Title
Estimation of the Invisible Z Background to Hadronic Supersymmetry Searches Performed With Proton-Proton Collision Data at 7 and 8 TeV Observed With the CMS Detector During the First run of the CERN Large Hadron Collider

Permalink
https://escholarship.org/uc/item/7cm2n3ks

Author
Sturdy, Jared

Publication Date
2013

Peer reviewed|Thesis/dissertation
Estimation of the Invisible Z Background to Hadronic Supersymmetry Searches Performed With Proton-Proton Collision Data at 7 and 8 TeV Observed With the CMS Detector During the First Run of the CERN Large Hadron Collider

A Dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Jared Todd Sturdy

August 2013

Dissertation Committee:

Professor J. William Gary, Chairperson
Distinguished Professor Gail Hanson
Professor Steve Wimpenny
The Dissertation of Jared Todd Sturdy is approved:

Committee Chairperson

University of California, Riverside
Acknowledgments

Obtaining a Ph.D. is not a solitary task. We truly build on the work of those who came before us. We gain from their insights, learn from their failings and are inspired by their discoveries. Occasionally we are even able to thank them personally. There are many people to be thanked upon the completion of this thesis. From undergraduate research advisors, graduate school battery mates, professors, research mentors, collaborators, friends, and family. All helped support me, guide me, push me and drive me to completion. To the Bethel Physics department, for giving me my start in physics and providing the foundation of my current career. To the University of Michigan CERN REU program (class of 2005) for giving me my first experience at CERN. To my advisor, for supporting my study and life abroad, and for pouring over my thesis pages, helping to greatly improve the quality of the document! Dick, your enthusiasm is surpassed only by your knowledge of calorimetry, thank you for imparting a fraction of your knowledge to me when we worked together! To Ken, thank you for your kind and tireless patience with my questions and mistakes during our many collaborations through the years. I’ve learned much from you! To Sudan, your cool head, eleven minute videos, and quick corrections of my mistakes were undeniably useful! Thank you for keeping me mostly on task. To my other collaborators, who in large part or small, have answered questions, guided discussions, and enabled me to complete this thesis. To my teatime and lunchtime companions over the years, our conversations about any topic (always related back to genes) were always interesting! To those who joined me on my many globe-trotting adventures, helping to scratch the travel itch that I frequently get and to refresh and restore sanity, your companionship and friendship were greatly appreciated!
To my many climbing partners, I will forever cherish our rock friendships, yip yip yip!

To softball teammates, I always looked forward to Sunday games (probable rainouts) and the camaraderie the plucky Quarks exhibited! To Saturday morning soccer games, in the heat of the Riverside sun, played on forbidden fields! To the Polish mafia, thank you for welcoming me into your ranks! I’m honored and blessed to have known you and lived with some of you! Thank you especially to Mateusz, for the many hikes and climbs, and even a few dangerous positions we put ourselves in! You’ve made this trip memorable! To my family, whose love and support cannot be overstated. Thank you also for the frequent free lodging and airport pickups and dropoffs!

Finally, to the amazing Frenchwoman I met in the middle of my journey, thank you for pushing me, you may not know, but you continually make me desire to be better! I hope you continue to push me!
To my parents who birthed me, gave me life, and taught me. Without their influence I would not have become a wanderer.

To all the wonderful people I met along the way, your friendships have truly enriched my journey.

To Mélanie, you make me want to be better.
Estimation of the Invisible Z Background to Hadronic Supersymmetry Searches Performed With Proton-Proton Collision Data at 7 and 8 TeV Observed With the CMS Detector During the First Run of the CERN Large Hadron Collider

by

Jared Todd Sturdy

Doctor of Philosophy, Graduate Program in Physics
University of California, Riverside, August 2013
Professor J. William Gary, Chairperson

In searches for supersymmetry (Susy) in all-hadronic channels, events with jets and a Z boson form an irreducible background when the Z boson decays to a pair of neutrinos. For R-parity conserving Susy models, every decay chain involving a superparticle must result in at least one lightest supersymmetric particle (LSP), which can be neutral and noninteracting. The detector signature for the LSP is identical to that of the two neutrinos from the Z, which appear as an imbalance in the transverse momentum of the event. The characterization of these types of events is crucially important in any search for new physics performed in a multi-jets+missing transverse momentum channel. A method for estimating this irreducible background using events with a measured photon is presented, and the results for a search performed in the all-hadronic multi-jet channel are shown to be consistent with standard model expectations. Limits are set on the mass of expected new particles in various models.
Table of Contents

Acknowledgements iv
Dedication vi
Abstract vii
Table of Contents viii
List of Figures xii
List of Tables xv

I Introduction and physics motivation 1

1 Introduction 2

2 Physics background 7
  2.1 Standard model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7
  2.1.1 QCD interactions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
  2.1.2 Electroweak interactions . . . . . . . . . . . . . . . . . . . . . . . . . 8
    2.1.2.1 Higgs mechanism . . . . . . . . . . . . . . . . . . . . . . . 11
  2.2 Supersymmetry . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
    2.2.1 Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
    2.2.1.1 R-parity and a candidate for dark matter . . . . . . . . . . 12
    2.2.2 Supersymmetric models . . . . . . . . . . . . . . . . . . . . . . . . . 13
    2.2.3 Search strategies . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14

II Experimental setup and event reconstruction 16

3 Experimental setup 17
  3.1 The LHC accelerator . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
    3.1.1 Particle acceleration . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
### 3.2 The Compact Muon Solenoid

- **3.2.1 The tracking system**
  - 3.2.1.1 Pixel detector
  - 3.2.1.2 Microstrip detector
- **3.2.2 The electromagnetic calorimeter**
  - 3.2.2.1 Design and layout
  - 3.2.2.2 Energy resolution
- **3.2.3 The Hadron calorimeter (HCal)**
  - 3.2.3.1 Design and layout
  - 3.2.3.2 Readout
  - 3.2.3.3 Energy resolution
- **3.2.4 The muon system**
- **3.2.5 The magnet system**
- **3.2.6 The trigger system**
  - 3.2.6.1 The level-1 trigger (L1)
  - 3.2.6.2 The high-level trigger (HLT)
- **3.2.7 Data-taking performance**

### 4 Event simulation and reconstruction

- **4.1 Event simulation**
- **4.2 Global event reconstruction (particle flow)**
  - 4.2.1 Muons ($\mu^\pm$)
  - 4.2.2 Electrons ($e^\pm$)
  - 4.2.3 Charged and neutral hadrons ($\pi^\pm, \pi^0, K^\pm, K^0, \eta$, etc.)
  - 4.2.4 Photons ($\gamma$)
  - 4.2.4.1 Isolation
  - 4.2.5 Jets
    - 4.2.5.1 Jet finding and clustering
    - 4.2.5.2 Pileup subtraction
    - 4.2.5.3 Jet energy corrections
  - 4.2.6 Missing energy
  - 4.2.7 Pileup

### III Analysis of all-hadronic events in the context of a search for supersymmetry

### 5 All-hadronic SUSY search

- **5.1 Motivation**
- **5.2 Signal regions and event selection**
  - 5.2.1 Datasets used and trigger selections
  - 5.2.2 Suppression of backgrounds to the search regions
    - 5.2.2.1 Suppression of the QCD background
    - 5.2.2.2 Suppression of top and electroweak processes ($tt\text{+jets}$ or $W^{\pm}\text{+jets}$)
5.3 Event cleaning .................................................. 59
5.4 Backgrounds to hadronic SUSY searches ..................... 59
  5.4.1 QCD background estimation .................................. 60
    5.4.1.1 Rebalance and smear (R+S) method ................. 60
  5.4.2 $W^\pm + \text{jets}$ and $t\bar{t}+\text{jets}$ background estimation .......... 61
    5.4.2.1 Lost lepton background .................................. 61
    5.4.2.2 Hadronic $\tau$ background ............................ 62

6 Estimating the background from $Z\to\nu\bar{\nu}+\text{jets}$ events using a $\gamma + \text{jets}$ sample 65
  6.1 Motivation .................................................. 65
  6.2 Photon event selection ....................................... 71
  6.3 Photon reconstruction efficiency measurement ............... 76
    6.3.1 Data/MC scale factor ..................................... 80
  6.4 Photon purity measurement ................................... 85
    6.4.1 7 TeV purity measurement: combined isolation variable ...... 86
    6.4.2 8 TeV purity measurement: $\sigma_{inj}$ Variable .......... 88
  6.5 Analysis of uncertainties ................................. 96
    6.5.1 7 TeV results ........................................... 96
    6.5.2 8 TeV analysis .......................................... 97
  6.6 Prediction of the $Z\to\nu\bar{\nu} + \text{jets}$ background ........... 100
    6.6.1 7 TeV results ........................................... 103
    6.6.2 8 TeV results .......................................... 103

7 Results and interpretation ..................................... 109
  7.1 Combined background estimates .............................. 109
    7.1.1 7 TeV analysis ......................................... 109
    7.1.2 8 TeV analysis ......................................... 110
  7.2 Interpretation of results ................................... 110
  7.3 Interpretation within the CMSSM ............................ 116
  7.4 Interpretation with simplified models ....................... 117
    7.4.1 Introduction to simplified models ....................... 117
      7.4.1.1 T1qqqq ($\tilde{g}$ pair production) .................. 119
      7.4.1.2 T2qq ($\tilde{q}$ pair production) .................... 120
    7.4.2 Exclusion curves ....................................... 120

IV Appendix, Glossary, Acronyms, Bibliography .................. 124

Glossary ......................................................... 125

Acronyms .......................................................... 127

Bibliography ........................................................ 129

Figure Sources ...................................................... 135
A Supplementary information

A.1 Event cleaning ................................................................. 136
A.2 $Z/\gamma$ Feynman diagrams ............................................. 140
A.3 Data/simulation comparisons ............................................. 144
   A.3.1 7 TeV analysis ............................................................ 144
   A.3.2 8 TeV analysis ............................................................ 148
A.4 Additional tables for the prediction of the $Z\rightarrow\nu\bar{\nu} + \text{jets}$ background .... 156

B Service work

B.1 HCAL logical map ............................................................ 164
B.2 CMS HCAL operations ..................................................... 165
B.3 CMS central operations ................................................... 167
B.4 Jet/\slashed{E}_T data quality monitoring ................................. 167
## List of Figures

1.1 The fundamental particles and interactions in the Standard Model . . . . . . 4

2.1 Some leading-order interactions in quantum chromodynamics. Here $g$ refers to a gluon, $q$ to a quark, and $\bar{q}$ to an antiquark. ................................. 9

2.2 Interactions in the electroweak standard model. ................................. 10

2.3 Feynman diagram showing how a SUSY (sfermion) term can cancel a divergent (fermion) term from the standard model. ................................. 12

3.1 The Large Hadron Collider (LHC) accelerator system, showing the four experiments and the different accelerators used to get protons into the LHC ring ................................................................. 18

3.2 An LHC dipole magnet cross section view, showing the two separate beam pipes as well as the cryogenic system ................................. 21

3.3 Layout and design of the Compact Muon Solenoid (CMS) detector, showing the location and size of the various subdetectors, the solenoid magnet, and the beam pipe ................................................................. 22

3.4 Tracking detector layout in the $r$-$z$ plane for one quarter of the detector . 24

3.5 Tracking detector layout showing the coverage in $\eta$ .......................... 25

3.6 Typical cascades as electrons ($e^\pm$) and photons ($\gamma$) pass through the ECAL. 27

3.7 Layout of the electromagnetic calorimeter (ECAL) ........................... 29

3.8 Layout of the hadron calorimeter (HCAL) ...................................... 32

3.9 Readout electronics system of the HCAL subdetector showing the data path from detector to readout electronics to counting house electronics to DAQ system ................................................................. 34

3.10 Muon detectors, showing the locations of the RPCs, CSCs and DTs . . . 35

3.11 The CMS solenoid magnet. ......................................................... 36

3.11 The CMS solenoid magnet. ......................................................... 37

3.12 Plots showing the luminosity performance of CMS and the LHC ........ 40

4.1 A transverse slice of the CMS detector, showing the various subdetectors, the orientation and strength of the magnetic field, and the trajectories for various particle types. ......................................................... 43

4.2 Pileup distribution for 2012 run .................................................. 51
5.1 τ response template ............................................. 64

6.1 Comparison of the leading-order Feynman diagram calculated with MAD-
Graph resulting in Z/γ production ............................................. 66
6.2 Comparison of some leading-order Feynman diagrams calculated with MAD-
Graph resulting in Z+4 jets and γ+4 jets (i.e., four additional partons) ............................................. 67
6.3 Comparison of the \( H_T \) distributions between \( Z\rightarrow\nu\bar{\nu}+\text{jets} \) events and γ+jets events ............................................. 67
6.4 Boson \( p_T \) and \( H_T \) distribution comparisons ............................................. 68
6.5 7 TeV data vs. simulation comparisons for event \( H_T \) and \( \not{H}_T \) ............................................. 76
6.6 7 TeV data vs. simulation comparisons for several photon variables ............................................. 77
6.7 8 TeV data vs. simulation comparisons for several event variables ............................................. 78
6.8 8 TeV data vs. simulation comparisons for several photon variables ............................................. 79
6.9 7 TeV γ acceptance and reconstruction efficiency vs. \( H_T \) and \( \not{H}_T \) ............................................. 81
6.10 8 TeV γ acceptance, reconstruction ID, pixel veto, and isolation efficiency vs. \( N_{\text{jets}} \) (in the 3 \( \not{H}_T \) bins). ............................................. 82
6.10 (Continued) 8 TeV γ acceptance, reconstruction ID, pixel veto, and isolation efficiency vs. \( N_{\text{jets}} \) (in the 3 \( \not{H}_T \) bins). ............................................. 83
6.11 Sample of tag-and-probe (T&P) fits performed for calculating efficiencies ............................................. 84
6.12 Pileup dependence of the efficiency of the pixel seed veto ............................................. 84
6.13 Signal template and data distribution with fitted sum, showing signal and
background contributions ............................................. 87
6.14 Photon selection purity as a function of \( N_{\text{jets}} \) ............................................. 91
6.15 Barrel photon purity fits for \( \not{H}_T > 200 \), in different \( N_{\text{jets}} \) bins ............................................. 92
6.15 (Continued) Barrel photon purity fits for \( \not{H}_T > 200 \), in different \( N_{\text{jets}} \) bins ............................................. 93
6.16 Endcap photon purity fits for \( \not{H}_T > 200 \), in different \( N_{\text{jets}} \) bins ............................................. 94
6.16 (Continued) Endcap photon purity fits for \( \not{H}_T > 200 \), in different \( N_{\text{jets}} \) bins ............................................. 95
6.17 \( \mathcal{P}_{Z(\mu^+\mu^-)/\gamma} \) as a function of \( N_{\text{jets}} \), \( \not{H}_T > 200 \text{ GeV}, H_T > 500 \text{ GeV} \) ............................................. 100

7.1 Data vs. data-driven background estimation for search variables and search bins, 7 TeV analysis ............................................. 112
7.2 Data vs. data-driven background estimation comparison with associated uncertainties for the 8 TeV analysis ............................................. 115
7.3 Exclusion contours in \( (M_0-M_{1/2}) \) and \( M_{\tilde{g}}-M_{\tilde{q}} \) planes ............................................. 118
7.4 Feynman diagrams for the two considered SMS topologies ............................................. 119
7.5 Exclusion contours for the T1qqqq and T2qq SMS topologies, showing the
improvement between the 7 TeV analysis and the 8 TeV analysis ............................................. 121

A.1 Comparison of some leading-order Feynman diagrams calculated with MAD-
Graph resulting in Z+1 jet and γ+1 jet (i.e., one additional parton) ............................................. 141
A.2 Comparison of some leading-order Feynman diagrams calculated with MAD-
Graph resulting in Z+2 jets and γ+2 jets (i.e., two additional partons) ............................................. 142
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.3</td>
<td>Comparison of some leading-order Feynman diagrams calculated with MadGraph</td>
<td>143</td>
</tr>
<tr>
<td>A.4</td>
<td>$H_T$ distribution for the 7 TeV analysis</td>
<td>144</td>
</tr>
<tr>
<td>A.5</td>
<td>$\slash H_T$ distribution for the 7 TeV analysis</td>
<td>145</td>
</tr>
<tr>
<td>A.6</td>
<td>$\gamma p_T$ distribution for the 7 TeV analysis</td>
<td>145</td>
</tr>
<tr>
<td>A.7</td>
<td>$\gamma \eta$ distribution for the 7 TeV analysis</td>
<td>146</td>
</tr>
<tr>
<td>A.8</td>
<td>$\gamma$ ISO$^{\text{Trk}}_{\Delta R=0.4}$ distribution for the 7 TeV analysis</td>
<td>146</td>
</tr>
<tr>
<td>A.9</td>
<td>$\gamma$ ISO$^{\text{Ecal}}_{\Delta R=0.4}$ distribution for the 7 TeV analysis</td>
<td>147</td>
</tr>
<tr>
<td>A.10</td>
<td>$\gamma$ ISO$^H_{\Delta R=0.4}$ distribution for the 7 TeV analysis</td>
<td>147</td>
</tr>
<tr>
<td>A.11</td>
<td>Legend for simulation samples</td>
<td>148</td>
</tr>
<tr>
<td>A.12</td>
<td>$N_{\text{vertices}}$</td>
<td>149</td>
</tr>
<tr>
<td>A.13</td>
<td>$N_{\text{jets}}(p_T &gt; 50 \text{ GeV, }</td>
<td>\eta</td>
</tr>
<tr>
<td>A.14</td>
<td>$N_{\text{jets}}(p_T &gt; 30 \text{ GeV, }</td>
<td>\eta</td>
</tr>
<tr>
<td>A.15</td>
<td>$H_T$</td>
<td>150</td>
</tr>
<tr>
<td>A.16</td>
<td>$\slash H_T$</td>
<td>151</td>
</tr>
<tr>
<td>A.17</td>
<td>$p_T^\gamma$</td>
<td>151</td>
</tr>
<tr>
<td>A.18</td>
<td>$p_T^\gamma$</td>
<td>152</td>
</tr>
<tr>
<td>A.19</td>
<td>$H_T$</td>
<td>152</td>
</tr>
<tr>
<td>A.20</td>
<td>$H_T$</td>
<td>153</td>
</tr>
<tr>
<td>A.21</td>
<td>$\eta^\gamma$</td>
<td>153</td>
</tr>
<tr>
<td>A.22</td>
<td>$\sigma_{\gamma\gamma}$</td>
<td>154</td>
</tr>
<tr>
<td>A.23</td>
<td>ISO$^{\text{HAD}^\pm}_{\Delta R=0.3}$</td>
<td>154</td>
</tr>
<tr>
<td>A.24</td>
<td>ISO$^{\text{HAD}^\gamma}_{\Delta R=0.3}$</td>
<td>155</td>
</tr>
<tr>
<td>A.25</td>
<td>ISO$^\gamma_{\Delta R=0.3}$</td>
<td>155</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Listing of all the known fundamental particles in the standard model. 3

2.1 Particle content in a minimal Susy model. 14

5.1 Primary datasets and triggers used in the 7 TeV analysis. 56

5.2 Primary datasets and triggers used in the 8 TeV analysis. 57

5.1 2011 simple cut-based γ ID and isolation requirements 73

5.2 2012 simple cut-based γ ID and isolation requirements 73

6.1 γ effective areas for PU subtraction on different isolation sums. 74

6.4 Datasets used in the 7 TeV analysis. All data were reconstructed using CMSSW_4.2.X version of the CMS reconstruction software. 74

6.5 Datasets used in the 8 TeV analysis. All data were reconstructed using CMSSW_5.3.5. 74

6.6 List of MC simulation samples used in the 2011 γ + jets analysis. 74

6.7 List of MC simulation samples used in the 2012 γ + jets analysis, together with cross section and total number of generated events. 75

6.8 Photon purity in bins of $p_T^\gamma$ and $|\eta|$ determined using the combined isolation template method with statistical and systematic uncertainties. 75

6.9 Sideband selection requirements for 8 TeV photon purity background template selection. 90

6.10 Correction factor and individual components, along with systematic errors, for the γ + jets method (cf. Table A.1). 98

6.12 Z→νν+jets uncertainties from γ + jets data for 3 ≤ $N_{jets}$ ≤ 5 for the 8 TeV analysis (cf. Table A.2). 101

6.12 (Continued) Z→νν+jets uncertainties from γ + jets data for 6 ≤ $N_{jets}$ ≤ 7 and $N_{jets}$ ≥ 8 for the 8 TeV analysis (cf. Table A.2). 102

6.13 Z→νν+jets prediction for 4.98 pb$^{-1}$ of integrated luminosity predicted with the γ + jets method in the 7 TeV exclusive search regions. Columns 3 and 4 show the expectation coming from the simulation, while the rightmost columns show the results of the data-driven prediction. 103
6.14 $Z \rightarrow \nu \bar{\nu} + \text{jets}$ predictions from $\gamma + \text{jets}$ simulation compared to the prediction coming from data ($19.4 \text{ pb}^{-1}$) for $3 \leq N_{\text{jets}} \leq 5$ for the 8 TeV analysis (cf. Table A.3). ................................................................. 105

6.14 (Continued) $Z \rightarrow \nu \bar{\nu} + \text{jets}$ predictions from $\gamma + \text{jets}$ simulation compared to the prediction coming from data ($19.4 \text{ pb}^{-1}$) for $6 \leq N_{\text{jets}} \leq 7$ and $N_{\text{jets}} \geq 8$ for the 8 TeV analysis (cf. Table A.3). ................................................................. 106

6.15 $Z \rightarrow \nu \bar{\nu} + \text{jets}$ predictions from $\gamma + \text{jets}$ data ($19.4 \text{ pb}^{-1}$) for $3 \leq N_{\text{jets}} \leq 5$ for the 8 TeV analysis (cf. Table A.4). ................................................................. 107

6.15 (Continued) $Z \rightarrow \nu \bar{\nu} + \text{jets}$ predictions from $\gamma + \text{jets}$ data ($19.4 \text{ pb}^{-1}$) for $6 \leq N_{\text{jets}} \leq 7$ and $N_{\text{jets}} \geq 8$ for the 8 TeV analysis (cf. Table A.4). ................................................................. 108

7.1 Predicted backgrounds from the 7 TeV dataset ......................................................... 111

7.2 Predicted backgrounds from the 8 TeV dataset $3 \leq N_{\text{jets}} \leq 5$ .............................................. 113

7.2 (Continued) Predicted backgrounds from the 8 TeV dataset $6 \leq N_{\text{jets}} \leq 7$ and $N_{\text{jets}} \geq 8$ ................................................................................. 114

A.1 Correction factor and individual components, along with systematic uncertainties expressed as a percentage of the correction, for the $\gamma + \text{jets}$ method (cf. Table 6.11). ................................................................................. 157

A.2 $Z \rightarrow \nu \bar{\nu} + \text{jets}$ uncertainties as a percentage of the prediction from $\gamma + \text{jets}$ data for $3 \leq N_{\text{jets}} \leq 5$ (cf. Table 6.12). (cf. Table A.2) ................................................................................. 158

A.2 (Continued) $Z \rightarrow \nu \bar{\nu} + \text{jets}$ uncertainties from $\gamma + \text{jets}$ data for $6 \leq N_{\text{jets}} \leq 7$ and $N_{\text{jets}} \geq 8$ (cf. Table 6.12). ................................................................................. 159

A.3 $Z \rightarrow \nu \bar{\nu} + \text{jets}$ predictions from $\gamma + \text{jets}$ simulation compared to the prediction coming from data ($19.4 \text{ pb}^{-1}$) for $3 \leq N_{\text{jets}} \leq 5$ with uncertainties expressed as a percentage of the prediction (cf. Table 6.14). ................................................................................. 160

A.3 (Continued) $Z \rightarrow \nu \bar{\nu} + \text{jets}$ predictions from $\gamma + \text{jets}$ simulation compared to the prediction coming from data ($19.4 \text{ pb}^{-1}$) for $3 \leq N_{\text{jets}} \leq 5$ with uncertainties expressed as a percentage of the prediction (cf. Table 6.14). ................................................................................. 161

A.4 $Z \rightarrow \nu \bar{\nu} + \text{jets}$ predictions from $\gamma + \text{jets}$ data ($19.4 \text{ pb}^{-1}$) for $3 \leq N_{\text{jets}} \leq 5$ with uncertainties taken as a percentage of the prediction (cf. Table 6.15). ................................................................................. 162

A.4 (Continued) $Z \rightarrow \nu \bar{\nu} + \text{jets}$ predictions from $\gamma + \text{jets}$ data ($19.4 \text{ pb}^{-1}$) for $6 \leq N_{\text{jets}} \leq 7$ and $N_{\text{jets}} \geq 8$ with uncertainties taken as a percentage of the prediction (cf. Table 6.15). ................................................................................. 163
Part I

Introduction and physics motivation
Chapter 1

Introduction

Experimental high energy particle physics has a rich and varied history. Springing from an unquenchable desire to understand nature at its most fundamental level, experimentalists have used increasingly higher energies to probe increasingly smaller distances. Scientists are continually attempting to break nature into its most indivisible units and understand how these units form the world we experience. Curiosity about what else could exist and about unknown principles drives scientists on their quest for deeper understanding.

The standard model of particle physics (SM) has proven to provide a remarkably consistent description of the physical world. Ever since it was proposed in the 1970’s [1, 2, 3, 4, 5] as a unified quantum field theory for three of the four known forces, the SM has provided the most reliable description of the physical world we inhabit. It has been tested to unprecedented precision and no experimental evidence has been seen to disagree with any of its predictions. The final piece of the model, the existence of the Higgs boson, has been confirmed within the past year [6, 7].

Part of its beauty lies in its simplicity, i.e., it attempts to describe all known matter
<table>
<thead>
<tr>
<th>particle class</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>quark ((q,q))</td>
<td>(3(n_{\text{colors}}) \times 3(n_{\text{families}}) \times 2(n_{\text{types}}) \times 2(n_{\text{particle/antiparticle}}))</td>
</tr>
<tr>
<td>lepton ((l,\bar{l}))</td>
<td>(3(n_{\text{families}}) \times 2(n_{\text{types}}) \times 2(n_{\text{particle/antiparticle}}))</td>
</tr>
<tr>
<td>gluon ((g))</td>
<td>(8(n_{\text{colors}}) )</td>
</tr>
<tr>
<td>photon (\gamma)</td>
<td></td>
</tr>
<tr>
<td>vector boson (W^\pm)</td>
<td></td>
</tr>
<tr>
<td>vector boson (Z^0)</td>
<td></td>
</tr>
<tr>
<td>Higgs boson (H^0)</td>
<td></td>
</tr>
<tr>
<td><strong>Total fundamental particles</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1.1:** Listing of all the known fundamental particles in the standard model.

and interactions using only a very small set of building blocks: fermion matter building blocks (grouped into three generations), the quarks and leptons, and boson force carriers (Fig. 1.1). Fermions are particles with half-integer spin, obeying Fermi–Dirac statistics, while bosons have integer spin and obey Bose–Einstein statistics. Gluons are the carrier of the strong (color) force and interact only between quarks and other gluons. They come in eight color combinations. The photon is the force carrier of the electromagnetic interaction between charged particles. As the photon has no electric charge, it does not self-interact. The weak vector bosons, \(W^\pm\) and \(Z^0\), mediate the weak force.

Each of the three families of fermion comes in particle/antiparticle pairs. In addition, the quarks, being strongly interacting particles, have a “color charge” and thus come in three colors. In total there are 61 fundamental particles, as listed in Table 1.1.

The complete formulation of the standard model has 19 free parameters, which must be determined from experiment (the 9 fermion masses, the Higgs mass, the three coupling constants, three weak mixing angles and an associated phase describing the CP violation, the QCD vacuum angle, and the Higgs vacuum expectation value \((\text{vev})\)).
(a) The building blocks of nature and the particles responsible for the fundamental interactions

(b) The relationships between the fundamental particles and the fundamental interactions

**Figure 1.1:** The fundamental particles and interactions in the standard model of particle physics (from Ref. [72, 73]).

However, despite the accurate description the standard model provides, it is known to be incomplete; at best, an effective theory, i.e., a low energy description of an underlying fundamental theory. Notably, there is no inclusion of the fourth fundamental force, gravity. Additionally, the particles described by the standard model make up only 4.9% of the observable energy density of the universe. It provides no explanation for dark matter, which is known to exist from measured galaxy rotation speeds [8, 9], or dark energy. Dark energy and dark matter are known, from large-scale cosmic surveys [10], to compose 68.3% and 26.8% (respectively) of the energy density of the universe. Recent experimental results suggest that the neutrinos have finite mass [11, 12, 13, 14, 15] and the standard model provides no explanation for the origin of this mass. Corrections to the Higgs boson mass, the value of which is not predicted by theory, are divergent (the hierarchy problem) leading to a problem known as the “fine-tuning” problem, i.e., in order for experimental results to
match with theory, the parameters of the theory must be tuned by hand to specific values, which is a philosophical problem for scientists who would like to have a physical reason for the fine tuning. As a result of these and other reasons, theorists have been working to extend the SM in attempts to account for its shortcomings. One of the more appealing theoretical extensions to the SM is supersymmetry (Susy), which provides possible explanations for some of the missing pieces.

Particle accelerator experiments provide one of the best ways to test theories of physics beyond the SM. In fact, for the past 60 years, as particle physics experiments have increased the energy reach, their experimental confirmation of the predictions of the SM have been realized time and again, culminating in 2012 with the discovery of the Higgs boson [6, 7].

The collision environment provides a high statistics sample of controlled events to measure the properties of particles and interactions as well as to search for evidence of new particles and new interactions.

From the fundamental particles listed in Table 1.1, more complex particles are formed, e.g., the proton and neutron. Both the proton and neutron are composed of three quarks, a combination known as a “baryon.” Another combination of quarks, a “meson,” is composed of a quark antiquark pair. The proton is the stable combination of two up-type quarks and one down-type quark (compared to the neutron, which is composed of two down-type quarks and one up-type quark). When colliding protons head-on, it is the constituent quarks or gluons that actually interact. In addition to the “valence” quarks mentioned above, due to the Heisenberg uncertainty principle, the proton has “sea” or “virtual” quarks and gluons that carry a fraction of the momentum. The distribution of the
momentum in a proton can be described by several parton distribution functions (PDFs), which are measured from experiments as a function of the energy of the proton and as a function of the momentum transfer involved in the collision. For this reason, when colliding protons, it is not possible to exactly know the momentum of the interacting partons, but over the course of many millions of collisions, the behavior can be averaged according to the PDFs. In addition, the transverse momentum of the initial-state partons will be zero, and thus, a measurement of the transverse momentum of the final-state products can be made and used to deduce the presence of undetected particles.

This thesis presents a search for new physics, specifically SUSY, in an all-hadronic final state with jets and missing transverse momentum, collected using the Compact Muon Solenoid (CMS) detector. The analysis was performed as a team\(^1\) and the author made substantial contributions to the estimation of the “invisible” \(Z^0\) background. The author was the primary person responsible for the estimation using \(\gamma+\text{jets}\), which is the primary focus of this thesis. In addition, work was performed to optimize the search regions of the analysis and to study the performance of some of the triggers used to collect the data. These efforts led to a PRL [16] publication in 2012 and to several public analysis summaries [17] of the results of the search.

The author was also involved in the commissioning and operation of the hadronic calorimeter subdetector, and the development and maintenance of several centrally used software tools (as discussed in Appendices B.1, B.2, B.3, and B.4).

\(^1\)CMS RA2 SUSY Group
Chapter 2

Physics background

2.1 Standard model

The standard model of particle physics represents our current best understanding of how the universe works at the subatomic scale. It is a quantum field theory and relies upon symmetry principles. The global symmetry group of the standard model is given by Eqn. 2.1.1. Here “C” corresponds to “color” charge, e.g., QCD, “L” corresponds to left-handed weak isospin fermions, and “Y” corresponds to weak hypercharge. All the interactions in the standard model are contained in the Lagrangian, Eqn. 2.1.2, subject to the constraint that they must be invariant under the symmetries of the gauge group (Eqn. 2.1.1).

\[ SU(3)_C \times SU(2)_L \times U(1)_Y \]  
\[ \mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} \]
2.1.1 QCD interactions

Quantum chromodynamics (QCD) is the quantum field theory of the strong interaction, i.e., the theory of quarks and gluons. QCD is a gauge theory represented by the symmetry group $SU(3)_C$ in the standard model symmetry group. The QCD Lagrangian describing all interactions is shown in Eqn. 2.1.3 and some lowest order QCD interactions are shown in Fig. 2.1. Some of the basic properties of QCD include the behavior of the interaction at high energies and long distances. At high energies, the strength of the interaction between quarks and gluons is very weak, a property known as “asymptotic freedom” [18, 4, 5]. However, as the distance between two quarks increases, the strength of the coupling increases, leading to a property called “confinement.” Confinement leads directly to the process of hadronization in particle collisions, e.g., showers of hadronic particles corresponding to the underlying parton. At a certain point, the energy stored in the color string connecting the partons becomes large enough to create a particle antiparticle pair from the vacuum. This process happens many times and the created particles group together to form color-neutral quark bound states, e.g., mesons or baryons. This collection of particles is collimated in the direction of the original parton and is called a “jet.”

$$\mathcal{L}_{QCD} = -\frac{1}{4} G_{\mu\nu} G^{\mu\nu}_A + \sum_f \bar{q}_f(i\not{D} - m_f)q_f$$  (2.1.3)

2.1.2 Electroweak interactions

The theory of electroweak (EW) interactions is the unified quantum field theory of the weak and electromagnetic interactions [1, 3, 2]. It is a gauge theory represented by
Figure 2.1: Some leading-order interactions in quantum chromodynamics. Here $g$ refers to a gluon, $q$ to a quark, and $\bar{q}$ to an antiquark.

the symmetry group $SU(2)_L \times U(1)_Y$ in the standard model symmetry group (Eqn. 2.1.1). Corresponding to $SU(2)_L$ and $U(1)_Y$ are vector fields $W^i_\mu$ and $B_\mu$, which transform under the symmetry operations of the respective group. The electroweak Lagrangian, showing the various interactions in the theory, is shown in Eqn. 2.1.4 [19]. The “gauge” terms describe the interactions of the gauge fields, while the “matter” terms describe the interactions of the gauge fields with the fermions. Finally, the “Higgs” term contains all the information about electroweak symmetry breaking (EWSB) and the Higgs mechanism, including the description of the interactions of the gauge fields with the Higgs field, while the “Yukawa” term describes the origin of fermion masses via coupling to the Higgs field.

\[
L_{EW} = L_{\text{Gauge}} + L_{\text{Matter}} + L_{\text{Higgs}} + L_{\text{Yukawa}}
\]

\[
L_{\text{Gauge}} = -\frac{1}{4} W_{A\mu\nu} W^{\mu\nu}_A - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}
\]

\[
L_{\text{Matter}} = \sum_{\text{generations}} i \left[ \bar{L} \not{\partial} L + \bar{Q} \not{D} Q + \bar{u}_R \not{D} u_R + \bar{d}_R \not{D} d_R + \bar{\epsilon}_R \not{D} \epsilon_R \right]
\]

\[
L_{\text{Higgs}} = |D_\mu|^2 + \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2
\]

\[
L_{\text{Yukawa}} = \sum_{\text{generations}} i \left[ -\lambda_e \bar{L} \cdot \phi e_R - \lambda_d \bar{Q} \cdot \phi d_R - \lambda_u \epsilon^{ab} \bar{Q}_a \phi_b^\dagger u_R + \text{h.c} \right]
\]
\[ \mathcal{L}_{\text{Yukawa}} = - \sum_{\text{generations}} \frac{\lambda_{f_i} \bar{f}_i f_i H_{SM}}{\sqrt{2}} \quad (2.1.4f) \]

After EWSB, the \( W \) and \( B \) gauge fields mix to create the \( W^\pm, Z^0, \) and \( \gamma \) fields as shown in Eqn. 2.1.5. From this it can be seen that the two particles arising from the neutral \( W_3 \) and \( B \) fields are related (see Figs. 2.2d and 2.2e), a fact which is directly exploited in this thesis.

\[
A_\mu (\gamma \text{ from QED}) = \sin \theta_W W_{3\mu} + \cos \theta_W B_\mu \quad (2.1.5a)
\]

\[
Z^0_\mu = - \cos \theta_W W_{3\mu} + \sin \theta_W B_\mu \quad (2.1.5b)
\]

\[
W^{\pm}_\mu = (W_{1\mu} \mp W_{2\mu}) / \sqrt{2} \quad (2.1.5c)
\]

(a) \( W^\pm \) coupling to leptons.  
(b) \( W^\pm \) coupling to quarks.  
(e) \( W^\pm \) coupling to \( Z^0/\gamma \).

(d) \( Z^0 \) coupling to fermions.  
(e) \( \gamma \) coupling to charged fermions.  
(f) \( W^\pm \) coupling to neutral pair of bosons.

**Figure 2.2:** Interactions in the electroweak standard model.
2.1.2.1 Higgs mechanism

Because explicit mass terms arising from the unbroken electroweak Lagrangian are not gauge invariant under the symmetry transformations of $SU(2)_L \times U(1)_Y$, a mechanism is needed to provide masses to the gauge bosons, as three of the four gauge bosons have nonzero mass ($M_{W^\pm} = 80.4$ GeV, $M_Z = 91.2$ GeV [20]). The Higgs mechanism [21, 22, 23] provides masses to the standard model particles through spontaneous symmetry breaking. With the introduction of a complex field corresponding to a massive scalar boson to the Lagrangian, the field obtains a nonzero vacuum expectation value ($vev$), the electroweak symmetry is broken, and the rewritten Lagrangian contains terms that correspond to mass terms for the gauge bosons. Additional gauge invariant terms involving the scalar field can be added to generate mass terms for the fermions as well. The mass of the Higgs boson is not given by the theory and must be determined experimentally. The recently observed boson [6, 7] suggests a mass $m_H \sim 126$ GeV.

2.2 Supersymmetry

2.2.1 Motivation

Supersymmetry is a theory of new physics that posits a new symmetry between bosons and fermions as an extension to the standard model. Susy is appealing for several reasons. It allows for cancellation of quadratic divergences that appear in higher-order corrections to the calculation of the Higgs boson mass, because terms of opposite sign are contributed by the bosonic and fermionic superpartners (Eqn. 2.2.1 in Fig. 2.3).
lightest supersymmetric particle (LSP) in R-parity conserving models is usually neutral, massive, and weakly interacting, SUSY thus provides a candidate for dark matter. Finally, SUSY provides a mechanism whereby the coupling constants from the three standard model symmetry groups can be unified at a high scale, leading to the possibility of an even more fundamental theory, i.e., a “grand unified theory” (GUT).

While SUSY can be tuned to produce many consistent models, the focus of experimentalists has been on models where the energy scale for new physics is on the order of 1 TeV. In addition to being potentially accessible to experimentalists, it is at about this energy that the standard model is expected to yield to a higher scale effective theory (SUSY), due to the relatively small mass of the Higgs boson.

2.2.1.1 R-parity and a candidate for dark matter

R-parity, sometimes known as matter parity, is an additional symmetry (similar to the parity symmetry of the standard model) that determines the possible decays available in SUSY models. R-parity is defined in Eqn. 2.2.2, where $B$ is the baryon number, $L$ is the lepton number, and $s$ is the spin of the particle in question. R-parity is a multiplicative quantum number, with $R = 1$ for ordinary standard model particles and $R = -1$ for SUSY
R-parity-conserving models are preferred due to consequences of R-parity violating (RPV) models, which allow for the decay of the proton, among other currently unobserved decays. While RPV models are not excluded, current theoretical prejudice favors R-parity-conserving models.

If R-parity is conserved, this necessarily implies that a sparticle must be produced in conjunction with another sparticle, i.e., sparticles will always pair produced at collider experiments. Another consequence is that the LSP must be stable. Additionally, if the LSP is neutral, it will be noninteracting and thus a good candidate for cold dark matter. The neutral LSP is preferred in many models being explored at the LHC due to cosmological constraints on exotic matter [24].

### 2.2.2 Supersymmetric models

The theoretical underpinnings of SUSY rely upon a symmetry that relates bosons to fermions. In SUSY, the fundamental particle content of the universe will at least double, as all standard model particles will have a SUSY partner. For a minimal extension of the standard model, the particle content, along with the relationship to standard matter, is shown in Table 2.1. The scalar superpartners of the fermions are called sfermions, and the fermionic superpartners of the bosons are called bosinos. The names of the individual particles can be constructed in a similar manner.

SUSY can be realized in many ways and it is important to choose a model that is
consistent and yet feasible to study. A completely unconstrained minimal parametrization of SUSY contains over 100 free parameters. It is necessary to reduce the number of free parameters in order to effectively test the predictions of the theory, so certain assumptions are made to construct a consistent model with fewer free parameters. As superpartners have not yet been observed, their masses must be higher than the corresponding standard model particles. This necessarily means that SUSY is a broken symmetry, and the unknown mechanism providing the SUSY breaking can determine the type of physics one can expect to observe. Some commonly studied models, such as mSUGRA (minimal SUper GRAvity) and the cMSSM (constrained Minimal Supersymmetric Standard Model), posit that gravity is responsible for providing the SUSY breaking.

### 2.2.3 Search strategies

There are many different signatures in supersymmetric models. This is important in order to confirm the discovery of SUSY, as evidence seen in different channels will strengthen the interpretation. Due to the production mechanisms available at the LHC it is sensible to choose an all hadronic search channel, as strong production modes, if kin-
matically accessible, will tend to dominate. However, due to the complicated nature of measuring hadronic activity, leptonic and photonic search modes can be simpler to analyze.

One of the common ingredients in searches focusing on R-parity conserving SUSY models is missing transverse momentum. As discussed in Section 2.2.1.1, the LSP will be stable, and in many models, it will be noninteracting and somewhat massive. The detector signature of this type of particle will be an energy imbalance, i.e., missing transverse momentum.

Due also to the expected high masses of the superpartners, the expected decay chains can involve long cascades of particles, often resulting in a large number of final-state particles. Many of these particles will be produced in the central portion of the detector. Thus searching for significant energetic activity in the central part of the detector is another key component of many SUSY search strategies.
Part II

Experimental setup and event reconstruction
Chapter 3

Experimental setup

3.1 The LHC accelerator

The CERN Large Hadron Collider (LHC) is a proton-proton (pp) accelerator and storage ring constructed in the tunnel that housed the Large Electron Positron (LEP) accelerator. The LHC delivers two counter-rotating proton beams that collide in bunches at four interaction points (IPs) (see Fig. 3.1) with a center of mass energy of 7-8 TeV. Each beam contributes 3.5-4 TeV, which is the equivalent of 180-206 MJ.

The design parameters of the LHC provide proton beams of 7 TeV each, colliding at a rate of 40 MHz, with bunches of protons colliding every 25 ns. However, at the present, due to a design flaw catastrophically discovered on September 19, 2008 [25], the operation has been limited to 4 TeV/beam with bunches spaced 50 ns apart. During this incident, an electrical connector between two magnets became nonsuperconducting, resulting in an electrical arc that explosively released helium and sent particulate matter into the beam.
Figure 3.1: The LHC accelerator system, showing the four experiments (clockwise from top: CMS, LHCb, ATLAS, ALICE) and the different accelerators used to get protons into the LHC ring (LINAC2 → Booster → PS → SPS → LHC) (from Ref. [74]).

pipe. As a result of this incident, operation of the LHC was delayed by more than a year and the maximum operating current was decreased by a factor of two.

Following the repair period, on November 30, 2009, the LHC became the highest energy accelerator in the world, circulating two beams with energies of 1.18 TeV [26] and shortly thereafter, providing collisions at a center of mass energy of 2.38 TeV. Following the 2009-2010 winter shutdown, on March 30, 2010, the LHC finally began full operation, with each beam accelerated to 3.5 TeV, providing collisions at $\sqrt{s} = 7$ TeV [27]. For operations in
2010-2011, the colliding beams remained 3.5 TeV each and the bunch spacing was decreased over the course of the run to 75 ns. In 2012, the beam energy was increased to 4 TeV and the bunch spacing decreased to 50 ns. Several special runs were taken to test the performance of the experiments at 25 ns, but the majority of data for analysis was taken at 50 ns.

The number of collision events delivered by the LHC to the various experiments is proportional to the instantaneous luminosity, \( L \). For a given physics process (with cross section \( \sigma_{\text{process}} \)), the number of expected events is expressed by Eqn. 3.1.1a. The luminosity depends on a number of parameters, as expressed in Eqn. 3.1.1b. Here \( N_b \) refers to the number of particles in each bunch while \( n_b \) is the number of bunches in each beam. Also important are several parameters related to the beam size, e.g., the emittance \( (\varepsilon_n) \) and the “beta-function” \( (\beta^*) \). These parameters describe the size, density, and uniformity of the beams at the IP. For a given number of protons, a smaller, i.e., more compact, bunch will result in more collisions, and thus in a higher luminosity.

\[
N_{\text{events/process}} = L \cdot \sigma_{\text{process}} \quad (3.1.1a)
\]

\[
L = \frac{N_b^2 \cdot n_b \cdot f_{\text{rev}} \cdot \gamma_r}{4 \cdot \pi \cdot \varepsilon_n \cdot \beta^* \cdot F} \quad (3.1.1b)
\]

Under ideal operating conditions, the LHC was intended to deliver \( L = 10^{34} \text{cm}^{-2} \text{s}^{-1} \) with 2808 bunches to the two main experiments, CMS and ATLAS. However, up to the end of the 2012 pp run, the maximum delivered luminosity has been \( L = 7 \cdot 10^{33} \text{cm}^{-2} \text{s}^{-1} \) with 1380 bunches.

### 3.1.1 Particle acceleration

The protons that will eventually collide in the various detectors begin their journey to nearly the speed of light as hydrogen atoms. In the presence of an electric field, the atoms
are stripped of the electrons and passed into the LINAC2. Here they are accelerated to 50 MeV and injected into the Proton Synchrotron (PS) Booster where they are accelerated to 1.4 GeV. From the PS the protons move to the Super Proton Synchrotron (SPS) where they are further accelerated to 450 GeV and finally injected into the LHC (Fig. 3.1).

The LHC uses a combination of RF cavities and magnets to accelerate and direct the proton beams. The purpose of the RF cavities is to add energy to the protons as they circulate in the accelerator. Additionally, the RF cavities ensure that the proton bunches maintain the proper structure, i.e., particles that are in front of or behind the bunch are “encouraged” to rejoin the bunch by either adding or removing energy.

The magnets bend the beams of particles to keep the bunches circulating and also “squeeze” the beam prior to collision at the various IPs. Various types of magnets are utilized to accomplish the different tasks. Dipoles (see Fig. 3.2) are primarily used to bend the beams, and quadrupole, sextupole, and octopoles to shape and focus the beams. Due to the high magnetic field required, the magnets are superconducting and must be held at a very low (∼ 2 K) temperature. This is accomplished with the help of the cryogenic system, which utilizes liquid helium to regulate the temperature.

3.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is one of the two general purpose experiments operating at CERN utilizing the LHC accelerator (the other being A Toroidal LHC ApparatuS (ATLAS) [28]). The experimental facility is located underground at interaction
Figure 3.2: An LHC dipole magnet cross section view, showing the two separate beam pipes as well as the cryogenic system (from Ref. [75]).

point 5 (IP5), near Cessy, France. A concise description\footnote{A more detailed description can be found in Ref. [29]} of the layout and operation of CMS is provided in this section.

CMS includes of an inner tracking system to provide measurements of charged particle position and momentum, a calorimeter to measure the energy of most particles, and a muon tracking system to provide extra measurements of the momenta and position of muons. In addition to the detectors, there is a 3.8 T superconducting solenoid magnet with a 2.7 T return field, which causes the trajectories of charged particles to bend into
Figure 3.3: Layout and design of the CMS detector, showing the location and size of the various subdetectors, the solenoid magnet, and the beam pipe (from Ref. [31]).
circle-like orbits, aiding in the momentum reconstruction and electric-charge identification.

CMS uses a right-handed coordinate system with the origin centered at the nominal interaction point. The $y$-axis points vertically upward and the $x$-axis points toward the center of the LHC ring. The radial distance from the origin is measured in the $x$-$y$ plane. The azimuthal angle, $\phi$, is measured with respect to the $x$-axis in the $x$-$y$ plane, while the polar angle, $\theta$, is measured with respect to the $z$-axis in the $y$-$z$ plane. Rather than using the polar angle, it is common to express this direction in terms of “pseudorapidity,” Eqn. 3.2.1, which has a value of zero along the $y$-axis and increases to infinity along the $z$-axis. Pseudorapidity is preferred due to the fact that particle production in hadron colliders is roughly flat as a function of pseudorapidity [30].

$$\eta = - \ln \left[ \tan \left( \frac{\theta}{2} \right) \right]$$ (3.2.1)

The separation of reconstructed objects is measured in the $\eta$-$\phi$ plane as in Eqn. 3.2.2a, while the separation in the transverse ($x$-$y$ plane) is measured as in Eqn. 3.2.2b, resulting in values in the range $0 \leq \Delta \phi_{ij} \leq \pi$

$$\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} \quad (3.2.2a) \quad \Delta \phi_{ij} = |\phi_i - \phi_j| \quad (3.2.2b)$$

3.2.1 The tracking system

The purpose of the tracker is to provide precise measurements of the trajectory of all charged particles exiting the IP. Knowing the trajectories of particles helps with vertex reconstruction and momentum calculation. Primary vertices, the points where pp interactions occur, and secondary vertices, the points where long-lived particles eventually...
decay, are critical in searches for new physics processes.

As a charged particle passes through matter, it can ionize the material around it. If the material is held at a voltage to create an electric field across the medium, the ionization charges will then drift and be detected at an electrode. From the drift electrons, the position of the ionizing particle is determined. If enough position measurements are made, the track of the ionizing particle can be reconstructed and its momentum determined.

The CMS tracking detector is built using two separate design layouts. Nearest the beam line, there are several layers of silicon pixel detectors. Outside this, there are layers of silicon microstrip detectors. CMS has the largest silicon tracking detector ever built, providing nearly 200 m² of silicon detector, and covering up to |η| < 2.5 in pseudorapidity.

**Figure 3.4:** Tracking detector layout in the $r$-$z$ plane for one quarter of the detector, showing the location of the pixel layers, the microstrip inner and outer barrel, inner detector and the endcap, and the coverage in $\eta$ (from Ref. [77]).
3.2.1.1 Pixel detector

The silicon pixel detector is the first part of the detector the particles will interact with. The sensors used in the pixel detector measure 100 $\mu$m $\times$ 150 $\mu$m. With a total of 48 million pixels in the barrel region and 18 million in the forward region, the pixel detector is able to provide precise measurements of particle position in the $r$-$\phi$ plane as well as in the $z$ direction.

In the barrel region, there are three layers 53 cm in length, covering radial distances between 4.4 cm to 10.2 cm from the beam line. The barrel layers are complemented by two layers, perpendicular to the beam direction, covering 6 cm to 15 cm from the beam line, at $z = \pm 34.5$ cm and $z = \pm 46.5$ cm, giving coverage in the forward detector region. The two regions allow for a minimum of three measurements of the particle’s position almost up to the complete range of $|\eta| < 2.5$.

The primary purpose of the pixel detector is to provide precise measurements of
particles near the interaction point to aid with secondary vertex reconstruction. The high resolution of the pixel detector provides very good position measurements.

3.2.1.2 Microstrip detector

Beyond the pixel detector lies the silicon microstrip tracker, which provides an additional 10 layers of detectors with ~24k sensors. The typical size of one of the sensors in this part of the detector is $6 \text{ cm}^2 \times 12 \text{ cm}^2$ or $10 \text{ cm}^2 \times 9 \text{ cm}^2$[29]. Similarly to the pixel detector, the strip tracker has barrel layers, providing radial coverage from $22.5 \text{ cm} \leq r \leq 108 \text{ cm}$. For endcap coverage, the strip tracker has layers from $z = \pm 80 \text{ cm}$ to $z = \pm 280 \text{ cm}$, providing sensors up to $|\eta| < 2.5$.

3.2.2 The electromagnetic calorimeter

The electromagnetic calorimeter (Ecal) provides precise measurements of the energy of electrons and photons. Electrons and photons interact with matter in several well understood ways, predominantly bremsstrahlung for electrons and positrons and pair production for photons. For high energy electrons and photons, secondary and tertiary electrons and photons will be produced by these processes and initiate an electromagnetic shower (Fig. 3.6). The energy loss as charged particles pass through matter is described by the Bethe-Bloch formula (Eqn. 3.2.3). The longitudinal length of the shower can be characterized by a parameter, radiation length (Eqn. 3.2.4), which determines how far an electron will travel in matter before it is left with $1/e = 0.368$ of its initial energy. This parameter is dependent on the material. The energy loss of a photon is related to $X_0$ as
Figure 3.6: Typical cascades as electrons ($e^\pm$) and photons ($\gamma$) pass through the Ecal. well as it will reduce to $1/e$ of its initial intensity after traveling a distance $x = \frac{9}{7}X_0$.

\[-\frac{dE}{dx} = K'Z^2 \frac{1 \ln \left( \frac{2m_e c^2 \beta^2 \gamma^2 T_{\text{max}}}{I^2} - \beta^2 - \frac{\delta(\beta \gamma)}{2} \right)}{\beta^2 Z A^{1/2}} \]  

(3.2.3)

\[X_0 = \frac{714.6 \text{ g cm}^{-2} A}{Z(Z+1) \ln \left( \frac{287}{\sqrt{Z}} \right)} \]  

(3.2.4)

The typical energy resolution of electromagnetic calorimeters can be parametrized as in Eqn. 3.2.5 [20]. Here the “stochastic” term (S) is mostly determined by the event-to-event fluctuations of the electromagnetic shower development and the constant term (C) is mostly due to calibration uncertainties. The final “noise” term (N) comes from the inherent noise in the electronic readout channels.

\[ \left( \frac{\sigma}{E} \right)^2 = \left( \frac{S}{\sqrt{E}} \right)^2 + \left( \frac{N}{E} \right)^2 + C^2 \]  

(3.2.5)
3.2.2.1 Design and layout

The CMS ECAL is composed of a barrel detector and endcap detector with a “preshower” silicon detector situated in front of the endcap, which aids with particle identification and improves energy resolution. The purpose of the ECAL is to measure the energy of electromagnetically interacting particles, i.e., electrons and photons, by measuring the energy released in the showers produced when such particles interact with heavy dense matter.

Lead tungstate (PbWO$_4$) crystals provide both the showering and scintillating material. The characteristics of PbWO$_4$ are well suited to the purpose of measuring the energy of electrons and photons. The material is dense (8.3 g/cm$^3$), which results in many interactions as the particle passes through the crystal. The short radiation length ($X_0 = 0.89$ cm), allows for a compact detector, as the thickness of the ECAL is only 23 cm, corresponding to $\sim 25X_0$, which should provide ample containment of electromagnetic showers. The small Molière radius (2.2 cm) allows for the showers to be narrow, resulting in a fine granular resolution of energy deposits with less leakage into neighboring crystals. Finally, the decay time for the scintillation light is fast enough to release 80% of the light before the next bunch crossing.

Because the light output is relatively low, the signal must be amplified in some way, and the chosen technologies for this task are avalanche photo diodes (APDs) in the barrel region and vacuum photo triodes (VPTs) in the endcap region. The amplified signals are digitized and sent to the front-end electronics where they are buffered and used to create trigger primitives (TPs).
The individual crystals are grouped into “super-crystals,” groups of 5 × 5 crystals arranged around the detector. In the endcap, the crystals are mounted into two half circular “D” shapes for easy assembly in the detector (Fig. 3.7). The super-crystals are arranged to collect energy deposits for triggering purposes, in an area adjacent in $\eta$ and $\phi$ to a corresponding HCal area (a calorimeter trigger tower).

**Figure 3.7:** Layout of the electromagnetic calorimeter (Ecal) showing the location of the barrel, endcap and preshower detector, as well as the organization of crystals into super modules and the “D” shape in the endcap (from Ref. [29]).
3.2.2.2 Energy resolution

The Ecal provides very good energy resolution for electrons and photons and is expressed in Eqn. 3.2.6 [29]

\[
\left( \frac{\sigma}{E} \right)^2 = \left( \frac{2.8\%}{\sqrt{E}} \right)^2 + \left( \frac{0.12}{E} \right)^2 + (30\%)^2
\] (3.2.6)

3.2.3 The Hadron calorimeter (Hcal)

The purpose of the hadronic calorimeter (Hcal) is to measure the energy deposited in hadronic showers from collision events as particles traverse the detector volume. The Hcal is a sampling detector, meaning it measures the energy of particles passing through the detector with a series of alternating absorber and active layer materials. The particles shower by means of nuclear interactions within the absorber material and the resulting particles interact with the active material to produce a signal proportional to the energy lost. For its active layer, Hcal uses a scintillator, which produces light output as electrons, excited by the shower particles, drop to lower energy levels and release energy in the form of photons. This optical signal is read out and transmitted to the off-detector electronics where it is combined with signals from nearby channels and sent to the regional calorimeter trigger (RCT). The RCT takes information from both Ecal and Hcal to make a decision on the event. If the event passes the level-1 accept (L1A), the information from the Hcal trigger and readout cards (HTRs) is sent to the data acquisition system (DAQ) for processing and writing to disk.
3.2.3.1 Design and layout

The physical construction of the HCal is sandwiched brass/scintillator. There are 17 scintillator layers, the first of which is directly after the ECal to provide a measurement of hadronic showers initiated within the ECal. The active layers are constructed as scintillator tiles surrounded by wavelength shifting fibers (WLSs) to take the light to the read out box (RBX).

The barrel hadronic calorimeter (HB) is physically situated outside the electromagnetic barrel calorimeter (EB) and inside the magnet coil. HB provides material between distances of 1.77 m to 2.95 m from the beam line and coverage in pseudorapidity from $0 < |\eta| < 1.392$. The 16 towers ($\eta$-$\phi$ segmented regions used for triggering) allow for 5.4 ($\eta = 0$) to 10.3 ($|\eta| = 1.3$) nuclear interaction lengths ($\lambda_I$, Eqn. 3.2.7) combined with the 1.1$\lambda_I$ provided by the ECal crystals.

$$\lambda_I \approx 35 \text{ g cm}^{-2} A^{1/3} \quad (3.2.7)$$

$\lambda_I$ is a measure of the distance in a material a particle will travel before reducing its energy by a factor of $1/e$.

As the HCal barrel is constrained by the ECal and magnet, the outer hadronic calorimeter (HO) was developed to be a “tail catcher” for hadronic showers that are not 100% contained by the ECal/HCal barrel. In the pseudorapidity region forward of the barrel ($1.3 < |\eta| < 3.0$), the endcap hadronic calorimeter (HE) provides coverage. At low pseudorapidity ($|\eta| < 1.6$), the $\phi$ segmentation is 5° and at higher values, the segmentation is doubled to 10°.
The pseudorapidity region forward of the endcap (up to $|\eta| < 5.0$) is covered by the forward hadronic calorimeter (HF). Due to the very high expected radiation in this region, the design of HF is different than that of HB and HE. In the case of HF, the active material is quartz and the absorber is steel. The Čerenkov light produced in the quartz fibers is read out using photomultiplier tubes (PMTs). The layout of the quartz fibers alternates, with one long (L) and the next short (S). The two fibers together allow for electromagnetic and hadronic showers to be distinguished: as the short fiber has more absorber material in front of it, the electromagnetic fraction of the shower can be determined by taking the difference between the output of the long fiber and the short fiber. The layout can be thought of as rings at constant pseudorapidity. The $\phi$ segmentation is $10^\circ$, except for the two rings nearest the beam line, where the segmentation is $20^\circ$.

**Figure 3.8:** Layout of the hadron calorimeter (HCal), showing the barrel and endcap tower layout (from Ref. [31]).
3.2.3.2 Readout

For HB and HE, the scintillation light is transmitted via WLS to optical decoder units (ODUs) where the signals of several WLSs are combined and directed onto hybrid photo diodes (HPDs). The HPD converts the optical signal to an electrical signal and sends the signal on to the counting house computers for processing. For HF the process is similar, with the PMTs taking the place of the HPDs. In the counting house computers the HTRs take all the signals from Hcal and combine them into an “event”. This event is then put into a buffer while a subset of the information is sent to the trigger system for a decision. If a L1A is received, the event is pushed out of the buffer to the DAQ. If no L1A is received, the event is dropped from the buffer.

3.2.3.3 Energy resolution

The energy resolution of the Hcal is generally worse than the Ecal due to the wildly varying nature of the hadronic shower and the inherent response fluctuations. This is due to the fact that hadronization can produce many different types of particles, i.e., \( \pi^\pm \) vs. \( \pi^0 \), each of which will have a different response in the detector. The resolution of the detector is largely characterized by the nonunity of the fraction \( h/e \), the ratio of the response of the detector to hadrons vs. electrons. Eqn. 3.2.8 [29] can be compared to Eqn. 3.2.6.

\[
\left( \frac{\sigma}{E} \right)^2 = \left( \frac{120\%}{\sqrt{E}} \right)^2 + (6.8\%)^2 \quad (3.2.8)
\]
Figure 3.9: Readout electronics system of the Hcal subdetector showing the data path from detector to readout electronics to counting house electronics to DAQ system (from Ref. [31]).

3.2.4 The muon system

CMS uses three different technologies for muon detection: drift tubes (DTs), resistive plate chambers (RPCs), and cathode strip chambers (CSCs) (Fig. 3.10). The DTs cover the barrel region of the detector while the CSCs are placed in the endcap region. RPCs are used in both regions to provide fast signal response.

The purpose of the muon systems is to provide measurements of the position and momentum of charged particles as they pass through the detector.
The operating principle of the muon systems is very similar to that of the silicon tracking system. As a charged particle passes through the muon system, it ionizes the active medium (in this case a gas) and the ions drift to electrodes where they are collected and read out to provide a position measurement.

![Figure 3.10: Muon detectors, showing the locations of the RPCs, CSCs and DTs as well as the coverage in η (from Ref. [31]).](image)

### 3.2.5 The magnet system

CMS utilizes a 3.8 Tesla superconducting solenoidal magnet to cause charged particles to bend. This aids in both momentum measurement as well as particle identification. The magnet is situated outside of the calorimeters in front of the muon chambers. The iron return yoke runs through the muon chambers serving as the medium through which the return flux of the solenoid is directed. The return field causes charged particles to bend.
with the opposite sign of curvature relative to their curvature inside the magnet bore. The high strength of the magnetic field allows for very good momentum resolution and is a key component in the excellent tracking resolution of the CMS detector. A map of the CMS magnetic field is shown in Fig. 3.11a, where the difference between the barrel field strength and the return field strength is clearly seen. Fig. 3.11b shows the installation of the magnet inside the iron return yoke and barrel muon system, with the Hcal barrel ready to be installed. To achieve the high field, the solenoid is superconducting and as such, is operated at a very low temperature of 4.5 K.

![Magnetic field map of the CMS solenoid, red shows the higher field strength of 3.8 T in the barrel, and 2 T in the return yoke](image.png)

**Figure 3.11:** The CMS solenoid magnet.
3.2.6 The trigger system

As the LHC provides bunches of $10^{11}$ protons every 50 ns, during every bunch crossing, it is very likely that something will happen. However, not everything that happens is “interesting” to particle physicists. In fact, as physicists are searching for new signals and rare phenomena, the cross sections of which are very small compared to “ordinary” standard model events, most of the collisions provide rather uninteresting events. Thus, it is the purpose of the trigger system to sift through the abundance of collision events and select only those that have potential interest for further study. This is an incredibly important task as of the 1 billion events that occur every second, the final trigger decision must select only the $\sim 100$ interesting events to keep for further study. The trigger system must have triggers
designed to capture all signals from any potentially interesting event, and the system must operate at a high efficiency. Due to the fact that collisions are happening very often, the trigger decisions must be made quickly. Thus, the first level of the trigger system utilizes dedicated logic chips that perform simple computations, e.g., summing energies and doing basic clustering of cells. After this level, more complicated trigger decisions are performed, seeded by the trigger results in the previous step.

### 3.2.6.1 The level-1 trigger (L1)

The CMS trigger system is composed of several parts. Most of the subdetectors combine the information sent from the detector to the front-end electronics via processing chips that create trigger primitives (TPs). The subdetector front-ends use specially designed electronic chips to perform the trigger calculations (FPGAs and ASICs). The processing at this stage is very granular and designed to be done quickly. These TPs are then combined and sent to their respective trigger processing units. The calorimeter triggers, i.e., the regional calorimeter trigger (RCT) and the global calorimeter trigger (GCT), process information from the Ecal and Hcal to produce trigger objects corresponding to electrons, photons, jets, and missing energy. The energies used to create these sums come from the calorimeter towers described earlier. The global muon trigger (GMT) processes TPs from the muon subsystems (i.e., DTs, RPCs, and CSCs) and is responsible for producing muon trigger objects.

At a global level, the L1 takes the information from the GCT and GMT and sends a decision to each subdetector (L1A), telling the front-end units whether to send an event to the DAQ for further processing or to drop it from the read-out buffer. This trigger is
necessary to reduce the rate of events processed to a manageable rate for the DAQ system. The LHC can provide collisions at a rate of 40 MHz (25 ns bunch spacing), however, in the first run, this has been limited to a maximum of 20 MHz (50 ns bunch spacing), which is far too high for the computers to process. The purpose of the global trigger (GT) is to reduce this rate to a maximum of 100 kHz, and more realistically 40 kHz, and it must do so with low latency.

3.2.6.2 The high-level trigger (HLT)

Within the DAQ the HLT performs additional selections to further reduce the rate for writing events to disk. The purpose of the HLT is to perform more complicated analysis that will select classes of events for specific analysis. The HLT is composed of complex algorithms, designed to collect all the events an analysis would like to analyze and save the events to disk. Each analysis would like to select 100% of the events in their signal region and not miss any events. In practice this is difficult to do, as there may be some inefficiency due to differences in the objects that form the trigger decisions from the corresponding objects used for offline analysis. Thus, the HLT attempts to create trigger objects as similar to the offline quantities as possible.

Due to the high rates of individual signatures in the final states of many searches, certain triggers must be prescaled, i.e., not every event that fires the trigger will be saved to disk. As such, complicated “cross triggers” have been designed to select events passing combinations of final-state signatures that individually would have rates that were too high to select with 100% efficiency. In this way, interesting events that share common signatures with uninteresting events can be selected while preserving disk space.
3.2.7 Data-taking performance

Overall, for the first run of the LHC (2010–2012), CMS was able to achieve \( \sim 92\% \) efficiency in recording the data delivered by the LHC. The recorded vs. delivered data sample was \( 40.76 \text{ pb}^{-1}/44.22 \text{ pb}^{-1} \) in 2010, \( 5.55 \text{ fb}^{-1}/6.13 \text{ fb}^{-1} \) in 2011, and \( 21.79 \text{ fb}^{-1}/23.30 \text{ fb}^{-1} \) in 2012 as seen in Figs. 3.12(a), (b), and (c). Figure 3.12(d) shows the increasing size of the data sample each year as the running conditions changed.

Figure 3.12: Plots showing the luminosity performance of CMS and the LHC (from Ref. [32]).
Chapter 4

Event simulation and reconstruction

4.1 Event simulation

Simulation of particle physics events is performed using various Monte Carlo generators and detector simulation software. The simulation takes place in several steps. The first step is to calculate the amplitudes for the various physics processes under study. This step is done using one of several “matrix element generators,” such as MadGraph [33, 34]. The program calculates the amplitudes based on the Feynman diagrams for all the contributing processes (up to the desired order) and then generates the hard-scatter events. At the end of this step, an event is described by initial-state partons and those partons produced in the matrix element describing the hard-scatter part of the event.

After the matrix element generation, the events are used as inputs to a parton showering program (usually PYTHIA [35, 36]), which takes the partons and generates addi-
tional radiation according to leading-order (LO) QCD calculations. 
PYTHIA also applies a 

hadronization model to generate final-state hadrons from the remaining partons.

Following the parton showering and hadronization, the simulated final-state hadrons 
are input into a detector simulation program (e.g., GEANT [37]), which models the behavior 
of the various final-state particles as they interact with the material of the detector. After 
this step, the output corresponds to the same signals as would be seen for data events in 
the detector.

4.2 Global event reconstruction (particle flow)

The current analysis utilizes several reconstructed detector observables that corre-

spond to different physics objects. Understanding how different particles appear as signals 
in the detector is important, as the signal produced in the detector is not identical for all 
particles (see Fig. 4.1). The different signals produced as different particles traverse the 
detector are combined using various algorithms designed to “reconstruct” the true object.

In CMS, the reconstruction and performance of different classes of physics observ-
ables (e.g., electromagnetic objects, muons, jets, and $\vec{E}_T$) is maintained and standardized by 
several “physics object groups,” or POGs. Each POG is responsible for ensuring that the 
reconstructed objects perform well for all analyses and that recommendations for standard 
object selections are provided.

This analysis makes use of the “global event reconstruction,” or particle-flow (PF) 
reconstruction [38, 39]. The PF reconstruction aims to provide a complete picture of 
the stable particles in the event using the full detector information. This is accomplished
Figure 4.1: A transverse slice of the CMS detector, showing the various subdetectors, the orientation and strength of the magnetic field, and the trajectories for various particle types (from Ref. [79]).

by matching energy deposits between subdetectors and attempting to classify the energy deposits based on the type of particle that left the deposits. The various types of particles the PF reconstruction attempts to identify are: muons, electrons, photons, charged hadrons, and neutral hadrons. Once the particles have been identified, composite objects can be reconstructed, e.g., jets, tau leptons, and missing energy. For instance, if a deposit is seen in a localized cluster of the ECAL but nowhere else, this is almost definitely a photon. If, in addition to the ECAL energy deposits, there are hits in the tracking system, then the particle was likely an electron. Other classes of particles reconstructed using the PF reconstruction are muons (hits in the tracker and muon system with minimal deposits in the HCal) and both charged and neutral hadrons (deposits in the HCal matched (or not) to tracks coming from the tracking detectors). After the particles are identified, they are grouped to form more complicated objects, such as jets and missing transverse energy.
4.2.1 Muons ($\mu^\pm$)

A muon, being a charged particle, will leave a track in the pixel and/or strip tracker. However, due to the fact that a muon is a minimum ionizing particle (MIP), it will not deposit much energy in the calorimeters, retaining most of its energy as it enters the muon systems. Here it will again leave hits as it passes through the system and usually escape the detector.

Knowing this, a muon can be reconstructed by looking for hits in the tracker that are matched to hits in the muon systems but have little or no corresponding energy deposits in the Ecal or Hcal. The energy and momentum of the muon can be computed with good precision using the extrapolation of the muon track through the system in the presence of a magnetic field. The curvature is related to the momentum of a particle passing through a magnetic field by Eqn. 4.2.1. Due to the direction of the magnetic field in CMS, once the muon enters the magnet return yoke, the curvature will change, as the magnetic field decreases and switches direction.

$$r = \frac{p_T}{0.3 \cdot B} \quad (4.2.1)$$

4.2.2 Electrons ($e^\pm$)

Electrons are reconstructed using the Ecal and tracker information, matching hits in the tracker to energy deposits in the Ecal. Due to the amount of material in the tracker, electrons lose a significant amount of energy before they encounter the Ecal. The
energy loss is due to bremsstrahlung, which is a non-Gaussian process. As such, CMS uses a Gaussian-sum filter (gsf) fit [40, 41, 42] to improve the track reconstruction of electrons.\textsuperscript{1}

4.2.3 Charged and neutral hadrons (\(\pi^\pm\), \(\pi^0\), \(K^\pm\), \(K^0\), \(\eta\), etc.)

The remaining particles in the event fall into one of three categories, charged hadrons, neutral hadrons, or photons. The first two are described in this section. Hadrons are the remnants of the hadronization of a parton as it moves away from the IP. These particles are what jets are composed of and identifying them separately will aid with jet reconstruction as well as \(\tau\) lepton reconstruction. Hadrons are expected to deposit most, if not all, of their energy in the hadron calorimeter.

Charged hadrons will leave hits in the tracking system as well as deposit energy in the calorimeters. Neutral hadrons will not have any measured track in the tracking detector and all their measured energy will be deposited in the HCal and ECal. To disentangle these two types of particle, a comparison is made between the measured calorimeter energy and the momentum of the reconstructed track. If the calorimeter-measured energy is significantly larger than the momentum of the reconstructed track, the deposit is considered to be a neutral hadron, otherwise, the track is considered to be a charged hadron.

4.2.4 Photons (\(\gamma\))

Photons (e.g., coming from \(\pi^0\) decays or from electron bremsstrahlung) are electromagnetically neutral objects, which mediate the electromagnetic force. Obviously they interact via the electromagnetic force, and due to their neutral charge they will not bend

\textsuperscript{1}A more complete description of the CMS electron reconstruction can be found in Ref. [43]
in the presence of a magnetic field. Thus, as a photon passes through the detector, it will deposit most of its energy in the ECal in a single, narrow shower. Additionally, photons will not leave hits in the tracking detector, so the ECal deposits can be checked to ensure that they do not align with the extrapolation of any hits in the tracker.

For the photons in this analysis, the reconstructed energy and momentum come directly from the energy deposited in the ECal. A further identification requirement is that there be no hits in the pixel tracker, i.e., the “pixel veto.”

4.2.4.1 Isolation

Prompt photons are those photons coming from the hard scattering process or final-state radiation rather than from $\pi^0$ decays or from other decays of particles arising during the hadronization of jets. In order to separate prompt photons from these common background photons, it is common to make use of “isolation” variables, as the background photons are usually accompanied by other nearby activity. Several isolation algorithms are utilized in this analysis and the construction of the isolation sums is described here.

For the analysis of 2011 7 TeV data, the isolation variables made use of the detector deposits in the form of simple ECal, HCal, and tracker energy sums within a hollow cone (i.e., excluding the energy of the photon candidate itself) around the reconstructed photon candidate of $\Delta R = 0.4$. The three isolation variables are computed separately from the deposits in the corresponding subdetector. The isolation sums are combined and the total has average pileup (see Section 4.2.7) energy removed.

For the analysis of 2012 8 TeV data, the isolation sums are computed from PF candidates in three categories: charged hadrons ($\text{Iso}_{\Delta R=0.3}^\text{HAD}^\pm$), neutral hadrons ($\text{Iso}_{\Delta R=0.3}^\text{HAD}^0$),
and photons ($\text{Iso}^{\gamma}_{\Delta R=0.3}$). The three isolation variables are computed by adding the $p_T$ of the candidates in each category within a cone of $\Delta R = 0.3$ around the reconstructed photon candidate. Each of the sums has pileup energy removed based on an “effective area” $\rho$-subtraction.

### 4.2.5 Jets

Jets are the detector signature of the showers of strongly interacting particles, i.e., quarks and gluons. As quarks and gluons are not color-singlet states, QCD stipulates that they may not exist in isolation in nature. As such, when they are produced in collisions, they leave the IP and cause other particles to be produced from the “sea” of virtual particles (quarks and gluons) through the process of hadronization (see Section 2.1.1). The resulting bound states typically then decay to pions and kaons. These particles all mostly interact in the HCAL and ECAL and the energy deposits in each are fed into a “jet-clustering” algorithm. These algorithms attempt to group the hits into collections that would be consistent with a set of collimated particles coming from a single point. This is accomplished by taking the highest energy “seed” hits and building clusters around these points within a pre-defined distance.

#### 4.2.5.1 Jet finding and clustering

In reconstructing jets from energy deposits or particle candidates, several considerations must be taken into account. The algorithm used must be both infrared- and collinear-safe, i.e., as the parton evolves, if it emits a soft gluon, or if the parton splits into two (or more) collinear subproducts, the reconstructed jet should be unchanged. Additional-
ally, when reconstructing jets from the set of PF candidates, isolated electrons, muons and photons will end up in the jet collection simply due to the fact that all particle candidates are run through the jet producer. The jet collection can then be cleaned at a later step, removing the objects one wishes not to have included in the jet collection.

The jets used in this analysis utilize the anti-$k_T$ jet clustering algorithm [44] with the FastJet [45, 46] jet finding implementation. The jet size parameter is set to $R = 0.5$. Jets are clustered by looking for high energy seeds and then clustering the PF candidates around them in a sequential recombination scheme. The algorithm makes use of two quantities computed from the PF candidates (Eqn. 4.2.2).

\[ d_{ij} = \min \{ k_{i,\perp}^{2p}, k_{j,\perp}^{2p} \} \frac{\Delta R_{ij}}{R} \]  \hspace{1cm} (4.2.2a) \hspace{1cm} \[ d_{iB} = k_{i,\perp}^{2p} \]  \hspace{1cm} (4.2.2b)

$k_{i,\perp}$ is the transverse momentum of the cluster (either calorimeter tower, particle, or cluster of particles), $\Delta R_{ij} = \sqrt{\Delta \eta_{ij}^2 + \Delta \phi_{ij}^2}$ is the separation (in azimuth and rapidity) between two clusters, and $p$ is a parameter that helps to determine the manner in which soft clusters are included.

During the clustering stage, two clusters $k_{i,\perp}$ and $k_{j,\perp}$ are compared using Eqn. 4.2.2. If $d_{ij} < d_{iB}$, the particles or clusters corresponding to $k_{i,\perp}$ and $k_{j,\perp}$ are combined and the step is repeated. In the other case, the cluster corresponding to $k_{i,\perp}$ is considered to be a jet, and the algorithm moves on to the next seed.

### 4.2.5.2 Pileup subtraction

Due to pileup (see Section 4.2.7) the reconstructed jet energy will necessarily contain energy deposited from “pileup” events. In order to be as insensitive to pileup as
possible, a method for removing this contribution from the reconstructed jet is needed. There are several methods employed in the jet CMS reconstruction, notably FastJet $\rho$-subtraction and PF “charged hadron subtraction” (PFchs) (discussed in Section 4.2.7).

FastJet $\rho$-subtraction [47, 48] calculates the event-by-event average energy density and removes the density times the area of the jet for each jet. The jet finding algorithm used for this calculation is the $k_T$ algorithm [49, 50] with size parameter $R = 0.6$.

### 4.2.5.3 Jet energy corrections

The recorded energy of the jet, using the information from the tracker and calorimeters, does not equal the “true” energy of the jet or of the parton that produced the jet. The energy resolution of the HCal is considerably larger than 100%. Making use of the PF reconstruction, and its inherent dependence on tracker measurements, greatly improves the resolution. Across a wide range of $p_T$ values and the whole detector acceptance, the jet energy resolution is found to be $\sim 15\%$ at low $p_T$ and $\sim 4\%$ for high $p_T$ jets [51]. Jet energy corrections are applied, based on the measured $p_T$ and $\eta$ of the reconstructed jet object, to account for the dependence of the response. The corrections are applied in different steps, or levels. Level 1 (L1) corrects the jet for pileup effects, while L2 and L3 correct the jet response to be flat vs. both $p_T$ and $\eta$. An additional residual correction is applied only to data, determined from data events, to account for differences in the response between the data and simulation.
4.2.6 Missing energy

In pp collisions, because protons are composite particles, their interactions are not clean and easily defined, unlike $e^+e^-$ collisions. This means that we cannot know the initial momentum of the two particles involved in the hard scattering process. However, we can use the transverse component of the total energy as we know very well that the total transverse momentum in an event will sum to zero for the initial-state partons, since the two beams are collided at zero angle. For perfectly measured events, the vectorial sum of the transverse momentum of all observed objects should be equal in magnitude and opposite in direction to the corresponding sum for all unobserved objects, where the unobserved objects are the particles that do not interact with the detector material.

Missing transverse momentum ($\vec{E}_T$) is reconstructed as the negative vectorial sum of the transverse momentum of all reconstructed physics objects in the event. Depending on the analysis performed, several types of corrections may need to be applied to the $\vec{E}_T$ in order for the $\vec{E}_T$ object used to accurately represent the genuine $\vec{E}_T$ coming from the invisible particles produced in the collision event. These corrections account for the imperfect resolution of the detector as well as the corrections applied to other physics objects in the event.

For the analysis performed in this thesis, the events considered are “all-hadronic” meaning there are no isolated leptons in the search signature. The principal observables in the event will be jets, the result of hadronic showers. In this case, a good approximation of the $\vec{E}_T$ is the quantity $\vec{H}_T$, the negative vectorial sum of all of the jets in an event, i.e., the missing transverse hadronic activity. This quantity is computed from all jets in the
event. Using the jets has the added benefit of including the jet energy corrections (JECs) automatically.

4.2.7 Pileup

In the LHC environment, each bunch crossing can produce more than one collision event. During the 2012 run, the average number of collision events was $\sim 21$ (see Fig. 4.2). Pileup is loosely defined as energy, i.e., activity, that does not originate from the initial proton-proton interaction of interest. Pileup can be due to an in-time process, such as a secondary hard scatter interaction occurring during the same bunch crossing. Additionally, due to the short spacing between colliding bunches ($50\,\text{ns}$), remnants from a previous interaction can still be interacting with the detector when the next collision occurs (out-of-time pileup, OOTPU). Both of these result in a degradation of the energy resolution of the

Figure 4.2: Pileup distribution for the 2012 run (from Ref. [32]).
detector and need to be mitigated. Pileup effects degrade the resolution of all measured objects and several methods have been developed to attempt to mitigate the adverse effects of pileup (see Section 4.2.5.2 for a discussion of the methods used to improve the jet reconstruction).

One of the solutions, unique to the particle-flow event reconstruction, is the removal of charged tracks not originating at the event primary vertex, i.e., “charged hadron subtraction” or “PFchs.” When tracks can be matched to a different primary vertex then the corresponding particle-flow candidates can be excluded from the algorithms used to create electrons, muons, photons and jets. PFchs removes, from the object reconstruction, all the charged hadrons identified in the PF reconstruction sequence originating from a primary vertex different from the leading primary vertex of the event. In this manner, energy deposits from different hard scattering processes are not included in the energy of the object.

The second method used in this analysis to remove pileup energy is “effective area ρ subtraction.” This method relies on measuring the average (per event) energy density due to pileup and subtracting the corresponding energy from the final reconstructed objects. The “effective area” is determined by the ratio of the dependence of the reconstruction of the object in question to the pileup density (ρ_{event}) and number of reconstructed primary vertices, N_{vtx}. In this way, the energy removed from each object will be different, as the reconstruction of each object depends differently on the pileup in the event. This method removes pileup contributions, as well as contributions from the “underlying event” (UE), not involved in the initial hard-scatter process. Here, the UE represents final-state decay products from the proton remnants.
Part III

Analysis of all-hadronic events in the context of a search for supersymmetry
Chapter 5

All-hadronic SUSY search

5.1 Motivation

In SUSY, if it is realized in nature, strong production channels (if kinematically accessible to experiments) are expected to dominate at the LHC and thus provide a good handle for potential discovery and setting stringent limits. For this reason, the analysis presented here has focused on signatures resulting from strong SUSY production, e.g., jets and missing transverse momentum, in order to be sensitive to a wide range of new physics models. Hadronic SUSY searches are characterized by signatures with jets, missing transverse momentum, and vetos on isolated leptons (electrons and muons).

5.2 Signal regions and event selection

This analysis requires hadronic activity by asking for at least three jets in a central production region ($|\eta| < 2.5$), where each jet is required to have $p_T > 50$ GeV. In addition to this requirement, the total hadronic scalar transverse momentum, $H_T$, is computed
from the same jets according to the formula in Eqn. 5.2.1a and is required to be greater than 500 GeV. The hadronic missing transverse momentum ($\vec{H}_T$) is computed from all jets satisfying $p_T > 30$ GeV and $|\eta| < 5.0$. In a fully hadronic event, the $\vec{H}_T$ measures the transverse momentum of the “invisible” particles, e.g., neutrinos and potential LSPs. For this search the $\vec{H}_T$ is required to be above 200 GeV.

\[
H_T = \sum_{j: p_T > 50, \atop |\eta| < 2.5} p_T, j \quad (5.2.1a) \quad \vec{H}_T = - \sum_{j: p_T > 30, \atop |\eta| < 5.0} \vec{p}_T, j \quad (5.2.1b)
\]

For the analysis presented in this thesis, the search regions were selected to be able to be sensitive to a variety of potential models, as well as to have sufficiently many events in each bin so as to be able to perform “data-driven” background estimations. In tail searches such as this, it is important to not rely too heavily on simulated events, which may not reliably reproduce the behavior in these regions.

In addition, trigger thresholds determine the allowable low thresholds for a baseline selection, where the analysis can be fully efficient for the accumulated data. For the analysis of the 2012 (8 TeV) data, the search was modified to add an additional binning in $N_{jets}$ to allow for increased sensitivity to different models of new physics.

### 5.2.1 Datasets used and trigger selections

Data collected by the CMS experiment is stored in different primary datasets. The primary datasets are defined by the triggers that fired in a given event. For this analysis the $H_T$ and $HTMHT$ datasets are used to define the signal regions. These datasets are fed by triggers based on $H_T$ or $\vec{H}_T$ or their combination, so called “cross triggers.” If an event fires one of the triggers that defines the primary dataset, this event is stored in that primary
dataset. A given event, because it can fire more than one trigger, can be stored in several primary datasets.

7 TeV analysis

The 7 TeV analysis used a baseline selection of $H_T > 500$ GeV, $MHT > 200$ GeV, and $N_{jets} \geq 3$, dictated by the loosest un-prescaled triggers that could be chosen such that they would be 100% efficient in this region. The jets were reconstructed using the particle-flow candidates and the anti-$k_T$ algorithm. The triggers used information from calorimeter jet algorithms rather than the particle-flow algorithms, which were used for the offline physics objects, but this did not affect the efficiency of the triggers in the search regions. The list of triggers and primary datasets used in the 2011 analysis is shown in Table 5.1

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Run Range</th>
<th>Trigger</th>
<th>$\int L dt$ (fb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/HT/Run2011A-May10ReReco-v1/AOD</td>
<td>160431–160578</td>
<td>HLT_HT160</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>160871–160943</td>
<td>HLT_HT240</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>160955–160956</td>
<td>HLT_HT260_MHT60</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>160957–161176</td>
<td>HLT_HT260</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>161217–163869</td>
<td>HLT_HT250_MHT60</td>
<td>-</td>
</tr>
<tr>
<td>/HT/Run2011A-PromptReco-v4/AOD</td>
<td>165088–166967</td>
<td>HLT_HT250_MHT70</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>167039–167913</td>
<td>HLT_HT250_MHT90</td>
<td>0.95</td>
</tr>
<tr>
<td>/HT/Run2011A-05Aug2011-v1/AOD</td>
<td>170249–172619</td>
<td>HLT_HT250_MHT90</td>
<td>0.39</td>
</tr>
<tr>
<td>/HT/Run2011A-PromptReco-v6/AOD</td>
<td>172620–173692</td>
<td>HLT_HT250_MHT90</td>
<td>0.71</td>
</tr>
<tr>
<td>/HT/Run2011B-PromptReco-v1/AOD</td>
<td>175832–176023</td>
<td>HLT_HT250_MHT100</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>176024–176470</td>
<td>HLT_HT300_MHT90</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>176545–178380</td>
<td>HLT_HT350_MHT90</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>178420–178708</td>
<td>HLT_HT350_MHT100</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>178712–180252</td>
<td>HLT_HT350_MHT110</td>
<td>2.71</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>4.98</td>
</tr>
</tbody>
</table>

Table 5.1: Primary datasets and triggers used in the 7 TeV analysis.

The signal region is further subdivided into 14 exclusive bins in $H_T$ and $MHT$. 

56
8 TeV analysis

For the analysis of 8 TeV data, the baseline selection again required $H_T > 500 \text{ GeV}$, $\not{E}_T > 200 \text{ GeV}$, and $N_{\text{jets}} \geq 3$. In addition, in 2012 the triggers more closely resembled the offline quantities used in the analysis as the particle-flow information was used in the triggers. The jets were reconstructed using the particle-flow candidates with the anti-$k_T$ algorithm, and utilized the PF charged hadron subtraction pileup removal method. For the second half of the 2012 run, the triggers included pileup removal information as well. The list of triggers and primary datasets used in the 2012 analysis is shown in Table 5.2.

<table>
<thead>
<tr>
<th>Primary Dataset</th>
<th>Trigger</th>
<th>$\int \mathcal{L} dt \ (\text{fb}^{-1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>/HT/Run2012A-13Jul2012-v1/AOD</td>
<td>HLT_PFHT350_PFMET100</td>
<td>0.809</td>
</tr>
<tr>
<td>/HT/Run2012A-recover-06Aug2012-v1/AOD</td>
<td>HLT_PFHT350_PFMET100</td>
<td>0.082</td>
</tr>
<tr>
<td>/HTMHT/Run2012B-13Jul2012-v1/AOD</td>
<td>HLT_PFHT350_PFMET100</td>
<td>4.04</td>
</tr>
<tr>
<td>/HTMHT/Run2012C-24Aug2012-v1/AOD</td>
<td>HLT_PFNopUHT350_PFMET100</td>
<td>0.495</td>
</tr>
<tr>
<td>/HTMHT/Run2012C-PromptReco-v2/AOD</td>
<td>HLT_PFNopUHT350_PFMET100</td>
<td>6.402</td>
</tr>
<tr>
<td>/HTMHT/Run2012D-PromptReco-v1/AOD</td>
<td>HLT_PFNopUHT350_PFMET100</td>
<td>7.274</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>19.466</td>
</tr>
</tbody>
</table>

Table 5.2: Primary datasets and triggers used in the 8 TeV analysis.

The signal region is further subdivided into 36 exclusive bins in $H_T$, $\not{E}_T$, and $N_{\text{jets}}$. The three jet multiplicities are $3 \leq N_{\text{jets}} \leq 5$, $6 \leq N_{\text{jets}} \leq 7$, and $N_{\text{jets}} \geq 8$ and allow for increased sensitivity to various models that have differing number of jets in the final state.

5.2.2 Suppression of backgrounds to the search regions

There are several standard model processes that can contribute to a final state with jets and missing transverse momentum, and additional cuts are imposed to minimize the contribution from these backgrounds.
5.2.2.1 Suppression of the QCD background

QCD events are characterized by many jets, and when the accurate reconstruction of a jet fails due to the imperfect resolution of the detector, artificial missing energy is seen, usually aligned with one of the jets. When the reconstruction fails to accurately reconstruct the true energy of a jet, the reconstructed jet will have a larger or smaller energy than the true jet while the measured direction will, for the most part, be accurately reconstructed. This leads to a momentum imbalance along the direction of one of the jets, either parallel, due to under-reconstructing the energy, or anti-parallel, due to over-reconstructing the jet energy. Thus, a cut is imposed on the absolute azimuthal angle between the $\vec{H}_T$ and each of the three leading (in $p_T$) jets (Eqn. 5.2.2). This serves to reject events, usually from QCD processes, wherein the energy of one of the jets has been mismeasured. In signal events, or other events with genuine missing energy, the LSP or neutrino will be well separated from the other partons in the decay and thus all of the jets will be well separated from the $\vec{H}_T$.

$$\Delta \phi (\text{jet}_{1,2,3}, \vec{H}_T) > 0.5, 0.5, 0.3 \quad (5.2.2)$$

5.2.2.2 Suppression of top and electroweak processes (t$\bar{t}$+jets or W$^{\pm}$+jets)

Another source of standard model events with jets and $\vec{H}_T$ is processes with the leptonic decay of a W boson ($W \rightarrow l\nu$), such as in the case of the semileptonic decay of a t$\bar{t}$ or in a W$^{\pm}$+jets event. In these cases the $\vec{H}_T$ is genuine, in that it corresponds to an invisible particle, the neutrino ($\nu$), produced in the interaction. To suppress such events, a cut is made to reject events where an isolated lepton (electron or muon) is reconstructed. In the hadronic decay channels, no isolated leptons are produced and the only leptons will
be coming from decays in jets, and these leptons should not be isolated.

5.3 Event cleaning

Pathological events due to detector misbehavior, misreconstruction of physics objects, and beam induced backgrounds, can still end up in the signal region. For this reason, several cleaning filters are applied to help reject these nonphysical events. The CMS $\vec{E}_T$ group has developed filters that characterize the various pathological event signatures and provides recommended filters for general use in all analyses. These common filters are applied to this analysis. In addition, pathological events have been identified in the course of performing the analysis and new filters have been developed and added. For a full list of the applied filters and their descriptions see Appendix A.1

5.4 Backgrounds to hadronic SUSY searches

There are several backgrounds to hadronic SUSY searches performed at LHC energies. A reducible background is one that has a signature different from the signal region, but is still able to end up in the signal region due to looser cuts, misidentification of objects, and detector effects. An irreducible background is one that is indistinguishable from the detector signature of a signal event and it is generally not possible to use any cut to reject this type of event. The dominant reducible backgrounds are $W^\pm/t\bar{t}$ +jets and QCD multi-jet events. As the focus of this thesis is the estimation of the irreducible $Z\rightarrow\nu\bar{\nu} +$ jets background, the methods used to estimate the other backgrounds will be briefly described in this section to set up the final picture when all backgrounds are combined to determine
the limit on new physics models. A more complete description of these methods can be found in Refs. [52, 53, 54, 55, 16].

5.4.1 QCD background estimation

As true QCD events have little or no “real” $\vec{E}_T$, the main source of $\vec{E}_T$ in QCD multi-jet events comes from poorly reconstructed jets. This can happen due to the inherent resolution effects in jet measurements. An additional source of $\vec{E}_T$ in QCD events is when a heavy flavor jet contains a high $p_T$ neutrino, which leads to a mismeasurement of the true jet $p_T$. The method developed to estimate this effect from the data is described below.

5.4.1.1 Rebalance and smear (R+S) method

Utilizing the fact that the “true” $\vec{E}_T$ in QCD events should be zero, a sample of multi-jet events is selected using an $H_T$ trigger. The jets in the event have their momenta kinematically rebalanced to create a seed sample to draw from later. The momenta are scaled using a maximum likelihood method on the jet resolution probability distribution, maximized based on the “true” $p_T$ of the jet, and subject to the constraint that $\vec{E}_T = 0$. After rebalancing, the event kinematics correspond roughly to those at the particle level, i.e., the hadrons produced from the partons during hadronization. The jets in these rebalanced events are then “smeared” by the jet response measured in simulated QCD events, which attempts to replicate the detector behavior on data. A method to increase the statistical power, known as a bootstrap method, is utilized, which takes a subset of events and samples from this set many times, using a randomly drawn jet response. After smearing, the event is then examined to determine whether it falls into one of the signal regions. The great
benefit of this method is that after smearing the rebalanced jets according to the measured jet response, the rebalanced events describe the kinematics of true QCD events very well and provide a reliable estimate of the QCD background.

5.4.2 $W^\pm + \text{jets}$ and $t\bar{t} + \text{jets}$ background estimation

Due to the genuine $\not{E}_T$ coming from the neutrino in leptonic decays of the $W^\pm$ boson ($W^\pm \rightarrow l^\pm \nu_l l \in \{ e, \mu, \tau \}$), processes involving $t\bar{t}$ ($q_t \rightarrow q_b + W^\pm)/W^\pm + \text{jets}$ have a chance to end up in the signal region when the lepton is not identified or reconstructed, when the lepton falls outside the detector acceptance, or when the lepton is a $\tau$ that subsequently decays hadronically. Methods have been developed to deal with each of these cases separately and are briefly described below.

5.4.2.1 Lost lepton background

In the event that a lepton (either electron or muon) is produced outside of the detector acceptance or fails either the reconstruction or isolation requirement, this event will not be rejected by the isolated lepton veto. To estimate this background, an isolated muon (electron) control sample is selected by inverting the isolated lepton veto on a muon-(electron-) triggered data sample. The method was performed in 2011 using only a muon control sample, and applying a correction to account for the different efficiency of the electron. These events are then weighted according to the acceptance, reconstruction, and isolation efficiencies (separately for electrons and muons, Eqns. 5.4.1) to estimate the probability that an electron or muon will fail to be reconstructed or isolated or be produced outside the detector acceptance.
Leptons will fall outside the detector acceptance if the $p_T$ is too low, as can happen in leptonic $\tau$ decays ($B(\tau \to l \bar{\nu}_l \nu_\tau : l \in \{e^\pm, \mu^\pm\}) = 35.24\%$) due to the extra momentum carried off by the neutrinos, or if the muon or electron traverses a part of the detector that has no active material ($|\eta| > 2.4$ for muons, and $|\eta| > 3$ or $1.4442 < |\eta| < 1.566$ for electrons). This acceptance efficiency is measured from simulated events.

The reconstruction and isolation efficiencies are measured using a tag-and-probe (T&P) technique in bins of the relevant search variables ($H_T, \not{H}_T, N_{jets}$) using $Z \to ll$ events in both data and simulation. In this way any differences between the efficiencies in data and simulation are accounted for with a systematic uncertainty. The T&P technique works by selecting a well identified lepton (tag) and a second lepton that passes a much looser selection (probe). The T&P pair are then combined into a potential $Z$ candidate according to a mass window around $M_Z = 91.2$ GeV. The probes are then categorized by whether or not they pass the tighter identification or isolation requirements. The efficiency is then obtained by taking the integral of fits performed on the dilepton invariant mass using the $Z$ lineshape. The efficiency is given by the integral performed on the distribution with a probe that passes the tighter requirement divided by the integral from all probes.

### 5.4.2.2 Hadronic $\tau$ background

The $\tau$ lepton typically decays via a one-pronged or a three-pronged mode, resulting in a low track multiplicity and a neutrino. The $\tau$ predominantly decays hadronically thus it
can appear as an extra jet in the event, and there is no possibility of rejecting this event by any selection requirement imposed to construct the signal region. In the absence of highly efficient $\tau$-lepton identification, this background is irreducible. To estimate the background coming from events of this type, a muon control sample is selected by inverting the isolated muon veto and requiring at least two jets. This control sample is then weighted according to the reconstruction and isolation efficiencies of the $\mu$, the ratio of $W^\pm$ branching fractions \( \frac{B(W^\pm \rightarrow \tau^\pm \nu_\tau)}{B(W^\pm \rightarrow \mu^\pm \nu_\mu)} \), and the branching fraction of $\tau$ to hadrons \( B(\tau \rightarrow \text{hadrons}) = 64.76\% \).

Several $p_T$ dependent $\tau$-response templates (Fig. 5.1) are taken from simulated events (both $W^\pm$+jets and $t\bar{t}$+jets) where the $\tau$ jet is matched to a generator-level tau lepton that has decayed hadronically. The response template models the visible fraction of energy in the $\tau$ decay, i.e., not coming from the neutrinos. The $p_T$ of the $\mu$ in the control sample is then scaled by the response for a $\tau$ jet randomly drawn from one the $\tau$-response templates. This event is then treated like any other event in the signal selection to arrive at an estimate of the hadronically decaying $\tau$ background.
Figure 5.1: $\tau$-response templates for $\tau$ jets normalized by the $\tau$ lepton $p_T$ in bins of the reconstructed $\tau$ jet (from Ref. [55]).
Chapter 6

Estimating the background from $Z \rightarrow \nu \bar{\nu} + \text{jets}$ events using a $\gamma + \text{jets}$ sample

6.1 Motivation

Exploiting the similarity in the kinematics between different vector boson+jets processes at high boson $p_T$, it should be possible to measure the photon $(\gamma) + \text{jets}$ event yield in data and apply a translation factor in order to obtain an estimate of the $Z \rightarrow \nu \bar{\nu} + \text{jets}$ background to an all-hadronic $\text{jets} + \slashed{E}_T$ search. After selecting $\gamma + \text{jets}$ events, the identified photon is removed and treated as an unmeasured particle. The event variables are then recomputed so that the event will mimic a $Z \rightarrow \nu \bar{\nu} + \text{jets}$ event. Comparing the Feynman diagrams for the two processes, it can be seen how this correspondence arises. For example,
in Fig. 6.1a when one ignores the neutrinos and the replacement $Z \rightarrow \gamma$ is performed, one obtains Fig. 6.1b. Similar replacements yield the correspondence in Fig. 6.2, showing how the jet (parton) content of the event does not affect the production of the $Z$ or $\gamma$. The simulated events used in this thesis were generated using the MadGraph matrix element calculator with Feynman diagrams including up to four partons. For sample diagrams of these additional processes with more computed partons, see Appendix A.2.

This method relies on the fact that at boson $p_T$ values well above the masses of the bosons, the differences in hadronic kinematics due to the mass of the $Z$ boson disappear between the two classes of events. This can be seen in Fig. 6.3a, where the hadronic portion of the event ($H_T$) has a flat ratio between the two sets of events when $H_T$ is computed from generator-level jets. Figure 6.3b shows the $H_T$ and ratio when computed using reconstructed jets that have had the jet matched to the generator-level photon removed, and the behavior is the same between the two calculations. Figure 6.4 shows the comparison between the two

Figure 6.1: Comparison of the leading-order Feynman diagram calculated with MadGraph resulting in $Z/\gamma$ production.
Figure 6.2: Comparison of some leading-order Feynman diagrams calculated with MAD-GRAPH resulting in $Z+4$ jets and $\gamma+4$ jets (i.e., four additional partons).

Figure 6.3: Comparison of the $H_T$ distributions between $Z\rightarrow\nu\bar{\nu}+\text{jets}$ events and $\gamma+\text{jets}$ events.
samples for the $p_T$ of the generator-level boson (Fig. 6.4a) and the event $H_T$ computed using generator-level jets (Fig. 6.4b). The two plots show that the $H_T$ represents the $p_T$ of the boson, Z or photon, and cutting on it will serve to select high $p_T$ Z boson events. Exploiting this fact, one can perform a data-driven, i.e., using a data control sample, prediction using a phenomenological ratio factor in this kinematic region defined by the reconstructed objects $H_T$, $H_T$, and $N_{jets}$.

First, a pure sample of isolated $\gamma + \text{jets}$ events is chosen to minimize the background from QCD (“secondary” photons, i.e., those coming from mesons in jets, e.g., $\pi^0$s) and $t\bar{t} + \text{jets}/W^{\pm} + \text{jets}$ (electroweak contamination due to misidentified electrons) events. This sample is then used to predict the $Z\rightarrow\nu\bar{\nu} + \text{jets}$ yield. The correspondence comes by removing the photon from the event to mimic the two undetected neutrinos, recomputing

![Graphs showing boson $p_T$ and $H_T$ distributions and ratios.](attachment:graphs.png)
the event variables \((H_T, H_T', \Delta \phi(H_T, \text{jet}))\), and finally correcting for experimental and phenomenological effects. The correspondence is expressed in Eqn. 6.1.1 where the first term describes the phenomenological ratio of \(Z/\gamma\) coming from simulation, and the second term describes the measured photon yield, corrected for acceptance, efficiency and purity.

\[
N_{\text{data}}^{Z \rightarrow \nu \bar{\nu} + \text{jets}} = \frac{\sigma(Z \rightarrow \nu \bar{\nu} + \text{jets})}{\sigma(\gamma + \text{jets})} \left( H_T, H_T', N_{\text{jets}} \right)_{\text{MC}} \times N_{\text{data}}^{\text{direct} \gamma + \text{jets}} \quad (6.1.1)
\]

Several types of photons can be encountered while analyzing the collision data. Not all of them have an analogue to \(Z \rightarrow \nu \bar{\nu} + \text{jets}\) production. Thus it is important to reduce the level of background photons through appropriately chosen cuts. Understanding the source of the different types of photons will aid in selecting appropriate cuts to define a good sample of photons.

**Prompt photons** are photons that are either produced during the hard scatter process, or from final-state QED radiation (FSR) during the hadronization process. Prompt photons can be divided into two categories: direct photons and fragmentation photons.

**Direct photons** are those photons arising directly from the hard scattering process. It is these photons that are of interest when considering the analogue of \(Z\) production, as these photons are produced through a very similar process as a \(Z\) boson.

**Fragmentation photons** are photons arising during the hadronization of quarks and gluons coming out of the hard scattering process, i.e., from final-state radiation. Since
this process has no analogous Z production mechanism, this type of photon is considered a background. As these photons are typically accompanied by other hadronic activity, the isolation cuts help to reduce photons arising from this process.

**Secondary photons** are photons arising from the decay of mesons produced in jets. It should be clear that this type of photon is not an appropriate candidate to translate to a Z event. In fact, Z+jets events can produce photons in the same way, but it is the Z that needs to be replaced, not the jets producing photons. Thus some selection criterion must be imposed to reduce the number of secondary photons in the data sample. Usually this can be done by applying some form of an isolation requirement on the reconstructed photon, requiring that the photon has little nearby activity, making it less likely that the photon came from a jet subproduct. Additionally, the transverse spread of the photon shower in the ECal is useful to help distinguish single photons from tightly collimated photon pairs, which can be produced from certain meson decays. In this analysis, two different isolation criteria are used. The first (Table 6.1) is based solely on the sum of the various detector energy deposits in a cone of predefined distance around the reconstructed photon object. The second (Table 6.2) uses information from the particle-flow event reconstruction around the reconstructed photon object.

**Mistagged electrons** are actually electrons that have been reconstructed as photons. In simulation, mistagged electrons can be studied by matching the reconstructed photon back to a generator-level electron. The primary sources for this type of contamination are semileptonic t̅t+jets events and W±+jets events. When the W± (either from a W± produced in the primary interaction, or from the decay of one of the top quarks in a t̅t event) decays...
to an $e^\pm\nu$ pair and the $e^\pm$ passes the photon selection, the reconstructed object will end up in the collection of photon objects, as well as the electron collection. The usual method for rejecting this type of event is to require no hit in the pixel detector, as electrons should leave a hit in this innermost part of the tracking system. However, some events of this type will fail to leave a hit in the pixel detector, and thus still end up in the photon sample. This is accounted for in $\varepsilon_{\text{ID}}$, which includes the efficiency of the pixel veto, $\varepsilon_{\text{pixel}}$.

The complete expression relating the direct photon yield in data to the measured and selected isolated prompt photon yield is shown in Eqn. 6.1.2

$$N_{\text{data}}^{\text{direct}\gamma+\text{jets}} = \frac{\mathcal{P} \cdot N_{\text{obs,}\gamma+\text{jets}}^{\text{data}}}{\varepsilon_{\text{Acc}} \cdot \varepsilon_{\text{ID}} \cdot \varepsilon_{\text{Iso}} \cdot \text{SF}_{\text{data/MC}}}$$

Here $\mathcal{P}$ represents the purity of the selected isolated photon sample observed in data and is discussed in Section 6.4. The other correction factors, discussed in Section 6.3, account for photons lost due to detector acceptance ($\varepsilon_{\text{Acc}}$), reconstruction ID ($\varepsilon_{\text{ID}}$) and isolation ($\varepsilon_{\text{Iso}}$) efficiencies, and differences in efficiency (both ID and isolation) between data and simulation ($\text{SF}_{\text{data/MC}}$).

### 6.2 Photon event selection

A pure photon sample is constructed by asking for a high $p_T$ photon ($p_T > 100$ GeV) within the detector acceptance of the ECAL ($0 < |\eta| < 1.4442$ or $1.566 < |\eta| < 2.5$). To reject contamination from misidentified electrons (semileptonic $t\bar{t}+\text{jets}$ and $W^\pm+\text{jets}$), reconstructed photon candidates having a hit in the pixel tracker are rejected. Additionally, a cut on the fraction of the photon energy measured in the HCAL to that
measured in the ECAL helps to reject fake photons arising from meson decays in jets. For the 7 TeV analysis, the algorithm used to compute this fraction, $\frac{\text{had EM}}{\text{EM}}$, takes the HCal towers within a cone of $\Delta R = 0.15$ around the supercluster of the reconstructed photon and computes the fraction of energy deposited in the HCal towers to that deposited in the ECAL supercluster. For 8 TeV running conditions, this algorithm was shown to be too sensitive to pileup effects and was replaced with $\frac{\text{hadTowOverEM}}{\text{EM}}$, which uses only the single HCal tower centered on the highest energy ECAL crystal in the photon’s supercluster. The final requirement for the identification (ID) of photons comes from the so-called “shower shape” variable, $\sigma_{\eta\eta}$, which is the log-weighted energy spread in $\eta$ of the photon (Eqn. 6.2.1) and measures the transverse width of the electromagnetic shower.

$$
\sigma_{\eta\eta} = \sqrt{\sum_{i}^{5 \times 5} \left( \frac{\eta_{i}^\text{crystal}}{w_{i}} \times 0.0175 + \frac{\eta_{i}^\text{seed} - \bar{\eta}^{5 \times 5}}{w_{i}} \right)^2 \times w_{i}}
$$

$$
\bar{\eta}^{5 \times 5} = \sum_{i}^{5 \times 5} \frac{\eta_{i} \cdot w_{i}}{w_{i}}, \quad w_{i} = \max \left\{ 0, 4.2 + \ln \left( \frac{E_{i}}{E_{5 \times 5}} \right) \right\}
$$

To further reject QCD contamination from fragmentation and secondary photons, an isolation requirement is imposed on the photon to ensure that there is minimal energy in the neighborhood of the photon.

For the analysis of 7 TeV data, a cut is applied on the combined energy sums for tracker, ECAL, and HCAL deposits around the photon candidate within a cone of $\Delta R = 0.4$. This sum is then corrected for pileup energy contributions using a simple $\rho$ subtraction, where $\rho$ is the average pileup energy density per event, computed using $k_T$ PF jets with a distance parameter of $R = 0.6$. The complete set of selection requirements for 7 TeV $\gamma + \text{jets}$
Table 6.1: γ selection ID and isolation requirements used in the 2011 γ + jets analysis.

In the analysis of 8 TeV data, cuts are applied on energy sums for charged, neutral and photon particle-flow candidates within a cone of ΔR = 0.3 around the reconstructed photon object. These isolation sums have pileup energy contributions removed using an “effective area” ρ-subtraction. The effective area is determined from fitting both the ρ value and the isolation sum vs. the number of reconstructed primary vertices in an event, N_vtx, and taking the ratio of the slopes. Table 6.2 shows the complete set of selection requirements for 8 TeV γ + jets events.

Table 6.2: 2012 simple cut-based γ ID and isolation requirements, tight working point. The effective areas for pileup subtraction of isolation sums are listed in Table 6.3.

The offline photon selection is combined with an online trigger selection (Tables 6.4 (7 TeV) and 6.5 (8 TeV)), requiring a loosely identified photon combined with hadronic energy
\[ \eta | \text{Isolation sum} \]
\[ \eta \text{ISO}_{\Delta R=0.3} \]
\[ \eta \text{ISO}_{\Delta R=0.3}^\pm \]
\[ \eta \text{ISO}_{\Delta R=0.3}^\gamma \]

0.0 - 1.0 0.012 0.030 0.148
1.0 - 1.479 0.010 0.057 0.130
1.479 - 2.0 0.014 0.029 0.112
2.0 - 2.2 0.012 0.015 0.216
2.2 - 2.3 0.016 0.024 0.262
2.3 - 2.4 0.020 0.039 0.260
> 2.4 0.012 0.072 0.266

Table 6.3: \( \gamma \) effective areas for PU subtraction on different isolation sums.

to select events similar to those in the search region. The standard analysis cuts are then applied to this photon sample. The sample is taken from the collision datasets listed in

Table 6.4: Datasets used in the 7 TeV analysis. All data were reconstructed using CMSSW_4_2_X version of the CMS reconstruction software.

Table 6.5: Datasets used in the 8 TeV analysis. All data were reconstructed using CMSSW_5_3_5.
Tables 6.4 and 6.5. The Monte Carlo simulation samples used for the analysis are listed in Tables 6.6 and 6.7.

### Table 6.6: List of MC simulation samples used in the 2011 $\gamma +$ jets analysis.

<table>
<thead>
<tr>
<th>Dataset (Summer11-PU_S4_START42_V11-v1/AODSIM production)</th>
<th>Generated events</th>
<th>$\sigma$ [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>/GJets_TuneZ2_200_HT_inf_7TeV-madgraph</td>
<td>9377170</td>
<td>798.3 (LO)</td>
</tr>
<tr>
<td>/ZJetsToNuNu_200_HT_inf_7TeV-madgraph</td>
<td>3067017</td>
<td>42.2 (NNLO)</td>
</tr>
<tr>
<td>/G_Pt-15to30_TuneZ2_7TeV_pythia6</td>
<td>2046119</td>
<td>172000.0 (LO)</td>
</tr>
<tr>
<td>/G_Pt-30to50_TuneZ2_7TeV_pythia6</td>
<td>2187260</td>
<td>16700.0 (LO)</td>
</tr>
<tr>
<td>/G_Pt-80to120_TuneZ2_7TeV_pythia6</td>
<td>1746637</td>
<td>447.0 (LO)</td>
</tr>
<tr>
<td>/G_Pt-120to170_TuneZ2_7TeV_pythia6</td>
<td>2073216</td>
<td>84.2 (LO)</td>
</tr>
<tr>
<td>/G_Pt-170to300_TuneZ2_7TeV_pythia6</td>
<td>2068650</td>
<td>22.6 (LO)</td>
</tr>
<tr>
<td>/G_Pt-300to47_TuneZ2_7TeV_pythia6</td>
<td>1791880</td>
<td>1.49 (LO)</td>
</tr>
<tr>
<td>/QCD_TuneZ2_HT-100To250_7TeV-madgraph</td>
<td>22580264</td>
<td></td>
</tr>
<tr>
<td>/QCD_TuneZ2_HT-250To500_7TeV-madgraph</td>
<td>20674219</td>
<td></td>
</tr>
<tr>
<td>/QCD_TuneZ2_HT-500To1000_7TeV-madgraph</td>
<td>14437469</td>
<td></td>
</tr>
<tr>
<td>/QCD_TuneZ2_HT-1000_7TeV-madgraph</td>
<td>6294851</td>
<td></td>
</tr>
<tr>
<td>/DYJetsToLL_TuneZ2_M-50_7TeV-madgraph-tauola</td>
<td>36277961</td>
<td></td>
</tr>
</tbody>
</table>

### Table 6.7: List of MC simulation samples used in the 2012 $\gamma +$ jets analysis, together with cross section and total number of generated events.

<table>
<thead>
<tr>
<th>Dataset (Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM production)</th>
<th>Generated events</th>
<th>$\sigma$[pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>/GJets_HT-200To400_8TeV-madgraph</td>
<td>10494617</td>
<td>960.5(LO)</td>
</tr>
<tr>
<td>/GJets_HT-400ToInf_8TeV-madgraph</td>
<td>1611963</td>
<td>107.5(LO)</td>
</tr>
<tr>
<td>/GJets_HT-400ToInf_8TeV-madgraph_ext</td>
<td>9534744</td>
<td></td>
</tr>
<tr>
<td>/ZJetsToNuNu_200_HT_400_TuneZ2Star_8TeV_madgraph</td>
<td>5055885</td>
<td>49.28(NNLO)</td>
</tr>
<tr>
<td>/ZJetsToNuNu_400_HT_inf_TuneZ2Star_8TeV_madgraph</td>
<td>1006928</td>
<td>6.26(NNLO)</td>
</tr>
<tr>
<td>/WJetsToNuNu_HT-400ToInf_8TeV-madgraph</td>
<td>1647804</td>
<td>30.08(NNLO)</td>
</tr>
<tr>
<td>/WJetsToNuNu_HT-400ToInf_8TeV-madgraph_v2</td>
<td>4971847</td>
<td></td>
</tr>
<tr>
<td>/TTJets_SemiLeptMGDecays_8TeV-madgraph</td>
<td>11081685</td>
<td>103.0(NNLO)</td>
</tr>
<tr>
<td>/QCD_HT-250To500_TuneZ2Star_8TeV-madgraph-pythia6</td>
<td>27002490</td>
<td>1.036E7(LO)</td>
</tr>
<tr>
<td>/QCD_HT-500To1000_TuneZ2Star_8TeV-madgraph-pythia6</td>
<td>30599239</td>
<td>842.6(LO)</td>
</tr>
<tr>
<td>/QCD_HT-1000ToInf_TuneZ2Star_8TeV-madgraph-pythia6</td>
<td>13829955</td>
<td>204.0(LO)</td>
</tr>
</tbody>
</table>

Figure 6.5 shows data vs. simulation ($\gamma +$ jets and QCD) comparisons for the recomputed event variables $H_T$ and $\bar{H}_T$ for the 7 TeV analysis. Figure 6.6 shows data vs. simulation ($\gamma +$ jets only) comparisons of several photon variables, $p_T$, $\eta$, and the individual
isolation variables used to compute the combined isolation variable.

Figure 6.5: 7 TeV data vs. simulation comparisons for event variables $H_T$ and $H_T^*$ with a baseline selection of $H_T > 500$ GeV and $H_T^* > 200$ GeV, and $N_{jets} \geq 3$ (from Ref. [55]).

Data vs. simulation comparisons of the recomputed event search variables (Fig. 6.7), as well as photon $p_T$, $\eta$, and isolation variables (Fig. 6.8) are shown for the 8 TeV analysis. Overall, good agreement is seen. For a complete set of data vs. simulation comparisons see Appendix A.3.

### 6.3 Photon reconstruction efficiency measurement

The efficiencies of the acceptance, reconstruction, identification, and isolation cuts for selected photons were obtained from Monte Carlo simulated events. Generator-level photons are selected if they can be classified as direct photons, i.e., photons coming from the matrix element calculation. The acceptance is defined as the fraction of generator-level photons that pass $p_T$ and $\eta$ requirements divided by the total number of generated photons.
Figure 6.6: 7 TeV data vs. simulation comparisons for several photon variables for a baseline selection of $H_T > 500$ GeV and $\not{H}_T > 200$ GeV, and $N_{jets} \geq 3$ (from Ref. [55]).
Figure 6.7: 8 TeV data vs. simulation comparisons for several event variables for a baseline selection of $H_T > 500$ GeV and $H_T > 200$ GeV, and $N_{jets} \geq 3$. 

(a) $N_{jets}(p_T > 50$ GeV, $|\eta| < 2.5)$

(b) $H_T$

(c) $H_T$
Figure 6.8: 8 TeV data vs. simulation comparisons for several photon variables for a baseline selection of $H_T > 500$ GeV and $H_T > 200$ GeV, and $N_{jets} \geq 3$. 

79
in the sample. For the 7 TeV analysis, the acceptance was parametrized as a function of $H_T$ and $HH_T$ in the selected search bins, as seen in Fig. 6.9a. For the 8 TeV analysis, the acceptance was parametrized as a function of $N_{\text{jets}}$ in three $HH_T$ bins corresponding to the parametrization of the phenomenological ratio.

The reconstruction efficiencies are measured in simulation by looking for a reconstructed and identified photon that is matched to a generator-level prompt photon. The isolation efficiency further requires that the reconstructed photon pass the isolation cuts and is defined as the efficiency of a photon passing the reconstruction and identification cuts to also pass the isolation cuts. For the 7 TeV analysis, the reconstruction efficiency combined the ID with the isolation and was parametrized as a function of both $H_T$ and $HH_T$ in the 14 defined search bins, as seen in Fig. 6.9b. Again, for the 8 TeV analysis, all the efficiencies are parametrized as a function of $N_{\text{jets}}$ in three $HH_T$ bins. The acceptance and different reconstruction efficiencies (ID, pixel seed veto, and isolation) in each of the three $HH_T$ bins are shown in Fig. 6.10.

6.3.1 Data/MC scale factor

Differences in the reconstruction efficiency between data and simulation are taken into account via a data-to-simulation scale factor ($SF_{\text{data/MC}}$). For the 7 TeV analysis the scale factor for selecting an identified and isolated photon was studied using a T&P method. The data sample utilized the same photon-triggered dataset as the main analysis, while the samples used for simulation were Drell–Yan dilepton plus jets samples with $M_{l^{+}l^{-}} > 50$ GeV. The tag object was chosen to be a well-identified electron, while the probe was a loosely
Figure 6.9: 7 TeV $\gamma$ acceptance and reconstruction efficiency vs. $H_T$ and $MHT$. 

(a) Geometric acceptance vs. $H_T$ and $MHT$.

(b) Reconstruction (ID+Iso) efficiency vs. $H_T$ and $MHT$. 

81
Figure 6.10: 8 TeV $\gamma$ acceptance, reconstruction ID, pixel veto, and isolation efficiency vs. $N_{\text{jets}}$ (in the 3 $H_T$ bins).
defined electron based primarily on information from the ECAL supercluster, which is the same for photon reconstruction. Probes were checked to determine whether they passed the full photon selection (Table 6.1 without the pixel seed veto) and the invariant mass of the T&P pair was required to lie within 20 GeV of the Z boson mass ($M_Z = 91.12$ GeV). The invariant mass distribution for passing, failing, and total probes is fit using the Z lineshape convoluted with a Gaussian and a linearly falling background (Fig. 6.11) in bins of some variable to determine the efficiency dependence on that same variable, e.g., $N_{vtx}$, $p_T$, or $|\eta|$. The resulting fitted function is then used to extract the efficiency, taking the integral of the passing probes, and dividing by the integral of the total probes. A similar method was also utilized to confirm that the efficiency of the pixel seed veto as a function of the pileup in
Figure 6.11: Sample of T&P fits performed in a single $N_{\text{vtx}}$ bin for calculating efficiencies vs. $N_{\text{vtx}}$.

the event was flat (Fig. 6.12).

Figure 6.12: Efficiency of the pixel seed veto as a function of the pileup in the event.

For the 8 TeV analysis, these scale factors were obtained from the official CMS EGamma POG, which used a T&P method similar to that described above.
6.4 Photon purity measurement

The photon selection is slightly contaminated by events in which a jet fragments in such a way as to pass the tight photon selection. This typically happens when the jet is composed of high-\(p_T\) neutral mesons, e.g., \(\pi^0\)s, which then decay to photons. These photons can be tightly collimated and thus appear as a single photon in the detector. A key variable to distinguish these events from true photon events is the “shower shape” variable, i.e., \(\sigma_{\text{iso}}\), which measures the transverse spread of the shower in the ECAL supercluster. This variable has a characteristic shape for true photons and is thus good for discriminating between true photons and fake photons.

To estimate the contribution of jets faking photons in the signal region, a “template method” is employed. A signal and a background template are fitted (using either RooFit [56] or the ROOT [57] TFractionFitter method) to a distinctive distribution in data for photons passing all cuts except for the cut on the fitting variable (combined isolation variable or \(\sigma_{\text{iso}}\) variable). The relative fractions are extracted and the templates are then scaled accordingly. The purity is obtained by integrating the two templates in the signal region of the variable and comparing the signal contribution to the total contribution (Eqn. 6.4.1).

\[
P = \frac{n_{\text{sig}}}{n_{\text{tot}}} \tag{6.4.1}
\]

The two different templates and methods are discussed in the following sections.
6.4.1 7 TeV purity measurement: combined isolation variable

For analysis of 7 TeV data, the measurement of the purity of the photon selection was performed using a method developed by the CMS QCD Photons group. The measurement is performed on the combined isolation variable by creating two templates, one corresponding to the signal and the other corresponding to the background. The signal template is taken from \( \gamma + \text{jets} \) simulated events, based on the nominal photon selection (Table 6.1). To create a background template, a sideband selection is performed on the data. The sideband region is defined by inverting the cuts on the shower shape variable, \( \sigma_{\text{iso}} \). In both cases, no cut is applied on the combined isolation variable.

Once the two templates are defined, their combination is fitted, using the RooFit package, to the signal distribution in data. This is done by first fitting the signal template with a Crystal Ball function \([58, 59]\). The Crystal Ball function consists of a Gaussian convoluted with an exponential tail to one side and models the behavior of the combined isolation variable rather well. This functional form is added to the raw histogram template from the sideband and their sum is used to fit to the \( \sigma_{\text{iso}} \) distribution in data, allowing some of the parameters of the signal function to float during the combined fitting stage to improve the fit quality.

Once the fit has converged, the final normalized functional forms are integrated in the region of interest (Iso\(_{\text{combined}} < 5.0\)). From this, the number of signal and background events in the signal region in data can be obtained, and the purity calculated as in Eqn. 6.4.1. Fig. 6.13 shows examples of the signal template, background template, and sum template, once it has been fitted to the data distribution.
Figure 6.13: Signal template and data distribution with fitted sum, showing signal and background contributions.
The purity is calculated in several bins in the \( \eta \) and \( p_T \) of the photon and the values are shown in Table 6.8. These values are used to obtain the purity in the search regions, defined by \( H_T \) and \( \mathbb{H}_T \), by taking the purity-weighted event yield for all events in each of the bins (Table 6.9). The systematic uncertainty is obtained by changing the range over which the fit is performed, as well as fixing several of the signal template parameters during the combined fitting step. The maximum variation around the nominal value is taken to be the uncertainty on the purity. The statistical uncertainty is taken as the uncertainty on the signal fraction returned by the combined fitting.

<table>
<thead>
<tr>
<th>( \eta )</th>
<th>( \mathcal{P} \pm (\text{stat.}) \pm (\text{syst.}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0 &lt;</td>
<td>\eta</td>
</tr>
<tr>
<td>( 0.9 &lt;</td>
<td>\eta</td>
</tr>
<tr>
<td>( 1.566 &lt;</td>
<td>\eta</td>
</tr>
<tr>
<td>( 2.1 &lt;</td>
<td>\eta</td>
</tr>
</tbody>
</table>

Table 6.8: Photon purity in bins of \( p_T^\gamma \) and \( |\eta| \) determined using the combined isolation template method with statistical and systematic uncertainties.

### 6.4.2 8 TeV purity measurement: \( \sigma_{inj\eta} \) Variable

For the analysis of 8 TeV data, the template variable is taken to be \( \sigma_{inj\eta} \) itself. Because the isolation variables used to select the 8 TeV data control sample are not combined, as they were in the analysis of the 7 TeV data, it was not straightforward to utilize the 7 TeV method. The signal template is taken from \( \gamma + \) jets simulated events passing the “tight” photon selection (Table 6.2) and matched to a generator-level photon, without the cut on \( \sigma_{inj\eta} \), while the background template is taken from a sideband selection performed on the data. The sideband region is defined to consist of the class of events that could fluc-
Table 6.9: Photon purity and event yields for the exclusive search regions determined using the combined isolation template method.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>500...800</td>
<td>200...350</td>
<td>97.79</td>
<td>$\pm$ 1.27</td>
<td>$\pm$ 1.60</td>
</tr>
<tr>
<td>500...800</td>
<td>350...500</td>
<td>99.03</td>
<td>$\pm$ 2.54</td>
<td>$\pm$ 0.94</td>
</tr>
<tr>
<td>500...800</td>
<td>500...600</td>
<td>99.20</td>
<td>$\pm$ 2.43</td>
<td>$\pm$ 0.80</td>
</tr>
<tr>
<td>500...800</td>
<td>600...</td>
<td>99.38</td>
<td>$\pm$ 2.26</td>
<td>$\pm$ 0.62</td>
</tr>
<tr>
<td>800...1000</td>
<td>200...350</td>
<td>97.98</td>
<td>$\pm$ 1.24</td>
<td>$\pm$ 1.48</td>
</tr>
<tr>
<td>800...1000</td>
<td>350...500</td>
<td>99.05</td>
<td>$\pm$ 2.45</td>
<td>$\pm$ 0.83</td>
</tr>
<tr>
<td>800...1000</td>
<td>500...600</td>
<td>99.37</td>
<td>$\pm$ 2.27</td>
<td>$\pm$ 0.63</td>
</tr>
<tr>
<td>800...1000</td>
<td>600...</td>
<td>99.30</td>
<td>$\pm$ 2.34</td>
<td>$\pm$ 0.70</td>
</tr>
<tr>
<td>1000...1200</td>
<td>200...350</td>
<td>97.66</td>
<td>$\pm$ 1.21</td>
<td>$\pm$ 1.72</td>
</tr>
<tr>
<td>1000...1200</td>
<td>350...500</td>
<td>99.15</td>
<td>$\pm$ 2.51</td>
<td>$\pm$ 0.85</td>
</tr>
<tr>
<td>1000...1200</td>
<td>500...</td>
<td>99.21</td>
<td>$\pm$ 2.44</td>
<td>$\pm$ 0.79</td>
</tr>
<tr>
<td>1200...1400</td>
<td>200...350</td>
<td>98.08</td>
<td>$\pm$ 1.43</td>
<td>$\pm$ 1.10</td>
</tr>
<tr>
<td>1200...1400</td>
<td>350...</td>
<td>99.31</td>
<td>$\pm$ 2.33</td>
<td>$\pm$ 0.69</td>
</tr>
<tr>
<td>1400...</td>
<td>200...</td>
<td>98.42</td>
<td>$\pm$ 1.58</td>
<td>$\pm$ 1.02</td>
</tr>
</tbody>
</table>

The photon candidates in the sideband must pass a very loose photon ID and isolation requirement and they must further fail at least one of the tight isolation requirements. This selection ensures that the objects in the background template are photon-like. The full sideband selection is shown in Table 6.10.
The purity is extracted by fitting, using the TFractionFitter method of ROOT, the signal and background templates to the \( \sigma_{\text{sign}} \) distribution taken from a data sample. The data sample is constructed by selecting photon events where the photon passes all of the tight photon cuts with the exception of the cut on \( \sigma_{\text{sign}} \) (Table 6.2). The fitting method scales the two templates up and down to find the proportion that most accurately matches the data distribution. When this is done, the sum template can be integrated in the signal region to find the total number of expected events \( n_{\text{tot}} \). The final signal template is also integrated in the same way to obtain the number of signal events \( n_{\text{sig}} \). The purity is then defined as in Eqn. 6.4.1.

The purity as a function of \( N_{\text{jets}} \) is shown in Fig. 6.14 for both barrel and endcap photons. The purity is computed in several exclusive \( N_{\text{jets}} \) bins and a final inclusive one and an extrapolation of the purity based on these points is used to perform the estimate. For reference, the fits leading to these plots are shown in Figs. 6.15 and 6.16.

The systematic uncertainty is obtained by scaling the signal template up or down.

Table 6.10: Sideband selection requirements for 8 TeV photon purity background template selection.
by the uncertainty on the signal fraction returned by the TFractionFitter. The signal yield is then recomputed with the scaled histogram and the adjusted purity is obtained from Eqn. 6.4.1. The statistical uncertainty is given by Eqn. 6.4.2.

\[
\sigma_{\text{stat.}} = \sqrt{\frac{P \cdot (1 - P)}{n_{\text{tot}}}} \tag{6.4.2}
\]

![Graphs showing photon selection purity as a function of Njets](image)

(a) Photon selection purity vs. Njets for barrel photons (b) Photon selection purity vs. Njets for endcap photons for \( \mathbb{H}_T > 200 \) GeV.

**Figure 6.14:** Photon selection purity as a function of Njets. The red bands show the systematic uncertainty while the blue bands show the statistical uncertainty.
Figure 6.15: Barrel photon purity fits for $H_T > 200$, in different $N_{\text{jets}}$ bins, the black line shows the value of the upper cut used for integrating the purity as well as selecting well-identified photons.
Figure 6.15: (Continued) Barrel photon purity fits for $H_T > 200$, in different $N_{\text{jets}}$ bins, the black line shows the value of the upper cut used for integrating the purity as well as selecting well-identified photons.
Figure 6.16: Endcap photon purity fits for $p_T > 200$, in different $N_{jets}$ bins, the black line shows the value of the upper cut used for integrating the purity as well as selecting well-identified photons.
(b) Endcap photon purity fits log scale y-axis.

Figure 6.16: (Continued) Endcap photon purity fits for $\not{E}_T > 200$, in different $N_{jets}$ bins, the black line shows the value of the upper cut used for integrating the purity as well as selecting well-identified photons.
6.5 Analysis of uncertainties

The primary sources of uncertainty in the estimation are due to the parametrization of the \( Z/\gamma \) ratio and the efficiencies vs. \( H_T, \frac{H_T}{H_T}, \) and \( N_{jets} \), the uncertainty on the purity measurement, the theoretical uncertainty on the \( Z/\gamma \) ratio, and the limited number of events in several of the control regions. The treatment of the uncertainties for the 7 TeV and 8 TeV analyses are discussed in this section.

6.5.1 7 TeV results

For the 7 TeV analysis, the theoretical uncertainty on the \( Z/\gamma \) ratio was obtained from the BLACKHAT theory group [60], who provided uncertainties in an inclusive \( (N_{jets} \geq 3) \) sample, in various \( H_T \) and \( \frac{H_T}{H_T} \) regions. The uncertainties were obtained by comparing leading-order matrix element with parton showering (LO ME+PS) calculation to the full next-to-leading-order (NLO) calculation.

The next significant uncertainty comes from the statistical uncertainty on the measurement of the \( Z/\gamma \) ratio taken from the simulation in each of the 14 search regions. This uncertainty is due to the low number of events passing the selection requirements in several of the bins.

The additional systematic uncertainties are taken to be:

- a 20% uncertainty on the 5% fragmentation component \((1-0.95)\) [61] to the prompt photon yield, resulting in an uncertainty of 1% on the phenomenological ratio
- the statistical uncertainty on the acceptance efficiency in each control region, taken from the simulation
- the statistical uncertainty on reconstruction ID/isolation efficiency in each control region, taken from the simulation
- a conservative 5% systematic uncertainty on the total efficiency
• a 1% uncertainty on the data-to-simulation scale factor measured using a T&P method (cf. Section 6.3.1)

• a 1% uncertainty on the mistag rate [62, 63]

• the uncertainty on the purity measurement using the combined isolation template method (cf. Section 6.4.1)

These individual uncertainties are combined under the assumption that they are uncorrelated. The results, detailing the various sources of uncertainty for the prediction, are shown in Table 6.11, along with the corresponding correction factor for each of the 14 search bins.

### 6.5.2 8 TeV analysis

The primary difference in the 8 TeV analysis with respect to the 7 TeV analysis was the additional dimension of $N_{\text{jets}}$ in the search regions. Due to this, a new treatment to constrain the uncertainty coming from the theoretical $Z/\gamma$ ratio was needed.

**Using $Z \rightarrow \mu^+ \mu^-$ events to constrain the theory uncertainty on the $Z/\gamma$ ratio**

The uncertainty coming from the theoretical $Z/\gamma$ ratio was evaluated by the BLACKHAT theory group for the 7 TeV analysis, which did not bin the search regions in the $N_{\text{jets}}$ variable. Their results were incorporated into the 7 TeV analysis. However, as the 8 TeV analysis introduced a binning of the search regions in the $N_{\text{jets}}$ variable, the inclusive uncertainty ($N_{\text{jets}} \geq 3$) obtained from the BLACKHAT group was not known to be sufficient. A method to estimate the uncertainty on the $Z/\gamma$ ratio as a function of $N_{\text{jets}}$ was developed to address this issue. A sample of $Z \rightarrow \mu^+ \mu^- + \text{jets}$ events is selected by requiring two isolated, opposite-sign muon candidates, with the composite $M_{\mu^+ \mu^-}$ within 20 GeV of the Z mass, $M_{\mu^+ \mu^-} < \ldots$
<table>
<thead>
<tr>
<th>$H_T$ [GeV]</th>
<th>$\bar{H}_T$ [GeV]</th>
<th>Correction factor</th>
<th>$Z/\gamma$ ratio $\pm$ (stat) $\pm$ (theoretical)</th>
<th>Acc. $\pm 5%$(syst.)</th>
<th>Reco/Iso $\pm 5%$(syst.)</th>
<th>Purity</th>
</tr>
</thead>
<tbody>
<tr>
<td>500...800</td>
<td>200...350</td>
<td>0.27 $^{+0.0157}_{-0.0157}$ (syst.)</td>
<td>0.21 $\pm 0.0041$ $\pm 0.0432$</td>
<td>0.92 $^{+0.0045}_{-0.0044}$</td>
<td>0.79 $^{+0.0069}_{-0.0068}$</td>
<td>0.98 $\pm 0.0127$</td>
</tr>
<tr>
<td>500...800</td>
<td>350...500</td>
<td>0.32 $^{+0.0239}_{-0.0239}$ (syst.)</td>
<td>0.24 $\pm 0.0093$ $\pm 0.0500$</td>
<td>0.91 $^{+0.0094}_{-0.0089}$</td>
<td>0.76 $^{+0.0143}_{-0.0139}$</td>
<td>0.99 $\pm 0.0254$</td>
</tr>
<tr>
<td>500...800</td>
<td>500...600</td>
<td>0.30 $^{+0.0354}_{-0.0349}$ (syst.)</td>
<td>0.24 $\pm 0.0216$ $\pm 0.0501$</td>
<td>0.94 $^{+0.0188}_{-0.0157}$</td>
<td>0.79 $^{+0.0317}_{-0.0295}$</td>
<td>0.99 $\pm 0.0243$</td>
</tr>
<tr>
<td>500...800</td>
<td>600...</td>
<td>0.37 $^{+0.0772}_{-0.0750}$ (syst.)</td>
<td>0.27 $\pm 0.0478$ $\pm 0.0894$</td>
<td>0.96 $^{+0.0370}_{-0.0236}$</td>
<td>0.73 $^{+0.0681}_{-0.0614}$</td>
<td>0.99 $\pm 0.0226$</td>
</tr>
<tr>
<td>800...1000</td>
<td>200...350</td>
<td>0.31 $^{+0.0265}_{-0.0263}$ (syst.)</td>
<td>0.23 $\pm 0.0136$ $\pm 0.0833$</td>
<td>0.91 $^{+0.0144}_{-0.0131}$</td>
<td>0.77 $^{+0.0212}_{-0.0203}$</td>
<td>0.98 $\pm 0.0124$</td>
</tr>
<tr>
<td>800...1000</td>
<td>350...500</td>
<td>0.36 $^{+0.0489}_{-0.0480}$ (syst.)</td>
<td>0.29 $\pm 0.0319$ $\pm 0.1029$</td>
<td>0.95 $^{+0.0233}_{-0.0181}$</td>
<td>0.80 $^{+0.0397}_{-0.0361}$</td>
<td>0.99 $\pm 0.0245$</td>
</tr>
<tr>
<td>800...1000</td>
<td>500...600</td>
<td>0.54 $^{+0.1396}_{-0.1347}$ (syst.)</td>
<td>0.38 $\pm 0.0805$ $\pm 0.1377$</td>
<td>0.97 $^{+0.0462}_{-0.0238}$</td>
<td>0.69 $^{+0.0875}_{-0.0792}$</td>
<td>0.99 $\pm 0.0227$</td>
</tr>
<tr>
<td>800...1000</td>
<td>600...</td>
<td>0.30 $^{+0.0639}_{-0.0622}$ (syst.)</td>
<td>0.24 $\pm 0.0443$ $\pm 0.0878$</td>
<td>0.98 $^{+0.0310}_{-0.0157}$</td>
<td>0.78 $^{+0.0649}_{-0.0570}$</td>
<td>0.99 $\pm 0.0234$</td>
</tr>
<tr>
<td>1000...1200</td>
<td>200...350</td>
<td>0.27 $^{+0.0351}_{-0.0345}$ (syst.)</td>
<td>0.22 $\pm 0.0226$ $\pm 0.0914$</td>
<td>0.91 $^{+0.0254}_{-0.0217}$</td>
<td>0.81 $^{+0.0353}_{-0.0322}$</td>
<td>0.98 $\pm 0.0121$</td>
</tr>
<tr>
<td>1000...1200</td>
<td>350...500</td>
<td>0.39 $^{+0.0917}_{-0.0884}$ (syst.)</td>
<td>0.30 $\pm 0.0592$ $\pm 0.1250$</td>
<td>0.94 $^{+0.0473}_{-0.0306}$</td>
<td>0.76 $^{+0.0761}_{-0.0661}$</td>
<td>0.99 $\pm 0.0251$</td>
</tr>
<tr>
<td>1000...1200</td>
<td>500...</td>
<td>0.73 $^{+0.2492}_{-0.2437}$ (syst.)</td>
<td>0.37 $\pm 0.0927$ $\pm 0.1538$</td>
<td>0.95 $^{+0.0637}_{-0.0337}$</td>
<td>0.50 $^{+0.1066}_{-0.0906}$</td>
<td>0.99 $\pm 0.0244$</td>
</tr>
<tr>
<td>1200...1400</td>
<td>200...350</td>
<td>0.31 $^{+0.0641}_{-0.0622}$ (syst.)</td>
<td>0.24 $\pm 0.0430$ $\pm 0.1026$</td>
<td>0.94 $^{+0.0406}_{-0.0286}$</td>
<td>0.78 $^{+0.0638}_{-0.0559}$</td>
<td>0.98 $\pm 0.0143$</td>
</tr>
<tr>
<td>1200...1400</td>
<td>350...</td>
<td>0.39 $^{+0.1132}_{-0.1079}$ (syst.)</td>
<td>0.36 $\pm 0.0971$ $\pm 0.1533$</td>
<td>1.00 $^{+0.0558}_{-0.0000}$</td>
<td>0.89 $^{+0.0829}_{-0.0562}$</td>
<td>0.99 $\pm 0.0233$</td>
</tr>
<tr>
<td>1400...200...</td>
<td>0.18 $^{+0.0411}_{-0.0402}$ (syst.)</td>
<td>0.13 $\pm 0.0268$ $\pm 0.0552$</td>
<td>0.96 $^{+0.0370}_{-0.0236}$</td>
<td>0.73 $^{+0.0681}_{-0.0614}$</td>
<td>0.98 $\pm 0.0158$</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.11: Correction factor and individual components, along with systematic errors, for the $\gamma +$ jets method (cf. Table A.1).
The kinematic cuts, $p_T > 20$ GeV and $|\eta| < 2.1$, applied to each muon are driven by the trigger requirement. The muons are then removed, as in the case for the normal photon analysis, and a reduced baseline selection of $H_T > 350$ GeV and $H_T > 100$ GeV is applied. This reduced selection was chosen to have enough events to see any potential trend. From this, the $Z/\gamma$ ratio was calculated using the $Z \rightarrow \mu^+ \mu^- + \text{jets}$ events and the $\gamma + \text{jets}$ events. The ratio was computed using data events as well as simulated events. The double ratio, data/simulation, was fit vs. $N_{\text{jets}}$ to a first degree polynomial and is shown in Fig. 6.17. Using the fit of the double ratio, Eqn. 6.1.1 was modified as shown in Eqn. 6.5.1 to account for differences in the ratio coming from data vs. simulation. The uncertainty on this fit, which is largely due to having very few events in the tails, is taken to be the uncertainty on the true ratio and is added to the other uncertainties in lieu of the theoretical uncertainty.

\[
N_{Z \rightarrow \nu \bar{\nu} + \text{jets}}^{\text{data}} = \left[ \frac{\sigma(Z \rightarrow \nu \bar{\nu} + \text{jets})}{\sigma(\gamma + \text{jets})} \right] (H_T, H_T, N_{\text{jets}})_{\text{MC}} \times R_{Z(\mu^+ \mu^-)/\gamma}^{\text{data/MC}} (N_{\text{jets}}) \times N_{\text{prompt} \gamma + \text{jets}}^{\text{data}}
\]

(6.5.1)

**Additional uncertainties coming from fitting the efficiencies vs. $N_{\text{jets}}$**

The uncertainties on the acceptance, reconstruction ID, pixel veto, and isolation cuts were all taken from the uncertainties on the fits of the respective efficiencies vs. $N_{\text{jets}}$. Correlations between the two parameters in the fit were taken into account using the correlation coefficient returned by the fit routine.

Tables listing the various sources of uncertainty for the prediction along with the
Figure 6.17: $R_{Z(\mu^+\mu^-)/\gamma}$ as a function of $N_{\text{jets}}$, $H_T > 200$ GeV, $H_T > 500$ GeV. The blue band shows the uncertainty on the fit, which is used as the uncertainty on the theoretical $Z/\gamma$ ratio.

corresponding raw uncertainty (Table 6.12) and percent uncertainty (Table A.2) for each of the 36 search bins are shown.

6.6 Prediction of the $Z\rightarrow\nu\overline{\nu}$ + jets background

The full $\gamma$ + jets data-driven prediction of the $Z\rightarrow\nu\overline{\nu}$ + jets background is presented in this section for both the 7 TeV analysis and the 8 TeV analysis. It is seen that the data-driven prediction is consistent with the expectation coming from the simulated $Z\rightarrow\nu\overline{\nu}$ + jets events passing the search selection.
\renewcommand{\arraystretch}{1.3}
\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|}
\hline
$3 \leq N_{\text{jets}} \leq 5$ & $H_T$ [GeV] & $H_T$ [GeV] & Data-driven & & & & & & & \\
\hline
\hline
500 & 800 & 200 & 300 & 450 & 600 & 2.63 & 4.04 & & & \\
\hline
500 & 800 & 300 & 450 & 600 & 2.63 & 4.04 & & & & \\
\hline
500 & 800 & 450 & 600 & 2.63 & 4.04 & & & & & \\
\hline
500 & 800 & 600 & 2.63 & 4.04 & & & & & & \\
\hline
800 & 1000 & 200 & 300 & 450 & 600 & 2.63 & 4.04 & & & \\
\hline
800 & 1000 & 300 & 450 & 600 & 2.63 & 4.04 & & & & \\
\hline
800 & 1000 & 450 & 600 & 2.63 & 4.04 & & & & & \\
\hline
800 & 1000 & 600 & 2.63 & 4.04 & & & & & & \\
\hline
1000 & 1250 & 1000 & 600 & 2.63 & 4.04 & & & & & \\
\hline
1000 & 1250 & 300 & 600 & 2.63 & 4.04 & & & & & \\
\hline
1000 & 1250 & 450 & 600 & 2.63 & 4.04 & & & & & \\
\hline
1000 & 1250 & 600 & 600 & 2.63 & 4.04 & & & & & \\
\hline
1250 & 1500 & 200 & 300 & 450 & 600 & 2.63 & 4.04 & & & \\
\hline
1250 & 1500 & 300 & 450 & 600 & 2.63 & 4.04 & & & & \\
\hline
1250 & 1500 & 450 & 600 & 2.63 & 4.04 & & & & & \\
\hline
1500 & 300 & 600 & 2.63 & 4.04 & & & & & & \\
\hline
\end{tabular}
\end{table}

(a) $3 \leq N_{\text{jets}} \leq 5$

\textbf{Table 6.12:} $Z \rightarrow \nu \bar{\nu} + \text{jets}$ uncertainties from $\gamma + \text{jets}$ data for $3 \leq N_{\text{jets}} \leq 5$ for the 8 TeV analysis (cf. Table A.2).
### Data-driven \( 6 \leq N_{\text{jets}} \leq 7 \) uncertainties

<table>
<thead>
<tr>
<th>( H_T ) [GeV]</th>
<th>( \not{H}_T ) [GeV]</th>
<th>Data-driven prediction</th>
<th>stat.</th>
<th>acc.</th>
<th>reco</th>
<th>pixel veto</th>
<th>iso.</th>
<th>data/MC</th>
<th>purity</th>
<th>ratio</th>
<th>pheno. fit</th>
<th>double</th>
</tr>
</thead>
<tbody>
<tr>
<td>500...800</td>
<td>200...300</td>
<td>22.72</td>
<td>± 2.78</td>
<td>+0.63</td>
<td>+0.34</td>
<td>+0.93</td>
<td>± 2.11</td>
<td>+0.81</td>
<td>+0.87</td>
<td>+0.37</td>
<td>+1.20</td>
<td>+4.51</td>
</tr>
<tr>
<td>500...800</td>
<td>300...450</td>
<td>9.92</td>
<td>± 1.88</td>
<td>−0.30</td>
<td>+0.22</td>
<td>+0.43</td>
<td>−2.11</td>
<td>−0.81</td>
<td>−0.93</td>
<td>−0.37</td>
<td>−1.20</td>
<td>−4.51</td>
</tr>
<tr>
<td>500...800</td>
<td>450...</td>
<td>0.70</td>
<td>± 0.50</td>
<td>+0.04</td>
<td>+0.03</td>
<td>+0.06</td>
<td>−1.88</td>
<td>+0.38</td>
<td>+0.43</td>
<td>+0.30</td>
<td>+0.07</td>
<td>+0.14</td>
</tr>
<tr>
<td>800...1000</td>
<td>200...300</td>
<td>9.09</td>
<td>± 1.75</td>
<td>+0.27</td>
<td>+0.14</td>
<td>+0.40</td>
<td>+1.08</td>
<td>+0.32</td>
<td>+0.40</td>
<td>−0.21</td>
<td>+0.51</td>
<td>+1.96</td>
</tr>
<tr>
<td>800...1000</td>
<td>300...450</td>
<td>4.23</td>
<td>± 1.22</td>
<td>+0.14</td>
<td>+0.10</td>
<td>+0.20</td>
<td>+0.50</td>
<td>+0.13</td>
<td>+0.17</td>
<td>+0.09</td>
<td>+0.23</td>
<td>+0.85</td>
</tr>
<tr>
<td>800...1000</td>
<td>450...</td>
<td>1.77</td>
<td>± 0.79</td>
<td>+0.08</td>
<td>+0.07</td>
<td>+0.13</td>
<td>+0.34</td>
<td>+0.05</td>
<td>+0.07</td>
<td>+0.06</td>
<td>+0.16</td>
<td>+0.35</td>
</tr>
<tr>
<td>1000...1250</td>
<td>200...300</td>
<td>4.37</td>
<td>± 1.21</td>
<td>+0.13</td>
<td>+0.07</td>
<td>+0.19</td>
<td>+0.43</td>
<td>+0.16</td>
<td>+0.17</td>
<td>+0.07</td>
<td>+0.24</td>
<td>+0.87</td>
</tr>
<tr>
<td>1000...1250</td>
<td>300...450</td>
<td>3.52</td>
<td>± 1.11</td>
<td>+0.11</td>
<td>+0.08</td>
<td>+0.16</td>
<td>+0.41</td>
<td>+0.11</td>
<td>+0.14</td>
<td>+0.08</td>
<td>+0.19</td>
<td>+0.70</td>
</tr>
<tr>
<td>1000...1250</td>
<td>450...</td>
<td>1.41</td>
<td>± 0.70</td>
<td>+0.07</td>
<td>+0.05</td>
<td>+0.11</td>
<td>+0.28</td>
<td>+0.05</td>
<td>+0.05</td>
<td>+0.05</td>
<td>+0.13</td>
<td>+0.28</td>
</tr>
<tr>
<td>1250...1500</td>
<td>200...300</td>
<td>3.33</td>
<td>± 1.05</td>
<td>+0.11</td>
<td>+0.06</td>
<td>+0.16</td>
<td>+0.36</td>
<td>+0.12</td>
<td>+0.13</td>
<td>+0.06</td>
<td>+0.21</td>
<td>+0.68</td>
</tr>
<tr>
<td>1250...1500</td>
<td>300...450</td>
<td>1.42</td>
<td>± 0.71</td>
<td>+0.04</td>
<td>+0.03</td>
<td>+0.06</td>
<td>+0.15</td>
<td>+0.04</td>
<td>+0.06</td>
<td>+0.03</td>
<td>+0.07</td>
<td>+0.28</td>
</tr>
<tr>
<td>1250...1500</td>
<td>450...</td>
<td>0.36</td>
<td>± 0.36</td>
<td>+0.02</td>
<td>+0.01</td>
<td>+0.02</td>
<td>+0.06</td>
<td>+0.01</td>
<td>+0.01</td>
<td>+0.03</td>
<td>+0.07</td>
<td>+0.07</td>
</tr>
<tr>
<td>1500...</td>
<td>200...300</td>
<td>1.33</td>
<td>± 0.67</td>
<td>+0.04</td>
<td>+0.02</td>
<td>+0.06</td>
<td>+0.14</td>
<td>+0.05</td>
<td>+0.05</td>
<td>+0.02</td>
<td>+0.08</td>
<td>+0.27</td>
</tr>
<tr>
<td>1500...</td>
<td>300...</td>
<td>1.06</td>
<td>± 0.61</td>
<td>+0.05</td>
<td>+0.04</td>
<td>+0.07</td>
<td>+0.19</td>
<td>+0.03</td>
<td>+0.04</td>
<td>+0.03</td>
<td>+0.09</td>
<td>+0.21</td>
</tr>
</tbody>
</table>

| \( N_{\text{jets}} \geq 8 \) uncertainties

<table>
<thead>
<tr>
<th>( H_T ) [GeV]</th>
<th>( \not{H}_T ) [GeV]</th>
<th>Data-driven prediction</th>
<th>stat.</th>
<th>acc.</th>
<th>reco</th>
<th>pixel veto</th>
<th>iso.</th>
<th>data/MC</th>
<th>purity</th>
<th>ratio</th>
<th>pheno. fit</th>
<th>double</th>
</tr>
</thead>
<tbody>
<tr>
<td>500...800</td>
<td>200...</td>
<td>0.00</td>
<td>± 0.58</td>
<td>+0.19</td>
<td>+0.15</td>
<td>+0.23</td>
<td>+0.37</td>
<td>+0.16</td>
<td>+0.18</td>
<td>+0.14</td>
<td>+0.26</td>
<td>+0.41</td>
</tr>
<tr>
<td>800...1000</td>
<td>200...</td>
<td>0.64</td>
<td>± 0.45</td>
<td>+0.04</td>
<td>+0.03</td>
<td>+0.06</td>
<td>+0.17</td>
<td>+0.02</td>
<td>+0.03</td>
<td>+0.03</td>
<td>+0.08</td>
<td>+0.14</td>
</tr>
<tr>
<td>1000...1250</td>
<td>200...</td>
<td>0.60</td>
<td>± 0.43</td>
<td>+0.04</td>
<td>+0.02</td>
<td>+0.05</td>
<td>+0.13</td>
<td>+0.02</td>
<td>+0.03</td>
<td>+0.02</td>
<td>+0.07</td>
<td>+0.14</td>
</tr>
<tr>
<td>1250...1500</td>
<td>200...</td>
<td>0.00</td>
<td>± 0.61</td>
<td>+0.25</td>
<td>+0.22</td>
<td>+0.31</td>
<td>+0.54</td>
<td>+0.15</td>
<td>+0.19</td>
<td>+0.20</td>
<td>+0.33</td>
<td>+0.43</td>
</tr>
<tr>
<td>1500...</td>
<td>200...</td>
<td>0.00</td>
<td>± 0.55</td>
<td>+0.19</td>
<td>+0.13</td>
<td>+0.23</td>
<td>+0.36</td>
<td>+0.19</td>
<td>+0.16</td>
<td>+0.12</td>
<td>+0.26</td>
<td>+0.39</td>
</tr>
</tbody>
</table>

(b) \( 6 \leq N_{\text{jets}} \leq 7 \) and \( N_{\text{jets}} \geq 8 \)

Table 6.12: (Continued) \( Z \rightarrow \nu \bar{\nu} \) + jets uncertainties from \( \gamma + \text{jets} \) data for \( 6 \leq N_{\text{jets}} \leq 7 \) and \( N_{\text{jets}} \geq 8 \) for the 8 TeV analysis (cf. Table A.2).
6.6.1 7 TeV results

Combining all correction factors in a given search bin yields a per-bin scale factor. When this overall scale factor is applied to the measured $\gamma + \text{jets}$ yield in data, as in Eqn. 6.1.1 combined with Eqn. 6.1.2, the prediction of the $Z\rightarrow\nu\bar{\nu} + \text{jets}$ background is obtained. The results are given in Table 6.13. As can be seen, the prediction is consistent with the expectation obtained from the simulation, within the estimated uncertainties on the method.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>500...800</td>
<td>200...350</td>
<td>387.10 ± 4.81</td>
<td>359.33 ± 10.80</td>
<td>+80.34</td>
<td>-80.34</td>
<td>+81.06</td>
<td>-81.06</td>
</tr>
<tr>
<td>500...800</td>
<td>350...500</td>
<td>112.61 ± 2.60</td>
<td>112.43 ± 6.68</td>
<td>+25.47</td>
<td>-25.46</td>
<td>+26.32</td>
<td>-26.32</td>
</tr>
<tr>
<td>500...800</td>
<td>500...600</td>
<td>21.19 ± 1.13</td>
<td>17.61 ± 2.35</td>
<td>+4.29</td>
<td>-4.28</td>
<td>+4.90</td>
<td>-4.89</td>
</tr>
<tr>
<td>500...800</td>
<td>600...</td>
<td>6.05 ± 0.60</td>
<td>5.48 ± 1.42</td>
<td>+2.16</td>
<td>-2.14</td>
<td>+2.59</td>
<td>-2.57</td>
</tr>
<tr>
<td>800...1000</td>
<td>200...350</td>
<td>49.33 ± 1.72</td>
<td>48.36 ± 3.91</td>
<td>+18.05</td>
<td>-18.05</td>
<td>+18.47</td>
<td>-18.47</td>
</tr>
<tr>
<td>800...1000</td>
<td>350...500</td>
<td>15.75 ± 0.97</td>
<td>16.01 ± 2.42</td>
<td>+6.21</td>
<td>-6.19</td>
<td>+6.66</td>
<td>-6.65</td>
</tr>
<tr>
<td>800...1000</td>
<td>500...600</td>
<td>5.45 ± 0.57</td>
<td>7.06 ± 1.97</td>
<td>+3.14</td>
<td>-3.10</td>
<td>+3.70</td>
<td>-3.67</td>
</tr>
<tr>
<td>800...1000</td>
<td>600...</td>
<td>5.33 ± 0.56</td>
<td>3.34 ± 1.01</td>
<td>-3.40</td>
<td>-3.39</td>
<td>+1.71</td>
<td>-1.72</td>
</tr>
<tr>
<td>1000...1200</td>
<td>200...350</td>
<td>15.09 ± 0.95</td>
<td>10.95 ± 1.74</td>
<td>+4.84</td>
<td>-4.83</td>
<td>+5.14</td>
<td>-5.14</td>
</tr>
<tr>
<td>1000...1200</td>
<td>350...500</td>
<td>5.09 ± 0.55</td>
<td>5.45 ± 1.46</td>
<td>+2.64</td>
<td>-2.61</td>
<td>+3.01</td>
<td>-3.00</td>
</tr>
<tr>
<td>1000...1200</td>
<td>500...</td>
<td>3.65 ± 0.47</td>
<td>2.18 ± 1.26</td>
<td>+1.19</td>
<td>-1.18</td>
<td>+1.73</td>
<td>-1.73</td>
</tr>
<tr>
<td>1200...1400</td>
<td>200...350</td>
<td>5.69 ± 0.58</td>
<td>3.11 ± 0.98</td>
<td>+1.46</td>
<td>-1.45</td>
<td>+1.76</td>
<td>-1.76</td>
</tr>
<tr>
<td>1200...1400</td>
<td>350...</td>
<td>3.29 ± 0.44</td>
<td>2.32 ± 0.95</td>
<td>+1.19</td>
<td>-1.17</td>
<td>+1.52</td>
<td>-1.51</td>
</tr>
<tr>
<td>1400...</td>
<td>200...</td>
<td>2.93 ± 0.42</td>
<td>3.34 ± 0.77</td>
<td>+1.61</td>
<td>-1.60</td>
<td>+1.79</td>
<td>-1.78</td>
</tr>
</tbody>
</table>

Table 6.13: $Z\rightarrow\nu\bar{\nu} + \text{jets}$ prediction for 4.98 pb$^{-1}$ of integrated luminosity predicted with the $\gamma + \text{jets}$ method in the 7 TeV exclusive search regions. Columns 3 and 4 show the expectation coming from the simulation, while the rightmost columns show the results of the data-driven prediction.

6.6.2 8 TeV results

The increased difficulty due to adding a binning in $N_{\text{jets}}$ notwithstanding, the overall prediction of the $Z\rightarrow\nu\bar{\nu} + \text{jets}$ background again proved to be robust. Different
from the 7 TeV method, the correction factor was applied as an event weight for the 8 TeV method according to Eqn. 6.6.1, where the weight, $W_j (H_T, \bar{H}_T, N_{jets})$, represents the ratio term from Eqn. 6.1.1 combined with the correction factors (efficiency and purity) from Eqn. 6.1.2, each calculated per event.

$$N_{\text{data } Z \rightarrow \nu \bar{\nu} + \text{jets}}^{\text{bin } i} = \sum_{j \in \text{events}} \sum_{\text{events} \in \text{bin } i} W_j (H_T, \bar{H}_T, N_{jets})$$  \hspace{1cm} (6.6.1)

Using simulated $\gamma + \text{jets}$ events to perform the prediction (middle block of Table 6.14) compared to the prediction obtained using the events in the data control sample (right block of Table 6.14), Table 6.14 shows good overall consistency in each of the 36 search regions. This indicates that there is little to no bias when applying the method on the data events. Shown in Table 6.15, the data-driven prediction in each of the 36 exclusive search bins is compared to the simulated $Z \rightarrow \nu \bar{\nu} + \text{jets}$ yield. The table shows that the prediction is, within the estimated uncertainty, in good agreement with the expected yield from simulation.
Table 6.14: $Z \rightarrow \nu\bar{\nu} + \text{jets}$ predictions from $\gamma + \text{jets}$ simulation compared to the prediction coming from data (19.4 pb$^{-1}$) for $3 \leq N_{\text{jets}} \leq 5$ for the 8 TeV analysis (cf. Table A.3).
<table>
<thead>
<tr>
<th>$6 \leq N_{\text{jets}} \leq 7$</th>
<th>Prediction</th>
<th>uncertainty</th>
<th></th>
<th></th>
<th>$N_{\text{jets}} \geq 8$</th>
<th>Prediction</th>
<th>uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>500...800</td>
<td>200...300</td>
<td>19.30 ±1.49</td>
<td>±1.20</td>
<td>±1.02</td>
<td>±1.58</td>
<td>22.72 ±2.78</td>
<td>±2.72</td>
</tr>
<tr>
<td>500...800</td>
<td>300...450</td>
<td>7.40 ±1.13</td>
<td>±0.55</td>
<td>±0.39</td>
<td>±0.67</td>
<td>9.92 ±1.88</td>
<td>±1.34</td>
</tr>
<tr>
<td>500...800</td>
<td>450...</td>
<td>0.89 ±0.28</td>
<td>±0.11</td>
<td>±0.08</td>
<td>±0.14</td>
<td>0.70 ±0.50</td>
<td>±0.17</td>
</tr>
<tr>
<td>800...1000</td>
<td>200...300</td>
<td>9.80 ±0.91</td>
<td>+0.62</td>
<td>±0.52</td>
<td>±0.81</td>
<td>9.09 ±1.75</td>
<td>±1.16</td>
</tr>
<tr>
<td>800...1000</td>
<td>300...450</td>
<td>4.70 ±0.65</td>
<td>±0.35</td>
<td>±0.25</td>
<td>±0.43</td>
<td>4.23 ±1.22</td>
<td>±0.61</td>
</tr>
<tr>
<td>800...1000</td>
<td>450...</td>
<td>1.24 ±0.33</td>
<td>±0.15</td>
<td>±0.11</td>
<td>±0.18</td>
<td>1.77 ±0.79</td>
<td>±0.39</td>
</tr>
<tr>
<td>1000...1250</td>
<td>200...300</td>
<td>4.36 ±0.61</td>
<td>+0.28</td>
<td>±0.24</td>
<td>±0.37</td>
<td>4.37 ±1.21</td>
<td>±0.55</td>
</tr>
<tr>
<td>1000...1250</td>
<td>300...450</td>
<td>2.31 ±0.45</td>
<td>±0.18</td>
<td>±0.13</td>
<td>±0.22</td>
<td>3.52 ±1.11</td>
<td>±0.50</td>
</tr>
<tr>
<td>1000...1250</td>
<td>450...</td>
<td>0.89 ±0.28</td>
<td>±0.11</td>
<td>±0.08</td>
<td>±0.14</td>
<td>1.41 ±0.70</td>
<td>±0.32</td>
</tr>
<tr>
<td>1250...1500</td>
<td>200...300</td>
<td>1.88 ±0.40</td>
<td>+0.12</td>
<td>±0.11</td>
<td>±0.16</td>
<td>3.33 ±1.05</td>
<td>±0.45</td>
</tr>
<tr>
<td>1250...1500</td>
<td>300...450</td>
<td>1.24 ±0.33</td>
<td>±0.09</td>
<td>±0.06</td>
<td>±0.11</td>
<td>1.42 ±0.71</td>
<td>±0.19</td>
</tr>
<tr>
<td>1250...1500</td>
<td>450...</td>
<td>0.71 ±0.25</td>
<td>±0.09</td>
<td>±0.06</td>
<td>±0.11</td>
<td>0.36 ±0.36</td>
<td>±0.07</td>
</tr>
<tr>
<td>1500...200...300</td>
<td>0.94 ±0.28</td>
<td>±0.06</td>
<td>±0.05</td>
<td>±0.08</td>
<td>1.33 ±0.67</td>
<td>±0.18</td>
<td>±0.28</td>
</tr>
<tr>
<td>1500...300...</td>
<td>1.34 ±0.34</td>
<td>±0.14</td>
<td>±0.10</td>
<td>±0.17</td>
<td>1.06 ±0.61</td>
<td>±0.22</td>
<td>±0.23</td>
</tr>
</tbody>
</table>

Table 6.14: Continued Z→ν$ar{ν}$+jets predictions from γ + jets simulation compared to the prediction coming from data (19.4 pb$^{-1}$) for $6 \leq N_{\text{jets}} \leq 7$ and $N_{\text{jets}} \geq 8$ for the 8 TeV analysis (cf. Table A.3).
<table>
<thead>
<tr>
<th>$3 \leq N_{\text{jets}} \leq 5$</th>
<th>Prediction</th>
<th>uncertainty</th>
<th>$Z\rightarrow\nu\bar{\nu}$</th>
<th>uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>500...800</td>
<td>200...300</td>
<td>1821.28</td>
<td>± 26.31</td>
<td>+86.99</td>
</tr>
<tr>
<td>500...800</td>
<td>300...450</td>
<td>993.58</td>
<td>± 19.78</td>
<td>+47.38</td>
</tr>
<tr>
<td>500...800</td>
<td>450...600</td>
<td>273.21</td>
<td>± 10.31</td>
<td>+16.99</td>
</tr>
<tr>
<td>500...800</td>
<td>600...</td>
<td>42.01</td>
<td>± 4.04</td>
<td>+2.56</td>
</tr>
<tr>
<td>800...1000</td>
<td>200...300</td>
<td>215.84</td>
<td>± 9.04</td>
<td>+10.81</td>
</tr>
<tr>
<td>800...1000</td>
<td>300...450</td>
<td>124.08</td>
<td>± 6.96</td>
<td>+6.77</td>
</tr>
<tr>
<td>800...1000</td>
<td>450...600</td>
<td>46.86</td>
<td>± 4.26</td>
<td>+3.25</td>
</tr>
<tr>
<td>800...1000</td>
<td>600...</td>
<td>35.29</td>
<td>± 3.70</td>
<td>+2.30</td>
</tr>
<tr>
<td>1000...1250</td>
<td>200...300</td>
<td>76.28</td>
<td>± 5.37</td>
<td>+3.87</td>
</tr>
<tr>
<td>1000...1250</td>
<td>300...450</td>
<td>39.34</td>
<td>± 3.92</td>
<td>+2.17</td>
</tr>
<tr>
<td>1000...1250</td>
<td>450...600</td>
<td>18.08</td>
<td>± 2.64</td>
<td>+1.41</td>
</tr>
<tr>
<td>1000...1250</td>
<td>600...</td>
<td>17.81</td>
<td>± 2.63</td>
<td>+1.23</td>
</tr>
<tr>
<td>1250...1500</td>
<td>200...300</td>
<td>25.34</td>
<td>± 3.10</td>
<td>+1.27</td>
</tr>
<tr>
<td>1250...1500</td>
<td>300...450</td>
<td>16.71</td>
<td>± 2.55</td>
<td>+0.95</td>
</tr>
<tr>
<td>1250...1500</td>
<td>450...</td>
<td>12.35</td>
<td>± 2.18</td>
<td>+0.94</td>
</tr>
<tr>
<td>1500...</td>
<td>200...300</td>
<td>10.51</td>
<td>± 1.99</td>
<td>+0.56</td>
</tr>
<tr>
<td>1500...</td>
<td>300...</td>
<td>10.93</td>
<td>± 2.07</td>
<td>+0.60</td>
</tr>
</tbody>
</table>

(a) $3 \leq N_{\text{jets}} \leq 5$

**Table 6.15:** $Z\rightarrow\nu\bar{\nu}$+jets predictions from $\gamma$+jets data (19.4 pb$^{-1}$) for $3 \leq N_{\text{jets}} \leq 5$ for the 8 TeV analysis (cf. Table A.4).
<table>
<thead>
<tr>
<th>$6 \leq N_{\text{jets}} \leq 7$</th>
<th>Prediction</th>
<th>uncertainty</th>
<th>$Z\rightarrow \nu\bar{\nu}$</th>
<th>uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>500...800</td>
<td>200...300</td>
<td>22.72</td>
<td>± 2.78</td>
<td>±2.72</td>
</tr>
<tr>
<td>500...800</td>
<td>300...450</td>
<td>9.92</td>
<td>± 1.88</td>
<td>±1.34</td>
</tr>
<tr>
<td>500...800</td>
<td>450...</td>
<td>0.70</td>
<td>± 0.50</td>
<td>±0.17</td>
</tr>
<tr>
<td>800...1000</td>
<td>200...300</td>
<td>9.09</td>
<td>± 1.75</td>
<td>±1.15</td>
</tr>
<tr>
<td>800...1000</td>
<td>300...450</td>
<td>4.23</td>
<td>± 1.22</td>
<td>±0.61</td>
</tr>
<tr>
<td>800...1000</td>
<td>450...</td>
<td>1.77</td>
<td>± 0.79</td>
<td>±0.39</td>
</tr>
<tr>
<td>1000...1250</td>
<td>200...300</td>
<td>4.37</td>
<td>± 1.21</td>
<td>±0.55</td>
</tr>
<tr>
<td>1000...1250</td>
<td>300...450</td>
<td>3.52</td>
<td>± 1.11</td>
<td>±0.50</td>
</tr>
<tr>
<td>1000...1250</td>
<td>450...</td>
<td>1.41</td>
<td>± 0.70</td>
<td>±0.32</td>
</tr>
<tr>
<td>1250...1500</td>
<td>200...300</td>
<td>3.33</td>
<td>± 1.05</td>
<td>±0.45</td>
</tr>
<tr>
<td>1250...1500</td>
<td>300...450</td>
<td>1.42</td>
<td>± 0.71</td>
<td>±0.19</td>
</tr>
<tr>
<td>1250...1500</td>
<td>450...</td>
<td>0.36</td>
<td>± 0.36</td>
<td>±0.07</td>
</tr>
<tr>
<td>1500...</td>
<td>200...300</td>
<td>1.33</td>
<td>± 0.67</td>
<td>±0.18</td>
</tr>
<tr>
<td>1500...</td>
<td>300...</td>
<td>1.06</td>
<td>± 0.61</td>
<td>±0.22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N_{\text{jets}} \geq 8$</th>
<th>Prediction</th>
<th>uncertainty</th>
<th>$Z\rightarrow \nu\bar{\nu}$</th>
<th>uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>500...800</td>
<td>200...</td>
<td>0.00</td>
<td>± 0.58</td>
<td>±0.41</td>
</tr>
<tr>
<td>800...1000</td>
<td>200...</td>
<td>0.64</td>
<td>± 0.45</td>
<td>±0.19</td>
</tr>
<tr>
<td>1000...1250</td>
<td>200...</td>
<td>0.60</td>
<td>± 0.43</td>
<td>±0.15</td>
</tr>
<tr>
<td>1250...1500</td>
<td>200...</td>
<td>0.00</td>
<td>± 0.61</td>
<td>±0.58</td>
</tr>
<tr>
<td>1500...</td>
<td>200...</td>
<td>0.00</td>
<td>± 0.55</td>
<td>±0.39</td>
</tr>
</tbody>
</table>

(b) $6 \leq N_{\text{jets}} \leq 7$ and $N_{\text{jets}} \geq 8$

Table 6.15: (Continued) $Z\rightarrow \nu\bar{\nu}$+jets predictions from $\gamma$+jets data (19.4 pb$^{-1}$) for $6 \leq N_{\text{jets}} \leq 7$ and $N_{\text{jets}} \geq 8$ for the 8 TeV analysis (cf. Table A.4).
Chapter 7

Results and interpretation

7.1 Combined background estimates

The separate background estimates for the $Z \rightarrow \nu \bar{\nu} + \text{jets}$, QCD and $t\bar{t} + \text{jets}/W^{\pm} + \text{jets}$ events are combined and presented below in both tabular and graphical form.

7.1.1 7 TeV analysis

The numerical results for each search bin are shown in Table 7.1. The total background uncertainty includes the individual uncertainties added under the assumption that the statistical uncertainties are completely uncorrelated, while the systematic uncertainties are fully correlated. Figure 7.1 shows the event search variables $H_T$ and $H_T^*$, as well as the yield in each exclusive search bin used in making predictions for the 7 TeV analysis. The event variables are compared to a specific point in the CMSSM parameter space, labeled LM5, corresponding to $M_0 = 230 \text{GeV}, M_{1/2} = 360 \text{GeV}, A_0 = 0, \tan \beta = 10, \mu > 0$ [64]. It can be seen from the last two columns in Table 7.1, as well as the figures in Fig. 7.1,
that the data are consistent with the data-driven standard model background prediction to
within the uncertainty on the prediction.

### 7.1.2 8 TeV analysis

Figure 7.2 shows the event count for each of the 36 exclusive search bins used in
the 8 TeV analysis and the estimated background components in each bin. Again, no excess
is observed in data.

### 7.2 Interpretation of results

No observed excess with respect to the estimated standard model backgrounds
has been observed. As a result, we turn our attention to setting limits on possible models
of new physics. Several common models have been investigated. The constrained minimal
supersymmetric standard model (CMSSM) is a parameterization of SUSY with five free
parameters. Also considered are various simplified model spectra (SMS) topologies, which
are constructed based on a simple event topology and can then be combined to describe a
complete theory. Each of the topologies represents some specific process and a limit can be
set on the cross section for each process as a function of the mass of the particles involved.

Limits are set using the $CL_s$ test statistic, defined as the ratio of confidence levels
from a signal plus background hypothesis to a background-only hypothesis (Eqn. 7.2.1). The
method was developed for Higgs searches at the LEP experiments [65, 66]. The confidence
level is defined in terms of the hypothesis, $H_a$, as the probability that a test statistic value
### Table 7.1: Predicted background yields obtained from the different data-driven estimation methods for the 14 search bins, defined in $H_T$ and $\mathbb{H}_T$, using the full 2011 7 TeV dataset. The total is the sum, with statistical uncertainties treated as uncorrelated and systematic uncertainties added assuming full correlation (from Ref. [16]).

<table>
<thead>
<tr>
<th>Selection</th>
<th>$H_T$</th>
<th>$\mathbb{H}_T$</th>
<th>$Z\rightarrow\nu\bar{\nu}$</th>
<th>$t\bar{t}/W^\pm$</th>
<th>$t\bar{t}/W^\pm$</th>
<th>QCD</th>
<th>Total</th>
<th>Observed in data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>from $\gamma + \text{jets}$</td>
<td>$\rightarrow e^\pm \mu^\pm + X$</td>
<td>$\rightarrow \tau^\pm_{\text{hadr}} + X$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500...800</td>
<td>200...350</td>
<td>359.3 ± 81.1</td>
<td>326.5 ± 47.0</td>
<td>348.5 ± 40.1</td>
<td>118.6 ± 76.9</td>
<td>1152.9 ± 127.7</td>
<td>1269</td>
<td></td>
</tr>
<tr>
<td>500...800</td>
<td>350...500</td>
<td>112.4 ± 26.3</td>
<td>47.8 ± 9.2</td>
<td>62.5 ± 8.7</td>
<td>2.2 ± 2.2</td>
<td>224.9 ± 29.3</td>
<td>236</td>
<td></td>
</tr>
<tr>
<td>500...800</td>
<td>500...600</td>
<td>17.6 ± 4.9</td>
<td>5.0 ± 2.2</td>
<td>8.7 ± 2.5</td>
<td>0.0 ± 0.1</td>
<td>31.3 ± 5.9</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>500...800</td>
<td>600...</td>
<td>5.5 ± 2.6</td>
<td>0.8 ± 0.8</td>
<td>2.0 ± 1.8</td>
<td>0.0 ± 0.0</td>
<td>8.3 ± 3.3</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>800...1000</td>
<td>200...350</td>
<td>48.4 ± 18.5</td>
<td>57.7 ± 15.3</td>
<td>56.3 ± 8.3</td>
<td>34.6 ± 24.0</td>
<td>197.0 ± 34.9</td>
<td>177</td>
<td></td>
</tr>
<tr>
<td>800...1000</td>
<td>350...500</td>
<td>16.0 ± 6.7</td>
<td>5.4 ± 2.3</td>
<td>7.2 ± 2.0</td>
<td>1.2 ± 1.3</td>
<td>29.8 ± 7.5</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>800...1000</td>
<td>500...600</td>
<td>7.1 ± 3.7</td>
<td>2.4 ± 1.5</td>
<td>1.3 ± 0.6</td>
<td>0.0 ± 0.2</td>
<td>10.8 ± 4.0</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>800...1000</td>
<td>600...</td>
<td>3.3 ± 1.7</td>
<td>0.7 ± 0.7</td>
<td>1.0 ± 0.3</td>
<td>0.0 ± 0.1</td>
<td>5.0 ± 1.9</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>1000...1200</td>
<td>200...350</td>
<td>11.0 ± 5.1</td>
<td>13.7 ± 3.8</td>
<td>21.9 ± 4.6</td>
<td>19.7 ± 13.3</td>
<td>66.3 ± 15.4</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>1000...1200</td>
<td>350...500</td>
<td>5.5 ± 3.0</td>
<td>5.0 ± 4.4</td>
<td>2.9 ± 1.3</td>
<td>0.4 ± 0.7</td>
<td>13.8 ± 5.5</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>1000...1200</td>
<td>500...</td>
<td>2.2 ± 1.7</td>
<td>1.6 ± 1.2</td>
<td>2.3 ± 1.0</td>
<td>0.0 ± 0.2</td>
<td>6.1 ± 2.3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>1200...1400</td>
<td>200...350</td>
<td>3.1 ± 1.8</td>
<td>4.2 ± 2.1</td>
<td>6.2 ± 1.8</td>
<td>11.7 ± 8.3</td>
<td>25.2 ± 8.9</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>1200...1400</td>
<td>350...</td>
<td>2.3 ± 1.5</td>
<td>2.3 ± 1.4</td>
<td>0.6 ± 0.8</td>
<td>0.2 ± 0.6</td>
<td>5.4 ± 2.3</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>1400...</td>
<td>200...</td>
<td>3.3 ± 1.8</td>
<td>2.7 ± 1.6</td>
<td>1.1 ± 0.5</td>
<td>12.0 ± 9.1</td>
<td>19.1 ± 9.4</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>597.0 ± 160.4</td>
<td>475.8 ± 76.2</td>
<td>522.4 ± 67.0</td>
<td>200.7 ± 82.7</td>
<td>1795.9 ± 258.4</td>
<td>1885</td>
<td></td>
</tr>
</tbody>
</table>
Figure 7.1: Data vs. data-driven background estimation for search variables and search bins, showing as a comparison, the LM5 ($m_0 = 230$ GeV, $m_{1/2} = 360$ GeV) point from the CMSSM parameter space with $A_0 = 0, \tan \beta = 10, \mu > 0$ for the 7 TeV analysis (from Ref. [16]).
<table>
<thead>
<tr>
<th>bin idx</th>
<th>$N_{\text{jets}}$</th>
<th>Selection</th>
<th>$Z \rightarrow \tau \tau$ from $\gamma + \text{jets}$</th>
<th>$t \bar{t}/W^\pm$ $\rightarrow e\mu^\pm + X$</th>
<th>$t \bar{t}/W^\pm$ $\rightarrow \tau^\pm_{\text{hadr}} + X$</th>
<th>QCD</th>
<th>Total background</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3-5</td>
<td>500...800 200...300</td>
<td>1823.2 ± 326.8</td>
<td>2210.7 ± 500.8</td>
<td>1683.7 ± 171.4</td>
<td>307.4 ± 219.4</td>
<td>6024.9 ± 659.7</td>
<td>6159</td>
</tr>
<tr>
<td>2</td>
<td>3-5</td>
<td>500...800 300...450</td>
<td>994.6 ± 178.1</td>
<td>660.1 ± 154.0</td>
<td>591.9 ± 62.5</td>
<td>34.5 ± 23.8</td>
<td>2281.0 ± 244.7</td>
<td>2305</td>
</tr>
<tr>
<td>3</td>
<td>3-5</td>
<td>500...800 450...600</td>
<td>273.5 ± 51.1</td>
<td>77.3 ± 20.5</td>
<td>67.6 ± 9.5</td>
<td>1.3 ± 1.5</td>
<td>419.7 ± 55.9</td>
<td>454</td>
</tr>
<tr>
<td>4</td>
<td>3-5</td>
<td>500...800 600...</td>
<td>42.1 ± 8.7</td>
<td>9.5 ± 4.1</td>
<td>6.0 ± 1.9</td>
<td>0.1 ± 0.3</td>
<td>57.6 ± 9.8</td>
<td>62</td>
</tr>
<tr>
<td>5</td>
<td>3-5</td>
<td>800...1000 200...300</td>
<td>216.1 ± 40.1</td>
<td>277.5 ± 67.5</td>
<td>191.6 ± 23.2</td>
<td>91.7 ± 65.5</td>
<td>776.9 ± 104.8</td>
<td>808</td>
</tr>
<tr>
<td>6</td>
<td>3-5</td>
<td>800...1000 300...450</td>
<td>124.2 ± 23.8</td>
<td>112.8 ± 29.2</td>
<td>83.3 ± 11.2</td>
<td>9.9 ± 7.4</td>
<td>330.2 ± 39.9</td>
<td>305</td>
</tr>
<tr>
<td>7</td>
<td>3-5</td>
<td>800...1000 450...600</td>
<td>46.9 ± 9.8</td>
<td>36.1 ± 11.0</td>
<td>23.6 ± 3.9</td>
<td>0.8 ± 1.3</td>
<td>107.5 ± 15.3</td>
<td>124</td>
</tr>
<tr>
<td>8</td>
<td>3-5</td>
<td>800...1000 600...</td>
<td>35.3 ± 7.5</td>
<td>9.0 ± 3.8</td>
<td>11.4 ± 3.2</td>
<td>0.1 ± 0.4</td>
<td>55.8 ± 9.0</td>
<td>52</td>
</tr>
<tr>
<td>9</td>
<td>3-5</td>
<td>1000...1250 200...300</td>
<td>76.4 ± 14.9</td>
<td>103.5 ± 27.9</td>
<td>66.8 ± 10.0</td>
<td>59.0 ± 24.7</td>
<td>305.7 ± 41.3</td>
<td>335</td>
</tr>
<tr>
<td>10</td>
<td>3-5</td>
<td>1000...1250 300...450</td>
<td>39.4 ± 8.2</td>
<td>52.4 ± 14.7</td>
<td>35.7 ± 6.2</td>
<td>5.1 ± 2.7</td>
<td>132.6 ± 18.1</td>
<td>129</td>
</tr>
<tr>
<td>11</td>
<td>3-5</td>
<td>1000...1250 450...600</td>
<td>18.1 ± 4.4</td>
<td>6.9 ± 3.3</td>
<td>6.6 ± 2.1</td>
<td>0.5 ± 0.7</td>
<td>32.2 ± 5.9</td>
<td>34</td>
</tr>
<tr>
<td>12</td>
<td>3-5</td>
<td>1000...1250 600...</td>
<td>17.8 ± 4.3</td>
<td>2.4 ± 1.8</td>
<td>2.5 ± 1.0</td>
<td>0.1 ± 0.3</td>
<td>22.8 ± 4.7</td>
<td>32</td>
</tr>
<tr>
<td>13</td>
<td>3-5</td>
<td>1250...1500 200...300</td>
<td>25.4 ± 5.5</td>
<td>31.0 ± 9.8</td>
<td>22.2 ± 3.9</td>
<td>31.2 ± 13.1</td>
<td>109.8 ± 17.7</td>
<td>98</td>
</tr>
<tr>
<td>14</td>
<td>3-5</td>
<td>1250...1500 300...450</td>
<td>16.7 ± 4.0</td>
<td>10.1 ± 4.5</td>
<td>11.1 ± 3.6</td>
<td>2.3 ± 1.6</td>
<td>40.2 ± 7.2</td>
<td>38</td>
</tr>
<tr>
<td>15</td>
<td>3-5</td>
<td>1250...1500 450...</td>
<td>12.4 ± 3.2</td>
<td>2.3 ± 1.7</td>
<td>2.8 ± 1.5</td>
<td>0.2 ± 0.5</td>
<td>17.6 ± 4.0</td>
<td>23</td>
</tr>
<tr>
<td>16</td>
<td>3-5</td>
<td>1500...200...300</td>
<td>10.5 ± 2.8</td>
<td>16.7 ± 6.2</td>
<td>15.2 ± 3.4</td>
<td>35.1 ± 14.1</td>
<td>77.6 ± 16.0</td>
<td>94</td>
</tr>
<tr>
<td>17</td>
<td>3-5</td>
<td>1500...300...</td>
<td>10.9 ± 2.9</td>
<td>9.7 ± 4.4</td>
<td>6.5 ± 2.0</td>
<td>2.4 ± 2.0</td>
<td>29.6 ± 5.9</td>
<td>39</td>
</tr>
</tbody>
</table>

(a) $3 \leq N_{\text{jets}} \leq 5$

Table 7.2: Predicted background yields obtained from the different data-driven estimation methods for the 36 search bins, defined in $H_T$, $H_T$ and $3 \leq N_{\text{jets}} \leq 5$, using the full 8 TeV dataset. The uncertainties on the different backgrounds are added in quadrature to obtain the total uncertainty on the background estimation (from Ref. [17]).
<table>
<thead>
<tr>
<th>bin idx</th>
<th>$N_{\text{jets}}$</th>
<th>Selection $\mu \bar{\nu}$ from $\gamma + \text{jets}$</th>
<th>$t\bar{t}/W^\pm \rightarrow e\nu\mu\bar{\nu} + X$</th>
<th>$t\bar{t}/W^\pm \rightarrow \tau^\pm_{\text{had}} + X$</th>
<th>QCD</th>
<th>Total background</th>
<th>Obs. data</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>6-7</td>
<td>500...800 200...300</td>
<td>22.7 ± 6.1</td>
<td>132.5 ± 58.2</td>
<td>127.1 ± 21.5</td>
<td>18.2 ± 9.2</td>
<td>300.5 ± 63.0</td>
</tr>
<tr>
<td>19</td>
<td>6-7</td>
<td>500...800 300...450</td>
<td>9.9 ± 3.1</td>
<td>22.0 ± 10.6</td>
<td>18.6 ± 4.3</td>
<td>1.9 ± 1.7</td>
<td>52.4 ± 12.0</td>
</tr>
<tr>
<td>20</td>
<td>6-7</td>
<td>500...800 450...600</td>
<td>0.7 ± 0.6</td>
<td>0.0 ± 1.6</td>
<td>0.1 ± 0.3</td>
<td>0.0 ± 0.1</td>
<td>0.8 ± 1.7</td>
</tr>
<tr>
<td>21</td>
<td>6-7</td>
<td>800...1000 200...300</td>
<td>9.1 ± 2.8</td>
<td>55.8 ± 25.3</td>
<td>44.6 ± 8.2</td>
<td>13.1 ± 6.6</td>
<td>122.6 ± 27.5</td>
</tr>
<tr>
<td>22</td>
<td>6-7</td>
<td>800...1000 300...450</td>
<td>4.2 ± 1.6</td>
<td>10.4 ± 5.5</td>
<td>12.8 ± 3.1</td>
<td>1.9 ± 1.4</td>
<td>29.3 ± 6.7</td>
</tr>
<tr>
<td>23</td>
<td>6-7</td>
<td>800...1000 450...600</td>
<td>1.8 ± 1.0</td>
<td>2.9 ± 2.5</td>
<td>1.3 ± 0.5</td>
<td>0.1 ± 0.4</td>
<td>6.1 ± 2.7</td>
</tr>
<tr>
<td>24</td>
<td>6-7</td>
<td>1000...1250 200...300</td>
<td>4.4 ± 1.6</td>
<td>24.1 ± 11.9</td>
<td>24.0 ± 5.5</td>
<td>11.9 ± 6.0</td>
<td>64.4 ± 14.5</td>
</tr>
<tr>
<td>25</td>
<td>6-7</td>
<td>1000...1250 300...450</td>
<td>3.5 ± 1.4</td>
<td>8.0 ± 4.8</td>
<td>9.6 ± 2.5</td>
<td>1.5 ± 1.5</td>
<td>22.6 ± 5.8</td>
</tr>
<tr>
<td>26</td>
<td>6-7</td>
<td>1000...1250 450...600</td>
<td>1.4 ± 0.8</td>
<td>0.0 ± 1.8</td>
<td>0.8 ± 0.5</td>
<td>0.1 ± 0.3</td>
<td>2.3 ± 2.1</td>
</tr>
<tr>
<td>27</td>
<td>6-7</td>
<td>1250...1500 200...300</td>
<td>3.3 ± 1.3</td>
<td>11.5 ± 6.7</td>
<td>6.1 ± 2.5</td>
<td>6.8 ± 3.9</td>
<td>27.7 ± 8.2</td>
</tr>
<tr>
<td>28</td>
<td>6-7</td>
<td>1250...1500 300...450</td>
<td>1.4 ± 0.8</td>
<td>3.5 ± 2.7</td>
<td>2.9 ± 1.5</td>
<td>0.9 ± 1.3</td>
<td>8.8 ± 3.4</td>
</tr>
<tr>
<td>29</td>
<td>6-7</td>
<td>1250...1500 450...600</td>
<td>0.4 ± 0.4</td>
<td>0.0 ± 1.2</td>
<td>0.1 ± 0.2</td>
<td>0.1 ± 0.3</td>
<td>0.5 ± 1.3</td>
</tr>
<tr>
<td>30</td>
<td>6-7</td>
<td>1500...2000 200...300</td>
<td>1.3 ± 0.8</td>
<td>10.0 ± 6.8</td>
<td>2.3 ± 1.3</td>
<td>7.8 ± 4.0</td>
<td>21.5 ± 8.0</td>
</tr>
<tr>
<td>31</td>
<td>6-7</td>
<td>1500...2000 300...450</td>
<td>1.1 ± 0.7</td>
<td>3.2 ± 2.8</td>
<td>2.9 ± 1.2</td>
<td>0.8 ± 1.1</td>
<td>8.0 ± 3.3</td>
</tr>
</tbody>
</table>

Table 7.2: (Continued) Predicted background yields obtained from the different data-driven estimation methods for the 36 search bins, defined in $H_T, \tilde{H}_T$ and $6 \leq N_{\text{jets}} \leq 7$ and $N_{\text{jets}} \geq 8$, using the full 8 TeV dataset. The uncertainties on the different backgrounds are added in quadrature to obtain the total uncertainty on the background estimation (from Ref. [17]).
Figure 7.2: Data vs. data-driven background estimation comparison with associated uncertainties for the 8 TeV analysis (from Ref. [17]).
(X) is less than or equal to the observed value \(X_{\text{observed}}\) (Eqn. 7.2.2).

\[
C_{L_{s}} = \frac{CL_{s+b}}{CL_{b}} \quad (7.2.1)
\]

\[
CL_{H_{a}} = P_{H_{a}}(X < X_{\text{observed}}) \quad (7.2.2)
\]

The exclusive search regions for both the 7 TeV analysis and the 8 TeV analysis are statistically independent samples and as such, the product of the individual likelihoods (Eqn. 7.2.3) can serve as the test statistic. For an estimated signal contribution \(s_{i}\), background contribution \(b_{i}\), and observed yield \(o_{i}\), the likelihood ratio for a given bin is expressed as in Eqn. 7.2.4.

\[
L = \prod_{i \in \text{bins}} L_{i} \quad (7.2.3)
\]

\[
L_{i} = \left( \frac{e^{-(s_{i}+b_{i})}(s_{i}+b_{i})^{o_{i}}}{o_{i}!} \right) / \left( \frac{e^{-b_{i}}b_{i}^{o_{i}}}{o_{i}!} \right) \quad (7.2.4)
\]

The limit is set using \(CL_{s} = 95\%\) as the cutoff. The test statistic is scanned across the whole available parameter space of the model in question. A curve can be drawn to show the excluded region, i.e., the points where the signal hypothesis can be rejected. The exclusion curve shows the minimum parameter values that could not be ruled out based on the observed data yield and measured background estimates.

7.3 Interpretation within the CMSSM

The CMSSM is a simplification of a general Susy theory with five free parameters (rather than the usual 108). The parameters are the universal gaugino mass \(M_{1/2}\), the universal scalar mass \(M_{0}\), the tri-linear coupling \(A_{0}\), the sign of the mass parameter \(\mu\), and the tangent of the mass breaking term \(\tan \beta\). In order to set limits that can be compared to other experiments, specific values of the five parameters are chosen. CMS has chosen to scan a parameter space in the \(M_{0}\cdot M_{1/2}\) plane, at fixed values of \(A_{0} = 0\), \(\tan \beta = 10\),
and $\mu > 0$. Results are then shown as an exclusion contour in this two-dimensional plane, showing the lower allowed limit for the two masses in order for the particular CMSSM model to remain valid. Limits were set for the 7 TeV analysis but not for the 8 TeV analysis because the CMSSM was deemed to be extremely unlikely based on the 7 TeV results.

For the CMSSM interpretation, the $M_0$-$M_{1/2}$ plane is divided into 20 GeV intervals for both masses, with $0 < M_{1/2}$[GeV] $< 1000$ and $0 < M_0$[GeV] $< 3000$, and with 10000 events generated at each signal point. The results are shown in Fig. 7.3 in terms of $M_0$ and $M_{1/2}$ and in terms of the physical masses, $M_{\tilde{g}}$ and $M_{\tilde{q}}$.

7.4 Interpretation with simplified models

7.4.1 Introduction to simplified models

Simplified model spectra (SMSs) [67, 68, 69] are specific processes in a supersymmetric model that are useful to place limits on the cross section of a process and the masses of the participating new physics particles in a model independent way. Due to the simplified nature of considering only a single process, rather than scanning over a large parameter space such as the CMSSM, one can simply scan over mass values of a few particles, and characterize the behavior of some sector of potential new physics. Additionally, these model topologies allow one to probe regions of parameter space not easily accessible to full SUSY models, and this can help to narrow down the possible specific global model.
Figure 7.3: Exclusion contours in the $M_0$-$M_{1/2}$ and $M_{\tilde{g}}$-$M_{\tilde{q}}$ planes (from Ref. [16]).
The T1 class of SMS (Fig. 7.4a) is characterized by strong gluino production (pp→˜g ˜g) with the gluino subsequently decaying to a quark, antiquark, and LSP, ˜g →q ˜χ_0^0. Models of this type are useful to set limits on the gluino mass and the mass of the lightest neutralino.
7.4.1.2 T2qq (̃q pair production)

The T2 class of SMS (Fig. 7.4b) is characterized by strong squark production (pp \rightarrow ̃q^* ̃q) with the squark subsequently decaying to a quark (antiquark) and the LSP, ̃q \rightarrow q \tilde{\chi}_1^0. Models of this type are useful to set limits on the squark mass and the mass of the lightest neutralino.

7.4.2 Exclusion curves

Similarly to the CMSSM case, limits are set by scanning over the two-dimensional plane of the two masses under consideration. The masses are scanned in 25 or 50 GeV intervals, depending on the model and the mass value, and each point has 10000 generated events. Figures 7.5(a) and (b) show the results for gluino pair production. For the 8 TeV analysis, shown in Fig. 7.5b, gluino masses up to 1.2 TeV are excluded for very low neutralino mass, while neutralino masses of up to 500 GeV are excluded for gluino masses \sim 1.1 TeV. The corresponding results for squark pair production are shown in Figs. 7.5c and (d). For the 8 TeV analysis, squark masses of up to 900 GeV have been excluded for very low neutralino mass, while neutralino masses of up to 350 GeV have been excluded for squark masses \sim 700 GeV.
(a) 7 TeV T1qqqq (gluino pair production) exclusion in the $M_{\tilde{g}}$-$M_{\tilde{g}^0}$ plane

(b) 8 TeV T1qqqq (gluino pair production) exclusion in the $M_{\tilde{g}}$-$M_{\tilde{g}^0}$ plane

(c) 7 TeV T2qq (squark pair production) exclusion in the $M_{\tilde{q}}$-$M_{\tilde{q}^0}$ plane

(d) 8 TeV T2qq (squark pair production) exclusion in the $M_{\tilde{q}}$-$M_{\tilde{q}^0}$ plane

Figure 7.5: Exclusion contours for the T1qqqq and T2qq SMS topologies, showing the improvement between the 7 TeV analysis and the 8 TeV analysis (from Refs. [16] and [17]).
Conclusions

The tremendous effort put forth by all collaborators and accelerator staff to commission and operate this fantastic machine cannot be overstated. During the spectacular run of the LHC for the first three years of operation, the collision data provided have been analyzed by many groups and exceptional results have been produced. The collision environment provided by the LHC is an ideal place to search for evidence of new physics phenomena, such as SUSY.

As presented in the previous chapter, no excess, with respect to the expectation of the standard model, has been seen in the data analyzed at the LHC, under some assumption of SUSY. As a result of this, limits have been placed on full SUSY models as well as on generic topologies, which are components of complete models. The limits on the mass values increased between the 7 TeV analysis and the 8 TeV analysis, as expected due to the increased energy and size of the data sample. The 7 TeV analysis excluded CMSSM models with low $M_0$ and $M_{1/2} < 560$ GeV, and with $M_0 > 1.5$ TeV and $M_{1/2} < 260$ GeV. In terms of the physically observable states, the squark ($\tilde{q}$) mass was excluded below $m_{\tilde{q}} \approx 1.1$ TeV and the gluino ($\tilde{g}$) was excluded below $m_{\tilde{g}} \approx 780$ GeV. For lower squark mass, the gluino was excluded below $m_{\tilde{g}} \approx 1.2$ TeV. For several simplified model topologies, the 8 TeV analysis
improved the exclusion of the 7 TeV analysis and excluded a neutralino mass of $m_{\tilde{\chi}^0_1} < 350$ GeV for models with squark pair production with squark masses up to $m_{\tilde{q}} \approx 700$ GeV. For gluino pair production, neutralino masses have been excluded up to $m_{\tilde{\chi}^0_1} < 500$ GeV for gluino mass up to $m_{\tilde{g}} \approx 1.1$ TeV.

An important method used to estimate the irreducible $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background was developed and extended and has been shown to be reliable. The added difficulty in binning the search region vs. $N_{\text{jets}}$ was managed and the uncertainty was constrained by several data-driven methods. This thesis shows that obtaining a data-driven estimate of the $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background to hadronic SUSY searches is possible, even in the extreme kinematic regions still allowed by SUSY models.
Part IV

Appendix, Glossary, Acronyms, Bibliography
**Glossary**

**ASIC** An application specific integrated circuit is a dedicated circuit that performs a specific task or calculation. 38

**bootstrap** A statistical method to draw from a seed sample multiple times and thus increase the statistical significance of a calculation. 60

**boson** A boson is a particle that obeys Bose statistics, also has whole integer spin. 3

**CERN** Organisation Européenne pour la Recherche Nucléaire. 17, 20

**pseudorapidity** Pseudorapidity is the lorentz invariant measure of the forward angle. 24, 31

**FastJet** FastJet is the name of a jet finding algorithm designed to quickly cluster energy deposits into a jet. 48, 49

**fermion** A fermion is a particle that obeys Fermi statistics, also has half integer spin, e.g., quark or lepton. 3

**FPGA** A field programmable gate array is a programmable logic device, which can be programmed to perform a variety of operations. 38

**gluon** Gluons are the force carrier for the strong nuclear force, also known as the color force. They come in 8 colors. 3

**jet** Jets are the signature of QCD interactions in the detector. As quarks and gluons separate from each other, the color lines connecting them gain energy until they break into a particle antiparticle pair, a process known as hadronization. 47

**λ** The nuclear interaction length is the distance in a material that a particle will travel before having its energy reduced by a factor of $1/e$. 31

**lepton** Leptons are one of the fundamental particle groups in the standard model. They come in three families and have two members per family. The families are electron, muon, and tau, with the second member of each family the corresponding neutrino. 3

**parton** A parton is the initial particle coming out of the hard scattering event that forms a jet in the final detector state. Typically a quark or gluon. 41

**quark** Quarks are one of the fundamental groups of fermions in the standard model. They come in three families with two members per family. The groups are $(u,d)$, $(s,c)$, and $(t,b)$. 3

**RF** radio-frequency. 20
Tag-and-probe is a method to estimate efficiencies in data using well understood processes, e.g., $Z \rightarrow ll$. A well-identified reconstructed object is selected as a “tag” and another object is selected if it passes a much looser selection (“probe”). The composite object is then created from the tag-and-probe pair and the efficiency is calculated using the fit to an invariant mass distribution using the probes that passed the tighter isolation vs. all selected probes. xiii, 62, 80, 83, 84, 97
## Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ECAL</strong></td>
<td>electromagnetic calorimeter. 26, 28, 30, 31, 38, 46</td>
</tr>
<tr>
<td><strong>HCal</strong></td>
<td>hadronic calorimeter. 30, 38, 127</td>
</tr>
<tr>
<td><strong>APD</strong></td>
<td>avalanche photo diode. 28</td>
</tr>
<tr>
<td><strong>ATLAS</strong></td>
<td>A Toroidal LHC ApparatuS. 20</td>
</tr>
<tr>
<td><strong>CMS</strong></td>
<td>Compact Muon Solenoid. xii, 6, 20–22, 24, 34–36, 38, 55, 127, 167</td>
</tr>
<tr>
<td><strong>CMSSM</strong></td>
<td>constrained minimal supersymmetric standard model. 110, 116</td>
</tr>
<tr>
<td><strong>CSC</strong></td>
<td>cathode strip chamber. 34, 38</td>
</tr>
<tr>
<td><strong>DAQ</strong></td>
<td>data acquisition system. 30, 33, 38, 39, 167</td>
</tr>
<tr>
<td><strong>DQM</strong></td>
<td>data quality monitoring. 165, 167, 168</td>
</tr>
<tr>
<td><strong>DT</strong></td>
<td>drift tube. 34, 38</td>
</tr>
<tr>
<td><strong>EB</strong></td>
<td>electromagnetic barrel calorimeter. 31</td>
</tr>
<tr>
<td><strong>EW</strong></td>
<td>electroweak. 8</td>
</tr>
<tr>
<td><strong>EWSB</strong></td>
<td>electroweak symmetry breaking. 9, 10</td>
</tr>
<tr>
<td><strong>GCT</strong></td>
<td>global calorimeter trigger. 38</td>
</tr>
<tr>
<td><strong>GMT</strong></td>
<td>global muon trigger. 38</td>
</tr>
<tr>
<td><strong>gsf</strong></td>
<td>Gaussian-sum filter. 45</td>
</tr>
<tr>
<td><strong>GT</strong></td>
<td>global trigger. 39</td>
</tr>
<tr>
<td><strong>HB</strong></td>
<td>barrel hadronic calorimeter. 31</td>
</tr>
<tr>
<td><strong>HE</strong></td>
<td>endcap hadronic calorimeter. 31</td>
</tr>
<tr>
<td><strong>HF</strong></td>
<td>forward hadronic calorimeter. 32</td>
</tr>
<tr>
<td><strong>HLT</strong></td>
<td>high-level trigger. ix, 39</td>
</tr>
<tr>
<td><strong>HO</strong></td>
<td>outer hadronic calorimeter. 31</td>
</tr>
<tr>
<td><strong>HPD</strong></td>
<td>hybrid photo diode. 33</td>
</tr>
<tr>
<td><strong>HTR</strong></td>
<td>HCal trigger and readout card. 30, 33</td>
</tr>
<tr>
<td><strong>IP</strong></td>
<td>interaction point. 17, 19, 20, 23, 47</td>
</tr>
<tr>
<td><strong>IP5</strong></td>
<td>interaction point 5. 20</td>
</tr>
<tr>
<td><strong>JEC</strong></td>
<td>jet energy correction. 51</td>
</tr>
<tr>
<td><strong>L1</strong></td>
<td>level-1 trigger. ix, 38</td>
</tr>
<tr>
<td><strong>L1A</strong></td>
<td>level-1 accept. 30, 33, 38</td>
</tr>
<tr>
<td><strong>LEP</strong></td>
<td>Large Electron Positron. 17</td>
</tr>
<tr>
<td><strong>LHC</strong></td>
<td>Large Hadron Collider. xii, 17–21, 37, 39, 51, 54</td>
</tr>
<tr>
<td><strong>LSP</strong></td>
<td>lightest supersymmetric particle. vii, 12, 13, 55, 58, 119, 120</td>
</tr>
<tr>
<td><strong>MIP</strong></td>
<td>minimum ionizing particle. 44</td>
</tr>
<tr>
<td><strong>ODU</strong></td>
<td>optical decoder unit. 33</td>
</tr>
<tr>
<td><strong>P5</strong></td>
<td>Point 5 (location of CMS). 167</td>
</tr>
</tbody>
</table>
PDF parton distribution function. 6
PF particle-flow. 42, 43
PMT photomultiplier tube. 32, 33
PS Proton Synchrotron. 20
QCD quantum chromodynamics. 8
RBX read out box. 31
RCT regional calorimeter trigger. 30, 38
RPC resistive plate chamber. 34, 38
RPV R-parity violating. 13
SM standard model of particle physics. 2, 5
SMS simplified model spectra. 110, 117, 119, 120
SPS Super Proton Synchrotron. 20
Susy supersymmetry. vii, 5, 6, 54, 116
T&P tag-and-probe. xiii, 62, 80, 83, 84, 97, 126
TP trigger primitive. 28, 38
vev vacuum expectation value. 11
VPT vacuum photo triode. 28
WLS wavelength shifting fiber. 31, 33
Bibliography


[53] CMS SUSY RA2 Group, “Inclusive search for new physics at CMS with the jets and missing momentum signature”, (2012). AN-11-247. 60


[55] CMS SUSY RA2 Group, “Inclusive search for new physics at CMS with the jets and missing momentum signature”, (2012). AN-11-398. 60, 64, 76, 77, 144, 145, 146, 147


133

[70] CMS Collaboration. See
https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFilters. 139

[71] CMS Collaboration. See
https://twiki.cern.ch/twiki/bin/view/CMS/TrackingPOGFilters. 139
Figure Sources


Appendix A

Supplementary information

A.1 Event cleaning

The complete list of event cleaning filters is listed and described here.

- Primary vertex and beam halo

  **OneGoodVertexFilter** - Requires at least one reconstructed primary vertex satisfying various quality cuts ($|z| < 24\,\text{cm}$, $\rho < 2\,\text{mm}$, $\text{ndof} > 4$

  **BeamHaloFilter** - Filters out events characteristic of the longitudinal passage of the proton beam through the detector rather than interaction at the IP. This happens when the remnants of the beam gas do not focus with the main beam.

- Anomalous calorimeter signals

  **HBHENoiseFilter** - Removes events with various types of HCal related noise inducing fake $\vec{E}_T$, e.g., a single HPD lighting up or a full RBX going hot
**EEBadScFilter** - Filters out events where an ECal endcap super cluster has a failed reconstruction

**ECALLaserCorrFilter** - Filters events where the ECal calibration laser coefficient is too large (8 TeV analysis only)

**HCALLaserFilter** - Filters out events where the HCal calibration laser fires in the collision sequence, lighting up large portions of the HCal (8 TeV analysis only)

- Dead ECal-cells, masked from readout produce fake $\vec{E}_T$

**EcalTPFilter** - Filter removing events with bad ECal energy based on trigger primitive information

**EcalBEFilter** - Filter removing events with bad ECal energy based on the “boundary energy” surrounding the masked cell

- Tracking failure

**PKAMFilter** (i.e., “Previously Known As Monster event”), removes events with anomalously high occupancy in the pixel detector, likely due to beam scraping

**TrackingFailureFilter** - Rejects events wherein the tracking reconstruction failed and the $\sum p_T$ of all the tracks is significantly less than the $H_T$

**InconsistentMuonFilter** - Filters out events with a muon that has inconsistent values for the reconstructed momentum between the PF case and the traditional reconstruction
**GreedyMuonFilter** - Filters out events where large calorimeter deposits are associated to a reconstructed muon

**ManyStripClustersFilter** - Filters out events with more strip clusters than expected based on the pixel hits due to coherent noise in the silicon strip tracker (8 TeV analysis only)

**TooManyStripClustersFilter** - Filters out additional events with more strip clusters than expected based on the pixel hits (8 TeV analysis only)

**LogErrorTooManyClustersFilter** - Filters out events where the track reconstruction was aborted due to the presence of too many clusters (8 TeV analysis only)

- Noise-induced jets

  **EENoiseFilter** - Filters out pathological events due to noise in the ECal endcap (7 TeV analysis only)

  **PBNRFilter** - i.e., “particle based noise rejection”, rejects events where one of the jets has too high a fraction of neutral or photon PF candidates (8 TeV analysis only)

- Anomalous HO signals and TOB/TEC tracking failure (8 TeV analysis only)
  
  - The energy deposits from HO were included in jet reconstruction starting with the CMSSW_5_2_X release. As HO was meant to be a tail catcher, a large fraction of energy deposited in HO implies a significant leakage of the hadronic shower beyond the HCal barrel. As the photodetectors and electronics used for HO are
the same as the rest of HCal, they are susceptible to similar detector related noise. To reject events with anomalous energy deposits in HO resulting in fake $H_T$, an event is rejected if the HO fraction is $>40\%$ for a jet.

- An additional track reconstruction failure mode was observed in the 2012 dataset and showed up as a large number of fake tracks reconstructed in the iteration step (6) when TOB/TEC seeds were used. An event is rejected if a jet within $0.9 < |\eta| < 1.9$ is reconstructed with more than 200 charged hadrons.

A detailed description of all the filters and their implementation in CMSSW provided by the MET and Tracking POG can be found in Refs. [70, 71].
A.2  Z/γ Feynman diagrams

MadGraph diagrams showing the similar production of Z and γ (denoted by ‘a’ in the following diagrams) boson production for one (Fig. A.1), two (Fig. A.2), and three (Fig. A.3) additional partons in the matrix element calculation.
Figure A.1: Comparison of some leading-order Feynman diagrams calculated with MADGRAPH resulting in $Z+1$ jet and $\gamma+1$ jet (i.e., one additional parton).
Figure A.2: Comparison of some leading-order Feynman diagrams calculated with MadGraph resulting in $Z+2$ jets and $\gamma+2$ jets (i.e., two additional partons).
Figure A.3: Comparison of some leading-order Feynman diagrams calculated with MADGRAPH resulting in $Z+3$ jets and $\gamma+3$ jets (i.e., three additional partons).
A.3 Data/simulation comparisons

A.3.1 7 TeV analysis

Data vs. simulation comparisons are shown for various distributions in the 7 TeV analysis. The MC contains all significant background processes and has been normalized to $4.98\,\text{pb}^{-1}$ integrated luminosity.

![Figure A.4: $H_T$ distribution for the 7 TeV analysis (from Ref. [55]).](image)
Figure A.5: $H_T$ distribution for the 7 TeV analysis (from Ref. [55]).

Figure A.6: $\gamma p_T$ distribution for the 7 TeV analysis (from Ref. [55]).
Figure A.7: $\gamma \eta$ distribution for the 7 TeV analysis (from Ref. [55]).

Figure A.8: $\gamma \text{ISO}_{\Delta R=0.4}^{\text{Trk}}$ distribution for the 7 TeV analysis (from Ref. [55]).
Figure A.9: $\gamma$ ISO$^{\text{ECAL}}_{\Delta R=0.4}$ distribution for the 7 TeV analysis (from Ref. [55]).

Figure A.10: $\gamma$ ISO$^{\text{HCAL}}_{\Delta R=0.4}$ distribution for the 7 TeV analysis (from Ref. [55]).
A.3.2 8 TeV analysis

Data vs. simulation comparisons are shown for various distributions in the 8 TeV analysis. The MC contains all significant background processes and has been normalized to 19.4 pb$^{-1}$ integrated luminosity.

![Legend for simulation samples with expected yields normalized to 19.37 fb$^{-1}$ integrated luminosity for the baseline selection.](image)

**Figure A.11:** Legend for simulation samples with expected yields normalized to 19.37 fb$^{-1}$ integrated luminosity for the baseline selection.
Figure A.12: $N_{\text{vertices}}$

Figure A.13: $N_{\text{jets}}(p_T > 50 \text{ GeV}, |\eta| < 2.5)$
Figure A.14: $N_{\text{jets}}(p_T > 30 \text{ GeV}, |\eta| < 5.0)$

Figure A.15: $H_T$
Figure A.16: $H_T$
Figure A.18: $p_T^\gamma$

Figure A.19: $H_T$
Figure A.20: $H_T$

Figure A.21: $\eta^\gamma$
Figure A.22: $\sigma_{\gamma_{\eta\eta}}$

Figure A.23: $\text{Iso}_{\Delta R=0.3}^{\text{HAD}}$
Figure A.24: $Iso^{HAD^0}_{\Delta R=0.3}$

Figure A.25: $Iso^\gamma_{\Delta R=0.3}$
A.4 Additional tables for the prediction of the $Z \rightarrow \nu \bar{\nu} + \text{jets}$ background
<table>
<thead>
<tr>
<th>$H_T$ GeV</th>
<th>$H_T$ GeV</th>
<th>Correction factor</th>
<th>$Z/\gamma$ ratio</th>
<th>Acc. [%]</th>
<th>Reco/Iso [%]</th>
<th>Purity [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>500...800</td>
<td>200...350</td>
<td>0.27 ±0.22% (syst.)</td>
<td>0.21 ±2.00% ±21%</td>
<td>91 ±0.49%</td>
<td>78 ±9.88%</td>
<td>97 ±1.30%</td>
</tr>
<tr>
<td>500...800</td>
<td>350...500</td>
<td>0.32 ±0.22% (syst.)</td>
<td>0.24 ±3.91% ±21%</td>
<td>90 ±1.04%</td>
<td>76 ±1.87%</td>
<td>99 ±2.56%</td>
</tr>
<tr>
<td>500...800</td>
<td>500...600</td>
<td>0.30 ±0.24% (syst.)</td>
<td>0.24 ±9.05% ±21%</td>
<td>94 ±2.00%</td>
<td>78 ±4.01%</td>
<td>99 ±2.45%</td>
</tr>
<tr>
<td>500...800</td>
<td>600...</td>
<td>0.37 ±0.39% (syst.)</td>
<td>0.27 ±17.64% ±33%</td>
<td>95 ±3.86%</td>
<td>73 +9.29%</td>
<td>99 ±2.27%</td>
</tr>
<tr>
<td>800...1000</td>
<td>200...350</td>
<td>0.31 ±0.37% (syst.)</td>
<td>0.23 ±5.88% ±36%</td>
<td>90 ±1.59%</td>
<td>77 ±2.74%</td>
<td>97 ±1.27%</td>
</tr>
<tr>
<td>800...1000</td>
<td>350...500</td>
<td>0.36 ±0.38% (syst.)</td>
<td>0.29 ±11.16% ±36%</td>
<td>94 ±2.46%</td>
<td>80 +4.96%</td>
<td>99 ±2.47%</td>
</tr>
<tr>
<td>800...1000</td>
<td>500...600</td>
<td>0.54 ±0.44% (syst.)</td>
<td>0.38 ±21.05% ±36%</td>
<td>96 ±4.78%</td>
<td>68 +12.69%</td>
<td>99 ±2.28%</td>
</tr>
<tr>
<td>800...1000</td>
<td>600...</td>
<td>0.30 ±0.41% (syst.)</td>
<td>0.24 ±18.16% ±36%</td>
<td>97 ±3.17%</td>
<td>77 +8.34%</td>
<td>99 ±2.36%</td>
</tr>
<tr>
<td>1000...1200</td>
<td>200...350</td>
<td>0.27 ±0.44% (syst.)</td>
<td>0.22 ±10.39% ±42%</td>
<td>91 ±2.79%</td>
<td>81 +4.35%</td>
<td>97 ±1.24%</td>
</tr>
<tr>
<td>1000...1200</td>
<td>350...500</td>
<td>0.39 ±0.48% (syst.)</td>
<td>0.30 ±19.89% ±42%</td>
<td>94 ±5.01%</td>
<td>76 +9.95%</td>
<td>99 ±2.53%</td>
</tr>
<tr>
<td>1000...1200</td>
<td>500...</td>
<td>0.73 ±0.54% (syst.)</td>
<td>0.37 ±25.31% ±42%</td>
<td>95 ±6.69%</td>
<td>50 +21.32%</td>
<td>99 ±2.46%</td>
</tr>
<tr>
<td>1200...1400</td>
<td>200...350</td>
<td>0.31 ±0.47% (syst.)</td>
<td>0.24 ±17.50% ±42%</td>
<td>93 ±4.32%</td>
<td>78 +8.15%</td>
<td>98 ±1.46%</td>
</tr>
<tr>
<td>1200...1400</td>
<td>350...</td>
<td>0.39 ±0.51% (syst.)</td>
<td>0.36 ±26.60% ±42%</td>
<td>100 ±5.58%</td>
<td>89 +9.27%</td>
<td>99 ±2.35%</td>
</tr>
<tr>
<td>1400...</td>
<td>200...</td>
<td>0.18 ±0.48% (syst.)</td>
<td>0.13 ±20.40% ±42%</td>
<td>95 ±3.86%</td>
<td>73 +9.29%</td>
<td>98 ±1.61%</td>
</tr>
</tbody>
</table>

Table A.1: Correction factor and individual components, along with systematic uncertainties expressed as a percentage of the correction, for the $\gamma +$ jets method (cf. Table 6.11).
<table>
<thead>
<tr>
<th>$N_{\text{jets}}$</th>
<th>Prediction</th>
<th>$H_T$ [GeV]</th>
<th>$\bar{H}_T$ [GeV]</th>
<th>stat.</th>
<th>acc.</th>
<th>reco.</th>
<th>pix.</th>
<th>iso.</th>
<th>data/MC</th>
<th>purity</th>
<th>ratio</th>
<th>pheno. fit</th>
<th>double</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 ≤ $N_{\text{jets}}$ ≤ 5</td>
<td>500...800 200...300</td>
<td>1821.28</td>
<td>± 1.44 %</td>
<td>+0.76 %</td>
<td>-0.76 %</td>
<td>+0.43 %</td>
<td>-0.43 %</td>
<td>+1.13 %</td>
<td>+2.46 %</td>
<td>+2.84 %</td>
<td>+2.36 %</td>
<td>+0.54 %</td>
<td>+1.52 %</td>
</tr>
<tr>
<td></td>
<td>500...800 300...450</td>
<td>993.58</td>
<td>± 1.99 %</td>
<td>+0.82 %</td>
<td>-0.82 %</td>
<td>+0.59 %</td>
<td>-0.59 %</td>
<td>+1.17 %</td>
<td>+2.79 %</td>
<td>+2.52 %</td>
<td>+2.24 %</td>
<td>+0.67 %</td>
<td>+1.70 %</td>
</tr>
<tr>
<td></td>
<td>500...800 450...600</td>
<td>273.21</td>
<td>± 3.77 %</td>
<td>+1.18 %</td>
<td>+0.94 %</td>
<td>+1.84 %</td>
<td>+4.50 %</td>
<td>+2.49 %</td>
<td>+2.22 %</td>
<td>+1.08 %</td>
<td>+2.51 %</td>
<td>+16.81 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>500...800 600...</td>
<td>42.01</td>
<td>± 9.62 %</td>
<td>+1.16 %</td>
<td>-0.92 %</td>
<td>-1.80 %</td>
<td>-0.18 %</td>
<td>-1.42 %</td>
<td>-2.41 %</td>
<td>-2.18 %</td>
<td>-1.06 %</td>
<td>-2.48 %</td>
<td>-16.79 %</td>
</tr>
<tr>
<td></td>
<td>800...1000 200...300</td>
<td>215.84</td>
<td>± 4.19 %</td>
<td>+0.83 %</td>
<td>+0.47 %</td>
<td>+1.23 %</td>
<td>+2.69 %</td>
<td>+2.86 %</td>
<td>+2.45 %</td>
<td>+0.58 %</td>
<td>+1.63 %</td>
<td>+17.09 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>800...1000 300...450</td>
<td>124.08</td>
<td>± 5.61 %</td>
<td>+0.99 %</td>
<td>+0.72 %</td>
<td>+1.41 %</td>
<td>+3.40 %</td>
<td>+2.56 %</td>
<td>+2.54 %</td>
<td>+0.80 %</td>
<td>+1.96 %</td>
<td>+17.14 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>800...1000 450...600</td>
<td>46.86</td>
<td>± 9.09 %</td>
<td>+1.35 %</td>
<td>-0.10 %</td>
<td>-2.10 %</td>
<td>-5.17 %</td>
<td>-2.48 %</td>
<td>-2.46 %</td>
<td>-1.23 %</td>
<td>-2.85 %</td>
<td>-16.98 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>800...1000 600...</td>
<td>35.29</td>
<td>± 10.48 %</td>
<td>+1.25 %</td>
<td>+1.00 %</td>
<td>+1.96 %</td>
<td>+4.80 %</td>
<td>+2.42 %</td>
<td>+2.22 %</td>
<td>+1.14 %</td>
<td>+2.66 %</td>
<td>+16.88 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000...1250 200...300</td>
<td>76.28</td>
<td>± 7.04 %</td>
<td>+0.85 %</td>
<td>-0.48 %</td>
<td>-1.27 %</td>
<td>-2.78 %</td>
<td>-2.77 %</td>
<td>-2.67 %</td>
<td>-0.60 %</td>
<td>-1.67 %</td>
<td>-17.15 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000...1250 300...450</td>
<td>39.34</td>
<td>± 9.95 %</td>
<td>+1.02 %</td>
<td>+0.74 %</td>
<td>+1.45 %</td>
<td>+3.50 %</td>
<td>+2.53 %</td>
<td>+2.57 %</td>
<td>+0.81 %</td>
<td>+2.02 %</td>
<td>+17.18 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000...1250 450...600</td>
<td>18.08</td>
<td>± 14.59 %</td>
<td>+1.55 %</td>
<td>-0.22 %</td>
<td>+2.40 %</td>
<td>+5.94 %</td>
<td>+2.46 %</td>
<td>+2.59 %</td>
<td>+1.37 %</td>
<td>+3.22 %</td>
<td>+17.17 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000...1250 600...</td>
<td>17.81</td>
<td>± 14.75 %</td>
<td>+1.36 %</td>
<td>-1.08 %</td>
<td>-2.12 %</td>
<td>-5.23 %</td>
<td>-2.34 %</td>
<td>-2.12 %</td>
<td>-1.23 %</td>
<td>+2.86 %</td>
<td>-16.08 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1250...1500 200...300</td>
<td>25.34</td>
<td>± 12.22 %</td>
<td>+0.84 %</td>
<td>+0.47 %</td>
<td>+1.24 %</td>
<td>+2.73 %</td>
<td>+2.71 %</td>
<td>+2.56 %</td>
<td>+0.59 %</td>
<td>+1.65 %</td>
<td>+17.13 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1250...1500 300...450</td>
<td>16.71</td>
<td>± 15.26 %</td>
<td>+1.08 %</td>
<td>+0.78 %</td>
<td>+1.54 %</td>
<td>+3.73 %</td>
<td>+2.41 %</td>
<td>+2.57 %</td>
<td>+0.86 %</td>
<td>+2.13 %</td>
<td>+17.28 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1250...1500 450...500</td>
<td>12.35</td>
<td>± 17.68 %</td>
<td>+1.53 %</td>
<td>+1.21 %</td>
<td>+2.37 %</td>
<td>+5.86 %</td>
<td>+2.34 %</td>
<td>+2.52 %</td>
<td>+1.35 %</td>
<td>+3.18 %</td>
<td>+17.15 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1500... 200...300</td>
<td>10.61</td>
<td>± 18.91 %</td>
<td>+0.94 %</td>
<td>-0.52 %</td>
<td>-1.39 %</td>
<td>-3.06 %</td>
<td>-2.76 %</td>
<td>-2.67 %</td>
<td>-0.65 %</td>
<td>-1.83 %</td>
<td>-17.27 %</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1500... 300...</td>
<td>10.93</td>
<td>± 18.90 %</td>
<td>-1.02 %</td>
<td>+0.76 %</td>
<td>+1.50 %</td>
<td>+3.62 %</td>
<td>+2.41 %</td>
<td>+2.37 %</td>
<td>+0.86 %</td>
<td>+2.09 %</td>
<td>+17.00 %</td>
<td></td>
</tr>
</tbody>
</table>

(a) $3 \leq N_{\text{jets}} \leq 5$

<p>| Table A.2: $Z\rightarrow \nu\bar{\nu}+$jets uncertainties as a percentage of the prediction from $\gamma$ + jets data for $3 \leq N_{\text{jets}} \leq 5$ (cf. Table 6.12). (cf. Table A.2). |</p>
<table>
<thead>
<tr>
<th>$N_{\text{jets}} \geq 8$</th>
<th>$H_T$ [GeV]</th>
<th>$B_T$ [GeV]</th>
<th>Prediction</th>
<th>stats.</th>
<th>acc.</th>
<th>reco.</th>
<th>pix.</th>
<th>iso.</th>
<th>data/MC</th>
<th>purity</th>
<th>ratio</th>
<th>pheno. fit</th>
<th>double</th>
</tr>
</thead>
<tbody>
<tr>
<td>$500 \ldots 800$</td>
<td>$200 \ldots 300$</td>
<td>4.00</td>
<td>±0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
</tr>
<tr>
<td>$800 \ldots 1000$</td>
<td>$200 \ldots 300$</td>
<td>0.64</td>
<td>±70.74 %</td>
<td>+6.62 %</td>
<td>+4.56 %</td>
<td>+10.69 %</td>
<td>+25.74 %</td>
<td>+3.28 %</td>
<td>+4.55 %</td>
<td>+3.97 %</td>
<td>+12.02 %</td>
<td>+22.58 %</td>
<td></td>
</tr>
<tr>
<td>$1000 \ldots 1250$</td>
<td>$200 \ldots 300$</td>
<td>0.60</td>
<td>±70.75 %</td>
<td>+6.01 %</td>
<td>+3.19 %</td>
<td>+8.80 %</td>
<td>+20.83 %</td>
<td>+4.07 %</td>
<td>+4.71 %</td>
<td>+2.83 %</td>
<td>+12.27 %</td>
<td>+23.44 %</td>
<td></td>
</tr>
<tr>
<td>$1250 \ldots 1500$</td>
<td>$200 \ldots 300$</td>
<td>0.00</td>
<td>±0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td></td>
</tr>
<tr>
<td>$1500 \ldots 2000$</td>
<td>$200 \ldots 300$</td>
<td>0.00</td>
<td>±0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td>+0.00 %</td>
<td></td>
</tr>
</tbody>
</table>

Table A.2: (Continued) $Z \rightarrow \nu \bar{\nu} + \text{jets}$ uncertainties from $\gamma + \text{jets}$ data for $6 \leq N_{\text{jets}} \leq 7$ and $N_{\text{jets}} \geq 8$ (cf. Table 6.12).
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>500...800</td>
<td>200...300</td>
<td>1847.28 ± 0.67%</td>
<td>±2.55%</td>
<td>±1.48%</td>
<td>±2.95%</td>
<td>1821.28 ± 1.44%</td>
<td>±4.78%</td>
<td>±17.05%</td>
<td>±17.74%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500...800</td>
<td>300...450</td>
<td>1027.23 ± 0.91%</td>
<td>±2.84%</td>
<td>±1.69%</td>
<td>±3.31%</td>
<td>993.58 ± 1.99%</td>
<td>±4.77%</td>
<td>±16.76%</td>
<td>±17.77%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500...800</td>
<td>450...600</td>
<td>292.22 ± 1.68%</td>
<td>±4.12%</td>
<td>±2.45%</td>
<td>±4.80%</td>
<td>273.21 ± 3.77%</td>
<td>±6.22%</td>
<td>±17.02%</td>
<td>±18.20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500...800</td>
<td>600...800</td>
<td>43.13 ± 4.38%</td>
<td>±3.98%</td>
<td>±2.34%</td>
<td>±4.61%</td>
<td>42.01 ± 9.62%</td>
<td>±6.09%</td>
<td>±16.99%</td>
<td>±18.12%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>800...1000</td>
<td>200...300</td>
<td>236.19 ± 1.86%</td>
<td>±2.69%</td>
<td>±1.60%</td>
<td>±3.13%</td>
<td>215.84 ± 4.19%</td>
<td>±5.01%</td>
<td>±17.18%</td>
<td>±17.93%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>800...1000</td>
<td>300...450</td>
<td>140.59 ± 2.45%</td>
<td>±3.14%</td>
<td>±1.90%</td>
<td>±3.67%</td>
<td>124.08 ± 5.61%</td>
<td>±5.45%</td>
<td>±17.27%</td>
<td>±18.17%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>800...1000</td>
<td>450...600</td>
<td>55.38 ± 3.88%</td>
<td>±4.65%</td>
<td>±2.82%</td>
<td>±5.44%</td>
<td>46.86 ± 9.09%</td>
<td>±6.94%</td>
<td>±17.25%</td>
<td>±18.70%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>800...1000</td>
<td>600...800</td>
<td>34.49 ± 4.91%</td>
<td>±4.42%</td>
<td>±2.65%</td>
<td>±5.15%</td>
<td>35.29 ± 10.48%</td>
<td>±6.52%</td>
<td>±17.12%</td>
<td>±18.43%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000...1250</td>
<td>200...300</td>
<td>94.71 ± 2.94%</td>
<td>±2.73%</td>
<td>±1.63%</td>
<td>±3.18%</td>
<td>76.28 ± 7.04%</td>
<td>±5.07%</td>
<td>±17.24%</td>
<td>±18.01%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000...1250</td>
<td>300...450</td>
<td>52.63 ± 4.01%</td>
<td>±3.17%</td>
<td>±1.92%</td>
<td>±3.71%</td>
<td>39.34 ± 9.95%</td>
<td>±5.53%</td>
<td>±17.31%</td>
<td>±18.23%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000...1250</td>
<td>450...600</td>
<td>21.70 ± 6.20%</td>
<td>±4.87%</td>
<td>±2.97%</td>
<td>±5.71%</td>
<td>18.08 ± 14.59%</td>
<td>±7.80%</td>
<td>±17.51%</td>
<td>±19.30%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000...1250</td>
<td>600...800</td>
<td>19.20 ± 6.58%</td>
<td>±4.41%</td>
<td>±2.64%</td>
<td>±5.14%</td>
<td>17.81 ± 14.75%</td>
<td>±6.93%</td>
<td>±17.26%</td>
<td>±18.76%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1250...1500</td>
<td>200...300</td>
<td>28.50 ± 5.36%</td>
<td>±2.65%</td>
<td>±1.60%</td>
<td>±3.13%</td>
<td>25.34 ± 12.22%</td>
<td>±5.01%</td>
<td>±17.22%</td>
<td>±17.97%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1250...1500</td>
<td>300...450</td>
<td>16.07 ± 7.26%</td>
<td>±3.12%</td>
<td>±1.88%</td>
<td>±3.65%</td>
<td>16.71 ± 15.26%</td>
<td>±5.68%</td>
<td>±17.41%</td>
<td>±18.39%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1250...1500</td>
<td>450...600</td>
<td>12.37 ± 8.22%</td>
<td>±5.07%</td>
<td>±3.11%</td>
<td>±5.95%</td>
<td>12.35 ± 17.68%</td>
<td>±7.65%</td>
<td>±17.49%</td>
<td>±19.25%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1500...2000</td>
<td>200...300</td>
<td>13.28 ± 7.86%</td>
<td>±2.79%</td>
<td>±1.69%</td>
<td>±3.26%</td>
<td>10.51 ± 18.91%</td>
<td>±5.29%</td>
<td>±17.38%</td>
<td>±18.23%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1500...2000</td>
<td>300...450</td>
<td>11.80 ± 8.45%</td>
<td>±4.00%</td>
<td>±2.45%</td>
<td>±4.69%</td>
<td>10.93 ± 18.90%</td>
<td>±5.49%</td>
<td>±17.14%</td>
<td>±18.09%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table A.3:** Z\to\nu\bar{\nu}+jets predictions from γ + jets simulation compared to the prediction coming from data (19.4 pb⁻¹) for $3 \leq N_{\text{jets}} \leq 5$ with uncertainties expressed as a percentage of the prediction (cf. Table 6.14).
Table A.3: (Continued) $Z \rightarrow \nu \bar{\nu} + \text{jets}$ predictions from $\gamma + \text{jets}$ simulation compared to the prediction coming from data (19.4 pb$^{-1}$) for $3 \leq N_{\text{jets}} \leq 5$ with uncertainties expressed as a percentage of the prediction (cf. Table 6.14).
<table>
<thead>
<tr>
<th>$3 \leq N_{\text{jets}} \leq 5$</th>
<th>Prediction on data</th>
<th>uncertainty [%]</th>
<th>$Z\rightarrow\nu\bar{\nu}$ uncertainty [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>500...800</td>
<td>200...300</td>
<td>1821.28 $\pm$ 1.44%</td>
<td>$\pm$ 4.78%</td>
</tr>
<tr>
<td>500...800</td>
<td>300...450</td>
<td>993.58 $\pm$ 1.99%</td>
<td>$\pm$ 4.77%</td>
</tr>
<tr>
<td>500...800</td>
<td>450...600</td>
<td>273.21 $\pm$ 3.77%</td>
<td>$\pm$ 6.23%</td>
</tr>
<tr>
<td>500...800</td>
<td>600...</td>
<td>42.01 $\pm$ 9.62%</td>
<td>$\pm$ 6.09%</td>
</tr>
<tr>
<td>800...1000</td>
<td>200...300</td>
<td>215.84 $\pm$ 4.19%</td>
<td>$\pm$ 5.61%</td>
</tr>
<tr>
<td>800...1000</td>
<td>300...450</td>
<td>124.08 $\pm$ 5.61%</td>
<td>$\pm$ 5.45%</td>
</tr>
<tr>
<td>800...1000</td>
<td>450...600</td>
<td>46.86 $\pm$ 9.09%</td>
<td>$\pm$ 6.94%</td>
</tr>
<tr>
<td>800...1000</td>
<td>600...</td>
<td>35.29 $\pm$ 10.48%</td>
<td>$\pm$ 6.52%</td>
</tr>
<tr>
<td>1000...1250</td>
<td>200...300</td>
<td>76.28 $\pm$ 7.04%</td>
<td>$\pm$ 5.67%</td>
</tr>
<tr>
<td>1000...1250</td>
<td>300...450</td>
<td>39.34 $\pm$ 9.95%</td>
<td>$\pm$ 5.14%</td>
</tr>
<tr>
<td>1000...1250</td>
<td>450...600</td>
<td>18.08 $\pm$ 14.59%</td>
<td>$\pm$ 7.80%</td>
</tr>
<tr>
<td>1000...1250</td>
<td>600...</td>
<td>17.81 $\pm$ 14.75%</td>
<td>$\pm$ 6.93%</td>
</tr>
<tr>
<td>1250...1500</td>
<td>200...300</td>
<td>25.34 $\pm$ 12.22%</td>
<td>$\pm$ 5.61%</td>
</tr>
<tr>
<td>1250...1500</td>
<td>300...450</td>
<td>16.71 $\pm$ 15.26%</td>
<td>$\pm$ 5.68%</td>
</tr>
<tr>
<td>1250...1500</td>
<td>450...</td>
<td>12.35 $\pm$ 17.68%</td>
<td>$\pm$ 5.71%</td>
</tr>
<tr>
<td>1500...2000</td>
<td>300...300</td>
<td>10.51 $\pm$ 18.91%</td>
<td>$\pm$ 5.29%</td>
</tr>
<tr>
<td>1500...2000</td>
<td>300...</td>
<td>10.93 $\pm$ 18.90%</td>
<td>$\pm$ 5.49%</td>
</tr>
</tbody>
</table>

(a) $3 \leq N_{\text{jets}} \leq 5$

Table A.4: $Z\rightarrow\nu\bar{\nu}$+jets predictions from $\gamma$+jets data (19.4 pb$^{-1}$) for $3 \leq N_{\text{jets}} \leq 5$ with uncertainties taken as a percentage of the prediction (cf. Table 6.15).
| $6 \leq N_{\text{jets}} \leq 7$ | Prediction | uncertainty [%] | $Z \rightarrow \nu \bar{\nu}$ | uncertainty [%] |
|---|---|---|---|
| | on data | stat. | syst. | pheno. | tot. | NNLO MC | stat. |
| 500...800 | 200...300 | 22.72 | ± 12.22% | +11.98% | ±20.53% | +23.79% | 20.16 | ± 4.05% |
| 500...800 | 300...450 | 9.92 | ± 18.90% | +13.46% | ±20.41% | +24.45% | 6.49 | ± 6.05% |
| 500...800 | 450... | 0.70 | ± 70.72% | +24.54% | ±22.53% | +33.35% | 0.67 | ± 18.90% |
| 800...1000 | 200...300 | 9.09 | ± 19.25% | +12.69% | ±20.90% | +24.48% | 7.49 | ± 5.63% |
| 800...1000 | 300...450 | 4.23 | ± 28.88% | +14.37% | ±20.76% | +25.30% | 4.57 | ± 7.22% |
| 800...1000 | 450... | 1.77 | ± 44.73% | +22.24% | ±21.84% | +31.21% | 1.45 | ± 12.80% |
| 1000...1250 | 200...300 | 4.37 | ± 27.74% | +12.59% | ±20.77% | +24.31% | 4.71 | ± 7.11% |
| 1000...1250 | 300...450 | 3.52 | ± 31.64% | +14.30% | ±20.67% | +25.18% | 2.55 | ± 9.67% |
| 1000...1250 | 450... | 1.41 | ± 50.01% | +22.79% | ±21.96% | +31.68% | 1.45 | ± 12.80% |
| 1250...1500 | 200...300 | 3.33 | ± 31.64% | +13.55% | ±21.28% | +25.34% | 1.90 | ± 11.18% |
| 1250...1500 | 300...450 | 1.42 | ± 50.00% | +13.26% | ±20.35% | +24.26% | 1.07 | ± 14.91% |
| 1250...1500 | 450... | 0.36 | ± 100.00% | +20.43% | ±21.40% | +29.58% | 0.74 | ± 17.96% |
| 1500... | 200...300 | 1.33 | ± 50.01% | +13.53% | ±21.30% | +25.26% | 1.09 | ± 14.74% |
| 1500... | 300... | 1.06 | ± 57.74% | +13.60% | ±20.46% | +30.11% | 0.88 | ± 16.44% |

| $N_{\text{jets}} \geq 8$ | Prediction | uncertainty [%] | $Z \rightarrow \nu \bar{\nu}$ | uncertainty [%] |
|---|---|---|---|
| | on data | stat. | syst. | pheno. | tot. | NNLO MC | stat. |
| 500...800 | 200... | 0.00 | ± 0.00% | ±0.00% | ±0.00% | ±0.00% | 0.10 | ± 50.00% |
| 800...1000 | 200... | 0.64 | ± 70.74% | ±29.67% | ±25.61% | ±39.53% | 0.17 | ± 37.80% |
| 1000...1250 | 200... | 0.60 | ± 70.75% | ±24.68% | ±26.49% | ±36.24% | 0.38 | ± 25.00% |
| 1250...1500 | 200... | 0.00 | ± 0.00% | ±0.00% | ±0.00% | ±0.00% | 0.31 | ± 27.74% |
| 1500... | 200... | 0.00 | ± 0.00% | ±0.00% | ±0.00% | ±0.00% | 0.12 | ± 44.72% |

Table A.4: (Continued) $Z \rightarrow \nu \bar{\nu}$+jets predictions from $\gamma$+jets data (19.4 pb$^{-1}$) for $6 \leq N_{\text{jets}} \leq 7$ and $N_{\text{jets}} \geq 8$ with uncertainties taken as a percentage of the prediction (cf. Table 6.15).
Appendix B

Service work

B.1 Hcal logical map

It is imperative in isolating misbehaving channels to be able to track signals from particles through the entire readout system. In this way, one can pinpoint the failure reason in a systematic way. To this end, my first task when I arrived at CERN was to write a piece of C++ code that mapped on-detector hardware channels to off-detector hardware. Due to the design structure of the Hcal, the majority of the channels are cyclical and the information can easily be put into an algorithm with simple loops. I took various pieces of previously written code (C++ and Fortran) and created a consolidated map project. Hcal software utilizes several detector ID hashes, i.e., unique identifiers for each piece of hardware, and during the course of designing the “logical map” I implemented a new hash, the HcalFrontEndId, which uniquely identified each channel based on information related to the Hcal front end, e.g., RBX (see Fig. 3.9). With this, one can now move between the various detector IDs to identify the complete information for a specific channel.
Additionally, during the course of running the detector, several issues with the physical mapping were discovered, e.g., fibers had been installed in swapped locations, and this was accounted for in several versions of the logical map code. The code was written to allow someone to take a snapshot of how the detector was connected at a given time. Eventually, the output of this code was loaded into one of the CMS databases, and the code itself integrated into the Hcal data quality monitoring (DQM) code.

B.2 CMS Hcal operations

The CMS Hcal is a large and complex machine and requires the time and expertise of many people to ensure quality running and efficient data taking. The people who are responsible for the maintenance of the machine are also the people who are actively analyzing data, and yet, both tasks consume a large amount of time.

I was involved with the operation of the Hcal subdetector in various capacities. I began my time as an Hcal operator during the early commissioning and early “global runs,” i.e., when the whole CMS detector was first run as a unit. During these times, it was necessary for each subdetector to provide people to be available all the time, to help debug problems as well as to take “local runs,” i.e., subdetector specific runs used for calibration and monitoring. Through being an HCal operator, I learned a great amount about how Hcal works and the various problems it can run into during typical operations. As a result of this, I eventually became an HCal expert, providing on-call support for later HCal operators, as well as support when the Hcal operator position was made obsolete and the whole detector was run with only ∼ 5 people in the control room.
One of the problems uncovered during my time as an Hcal expert was the sudden transition of a single HB cell to continuously saturate. This problem appeared as a spike in several of the trigger rates and I was called at 4 a.m. Although the root cause of the problem was not immediately clear, after consultation with additional experts, I was able to recommend that the central trigger shifter mask the area from contributing to the trigger.

Although I encountered no other problems of this sort, other interactions with the central shift crew were similar. I would receive a call about some issue, consult with the monitoring plots and logs to determine if there was a problem and either diagnose it or call the person responsible for the misbehaving component. The aim was to minimize the amount of “downtime,” i.e., the amount of time that the accelerator is delivering quality beam data that is bad or corrupt due to a misbehaving subdetector.

Another task I was involved in was to provide “prompt feedback” on issues discovered. For example, when the hot HB cell was discovered at 4 in the morning, the decision was to mask it and further investigate in the day that followed. These types of tasks were assigned to the same experts who provided on-call support, and were undertaken on a longer term than required by the need to get the detector operational again, which was the main consideration during data taking. In these cases, an analysis was performed, designed to isolate the source of the problem, e.g., a specific piece of misbehaving hardware. After the analysis was performed, the results were presented in the Hcal operations meeting and a plan of action was taken to correct the issue.
B.3 CMS central operations

Running a large scale experiment such as CMS requires a large number of people to monitor the detector to ensure that it functions with a high efficiency. To do my part I was involved in the central operation of the CMS detector, specifically, serving as a DAQ system operator, or shifter. This position is important as this person is responsible for starting and stopping physics taking runs and ensuring that the data being written out to disk is not encountering any problems. This task required sitting in the control room at Point 5 (location of CMS) (P5) for eight straight hours, regularly monitoring up to six different computer screens. Audio alarms aided the shifter in knowing that a transition, whether intended or otherwise, was happening.

In cases where a specific portion of the DAQ system was misbehaving, the expert would be called, and based on the recommendation of the expert, the misbehaving portion could be masked and investigated later, or the run could be stopped to diagnose further. As always, the goal of the operation is to collect proton-proton collision data with a high efficiency and little downtime, and reliable and fast action is required while beam is in the LHC machine.

B.4 Jet/$\vec{E}_T$ data quality monitoring

Ensuring that the reconstructed physics observables are behaving properly is paramount in ensuring that analyses are robust. In addition, understanding why basic distributions change helps track down systematic issues in various subdetectors. To this end, I spent several years developing and maintaining the DQM code for the JetMET group. The code
monitored the performance of the different jet collections, as well as the different $\vec{E}_T$ collections. In serving as the lead developer, I was involved in the certification of data for the purposes of JetMET. This consisted in writing code to perform tests directly within the DQM code and visually scanning selected distributions for problems. At the end of the certification, the run was considered either good or bad from a JetMET perspective, meaning that the jet and $\vec{E}_T$ objects are useful (or not) for physics analysis. If problems were discovered during certification, the problem is almost invariably due to some portion of the detector misbehaving, and the JetMET DQM code provides an additional check for problems that escape the attention of the specific subdetector certification teams. To this end, we coordinated with the different subdetectors in attempting to specify any issues encountered during certification.

One of the developments I was directly responsible for was the ability to monitor distributions taken with only a specific trigger. This became a necessity once triggers began to be highly prescaled and the inclusive distributions appeared distorted. Selecting a specific trigger, or only an un-prescaled trigger, allowed us to more accurately determine the performance of the various reconstruction algorithms.