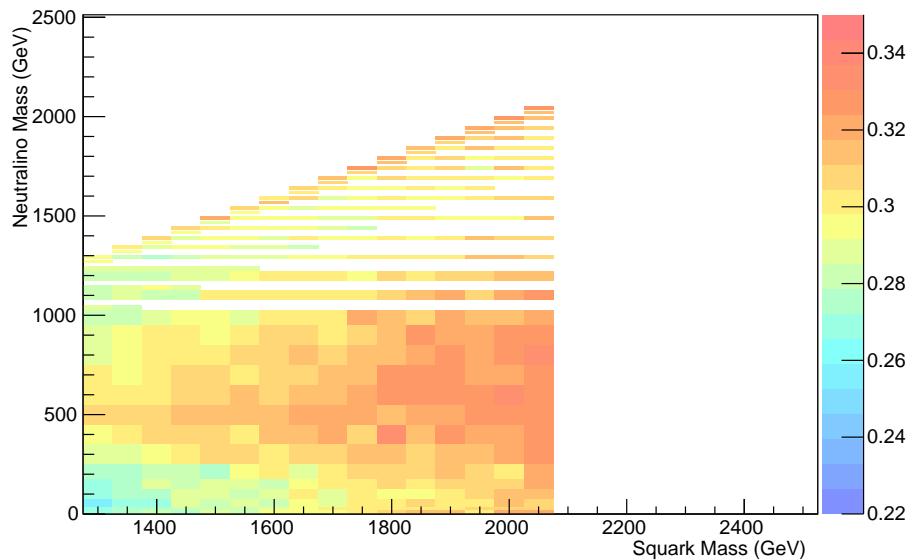


Figure 9.6: Acceptance \times efficiency for the T6gg simplified model.

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. Technical report.
2. 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.
3. 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.
4. R. D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015. doi: 10.1007/JHEP04(2015)040.
5. 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.
6. 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.
7. M. Cacciari, G. P. Salam, and G. Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 04:063, 2008. doi: 10.1088/1126-6708/2008/04/063.
8. CERN. Cern document server (cds). Technical report. URL <http://cds.cern.ch>.
9. S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.
10. CMS. Twiki. . URL <https://twiki.cern.ch/twiki/bin/view/CMSPublic/EcalDPGResultsCMSDPS2017003>.
11. CMS. Twiki. . URL <https://twiki.cern.ch/twiki/bin/view/CMSPublic/EcalDPGResultsCMSDPS2015057>.
12. CMS. Twiki. . URL <https://twiki.cern.ch/twiki/bin/view/CMS/L1TriggerPhaseIDiagrams>.

13. CMS Collaboration. Changes to CMS ECAL electronics : addendum to the Technical Design Report. CMS Technical proposal CERN-LHCC-2002-027, CMS-TDR-4-add-1, CERN, 2002. URL <http://cds.cern.ch/record/581342>.
14. CMS Collaboration. Cms luminosity measurements for the 2016 data-taking period. CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017. URL <http://cds.cern.ch/record/2257069>.
15. 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>.
16. 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>.
17. 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>.
18. 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>.
19. L. Evans and P. Bryant. Lhc machine. *Journal of Instrumentation*, 3(08):S08001, 2008. URL <http://stacks.iop.org/1748-0221/3/i=08/a=S08001>.
20. C. W. Fabjan and F. Gianotti. Calorimetry for particle physics. *Rev. Mod. Phys.*, 75:1243–1286, Oct 2003. doi: 10.1103/RevModPhys.75.1243. URL <https://link.aps.org/doi/10.1103/RevModPhys.75.1243>.
21. 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.
22. V. Khachatryan et al. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST*, 12:P02014, 2017. doi: 10.1088/1748-0221/12/02/P02014.
23. 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.
24. 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.

25. A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the cms detector. *JINST*, 12:P10003, 2017. doi: 10.1088/1748-0221/12/10/P10003.
26. 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.
27. S. Van der Meer. Calibration of the Effective Beam Height in the ISR. 1968.

This document was prepared & typeset with pdfL^AT_EX, and formatted with NDDiss2_ε classfile (v3.2013[2013/04/16]) provided by Sameer Vijay and updated by Megan Patnott.