Measurements of Z Transverse Momentum Shape Using Novel Variables With the CMS Detector

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF THE UNIVERSITY OF MINNESOTA
BY

Kevin H. Klapoetke

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
Doctor of Philosophy

June, 2014
Acknowledgements

I would like to thank my friends and colleagues in the CMS group, past and present. My Adviser Jeremiah Mans for countless things, Sinjini Sengupta for being an enormous help in getting started in CMS and a good friend as well, the original group of fellow grad students who have graduated before me (Phil Dudero, Jason Haupt, Abe Debenedetti and Seth Cooper), the new group I will leave behind (Alex Gude, Joe Pastika, Tambe Norbert, Jared Turkewitz, Alexey Finkel), the very new group that I haven’t gotten to know as well as I’d like (Sean, Pete, Zach, Andrew, and others), and the post-docs and other CMS associates (Bryan Dahmes, Nicole Ruckstuhl, Giovanni Franzoni, and the rest).

I thank my wife, Brianna, for giving me a life so much better than writing a thesis could ever be, and my family for always being there; Mom, Dad, Karen, Mark, and Sam. A further thanks to all my friends I’ve made living here from many different places, at many different times.

I am grateful to the students I’ve taught to, the TAs I’ve taught with, the professors I’ve taught for, and the University I’ve taught at. A special thanks to Dodi and Dr. Hermann for help in a very difficult time.

I’d also like to thank Kevin Klapoetke for finally finishing this thesis with me. I could not have done it without you, though at times I feel you were less help than you could have been.
Dedication

To the Klapoetkes and the Humphreys in my life, and to those that come along with them. All of them.
Abstract

The understanding of transverse momentum in a proton-proton collision is important. Particularly at low values, measurements of transverse momentum are not in agreement with proposed models. This thesis describes a measurement of the physics behind the transverse momentum of the Z boson using a method that reduces the effect of energy resolution and capitalizes on the positional accuracy of the CMS detector. This method is a differential cross-section measurement with respect to the presented novel variable, $\phi^*$. The data used was from the 2011 run at the LHC experiment, totaling 216pb$^{-1}$ of integrated luminosity at $\sqrt{s} = 7$ TeV.
2.4.2 The Hadronic Calorimeter ........................................ 16
2.4.3 The Forward Hadron Calorimeter ......................... 16
2.4.4 Triggers ..................................................... 19

3 HF Electrons .................................................... 24

3.1 Introduction .................................................. 24

3.1.1 HF Electron Reconstruction ................................. 24

3.2 Finding HF Electron Candidates ............................... 25

3.2.1 PMT Hits .................................................. 25

3.2.2 The HF Cluster ........................................... 27

3.2.3 HF Cluster Variables ..................................... 28

3.3 HF Electron Candidate Corrections ............................. 29

3.3.1 Correcting Position ...................................... 30

3.3.2 Correcting Energy ....................................... 33

3.4 Identification of HF Electrons ................................. 36

4 Physics of Z Transverse Momentum ............................. 43

4.1 Introduction .................................................. 43

4.2 Standard Model ................................................ 43

4.2.1 Quarks .................................................... 44

4.2.2 Gluons and Strong Interactions ............................ 46

4.2.3 Hadrons: Baryons and Mesons ............................ 46

4.2.4 The Weak Bosons ....................................... 47

4.2.5 Electrons and Leptons .................................... 48

4.3 Proton-Proton Collisions ...................................... 48

4.3.1 Protons at the LHC ....................................... 48

4.3.2 Proton Parton Model ..................................... 48

4.3.3 Probing Proton Partons .................................. 49

4.4 The Transverse Momentum of Z bosons ...................... 51

4.4.1 Transverse Momentum ................................... 51

4.4.2 Z Boson Production with Transverse Momentum .......... 52

4.4.3 Z Boson Differential Cross-Section ................. 54

4.5 QCD Predictions of Transverse Momentum ............. 54
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5.1</td>
<td>Perturbative Calculations of the Differential Cross-section</td>
<td>55</td>
</tr>
<tr>
<td>4.5.2</td>
<td>The Differential Form at Low $Q_T$</td>
<td>57</td>
</tr>
<tr>
<td>4.5.3</td>
<td>Non-Perturbative Sudakov Method</td>
<td>59</td>
</tr>
<tr>
<td>4.5.4</td>
<td>Leading Logarithm Method</td>
<td>60</td>
</tr>
<tr>
<td>4.6</td>
<td>Measurements</td>
<td>61</td>
</tr>
<tr>
<td>4.6.1</td>
<td>Resolution</td>
<td>61</td>
</tr>
<tr>
<td>4.6.2</td>
<td>Novel Variables</td>
<td>62</td>
</tr>
<tr>
<td>5</td>
<td>Measuring a Z Boson Differential Cross-section</td>
<td>66</td>
</tr>
<tr>
<td>5.1</td>
<td>Electron Selection</td>
<td>66</td>
</tr>
<tr>
<td>5.1.1</td>
<td>ECAL Electron Identification</td>
<td>67</td>
</tr>
<tr>
<td>5.2</td>
<td>Acceptance</td>
<td>68</td>
</tr>
<tr>
<td>5.3</td>
<td>Differential Cross-section</td>
<td>69</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Data Driven Definition of Differential Cross-section</td>
<td>69</td>
</tr>
<tr>
<td>5.4</td>
<td>Z Boson Detection</td>
<td>71</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Efficiency</td>
<td>71</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Backgrounds</td>
<td>72</td>
</tr>
<tr>
<td>5.5</td>
<td>Example Differential Cross-section Measurements</td>
<td>72</td>
</tr>
<tr>
<td>5.5.1</td>
<td>Data Set and Event Reconstruction</td>
<td>73</td>
</tr>
<tr>
<td>5.5.2</td>
<td>Results</td>
<td>75</td>
</tr>
<tr>
<td>5.5.3</td>
<td>Transverse Momentum Results</td>
<td>77</td>
</tr>
<tr>
<td>6</td>
<td>Differential $\phi^*$ Analysis</td>
<td>80</td>
</tr>
<tr>
<td>6.1</td>
<td>Differential Cross-section with $\phi^*$</td>
<td>80</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Z Boson Definitions</td>
<td>80</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Binning</td>
<td>83</td>
</tr>
<tr>
<td>6.2</td>
<td>Electron Efficiencies</td>
<td>84</td>
</tr>
<tr>
<td>6.2.1</td>
<td>ECAL Electron Efficiency</td>
<td>86</td>
</tr>
<tr>
<td>6.2.2</td>
<td>HF Electron Efficiency</td>
<td>89</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Final Efficiency</td>
<td>91</td>
</tr>
<tr>
<td>6.3</td>
<td>Background Estimation</td>
<td>94</td>
</tr>
<tr>
<td>6.4</td>
<td>Systematic Uncertainties</td>
<td>94</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Bin Migration Uncertainty</td>
<td>96</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Table of cell position correction values.</td>
<td>33</td>
</tr>
<tr>
<td>3.2</td>
<td>Parameters used for data and Summer11 Monte Carlo simulation in the transformation $E_{S/L} \rightarrow E_{\text{cor}}^{S/L}$.</td>
<td>41</td>
</tr>
<tr>
<td>3.3</td>
<td>Systematically optimized $C_{2d,m_{\text{cor}}}$ values used in the new two-dimensional requirement for the efficiency for electron reconstruction. The requirements are 50%, 75% and 95% efficiency correspond to the tight, medium and loose electron ID requirements.</td>
<td>41</td>
</tr>
<tr>
<td>4.1</td>
<td>A table of linearly independent color-anticolor superpositions for gluons, forming a complete color-octet (though other complete, linearly independent octets are equally correct). Colors represented are red ($r$), green ($g$), blue ($b$), anti-red ($\bar{r}$), anti-green ($\bar{g}$), and anti-blue ($\bar{b}$).</td>
<td>46</td>
</tr>
<tr>
<td>4.2</td>
<td>Selected decay modes of the Z boson and the experimental probability of each.</td>
<td>47</td>
</tr>
<tr>
<td>5.1</td>
<td>A table of values used for the working point requirements for WP80 for each ECAL detector as used for 2011 data.</td>
<td>69</td>
</tr>
<tr>
<td>6.1</td>
<td>The sub-detectors and corresponding electron pseudorapidity acceptance ($\eta_e$) used for reconstruction of electrons, along with the corresponding Z boson definition with their resulting $y$ ranges. A further requirement is that all electrons have $p_T &gt; 20$ GeV and the Z boson mass is between 60 and 120 GeV.</td>
<td>81</td>
</tr>
<tr>
<td>6.2</td>
<td>Monte Carlo Simulations used for signal, background estimation, and alternative tunes for signal (for use at the generator level).</td>
<td>83</td>
</tr>
<tr>
<td>6.3</td>
<td>The binning in $\phi^*$ with a the expected events per bin at a luminosity of 216pb$^{-1}$, determined from simulation at each generated acceptance level.</td>
<td>85</td>
</tr>
</tbody>
</table>
7.1 The Final Differential measurement for the EB-EB definition, with statistical and systematic uncertainties given in units of $\phi^*$. .......................... 110

7.2 The Final Differential measurement for the EE-EB definition, with statistical and systematic uncertainties given in units of $\phi^*$. .......................... 110

7.3 The Final Differential measurement for the EE-HF definition, with statistical and systematic uncertainties given in units of $\phi^*$. .......................... 112

7.4 The $\chi^2$/ndof for the various tunes of PYTHIA (and the solitary POWHEG+PYTHIA SignalMC tune). ......................................................... 113

A.1 Bin Migration Matrix for the EB-EB acceptance of Z bosons determined from simulation. ................................................................. 122

A.2 Bin Migration Matrix for the EE-EB acceptance of Z bosons determined from simulation. ................................................................. 123

A.3 Bin Migration Matrix for the EE-HF acceptance of Z bosons determined from simulation. ................................................................. 123

A.4 The Systematic uncertainties for the EB-EB Z boson acceptance, in units of $\phi^*$. ................................................................. 124

A.5 The Systematic uncertainties for the EE-EB Z boson acceptance, in units of $\phi^*$. ................................................................. 125

A.6 The Systematic uncertainties for the EE-HF Z boson acceptance, in units of $\phi^*$. ................................................................. 125

B.1 The Final Differential measurement for the ATLAS Comparison definition, with statistical and systematic uncertainties given in units of $\phi^*$. Also the ATLAS measurement and approximate uncertainty. ............... 126

B.2 The $\chi^2$/ndof between the ATLAS comparison definition and the ATLAS measurement (and the various tunes of PYTHIA) for the range of best agreement (the first thirteen bins, $\phi^*$ of 0.000-0.391) and the whole range ($\phi^*$ of 0.000-1.000). ................................................................. 127
List of Figures

2.1 A detailed site diagram of the LHC. The paths through initial proton accelerators (LINAC 2, BOOSTER, PS, and SPS) and the four detectors (CMS, ATLAS, LHCb, and ALICE) are shown, as well as the paths of other nearby or dependent experiments. ............................................. 5

2.2 A cut-away view of CMS with sub-detectors labeled. The Very-forward Calorimeter is more commonly referred to as the Forward Hadron Calorimeter. ................................................................. 6

2.3 A side view of the silicon tracking system showing the \( \eta \) location of the pixel track and the subsections of the strip tracker: Inner Barrel (TIB), Outer Barrel (TOB), Inner Disk (TID), and End Cap (TEC). .......... 11

2.4 A side view of a quadrant the Muon sub-detectors (DT, RPC and CSC) and their boundaries in \( \eta \). ........................................................................................................ 12

2.5 The structure and \( \eta \) boundaries of a quadrant of ECAL. ..................... 15

2.6 A side view of a quadrant of HCAL, showing the internal segmentation and boundaries (in terms of \( i\eta \)) .................................................. 17

2.7 A section of HF in the \( \eta - \phi \) plane. .................................................. 20

2.8 A view of a quadrant of HF from the side \( (r - z \) plane). The absorber contains the long and short quartz fibers which channel Cerenkov radiation into the photo-multipliers (PMT) in the readout boxes. The dashed vertical red line shows the depth at which the short fibers start. .......... 21
3.1 A 60° section of HF in r-φ space. The values along the upper and outer edges represent the \(i\eta\) and \(i\phi\) indexes respectively. The values along the lower edge represent the \(|\eta|\) values of the tower boundaries. The shading indicates the seed cell (yellow), 3x3 region (green), and 5x5 region (blue) for an EM cluster.

3.2 Plots of the generated position minus the reconstructed position for (a) \(\eta\) and (b) \(\phi\). The red dashed line represents the resolution of uncorrected position, while the solid black line represents the resolution after both the energy and cell position correction. Gaussians are fitted to each peak for reference and comparison. The small change in width is due to a majority of events occurring in the center of the cell, where the cell position corrections are an order lower than the width of the resolution.

3.3 A plot of the mean value for generated position minus the reconstructed position as a function of \(\log\frac{E}{(100 \text{ GeV})}\) for \(\eta\). The red line represents the piecewise correction function.

3.4 Plots of the mean value for generated position minus the reconstructed position as a function of Cell position for (a) \(\eta\) and (b) \(\phi\). The red line represent the fit from table 3.1. In the center a bias is seen due to electrons being preferentially reconstructed at the exact center of the the tower. As it is an order of magnitude smaller than the overall position resolution, it is not a noticeable effect on the final HF electron.

3.5 A plot displaying the average difference between the reconstructed electron \(E_T^{\text{reco}}\) and predicted electron \(E_T^{\text{pred}}\) as a function of the number of pileup events (vertex count), at an \(|\eta|\) of 32, using 2011 data. The red dashed line represents a fit whose slope is \(m_{\eta\phi} (\approx 200 \text{ MeV per vertex})\). \(E_T^{\text{pred}}\) is a prediction using an ECAL electron and the mass of the Z boson. In the equation: \(M_Z\) is the mass of the Z boson, \(E^{\text{ECAL}}\) is the measured energy of the matched ECAL electron, \(\eta^{\text{ECAL}}/\phi^{\text{ECAL}}\) are the \(\eta/\phi\) of the matched ECAL electron, and \(\eta^{\text{HF}}/\phi^{\text{HF}}\) are the \(\eta/\phi\) of the HF electron.
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.6</td>
<td>Plot of the $E_{9/25}$ isolation variable. “Data:Signal” (black line) and “Monte Carlo:signal” (Summer11, the green line) are contrasted to “Data:Background” (red shaded region). Excess “Monte Carlo:signal” over “Data:Signal” in the highest bin is due to less pileup present in the simulation, but otherwise there is agreement between data and simulation.</td>
</tr>
<tr>
<td>3.7</td>
<td>Plot of the $E_{C/25}$ transverse variable. “Data:Signal” (black line) and “Monte Carlo:signal” (Summer11, the green line) agree and are contrasted to “Data:Background” (red shaded region).</td>
</tr>
<tr>
<td>3.8</td>
<td>Plot of the $E_{S/L}$ longitudinal variable. “Data:Signal” (black line) and “Monte Carlo:signal” (Summer11, the green line) agree and are contrasted with “Data:Background” (red shaded region). The overlap between signal and background is greater than in other variables, due to high electromagnetic content in background events. This overlap strongly motivates using the two-dimensional method rather than a ‘square cut’ in the two variables.</td>
</tr>
<tr>
<td>3.9</td>
<td>(a) Plot of the uncorrected $E_{S/L}$ vs $\log \frac{E_L}{(100 , \text{GeV})}$ (b) plot of $E_{cor}^{S/L}$ vs $\log \frac{E_L}{(100 , \text{GeV})}$. The dots represent the mean value for the Y-axis in each plot. An energy dependence is seen in (a) and removed in (b).</td>
</tr>
<tr>
<td>3.10</td>
<td>A plot of $E_{cor}^{S/L}$ vs $E_{C/9}$ (“the 2D requirement”) for electrons. The red line represents a “Loose” requirement, where all events below the line are accepted.</td>
</tr>
<tr>
<td>4.1</td>
<td>The standard model</td>
</tr>
<tr>
<td>4.2</td>
<td>A sample PDF for a proton at 10 GeV</td>
</tr>
<tr>
<td>4.3</td>
<td>Feynman diagrams showing (a) $s$-channel $Z \rightarrow e^+e^-$ decay, (b) $t$-channel $Z$ boson production, and (c) $Z$ boson production diagram with initial state radiation.</td>
</tr>
<tr>
<td>4.4</td>
<td>Comparison of the growth of the terms of $Q_T$ within the perturbation on an arbitrary scale.</td>
</tr>
<tr>
<td>4.5</td>
<td>A view of a $Z \rightarrow e^+e^-$ event in the transverse plane of the lab frame. Novel and related variables are indicated.</td>
</tr>
</tbody>
</table>
4.6 Toy Monte Carlo Z bosons with comparison of generated values to smeared reconstruction for (a) $Q_T$ and (b) $\phi^*$ over 10,000 events. This demonstrates that $\phi^*$ variable is less effected by detector resolution, especially in the 'non-perturbative' (low $Q_T/\phi^*$) region. The plots of $Q_T$ and $\phi^*$ are over equivalent ranges.

5.1 The spectrum of di-electron events seen in the CMS experiment in 2010. Different peaks correspond to the labeled particle.

5.2 The bin migration for the differential $Q_T$ measurement. Note that it is especially nontrivial at low values where $\phi^*$ provides a much higher resolution.

5.3 The differential cross-section for Z bosons as a function of absolute rapidity. The combined results extends only to $|Y| < 2$. Results beyond this point are electron only, using HF electrons. The shaded area represents the predicted simulation with the the uncertainties of the CT10 PDFs.

5.4 The change in $\chi^2$ values between the differential rapidity results and predictions from the NLO HERAPDF 1.5 PDF set for the eigen vectors (upper plot) and model parameters (lower plot). The greatest disagreement between the two is in the modeling of strange-quark PDF as a fraction of the down-quark-sea PDF (that is, the relative amplitude of strange quarks in the sea). Other PDFs compared in the paper (but not shown here) were CT10 PDF, MSTW2008, and NNPDF.

5.5 The combined differential cross-section for transverse momentum, focusing on $Q_T < 30$ GeV. The colored lines represent various simulation configurations that have been used. The lower plot represents the comparison of these simulations to data divided by the uncertainty ($\delta$) on the data. The outer and inner bands represent $\pm 1\delta$ and $\pm 2\delta$ respectively.

5.6 The combined differential cross-section for transverse momentum, focusing on $Q_T > 20$ GeV. The shaded regions represent proposed simulation configurations. The lower plot shows the ratio between the data and the theory predictions. The bands in the lower plot represent the one standard deviation combined theoretical and experimental uncertainties.
6.1 The number of events per pile-up for the 2011A DoubleElectron 216pb⁻¹ data set.

6.2 A typical efficiency calculation plot used in the WP80-data efficiency calculation for probe electrons with $p_T$ between 45 and 50 GeV and a pile-up less than 5. The graph is represents the data (the circle points) matched with a combined fit (the black solid line) made up of combined background (black dashed line) and signal (red dashed line). The graph on the left is before the WP80 requirements are made, the one on the right is after. The number of electrons in each graph is taken to be the average between the integral of the signal and the number of events minus the integral of the background. The single electron efficiency for this bin is the ratio of number of electrons on the right to the number of electrons on the left.

6.3 The correction factors for the Trigger Efficiency stage of ECAL electron efficiency calculation, binned in $p_T$ and $\eta$ for (a) low (0-4) and (b) high (5-100) pile-up.

6.4 The correction factors for the Electron Identification stage of ECAL electron efficiency calculation, binned in $p_T$, $\eta$ for (a) low (0-4) and (b) high (5-100) pile-up.

6.5 The total correction factor applied to ECAL electron efficiency to match the simulated efficiency to experimental data, binned in $p_T$, $\eta$ for (a) low (0-4) and (b) high (5-100) pile-up.

6.6 The efficiency for HF TID requirements, binned in $\phi^*$ and $\eta$ for (a) low (0-4) and (b) high (5-100) pile-up.

6.7 The efficiency for HF electron ID requirements, binned in $p_T$ and $\eta$ for (a) low (0-4) and (b) high (5-100) pile-up.

6.8 Final cumulative efficiencies per stage for (a) EB-EB, (b) EE-EB, and (c) EE-HF in $\phi^*$.

6.9 Estimated background as a fraction of total events per bin in the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions.
6.10 The statistical fractional uncertainties on the uncertainty calculations made using the SignalMC simulation set for the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions. .............................................. 97

6.11 The systematic uncertainties from bin migration for the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions. .............................................. 99

6.12 The systematic uncertainties from efficiencies per stage for the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions. .............................................. 100

6.13 The statistical uncertainties from efficiency per stage for the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions. .............................................. 102

6.14 Uncertainties on the final measurement for the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions due to energy scale uncertainties in ECAL and (d) for EE-HF Z bosons due to energy scale uncertainties in HF. .............................................. 104

6.15 Uncertainties on the final measurement for the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions due to positional uncertainties in ECAL and (d) for EE-HF Z bosons due to alignment uncertainties in HF. .............................................. 105

6.16 Comparisons of the contributing sources of uncertainty for the EB-EB Z boson definition. The Total Systematic (black line) is the sum in quadrature of the other systematic uncertainties (the colored, solid lines). ........ 106

6.17 Comparisons of the contributing sources of uncertainty for the EE-EB Z boson definition. The Total Systematic (black line) is the sum in quadrature of the other systematic uncertainties (the colored, solid lines). ........ 107

6.18 Comparisons of the contributing sources of uncertainty for the EE-HF Z boson definition. The Total Systematic (black line) is the sum in quadrature of the other systematic uncertainties (the colored, solid lines). ........ 108

7.1 The final results of the differential measurement in the (a) EB-EB, (b) EE-EB, and (c) EE-HF channels. The red line is the SignalMC simulation for comparison. .............................................. 111

7.2 A comparison of of the ratios of multiple PYTHIA tunes (and data) to the POWHEG+PYTHIA Z2 simulation in EB-EB the Z boson definition .............................................. 114

7.3 A comparison of of the ratios of multiple PYTHIA tunes (and data) to the POWHEG+PYTHIA Z2 simulation in EE-EB the Z boson definition .............................................. 115
7.4 A comparison of the ratios of multiple PYTHIA tunes (and data) to the POWHEG+PYTHIA Z2 simulation in EE-HF the Z boson definition.

7.5 A comparison of the ratios of the ATLAS Comparison definition and ATLAS measurement to the POWHEG+PYTHIA Z2 simulation (as well as the other tunes). Note the strong agreement in all but the last two bins.

B.1 (a) The final measurement of the ATLAS Comparison definition with the ATLAS measurement and Summer11 shown for comparison. Further plots of the ATLAS Comparison definition for (b) the final cumulative efficiencies per stage, (c) the estimated background as a fraction of total events per bin, and (d) the systematic uncertainties from bin migration.

B.2 Plots of the ATLAS Comparison definition for (a) the systematic uncertainties from efficiencies per stage, (b) the statistical uncertainties from efficiency per stage, (c) the uncertainties on the final measurement due to energy scale in ECAL, and (d) the uncertainties on the final measurement for due to positional uncertainties in ECAL.
Chapter 1

Introduction

1.1 Introduction

High energy particle physics seeks to describe the types of fundamental particles and their interactions with each other. The current compilation of proven and accepted knowledge on this subject is the Standard Model. While the Standard Model does not explain everything that is seen (the presence of dark matter, for example) or hoped to be seen (Supersymmetry, or perhaps right-handed heavy neutrinos), it provides a reliable basis from which future progress can be staged from. The discovery of the Higgs boson at the Large Hadron Collider (LHC) marks another triumph of the Standard Model in both predicting the existence Higgs bosons and understanding the signal of the Higgs boson when it was seen. While discovering new particles is perhaps the most glamorous activity for the Standard Model and high energy physics, it is the hard work of understanding the Standard Model in as full, precise detail as possible that makes this possible. To this end, a preliminary step in searching for the new is understanding the old. One ‘old’ piece of the Standard Model that merits further investigation is the Z boson. It has a high rate of production and can produce relatively clear signals of electrons or muons within a general purpose particle detector for a hadron collider, such as the LHC. Detailed investigation of the behavior of the Z boson provides insight not only on this one particle, but on the particles that participated in creating it. The very nature of proton-proton collisions makes the kinematics of a created Z boson a window into the complex behavior and probabilities within the protons during high energy collisions. In
particular, the transverse momentum given to a $Z$ bosons can be measured to further the understanding of complicated predictions made by theory and investigate the low end of this spectrum where theory and experiment do not yet agree.

1.1.1 Breakdown of the Thesis

The goal of this thesis is to perform a differential cross-section measurement of $Z \rightarrow e^+e^-$ with respect to $\phi^*$, a novel variable closely related to transverse momentum. There will be a specific focus on the regions analogous to non-perturbative regions of transverse momentum. Accurately matching this measurement to theoretical models will present information on the behavior of this region that will inform theory and simulations of particle collisions.

To begin, Chapter 2 describes the experiment and apparatus that provided the data to be studied, the LHC and the Compact Muon Solenoid (CMS), along with their general abilities and relevance to this thesis. Continuing from the presentation of the CMS experiment, Chapter 3 presents a CMS specific method of detecting electrons in the outer limits of the detector. This method is an original work of this author and will be used within the final differential measurement. Chapter 4 provides theoretical information, beginning with a general overview of high energy physics before focusing on the specific physics behind the transverse momentum and introducing the novel variable $\phi^*$ that will be the target of the differential measurement.

Chapter 5 describes how a differential measurement is set up in general and presents previous differential measurements of rapidity and transverse momentum made with CMS data. The actual $\phi^*$ differential measurement will be the focus of Chapter 6, discussing in detail the methods of obtaining data, rejecting background, and determining systematics. Finally, Chapter 7 will present the final measurements, along with an analysis of comparison to theory and conclusions on the results.
Chapter 2

The LHC and CMS

2.1 The LHC

The Large Hadron Collider (LHC) (fig. 2.1) is the largest synchrotron particle accelerator ever completed [1]. It is a large ring of magnets and RF cavities, 27 kilometers in circumference, that accelerates and collides groups of protons for scientific research. The LHC was built by the European Organization for Nuclear Research (commonly named CERN) in a circular tunnel $\approx 100$ m below the surface of the earth for the purpose of studying high energy particle collisions. Using superconducting magnets as guides, large groups, or “bunches”, of protons (over 100 billion in a group) can be accelerated to momenta of up to 7 Tera-electron Volts (TeV), though in the data of this thesis the actual momenta of single protons were 3.5 TeV. These proton bunches pass through other bunches traveling in the opposite direction at four locations along the LHC at a maximum rate of one crossing every 25 nanoseconds. Despite the great numbers of protons, only a few collisions take place every crossing. At these crossing points are the main experiments of the LHC, used to observe the results of the proton-proton collisions. These experiments are the Compact Muon Solenoid 2 (CMS), A Toroidal LHC Apparatus 3 (ATLAS), the LHC-beauty 4 (LHCb) and A Large Ion Collider Experiment 5 (ALICE). CMS is the experiment utilized in this paper, and is located at the halfway point around the LHC from its technical origin (where ATLAS is housed).

Due to the size, curvature, and mechanics of the LHC, it cannot start with protons at rest and accelerate them up to the collisional energy. Proton synchrotron accelerators
have a minimum and maximum momentum they can circulate, set by the radius of the accelerator and the highest and lowest stable magnetic fields achievable by their dipole magnets. To get a proton up to LHC energy requires a trip through several different accelerators first. The protons begin initial acceleration in a linear accelerator (LINAC 2) which gets them up to 50 MeV. The protons then enter three synchrotron pre-accelerators of increasing radius and energy: the Proton Synchrotron Booster (PSB) accelerates up to 1.4 GeV, the Proton Synchrotron (PS) accelerates up to 25 GeV, and the Super Proton Synchrotron (SPS) accelerates up to 450 GeV and injects the protons (both clockwise and counter-clockwise) into the LHC. The LHC then accelerates the protons for several minutes to get them to collision energy. Bunches of protons can be kept in the ring for ten hours or more at a time.

There are other experiments that use the LHC for part or all of their activities. Lead ion collisions are also performed in the LHC, the difference only in initial pre-acceleration path (LINAC 3 and Low Energy Ion Ring (LEIR)) and final energy within the LHC (2.64 TeV per nucleon pair or 574 TeV total for the ion). CERN also produces neutrinos from the SPS beam which are sent to Gran Sasso, Italy to neutrino experiments such as ICARUS [6] and OPERA [7].

2.2 CMS

CMS is a general purpose detector designed to detect the presence of as many types of fundamental particles with as much physical coverage as possible (fig. 2.2). While the particles CMS is designed to detect are sub-atomically small, they are also extremely energetic, approaching the speed of light in velocity. A near-relativistic particle can travel a long distance, despite having a short lifetime. In order to observe high energy particles, a detector must either be very large or very dense, often both. CMS is 15 meters in diameter and 21 meters long and weighs 12,500 tonnes, making it smaller (but much denser) than the other general purpose experiment, ATLAS (25 meters in diameter, 46 meters in length, and 7,000 tonnes). This smallness in size with comparison to ATLAS (and other proposed detectors) motivates the name ‘Compact’ in CMS.
Figure 2.1: A detailed site diagram of the LHC. The paths through initial proton accelerators (LINAC 2, BOOSTER, PS, and SPS) and the four detectors (CMS, ATLAS, LHCb, and ALICE) are shown, as well as the paths of other nearby or dependent experiments.
Figure 2.2: A cut-away view of CMS with sub-detectors labeled. The Very-forward Calorimeter is more commonly referred to as the Forward Hadron Calorimeter.
2.2.1 CMS Geometry

The distinctive shape of CMS is a cylindrical prism with a dodecagonal (12-sided) outer cross-section with the length of the cylinder parallel to the beam, this being the geometry of the outer most sub-detector, the muon chambers. The solenoid is a true cylinder in shape, while the inner-most sub-detectors are octadecagonal (18-sided) in cross-section. Because of these roughly circular cross-sections, cylindrical co-ordinates are used to describe locations within the experiment. The origin of the co-ordinates is located at the center of the beam line within the experiment, at the nominal collision point. Using a right handed co-ordinate system, the \( z \)-axis points along the beam line, the \( y \)-axis points vertically up and the \( x \)-axis points toward the center of curvature of the LHC. The standard azimuthal angle in the \( x-y \) plane, \( \phi \), starts at zero at the \( x \)-axis, measured in radians. The polar angle, \( \theta \), is measured from the \( z \)-axis, though it is not commonly used. Instead, psuedorapidity (defined as \( \eta = -\ln(\tan(\theta/2)) \)), is used to describe positions along the \( z \)-axis. In this variable, \( \eta=0 \) corresponds to a position perpendicular to the beam at \( z = 0 \), while an \( \eta \) of infinity corresponds to a particle traveling within the beam itself. The \( \eta \) variable is the common co-ordinate used when defining boundaries within experiment with a majority of the sub-detectors covering a region out to \( |\eta| < 2.5 \). Two important pseudo-coordinates are \( i\eta \) and \( i\phi \). These are integer valued variables that correspond to \( \eta/\phi \) sections (respectively) within a detector, most often aligning with a distinct tower or module.

When looking from a side view one notices that most sub-detectors have two distinct regions. The first is a barrel region, which is where the detectors are a constant radius from the beam line, and the second is the end-cap region where the detectors act to close off the ends of CMS. For example, in the previously mentioned muon detectors the barrel of the muon detector is at \( |\eta| < 1.2 \), while the end-caps cover the region of \( 1.2 < |\eta| < 2.4 \).

Transverse Variables

As will be discussed at later points in this thesis, the total momentum of a collision along the \( z \)-axis is an unknown, albeit constrained, quantity, but momenta in perpen-
dicular (transverse) directions should have a net value of zero. This symmetry causes
the transverse components of quantities such as energy ($E$) and momentum ($Q$) to be important observables. The transverse energy ($E_T$) and transverse momentum ($Q_T$) can be calculated by:

\[
E_T = \frac{E}{\cosh(\eta)} \quad Q_T = \frac{Q}{\cosh(\eta)}
\]

In this thesis, the focus will be on the transverse momentum of Z bosons.

### 2.2.2 Particle Detection

Detection and reconstruction of particles is the sole purpose of the CMS apparatus. It is designed for the direct detection of light or stable particles, and the indirect detection of heavy or unstable particles through the light or stable particles they decay into. The only currently-known elementary particle that does not have a clear response in CMS is the neutrino.

An imbalance in transverse variables can provide a clue that neutrinos were present, although much information (e.g. full energy, trajectory, and total number of neutrinos) is not detectable.

The focus of this paper will be on the indirect detection of Z bosons through direct electron detection. The sub-detectors used for electron detection are the Silicon Tracker, the Electromagnetic Calorimeter, and the Forward Hadron Calorimeter.

### 2.3 The Tracking Systems of CMS

Tracking detectors are designed to follow the path of a particle as it travels through the detector, but not to significantly interact with the particle or disturb its travel. Tracking detectors are sensitive only to electrically-charged particles as the presence of electric charge is easy to detect with minimal interaction. Neutral particles, such as photons, neutrinos, and neutral kaons, do not show up within a tracking detector. Within a magnetic field, the path of a charged particle will curve inversely proportionally to the amount of momentum perpendicular to the field. The charge of a particle can be determined by the direction of curvature: a positive charge will curve counter-clockwise...
in the plane transverse to the magnetic field, clockwise for a negative charge. The momentum of a particle \((p)\) is determined by the radius of curvature \((r)\),

\[
p = \frac{qB}{r}
\]

\[
\frac{r}{L} \sim \phi_B
\]

\[
p = \frac{qLB}{\phi_B}
\]

where \(q\) is the charge of the particle, \(B\) is the magnitude of the magnetic field, \(L\) is the length of the path, and \(\phi_B\) is the small angle approximation of the deflection angle from the transverse distance traveled, nearly equivalent to \(L\) at higher momenta. In simple terms, as momentum increases, the curvature becomes increasingly straight along length \(L\). Resolution within trackers is measured in terms of positional error \((d\sigma)\) such that

\[
d\sigma \sim d\phi_B L
\]

\[
= d\left(\frac{1}{p}\right)
\]

\[
\frac{d\sigma}{dp} = \frac{qL^2B}{p^2}
\]

\[
\frac{dp}{p} = \frac{pd\sigma}{eL^2B}
\]

This demonstrates that as momentum increases, fractional momentum resolution degrades, as \(e, L, B,\) and \(d\sigma\) are fixed value (or nearly so, in the case of \(L\)). This effect can be minimized by providing large magnetic fields and long paths of flight.

In CMS, tracking is made possible by the large field provided by the solenoid, and the tracking detectors are the silicon tracker and the muon detectors.

### 2.3.1 Solenoid

The central feature of the detector is the solenoid, referenced in the name of CMS. It is a powerful superconducting magnet that creates a 3.8 Tesla field parallel to the colliding beams. The current flowing through the super-cooled niobium-titanium coils creating the field is 18,160 A, which causes it to store 2.3 GJ of energy, enough to boil 6 tons of ice water. Extensive cooling systems protect the solenoid from damaging itself and surroundings in case of rapid discharge of this energy. The return yoke for the
magnetic field is within the muon systems. The solenoid is the largest non-detecting system within CMS.

2.3.2 Silicon Tracker

When protons collide, the resulting particles travel radially into CMS. The first detector they encounter is the silicon tracker (fig. 2.3), a tracking detector made of silicon pixels (inner radius) and silicon strips (outer radius). As mentioned, the intense magnetic field ensures that charged particles follow a curved path that will be seen in the tracker, measuring electric charge and momentum. The particles leave “hits” which can be combined into tracks via software. These tracks are the first signals seen by CMS of a particle, and can be matched to other signals in other subdetectors. For electrons, the silicon tracker provides a track that matches the location and momentum of a signal from the within the Electromagnetic calorimeter. Due to the high probability of an electron radiating photons in the tracker, special reconstruction software is needed to properly find the path it follows.

The tracker has spatial resolution of the order of 10 $\mu$m and has better than 99% efficiency in spatial reconstruction. It has an energy resolution that is 0.1% per GeV/c of transverse momenta measured [8]. The tracker covers the region of $|\eta| < 2.5$, giving coverage of the majority of interesting events, while avoiding regions of dangerous radiation.

2.3.3 Muon Systems

The eponymous Muon detectors form the outer shell of CMS (and the middle of its name), as muons are able to travel through the entire radius of CMS. While most particles are very short-lived and/or highly interactive with detector material, muons have a lifetime long enough to travel though the entirety of CMS without decaying and a suppressed likely-hood to radiate or interact while within. Fortunately, a muon has an electric charge, so sensitive tracking electronics can still detect the presence of a muon.

The Muon detector consists of three different sub-detectors (fig. 2.4): Drift Tubes (DT, located in the barrel region), Resistive Plate Chambers (RPC, located in the barrel and end-cap, $|\eta| < 1.6$), and Cathode Strip Chamber (CSC, located in the end...
cap, $1.2 < |\eta| < 2.4$). The DTs and CSCs are designed for high resolution of position (on the order of 100 $\mu$s), while the RPCs provide high timing resolution (on the order of 1 ns). Between the muon chambers are iron plates that serve as the return yoke of the solenoid. Because of this, the magnetic field in the muon detectors is quite low, especially in the barrel region, and muon responses are straight vectors, not curved as in the silicon tracker. Muons do curve while in the iron, but with opposite bending than while within the solenoid as the magnetic field has reversed directions. The result is that in each section of the detector the path of the muon is bent by the interspersing iron, so there is still a curvature with which to calculate momentum and charge. Typically, a track in the muon systems can be matched to a track of similar momentum, path, and charge within the silicon tracker, providing more precise information on the muon.
2.4 The Calorimeters of CMS

The tracker and muon detectors are similar in that they are designed to interact only minimally with a particle, but to map the path of the particle. Immediately outside the silicon tracker are calorimeters which are designed to interact with particles and to measure their energy by doing so. CMS has two types of calorimeters for two types of particles which have different characteristic shower depths. The first calorimeter is the electromagnetic calorimeter (ECAL) which is designed to capture the energy of the short length showers of electromagnetic (EM) particles: photons and electrons. The second calorimeter is the hadronic calorimeter (HCAL), which is designed to contain
and measure the much longer showers of hadronic particles.

Calorimeters measure energy. In particle physics, the method of energy measurement relies on high energy particles forming showers of many less-energetic particles that will interact with an activating material. As long as the initial energy of a particle ($E_0$) is above a certain critical energy ($E_c$) it is more likely to shower into $n$ daughter particles than deposit its energy into the calorimeter. The characteristic length a particle travels before showering is determined by the particle type and detector material. For an electromagnetic particle this is a radiation length, for a hadron it is an interaction length. Assuming an electromagnetic particle (as electrons are the focus of this paper), a simple but useful model can be built by taking $n = 2$ and $E_c$ is the energy below which electrons will not radiate photons and photons will not pair produce electron-positron pairs (both of which are determined by the Bremsstrahlung cross-section). With this assumption the number of particles ($N$) and energy of one particle ($E_p$) in the $t^{th}$ generation (i.e. the number of radiation lengths traveled) of the shower are

$$N = 2^t$$
$$E_p = \frac{E_0}{N}$$

When the energy of a particle reaches $E_c$, the shower has reached its maximum number $N_{max}$.

$$E_c = \frac{E_0}{N_{max}}$$
$$N_{max} = \frac{E_0}{E_c}$$

At this point the number of particles is linearly related to the initial energy. As each particle below $E_c$ gives off the last of its energy through ionization, the light read out from the resulting scintillation provides a count of each one. As the error on the final number of particles is stochastic and thus $\sigma(N) \sim \sqrt{N}$. As $E_0$ is linear with $N$,

$$\frac{\sigma(N)}{N} \sim \frac{1}{\sqrt{N}} \Rightarrow \frac{\sigma(E)}{E} \sim \frac{1}{\sqrt{E}}$$

and thus the fractional error of measured energy decreases with higher energies (compare to tracking detectors which increases in error as energy/momentum increases).

With a hadronic particle, $n$ varies from generation to generation, with $\frac{2}{3}n$, on average, being charged hadrons that continue the shower while $\frac{1}{3}n$ are neutral hadrons that
rapidly decay electromagnetically, producing showers which proceed as described above. The shower continues to deposit one third (on average, with large variation) of its remaining energy electromagnetically during each interaction length, until the remaining charged hadrons have energy less than $E_c$. In the end, the primary energy seen is that which is deposited by the neutral hadrons (in total, $E_n$). The fraction $E_n/E_0$ depends on the the maximum number of interaction depths which in turn depends on the $n$-value for each generation, but as energy increases, it approaches one. Further information on hadronic showers and calorimetry can be found elsewhere [9].

2.4.1 The Electromagnetic Calorimeter

ECAL is made of lead tungstate, a heavy clear crystal chosen for its density and short radiation length. Electrons and photons that travel into these crystals will radiate approximately once every 0.89 cm traveled. As mentioned previously, for a photon this means splitting into an electron-positron pair, while for an electron this mean radiating a photon. These produced particles then radiate further and an electromagnetic shower of particles occurs. As the particles reach their maximum shower-depth, they ionize the crystals, which causes scintillation. The light from the scintillation travels through the crystals to Avalanche Photo-Diodes (APDs) in the barrel region and Vacuum Photo-Triodes (VPTs) in end-cap region, which convert the light into electrical signals that are sent out to be stored as data. In total there are more than 75,000 lead tungstate crystals in ECAL. ECAL is comprised of two sub-detectors, the ECAL Barrel (EB) which covers a region of $|\eta| < 1.479$ and the ECAL End-caps (EE) at $1.479 < |\eta| < 3.0$ (fig. 2.5).

Separating Electrons from Photons

An electron entering ECAL showers electromagnetically (as previously described). As described earlier, photons shower by pair producing electrons and electrons shower by radiating photons. In this way, a few generations into a shower of either particle will be essentially identical to the other, as will the resulting scintillation seen by the readout of ECAL. This is where the synergy between calorimeters and tracking detectors becomes apparent. An electron will leave a curving track in the silicon tracker, allowing charge identification and momentum measurement, and will point to a energy cluster in
Figure 2.5: The structure and η boundaries of a quadrant of ECAL.
ECAL which should correspond to the measured momentum (a photon will not leave any track). Thus while ECAL alone has very good measurement of energy of electrons and photons, the tracker is need to to distinguish electrons (and positrons) from photons. A further complication can come from an electron or photon that radiates before entering ECAL. For example, if an electron radiates a photon before hitting ECAL, some of the energy of the electron will go to the photon and, while the electron will continue to curve in the magnetic field, the photon will travel straight from its emission point. In this case, two signals in ECAL will be observed, each with lower energy than the track seen in the tracker. Because of this, electrons appear as a group of signals in ECAL that lie along a path of constant $\eta$ but changing $\phi$ (the magnetic field only exerts a force in the $\phi$ direction).

2.4.2 The Hadronic Calorimeter

The main section of HCAL is made of brass and stainless steel absorber plates and plastic scintillators (fig. 2.6). The absorber plates interact with hadronic particles over a distance much greater than the radiation length of EM particles. Thus while most of an EM shower will be contained in ECAL, a hadron has only about a 50% chance of interacting once before it reaches HCAL. In the brass plates the hadrons form a shower of particles, and as they pass through the scintillators the number of shower particles (which is proportional to the energy of the shower) is read out as light.

The HCAL detector contains the HCAL Barrel (HB, $|\eta| < 1.4$) and HCAL End-cap (HE, $1.4 < |\eta| < 3.0$). HCAL also has two other sub-detectors, the Outer HCAL (HO) and the Forward HCAL (HF). HO sits on the outer side of the solenoid, and observes showers that penetrate deep enough within CMS to pass through the solenoid. HO is mechanically the same as HB and HE, except the solenoid is used for absorber material rather than brass. The HF sub-detector, while considered part of the HCAL system, is a unique detector in design and function.

2.4.3 The Forward Hadron Calorimeter

The HF detector consists of two end-cap detectors that sit at the highest $\eta$ regions of CMS in HF (2.8 < $|\eta| < 5.2$, referred to as HF Plus and HF Minus in the corresponding
Figure 2.6: A side view of a quadrant of HCAL, showing the internal segmentation and boundaries (in terms of $i\eta$).

$\eta$ region. These regions are full of intense radiation, so HF needs to be very resistant versus radiation. Due to the high radiation, there is no tracking detector in front of HF (this means there is no charge information or EM filtering before a particle enters the detector). The detector is made of stainless steel and quartz fiber and relies on reading out Cerenkov radiation from the shower for the detection of energy. Cerenkov light is produced when a particle travels through a material at a velocity ($v$) faster than the speed of light in a material ($c/n$, where $c$ is the speed of light in a vacuum and $n$ is the index of refraction for the material). For quartz ($n = 1.5$) this occurs for particles traveling above two-thirds the speed of light, a speed an electron can reach with a trivial amount of energy but requiring much higher energies for particles with higher mass. In
HF, Cerenkov light is dominantly from electrons. In the case of a hadronic shower, these electrons are from the neutral hadron decays and the principal for determining initial energy closely follows that describe for hadron calorimeters. In the case of an electron or photon, the electrons in the shower will be seen, while the photons will not. Because some of the shower will occur at points where there is no quartz fiber, only a sample of the total shower will be seen. This effect is correctable by testing particles of known energy in the detector and finding the appropriate scaling factor. The scaling of the energy response in HF is tuned for 100 GeV charged pions to have a read-out energy of 100 GeV. The quartz fibers come in two lengths as they run the longitudinal length of the detector: long fibers (1.65 m) begin at the face of the detector (where particles enter), and short fibers (1.43 m) which start 22 cm from the detector face. The fibers are spaced transversely 55 mm apart within the detector and alternate types so that a long fiber is only adjacent to short fibers in any direction you go, and vice-versa. Each HF is made up of eighteen wedges, with each wedge containing twenty-four towers. One tower takes up 0.175 units in both \( \eta \) and \( \phi \) space, except for the highest \( \eta \) towers which are twice as wide in \( \phi \) (fig. 2.7). The towers sections are parallel to the beam line, as are the quartz fibers, and thus are not quite projective along the path of a particle coming from the center of CMS. This is done so the particles cannot pass through the gaps between fibers and fail to be observed.

**Long and Short Fibers**

The difference in the length of the quartz fibers is so that HF can gather information on both hadronic showers and electromagnetic showers. The long quartz fibers reach all the way to the face of the detector where incoming particles will arrive (fig. 2.8). When an EM particle hits HF and starts to shower, only the long fibers will pick up that energy. A majority of the energy of the shower will be measured in the first few centimeters of the detector, so when the short fibers start 10 radiation lengths later, a much smaller fraction of the energy will be measured. Thus, for EM showers there will be a noticeably higher amount of energy in the long fibers than in the short fibers. In the case of hadronic showers, the 22 cm gap is only one interaction length into the stainless steel, that means the start of the hadronic shower should nearly coincide with the start of the short fibers. Thus the the energy seen in short fibers should be roughly
equal to the energy in the long fibers. In this way there is basic discrimination between
electromagnetic and hadronic showers.

**Electronic Read-out**

The HF has a total of 1728 channels through which data is read out, one for each length
of fiber in each tower. The collected Cerenkov light is channeled along light guides into
Photo-Multiplier Tubes (PMTs) that are located at the back of the detector in the ‘top’
(lowest \(|\eta|\)) 60 cm in readout boxes. From the readout box, information is transmitted
to data acquisition. Much of the rest of the material in HF is shielding that tries to
prevent the immense radiation at this \(|\eta|\) from leaving the detector (see again fig. 2.8).

2.4.4 Triggers

The optimal rate of proton collisions at the LHC is 40 MHz (one proton bunch crossing
every 25 ns). When data from an event is stored, the information of hundreds of
thousands of electronic channels must be written to a storage disk. The practical rate
for writing that much information is around 1 kHz. Furthermore, a majority of events
will not be useful for the types of analyses the CMS collaborators are focusing on. In
order to only spend processing time storing data from events that have are potentially
the most interesting, a triggering system is set in place to keep desired data, but ignore
the uninteresting events. The triggering system is split into two parts, the Level-1
trigger (L1) and the High Level Trigger (HLT).

**Level-1 Trigger**

The L1 system is a fast set of custom electronics that looks at broad patterns, called
‘trigger primitives’, that are easily computed and sorted in a short amount of time.
Trigger primitives are specific to each sub-detector (excluding the silicon tracker) and are
passed into a series of trigger groupings/tests, culminating in a global trigger, which has
final say in whether an event passes L1. Trigger objects include electromagnetic objects
(from ECAL, with no track information), muon objects (from the muon chambers only),
and jet objects (using fixed \(\eta/\phi\) regions in the calorimeters). The L1 trigger is designed
to reach a pass or fail conclusion as rapidly as possible, keeping the maximum rate of
Figure 2.7: A section of HF in the $\eta - \phi$ plane.
Figure 2.8: A view of a quadrant of HF from the side (r – z plane). The absorber contains the long and short quartz fibers which channel Cerenkov radiation into the photo-multipliers (PMT) in the readout boxes. The dashed vertical red line shows the depth at which the short fibers start.
passing events to 100 kHz. If a set of trigger objects fails L1 acceptance, all the data from the event is ‘thrown-out’, never to be seen. If an event passes the global trigger, it sends a Level-1 Accept (L1A) signal and the event is passed to the next level of triggering.

**High Level Trigger**

The HLT system is a large set of filtering triggers that works on the full data from an event. By running in parallel, multiple triggers can simultaneously look at the same data and more detailed parameters can be looked at. Each path is required to take less than 10 ms to reach a pass/fail decision, but as long as one path passes, the data is kept. The final data rate of the HLT (and by extension, storage of data) is between 200-500 Hz.

If a trigger is found to activate too often, then only a sampling of that trigger is looked at. This method, called prescaling, allows only a random selection of the trigger be accepted, such that the statistical effect is only the desired percentage make it through the trigger. A trigger with a prescale of 100 would only let 1-in-100 events actually be accepted.

**Electron Triggers**

For the purpose of this thesis, the relevant triggers are the electron triggers. In L1, ECAL and HF provide input to trigger primitive generators, which pass to a regional calorimeter trigger, and a global calorimeter triggers before receiving the L1A from the global trigger. These triggers are very efficient at accepting events with electrons, but only in the ECAL region. To have a forward-calorimeter L1 trigger on electrons, jet primitives and triggers must be used. A jet primitive is a 2x3 cluster in $\eta/\phi$ space, and 3x3 cluster of these primitives makes up the global jet trigger in HF. As will be discussed in the next chapter, the expected size of an electron in HF is contained within a 3x3 $\eta/\phi$ cluster. Thus, a L1 trigger on an HF electron will be based on a region at least six times larger than needed. In such a region even a fairly isolated electron could be a minority contributor to the energy response. Currently, accuracy with a L1 trigger is low for HF electrons (this will be accounted for in the analysis when using them).
In HLT, single electron triggers often need to be heavily prescaled, as they are very common. To maximize statistics for di-lepton studies, triggers requiring two electrons can be used to cut down the need for prescaling, as their rate is much lower than similar single electron triggers. These di-electron triggers include triggers with two electrons in ECAL and one in ECAL/one in HF.
Chapter 3

HF Electrons

3.1 Introduction

The coverage of CMS’s standard electron detection system ends with the joint ECAL and tracker coverage, at an $|\eta|$ of 2.5 (ECAL has trackless coverage out to $|\eta| < 3.0$). At higher $|\eta|$ the high radiation would decrease the lifetime of the tracker and ECAL. For this reason, most experiments (CMS included) simply accept that electrons in this region are outside of the detector acceptance. However, a reasonable number of electrons from physically interesting processes, specifically Z-decay leptons, will end up in the region of $|\eta| > 3$, which coincides with the coverage of HF. While HF was not designed with electron detection as its primary goal, its functionality still allows the detection and reconstruction of electrons. With the inclusion of the coverage of HF to electron detection, physics measurements can include a broader acceptance of electrons.

3.1.1 HF Electron Reconstruction

There are three steps to reconstructing electrons in LHC collision events using HF. The first step searches through the signals in HF to find potential candidates for electrons by searching for clusters of energy in the calorimeter. The second applies corrections to adjust the energy and position of the candidates to account for detector response. The last step applies identification requirements that compare the characteristics of the candidates to the profile of an electron shower, and accept or reject the candidate accordingly in order to remove non-electron responses.
Monte Carlo simulation from the Summer 2011 (referred to as Summer11 or “Monte Carlo:Signal”) production of $\text{DYToEE}_\text{M-20}_\text{CT10}_\text{TuneZ2}_\text{7TeV-powheg-pythia}$ was used for simulating electrons in HF. Data used were from the 2011A run at the LHC with center of mass $(\sqrt{s})$ of 7 TeV and selected runs totaling 216 pb$^{-1}$ in luminosity. The data sample specific to HF electrons consists of events that pass a high-level trigger that requires a loosely defined ECAL electron (described in [10]) with $p_T > 17/\text{GeV}$ and a very loosely defined HF electron (described later in this chapter) with $p_T > 15/\text{GeV}$ named HLT_Ele17_CaloIdL_CaloIsoVL_Ele15_HFL. This set of data was split into samples of “Data:Signal” and “Data:Background”. Data:Signal consists of data that pass a series of Z identification requirements: One electron in ECAL that passes set of requirements tuned to 80% acceptance of signal (referred to as ‘working point 80’ or ‘WP80’), one electron in HF that passes the loose identification requirements, and an invariant mass of the two electrons between 60 and 120/GeV. Data:Background is made of electrons that fail the ECAL and HF electron requirements, but are still in the mass window.

Without a tracker in front of HF, charge detection is not possible. This means that electrons, positrons and photons will all look very similar. For this reason, the particles found in HF with the following techniques are accurately called HF electromagnetic (HFEM) objects. However, the development of this technique is part of a larger search for $Z \rightarrow e^+e^-$ events and by requiring a second ECAL detected electron most HFEM objects used can be constrained (via invariant mass) to be electrons. From this point on, HFEM objects will simply be called HF electrons.

### 3.2 Finding HF Electron Candidates

To reconstruct HF electron candidates first the cell level responses are cleaned of known issues, then potential electron clusters are formed, and finally corrections are applied.

#### 3.2.1 PMT Hits

Due to positioning the PMTs directly behind the low $|\eta|$ region of HF (see fig. 2.8), there is a noticeable rate of events where a charged particle (typically a muon) travels through all of HF and impacts a PMT. The charged particle produces Cherenkov light in the window or envelope of the PMT and creates an anomalous response. The position
of the tower associated with the PMT does not correspond to the actual location of
the PMT; what is seen is an isolated tower displaying significant energy in only one of
its channels, when no particle is there. The HF detector team has developed several
methods for finding and rejecting PMT hits \[11, 12\]. The first method is named the
Polynomial Energy Threshold (PET) cleaning algorithm, which rejects PMT hits based
on the behavior of the long and short fibers within the same cell and checks to see if
only one set of fiber lengths was activated. The second is a pulse shape based cleaning
algorithm, which uses the timing of signal in the PMT to reject anomalous hits. Light
from true events must travel through the quartz fibers to reach the PMTs, while muons
causing PMT hits are traveling at relativistic speed and create the light right at the
PMT. As a result, PMT hit signals arrive 4-5 nanoseconds earlier than true detector
signals, and can be removed by looking at the arrival time of the signal.

**PET Algorithm**

The PET algorithm depends on the variable R:

\[
R = \frac{L - S}{L + S}
\]

where \(L\) is the energy in the long fiber of a tower, and \(S\) is the energy in the short
fiber. For a PMT hit, one fiber in the long/short pair will have very high energy while
the other will have little to none. Thus, \(R \to 1\) for a long fiber PMT hit and \(R \to -1\)
for a short fiber PMT hit. The effectiveness of this requirement depends on the energy
of the fiber and its tower location in \(\eta\), especially for long fibers. At lower energies, \(R\)
will often give values close to 1 even when it is a legitimate HF event because the low
energy is not enough to reach the short fiber length. How low this energy is depends on
the \(\eta\) location within the detector, so the energy is defined as a quadratic equation as a
function of \(i\eta\). This is the eponymous Polynomial Energy Threshold. Short fibers have
similar but lower requirements, as it is not normal for true events to have an \(R\) value
of -1 at any energy level. From the PET method the requirements for classifying a hit
in HF as a PMT hit are as follows:

\[
E_L > 162.4 - 10.19|i\eta| + 0.21|i\eta|^2
\]

and \(R > 0.98\)
for long fibers and

\[ E_S > 129.9 - 6.61|\eta| + 0.1153|\eta|^2 \]  \hspace{1cm} (3.3)

and \( R < -0.8 \)

for short fibers.

**Pulse Shape Based Cleaning**

The pulse shape refers the the charge vs time of the PMT output and relies on the variable

\[ Q = \frac{q_4}{\sum_{i=3}^{6} q_i} \]  \hspace{1cm} (3.4)

where \( q_i \) is the amount of charge in time sample \( i \). Proper events put most of their charge in time sample 4 (\( q_4 \)) leading to a \( Q \) of around unity. Out of time events (e.g. PMT hits) will appear early or late, lowering their \( Q \) value. Cell energies with energy \( E \) (with \( E > 40 \text{ GeV} \)) are accepted if they pass

\[ Q \geq 0.93 - e^{(-0.38275 - 0.012667 \times E)} \]  \hspace{1cm} (3.5)

By applying these requirements on HF cell energies, PMT hits are removed from polluting the selection of clusters.

### 3.2.2 The HF Cluster

The smallest division of calorimetric response is the signal from a single depth segment of a single tower in a detector, referred to here as 'cell energy'. In HF this represents the energy in one fiber length (long or short) of one tower. To find a candidate cluster of energy, all long fiber cell energies with transverse energy (\( E_T \)) greater than 5 GeV are sorted, greatest \( E_T \) to least, into a list of potential 'seeds'. The seed by itself is named a 1x1 region, referring to its \((\eta, \phi)\) size (fig. 3.1). These seeds become the center of a cluster of cell energies, including all signals within two towers of the seed in the plus and minus direction of \( \eta \) and \( \phi \). The total cluster is a square five towers long on each side and is called a 5x5 region. Within the 5x5 region, the square with three towers per side centered on the seed tower is called the 3x3 region. As 5x5 regions are made around
the seeds, in order of descending $E_T$, lower energy seeds included within a constructed cluster are removed from the list of seeds, ensuring the isolation and uniqueness of the HF cluster. Seed towers are further required to be in the $|\eta|$ range of [30,39] due to geometric limitations to making clusters in towers of $i\eta$ 29, 40 and 41.

### 3.2.3 HF Cluster Variables

Once the list of seeds has been made, a set of associated shape variables is constructed and its energy and position are calculated. The shape variables are based on the various groupings of towers one can make around the seed tower, and which fibers (long or short) are used. Two groups of variables are made: one for the long fibers and one for the short fibers. In each group the 1x1 region (the seed tower), the 3x3 region (the seed tower, and every tower adjacent to it in $\eta/\phi$) and the 5x5 region are examined. There is also one special variable, the ‘core’, which is only computed for the long fibers. It
is defined as the seed tower and the highest energy neighboring tower, provided it has energy greater than half the energy of the seed tower (if one exists; otherwise the core is the same as the long fiber 1x1 region). These variables are used in the definition of cluster energy, as well as for identification variables (see discussion later).

The energy of the cluster is defined as the energy in the 3x3 long fiber region.

\[ E_{\text{raw}} = \sum_{i\in 3\times 3} L_i, \quad (3.6) \]

where \( L_i \) is the energy of the long fibers in the i-th tower of the 3x3 region. The 3x3 is chosen because the transverse size of a tower is between 2.3 and 10.4 Moliere radii (depending on \( |\eta| \), so it can contain the transverse size of a electron-based shower. A large amount of electron energy may escape into one adjacent tower (this is the logic behind the 'core' variable) and a reasonable amount of energy may escape into the 3x3 region. However, energy from the shower will not significantly reach the full 5x5 region. For an electron, the energy in the 3x3 and 5x5 are nearly equivalent, a fact that will be used later in the electron identification section.

The position (in terms of \( \eta/\phi \)) is calculated using the average of the individual tower coordinates weighted by the log of the energy of that tower, because shower energy drops exponentially as it spreads transversely through the towers:

\[ \eta_{\text{raw}} = \frac{\sum_{i\in 3\times 3} \eta_i \times \log\left(\frac{L_i}{1 \text{ GeV}}\right)}{\sum_{i\in 3\times 3} \log\left(\frac{L_i}{1 \text{ GeV}}\right)}, \quad (3.7) \]

\[ \phi_{\text{raw}} = \frac{\sum_{i\in 3\times 3} \phi_i \times \log\left(\frac{L_i}{1 \text{ GeV}}\right)}{\sum_{i\in 3\times 3} \log\left(\frac{L_i}{1 \text{ GeV}}\right)}, \quad (3.8) \]

where \((\eta_i, \phi_i)\) are the coordinates of the center of the i-th tower in a 3x3 region.

After the list of seeds has been converted into a list of clusters and associated shape variables, that list of HF electron candidates is passed on to receive corrections and identification requirements.

### 3.3 HF Electron Candidate Corrections

Once the HF cluster is formed, its position and energy require corrections. The construction of the detector and energy dependent behavior of showers create biases in
the position reconstruction. The general calibration of energy regarding the long/short fibers for HF and the impact of multiple collisions in one event require a rescaling of the energy of the candidate cluster. In order to ensure the electron candidates are given accurate position and energy values, these effects were studied and corrections were derived. For correction studies, Monte Carlo simulations of Z decaying to electrons were used, comparing the true values from the simulator with the values reconstructed from the same electron after full simulation. The differences between truth and reconstruction in position and energy are systematic in nature, and therefore can be reliably corrected.

3.3.1 Correcting Position

When observing the resolution plots (generated-raw reconstructed) of $|\eta|$ (fig. 3.2) a bias appears in $|\eta|$ for the mean of the resolution to be noticeably above zero (around 0.01, or $\approx 5\%$ of the tower size in $\eta$). This is because while the particle travels along a path that is radially projective from the center of the detector, the towers of HF are instead parallel to the beam axis. Due to the penetration of the shower into the tower, the path of the shower will travel into a neighboring lower-$|\eta|$ tower, depositing energy and biasing the weighted position reconstruction. As this effect depends on shower depth and shower depth is dependent logarithmically on energy, the correlation between position bias and logarithmic energy is examined by plotting the $\eta$ resolution versus the log of energy (fig. 3.3). This bias can be fixed with a piecewise function:

\[
\begin{align*}
\text{If: } \log \frac{E}{(100 \text{ GeV})} & \leq 1.25; \quad \eta_{\text{reco}} = \eta_{\text{raw}} \pm C_0 \\
\text{If: } 1.25 < \log \frac{E}{(100 \text{ GeV})} & \leq 2.0; \quad \eta_{\text{reco}} = \eta_{\text{raw}} \pm [A_0 + B_0 \log \frac{E}{(100 \text{ GeV})}] \\
\text{If: } \log \frac{E}{(100 \text{ GeV})} & > 2.0; \quad \eta_{\text{reco}} = \eta_{\text{raw}} \pm C_1
\end{align*}
\]

where the $\pm$ matches the sign of $\eta$. The function is piecewise because at low energies the shower does not penetrate and has no bias, at medium energies the bias slowly develops with increasing energy, and at high energies the bias effect has saturated.

A second type of positional correction needed is to correct a positional bias that develops inside the HF towers in $\eta$ and $\phi$. The log-weighting method used to define the position of the cluster is an approximation only, and a systematic discrepancy is
Figure 3.2: Plots of the generated position minus the reconstructed position for (a) \( \eta \) and (b) \( \phi \). The red dashed line represents the resolution of uncorrected position, while the solid black line represents the resolution after both the energy and cell position correction. Gaussians are fitted to each peak for reference and comparison. The small change in width is due to a majority of events occurring in the center of the cell, where the cell position corrections are an order lower than the width of the resolution.

seen near the boundary of towers. To correct this, the average effect of this discrepancy has on the position of a cluster is measured as a function of the showers reconstructed location within a tower. This introduces a new set of variables:

\[
\eta_{\text{cell}} = \frac{|\eta|_{\text{raw}} - |\eta|_{\text{min}}}{|\eta|_{\text{max}} - |\eta|_{\text{min}}},
\]

(3.10)

\[
\phi_{\text{cell}} = \frac{\phi_{\text{raw}} - \phi_{\text{min}}}{\phi_{\text{max}} - \phi_{\text{min}}},
\]

(3.11)

where \((\eta_{\text{min/max}}, \phi_{\text{min/max}})\) are the minimum/maximum values of \((\eta, \phi)\) for a given tower. Thus the \(\eta_{\text{cell}}/\phi_{\text{cell}}\) gives the cluster position within the tower as a percentage from the edge of minimum value. Values outside the range of [0,1] are being reconstructed outside the seed tower, which certainly needs correcting.

When the resolution position bias is plotted as a function of cell position, there is a correlation between the two (fig. 3.4). This allows a positional correction to \(\eta/\phi\) to be
Figure 3.3: A plot of the mean value for generated position minus the reconstructed position as a function of \( \log \frac{E}{100 \text{ GeV}} \) for \( \eta \). The red line represents the piecewise correction function made using \( \eta_{cell}/\phi_{cell} \)

\[
X_{\text{reco}} = X_{\text{raw}} \pm [p_0 + p_1 \cdot X_{\text{cell}} + p_2 \cdot X_{\text{cell}}^2 + p_3 \cdot X_{\text{cell}}^3]
\]  

(3.12)

where \( X \) is \( \eta \) or \( \phi \) and the \( \pm \) matches the sign of \( \eta \) (and is always positive for \( \phi \)). The values for each variable are in table 3.1. The effects of this correction are most significant at the edges of the tower (reconstructed \( X_{\text{cell}} \) of -0.2 to 0.2 and 0.8 to 1.2). In the center of the tower (\( X_{\text{cell}} \) of 0.4 to 0.6) there is a noticeable shape that the correction does not attempt to fix. This shape (most clearly seen in the \( \phi \) version), is from the tendency to reconstruct events near the center of the tower as being at the exact center of the tower. This effect is an order of magnitude smaller than the overall position resolution, and any applied corrections to it did not have any noticeable impact on final results.
Figure 3.4: Plots of the mean value for generated position minus the reconstructed position as a function of Cell position for (a) $\eta$ and (b) $\phi$. The red line represent the fit from table 3.1. In the center a bias is seen due to electrons being preferentially reconstructed at the exact center of the the tower. As it is an order of magnitude smaller than the overall position resolution, it is not a noticeable effect on the final HF electron.

<table>
<thead>
<tr>
<th>X</th>
<th>$p_0$</th>
<th>$p_1$</th>
<th>$p_2$</th>
<th>$p_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta$</td>
<td>0.0125</td>
<td>-0.0475</td>
<td>0.0732</td>
<td>-0.0506</td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.0115</td>
<td>-0.0394</td>
<td>0.0486</td>
<td>-0.0335</td>
</tr>
</tbody>
</table>

Table 3.1: Table of cell position correction values.

### 3.3.2 Correcting Energy

The default calibration of HF is set so that the energy response in long and short fibers are equal for high energy protons and charged pions. For electrons, this calibration results in an energy reconstructed in the short fibers around 30% of the energy in the long fibers. Since electron reconstruction only uses the long fibers for energy measurement, a scale correction to the energy was added. This correction ($C_{El}$) is constant throughout HF. An additional correction that is needed is due to dead material in front of HF, mostly support structure for an experiment named TOTEM (TOTal Elastic and diffractive cross-section Measurement, located along the beam line outside CMS) and
the ECAL preshower. Dead material initiates and absorbs particle showers before the event reaches HF, thus decreasing the observed energy. The amount of dead material varies with \( i_{\eta} \), and it enters the correction equation as \( \omega_{i_{\eta}} \). Also in the \( \omega_{i_{\eta}} \) variable are other \( \eta \) dependent corrections that appear in simulation studies. Raw energy \( (E_{\text{raw}}) \) is corrected to the final reconstructed energy \( (E_{\text{reco}}) \) by the equation:

\[
E_{\text{reco}} = \omega_{i_{\eta}} \times C_{El} \times E_{\text{raw}}
\]

where \( C_{El} = 1.352 \pm 0.001 \) and \( \omega_{i_{\eta}} \) ranges within \([0.8, 1.02]\)

**Pileup Corrections to Energy**

An important energy consideration in HF occurs when more than one collision occurs at the same time. The LHC collides groups (clouds) of protons, because attempting one-on-one proton collisions would be prohibitively inefficient. As the number of protons in a cloud increases, the likelihood of an interesting collision rises. Eventually a point is reached where multiple events occur every time the proton beams collide. Because the events are ‘piling up’ on top of each other, the effect of multiple events is called ‘pileup’. Pileup corrections were unneeded in early LHC data because pileup was low, but in the 2011 data set the number of vertices often reaches 10-16 and the corrections are necessary. As seen in [13], there are on average six charged particles per ring of \( \eta \) per vertex, each with an average \( E_T \) of 0.5 GeV. If equal numbers of positive/neutral/negative particles is assumed, the amount of extra \( E_T \) in the 3x3 region that defines the energy, \( E_{pu}^T \), can be estimated by:

\[
E_{pu}^T = \frac{3}{2} N_{\pm} \times A_{\text{tower}} \times N_{\text{towers}} \times \overline{E}_{T}^\text{ave}
\]

\[
= \frac{3}{2} \times 0.00306 \times 9 \times 0.5
\]

\[
= 0.198 \text{ GeV} \approx 200 \text{ MeV}
\]

Where \( N_{\pm} \) is the number of charged particles per ring of \( \eta \) \((6)\), \( A_{\text{tower}} \) is the area in \( \eta/\phi \) of one HF tower \((0.175 \times 0.175)\), \( N_{\text{towers}} \) is the number of towers in the energy calculation of a cluster \((9)\), and \( \overline{E}_{T}^\text{ave} \) in the average \( E_T \) of a pileup particle \((0.5 \text{ GeV})\). This means that as pileup increases, HF electrons are expected to gain an average of 0.2 GeV in transverse energy per additional event (fig. 3.5).
\[ E_{\text{pred}}^{\text{T}} = \frac{M_Z^2 \cosh(\eta^{\text{ECAL}})}{2E_{\text{ECAL}}[\cosh(\eta^{\text{ECAL}} - \eta^{\text{HF}}) - \cos(\phi^{\text{ECAL}} - \phi^{\text{HF}})]} \]  

(3.14)

Figure 3.5: A plot displaying the average difference between the reconstructed electron \( E_{\text{T}}^{\text{reco}} \) and predicted electron \( E_{\text{T}}^{\text{pred}} \) as a function of the number of pileup events (vertex count), at an \(|\eta|\) of 32, using 2011 data. The red dashed line represents a fit whose slope is \( m_{\eta} \) (\( \approx 200 \text{ MeV per vertex} \)). \( E_{\text{T}}^{\text{pred}} \) is a prediction using an ECAL electron and the mass of the Z boson. In the equation: \( M_Z \) is the mass of the Z boson, \( E_{\text{ECAL}} \) is the measured energy of the matched ECAL electron, \( \eta^{\text{ECAL}}/\phi^{\text{ECAL}} \) are the \( \eta/\phi \) of the matched ECAL electron, and \( \eta^{\text{HF}}/\phi^{\text{HF}} \) are the \( \eta/\phi \) of the HF electron.
Thus a value for energy corrected for pileup ($E_{\text{reco}}^{\text{pileup}}$) can be calculated:

$$E_{\text{reco}}^{\text{pileup}} = E_{\text{reco}} - [m_{\text{pileup}} \cosh(\eta_{i\eta}) \times (N_{\text{vtx}} - 1)] \quad (3.15)$$

where $m_{\text{pileup}}$ is a slope value that corresponds to the theoretical $\approx 0.2$ GeV = 200 MeV per vertex, $\eta_{i\eta}$ is the $\eta$ that corresponds to the $i\eta$ of the seed tower, and $N_{\text{vtx}}$ is the number of vertices. The calculated value of $m_{\text{pileup}}$ used for data corrections is 0.1902 GeV.

### 3.4 Identification of HF Electrons

At this point, a candidate cluster has been created and corrected, as one would an electron in HF, but no effort has been made to ensure this candidate is in fact an electron. To this point in the procedure, any impact in HF that meets the energy threshold will be contained in the collection of HF clusters. To eliminate the non-electron events, a set of requirements is created that true electron clusters will pass, and non-electron clusters will fail.

The first requirement applied is via an isolation variable to eliminate all the clusters that are obvious jets. As a jet is a large group of particles, its transverse shape in $\eta/\phi$ is comparatively wide. Conversely the transverse size of an electron is quite narrow. The isolation variable uses the ratio between the long fiber energy of the 3x3 cluster and the long fiber energy of the 5x5 cluster ($E_{9/25}$).

$$E_{9/25} = \frac{\sum_{i\in3x3}(L_i)}{\sum_{i\in5x5}(L_i)} \quad (3.16)$$

where $L_i$ is the energy of the long fibers in the $i$-th tower of the relevant cluster.

For the electron, all the energy should be completely contained within the 3x3 cluster; therefore this ratio should trend to unity. A broad jet, however, should have a measurable amount of energy spill out of the 3x3 into the 5x5, and it will have a ratio noticeable lower than unity. In practice, the isolation variable requirement is set to $E_{9/25} > 0.94$ (fig. 3.6).

Our second variable is similar as it is a further transverse/isolation classification. As electron showers are transversely narrow, most of their energy is not just within the

---

1 Addition of short fibers has a negligible effect on transverse requirements, and is not needed.
Figure 3.6: Plot of the $E_{9/25}$ isolation variable. “Data: Signal” (black line) and “Monte Carlo: signal” (Summer11, the green line) are contrasted to “Data: Background” (red shaded region). Excess “Monte Carlo: signal” over “Data: Signal” in the highest bin is due to less pileup present in the simulation, but otherwise there is agreement between data and simulation.

3x3, but within the core (described in an earlier section). By looking at the ratio of energy in the core to the 3x3 long energy ($E_{C/9}$), a profile of what percentage of energy is in the core towers is seen:

$$E_{C/9} = \frac{\sum_{i\in\text{core}} L_i}{\sum_{i\in3\times3} L_i},$$  

(3.17)

where $L_i$ is the energy of the long fibers in the i-th tower of the relevant cluster. For electrons this variable will be close to unity, and for non-electrons less than one (fig. 3.7).

A second ID variable that is the longitudinal (depth) behavior of the cluster. Electrons are expected to travel only a short depth into HF (depositing only a small fraction of energy in the short fibers), while hadrons will penetrate further within (depositing nearly equal energy in long and short fibers). Thus the ratio of energy in the short
fibers of the 3x3 cluster to the energy in the long fibers of the 3x3 cluster \( E_{S/L} \) is a measure of how deep a shower travels:

\[
E_{S/L} = \frac{\sum_{i \in 3 \times 3} S_i}{\sum_{i \in 3 \times 3} L_i},
\]  

(3.18)

where \( L_i \) is the energy of the long fibers in the \( i \)-th tower of the 3x3 cluster, and \( S_i \) is the energy of the \( i \)-th short fiber. For electrons, this variable skews towards 0.2-0.3, while for charged hadrons it trends towards unity. As most jets will be a mix of charged hadrons, photons, and electrons, the usual background value in this variable is closer to one half (fig. 3.8).

There is an energy-dependent effect in \( E_{S/L} \). The energy of an electron shower determines its shower depth, shallow though it may be. Higher energy electrons will penetrate further into the short fibers, giving the \( E_{S/L} \) a higher value. If the \( E_{S/L} \) variable is treated as energy independent, the requirements will either unintentionally
Figure 3.8: Plot of the $E_{S/L}$ longitudinal variable. “Data:Signal” (black line) and “Monte Carlo:signal” (Summer11, the green line) agree and are contrasted with “Data:Background” (red shaded region). The overlap between signal and background is greater than in other variables, due to high electromagnetic content in background events. This overlap strongly motivates using the two-dimensional method rather than a ‘square cut’ in the two variables.
remove high energy electrons, or let in an unacceptable amount of low-energy background events. To account for this energy dependence, a corrected, energy-dependent version of $E_{S/L}$ is created:

$$E_{S/L}^{\text{cor}} = \alpha + \beta \log \frac{E \text{ GeV}}{100 \text{ GeV}} + \gamma E_{S/L}$$

Where:

$$\alpha = \frac{-b}{\sqrt{1 + m^2}}$$
$$\beta = \frac{-m}{\sqrt{1 + m^2}}$$
$$\gamma = \frac{1}{\sqrt{1 + m^2}}$$

The new $E_{S/L}^{\text{cor}}$ variable is the perpendicular distance between the point $(\log \frac{E \text{ GeV}}{100 \text{ GeV}}, E_{S/L})$ to the line $E_{S/L} = m \log E + b$, in a log $\frac{E \text{ GeV}}{100 \text{ GeV}} \times E_{S/L}$ space (fig. 3.9). The values of $m, b, \alpha, \beta, \gamma$ can be seen in table 3.2.

While the $E_{C/9}$ and $E_{S/L}^{\text{cor}}$ requirements could be applied separately, this is not nearly as efficient at keeping electrons and excluding background as applying the requirements together as one two dimensional requirement (fig. 3.10):

$$E_{C/9} - m_{\text{cor}} \cdot E_{S/L}^{\text{cor}} > C_{2d}, \quad (3.19)$$

With this requirement, optimization can be made for either electron efficiency or background rejection (see table 3.3). For a tight electron definition background rejection is maximized, leading to a selection of $m_{\text{cor}} = 0.20$ and $C_{2d} = 0.92$ for a 50%/93% signal acceptance/background rejection effect. For a loose electron definition efficiency is maximized and requires $m_{\text{cor}} = 0.475$ and $C_{2d} = 0.815$ for a 95%/75% signal acceptance/background rejection. For a medium definition $m_{\text{cor}} = 0.275$ and $C_{2d} = 0.875$, creating a balanced 75%/86% signal acceptance/background rejection. A further very loose electron definition of $m_{\text{cor}} = 0.475$ and $C_{2d} = 0.7$ is used for High Level Triggering, such as in the aforementioned HLT_Ele17_CaloIdL_CaloIsoVL_Ele15_HFL trigger.
Figure 3.9: (a) Plot of the uncorrected $E_{S/L}$ vs log $\frac{E_L}{100 \text{ GeV}}$ (b) plot of $E_{S/L}^{\text{cor}}$ vs log $\frac{E_L}{100 \text{ GeV}}$. The dots represent the mean value for the Y-axis in each plot. An energy dependence is seen in (a) and removed in (b).

Table 3.2: Parameters used for data and Summer11 Monte Carlo simulation in the transformation $E_{S/L} \mapsto E_{S/L}^{\text{cor}}$.

<table>
<thead>
<tr>
<th></th>
<th>$b$</th>
<th>$m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data: Signal</td>
<td>0.008 ± 0.0042</td>
<td>0.221 ± 0.0030</td>
</tr>
<tr>
<td>Monte Carlo: Signal</td>
<td>0.037 ± 0.0032</td>
<td>0.200 ± 0.0022</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data: Signal</td>
<td>-0.008 ± 0.0041</td>
<td>-0.216 ± 0.0028</td>
<td>0.9764 ± 0.00061</td>
</tr>
<tr>
<td>Monte Carlo: Signal</td>
<td>-0.036 ± 0.0031</td>
<td>-0.196 ± 0.0021</td>
<td>0.9806 ± 0.00042</td>
</tr>
</tbody>
</table>

Table 3.3: Systematically optimized $C_{2d}, m_{\text{cor}}$ values used in the new two-dimensional requirement for the efficiency for electron reconstruction. The requirements are 50%, 75% and 95% efficiency correspond to the tight, medium and loose electron ID requirements.

<table>
<thead>
<tr>
<th>Efficiency</th>
<th>50%</th>
<th>55%</th>
<th>60%</th>
<th>65%</th>
<th>70%</th>
<th>75%</th>
<th>80%</th>
<th>85%</th>
<th>90%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{2d}$</td>
<td>0.92</td>
<td>0.91</td>
<td>0.90</td>
<td>0.895</td>
<td>0.885</td>
<td>0.875</td>
<td>0.865</td>
<td>0.845</td>
<td>0.835</td>
<td>0.815</td>
</tr>
<tr>
<td>$m_{\text{cor}}$</td>
<td>0.20</td>
<td>0.25</td>
<td>0.30</td>
<td>0.225</td>
<td>0.275</td>
<td>0.275</td>
<td>0.275</td>
<td>0.425</td>
<td>0.375</td>
<td>0.475</td>
</tr>
</tbody>
</table>
Figure 3.10: A plot of $E_{S/L}^{cor}$ vs $E_{C/9}$ (“the 2D requirement”) for electrons. The red line represents a “Loose” requirement, where all events below the line are accepted.
Chapter 4

Physics of Z Transverse Momentum

4.1 Introduction

To properly explain the central analysis of this thesis, the foundational physics upon which it builds is a necessary starting point. The aptly-named standard model serves as the cornerstone of particle physics and this chapter will briefly describe it and discuss in detail the components related to this thesis. Using this context, the specifics of high energy proton-proton collisions and Z boson production can then be meaningfully presented. The focus of this chapter then narrows down to the analysis of the transverse momentum of Z bosons and the theory that predicts this behavior. The final section will discuss the challenges of measuring transverse momentum and the motivation to use the novel variable $\phi^*$, and its advantages over conventional $Q_T$.

4.2 Standard Model

The standard model (fig. 4.1) is the most complete, fundamental and accepted explanation of physics at the ‘High Energy’ level. The model contains six quarks and six leptons that make up matter-creating fermions and four force-carrying bosons. All particles have an antiparticle identical in every way except charge (except for photons, Z
bosons, and gluons which are their own anti-particle). The model also explains the behavior and interactions of these particles at the high energy scale. The completeness of the model is such that it can explain nearly all of the strong, weak, and electromagnetic interactions observed and answers most of the questions asked of it. In the past the standard model predicted the existence and properties of many of its members (notably the Z boson) before they were directly observed, and in the present the motivation it provided to search for the Higgs boson has lead to the recent discovery of the Higgs boson with a mass around 125 GeV [14]. As such, the particles and processes described within the standard model will be the basis this thesis is given from. Following is a brief discussion on model members pertinent to this thesis.

4.2.1 Quarks

Quarks are the building blocks of the standard model as they can be bound together to make particles called hadrons. The lightest quarks are the up- and down-quark, and make up all the stable hadronic matter present in the known universe (protons and neutrons). Heavier generations of quarks are unstable and will quickly decay to lighter versions.

While heavy quarks do not dominate many natural situations, their existence is observed in particles, such as protons, where they have a quantum probability of existing and thus affect the properties of the particle. Within a collider these heavy quarks are produced more frequently and much effort has been put into studying and quantifying their properties and behaviors.

Quarks have fractional electric charge (2/3 for up-type quarks, -1/3 for down-type quarks). This means that a quark-antiquark pair have a net 0 or 1 electric charge, while combinations of three quarks (such as protons) can have 0, 1, or 2 net electric charge. This electric charge means that quarks can interact via the electroweak force (photons and W/Z bosons).

Another important quark property is color charge. Color charge is the name for the property of quarks that interacts with the strong force (via gluons). Quarks come in three different color charges: red, blue and green (with anti-red, anti-blue, and anti-green for antiquarks). It is the strong force that allows quarks to bind together with enough strength to form hadrons.
Figure 4.1: The standard model
4.2.2 Gluons and Strong Interactions

Gluons are massless carriers of the strong force that bind quarks together into pairs or triplets. While quarks carry a single color of the three color charges gluons are defined as having eight possible combinations of color-anticolor charge superpositions (table 4.1) that form a color-octet. In addition to quarks, gluons also interact with each other. If quarks bound by gluons move apart the gluon-gluon interactions increase and keep the strength of the bond essentially constant. In this way, the bond is strengthened enough that the energy required to break the bond is greater than the energy needed to create a pair of new quarks, so the pair-production occurs first. The new quarks match up with the separating quarks into two new pairs. This phenomena, called quark confinement, prevents a single quark from being isolated and is one method of making new hadrons.

\[
\begin{align*}
(r\bar{b} + b\bar{r})/\sqrt{2} & - i(r\bar{b} - b\bar{r})/\sqrt{2} \\
(r\bar{g} + g\bar{r})/\sqrt{2} & - i(r\bar{g} - g\bar{r})/\sqrt{2} \\
(b\bar{g} + g\bar{b})/\sqrt{2} & - i(b\bar{g} - g\bar{b})/\sqrt{2} \\
(r\bar{r} - 2b\bar{b} + g\bar{g})/\sqrt{6} &
\end{align*}
\]

Table 4.1: A table of linearly independent color-anticolor superpositions for gluons, forming a complete color-octet (though other complete, linearly independent octets are equally correct). Colors represented are red (r), green (g), blue (b), anti-red (\(\bar{r}\)), anti-green (\(\bar{g}\)), and anti-blue (\(\bar{b}\)).

4.2.3 Hadrons: Baryons and Mesons

Via the strong force, quarks form into two types of hadrons: baryons (made of three quarks) and mesons (made of two quarks). The color charge of any hadron has to be ‘colorless’, which means that in mesons every color charge present is balanced by an anti-color charge and in baryons one of each color must be present (conceptually making ‘white’, which is colorless). Because of this, hadrons, as a whole, do not interact via the strong force with each other, but are capable of interacting via the residual strong force (which holds together elemental nuclei) by exchanging mesonic ‘residue’ from internal strong interactions. While in a very simple model of a hadron there are only as many gluons as needed to bind the quarks colorlessly, gluon-gluon interactions and quark confinement allow for a much more complex reality. Thus, at any given
energy there are unknowable numbers of low energy gluons and short-lived quark pairs within a hadron. As energy increases, so does the number of these extra particles. As a baryon, a proton is subject to this internal uncertainty which turns out to be very important in proton-proton collisions.

### 4.2.4 The Weak Bosons

The weak force is mediated by the W and Z bosons. The W boson can have either a positive or negative electric charge, and mediates interactions between up-type and down-type quarks and between leptons and neutrinos, as can be seen by its common decay modes. More importantly, with regard to this paper, is the Z boson, which is electrically neutral and interacts quarks and leptons of the same flavor and neutral net electric charge.

The Z boson most often decays into quarks (table 4.2) which is the common channel studied in lepton colliders (such as LEP), but the lepton channels of electrons and muons are more commonly measured in hadron collision experiments (such as the LHC) because of the clearer signal these leptons have in detectors versus hadronic decay signals. Previous experiments (LEP) have measured the Z boson and its couplings in great detail. Due to the precision of these measurements (the mass measurement has a precision of 0.00023%) the Z boson is now an effective tool that aids experiments in searching for and analyzing other objects. Current efforts at the LHC to measure the properties of Z bosons are in hopes to further tune the ability to use it as a probe of the unknown.

<table>
<thead>
<tr>
<th>Channel</th>
<th>example</th>
<th>probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>quarks</td>
<td>$Z \rightarrow q\bar{q}$</td>
<td>69.91%</td>
</tr>
<tr>
<td>lepton</td>
<td>$Z \rightarrow \ell\ell$</td>
<td>10.097%</td>
</tr>
<tr>
<td></td>
<td>$Z \rightarrow ee$</td>
<td>3.363%</td>
</tr>
<tr>
<td></td>
<td>$Z \rightarrow \mu\mu$</td>
<td>3.366%</td>
</tr>
<tr>
<td>invisible</td>
<td>$Z \rightarrow \nu\nu$</td>
<td>20.00%</td>
</tr>
</tbody>
</table>

Table 4.2: Selected decay modes of the Z boson and the experimental probability of each.
4.2.5  Electrons and Leptons

Leptons are matter constituents that interact electromagnetically with other fermions and weakly with each other. They are a common final product in particle interactions. Of particular importance is the electron, which is as previously mentioned is a good signal for observing the $Z$ boson. Electrons are completely stable and their interactions within the standard model are very well understood, making them highly observable in experiments. When a $Z$ boson decays into two electrons (called ‘daughter’ particles), the electrons carry all the information about the $Z$ boson that can be determined.

4.3  Proton-Proton Collisions

4.3.1  Protons at the LHC

As previously mentioned, the LHC collides two clouds of protons at a center of mass energy ($\sqrt{s}$) of 7 TeV (3.5 TeV for each proton). At this energy, the simple model of protons as three quarks is inadequate. Once a proton has been given this much energy, a more complex and comprehensive model is needed.

4.3.2  Proton Parton Model

The parton model of protons (and other hadrons) creates an image of a proton not as an individual distinct object but a collection of many quark-and-gluon point-objects called partons. While at lower energies the three valence quarks of the proton dominate the partons, at higher energies a ‘sea’ of other quarks, along with an indeterminable number of gluons, are present. The sea-quarks are a superposition of the quantum states of quark-antiquark pairs for each flavor. Likewise the gluons contributes to the superposition with quantum states of fusing, radiating an pair producing quarks. Upon a proton-proton collision the superposition collapses to a single state (as if measured) and it is the individual partons of that state that hit each other, and scatter or interact accordingly. For a specific process, such as the production of a $Z$ boson, the results depend only on the specific partons involved, not the whole proton. For example, a valence quark and an anti-quark from the sea can annihilate to form a $Z$ boson, something a non-parton model cannot explain.
4.3.3 Probing Proton Partons

Since it is made up of many partons, the total momentum of the proton is made up of the net momenta of the partons, as each parton has a piece of the total momentum of the proton. The momentum of a single parton can be described by what fraction of the total momentum it carries. This value, called the Bjorken $x$ (or momentum fraction) is defined as:

$$x = \frac{p_x}{P_{\text{tot}}}$$  \hspace{1cm} (4.1)

where $p_x$ is the momentum of the parton in question, and $P_{\text{tot}}$ is the total momentum of the proton, which is also the sum of all the parton momenta.

Proton Parton Distribution Functions

The $x$ value of any given parton is probabilistically obtained and different flavors of partons have different likelihoods of having a specific momentum fraction. The valence quarks have a high probability of receiving a large momentum fraction, while a gluon has a high probability of receiving a small fraction of the total momentum (fig. 4.2). The distribution of the probabilities of having a specific momentum fraction are collected for each flavor of quark, anti-quark and gluon and form a Parton Distribution Function (PDF). The precise values for the proton PDF are not currently provided by QCD. Whether PDFs are at all predictable by any model is not yet known. As such, current PDF models rely on fits to data and assumptions that are reasonable but not fully-justified, for their creation. There are currently many different groups of scientists creating PDFs using a broad spectrum of data, theories, and assumptions. The different PDFs effectively compete with each other to best describe reality with the goal of eventually answering the question of how PDFs work, or at least which assumptions/fits are the best when compared to measurements.

One method of measuring PDFs is relating the $x$ of the parent partons of the Z boson to the rapidity ($y$) of the Z boson such that

$$x_\pm = \frac{M_Z}{\sqrt{s}} e^{\pm y}$$  \hspace{1cm} (4.2)

where $M_Z$ is the invariant mass of the Z, $s$ is the center of mass energy of the protons, $y$ is the rapidity of the Z. The sign before the rapidity in the exponent depends on which
Figure 4.2: A sample PDF for a proton at 10 GeV
of the partons the momentum fraction is from, $x_+$ for the momentum fraction of the parton that came from proton heading in the ‘positive’ rapidity direction, or $x_-$ for the parton heading in towards negative rapidity. Thus each rapidity measurement provides the $x$ values for both incident partons.

4.4 The Transverse Momentum of Z bosons

Another important measurement of Z bosons is their transverse momentum. The transverse momentum of the Z boson provides information on its parent protons, though not as simply as rapidity does on momentum fraction.

4.4.1 Transverse Momentum

In its simplest concepts, analysis of the transverse momentum of Z bosons is the study of how a Z boson gains momentum transverse to the trajectory of the protons that formed it. The momentum of the protons is entirely in the direction of beam path and so in any collision in CMS the final total transverse momentum will be zero. Thus, any transverse momentum given to the Z boson must be balanced out somewhere else in the interactions creating it. By studying transverse momentum, a greater understanding of several of these processes can be obtained.

Notation

the transverse momentum of a Z boson is specifically referred to as $Q_T$, similarly the four-momentum of the Z boson is $Q$. The transverse momentum of either of the daughter electrons of the Z boson are referred to as $p_T$ (or $p_{T1}$ and $p_{T2}$ when being specific), and $P_T$ is used to refer to transverse momentum in general terms or for a particle other than the Z bosons or its daughter electrons. Some other values to soon be used are $b$ the impact parameter, $M$ the mass of the Z boson (where $Q^2 = M^2$), $\alpha_s$ the strong coupling constant, and $x_{1,2}$ the momentum fraction of the specific parent parton.
4.4.2 Z Boson Production with Transverse Momentum

The simplest form of Z boson production is the s-channel (fig. 4.3 a). In this situation of simple quark-antiquark annihilation there is no method for $Q_T$ to enter. Were this the only process for Z boson production, this study would be trivial, with all Z bosons having a $Q_T$ of zero (within experimental uncertainty). Expectedly, there are variations on this process that can add $Q_T$. One simple variation on the s-channel is the t-channel (fig. 4.3 b). In this channel the final products include the Z and a quark which can have equal and opposite transverse momentum. In this channel, significant amounts of $P_T$ can go into the quark to balance out the $Q_T$. This channel is dominant in the production of Z bosons at the LHC because the gluon required in the t-channel is more probable to be present than the antiquark needed for the s-channel.

The s- and t-channel diagrams are the simplest way to take partons and make Z bosons, but they are not the complete story. These diagrams show the minimum requirements for the production of Z bosons, but there can be other processes occurring that are not necessary to the production but have a kinematic impact on the final product, such as contributing to the transverse momentum of the Z boson. One common sub-process that can contribute to $Q_T$ is neutral radiation. When a branch radiates, it ejects a boson without changing the original particles type, just its momentum.

An example of this is already seen in the t-channel, where the quark emits a Z boson, altering the quarks momentum, but not its flavor. Since the Z is the target particle, it is considered the product, not just radiation, but quarks can also emit gluons before or after the desired process, which will alter the final $Q_T$ (fig. 4.3 c).

Initial State Radiation (ISR) is a conceptual way of explaining how partons can get an initial $P_T$, the idea being that the parent partons emit gluons before interacting to make the Z (or other desired particle). In this case the parent parton and the radiation can have equal and opposite $P_T$, thus giving the parent some $P_T$ to pass on to the Z when it is created. As the probability of radiation increases as the energy of the radiated gluon decreases, ISR can introduce very low amounts of $Q_T$ which is why ISR is used to account for the majority of the lower end of the $Q_T$ spectrum.
Figure 4.3: Feynman diagrams showing (a) $s$-channel $Z \rightarrow e^+e^-$ decay, (b) $t$-channel $Z$ boson production, and (c) $Z$ boson production diagram with initial state radiation.
4.4.3 Z Boson Differential Cross-Section

To closely measure the behavior of $Q_T$, this thesis will use a differential cross-section measurement. The term ‘cross-section’ in particle physics refers to the probability of a process occurring, expressed in units of area:

$$\sigma(X) = \frac{N_X}{\mathcal{L}}$$  \hspace{1cm} (4.3)

where $\sigma(X)$ is the cross-section of process $X$, $N_X$ is the number of events $X$ that occur and $\mathcal{L}$ is the luminosity (the measure of how many interaction opportunities have occurred in units of inverse area). For example, the inclusive $Z$ cross-section ($\sigma(Z+X)$) is a measurement of how likely a $Z$ boson is to be made by any means. When measuring the $Q_T$ of a $Z$ boson, it is logical to put the measurement in terms of a cross-section. This can be done ‘differentially’ by counting the number $Z$ bosons with a specific $Q_T$ value (or range of values, in the case of non-infinite statistics) and knowing the luminosity over which you are counting, one can make a measurement of cross-section as a function of $Q_T$. Thus the differential cross-section is defined as:

$$\frac{d\sigma(Z \rightarrow e^+e^-)}{dQ_{T,k}} = \frac{(N_{Q_{T,k}})}{\Delta Q_{T,k} \times \mathcal{L}}$$  \hspace{1cm} (4.4)

where $q_{T,k}$ is a specific range of $Q_T$, $N_{q_{T,k}}$ are the number of $Z \rightarrow e^+e^-$ events that are in $Q_T$ range $k$, and $\Delta q_{T,k}$ is the size of $Q_T$ range $k$. By intelligently binning over the range of the data a spectrum of differential cross-sections can be created for $Q_T$, or other similar variables. The dependence on luminosity can be removed by normalizing the differential cross-section to the total cross-section, creating a generalized measurement called the normalized differential cross-section, $\frac{1}{\sigma} \frac{d\sigma}{dQ_T}$.

4.5 QCD Predictions of Transverse Momentum

Quantum Chromo Dynamics (QCD) is a branch of Quantum Field Theory that describes the effects of color charge (i.e. the strong force) on quarks and gluons. Since quarks (often with some influence from gluons) are responsible for the creation of $Z$ bosons, QCD can provide theoretical input on what the differential cross-section of the $Z$ boson should look like, and in-turn measurements of the differential cross-section can inform
and improve the understanding of QCD. While a rigorous explanation of the QCD modeling of the differential cross-section of Z bosons is far beyond the scope of this thesis, a simplified overview of the basics as they apply to the differential cross-section measurement will be presented. The greatest level of detail (though still very simplified) will be given to QCD calculations in the low end of the transverse momentum spectrum, for it is here where theoretical calculations run into problems and the most disagreement with experimental measurements can be found.

4.5.1 Perturbative Calculations of the Differential Cross-section

Simply put, to predict the differential cross-section of Z bosons, QCD collects the various processes that can create Z bosons and combines them as a collection of equations that depend on kinematic variables of the Z boson. A very conceptualized basic form of the differential cross-section is:

$$\frac{d\sigma}{dQ^2 dy dQ_T^2} = \Sigma_i F_i(Q, Q_T, y)$$

(4.5)

where the $dQ^2 dy dQ_T^2$ term illustrate the three differential variables of the Z boson that the cross-section depends on: momentum($Q$), rapidity ($y$), and transverse momentum ($Q_T$), with the $Q$ and $Q_T$ squared to render them scalar (i.e. independent of spacial variables like $\phi$ or $\eta$). The functions $F_i(Q, Q_T, y)$ represent the different sources that can create, or suppress, Z boson production based on the differential variables. To get down to a differential cross-section that depends only on $Q_T$, the function is integrated over the other differential variables. The functions, which could now be represented simply as $F_i(Q_T)$, are actually very complex and occasionally contain infinite terms or other such elements that make fully determining the terms prohibitively difficult, if not impossible, at this stage of scientific understanding. In cases like this, perturbative theory is used to simplify calculations. Perturbation theory is a method a calculating complex, finely-tuned equations by starting with a simple, roughly-accurate equation and adding on layers of complexity until the result is satisfyingly close to what the real equation would give. Ideally the first, simplest piece has the largest impact on the final result, and the additional corrections are each smaller and smaller, resulting in a convergence to the right behavior. If this is not the case, if the attempts to fine-tune the
simple equation begin to get larger or non-convergent, then the equation is considered non-perturbative, and is generally much harder to accurately calculate.

In perturbative calculations of the $Q_T$ differential cross-section, the variable that most often effects the magnitude of the term is a factor of $\alpha_s$, the strong coupling constant, that appears in most terms [15]. The $\alpha_s$ factor is a ‘running constant’, in that it varies with $Q$. As $Q$ approaches zero, $\alpha_s(Q) \approx 1$, but otherwise $\alpha_s(Q) \ll 1$ and can be used perturbatively. When all the contributing functions to the differential cross-section are determined perturbatively, their contributions to the final sum can be arrange by order of $\alpha_s$ they contain. In this way, the sum over all the contributing functions can be rewritten in the form:

$$\frac{d\sigma}{dQ_T^2} = \alpha_s u_1 + \alpha_s^2 u_2 + \alpha_s^3 u_3 \ldots$$

where $u_i$ is a function, often quite complex, made by collecting the terms from the contributing processes with a factor of $\alpha_s^i$. As long as the $u_i$ do not grow larger by orders of magnitude as $i$ increases and $\alpha_s(Q) \ll 1$, the first terms will be the ones with greatest effect and equation will converge. Selecting $O(\alpha_s)$ (effectively choosing only $u_1$) will give the Leading Order (LO) of the perturbation, likewise $O(\alpha_s^2)$ (Next to Leading Order or NLO) or $O(\alpha_s^3)$ (Next-to-Next-to Leading Order or NNLO) can be the last term chosen. From a conceptual standpoint, the order chosen is arbitrary, determined mostly by how exhaustive the discussion needs to be. From a theory standpoint, the order chosen is often limited by the scale of the uncertainty within the theory itself and the time it takes to calculate out the terms. Likewise simulations often are limited to NLO or NNLO by the complexity and uncertainty of the higher order contributions. In order to compare theory to an experimental measurement, the order must be chosen to coincide with the precision of the apparatus: too low of an order and the experiment will disagree due to unaccounted perturbative effects, too high of an order and resources are wasted simulating detailed effects that the experiment cannot possibly detect.

When applying the theory to an experimental measurement the order chosen is further constrained by the precision of the apparatus.

In the LO case of the differential cross-section, the $u_1$ term can be broken down to
show its dependence on $Q_T$ [16]:

$$u_1 = A \left[ \ln\left( \frac{Q^2}{Q^2_T} \right) \right] + B \left[ \frac{1}{Q^2_T} \right] + C(Q^2_T) \quad (4.7)$$

Where $A$ and $B$ are simply coefficients, but $C$ is a function. Higher orders will continue to look similarly, but will begin including higher orders of $Q^2_T$, $\frac{1}{Q^2_T}$, and $\ln\left( \frac{Q^2}{Q^2_T} \right)$.

The perturbative assumption that $\alpha_s$ will be the dominant factor in the calculation is accurate when $Q_T$ is similar to $Q$ in magnitude (this will be considered the “high $Q_T$” region). In this region the sum of $u_1$ and its higher orders will converge to form a calculation of the differential cross-section that has good agreement with experimental results. However, at values of $Q_T \ll Q$ (the “low $Q_T$” region) the values of $\ln\left( \frac{Q^2}{Q^2_T} \right)$ and $\frac{1}{Q^2_T}$ become large enough that higher orders of them will overpower the effects of $\alpha_s$, eventually becoming singular as $Q_T$ approaches zero (fig. 4.4). This creates regions of $Q_T$ where this perturbative calculation diverges and does not give accurate results. The solution is to re-derive the differential cross-section at this level in a different manner, with a focus on how to handle these singular terms.

### 4.5.2 The Differential Form at Low $Q_T$

The specialized recalculation of the differential cross-section for vector bosons at low $Q_T$ is described in detail in [17]. The form this thesis will refer to is:

$$\frac{d\sigma}{dQ^2dydQ^2_T} \approx \int_0^\infty \frac{bd}{2\pi} J_0(bQ_T)e^{-R(b,Q)} \times \Sigma(b, Q, x_1, x_2) + Y(Q_T, Q, x_1, x_2) \quad (4.8)$$

where $J_0(bQ_T)$ is the Bessel function, $R(b, Q)$ is term that accounts for radiation, $\Sigma(b, Q, x_1, x_1)$ is a term that contains the Born-level behaviors, and $Y(Q_T, Q, x_1, x_2)$ is a term that connects this low $Q_T$ calculation with the high $Q_T$ region (it is negligible in the low $Q_T$ region, but is important for matching low $Q_T$ behavior to the high $Q_T$ perturbative results within the transitional $Q_T$ region), $b$ is the impact parameter which defines the geometry of the collision, while $x_1$ and $x_2$ are the momentum fractions of the parent partons. The exact form of most of the equations ($\Sigma(b, Q, x_1, x_1)$, $Y(Q_T, Q, x_1, x_2)$, $J_0(bQ_T)$) are not relevant to the discussion in this chapter, only the $R(b, Q)$ radiator term will receive much attention as it contains the soft gluon radiation behavior (part of the transverse momentum creating process referred to earlier as
Figure 4.4: Comparison of the growth of the terms of \( Q_T \) within the perturbation on an arbitrary scale.
ISR), which will be the source of the \( \ln\left(\frac{Q^2}{Q_T^2}\right) \) terms that cause the low \( Q_T \) singularities. Thus the form and treatment of the \( R(b, Q) \) will be the focus of determining how to non-divergently calculate the differential cross-section in this region.

**Reforming Divergent Equations**

There are several phenomena, such as gravity, on which the standard model does not include any information. Other phenomena may be describable by the standard model but are not yet computable by even the best modern methods. When approaching parts of experimental high energy physics, it is sometimes difficult to know if the observations are influenced by something the standard model doesn’t claim to explain, or by something the standard model can explain but has not yet been fully modeled or understood by physicists. In the case of \( R(b, Q) \), its singular behavior at low \( Q_T \) (leading to obvious disagreements with experimental data) may imply that it is not definable by QCD theories because it is not ruled by the standard model as we know it. Alternatively it could be that, while ruled by the standard model, the current modeling of it is incorrect because of computational limitations. If the radiator is not fully describable by the standard model (and by extension QCD), then it is necessary to treat its behavior in this region as non-perturbative. The form of \( R(b, Q) \) in this case must be based on data, not theory, and requires that a series of different possible parametrized forms be tried until a form and parametrization of best fit can be found, a so-called Sudakov term. On the other hand, if the radiator is fully dependent on the standard model and the singular nature is from modern computational limitations (dependence on order of \( \alpha_s \) being the dominant term, for example) then it is appropriate to reassess how else it can be calculated, essentially keeping it a perturbative term, but coming at it from a different mind-set of what the scaling is, often called a summation. As there is reasonable scientific support for both methods, both will be presented for consideration in this thesis.

**4.5.3 Non-Perturbative Sudakov Method**

When approaching \( R(b, Q) \) as non-perturbative, it is important to note that not all of its contribution becomes singular, which is to say that not all of \( R(b, Q) \) need be
non-perturbative. Rather \( R(b, Q) \) can be broken into two parts

\[
e^{-R(b, Q)} = e^{-S_P(b, Q)} e^{-S_{NP}(b, Q)}
\]

\[
S_{NP}(b, Q) = g_s(b, \ln Q) + g_a(b, x_1) + g_b(b, x_2)
\]

(4.9)

where \( S_P(b, Q) \) is perturbative and calculable as with other such terms, but \( S_{NP}(b, Q) \) is non-perturbative Sudakov term, effectively a dampening term that counter-acts the singularities, and must be determined by fitting experimental data and \( g_s, g_a, g_b \) are whichever functions one can get to fit the data. As a result, there are several suggested forms for \( S_{NP}(b, Q) \), but [18] offers up compelling evidence that their BNLY parametrization has an excellent fit to data:

\[
g_s(b, \ln Q) = g_2 \ln \left( \frac{Q}{2Q_0} \right) b^2
\]

\[
g_a(b, x_1) = \frac{g_1 b^2}{2} (1 + 2g_3 \ln(10x_1))
\]

\[
g_b(b, x_2) = \frac{g_1 b^2}{2} (1 + 2g_3 \ln(10x_2))
\]

\[
S_{NP}^{BLNY}(b, Q) = \left[ g_1 + g_2 \ln \left( \frac{Q}{2Q_0} \right) + g_1 g_3 \ln(100x_1x_2) \right] b^2
\]

where \( Q_0 \) is the energy scale over which \( S_{NP}^{BLNY} \) is needed (when \( Q \geq Q_0 \), the logarithms drop out), and \( g_1, g_2, g_3 \) are parameters determined by experiment. The term \( x_1x_2 = M^2 s = Q^2 s \), effectively removing rapidity dependence from the equation and causing \( S_{NP}^{BLNY} \) to depend only on \( Q \). The crux of this method of handling the low \( Q_T \) region is performing \( Q_T \) measurements with enough sensitivity to make measurements of the \( g_i \) parameters with significant certainty.

### 4.5.4 Leading Logarithm Method

To keep the \( R(b, Q) \) term perturbative, it is beneficial to perform a re-summation of the terms, ordering them not simply by \( \alpha_s \), but the terms that actually dominate at low \( Q_T \), which will necessarily contain factors of the singular \( \ln \left( \frac{Q^2}{Q_T^2} \right) \) and \( \frac{1}{Q_T} \) [19]. If just the singular terms of the differential cross-section are collected (and all coefficients ignored) they will collectively have the form

\[
\sum_n \sum_m (2n-1) \alpha_s^n Q_T^{-m} \ln^m \left( \frac{Q^2}{Q_T^2} \right)
\]

(4.10)
where \( n \) is summed up to sufficiently high order. The singular terms will now appear in the differential calculation as

\[
\frac{d\sigma}{dQ_T^2} \sim \frac{\alpha_s}{Q_T^2} (L + 1) + \frac{\alpha_s^2}{Q_T^2} (L^3 + L^2 + L + 1) + \frac{\alpha_s^3}{Q_T^2} (L^5 + \ldots + 1) \ldots
\]  

\((4.11)\)

where \( L = \ln \left( \frac{Q^2}{Q_T^2} \right) \). As before, the end desire is to order this equation in terms of dominant order. In this case terms of the form \( \alpha_s^n L^m \) (setting \( m = 2n - 1 \)) will be the largest contributors to the differential cross-section (described as the Leading Logarithm or LL), \( \alpha_s^n L^{m-1} \) will make up the Next-to Leading Logarithms (NLL), and \( \alpha_s^n L^{m-2} \) will make up the Next-to-Next-to Leading Logarithms (NNLL).

\[
\frac{d\sigma}{dQ_T^2} \sim (\alpha_s L + \alpha_s^2 L^3 + \alpha_s^3 L^5 + \ldots)_{\text{LL}} + (\alpha_s + \alpha_s^2 L^2 + \alpha_s^3 L^4 + \ldots)_{\text{NLL}}
\]

\[
+ (\alpha_s^2 L + \alpha_s^3 L^3 + \ldots)_{\text{NNLL}}
\]

\((4.12)\)

The (NN,N)LL terms are a re-summation from the convolution of \( J_0(bQ_T)e^{-R(b,Q)} \Sigma(b, Q, x_1, x_2) \) when integrated over \( b, y, \) and \( Q \). From this, the form of \( R(b,Q) \) is found by the coefficients of the singular terms. In \([20]\) it is shown

\[
R(b,Q) = L_R g_1(\alpha_s L_R) + g_2(\alpha_s L_R) + \frac{\alpha_s}{\pi} g_3(\alpha_s L_R) + \ldots
\]

\((4.13)\)

where \( L_R = \ln(b^2 Q^2) \) (not to be confused with \( \ln \left( \frac{Q^2}{Q_T^2} \right) \)), and \( g_i \) are functions of LL, NLL, NNLL, \ldots respectively, whose forms can be found in \([21]\). The goal of using the leading logarithm re-summation is to show that current theory can successfully describe experimental data without the need to introduce non-perturbative terms.

### 4.6 Measurements

#### 4.6.1 Resolution

Precision is an important part of any measurement. The initial barrier to most measurements is statistical, but as data increases the true limit to precision is resolution systematics. There are technical and physical limitations to any detector. As mentioned in the previous chapter, the calorimeters used to detect the daughter electrons of the Z boson rely heavily on the probabilistic behavior of showers to measure the energy of the initial particle. Within ECAL and HF the energy resolution has a definite limit that
is impossible to remove with more data or better calibration (though these too should be optimized). As the calculation of $Q_T$ includes the measured energy of the daughter electrons, this introduces similar limits on the resolution of measuring $Q_T$. For a search for the $Z$ bosons or a measurement of its total cross-section this would be easily ignored, but for a study of differential cross-section with the intention of high precision, it is an effect that should be removed if possible. Fortunately $Q_T$ is a measurement not only of energy, but direction. The positional resolution of calorimeters is quite high for electrons, as their shower width is quite narrow and detector segmentation is designed with high spatial resolution in mind. If the dependence of measurements of $Q_T$ on energy can be minimized, the positional information with $Q_T$ may still remain.

4.6.2 Novel Variables

At times when it is desirable to combine existing variables into a new quantity that includes information from one physics source, but not another. A relevant example of this is the combination of momentum and $\eta$ to form transverse momentum. Because each parton in the collision has an unknowable, largely unconstrained, $x$, the total momentum of a collision has a very broad range of possible values. Thus, while momentum is a basic physics concept and a fairly simple measurement, there are not enough constraints to make it a meaningful one at hadron colliders. Fortunately, it is known that all the net momentum, whatever it may be, is in the longitudinal direction, and that in the transverse directions net momentum is zero. So by removing part of momentum and leaving only the usefully constrain transverse portion, a variable that is very useful for analyzing physics is created, one which can be meaningfully predicted and interpreted. Occasionally a variable that is new, or ‘novel’, is created to achieve a similar result. As mentioned, electron energy resolution is a quantity that it would be advantageous to remove from the $Q_T$ analysis, while still making use of the high position resolution. What follows is the description of a set of novel variables constructed for this purpose [21].

By looking at $Q_T$ in the $x - y$ plane, $Q_T$ can be deconstructed into components with respect to the axis around which the transverse component of the momenta of the
daughter electrons are equal (fig. 4.5). This ‘thrust axis’, $\hat{t}$, is defined as

$$\hat{t} = \frac{\overrightarrow{p}_T^{(1)} - \overrightarrow{p}_T^{(2)}}{|\overrightarrow{p}_T^{(1)} - \overrightarrow{p}_T^{(2)}|}$$

(4.14)

where $\overrightarrow{p}_T^{(1,2)}$ is the $p_T$ of the (first,second) lepton. The component of $Q_T$ parallel to $\hat{t}$ is called $a_L$, while the component transverse to $\hat{t}$ is called $a_T$, mathematically stated as

$$a_L = \overrightarrow{Q}_T \cdot \hat{t}$$

$$a_T = |\overrightarrow{Q}_T \times \hat{t}|.$$  

(4.15)

---

Figure 4.5: A view of a $Z \rightarrow e^+e^-$ event in the transverse plane of the lab frame. Novel and related variables are indicated.

When the opening angle between the leptons, $\Delta\phi^{ll}$, is approximately $\frac{\pi}{2}$, the uncertainty of $a_T$ (da_T) is approximate to $d p_T \times \sin \theta_{tiny}$, where $\theta_{tiny}$ is very small. In the same situation both da_L and dQ_T are approximate to the $d p_T \times \cos \theta_{tiny} \approx d p_T$. This property carries over for all values of $\Delta\phi^{ll} > \frac{\pi}{2}$, which corresponds to approximately 99% of the total cross-section, and is the range over which $a_T$ is defined. Thus $a_T$ is a novel variable that reduces the impact of lepton $p_T$ resolution on $Q_T$ physics.

To further reduce the dependence on $p_T$ resolution $a_T$ can be divided by $Q$. Many uncertainties present in $Q$ are also in $a_T$ and thus this division removes them. From
this, a new novel variable is created ($\phi^*$). $\phi^*$ is designed to mimic $\frac{a_T}{Q}$ in the region where $\vec{p}_{T}^{(1)} \approx \vec{p}_{T}^{(2)}$, providing the relationship:

$$\frac{a_T}{Q} \approx \phi^* \equiv \tan \frac{\phi_{\text{acop}}}{2} \sin \theta^*$$  (4.16)

where $\phi_{\text{acop}} = \pi - \Delta \phi$ is the acoplanarity of the electrons ($\Delta \phi$ is the difference in $\phi$ between the electrons), and $\theta^*$ is the scattering angle, defined in terms that only use positional information,

$$\cos \theta^* = \tanh \left( \frac{\eta^- - \eta^+}{2} \right)$$  (4.17)

with $\eta^-$ and $\eta^+$ the $\eta$ of the negative and positive electrons respectively (from now on refereed to as $\Delta \eta$). While the above is the scientifically motivated definition of $\phi^*$, a simpler, equivalent definition may more clearly demonstrate the dependence only on positional resolution of the electrons

$$\phi^* = \cot \frac{\Delta \phi}{2} \frac{\Delta \eta}{\text{sech} \frac{\Delta \eta}{2}}$$  (4.18)

Using this novel variable it is possible to make transverse physics measurements with higher resolution than transverse momentum. Measurements using this variable have already been made at other experiments, such as D0 [22] and ATLAS [23].

**Comparison of $Q_T$ and $\phi^*$**

To demonstrate the advantages of $\phi^*$ over $Q_T$, a toy Monte Carlo simulation of Z bosons was created. To begin with, Z bosons were generated with a randomly assigned $Q_T$, weighted according to a Landau function fitted to data. The decay of the electrons was simulated in the rest frame, each having equal and opposite $p_T$, with a decay angle generated randomly from the Colin-Soper frame distribution $A(1 + \cos^2 \theta) + B \cos \theta$ ($A = 0.4$, $B = 0.02$). The electrons were boosted into the lab frame and the $Q_T$ and $\phi^*$ of the parent Z boson were reconstructed, to be reported as the ‘generated’ values.

Then the effects of calorimetry resolution were simulated by smearing the position and energy values of the electrons via Gaussian generated values (with a width equal to the resolution of the variable). The position resolution used was $\sigma_{\eta/\phi} = 0.03$ in $\eta/\phi$ which
is based on the position resolution of HF electrons (which far worse than the resolution in ECAL, essentially a worst case scenario). The energy resolution ($\sigma_E$) was modeled after the ECAL resolution as determined from test beam results:

$$\frac{\sigma_E}{E_e} = \sqrt{\frac{N^2}{E_e} + \frac{S^2}{\sqrt{E_e}} + C^2}$$

(4.19)

where $E_e$ is the energy of the electron, $N$ is the, $S$ is the stochastic term, and $C$ is the constant term. The values from ECAL test beam are $N = 12.4\%$, $S = 2.8\%$, and $C = 0.26\%$, which represent the best case scenario within the CMS detector (HF energy resolution is far worse). After the smearing the $Q_T$ and $\phi^*$ values recalculated, providing ‘reconstructed’ values. This was done 10,000 times. The results (fig. 4.6) show that while $\phi^*$ shows reconstruction within the statistical error bars of the generated values, $Q_T$ shows significant deviation from generated values, especially in the low $Q_T$ region.

Figure 4.6: Toy Monte Carlo Z bosons with comparison of generated values to smeared reconstruction for (a) $Q_T$ and (b) $\phi^*$ over 10,000 events. This demonstrates that $\phi^*$ variable is less affected by detector resolution, especially in the 'non-perturbative' (low $Q_T/\phi^*$) region. The plots of $Q_T$ and $\phi^*$ are over equivalent ranges.
Chapter 5

Measuring a Z Boson Differential Cross-section

As measuring a differential cross-section is the goal of this thesis, the method of doing so in an experimental context will now be presented. This chapter begins with a discussion on the electron reconstruction methods used in this analysis and the acceptance conditions placed on those electrons. This is followed by a general account of the differential cross-section measurement method and a conclusion with two published differential measurements as examples of the techniques. For the purpose of these last chapters, it will be assumed that the Z bosons will be found using only the decay to two electrons, unless otherwise stated.

5.1 Electron Selection

The identification and reconstruction of electrons must be studied and understood before the Z boson can be reliably measured. The basics of electron detection in ECAL are described in Section 2.4.1. This section will present in more detail the specific characteristics an electron candidate must have to be considered a true electron in this analysis. The reconstruction and identification of HF electrons is described in detail in Chapter 3.
5.1.1 ECAL Electron Identification

As described earlier, electrons are reconstructed as a cluster of signals in the ECAL crystals that lie along a path of constant $\eta$ but changing $\phi$ (due to the magnetic field). The grouping of these clusters of crystals forms a ‘supercluster’ in ECAL. An electron will also leave a track in the silicon tracker, the curvature of which provides a measurement of momentum and charge. By matching the momentum and path of a track from the tracker to the energy and location of a supercluster in ECAL, an electron can be reconstructed in terms of position, charge, and energy.

This analysis will place further identification requirements on electron candidates to ensure their purity. The first requirement is that it falls within the detectable regions of either EB or EE. A second requirement is that it has $p_T > 20$ GeV (as at lower energies, electrons and not-electrons are harder to differentiate). This sets the general acceptance of the electron.

Next an identification process based on several requirements is applied to ensure the electron is not only an electron but is furthermore a lone electron (as would be part of a $Z$ boson decay) and not part of a group of other particles.

The selection begins by checking to ensure that the electron is from a ‘prompt’ electron and not an electron that converted from a photon and looks for missing hits within the tracker between the electron and the vertex (indicating a photon traveled into the tracker some distance, leaving no track, before converting to an electron). Another form of conversion within the tracker is when an electron radiates a photon, which then converts into electrons. This causes a set of conversion tracks to appear near the track of the original electron. The information from these tracks (specifically the $p_T$) is needed to properly reconstruct the electron candidate. To ensure that any included conversion tracks are truly from this process, and not unrelated tracks that happen to be nearby, there is a maximum distance and angle from the original track these conversion tracks are required to be within (the $Dist$ and $\Delta \cot \theta$ variables). Further track-related variables measure the angular proximity of the ECAL supercluster to the end of the track in the tracker ($\Delta \eta_t, \Delta \phi_t$).

A final positional selection tests how narrow in $\eta$ the supercluster is. While the supercluster can have significant spread in the $\phi$ direction (due to curving in the magnetic field), an electron will have a smaller size in $\eta$ ($\sigma_{\eta_{\text{in}}}$).
A single energy-based requirement tests the relative isolation of an electron. These isolation variables look at how much energy is reconstructed in the tracker, ECAL, and HCAL near the location of the electron candidate. In an ideal isolation situation, within the localized area of the electron, the tracker would only contain the track from the electron (measurable by looking at the total $p_T$ of the tracks), ECAL would only contain transverse energy from the electron supercluster, and HCAL would have no appreciable transverse energy in it at all. An unisolated electron would have surplus of $p_T$ in the tracker and $E_T$ in ECAL and HCAL. Thus, the isolation variable is defined as:

$$I = \frac{\sum_{\Delta R<0.3} p_{T \text{track}} - p_T + \sum_{\Delta R<0.3} E_{T \text{ECAL}} - E_{T \text{SC}} + \sum_{\Delta R<0.3} E_{T \text{HCAL}} - \pi(0.3)^2 \rho}{p_T}$$  \hspace{1cm} (5.1)$$

where $I$ is the isolation variable that will be selected on, $\sum_{\Delta R<0.3}$ is the sum over a cone with radius $R = 0.3$ of the following values: the transverse momentum in the tracker ($p_{T \text{track}}$) less the transverse momentum of the electron ($p_T$), the transverse energy in ECAL ($E_{T \text{ECAL}}$) less the transverse energy of the electron supercluster in ECAL ($E_{T \text{SC}}$), and the transverse energy in HCAL ($E_{T \text{HCAL}}$). The Isolation variable is very sensitive to the additional energy in the detectors due to pile-up. The solution is to subtract a pile-up term from the summations before dividing by $p_T$. This term is $\pi(0.3)^2 \rho$ where $\pi(0.3)^2$ is the area of the cone, and $\rho$ is a term representing the energy per area due to pile-up. The value of $\rho$ is found from a collection of jets (labeled hltKT6CaloJets). The symbol $\rho$ is the average unclustered energy in the region of the jets and provides a measurement of the average energy density from pile-up.

All of these variables and requirements are collected and the values of each are tuned in concert to produce a set of ‘working points’ that are defined by the efficiency of electron reconstruction. For example, working point 80 (WP80) is a working point that nominally keeps 80% of true electrons. The values used for WP80 with 2011 data are reported in Table 5.1.

### 5.2 Acceptance

In order to compare an analysis with theory or other experimental measurements, it is important to define what types of “ideal” events are a part of the measurement,
Table 5.1: A table of values used for the working point requirements for WP80 for each ECAL detector as used for 2011 data.

Acceptance can be further set by requirements placed on electron kinematics (e.g. $p_T > 20 \text{ GeV}$) and by the physical location in $\eta$ and $\phi$ of the electrons within the detectors. By including electron properties in the acceptance, the definition of a $Z \to e^+e^-$ event can be further refined for a more direct (and hopefully precise) comparison to theory. For the final $\phi^*$ analysis, acceptance is defined by the mass requirement, $p_T$ limitations on electrons, and $\eta$ requirements (which will be directly related to which subdetectors detect the electrons).

### 5.3 Differential Cross-section

#### 5.3.1 Data Driven Definition of Differential Cross-section

The measurement of differential cross-section with data (or simulation) is performed via the bin-by-bin equation

$$
\frac{1}{\sigma} \frac{d\sigma(Z \to e^+e^-)}{dY_k} = \frac{(\epsilon_{\text{Acc}})}{(N - B)} \left[ \frac{\Sigma_i M_i^k (N_i - B_i)}{\Delta_i (\epsilon_{\text{Acc}})_i} \right]
$$

(5.2)
where $k$ is the enumeration of bins in a generalized variable $X$, $M_k^i$ is a bin-migration correction term that corrects for the probability of $Z$ boson in the $k$th bin of the variable being instead reconstructed into the $i$th bin of the variable, $\Delta_i$ is the bin width, $N_i$ is the number of events measured in the $i$th bin, $B_i$ is the estimated number of background events in the $i$th bin, and $(\epsilon_{Acc})_i$ is the efficiency times acceptance of the bin. The symbols $N$, $B$, and $\epsilon_{Acc}$ represent those same values evaluated over the total range of $X$. Subsequently, while the reconstruction of $X$ (which determines the $N$ variable) is of most importance, the determination of efficiency times acceptance and estimation of background will have a significant effect on the final measurement. While the values of $N$ will be taken from data, $B$ and $\epsilon_{Acc}$ will make use of a combination of data and simulated events. The cross-section, $\sigma$, will not be the total $Z \rightarrow e^+e^-$ cross-section, but will instead be determined by the acceptance used.
5.4 Z Boson Detection

5.4.1 Efficiency

The ability to find electrons in the CMS detector depends on them arriving in parts of the detector that can be detected (the fiducial acceptance) and the percentage of detectable electrons in CMS that pass identification requirements (the efficiency). Accurately understanding and applying these values allow a comparison of data (subject to these limitations) and theory (which fundamentally has perfect acceptance and efficiency). For the purposes of this analysis, within the $Z \rightarrow e^+e^-$ acceptance chosen (section 5.2) the fiducial acceptance is always 100%.

Conceptually, the efficiency for a given acceptance ($\epsilon_{\text{Acc}}$) provides what percentage of Z bosons that will be seen as a function of the differential variable, $X$.

$$\epsilon_{\text{Acc}}(X) = \int P(e_1, e_2; X) \times \epsilon(e_1, e_2) \times T(e_1, e_2) \, de_1 \, de_2 \quad (5.3)$$

where $P(e_1, e_2; X)$ is the probability of a Z with a specific $X$ creating daughter electrons with properties $e$ (which can include $\eta, \phi, p_T$, etc.), $\epsilon(e_1, e_2)$ is the efficiency reconstructing an electron pair with such properties, and $T(e_1, e_2)$ is the efficiencies of the triggers of the detector at recording a specific event.

$P(e_1, e_2; X)$ is present in the generation of events, either literally in the simulation code or as an abstraction of reality for data. For simulation, simply producing a generated plot of $X$ with weighting each event by its corresponding efficiencies will produce the $\epsilon_{\text{Acc}}$ plot. For data, $\epsilon_{\text{Acc}}$ can be “removed” to give a presentation of ideal results by either dividing the data bin-by-bin from by the simulated $\epsilon_{\text{Acc}}$ per said bin. Often the $\epsilon_{\text{Acc}}$ will not be calculated from initial acceptance to final identification in one step. Instead, $\epsilon_{\text{Acc}}$ calculation will usually be broken up in several progressive stages (usually because different data sets were needed to calculate individual types of efficiencies), and will include correction factors when simulation is used in part of the calculation.

Tag-and-Probe

The goal of an electron efficiency is to find how likely a single electron with a specified $p_T, \eta$, and pile-up is to be reconstructed into a passing $Z \rightarrow e^+e^-$ event. However, there are always two electrons in any such event. To find the single electron efficiency
of a di-electron object, a method of testing efficiencies call ‘Tag-and-Probe’ is used. In Tag-and-Probe, a $Z \rightarrow e^+e^-$ event is divided in two: one electron is used as a reference because it is a good electron (the Tag) and one is the actual target of the test (the Probe). Tag electrons are required to pass all the electron reconstruction requirements, proving that it is a good electron to test with (sometimes called a ‘golden’ electron), while the probe only needs to pass the requirements leading up to the one being tested. For an efficiency calculation of a general requirement, $N_{\text{total}}$ is the collection of probes (with matching tags) that have survived all previous requirements, and $N_{\text{pass}}$ is the subset of $N_{\text{total}}$ that passes that requirement, with the efficiency being $\epsilon = \frac{N_{\text{pass}}}{N_{\text{total}}}$. By grouping probe electrons into ‘bins’ of $p_T$, $\eta$, and pile-up, the average efficiency of an electron in that bin can be calculated.

### 5.4.2 Backgrounds

Sources of background are events that produce reconstructed electrons that pass the identification requirements, while not being legitimate $Z \rightarrow e^+e^-$. These may be actual electrons or other particles that fake electron response in the calorimeters. It is impossible to isolate this from the data directly, so the contributions to background must be studied separate from $Z \rightarrow e^+e^-$ to determine its impact. The most prevalent sources of background are from $Z \rightarrow \tau\tau$, $W \rightarrow l\nu$, diboson events ($ZZ, ZW, WW$), $t\bar{t}$-bar, and QCD multi-jets. Monte Carlo simulations are used to estimate all backgrounds except for the QCD. QCD background is found using experimental data fitted with a background shape function, and will be described in further detail in Section 6.2. The background events are fit into bins in $X$ and scaled to the luminosity of the data. These form the value of $B$ that will be subtracted bin by bin in the final calculation.

### 5.5 Example Differential Cross-section Measurements

CMS has published measurements of the single differential cross-section of $Z$ bosons as a function of rapidity and transverse momentum. For these measurements, both $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ events were used and their results combined. The results demonstrate both the usefulness of differential cross-sections and the motivation to perform a $\phi^*$ variable analysis. The paper also illustrates the techniques that will be
used in the $\phi^*$ measurement.

5.5.1 Data Set and Event Reconstruction

The data consisted of 36 pb$^{-1}$ of events taken in 2010 by the CMS experiment at center-of-mass energy of 7 TeV. Both measurements required that the leptons (electrons or muons) had $p_T > 20$ GeV and an invariant mass between 60 and 120 GeV. The muons had a geometric acceptance of $|\eta| < 2.1$, while electrons were reconstructed over the full range of ECAL and HF ($|\eta| < 4.6$), excluding trackless regions and detector gaps. The methods and requirements of reconstruction and identification were the same as for 2011 data, with a few simplifications due to lack of pile-up. Bin migration matrices were used to correct for shifts between expected and reconstructed values, as seen in simulation (fig. 5.2).

Figure 5.2: The bin migration for the differential $Q_T$ measurement. Note that it is especially nontrivial at low values where $\phi^*$ provides a much higher resolution.
Figure 5.3: The differential cross-section for Z bosons as a function of absolute rapidity. The combined results extends only to $|Y| < 2$. Results beyond this point are electron only, using HF electrons. The shaded area represents the predicted simulation with the the uncertainties of the CT10 PDFs.
5.5.2 Results

Rapidity results

Due to $\eta$ restrictions on muons, the region for the combination of muon and electron measurements was limited to $|y| < 2.0$. For this combined region, muons and electrons agreed favorably ($\chi^2 = 0.85$). The use of HF electrons allowed the range of the electron-only measurement to be extended to $|y| < 3.5$, which covers 95% of the total expected distribution. The complete rapidity results were then normalized to the cross-section of all $Z \rightarrow e^+e^-, \mu^+\mu^-$ events within the rapidity range (fig. 5.3).

As previously discussed (in section 4.3.3), the rapidity of a $Z$ boson directly probes the PDF of the parent partons. The results of this measurement was compared to several different sets of model PDFs and was able to provide feedback on them. In general, PDF sets rely on eigenvectors and/or model dependencies for their tuning. Of the PDF sets tested, CT10 and MSTW2008 depend only on eigenvectors, while HERAPDF 1.5 uses both eigenvectors and model dependencies. To test a PDF sets, the $\chi^2$ between the base set and the rapidity measurement was found. Then the each component (eigenvector or model dependency) was varied by one standard deviation around its default variable. The difference in each separate $\chi^2$ from the base indicates to which parts of the model set the measurement is sensitive to (fig. 5.4). The $\chi^2$ from the base sets of CT10, MSTW2008, and HERAPDF 1.5 were 18.5, 18.3, and 18.4 respectively. For all sets, several eigenvectors showed sensitivity to the result, CT10 showed greatest. For HERAPDF, the result was seen to be sensitive to model dependencies as well, most notably the Strange-quark PDF as a fraction of the Down-quark PDF.

A slightly different type of PDF model, NNPDF 2.0 was also considered. It functions by making various replicas of PDF sets sampled from the same space. The $\chi^2$ with the base prediction was 18.4, and a majority of the 100 standard replica sets were similar. There were several results with higher values of $\chi^2$ (up to 34.5), which indicates that these replicas significantly disagree with the rapidity measurement.
Figure 5.4: The change in $\chi^2$ values between the differential rapidity results and predictions from the NLO HERAPDF 1.5 PDF set for the eigen vectors (upper plot) and model parameters (lower plot). The greatest disagreement between the two is in the modeling of strange-quark PDF as a fraction of the down-quark-sea PDF (that is, the relative amplitude of strange quarks in the sea). Other PDFs compared in the paper (but not shown here) were CT10 PDF, MSTW2008, and NNPDF.
5.5.3 Transverse Momentum Results

For the transverse momentum measurement, only the combined results were reported and thus to match the range of muons, only electrons with $|\eta| < 2.1$ were used. The results were reported for $Q_T < 600$ GeV, and are normalized to the cross-section integrated over the phase-space of the di-lepton acceptance (i.e. both leptons having $p_T > 20$ GeV, $|\eta| < 2.1$). The two particles were found to be compatible over the $Q_T$ range measured (reduced $\chi^2 = 0.74$). The total result was compared to the prediction of POWHEG+PYTHIA with parameters set to the Z2 tune with a $\chi^2$/ndof=19.1/9.

At low values of $Q_T$ ($Q_T < 30$, fig. 5.5), where nonperturbative QCD dominates, POWHEG+PYTHIA had poor agreement with the measurement ($\chi^2$/ndof=76.3/8). To search for models with better agreement, three PYTHIA tunes were investigated: Z2, ProQ20, and Perugia 2011. The Z2 tune is based on LHC data and strongly depends on the underlying event to generate transverse momentum, Perugia 2011 relies strongly on ISR and FSR as well as color reconnection, while ProQ20 uses modeled ‘inherent’ transverse momentum of the partons for simulating $Q_T$. Of the tunes, the PYTHIA Z2 and ProQ20 tunes provided greatest agreement ($\chi^2$/ndof = 9.4/8 and 13.3/8, respectively), while the Perugia 2011 tune did not agree well ($\chi^2$/ndof = 48.8/8).

At high $Q_T$ (fig. 5.6), where the calculations are perturbative, the alternative package ‘Fully Exclusive W,Z Production’ (FEWZ) was also compared to the measurement. FEWZ bases its calculations on a dynamic scale ($\sqrt{M_Z^2 + Q_T^2}$) rather than a fixed scale based on the mass of the Z boson. The FEWZ prediction had a much poorer agreement ($\chi^2$/ndof = 30.5/8) than that of POWHEG+PYTHIA.

In the end, no single tune was found to correctly describe both low and high regions simultaneously. Furthermore, in the low $Q_T$ regions, resolution is particularly low, while the various tunes vary greatly. By probing this region with a higher resolution variable ($\phi^*$) a better ability to compare to theory will be gained.
Figure 5.5: The combined differential cross-section for transverse momentum, focusing on $Q_T < 30$ GeV. The colored lines represent various simulation configurations that have been used. The lower plot represents the comparison of these simulations to data divided by the uncertainty ($\delta$) on the data. The outer and inner bands represent $\pm 1\delta$ add $\pm 2\delta$ respectively.
Figure 5.6: The combined differential cross-section for transverse momentum, focusing on $Q_T > 20$ GeV. The shaded regions represent proposed simulation configurations. The lower plot shows the ratio between the data and the theory predictions. The bands in the lower plot represent the one standard deviation combined theoretical and experimental uncertainties.
Chapter 6

Differential $\phi^*$ Analysis

This chapter deals with the specific methods used to make the measurement of the $\phi^*$ observable. The data and simulation used to calculate the measurement and related values will be presented, along with the definitions of the $Z \rightarrow e^+e^-$ acceptances over which the measurements will be made. This chapter then provides a focused look at the calculations of both the electron efficiencies and the systematic uncertainties of the measurement in each of the acceptance regions.

6.1 Differential Cross-section with $\phi^*$

For the differential cross-section of $\phi^*$, the matrix term (see Equation 5.2) is removed due to the higher resolution of the $\phi^*$ variable.

$$\frac{1}{\sigma} \frac{d\sigma(Z \rightarrow e^+e^-)}{d\phi^*_k} = \frac{(\epsilon_{Acc})}{(N - B)} \frac{(N_k - B_k)}{\Delta_k(\epsilon_{Acc})_k}$$  \hspace{1cm} (6.1)

We consider the residual bin migration as a systematic uncertainty.

6.1.1 Z Boson Definitions

There are three sub-detectors capable of finding electrons, the ECAL Barrel (EB), the ECAL End-cap (EE) and the Forward HCAL (HF). Of the possible combinations, three definitions of Z bosons are chosen for this analysis for the coverage of rapidity they provide. These three definitions are Z bosons with both electrons in EB, Z bosons...
with one electron in EB and one in EE, and Z bosons with one electron in EE and one in HF (EB-EB, EE-EB, and EE-HF respectively). The range of $\eta$ allowed for electrons in each sub-detector and the rapidity range for each Z boson definition can be seen in Table 6.1. The two ECAL Z boson definitions provide high resolution reconstruction in overlapping rapidity coverage. The resolution in HF is less than its ECAL counterparts, but it also has the greatest rapidity coverage. With these divisions in the $\phi^*$ analysis, different parts of the Z spectrum can be probed and possible differing behaviors can be seen. Once constructed, the $\phi^*$ of the Z boson is calculated from the daughter electron properties as discussed in the previous chapter.

<table>
<thead>
<tr>
<th>Sub-detector</th>
<th>Range in $\eta_e$</th>
<th>Z Defintion</th>
<th>Range in $y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECAL Barrel (EB)</td>
<td>$</td>
<td>\eta_e</td>
<td>&lt; 1.444$</td>
</tr>
<tr>
<td>ECAL Endcap (EE)</td>
<td>$1.566 &lt;</td>
<td>\eta_e</td>
<td>&lt; 2.50$</td>
</tr>
<tr>
<td>Forward HCAL (HF)</td>
<td>$3.10 &lt;</td>
<td>\eta_e</td>
<td>&lt; 4.60$</td>
</tr>
</tbody>
</table>

Table 6.1: The sub-detectors and corresponding electron pseudorapidity acceptance ($\eta_e$) used for reconstruction of electrons, along with the corresponding Z boson definition with their resulting $y$ ranges. A further requirement is that all electrons have $p_T > 20$ GeV and the Z boson mass is between 60 and 120 GeV.

**Electron Selection**

The selection of ECAL electrons used in the Z boson definitions requires electrons that pass the WP80 ID requirements and trigger the `HLT_Ele17_CaloIdL_CaloIsoVL_Ele8_CaloIdL_CaloIsoVL` high level trigger. The selection of an HF electron for the EE-HF definition requires HF electron that passes the medium requirements for HF electron identification and that matches the HF trigger in the HLT trigger of `HLT_Ele17_CaloIdL_CaloIsoVL_Ele15_HFL` with the ECAL trigger matching the selected EE electron. All electrons must have $p_T > 20$ GeV and fall within the geometric acceptance of their detector.

**Data and Simulation**

Data used were from the 2011A run at the LHC with center of mass ($\sqrt{s}$) of 7 TeV. These runs were chosen from the data set `DoubleElectron` at times of minimal pile-up (see fig. 6.1) and total 216 pb$^{-1}$ in integrated luminosity. The `DoubleElectron` set was
formed using unprescaled double-electron triggers with prescaled single electron triggers for study purposes.

Figure 6.1: The number of events per pile-up for the 2011A DoubleElectron 216pb$^{-1}$ data set.

The simulation used was from the Summer 2011 Monte Carlo production by the CMS group. The signal is a fully simulated sample of $Z \rightarrow e^+e^-$ generated with POWHEG [25] and interfaced with PYTHIA (v. 6.422) [26] for the underlying event, using the CT10 PDF parametrization [27] and the Z2 tune [28] (DYToEE_M-20_CT10_TuneZ2_7TeV-powh
eg-pythia, or simply SignalMC). Several generator-only samples of signal were made using only PYTHIA (v. 6.422) for various tunes. The tunes for the alternative samples are Z2, D6T (the default tune, pre-2010), Perugia [29], ProPT, and ProQ2 [30]. Background event simulation were productions of $t\bar{t}$ and $Z \rightarrow \tau\tau$ generated via PYTHIA. Table 6.2 includes the relevant details for each simulation set.

<table>
<thead>
<tr>
<th>Process</th>
<th>Simulation Set</th>
<th>Cross-Section</th>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>SignalMC</td>
<td>DYToEE_M-20_CT10_TuneZ2_7TeV-powheg-pythia</td>
<td>1300 pb</td>
<td>4524068</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>TT_TuneZ2_7TeV-pythia6-tauola</td>
<td>94 pb</td>
<td>1089625</td>
</tr>
<tr>
<td>$Z \rightarrow \tau\tau$</td>
<td>DYToTauTauM-20_TuneZ2_7TeV-pythia6-tauola</td>
<td>1300 pb</td>
<td>2032536</td>
</tr>
<tr>
<td>Alternative Tunes</td>
<td>Generator-Level: Z2, D6T, Perugia, ProPT, ProQ2</td>
<td>1300 pb</td>
<td>10000000 (for each)</td>
</tr>
</tbody>
</table>

Table 6.2: Monte Carlo Simulations used for signal, background estimation, and alternative tunes for signal (for use at the generator level).

### 6.1.2 Binning

Defining the binning (the ranges of $\phi^*$ that count as one differential value) for a differential cross-section is a very important step, one that depends on several decisions made by the experimenters. There is no singular, optimal way to chose a binning, but many wrong ones if done carelessly or even too cautiously. The ranges of $\phi^*$ chosen for the binning depends on several factors: the number of events, the resolution of the measurement, and the behavior of the variable.

The number of events (both per bin and in total) is the dominant factor to consider. If there are too few events total, any binning chosen will be dominated by statistical uncertainty. Likewise, even if there are a large number of total events, too small of binning will create a similar statistical error in data presentation. Too minimize statistical impact, bins should be chosen to minimize $\sqrt{\frac{N}{E}}$, ideally to be trivial compared to other uncertainties. In this analysis, the total number of events is set by the data set chosen, so only the number of events per bin can be varied by changing the bin size.

Resolution of a measurement plays an important role due to systematic uncertainties in positional resolution that cause the true $\phi^*$ of $Z$ boson to not be precisely reconstructed. If binned too finely, these uncertainties will randomly spread events out from their proper bin and into neighboring bins (this is called ‘bin-migration’). Solutions
to bin-migration are to create bins large enough to reduce this effect to a negligible percentage or to create a matrix of bin-migration probabilities and include this effect in the final measurements. A large motivation of developing the $\phi^*$ novel variable is because it will have much better energy resolution than $Q_T$ would have, increasing its overall resolution greatly.

The shape of the distribution is important because too coarse of a binning will flatten out shapes, losing important information. Ideally, sections with high slopes, peaks or other interesting phenomena should be binned finely-enough to see the relevant amount of detail about said slope or peak. If binning is too coarse, detailed study of the underlying physics is smoothed away, removing the ability to test theory. This balances out the motivation for larger bin ranges by statistic and resolution effects. As the most interesting shape in $\phi^*$ occurs at low values (and also the greatest statistics) the low values of $\phi^*$ will be finely binned, while the greater values (both flatter in shape and lower in statistics) will be more coarsely binned.

Finally, a practical consideration when choosing binning is what binnings have been used in similar measurements of the same phenomena. This is helpful to the community for the purpose of easily comparing results of equivalent measurements by different groups.

Ideally, the binning chosen will have low $\sqrt{N}$, no need for a bin migration matrix, and still present fine enough detail to perform the desired level of detail in the analysis of $\phi^*$. The final binning used in this analysis can be seen in Table 6.3. The same binning is used in all three acceptance regions.

6.2 Electron Efficiencies

In the 2011 data sample, the use of di-lepton triggers introduces a correlation between the two leptons of a Z boson decay and makes the $\epsilon_{\text{Acc}}$ dependent on the joint characteristics of the di-lepton pair. Due to this, simulation was used more heavily in the $\epsilon_{\text{Acc}}$ calculation than in the previous analyses described in Chapter 5. The single electron efficiencies are broken up into bins by electron $\eta$, $p_T$, and pile-up.
<table>
<thead>
<tr>
<th>$\phi^*$ Range</th>
<th>Events: EB-EB</th>
<th>EE-EE</th>
<th>EE-HF</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000-0.008</td>
<td>2325</td>
<td>1741</td>
<td>836</td>
</tr>
<tr>
<td>0.008-0.016</td>
<td>2290</td>
<td>1732</td>
<td>818</td>
</tr>
<tr>
<td>0.016-0.024</td>
<td>2215</td>
<td>1658</td>
<td>788</td>
</tr>
<tr>
<td>0.024-0.034</td>
<td>2595</td>
<td>1923</td>
<td>908</td>
</tr>
<tr>
<td>0.034-0.045</td>
<td>2597</td>
<td>1926</td>
<td>887</td>
</tr>
<tr>
<td>0.045-0.057</td>
<td>2515</td>
<td>1809</td>
<td>835</td>
</tr>
<tr>
<td>0.057-0.072</td>
<td>2660</td>
<td>1887</td>
<td>860</td>
</tr>
<tr>
<td>0.072-0.091</td>
<td>2710</td>
<td>1923</td>
<td>812</td>
</tr>
<tr>
<td>0.091-0.114</td>
<td>2515</td>
<td>1727</td>
<td>714</td>
</tr>
<tr>
<td>0.114-0.145</td>
<td>2487</td>
<td>1687</td>
<td>649</td>
</tr>
<tr>
<td>0.145-0.189</td>
<td>2439</td>
<td>1569</td>
<td>588</td>
</tr>
<tr>
<td>0.189-0.258</td>
<td>2362</td>
<td>1487</td>
<td>509</td>
</tr>
<tr>
<td>0.258-0.391</td>
<td>2317</td>
<td>1401</td>
<td>441</td>
</tr>
<tr>
<td>0.391-0.524</td>
<td>1093</td>
<td>659</td>
<td>180</td>
</tr>
<tr>
<td>0.524-1.000</td>
<td>1275</td>
<td>747</td>
<td>184</td>
</tr>
</tbody>
</table>

Table 6.3: The binning in $\phi^*$ with a the expected events per bin at a luminosity of 216pb$^{-1}$, determined from simulation at each generated acceptance level.

When determining efficiencies from data the equation given in section 5.4.1 is modified to be

$$\epsilon = \frac{N_{\text{pass}} - B_{\text{pass}}}{N_{\text{total}} - B_{\text{total}}}$$

(6.2)

where the variable $B$ represents the background events that pass the requirements to be in that set (in general the $B$ term will not be shown in future equations, but is assumed to be there in all data based calculations). The number of events included in a group are calculated using the Tag-and-Probe method described in section 5.4.1. Often total efficiencies must be broken up into efficiencies of several separate cuts. This is most often done when multiple data sets are needed to calculate different steps in the efficiency process. These individual efficiencies will form a chain of terms that will conceptually cancel to re-form the original efficiency ($\frac{N_{\text{pass}}}{N_0} \times \frac{N_0}{N_{\text{total}}} = \frac{N_{\text{pass}}}{N_{\text{total}}}$).

To properly account for the effect of di-lepton correlations, the total ECAL electron efficiency will be fully simulated, and then matched to data by means of correction factors. Correction factors are very similar to efficiencies in how they are constructed and applied, but may have values greater than one. A correction factor includes a term
derived from data and a term derived from simulation in the form

$$F_{\text{cor}} = \left[ \frac{N_{\text{pass}}}{N_{\text{total}}} \right]_{\text{data}} \times \left[ \frac{N_{\text{pass}}}{N_{\text{total}}} \right]^{-1}_{\text{MC}}$$  \hspace{1cm} (6.3)

with the terms conceptually canceling with each other leaving only the factor needed to convert an efficiency based on $N_{\text{pass}}$ from simulation to data.

**Calculating QCD Background**

To find the number of background events in a specific efficiency bin, the invariant mass plot is made, and a two-part function is fit to it (fig. 6.2). The first function is the distribution of signal events within a defined efficiency bin, normalized over the range of 60 to 120 GeV. The second is a background estimation function (based on typical QCD background shapes) of the form:

$$\alpha \left( \text{erfc} \left( \frac{\beta - m}{\delta} \right) e^{-\gamma m} \right)$$  \hspace{1cm} (6.4)

where $m$ is the invariant mass and $\alpha$, $\beta$, $\gamma$, and $\delta$ are variables that are used to fit the function. The number of signal events ($N - B$) can be computed two ways from this fit; it can be the integral of the signal function or the number of total events minus the integral of the background function. This analysis uses the average of these two methods for the final number of signal events.

### 6.2.1 ECAL Electron Efficiency

For the ECAL electrons, the calculation for the efficiency of reconstructing and identifying a single electron ($\epsilon_e$) is determined by fully simulating the efficiency of the WP80 identification requirement, with a correction factor to match it to data accounting for triggering efficiencies and differences in WP80 identification between data and simulation. This is represented by:

$$\epsilon_e = \left[ \frac{N(\text{WP80})}{N(\text{Acceptance})} \right]_{\text{MC}} \times \left[ \frac{N(\text{TID})}{N(\text{GSF})} \right]_{\text{data}} \times \left[ \frac{N(\text{WP80})}{N(\text{TID})} \right]_{\text{data}} \times \left[ \frac{N(\text{WP80})}{N(\text{GSF})} \right]^{-1}_{\text{MC}}$$  \hspace{1cm} (6.5)

where $N(X)$ is the number of electrons in category $X$, as found by Tag-and-Probe. ‘Acceptance’ is the group of electrons passing acceptance into the specific Z boson
Figure 6.2: A typical efficiency calculation plot used in the WP80-data efficiency calculation for probe electrons with $p_T$ between 45 and 50 GeV and a pile-up less than 5. The graphs represent the data (the circle points) matched with a combined fit (the black solid line) made up of combined background (black dashed line) and signal (red dashed line). The graph on the left is before the WP80 requirements are made, the one on the right is after. The number of electrons in each graph is taken to be the average between the integral of the signal and the number of events minus the integral of the background. The single electron efficiency for this bin is the ratio of number of electrons on the right to the number of electrons on the left.

definition, ‘GSF’ is the group of electrons that pass GSF electron identification, ‘TID’ is the group of electrons that pass trigger requirements, ‘WP80’ is the group of electrons that pass WP80 requirements. The $\left[ \frac{N(\text{WP80})}{N(\text{Acceptance})} \right]_{MC}$ term represents the full simulation of the total electron reconstruction and identification efficiency. As the single electron triggers are prescaled, insufficient events are available to measure $\left[ \frac{N(\text{WP80})}{N(\text{GSF})} \right]_{data}$ directly. We instead break the correction factor into two stages, each derived separately from data, Trigger Efficiency ($\left[ \frac{N(\text{TID})}{N(\text{GSF})} \right]_{data}$) and Electron Identification ($\left[ \frac{N(\text{WP80})}{N(\text{TID})} \right]_{data}$).

**ECAL Trigger Efficiency Correction Factor**

The Trigger Efficiency correction factor of ECAL electrons accounts for the probability that an electron in data will activate a high level di-electron trigger. If the trigger is
not passed, the event is not stored, which leading to a loss in efficiency. As the passing of triggers is not a part of the full simulation used, a data driven correction factor is applied.

To find the trigger efficiency from data a set is formed from events passing a lax, single electron trigger. Using the tag-and-probe, the electron that matches the trigger is used as the tag and the second, untriggered electron (if there is one) that passes GSF requirements is the probe. This forms the base group \( N(GSF) \). The group is then passed through the di-lepton triggers used for the analysis and the probes that survive form \( N(TID) \). This provides the correction for the effects of trigger efficiency to the calculation (fig. 6.3).

![Correction Factors](image)

**Figure 6.3:** The correction factors for the Trigger Efficiency stage of ECAL electron efficiency calculation, binned in \( p_T \) and \( \eta \) for (a) low (0-4) and (b) high (5-100) pile-up.

**ECAL Electron Identification Correction Factor**

The Electron Identification correction factor contains two terms: a data term which covers the single electron efficiency of WP80 identification from triggered electrons, and a simulated term which is the inverse of the single electron efficiency of simulated WP80 identification from GSF electrons. When combined, the dilepton biases in each of the WP80 identification efficiencies will effectively cancel, leaving a ratio of data-based
electron identification efficiency to simulation-based efficiency (fig. 6.4). By combining with the Trigger Efficiency stage, the respective terms of GSF and TID “cancel” creating a single total correction factor that will convert the fully simulated efficiency to properly match the total reconstruction and identification efficiency of data (fig. 6.5).

Figure 6.4: The correction factors for the Electron Identification stage of ECAL electron efficiency calculation, binned in $p_T$, $\eta$ for (a) low (0-4) and (b) high (5-100) pile-up.

### 6.2.2 HF Electron Efficiency

For the ECAL-HF Z bosons, the reconstruction efficiency of a single HF electron ($\epsilon_e$) is determined completely from data as

$$\epsilon_e = \left[ \frac{N(\text{TID})}{N(\text{Acceptance})} \right]_{\text{data}} \times \left[ \frac{N(\text{HFID})}{N(\text{TID})} \right]_{\text{data}}$$

(6.6)

where TID is the electrons that pass HF triggering requirements, and HFID is the group of electrons that pass the tight HF electron requirements. As HF electron efficiency is calculated directly from data, no corrections terms are used, rather it is electron efficiency itself that is made up of two stages, Trigger Efficiency ($\left[ \frac{N(\text{TID})}{N(\text{Acceptance})} \right]_{\text{data}}$) and Electron Identification ($\left[ \frac{N(\text{HFID})}{N(\text{TID})} \right]_{\text{data}}$).
Figure 6.5: The total correction factor applied to ECAL electron efficiency to match the simulated efficiency to experimental data, binned in $p_T$, $\eta$ for (a) low (0-4) and (b) high (5-100) pile-up.

**HF Trigger Efficiency**

The trigger efficiency of HF electrons is the probability that a supercluster in HF will pass the trigger requirements to become a stored event. Not every super-clustered HF electron passes the di-electron HF trigger. If that trigger is not passed, the event will not be stored, and thus there is a potential efficiency term in HF electron reconstruction due to the trigger. The method of calculating this efficiency is to find a super-clustered HF electron present in an event that passes a single (ECAL) electron trigger, and then apply the requirement that it passes the double electron trigger and match the HF object passing the trigger. When determining which single electron trigger to use, a difficulty arose in finding enough data statistics as many of the single electron triggers were greatly suppressed (prescaled) in the 2011A runs. The trigger used was $\text{HLT}_e27\_\text{CaloIdVT}\_\text{CaloIsoT}\_\text{TrkIdT}\_\text{TrkIsoT}$ which had the highest statistics due to not being prescaled, but did place a requirement of $p_T > 27$ GeV on the ECAL electron, which eliminates some of the acceptance space of EE-HF. Other triggers with lower $p_T$ requirements were considered, but in the end they did not provide enough statistics. To ameliorate the effects that the ECAL $p_T$ requirements would have on the efficiency of
the HF electron, the calculation for HF Trigger efficiency was binned in $\phi^*$ rather than $p_T$ (fig. 6.6).

![Figure 6.6](image)

Figure 6.6: The efficiency for HF TID requirements, binned in $\phi^*$ and $\eta$ for (a) low (0-4) and (b) high (5-100) pile-up.

**HF Electron Identification Efficiency**

The identification efficiency of HF electrons is the probability that a supercluster passes the tight HF electron identification requirements (fig. 6.7). This efficiency was taken directly from the 2011A data set, simply using tag-and-probe to compare events that passed TID to events that passed HFEID.

6.2.3 Final Efficiency

The cumulative efficiencies at each stage can be seen in Figure 6.8. The acceptance level includes the positional requirements and $p_T > 20$ GeV cut on the generated values of the electrons. The reconstructed $p_T > 20$ GeV line simply shows the amount lost over the course of full simulation and reconstruction. For the ECAL electrons, the Trigger Efficiency stage has that correction factor applied by itself to the cumulative efficiency but the Electron Identification stage convolutes that correction factor with the fully
Figure 6.7: The efficiency for HF electron ID requirements, binned in $p_T$ and $\eta$ for (a) low (0-4) and (b) high (5-100) pile-up.
Figure 6.8: Final cumulative efficiencies per stage for (a) EB-EB, (b) EE-EB, and (c) EE-HF in $\phi^*$. 
simulated identification as part of the calculation method of the final $\epsilon_{\text{Acc}}$.

The efficiencies are almost independent of $\phi^*$, varying notably only between the different Z boson definitions. This stability in efficiency is an additional benefit of the $\phi^*$ variable.

### 6.3 Background Estimation

Background events can come from jets that fake their way past the identification requirements, electrons that are not from $Z \rightarrow e^+e^-$, or a combination of the two. Possible background signals that were investigated were $t\bar{t}$, $Z \rightarrow \tau\tau$, and QCD di-jets. QCD di-jets are found to be the most common source of background and an estimate of its numbers are obtainable from data in the same manner as described in section 6.2. The $t\bar{t}$ and $Z \rightarrow \tau\tau$ backgrounds are simulated from PYTHIA generated Monte Carlo (presented previously in Table 6.2). The dominant background is QCD, with $t\bar{t}$ becoming a contributing factor in higher bins of $\phi^*$ (fig. 6.9).

### 6.4 Systematic Uncertainties

Uncertainties on the values of the differential measurement come from several systematic sources: bin migration effects, uncertainties of the efficiencies used (from both systematic and statistical sources), uncertainties in background estimation, and uncertainties in the energy scale, position, and alignment of the detector. As is appropriate for a differential measurement, the fractional uncertainty per bin ($\delta_i$) determined from simulation is:

$$\delta_i = \frac{|N_i^+ - N_i^0|}{N_i^0}$$

where $N_i^+$ is the number of events in the $i$th bin of the plot that has been varied by the uncertainty, $N_i^+$ is the total events in that plot, $N_i^0$ is the number of events in the $i$th bin of the base plot, and $N_T^0$ is the total events in the base plot. The total systematic uncertainty is the sum in quadrature of all individual components.
Figure 6.9: Estimated background as a fraction of total events per bin in the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions.
Statistical Uncertainties in Simulation

Because of the finite size of the SignalMC set, the determination of the systematic uncertainties themselves have an uncertainty. The resulting fractional uncertainty (fig. [6.10]) is present in all systematic uncertainty calculations made using SignalMC, but will not be shown on the related plots as it has significant effect in the final uncertainty calculations.

6.4.1 Bin Migration Uncertainty

Due to the resolution of the detector the $\phi^*$ reconstructed may not be in the same bin as its true $\phi^*$, the number of events in a bin will change from its ‘correct’ value as some events are incorrectly reconstructed outside of it, and some that should be outside are incorrectly reconstructed inside. The net effect of this migration (number of events gained less the number lost) is ideally zero, or at least a very low fraction of events. If there is a bin correlated effect to the migration, certain bins gain significant numbers of events while other bins will lose events, there becomes an impacting uncertainty to which bin an event should be reconstructed in. Often measurements will use a matrix-based correction term to account for this effect, but as the statistical limitations are greater than these effects, they will be treated simply as systematic effect, though for completeness the bin migration matrices will be provided in Appendix A.1 Tables A.1, A.2 and A.3.

As this uncertainty is directly related to the the $\phi^*$ variable, it is thus greatly affected by the underlying physics of $\phi^*$ (and thus fundamentally on the physics of $Q_T$), and the determination of bin migration needs to take into account the fact that the method of generating $\phi^*/Q_T$ in SignalMC simulation is itself in question (which is why this measurement from data is worth executing). Thus, the behavior of other simulation tunes need to be considered to measure this uncertainty. The SignalMC simulation uses the Z2 tune, and the bin migration effects from the D6T, Perugia, ProPT and ProQ2 tunes will also be examined. The maximum value of the fractional uncertainty of a bin from any of the tunes is taken as the final systematic for that bin. The final uncertainty due to bin migration is reasonably low in EB-EB and EE-EB, while it is a dominant factor (as expected, due to lower resolution) in EE-HF (fig. [6.11]). While
Figure 6.10: The statistical fractional uncertainties on the uncertainty calculations made using the SignalMC simulation set for the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions.
this is a systematic effect (predominantly from detector resolution and binning choice) higher statistics would allow the use of an unfolding matrix to remove the bin migration effect while introducing less uncertainty than it corrected (which is not the case with current statistics).

6.4.2 Background Estimation

The uncertainties of background estimation come from statistical uncertainties of the number of background seen, and in QCD from statistical effects of in the fitting of the equations that determine the amount of QCD (see Section 6.3). The uncertainties from background are already represented in figure 6.9

6.4.3 Efficiency Uncertainties

Within the calculations of the electron reconstruction efficiencies, there are uncertainties from both the systematics of efficiency calculated and statistics used in the calculation.

Systematic Efficiency Uncertainties

The systematic uncertainty is calculated as part of the efficiency calculation (see section 6.2), and the values of plus or minus that uncertainty are stored. The effect of the efficiency systematic error on the final measurement is calculated by replacing the efficiency with the plus (or minus) uncertainty version one efficiency stage at a time. The fractional error for each is found and the average between the plus and minus variation is taken as the total fractional uncertainty in that stage due to efficiency systematics (fig. 6.12).

Statistical Efficiency Uncertainties

Because of the finite number of data events used in calculating the efficiencies, statistical uncertainties will arise in the calculations. To simulate the effect data statistics has on the efficiency, the binomial probability function \( P(\epsilon) \) is used to vary the efficiency and then final results recalculated.

\[
P(\epsilon) = \frac{(D + 1)!}{N!(D - N)!} \epsilon^N (1 - \epsilon)^{D-N}
\]
Figure 6.11: The systematic uncertainties from bin migration for the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions.
Figure 6.12: The systematic uncertainties from efficiencies per stage for the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions.
where \( \epsilon \) is the binned efficiency being varied, \( D \) is the number of events in the denominator used to calculate \( \epsilon \), and \( N \) is the number of events in the numerator used to calculate \( \epsilon \). By performing this many times (100 in this analysis) a collection of results subject to statistical variation of efficiency is formed. The average fractional uncertainty per bin of these results is taken to be the statistical impact on the final results (fig. 6.13).

### 6.4.4 Energy and Position Uncertainties

Another source of systematic error can come from the uncertainty of the absolute calibration and scale of energy, ECAL position, and HF alignment. Energy and position scale refer to the extent that the reported energy or position varies from its true value. Alignment refers to the possibility that HF is rotated or shifted from its expected orientation (cross checks with the silicon tracker remove this effect from ECAL).

As insensitivity to energy resolution is a motivation to using \( \phi^* \) in this differential measurement, it is expected that the effects of energy scale should be minimal. For energy scale, the energies of EB, EE, and HF were scaled by 0.5\%, 1.5\%, and 3.0\% respectively. These are highly conservative values for the actual calibration status of these detectors. As expected, the reconstruction of \( \phi^* \) was resistant to energy scale errors, (fig. 6.14).

The positional scale of ECAL is an uncertainty on which \( \phi^* \) should be sensitive. The variance in actual position possible in ECAL is restricted by tracker, as tracks are matched to events in ECAL. The \( \Delta \eta \) and \( \Delta \phi \) variables in the WP80 requirements (see Table 5.1) are a measurement of differences between tracker and ECAL positional data. To estimate a possible limit on positional scale, the values of the \( \Delta \eta \) requirement for each detector (0.005 for EB, 0.006 for EE) were used to perform a flat smearing of the position reconstruction in both \( \eta \) and \( \phi \) (the value of the \( \Delta \phi \) requirement is not used for this scale estimation as it is much larger than a possible ECAL position uncertainty, as electrons are expected to curve in \( \phi \) between the tracker and ECAL). While this is potentially an overestimation of the possible positional uncertainty in ECAL, the resulting fractional uncertainty on \( \phi^* \) is still low (fig. 6.15).

The second positional uncertainty is a concern of alignment of HF, which cannot benefit from constraints from tracker matching. The possible HF misalignment parameters are a possible x-y plane shift of 10 mm and a possible rotation of 15 mm (at the
Figure 6.13: The statistical uncertainties from efficiency per stage for the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions.
face of the detector). This effect only affects EE-HF Z bosons, but with less restrictions on possible errors, the calculated uncertainty is notably higher (fig. 6.15d).
Figure 6.14: Uncertainties on the final measurement for the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions due to energy scale uncertainties in ECAL and (d) for EE-HF Z bosons due to energy scale uncertainties in HF.
Figure 6.15: Uncertainties on the final measurement for the (a) EB-EB, (b) EE-EB, and (c) EE-HF Z boson definitions due to positional uncertainties in ECAL and (d) for EE-HF Z bosons due to alignment uncertainties in HF.
Figure 6.16: Comparisons of the contributing sources of uncertainty for the EB-EB Z boson definition. The Total Systematic (black line) is the sum in quadrature of the other systematic uncertainties (the colored, solid lines).
Figure 6.17: Comparisons of the contributing sources of uncertainty for the EE-EB Z boson definition. The Total Systematic (black line) is the sum in quadrature of the other systematic uncertainties (the colored, solid lines).
Figure 6.18: Comparisons of the contributing sources of uncertainty for the EE-HF Z boson definition. The Total Systematic (black line) is the sum in quadrature of the other systematic uncertainties (the colored, solid lines).
Chapter 7

Results and Conclusion

7.1 Results

The final measurements of the differential cross-section with respect to $\phi^*$ for 216 pb$^{-1}$ were made by subtracting background estimations from the raw data and correcting by $\epsilon_{\text{Acc}}$. The results for each Z boson acceptance can be seen in figure 7.1 and tables 7.1-7.3.

7.1.1 Uncertainties

A comparison of the systematic uncertainties can be seen figures 6.16-6.18 and as tables in Appendix A, tables A.4-A.6. The greatest sources of uncertainties are statistically-limited systematics: bin migration effects, background estimation, and efficiency calculation. The bin migration being treated as an uncertainty is due to the statistics limiting the effectiveness of an unfolding matrix, and in many bins (especially at low $\phi^*$) it is the dominant uncertainty. The systematics contribution due to background estimation (the common dominant uncertainty in the middle $\phi^*$ ranges) is almost solely from the uncertainty on QCD calculation, which is from uncertainty in fitting the background shape and largely dependent on the bin-by-bin statistics. Likewise, the efficiency calculation contribution to uncertainty is mostly from the statistical uncertainties on calculation efficiencies, specifically the Electron Identification correction factor in ECAL and the Trigger Efficiency in HF which had the lowest statistics.
<table>
<thead>
<tr>
<th>$\phi^*$ Range</th>
<th>$\frac{1}{\sigma} \frac{d\sigma(Z\rightarrow e^+e^-)}{d\phi^*_k}$</th>
<th>Statistical</th>
<th>Systematic</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000-0.008</td>
<td>9.586</td>
<td>0.017</td>
<td>0.052</td>
</tr>
<tr>
<td>0.008-0.016</td>
<td>9.379</td>
<td>0.017</td>
<td>0.070</td>
</tr>
<tr>
<td>0.016-0.024</td>
<td>8.335</td>
<td>0.014</td>
<td>0.060</td>
</tr>
<tr>
<td>0.024-0.034</td>
<td>7.742</td>
<td>0.014</td>
<td>0.044</td>
</tr>
<tr>
<td>0.034-0.045</td>
<td>6.829</td>
<td>0.012</td>
<td>0.039</td>
</tr>
<tr>
<td>0.045-0.057</td>
<td>5.562</td>
<td>0.0095</td>
<td>0.029</td>
</tr>
<tr>
<td>0.057-0.072</td>
<td>4.565</td>
<td>0.0079</td>
<td>0.031</td>
</tr>
<tr>
<td>0.072-0.091</td>
<td>3.857</td>
<td>0.0069</td>
<td>0.025</td>
</tr>
<tr>
<td>0.091-0.114</td>
<td>2.929</td>
<td>0.0051</td>
<td>0.020</td>
</tr>
<tr>
<td>0.114-0.145</td>
<td>2.199</td>
<td>0.0038</td>
<td>0.012</td>
</tr>
<tr>
<td>0.145-0.189</td>
<td>1.580</td>
<td>0.0028</td>
<td>0.010</td>
</tr>
<tr>
<td>0.189-0.258</td>
<td>0.9027</td>
<td>0.0015</td>
<td>0.0050</td>
</tr>
<tr>
<td>0.258-0.391</td>
<td>0.4874</td>
<td>0.0008</td>
<td>0.0042</td>
</tr>
<tr>
<td>0.391-0.524</td>
<td>0.2331</td>
<td>0.0003</td>
<td>0.0039</td>
</tr>
<tr>
<td>0.524-1.000</td>
<td>0.0754</td>
<td>0.0001</td>
<td>0.0020</td>
</tr>
</tbody>
</table>

Table 7.1: The Final Differential measurement for the EB-EB definition, with statistical and systematic uncertainties given in units of $\phi^*$.

<table>
<thead>
<tr>
<th>$\phi^*$ Range</th>
<th>$\frac{1}{\sigma} \frac{d\sigma(Z\rightarrow e^+e^-)}{d\phi^*_k}$</th>
<th>Statistical</th>
<th>Systematic</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000-0.008</td>
<td>9.913</td>
<td>0.024</td>
<td>0.089</td>
</tr>
<tr>
<td>0.008-0.016</td>
<td>9.858</td>
<td>0.024</td>
<td>0.068</td>
</tr>
<tr>
<td>0.016-0.024</td>
<td>9.864</td>
<td>0.024</td>
<td>0.064</td>
</tr>
<tr>
<td>0.024-0.034</td>
<td>7.571</td>
<td>0.018</td>
<td>0.052</td>
</tr>
<tr>
<td>0.034-0.045</td>
<td>6.888</td>
<td>0.016</td>
<td>0.047</td>
</tr>
<tr>
<td>0.045-0.057</td>
<td>5.940</td>
<td>0.014</td>
<td>0.036</td>
</tr>
<tr>
<td>0.057-0.072</td>
<td>4.611</td>
<td>0.011</td>
<td>0.044</td>
</tr>
<tr>
<td>0.072-0.091</td>
<td>3.616</td>
<td>0.0082</td>
<td>0.041</td>
</tr>
<tr>
<td>0.091-0.114</td>
<td>2.711</td>
<td>0.0059</td>
<td>0.035</td>
</tr>
<tr>
<td>0.114-0.145</td>
<td>2.135</td>
<td>0.0047</td>
<td>0.025</td>
</tr>
<tr>
<td>0.145-0.189</td>
<td>1.427</td>
<td>0.0031</td>
<td>0.010</td>
</tr>
<tr>
<td>0.189-0.258</td>
<td>0.927</td>
<td>0.0020</td>
<td>0.011</td>
</tr>
<tr>
<td>0.258-0.391</td>
<td>0.4518</td>
<td>0.0009</td>
<td>0.0055</td>
</tr>
<tr>
<td>0.391-0.524</td>
<td>0.2082</td>
<td>0.0003</td>
<td>0.0029</td>
</tr>
<tr>
<td>0.524-1.000</td>
<td>0.0746</td>
<td>0.0001</td>
<td>0.0046</td>
</tr>
</tbody>
</table>

Table 7.2: The Final Differential measurement for the EE-EB definition, with statistical and systematic uncertainties given in units of $\phi^*$. 
Figure 7.1: The final results of the differential measurement in the (a) EB-EB, (b) EE-EB, and (c) EE-HF channels. The red line is the SignalMC simulation for comparison.
\[
\phi^* \text{ Range} \quad \frac{d\sigma(Z \rightarrow e^+e^-)}{d\phi^*_h} \quad \text{Statistical} \quad \text{Systematic}
\]

<table>
<thead>
<tr>
<th>$\phi^*$ Range</th>
<th>(\frac{d\sigma(Z \rightarrow e^+e^-)}{d\phi^*_h})</th>
<th>Statistical</th>
<th>Systematic</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000-0.008</td>
<td>11.15</td>
<td>0.053</td>
<td>0.78</td>
</tr>
<tr>
<td>0.008-0.016</td>
<td>11.20</td>
<td>0.053</td>
<td>0.62</td>
</tr>
<tr>
<td>0.016-0.024</td>
<td>9.84</td>
<td>0.043</td>
<td>0.30</td>
</tr>
<tr>
<td>0.024-0.034</td>
<td>9.17</td>
<td>0.043</td>
<td>0.37</td>
</tr>
<tr>
<td>0.034-0.045</td>
<td>8.02</td>
<td>0.037</td>
<td>0.23</td>
</tr>
<tr>
<td>0.045-0.057</td>
<td>6.29</td>
<td>0.027</td>
<td>0.29</td>
</tr>
<tr>
<td>0.057-0.072</td>
<td>5.16</td>
<td>0.023</td>
<td>0.30</td>
</tr>
<tr>
<td>0.072-0.091</td>
<td>4.21</td>
<td>0.019</td>
<td>0.16</td>
</tr>
<tr>
<td>0.091-0.114</td>
<td>2.59</td>
<td>0.010</td>
<td>0.12</td>
</tr>
<tr>
<td>0.114-0.145</td>
<td>1.767</td>
<td>0.0065</td>
<td>0.067</td>
</tr>
<tr>
<td>0.145-0.189</td>
<td>1.454</td>
<td>0.0058</td>
<td>0.043</td>
</tr>
<tr>
<td>0.189-0.258</td>
<td>0.764</td>
<td>0.0029</td>
<td>0.036</td>
</tr>
<tr>
<td>0.258-0.391</td>
<td>0.392</td>
<td>0.0015</td>
<td>0.017</td>
</tr>
<tr>
<td>0.391-0.524</td>
<td>0.1544</td>
<td>0.0004</td>
<td>0.0072</td>
</tr>
<tr>
<td>0.524-1.000</td>
<td>0.0479</td>
<td>0.0001</td>
<td>0.0048</td>
</tr>
</tbody>
</table>

Table 7.3: The Final Differential measurement for the EE-HF definition, with statistical and systematic uncertainties given in units of $\phi^*$. 
7.1.2 Comparison with Models and Theories

In addition to the SignalMC simulated data set (which is specifically a POWHEG+PYTHIA simulation using the Z2 tune), five additional (Pythia only) data sets were compared. These sets are the Z2 tune (Pythia only), the D6T tune, the Perugia 2010 tune, the ProPt tune, and ProQ2 tune (figs. 7.2 [7.4]). In general, the PYTHIA-only tunes show visibly better agreement to the experimental data. The individual $\chi^2$/ndof for each (Table 7.4) support this. While there is no conclusive best fit, the EB-EB acceptance showing best agreement with ProQ2, the EE-EB acceptance shows best agreement with D6T, and EE-HF shows best agreement to the Z2 only tune, while SignalMC and ProPt (to a lesser degree) are consistently in larger disagreement.

<table>
<thead>
<tr>
<th>Channel</th>
<th>SignalMC</th>
<th>Z2</th>
<th>D6T</th>
<th>Perugia</th>
<th>ProPt</th>
<th>ProQ2</th>
</tr>
</thead>
<tbody>
<tr>
<td>EB-EB</td>
<td>2343.6/15</td>
<td>364.0/15</td>
<td>367.5/15</td>
<td>944.8/15</td>
<td>1330.5/15</td>
<td>342.1/15</td>
</tr>
<tr>
<td>EE-EB</td>
<td>1360.3/15</td>
<td>401.2/15</td>
<td>340.7/15</td>
<td>562.6/15</td>
<td>786.7/15</td>
<td>360.9/15</td>
</tr>
<tr>
<td>EE-HF</td>
<td>83.7/15</td>
<td>41.4/15</td>
<td>46.6/15</td>
<td>69.3/15</td>
<td>70.5/15</td>
<td>54.9/15</td>
</tr>
</tbody>
</table>

Table 7.4: The $\chi^2$/ndof for the various tunes of PYTHIA (and the solitary POWHEG+PYTHIA SignalMC tune).

Comparing to QCD Theories

Due to the highly specialized nature of what the Non-perturbative Sudakov Theory and Perturbative NLL Theory focus on, the precise mechanics they depend on are not present in common simulation methods, such as PYTHIA. While the method PYTHIA uses to simulate $Q_T$ distributions is similar to the Sudakov method, the values used in a particular tune are not directly comparable to the coefficients in the BNLY term, and NLL is even more divergent in comparability. To truly study the two theories, very specialized simulations built around the tested variables are need. While such simulations are available, the time it would take to adapt such simulations to this specific measurement would not likely be rewarding at the current level of statistics. It is expected that a differential cross-section measurement using data from CMS taken in 2012 with much greater statistics will be available soon, and will build upon this measurement to be a more appropriate tool for comparison directly to QCD theories.
Figure 7.2: A comparison of the ratios of multiple PYTHIA tunes (and data) to the POWHEG+PYTHIA Z2 simulation in EB-EB the Z boson definition.
Figure 7.3: A comparison of the ratios of multiple PYTHIA tunes (and data) to the POWHEG+PYTHIA Z2 simulation in EE-EB the Z boson definition.
Figure 7.4: A comparison of the ratios of multiple PYTHIA tunes (and data) to the POWHEG+PYTHIA Z2 simulation in EE-HF the Z boson definition.
7.1.3 Comparing to ATLAS

The most comparable measurement of $\phi^*$ from other sources is from a measurement performed by the ATLAS detector [23]. A comparison between the two measurements can be made by creating a new Z boson definition (referred to here as the ‘ATLAS Comparison’ definition) that matches the kinematic acceptances of the ATLAS measurement. This acceptance is an electron requirement of $p_T < 20$, $|\eta_e| < 2.4$, and a dilepton mass between 66 and 116 GeV. The binning is not identical for the ATLAS measurement, in general the ATLAS measurement has bins half as large, requiring a minor recalculation of the ATLAS measurement to compare. The full set of plots mirroring the treatment of the initial three Z boson definitions can be found in Appendix B. Agreement between the ATLAS comparison definition and the ATLAS measurement is strong in all but the last two bins of the measurement (fig. 7.5), with a $\chi^2$/ndof of 74.5/13 in that range of 0.0 to 0.381 in $\phi^*$, which becomes 660.4/15 when the last two bins are added.

7.2 Conclusion

This thesis has presented the relevant components of the CMS detector and LHC experiment, methods of detection and identification of electrons, and physics based motivations for a differential cross-section measurement with respect to the novel variable $\phi^*$. A strategy for making this measurement was described and carried out. The strategy involved calculations of reconstruction efficiencies, bin migration effects, background population using both experimental data and simulation.

The final measurement for the differential cross-section with respect to $\phi^*$ at 216pb$^{-1}$ was reported and initial comparisons towards predictions made. The predominant uncertainties were statistically dependent systematics, most notably the bin migration effects and QCD background estimation. This measurement also showed good agreement with other similar measurements.

The $\phi^*$ variable shows sensitivity to underlying physics at low values, and with greater statistics has excellent potential for making precision measurements of $Q_T$ physics and QCD theories.
Figure 7.5: A comparison of the ratios of the ATLAS Comparison definition and ATLAS measurement to the POWHEG+PYTHIA Z2 simulation (as well as the other tunes). Note the strong agreement in all but the last two bins.
References


Appendix A

Extra Tables

A.1 Bin Migration Matrices

The bin migration matrices for the various Z boson acceptances. Rows represent constant true simulated $\phi^*$, columns are at constant reconstructed $\phi^*$.

\[
\begin{bmatrix}
0.950 & 0.037 & 0.005 & 0.002 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.037 & 0.918 & 0.034 & 0.002 & 0.001 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.005 & 0.032 & 0.919 & 0.029 & 0.003 & 0.001 & 0.001 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.004 & 0.006 & 0.033 & 0.928 & 0.027 & 0.003 & 0.001 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.003 & 0.003 & 0.005 & 0.028 & 0.934 & 0.028 & 0.003 & 0.002 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.002 & 0.001 & 0.002 & 0.004 & 0.026 & 0.935 & 0.022 & 0.003 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.001 & 0.002 & 0.003 & 0.022 & 0.947 & 0.021 & 0.002 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.001 & 0.002 & 0.003 & 0.022 & 0.953 & 0.017 & 0.002 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.018 & 0.962 & 0.016 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.002 & 0.016 & 0.966 & 0.012 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.015 & 0.976 & 0.009 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.010 & 0.984 & 0.007 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.006 & 0.990 & 0.005 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.003 & 0.992 & 0.003 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.003 & 0.997
\end{bmatrix}
\]

Table A.1: Bin Migration Matrix for the EB-EB acceptance of Z bosons determined from simulation.
Table A.2: Bin Migration Matrix for the EE-EB acceptance of Z bosons determined from simulation.

\[
\begin{bmatrix}
0.935 & 0.050 & 0.006 & 0.003 & 0.002 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.052 & 0.090 & 0.050 & 0.005 & 0.002 & 0.001 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.007 & 0.050 & 0.888 & 0.042 & 0.003 & 0.001 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.003 & 0.005 & 0.049 & 0.906 & 0.030 & 0.004 & 0.002 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.001 & 0.003 & 0.004 & 0.038 & 0.911 & 0.038 & 0.003 & 0.001 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.001 & 0.001 & 0.002 & 0.004 & 0.039 & 0.915 & 0.031 & 0.002 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.002 & 0.004 & 0.034 & 0.929 & 0.028 & 0.002 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.001 & 0.0 & 0.003 & 0.032 & 0.943 & 0.026 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.001 & 0.002 & 0.023 & 0.948 & 0.017 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.002 & 0.023 & 0.965 & 0.016 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.016 & 0.973 & 0.011 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.011 & 0.980 & 0.007 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.008 & 0.990 & 0.007 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.003 & 0.987 & 0.003 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.005 & 0.997 & 0.000 & 0.0 & 0.0 & 0.0 & 0.0 \\
\end{bmatrix}
\]

Table A.3: Bin Migration Matrix for the EE-HF acceptance of Z bosons determined from simulation.

\[
\begin{bmatrix}
0.476 & 0.318 & 0.133 & 0.044 & 0.010 & 0.001 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.312 & 0.311 & 0.229 & 0.097 & 0.024 & 0.004 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.144 & 0.211 & 0.279 & 0.196 & 0.074 & 0.016 & 0.002 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.054 & 0.120 & 0.246 & 0.336 & 0.214 & 0.069 & 0.010 & 0.001 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.010 & 0.033 & 0.091 & 0.232 & 0.362 & 0.221 & 0.055 & 0.005 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.002 & 0.006 & 0.019 & 0.080 & 0.230 & 0.388 & 0.217 & 0.031 & 0.001 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.002 & 0.002 & 0.014 & 0.070 & 0.256 & 0.457 & 0.215 & 0.017 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.002 & 0.007 & 0.043 & 0.237 & 0.527 & 0.190 & 0.009 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.001 & 0.003 & 0.022 & 0.210 & 0.605 & 0.167 & 0.002 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.001 & 0.011 & 0.183 & 0.687 & 0.133 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.002 & 0.136 & 0.775 & 0.096 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.890 & 0.850 & 0.063 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.053 & 0.908 & 0.063 & 0.000 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.029 & 0.910 & 0.039 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.027 & 0.961 & 0.0 \\
\end{bmatrix}
\]
## A.2 Final Systematic Uncertainties Tables

<table>
<thead>
<tr>
<th>$\phi^*$ Range</th>
<th>Bin Migration</th>
<th>Background Estimation</th>
<th>Efficiency Calculation</th>
<th>Energy Scale</th>
<th>Position and Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000-0.008</td>
<td>0.042</td>
<td>0.016</td>
<td>0.025</td>
<td>0.0020</td>
<td>0.0057</td>
</tr>
<tr>
<td>0.008-0.016</td>
<td>0.040</td>
<td>0.044</td>
<td>0.023</td>
<td>0.0023</td>
<td>0.029</td>
</tr>
<tr>
<td>0.016-0.024</td>
<td>0.024</td>
<td>0.051</td>
<td>0.021</td>
<td>0.0001</td>
<td>0.0034</td>
</tr>
<tr>
<td>0.024-0.034</td>
<td>0.032</td>
<td>0.025</td>
<td>0.017</td>
<td>0.0012</td>
<td>0.0029</td>
</tr>
<tr>
<td>0.034-0.045</td>
<td>0.032</td>
<td>0.020</td>
<td>0.011</td>
<td>0.0012</td>
<td>0.0025</td>
</tr>
<tr>
<td>0.045-0.057</td>
<td>0.014</td>
<td>0.023</td>
<td>0.0082</td>
<td>0.0007</td>
<td>0.0037</td>
</tr>
<tr>
<td>0.057-0.072</td>
<td>0.023</td>
<td>0.019</td>
<td>0.0063</td>
<td>0.0008</td>
<td>0.0051</td>
</tr>
<tr>
<td>0.072-0.091</td>
<td>0.0067</td>
<td>0.024</td>
<td>0.0036</td>
<td>0.0006</td>
<td>0.0007</td>
</tr>
<tr>
<td>0.091-0.114</td>
<td>0.013</td>
<td>0.016</td>
<td>0.0018</td>
<td>0.0009</td>
<td>0.0007</td>
</tr>
<tr>
<td>0.114-0.145</td>
<td>0.0043</td>
<td>0.011</td>
<td>0.0012</td>
<td>0.0002</td>
<td>0.0013</td>
</tr>
<tr>
<td>0.145-0.189</td>
<td>0.0034</td>
<td>0.0094</td>
<td>0.0022</td>
<td>0.0002</td>
<td>0.0008</td>
</tr>
<tr>
<td>0.189-0.258</td>
<td>0.0011</td>
<td>0.0038</td>
<td>0.0028</td>
<td>0.0003</td>
<td>0.0009</td>
</tr>
<tr>
<td>0.258-0.391</td>
<td>0.0003</td>
<td>0.0025</td>
<td>0.0033</td>
<td>0.0002</td>
<td>0.0003</td>
</tr>
<tr>
<td>0.391-0.524</td>
<td>0.0006</td>
<td>0.0019</td>
<td>0.0034</td>
<td>0.0001</td>
<td>0.0003</td>
</tr>
<tr>
<td>0.524-1.000</td>
<td>0.0001</td>
<td>0.0005</td>
<td>0.0019</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table A.4: The Systematic uncertainties for the EB-EB $Z$ boson acceptance, in units of $\phi^*$. 
Table A.5: The Systematic uncertainties for the EE-EB Z boson acceptance, in units of $\phi^*$. 

<table>
<thead>
<tr>
<th>$\phi^*$ Range</th>
<th>Bin Migration</th>
<th>Background Estimation</th>
<th>Efficiency Calculation</th>
<th>Energy Scale</th>
<th>Position and Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000-0.008</td>
<td>0.077</td>
<td>0.015</td>
<td>0.025</td>
<td>0.0061</td>
<td>0.034</td>
</tr>
<tr>
<td>0.008-0.016</td>
<td>0.048</td>
<td>0.037</td>
<td>0.022</td>
<td>0.019</td>
<td>0.014</td>
</tr>
<tr>
<td>0.016-0.024</td>
<td>0.050</td>
<td>0.021</td>
<td>0.027</td>
<td>0.0041</td>
<td>0.021</td>
</tr>
<tr>
<td>0.024-0.034</td>
<td>0.026</td>
<td>0.042</td>
<td>0.015</td>
<td>0.0039</td>
<td>0.0043</td>
</tr>
<tr>
<td>0.034-0.045</td>
<td>0.028</td>
<td>0.031</td>
<td>0.012</td>
<td>0.0017</td>
<td>0.019</td>
</tr>
<tr>
<td>0.045-0.057</td>
<td>0.014</td>
<td>0.022</td>
<td>0.0075</td>
<td>0.0040</td>
<td>0.024</td>
</tr>
<tr>
<td>0.057-0.072</td>
<td>0.010</td>
<td>0.031</td>
<td>0.0055</td>
<td>0.0008</td>
<td>0.029</td>
</tr>
<tr>
<td>0.072-0.091</td>
<td>0.020</td>
<td>0.030</td>
<td>0.0032</td>
<td>0.0028</td>
<td>0.020</td>
</tr>
<tr>
<td>0.091-0.114</td>
<td>0.021</td>
<td>0.027</td>
<td>0.0017</td>
<td>0.0014</td>
<td>0.0044</td>
</tr>
<tr>
<td>0.114-0.145</td>
<td>0.015</td>
<td>0.020</td>
<td>0.0013</td>
<td>0.0016</td>
<td>0.0037</td>
</tr>
<tr>
<td>0.145-0.189</td>
<td>0.0008</td>
<td>0.0093</td>
<td>0.0017</td>
<td>0.0000</td>
<td>0.0021</td>
</tr>
<tr>
<td>0.189-0.258</td>
<td>0.0011</td>
<td>0.010</td>
<td>0.0025</td>
<td>0.0004</td>
<td>0.0008</td>
</tr>
<tr>
<td>0.258-0.391</td>
<td>0.0011</td>
<td>0.0042</td>
<td>0.0033</td>
<td>0.0005</td>
<td>0.0005</td>
</tr>
<tr>
<td>0.391-0.524</td>
<td>0.0003</td>
<td>0.0002</td>
<td>0.0028</td>
<td>0.0002</td>
<td>0.0003</td>
</tr>
<tr>
<td>0.524-1.000</td>
<td>0.0001</td>
<td>0.0010</td>
<td>0.0045</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table A.6: The Systematic uncertainties for the EE-HF Z boson acceptance, in units of $\phi^*$. 

<table>
<thead>
<tr>
<th>$\phi^*$ Range</th>
<th>Bin Migration</th>
<th>Efficiency Calculation</th>
<th>Background Estimation</th>
<th>Energy Scale</th>
<th>Position and Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000-0.008</td>
<td>0.70</td>
<td>0.21</td>
<td>0.24</td>
<td>0.033</td>
<td>0.15</td>
</tr>
<tr>
<td>0.008-0.016</td>
<td>0.55</td>
<td>0.20</td>
<td>0.21</td>
<td>0.015</td>
<td>0.045</td>
</tr>
<tr>
<td>0.016-0.024</td>
<td>0.13</td>
<td>0.18</td>
<td>0.14</td>
<td>0.009</td>
<td>0.14</td>
</tr>
<tr>
<td>0.024-0.034</td>
<td>0.18</td>
<td>0.14</td>
<td>0.15</td>
<td>0.024</td>
<td>0.25</td>
</tr>
<tr>
<td>0.034-0.045</td>
<td>0.11</td>
<td>0.11</td>
<td>0.06</td>
<td>0.010</td>
<td>0.15</td>
</tr>
<tr>
<td>0.045-0.057</td>
<td>0.23</td>
<td>0.14</td>
<td>0.10</td>
<td>0.0052</td>
<td>0.047</td>
</tr>
<tr>
<td>0.057-0.072</td>
<td>0.24</td>
<td>0.097</td>
<td>0.14</td>
<td>0.0072</td>
<td>0.013</td>
</tr>
<tr>
<td>0.072-0.091</td>
<td>0.065</td>
<td>0.095</td>
<td>0.064</td>
<td>0.0047</td>
<td>0.085</td>
</tr>
<tr>
<td>0.091-0.114</td>
<td>0.075</td>
<td>0.064</td>
<td>0.071</td>
<td>0.0046</td>
<td>0.014</td>
</tr>
<tr>
<td>0.114-0.145</td>
<td>0.047</td>
<td>0.013</td>
<td>0.043</td>
<td>0.0016</td>
<td>0.017</td>
</tr>
<tr>
<td>0.145-0.189</td>
<td>0.033</td>
<td>0.017</td>
<td>0.021</td>
<td>0.0016</td>
<td>0.0066</td>
</tr>
<tr>
<td>0.189-0.258</td>
<td>0.0030</td>
<td>0.020</td>
<td>0.029</td>
<td>0.0023</td>
<td>0.0078</td>
</tr>
<tr>
<td>0.258-0.391</td>
<td>0.0023</td>
<td>0.0077</td>
<td>0.015</td>
<td>0.0008</td>
<td>0.0017</td>
</tr>
<tr>
<td>0.391-0.524</td>
<td>0.0046</td>
<td>0.0027</td>
<td>0.0047</td>
<td>0.0005</td>
<td>0.0005</td>
</tr>
<tr>
<td>0.524-1.000</td>
<td>0.0011</td>
<td>0.0031</td>
<td>0.0036</td>
<td>0.0000</td>
<td>0.0003</td>
</tr>
</tbody>
</table>
Appendix B

Comparison to Atlas Results

Plots and Tables relating to the ATLAS Comparison definition.

<table>
<thead>
<tr>
<th>$\phi^*$ Range</th>
<th>$\frac{1}{\sigma} \frac{d\sigma(Z\rightarrow e^-e^-)}{d\phi^*}$</th>
<th>Stat.</th>
<th>Syst.</th>
<th>ATLAS Meas.</th>
<th>ATLAS Uncert.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000-0.008</td>
<td>9.982</td>
<td>0.014</td>
<td>0.093</td>
<td>9.750</td>
<td>0.029</td>
</tr>
<tr>
<td>0.008-0.016</td>
<td>9.678</td>
<td>0.014</td>
<td>0.062</td>
<td>9.310</td>
<td>0.029</td>
</tr>
<tr>
<td>0.016-0.024</td>
<td>8.755</td>
<td>0.012</td>
<td>0.058</td>
<td>8.655</td>
<td>0.028</td>
</tr>
<tr>
<td>0.024-0.034</td>
<td>7.619</td>
<td>0.011</td>
<td>0.048</td>
<td>7.785</td>
<td>0.023</td>
</tr>
<tr>
<td>0.034-0.045</td>
<td>6.834</td>
<td>0.0095</td>
<td>0.037</td>
<td>6.770</td>
<td>0.021</td>
</tr>
<tr>
<td>0.045-0.057</td>
<td>5.769</td>
<td>0.0078</td>
<td>0.033</td>
<td>5.745</td>
<td>0.018</td>
</tr>
<tr>
<td>0.057-0.072</td>
<td>4.580</td>
<td>0.0062</td>
<td>0.021</td>
<td>4.764</td>
<td>0.015</td>
</tr>
<tr>
<td>0.072-0.091</td>
<td>3.792</td>
<td>0.0052</td>
<td>0.021</td>
<td>3.783</td>
<td>0.012</td>
</tr>
<tr>
<td>0.091-0.114</td>
<td>2.838</td>
<td>0.0038</td>
<td>0.018</td>
<td>2.9314</td>
<td>0.0094</td>
</tr>
<tr>
<td>0.114-0.145</td>
<td>2.185</td>
<td>0.0029</td>
<td>0.012</td>
<td>2.1545</td>
<td>0.0069</td>
</tr>
<tr>
<td>0.145-0.189</td>
<td>1.523</td>
<td>0.0020</td>
<td>0.012</td>
<td>1.5014</td>
<td>0.0048</td>
</tr>
<tr>
<td>0.189-0.258</td>
<td>0.9114</td>
<td>0.0012</td>
<td>0.0072</td>
<td>0.9370</td>
<td>0.0031</td>
</tr>
<tr>
<td>0.258-0.391</td>
<td>0.4762</td>
<td>0.0006</td>
<td>0.0048</td>
<td>0.4768</td>
<td>0.0016</td>
</tr>
<tr>
<td>0.391-0.524</td>
<td>0.2294</td>
<td>0.0002</td>
<td>0.0035</td>
<td>0.1656</td>
<td>0.0007</td>
</tr>
<tr>
<td>0.524-1.000</td>
<td>0.0746</td>
<td>0.0001</td>
<td>0.0013</td>
<td>0.0431</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

Table B.1: The Final Differential measurement for the ATLAS Comparison definition, with statistical and systematic uncertainties given in units of $\phi^*$. Also the ATLAS measurement and approximate uncertainty.
Table B.2: The χ²/ndof between the ATLAS comparison definition and the ATLAS measurement (and the various tunes of PYTHIA) for the range of best agreement (the first thirteen bins, φ* of 0.000-0.391) and the whole range (φ* of 0.000-1.000).

<table>
<thead>
<tr>
<th>ATLAS</th>
<th>Summer11</th>
<th>Z2</th>
<th>D6T</th>
<th>Perugia</th>
<th>ProPT</th>
<th>ProQ2</th>
</tr>
</thead>
<tbody>
<tr>
<td>74.5/13</td>
<td>2456.3/13</td>
<td>285.5/13</td>
<td>304.8/13</td>
<td>407.0/13</td>
<td>1274.2/13</td>
<td>200.3/13</td>
</tr>
<tr>
<td>660.4/15</td>
<td>2457.9/15</td>
<td>291.0/15</td>
<td>326.4/15</td>
<td>431.5/15</td>
<td>1334.5/15</td>
<td>200.6/15</td>
</tr>
</tbody>
</table>
Figure B.1: (a) The final measurement of the ATLAS Comparison definition with the ATLAS measurement and Summer11 shown for comparison. Further plots of the ATLAS Comparison definition for (b) the final cumulative efficiencies per stage, (c) the estimated background as a fraction of total events per bin, and (d) the systematic uncertainties from bin migration.
Figure B.2: Plots of the ATLAS Comparison definition for (a) the systematic uncertainties from efficiencies per stage, (b) the statistical uncertainties from efficiency per stage, (c) the uncertainties on the final measurement due to energy scale in ECAL, and (d) the uncertainties on the final measurement for due to positional uncertainties in ECAL.