Search for Supersymmetry with Diphoton Events in 7 TeV pp Collisions at the Compact Muon Solenoid

by

B. Heyburn

B.S., California Institute of Technology, 2005

A thesis submitted to the Faculty of the Graduate School of the University of Colorado in partial fulfillment of the requirements for the degree of Doctor of Philosophy Department of Physics

2011
This thesis entitled:
Search for Supersymmetry with Diphoton Events
in 7 TeV pp Collisions at the Compact Muon Solenoid
written by B. Heyburn
has been approved for the Department of Physics

Prof. Uriel Nauenberg

Prof. Kevin Stenson

Date ____________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.
This dissertation describes the search for General Gauge Mediated (GGM) supersymmetry breaking in 7 TeV center-of-mass proton-proton collisions from the Large Hadron Collider (LHC); the total integrated luminosity of data analyzed was $35.5 \pm 3.9 \text{ pb}^{-1}$ from the 2010 data taking. In GGM scenarios, a pair of supersymmetric particles is formed through strong production; each is expected to decay into jets, a photon, and a gravitino. Hence, searches for events with two photons and at least one jet were carried out; one observed event with missing transverse energy ($MET > 50 \text{ GeV}$) was found. This is consistent within statistical fluctuations of the predicted Standard Model background from QCD and EW sources, $1.11 \pm 0.76$ events. For one set of GGM parameters, the predicted cross-section (in leading order and next-to-leading order) is $2.309 \pm 0.413 \text{ pb}$, leading to a predicted number of events with $MET > 50 \text{ GeV}$ of $17.1 \pm 3.7$. Given the number of observed events and the expected background, an upper limit cross section of $0.558 \text{ pb}$ was set and the model point was excluded. Upper limit calculations and exclusion regions were calculated for gluino and squark masses of $400 - 2000 \text{ GeV}$ and neutralino masses of $50 \text{ GeV}$, $150 \text{ GeV}$, and $500 \text{ GeV}$.
Dedication

To Mom and Dad
Acknowledgements

I would like to acknowledge my advisor, Professor Uriel Nauenberg, who has been a tremendous support and resource during the last several years. He has always looked out for my best interest and I appreciate his confidence in me. Furthermore, I would like to thank my committee, Professors John Cumalat, Oliver DeWolfe, J. Michael Shull, and Kevin Stenson. In particular, I thank Kevin for his thorough reading of this thesis, as well as for his support and friendship while I was living at CERN. I would also like to acknowledge my analysis group at CMS, especially Michael Balazs, Rachel Yohay, Alexander Ledovskoy, and Shilei Zang from whom I have learned interesting physics and valuable skills. I thank the thousands of physicists and engineers who made the LHC and the CMS experiment possible and I want to acknowledge many useful discussions with Michael Betancourt, Freya Blekman, Mauro Dinardo, Brian Drell, Chris Edelmeier, Kevin Givens, Lauren Kost-Smith, Robynne Lock, Eduardo Luiggi, James McCollough, and Keith Ulmer. I thank Susan Spika and Jeanne Nijhowne for making my life infinitely easier.

I am lucky enough to have marvelous friends who double as brilliant scientists and engineers and I thank them for their constant inspiration: Kevin Duncklee, Vicki Dydek, Ruth Griswold, Jaclyn Homnick, Sam Lawler, Manisha Lotlikar, Karin Oberg, Jason Raycroft, and Hesper Rego. I want to thank the inventive and talented Curtis Smith for making my life one happy surprise after another. Finally, I thank my family without whom I would be homeless (literally) and mirthless (honestly). I love you Mom, Dad, Bobby, Laura, Guy, Rivers, and Grandma.
Contents

Chapter

1 Introduction ................................. 1

2 Supersymmetry and Gauge Mediated Supersymmetry Breaking .................. 4
  2.1 Introduction ........................................ 4
  2.2 Minimal Supersymmetric Standard Model (MSSM) .............................. 8
    2.2.1 Considerations when building the MSSM .................................. 8
    2.2.2 The MSSM Matter Fields ................................................. 9
    2.2.3 The MSSM Lagrangian .................................................. 11
    2.2.4 The MSSM Parameter Space ............................................. 12
  2.3 Gauge Mediated Supersymmetry Breaking (GMSB) .............................. 13
    2.3.1 Messenger Sector and Hidden Sector Interactions ........................ 14
    2.3.2 Gaugino and Scalar Masses .............................................. 15
    2.3.3 Super-Higgs mechanism and Gravitino as LSP ............................ 17
    2.3.4 Fundamental GMSB Parameters ......................................... 18
    2.3.5 Theoretical Issues ..................................................... 19
  2.4 General Gauge Mediation (GGM) ............................................. 20
  2.5 GGM Phenomenology .............................................. 21
    2.5.1 Collider Signatures: Neutralino as NLSP ............................... 22
    2.5.2 Collider Signatures: Stau as NLSP ...................................... 24
2.5.3 Benchmark Search Points .............................................................. 26
2.5.4 Current Experimental Limits ......................................................... 27
2.6 Conclusion .................................................................................. 27

3 The Compact Muon Solenoid Detector and the Large Hadron Collider 29

3.1 Large Hadron Collider ................................................................. 29
  3.1.1 LHC Layout and Overview ....................................................... 30
  3.1.2 The LHC Magnets ................................................................. 32
  3.1.3 The Vacuum System ............................................................. 34
  3.1.4 The Injection System ............................................................. 35
  3.1.5 The LHC Dumping System ..................................................... 36

3.2 The Compact Muon Solenoid ...................................................... 38
  3.2.1 The Tracking System ........................................................... 40
  3.2.2 The Electromagnetic Calorimeter ......................................... 43
  3.2.3 The Hadron Calorimeter ....................................................... 51
  3.2.4 The Superconducting Magnet ............................................. 55
  3.2.5 The Muon System ............................................................... 58
  3.2.6 The Trigger ....................................................................... 60
  3.2.7 The Data Acquisition System ............................................ 63

4 Software and Data Reconstruction ................................................. 65
  4.1 Introduction ........................................................................... 65
  4.2 Basic Kinematic Variables ....................................................... 65
  4.3 CMSSW ............................................................................... 66
  4.4 High Level Trigger ............................................................... 66
  4.5 Photon and Electron Reconstruction ....................................... 67
    4.5.1 Clustering ................................................................... 67
    4.5.2 ECAL Spikes ................................................................. 69
## Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 SM elementary particles and their properties.</td>
<td>5</td>
</tr>
<tr>
<td>2.2 SM elementary particles and their properties.</td>
<td>7</td>
</tr>
<tr>
<td>3.1 Lengths of vacuum system sectors at the LHC.</td>
<td>34</td>
</tr>
<tr>
<td>3.2 DAQ parameters</td>
<td>63</td>
</tr>
<tr>
<td>5.1 Datasets used in analysis.</td>
<td>76</td>
</tr>
<tr>
<td>5.2 Triggers used in analysis.</td>
<td>76</td>
</tr>
<tr>
<td>5.3 Final Event Count.</td>
<td>87</td>
</tr>
<tr>
<td>5.4 QCD Background Estimation and Candidate Events for ( \text{MET} &gt; 50 \text{ GeV} ).</td>
<td>94</td>
</tr>
<tr>
<td>5.5 Final Estimation of Background and Candidate Events for ( \text{MET} &gt; 50 \text{ GeV} ).</td>
<td>103</td>
</tr>
<tr>
<td>5.6 Breakdown of cuts and their effects on GGM signal efficiency.</td>
<td>107</td>
</tr>
<tr>
<td>5.7 Photon veto loss in efficiency in different material budget scenarios.</td>
<td>117</td>
</tr>
<tr>
<td>5.8 Variation of GGM Signal Efficiency Due to Jet Energy Scale Uncertainties.</td>
<td>118</td>
</tr>
<tr>
<td>5.9 Results for Upper Limit Example Point.</td>
<td>122</td>
</tr>
<tr>
<td>5.10 Critical parameters for determining upper limits and exclusions on GGM models.</td>
<td>123</td>
</tr>
<tr>
<td>6.1 Event counts for sample objects failing or passing shower shape cut.</td>
<td>132</td>
</tr>
<tr>
<td>6.2 Diphoton candidate sample composition using standard selection and loosened selection.</td>
<td>133</td>
</tr>
</tbody>
</table>
6.3 Diphoton candidate sample composition using standard selection and loosened se-
lection with additional loosened track isolation. ........................................... 133
6.4 Datasets used for MC Analysis ................................................................. 142
6.5 MC Analysis: Estimation of Background and Candidate Events for $MET > 50$ GeV 143
6.6 Composition of MC $\gamma\gamma$ sample for no jet requirement. ..................... 143
6.7 Composition of MC $\gamma\gamma$ Sample for $\geq 1$ jet. ................................. 146
Figures

Figure

2.1 Unification of gauge couplings. .................................................. 7
2.2 Schematic of GMSB scenario. ..................................................... 14
2.3 Sparticle mass Feynman diagrams. ............................................. 15
2.4 Evolution of sparticle masses from the messenger scale to the electroweak scale. ... 17
2.5 Cross sections for GGM models based on strong-production. ................. 21
2.6 GGM events at a proton-proton collider. ....................................... 23
2.7 Ratio of squark to gluino masses. ............................................... 24
2.8 Various collider GMSB experimental signatures. ................................ 25
2.9 Cross sections times branching fractions for diphoton signatures. ............ 26

3.1 The layout of the LHC including the major experiments, the injection, cleaning, RF
cavity, and dump locations. ......................................................... 31
3.2 The cross-sectional schematic of a typical LHC magnet (dipole). ............ 33
3.3 The screen for the LHC beam pipe. ............................................ 36
3.4 The CERN accelerators. .......................................................... 37
3.5 A schematic of the LHC dumping system. ..................................... 37
3.6 The shape of the beam as deposited onto the graphite and concrete absorber during
the process of beam dumping. ..................................................... 38
3.7 3D blown-apart drawing of CMS. ................................................. 40
3.8 Schematic of the silicon strip tracker and pixel system. ............................. 41
3.9 Schematic view of the CMS Pixel Detector. ................................................ 42
3.10 Diagram of ECAL arrangement. ................................................................. 45
3.11 ECAL lead tungstate crystal with two APDs (left panel) or one VPT (right panel). 47
3.12 ECAL crystal transparency as a function of time. ......................................... 49
3.13 Drawing of HB wedges. Units are in mm. .................................................. 53
3.14 Schematic of HE mounted on endcap yoke of the muon system. .................. 54
3.15 Layout of the HCAL in longitudinal view. .................................................. 55
3.16 A comparison of the energy/mass ratio versus stored energy for various detector magnets. ................................................................. 57
3.17 Drawing of the five barrel solenoid modules and support structures. ............ 57
3.18 A DT chamber in position inside the iron yoke. ......................................... 59
3.19 RPC double-gap design. ............................................................................. 60
3.20 L1 trigger decision flow. ............................................................................. 62
4.1 Schematic of ECAL clustering. ..................................................................... 68
5.1 Spike cleaning topologies. ........................................................................... 75
5.2 Cross sections in the squark-gluino mass phase space. ................................. 78
5.3 k-factors in the squark-gluino mass phase space. ......................................... 79
5.4 Number of generated jets in GGM events. ................................................... 80
5.5 ECAL and HCAL isolation cones. ................................................................. 82
5.6 Sketch of beam halo events. ........................................................................ 84
5.7 Position vs. Time of Beam Halo events. ..................................................... 85
5.8 $W\gamma\gamma$ and $Z\gamma\gamma$ Feynman diagrams. ............................................ 89
5.9 $diEM_{pT}$ spectra for $\gamma\gamma$ and $Z \rightarrow ee$ events. ............................... 90
5.10 $diEM_{pT}$ spectra for $\gamma\gamma$ and fake-fake events. .................................... 91
5.11 $diEM_{pT}$ ratio of $\gamma\gamma$ to $Z \rightarrow ee$ events. ......................................... 92
5.12 $diEMp_T$ ratio of $\gamma\gamma$ to fake-fake events. ........................................... 93
5.13 The invariant mass of two electron candidates in $Z \rightarrow ee$ sample. ........................ 97
5.14 The invariant mass of one electron candidates and one photon candidate in $e\gamma$ sample. 98
5.15 The invariant mass of two photon candidates in diphoton sample. ................................. 99
5.16 The electron/photon reconstruction efficiency as a function of pseudorapidity. ............ 99
5.17 The electron/photon reconstruction efficiency as a function of transverse momentum. 100
5.18 $MET$ Spectrum using $Z \rightarrow ee$ events as QCD background control sample. ........ 101
5.19 $MET$ Spectrum using fake-fake events as QCD background control sample. .......... 102
5.20 Event displays of diphoton candidate with $MET > 50$ GeV ................................. 103
5.21 Analysis Method Check: $e\gamma$ $MET$ Spectrum. .................................................. 105
5.22 Leading photon $E_T$ versus trailing photon $E_T$ for GGM example point. ................. 108
5.23 Fits for all tag and probe pairs (including divided by passing probes and failing probes) for $Z$ MC electrons (a) and $Z$ data electrons (b). .................. 111
5.24 The efficiency scale factor $\frac{\epsilon_{data}}{\epsilon_{MC}}$ as a function of $p_T$ (a), $\eta$ (b), and $\Delta R$ between the nearest photon (electron) and jet (c). .................................................. 112
5.25 Efficiency for electron reconstruction in data as a function of primary vertices. .... 113
5.26 Comparison of electron efficiency using Tag and Probe and electron efficiency using MC Truth. ................................................................. 114
5.27 Comparison of isolation distributions for electrons and photons. ......................... 116
5.28 Uncertainties on GGM signal efficiencies due to PDF in percent. .......... 119
5.29 GGM signal efficiencies. .......................................................... 120
5.30 Upper Limits in the squark-gluino mass plane for $m_{\tilde{\chi}_0} = 50$ GeV (a), $m_{\tilde{\chi}_0} = 150$ GeV (b), and $m_{\tilde{\chi}_0} = 500$ GeV (c). .................. 124
5.31 Uncertainties on calculated cross sections due to PDF (in percent). ................. 125
5.32 Exclusion values in the squark-gluino mass plane for $m_{\tilde{\chi}_0} = 50$ GeV (a), $m_{\tilde{\chi}_0} = 150$ GeV (b), and $m_{\tilde{\chi}_0} = 500$ GeV (c). .................. 126
5.33 Exclusion contours for the three neutralino masses. ........................................... 127
6.1 $\sigma_{in\eta}$ plot of leading vs. trailing object. ................................. 129
6.2 $\sigma_{in\eta}$ plot of leading vs. trailing photon for the diphoton sample (with jet required). 130
6.3 pFMET vs. tcMET for $\gamma\gamma$ events with $\geq 1$ jet requirement. .................. 134
6.4 QCD MET distribution from $Z \rightarrow ee$ and fake-fake samples. ......................... 135
6.5 $\eta$ distributions for leading jet in various samples. ........................................... 136
6.6 Scatter plot of $\Delta R$ between the leading photon and the jets versus the trailing photon and the jets for the $\gamma\gamma$ candidate data sample. ................................. 138
6.7 $\Delta R$ between leading generated photon and the closest jet. ............................. 139
6.8 Photon reconstruction efficiency versus $\Delta R$ for the leading generated photon in the event. ................................................................. 140
6.9 GGM acceptance versus number of generated jets. ............................................ 141
6.10 MC Analysis: MET Spectrum for $Z \rightarrow ee$ QCD control sample. .................... 144
6.11 MC Analysis: MET Spectrum for fake-fake QCD control sample. ....................... 145
6.12 Data versus MC $E_T$ distributions for QCD control samples. ............................. 147
6.13 Data versus MC $\eta$ distributions for QCD control samples. .............................. 148
6.14 Data versus MC $\phi$ distributions for QCD control samples. .............................. 149
6.15 Data versus MC ECAL Isolation distributions for QCD control samples. ................. 150
6.16 Data versus MC HCAL Isolation distributions for QCD control samples. .............. 151
6.17 Data versus MC Track Isolation distributions for QCD control samples. ............... 152
6.18 Data versus MC $H/E$ distributions for QCD control samples. ............................. 153
6.19 Data versus MC $\sigma_{in\eta}$ distributions for QCD control samples. .................... 154
6.20 Data versus MC diEM$\eta$ distributions for QCD control samples. ....................... 155
B.1 The invariant mass of two electron candidates for $|\eta| < 0.3$. ............................. 163
B.2 The invariant mass of two electron candidates for $0.3 < |\eta| < 0.6$. ..................... 164
B.3 The invariant mass of two electron candidates for $0.6 < |\eta| < 0.9$. .................... 164
B.4 The invariant mass of two electron candidates for $0.9 < |\eta| < 1.2$. .................... 165
B.5 The invariant mass of two electron candidates for $1.2 < |\eta| < 1.5$. ................. 165
B.6 The invariant mass of two electron candidates for $30 \text{ GeV} < p_T < 45 \text{ GeV}$. ................. 166
B.7 The invariant mass of two electron candidates for $45 \text{ GeV} < p_T < 60 \text{ GeV}$. ................. 166
B.8 The invariant mass of two electron candidates for $60 \text{ GeV} < p_T < 75 \text{ GeV}$. ................. 167
B.9 The invariant mass of two electron candidates for $75 \text{ GeV} < p_T$. ................. 167
B.10 The invariant mass of one electron and one photon candidate for $|\eta| < 0.3$. ................. 168
B.11 The invariant mass of one electron and one photon candidate for $0.3 < |\eta| < 0.6$. ............. 168
B.12 The invariant mass of one electron and one photon candidate for $0.6 < |\eta| < 0.9$. ............. 169
B.13 The invariant mass of one electron and one photon candidate for $0.9 < |\eta| < 1.2$. ............. 169
B.14 The invariant mass of one electron and one photon candidate for $1.2 < |\eta| < 1.5$. ............. 170
B.15 The invariant mass of one electron and one photon candidate for $30 \text{ GeV} < p_T < 45 \text{ GeV}$. ................. 170
B.16 The invariant mass of one electron and one photon candidate for $45 \text{ GeV} < p_T < 60 \text{ GeV}$. ................. 171
B.17 The invariant mass of one electron and one photon candidate for $60 \text{ GeV} < p_T < 75 \text{ GeV}$ and $p_T > 75 \text{ GeV}$. ................. 171
C.1 Leading electron vs. trailing electron distributions for $E_T$, $\eta$, and $\phi$. ................. 173
C.2 Leading photon vs. trailing photon distributions for $E_T$, $\eta$, and $\phi$. ................. 174
C.3 Leading electron or photon vs. trailing electron or photon distributions for $E_T$, $\eta$, and $\phi$. ................. 175
C.4 Leading fake vs. trailing fake distributions for $E_T$, $\eta$, and $\phi$. ................. 176
C.5 Leading electron vs. trailing electron distribution for $\sigma_{\eta\eta\eta}$, Time (ns) (middle), and Energy (GeV). ................. 177
C.6 Leading photon vs. trailing photon distributions for $\sigma_{\eta\eta\eta}$, Time (ns), and Energy (GeV). ................. 178
C.7 Leading electron or photon vs. trailing electron or photon distributions for $\sigma_{\text{ij}}$, 
Time (ns), and Energy (GeV). ................................. 179
C.8 Leading fake vs. trailing fake distributions for $\sigma_{\text{ij}}$, Time (ns), and Energy (GeV). . 180
C.9 Leading object ECAL isolation distributions for candidate and control samples. . . 181
C.10 Leading object HCAL isolation distributions for candidate and control samples. . . 182
C.11 Leading object Track isolation distributions for candidate and control samples. . . 183
C.12 Leading object Sum of isolation distributions for candidate and control samples. . . 184
C.13 Trailing object ECAL isolation distributions for candidate and control samples. . . 185
C.14 Trailing object HCAL isolation distributions for candidate and control samples. . . 186
C.15 Trailing object Track isolation distributions for candidate and control samples. . . 187
C.16 Trailing object Sum of isolation distributions for candidate and control samples. . . 188
C.17 Number of jets in candidate and control samples. .................................................. 189
C.18 Invariant mass in candidate and control samples. .................................................. 190
C.19 Number of vertices in candidate and control samples. .......................................... 191
C.20 $p_T$ of leading jet in candidate and control samples. ............................................ 192
C.21 $H_T$ in candidate and control samples. ................................................................. 193
C.22 tcMET in candidate and control samples. .............................................................. 194
C.23 $\Delta R$ between objects in candidate and control samples. ..................................... 195
C.24 $\Delta \phi$ between objects in candidate and control samples. .................................. 196
C.25 $\Delta \eta$ between objects in candidate and control samples. .................................. 197
C.26 diEM $p_T$ of objects in candidate and control samples. ....................................... 198
C.27 $\Delta \phi$ between leading jet and MET in candidate and control samples. .............. 199
C.28 $\Delta \phi$ between leading object and MET in candidate and control samples. ........... 200
Chapter 1

Introduction

The Standard Model (SM), developed in the 1960s and 1970s, has enjoyed enormous success as an encompassing theory of electroweak and strong interactions. The SM predicts the existence of the Z boson and its interactions (from the Lagrangian developed in the electroweak unification) as well as of the charm quark (due to the absence of flavor-changing neutral currents) and of the tau neutrino (after the discovery of the tau lepton). This was followed in the subsequent decades by the experimental confirmation of the existence of these particles as well as of their predicted properties. Despite the success of the SM, modern particle physicists are still challenged by remaining puzzles. For example, the Higgs boson, the last of the SM predicted particles, has not yet been discovered, and the current SM prescription contains no room for neutrino masses for which there is experimental evidence. Still, most physicists are convinced that the Higgs boson will be discovered by the experiments at the Large Hadron Collider (LHC) and the SM can easily be modified to include neutrino masses. More significant problems are highlighted through the so-called “Hierarchy Problem,” in which quadratic divergences to the Higgs mass corrections lead to enormous masses at energies of $M_{GUT} \sim 10^{16}$ GeV or $M_{Planck} \sim 10^{19}$ GeV. The Hierarchy Problem is related to issues of fine-tuning and naturalness which are philosophically distasteful. Other problems with the SM include the lack of incorporation of gravitational interactions, and the lack of an explanation of dark matter and dark energy.

Theorists have been proposing a number of solutions to these problems. One of the most popular solutions is Supersymmetry, a theory that doubles the number of elementary particles
by introducing a bosonic “superpartner” for every fermion and vice-versa. This theory, along with how it addresses issues from the SM is described in-depth in Chapter 2. Supersymmetry must be a broken symmetry or the superpartners would have already been seen, as they would have the same masses as the regular SM particles. Several symmetry breaking mechanisms have been proposed and all have different experimental predictions. In one popular model known as Gauge Mediated Supersymmetry Breaking, and its more generalized form General Gauge Mediation (GGM), supersymmetry breaking takes place in a complicated manner in which Supersymmetry is broken in a “hidden” sector and the breaking is communicated to the “visible” sector via interactions with a “messenger” sector.\(^1\) This breaking is then communicated to the normal or visible sector (where SM particles and interactions are observed) via gauge interactions. GGM models predict a certain type of event emerging from proton-proton interactions at the Large Hadron Collider (LHC). The search for experimental evidence of this breaking mechanism, also described in Chapter 2, is the subject of this dissertation.

On March 30, 2010, the LHC began colliding protons at an unprecedented center-of-mass energy of 7 TeV, and data was immediately recorded with high efficiency by the Compact Muon Solenoid (CMS) experiment. CMS, one of two general-purpose detectors built to study LHC interactions, features an almost hermetic design of all-silicon particle tracking detectors, high-resolution electromagnetic and hadronic calorimeters, a high magnetic field superconducting solenoid, and a segmented muon detector. The LHC and the CMS detector are described in Chapter 3. The synthesis of information coming from the various subdetectors is known as “data reconstruction” and is the process of taking detector signals and forming physics objects with them (photons, electrons, etc.). The process of data reconstruction is described in Chapter 4.

In GGM, the lightest supersymmetric particle is always the gravitino. The next-to-lightest supersymmetric particle is either the neutralino or the stau. Neutralinos are electrically neutral combinations of gauginos and the Higgsino. Most often, the bino-like neutralino decays immediately

\(^1\) A “hidden” or “secluded” sector is the term used to described unobserved or hypothetical quantum fields and particles (in high energy regimes) that do not interact directly with the SM particles and fields. A “messenger sector” is similar, but in this case describes a different hidden sector than the one in which supersymmetry is broken.
into a photon and gravitino (for short neutralino lifetimes). Because supersymmetric particles are created in pairs, GGM models thus predict LHC events that have two photons originating from the proton-proton interaction point, jets (hadronic objects), and missing transverse energy (carried away by undetectable gravitinos). This dissertation describes the search for these types of events using the first year of LHC data taken at 7 TeV center-of-mass energy, which is the subject of Chapter 5 and is the true heart of this dissertation. This chapter includes: 1) descriptions of the datasets used (both real data and Monte Carlo simulation used for comparison), 2) event selection and object (e.g., photon) identification, 3) determination of the expected background, 4) calculation of the expected number of signal events (using simulation), including the acceptance and efficiency of selecting those types of events and 5) calculation of the upper-limits of cross-sections of GGM type events and calculation of exclusion of certain models. Chapter 6 features various related investigations which support the choices of event selection, background determination methodology, etc. Finally, the conclusion, Chapter 7 summarizes the results and provides an outlook for the future.
Chapter 2

Supersymmetry and Gauge Mediated Supersymmetry Breaking

2.1 Introduction

The Standard Model (SM) describes the electromagnetic, weak, and strong interactions of elementary particles. These elementary particles are the quarks, leptons, and gauge bosons. The quarks and leptons are made up of three generations and the fundamental interactions between them are mediated by the gauge bosons. Their properties are given in Table 2.1. The photon is the gauge boson that mediates the electromagnetic interaction and it couples to particles with electromagnetic charge (including quarks and the charged leptons). The $W^\pm$ and Z are the gauge bosons that mediate the weak interaction and they couple to particles with weak isospin and weak hypercharge (including quarks and all leptons). The (eight) gluons are the gauge boson that mediate the strong interaction and they couple to particles with “color” (red, green, blue, and their anti-color counterparts). The only particles that have color charge are quarks and gluons and only bound states of colorless particles can exist. Combinations of all three colors form baryons while color-anticolor pairs form mesons. The strong force between quarks increases with their separation (enough to pull quark-antiquark pairs out of the vacuum), so there are no free quarks. Finally, most of the mass of the protons and neutrons comes from the strong force, but the masses of the $W$ and Z bosons, as well as of the charged leptons comes from spontaneous electroweak symmetry breaking (resulting in a nonzero vacuum expectation value). The result of this is the Higgs field.

---

1 Electromagnetic and weak interactions can be unified into one “electroweak” (EW) theory [1]. Charge (Q), weak isospin ($I_3$), and weak hypercharge (Y) are related by $Q = I_3 + \frac{Y}{2}$. $W^+$ has $I_3 = +1$, $W^-$ has $I_3 = -1$, and Z has $I_3 = 0$. 

and Higgs boson, which is the only SM particle not yet discovered and it is the subject of intense experimental investigation.

Table 2.1: SM elementary particles and their properties.

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Charge</th>
<th>Spin</th>
<th>Mass (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>up (u)</td>
<td>Quark</td>
<td>+2/3</td>
<td>1/2</td>
<td>0.002</td>
</tr>
<tr>
<td>down (d)</td>
<td>Quark</td>
<td>−1/3</td>
<td>1/2</td>
<td>0.005</td>
</tr>
<tr>
<td>charm (c)</td>
<td>Quark</td>
<td>+2/3</td>
<td>1/2</td>
<td>1.5</td>
</tr>
<tr>
<td>strange (s)</td>
<td>Quark</td>
<td>−1/3</td>
<td>1/2</td>
<td>0.1</td>
</tr>
<tr>
<td>top (t)</td>
<td>Quark</td>
<td>+2/3</td>
<td>1/2</td>
<td>173.1</td>
</tr>
<tr>
<td>bottom (b)</td>
<td>Quark</td>
<td>−1/3</td>
<td>1/2</td>
<td>5</td>
</tr>
<tr>
<td>electron (e)</td>
<td>Lepton</td>
<td>−1</td>
<td>1/2</td>
<td>0.000511</td>
</tr>
<tr>
<td>electron neutrino (ν_e)</td>
<td>Lepton</td>
<td>0</td>
<td>1/2</td>
<td>~ 0</td>
</tr>
<tr>
<td>muon (μ)</td>
<td>Lepton</td>
<td>−1</td>
<td>1/2</td>
<td>0.106</td>
</tr>
<tr>
<td>muon neutrino (ν_μ)</td>
<td>Lepton</td>
<td>0</td>
<td>1/2</td>
<td>~ 0</td>
</tr>
<tr>
<td>tau (τ)</td>
<td>Lepton</td>
<td>−1</td>
<td>1/2</td>
<td>1.777</td>
</tr>
<tr>
<td>tau neutrino (ν_τ)</td>
<td>Lepton</td>
<td>0</td>
<td>1/2</td>
<td>~ 0</td>
</tr>
<tr>
<td>photon (γ)</td>
<td>Gauge boson</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>W±</td>
<td>Gauge boson</td>
<td>±1</td>
<td>1</td>
<td>80.4</td>
</tr>
<tr>
<td>Z</td>
<td>Gauge boson</td>
<td>0</td>
<td>1</td>
<td>91.2</td>
</tr>
<tr>
<td>gluon (g)</td>
<td>Gauge boson</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Higgs (H)</td>
<td>Boson</td>
<td>0</td>
<td>0</td>
<td>114 &lt; m &lt; 215</td>
</tr>
</tbody>
</table>

In the SM, quantum corrections to the Higgs mass are computed in perturbation theory. At energies up to \( M_{GUT} \sim 10^{16} \) GeV or \( M_{Planck} \sim 10^{19} \) GeV, because the Higgs is a scalar boson, the Standard Model includes large quadratic divergences in the Higgs mass corrections which implies an enormous mass for the Higgs at these interaction energies (\( \sim 10^{32} \) to \( \sim 10^{38} \) GeV). However, in the context of electroweak symmetry breaking, the Higgs mass must be on the order of \( \sim 100 \) GeV. This is known as the “Hierarchy Problem.” Note that in the corrections to the Higgs mass, due to its scalar nature, bosons and fermions contribute with opposite signs. A solution to the quadratic divergence is found in Supersymmetry (SUSY), a proposed extension to the Standard Model. Supersymmetric models introduce a symmetry between bosons and fermions: Every ordinary particle has a “superpartner” which is identical to the original particle except that the spin
differs by $\frac{1}{2}$, and these particles are called “sparticles”.

A table summarizing the properties of sparticles is shown in Table 2.2. This doubles the number of particles, and leads to a cancellation between the sparticle and particle quadratic contributions to the (spin 0) Higgs mass corrections, keeping the Higgs mass at finite values (because boson and fermion interactions contribute with opposite signs to these corrections). Thus, now that there are even numbers of each in SUSY, the quadratic divergences cancel out. Besides giving a solution to the Hierarchy Problem, there are several inadvertent theoretical benefits of supersymmetry. For one, supersymmetric theories provide good candidates for the dark matter that has been cosmologically observed. Dark matter could be the lightest supersymmetric particle, or LSP, and is usually the lightest neutralino or gravitino (see Section 2.3 for more detail on the gravitino as LSP). Neutral LSPs are stable if R-parity is conserved (see Equation 2.2). R-parity conservation, discussed in detail later on, is imposed to increase the lifetime of the proton (since protons are not observed to decay), and it also imposes conservation of lepton and baryon number. Because proton decay involves violating both lepton and baryon number simultaneously, no renormalizable R-parity conserving coupling leads to proton decay. Additionally, with SUSY, there is a unification of gauge couplings at $\sim 10^{16}$ GeV, hinting at higher symmetries. Figure 2.1 shows such a unification of gauge couplings for a theory with SUSY-breaking at $M_{\text{GUT}}$. Such a unification does not occur with SM particles alone.

Spontaneous breaking of SUSY is necessary because otherwise the superpartners would have the same masses as ordinary particles (and thus they would have already been produced and detected). Due to theoretical constraints, in all SUSY breaking models, SUSY breaking takes place in a high-energy sector known as the “hidden” or “secluded” sector, and the breaking is transmitted to the “ordinary” sector (which is the sector that contains the usual particles and their superpartners) via a “messenger” sector (which is the sector that transmits the SUSY breaking) (see Section 2.3.1).

---

2 The naming convention is as follows: To get the name of bosonic superpartner of a fermion, an “s” is added to the beginning of the ordinary fermion name. For instance, electron $\rightarrow$ selectron, tau $\rightarrow$ stau. To get the name of the fermionic superpartner of a boson, an “-ino” is added to the end of the ordinary boson name. For instance, gluon $\rightarrow$ gluino, and W $\rightarrow$ Wino.

3 “Spontaneous breaking” means that the vacuum (no-particle) state acquires a non-zero expectation value.
Table 2.2: SM elementary particles and their properties.

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Charge</th>
<th>Spin</th>
</tr>
</thead>
<tbody>
<tr>
<td>sup ($\tilde{u}$)</td>
<td>Squark</td>
<td>$+2/3$</td>
<td>0</td>
</tr>
<tr>
<td>sdown ($\tilde{d}$)</td>
<td>Squark</td>
<td>$-1/3$</td>
<td>0</td>
</tr>
<tr>
<td>scharm ($\tilde{c}$)</td>
<td>Squark</td>
<td>$+2/3$</td>
<td>0</td>
</tr>
<tr>
<td>sstrange ($\tilde{s}$)</td>
<td>Squark</td>
<td>$-1/3$</td>
<td>0</td>
</tr>
<tr>
<td>stop ($\tilde{t}$)</td>
<td>Squark</td>
<td>$+2/3$</td>
<td>0</td>
</tr>
<tr>
<td>sbottom ($\tilde{b}$)</td>
<td>Squark</td>
<td>$-1/3$</td>
<td>0</td>
</tr>
<tr>
<td>selectron ($\tilde{e}$)</td>
<td>Slepton</td>
<td>$-1$</td>
<td>0</td>
</tr>
<tr>
<td>selectron sneutrino ($\tilde{\nu}_e$)</td>
<td>Slepton</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>smuon ($\tilde{\mu}$)</td>
<td>Slepton</td>
<td>$-1$</td>
<td>0</td>
</tr>
<tr>
<td>smuon sneutrino ($\tilde{\nu}_\mu$)</td>
<td>Slepton</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>stau ($\tilde{\tau}$)</td>
<td>Slepton</td>
<td>$-1$</td>
<td>0</td>
</tr>
<tr>
<td>stau sneutrino ($\tilde{\nu}_\tau$)</td>
<td>Slepton</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>photino ($\tilde{\gamma}$)</td>
<td>Gaugino</td>
<td>0</td>
<td>1/2</td>
</tr>
<tr>
<td>Wino ($\tilde{W}^\pm$)</td>
<td>Gaugino</td>
<td>$\pm1$</td>
<td>1/2</td>
</tr>
<tr>
<td>Zino ($\tilde{Z}$)</td>
<td>Gaugino</td>
<td>0</td>
<td>1/2</td>
</tr>
<tr>
<td>gluino ($\tilde{g}$)</td>
<td>Gaugino</td>
<td>0</td>
<td>1/2</td>
</tr>
<tr>
<td>Higgsino ($\tilde{H}$)</td>
<td>Sfermion</td>
<td>0</td>
<td>1/2</td>
</tr>
</tbody>
</table>

In Minimal Supergravity (mSUGRA) models, supersymmetry breaking is communicated to the visible sector by gravitational interactions. mSUGRA is an attractive model because, for the

Figure 2.1: Unification of gauge couplings assuming $M_{\text{GUT}}$ SUSY breaking and suitable normalization of $U(1)_Y$, as given in [2]. In the $U(1) \times SU(2) \times SU(3)$ gauge group description of the Standard Model, $\alpha_1$, $\alpha_2$ and $\alpha_3$ are the strengths of the gauge couplings.
first time, gravity plays a role in electroweak physics, and it is the most popular category of SUSY models studied. There are also gaugino-mediated and anomaly-mediated SUSY breaking theories. Gauge Mediated Supersymmetry Breaking (GMSB) is a model in which supersymmetry is broken in the messenger sector by interactions with the secluded sector, and the breaking is transmitted to the visible sector by ordinary gauge interactions. Here, the messenger sector couples directly with the hidden sector. Even though GMSB theories postpone the inclusion of gravity in electroweak physics, it makes it possible to use field theory tools entirely to describe the theory. Thus, the supersymmetry breaking parameters of the Minimal Supersymmetric Standard Model (MSSM) can be calculated, including the supersymmetric mass spectrum and distinctive phenomenological features. Additionally, without fine-tuning, GMSB predicts flavor-changing neutral current (FCNC) suppression \[3\] and a small neutron electron dipole moment (nEDM), consistent with experimental results. This chapter describes the formalism of the Minimal Supersymmetric Standard Model and Gauge Mediated Supersymmetry Breaking (including the cosmological constant problem), extension to General Gauge Mediation (GGM), phenomenology of GMSB/GGM models, and recent experimental results.

2.2 Minimal Supersymmetric Standard Model (MSSM)

2.2.1 Considerations when building the MSSM

There are a few constraints on the development of a supersymmetric theory. As discussed above, supersymmetry must be a broken theory, and the terms which break supersymmetry should not introduce quadratic divergences. Otherwise, the motivation behind supersymmetry — the solution to the hierarchy problem — becomes irrelevant. Terms that break supersymmetry but do not introduce quadratic divergences are called “softly-breaking” terms \[4\]. Furthermore, because of the SuperTrace Theorem, or General Mass Formula, supersymmetry breaking cannot be communicated to the ordinary sector by regular tree-level interactions. This formula, first developed in 1977 by
Ferrera, Girardello, and Palumbo [5], states:

\[ \sum_{J=0}^{All\ Spins} (-1)^{2J}(2J + 1)m_J^2 = 0. \] (2.1)

Here, \( J \) is the spin and \( m_J \) is the mass associated with a particle with spin \( J \). Examination of this equation shows that the sum of masses squared of bosons and fermions must be equal. However, a SUSY breaking theory could be developed which has no renormalizable tree-level couplings in the ordinary sector [4]. This is the case for GMSB (discussed later). Finally, supersymmetric theories should include the conservation of \( R \)-parity where \( R \) is defined as

\[ R = (-1)^{2J+3B+L}. \] (2.2)

Here, \( J \) is the spin, \( B \) is the baryon number, and \( L \) is the lepton number. In this way, all ordinary Standard Model particles have a value of \( R = +1 \) while supersymmetric particles have a value of \( R = -1 \). Inspection of this equation yields a number of interesting consequences. For one, both baryon number and lepton number are conserved within \( R \)-parity conservation, which suppresses proton decay. \( R \)-parity conservation also prevents mixing between Standard Model and supersymmetric particles. Furthermore, any initial state of particles is going to have \( R = +1 \) since they are produced by Standard Model particles. Therefore, any SUSY particles created will be created in pairs (as \((-1)(-1) = +1\)), and each supersymmetric particle will decay into another supersymmetric particle, until one reaches the Lightest Supersymmetric Particle (LSP), which must be stable. Stable LSPs are candidates for dark matter if they are neutral, thereby possibly providing a solution to the dark matter problem.

### 2.2.2 The MSSM Matter Fields

Each regular field in the Standard Model is promoted to a supersymmetric field. A supersymmetric field depends both on the normal space-time coordinates \((x^\mu)\) and also on \( \theta \) and \( \bar{\theta} \), which are two-component spinors and the anti-commutator \( \{\theta_a, \theta_b\} = 1 \) [6] (also known as Grassmann numbers). In this case, the fields are selected to be left-chiral scalar superfields [3]. A general
The left-chiral superfield can be written as [7]:

\[ \Phi_L(x, \theta) = \phi(x) + \sqrt{2}\psi(x)\theta + \theta\theta F(x). \]  

(2.3)

The dimension of the superfield is \([\Phi] = 1\) and the Grassmann numbers have dimension \([\theta] = -\frac{1}{2}\) (where the square brackets indicate “dimension of” the contents). Thus, the left-chiral superfield is composed of a scalar component \([\phi] = 1\), a fermionic component with \([\psi] = \frac{3}{2}\), and an auxiliary scalar component with \([F] = 2\). In the superfield, the superpartners of ordinary particles are introduced, with each superpartner having the same quantum numbers as its regular partner, but differing in spin by \(1/2\). The matter content of the MSSM is given by the following promotion of regular fields to superfields ([3]):

\[
\begin{pmatrix}
\nu_{iL} \\
e_{iL}
\end{pmatrix}
\rightarrow
\hat{L}_i \equiv
\begin{pmatrix}
\hat{\nu}_i \\
\hat{e}_i
\end{pmatrix}
\]  

(2.4)

\[
\begin{pmatrix}
u_{iL} \\
\xi_{iL}
\end{pmatrix}
\rightarrow
\hat{Q}_i \equiv
\begin{pmatrix}
\hat{u}_i \\
\hat{d}_i
\end{pmatrix}
\]  

(2.5)

\[
\begin{pmatrix}
\xi_{iL} \\
\eta_{iL}
\end{pmatrix}
\rightarrow
\hat{E}_i
\]  

(2.6)

\[
\begin{pmatrix}
u_{iR} \\
\xi_{iR}
\end{pmatrix}
\rightarrow
\hat{U}_i
\]  

(2.7)

\[
\begin{pmatrix}
u_{iR} \\
\xi_{iR}
\end{pmatrix}
\rightarrow
\hat{D}_i
\]  

(2.8)

where \(i\) runs over the three generations. So, for example:

\[ \hat{u} = \hat{u}_L(\hat{x}) + i\sqrt{2}\theta\psi_u(\hat{x}) + i\tilde{\theta}\theta L F_u(\hat{x}). \]  

(2.9)

Similarly, the Higgs doublet is also promoted to a superfield:

\[ \phi = \begin{pmatrix} \phi^+ \\ \phi^- \end{pmatrix} \rightarrow \hat{H}_u = \begin{pmatrix} \hat{h}^+_u \\ \hat{h}^0_u \end{pmatrix}. \]  

(2.10)
The vacuum expectation value of the scalar component of this doublet gives rise to the mass of up-type quarks, a second Higgs doublet gives rise to the mass of down-type quarks:

\[
\hat{H}_d = \begin{pmatrix}
\hat{h}_d^- \\
\hat{h}_d^0
\end{pmatrix}.
\]  

(2.11)

2.2.3 The MSSM Lagrangian

In the Minimal Supersymmetry Standard Model (MSSM), the gauge symmetry group is the same as that of the Standard Model: \( SU(3)_C \times SU(2)_L \times U(1)_Y \) [3], and the Lagrangian can be simply written in terms of a supersymmetric part and a SUSY-breaking part:

\[
\mathcal{L}_{\text{MSSM}} = \mathcal{L}_{\text{SUSY}} + \mathcal{L}_{\text{breaking}}
\]  

(2.12)

\( \mathcal{L}_{\text{breaking}} \) contains only terms which involve SUSY particles (thus breaking the symmetry), and the method for mediating this breaking in GMSB is discussed later on. \( \mathcal{L}_{\text{SUSY}} \) contains three fundamental components:

- **The Gauge Kinetic Component** This piece includes the kinetic terms for gauge bosons and gaugino fermions and contains terms including the gauge stress tensor, CP-violating parameters, and the gauge couplings.

- **The Matter and Higgs Kinetic Component** This piece can be divided into two parts, a kinetic term and a superpotential. The kinetic terms come from the introduction of a Kähler Potential, which is a real valued potential [3]. For instance, if \( K = \Phi_L^\dagger \Phi_L \), \( K \) is real although \( \Phi_L \) is not.

- **The Superpotential**: This component introduces the regular Standard Model Yukawa couplings as well as the mass terms for the Higgs doublets.

In order to preserve electroweak symmetry breaking, one must minimize the scalar (Higgs) potential and determine the conditions in which electroweak symmetry breaking takes place. For explicit representation and minimization for the scalar potential, see [3]. With proper minimization, one arrives at the following relationships:

\[
B_\mu = \frac{(m_{H_u}^2 + m_{H_d}^2 + 2\mu^2) \sin 2\beta}{2}
\]  

(2.13)
Here, $\mu$ and $B$ are coupling constants from the scalar Higgs potential which contains terms such as $B\mu(H_d H_u + \text{Hermitian Conjugate})$ and $\mu^2 |h_0^0|^2$. $\tan \beta$ is a useful new parameter and is defined as:

$$
\tan \beta \equiv \frac{v_u}{v_d}
$$

where $\langle h_0^0 \rangle \equiv v_u$ and $\langle h_0^0 \rangle \equiv v_d$ are the real-valued vacuum expectation values (VEVs).

If Equations 2.13 and 2.14 are satisfied, there is a well-defined minimum in the scalar potential and then electroweak symmetry breaking occurs just as in the Standard Model. Note that in most MSSM models, $\tan \beta$ and $\text{sign}(\mu)$ are free parameters and have important consequences in the phenomenology of GMSB models, discussed later on. Equation 2.14 also illustrates the “$\mu$ problem.” The supersymmetry breaking terms ($\mu$, $\tan \beta$, Higgs masses) need to be on the order of the electroweak symmetry breaking scale ($\sim M_Z$) or there is fine tuning involved to have the terms cancel out to the small $Z$ mass.

### 2.2.4 The MSSM Parameter Space

The Standard Model contains nineteen free parameters: nine fermion masses (excluding neutrinos in this context), three gauge couplings, one CP-violating phase and three mixing angles in the Cabibbo–Kobayashi–Maskawa matrix (which describes flavor-changing weak decays), $\theta_{QCD}$ (the QCD vacuum angle)$^4$, and the $\mu$ and $\lambda$ from the Higgs potential. The MSSM has far more parameters than that due to the extra matter fields, mixing angles, and phases. These can be summarized as (follows discussion in [3]):

- **Gauge Sector:** Similarly to the Standard Model, there are three gauge couplings and $\theta_{QCD}$, but add in five gaugino masses (one of the six gauginos mass can be removed by performing a chirality transformation). This gives nine free parameters.
- **Higgs Sector:** Here there are two real mass terms for the Higgs, $m_{H_u}^2$ and $m_{H_d}^2$, the $\mu$ from the superpotential and the $B$ from the soft-breaking term. The last two terms are complex, but one of the phases can be reabsorbed into the overall phase of the Higgs field. Thus, there are five free parameters here.

---

$^4$ QCD, or Quantum Chromodynamics, is the theoretical framework that describes the strong interaction.
• **Matter Sector:** There are three $3 \times 3$ complex Yukawa coupling matrices, leading to 54 parameters. This is doubled because there are another 54 that come from the matrices that are coupling-constants in the trilinear scalar interactions (interactions between the Higgs doublets and squarks and sleptons). Furthermore, there are five soft-breaking mass matrices for each scalar partner of the quarks and leptons, and each has six real parameters and three phases. This leads to 45 parameters and a total of 153 parameters.

This leads to a total of 167 free parameters in the MSSM. However, this number can be decreased by performing field redefinitions. This removes 43 parameters, leaving 124 parameters of the MSSM. This is quite a large number of free parameters, but they can be greatly reduced to a few fundamental parameters by defining the method of supersymmetry breaking. In GMSB, it will turn out that there are just six free parameters.

### 2.3 Gauge Mediated Supersymmetry Breaking (GMSB)

Supergravity mediated SUSY breaking mechanisms are probably the most studied (experimentally and theoretically) in contemporary particle physics and combines supersymmetry with gravity. In fact, a local supersymmetric model implies supergravity as it requires massive gravitons and gravitinos. In minimal supergravity models (mSUGRA), supersymmetry is broken with a super-Higgs mechanism in a “hidden” sector which is coupled to the visible (MSSM) sector through gravitational interactions. In contrast, Gauge Mediated Supersymmetry Breaking suppresses supergravity by taking the gravitino mass to be very small. This means that GMSB is therefore a global SUSY symmetry (which has implications for the cosmological constant, discussed later).

Furthermore, there is also an additional sector, the “messenger” sector. In GMSB, supersymmetry is broken in the hidden sector through a chiral superfield $\hat{X}$. Then the “messengers” couple to the hidden sector through Yukawa interactions at the tree level. The messengers then transmit this breaking to the ordinary sector (which contains the usual MSSM particles and sparticles) via the normal gauge interactions. A schematic of these processes is shown in Figure 2.2.
2.3.1 Messenger Sector and Hidden Sector Interactions

The simplest interaction between the messenger sector and the hidden sector can be given by the following addition to the MSSM superpotential:

\[
\hat{W}_M = \lambda_\ell \hat{X} \hat{\Phi}^\ell \hat{\Phi}^\ell + \lambda_q \hat{X} \hat{\Phi}^q \hat{\Phi}^q.
\]  

Here \( \hat{X} \) is the left-chiral superfield belonging to the hidden sector, and \( \hat{\Phi}^\ell \) and \( \hat{\Phi}^q \) are the superfields belonging to the lepton and quark messenger sectors. Unspecified interactions in the hidden sector let \( X \) acquire a non-zero vacuum expectation value of its scalar (\( \langle x \rangle \)) and auxiliary (\( \langle F_S \rangle \)) parts given by:

\[
\langle X \rangle = \langle x \rangle + \theta^2 \langle F_S \rangle.
\]  

Therefore, due to the tree-level Yukawa interactions, the masses of the components of the messenger are of order \( \langle x \rangle \), with splittings of order \( \sqrt{\langle F_S \rangle} \):

\[
m^2_{\tilde{\ell}_M} = | \lambda_\ell \langle x \rangle |^2 \pm | \lambda_\ell \langle F_S \rangle | 
\]  

\[
m^2_{\tilde{q}_M} = | \lambda_q \langle x \rangle |^2 \pm | \lambda_q \langle F_S \rangle | 
\]

If \( \lambda \simeq \lambda_\ell \simeq \lambda_q \), then let \( \lambda \langle x \rangle = M_M \). Also let \( \langle F_S \rangle = F_M \). \( M_M \) represents the messenger mass scale and \( F_M \) represents the extent of SUSY breaking in the messenger sector. Keeping in mind the dimensionality of \( \theta \), \([F_M] = \text{mass}^2 \). So, this is simplified to:

\[
M^2_{\Phi^+\Phi^-} = M^2_M \pm F_M.
\]
Thus, the messenger fermions and scalars have different masses. In order to have mass squared terms that are greater than zero, $M_M > \sqrt{F_M}$ and to have negligible gravitational interactions $M_M \ll M_{Planck}$. For convenience, $\Lambda = \frac{F_M}{M_M}$, which is on the order of the weak scale ($\sim 100$ GeV).

2.3.2 Gaugino and Scalar Masses

The particles and sparticles in the ordinary sector are degenerate at tree-level because they do not couple directly with $\tilde{X}$, the left-chiral superfield. Instead, they interact via gauge interactions with the messenger particles and therefore gaugino masses come from one-loop level interactions and the scalar (squark/slepton) masses arrive at the two-loop level. Such diagrams are shown in Figure 2.3.

![Feynman diagrams contributing to sparticle masses](image)

Figure 2.3: Feynman diagrams contributing to sparticle masses from [4]. Gaugino masses are via the one-loop diagrams ($\lambda$), and scalar (sfermion) masses are via the two-loop diagrams ($\tilde{f}$). Messenger fields are denoted by $\Phi$ where the fermionic components are solid lines and the scalar components are dashed lines.

For the approximation that the SUSY breaking scale is smaller than the messenger mass scale, $F_M \ll M_M^2$, the gaugino particles for each gauge group $i$ get masses, $M_i$ (as given in [3]):

$$M_i = \frac{\alpha_i}{4\pi} N \Lambda.$$  (2.21)
\( N \) is the number of messenger generations, \( \Lambda = \frac{F_{\mu}}{M_M} \), and \( \alpha_i \) are the gauge couplings. Using the same approximation as in equation 2.21 and noting that the scalar masses (supersymmetric bosons) squared scale with \( N \),

\[
m^2_i = 2N \Lambda^2 \left[ C^i_1 \frac{\alpha_1}{4\pi} \right]^2 + C^i_2 \left( \frac{\alpha_2}{4\pi} \right)^2 + C^i_3 \left( \frac{\alpha_3}{4\pi} \right)^2 \tag{2.22}\]

Note that the scalar masses are denoted by \( m_i \), as opposed to \( M_i \) for the gluino masses above. Here, \( C_i \) are different for each coupling constant term and are given by:

\[
C^i_1 = \frac{3}{5} Y_i^2 \tag{2.23a}
\]

\[
C^i_2 = \begin{cases} 
\frac{4}{9} & \text{for doublets} \\
0 & \text{for singlets} 
\end{cases} \tag{2.23b}
\]

\[
C^i_3 = \begin{cases} 
\frac{4}{3} & \text{for triplets} \\
0 & \text{for singlets} 
\end{cases} \tag{2.23c}
\]

where \( Y_i \) is the weak hypercharge. Several observations can be made. One, because both the gaugino and sfermion masses are proportional to the gauge couplings, sparticles which have color will be heavier than sparticles without color (because the strong interaction is the strongest interaction). Two, any sparticles with identical quantum numbers will have identical masses (masses do not depend on generation/flavor). This means that there is no violation of the approximate flavor symmetries, leading naturally to Flavor Changing Neutral Current (FCNC) suppression found in the Standard Model. This is a virtue over other supersymmetry breaking theories which, in principal, do not have to conserve Standard Model flavor symmetries. Finally, because the masses are acquired at the one- and two-loop level, there is no violation of Equation 2.1. The final mass spectrum is obtained by using one-loop Renormalization Group Equations to evolve these mass spectra at the \( M_M \) scale to the scale of the ordinary sector. For an example, see Figure 2.4. For more discussion on physical masses, see Section 2.5.
2.3.3 Super-Higgs mechanism and Gravitino as LSP

The Higgs mechanism occurs when there is spontaneous electroweak local symmetry breaking. Here, the gauge fields acquire a nonzero mass, where the longitudinal component is the Goldstone boson. Goldstone bosons are the massless consequences to spontaneous electroweak global symmetry breaking. The “Super-Higgs mechanism” works similarly: a massless goldstino is created when global supersymmetry is broken spontaneously. When the goldstino interacts gravitationally, and the theory is promoted to a local symmetry, the goldstino becomes absorbed in the longitudinal component of the gravitino [3] (the gravitino is analogous to the gauge fields in local gauge theories, and is the spin $\frac{3}{2}$ superpartner of the spin 2 graviton). The gravitino mass is:

$$m_{\tilde{G}} = \frac{F}{\sqrt{3} M_{\text{Planck}}}.$$  \hspace{1cm} (2.24)

$F$ is the fundamental SUSY breaking scale (in the hidden sector), as opposed to $F_M$, the breaking scale in the messenger sector. It is useful to introduce the parameter $c_{\text{grav}}$, where $c_{\text{grav}} = \frac{F}{F_M}$. Sparticle masses should be on the order of 100 GeV (for proper electroweak symmetry breaking).
This means that for relevant values of $F_M$ ($M_M \ll M_{\text{Planck}}$), $m_\tilde{G}$ is on the order of a few GeV or less. Therefore, the gravitino is the lightest supersymmetric particle (LSP), and is consequently important phenomenologically. Gravitinos can be a candidate for dark matter and they would have come from reheating after inflation or from sparticle decays. However, constraints from Big Bang Nucleosynthesis put the gravitino mass even lower — less than a keV (but not massless)[9].

Note that a particular set of MSSM parameters does not necessarily need to predict the correct relic dark matter abundance. Simple extensions to the model can lead to the correct prediction. For instance, if the prediction is too small, one can add another sector of interactions with stable neutral particles. If the prediction is too large, one can assume that $R$-parity is slightly broken so that the LSP can decay [10].

2.3.4 Fundamental GMSB Parameters

There are six fundamental GMSB parameters (far fewer than the 124 of the MSSM), all of which have previously been mentioned. What follows is a summary of each of the parameters and a discussion of the reasonable values of each.

- $M_M$: This is the messenger mass scale (units of mass). $M_M > \sqrt{F_M}$ and $M_M \ll M_{\text{Planck}}$. Here $F_M$ is the scale of SUSY breaking in the Messenger Sector (in units of mass$^2$). $F_M$ is wrapped into the definition of $\Lambda$ below.

- $\Lambda$: $\Lambda = \frac{F_M}{M_M}$, thereby, with a given value of $M_M$, defines the SUSY breaking scale in the messenger sector (units of mass). For sparticles to be on the order of the weak scale, $10 \lesssim \Lambda \lesssim 150$ TeV [3].

- $N$: $N$ is the number of generations (dimensionless). Typically, $N \leq 4$ due to gauge coupling constraints. Values larger than this are possible only if $M_M$ is larger. Note that from Equations 2.21 and 2.22, gaugino masses scale linearly with $N$ while scalar masses scale with $\sqrt{N}$. Large values of $N$ tend to indicate that the lightest stau is the Next-to-Lightest Superparticle (NLSP) rather than the lightest neutralino being the NLSP. This is because the scalar masses increase less than gaugino masses as $N$ increases.

- $\tan \beta$: $\tan \beta \equiv \frac{v_u}{v_d}$ (units are dimensionless). This is the ratio of the vacuum expectation values of the Higgs doublets. The mass of the Standard Model Higgs
can be represented by the hypotenuse of the suggested triangle. Values of \( \tan \beta \) range from 1 to 50 and large values of \( \tan \beta \) indicate the stau is the NLSP.

- **sign\((\mu)\)**: This is the sign \((\pm 1)\) of \( \mu \) from the Higgs sector (units are dimensionless).

- **\( c_{grav} \)**: \( c_{grav} = \frac{F}{F_M} \) (dimensionless). Here \( F \) is the overall SUSY breaking scale and \( F_M \) represents the SUSY breaking scale in the messenger sector, so \( c_{grav} \geq 1 \) and can have a large range. The lifetime of the NLSP, \( c_{\tau_{NLSP}} \propto c_{grav}^2 \), so this parameter also sets the NLSP lifetime, and therefore can lead to very different signatures in collider experiments.

### 2.3.5 Theoretical Issues

Note that there is a major issue with GMSB due to the necessity of fine-tuning the cosmological constant, as discussed in [11] and other sources. The vacuum energy density is given by:

\[
\langle V \rangle = F_M^2 - 3 \frac{\langle W \rangle}{M_{Planck}^2}
\]

where \( \langle W \rangle \) is the expectation value of the superpotential \( W \) at the vacuum. For the cosmological constant to be zero, these terms must cancel each other out. However, \( F_M \sim \Lambda^2 \) and \( \langle W \rangle \sim \Lambda^3 \), and \( \Lambda \ll M_{Planck} \) as discussed before, so they cannot cancel. Only with the addition of a large constant to the superpotential would the cosmological constant then be consistent with zero. While this is possible, it indicates a fine-tuning problem[12].

This is directly related to the so-called “\( \mu \) problem” (see Equation 2.14): the scale for the supersymmetry breaking terms should be at the same order of magnitude as the electroweak scale. However, \( \frac{F}{M_{Planck}} \ll M_{Weak} \), which suggests large fine-tuning of the Higgs parameters. Extensions to GMSB which address the cosmological problem also address the \( \mu \) problem [4].

In addition to these issues, there are also issues having to do with the supersymmetry breaking vacuum being metastable, and possible cosmological problems having to do with stable light gravitinos [7].
2.4 General Gauge Mediation (GGM)

There are many generalizations of the GMSB models and they are collectively known as General Gauge Mediation (GGM). In these models, there can be multiple messenger generations, multiple breaking scales, etc. GGM models are defined as those in which, in the limit that the MSSM gauge couplings \( \alpha_i \to 0 \), the theory decouples into the MSSM and a separate hidden sector. Thus, in these models, SUSY breaking is transmitted to the MSSM sector via gauge mediation only; when the gauge couplings go to zero, SUSY breaking is no longer transmitted \[13\]. GGM thus defined includes several different types of models. The first class of models are those described previously in this chapter - those in which the SUSY breaking is done in a hidden sector via left-chiral superfields \( \hat{X} \) which couple directly to superfield messengers, \( \Phi \), which then interacts with MSSM fields via the SM gauge interactions. Any given model can contain any number of messengers of left-chiral SUSY breaking superfields. The second class of models are those which exhibit direct gauge mediation — those in which the messenger fields \( \Phi \) participate in the SUSY breaking process. This includes both weakly-coupled SUSY breaking messenger fields and strongly-coupled models which may not have explicit messenger fields (but still fall under the definition of GGM above). All models exhibit flavor universality among the sfermions and gravitino as LSP. A large number of them also exhibit gaugino mass unification, a neutralino or stau NLSP, and large mass hierarchies among sfermions with different gauge quantum numbers \[13\]. GGM leaves a large parameter space but there are models which span this space without additional hidden relations \[14\].

Many of these models follow Equations 2.21 and 2.22 and hence feature heavier strongly interacting sparticles (squarks and gluinos) than the weakly interacting sparticles (sleptons, winos, etc.) \[10\]. However, some models do not have this theoretical constraint and hence the squarks and gluinos can be very light. This is attractive from an experimental point of view because it indicates that some GGM models will have large production cross sections for squarks and gluinos, and thus allow for its discovery or exclusion with early LHC data. In fact, GGM models are gaining in popularity because of this and papers illustrating different GGM benchmark scenarios have been
published [15]. A comparison of strong-production GGM cross sections at the Tevatron compared to the LHC can be found in Figure 2.5.

![Cross sections for GGM models based on strong-production at the Tevatron (√s = 1.96 TeV) and LHC (√s = 7 TeV), from [15]. Note the different scale for the Tevatron (fb) and the LHC (pb). A barn (b) is a unit of area and describes the cross sections for a scattering process. 1 b = 10^{-28} m^2.](image)

2.5 GGM Phenomenology

Many key pieces of GGM phenomenology have already been mentioned. Due to $R$-parity conservation, sparticles are always produced in pairs, and the lightest superpartner (LSP) is stable. Furthermore, in GGM, this LSP is the gravitino, a very light ($\sim$ few keV) neutral particle. The parameter $c_{grav}$ determines the lifetime of sparticles decaying into the LSP (but does not affect the masses of the sparticles). Heavier sparticles will decay down in chains to another sparticle and a regular particle (preserving $R$-parity) until finally the next-to-lightest SUSY particle (NLSP) is reached. For example, this means that at the Large Hadron Collider, this decay is possible:

$$p + p \rightarrow \tilde{S}_1 + \tilde{S}_2 + X$$

$$\rightarrow \tilde{S}_{NLSP} + X_1 + \tilde{S}_{NLSP} + X_2 + X$$

$$\rightarrow \tilde{S}_{LSP} + X_{NLSP} + \tilde{S}_{LSP} + X_{NLSP} + X$$

$$= 2\tilde{S}_{LSP} + 2X_{NLSP} + X$$ (2.26)
Here, $\tilde{S}_{1,2}$ are generic supersymmetric particles. These then cascade down to the NLSP ($\tilde{S}_{NLSP}$) and its standard model partner ($X_{1,2}$). Then the NLSP decays to the LSP ($\tilde{S}_{LSP}$) and its partner ($X_{NLSP}$). In this case, $X$ represents unspecified Standard Model particle(s). For a schematic illustrating this pair production, see Figure 2.6. For phenomenologically consistent values of GGM parameters, the NLSP is either the stau or the lightest neutralino (it can also be the sneutrino, but in a very limited number of cases [4]. Neutralinos are electrically neutral combinations of gauginos and/or the Higgsino. There are neutralinos, and the NLSP is the lightest of them ($\tilde{\chi}_1^0$) (there are also four “charginos”). There are also two staus ($\tilde{\tau}_1$ and $\tilde{\tau}_2$, $\tilde{\tau}_1 = \tilde{\tau}$ will be used here).

$$\tilde{\tau} \rightarrow \tau + \tilde{G} \quad (2.27a)$$

$$\tilde{\chi}_1^0 \rightarrow \begin{cases} 
\gamma + \tilde{G} \\
Z + \tilde{G} \\
H + \tilde{G} \\
e^+e^- + \tilde{G}
\end{cases} \quad (2.27b)$$

where $\tilde{G}$ is the gravitino. Each category of decay (neutralino or stau) has different phenomenological consequences and experimental signatures (at colliders), and will be discussed separately (with more emphasis on the neutralino as the NLSP scenario). As discussed in Section 2.3.4, large values of $N$, low values of $\Lambda$, and large values of $\tan \beta$ tend to indicate the lightest stau as the NLSP because of their effect on scalar masses and mixing.

### 2.5.1 Collider Signatures: Neutralino as NLSP

The neutralino, as discussed above, can decay into a number of Standard Model products (and missing transverse energy or $MET$) depending on the content of the neutralino.\(^5\) The branching fraction of the neutralino to a photon and gravitino is generally around 80% for most values of $\Lambda$ [4]. Although this is a large branching fraction, the equation 2.26 is not necessarily correct as

\(^5\) The “transverse” indicates that only the energy perpendicular to the beam pipe is measured. This quantity is discussed in detail in Section 4.2.
the two neutralinos do not have to decay to the same final state. Still, most searches for GGM with a neutralino NLSP search for one or two high-momentum photons since these events are most common and more easily distinguishable from Standard Model events. This signal which features two high-energy photons and $MET$ is the focus of this thesis.

It is necessary to identify possible signatures at colliders based on the NLSP lifetime. For the neutralino as NLSP case, and $c\tau \sim 0$, two high-energy photons, multiple jets (from sparticle decay chains) and $MET$ (from the gravitinos) is the experimental signature; see the top half of Figure 2.8. For intermediate lifetimes, there may be one or two photons which do not point back to the main vertex. If events are selected with two high-energy photons, both the first and last type of events are clear signatures of GGM if the $MET$ distribution is significantly in excess of the Standard Model $MET$ distribution. Major backgrounds to this process include QCD processes with many jets, some of which have a mismeasured energy, leading to fake $MET$. Other backgrounds include electroweak processes with one or more photons, and/or one or more electrons misreconstructed as photons. These events have true $MET$ (due to neutrinos). Robust analyses use data-driven methods for reducing these backgrounds.

For lifetimes that are very long, the signature contains no photons and contains only jets and $MET$, and is indistinguishable from SUGRA mediated SUSY breaking cases. It is interesting to note that if the mass spectrum of the various sparticles could be measured, it would easily
differentiate between GGM and mSUGRA scenarios. They have different spectra due to the fact that mSUGRA models feature two different scales for scalar and gaugino masses. For example, see Figure 2.7.

![Figure 2.7](image-url)

Figure 2.7: The ratio of squark to gluino masses (left) and left selectron to right selectron masses (right) as a function of the ratio of the scalar mass to gaugino mass scales. The dots are mSUGRA models while the bands are GMSB models (with $N = 1, 2, 3, 4$). Clearly the ratio of gaugino to scalar masses are fixed in GMSB and the spectra are generally different amongst the two classes of models. From [16].

Despite this, theorists have identified a case where the masses between the two models are essentially the same, and that the measurement of squark, gluino, slepton and chargino masses may not be enough to differentiate between the two models [16]. Still, the parameters that led to these same masses are very different, and additional measurements could easily rule out one of the models.

2.5.2 Collider Signatures: Stau as NLSP

For regions where the stau is the NLSP, there are different possibilities for decay signatures. For cases where the zino is heavier than the stau, but lighter than other tau-like sleptons, then the decay $\tilde{Z} \rightarrow \tilde{\tau}\tau$ occurs. If, however, the Zino is heavier than other sleptons, then $\tilde{Z} \rightarrow \tilde{l}l$ can occur,
where $\tilde{l}$ is either a right-handed selectron or right-handed smuon. In the case where the selectron or smuon are heavier than the $m_{\tilde{\tau}} + m_{\tau}$, then the slepton would cascade to the stau via a three body decay (for example, $m\tilde{e} \rightarrow e\tau\tilde{\tau}$), which would decay to a tau and gravitino. However, in the case where this is not true, the slepton would then decay directly to the LSP.

The bottom half of Figure 2.8 shows various detector signatures for the case where the stau is the NLSP. When $c\tau \sim 0$, there are two charged leptons (with curved tracks in a magnetic field) and $MET$ from the gravitinos. For long lifetimes, the only signal is two heavy charged particles. For intermediate times, the signal would show up as “kinked” tracks due to the stau decaying to the tau, and $MET$ from the gravitino.

Figure 2.8: Different collider experimental signatures for different NLSP types and decay times. See text for more in-depth discussion.
2.5.3 Benchmark Search Points

Reference [15] outlines a number of benchmarks for the search for GGM at the LHC for different types of neutralino NLSPs. It points out that for photino-like NLSP neutralinos, the dominant production method comes from two sources. This is the relevant NLSP for this analysis because they decay to either a photon or Z with a 0.77 and 0.23 branching ratio (Br), respectively. The first production method is gluino pair production ($\tilde{g}\tilde{g}$) from minimal GGM models; here the cross-section ($\sigma$) is determined by the gluino mass. The second is from gluino-squark ($\tilde{g}\tilde{q}$) and squark-squark ($\tilde{q}\tilde{q}$) production from Gauge Mediation with Split Messengers (GMSM) theories, one of a class of GGM type theories. For $\sigma \times Br$ see Figure 2.9.

![Figure 2.9: Cross sections times branching ratios ($\sigma \times Br$) leading to diphoton signatures at the LHC for GGM models based on strong-production, from [15].](image)

The main free parameters in GGM models are the gaugino and the scale (squark) masses [17] and so typical Monte Carlo production focuses on these parameter spaces. For specific information about the Monte Carlo simulation used in this analysis, see Section 5.2.1.
2.5.4 Current Experimental Limits

Until the recent LHC era, GGM had been primarily experimentally studied from the minimal GMSB viewpoint. The best limit on GMSB for the neutralino as NLSP channel comes from analyzing $\tilde{\chi}_1^0 \rightarrow \gamma + \tilde{G}$ channel (yielding two photons and missing energy). The limit is provided by the D0 experiment at the Tevatron collider [18]. With proton-antiproton collisions at a center of mass energy of 1.96 TeV and $6.3 \pm 0.4 \text{ fb}^{-1}$ worth of data. They found no evidence for GMSB and set a lower limit of 175 GeV on the $\tilde{\chi}_1^0$ mass for $\tau_{NLSP} \ll 1 \text{ ns}$.

The best limit on GMSB for the stau as the NLSP comes from analyzing a channel with two or three tau leptons at high momentum and missing energy. The limit is provided by the OPAL experiment at the LEP collider [19]. With electron-positron collisions at a center of mass energy of 209 GeV, the OPAL collaboration determined that there was no evidence for GMSB and set a lower limit of 87.4 GeV on the $\tilde{\tau}$ mass (for all NLSP lifetimes) by scanning the parameter space.

As for GGM scenarios, ATLAS (another all-purpose detector at the LHC) recently published a result with 3 pb$^{-1}$ of data and found no excess high-\textit{MET} events over the SM prediction [20]. The results were not interpreted within the context of a GGM model, but rather one of Universal Extra Dimensions (UED) which predicts a similar signature. It is safe to say that this analysis, and the corresponding CMS paper puts a far more stringent limit on GGM models given that over 10 times the amount of data is used in the analysis.

2.6 Conclusion

GMSB (and GGM) provides for interesting new physics beyond the Standard Model and is very theoretically motivated. However, despite all the theoretical benefits of GMSB there has been no direct evidence for GMSB, or for any supersymmetric theory. Furthermore, this discussion has been based on a large number of assumptions about physics at the scale of SUSY breaking and beyond and there are also a number of theoretical issues (as discussed). This has led to the development of many non-minimal GMSB models which include things like: non-gauge corrections to the
Higgs masses, non-SU(5) multiplet messengers, messenger threshold corrections, gauge multiplet messengers, non-zero D-term vacuum expectation values leading to non-zero contributions to scalar masses, and strongly-coupled messengers. For a good review, see [21].

Now that the Large Hadron Collider has turned on, physicists can tests predictions of GGM models over a very large phase space, as well as numerous other Beyond the Standard Model ideas, hopefully illuminating a number of these theoretical issues and providing for an exciting new era of particle physics.
Chapter 3

The Compact Muon Solenoid Detector and the Large Hadron Collider

The Large Hadron Collider (LHC) is the highest energy collider built to date. The Compact Muon Solenoid (CMS) Experiment is a general-purpose experiment designed to study as many physics processes as possible, but particularly Higgs and Beyond the Standard Model searches. This chapter describes the LHC and CMS, including the various subdetectors.

3.1 Large Hadron Collider

The Large Hadron Collider was designed to provide proton-proton collisions at a center-of-mass energy ($\sqrt{s}$) of 14 TeV at luminosities of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. However, due to issues with the accelerator, during 2010 it operated at 7 TeV. The instantaneous luminosity during the period of data taking ranged from $10^{29} \text{ cm}^{-2}\text{s}^{-1}$ to $10^{32} \text{ cm}^{-2}\text{s}^{-1}$. The LHC is also capable of providing lead (Pb) heavy ion collisions at 2.8 TeV per nucleon at $10^{27} \text{ cm}^{-2}\text{s}^{-1}$. This section on the LHC describes the layout of the LHC collider, LHC performance, information about the magnets, the vacuum system, the injection system, and the beam dumping system. The entire system is supported by well-developed beam instrumentation and control system tools which will not be discussed here in the interest of space, but the inquiring reader is referred to [22]. Unless otherwise noted, the figures and results come from [22] or the LHC website.
3.1.1 LHC Layout and Overview

The LHC was built inside the previously existing tunnel used by the Large Electron Positron (LEP) collider, a 26.7 km long tunnel that lies underneath the French-Swiss border, and is operated by the European Organization for Nuclear Research (CERN). The tunnel is located between 45 m and 170 m underground at an incline of 1.4%. There are eight straight sections (∼130 m each) and eight arcing sections (∼2.45 km each). The straight sections were built for LEP because of the large synchrotron radiation losses when accelerating (bending) electrons. These straight sections are not needed for the LHC but the re-use of the tunnel obviously requires their presence. The eight straight sections provide for eight potential interaction points; however, only four interaction points are used at the LHC in order to minimize beam disruption. Of the four LHC experiments, two are general-purpose hermetic experiments designed to take advantage of the high luminosity the LHC provides for proton-proton collisions. These are the Compact Muon Solenoid (CMS) experiment and the A Toroidal LHC Apparatus (ATLAS) experiment. A third experiment, LHCb, is designed specifically for B physics and is operated at a lower luminosity. Finally, the fourth experiment, A Large Ion Collider Experiment (ALICE) is designed for heavy ion physics for the Pb-Pb collisions the LHC is also capable of providing. A layout of the LHC including the four interaction points is given in Figure 3.1. The LHC receives its protons (at 450 GeV per beam) from the Proton Synchrotron (PS) combined with the Super Proton Synchrotron (SPS).

3.1.1.1 LHC Performance

The number of events expected for a given physical process is:

\[ N = L\sigma \]  

(3.1)

where \( \sigma \) is the cross-section of the physical process (usually given in pb, or \( 10^{-36} \text{ cm}^2 \)) and \( L \) is the luminosity (given in \( \text{pb}^{-1} \)). As mentioned above, the LHC is designed to deliver proton-proton collisions at a luminosity of \( L = 10^{34} \text{ cm}^{-2}\text{s}^{-1} \); note that this luminosity requirement prevents the use of antiprotons (due to difficulty in producing large numbers of them). The luminosity is given
Figure 3.1: The layout of the LHC including the major experiments, the injection, cleaning, RF cavity, and dump locations.

by:

\[
L = \frac{N_b^2 n_b f \gamma F}{A}
\]  

(3.2)

where \(N_b\) is the number of protons per bunch, \(n_b\) is the number of bunches in the ring, \(f\) is the frequency of revolution of the bunches, \(\gamma\) is the Lorentz factor, \(F\) is a unitless geometric reduction factor that takes into account the angle of crossing at the interaction point (IP), and \(A\) is the cross-sectional area of the beam that takes into account the beam emittance (the extent of space occupied by particles in the beam). For the given design luminosity, and constraints due to the mechanical beam screen, magnetic alignment and geometry, and the minimum acceptable aperture (in terms of RMS of beam widths), the nominal beam size is 1.2 mm. Interacting beams experience a non-linear interaction that results in a shift of protons with respect to the nominal beam. Taking
into account the beam size and this shift, the maximum number of protons in a bunch is limited to $N_b = 10^{11}$. A bunch spacing of 25 ns (the space between each bunch of protons) corresponds to 2808 bunches in the 26.7 km ring (including bunches intentionally left empty). There are a number of considerations that are taken into account when determining these performance capabilities in limitations and they include: beam instabilities, maximum magnetic field and quench limits, the heat load, the luminosity lifetime, and the large energies stored within the beam (that require careful dumping systems, see 3.1.5). As an aside, note that the luminosity lifetime is expected to be 15 hours. Considering the design luminosity, the turnaround time of the LHC (up to 7 hours), and the expected number of running days per year (200), this means that if there are no major issues, the LHC can theoretically deliver 80-120 fb$^{-1}$ of data per year.

Unlike particle-antiparticle colliders which can use the same set of magnets and vacuum chambers, particle-particle colliders must have two separate rings with counter rotating beams. The LHC tunnel in the arcing sections has an internal diameter of only 3.7 m. Because of the small tunnel diameter, a twin-bore magnet design was chosen in which the separate magnetic coils and beam pipes use the same mechanical structure and cryostat.

### 3.1.2 The LHC Magnets

There are over 50 different kinds of magnets at use in the LHC. The magnets all use the twin-bore design in which there are separate magnetic coils and beam pipes but the same mechanical structure and cryostat. The twin-bore design has the benefit of being physically small and relatively inexpensive. However, the rings are magnetically coupled to each other which reduces the flexibility of the collider.

These “cryomagnets” all use supercooled (to 2 K by superfluid helium) Niobium-Titanium (NbTi) cables to create the magnetic fields. Compared to other colliders, which have operated magnets at 4 – 5 K, the heat capacity in the cables is drastically reduced. Thus, for a given temperature margin (the difference between the operating temperature and the quench temperature), only a small amount of energy can be deposited in the cable before a quench will occur. Therefore,
that temperature margin must be kept as high as possible. Furthermore, since the stress from the Lorentz force increases with the square of the magnetic field (which is higher than previous collider magnets), the surrounding mechanical structures must be much stronger than those used in previous colliders.

Each magnet contains a “cold mass” which is the core of the magnet and contains all materials which are cooled by the superfluid helium, including an iron yoke, the beam pipes, and the magnetic coils. The cold mass is surrounded by insulation, thermal shields, and mechanical support. For a schematic view of the cross section of a typical magnet, see Figure 3.2.

![Figure 3.2: The cross-sectional schematic of a typical LHC magnet (dipole).](image)

The two most important types of magnets are the dipole magnets (which bend the proton bunches around the arcs) and the quadrupole magnets (which focus the beam), and the “inner triplet” magnets (which steer the beam at collision points). Each of the eight LHC arcs contains 154 dipole magnets which bend the path of the protons, and operate at a peak magnetic field of 8.33
T (leading to the maximum 7 TeV energy per proton beam). There are 858 quadrupole magnets which focus the beam by narrowing the width and height of the beam in sequential stages. Other types of multipoles also help to focus the beam in addition to counteracting negative effects due to gravity, electromagnetic interactions between bunches, and other beam instabilities. Finally, there are the inner triplet magnets which are carefully designed to focus and steer the two beams into each other. These sets of magnets are only located in the LHC sections with experiments.

3.1.3 The Vacuum System

There are three separate vacuum systems for the LHC — one is an insulation vacuum for the cryomagnets described above, one is the insulation vacuum for helium distribution, and one is for the beam pipe itself. For the first two vacuum systems, the pressure is designed to be $10^{-4}$ bar at room temperature and $10^{-9}$ bar when cooled down to cryogenic temperatures. The vacuum system for the beam pipe is much more stringent because the beam lifetime and non-beam backgrounds at the experiments are significantly affected by gas within the beam pipe. At room temperature, the vacuum system is designed to be at $10^{-13}$ bar. At cryogenic temperatures, the pressure is instead quoted in gas densities (normalized to hydrogen), and is expected to be less than $10^{15} \text{ H}_2 \text{ m}^{-3}$ in normal areas of the LHC, and less than $10^{13} \text{ H}_2 \text{ m}^{-3}$ in areas around the LHC experiments (to minimize backgrounds). The vacuum systems are divided into manageable lengths, given in Table 3.1.

<table>
<thead>
<tr>
<th>Vacuum Type</th>
<th>Section Length</th>
<th>Total Sections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insulation (cryomagnets)</td>
<td>214 m</td>
<td>112</td>
</tr>
<tr>
<td>Insulation (helium distribution)</td>
<td>428 m</td>
<td>56</td>
</tr>
<tr>
<td>Beam Pipe</td>
<td>Various (up to 2900 m)</td>
<td>278</td>
</tr>
</tbody>
</table>

Table 3.1: Lengths of vacuum system sectors at the LHC.
3.1.3.1 Insulation Vacuum Systems

The insulation vacuum system designs, one for the cryomagnets and one for the helium distribution, are most significantly affected by the large volume needed at vacuum (80 m$^3$ per section for the cryomagnets, 85 m$^3$ for the helium distribution) and by the multiple layers of insulating material (200 layers for the cryomagnets, 140 for the helium distribution). This causes a large amount of gas load on the system which requires high capacity pumping systems and a good strategy for leak detection so that the volumes can be pumped down in an acceptable time. The two systems are typically separate but can be combined in the longitudinal direction. The configuration of the system is designed such that any one LHC cell (two dipoles and two quadrupoles) can be individually brought up to room temperature.

3.1.3.2 Beam Pipe Vacuum Systems

The beam pipe vacuum design is affected by beam instabilities, the background conditions at the interaction points, and by various heat inducing phenomena including synchrotron radiation and energy loss from nuclear scattering. The heat phenomena affect the vacuum design because the cryogenic elements of the LHC must be kept cold — hence a beam screen was developed which allows gas to condense on surfaces where they are protected from collisions with more energetic particles as seen in Figure 3.3. The vacuum lifetime itself is most strongly affected by nuclear scattering of protons on residual gas. To further reduce the presence of beam gas, the beam pipe is baked using heating tape and heating jackets so that every component reaches at least 250 $^\circ$C causing gases to separate from the beam pipe walls so that they may be removed by the vacuum pump system.

3.1.4 The Injection System

The process of getting protons into the LHC involves the use of several smaller accelerators as seen in Figure 3.4. The essential idea is that the protons are increasingly accelerated using previous CERN accelerators until they reach an energy of 450 GeV at which point they are injected into the
LHC at two LHC sections (two and eight). A transfer line brings the protons to within 150 m of the injection sites at these two points, and the beam approaches from underneath and outside the main LHC rings. Five “septum” magnets (which apply magnetic fields in short pulses) kick the beam horizontally into place underneath the main rings of the LHC. Then a series of four “kicker” magnets deflect the beam vertically into the closed-orbit LHC main rings. After the bunches are in the main rings, a series of focusing quadrupole magnets shapes the beam.

### 3.1.5 The LHC Dumping System

The LHC beam current at design luminosity is 0.584 A which translates to a stored energy in the beam of 362 MJ. Combined with the stored energy in the LHC magnets, 600 MJ, there is over 1 GJ of energy that must be safely and reliably dumped when required. The beam dumping system is located at Octant 6, as seen in Figure 3.1, and a schematic of the dumping system is shown in Figure 3.5. There are 15 kicker magnets which kick the beam horizontally into the septum magnets, which provide a vertical deflection (above the LHC cryostat). Afterward come the “dilution” kickers which sweep the beam into an “e” shape, which will eventually strike the
Figure 3.4: The CERN accelerators. The LHC proton injection comes from the SPS and PS, which are in turn fed by smaller linear accelerators.

absorber located 750 m away from the interaction point at section 6 (see 3.6). The absorber is a cylinder of graphite that is 1 m in diameters and 8 m long, which is encased in concrete. Such a dump can withstand the energy deposited by the dumped beam.

Figure 3.5: A schematic of the LHC dumping system. Bending magnets are labeled by “B” and focusing quadrupole magnets are labeled by “Q.” Distances are in meters.
Figure 3.6: The shape of the beam as deposited onto the graphite and concrete absorber during the process of beam dumping.

3.2 The Compact Muon Solenoid

The design of the LHC is motivated by the desire to study physics at the TeV scale including studies of electroweak symmetry breaking (presumed to be the Higgs mechanism), theories beyond the Standard Model, and studies of the strong interaction at extreme densities, temperatures and parton momenta. The CMS detector is, in turn, designed to meet the goals of the LHC physics program. Furthermore, at the design luminosity of the LHC ($10^{34} \text{ cm}^{-2}\text{s}^{-1}$), approximately 1000 charged particles will emerge from the interaction point every 25 ns and a mean of 20 inelastic collisions are superimposed on every event of interest. In order to reduce the effects of this “pileup”, CMS design requirements included highly granular detectors with excellent time resolution. This results in low occupancy (a small amount of signal per unit time per detector unit), allowing the detectors to make accurate and precise measurements. Furthermore, the resulting large number of channels indicates that CMS must also have good time synchronization among the detector elements.

The LHC physics program also motivates a number of detector-specific design requirements. First, CMS requires good muon identification and resolution over a wide range of momenta and angles, and must possess the ability to determine the charge of the muon at momenta less than 1 TeV. Second, the tracker must be designed with good momentum resolution and reconstruction
efficiency. The desire for $\tau$ and $b$-jet tagging requires a very high resolution pixel detector close to the interaction point in order to identify secondary vertices. Third, the electromagnetic calorimeter must have good energy resolution as well as good diphoton and dielectron mass resolution; it must cover a wide geometric range (nearly hermetic), allow for $\pi^0$ rejection and provide for efficient photon and lepton isolation at high energies. Finally, for good missing transverse energy and dijet resolution, CMS also requires a hermetic hadron calorimeter with fine segmentation. As an aside, because design and construction of CMS was started long before the experimental cavern was available, CMS is a modular detector.

The CMS detector is 22 m long, 15 m wide and weighs approximately 14000 tons. It is installed at Point 5 along the LHC ring, approximately 100 meters underground and close to the town of Cessy, France. From the inside out, CMS is composed of: the tracker (with high-resolution inner pixel tracking system), the nearly hermetic Electromagnetic Calorimeter (ECAL), and the nearly hermetic Hadronic Calorimeter (HCAL), the solenoid magnet, and finally the muon system (interspersed with the iron return yoke for the solenoid) (see Figure 3.7).

The coordinate system of CMS is oriented about the origin which is centered at the nominal collision point. The axes are defined as follows: $\hat{z}$ is along the beam line toward the Jura mountains from P5, $\hat{y}$ is upward toward the surface of the earth, and $\hat{x}$ is toward the center of the LHC ring. Alternative coordinates are often used: $\hat{r}$ is the distance from the beam line, $\hat{\phi}$ is the angle sweeping a circle around the beam pipe, and $\eta$ is the “pseudorapidity” defined as:

$$\eta = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right] \quad (3.3)$$

where $\theta$ is the angle from the beam line. Thus, a particle traveling perpendicular to the beam direction ($\theta = \frac{\pi}{2}$) has $\eta = 0$ while a particle traveling down the beam line ($\theta = 0$) has $\eta = \infty$. In the high momentum limit, this reduces to the rapidity:

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

where $E$ is the energy of the particle and $p_L$ is the momentum component parallel to the beam line.
This chapter is divided into subsections concerning each subdetector; because the analysis in this thesis primarily uses photons, the emphasis is on the electromagnetic calorimeter (ECAL). Unless otherwise referenced, results are obtained from Reference [23].

3.2.1 The Tracking System

The tracking system is designed to provide excellent measurements of the trajectories of charged particles coming from the interaction point. The tracker is composed of over 200 m$^2$ of active silicon sensors, making it the largest detector of its kind. Three major considerations are paid attention to during the design of the detector: the need for cooling, the need for radiation hardness, and the need for a low material budget. The inner tracking system of the CMS detector must be very fast and very granular because of the high LHC particle fluence. The fast and granular nature of the detector suggests high power density which in turn requires efficient cooling. The cooling in the tracker system is done with perfluorohexane (C$_6$F$_{14}$) coolant, which has a very low viscosity,
is good under irradiation, and is very volatile. The tracking system is kept at approximately \(-10\) °C. Finally, due to its proximity to the interaction point, the tracker must also be very radiation hard. Radiation damage comes in several forms including bulk and surface damage to the silicon, and transient phenomena. Radiation damage increases leakage current which leads to self-heating, thereby enforcing the need for efficient cooling. Furthermore, because of the inevitable radiation damage, the tracking system is designed with easy access for replacement in mind (particularly for the pixels, where the cabling allows for yearly access if needed). Finally, the material budget of the tracker is kept as low as possible because of the adverse effects of multiple scattering on tracking resolution, as well as the adverse effects of bremsstrahlung and photon conversion on the ability of the calorimeters to measure electron and photon energies.

The tracking system is composed of two distinct parts: the pixel detector and the strip tracker as depicted in Figure 3.8.

Figure 3.8: Schematic of the silicon strip tracker and pixel system. For acronyms, see text or glossary.
3.2.1.1 The Pixel Detector

The CMS Pixel Detector is the innermost subdetector of the experiment, covering a pseudo-rapidity of $|\eta| < 2.5$. The high resolution allows it to track particles with extreme accuracy thereby providing for precise measurement of secondary vertices and impact parameters. It consists of two major components, the barrel pixels (BPIX) and the forward pixels (FPIX). The BPIX consists of 768 modules (varying-sized arrays of Read Out Chips, or ROCs), arranged into three layers at radii of 4.3 cm, 7.2 cm, and 11.0 cm. The FPIX has 672 modules arranged into two disks on either end of the BPIX, 34.5 cm and 46.5 cm from the middle of the BPIX (see Figure 3.9). The geometry allows for at least three tracking points over the full $\eta$ range. Each pixel sensor is bump-bonded to the ROC. There are 52 pixel sensors per row and 80 pixel sensors per column, for a total of 4160 per ROC. With 15,840 ROCs, there are a total of 66 million pixels [24][25] which results in an occupancy of less than 1% per event. The pixel detector features automatic zero suppression and programmable thresholds for each individual pixel. Each silicon pixel sensor is 100 $\mu$m $\times$ 150 $\mu$m. However, due to the analog read out, the resolution is even finer at 15 – 20 $\mu$m. This is because of the effect of charge-sharing: the 4 T magnetic field causes released charge to drift in the BPIX leading to sharing amongst neighboring pixels. There is also a charge-drift effect for the FPIX and the FPIX sensors are purposefully tilted with respect to incoming tracks to exploit charge sharing.

Figure 3.9: Schematic view of the CMS Pixel Detector with three barrel pixel (BPIX) layers and two forward endcaps (FPIX) on either end.
3.2.1.2 The Strip Tracker

The strip tracker is composed of four main parts: the Tracker Inner Barrel (TIB, ten layers), the Tracker Inner Disks (TID, three disks per side), the Tracker Outer Barrel (TOB, six layers) and the Tracker End Cap (TEC, nine disks per side). Like the pixel system, it covers the pseudorapidity region of $|\eta| < 2.5$. For the TIB and the TID, which extend radially from 20 cm to 55 cm, the strips are 10 cm by 90 $\mu$m and have an occupancy of $2 - 3\%$. For the TOB, which extends from $r = 55$ cm to $r = 110$ cm, the strips have a dimension of 25 cm by 180 $\mu$m. These strip dimensions are the same dimensions as for the TEC, which extends radially from 22.5 cm to 113.5 cm. Because of the smaller particle fluence at these farther radii, the tracker can afford to reduce the number of read-out channels. These longer strips have an increased capacitance which increases the noise. The pitch of the strips ranges from 80 $\mu$m to 183 $\mu$m. For increased resolution, 40% of the outer layers are double layered with a “stereo” angle of 5.7°. The geometry of the strip tracker ensures at least 9 hits (and an average of 14 hits) in the tracker for $|\eta| < 2.4$. The tracker itself is supported by a tube which is suspended from the HCAL barrel.

Finally, it is important to note that the neighboring electromagnetic calorimeter (ECAL) must be operated at +18 °C, whereas the tracker is operated at −10 °C. In order to achieve this thermal gradient over a short distance, an active thermal screen is employed which features feedback control. This allows both detectors to be operated at the necessary temperatures and also prevents thermal stress on the structural support.

3.2.2 The Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) is a homogeneous, hermetic detector meant to provide full coverage for accurate measurement of missing transverse energy ($MET$) and photon and electron energy. Design requirements include the need to be fast, to be radiation hard, and to have fine granularity. It is composed of 61200 lead tungstate ($PbWO_4$) crystals in the barrel section (covering a pseudorapidity of $|\eta| < 1.479$) and 7324 crystals in each of two endcaps (covering
a pseudorapidity of $1.479 < |\eta| < 3.0$). Lead tungstate is chosen because of its high density (8.28 g/cm³), short radiation length (0.89 cm) and small Molière radius (2.2 cm), thus ensuring scintillation can take place in a small area without leakage to neighboring crystals.

### 3.2.2.1 Composition and Geometry

The ECAL Barrel (EB) is segmented into 360 crystals in $\phi$ by 170 in $\eta$. The crystals are grouped into submodules ($5 \times 2$ modules), modules (of different types, 400-500 crystals), and, finally, supermodules. The supermodules are 20 crystals in $\phi$ by 85 crystals in $\eta$, spanning a 20° arc. There are 17 different types of crystals in the EB, which vary according to the location along $\eta$ of the crystal. Each crystal covers approximately $0.0174 \times 0.0174$ in $\Delta \eta \times \Delta \phi$ and is 23 cm long. Each crystal is tapered slightly (484 mm² at the front face and 676 mm² at the rear face); this allows the front and rear faces of each crystal to face the interaction point, allowing for a precise energy measurement.\(^1\)

The ECAL Endcap (EE) contains crystals of identical shape (819.10 mm² at the front face and 900 mm² at the rear face) arranged into $5 \times 5$ crystals called a supercrystal. Unlike the crystals in the EB, these have a non-pointing geometry. They are instead pointing at a focus 1300 mm beyond the interaction point. Each crystal is tilted lengthwise at an angle of $(2 - 8)^\circ$ with respect to the vector formed by pointing from the crystal face to the interaction point and covers an area ranging from $0.0175 \times 0.0175$ to $0.05 \times 0.05$ in $\Delta \eta \times \Delta \phi$.

The ECAL also contains a pre-shower intended to improve the photon $\pi^0$ separation and covers $1.653 < \eta < 2.6$. The preshower consists of lead absorbers which cause hadronic particles to shower over 2-3 radiation lengths, and two layers of silicon strip detectors. The preshower uses silicon detectors in this instance because of the compact shape and high detection efficiency. It allows for the separation of $\pi^0$ and photon signals before reaching the EE. Note that photons in this analysis are limited to the barrel region and thus do not utilize the preshower. A diagram of

---

\(^1\) Actually, to avoid particles passing through the cracks in the crystals, each crystal is aligned such that they make a small angle ($3^\circ$) with the vector pointing from the crystal face to the interaction point.
the ECAL layout is shown in Figure 3.10.

![Diagram of ECAL arrangement.](image)

**Figure 3.10: Diagram of ECAL arrangement.**

### 3.2.2.2 ECAL Lead Tungstate Crystals

Lead tungstate crystals have improved significantly over the last few years. They are desirable due to their high density (allowing for a compact calorimeter), and short radiation length and Molière radius (which allows for scintillation to take place in a small area, leading to fine granularity). These crystals emit blue-green scintillation light at 420 – 430 nm. Lead tungstate crystals are also very fast: 80% of the light is emitted within 25 ns, the design LHC bunch crossing time. They generally have a low light output, approximately 4.5 photoelectrons/MeV (as collected in the photodetectors). To increase the internal reflection, thereby increasing the light collection on the photodetectors, the crystals are polished. However, because the crystals have a tapered shape and high index of fraction ($n = 2.229$ around the scintillation wavelength), this causes nonuniformity in the light collection on the photodetectors. To account for this, one of the faces of the crystals is depolished. This effect is not as pronounced in the endcap where the crystal faces are nearly parallel.

Ionizing radiation damage from the LHC collisions causes absorption bands to form due to impurities in the crystal lattice. This effect is monitored and corrected using injected laser light to monitor the optical transparency, which is discussed in the section on laser monitoring. Hadronic
radiation has also been shown to cause a reduction of light output, but this is expected to stay within the limits of good ECAL performance during LHC operation.

3.2.2.3 ECAL Photodetectors

The photodetectors must be fast, radiation hard, and able to operate in the strong magnetic field provided by the solenoid. Furthermore, because of the low light yield of the lead tungstate crystals, they must amplify the signal and be insensitive to other particles traversing them. The CMS ECAL employs two different photodetectors: avalanche photodiodes (APDs) in the barrel, and vacuum phototriodes (VPTs) in the endcap. The VPTs are more radiation hard (hence their usage in the endcap), but have poorer quantum efficiency and gain. This is offset by the large surface area on the back of the crystals.

The APDs are silicon semiconductor devices which are operated at high bias voltages in order to achieve avalanche multiplication. They are $5 \times 5 \, \text{mm}^2$ and two APDs are connected to each crystal in order to increase acceptance. They are operated at a voltage of $340 - 430$ V, providing a gain of 50 (although, this depends on the voltage and the temperature). The gain directly influences the resolution of the ECAL, and because the gain is influenced by the operating voltage, this indicates that a very stable power supply is required. Such a power supply is custom built by the CAEN company and has a stability of a few tens of mV. The APDs have a high detection efficiency for the scintillation light from the lead tungstate crystals, $75 \pm 2\%$ at 430 nm. Furthermore, the dark current is $< 50 \, \text{nA}$. With radiation damage this is expected to rise to 5 $\mu\text{A}$. Each APD is thoroughly tested and screened. They exhibit no significant noise up to a gain of 300 and should operate reliably for 10 years of LHC running.

The VPTs are photomultipliers with a single gain stage. They have a diameter of 25 mm and an active area of 280 mm$^2$ and one VPT is connected to each crystal. The photocathode is at ground, while the dynode is operated at a voltage of $+600$ V and the anode is operated at $+800$ V. This leads to a gain of 10.2 in the absence of a magnetic field. When placed in a strong axial magnetic field, this gain is reduced. Still, the response is $> 90\%$ at a $15^\circ$ angle to the 4 T magnetic
field. Because the gain is saturated at the operating voltages, there is no need for the voltage to be precisely controlled. The VPTs are also thoroughly tested and transmission loss is expected to be \(< 10\%\) over the operation period for the LHC. See Figure 3.11 for pictures of crystals with a mounted photodetector.

![ECAL lead tungstate crystal with two APDs (left panel) or one VPT (right panel).](image)

**Figure 3.11**: ECAL lead tungstate crystal with two APDs (left panel) or one VPT (right panel).

### 3.2.2.4 ECAL Electronics

The on-detector electronics must accurately pick up the small signals from the detector with high speed. Furthermore, they must be radiation-hard because they are located on the crystals. The basic unit for the electronics is the “trigger tower” which is the unit used for triggering (see Section 3.2.6). The trigger towers are composed of \(5 \times 5\) crystals in \(\eta \times \phi\) for the barrel or identical \(5 \times 5\) crystals in the endcap (also known as a supercrystal). Each trigger tower unit has 5 Very Front End (VFE) boards which contain a Multi-Gain Pre-Amplifier (MGPA), an Analog to Digital Converter (ADC), and a radiation-hard buffer. The MGPA shapes the signal before passing it on to the ADC where it is digitized, and then finally the buffer adopts the ADC output for the Front End (FE) card. Each trigger tower unit contains one FE card which both creates the Trigger Primitive Generators (TPG) to send to the level-one (L1) trigger and also stores data to be transmitted upon receipt of the L1 trigger decision (again, see Section 3.2.6). The data is sent to the off-detector electronics using Gigabit Optical Hybrids (GOH) which are composed of electronics that serialize
the data and fiber optic cables. There are two for the barrel and six for the endcap. Each electronics unit also contains one Low Voltage Regulator (LVR) to power the FE and VFE cards.

The off-detector electronics is composed of a variety of different electronics boards stored in a VME crate (VERSA Module Eurocard, a standard computer architecture). These electronics serve both the Data Acquisition (DAQ) and trigger systems. For the DAQ system, this includes collecting crystal data from multiple FE boards, data suppression (using a programmable selective read-out algorithm), checking the integrity of the data, and monitoring memory occupancy. For the trigger system, this includes finalizing and synchronizing local triggers before sending the data to the regional calorimeter trigger.

3.2.2.5 ECAL Laser Monitoring System

Ionizing radiation damage due to the LHC collisions causes the formation of absorption bands due to impurities in the crystal lattice. The damage is limited but occurs rapidly. Annealing fixes the damage, so that, at a constant dose-rate, there is an equilibrium between the damage and the repair, resulting in a constant dose-rate dependent reduction in transparency. The LHC provides for varying dose-rates as the machine provides collisions or refills (see Figure 3.12). The magnitude of the changes in optical transmission are about 1-2% for the barrel (EB) and up to tens of percent for the endcap (EE). This effect is monitored and corrected using injected laser light to monitor the optical transparency. Laser pulses are injected into the crystals with optical fibers during opportunities such as beam gaps (periods with no proton bunches present). It takes approximately 30 minutes to scan all of the ECAL. Two kinds of light are used — one at 440 nm (the scintillation wavelength) and one at 796 nm (far from the realm of transparency changes, used to monitor the stability of the rest of the system). The response of the APDs is normalized by the laser pulse magnitude as measured by PN photodiodes. Thus, the transparency is given by:

$$ R(t) = \frac{APD(t)}{PN(t)} $$

Because of the different optical paths and spectra involved, the change in transparency does
Figure 3.12: Simulation of crystal transparency as a function of time assuming an instantaneous luminosity of $L = 2 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ and a machine cycle of 10 hours beam time and 2 hours filling time.

not affect the laser light the same as scintillation light. However, the relationship between the two can be described by a power law where the exponent is a characteristic of the crystal.

### 3.2.2.6 ECAL Calibration

The accurate calibration of the ECAL is of great importance as it influences the constant term of the energy resolution (see Section 3.2.2.7). Calibration to the precision of a few parts per thousand is difficult because of many small effects. There are two components in the calibration: a global component (giving the absolute energy scale) and the channel-to-channel relative component (known as “intercalibration”). The global component can be relatively easily determined by comparison to known physics processes. The channel-to-channel component is more difficult. It comes primarily from crystal-to-crystal variation in the scintillation light yield, yielding a total variation of 15% in the EB (but only 8% within a supermodule). Due to effects from the VPT signal yield,
gain variation and quantum efficiency, this rises to 25% in the EE. Estimates of intercalibration coefficients can come from a variety of sources. The first comes from laboratory measurements of crystal light yields, which brings the variation down to < 5% in the EB and < 10% in the EE. This can further be improved upon using cosmic rays and high energy electrons, bringing the calibration resolution to better than 1.5%. Final intercalibration is done with physics events from LHC operation. One type of event that takes advantage of the silicon tracker is to use the momentum of isolated electrons (for instance, from $W \rightarrow e\nu$ decays, which have the benefit of having a similar momentum as the photons from the benchmark $H \rightarrow \gamma\gamma$ process). Another type uses photons from $\pi^0$ events. Because there is no $\phi$ dependence on the energy deposition in the ECAL, calibration can occur within fixed $\eta$ regions.

### 3.2.2.7 ECAL Energy Resolution

Above 500 GeV, shower leakage from the rear of the calorimeter becomes significant. For energies $< 500$ GeV, the energy resolution can be described by the following equation:

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

Here, $S$ is the stochastic (random) term, $N$ is the noise term, and $C$ is the constant term.

Three basic sources contribute to the stochastic term: event-to-event fluctuations in the (lateral) shower containment (approximately 1.5% when using an array of $5 \times 5$ crystals), fluctuations in the energy deposited in the preshower or silicon tracker with respect to what is measured (parametrized as $\frac{5\%}{E_{\rm{Ecal}}}$), and a photostatistics contribution (approximately 2.1%) which depends on the number of primary photoelectrons released in the photodetector per GeV and the noise factor which parameterizes fluctuations in the gain process.

The noise term ($N$) is affected by electronics noise and pileup noise.\footnote{“pileup” is the term used to describe the situation in which more than one proton-proton interaction happens per bunch crossing. Thus, signals from unwanted interactions contribute to the noise in this situation.} Electronics and digitization noise can be determined by reconstructing the amplitude of a signal in the test beam — it is found to be 40 MeV/channel in the barrel. APD irradiation by neutrons will also contribute to
the electronic noise and is expected to be 8 MeV/channel after one year of operation at a luminosity of \(10^{33} \text{ cm}^{-2}\text{s}^{-1}\). The noise from the VPTs is expected to be constant, at 50 MeV. The pileup noise is studied using simulations and is shown to be small.

Finally, the constant term (C) is affected by 1) non-uniformity of the longitudinal light collection (0.35% per radiation length, which is achieved by the depolishing of one of the crystal faces) 2) intercalibration errors and 3) leakage of energy from the back of the crystal (which is negligible).

In 2004 the energy resolution was studied in the CERN H4 beam. Electron beams with a momentum of \(20 - 250 \text{ GeV}\) confirmed the above expectations. A typical energy resolution is found to be:

\[
\left( \frac{\sigma}{E} \right)^2 = \left( \frac{2.8}{\sqrt{E}} \right)^2 + \left( \frac{0.12}{E} \right)^2 + (0.30)^2
\]

where \(E\) is in GeV [26].

### 3.2.3 The Hadron Calorimeter

The Hadron Calorimeter (HCAL), in conjunction with the ECAL, is important for measuring hadronic jet energy and for determining missing transverse energy (MET) coming from neutrinos or exotic particles (for instance, neutralinos from supersymmetric theories). The HCAL is nestled between the ECAL and the magnet solenoid, between a radius of 1.77 m and 2.95 m. In this area is the HB (HCAL Barrel) and HE (HCAL endcap). Because significant calorimeter depth is needed for hadronic showering, there is also an HCAL Outer component (HO) outside the solenoid. The HO decreases shower leakage, improving the MET resolution. Furthermore, there is an HCAL Forward (HF) detector in the very forward region, 11.2 m from the interaction point and covering a pseudorapidity of \(4.5 < |\eta| < 5.2\); the HF can be used to help determine the luminosity collected...
from the LHC. The HCAL is composed of absorbing regions (usually brass or stainless steel) and active scintillating regions (made of plastic scintillators or quartz fibers). The light from the active region is collected into wavelength shifting fibers before heading to the electronic system. The light yield is measured with Hybrid Photodetectors (HPDs) which have a maximum voltage of 10 V. 68% of the pulse is contained within the 25 ns window; timing synchronization is achieved with a UV laser which can illuminate large sections of the HCAL at once.

The HB covers a pseudorapidity of $|\eta| < 1.3$ and contains two separate barrel sections divided into 36 “wedges.” The wedges are divided into 4 segments in $\phi$ and sixteen segments in $\eta$ and arranged such that there is no projected dead material from the interaction point. Thus, one segment or HCAL “cell” covers a range of $0.087 \times 0.087$ in $\delta \eta \times \delta \phi$. This cell is the basic geometric unit of a calorimeter “tower” (one HCAL cell and the geometrically corresponding ECAL crystals) used in jet reconstruction and the Level One trigger. Each wedge has alternating layers of absorber and “tile and wavelength-shifting fiber” active region (see Figure 3.13). The innermost and outermost (in the radial direction) absorbing plates are steel while the remaining plates are brass. This leads to 10.6 interaction lengths at $|\eta| = 1.3$, with an additional 1.1 interaction lengths provided from the ECAL. The plastic scintillator tiles are arranged in a tray which can be replaced without disassembling the absorber.

The HE, shown in Figure 3.14 covers a pseudorapidity of $1.3 < |\eta| < 3$, which contains 32% of the particles in the final state. Because of this, it must handle high counting rates and have a high radiation tolerance. Furthermore, because it is inserted into the end of the solenoid, it must be composed of non-magnetic materials. Due to this constraint, as well as cost and interaction

---

3 The luminosity is the number of particles per unit area per unit time and it is difficult to measure. The HF can help measure this because it collects a large fraction of the particles coming from the interaction point. At low luminosities the number of (minimum bias) energy depositions are counted and used as a baseline. An energy threshold just below 1 GeV will detect essentially all interactions. The energy depositions increase linearly with the luminosity. At high luminosities where far fewer minimum bias events are recorded (see Section 3.2.6) and almost the entire HF region is illuminated, the regions which do not have energy depositions are instead counted and used to infer the luminosity.

4 Although “absorbing” is the parlance of the field, it is a bit of a misnomer. In these regions of dense material, hadrons interact with the nuclei in the material which produces more particles and destroys the initial hadron. The newly produced particles have less energy than the initial hadron and also interact with the material, producing yet more particles at less and less energy. Eventually all particles produced are stopped.

5 An interaction length is defined as the average path length required to reduce the energy of a particle by $1/e$. 

length desires, brass is used for the absorbing plates (the same plastic scintillator is used as in the HB). The structural materials must also be non-magnetic to not disturb the field. Here, the EE and HE calorimeters combined provide approximately 10 interaction lengths.

The HO, situated outside of the solenoid, covers a pseudorapidity of $|\eta| < 1.3$ and is designed to catch the tails of hadronic showers and complement the HB. The HO takes advantage of the solenoid as an additional absorbing region, which provides $\frac{1.4}{\sin(\theta)}$ interaction lengths. The HO consists of 1 or 2 layers of scintillator on either side of a 19.5 cm thick piece of iron. The HO is geometrically constrained by the muon system and the mechanical structure of CMS. In all, the total calorimeter depth is approximately 11.8 interaction lengths (less at the barrel-endcap junction). The addition of the HO decreases shower leakage, thereby improving the MET resolution.

Figure 3.13: Drawing of HB wedges. Units are in mm.
Finally, the HF is the very forward HCAL component. It is designed to survive in unrivaled high particle fluxes for a minimum of 10 years. During this time, if 10 MGy are delivered to the HF, the optical transmission is cut in half. With this consideration in mind, the active material is chosen to be quartz fibers which are very radiation hard. Furthermore, steel is chosen as the absorber. The HF itself is broken into two longitudinal segments in order to differentiate electromagnetic particles from hadrons (electrons and photons deposit most of their energy in the first 22 cm). The components of the HF are shielded behind lead, steel, and borated polyethylene slabs. The HF is used as a luminosity monitor. Two different methods are used to determine the luminosity, including 1) a method which correlates the HCAL tower energy with luminosity and 2) a method in which the average fraction of empty towers is used to infer the mean number of interactions per bunch crossing. See Figure 3.15 for a layout of the HCAL.
3.2.3.1 HCAL Resolution

Similar to the ECAL, the HCAL Resolution is given as:

$$\left( \frac{\sigma}{E} \right)^2 = \left( \frac{S}{\sqrt{E}} \right)^2 + C^2 \quad (3.8)$$

The typical HCAL resolution is:

$$\left( \frac{\sigma}{E} \right)^2 = \left( \frac{70\%}{\sqrt{E}} \right)^2 + (8\%)^2 \quad (3.9)$$

where $E$ is in GeV [26].

The ECAL resolution is much better than the HCAL resolution (see Section 3.2.2.7) which will influence the method to determine the QCD background in this diphoton analysis (see Section 5.4.1).

3.2.4 The Superconducting Magnet

The CMS superconducting magnet is designed to reach a 4 T field, although it is routinely operated at 3.8 T in order to increase the lifetime of the magnet. It is comprised of two main parts:
a superconducting solenoid or “cold mass” and a large iron return yoke.

The superconducting solenoid is composed of four winding layers of NiobiumTitanium (NbTi) conductor which is mechanically reinforced with an aluminum alloy ( unlike previous magnets, which have a maximum one or two layers and are not reinforced). At zero magnetic field, NbTi has a critical temperature $T_c = 9.25$ K, and at a magnetic field of 4.6 T, $T_c = 7.3$ K. The solenoid is operated at a temperature $T = 4.5$ K and at an operating current of $\sim 19$ kA. At this temperature and current, the maximum temperature for which current can flow freely ($T_g$) is calculated as $T_g = 6.44$ K, which means there is a temperature margin of 1.94 K. The radial extent of the coil is kept small for physics considerations. Unlike previous thin detector solenoids, the CMS magnet has a large ratio between stored energy and cold mass (11.6 KJ/kg) (see Figure 3.16), which causes large deformations during the energizing of the magnet. Thus, the coil itself must have a structural function as well as a magnetic function — the support cylindrical mandrels designed for this purpose carry 30% of the magnetic hoop stress. Other associated components include the vacuum and cryogenic systems, as well as the grounding circuits, current leads, and the quench detection system.

The iron yoke is composed of five barrel wheels and six endcap disks, ranging in weight from 400 tons to 1920 tons (the central wheel). The easy relative movement of these pieces allows for easy subdetector assembly and insertion. The displacement of the wheels or disks is done by air and grease pads. Once all elements are placed next to each other, they are pre-stressed with 100 tons of force, which ensures good contact when the magnet is turned off. After the magnet is on, the total compressive force is 8900 tons.

The support system is designed to withstand forces created by a 10 mm magnetic misalignment in any direction of the cold mass with respect to the iron yoke (see Figure 3.17). Measurement of the misalignment can be done by measuring the displacement of the cold mass or by the stress on supporting rods. Tests done during the surface hall tests in 2006 showed a displacement of the cold mass of 0.4 mm in the $+z$ direction, indicating a misalignment of less than 2 mm off the magnetic center in $z$. 
Figure 3.16: A comparison of the energy/mass ratio versus stored energy for various detector magnets.

Figure 3.17: Drawing of the five barrel solenoid modules and support structures.
3.2.5 The Muon System

Muon identification is a very useful tool in many physics analyses because of the ease of identification of muons and because muons have much smaller radiative losses than electrons. Because of this, the muon system is designed for excellent muon identification (including charge identification), momentum measurement, and triggering. The CMS muon system uses three different types of gaseous detector technologies: drift tubes (DTs) in the barrel, cathode strip chambers (CSCs) in the endcap, and resistive plate chambers (RPCs) which are used primarily for triggering and are present in both the endcap and barrel. The entire muon system must be very well aligned both with itself and also with the tracker (the alignment is accurate to 75 µm in the barrel and 150 µm in the endcap). The alignment is done through several methods including survey and photogrammetry measurements, measurements from an opto-mechanical system, cosmic ray muon measurements, and results of alignment algorithms based on reconstructed tracks. The entire muon system covers a pseudorapidity range of $|\eta| < 2.4$.

The DTs cover a pseudorapidity range $|\eta| < 1.2$. This system is composed of four “stations” which are interspersed among the iron yoke (see Figure 3.18). Three of the stations contain 8 “chambers”, four of which measure the muon coordinate in the $r-\phi$ plane, and four of which measure the muon coordinate in the $z$ direction. The fourth station contains no $z$ direction chambers. Each chamber is composed of two or three “superlayers” each of which contain four layers of rectangular drift cells (for redundancy). There are approximately 172,000 sensitive wires in the DT system. The wire length is around 2.4 m in the $r-\phi$ plane and is located in a gas mixture of 85% Ar and 15% CO$_2$. The tube geometry is selected because it protects against damage from a broken wire and also decouples neighboring cells from electromagnetic showers resulting from the passing muon traveling through dense materials.

The CSC system covers a rapidity range $0.9 < |\eta| < 2.4$. There are 468 CSCs divided among two endcaps with four stations each. Each CSC contains 6 anode wire planes and 7 cathode panels (leading to 6 gas gaps per CSC), for a total of 2 million wires. The cathode strips run
radially outward and allows for precision measurement in the $r - \phi$ plane while the anode wires are perpendicular to the cathode strips, and provide for a measurements in $\eta$ (as well as a beam-crossing time). The CSC system provides for robust pattern recognition for the rejection of non-muon backgrounds and also for efficient matching of hits to the other muon system and to the inner tracking system.

The RPCs are a complementary system to the DTs and CSCs; they are present in both the barrel and endcap, and are designed specifically for the trigger system. The RPCs cover a rapidity range $|\eta| < 1.6$. The RPCs consist of a double-gap module (with read-out strips in between, see Figure 3.19) that are operated in avalanche mode (for high rates). They have poorer spatial resolution than the DTs and CSCs, but much better time resolution (the time can be measured to a length much shorter than the 25 ns bunch crossing interval). Therefore, they can identify the relevant bunch crossing for a particular muon track. There are six RPC layers in the barrel, two in the first two DT stations, and one each in the last two. The reason for the redundancy is that the first two stations are for the proper trigger treatment of low-momentum tracks that may not
reach the outer stations. The endcap has three planes of RPCs in each of the first three stations. In 10 years of data taking, there should be no efficiency degradation of the RPCs.

![Figure 3.19: RPC double-gap design.](image)

3.2.6 The Trigger

Because of the enormous design luminosity of the LHC (including the rate of bunch crossings — every 25 ns or 40 MHz), CMS must implement a trigger system which reduces the amount of data recorded, saving only the most interesting events. Per crossing, the detectors produce about 1 MB of data. Without the trigger, this translates to 40 TB/s of data, which is impossible to store. The trigger employs two steps: the level one (L1) trigger, which is a set of custom designed electronics, and the High Level Trigger (HLT) which is software that is installed on a filter farm of approximately 1000 CPUs.

3.2.6.1 Level-1 Trigger

The L1 trigger is designed for an output rate limit of 100 kHz (reduced from 40 MHz), and in action operated at about 30 kHz (a factor of three for safety). The L1 triggers takes coarse data from the calorimeters and muon system (but not the tracking system) and performs an analysis that determines whether or not the event is kept. During this time, higher resolution data is
in the memory pipeline. The latency for the L1 trigger is 3.1 $\mu$s, so the processing is also pipelined. The L1 trigger is composed of Field-Programmable Gate Arrays (FPGAs) where possible (which are very flexible), or Application-Specific Integrated Circuits (ASICs) and programmable memory lookup tables (LUTs) in cases where speed, radiation hardness, or density is an issue. These electronics are housed either on the detector or in the underground control room.

The L1 trigger is composed of a number of steps, as shown in Figure 3.20. First, there is the local trigger, or Trigger Primitive Generator (TPG), which takes energy measurements in the calorimeter or track segments in the muon system and passes these to the Regional Trigger. Here, logic is used to establish, rank, and sort various trigger objects (electrons/muons). The rank of an object is determined as a function of its energy, momentum, and quality. This regional information is then passed to the Global Calorimeter or Global Muon trigger, which determines the very highest ranking objects over the entire subdetector. This information is passed on to the Global Trigger, which accepts or rejects an event based on the subdetector triggers. The global trigger also takes into account the readiness of the data acquisition system and the subdetectors. This is known as the Trigger Control System (TCS). Once a decision is made, it is passed to the subdetectors with the Timing, Trigger, and Control system (TTC). This system tells the subdetectors whether or not to forward the more high-resolution information from the event to the Data Acquisition System (DAQ) (see Section 3.2.7).

**The L1 Calorimeter Triggers.** The calorimeters are divided into trigger tower areas. The towers have an expanse in $(\eta, \phi)$ of (0.087, 0.087), for $|\eta| < 1.74$ (corresponding to one HCAL cell). For larger values of $\eta$, the towers are larger. The TPGs are integrated with the calorimeter read-out and sum the transverse energies measured in the ECAL crystals or HCAL trigger-towers. This information is passed on to the regional trigger via a high-speed serial link. The regional trigger then ranks the local candidates before forwarding it on to the global calorimeter trigger and, finally, the global trigger.

**The L1 Muon System Trigger.** The muon system, which is composed of three different types of detectors, has a trigger system which uses all three technologies. In the local trigger,
both the Drift Tubes and Cathode Strip Chambers. Forward information about track segments and their corresponding bunch crossings to the regional trigger. At the regional level, the DT and CSC information is combined and the track segments are formed into muon tracks. In addition, the Resistive Plate Chambers, with their excellent timing information, have their own tracks at the regional level. Then, information from all three is forwarded to the global muon trigger, which ranks and sorts candidates before sending it to the Global Trigger for a final event decision. Using information from all three muon systems improves trigger performance.

### 3.2.6.2 The High Level Trigger

The High Level Trigger is a software package that uses input from the L1 trigger to determine whether or not a single event should finally be fully reconstructed and the information stored. More information on the HLT can be found in Chapter 4 which describes the software used at CMS as well as how data is reconstructed into physics objects.
3.2.7 The Data Acquisition System

The Data Acquisition system (DAQ) collects and processes data coming from the CMS subdetectors. It is designed to do this according to the design LHC bunch crossing frequency (40 MHz). It must receive information at the L1 trigger rate of 100 kHz (corresponding to 100 GB/s) and provide enough computing power for the HLT, as well as be able to output \( \sim 100 \) Hz of event data for offline reconstruction and analysis.

All subdetectors store their data in a 40 MHz buffer. Upon arrival of a L1 trigger (synchronized among all subdetectors), the Front-End Drivers (FEDs) push the data into the DAQ system from the buffers. Data from the FEDs are read out into what is known as Front-end Read-out Links (FRLs) (which can merge data from two FEDs). The DAQ is designed for 512 FRLs, which is above the number needed by the subdetectors (\( \sim 450 \)) and allows for inputs from local trigger units and other contingencies. A list of FED numbers, FRL numbers, and data channels for the various subdetectors are available in Table 3.2. While the number of FRLs is typically equal to the number of FEDs (except for the tracker), the number of FEDs is not an obvious function of the number of channels of the subdetector — it depends on zero-suppression ability and other detector-specific information. The FRL electronics are located in the underground electronics room.

<table>
<thead>
<tr>
<th>Subdetector</th>
<th>Number of Channels</th>
<th>Number of Data Sources (FEDs)</th>
<th>Number of FRLs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel Tracker</td>
<td>66 M</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Strip Tracker</td>
<td>9.3 M</td>
<td>440</td>
<td>250</td>
</tr>
<tr>
<td>ECAL</td>
<td>76 k</td>
<td>54</td>
<td>53</td>
</tr>
<tr>
<td>ECAL Preshower</td>
<td>144 k</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>HCAL</td>
<td>9 k</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>Muons CSC</td>
<td>500 k</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Muons RPC</td>
<td>192 k</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Muons DT</td>
<td>195 k</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

Software known as the event builder assembles the data from all FRLs which belong to a single L1 trigger. This event data is then sent to the filter farm where the HLT analysis is performed. The filter farm is located above ground, so the event builder is also responsible for transferring this
data. The “Event Filter” process is responsible for performing the HLT selection but also, if an event is selected by the HLT, serves a subset of events to local (or remote) shifters studying data quality and transfers all data from the local storage at the CMS site to the CERN data center for more permanent storage.
Chapter 4

Software and Data Reconstruction

4.1 Introduction

This chapter describes the software used by CMS, known as CMSSW, and how this software is used to reconstruct the relevant physics objects for this analysis: photons, jets, and missing transverse energy ($MET$).

4.2 Basic Kinematic Variables

CMS is not a completely hermetic detector because if it were, then it would (obviously) block the beam pipe used to deliver the proton-proton collisions. For this reason, particles and energy from collisions are lost “down the beam pipe.” Thus, this analysis (and most analyses) focus only on the component of kinematic quantities that are transverse to the beam pipe so that conservation of transverse momentum can be exploited. These quantities are known as “transverse” quantities — such as “transverse momentum” or “transverse energy.” Transverse momentum ($p_T$) is the component of a particle’s momentum that is transverse to the beam line, and transverse energy ($E_T$) is the magnitude of this component. $E_T$ can be modified by the mass of the particle, but because all regular particle masses are very small compared to the energies being studied, $E_T$ and $p_T$ are essentially the same quantity. For the (massless) photons under study, $E_T$ and $p_T$ are equivalent. Another common variable used is missing transverse energy ($MET$) and is described below. Finally, the pseudorapidity is a very commonly used variable (although more geometric in nature than kinematic). It was described in the introduction to CMS in Chapter 3. In
short, pseudorapidity, $\eta$ describes the location of its energy deposits in the CMS subdetectors, and subdetectors themselves cover certain ranges in $\eta$ (again, see the chapter on the CMS experiment). A particle traveling perpendicular to the beam direction has $\eta = 0$ while a particle traveling down the beam line has $\eta = \infty$. Thus, those with smaller $\eta$ values are in the barrel region of the detectors as opposed to those with larger values which are in the endcap region of the subdetectors. Because of the initial momenta of the colliding particles (parallel to the beam line), there are far higher backgrounds in high-$\eta$ regions and these subdetectors must be particularly radiation hard.

4.3 CMSSW

CMSSW is the software framework used by CMS to perform a variety of functions: to select events during high level triggering (HLT), to deliver the results to experimenters, to reconstruct events, to simulate events and detector response, and to provide the tools necessary for analysis. The process of reconstruction is the process of taking signals in the various subdetectors and putting together that information to form physics objects (photons, electrons, jets, etc.) and their kinematic properties, as well as their geometric location. The objects in an event can indicate a variety of physics processes that happened during that event. It is an object-oriented software which primarily uses C++ [26]. The version used in this analysis is CMSSW_3_8_3.

4.4 High Level Trigger

The High Level Trigger software package takes input from the L1 trigger (described in Section 3.2.6) and performs partial reconstruction of an event in order to determine if the event should be kept and fully reconstructed. There are several types of triggers depending on the analysis one wants to do. For instance, if one wants to look at photons, one uses the HLT_PhotonX trigger, which means that the software looks for events with photons. A variety of photon triggers can be present in any one incarnation of the HLT software and may have various requirements on energy, isolation, or shape of the detector signal. A list and description of the particular triggers used in this analysis can be found in Chapter 5. Because of limited processing power and storage space,
not every event that passes through CMS can be recorded and reconstructed — only a limited rate of data can be taken. Therefore, triggers must have a relatively tight selection, and the needs of the various CMS physics groups must be balanced. Furthermore, because the HLT selection is done online and before all detector responses are stored, it must be very fast and work well with a subset of information.

4.5 Photon and Electron Reconstruction

An electromagnetic object (photon or electron) passing through the ECAL will leave energy in several neighboring crystals in the ECAL. The only information about photons originating from a proton-proton collision comes from the ECAL detector as they are chargeless (so they do not leave tracks in the Tracker) and do not deposit energy in the HCAL or Muon detectors. Electrons also do not deposit energy in the HCAL or Muon detectors, but they do leave charged tracks in the Tracker. For this analysis, the energy measurement of the photons and electrons is performed identically. Electrons are different from photons in that they are matched to a pixel track stub, or “pixel seed”, a process described below.

4.5.1 Clustering

Crystals which have energy deposits in them are grouped together to form (basic) clusters, and clusters are grouped together to form superclusters (in order to ensure all energy is recorded). The software algorithms which perform this clustering are different depending on whether or not the crystals in question are located in the barrel (EB) or endcap (EE). Only photons located in the EB are used in the analysis, so the discussion of EE photon reconstruction is disregarded. Also note that clustering is the same for a photon as it is for an electron.

The algorithm that creates the superclusters, known as the “hybrid clusterizing algorithm” starts with searching for individual crystals with large energies [26]. This seed crystal has the highest energy of the crystals surrounding it. After a seed crystal is located, crystals are added to the seed crystal in the $\eta$ direction until a $1 \times 5$ or $1 \times 3$ array is formed (the choice of $1 \times 3$ or
$1 \times 5$ is made by whether or not the outer crystals in the strip have energy deposits). Once an array is created, clustering moves to the $\phi$ direction, extending both in $+\phi$ and $-\phi$ until the effect of adding additional strips in $\phi$ no longer adds significant energy to the cluster. The end result is called a supercluster. For a visual explanation, see Figure 4.1.

The clustering over an extended $\phi$ region is done because the strong magnetic field from the $3.8$ T solenoid causes the energy flow of primary electrons or converted primary photons to spread in $\phi$.

The clustering over an extended $\phi$ region is done because the strong magnetic field from the $3.8$ T solenoid causes the energy flow of primary electrons or converted primary photons to spread in $\phi$.

After clustering is completed, several energy corrections are applied, which total approximately $1\%$ of the uncorrected supercluster energy [36]. The first of these is an $\eta$ dependent energy correction and is meant to compensate for losses due to lateral energy leakage due to the $3^\circ$ offset of the EB crystals (see section 3.2.2). The second is a correction which compensates for material in front of the ECAL (the tracker). These interactions spread only in the $\phi$ direction due to the magnetic field, so the correction is a function of the size of the supercluster in the $\phi$ direction compared to its size in the $\eta$ direction. Finally, there is a correction, also due to material in front of the ECAL, that is $\eta$ and transverse energy ($E_T$) dependent due to the varying amount of material along $\eta$ and the dependence on $E_T$ of bremsstrahlung and conversion.

All superclusters are considered to be potential photon candidates and are stored as photon...
objects in the software. If the ratio of the energy of the $3 \times 3$ crystals around the seed crystal to the energy of the entire supercluster is $\leq 0.94$, then the photon is assumed to be unconverted (i.e., not from $\gamma \rightarrow e^+e^-$ decays). This ratio, called $R9$, essentially measures how spread the cluster is in the $\phi$ direction and is sensitive to conversions [36]. For an unconverted photon, the energy of the photon is assigned the energy and position of the $5 \times 5$ crystals around the supercluster seed crystal. For a converted photon, the energy and position is assigned based on the entire supercluster.

4.5.2 ECAL Spikes

During the 2010 LHC running, it was discovered that the ECAL suffers from anomalous energy deposits in the APDs that can fake high energy photons. A variety of methods to protect against these spikes were employed, all of which take advantage of the unphysical shower shape of the spikes. Detailed information about the spike cleaning can be found in the chapter that describes the data analysis, Section 5.2.

4.5.3 Tracking and Pixel Seed Matching

Track reconstruction in CMS is performed by the Combinatorial Track Finder (CTF) [27]. The first step in tracking charged particles is the identification of “seeds” from which to form a track. A seed is any set of three nearby hits in consecutive layers in the tracker (or two hits which are additionally constrained by the beamspot or a pixel vertex) that would connect in a physically valid straight or curved line. The collection of seeds found in an event is then cleaned to avoid redundancy. Using these seeds, tracks are then grown layer-by-layer from the innermost part of the detector (starting with the seed) toward the outer layers of the tracker. As hits are found in each subsequent layer, the positions are added to the track and the track parameters (and associated uncertainties) are recalculated. For each subsequent layer, a track in which no hit is measured is propagated to account for the possibility of the track not leaving a hit [28]. The search continues until the outer boundary of the tracker is reached. The CTF uses a Kalman Filter algorithm to determine potential track paths. The Kalman Filter is mathematically equivalent to a global
least-squares minimization [28].

Tracking is an iterative process. In the first iteration only triplets of pixel hits are considered for seeding. In the first three iterations, only seeds formed from hits in the pixel are considered. Subsequent iterations use combinations of pixel and tracker hits. Seeds which are exclusively formed from pixel hits are known as pixel seeds. Pixel seeds are very relevant for this analysis as they are used to distinguish between electrons and photon (see Chapter 5).

One or more pixel seeds can be matched to a supercluster. For pixel seed matching, superclusters are broken down into basic \((5 \times 5)\) clusters [36]. For each basic cluster, the pixel seed matching code loops over the collection of pixel seeds identified by performing track reconstruction. Pixel seeds that form tracks intersecting a basic cluster are considered matched to the basic cluster. Basic clusters which have one or more seeds are considered to “have a pixel seed”. Superclusters formed from these basic clusters would also be considered to “have a pixel seed” or “be pixel seed matched” — a strong indication that the particle that left the energy deposit in the ECAL was, in fact, charged.

4.6 Jet Reconstruction

There are four types of jet reconstruction used in CMS [29]. This analysis uses the Jet Plus Tracks (JPT) algorithm which combines information from the calorimeters with information from the tracker (which has excellent resolution). The first step in reconstructing JPT jets is to first construct Calorimeter (Calo) Jets. Calorimeter jets are formed by combining information about energy deposits in the ECAL and HCAL. A calorimeter “tower” is formed by taking an HCAL cell (see Chapter 3) and the geometrically corresponding ECAL crystals (in the barrel, this is a simple \(5 \times 5\) basic cluster). This tower covers an area of \(0.087 \times 0.087\) in \(\delta \eta \times \delta \phi\). In order to suppress noise, individual ECAL crystal and HCAL cell thresholds are enforced.

The calorimeter towers are then clustered into jets using the “anti-\(k_T\)” algorithm with a jet size parameter of \(R = \sqrt{\eta^2 + \phi^2} = 0.5\). The anti-\(k_T\) algorithm combines calorimeter towers by starting with the highest energy tower [30] and searching for nearby calorimeter towers with energy
depositions, and adding them to the jet. This algorithm guarantees a cone-like geometry.

Once calorimeter jets are reconstructed, to create JPT jets, one associates charged tracks with the calorimeter jets based on the spatial separation in $\eta$ and $\phi$ between the jet axis and the track momentum (i.e., the tracks originate from the same vertex as the jet). The associated tracks are projected onto the surface of the calorimeter. If the track is in the cone of the calorimeter jet, it is classified as an “in-cone” track. For in-cone tracks, the expected energy deposition in the calorimeter is subtracted from the jet and the momentum of the track is added back in (improving resolution). For tracks in which the projection falls out of the cone of the calorimeter jet, the momentum of the track is also added to the jet (thereby allowing for the inclusion of energies not measured in the calorimeters alone). The result is known as a JPT jet.

4.6.1 Jet Energy Corrections

A number of energy corrections are applied to jets after their reconstruction because the response of the calorimeters to jets (“jet response”) is non-linear in $p_T$ and not uniform in $\eta$. These are known as the Level 1, Level 2, Level 3, etc. corrections and are applied sequentially [29]. The L1 correction is an overall energy offset correction (applied to all jets in an event). The correction is meant to account for calorimeter noise and effects from pile-up (extra proton-proton collisions in the same bunch-crossing). The L2 correction corrects a jet’s pseudorapidity ($\eta$). This correction is designed to make the jet response flat versus $\eta$ and is determined either using Monte Carlo simulation\(^1\) information or data information from dijet events — such events should have jets balanced in $\eta$. Finally, the L3 correction corrects the jet transverse momentum $p_T$. This correction is designed to make the jet response flat versus $p_T$. The correction is again determined either by Monte Carlo information or by using $Z + \gamma + jet$ events — where such events should balance in $p_T$.

The energies of jets used in this analysis are calculated using all L1+L2+L3 corrections. However, the calculation of the jet position in $\eta$ does not use any corrections because the L2 correction was set to zero in CMSSW$_{3.8.3}$.

\(^1\) Monte Carlo simulation refers to algorithms which simulate physical processes by random sampling.
4.7 Missing Transverse Energy ($MET$) Reconstruction

Accurate $MET$ measurements can be made at CMS because the detector covers a solid angle of nearly $4\pi$. Thus, the presence of neutrinos and hypothetical neutral weakly interacting particles can be inferred by the presence of significant amounts of missing transverse momentum which is the (apparent) imbalance of momentum in the direction perpendicular to the beam direction ($\vec{E}_T$). The magnitude of this momentum is known as Missing Transverse Energy ($MET$). The $MET$ resolution depends strongly on the calorimeter resolution, and is therefore determined by the hadronic energy in the event (because the ECAL resolution is much better than the HCAL resolution).

$MET$ is the magnitude of the negative vector sum of the momentum transverse to the beam axis of all final-state particles [31]. The traditional method to measure $MET$ is to use the calorimeter tower energies, assume massless particles, and use the angles defined by a vector from the primary vertex of the interaction to the location of the tower. This corresponds to the standard “calorimeter” or “caloMET”. This analysis uses “track-corrected $MET$”, or “tcMET.” It is similar to caloMET, but instead of using the calorimeter tower energies, the corresponding charged-track momenta are used instead — leading to a much more accurate description of $MET$ and one that is compatible with the use of JPT jets. After tcMET is calculated, it is corrected for anomalous signals in the calorimeter (in particular for spikes in the ECAL) and beam halo muons which can leave substantial unbalanced energy in the calorimeters. Methods for applying these corrections are found in [31].
Chapter 5

Analysis with 35.5 pb$^{-1}$ of CMS Data at 7 TeV

5.1 Introduction

This analysis searches for evidence of General Gauge Mediated Supersymmetry Breaking (GGM) where the neutralino is the next-to-lightest supersymmetric particle and the gravitino is the lightest supersymmetric particle. Prompt decays of the neutralino would yield events with two high-momentum photons, large missing transverse energy ($MET$) associated with gravitinos and jets at the LHC. The analysis compares the $MET$ distribution from events with two high momentum photons and at least one jet compared to the background distribution expected from Standard Model processes. The data used was collected by CMS from $\sqrt{s} = 7$ TeV proton-proton collisions at the LHC during 2010. The integrated luminosity (the amount of data collected) is $35.5 \pm 3.9$ pb$^{-1}$[32].

In order to put limits on (or discover) particular GGM models, several areas must be studied. The expected background must be determined, the efficiency to find the signal must be studied, and the luminosity must be measured. To exclude particular GGM models, the cross sections and associated errors must also be established. In this chapter, the datasets and event selection are discussed, which leads to the determination of the background (Section 5.4). The GGM signal efficiency (and systematic uncertainties) of the GGM signal events (from Monte Carlo simulation) are also determined (Section 5.6). The luminosity and its associated uncertainty is given by [32], and the cross sections and associated uncertainties are calculated using PROSPINO [33].
5.2 Datasets and Triggers

Proper event selection is crucial for any analysis. In particular for this analysis, the photon identification is an important part of the process. The CMS datasets used in this analysis are listed in Table 5.1 along with the total number of events recorded (any event that passes any High Level Trigger). A subset of the data is preselected or “skimmed.” Initially, the skim required only that there be one ECAL supercluster with \( E_T > 30 \) GeV and one with \( E_T > 20 \) GeV (for a description of “supercluster”, essentially an energy deposit, please see the chapter on data reconstruction, Chapter 4). This skim corresponds to the first dataset listed in Table 5.1. The skim then evolved to require that both ECAL superclusters have an energy \( E_T > 30 \) GeV (this skim corresponds to the second dataset listed in Table 5.1). The number of events remaining after the skim is also given in Table 5.1.

The data is reconstructed using CMSSW 3.8.3 (see chapter on data reconstruction, Chapter 4). During early running, it was discovered that there are spikes generated in the ECAL due to anomalous deposits in the APDs that can fake high energy photons. These are easily removed during the standard CMSSW 3.8.3 reconstruction. Unlike electromagnetic objects which shower over several crystals, the spikes appear in only one or two crystals. Because of this feature, they are removed from the ECAL reconstructed hit (RecHit) collection by applying a “Swiss cross” veto. If the ratio of energy of the sum of the four neighboring crystals to the energy of the seed crystal (thus forming a “Swiss cross” shape) is \( \leq 0.05 \) of the total energy, then the reconstructed hit is rejected. There is another approach to spike rejection, and it is known as the “e2/e9” veto. In this case, the sum of the highest two energy crystals in the \( 3 \times 3 \) supercluster (e2) is compared to the sum of the energy of the entire \( 3 \times 3 \) supercluster (e9). If the ratio \( e2/e9 \) is \( \geq 0.95 \), then the cluster is identified as a spike and removed. Unlike the Swiss cross veto, the e2/e9 veto is not applied during the standard reconstruction, and is instead used at the analysis level. See Figure 5.1 for a visual description of these vetoes.

A variety of triggers are used as the LHC running evolved and are listed in Table 5.2, along
with the raw event counts from each trigger after the skimming is completed. All triggers require one photon with transverse energy of $E_T > 30$ GeV with the exception of one trigger which requires $E_T > 22$ GeV and a match to an ECAL supercluster (HLT-Photon22-SC22HE-L1R-v1). With the exception of HLT-Photon30-L1R, the triggers are also all cleaned of spikes (described above). Triggers that came into use later in the run also have a track isolation requirement (HLT-Photon30-Isol-EBOnly-Cleaned-L1R-v1 and HLT-Photon22-SC22HE-L1R-v1). This requires that the photon not be matched to a track within a radius of $\Delta R = \sqrt{(\phi_{\text{photon}} - \phi_{\text{track}})^2 + (\eta_{\text{photon}} - \eta_{\text{track}})^2} \leq 0.4$.

The offline selection requires photons to have $E_T > 30$ GeV. The fact that the offline selection and trigger cut are identical is not an issue. First, the signal (GGM Monte Carlo) efficiency is minimally affected by raising the $E_T$ cut to 35 GeV. Second, both photons in our event selection (described below) must meet the offline $E_T > 30$ GeV cut; thus, not just one object of interest but two objects of interest are capable of triggering. None of the triggers were prescaled when used.

\footnote{In order to keep the event rate low enough such as not to overwhelm the computing system during the high instantaneous luminosity running during the last part of the 2010 run, this last 22 GeV trigger required a match to the supercluster. However, that lowered the rate significantly enough that the energy requirement could also be lowered from the typical 30 GeV.}
Table 5.1: Datasets used in analysis.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Total Events</th>
<th>Skimmed Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>EG-Run2010A-Sept17ReReco-v2-RECO</td>
<td>52257480</td>
<td>1100267</td>
</tr>
<tr>
<td>Photon-Run2010B-PromptReco-v2-RECO</td>
<td>169592333</td>
<td>2750303</td>
</tr>
</tbody>
</table>

The total luminosity for this sample was $35.5 \pm 3.9 \text{ pb}^{-1}$ [32]. The instantaneous luminosity during the period of data taking ranged from $10^{29} \text{ cm}^{-2}\text{s}^{-1}$ to $10^{32} \text{ cm}^{-2}\text{s}^{-1}$.

5.2.1 General Gauge Mediation Signal Monte Carlo

In order to establish upper limits on cross sections or discover General Gauge Mediation Supersymmetry Breaking (GGM) models, it is helpful to generate “signal” Monte Carlo (MC) for comparison with data. As discussed in the theory section, in GGM models the neutralino is the next-to-lightest supersymmetric partner (NLSP) and the gravitino is the lightest supersymmetric partner (NLSP). The MC samples for these models are generated using PYTHIA [34] for production followed by GEANT [35] to simulate the detector and model the detector response. The simulated events were run through the standard data reconstruction software (CMSSW). A grid of signal points was generated where the masses of gluinos ($m_{\tilde{g}}$) and the masses of squarks ($m_{\tilde{q}}$) are assumed to be degenerate and range from 400 GeV to 2000 GeV. For each combination of masses, three different neutralino masses ($m_{\tilde{\chi}_0}$) were considered: 50 GeV, 150 GeV, and 500 GeV. This grid corresponds to the LHC Benchmark for GGM studies [15]. The dominant modes of production are gluino-gluino, squark-gluino, and squark-squark.

Leading order cross sections for these models are calculated using PROSPINO [33] and shown

Table 5.2: Triggers used in analysis.

<table>
<thead>
<tr>
<th>Triggers</th>
<th>Events Yielded</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLT-Photon30-L1R</td>
<td>71734</td>
</tr>
<tr>
<td>HLT-Photon30-Cleaned-L1R</td>
<td>1045396</td>
</tr>
<tr>
<td>HLT-Photon30-Isol-EBOOnly-Cleaned-L1R-v1</td>
<td>117694</td>
</tr>
<tr>
<td>HLT-Photon22-SC22HE-L1R-v1</td>
<td>379371</td>
</tr>
</tbody>
</table>
in Figure 5.2. Next-to-leading order (NLO) cross sections were also calculated, leading to k-factors (the k-factor is the factor one must multiply the leading-order cross section to account for the NLO calculation). k-factors are shown in Figure 5.3.

In order to compare distributions from a typical GGM point to distributions from data, throughout this thesis the point corresponding to \( m_{\tilde{q}} = 640 \text{ GeV} \), \( m_{\tilde{g}} = 640 \text{ GeV} \), and \( m_{\tilde{\chi}_0} = 150 \text{ GeV} \) was used. This point has a cross section of 1.6 pb and a k-factor of 1.44.

As discussed in Section 2.5, this grid of points features prompt photons from neutralino decays, \( MET \) from gravitinos, and the presence of jets. The average number of jets per signal point is given in Figure 5.4. Gluino decays yield two jets (from quark-antiquark production) and a gaugino (the neutralino in the GGM case is the lightest gaugino). Squark decays yield one jet and a gaugino if the squark is lighter than the gluino. If the squark is heavier than the gluino, it will emit a gluino and a quark, leading to three jets and a gaugino. Thus, GGM events have between two and six jets from SUSY cascades, with more jets in the \( m_{\tilde{q}} > m_{\tilde{g}} \) case.

5.3 Event Selection and Object Identification

The events of interest contain two high energy photons, large \( MET \), and jets. Correspondingly, two samples are selected with two photons, one that requires one jet, and one that does not (in order to increase statistics). All significant calculations are done requiring at least one jet. Furthermore, in order to do the QCD and EW background determination, three control samples are identified. Two of the control samples are used for independently determining the QCD background. These samples are orthogonal and can thus be considered as cross-checks. Furthermore, the results of the predictions from the two distinct samples can be combined to give an even better estimate of the background. One of these two samples has two electrons in it (essentially, \( Z \rightarrow ee \) events). The other has two “fakes” and is called “fake-fake.” Fakes are essentially “fake photons”, that are electromagnetically rich jets, defined in Section 5.3.3. Finally, one sample is selected that has one electron and one photon in it (\( e\gamma \) sample), and it is used for EW background determination (see Section 5.4.2) The \( Z \rightarrow ee \) sample is also used in the EW background determination because
Figure 5.2: Cross sections (in pb) in the squark-gluino mass phase space for $m_{\tilde{\chi}_0} = 50$ GeV (a), $m_{\tilde{\chi}_0} = 150$ GeV (b), $m_{\tilde{\chi}_0} = 500$ GeV (c).
Figure 5.3: k-factors in the squark-gluino mass phase space. The k-factor is the factor one must multiply the leading-order cross section to account for the NLO calculation.

it is used to determine the electron-to-photon fake rate (Section 5.4.2). More details about the object selection and final data samples are given in this section.

5.3.1 Photon Identification

The identification of photons follows the official recommended prescription from CMS, and hence is a well-studied and well-defined selection. Each photon in an event must pass the following criteria, which will be described more fully below:

- $E_T > 30$ GeV
- $\eta < 1.479$
- The seed crystal cannot be within $\sim 6$ crystals of the edge of the barrel which is equivalent to $\Delta \eta = 0.1$
Figure 5.4: Number of generated jets with $p_T > 30$ GeV and $|\eta| < 2.6$ for $m_{\tilde{\chi}} = 50$ GeV (a), $m_{\tilde{\chi}} = 150$ GeV (b), and $m_{\tilde{\chi}} = 500$ GeV (c).

- $\sigma_{\text{iso}} < 0.013$

- Isolation Criteria

  * ECAL $E_T^{\text{Sum}}$ in the isolation cone $< 0.006 \cdot E_T + 4.2$ GeV
  * HCAL $E_T^{\text{Sum}}$ in the isolation cone $< 0.0025 \cdot E_T + 2.2$ GeV
- $E_T^{Sum}$ of tracks in track isolation cone $< 0.001 \cdot E_T + 2.0$ GeV

- $H/E < 0.05$ (where $H$ is the hadronic energy in the HCAL tower directly behind the ECAL supercluster, and $E$ is the energy of the ECAL supercluster)

- $e2/e9 \leq 0.95$ where $e2$ is the energy of the two highest-energy crystals in the $3 \times 3$ supercluster, and $e9$ is the energy of the entire $3 \times 3$ supercluster.

- The time of the signal in the ECAL must be within 3 ns of the interaction time (only applied in events with no jet requirement).

- Has no pixel seed match (i.e., has no pixel track stub that is associated with the supercluster, see Section 4.5.3.).

The reason for the minimum transverse energy requirement on the photon is simple enough - the signature features high energy photons. The requirement that the photon be in the barrel ($|\eta| < 1.479$) is related to two things. One, the barrel has better energy resolutions because of the crystal geometry (see Section 3.2.2). In the barrel, all crystals subant an equal amount of $\phi - \eta$ space. Furthermore, there are larger backgrounds in the endcaps due to the momentum of the incident protons in the collision. The reason for avoiding seed crystals close to the barrel/endcap gap is that some energy may be lost in the gap (or endcap ECAL) and therefore does not provide an accurate energy measurement. The cut in combination with the requirement that the photon be located in the barrel reduces to requiring that the photon $|\eta| < 1.469$.

The variable $\sigma_{i_{\eta}i_{\eta}}$ is used to take into account the shape of the energy deposit. It is the width of the shower shape in the $\eta$ direction and is calculated using a log energy weighted width in $\eta$. It is calculated only within the central $5 \times 5$ crystal array of the supercluster (around the maximum energy crystal). As opposed to the $\sigma_{\eta\eta}$ variable which is calculated using the crystal coordinates, $\sigma_{i_{\eta}i_{\eta}}$ is calculated in terms of the crystal index and is given by the equation:

$$\sigma_{i_{\eta}i_{\eta}}^2 = \frac{\sum_{i = 1}^{5 \times 5} w_i (i_{\eta} - i_{\eta_{seed}})^2}{\sum_{i = 1}^{5 \times 5} w_i^2}$$ (5.1)
where \( w_i = \ln (E_i/ \sum_i^{5\times5} E_i) \), \( E_i \) is the energy of the \( i^{th} \) crystal and \( i \eta_i \) is the \( \eta \) index of the \( i^{th} \) crystal within the \( 5 \times 5 \) cluster around the crystal seed, and \( i \eta_{\text{seed}} \) is the \( \eta \) index of the seed crystal and \( E_{5\times5} \) is the energy of the \( 5 \times 5 \) crystals around the seed. \( \sigma_{i \eta i \eta} \) is unaffected by gaps between crystals so it is more regular across boundaries (what matters is the spread in the crystals in the \( \eta \) direction while the object is showering, the object does not shower in the gaps). This variable is used to discriminate against jets.

The isolation criteria are also straightforward - the photon must be isolated in order to discriminate against jets, and to ensure an accurate measurement of the photon energy. The isolation cones are described in Figure 5.5.\(^2\)

![Figure 5.5: Figure illustrating the ECAL and HCAL isolation cones. The track isolation cone differs from the ECAL isolation cone by the width of the \( \eta \) strip (0.015) and the radius of the central hole (0.04) (Figure not to scale).](image)

The isolation is based on a sum of reconstructed hits in the isolation cones. The inner cone excludes real energy for a given supercluster which is not clustered (leakage). Note the removal of the fixed-width \( \eta \) strip in the ECAL isolation cone. This type of isolation, where the \( \eta \) strip is excluded, is called “Jurassic isolation” and is designed to remove the unclustered energy in the cone as well as the clustered energy. A strip of fixed \( \eta \) along \( \phi \) completely excludes the region that

\(^2\) Note that for the isolation requirements, \( E_T^{\text{Sum}} \) refers to the sum of all ECAL (or HCAL or Track) energy in the isolation cone, whereas \( E_T \) refers to the energy of the object under study.
could be clustered (see Section 4.5.1). Clustering is done along $\phi$ at fixed $\eta$ (because the spread due to the magnetic field is in $\phi$).

The $e2e9$ cut is used to further discriminate against spikes and anomalous signals.

The timing cut is implemented to reject non-beam background associated with beam halo (where signals arrive early compared to photons originating from the interaction point) and muons from cosmic ray interactions in the atmosphere (which are random in time). Beam halo muons travel in time with bunches from the beam, but photons emitted from these muons are closer to the ECAL detectors than photons coming from proton-proton collisions, and for this reason have a negative timing with respect to the main collision event time. For a visual aid to understanding the timing of the Beam Halo, refer to Figure 5.6. Typical timing distributions of photons arriving from proton-proton collisions versus non-beam backgrounds can be found in Figure 5.7. This cut is not applied for events in which a jet is required because a reconstructed jet implies an event originating from the interaction point (namely, there are tracks associated with the interaction point).

The pixel seed veto ensures that there is no pixel track stub (used to generate tracks) that is associated with the supercluster (in space). A pixel seed would indicate that the supercluster is more likely associated with an electron (or jet) and is therefore not an isolated photon.

5.3.2 Electron Identification

The identification of electrons is identical to that of photons except that in this case there is a pixel seed requirement. Note that the electron and photon identifications are completely orthogonal (distinct), and that the efficiencies for passing all cuts should be similar for both isolated electrons and isolated photons. Thus, by determining the efficiencies for electron identification (primarily using $Z \rightarrow ee$ decays), the efficiencies for photons are also determined.

5.3.3 Fake Identification

A fake is defined identically to a photon except that it is allowed to have a pixel seed and is required to fail either of the following two cuts:
Figure 5.6: Sketch illustrating that photons arriving from proton-proton collisions and from beam halo will have different timing distributions. Muons from beam halo travel in time with the bunches and will emit photons that have a distinct timing difference from interaction photons due to the difference in travel time necessary to reach the ECAL.

\[
\Delta t = \frac{(Z + \sqrt{Z^2 + R^2})}{c}
\]

- \( \sigma_{\text{miss}} < 0.013 \) (see discussion below)
- \( E_T \) of tracks in the track isolation cone < \( 0.001 \cdot E_T^{\text{fake}} + 2.0 \text{ GeV} \)

Thus, the “fakes” are very similar to photons and should therefore have similar energy resolution. These objects in reality are a very select group of jets which are mainly electromagnetic, with little to no hadronic energy associated with them other than that they may have tracks in their vicinity.

5.3.4 Jet Selection

Jets are found using the JPT algorithm with energy corrections applied (For JPT description, see Section 4.6). They must satisfy the following criteria, officially recommended by CMS:
Figure 5.7: Supercluster position ($\eta_{SC}$) vs. time of photon seed ($t_{seed}$). Photons originating from beam halo type events (yellow crosses) are identified by their association with events that have hadronic calorimeter and muon signals, no reconstructed tracks, and missing transverse energy $MET > 25$ GeV (since the events are expected to be unbalanced). Prompt photons (blue squares) are identified by having no association with hadron calorimeter or muon signals, photon seed time within 3 ns in events requiring $MET < 15$ GeV, and at least three reconstructed tracks in the tracker. Candidate photons (black circles) are similar to prompt photons, but require $MET > 30$ GeV (indicating events of interest). See [36].

- $p_T > 30$ GeV
- $|\eta| \leq 2.6$ (which excludes the forward hadron calorimeter)
- $f_{HPD} \leq 0.98$, where $f_{HPD}$ is the fraction of energy contributed by the highest energy HCAL hybrid photodetector readout (see Section 3.2.3).
- $N_{90} \geq 2$, where $N_{90}$ is the minimum number of ECAL and HCAL cells required to contain
90% of the jet energy

- $EMF \geq 0.01$ where $EMF$ is the electromagnetic fraction of the energy of the jet

- Must be separated from all photons in the event that pass the photon identification by

$$\Delta R = \sqrt{(\phi_{\text{photon}} - \phi_{\text{jet}})^2 + (\eta_{\text{photon}} - \eta_{\text{jet}})^2} \geq 0.9$$

(this cut is derived from the 0.4 isolation cone of the photon plus the 0.5 reconstruction cone of the jet - see Chapter 4 for a description of the jet algorithm).

### 5.3.5 Data samples

Four different data samples are used in this analysis, described now in detail:

- $\gamma\gamma$: the candidate sample which contains at least two photons meeting the photon identification requirements above, separated by $\Delta R \geq 0.8$ (which represents two isolation cones for $\Delta R = 0.4$), and $\delta\phi \geq 0.05$ for the no jet requirement case. These events are used to search for the signal but are most likely from QCD multi-jet, $\gamma + \text{jet}$, and direct diphoton production.

- $e\gamma$: the control sample which contains one photon and one electron meeting the identification requirements above, and which are separated by $\Delta R \geq 0.8$ (and $\delta\phi \geq 0.05$ for the no jet requirement case). This sample is used for the EW background determination, and for determining the electron-photon fake rate. These events are most likely from $W\gamma$ and $W\text{jet}$ events (where the jet in the latter case fakes a photon) and where the $W$ decays to an electron and a neutrino.

- $ee$: the control sample which contains at least two electrons meeting the identification requirements above, and which are separated by $\Delta R \geq 0.8$ (and $\delta\phi \geq 0.05$ for the no jet requirement sample). This sample is used for QCD background determination, and for determining the electron-photon fake rate. These events are most likely from $Z \rightarrow ee$ decays.
• fake-fake (ff): the control sample containing at least two fakes meeting the identification requirements above, and which are separated by $\Delta R \geq 0.8$ (and $\delta \phi \geq 0.05$ for the no jet requirement case). These events are most likely from QCD multi-jet production.

For each sample above there are actually two samples - one which requires $\geq 1$ jet, and one which makes no jet requirement. Because jets are expected in GGM events, the samples with jets will be used for the final analysis and limit setting, but the no jet requirement case (which has larger statistics) is also used as a check. In the case where an event has more than two valid electrons, photons, or fakes, then the two highest $E_T$ objects are used in order to classify the event. In this way, one event does not fall into multiple samples. The reason for the $\delta \phi$ requirement in the no-jet requirement sample is to further discriminate against photons from beam halo muon bremsstrahlung. Halo muons travel parallel to the beam line and therefore photons originating from these muons would not be well separated in $\phi$. For a large number of kinematic plots comparing these samples, see Appendix C.

The final event count for these samples is given in Table 5.3.

<table>
<thead>
<tr>
<th>Sample</th>
<th>No Jet Requirement</th>
<th>$\geq 1$ Jet Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma\gamma$</td>
<td>404</td>
<td>87</td>
</tr>
<tr>
<td>$e\gamma$</td>
<td>234</td>
<td>43</td>
</tr>
<tr>
<td>$ee$</td>
<td>4348</td>
<td>588</td>
</tr>
<tr>
<td>fake-fake</td>
<td>1950</td>
<td>425</td>
</tr>
</tbody>
</table>

5.4 Standard Model Background to MET shape

There are four different sources of background to GGM-type events. The first and most significant background is called the “QCD” background and is background with no true MET; the MET in these events comes from mismeasured jets. The second most significant type of background is called the “EW” background and comes from events with true MET (neutrinos from W decays). The third type of background is “non-beam” background, and comes from beam halo or cosmic
ray muons that have undergone bremsstrahlung; as these are non-beam backgrounds there is also MET associated with them because the event is unbalanced. These backgrounds are removed by the photon selection itself, which requires the photon to have a specific shower shape and timing relative to the collision time (see Section 5.3.1 for more detail). The final (and least significant) background is the “irreducible” background and comes from true diphoton plus MET events, such as $W\gamma\gamma$ and $Z\gamma\gamma$ events (see Figure 5.8). MET in $W\gamma\gamma$ events comes from neutrinos from the $W \rightarrow e\nu$ decay and in $Z\gamma\gamma$ events from the neutrinos in $Z \rightarrow \nu\bar{\nu}$ decays. These types of events have very small cross sections: 0.1 pb for $W\gamma\gamma$ and 0.02 pb for $Z\gamma\gamma$ as calculated by MADGRAPH [37]. The branching ratio for $W \rightarrow \ell\nu$ is $10.80 \pm 0.09\%$ and the branching ratio for $Z \rightarrow \nu\bar{\nu}$ is $20.00 \pm 0.06\%$ [38]. Thus, even with 100% acceptance, these backgrounds are negligible compared to the EW and QCD backgrounds. Therefore, this section addresses only the QCD and EW type backgrounds. The QCD background can be determined from both the fake-fake and $Z \rightarrow ee$ samples (the determination of the MET shape will be described), and the EW background is estimated from the $e\gamma$ sample (scaled appropriately to account for the $e - \gamma$ fake rate). At $MET > 50$ GeV, a quantitative comparison between the expected background (from Standard Model processes) and the candidate sample can be determined. If there are excess events in the candidate sample for $MET > 50$ GeV, that would be a sign of GGM-like events.

5.4.1 QCD Background

The QCD background is defined as Standard Model events from proton-proton collisions that have no true missing transverse energy ($MET$). The $MET$ resolution is dominated by the hadronic activity in the event because the ECAL resolution is much better than the HCAL resolution (see Chapter 3). The selection of fake-fake or $Z \rightarrow ee$ events are chosen as separate control samples because they both have similar hadronic activity compared to the candidate $\gamma\gamma$ sample and also have no true $MET$. Two control samples are used to ensure that the analysis methodology gives the same result in both cases.

Hadronic recoil in events with no true $MET$ should balance the electromagnetic energy of
the two photons, fakes, or electrons. Therefore, it should be the same in all samples provided the fake-fake or $ee$ energy spectra are reweighted to correspond to the photon energy spectrum in the $\gamma\gamma$ sample. Hence, the $MET$ resolution is correlated with the vector sum of the transverse momentum of the two leading electromagnetic objects ($diEM_p_T$) (for example, consider a Mercedes-Benz type event with two photons balancing a jet). In order to correct for the different $diEM_p_T$ spectra, the two QCD control samples are (separately) scaled to match the $diEM_p_T$ spectrum from the candidate $\gamma\gamma$ sample. First, the ratio of the $diEM_p_T$ spectrum from the candidate $\gamma\gamma$ to the $diEM_p_T$ spectrum from the control sample (either fake-fake or $Z \rightarrow ee$) is taken and a linear fit to this ratio is performed. Then, the control sample events are individually scaled according to the value of the fit at each event’s $diEM_p_T$ value (the “$diEM$ scale factor”), making the $diEM_p_T$ spectra agree. The (uncorrected) $diEM_p_T$ spectra are shown in Figures 5.9 and 5.10, while the
ratio plots are shown in Figures 5.11 and 5.12. The $MET$ distributions of the QCD events are also scaled using this same $diEMpT$ scale factor. The $MET$ distribution (for reasons which will become clear in the next paragraph) is called the “Central $MET$ distribution”.

![Normalized (to one) unweighted $diEMpT$ spectra for the $\gamma\gamma$ and $Z \rightarrow ee$ events with no jet requirement (a) and $\geq 1$ jet requirement (b). The Monte Carlo GGM example point is also shown for comparison.](image)

To calculate the uncertainty on the ratio, each bin of the original $diEMpT$ spectrum is varied. To determine the new value of the bin, a random number is chosen by varying a Gaussian distribution with a mean of the bin’s value and a sigma equal to the statistical uncertainty in that bin. This is repeated for all bins (so that the shape is varied, but the statistical uncertainty is the same) for both the candidate diphoton sample and the control QCD samples, and the same
process of performing the ratio plot and fitting is completed. This process is repeated 1000 times so that there are 1000 new MET histograms with values distributed relative to the original MET histogram (fluctuating above and below the central value). Then, the systematic uncertainty in each bin is calculated by determining the number of fluctuations below and above the value in the Central MET Distribution. The negative uncertainty for the bin is taken to be the entry value that is one sigma below the value from the Central MET Distribution; similarly, the positive uncertainty for the bin is taken to be the value that is one sigma above the value from the Central MET Distribution. Thus, the uncertainties due to the MET shape can be asymmetric in each bin.
Figure 5.11: The $d\text{iEM}_{p_T}$ ratio of $\gamma\gamma$ to $Z \rightarrow ee$ events with no jet requirement (a) and $\geq 1$ jet requirement (b). Note the different binning with respect to Figure 5.9.

Finally, it is assumed that there is no new physics at low values of $MET$. The $MET$ distributions of the QCD control samples are normalized such that the total number of entries in the QCD sample for $MET < 20$ GeV is equal to the total number of entries in the candidate $\gamma\gamma$ (less the EW background obtained for the $e\gamma$ sample for $MET < 20$ GeV, see section 5.4.2).

Using the fake-fake sample, there are, for $MET > 50$ GeV, $0.50^{+0.35}_{-0.35}$ predicted events with $\geq 1$ jet requirement, and $0.76^{+0.41}_{-0.41}$ (with no jet requirement). Using the $Z \rightarrow ee$ sample, there are,
Figure 5.12: The $diEM_{p_T}$ ratio of $\gamma\gamma$ to fake-fake events with no jet requirement (a) and $\geq 1$ jet requirement (b). Note the different binning with respect to Figure 5.10.

for $MET > 50$ GeV, $1.75^{+0.77}_{-0.73}$ predicted events with $\geq 1$ jet requirement, and $2.98^{+0.70}_{-0.68}$ with no jet requirement. These results are summarized in Tables 5.4 and 5.5 (also with EW background estimation). Final $MET$ plots (with EW background estimation included, described below) are given in Figures 5.18 and 5.19.
Table 5.4: QCD Background Estimation and Candidate Events for $MET > 50$ GeV.

<table>
<thead>
<tr>
<th>Type</th>
<th>$\geq 1$ jet</th>
<th>no jet cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Candidate Signal Events</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Fake-Fake QCD Background Estimate</td>
<td>$0.50^{+0.35}_{-0.33}$</td>
<td>$0.76^{+0.41}_{-0.40}$</td>
</tr>
<tr>
<td>$Z \rightarrow ee$ QCD Background Estimate</td>
<td>$1.75^{+0.35}_{-0.73}$</td>
<td>$2.98^{+0.41}_{-0.68}$</td>
</tr>
</tbody>
</table>

5.4.2 Electroweak (EW) Background

EW backgrounds contain true $MET$, and in this case come from $W$ decays into an electron and neutrino, where the electron fakes a photon (from $W\gamma$ or $Wjet$ events, the latter having also a jet misidentified as a photon). To determine the number of events in our candidate signal sample that actually come from events with $W$s having an electron faking a photon, the $e\gamma$ sample $MET$ distribution is multiplied by $N_{e\gamma} \cdot \frac{f_{e\gamma}}{1-f_{e\gamma}}$, where $N_{e\gamma}$ is the number of $e\gamma$ events in each bin and $f_{e\gamma}$ is the rate that an electron fakes (looks like) a photon. If there are $N_{W\gamma\rightarrow\gamma e\nu}^{true}$ “true” $W\gamma \rightarrow e + \gamma + \nu$ events, then, after reconstruction:

$$N_{W\gamma\rightarrow\gamma e\nu}^{true} \cdot (1 - f_{e\gamma}) = N_{e\gamma}$$

$$N_{W\gamma\rightarrow\gamma e\nu}^{true} \cdot f_{e\gamma} = N_{\gamma\gamma}^{fake}.$$ (5.3)

Solving this system of equations yields

$$N_{\gamma\gamma}^{fake} = N_{e\gamma} \cdot \frac{f_{e\gamma}}{1-f_{e\gamma}}.$$ (5.4)

Therefore, to determine the number of EW-type background events, we simply scale the $MET$ distribution of the $e\gamma$ sample by $\frac{f_{e\gamma}}{1-f_{e\gamma}}$. Determination of $f_{e\gamma}$ must then be done. If the misidentification rate is high, then there will also be contributions from Drell-Yan and $t\bar{t}$ events where both electrons are misidentified as photons (this is not the case as shown by the results below).

As described previously, the electron and photon selection are identical except for the identification of a pixel seed for the electron. An electron is an ECAL supercluster with identical isolation and shape variables as that of a photon and, in addition, is associated with a reconstructed pixel match from the silicon tracker. A photon has a veto on this same pixel match. Therefore, the photon and electron selections are orthogonal and contain distinct and separate events. An electron
that is misidentified as a photon is one in which this ECAL supercluster has no pixel seed match, but is in reality an electron. If there are $N_{Z}^{\text{true}}$ true $Z \rightarrow ee$ events, then the number of $Z$ events in the dielectron sample is given by:

$$N_{ee} = N_{Z \rightarrow ee}^{\text{true}} (1 - f_{e\gamma})^2$$

(5.5)

Similarly, the number in the electron-photon sample is:

$$N_{e\gamma} = 2N_{Z \rightarrow ee}^{\text{true}} (f_{e\gamma})(1 - f_{e\gamma})$$

(5.6)

The fake rate can be determined by solving the two previous equations:

$$1 - f_{e\gamma} = \frac{2N_{ee}}{2N_{ee} + N_{e\gamma}} = \epsilon_{e\gamma}$$

(5.7)

where $\epsilon_{e\gamma}$ is the efficiency for properly identifying a photon. $N_{e\gamma}$ and $N_{ee}$ are determined by fitting the reconstructed $Z$ in the invariant mass histograms for the $Z \rightarrow ee$ and $e\gamma$ samples. The fake rate describing how often a true electron is misidentified as a photon can be determined by comparing the integrals of these fits. By also looking at the diphoton invariant mass, this method is overdetermined. If there are any odd correlations with the electron misidentification rate then the predicted number of $Z$ events in the diphoton sample using the fake rate will not match the actual number of $Z$ events in the diphoton sample determined using the fit. This is not the case, as results show that the number of actual $N_{\gamma\gamma}$ events matches what is predicted by using $N_{ee}$, $N_{e\gamma}$ and the determined $f_{e\gamma}$.

A Crystal Ball function is used to fit the invariant mass distribution of each sample. The fit is done twice, once using a linear background and once using a constant background. The statistical uncertainties are determined by the integrals of the fits, and the systematic uncertainties are determined by comparing the differences in the integrals between the linear and constant background.

\begin{equation}
 f(x) = \begin{cases} 
 A \cdot e^{-\frac{1}{2} \left(\frac{x-x_0}{\sigma}\right)^2} & \frac{\frac{x-x_0}{\sigma}}{\alpha} > \alpha \\
 A \cdot (\frac{\alpha}{\sigma})^n \cdot e^{-\frac{\alpha^2}{2}} \cdot (\frac{\alpha}{\sigma} - \frac{x-x_0}{\sigma})^{-n} & \frac{\frac{x-x_0}{\sigma}}{\alpha} < \alpha 
 \end{cases}
\end{equation}

(5.8)

where $A$ is the amplitude, $x_0$ is the mean, $\sigma$ is the width, $\alpha$ is a Gaussian tail parametrization and $n$ is a normalization of the tail.

---

\footnotesize

3 The Crystal Ball function, which helps describe regions with low-mass tails due to bremsstrahlung, is given as:
fits (higher order polynomial fits do not change the integral significantly from the linear fit). The electron-photon misidentification rate, \( f_{e\gamma} \), is determined using Equation 5.7, where \( N_{ee} \) and \( N_{e\gamma} \) are the number of events in the Z peak as determined by the fit.

The invariant mass histograms were further separated by leading object \( p_T \) and \( |\eta| \) to determine the efficiency as a function of transverse momentum and pseudorapidity. In this case, the fit was done with a Breit-Wigner function convoluted with a Gaussian, and again with linear and constant backgrounds to determine the systematic error. The normalization, Breit-Wigner mean and width, Gaussian width, and linear and constant backgrounds were allowed to float. The reason for the change in fit is that the Breit-Wigner/Gaussian function converges more reliably on these very low-entry histograms (correspondingly, the uncertainties on these fits are largely statistical). All parameters of the fits were allowed to float.

Every fit was done with the log likelihood method, preferred for histograms with small numbers of entries as it does not ignore the bins with zero entries (also there is no background subtraction that would cause the uncertainties to no longer be Poissonian). For binned fit results, see Appendix B.

The fits, shown in Figures 5.13 and 5.14 yield \( N_{ee} = 3823.0 \pm 1273.6(85.9\text{sys}) \) and \( N_{e\gamma} = 98.1 \pm 48.9(5.2\text{sys}) \), leading to a measured efficiency of \( 1 - f_{e\gamma} = 0.987 \pm 0.008 \). The errors are high due to the complexity of the Crystal Ball function. Using this to predict the number of Z events in the diphoton sample, \( N_{\gamma\gamma\text{FromZ}} = 0.63 \pm 0.77 \), which is consistent with zero. Similarly, a fit to the invariant mass yields a normalization of \( 2.1 \pm 1.3 \) for the diphoton sample, which is consistent with the prediction using the \( f_{e\gamma} \), within uncertainties (see Figure 5.15). Therefore, there are no correlations between electron reconstruction efficiencies. Furthermore, the efficiency is not dependent on \( p_T \) or \( \eta \) (see Figures 5.16 and 5.17). Constant fits were performed on the plots of efficiency versus \( \eta \) or \( p_T \) and the results are consistent with the unbinned result (again, see Figures

\[ f(x) = \frac{\Gamma}{(x - x_0)^2 + \frac{\Gamma^2}{4}} \]  

where \( x_0 \) is the mean and \( \Gamma_0 \) is the width.

---

\( ^4\) The Breit-Wigner function used is: 

\[ f(x) = \frac{\Gamma}{(x - x_0)^2 + \frac{\Gamma^2}{4}} \] 

(5.9)
Figure 5.13: The invariant mass of two electron candidates in $Z \rightarrow ee$ sample. The $Z$ peak is fit using a Crystal Ball function with constant background (a) and linear background (b).
Figure 5.14: The invariant mass of one electron candidates and one photon candidate in $e\gamma$ sample. The $Z$ peak is fit using a Crystal Ball function with constant background (a) and linear background (b).
Figure 5.15: The invariant mass of two photon candidates in diphoton sample. The $Z$ peak is fit using a Crystal Ball function with a linear background. The mass, width, and $\alpha$ parameters were set equal to the fit results for the $Z \to ee$ sample.

Figure 5.16: The electron/photon reconstruction efficiency as a function of pseudorapidity.
Figure 5.17: The electron/photon reconstruction efficiency as a function of transverse momentum.

Once the value of $f_{e\gamma}$ is determined, the MET background can be estimated by scaling EW ($e\gamma$) events according to Equation 5.4. Using the found $e\gamma$ fake rate, $f_{e\gamma} = 0.013 \pm 0.007$, the corresponding weight by which to scale the $e\gamma$ sample is found to be $0.013 \pm 0.008$ (see Equation 5.4). After scaling, the number of events with $MET > 50$ GeV is $0.04 \pm 0.03$ for both the $\geq 1$ jet requirement case and the no jet requirement case. The distribution (along with the QCD distributions) can be found in Figures 5.18 and 5.19.

5.4.3 Final Background (QCD+EW) Summary

The MET distribution including all predicted backgrounds is given in Figure 5.18 for the $Z \rightarrow ee$ QCD control sample and Figure 5.19 for the fake-fake QCD control sample. The number of observed candidate events and the predicted backgrounds are given in Table 5.5. Only one candidate event with $MET \geq 50$ was found. The different event displays for this event can be seen in Figure 5.20.
Figure 5.18: The diphoton candidate MET spectrum, the $e\gamma$ MET spectrum (rescaled according to the $e\gamma$ fake rate) for the EW MET background prediction, and the $Z \rightarrow ee$ MET spectrum (reweighted and normalized to the $MET < 20$ GeV region) for the QCD MET background prediction for not jet requirement (a) and $\geq 1$ jet requirement (b). The simulated GGM spectrum is also shown for comparison. Last bin contains overflows.
Figure 5.19: The diphoton candidate MET spectrum, the $e\gamma$ MET spectrum (rescaled according to the $e\gamma$ fake rate) for the EW MET background prediction, and the fake-fake MET spectrum (reweighted and normalized to the $MET < 20$ GeV region) for the QCD MET background prediction for not jet requirement (a) and $\geq 1$ jet requirement (b). The simulated GGM spectrum is also shown for comparison. Last bin contains overflows.
Figure 5.20: The XY event display of the one observed candidate event with $MET > 50 \text{ GeV}$ (a), the 3D display (b), the lego event display (c), and the $\rho - Z$ display (d).

Table 5.5: Final Estimation of Background and Candidate Events for $MET > 50 \text{ GeV}$.

<table>
<thead>
<tr>
<th>Type</th>
<th>$\geq 1 \text{ jet}$</th>
<th>no jet cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Candidate Signal Events</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Fake-Fake QCD Background Estimate</td>
<td>$0.50^{+0.35}_{-0.35}$</td>
<td>$0.76^{+0.41}_{-0.41}$</td>
</tr>
<tr>
<td>$Z \rightarrow ee$ QCD Background Estimate</td>
<td>$1.75^{+0.77}_{-0.73}$</td>
<td>$2.98^{+0.70}_{-0.68}$</td>
</tr>
<tr>
<td>$e\gamma$ EW Background Estimate</td>
<td>$0.04 \pm 0.03$</td>
<td>$0.04 \pm 0.03$</td>
</tr>
<tr>
<td>Total Background (using fake-fake)</td>
<td>$0.54^{+0.35}_{-0.35}$</td>
<td>$0.80^{+0.41}_{-0.41}$</td>
</tr>
<tr>
<td>Total Background (using $Z \rightarrow ee$)</td>
<td>$1.79^{+0.77}_{-0.73}$</td>
<td>$3.02^{+0.70}_{-0.68}$</td>
</tr>
<tr>
<td>Predicted GGM Events</td>
<td>$10.27 \pm 2.20$</td>
<td>$10.28 \pm 2.20$</td>
</tr>
<tr>
<td>Predicted GGM Events (with K factor of 1.66)</td>
<td>$17.1 \pm 3.7$</td>
<td>$17.1 \pm 3.7$</td>
</tr>
</tbody>
</table>
5.5 Analysis Method Check — Evidence of $W\gamma$ and $Wjet$

It is possible to confirm that the method gives reasonable results by using it to “discover” a known process with true MET. For example, $W\gamma$ and $Wjet$ events have true MET resulting from the decay of the W into an electron and neutrino. Thus, if the analysis is performed on the $e\gamma$ sample described above, there should be excess events at high MET (the photon for the $Wjet$ sample would come from a jet misidentified as a photon). It should be noted that this is in no way a fine-tuned analysis as the electron identification would be different in an analysis truly searching for $W\gamma$ or $Wjet$ events. Never the less, it is a good check to determine whether the analysis works. Shown in Figure 5.21 is the MET spectrum for the $e\gamma$ sample (no jet requirement), along with the predicted QCD background (using $Z \rightarrow ee$ events as the control sample). Also plotted are contributions from $W\gamma$ and $Wjet$ Monte Carlo, and the total predicted spectrum (QCD + $W\gamma$ and $Wjet$ MC). The first order cross section for $W\gamma$ analysis is given as 54.5 pb for 7 TeV collisions at CMS, and is multiplied by a k-factor of 1.8 to account for next to leading order production. The cross section for $Wjet$ is 26.5 nb for 7 TeV proton-proton collisions. For MET $> 30$ GeV, there are 19 events in the $e\gamma$ sample but only 7.9 ± 1.0 events for the QCD background. Thus, there is an excess of events for MET $> 30$ GeV indicating the presence of $W\gamma$ and $Wjet$ events.

One can also perform a Kolmogorov Smirnov (KS) test\textsuperscript{5} to determine the level of agreement between the $e\gamma$ candidate spectrum and the two different background predictions. Comparing the $e\gamma$ spectrum to just the QCD predicted background yields a KS value of 0.30, while comparing the $e\gamma$ spectrum to the total predicted background (QCD+MC) yields a KS value of 0.88, indicating that the agreement is much better with the MC signal events.

\textsuperscript{5} A Kolmogorov Smirnov test is a test for the equality of two one-dimensional continuous distributions. It is calculated by comparing the separation between the Empirical Distribution Functions of the two distributions. The empirical distribution function describes the probability that a random variable X with a given probability distribution will be found at a value less than or equal to x. Identical histograms would yield a KS value of 1 while completely disjoint histograms would yield a KS value of 0.
Figure 5.21: The \( \text{MET} \) distribution of the \( e\gamma \) sample (no jet required) compared to the QCD background prediction using \( Z \rightarrow ee \) events, \( W\gamma \) and \( W\text{jet} \) MC, and the total \( (\text{QCD} + \text{MC}) \) predicted \( e\gamma \) spectrum. Photons from \( W\text{jet} \) production come from when the jet is misidentified as a photon.

### 5.6 Efficiency of GGM Events

Once the background is determined, in order to exclude regions of GGM model phase space, the efficiency of the GGM space must be determined. By performing an identical diphoton event selection on the grid of GGM points described in Section 5.2.1 as on the data, the expected number of GGM events can be determined. For each model in the phase space a signal efficiency was determined and is equal to the number of events that pass the diphoton selection (and have \( \text{MET} > 50 \text{ GeV} \)) over the total number of events generated. This ratio (the GGM signal efficiency) has an associated statistical uncertainty.

The signal efficiency of the GGM events is further affected by several factors associated with the event selection. These include the jet efficiency, the photon efficiency, the jet energy scale uncertainty, and the parton distribution function uncertainty.
5.6.1 Breakdown of Cuts on Efficiency

For the example GGM model used throughout this thesis (corresponding to $m_{\tilde{q}} = 640$ GeV, $m_{\tilde{g}} = 640$ GeV, $m_{\tilde{\chi}_0} = 150$ GeV), the breakdown of the various cuts on the signal efficiency was studied. For 9996 total generated events, 6207 contained two generated photons from neutralinos with $p_T > 30$ GeV and $|\eta| < 1.379$. Of these, 6051 events contained two generated photons that were matched to a reconstructed photon with $\Delta R \leq 0.1$ to (this corresponds to a $\sim 98\%$ photon reconstruction efficiency). After the application of the photon identification requirements, 2914 events remained, and 2795 remained after the requirement of at least one isolated jet. Finally, 2316 events were accepted after the final selection (the most significant of which is the separation of the two photons by $\Delta R = 0.8$; this cut is implemented to avoid overlap of the isolation cones). This breakdown is summarized in Table 5.6.

The greatest loss of events is the requirement that there be two photons in the barrel and with $p_T > 30$ GeV. However, this requirement heavily reduces the background and so therefore a loss in signal efficiency is acceptable. The other large drop in signal efficiency is due to the photon identification requirements outlined in Section 5.3.1 and again these requirements serve to drastically reduce background and discriminate from jets. The track isolation requirement leads to the greatest loss in signal efficiency; by requiring only the track isolation requirement (and not the other photon identification requirements), the number of events drops from 6051 to 3827. This is due to the large number of jets in the signal events. The shower shape ($\sigma_{\text{iinj}}$) requirement and the $H/E$ requirement reduce the GGM signal efficiency the least. Note that the requirement of an isolated jet does not affect the signal efficiency very much. This is studied in detail in Section 6.6.

5.6.2 Jet Efficiency

The efficiency for identifying (JPT) jets is given in [29] as $0.961 \pm 0.011$. GGM events typically have three or more jets in them and the event selection requires only one jet (isolated from photons). For more information on photon reconstruction efficiency, particularly as a function of the number of jets in the event, see Section 6.6.
Table 5.6: Breakdown of cuts and their effects on GGM signal efficiency.

<table>
<thead>
<tr>
<th>Cut</th>
<th>Number of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Generated Events</td>
<td>9996</td>
</tr>
<tr>
<td>Two Generated photons</td>
<td>6207</td>
</tr>
<tr>
<td>Matched to reco photons</td>
<td>6051</td>
</tr>
<tr>
<td>Reco photons passing photon ID</td>
<td>2914</td>
</tr>
<tr>
<td>At least one isolated jet</td>
<td>2795</td>
</tr>
<tr>
<td>Final Selection</td>
<td>2316</td>
</tr>
</tbody>
</table>

to be present. This indicates that the efficiency for selecting GGM events due to the presence of a jet is over 99.9%. Thus, the effect of the \( \geq 1 \) jet requirement on the GGM signal efficiency is negligible. Note, however, that the presence of jets does affect the efficiency for photons and thus indirectly affects the GGM signal efficiency. Sections 5.6.3 and 6.6 contain more details on this effect.

5.6.3 Photon Efficiency

There are two studies associated with the photon efficiency. One is a study of the efficiency of photons under the identification cuts (isolation cuts as well as the \( H/E \) cut and the shower shape cut). In this case one assumes that because the electron and photon efficiencies should be similar under their nearly identical selection, one can use the efficiency of electrons to determine the efficiencies of photons. The other study of the efficiency of photons has to do with the pixel seed veto (as it is the one difference between electrons and photons and hence cannot be studied using the Tag and Probe method).

Both measurements will be discussed in the sections that immediately follow. Note that the efficiency of the HLT selection on the GGM MC sample is assumed to be nearly 100% due to the high energy of the photons coupled with the offline selection of \( E_T > 30 \text{ GeV} \). Figure 5.22 shows the the energy distributions for the leading and trailing photons in the example GGM model.
5.6.3.1 Photon Identification Efficiency Using Tag and Probe

“Tag and probe” is a widely-used method to find the efficiencies of electrons passing certain identification requirements. In this case, the efficiencies of photons in the sample is assumed to be the same as the efficiency for electrons because they differ only by the pixel seed veto requirement. A “tag” is a well-identified electron coming from a $Z$ decay. Once a tag is identified there is a high probability of finding another electron in the event (also from the $Z$ decay). This electron is known as the “probe.” A very loose selection is defined to identify the probe. Once a tag and probe are identified, their invariant mass is calculated and a fit to the $Z$ mass is performed using all tag and probe events found in the data (or Monte Carlo) sample. This fit yields an integral that gives the number of $Z$ events. After this is performed, the probe selection is made tighter by applying the identification cuts under study (in this case, HCAL, ECAL and Track isolation as well as the $H/E$ cut and the shower shape ($\sigma_{i\eta i\eta}$) cut). Again the invariant masses are calculated and a fit is performed. The efficiency for electrons (photons) passing this cut is equal to the ratio of the latter integral to the former integral.

It is important to note that these selections must be carefully defined. For instance, the passing probe selection (those probes passing the tighter selection) should be similar to the tag selection. Then, the probability of locating a tag electron in the first place is uncorrelated to
locating a passing probe. Thus, one can truly say that locating a tag electron only means that it is likely that another electron is in the sample. However, in this case the tag selection would be too loosely defined, leading to a lot of background (making the fit to the invariant $Z$ mass less precise). So, the tag selection is tighter than the standard photon identification selection used in this analysis. However, the additional cuts chosen should not affect our ability to identify the other electron from the $Z$. Similarly, the very loose probe selection should be loose in order not to cut out any $Z$ electrons, but it should be tight enough to correctly identify electrons. Otherwise, again, the invariant $Z$ mass histogram would have too much background to yield a good fit. If the selection is too tight it could bias the results because the photon identification requirements would be applied on top of a possibly biased sample in which the loose selection cuts are heavily correlated with the photon identification cuts. Thus, the following tag and probe selections were defined:

**Tag selection:**

- $p_T > 30$ GeV
- $|\eta| < 1.479$
- $\sigma_{\eta\eta} < 0.009$
- $H/E < 0.05$
- Matched to a pixel seed
- Matched to a track with $p_T > 15$ GeV with $\Delta R \leq 0.004$

**Probe selection:**

- $p_T > 30$ GeV
- $|\eta| < 1.479$
- $H/E < 0.5$
- Loosely matched to a track with $p_T > 15$ GeV with $\Delta R \leq 0.1$
Separated from a jet by $\Delta R \geq 0.9$

The efficiency for electrons to pass the tag and probe selection is defined as $\epsilon_{\text{data}}^{e}$. To get the efficiency for photons in data, this factor must be corrected by the efficiency for electrons to pass the tag and probe selection in Monte Carlo $Z \rightarrow ee$ events ($\epsilon_{\text{MC}}^{e}$) and by the efficiency for photons to pass the selection cuts in Monte Carlo photon + jet events ($\epsilon_{\gamma}^{\text{MC}}$). This latter efficiency is determined by utilizing the information in simulated Monte Carlo (“MC truth”) that indicates whether or not the reconstructed photon corresponds to photon generated from a physics process. The overall photon efficiency is:

$$\epsilon_{\gamma} = \epsilon_{\gamma}^{\text{MC}} \times \frac{\epsilon_{\text{data}}^{e}}{\epsilon_{\text{MC}}^{e}}.$$  \hspace{1cm} (5.10)

The ratio $\frac{\epsilon_{\text{data}}^{e}}{\epsilon_{\text{MC}}^{e}}$ is defined as the “scale factor” by which to modify the signal efficiency found in the MC. The signal efficiency is modified by multiplying it by the square of the scale factor, $\left(\frac{\epsilon_{\text{data}}^{e}}{\epsilon_{\text{MC}}^{e}}\right)^2$. For the numerator (electrons from data) and denominator (electrons from $Z \rightarrow ee$ MC), the tag and probe analysis is done. The invariant mass histograms of $Tag - \text{PassingProbe}$ and $Tag - \text{FailingProbe}$ are fit simultaneously with a Breit-Wigner function convoluted with a Gaussian and an exponential background shape (which is allowed to be different for $Tag - \text{PassingProbe}$ and for $Tag - \text{FailingProbe}$). A Crystal Ball fit is not used in this case due to the low statistics. The fit is performed using an extended maximum likelihood method.

The efficiency of the data electrons to pass all photon identification requirements (except the veto on pixel seed) was determined to be: $\epsilon_{\text{data}}^{e} = 0.887 \pm 0.007$. The efficiency of the MC electrons to pass the photon identification selection was determined to be $\epsilon_{\text{MC}}^{MC} = 0.900 \pm 0.001$. Plots showing the histograms and fits can be found in Figure 5.23. Hence, the global scale factor is $\frac{\epsilon_{\text{data}}^{e}}{\epsilon_{\text{MC}}^{e}} = 0.986 \pm 0.008$.

The scale factor was studied as a function of $p_T$, $\eta$, and $\Delta R$ between the photon (electron) and the nearest jet. The scale factor is flat as a function these variables and thus the global scale factor will be used in the calculation of the GGM signal efficiency. Figure 5.24 shows the scale factor as a function of these four variables.
Figure 5.23: Fits for all tag and probe pairs (including divided by passing probes and failing probes) for Z MC electrons (a) and Z data electrons (b).
5.6.3.2 Effect of pileup on $\epsilon^\text{data}_e$

To investigate the effect of pileup (more than one proton-proton interaction per crossing), the ratio $\frac{\epsilon^\text{data}_e}{\epsilon^\text{MC}_e}$ was computed for a data sample in which the number of primary vertices per event, $N_{PV}$, was equal to one, and again for all values of $N_{PV}$. The efficiency for when $N_{PV} = 1$ is $\epsilon^\text{data}_e =$
0.920 ± 0.010, which differs from the case in which all numbers of primary vertices were considered (as given above, $\epsilon_e^{\text{data}} = 0.887 ± 0.007$). For $N_{PV} = 1$, the scale factor is $\frac{\epsilon_e^{\text{data}}}{\epsilon_e^{\text{MC}}} = 1.022 ± 0.036$. Again, for all values of $N_{PV}$, $\frac{\epsilon_e^{\text{data}}}{\epsilon_e^{\text{MC}}} = 0.986 ± 0.008$. The efficiency for data as a function of $N_{PV}$ is given in Figure 5.25, along with the $Z \rightarrow ee$ MC point (for which pileup was not simulated, so there is always only one primary vertex). The difference, 0.036, is taken as a systematic uncertainty on the efficiency.

![Figure 5.25](image)

Figure 5.25: Efficiency for electron reconstruction in data as a function of primary vertices. MC $Z \rightarrow ee$ does not have any pileup simulated, hence $N_{PV} = 1$.

### 5.6.3.3 Comparison between MC Truth Efficiency and Tag and Probe Efficiencies

A comparison between efficiencies using the Tag and Probe method compared to using Monte Carlo truth information is shown in Figure 5.26. The efficiency using Monte Carlo truth information is obtained by counting the total number of reconstructed electrons that pass the photon (electron) selection and are matched to a generator electron resulting from a $Z \rightarrow ee$ decay (with $\Delta R \leq 0.3$) and dividing by the total number of reconstructed electrons matched to a generator electron (again
with $\Delta R \leq 0.3$). In Figure 5.26, the efficiency found using MC truth information is subtracted from the efficiency using Tag and Probe, and then divided by the efficiency using Tag and Probe. The data fit well to a constant in all cases, and the difference is less than 2%, which is taken as an additional systematic uncertainty on the scale factor.

![Graphs](https://via.placeholder.com/150)

Figure 5.26: Efficiency $\epsilon_e^{MC}$ determined by Tag and Probe subtracted by the efficiency $\epsilon_e^{MC}$ determined by using Monte Carlo truth information, divided by the efficiency using Tag and Probe, as a function of $p_T$ (a), $\eta$ (b), and $\Delta R$ between photon and nearest jet (c).
5.6.3.4 Check that $\epsilon_e^{MC}/\epsilon_e^{data} = \epsilon_\gamma^{MC}/\epsilon_\gamma^{data}$

The use of Tag and Probe works under the assumption (described earlier in this section) that the photon efficiency and electron efficiency should be similar under the selection cuts because they only differ by a pixel veto. This assumption is supported by Figure 5.27 which shows that the distributions of photons and electrons (from Monte Carlo $Z \rightarrow ee$ and photon + jet samples) for the selection variables under study. Again, the reconstructed electrons were matched to generator electrons with $\Delta R \leq 0.3$ coming from $Z$ decays. The photons were matched to generator photons with $\Delta R \leq 0.3$ coming from a hard scatter. The photons and electrons were required to have $p_T > 30$ GeV and $H/E < 0.5$ (to distinguish from jets). The distributions are flat in the region of interest indicating that the assumption is valid.

5.6.3.5 Pixel Seed Veto Efficiency

True photons may have pixel seeds incorrectly associated to them due to uncertainties in reconstruction. Because the amount of material that a particle passes through before reaching the calorimeters significantly affects its reconstruction efficiency, care should be made to determine the systematic uncertainty associated to the pixel seed veto efficiency due to differing material scenarios. A large amount of work has gone into accurately modeling the CMS tracker (through which the particles must travel before reaching the calorimeter). However, there are still some uncertainties associated with this modeling due to, in particular, approximations of the actual geometry of the physical volumes and of the material distributions. A study has been done to accurately determine the uncertainty associated with these issues. The relevant quantity for determining how the material in the tracker affects reconstruction is the thickness, in terms of radiation lengths $X_0$, that a particle would traverse coming from the nominal interaction point. Thus, a low-material scenario ($X_0^\text{min}$) and a high-material scenario ($X_0^\text{max}$) were used to simulate simple particle scenarios in addition to the nominal scenario ($X_0^\text{nom}$). For this study, photons in the range of $10$ GeV $< p_T < 100$ GeV are simulated in the three different material scenarios. Results for the fraction of photons having
Figure 5.27: Ratio of the electron distributions from $Z \rightarrow ee$ MC events to the photon distributions from photon + jet events for ECAL isolation (a), HCAL isolation (b), Track Isolation (c), $H/E$ (d) and $\sigma_{\eta\eta}$ (e). Photon identification cuts are given in Section 5.3.1 and can be dependent on the $p_T$ of the photon.
a mistaken pixel seed associated with them are given in Table 5.7.

Table 5.7: Photon veto loss in efficiency in different material budget scenarios.

<table>
<thead>
<tr>
<th>Geometry</th>
<th>Fraction of Photons with Seeds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum Material ($X_0^{min}$)</td>
<td>$0.0353 \pm 0.0004$</td>
</tr>
<tr>
<td>Maximum Material ($X_0^{max}$)</td>
<td>$0.0350 \pm 0.0004$</td>
</tr>
<tr>
<td>Nominal Material ($X_0^{nom}$)</td>
<td>$0.0339 \pm 0.0004$</td>
</tr>
</tbody>
</table>

The efficiency of the pixel seed veto is then given as $0.9661 \pm 0.0004 (stat) \pm 0.0014 (sys)$ where the systematic uncertainty is taken as the greater of differences between the nominal budget scenario and the alternative scenarios (in this case, the difference between the $X_0^{min}$ scenario and the $X_0^{nom}$ scenario). Note that all other photon identification cuts are identical to what is used in the analysis. Note that the efficiency of the pixel seed veto is already included in the signal efficiency calculation (because it is a requirement on the photons), but an additional systematic uncertainty due to the material budget is added to the uncertainty on the photon efficiency.

5.6.3.6 Signal Efficiency Uncertainty due to Photon Efficiency

The final photon efficiency is given as the tag and probe efficiency and its associated uncertainties ($0.986 \pm 0.008 (stat) \pm 0.041 (sys)$), along with the associated systematic uncertainty from the pixel seed veto efficiency study, 0.0014, which is negligible.

5.6.4 Jet Energy Scale Systematic

The uncertainty on the energy scale of calorimeter jets is 10% [30]. Because at least one jet of $p_T > 30$ GeV is required in the event selection, the jet energy scale uncertainty affects the uncertainty on the GGM signal efficiency. In order to study the effect on the signal efficiency, the $p_T$ cut was varied by 10% and the number of GGM MC events passing the selection was studied. Results for a few example points are given in Table 5.8. Clearly there is very little effect on the signal efficiency. For limit calculations, a conservative 2% was assigned as a systematic uncertainty on the signal efficiency due to the jet energy scale uncertainty.
Table 5.8: Variation of GGM Signal Efficiency Due to Jet Energy Scale Uncertainties. Signal Point 1 corresponds to $m_{\tilde{q}} = 640$ GeV, $m_{\tilde{g}} = 640$ GeV, $m_{\tilde{\chi}_0} = 150$ GeV. Signal Point 2 corresponds to $m_{\tilde{q}} = 400$ GeV, $m_{\tilde{g}} = 2000$ GeV, $m_{\tilde{\chi}_0} = 50$ GeV. Signal Point 3 corresponds to $m_{\tilde{q}} = 600$ GeV, $m_{\tilde{g}} = 960$ GeV, $m_{\tilde{\chi}_0} = 500$ GeV.

<table>
<thead>
<tr>
<th>Signal Point</th>
<th>Acc. (jet $p_T &gt; 27$)</th>
<th>Acc. (jet $p_T &gt; 30$ GeV)</th>
<th>Acc. (jet $p_T &gt; 33$ GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.245 ± 0.004</td>
<td>0.249 ± 0.004</td>
<td>0.249 ± 0.004</td>
</tr>
<tr>
<td>2</td>
<td>0.185 ± 0.004</td>
<td>0.184 ± 0.004</td>
<td>0.184 ± 0.004</td>
</tr>
<tr>
<td>3</td>
<td>0.141 ± 0.004</td>
<td>0.140 ± 0.004</td>
<td>0.140 ± 0.004</td>
</tr>
</tbody>
</table>

5.6.5 Parton Distribution Function Systematic Uncertainty on GGM Signal Efficiency Systematic

There are theoretical and experimental uncertainties associated with parton distribution functions (PDFs), which describe the momentum of partons inside of hadrons. The use of different PDFs affects the calculation of cross sections, cross section uncertainties, signal efficiencies and signal efficiency uncertainties. There are a number of available PDFs for use that exploit fits to data from Drell-Yan events, multijet events, and deep-inelastic scattering. The LHC has recommended a particular method for calculating the uncertainties at NLO and next-to-next-to-leading-order (NNLO) due to differences in PDFs [39]. The method involves taking the three most common PDFs that use results from the Tevatron and fixed target experiments: NNPDF [40], MRST2001 [41], CTEQ6 [42] and assigning each event in the Monte Carlo a weight corresponding to the values and uncertainties in the PDF associated with the initial parton distribution (using the method in [43]). These weights are then used to determine the rate and GGM signal efficiency by determining the number of reweighted events before and after the selection. For each PDF, these weights yield an uncertainty on the cross section and the GGM signal efficiency. An “envelope” method is used where the uncertainty on the GGM signal efficiency is given by:

$$s = \frac{1}{2}(max(x_1 + s_1, x_2 + s_2, x_3 + s_3) - min(x_1 - s_1, x_2 - s_2, x_3 - s_3))$$ (5.11)

where $x_1$, $x_2$, and $x_3$ are the GGM signal efficiencies corresponding to each of the three different PDFs, $s_1$, $s_2$ and $s_3$ are their uncertainties, and “min” and “max” specify taking the maximum or minimum of the three values within the parenthesis.
The selection was made at the generator level and based on two photons with \( p_T > 30 \text{ GeV} \) and one jet with \( p_T > 30 \text{ GeV} \). Uncertainties on the GGM signal efficiencies were calculated for each GGM grid point and can be seen in Figure 5.28. For the majority of the squark and gluino phase space the uncertainties are < 5\%. Only for \( m_{\tilde{q}} > 1600 \text{ GeV} \) and \( m_{\tilde{g}} > 1000 \text{ GeV} \) do these uncertainties become significant. Regardless, uncertainties on the GGM signal efficiencies due to the PDFs for each grid point were added to the total uncertainty on the GGM signal efficiency used in setting limits.

![Graph showing uncertainties on GGM signal efficiencies due to PDF in percent.](image)

Figure 5.28: Uncertainties on GGM signal efficiencies due to PDF in percent.

### 5.6.6 Final GGM Signal Efficiencies

The final GGM signal efficiencies are given in Figure 5.29. Note the signal efficiencies are correlated with the number of jets, as shown in Figure 5.4. This is due to the effect of the jets on photon isolation, as discussed in Section 5.6.1 and more thoroughly in Section 6.6.
Figure 5.29: GGM Signal Efficiencies for $m_{\tilde{\chi}_0} = 50$ GeV (a), $m_{\tilde{\chi}_0} = 150$ GeV (b), and $m_{\tilde{\chi}_0} = 500$ GeV (c).

5.7 Determination of Upper Limit Contours for GGM

In order to set limits or exclusions on GGM models it is necessary to know the expected background, the luminosity ($35.5 \pm 3.9$ pb$^{-1}$) [32], and the GGM efficiency. From Bayes’s theorem,
the probability of the cross section $\sigma$ as a function of $N$ observed events can be written:

$$p(\sigma|N) = \frac{L(N|\sigma) \pi(\sigma)}{\int L(N|\sigma') \pi(\sigma') d\sigma'}$$

(5.12)

where $L(N|\sigma)$ is the likelihood to observe $N$ events as a function of $\sigma$ and $\pi(\sigma)$ is the prior probability distribution function (pdf)$^7$ that describes the probability of $\sigma$. The denominator normalizes $p(\sigma|N)$ to one. A flat prior pdf is used which describes the state of knowledge of $\sigma$:

$$\pi(\sigma) = \begin{cases} 
0 & \sigma < 0 \\
1 & \sigma \geq 0 
\end{cases}$$

(5.13)

The shape of this prior (flat) indicates that cross sections less than 0 have zero probability of being true (they are unphysical) and all greater than 0 are equally possible.

$L(\sigma|N)$ depends on the three parameters mentioned above: luminosity ($l$), background ($b$), and GGM signal efficiency ($\epsilon$). These parameters are called “nuisance” parameters because while they are not of direct interest to the analysis, they must be correctly accounted for in the analysis. Each of these nuisance parameters is a positive value and is modeled by a pdf centered about its nominal value and with a standard deviation equal to the parameter uncertainty ($\pi(l)$, $\pi(b)$, and $\pi(\epsilon)$, respectively). In order to determine the sensitivity of the upper limit to these nuisance parameter pdfs, three different shapes were used: Gaussian, log-normal, and gamma. $L(\sigma|N)$ can be written as:

$$L(N|\sigma) = \int L(N|\sigma, l, b, \epsilon) \pi(l) \pi(b) \pi(\epsilon) \, dl \, db$$

(5.14)

and

$$L(N|\sigma, l, b, \epsilon) = \frac{(\sigma l + b)^N}{N! e^{-(\sigma l + b)}}$$

(5.15)

which is the Poisson distribution. With one signal candidate observed after all selection cuts, a 95% confidence level upper limit on the cross section can be determined by the equation:

$$0.95 = \int_0^{CL95} p(\sigma|1) d\sigma = \frac{\int_0^{CL95} (\sigma l + b) e^{-(\sigma l + b)} \pi(l) \pi(b) \pi(\epsilon) \, dl \, db \, d\sigma}{\int_0^\infty (\sigma l + b) e^{-(\sigma l + b)} \pi(l) \pi(b) \pi(\epsilon) \, dl \, db \, d\sigma'}$$

(5.16)

where $CL95$ is the upper limit for the cross section.

7 A “prior” pdf is one reflects the knowledge and uncertainty of a variable before the data is taken into account and it is the cornerstone of a Bayesian analysis.
5.7.1 Example of Upper Limit Calculation

Taking the example of the typical GGM point with $m_{\tilde{q}} = 640$ GeV, $m_{\tilde{g}} = 640$ GeV, and $m_{\tilde{\chi}^0} = 150$ GeV, the GGM signal efficiency after all efficiency corrections and uncertainty estimation, is $\epsilon = 0.1944 \pm 0.0036 \text{(stat)} \pm 0.0071 \text{(sys)}$. The total background estimation using fake-fake as a QCD control sample is $0.54 \pm 0.35$, while the background estimation using $Z \rightarrow ee$ as the QCD control sample is $1.79 \pm 0.77$. To get the best estimate on the background, an average of the results for the $Z \rightarrow ee$ and fake-fake is used. Note that the uncertainty on the two backgrounds are correlated because they are both normalized to the same diphoton signal candidate sample. This uncertainty is 14% on the background. Before the estimations are averaged, this uncertainty must be subtracted from each sample before they are averaged and then added back in. By modeling the two different (non-negative) backgrounds with a log-normal distribution, the combined background estimation is $1.11 \pm 0.76$. Note that a conservative systematic uncertainty of 0.68 was added to the uncertainty on this value and is the maximum of the difference between the combined estimate and the two individual estimates. Using the luminosity of $35.5 \pm 3.9 \text{ pb}^{-1}$, the example upper limit is shown in Table 5.9 for the different pdf and background estimations.

Table 5.9: Upper Limit Example for GGM Example Point using three different prior pdfs for the nuisance parameters: Gaussian, log-normal, and Gamma. Cross sections are in pb.

<table>
<thead>
<tr>
<th>Type of Background Sample</th>
<th>Background Events</th>
<th>Gaussian pdf</th>
<th>log-normal pdf</th>
<th>Gamma pdf</th>
</tr>
</thead>
<tbody>
<tr>
<td>fake-fake</td>
<td>$0.54 \pm 0.35$</td>
<td>$0.603 \text{ pb}$</td>
<td>$0.587 \text{ pb}$</td>
<td>$0.577 \text{ pb}$</td>
</tr>
<tr>
<td>$Z \rightarrow ee$</td>
<td>$1.79 \pm 0.77$</td>
<td>$0.552 \text{ pb}$</td>
<td>$0.531 \text{ pb}$</td>
<td>$0.527 \text{ pb}$</td>
</tr>
<tr>
<td>Combined</td>
<td>$1.11 \pm 0.76$</td>
<td>$0.577 \text{ pb}$</td>
<td>$0.558 \text{ pb}$</td>
<td>$0.550 \text{ pb}$</td>
</tr>
<tr>
<td>No background</td>
<td>$0.0 \pm 0.0$</td>
<td>$0.658 \text{ pb}$</td>
<td>$0.643 \text{ pb}$</td>
<td>$0.641 \text{ pb}$</td>
</tr>
</tbody>
</table>

The differences among the different prior pdfs can be attributed to their shapes: the Gaussian distribution has the sharpest tails while the log-normal distribution has longer tails. The longer the tails the more variation in the uncertainties; hence the less conservative limits. Still, the variation between the different prior pdfs for the nuisance parameters is negligible, so to calculate the upper

\[^8\] If a random variable $x$ is log-normally distributed, then $\ln x$ is normally distributed. For a normal distribution, the average of two variables is given by $\mu = (\mu_1/w_1 + \mu_2/w_2)/(w_1 + w_2) \pm 1/\sqrt{w_1 + w_2}$ where $w_1 = 1/\sigma_1^2$ and likewise for $w_2$ [38]
limits for the entire grid of GGM points, the log-normal pdf will be used. Furthermore, in order to make the most accurate limit, the combined measurement of the $Z \rightarrow ee$ and fake-fake background will be used as the prediction on the background. A summary of the critical parameters for the upper limit calculations is given in Table 5.10. The most significant uncertainty associated with the upper limit calculation comes from the uncertainty on the luminosity. Results for the upper limits are given in Figure 5.30.

Table 5.10: Critical parameters for determining upper limits and exclusions on GGM models.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminosity</td>
<td>35.5 pb$^{-1}$</td>
<td>3.9 pb$^{-1}$</td>
<td>0.04%</td>
<td>0.14%</td>
</tr>
<tr>
<td>Pixel seed veto efficiency</td>
<td>96.53%</td>
<td>0.15%</td>
<td>0.042</td>
<td>0.041</td>
</tr>
<tr>
<td>Photon scale factor</td>
<td>0.986</td>
<td>2.0%</td>
<td>0.008</td>
<td>0.041</td>
</tr>
<tr>
<td>JET/MET energy scale</td>
<td>0.216</td>
<td>0.019</td>
<td>0.004</td>
<td>0.019</td>
</tr>
<tr>
<td>Typical GGM Signal Efficiency</td>
<td>1.11</td>
<td>0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Candidate Signal Events</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.8 Exclusions of GGM Regions

Depending on the upper limit found for the particular points in the GGM squark-gluino-neutralino mass phase space, certain exclusions can be made given the theoretical cross section for these points. Consider the “GGM hypothesis” that signal events come from two sources: background with expected rate $1.11\pm0.76$ and GGM SUSY with an expected rate based on the calculated cross section (with uncertainties). The number of observed events follows a Poisson distribution. The probability to observe one or fewer events (the number in the candidate sample) under the GGM hypothesis is calculated. If this probability is less than 5% the hypothesis is excluded. The necessary cross sections and k-factors for the GGM phase space needed for this calculation were given in Figures 5.2 and 5.3. To properly calculate the exclusion regions, the uncertainties on the cross sections must be included.
Figure 5.30: Upper Limits in the squark-gluino mass plane for $m_{\tilde{\chi}_0} = 50 \text{ GeV}$ (a), $m_{\tilde{\chi}_0} = 150 \text{ GeV}$ (b), and $m_{\tilde{\chi}_0} = 500 \text{ GeV}$ (c).
5.8.1 PDF Uncertainties on Cross Section

The uncertainties on the cross section due to the PDF uncertainty are calculated in a similar way as for the GGM signal efficiency (see Section 5.6.5) and shown in Figure 5.31.

![Figure 5.31: Uncertainties on calculated cross sections due to PDF (in percent).](image)

5.8.2 Normalization Uncertainty on Cross Section

There are additional uncertainties on the cross sections due to uncertainties in the renormalization of the QCD Lagrangian. These uncertainties are determined by changing the renormalization scale in PROSPINO by a factor of 2 and comparing the differences in theoretical cross sections. The value obtained for this uncertainty is 13% on the cross sections [33].

5.8.3 Exclusion Results

For the example GGM point, the upper limit was found to be 0.647 pb. This particular GGM point has a theoretical cross section of 1.6 pb with a k-factor of 1.44. The uncertainty on the cross section due to the PDF uncertainties is 12%. Thus, the exclusion was calculated using a cross section value of $\sigma_{\text{example}} = 2.31 \pm 0.41$ pb. With the corresponding GGM signal efficiency, this means an expected number of events of $17.1 \pm 3.7$ events. The probability of observing 1 or fewer events was calculated as $9.5 \times 10^{-5}$, well below 5%, so this point is excluded. In a similar
manner the rest of the GGM points were calculated. Results for the exclusion values are given in Figure 5.32. The exclusion contours are given in Figure 5.33.

Figure 5.32: Exclusion values in the squark-gluino mass plane for $m_{\tilde{\chi}_0} = 50$ GeV (a), $m_{\tilde{\chi}_0} = 150$ GeV (b), and $m_{\tilde{\chi}_0} = 500$ GeV (c).
Figure 5.33: Exclusion contours for the three neutralino masses. Anything below the contour lines is excluded.
Chapter 6

Investigations Related to Analysis

6.1 Introduction

During the course of a standard analysis it becomes necessary to investigate certain effects to determine if the sample selection and analysis methods are appropriate. The contents of this chapter explore issues that were brought up during the analysis and the subsequent approval period. In addition, the last section describes the results of doing this analysis on Monte Carlo to further check the validity of the method and to determine how well the Monte Carlo reproduces the data.

6.2 Investigation of Shower Shape on sample composition

In this analysis there are very few candidate events. The “tight” photon identification selection described in Section 5.3.1 differs from the “loose” CMS photon identification in two ways: the shower shape cut is not imposed and the track isolation is slightly loosened:

- $E_{T}^{\text{Sum}}$ of Tracks in track isolation cone $< 0.001 \cdot E_{T} + 3.5 \text{ GeV}$
- $\sigma_{\eta\eta} < 0.013$ dropped

The track isolation cut and shower shape cut serve to discriminate against jets which, unlike photons, contain tracks and have larger electromagnetic showers. To understand if loosening the photon identification requirements leads to a gain in signal events, the signal sample composition is analyzed to determine if it includes primarily real diphoton events, real photon plus fake (jet)
events, or real fake-fake (jet-jet) events. By comparing the signal composition with the tight cut and the loose cut, it can be determined if loosening the cut improves the analysis. Because the most stringent cut imposed is the shower shape ($\sigma_{\eta \eta}$) cut (as opposed to the track isolation cut), that is investigated first. Then, a loosening of the track isolation requirement is done and again the signal composition can be determined.

As a first look, Figure 6.1 shows the value of $\sigma_{\eta \eta}$ for the leading and trailing electromagnetic objects $Z \rightarrow ee$ sample and the fake-fake sample. Similarly, Figure 6.2 shows the value of $\sigma_{\eta \eta}$ for the leading and trailing photons in the data candidate sample. Because the diphoton distribution in the data candidate sample is similar to the dielectron sample in the $Z \rightarrow ee$ distribution, it is likely that loosening the cut will lead to a significant increase in background with only a few additional signal events.

![Figure 6.1: $\sigma_{\eta \eta}$ plot of leading vs. trailing electron for the $Z \rightarrow ee$ control sample (a) and $\sigma_{\eta \eta}$ plot of leading vs. trailing fake for fake-fake sample (b). In both cases the presence of a jet is required.](image)

To quantify the composition of the candidate diphoton data sample, a system of equations is used to represent event counts and efficiencies, employing three different data samples: the signal diphoton sample, the $Z \rightarrow ee$ control sample, and the fake-fake control sample. The situation is slightly complicated by the fact that the leading and trailing objects may have different efficiencies for passing the shower shape cut, but this just leads to a larger system of equations. Three
reasonable assumptions are capitalized on: that the $Z \rightarrow ee$ sample is dominated by $Z$ production with negligible contribution from backgrounds, that the fake-fake sample is completely dominated by QCD with negligible contamination from direct diphoton production (or other sources of true electromagnetic objects), and that the efficiency for a cut on $\sigma_{\eta \eta}$ is the same for electrons and photons.

Each sample can be divided into four sub samples: $N_{pp}$ (both objects pass the shower shape cut), $N_{pf}$ (only the leading object passes the shower shape cut), $N_{fp}$ (only the trailing object passes the shower shape cut), $N_{ff}$ (both objects fail the shower shape cut). Here, the first subscript refers to the leading object, and the second subscript to the trailing object. Considering the $Z \rightarrow ee$ sample, the following system of equations can be written:

- $N_{pp} = \epsilon_1 \epsilon_2 N_{Z-true}$
- $N_{pf} = \epsilon_1 (1 - \epsilon_2) N_{Z-true}$
- $N_{fp} = (1 - \epsilon_1) \epsilon_2 N_{Z-true}$
- $N_{ff} = (1 - \epsilon_1)(1 - \epsilon_2) N_{Z-true}$
\( N_{ff} = (1 - \epsilon_1)(1 - \epsilon_2)N_{Z\text{-true}} \)

where \( \epsilon_1 \) is the efficiency for the leading electron to pass the shower shape cut, and \( \epsilon_2 \) is the efficiency for the trailing electron to pass the shower shape cut (these two efficiencies will be similar). \( N_{Z\text{-true}} \) is the number of actual \( Z \) decays before applying the cut. This system of equations can be solved to determine \( \epsilon_1 \) and \( \epsilon_2 \). Similarly, for the fake-fake sample:

\( N_{pp} = f_1 f_2 N_{ff\text{-true}} \)
\( N_{pf} = f_1 (1 - f_2) N_{ff\text{-true}} \)
\( N_{fp} = (1 - f_1) f_2 N_{ff\text{-true}} \)
\( N_{ff} = (1 - f_1)(1 - f_2) N_{ff\text{-true}} \)

where \( f_1 \) and \( f_2 \) are the efficiencies for the leading and trailing fakes (respectively) to pass the shower shape cut and \( N_{ff\text{-true}} \) is the number of true fake-fake events. Unlike the case of electrons, it is not likely that these two efficiencies will be similar, but again, \( f_1 \) and \( f_2 \) can be individually solved for.

Now consider the candidate diphoton sample. Events with two photons that both pass the shower shape cut can actually be composed of real diphoton events, photon-fake events, and fake-fake events depending on whether each object failed or passed the shower shape cut. The same statement applies for events with one leading photon that passes the shower shape cut and one trailing photon that does not (and hence is a fake), and one leading photon that does not pass the shower shape cut and one trailing that does pass the cut, and for the sample in which both photons fail the cut (and are hence identified as fakes). Thus,

\( N_{pp} = \epsilon_1 \epsilon_2 N_{\gamma\gamma} + \epsilon_1 f_2 N_{\gamma \text{-fake}} + f_1 \epsilon_2 N_{\text{fake-}\gamma} + f_1 f_2 N_{\text{fake-fake}} \)
\( N_{pf} = \epsilon_1 (1 - \epsilon_2) N_{\gamma\gamma} + \epsilon_1 (1 - f_2) N_{\gamma \text{-fake}} + f_1 (1 - \epsilon_2) N_{\text{fake-}\gamma} + f_1 (1 - f_2) N_{\text{fake-fake}} \)
\( N_{fp} = (1 - \epsilon_1) \epsilon_2 N_{\gamma\gamma} + (1 - \epsilon_1) f_2 N_{\gamma \text{-fake}} + (1 - f_1) \epsilon_2 N_{\text{fake-}\gamma} + (1 - f_1) f_2 N_{\text{fake-fake}} \)
\[ N_{ff} = (1 - \epsilon_1)(1 - \epsilon_2)N_{\gamma\gamma} + (1 - \epsilon_1)(1 - f_2)N_{\gamma\text{--fake}} + (1 - f_1)(1 - \epsilon_2)N_{\text{fake-\gamma}} + (1 - f_1)(1 - f_2)N_{\text{fake-fake}}. \]

Here, \( N_{\gamma\gamma} \) is the number of true diphoton events, \( N_{\gamma\text{--fake}} \) is the number of true photon-fake events, and \( N_{\text{fake-fake}} \) is the number of true fake-fake events. Do not confuse this with \( N_{pp} \) which is the number of diphoton events in our candidate sample (where both pass the shower shape cut) (which corresponds to 87 events in the standard analysis). Similarly, \( N_{pf} \) is the number of events in which the leading object passes all photon identification cuts and the trailing object passes all photon identification cuts except for the shower shape cut and similarly for \( N_{fp} \) and \( N_{ff} \). Note that this notation can be slightly confusing as a photon (or electron) which fails the shower shape cut is identified as a fake.

The above system of equations can be solved for the four unknowns \( (N_{\gamma\gamma}, N_{\gamma\text{--fake}}, N_{\text{fake-\gamma}}, N_{\text{fake-fake}}) \) and thus used to determine the fraction of true diphoton events in the candidate sample (as well as the fraction of photon-fake and fake-fake). This can be compared to the fractions in an event sample composed not of two “tight” photon identification but of one “tight” photon and one “loose” photon. Thus, it can be determined if loosening the cut helps increase true diphoton event yields without significantly increasing the background in the sample. The raw numbers for event counts are given in Table 6.1. Using these numbers, the efficiencies are given as \( \epsilon_1 = 1.00 \pm 0.00 \), \( \epsilon_2 = 0.99 \pm 0.00 \), \( f_1 = 0.65 \pm 0.02 \), and \( f_2 = 0.69 \pm 0.02 \). Thus, the final event composition can be determined and is given in Table 6.2.

Table 6.1: Raw event counts for different samples in which the leading and trailing objects either pass or fail the \( \sigma_{\text{inj}} < 0.013 \) cut.

<table>
<thead>
<tr>
<th>Category</th>
<th>ee sample</th>
<th>ff sample</th>
<th>( \gamma\gamma ) sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both objects pass ( \sigma_{\text{inj}} ) cut</td>
<td>588</td>
<td>167</td>
<td>87</td>
</tr>
<tr>
<td>Only leading object passes ( \sigma_{\text{inj}} ) cut</td>
<td>5</td>
<td>74</td>
<td>8</td>
</tr>
<tr>
<td>Only trailing objects passes ( \sigma_{\text{inj}} ) cut</td>
<td>0</td>
<td>90</td>
<td>0</td>
</tr>
<tr>
<td>Both objects fail ( \sigma_{\text{inj}} ) cut</td>
<td>0</td>
<td>36</td>
<td>0</td>
</tr>
</tbody>
</table>

Thus, by loosening the requirement on the photon identification by dropping the shower shape cut, the background composition of the diphoton candidate sample increases. Of the eight
Table 6.2: Composition of diphoton candidate sample using standard selection (both photons must pass shower shape cut) and loosened selection (one photon may fail shower shape cut).

<table>
<thead>
<tr>
<th>Photon Cuts</th>
<th>True $\gamma\gamma$ fraction</th>
<th>True $\gamma$-fake fraction</th>
<th>True fake-fake fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tight-Tight</td>
<td>0.808</td>
<td>0.192</td>
<td>0.0</td>
</tr>
<tr>
<td>Tight-Loose</td>
<td>0.746</td>
<td>0.254</td>
<td>0.0</td>
</tr>
</tbody>
</table>

additional events gained in the diphoton candidate sample, only one is likely to be a true diphoton event. This small increase in signal comes at the expense of large additional background and thus there is no advantage to loosening the photon identification requirements.

This analysis was repeated using the looser track isolation requirement. In this case, dropping the shower shape requirement again only serves to increase the background composition of the sample as shown in Table 6.3. Furthermore, loosening the track isolation requirement dramatically increases the background composition of the sample.

Table 6.3: Composition of diphoton candidate sample using standard selection (both photons must pass shower shape cut) and loosened selection (one photon may fail shower shape cut). In both cases, the track isolation requirement was loosened to $E_T$ of Tracks in track isolation cone $< 0.001 \cdot E_T + 3.5$ GeV.

<table>
<thead>
<tr>
<th>Photon Cuts</th>
<th>True $\gamma\gamma$ fraction</th>
<th>True $\gamma$ - fake fraction</th>
<th>True fake-fake fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tight-Tight</td>
<td>0.746</td>
<td>0.254</td>
<td>0.0</td>
</tr>
<tr>
<td>Tight-Loose</td>
<td>0.660</td>
<td>0.340</td>
<td>0.0</td>
</tr>
</tbody>
</table>

6.3 Investigation of different MET Algorithms

This analysis uses tcMET, as described in the Section 4.7. A different type of MET can also be calculated, called “pfMET” and is based on a system of reconstruction known as “particle flow.” Unlike normal reconstruction, particle flow reconstructions aims to take the entire event topology into consideration before objects are ultimately identified. The goal of particle flow reconstruction is to provide a single list of reconstructed particles in an event, which can be of type: photon, hadron (neutral or charged), muon, or electron. These objects constitute a complete description of an event and are used as input to higher-level reconstructed objects such as jets and
MET. The particle flow algorithm starts with calorimeter clustering, tracking with extrapolation to calorimeters, and standard muon identification and electron pre-identification (from standard reconstruction). This information is then combined to topologically link the elements before a final particle list is produced. Because this type of reconstruction process can yield different kinematic values for varying objects, one should compare values of pfMET to values of tcMET to see if there are any large differences. This comparison is shown in Figure 6.3 where the two different MET values for the candidate diphoton sample are reasonably correlated for this level of statistics. Thus, there are likely no large systematic errors introduced by using standard reconstruction as opposed to particle flow reconstruction. In addition, the single event which passed tcMET > 50 GeV also passed pfMET > 50 GeV and no other events passed pfMET > 50 GeV.

![pfMET vs. tcMET for γγ events with ≥ 1 jet requirement](image)

Figure 6.3: pfMET vs. tcMET for γγ events with ≥ 1 jet requirement.
6.4 Comparison of MET Spectra for fake-fake and $Z \to ee$ events

Because there are two different control samples used for determining the QCD MET spectrum from the data (using either fake-fake or $Z \to ee$ events), a natural question is how well the MET shapes for these two control samples agree. The left plot in Figure 6.4 shows the MET distribution for these two samples, normalized and overlaid (without any diEM $p_T$ reweighting). A Kolmogorov-Smirnov test yields a value of 0.685, indicating that the two samples are in relatively good agreement. The right plot in Figure 6.4 is the ratio of the normalized MET distribution of the $Z \to ee$ sample to the fake-fake sample. A fit of this ratio to a constant yields a value of $0.96 \pm 0.06$ indicating good agreement (with a $\chi^2$ per degree of freedom of 0.35).

Figure 6.4: The QCD MET distribution from $Z \to ee$ and fake-fake samples normalized to unity and overlaid (a) and the ratio of the normalized MET distribution of the $Z \to ee$ sample to the fake-fake sample (b).

6.5 Leading Jet $\eta$ Distributions

The analysis selection requires photons to be well-separated from jets ($\Delta R$ between photon and jet must be $\geq 0.9$). As the photons are also required to be central ($|\eta| < 1.479$), the concern arises that events in which the jets are located in the endcaps are being preferentially selected. This could be a problem because GGM jets are centrally located. To determine if this is the case, the leading jet $\eta$ distributions between the candidate data sample ($\gamma\gamma$), the GGM MC sample, and the
two different QCD control samples ($Z \rightarrow ee$ and fake-fake) are compared in Figure 6.5. There is no evidence that the cuts are forcing the jets to be in the endcaps. Furthermore, the distributions among the various samples are similar, particularly for the $\gamma\gamma$ candidate sample as opposed to the signal GGM MC.

![Figure 6.5: Top left: Leading jet $\eta$ for the fake-fake sample. Top right: Leading jet $\eta$ for the $Z \rightarrow ee$ sample. Bottom left: Leading jet $\eta$ for the $\gamma\gamma$ sample. Bottom right: Leading jet $\eta$ for the GGM example point ($\gamma\gamma$ selection).](image)

6.6 Study of $\Delta R$ between Jet and Photons

One concern for the analysis is whether or not the photon reconstruction efficiency is affected by the condition that at least one jet be present in the event, and that this jet be separated by $\Delta R \geq 0.9$ from both the leading and trailing photon. GGM events have many jets in them, so
the concern is not that the signal efficiency will drop due to the requirement of the presence of a jet, but that the signal efficiency will drop because the jet is required to be separated from the photons. Note that this cut does not require all jets to be separated from the photons — only one separated jet needs to exist in the event. As the number of jets increases in an event, the probability of at least one separated jet being in the event will only increase. The real issue is how the $\Delta R$ requirement affects photon reconstruction efficiency, and whether or not this leads to a loss in acceptance of GGM events because photons are lost.

First, for reference, in Figure 6.6 is a scatter plot for the $\Delta R$ between the leading photon and all jets versus the trailing photon and all jets for the $\gamma\gamma$ candidate data sample (without the isolation requirements). There is a clustering on both axes - for these small values of $\Delta R$ between the photon and the jet, the photon will most likely fail the isolation requirements. The clustering at $\Delta R_1 = 0, \Delta R_2 = \pi$ are due to back-to-back dijet events. The region of $\Delta R = 0.9$ is also shown where the photons will likely pass the selection.

The photon reconstruction efficiencies as a function of $\Delta R$ to jets can be examined by looking at various distributions for a typical signal MC sample ($m_{\tilde{q}} = 640$ GeV, $m_{\tilde{g}} = 640$ GeV, $m_{\tilde{\chi}_0} = 150$ GeV). As shown in Figure 6.7, the proximity of the leading photon in the event to the closest jet depends on the multiplicity of jets in the event (the more jets, the more likely one will be close to a photon). Basic $p_T$ and $\eta$ cuts were imposed to mimic the final selection. Jet selection is described in Section 5.3.4. Reconstructed jets were matched to generated jets by requiring that they be matched to within $\Delta R \leq 0.4$ and $\Delta p_T/p_T \leq 3.0$.

Furthermore, the photon reconstruction efficiency depends on the proximity to the nearest jet. However, as shown in Figure 6.8, it is largely flat for $\Delta R \geq 0.7$ where it falls off sharply because of the isolation requirements on the reconstructed photon, see Figure 6.8. In this case, if the generated photon matched a reconstructed photon within $\Delta R = 0.1$ and the reconstructed photon passed the isolation requirements, it is considered properly reconstructed. Thus, the requirement that a jet must be separated from the photons by $\Delta R = 0.9$ ensures that the photon is properly isolated. This value of $\Delta R$ is chosen because although there is significant efficiency for $\Delta R \leq 0.9$, a cut of
Figure 6.6: Scatter plot of $\Delta R$ between the leading photon and the jets versus the trailing photon and the jets for the $\gamma\gamma$ candidate data sample.

$\Delta R \geq 0.9$ ensures that the photon isolation cones and the jet cones do not overlap at all. Also, equally important, note that as the number of jets increase, the photon reconstruction efficiency for a given proximity to a jet does not depend on the multiplicity of jets in the event as can be seen by comparing the various sub-figures of Figure 6.8. The efficiency versus $\Delta R$ is independent of the number of jets. however, the higher the jet count, the more likely there will be a low $\Delta R$ between the photons and a jet, and are therefore less likely to pass the selection. Thus when the probability that a photon is near a jet is high, the efficiency for selecting that event is low, regardless of what kind of event it is (supersymmetric, QCD, etc.).

Finally, in Figure 6.9, the acceptance of the typical GGM signal region as a function of
The number of jets is given for both the separation requirement of $\Delta R = 0.9$ of the jet from the photons and for no separation requirement. As is clear, the acceptance depends very little on the separation requirement except in the case where the number of generated jets is equal to one. This is because in the case of only one jet in the event, if this jet is not separated from the photons, there are no other jets in the event that could satisfy the jet requirement. With the higher numbers of jets in the event, the higher the probability that there will be at least one jet that is separated from photons in the event. However, the photons in these events have a lower probability of passing the
Figure 6.8: Photon reconstruction efficiency versus $\Delta R$ for the leading generated photon in the event. Top row is $N_{jets} = 1$ (left) and $N_{jets} = 2$ (right). Middle row is $N_{jets} = 3$ (left) and $N_{jets} = 4$ (right). Bottom row is $N_{jets} = 5$ (left) and $N_{jets} = 6$ (right).

... photon isolation requirements due to their proximity to other (non-separated) jets in the events; thus the overall signal efficiency drops for higher numbers of jets.

6.7 Monte Carlo Analysis

Repeating the analysis on Monte Carlo (MC) is a useful tool for further validating the analysis method and to determine how well the MC reproduces the data. A full suite of Standard Model MC samples were used for this study; the samples were simulated and reconstructed in the same software release as used for the data analysis (CMSSW_3.8.3). The samples are listed in Table 6.4.
along with the number of events generated and the cross section which yields a weight for each event that passes the selection; because of this there are occasional spikes in a distribution if a high-weight event passes the selection.

The determination of the background for the $MET$ distribution was repeated on this MC sample as in Section 5.4.\footnote{Note that the electron-photon fake rate as determined from data in Section 5.4.2 was used. It is not expected to vary significantly from the corresponding MC result.} The $MET$ distributions for both the $Z \rightarrow ee$ QCD control sample and the fake-fake QCD control sample are shown in Figure 6.10 and 6.11 respectively. The corresponding predictions of events for $MET > 50$ GeV are given in Table 6.5. Comparing this to Table 5.5 shows that the results are in good agreement within errors and that the MC is a good representation of the data for this method, particularly for the requirement of $\geq 1$ jet. Note also that there is no signal GGM MC in the plot other than the separate distribution (shown). The $MET$ distribution of the $\gamma\gamma$ sample matches that of the QCD control samples. For the case of the requirement of
Table 6.4: Datasets used for MC Analysis

<table>
<thead>
<tr>
<th>Monte Carlo Sample</th>
<th>Events Generated</th>
<th>Cross-Section (pb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCD (30 – 50 GeV)</td>
<td>3264660</td>
<td>5312370</td>
</tr>
<tr>
<td>QCD (50 – 80 GeV)</td>
<td>3191546</td>
<td>6359119</td>
</tr>
<tr>
<td>QCD (80 – 120 GeV)</td>
<td>3208299</td>
<td>784265</td>
</tr>
<tr>
<td>QCD (120 – 170 GeV)</td>
<td>3045200</td>
<td>115134</td>
</tr>
<tr>
<td>QCD (170 – 300 GeV)</td>
<td>3220080</td>
<td>24263</td>
</tr>
<tr>
<td>QCD (300 – 470 GeV)</td>
<td>3171240</td>
<td>1168</td>
</tr>
<tr>
<td>QCD (470 – 600 GeV)</td>
<td>2019732</td>
<td>70.2</td>
</tr>
<tr>
<td>QCD (600 – 800 GeV)</td>
<td>1979055</td>
<td>15.55</td>
</tr>
<tr>
<td>QCD (800 – 1000 GeV)</td>
<td>2084404</td>
<td>1.84</td>
</tr>
<tr>
<td>QCD (1000 – 1400 GeV)</td>
<td>1086966</td>
<td>0.332</td>
</tr>
<tr>
<td>QCD (1400 – 1800 GeV)</td>
<td>1021510</td>
<td>0.0109</td>
</tr>
<tr>
<td>QCD (&gt; 1800 GeV)</td>
<td>529360</td>
<td>0.00036</td>
</tr>
<tr>
<td>Photon + Jets (20 – 40 GeV)</td>
<td>8194252</td>
<td>31910</td>
</tr>
<tr>
<td>Photon + Jets (40 – 100 GeV)</td>
<td>2217101</td>
<td>23620</td>
</tr>
<tr>
<td>Photon + Jets (100 – 200 GeV)</td>
<td>1061602</td>
<td>3476</td>
</tr>
<tr>
<td>Photon + Jets (&gt; 200 GeV)</td>
<td>1142171</td>
<td>485</td>
</tr>
<tr>
<td>Diphoton + Jets</td>
<td>1080060</td>
<td>134</td>
</tr>
<tr>
<td>Diphoton (Box) (10 – 25 GeV)</td>
<td>792710</td>
<td>358</td>
</tr>
<tr>
<td>Diphoton (Box) (25 – 250 GeV)</td>
<td>768815</td>
<td>12.4</td>
</tr>
<tr>
<td>Diphoton (Box) (&gt; 250 GeV)</td>
<td>790685</td>
<td>0.000208</td>
</tr>
<tr>
<td>Diphoton (Born) (10 – 25 GeV)</td>
<td>523270</td>
<td>236</td>
</tr>
<tr>
<td>Diphoton (Born) (25 – 250 GeV)</td>
<td>536230</td>
<td>22.4</td>
</tr>
<tr>
<td>Diphoton (Born) (&gt; 250 GeV)</td>
<td>541900</td>
<td>0.008</td>
</tr>
<tr>
<td>W + Photon (to eν)</td>
<td>541698</td>
<td>100</td>
</tr>
<tr>
<td>W + Photon (to µν)</td>
<td>537351</td>
<td>100</td>
</tr>
<tr>
<td>W + Photon (to τν)</td>
<td>543779</td>
<td>100</td>
</tr>
<tr>
<td>W + Jets (to lν)</td>
<td>14818245</td>
<td>24380</td>
</tr>
<tr>
<td>Z + Photon (to eeγ)</td>
<td>325267</td>
<td>27</td>
</tr>
<tr>
<td>Z + Photon (to ττγ)</td>
<td>319541</td>
<td>27</td>
</tr>
<tr>
<td>Z + Photon (to ννγ)</td>
<td>382126</td>
<td>25</td>
</tr>
<tr>
<td>TT + Jets</td>
<td>1394548</td>
<td>94</td>
</tr>
<tr>
<td>Drell-Yan (to ee)</td>
<td>2127607</td>
<td>1300</td>
</tr>
</tbody>
</table>

≥ 1 jet, a Kolmogorov-Smirnov test on the agreement between the MC candidate diphoton sample and the MC fake-fake QCD control sample gives a value of 0.959 and a Kolmogorov-Smirnov test on the agreement between the MC candidate diphoton sample and the MC $Z \rightarrow ee$ sample gives a value of 0.575, indicating good agreement in both cases. This is a further indication of the validity of the method. The agreement for the case of no jet requirement is not as good. Because the QCD samples are more poorly modeled in low energy regions, for the the very high cross section events,
there are not enough MC events to make the MC weighting valid.

The composition of the $\gamma\gamma$ sample is given in Tables 6.6 and 6.7. The composition was determined by noting how many events passed the selection for each given sample (although no other MC generator information was used). For all values of $MET$, the primary components are Diphoton + Jets and Photon + Jets events. For $MET > 50$ GeV, the primary component of the sample is QCD. The reason for this is the mismeasurement of the jets, as described in Section 5.4.1.

Table 6.5: MC Analysis: Estimation of Background and Candidate Events for $MET > 50$ GeV

<table>
<thead>
<tr>
<th>Type</th>
<th>$\geq 1$ jet</th>
<th>no jet cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Candidate (diphoton) Events</td>
<td>0.38 ± 0.27</td>
<td>0.44 ± 0.28</td>
</tr>
<tr>
<td>Fake-Fake QCD Background Estimate</td>
<td>1.55±1.33</td>
<td>3.05±1.80</td>
</tr>
<tr>
<td>$Z \to ee$ QCD Background Estimate</td>
<td>0.69±0.85</td>
<td>0.87±0.95</td>
</tr>
<tr>
<td>$e\gamma$ EW Background Estimate</td>
<td>0.03±0.02</td>
<td>0.05±0.04</td>
</tr>
<tr>
<td>Total Background (using fake-fake)</td>
<td>1.58±1.33</td>
<td>3.10±1.80</td>
</tr>
<tr>
<td>Total Background (using $Z \to ee$)</td>
<td>0.72±0.85</td>
<td>0.93±0.95</td>
</tr>
<tr>
<td>Predicted GGM1 Events</td>
<td>10.27±2.20</td>
<td>10.28±2.20</td>
</tr>
<tr>
<td>Predicted GGM1 Events (with K factor of 1.66)</td>
<td>17.1±3.7</td>
<td>17.1±3.7</td>
</tr>
</tbody>
</table>

Table 6.6: Composition of MC $\gamma\gamma$ sample for no jet requirement as determined by number of events from each sample passing the selection requirements.

<table>
<thead>
<tr>
<th>Monte Carlo Sample</th>
<th>$\gamma\gamma$ Events ($MET &gt; 0$ GeV)</th>
<th>$\gamma\gamma$ Events ($MET &gt; 50$ GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCD</td>
<td>6.50</td>
<td>0.31</td>
</tr>
<tr>
<td>Photon + Jets</td>
<td>117.11</td>
<td>0.02</td>
</tr>
<tr>
<td>Diphoton + Jets</td>
<td>120.33</td>
<td>0.05</td>
</tr>
<tr>
<td>Diphoton (Box and Born)</td>
<td>71.63</td>
<td>0.02</td>
</tr>
<tr>
<td>W + Photon</td>
<td>0.20</td>
<td>0.03</td>
</tr>
<tr>
<td>W + Jets</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>Z + Photon</td>
<td>0.21</td>
<td>0.01</td>
</tr>
<tr>
<td>TT + Jets</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>Drell-Yan (to ee)</td>
<td>0.14</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>316.19</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Comparing various distributions between the data and the MC for various selections (either fake-fake or $Z \to ee$) can indicate how well the QCD MC models the data. These distributions for the $\geq 1$ jet case are shown in Figures 6.12, 6.13, 6.14, 6.15, 6.16, 6.17, 6.18, 6.19, 6.20. Note that
Figure 6.10: Monte Carlo Analysis: The $Z \rightarrow ee$ MET spectrum (rewighted and normalized) for the QCD MET background prediction, the $e\gamma$ MET spectrum (rescaled according to the $e\gamma$ fake rate) for the EW MET background prediction, the diphoton candidate MET spectrum with no jet requirement (a) and $\geq 1$ jet requirement (b). The Monte Carlo GGM example point spectrum is also shown for comparison. The last bin contains the overflow events.
Figure 6.11: Monte Carlo Analysis: The fake-fake MET spectrum (rewighted and normalized) for the QCD MET background prediction, the $e\gamma$ MET spectrum (rescaled according to the $e\gamma$ fake rate) for the EW MET background prediction, the diphoton candidate MET spectrum with no jet requirement (a) and $\geq 1$ jet requirement (b). The Monte Carlo GGM example point spectrum is also shown for comparison. The last bin contains the overflow events.
Table 6.7: Composition of MC $\gamma\gamma$ sample for no jet requirement as determined by number of events from each sample passing the selection requirements.

<table>
<thead>
<tr>
<th>Monte Carlo Sample</th>
<th>$\gamma\gamma$ Events ($MET &gt; 0$)</th>
<th>$\gamma\gamma$ Events ($MET &gt; 50$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCD</td>
<td>5.45</td>
<td>0.28</td>
</tr>
<tr>
<td>Photon + Jets</td>
<td>21.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Diphoton + Jets</td>
<td>33.72</td>
<td>0.04</td>
</tr>
<tr>
<td>Diphoton (Box and Born)</td>
<td>6.52</td>
<td>0.00</td>
</tr>
<tr>
<td>W + Photon</td>
<td>0.10</td>
<td>0.02</td>
</tr>
<tr>
<td>W + Jets</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Z + Photon</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>TT + Jets</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Drell-Yan (to $ee$)</td>
<td>0.20</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>67.05</td>
<td>0.38</td>
</tr>
</tbody>
</table>

unphysical spikes in the MC distributions are due to the necessity of weighting the events. The agreement is fairly good, in particular for the $Z \rightarrow ee$ sample.
Figure 6.12: For events with $\geq 1$ jet, data versus MC $E_T$ distribution of leading two fakes for fake-fake sample (a) and leading two electrons for $Z \rightarrow ee$ (b).
Figure 6.13: For events with $\geq 1$ jet, data versus MC $\eta$ distribution of leading two fakes for fake-fake (a) and leading two electrons for $Z \rightarrow ee$ (b).
Figure 6.14: For events with $\geq 1$ jet, data versus MC $\phi$ distribution of leading two fakes for fake-fake (a) and leading two electrons for $Z \to ee$ (b).
Figure 6.15: For events with $\geq 1$ jet, data versus MC ECAL Isolation distribution of leading two fakes for fake-fake (a) and leading two electrons for $Z \rightarrow ee$ (b).
Figure 6.16: For events with $\geq 1$ jet, data versus MC HCAL isolation distribution of leading two fakes for fake-fake (a) and leading two electrons for $Z \rightarrow ee$ (b).
Figure 6.17: For events with $\geq 1$ jet, data versus MC Track Isolation distribution of leading two fakes for fake-fake (a) and leading two electrons for $Z \to ee$ (b).
Figure 6.18: For events with $\geq 1$ jet, data versus MC $H/E$ distribution of leading two fakes for fake-fake (a) and leading two electrons for $Z \rightarrow ee$ (b).
Figure 6.19: For events with \( \geq 1 \) jet, data versus MC \( \sigma_{\eta \eta} \) distribution of leading two fakes for fake-fake (a) and leading two electrons for \( Z \rightarrow ee \) (b).
Figure 6.20: For events with $\geq 1$ jet, data versus MC $d\text{i}EMpT$ distribution for fake-fake (a) and leading two electrons for $Z \rightarrow ee$ (b).
Chapter 7

Conclusion

The previous chapters described the search for evidence of General Gauge Mediation by an examination of $35.5 \pm 3.9 \text{ pb}^{-1}$ of LHC proton-proton collision data recorded by CMS. The search focused on diphoton events with at least one jet and comparing the \textit{MET} distribution of these events to one predicted by background. No evidence of GGM was found. It is shown that only one candidate diphoton event with $\text{MET} > 50 \text{ GeV}$ passed the selection which is consistent with the predicted background ($0.54 + 0.35 - 0.35$ events using the difake control sample or $1.79 + 0.77 - 0.73$ using the $Z \rightarrow ee$ control sample). Combining these results yields an estimated background of $1.11 \pm 0.76$. Given the predicted background and number of observed events, upper limits were calculated for a number of GGM signal points corresponding to different squark, gluino, and neutralino masses as shown in Figure 5.30. Given the theoretical cross-sections (and their corresponding errors), certain GGM signal points can be excluded; the exclusion contours are given in Figure 5.33.

While discovery of GGM would certainly have been exciting — putting upper limits on cross-sections for models and the ability to exclude certain models is an important step in the search for new physics. The LHC intends to deliver another 1000 pb$^{-1}$ or more of data to the experiments in 2011. With this much data, perhaps GGM can be discovered or excluded entirely. Regardless of the outcome, experimentalists and theorists will continue to collaborate for years to come in their quest to find an accurate model of the four fundamental interactions.
Bibliography

Appendix A

Glossary

**Absorption band.** Range of wavelengths which are able to excite a particular energy transition in a material. See Section 3.2.2.2.

**APD.** Avalanche Photodiode. Used as the photodetector for the barrel section of the Electromagnetic Calorimeter. See Section 3.2.2.3.

**ASIC.** Application-specific integrated circuits: a type of integrated circuit used in the L1 trigger. See Section 3.2.6.

**Avalanche multiplication.** A behavior in which an electron, upon being freed in an avalanche photodiode, is accelerated by the high applied bias voltage, freeing other electrons in its path, thereby causing more multiplication. See Section 3.2.2.3.

**BPIX.** Barrel pixel detector, part of the tracking system. See Section 3.2.1.1.

**C26000 brass.** A standard brass alloy used for absorbing plates in the HE. See Section 3.2.3.

**CSC.** Cathode Strip Chambers. Part of the muon detector system. See Section 3.2.5.

**Dark current.** The response of a receptor of radiation during periods where it is not exposed to radiation. See Section 3.2.2.3.

**DAQ.** Data Acquisition system. See Section 3.2.7.

**DT.** Drift Tubes. Part of the muon detector system. See Section 3.2.5.

**EB.** The Barrel section of the ECAL. See Section 3.2.2.1.

**ECAL.** The CMS Electromagnetic Calorimeter. See Section 3.2.2.
**EE.** The Endcap section of the ECAL. See Section 3.2.2.1.

**EMF.** The electromagnetic fraction of energy of a jet. This quantity is used in jet selection, see Section 5.3.4.

**EW.** Electroweak. The framework or interaction the describes the electromagnetic and weak interactions. See Chapter 2.

**fHPD.** Fraction of energy contributed by the highest energy hybrid photodetector readout to the total HCAL energy of the jet. For a description of hybrid photodetectors, see Section 3.2.3. This quantity is used in jet selection, see Section 5.3.4.

**FPIX.** Forward pixel detector, part of the tracking system. See Section 3.2.1.1.

**FPGA.** Field Programmable Gate Array: a type of integrated circuit used in the L1 trigger. See Section 3.2.6.

**GGM.** General Gauge Mediation Supersymmetry Breaking. A more general theory than GMSB than can contain more than one messenger generation, more than one breaking scale, etc. See Chapter 2.

**GMSB.** Gauge Mediated Supersymmetry Breaking. See Chapter 2.

**HB.** Hadron Barrel calorimeter. See Section 3.2.3.

**HE.** Hadron Endcap calorimeter. See Section 3.2.3

**HF.** Hadron Forward calorimeter. See Section 3.2.3

**HO.** Hadron Outer calorimeter. See Section 3.2.3

**HCAL.** The CMS Hadronic Calorimeter. See Section 3.2.3

**IP.** Interaction Point, the location where two colliding LHC protons interact.

**JPT.** Jets Plus Tracks. Refers to the type of jet algorithm used to reconstruct jets on CMS. See Section 4.6.

**L1.** The Level 1 trigger. See Section 3.2.6

**LHC.** The Large Hadron Collider. See Section 3.1.

**LUT.** Programmable Memory Lookup table: a type of integrated circuit used in the L1 trigger. See Section 3.2.6
MET. Missing Transverse Energy. MET is the magnitude of the negative vector sum of the momentum transverse to the beam axis of all final-state particles. See Section 4.7.

Moliere radius. The transverse dimension of an energetic shower, defined by the radius of a cylinder containing 90% of the shower’s energy deposition. See Section 3.2.2.

N90. The minimum number of ECAL and HCAL cells required to contain 90% of the jet energy. This quantity is used in jet selection, see Section 5.3.4. For a description of ECAL and HCAL cells, see Sections 3.2.2 and 3.2.3.

NbTi. Niobium Titanium, the material used for the superconducting solenoid and also the LHC magnets. See Chapter 3.

PDF. Parton Distribution Function. See Section 5.6.5.


QCD. Quantum Chromodynamics, the framework that describes the strong interaction. See Chapter 2.

Photoelectrons. Electrons emitted by the photoelectric effect. See Section 3.2.2.2.

Radiation length. A characteristic of a material that relates to a characteristic amount of matter transversed by electromagnetic interacting particles. Electrons generally lose energy in matter through bremsstrahlung while photons loose energy by pair production. See Section 3.2.2.2.

RPC. Resistive Plate Chambers. Part of the muon detector system. See muon section of Detector Chapter.

SM. The Standard Model, the theoretical framework that describes the electromagnetic, weak, and strong interactions of elementary particles. See Chapter 2.

SUSY. Supersymmetry, a proposed extension to the Standard Model which introduces a symmetry between bosons and fermions. See Chapter 2.

TCS. Trigger Control System. The part of the trigger system which takes into account DAQ and subdetector readiness. See Section 3.2.6.

TEC. Tracker Endcap. See Section 3.2.1.

TIB. Tracker Inner Barrel. See Section 3.2.1.
**TID.** Tracker Inner Disks. See Section 3.2.1.

**TOB.** Tracker Outer Barrel. See Section 3.2.1.

**Tower.** A calorimeter object that consists of one or more HCAL cells and the ECAL crystals that correspond geometrically. See Section 4.

**TPG.** Trigger Primitive Generator. Local (detector-specific) trigger. See Section 3.2.6.

**TTC.** Timing, Trigger, Control system. The part of the trigger system tells the subdetectors whether or not to pass on more high-resolution information. See Section 3.2.6.

**VME Crate.** A VME crate is an enclosure that uses VME bus computer architecture. It stands for “VERSA Module Eurocard,” a standard computer data path. See Section 3.2.2.4.

**VPT.** Vacuum Phototriode. Used as the photodetector for the endcap section of the Electromagnetic Calorimeter. See Section 3.2.2.3.
Appendix B

\textit{ee and }e\gamma\textit{ Invariant Mass Histograms in Various Bins}

The plots on the following pages are invariant mass histograms of the di-electron and electron-photon sample data samples in a variety of $\eta$ and $p_T$ bins, as discussed in Section 5.4.2. The integrals of fits to these histograms resulted in Figures 5.16 and 5.17. For $p_T > 60$ GeV in the $e\gamma$ sample, the integrals were determined by counting events with invariant mass between 80 GeV and 100 GeV (due to very low statistics).

Figure B.1: The invariant mass of two electron candidates for $0.0 < |\eta| < 0.3$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).
Figure B.2: The invariant mass of two electron candidates for $0.3 < |\eta| < 0.6$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).

Figure B.3: The invariant mass of two electron candidates for $0.6 < |\eta| < 0.9$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).
Figure B.4: The invariant mass of two electron candidates for $0.9 < |\eta| < 1.2$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).

Figure B.5: The invariant mass of two electron candidates for $1.2 < |\eta| < 1.5$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).
Figure B.6: The invariant mass of two electron candidates for $30 \text{ GeV} < p_T < 45 \text{ GeV}$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).

Figure B.7: The invariant mass of two electron candidates for $45 \text{ GeV} < p_T < 60 \text{ GeV}$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).
Figure B.8: The invariant mass of two electron candidates for 60 GeV < \( p_T < 75 \) GeV. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).

Figure B.9: The invariant mass of two electron candidates for 75 GeV < \( p_T < 75 \) GeV. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).
Figure B.10: The invariant mass of one electron and one photon candidate for $0.0 < |\eta| < 0.3$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).

Figure B.11: The invariant mass of one electron and one photon candidate for $0.3 < |\eta| < 0.6$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).
Figure B.12: The invariant mass of one electron and one photon candidate for $0.6 < |\eta| < 0.9$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).

Figure B.13: The invariant mass of one electron and one photon candidate for $0.9 < |\eta| < 1.2$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).
Figure B.14: The invariant mass of one electron and one photon candidate for $1.2 < |\eta| < 1.5$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).

Figure B.15: The invariant mass of one electron and one photon candidate for $30 \text{ GeV} < p_T < 45 \text{ GeV}$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).
Figure B.16: The invariant mass of one electron and one photon candidate for $45 \text{ GeV} < p_T < 60 \text{ GeV}$. The Z peak is fit using a Breit-Wigner convoluted with a Gaussian with constant background (a) and linear background (b).

Figure B.17: The invariant mass of one electron and one photon candidate for $60 < p_T < 75$ (a) and $p_T > 75 \text{ GeV}$ (b). Due to low statistics, the integrals of these fits were taken to be the total number of entries between 80 GeV and 100 GeV.
Appendix C

Kinematic Plots

The plots on the following pages show many kinematic distributions for the three data samples (γγ, fake-fake, ee) and, in some cases, for the Monte Carlo GGM example sample (with γγ selection applied). Such comparisons are useful for understanding the candidate and control samples.
Figure C.1: Leading electron vs. trailing electron distributions in the $Z \rightarrow ee$ data sample with no jet requirement (left) and $\geq 1$ jet requirement (right) for $E_T$ (top), $\eta$ (middle), and $\phi$ (bottom).
Figure C.2: Leading photon vs. trailing photon distributions in the candidate diphoton data sample with no jet requirement (left) and $\geq 1$ jet requirement (right) for $E_T$ (top), $\eta$ (middle), and $\phi$ (bottom).
Figure C.3: Leading electron or photon vs. trailing electron or photon distributions in the $e\gamma$ data sample with no jet requirement (left) and $\geq 1$ jet requirement (right) for $E_T$ (top), $\eta$ (middle), and $\phi$ (bottom).
Figure C.4: Leading fake vs. trailing fake distributions in the fake-fake data sample with no jet requirement (left) and ≥1 jet requirement (right) for $E_T$ (top), $\eta$ (middle), and $\phi$ (bottom).
Figure C.5: Leading electron vs. trailing electron distributions in the $Z \rightarrow ee$ data sample with no jet requirement (left) and $\geq 1$ jet requirement (right) for $\sigma_{\eta\eta}$ (top), Time (ns) (middle), and Energy (GeV) (bottom).
Figure C.6: Leading photon vs. trailing photon distributions for the candidate diphoton data sample with no jet requirement (left) and ≥ 1 jet requirement (right) for $\sigma_{\eta \eta}$ (top), Time (ns) (middle), and Energy (GeV) (bottom).
Figure C.7: Leading electron or photon vs. trailing electron or photon in the $e\gamma$ data sample with no jet requirement (left) and $\geq 1$ jet requirement (right) for $\sigma_{\eta \eta}$ (top), Time (ns) (middle), and Energy (GeV) (bottom).
Figure C.8: Leading fake vs. trailing fake distributions in the fake-fake data sample with no jet requirement (left) and $\geq$ 1 jet requirement (right) for $\sigma_{\eta_{\text{HI}}}$ (top), Time (ns) (middle), and Energy (GeV) (bottom).
Figure C.9: ECAL isolation for leading object in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.10: HCAL isolation for leading object in ee data sample (top), eγ data sample (second from top), fake − fake data sample (middle), γγ data sample (second from bottom) and γγ GGM MC sample (bottom) with no jet requirement (left) and ≥ 1 jet requirement (right).
Figure C.11: Track isolation for leading object in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.12: Sum of isolations for leading object in $ee$ data sample (top), $e\gamma$ data sample (second from top), fake − fake data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.13: ECAL isolation for trailing object in ee data sample (top), eγ data sample (second from top), fake$-$fake data sample (middle), γγ data sample (second from bottom) and γγ GGM MC sample (bottom) with no jet requirement (left) and ≥ 1 jet requirement (right).
Figure C.14: HCAL isolation for trailing object in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake-fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.15: Track isolation for trailing object in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.16: Sum of isolations for trailing object in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.17: Number of jets in ee data sample (top), e\(\gamma\) data sample (second from top), fake–fake data sample (middle), \(\gamma\gamma\) data sample (second from bottom) and \(\gamma\gamma\) GGM MC sample (bottom) with no jet requirement (left) and \(\geq 1\) jet requirement (right).
Figure C.18: Invariant mass in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.19: Number of vertices in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma \gamma$ data sample (second from bottom) and $\gamma \gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.20: $p_T$ (GeV) of leading jet in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake$–$fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.21: $H_T$ (GeV) in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.22: $t_{c\text{MET}}$ (GeV) of leading jet in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake-fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.23: $\Delta R$ between the two electromagnetic objects in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.24: $\Delta \phi$ between the two electromagnetic objects in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.25: $\Delta \eta$ between the two electromagnetic objects in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.26: diEM $p_T \left( \sum \vec{p}_T \right)$ of the two electromagnetic objects in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.27: $\Delta \phi$ between the leading jet and MET in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).
Figure C.28: $\Delta \phi$ between the leading electromagnetic object and MET in $ee$ data sample (top), $e\gamma$ data sample (second from top), $fake - fake$ data sample (middle), $\gamma\gamma$ data sample (second from bottom) and $\gamma\gamma$ GGM MC sample (bottom) with no jet requirement (left) and $\geq 1$ jet requirement (right).