Measurement of the phistar distribution of Z bosons decaying to electron pairs with the CMS experiment at a center-of-mass energy of 8 TeV

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

Alexander Erling Gude

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTORATE OF PHILOSOPHY

Jeremiah Mans

May, 2015

Acknowledgements

No scientific work is created in isolation and this thesis is no exception. Many people were instrumental in ensuring its completion, either directly through their work, or indirectly through their influences on my life.

To begin, I would like to offer thanks to my parents, Hans and Laurel Gude, who always encouraged my love of science. Whether it was when I wanted to study dinosaurs, or planets and stars, or whales, or finally supernovae and sub-atomic particles, they were always there cheering me on.

Thanks to my science and math teachers; they saw something in me when I was young and helped to nurture the skills and interests I would need to be a scientist, and for that I am forever in their debt.

A big thanks to the members of the Supernova Cosmology Project, especially Saul Perlmutter, Anthony Spadafora, Nao Suzuki, and David Rubin. They made me part of their group when I was a student at Berkeley and taught me how to do precision physics. Working with these world-class scientists, and being able to contribute, gave me the confidence I needed when I otherwise felt like I was inadequate as a scientist.

A special thanks to the graduate students at the University of Minnesota: Michael Albright, Dan Endean, Charles McEachern, Roxanne Radpour, Dominic Rocco, Allan and Mandy Straub, and Jan Zirnstein. They offered their friendship to a Californian sorely out of place in the Midwest and so filled the last six years with joy.

I would especially like to thank the members of the CMS group at the University of Minnesota: Roger Rusack, Yuichi Kubota, Bryan Dahmes, Nathaniel Joseph Pastika, and Jared Turkewitz. They offered invaluable advice about difficulties I encountered when I was just starting out in high energy physics. Thanks to Kevin Klapoetke, who worked on this measurement before me, to Alexey Finkel, who assisted me with this

measurement, and Zachary Lesko, who also assisted me and will take over now that I have graduated. A special thanks to my adviser, Jeremiah Mans, who guided me through my formative graduate school years and who made me the scientist I am today. Also thanks to Nicole Ruckstuhl, who worked side-by-side on the analysis with me and without whom I would have been completely lost. If this work is of any value, it is because of the contributions from those two.

Of course, none of this work would have been possible without the thousands of scientists, engineers, and other people working hard at CERN, the LHC, and CMS. *Merci beaucoup pour votre travail acharné*!

Thanks to Alan Wu for his extensive edits and for letting me bounce ideas off him when I was not sure how to phrase things. Without his help, this work would have been even more impenetrable than it already is.

Finally, I would like to thank my wife, Connie Lam. She offered me her love and support through the entire process, and never wavered even when the thesis began to take over all of my free time. I love you!

Dedication

To my wife, Connie, and to our daughter; I can not wait to share the wonder that is the exploration of the natural world with you!

Abstract

Measurements of the Z boson transverse momentum (Q_{T}) spectrum serves as both a precision test of non-perturbative QCD and helps to reduce the uncertainty in the measurement of the W boson mass. However, Q_{T} is limited at its lowest values by detector resolution, and so a new variable, ϕ^*, which performs better in the low Q_{T} region, is used instead. This thesis presents the first measurement of the normalized differential cross section of Z bosons decaying to electron pairs in terms of ϕ^* at $\sqrt{s} = 8\,\mathrm{TeV}$. The data used in this measurement were collected by the CMS detector at the LHC in 2012 and totaled $19.7\,\mathrm{fb}^{-1}$ of integrated luminosity. The results are compared to predictions from simulation, which are found to provide a poor description of the data.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

High energy particle physics is the study of the properties and interactions of the basic building blocks of all matter. These properties and interactions are described by the Standard Model, the best tested and most accurate scientific theory to date. Even so, there are regions of the Standard Model where calculations are difficult to perform, as is the case with low energy quantum chromodynamics (QCD) interactions which are not calculable via perturbation theory.

This thesis describes the measurement of a process which probes this region of the Standard Model using Z bosons decaying to electron pairs. These decays are a very clean probe of QCD as neither the Z nor the electrons carry color charge and so all of the affect from QCD is isolated in the initial interaction. These initial interactions sometimes give the Z boson a non-zero momentum transverse to the beamline (Q_{T}), which is what this thesis measures. The novel variable ϕ^* is used in place of Q_{T} because it is less susceptible to detector resolution effects and systematic uncertainties while providing a probe of the same physics.

The data used in this thesis were collected in 2012 with the Compact Muon Solenoid, one of four particle detectors at the Large Hadron Collider. In total, $19.7\,\mathrm{fb}^{-1}$ of integrated luminosity was recorded at a center-of-mass energy of 8 TeV, making this thesis the first measurement of ϕ^* at that energy.

The primary measurement made in this thesis is presented as a normalized differential cross section; an additional absolute cross section measurement is also provided. These result are compared to predictions from several sets of simulated events.

This thesis is organized as follows:

Chapter 2 presents the history of the Standard Model and the motivation for the measurement.

Chapter 3 gives an overview of the Large Hadron Collider (LHC) and the Compact Muon Solenoid (CMS).

Chapter 4 describes the way in which electrons are measured at CMS.

Chapter 5 introduces the simulated data samples and the actual data samples used in the measurement, as well as the scale factors used to correct the simulated data.

Chapter 6 discusses the method of selecting events and correcting them for various detector effects.

Chapter 7 presents the analysis including the uncertainties, the final results of the measurement, and a short discussion on the results.

Appendix A presents additional measurements.

Appendix B contains tables detailing the uncertainties on the final measurements.

Appendix C shows the plots from the QCD multi-jet and W + jets background fits.

Appendix D reviews the terms and acronyms used in this thesis.

Chapter 2

Physics of Z Transverse Momentum

2.1 The Standard Model

Our current understanding of how matter interacts at high energies is entirely described by the Standard Model as constructed by Weinberg, Glashow, and Salam [1, 2, 3]. The model combines three of the four fundamental forces (leaving out only gravity, which is so weak as to be negligible) and is the most accurate scientific theory ever formulated.

2.1.1 The Electromagnetic Force

The modern theory of electromagnetism began with Maxwell's theory developed in the middle of the 19th century [4]. Maxwell was the first to conclude that light was an electromagnetic wave, the full importance of which was only later understood when it was discovered that the photon was the force carrier of the electromagnetic force [5].

In the early 20th century, Lorentz and Einstein developed relativistic mechanics and showed that Maxwell's theory was Lorentz invariant [6, 7]. Dirac updated the theory in 1920 when he was able to quantize the electromagnetic field as an ensemble of harmonic oscillators [8]. Dirac would go on to discover that antiparticles were a natural consequence of his equations [9, 10]. These antiparticles were found by Anderson in 1932 as he observed cosmic rays in a cloud chamber [11].

As microwave technology improved in the 1940s, more accurate measurements of the energy level shifts in hydrogen were made, resulting in the discovery of the Lamb shift by Lamb and Rutherford [12]. This shift was not immediately explainable, but Bethe would resolve this discrepancy by showing that electromagnetic theory could account for the Lamb shift using non-relativistic calculations [13]. Bethe's work inspired multiple other physicists including Dyson, Feynman, Schwinger, and Tomonaga to work along similar lines. The result of their efforts was quantum electrodynamics (QED), a fully relativistic and self-consistent theory of electromagnetic interactions [14, 15, 16, 17].

QED is a perturbation theory with expansions performed in terms of the fine structure constant, α. As $\alpha \approx 7.297 \cdot 10^{-3}$, the higher order terms contribute smaller and smaller corrections and so only a few orders need to be computed to make very accurate predictions. For example, the anomalous magnetic moment of the electron is calculated to $\mathrm{O}(\alpha^4)$, and agrees with experiment to more than 10 significant figures.

2.1.2 The Weak Force

The need for a weak force, and hence a theory describing it, was first hinted at by beta decay experiments in the early 1900s. These experiments culminated in Chadwick's 1914 discovery that the energy spectrum of electrons ejected in beta decay was continuous instead of a delta function as would be expected for a two-body decay [18]. While some believed that this discovery indicated that momentum and energy were not conserved, Pauli proposed an alternative: that there was a neutral and invisible particle (known today as the neutrino), which carried away some of the energy [19]. Fermi began working on this idea and invented a four fermion contact interaction in which a neutron decayed into a proton, an electron, and a neutrino [20].

In 1947, Rochester and Butler discovered a particle that decayed to two pions which they called the θ; in 1949, Brown and Powell discovered a particle that decayed to three pions which they called the τ [21, 22]. It was soon discovered that these particles had the same mass and lifetime—indicating that they were the same particle—but, based on their decay products, they must have different parity. Lee and Yang proposed that the θ and the τ were in fact the same particle, but that it was undergoing a parity violating decay [23]. Their idea was confirmed by Wu in 1956, who showed that electrons were preferentially emitted from Co_{60} in one direction, and also by Garwin, Lederman, and

Weinrich in 1957, who studied $\pi^+ \to \mu^+ + \nu_\mu$ decays in a storage ring [24, 25].

In 1954, Yang and Mills replaced Fermi's contact interaction with a non-Abelian gauge theory that contained a spin-1 boson to mediate the force [26]. However, this boson was massless, which if true would have given the weak force infinite range, a feature that was not observed. In 1960, Glashow was able to modify Yang and Mill's theory by adding Sudarshan and Marshak's vector minus axial ($V-A$) model to produce a unified electroweak force described by the $\mathrm{SU(2)_L} \times \mathrm{U(1)}$ gauge group [1, 27]. $\mathrm{SU(2)_L}$ is a left-handed interaction and so violates parity as expected. Weinberg and Salamn finished up the model in 1967 when they added the Brout–Englert–Higgs mechanism which gave the vector bosons mass and so explained the short ranged nature of the weak interaction [2, 3, 28, 29].

2.1.3 The Strong Force

The theory of the strong force grew out of studies of atomic nuclei. In 1911, Rutherford discovered that atomic nucleus was a compact, positively charged object [30]. In 1917, Rutherford showed that larger nuclei were composed of hydrogen nuclei and so discovered the proton [31]. The discovery of the uncharged neutron in 1932 by Chadwick indicated that the atomic nucleus was made up of multiple types of nucleons, and that it could not be held together by the electromagnetic force [32]. In 1934, Yukawa—having noted that Fermi's contact interaction was too weak to hold nuclei together—tried to explain this nuclear force using meson exchange [33]. In 1947, Lettes, Occhialini, and Powell discovered the pion which seemed to confirm Yukawa's theory [34].

In the 1950s and early 1960s, dozens of new mesons were discovered, indicating the need for a new theory. Some of these new mesons seemed to have a new type of quantum number that limited their available decays, leading them to be called "strange" particles. An effort to explain these particles lead to the Gell-Mann–Nishijima formula [35, 36, 37]. This theory lead Gell-Mann and Ne'eman to come up with a classification scheme for mesons and baryons based on the $\mathrm{SU(3)_C}$ group which Gell-Mann named the Eightfold Way. In 1964, Gell-Mann and Zweig realized that the Eightfold Way implied that mesons were composed of sub-atomic particles which became known as quarks [38, 39]. Their model explained both isospin and the "strangeness" observed earlier as a consequence of the existence of three quarks, the up, down, and strange, which

were discovered at the Stanford Linear Accelerator Center (SLAC) in 1968. The model was modified in 1973 by Gross, Wilczek, and Politzer to include asymptotic freedom in which quarks interact weakly at high energies but strongly at low energies [40, 41]. This new combined model became known as quantum chromodynamics (QCD). QCD was used to predict the existence of the charm quark and, upon its success, was incorporated with electroweak theory to form the full $\mathrm{SU(3)_C \times SU(2)_L \times U(1)}$ symmetry group of the Standard Model.

QCD, like QED, can also be expanded in terms of a coupling constant, α_s. However, unlike α in QED, the value of α_s is dependent on the energy of the interaction as follows:

$$\alpha_\mathrm{s}\left(Q\right) \approx \frac{1}{\ln\left(Q/\Lambda_\mathrm{QCD}\right)} \tag{2.1}$$

where Λ_QCD is the QCD scale and Q is the interaction energy. This means that QCD can be expanded perturbatively for large Q, but not for low energy interactions.

2.1.4 Experimental Verification

The Standard Model has made numerous predictions which have been borne out by experiment. The neutral current interaction was observed by the Gargemelle experiment at CERN in 1973 shortly after it was predicted by Weinberg, Glashow, and Salam [42]. This confirmation of their theory won the trio the Nobel Prize in 1979.

The remaining quarks (the up, down, and strange having been discovered in 1968) were found over the next twenty years. The charm was found at the SLAC and Brookhaven National Laboratory in 1974 via J/ψ decays [43, 44]. The bottom was discovered at Fermi National Accelerator Laboratory (FNAL) in 1977 by the E288 experiment [45]. The final quark, the incredibly heavy top, had to wait until 1995 to be discovered by the CDF and D0 experiments running at the Tevatron at FNAL [46, 47].

The W and Z bosons were discovered at CERN in 1983 at the UA1 and UA2 experiments running on the Super Proton Synchrotron [48, 49, 50, 51]. These bosons were an excellent test of the Standard Model as their masses could be very exactly calculated, and the newly discovered bosons had masses that matched the calculations to high precision. The final piece of the Standard Model, the Higgs bosons, was discovered by ATLAS and CMS in 2012 using the LHC at CERN [52, 53]. Although precision

measurements of the Higgs are still ongoing, it so far precisely matches the predictions of the Standard Model.

2.1.5 Components of the Standard Model

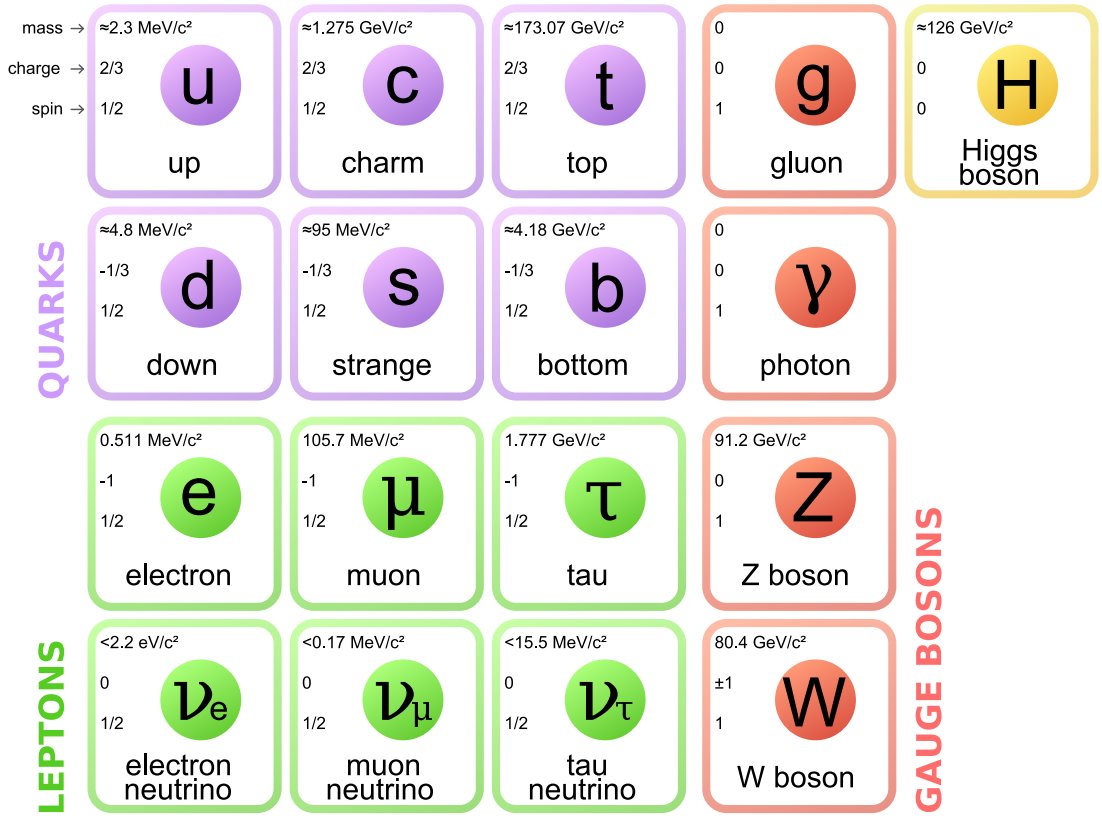


Figure 2.1: The particles of the Standard Model with information about their type, mass, charge, and spin. The quarks and leptons make up matter, while the gauge bosons mediate interactions. The Higgs gives mass to the W and Z bosons.

There are two types of particles in the Standard Model: fermions—with half integer spin—and bosons—with integer spin. The fermions, which make up all of the matter in the universe, are further subdivided into two groups, leptons and quarks, while the bosons are divided into gauge bosons and the Higgs. The various particles are shown schematically in figure 2.1.

Leptons have spin–1/2 and have charge $q = -1$ for the electron, muon, and tau, or 0

for the neutrinos. They interact electromagnetically (if they have a non-zero charge) and weakly. There are three generations, or flavors, of leptons and each generation consists of a charged lepton, and an uncharged, nearly massless neutrino. The lightest of the charged lepton generations is the electron, which is stable. The next two generations contain the muon and the tau, which are unstable and eventually decay to electrons. The tau, being very heavy, decays quickly, while the muon is stable long enough to escape a particle detector. Leptons are important in particle detection as they provide a very clean decay signature. Neutrinos are very difficult to detect in particle detectors, and so their presence is inferred from the missing energy in the vector sum of all particles in the collision. This analysis uses electrons to make its measurement.

Quarks also have spin–1/2 although, unlike the leptons, their charge is fractional and so takes values of $q = 2/3$ or $-1/3$. They interact strongly, electromagnetically, and weakly. There are three generations of quarks, with each successive generation having higher mass constituents. There are six flavors of quarks, with each generation containing two. The first generation consists of the up and down (u and d) quarks. These quarks are stable and make up protons, neutrons, and pions. The next generation of quarks contains the charm and the strange (c and s). They form heavier states that decay quickly, such as the J/ψ and kaons. The final generation consists of the heavy bottom (b) and the extraordinarily heavy top (t)—the most massive particle in the Standard Model. Bottom quarks can form bound states, but top quarks are so heavy they decay before any bound states can form.

Quarks carry color charge, of which there are three: red (r), blue (b), and green (g). Strongly interacting objects obey confinement, which means that only color neutral (colorless) states are allowed. Because of confinement, quarks bind together into colorless composite particles. These particles are called mesons—with two quarks ($q\overline{q}$)—and baryons—with three quarks (qqq). When an object containing quarks breaks up, the individual colored fragments will create additional colored objects in order to remain color neutral. This leads to the formation of "jets", sprays of high energy particles that originate from one of these fragments as it tries to maintain its colorless state.

Bosons are the second type of particle in the Standard Model. They are further subdivided into gauge bosons—which mediate the three forces—and the Higgs boson—which gives mass to the W and Z bosons.

The gauge boson which mediates the strong force is the gluon. Gluons interact with objects that carry color; as carriers of color themselves, this means that gluons interact with both quarks and other gluons. Gluons can have any one of eight different possible color-anticolor superpositions that form a color-octet. This number comes from the number of generators of $\mathrm{SU(3)_C}$. Such octets are not unique, but a commonly used definition is listed in table 2.1.

$$\begin{array}{ll} \left(r\bar{b}+b\bar{r}\right)/\sqrt{2} & -i\left(r\bar{b}-b\bar{r}\right)/\sqrt{2} \\ \left(r\bar{g}+g\bar{r}\right)/\sqrt{2} & -i\left(r\bar{g}-g\bar{r}\right)/\sqrt{2} \\ \left(b\bar{g}+g\bar{b}\right)/\sqrt{2} & -i\left(b\bar{g}-g\bar{b}\right)/\sqrt{2} \\ \left(r\bar{r}-b\bar{b}\right)/\sqrt{2} & \left(r\bar{r}-2b\bar{b}+g\bar{g}\right)/\sqrt{6} \end{array}$$

Table 2.1: One of the possible color-octets. The colors are red (r), blue (b), green (g), and their anticolors ($\bar{r}$, $\bar{b}$, and $\bar{g}$) .

There are four gauge bosons that mediate the electroweak interaction: the photon (γ), the Z, and the $\mathrm{W^{\pm}}$. The photon and the Z are uncharged, while the $\mathrm{W^{\pm}}$ carry charge of ± 1. The W and Z are not the particles described by the $\mathrm{SU(2)_L \times U(1)}$ group, but are instead linear combinations of these fields created through combination with the Brout–Englert–Higgs mechanism. The W participates in interactions that change quark and lepton flavor, for example $\mathrm{t \to Wb}$ or $\mu \to \mathrm{W}\nu_\mu$, while the Z does not.

The Z boson is an excellent probe of precision physics as its well measured mass ($91.1876 \pm 0.0021\,\mathrm{GeV}$) and its sharp width ($2.4952 \pm 0.0023\,\mathrm{GeV}$) make it easy to identify from its decay products [54]. In a hadron collider, the most common $\mathrm{Z \to q\bar{q}}$ decay mode is difficult to select as there are many hadronic jets in each event, and so $\mathrm{Z} \to \ell^+\ell^-$ decay modes are preferred. In this analysis we look at the $\mathrm{Z \to e^+e^-}$ decay mode; our collaborators are working on performing a similar measurement with $\mathrm{Z} \to \mu^+\mu^-$ decays. A few common decay modes and their branching fractions are listed in table 2.2.

The final piece of the Standard Model, the Higgs boson, gives mass to the weak bosons through interactions with the Higgs field. In the Lagrangian of the Standard Model, it is not possible to write gauge invariant mass terms for the W and Z bosons. The Brout–Englert–Higgs mechanism adds a complex scalar field whose symmetry is spontaneously broken leading to the W and Z masses. This rotates the primordial

Mode	Fraction (Γ_i/Γ)
$Z \to q\overline{q}$	$69.91 \pm 0.06\,\%$
$Z \to e^+e^-$	$3.363 \pm 0.004\,\%$
$Z \to \mu^+\mu^-$	$3.366 \pm 0.007\,\%$
$Z \to \tau^+\tau^-$	$3.370 \pm 0.008\,\%$
$Z \to \nu\overline{\nu}$	$20.00 \pm 0.06\,\%$

Table 2.2: Selected decay modes of the Z boson.

vector boson plane consisting of the massless B^0 and W^0 fields into the Z boson and the photon. The angle of this rotation is called the Weinberg angle, θ_W. This symmetry breaking also produces a spin–0 boson, the Higgs.

The Weinberg angle is defined by the mass of the Z and W bosons as follows:

$$\sin\theta_W \equiv \frac{M_W}{M_Z} \tag{2.2}$$

This angle is present in many terms of the electroweak Lagrangian, and so controls the rates of electromagnetic decays as well as the couplings of the Higgs to the vector bosons. It can not be calculated and so must be measured experimentally via the masses of the vector bosons.

2.2 Z Boson Transverse Momentum

The leptonic decays of the Z are an excellent probe of QCD as neither the Z nor the leptons carry color charge. This means that there is no color flow between the initial and final states, and so the only QCD signature encoded in the decay of the Z is that of the initial interaction. One particularly useful variable for probing QCD is the transverse momentum of the Z boson, Q_T, where transverse is defined relative to the beamline in the collider [55, 56, 57]. As the beam protons have near-zero momentum transverse to the beamline, the Q_T of the Z boson comes from QCD process like initial state radiation, which are discussed in section 2.3. Low values of Q_T probe the non-perturbative regions of QCD as, due to asymptotic freedom, it is these low momentum transfer interactions

where α_s is large.

In addition to aiding in the understanding of QCD processes, measuring the Z Q_T spectrum helps improve measurements of the W mass (M_W). The W mass is an interesting quantity to measure because of the role it plays in determining the Weinberg angle and because it can be calculated once all the other inputs to the Standard Model are known. With the discovery of the Higgs, it is now possible to compared the measured value of M_W with the theoretical value. Performing such a fit to determine M_W leads to a theory result that is better constrained then the experimental measurement [58]. There is some disagreement between the fit and the measurement of M_W, as shown in figure 2.2, which could be indicative of new physics, but the comparison is inconclusive; a better measurement is needed to resolve the issue.

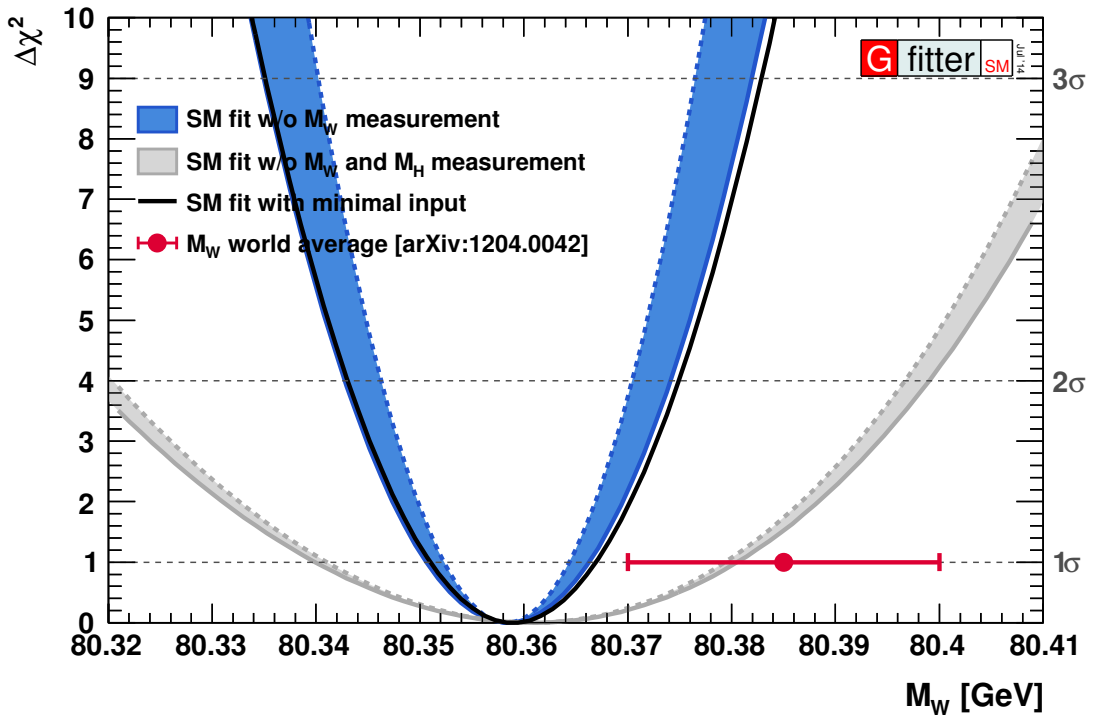


Figure 2.2: The measured value (point) of M_W compared to fit results from Gfitter (band). There is some tension between the measurement and the fit, but a better measurement is needed to determine if it is significant.

The W mass is difficult to measure. In an electron collider like LEP, Z bosons are

easy to produce at rest as the $e^+e^- \to Z$ channel is available allowing a high precision mass measurement. Unfortunately, it is not possible to build a high luminosity collider to produce W bosons at rest through the $e\nu_e \to W$ channel, and so instead the $q\overline{q} \to W$ channel is used at hadron colliders like the LHC. However, because of the large numbers of hadronic jets at a hadron collider, only the $W \to \ell\nu$ decay modes can be selected, and so the mass must be measured using the single observable lepton because the neutrino escapes unseen. This is done by creating a template with simulated data and then using the template to fit for the mass from lepton's momentum. As both the mass and Q_T of the W affect the lepton's momentum, any uncertainty in the Q_T distribution of the simulation translates to an uncertainty in the final mass measurement. As the W is subject to the same processes described below for the Z, the measurement of the Z Q_T spectrum constrains the W Q_T spectrum.

2.2.1 Z Boson Differential Cross Section

The measurement that will be performed in this analysis is a differential cross section measurement. A cross section measurement is, at its heart, a simple counting experiment. The cross section of a process, Z production for instance, is given by:

$$\sigma(Z) = \frac{N_Z}{\mathcal{L}} \tag{2.3}$$

where N_Z is the number of Z bosons that were created and $\mathcal{L}$ is the integrated luminosity of the data which is a measurement of how many interaction opportunities took place. A differential cross section is used to measure the cross section of a process as a function of a variable, for instance Q_T. Then the cross section becomes:

$$\frac{d\sigma(Z)}{dQ_{T,k}} = \frac{N_{Z_{Q_{T,k}}}}{\Delta Q_{T,k} \cdot \mathcal{L}} \tag{2.4}$$

where $Q_{T,k}$ is a range of Q_T, $N_{Z_{Q_{T,k}}}$ is the number of Z events with Q_T that fall within the range, and $\Delta Q_{T,k}$ is size of the range. This measurement is still dependent on the luminosity which can be difficult to accurately measure. This dependency can be removed by dividing by the total cross section:

$$\frac{1}{\sigma(\mathrm{Z})}\frac{\mathrm{d}\sigma(\mathrm{Z})}{\mathrm{d}Q_{\mathrm{T},k}} = \frac{\mathcal{L}}{N_Z}\frac{N_{\mathrm{Z}_{Q_{\mathrm{T},k}}}}{\Delta Q_{\mathrm{T},k}\cdot\mathcal{L}} = \frac{N_{\mathrm{Z}_{Q_{\mathrm{T},k}}}}{\Delta Q_{\mathrm{T},k}\cdot N_Z} \tag{2.5}$$

2.2.2 The Proton Parton Model

At the very high energies of the LHC ($\sqrt{s} = 8\,\mathrm{TeV}$ in 2012), a proton is not well described by assuming that it is composed of only three valence quarks as the energy of the proton is much higher than its own binding energy ($m_{\mathrm{p}} \approx 1\,\mathrm{GeV}$). Instead, it is described by a parton model—a model in which there are three valence quarks, but also numerous gluons and "sea quarks", where the sea quarks are a superposition of quark and antiquark same-flavor pairs. These constituents of the protons are referred to as "partons". The parton model is formulated in the infinite momentum frame, a frame where the proton has very high energy compared to its rest mass ($E \gg m_{\mathrm{p}}$). In this frame the mass of the proton can be neglected.

When two protons collide at the LHC, it is a parton from each that interacts in the hard scattering process. This parton-parton interaction can be considered as independent of the other partons because the internal quantum state of a proton in the infinite momentum frame is "frozen" due to relativistic time dilation. Each of the partons that takes part in the interaction will carry only a faction of their parent proton's total momentum. This fraction is parameterized by the Bjorken x_i variable, defined as $x_i P = p_i$, where P is the total proton momentum and p_i is the momentum of a parton. In this framework, all of the momentum is along the beamline as the only possible source of p_{T} is from the binding energy of the proton, which is approximately m_{p}, and so negligibly small.

The x_i value of a parton is not fixed, but is instead a probability distribution that is dependent on the flavor of the parton and the energy scale (Q^2) of the interaction. If there were only three, non-interacting quarks within the proton, then the PDF would simply be a Dirac delta-function at $x_i = 1/3$, but as the quarks are strongly interacting they are constantly exchanging momentum with each other via gluons. The gluons carry some of the momentum, and so the delta function for each quark is smeared out. Additionally, higher order processes tend to enhance the lowest momentum regions, causing the distributions to rise at low x_i.

Collections of these distributions for the various flavors are called parton distribution functions (PDF). PDFs can not currently be calculated from QCD, and it is not known if doing so is even possible. Instead, PDFs are models which are fit to data and extrapolated to new interaction scales using perturbative QCD. An example set of PDFs from the MSTW collaboration[59] is shown in figure 2.3. In the low interaction energy case, shown in the left plot, the u and the d quark distributions have peaks at high x_i, as the valence quarks carry most of the momentum. In the high interaction energy case, shown on the right plot, a small peak is still seen in the u and d quark distributions, but the sea quarks and gluons also carry a large amount of the total momentum.

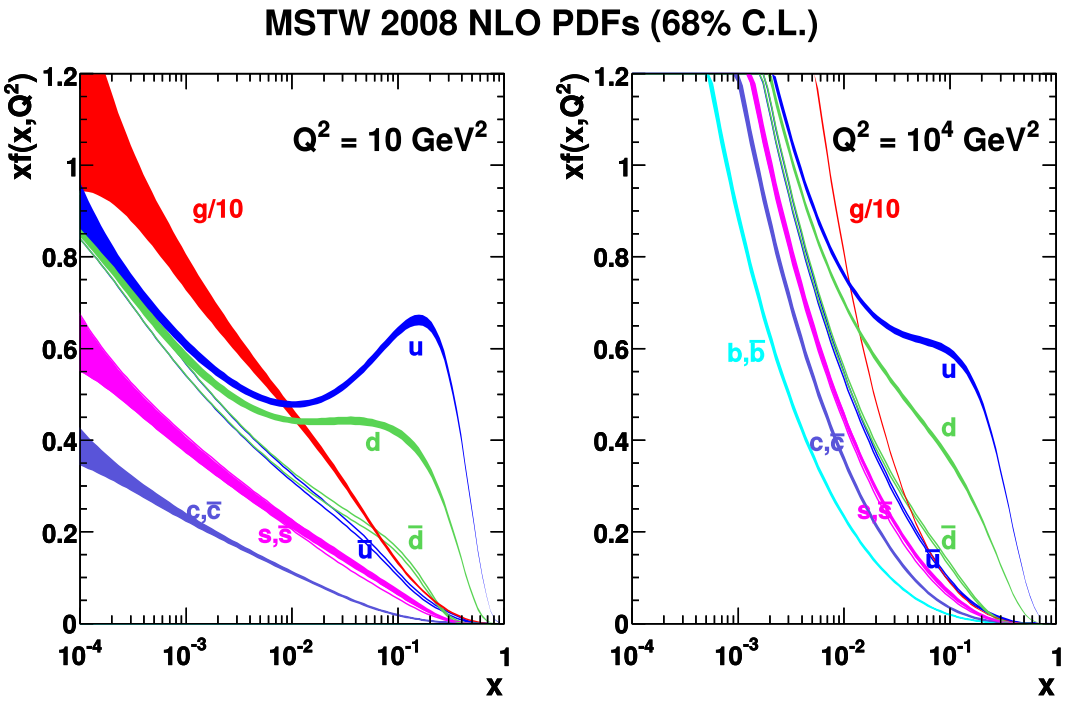


Figure 2.3: Example PDFs from the MSTW collaboration for interaction energy scales of $Q^2 = 10\,\mathrm{GeV}^2$ (left) and $Q^2 = 10^4\,\mathrm{GeV}^2$ (right). The bands represent the x_i distributions for the various partons.

2.3 Production of Z Bosons In Proton-Proton Collisions

In order to understand where the transverse momentum of a Z boson comes from, it is important to understand how they are produced at the LHC.

2.3.1 Drell–Yan Production

Drell–Yan (DY) production is the lowest order in α_s process of producing dilepton pairs via the Z boson [60, 61]. The Feynman diagram of this process is shown in figure 2.4. As discussed previously, the two quarks will have $p_T \approx 0$ as most of the proton's momentum (and hence, the parton's momentum) is along the beamline, and so the Z will also have $Q_T \approx 0$. The cross section for the DY process, in terms of the x_i variables of the incoming partons, is:

$$\frac{d^2\sigma}{dx_1 dx_2} = \frac{4\pi\alpha^2}{9x_1x_2s} f(x_1, x_2) \tag{2.6}$$

$$f(x_1, x_2) = \sum_a Q_a^2 \left[f_1^a(x_1) f_2^{\bar{a}}(x_2) + f_1^{\bar{a}}(x_1) f_2^a(x_2) \right] \tag{2.7}$$

where x_i is the momentum fraction carried by the parton from the ith proton, $f_i^a(x_i)$ is the individual PDF for a quark of flavor a from the ith proton, s is the Mandelstam variable that is the square of the center-of-mass energy, and the sum is over all quark flavors.

2.3.2 Higher Order Production

In addition to the DY process discussed previously, there are higher order terms in α_s that contribute to Z boson production. Some of these terms are shown in figure 2.5.

The first type of term, shown in figures 2.5a and 2.5b, contains initial state radiation (ISR). As discussed earlier, protons are highly dynamic systems with their quarks interacting through multiple gluon exchanges. When a collision happens, one of the partons is "removed" from the proton, leaving a "beam remnant" behind that is unbalanced in both momentum and energy. This most often tears the remnant apart, leading to colored objects escaping and hadronizing. These colored objects are ISR. The diagrams shown here have a gluon radiating from the incoming quark, one of the lowest order in

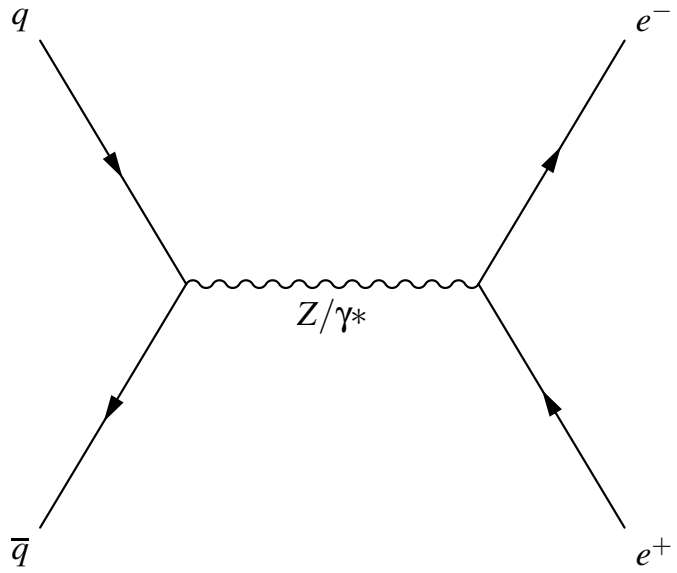


Figure 2.4: Feynman diagram of Drell–Yan $Z \rightarrow e^+e^-$ production.

α_s ISR processes. The quark can be left with some non-zero p_T as it gains momentum opposite that of the gluon, and so it may impart some Q_T on the Z. The probability of radiating a gluon increases as the energy of the radiated gluon decreases, and so ISR often introduces a small amount of Q_T.

The second type of term, shown in figures 2.5c and 2.5d, is one where the quark interacts with a gluon from the other proton and radiates a Z boson. Gluons are common in LHC interactions and so these terms are non-negligible, although their exact contribution is difficult to calculate. These interactions have quarks with energies close to the Z mass, and so they have the potential to generate higher p_T Z bosons.

2.3.3 Final State Radiation

Although the $Z \rightarrow e^+e^-$ decay is a very clean decay with no color interaction, there is one final state process that can impact the measurement of Z Q_T. In this process, known as final state radiation (FSR), the decay leptons radiate photons. The diagram of this process is shown in figure 2.6. Measurements and theories treat FSR differently and so we define three types of generator level electron that each account for FSR in a different manner: Born, bare, and dressed. Generator level refers to the variables

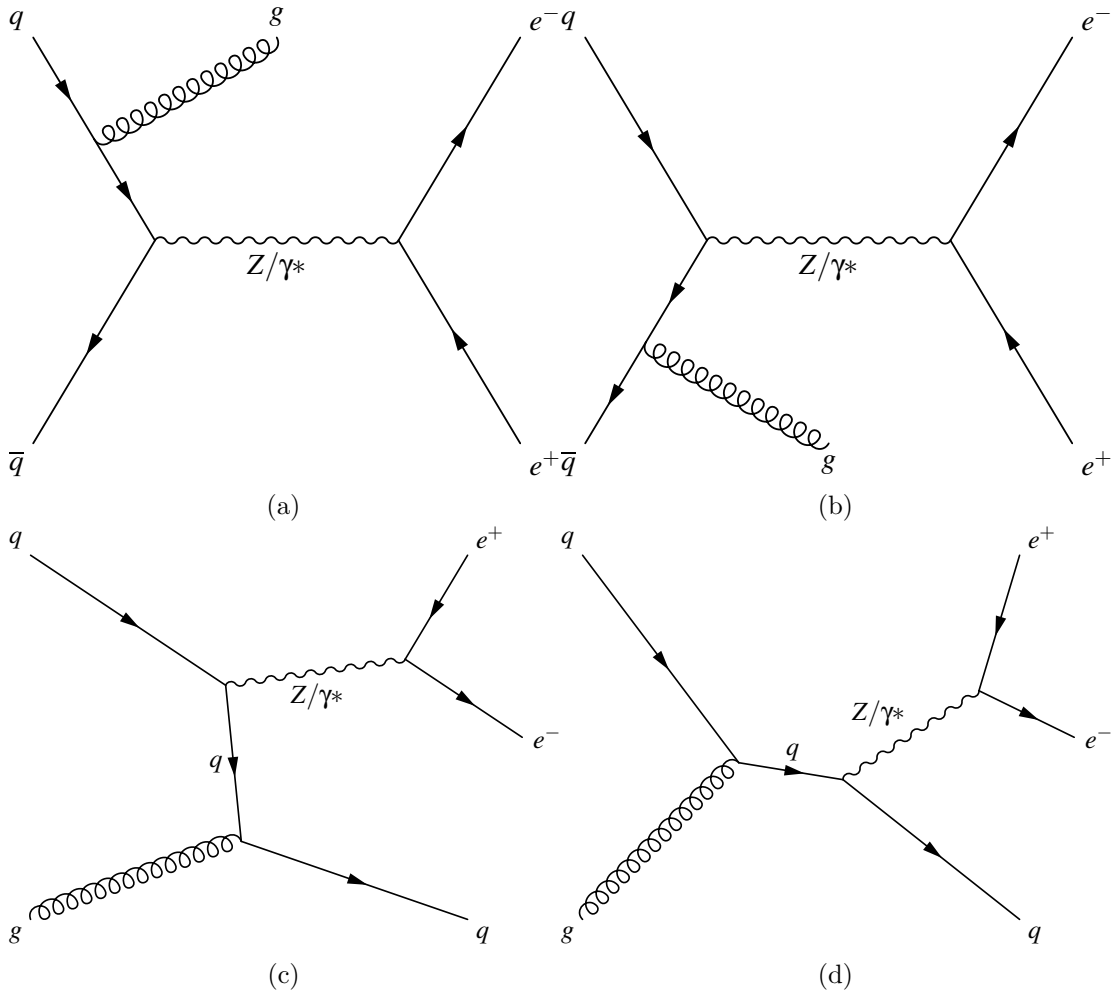


Figure 2.5: Higher order in α_s $Z \to e^+e^-$ Feynman diagrams. Figures 2.5a and 2.5b are ISR where one of the incoming quarks radiates a gluon. In figures 2.5c and 2.5d the quark radiates a Z.

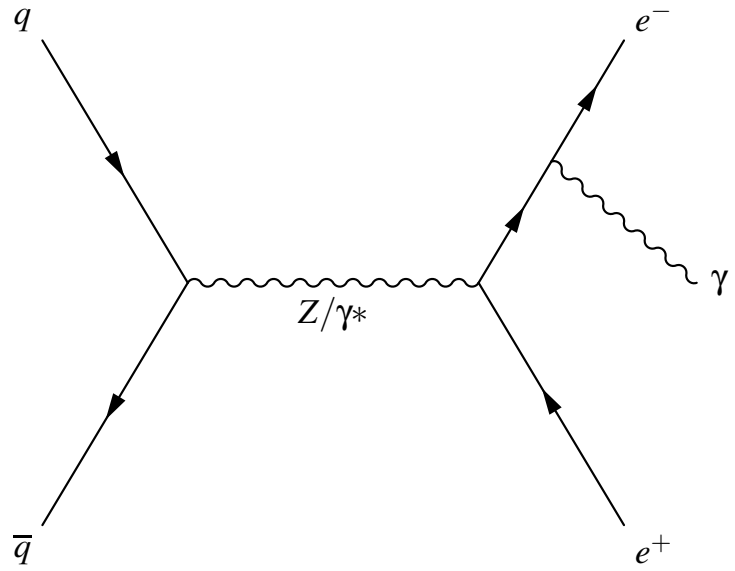


Figure 2.6: Feynman diagram of $Z \to \ell^+\ell^-$ with FSR.

describing a collision event as computed by a Monte Carlo (MC) generator program without considering interaction with the detector or the analysis reconstruction chain. Generator level quantities are, therefore, free of detector and measurement effects.

Born generator level electrons are electrons directly after the $Z \to e^+e^-$ decay but before the electrons have radiated any FSR photons. This definition is what most theoretical results will provide.

Bare generator level electrons are Born electrons after they have radiated all of their FSR photons. This definition most closely matches how muons are measured in a detector as the momentum of a muon is measured using only its track, ignoring any photons.

Dressed generator level electrons are bare electrons with their FSR photons added back in vector sum. The photons are only added if they are within a cone of size $\Delta R < 0.1$ around the electron. This definition most closely matches how electrons are measured in a detector; the energy of the electron is measured using a calorimeter which also integrates the nearby photons into the measurement.

Whenever generator level quantities are used, for example when performing the unfolding as discussed in section 6.4, dressed electrons are used.

2.4 A New Variable: ϕ^*

In the low Q_{T} region—the region which contains most of the Z events and is also the region governed by non-perturbative QCD—the measurement is dominated by the systematic uncertainties and the experimental resolution. Theorists have proposed a new variable, ϕ^*, that depends only on the direction of the leptons and not on their energy in order to reduce the systematic uncertainties associated with the measurement [62]. ϕ^* is correlated to $Q_{\mathrm{T}}/m_{\ell\ell}$, where $m_{\ell\ell}$ is the invariant mass of the two leptons, so it probes the same physics as Q_{T}. This variable is less susceptible to detector resolution effects as the energy and momentum resolutions of a detector are, in general, worse than the position resolution. This is due to the fact that the position of leptons is measured with a very fine-grained tracking device whereas the energy is measured with a calorimeter.

The definition of ϕ^* is:

$$\phi^* = \cot \frac{\Delta\phi}{2} \operatorname{sech} \frac{\Delta\eta}{2} \tag{2.8}$$

where $\Delta\phi$ is the azimuthal opening angle between the leptons, and $\Delta\eta$ is a measure of the scattering angle of the leptons with respect to the beam direction in the rest frame of the dilepton system.

A simplified simulation of $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ decays was run in order to demonstrate the insensitivity of ϕ^* to detector effects as compared to Q_{T}. The p_{T} of the electrons in this simulation was smeared by 10 % while the position of the electrons was smeared by 1 %. The ϕ^* distribution and Q_{T} distributions were then constructed using these electrons and compared to the input distributions. The results are shown in figure 2.7. As can be seen, the ϕ^* distribution after smearing is much closer to the input ϕ^* distribution than the Q_{T} distribution is to the input Q_{T} distribution.

ATLAS, one of the other experiments on the Large Hadron Collider (LHC), has measured ϕ^* at a center-of-mass energy of 7 TeV ($\sqrt{s} = 7\,\mathrm{TeV}$) [63]. D0 has also measured ϕ^* at the Tevatron at $\sqrt{s} = 1.96\,\mathrm{TeV}$ [64, 65]. This thesis presents the first measurement of ϕ^* at $\sqrt{s} = 8\,\mathrm{TeV}$.

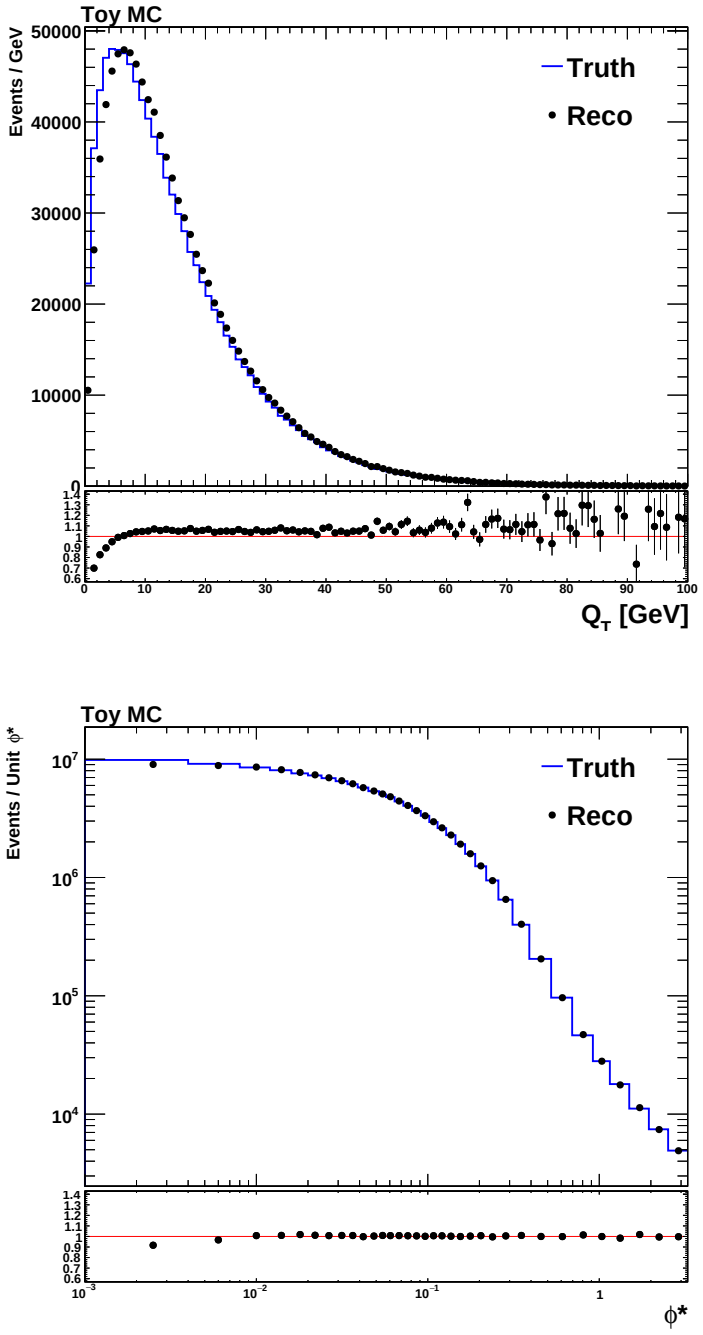


Figure 2.7: The top figure shows the Q_{T} distributions as input into a simulation (histogram) and as recovered after smearing of lepton's p_{T} (points). The bottom figure shows the ϕ^* distributions as input (histogram) and as recovered after smearing (points).

Chapter 3

The CMS Experiment

3.1 The Large Hadron Collider

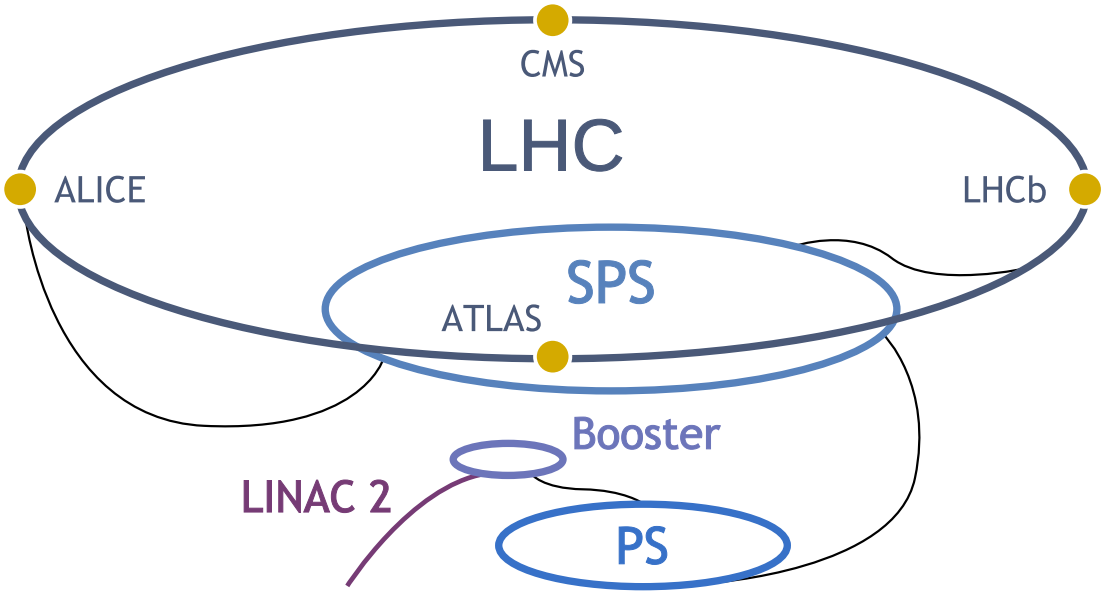


Figure 3.1: A schematic view of the LHC and the location of the four experiments. Also shown are the accelerators that feed protons into the LHC.

The Large Hadron Collider (LHC) is the world's highest energy and largest particle accelerator with a maximum design center-of-mass energy of 14 TeV and a radius of

2804 m [66]. It collides protons on protons, as well as protons on ionized lead (Pb) and Pb on Pb. In 2012, when LHC most recently produced collisions, it had a center of mass energy of 8 TeV; when it turns back on in 2015 after upgrades, it will run at 13 TeV. The LHC is located near Geneva, Switzerland, although it is so large that most of the accelerator (including Point 5, where the CMS detector is located) is in France.

A number of smaller accelerators are used together in series to accelerate protons to the energies necessary to be injected into the LHC. The first step is a linear accelerator, Linac 2, which accelerates protons from rest to 50 MeV. These protons are then injected into a chain of three circular accelerators, each injecting into the next. The first of these accelerators is the Proton Synchrotron Booster (PSB) which accelerates the protons to 1.4 GeV. The second is the Proton Synchrotron (PS) which accelerates the protons to 26 GeV. The third and final accelerator is the Super Proton Synchrotron (SPS) which accelerates the protons to 450 GeV and injects directly into the LHC.

Bunches of protons are accelerated using this system and injected into the LHC to form two counter-rotating beams. When the desired number of bunches have been injected into the LHC, the LHC accelerates them to 4 TeV. When the beams have reached their nominal energy, they are focused and brought into collision at four different points on the ring where the various experiments (ALICE, ATLAS, LHCb, and CMS) are located. A cartoon layout of the LHC and its accelerator chain is shown in figure 3.1.

The beams are steered around the accelerator ring by a series of superconducting, dipole magnets. When running at a center-of-mass energy of 8 TeV, these magnets operate at roughly 7.5 T. There are also quadrupoles to focus the beam and some higher order magnets around the ring to correct for lattice defects. The bunches of protons are accelerated by 16 superconducting radio frequency cavities. These cavities accelerate slower protons while slowing faster ones, thereby keeping the proton bunches compact in both real and momentum space. There is room for 2808 bunches in LHC separated by 25 ns, although in 2012 50 ns spacing was used and so there were only 1374 bunches of which 1368 were brought into collision at CMS and ATLAS.

The LHC is also the highest luminosity collider in the world. The instantaneous luminosity is given by:

$$\mathcal{L} = fn\frac{N^2}{\sigma} \tag{3.1}$$

where f is the frequency of interaction (which is fixed by the LHC's circumference), n is the number of bunches in a beam, N is the number of protons per bunch (with the N^2 coming from the assumption that there are the same number of protons in the two colliding bunches), and σ is the area profile of the beams.

Although a higher luminosity means more particles are produced and more data can be collected, the maximum luminosity is limited by several practical factors. The first factor is cost; a higher luminosity general requires a more expensive machine as the ring is either made larger or the technology needed to run the machine is made more complex. The cost of the detectors also increases as they require more channels to separate the larger number of particles, faster readout to deal with the increased event rate, and higher bandwidth to read out the larger numbers of channels. The second is the challenge that higher luminosities present to the analyzers. The luminosity can be increased by increasing the number of protons in a bunch or by squeezing the bunches more tightly, but eventually the probability of getting multiple proton-proton interactions per bunch crossing becomes large, leading to a phenomenon known as pileup. These extra interactions add additional particles to the detector and can make it difficult to separate interesting events from uninteresting background. The luminosity can also be increased by increasing the number of bunches in the machine, but this decreases the time between the collisions and leads to a phenomenon known as out-of-time pileup which can also obscure interesting events. The third factor is that high radiation doses damage the detectors. Plastic and crystal scintillators darken while silicon detectors become noisy. This forces the detectors to replace their components more frequently.

In 2012, the optimal luminosity was achieved by running with bunches spaced by 50 ns instead of the design nominal bunch spacing of 25 ns. This larger bunch spacing was chosen because at smaller spacing the bunches were destabilized by electron clouds—clouds of electrons knocked out of the beam pipe by the beam's synchrotron radiation, as well as secondary electrons freed when the photoelectrons impact their surroundings.

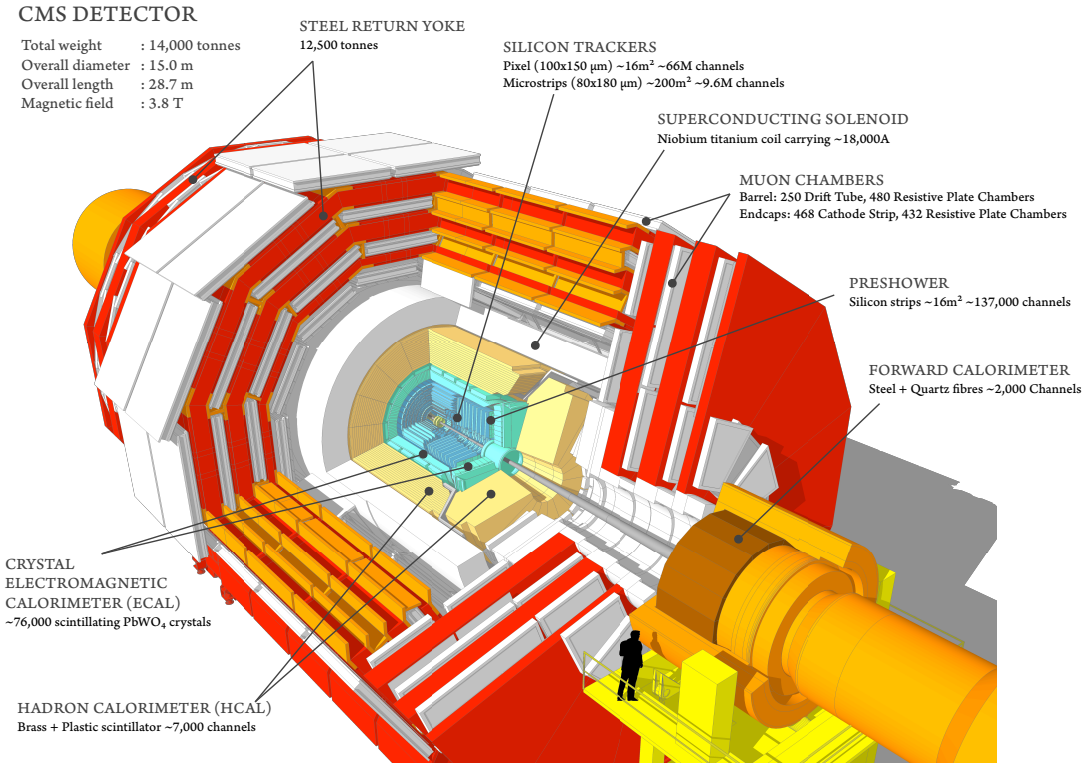


Figure 3.2: A cut-away view of the CMS experiment.

3.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is one of two general-purpose particle detectors built on the LHC ring [67, 68]. CMS is designed to detect the very high energy, subatomic particles that are produced in the LHC's proton-proton collisions. CMS is designed to have a high acceptance and efficiency for these collisions. Here acceptance indicates the area in both physical space as well as the area in energy and momentum space in which the detector can detect particles. Efficiency is the probability that a particle in CMS's acceptance region is properly measured. In order to have a high acceptance and efficiency, CMS must be both large—to cover a large area of physical space—and dense—to cover a large area in momentum and energy space. CMS is 21.6 m in length, 14.6 m in diameter, and weighs 14 kt.

CMS is built as a series of nested, finite cylinders, where each cylinder is a separate subdetector. The beams enter the detector along the axis of the cylinder. The collision point is in the center. There are a pair of endcaps on either side of the cylinder to increase the acceptance of the detector. The endcaps and central region of the cylinder (called the barrel) overlap to prevent particles from escaping undetected through the crack. A cutaway of the detector is shown in figure 3.2.

The coordinate system used by CMS is as follows: the origin is the nominal interaction point at the center of CMS, the x-axis is defined to point to the center of the LHC ring, the y-axis is defined as vertically up, and the z-axis points counter-clockwise and tangent to the LHC ring such that it forms a right-handed coordinate system with the x-axis and y-axis. CMS uses a cylindrical coordinate system with coordinates (η, ϕ). The azimuthal angel ϕ is in the x–y plane measured from the x-axis so that $\phi = \pi/2$ at that y-axis, while η is the pseudorapidity defined by:

$$\eta = -\ln \tan \frac{\theta}{2} \tag{3.2}$$

where θ is the polar angle with respects to the positive z-axis. The magnetic field points along the z-axis.

3.2.1 Inner Tracking System

The inner tracking system, referred to as the tracker, is the subdetector closest to the interaction point. The tracker's primary purpose is to measure the charge of particles, the momentum of these same particles, and the location of interaction vertices—both the primary vertex, the various additional proton-proton vertices from pileup, and the vertices of long-lived particles like b mesons. The tracker consists of two types of silicon detectors: silicon pixels and silicon strips. The tracker covers a pseudorapidity range of $|\eta| < 2.4$. The pseudorapidity coverage of the tracker is the determining factor behind the $|\eta|$ bounds used in our acceptance definition, discussed in section 6.1.

Pixel Tracker

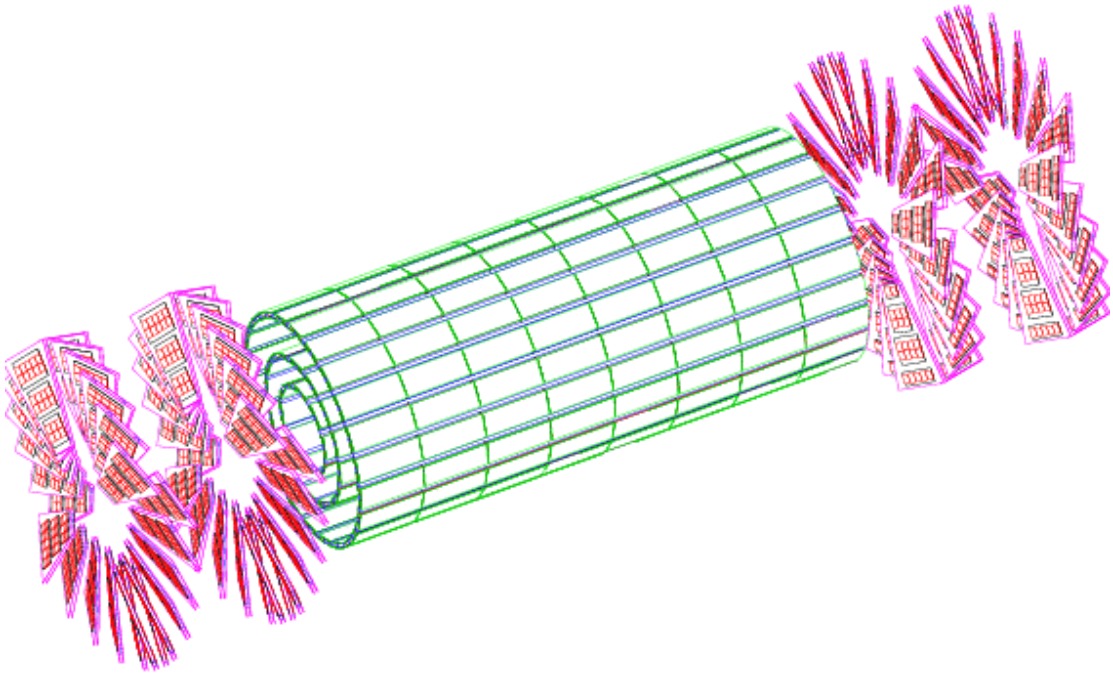


Figure 3.3: A rendering of the CMS pixel Tracker.

The silicon pixels are used in the region closest to the beam pipe where the particle flux is the highest and hence the finest granularity is needed. Their primary purpose is to very accurately locate the primary and secondary vertices in a collision.

There are three barrel layers of the pixel tracker at radii of 4.3 cm, 7.3 cm, and 10.2 cm, each with a length of 53 cm. At each end, there are two endcap annular disks as well placed at $|z| = 34.5$ cm and 46.5 cm. Each of these disks has an inner radius of 6 cm and an outer radius of 15 cm.

Each pixel has an area of $100 \times 150\,\mu m^2$, but the resolution of the tracker is better than that because of charge sharing. If a charged particle ionizes multiple pixels, then a weighted average of the charges can be used to get sub-pixel resolution on the location of the hit. In order to increase the charge sharing, a large Lorentz angle (23°) is used. The blades which make up the endcap disks are fanned out in a turbine-like geometry with a rotation of 20° to benefit from the same effect. The layout of the pixel detector is shown in figure 3.3.

Strip Tracker

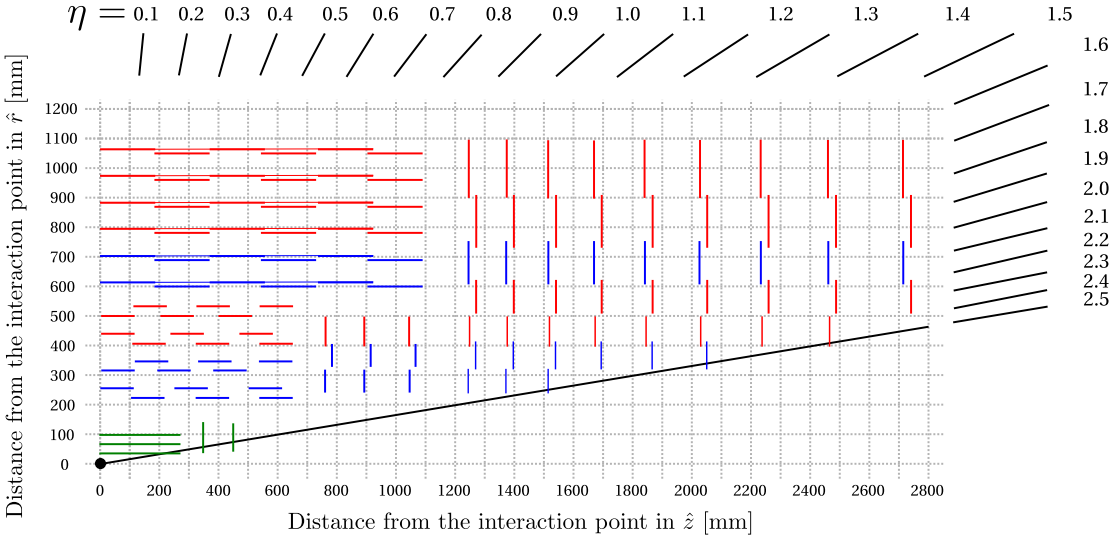


Figure 3.4: A quarter cross-sectional view of the CMS tracker.

The silicon strips are used further from the beam pipe than the pixels and cover a much large radius from the beam pipe. The strip tracker has 200 times the area of the pixel tracker, and so it was not economically feasible to use the more expensive pixel detectors in this region. The strips' primary purpose is to measure the momentum and curvature of the charged particles by extending the tracker's coverage over a larger

radius.

Although the strips provide only two points in space to locate a hit (as compared to three for the pixels), this coarser geometry is adequate because of the much lower particle flux in this region of the detector. For the momentum, the location of the hit in the r–ϕ plane is the most important information, and so the strips are aligned parallel to the magnetic field.

The strip tracker is itself divided into multiple components. In the barrel there is the TIB (Tracker Inner Barrel) and the TOB (Tracker Outer Barrel). The TIB consists of four layers and has coverage up to $|z| < 65\,\text{cm}$. The silicon sensors making up the TIB have a strip pitch which varies between 80 µm to 120 µm and have a thickness of 320 µm. The first two layers are constructed with a double layer of modules with a stereo angle of 100 mrad, providing information about the location of the hit in both the r–ϕ and r–z plane. The TOB consists of six layers and has coverage up to $|z| < 110\,\text{cm}$. The TOB has a strip pitch which varies from 120 µm to 180 µm and have a thickness of 500 µm. The thicker sensors are able to be used in this region because the radiation levels are lower. Having thicker sensors helps to maintain a high signal-to-noise ratio. Just like the TIB, the first two layers of the TOB are also built with a double layer of modules with a stereo angle of 100 mrad.

The ends of the strip tracker consists of the TEC (Tracker Endcap) and the TID (Tracker Inner Disk). The TEC consists of nine disks in the region $120\,\text{cm} < |z| < 280\,\text{cm}$. The TID consists of three disks in between the end of the TIB and the start of the TEC. The TEC and TID consist of modules arrayed in rings around the beam line, with the face of each module pointed towards the interaction point so that they have varying orientations depending on their distance from the interaction point. The first two rings of the TID and the first, second, and fifth rings of the TEC are built with a double layer of modules. The modules in the TID and the first three disks of the TEC have a thickness of 320 µm while the rest of the modules in the TEC have a thickness of 500 µm. The TID and the first three disks of the TEC have a strip pitch which varies from 97 µm to 143 µm while the rest of the modules in the TEC have a strip pitch from 143 µm to 183 µm. The layout of the entire tracker, including the strip tracker, is shown in figure 3.4.

3.2.2 Electromagnetic Calorimeter

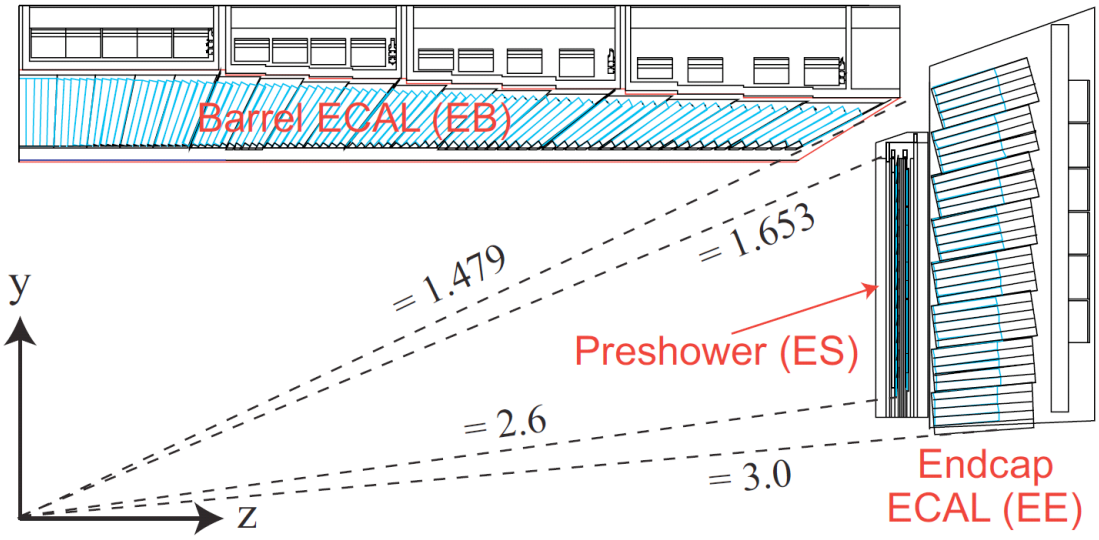


Figure 3.5: A quarter cross-sectional view of the CMS Electromagnetic calorimeter showing the barrel (EB), endcap (EE), and preshower (ES) components.

The electromagnetic calorimeter (ECAL) in CMS is built around the tracker. Its primary purpose is to measure the energy of electrons and photons. ECAL's design was motivated by the need to measure the $H \to \gamma\gamma$ decay. In order to have suitable energy resolution, most of the energy of the particle must be contained within the calorimeter, and so ECAL had to be very dense. In order to separate the highly boosted photons, ECAL needed to have fine granularity. Due to the high luminosity of the LHC, ECAL had to be radiation hard.

In order to meet the three design goals, ECAL is made out of 75,848 scintillating lead tungstate ($PbWO_4$) crystals: 61,200 in the barrel, and 7324 in each of the two endcaps. The radiation length in lead tungstate is short ($X_0 = 0.89\,\text{cm}$) allowing ECAL to be compact (a necessity since it must fit within the solenoid) but still contain $\approx 25X_0$ of scintillator. The Molière radius is also small (2.2 cm) allowing showers to be contained within each crystal. Lead tungstate is also radiation hard (up to 100 kGy). While lead tungstate has a fast response (80 % of light is given off within 25 ns), it does not produce very much light (30 γ/MeV) and so sensitive photodetectors that can operate within a

strong magnetic field are required. In the barrel avalanche photodiodes are used, and in the endcaps vacuum phototriodes are used.

Electromagnetic Calorimeter Barrel

The ECAL barrel detector (EB) has an inner radius of 129 cm and covers a pseudorapidity range from $0 < |\eta| < 1.479$. It is composed of 36 identical "supermodules" each covering half the barrel's length and 1/18th of the barrel's circumference. The crystals used in the EB have a front face cross section of $22 \times 22\,\mathrm{mm}^2$ and a length of 230 mm. They cover 1° in $\Delta\eta$ and $\Delta\phi$. The crystals are tilted with their axis 3° off from the vertex region in order to obscure the small gaps between crystals from particles leaving the interaction point.

Electromagnetic Calorimeter Endcap

The ECAL endcap detectors (EE) are located with their front face at $|z| = 314\,\mathrm{cm}$ and cover a pseudorapidity range of $1.479 < |\eta| < 3.0$. Each endcap is composed of two "Dees" consisting of semi-circular aluminum plates with crystals mounted on them. The crystals used in the EE have a front face cross section of $28.6 \times 28.6\,\mathrm{mm}^2$ and a length of 220 mm. Unlike in EB, where the crystals are arranged in an η–ϕ grid, the crystals in EE are arranged in an x–y grid. Like the EB crystals, the EE crystals also off-point from the vertex region.

Electromagnetic Calorimeter Preshower

In addition to EB and EE, there is a smaller subdetector, the preshower (ES), placed in front of EE. It is made of lead and silicon and has higher granularity then EE. It was designed to help differentiate $\pi^0 \to \gamma\gamma$ decays from $\mathrm{H} \to \gamma\gamma$ decays.

The layout of ECAL is shown in figure 3.5.

3.2.3 Hadronic Calorimeter

The hadron calorimeter (HCAL) in CMS is built around ECAL and inside the solenoid. HCAL's primary purpose is to measure the energy of the various strongly-interacting particles created by the collision. HCAL must have good containment of hadronic

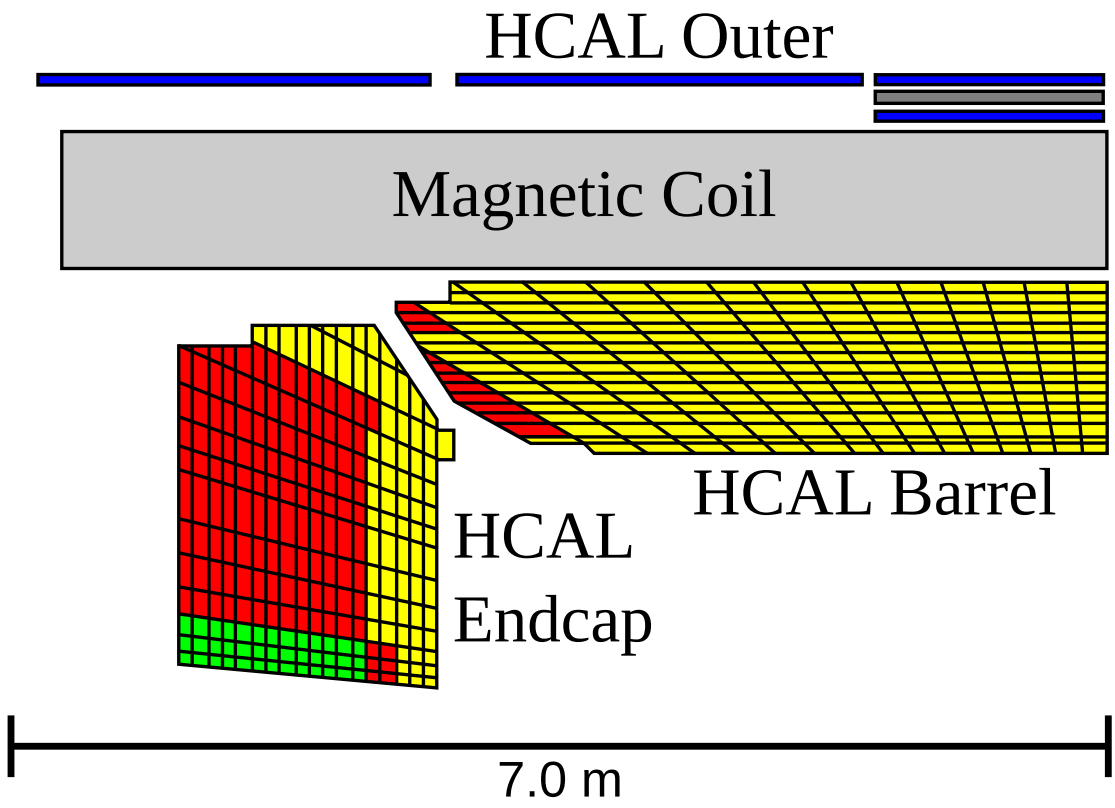


Figure 3.6: A quarter cross-sectional view of the CMS hadronic calorimeter showing the barrel (HB), endcap (HE), and outer (HO) components. Not shown is the forward calorimeter (HF).

particles so that the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, of an event can be accurately measured. This variable is useful for identifying non-interacting particles like neutrinos, dark matter, or other long-lived exotic particles.

HCAL is a sampling calorimeter which is appealing because that allowed it to be made compact (HCAL must fit within the magnet), cheap (ECAL was expensive, so savings elsewhere were necessary), and with good enough resolution, since the resolution is dominated by the hadronic over electromagnetic (H/E) correction. This correction arises from the fact that the detector response from electromagnetic interactions is different than for hadronic interactions, and from the fact that there are statistical fluctuations in the number of neutral pions—which decay to two photons and so interact electromagnetically—and charged pions—which interact hadronically—which make up a hadronic shower in HCAL on an event-by-event basis. These differences are corrected for in aggregate, with the effect of broadening the calorimeter's resolution.

Hadron Barrel, Hadron Endcaps, and Hadron Outer

The hadron barrel (HB) and hadron endcaps (HE) detectors are constructed of alternating layers of absorber and scintillators. The absorber is made of brass because it has a short interaction length, is easy to machine, and is nonmagnetic. In between, the showers are sampled by plastic scintillator tiles read out with embedded wavelength-shifting fibers. The wavelength-shifting fibers are spliced to clear fibers outside the scintillator which carry the signal to the readout system. The light is readout by multi-channel hybrid photodiodes. The hadron outer (HO) detector is a layer of plastic scintillator on the outside of the solenoid before the muon systems begin. It serves as a "tail catcher" by measuring the hadronic energy leaking through HCAL and interacting with the magnet.

Hadron Forward

The hadron forward (HF) calorimeter is the most forward of of the detectors that make up HCAL, and the one of the most forward detectors in CMS. HF is located 11.2 m from the interaction point. It covers a pseudorapidity range of $3 < |\eta| < 5$ and so receives a very high particle flux. HF must be very radiation hard in order to survive in this environment. HF is constructed of steel and quartz. The signal in HF originates from Cherenkov light in the quartz.

The steel absorb is 1.65 m thick. Placed within the absorber are quartz fibers which are 0.6 mm in diameter and oriented parallel to the z-axis. These fibers are arranged in a square grid 5 mm apart. There are two lengths of fibers: "long" 1.65 m ones and "short" 1.43 m ones. These different lengths allow HF to sample showers at two different depths.

The layout of HCAL is shown in figure 3.6.

3.2.4 Magnet

The central feature of CMS, both from a design standpoint and structurally, is the large superconducting solenoid magnet. A high strength magnetic field is required in CMS in order to achieve a momentum resolution of $\Delta p/p \approx 10\,\%$ for muons with $p = 1\,\text{TeV}$. The momentum of these muons must be measured using their curvature in the tracker and the muon chambers. Additionally, the magnetic field allows the charge of particles to be measured by the direction in which their tracks curve. The solenoid provides a magnetic field for all of the components inside the bore of the magnet. The fringing fields are collected and returned by the iron return yokes which also support the muon chambers. In this way a field is provided for the muon chambers without requiring an additional magnet.

The solenoid is 12.9 m long with an inner bore of 5.9 m. It generates a 3.8 T field using 2168 turns of conductor which carry 19.5 kA. The total stored energy is 2.7 GJ. The solenoid is encased in aluminum to help dissipate heat and add strength. This entire assembly is enclosed in a stainless steel cryostat. The cryostat must be exceptionally strong because it not only supports the solenoid, but also all of the barrel subdetectors (HCAL, ECAL, and the tracker) that are mounted within it. The calorimeters and the tracker are mounted within the solenoid in order minimize the uninstrumented mass—where particles could lose energy due to interactions—between the calorimeters and the collision region.

3.2.5 Muon System

The muon system is the outer most subdetector of CMS. Its primary purpose is to assist in measuring high momentum muons and to allow triggering on muons. Without

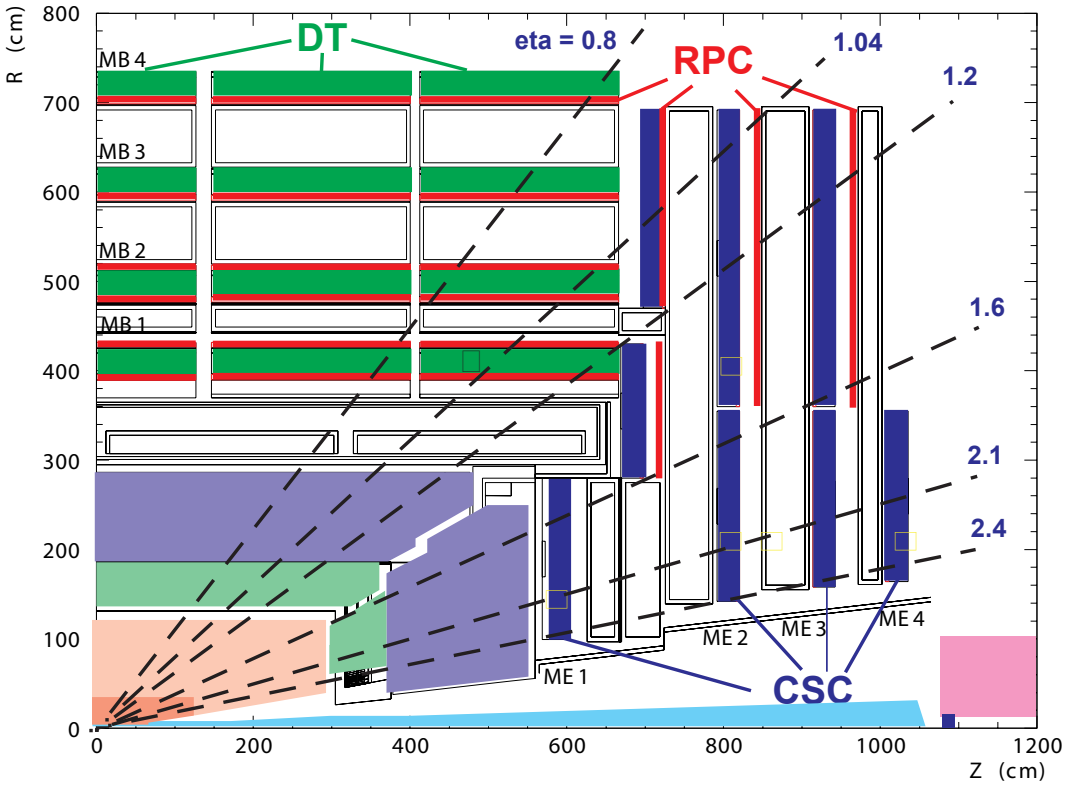


Figure 3.7: A quarter cross-sectional view of the CMS muon system showing the drift tubes (DT), cathode stripe chambers (CSC), and resistive plate chambers (RPC).

the muon system, no triggering on muons would be possible because they are minimum ionizing particles in the calorimeters and the tracker can not be read out quickly enough to use in triggering. For high energy muons ($p_{\mathrm{T}} > 200\,\mathrm{GeV}$), a combination of the tracker and muon system provides the most accurate measurement because of the increased lever arm provided by the muon system. The muon chambers are built within the iron return yokes and as such are outside the primary magnetic field provided inside the solenoid.

The muon systems must cover a very large surface area, $25{,}000\,\mathrm{m}^2$, and so gaseous detectors are used in order to keep the cost down. Three types of gaseous detectors are used in the muon system: drift tube (DT) chambers, resistive plate chambers (RPC), and cathode strip chambers (CSC),

Drift Tubes

DTs are gas ionization detectors with an anode wire strung down the center of each tube which act as the cathode. The wire collects the ionization charged caused by a muon passing through the tube. By using a timing and pulse shape measurement, the DTs can achieve a point resolution of $\approx 200\,\mu\mathrm{m}$ leading to a position resolution in ϕ of $100\,\mu\mathrm{m}$ and a direction measurement good to approximately 1 mrad. The maximum drift length in the DTs is 2 cm.

Resistive Plate Chambers

The RPCs are constructed of two parallel high-voltage plates. When a muon passes through it causes an electron cascade that is collected on the anode, which is divided into strips to allow the position of the cascade to be measured. RPCs have worse position resolution than the DTs, but are much faster because they have a shorter drift length, with a timing response of $\approx 1\,\mathrm{ns}$.

Cathode Strip Chambers

The CSCs are flat gas chambers with parallel anode wires running the length on one side, and cathode wires on the other side running orthogonal to the anodes. A prompt signal is provided by the anode wires, while a weighted average of the image charge on the cathode provides a slower but more precise location. The anode signal is used in

triggering at Level-1 while the cathodes are used for final reconstruction. The spatial resolution of $\approx 200\,\mu m$ ($100\,\mu m$ in the first ring of the first endcap disk) and the angular resolution in ϕ is $\approx 2\,mrad$.

Muon System Barrel

In the barrel region ($|\eta| < 1.2$) DTs are used because the induced neutron background is low, the muon rate is low, and the magnetic field is low. RPCs are also used in the barrel to augment the DTs. The barrel consists of four layers of muon stations in five wheels corresponding to the five wheels making up the iron return yoke. Each wheel has 12 sectors that cover 30°. The chambers are staggered such that a high energy muon produced near the boundary must cross 3 of the 4 layers. The inner three layers have 12 chambers, while the fourth layer has an additional chamber in both the bottom and top sectors, for a total 14 chambers in that layer. The first three layers are aligned to measure in r–ϕ and z, while the 4th layer measures only in r–ϕ. Each chamber consists of 12 planes of DTs with 2 RPCs—1 on the front and 1 on the back—in the first two layers, and 1 RPC—on the front—for the second two layers.

Muon System Endcap

In the endcap region, the higher magnetic field, neutron background, and muon flux make DTs ineffective and so CSCs are used instead. There are three disks of CSC chambers with RPCs sandwiched between disks of the iron return yoke in the rapidity range $|\eta| < 1.6$. After that range there are only CSCs chambers. There are 6 CSCs per chamber.

The layout of the muon system is shown in figure 3.7.

3.2.6 The Trigger

The trigger's job is to select the interesting physics events from all of the collisions happening in CMS. The trigger must be able to handle the full bunch crossing rate at the LHC of 40 MHz (although it only ran at 20 MHz in 2012). The trigger accomplishes this by using a two level design. The first level, the Level-1 (L1) trigger, consists of custom hardware cuts the 20 MHz of collisions down to about 100 kHz. These events

are then passed to the second level of the trigger, the high-level trigger (HLT), which uses software running on commodity hardware to select 1 kHz of events.

The Level-1 Trigger

The L1 trigger is implemented in custom hardware because it must be incredibly fast. Each event is given just 3.2 µs to travel from the detector to the service cavern and be accepted or rejected by the L1 trigger. Of this time, less than 1 µs is given for the L1 trigger to make its decision. The L1 trigger only uses information from the muon system and the calorimeters; it does not use information from the tracker as it takes too long for this data to be assembled and transferred. From the calorimeters the L1 trigger gets the sum of energy in a fixed array of 5x5 crystals in ECAL as well as the energy behind each array in HCAL. From the muon chambers the L1 trigger gets a short vector from each detector indication the location of a hit in the chamber and the direction the particle was traveling in. The L1 trigger also makes use of a few global sums including E_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$.

While a decision is being made, the data from the collision are stored in a hardware buffer. If the L1 trigger selects the event, the data are moved from the buffer and sent to the central DAQ system where it is held until all the data are assembled, at which point it is sent to one of the computers that makes up HLT.

The High-Level Trigger

The HLT is implemented using the same software that is used for offline analyses. Running the same analysis software as is used offline allows HLT to access the full range of data from every part of the detector and use any reconstruction or selection algorithm available offline. This allows very sophisticated triggers to be written, and additionally allows triggers to be easily modified as conditions change. The HLT is designed so that the simplest triggers are run first, with more sophisticated—and hence computationally expensive—triggers running later if needed.

HLT runs on a server farm consisting of commodity hardware. There are currently around 10,000 processor cores in the HLT farm. By using commodity server hardware, HLT benefits from the rapid speed and power advances that are being made in commercial computing.

Chapter 4

Event Reconstruction

4.1 Electron Reconstruction

Reconstruction of electrons in CMS is complicated by the fact that the tracker contains a large amount of material and so electrons must pass through up to two radiation lengths of material before reaching ECAL. Information about the material in front of ECAL is given in figure 4.1[69]. Electrons emit photons when they interact with matter in a process known as bremsstrahlung. These photons are then separated from the electron as they do not bend in the magnetic field while the electron does. In order to accurately reconstruct the energy of the electron, these photons must be accounted for in the final energy sum, and this is what the electron reconstruction algorithm attempts to do.

The reconstruction of electrons with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ in CMS begins with a "seed" cluster of energy deposits in ECAL [70]. As an electron bends in the magnetic field, it takes a helical path at constant η but changing ϕ. Any photons radiated by the electron will lie along this path. These photons must be added together with the electron in order to accurately measure its energy, and so additional clusters on the same path as the seed cluster are connected to form a "supercluster" [71]. A supercluster, therefore, includes the energy deposit from the electron as well as the energy from its nearby photons at constant η.

From these superclusters, a volume in the tracker where the electron is likely to have come from is determined by propagating the energy-weighted mean position of the

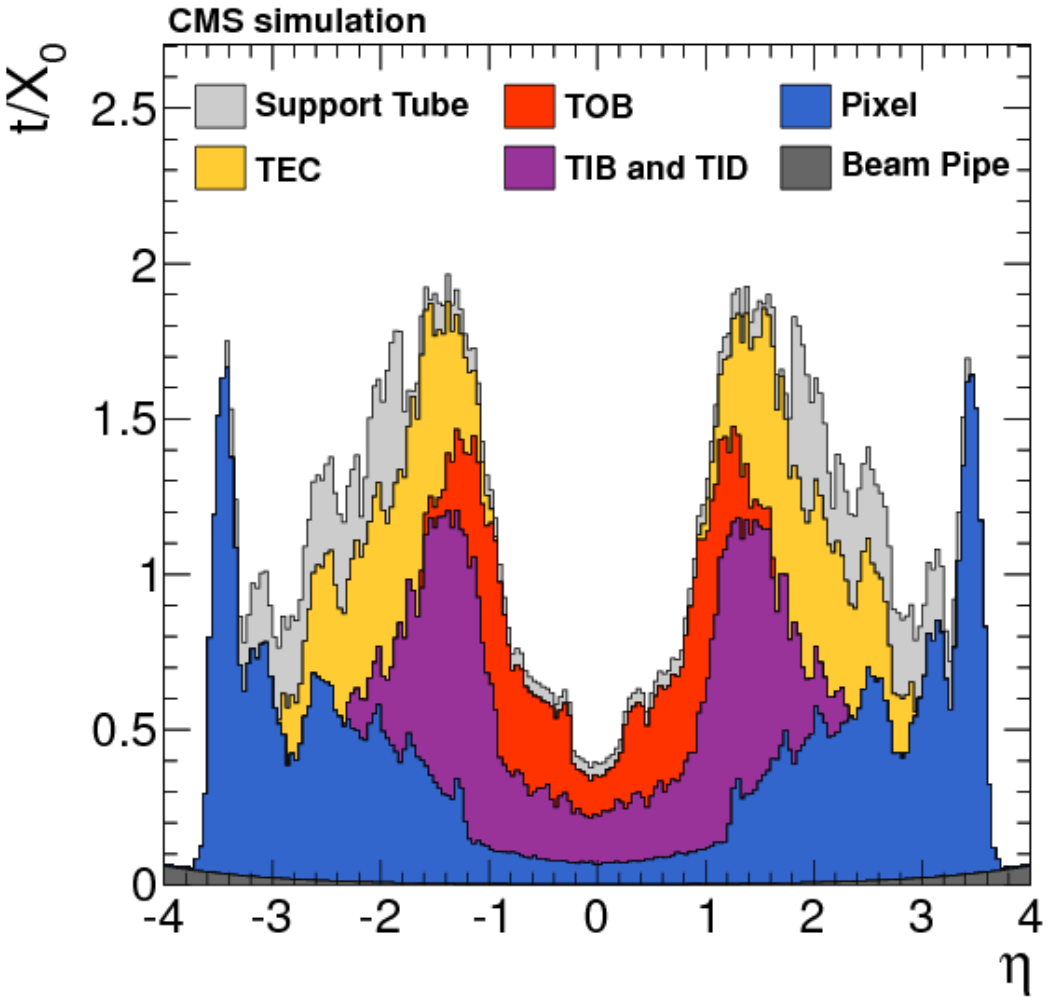


Figure 4.1: The thickness of material, t, divided by the radiation length, X_0, encountered by particles leaving the nominal interaction point before reaching ECAL.

supercluster back through the magnetic field. This spot is then used to seed a track-finding algorithm in the pixel layer. Hits in the tracker are searched for starting at the innermost layer and working outward. In order to account for the changing shape of the track as the electron loses energy from interacting with the tracker material, a "Gaussian Sum Filter" (GSF) [72] is used rather than the simpler Kalman filter used for muons and hadrons. Low energy electrons ($E_{\mathrm{T}} < 15\,\mathrm{GeV}$) are constructed with p_{T} from the tracker and E_{T} from ECAL. For higher energy electrons, the ECAL energy is used without the tracker p_{T} to avoid issues introduced by the possible poor fits in the tracker. The η and ϕ of all electron candidates, regardless of energy, is taken from the track by projecting back to the interaction point.

4.2 Additional Corrections

Although the measurement of ϕ^* is relatively insensitive to the energy of the electrons, the energy still plays a role in determining the electron p_{T} and hence whether an event passes the selection criteria used in this analysis (discussed in section 6.2.1). Therefore, it is important to accurately measure this quantity, even if it does not directly change the final observable. To this end, two sets of energy and momentum corrections, centrally produced by the CMS collaboration, are applied to both the data and the reconstructed quantities in the simulated data. A summary of the method used to derive these corrections follows because the papers detailing them are not public.

4.2.1 Regression

The first set of corrections were calculated using a multivariate regression trained on $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ and $\mathrm{H} \to \mathrm{ZZ}$ MC [73]. The regression used a boosted decision tree trained on 41 different variables parameterizing electron shower shape, the electron track, and the location of the shower in EB. The algorithm was trained separately for EB and EE electrons. In order to prevent over-training the MC samples were split in half, with one half used for training and the other half used for validation. Electrons used in the regression were required to have low radiated energy fraction (< 0.01) as determined by the generator level MC and $p_{\mathrm{T}} > 7\,\mathrm{GeV}$. The target variable was the ratio of the generator level bare electron energy over the reconstructed energy. This correction was

applied to both the data and the reconstructed MC electrons used in this analysis.

4.2.2 Energy Scale and Resolution

The second set of corrections were calculated with $Z \to e^+e^-$ MC using two independent methods [74]. The first method was used to correct for the energy scale while the second method was used to correct for the resolution.

In the first method, the MC sample and the data were fit with the convolution of a Breit–Wigner with a Crystal Ball (CB) function. The CB function is used to model the resolution of the detector and losses due to bremsstrahlung from the material in front of ECAL. It consists of a Gaussian with a power-law low-side tail. It was first used by the Crystal Ball Collaboration [75]. The Breit–Wigner function models the analytic shape predicted for the Z mass resonance. The parameters of the Breit–Wigner were fixed to the nominal values from the Particle Data Group: $M_Z = 91.1876 \pm 0.0021\,\text{GeV}$, $\Gamma_Z = 2.4952 \pm 0.0023\,\text{GeV}$. The parameters of the CB function were free parameters in the fit.

The scale correction is both time dependent and η dependent and so fits were performed separately in four pseudorapidity bins in various run ranges. The peak of the CB function in data and MC were compared with the relative shift taken as the scale correction, ΔP, given by equation (4.1).

$$\Delta P = \frac{\Delta m_{\text{data}} - \Delta m_{\text{MC}}}{M_Z} \tag{4.1}$$

In the second method, two categories of electrons are defined: showering electrons ($R_9 < 0.94$) and non-showering electrons ($R_9 > 0.94$). R_9 is the ratio of energy in the 3x3 square of crystals in ECAL where the electron impacted over the energy in the supercluster and so a larger number means a narrower shower. A probability density function (PDF) is created using the $Z \to e^+e^-$ MC. For each event in the PDF, the supercluster energy is modified by applying a Gaussian multiplicative factor $1 + \Delta P$ with a standard deviation of $\Delta\sigma$, where ΔP is the scale correction and $\Delta\sigma$ models the resolution. The resolution parameter was selected for each of the two types of electrons, the four pseudorapidity bins, and various run ranges using a likelihood maximization. In the EB pseudorapidity bins, there were enough events to bin in E_T and so these

corrections are also E_{T} dependent.

4.3 Cut Based Identification

There are several centrally-defined "cut based identification" requirements used to select high quality electrons at CMS. These requirements make use of the variables defined in section 4.4 in order to reject low quality and unisolation electrons. We use two of requirements, referred to as `Medium` and `Tight`, with `Tight` having stricter requirements than `Medium`. The exact definition of these requirements are given in table 4.1.

Variable	`Tight`		`Medium`	
	EB	EE	EB	EE
$\Delta\eta_{\mathrm{in}} <$	0.004	0.005	0.004	0.007
$\Delta\phi_{\mathrm{in}} <$	0.03	0.02	0.06	0.03
$\sigma_{i\eta i\eta} <$	0.01	0.03	0.01	0.03
H/E $<$	0.12	0.10	0.12	0.10
d0 $<$	0.02	0.02	0.02	0.02
dz $<$	0.1	0.1	0.1	0.1
$\mid (1/E - 1/p) \mid <$	0.05	0.05	0.05	0.05
$\mathrm{P}_{\mathrm{vtx}} <$	10^{-6}	10^{-6}	10^{-6}	10^{-6}
$\mathrm{N}_{\mathrm{miss}} \leq$	0	0	1	1
$\mathrm{Iso}_{\mathrm{PF}} <$	0.10	0.10	0.15	0.15

Table 4.1: Identification and isolation requirements for `Tight` and `Medium` requirements in the ECAL barrel (EB) and ECAL endcap (EE).

4.4 Electron Variables

Not everything reconstructed as an electron in CMS is a real electron. One source of fake electrons is a process referred to as charge exchange: $\pi^+ + \mathrm{n} \rightarrow \pi^0 + \mathrm{p} \rightarrow 2\gamma + \mathrm{p}$. In this process, a charged pion interacts with a nuclear neutron which convert to a neutral pion and a proton where the neutral pion quickly decays to two photons. If this

process happens in ECAL, it can result in a track (from the charged pion) pointing at a supercluster with a shape similar to a photon or electron shower (from the two photons) which will often be reconstructed as an electron. Another source of fake electrons is coincidence events where a photon impacts ECAL near the track left by another charged particle.

Not every electron is equally likely to be from a $Z \to e^+e^-$ decay; many electrons detected in CMS come from other processes. One such process is photon conversion, $\gamma \to e^+e^-$, where a photon in the presence of a nucleus converts to two electrons. Another source of electrons is from QCD jets which can sometimes produce an electron in their numerous decays.

There are several electron variables defined in order to allow the separation of these lower quality and fake electrons from the electrons from $Z \to e^+e^-$ decays. These variables are broken down into three categories:

Identification (ID):

These variables quantify how much like an electron the reconstructed particle looks and are used to reject fake electrons like those from charge exchange.

Conversion Rejection:

These variables are used to reject electrons from $\gamma \to e^+e^-$ conversions.

Isolation:

These variables measures how much energy from other particles is deposited near the electron in the detector and are used to help reject electrons within jets.

The following subsections contain plots of some of the most important electron variables as well as a discussion of their use in selecting electrons from $Z \to e^+e^-$ decays. The data shown on the plots consists of every electron candidate in the "SingleMuon" dataset with $|\eta| < 2.4$ and $p_T > 20\,\text{GeV}$. This dataset was used because the trigger that selects events for it makes no selection on the electron variables and therefore the electron candidates in the set represent a sample of minimally-biased electrons. The MC electrons are from the MadGraph DY $\to \ell^+\ell^-$ sample with the same kinematic requirements are applied as are used on the data, with the additional requirement that the event must contain a generator level $Z \to e^+e^-$ decay. These selections give us a

data sample that is indicative of the general distribution of all electrons in CMS events, while the MC sample shows what the distribution looks like in signal events.

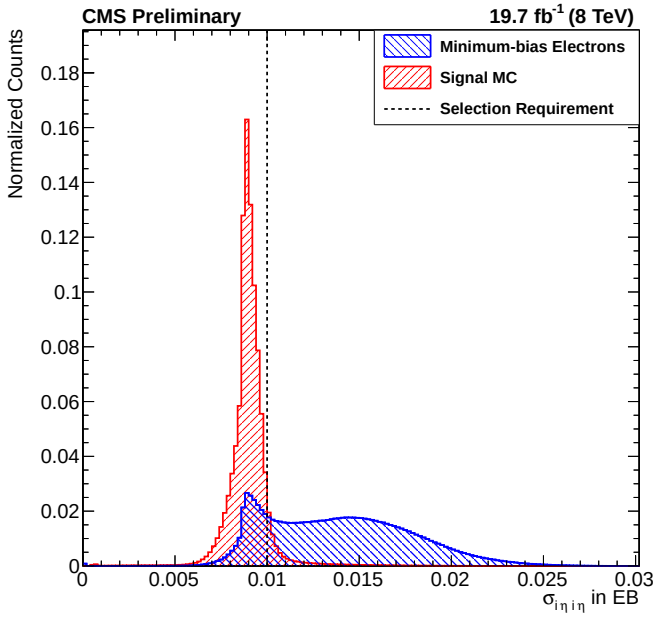
4.4.1 Identification

The shape of the electromagnetic shower in the calorimeters is used to discriminate between electrons and other particles. Electrons generally have very narrow showers whereas hadronic particles have wide showers. The size of the shower in η is characterized by $\sigma_{i\eta i\eta}$. Charge exchange events often have a higher $\sigma_{i\eta i\eta}$ than real electrons because the proton is can be knocked loose by the interaction and leaves a broader shower. Comparisons of the $\sigma_{i\eta i\eta}$ distributions between all electrons and for signal electrons for both EB and EE are shown in figure 4.2.

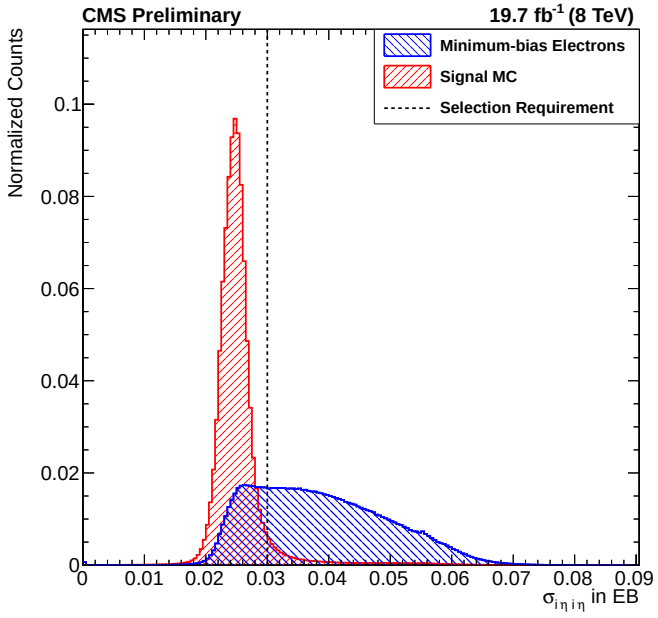
Electron showers are mostly contained within ECAL and so the ratio of energy around the hit in HCAL over the energy around the hit in ECAL, H/E, is also used to parameterize the shower shape. Charge exchange events often have higher H/E than electrons because the loose proton can escape and enter HCAL. A comparison of the H/E distributions between all electrons and for signal electrons is shown in figure 4.3.

The distance between the track and the supercluster in η–ϕ space is given by $\Delta\phi_{\text{in}}$ and $\Delta\eta_{\text{in}}$. Real electrons have small values of $\Delta\phi_{\text{in}}$ and $\Delta\eta_{\text{in}}$ because the electron caused both the track and the supercluster, but coincidence events are equally as likely to have a large track separation as a small one because each element comes from an independent object. Comparisons of the $\Delta\phi_{\text{in}}$ and $\Delta\eta_{\text{in}}$ distributions between all electrons and for signal electrons are shown in figures 4.4a and 4.4b, respectively.

The compatibility of energy of the supercluster and the momentum of the track is parameterized by $(1/E - 1/p)$. Electrons will have $(1/E - 1/p)$ near 0 because their momentum and energy, having been measured for the same object, will agree. Charge exchange events will have positive $(1/E - 1/p)$ as their momentum is measured in the tracker before the interaction (and hence measures the full momentum of the π^+) but their energy is measured in ECAL after the interaction and hence after the proton has carried away some of the energy. Coincidence events, because the energy and momentum are measured from independent objects, can take both positive and negative values of $(1/E - 1/p)$. A comparison of the $(1/E - 1/p)$ distributions between all electrons and for signal electrons is shown in figure 4.5.



(a)



(b)

Figure 4.2: The $\sigma_{i\eta i\eta}$ variable distribution in EB (top) and EE (bottom) for all electrons with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.4$ in a set of events selected with a muon trigger and in MadGraph $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ MC.

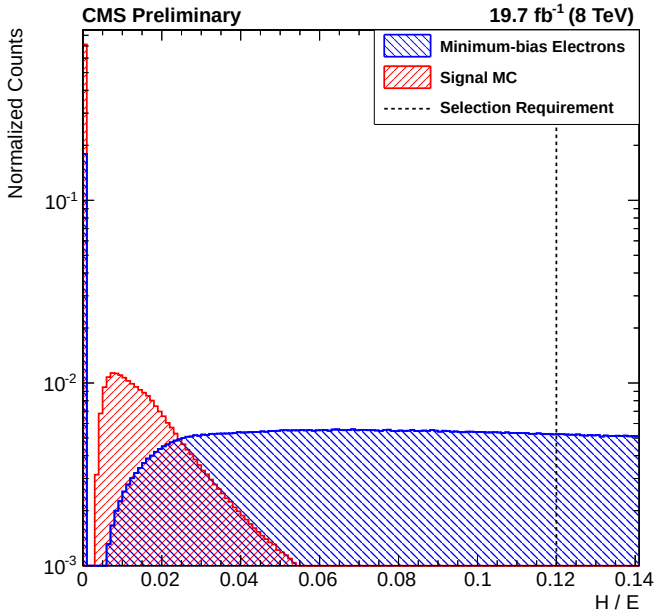
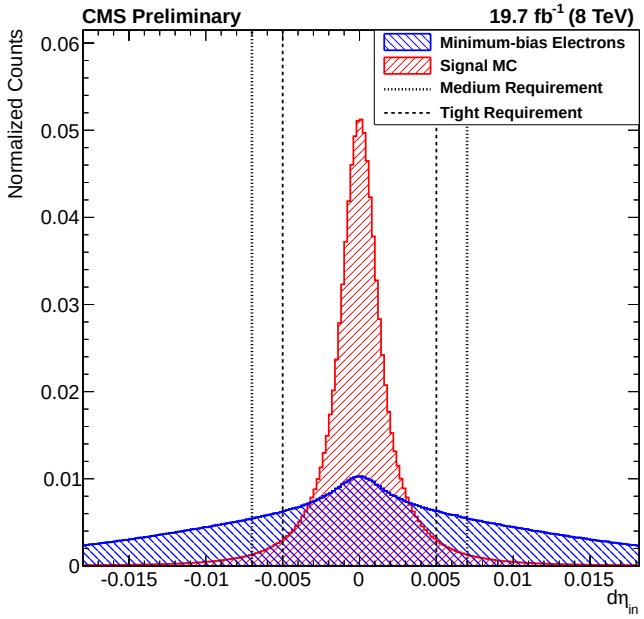


Figure 4.3: The H/E distributions for all electrons with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.4$ in a set of events selected with a muon trigger and in MadGraph $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ MC.

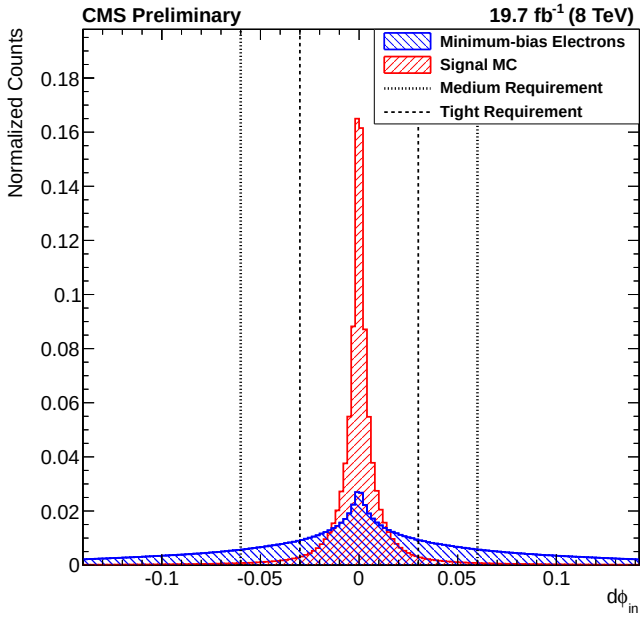
4.4.2 Conversion Rejection

Photon conversions generally happen away from the primary vertex and so the distances of the hits in the track from the vertex are useful quantities to reject conversions. The transverse and longitudinal separation between the track and the primary vertex are given by d0 and dz. In addition to the raw distance of the track from the primary vertex, there is also a fit probability, $\mathrm{P_{vtx}}$, which indicates the probability that a track came from the primary vertex. This variable is also used to reject conversions. Comparisons of the d0 and dz distributions between all electrons and for signal electrons are shown in figures 4.6a and 4.6b, respectively.

Conversions generally happen after the photon passes through several layers of the tracker and so their tracks will have missing layers, the number of which is given by $\mathrm{N_{miss}}$. A comparison of the $\mathrm{N_{miss}}$ distributions between all electrons and for signal electrons is shown in figure 4.7.



(a)



(b)

Figure 4.4: The $\Delta\eta_{\text{in}}$ (top) and $\Delta\phi_{\text{in}}$ (bottom) variable distributions for all electrons with $p_{\text{T}} > 20\,\text{GeV}$ and $|\eta| < 2.4$ in a set of events selected with a muon trigger and in MadGraph $\text{Z} \to \text{e}^{+}\text{e}^{-}$ MC.

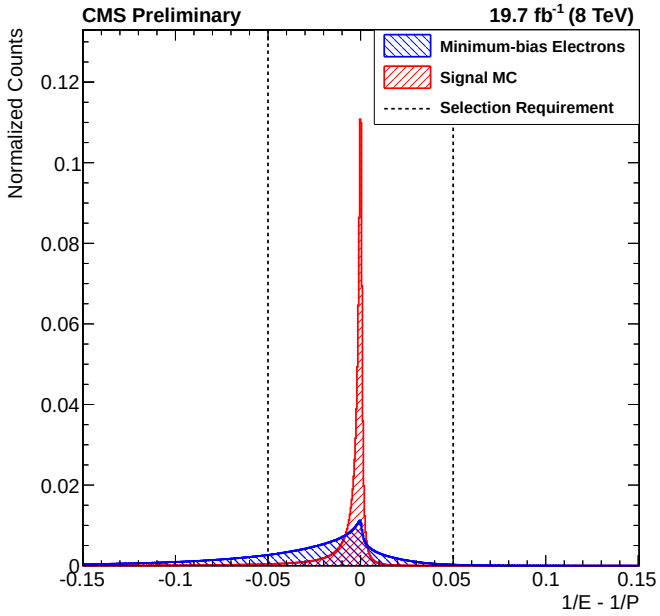


Figure 4.5: The $(1/E - 1/p)$ distributions for all electrons with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.4$ in a set of events selected with a muon trigger and in MadGraph $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ MC.

4.4.3 Isolation

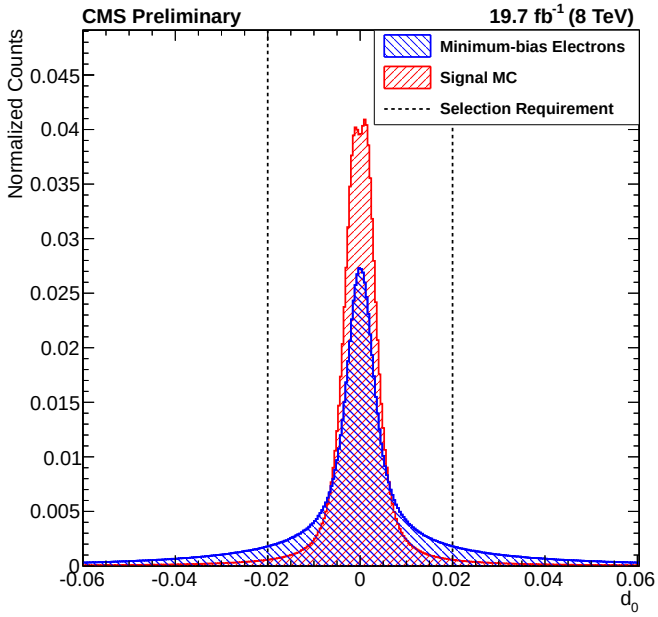
Hadronic jets sometimes produce electrons in the numerous decays happening within them. These electrons can be rejected by looking at the sum of the energy in the tracker, ECAL, and HCAL around the electron, as electrons in jets will have a large amount of energy surrounding them, whereas electrons from Z decays will tend to be isolated.

The isolations used in the trigger (`HLT_Ele27_WP80`) are defined as follows:

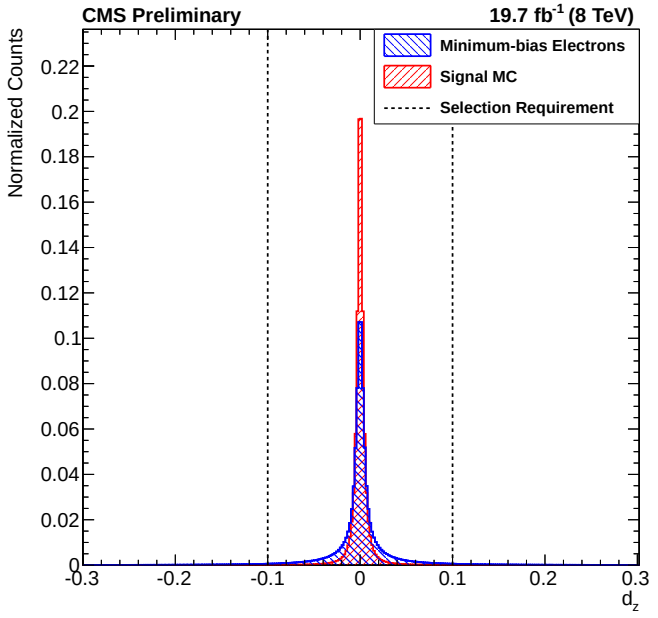
$$\mathrm{Iso}_{\mathrm{HCAL}} = \frac{\sum_{\Delta \mathrm{R}<0.3} E^{\mathrm{HCAL}}}{E_{\mathrm{T}}^{\mathrm{Electron}}} \tag{4.2}$$

$$\mathrm{Iso}_{\mathrm{ECAL}} = \frac{\sum_{\Delta \mathrm{R}<0.3} E^{\mathrm{ECAL}} - E^{\mathrm{SC}}}{E_{\mathrm{T}}^{\mathrm{Electron}}} \tag{4.3}$$

where $\sum_{\Delta \mathrm{R}<0.3}$ is a sum on the energy in a $\Delta R < 0.3$ cone around the supercluster location, E^{HCAL} is the energy in HCAL, E^{ECAL} is the energy in ECAL, and E^{SC} is the energy of the supercluster, which is subtracted out of the ECAL isolation sum. No subtraction is applied to the HCAL isolation as we expect all of the electron's energy to



(a)



(b)

Figure 4.6: The d0 (top) and dz (bottom) variable distributions for all electrons with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.4$ in a set of events selected with a muon trigger and in MadGraph $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ MC.

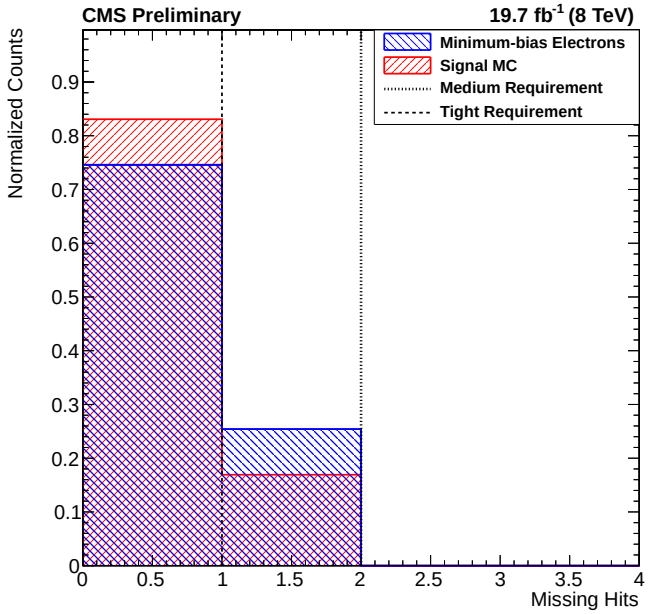
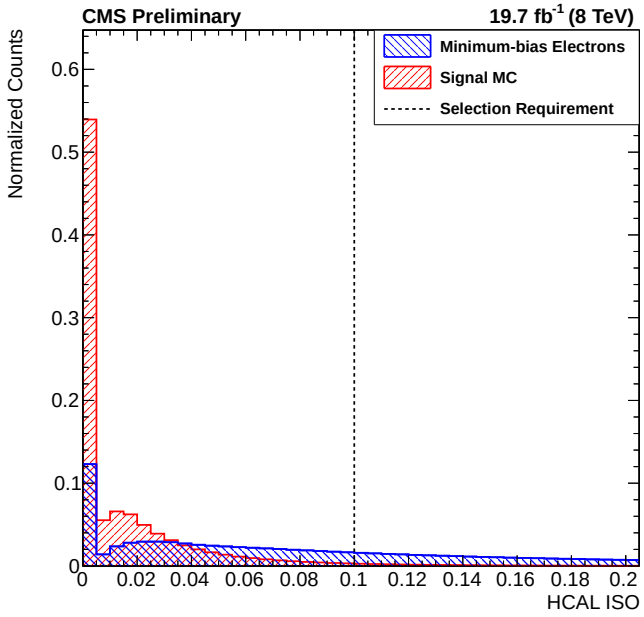


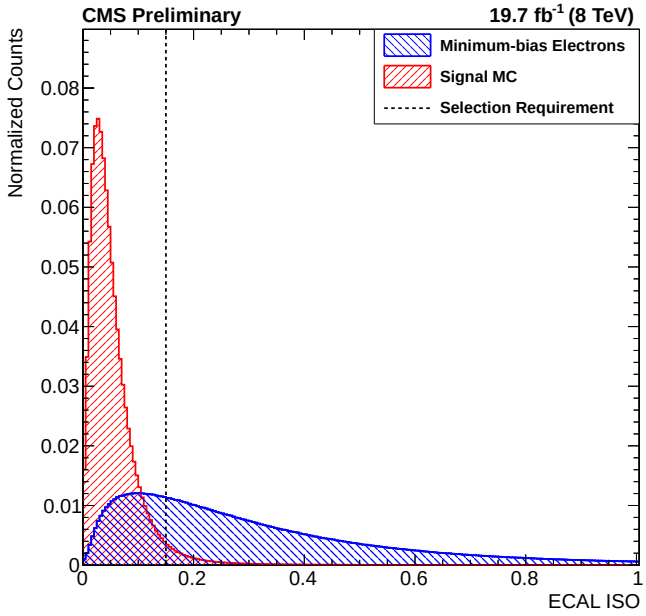
Figure 4.7: The N_{miss} distribution for all electrons with $p_T > 20\,GeV$ and $|\eta| < 2.4$ in a set of events selected with a muon trigger and in MadGraph $Z \to e^+e^-$ MC.

be contained in ECAL. Both isolation values are normalized by the E_T of the electron. Comparisons of the Iso_{HCAL} and Iso_{ECAL} distributions between all electrons and for signal electrons are shown in figures 4.8a and 4.8b, respectively.

A different isolation variable is defined for selection of events in this analysis that is more expensive to compute but takes advantage of the tracker as well as ECAL and HCAL. This isolation uses a "particle flow" [76, 77] technique which is a method of reconstructing jets that uses information from every subdetector and tries to reconstruct the individual particles in a jet by matching them to their responses in the various subdetectors. To keep the algorithm simple, particle flow categorizes every particle into one of five types: photons, electrons, muons, charged hadrons, and neutral hadrons. A photon is a particle with energy only deposited in ECAL. An electron is a particle with an ECAL energy deposit and a track. A muon is a track in the central tracker matched to a track in the muon system. A charged hadron is any energy cluster in HCAL with a possible matching ECAL cluster and track. A neutral hadron is any energy cluster in HCAL with a possible matching ECAL cluster without a matching track. These particle



(a)



(b)

Figure 4.8: The HCAL (top) and ECAL (bottom) isolation variable distributions for all electrons with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.4$ in a set of events selected with a muon trigger and in MadGraph $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ MC.

flow jets are used to calculate an energy density due to pileup, ρ, in the detector which is used to remove the pileup contribution from the isolation sum. The particle flow isolation, $\mathrm{Iso_{PF}}$, is given by:

$$\mathrm{Iso_{PF}} = \sum_{\Delta \mathrm{R}<0.3} \frac{\left(p_{\mathrm{T}}^{\mathrm{Track}} + E^{\mathrm{ECAL}} + E^{\mathrm{HCAL}}\right) - p_{\mathrm{T}}^{\mathrm{Electron}} - E^{\mathrm{SC}} - 0.3^2\pi\rho}{p_{\mathrm{T}}^{\mathrm{Electron}}} \tag{4.4}$$

where the variables are the same as above, with the addition of $p_{\mathrm{T}}^{\mathrm{Track}}$, which is the p_{T} of all tracks in the tracker, $p_{\mathrm{T}}^{\mathrm{Electron}}$, which is the p_{T} of the electron's track, and $\left(0.3^2\pi\rho\right)$, which is the energy around the electron due to pileup calculated from particle flow. The particle flow isolation is normalized by the p_{T} of the electron. A comparison of the $\mathrm{Iso_{PF}}$ distributions between all electrons and for signal electrons is shown in figure 4.9.

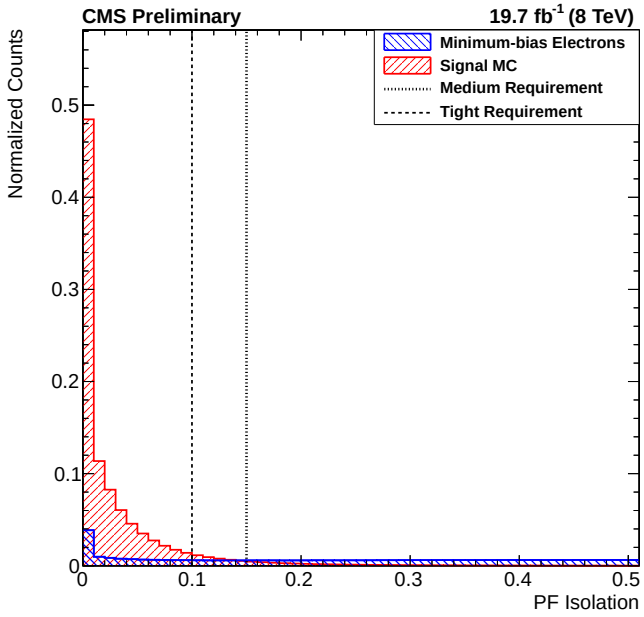


Figure 4.9: The particle flow isolation variable distribution for all electrons with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.4$ in a set of events selected with a muon trigger and in MadGraph $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ MC.

Chapter 5

Data and Simulation Samples

5.1 Data

The data used in this analysis were collected by the CMS detector in 2012 at a center-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$. The LHC delivered $23\,\text{fb}^{-1}$ of integrated luminosity during the year as seen in figure 5.1. This period was divided into four run eras referred to as 2012A, B, C, and D. During an era, the LHC run parameters are kept roughly static to allow for consistent data taking conditions. In between eras, maintenance and beam adjustments were performed on the LHC in order to deliver higher luminosity.

The data collected by CMS are split into smaller datasets based on the physics objects contained within the events. This allows analyses to use only one or two datasets, instead of requiring them to deal with the entirety of the CMS data (which is many petabytes, and hence too large for most institutes to store locally). The HLT sorts events into the various datasets based on the triggers that the event fired. In this manner, an event can end up in multiple datasets if it fired multiple triggers. This analysis uses the "SingleElectron" dataset which was collected with the HLT trigger `HLT_Ele27_WP80`. These datasets were reconstructed—converted from raw detector response into physics objects—in January 2013, in order to make use of the most recent calibrations derived from the entire 2012 run. A summary of the datasets used is provided in table 5.1.

Although there is a "DoubleElectron" dataset which uses a trigger designed to find Z bosons, this analysis uses the "SingleElectron" dataset selected with the `HLT_Ele27_WP80`

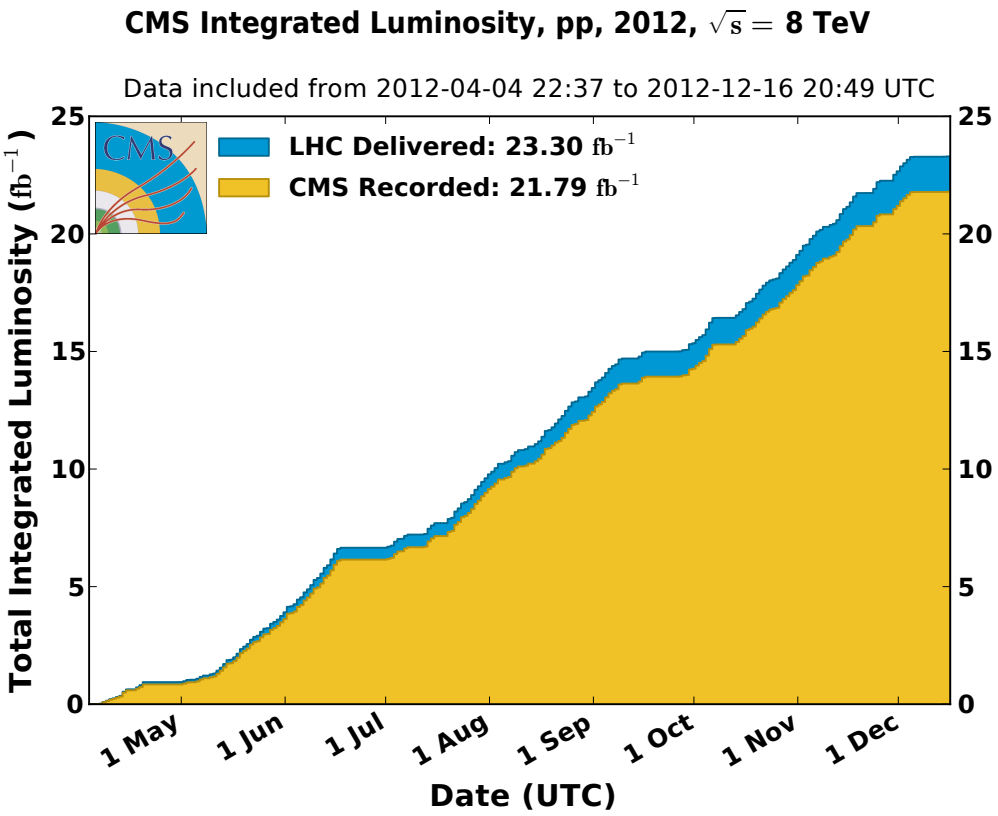


Figure 5.1: The integrate luminosity delivered and recorded by CMS in 2012. The flat periods in May, July, and September correspond to the boundaries between the run eras.

Dataset Name	Run Dates	Luminosity
/SingleElectron/Run2012A-22Jan2013-v1	2012-04-05 to 2012-05-08	$876.225\,\mathrm{pb}^{-1}$
/SingleElectron/Run2012B-22Jan2013-v1	2012-05-10 to 2012-06-18	$4.412\,\mathrm{fb}^{-1}$
/SingleElectron/Run2012C-22Jan2013-v1	2012-07-01 to 2012-09-27	$7.055\,\mathrm{fb}^{-1}$
/SingleElectron/Run2012D-22Jan2013-v1	2012-09-28 to 2012-12-06	$7.369\,\mathrm{fb}^{-1}$

Table 5.1: The name, run dates, and integrated luminosity of the datasets used in this analysis

trigger. The primary motivation behind using this trigger was to allow a direct comparison with a similar ϕ^* analysis being performed by CMS which used $\mathrm{Z} \to \mu^+\mu^-$ events selected with a single muon trigger. The single electron trigger requires an electron with $p_\mathrm{T} > 27$ which passes Working Point 80 (`WP80`). `WP80` is a set of selection requirements on lepton isolation and shower shape variables designed to be 80 % efficient in selecting real electrons. The requirements that make up `WP80` are listed in table 5.2. This trigger had the lowest p_T threshold of any single electron trigger that was unprescaled run during 2012. To prescale a trigger means to apply a rate reduction by randomly throwing out a certain fraction of events in order to keep the total trigger rate manageable; as this trigger was unprescaled, no events were discarded in this manner.

The events from the "SingleElectron" sample are further filtered for quality. A centrally-produced list of good luminosity segments is used to select only events in which no part of the detector was malfunctioning or disabled. After accounting for detector dead time and beam quality, $19.7\,\mathrm{fb}^{-1}$ of integrated luminosity are used for physics analysis.

5.2 Monte Carlo

This analysis makes use of numerous simulated data samples in order to estimate backgrounds and signal yields, derive scale factors, and correct for the effects of bin migration on the final measurement. This simulated samples are referred to as Monte Carlo, or MC.

Value	EB	EE
$\lvert\eta\rvert <$	1.4791	2.65
$p_{\mathrm{T}} >$	27	27
$\sigma_{i\eta i\eta} <$	0.1	0.03
H/E $<$	0.1	0.05
$\mathrm{Iso}_{\mathrm{ECAL}} <$	0.15	0.1
$\mathrm{Iso}_{\mathrm{HCAL}} <$	0.1	0.1
Pixel Matching $\geq$	1	1
$\lvert\,(1/E - 1/p)\,\rvert <$	0.05	0.05
$\lvert\Delta\eta_{\mathrm{in}}\rvert <$	0.007	0.007
$\lvert\Delta\phi_{\mathrm{in}}\rvert <$	0.06	0.03

Table 5.2: The selection requirements for the `HLT_Ele27_WP80` trigger for electrons which end up in the barrel region or the endcap region of ECAL. The variables used are detailed in section 4.4.

5.2.1 Monte Carlo Generators

There are multiple MC generators, each of which makes its own assumptions about the behavior of the various particles and interactions it simulates. In general, the process of generating a simulated proton-proton collision event is broken up into multiple steps. These steps can all be performed by one program, but in general are handled by several. These steps are summarized below:

Parton distributions: In order to simulate the collision of two protons, a set of PDFs is needed to calculate the probability of finding a parton with a specific x_i. PDFs are provided by various theory groups who produce them by fitting their parton models to collider data. MC software can generally run with multiple different PDF sets. PDFs are discussed in detail in section 2.2.2.

Initial state shower: A shower-initiating parton from each beam of protons initiates a sequence of parton branchings (for example, q $\rightarrow$ qg). One parton from each of these showers will go on to interact in the hard scattering process, the others will become initial state radiation. These shower initiators are given non-zero p_{T}

to account for the primordial transverse momentum of the partons in the proton. The amount of p_{T} given to the shower initiator is one of the parameters that can be set in the MC software; it is often set to be larger than would be expected physically (about the same order as the proton mass) in order to account for other hard-to-model processes.

Hard scattering process: In the hard scattering process, the two partons collide and a set of outgoing particles is generated based on the matrix element level calculations.

Resonances and decays: Some of the outgoing particles may be resonances (like the W and Z) or otherwise unstable particles which decay, while others may be stable (but which may still radiate, for example, $e \to e\gamma$). In either case, these particles and their decay chains are simulated in detail until all the remaining particles are stable or long-lived enough to escape the simulation volume.

Multiple interactions and the underlying event: In addition to the hard scattering, there are further interactions from other partons in the shower. After the partons from these interactions are removed from the protons, what is left are "beam remnants" that may have internal color and charge. Additionally, these remnants can have non-zero p_{T} as the extracted partons have non-zero p_{T}. The remnants are still color connected to the hard interaction and must decay accordingly.

Hadronization: Bare quarks and gluons can not be observed so they must be combined into colorless objects in a process known as hadronization. The exact manner in which this happens is not understood from QCD, and so the different MC software suites use different internally developed models to do so.

In the MC samples that we use, the parton distributions, initial state shower, hard scattering, and resonances and decays are handled by MadGraph v5 [78], powheg [79, 80, 81], and Pythia6 v6.4.24 [82], as indicated in table 5.3. All of the samples use Pythia6 for the underlying event and hadronization, except for two locally produced samples which use Pythia8.

MC programs have multiple different parameters that control their behavior in each of these steps. As not all of these parameters can be derived from QCD, many of them must be "tuned" to best fit the data. Parameters that control hadronization and final-state showers are fit using data from LEP, which offers the cleanest environment as the incoming electrons have no color charge. Parameters that control multiparton interactions and initial state radiation are fit using earlier LHC data. In the case of this analysis, only the parameters used in PYTHIA6 and PYTHIA8 are tuned.

5.2.2 Centrally Produced Monte Carlo Datasets

All of the MC samples used in this analysis were centrally-generated by the CMS collaboration except for two POWHEG DY $\rightarrow e^+e^-$ samples which were generated locally. A DY $\rightarrow \ell^+\ell^-$ signal sample and a $t\bar{t}$+jets background sample were generated with MADGRAPH. Diboson (ZZ, WZ, WW) background samples were generated with PYTHIA6. Background samples consisting of $\bar{t}W$, tW, and DY $\rightarrow \tau^+\tau^-$ were generated using POWHEG. A secondary signal MC was also generated with POWHEG. The details of these samples are listed in table 5.3.

All of these MC samples use PYTHIA6 with the `Z2star` tune for modeling the underlying event and hadronization except for the diboson samples. The `Z2star` tune was centrally-produced by the CMS collaboration by tuning PYTHIA6 to match the 2011 CMS data. TAUOLA is used for tau decays in all of the samples except for $t\bar{t}$ and POWHEG DY $\rightarrow e^+e^-$ MC samples [83].

After the generation step, MC is sent through a full detector simulation which uses GEANT4 [84] to mimic the detector response. This detector response is reconstructed using the full CMS reconstruction chain to produce MC files in a format identical to actual data.

MC events have additional simulated minimum-bias events overlaid on top of them to better match the conditions found in actual running. These minimum-bias events attempt to simulate what a typical proton-proton interaction looks like in the detector and are used to mimic pileup. The number of pileup events to be added is drawn from a distribution decided upon before data is taken and so it events must be reweighted after the data taking period is over to make the number of pileup in MC match the same

Process	Requirements	Generator	σ (pb)	Events ($\times 10^6$)
DY $\to \ell^+\ell^-$	$m_{\ell\ell} > 50\,\text{GeV}$	MadGraph	3531.9 (NNLO)	30.460
DY $\to e^+e^-$	$m_{\ell\ell} > 20\,\text{GeV}$	powheg	1966.7	3.297
DY $\to \tau^+\tau^-$	$m_{\ell\ell} > 20\,\text{GeV}$	powheg	1966.7	3.297
$t\bar{t}$		MadGraph	23.64	3.984
$t \to Wb$		powheg	11.1	0.498
$\bar{t} \to W\bar{b}$		powheg	11.1	0.493
WW		Pythia6	54.84	10.000
WZ		Pythia6	33.21	10.000
ZZ		Pythia6	17.7	9.800

Table 5.3: Summary of the MC samples used in this analysis. The Drell-Yan MC samples have mass requirements on the events while the other MC samples have none. All cross sections are NLO unless otherwise stated.

quantity in data. The data distribution is calculated based on the instantaneous luminosity and the inelastic proton-proton cross section. The ratio of these measurements is used to reweight the MC.

5.2.3 Locally Produced Monte Carlo Datasets

In addition to the centrally produced samples list above, two custom samples were produced in order to test the effect of different hadronizers and tunes on the generated ϕ^* distribution. The exact same events from the centrally produced powheg DY $\to e^+e^-$ sample were run through Pythia8 using two new tunes: `Tunepp5` and `Tunepp14`. These samples were not run through Geant4 as only the generator level ϕ^* distribution was needed from them because they are used only for comparison to the final result.

5.3 Scale Factors

The detector response to various signals is not always perfectly simulated in MC and so the efficiencies of various selection requirements are not the same in data and MC. In order to correct for this difference, each event in MC is reweighted with a series of

scale factors which are the efficiency of some selection requirement in data divided by the same efficiency as measured on MC, as follows:

$$\mathrm{SF} = \frac{\epsilon^{\mathrm{Data}}}{\epsilon^{\mathrm{MC}}} \tag{5.1}$$

Three scale factors are applied to each event: trigger, reconstruction, and identification. The trigger scale factors were measured by us and are detailed in section 5.3.2 while the reconstruction and identification scale factors were measured centrally by the CMS collaboration. The centrally-produced values are used as doing so is a requirement of passing the internal analysis review. The methods used to measure the reconstruction and isolation scale factors are summarized in sections 5.3.3 and 5.3.4 because the papers detailing them are not public

5.3.1 Tag and Probe

Tag and Probe (T&P) is a minimally-biased method of calculating the efficiency of some analysis selection requirement. T&P takes advantage of the well-known mass and narrow width of the Z boson to select a set of electrons for which very few selection requirements have been applied. This is done by finding one high-quality electron, the tag, and another minimally-biased object, the probe, that could be an electron, such as a supercluster. The invariant mass of these objects is computed and if it is near the Z mass peak, it is very likely that the probe is also an electron.

Once a set of minimally-biased probe electrons is constructed, the selection requirement can be applied to them. The efficiency of that requirement is then the number of probes that pass divided by the total number in the sample as follows:

$$\epsilon = \frac{\mathrm{n}^{\mathrm{obs.}}_{\mathrm{pass}}}{\mathrm{n}^{\mathrm{obs.}}_{\mathrm{total}}} \tag{5.2}$$

5.3.2 Single Electron Trigger

The efficiency of the HLT trigger used in this analysis, `HLT_Ele27_WP80`, is measured using T&P on the primary dataset. The efficiency is measured in bins of probe p_{T} and probe η with bin boundaries of {30, 40, 50, 70, 250} in p_{T} and {-2.1, -2.0, -1.556, -1.442, -0.8, 0., 0.8, 1.442, 1.556, 2.0, 2.1} in η.

Both the tag electron and the probe electron are required to satisfy $|\eta| < 2.1$, $p_{\mathrm{T}} > 30$, and to pass `Tight` requirements. These requirements are the same as required of the central electron in the full analysis selection, and hence the efficiency is measured relative to that selection. The pair must have an invariant mass such that $60\,\mathrm{GeV} < m_{\mathrm{ee}} < 120\,\mathrm{GeV}$. The tag electron is required to be matched to an electron that fired the trigger with $\Delta R < 0.3$. There is no requirement placed on the charge of the electron pair. Events with three or more electrons that pass these requirements are rejected.

Probes are considered passing if they are also matched to an electron that fired the trigger with $\Delta R < 0.3$, and failing otherwise. The efficiency in each bin is the number of passing probes divided by the number of failing probes, where the number of passing and failing probes is determined by using a simple count. In an individual event, both electrons are tried as a tag so that an event may contribute to the efficiency measurement twice if both electrons pass the tag requirements.

The efficiency is computed in exactly the same way on the MadGraph signal sample. The MC events are reweighted for pileup, reconstruction efficiency, and identification efficiency before the trigger efficiency is measured. The measured efficiencies for data and MC are listed in tables 5.4 and 5.5, respectively. The ratio of the data efficiency to the MC efficiency is shown in figure 5.2.

In the case of the trigger, because either electron could cause the event to pass, the scale factors can not be computed for each bin, but instead must be computed for each pair of bins. If only one electron in the event has $p_{\mathrm{T}} > 30$ and $|\eta| < 2.1$ then the scale factor is simply that given by equation (5.1), but if both electrons pass the requirements then either could have fired the trigger and so the scale factor is given by:

$$\mathrm{SF}_{1\text{ or }2} = \frac{1 - \left(1 - \epsilon_0^{\mathrm{Data}}\right)\left(1 - \epsilon_1^{\mathrm{Data}}\right)}{1 - \left(1 - \epsilon_0^{\mathrm{MC}}\right)\left(1 - \epsilon_1^{\mathrm{MC}}\right)} \tag{5.3}$$

where $\epsilon_{0,1}^{\mathrm{Data}}$ is the efficiency as measured in data for the 0th and 1st electrons, and $\epsilon_{0,1}^{\mathrm{MC}}$ is the efficiency as measured in MC. Equation (5.3) is just the probability that one or both of the electrons fired the trigger divided by the same quantity in MC. This equation assumes that the probability of one electron firing the trigger is uncorrelated with the probability of the other electron firing the trigger.

Although the agreement between the efficiency in data and in MC is good, the

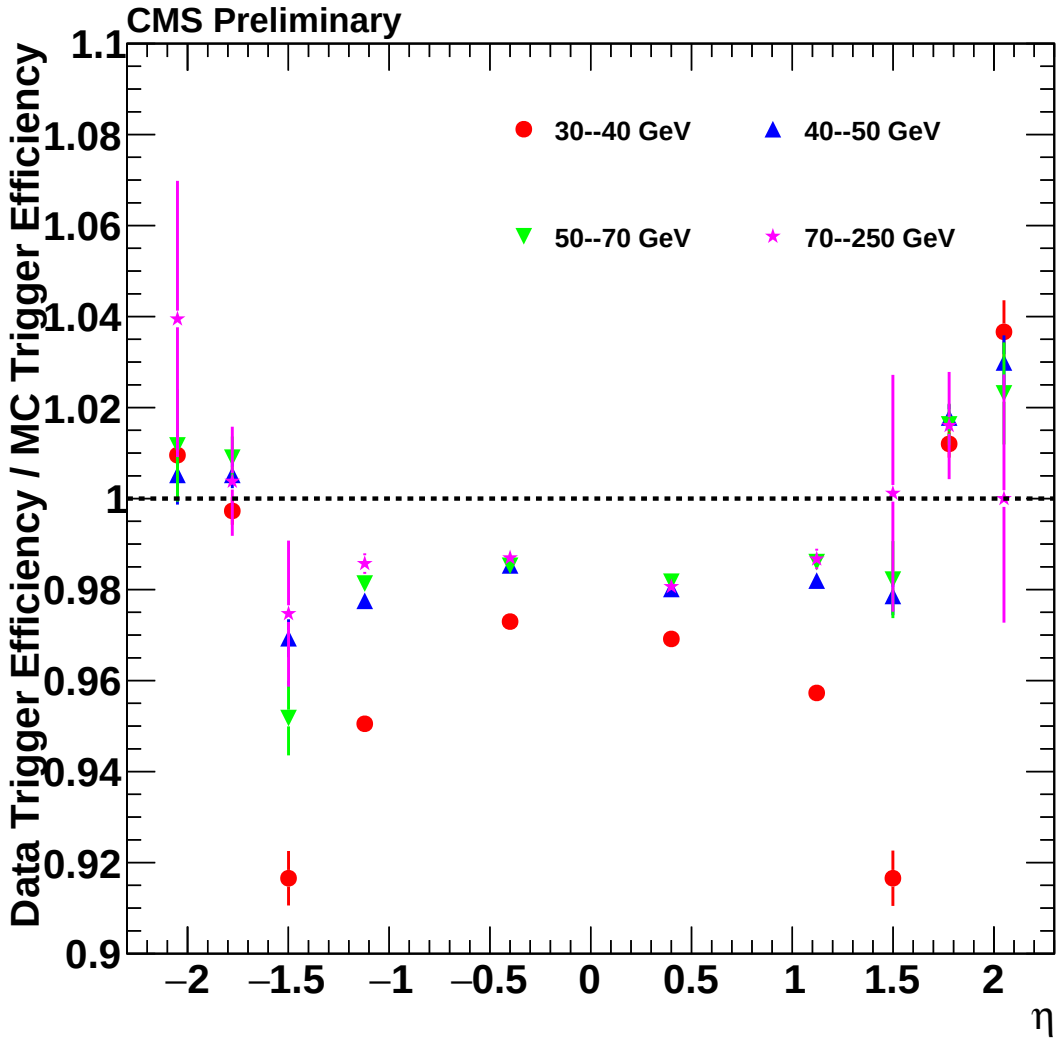


Figure 5.2: The ratio of the trigger efficiencies in data from table 5.4 over the trigger efficiencies in MadGraph MC from table 5.5.

η	30–40 GeV	40–50 GeV	50–70 GeV	70–250 GeV
-2.1 to -2	$0.741^{+0.003}_{-0.003}$	$0.773^{+0.003}_{-0.003}$	$0.780^{+0.005}_{-0.005}$	$0.79^{+0.01}_{-0.01}$
-2 to -1.556	$0.734^{+0.001}_{-0.001}$	$0.772^{+0.001}_{-0.001}$	$0.786^{+0.002}_{-0.002}$	$0.792^{+0.005}_{-0.005}$
-1.556 to -1.442	$0.725^{+0.003}_{-0.003}$	$0.821^{+0.002}_{-0.002}$	$0.809^{+0.004}_{-0.004}$	$0.848^{+0.010}_{-0.010}$
-1.442 to -0.8	$0.8930^{+0.0005}_{-0.0005}$	$0.9396^{+0.0003}_{-0.0004}$	$0.9509^{+0.0006}_{-0.0006}$	$0.966^{+0.001}_{-0.001}$
-0.8 to 0	$0.9213^{+0.0004}_{-0.0004}$	$0.9528^{+0.0002}_{-0.0002}$	$0.9601^{+0.0004}_{-0.0004}$	$0.9692^{+0.0010}_{-0.0010}$
0 to 0.8	$0.9174^{+0.0004}_{-0.0004}$	$0.9473^{+0.0003}_{-0.0003}$	$0.9561^{+0.0004}_{-0.0004}$	$0.963^{+0.001}_{-0.001}$
0.8 to 1.442	$0.8964^{+0.0005}_{-0.0005}$	$0.9424^{+0.0003}_{-0.0003}$	$0.9533^{+0.0006}_{-0.0006}$	$0.966^{+0.001}_{-0.001}$
1.442 to 1.556	$0.714^{+0.003}_{-0.003}$	$0.823^{+0.002}_{-0.002}$	$0.827^{+0.004}_{-0.004}$	$0.861^{+0.009}_{-0.010}$
1.556 to 2	$0.758^{+0.001}_{-0.001}$	$0.800^{+0.001}_{-0.001}$	$0.811^{+0.002}_{-0.002}$	$0.823^{+0.005}_{-0.005}$
2 to 2.1	$0.764^{+0.003}_{-0.003}$	$0.792^{+0.002}_{-0.002}$	$0.797^{+0.005}_{-0.005}$	$0.82^{+0.01}_{-0.01}$

Table 5.4: The electron trigger efficiency in data.

ratio of these efficiencies is not identically one, as might be expected if the selection used in T&P was strictly tighter than the trigger. There are several reasons for this. First, the trigger requires a lower value of H/E than `Tight` does. Second, the trigger uses simplified isolation variables while `Tight` uses particle flow isolation. Third, the detector calibrations used in the trigger (which ran in 2012 and so was calibrated with 2011 data) and the reconstructed data (reconstructed in 2013 using the data from the 2012 to calibrate) are different. Fourth and finally, any differences in the MC simulation of the trigger or the detector would lead to differences in the efficiencies.

5.3.3 Electron Reconstruction

Electron reconstruction begins with the assembly of a supercluster in ECAL and ends with the matching of a supercluster to a track in the tracker. The details of electron reconstruction are described in section 4.1. The efficiency of an electron with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ depositing enough energy in ECAL to be reconstructed into a supercluster is very high, although the exact efficiency must be measured in MC as there is no more basic

η	30–40 GeV	40–50 GeV	50–70 GeV	70–250 GeV
-2.1 to -2	$0.734^{+0.004}_{-0.004}$	$0.769^{+0.004}_{-0.004}$	$0.771^{+0.008}_{-0.008}$	$0.76^{+0.02}_{-0.02}$
-2 to -1.556	$0.736^{+0.002}_{-0.002}$	$0.768^{+0.002}_{-0.002}$	$0.779^{+0.003}_{-0.003}$	$0.789^{+0.008}_{-0.008}$
-1.556 to -1.442	$0.791^{+0.004}_{-0.004}$	$0.847^{+0.003}_{-0.003}$	$0.850^{+0.006}_{-0.006}$	$0.87^{+0.01}_{-0.02}$
-1.442 to -0.8	$0.9395^{+0.0006}_{-0.0006}$	$0.9612^{+0.0004}_{-0.0004}$	$0.9690^{+0.0007}_{-0.0008}$	$0.980^{+0.002}_{-0.002}$
-0.8 to 0	$0.9469^{+0.0005}_{-0.0005}$	$0.9670^{+0.0003}_{-0.0003}$	$0.9745^{+0.0005}_{-0.0005}$	$0.982^{+0.001}_{-0.001}$
0 to 0.8	$0.9466^{+0.0005}_{-0.0005}$	$0.9665^{+0.0003}_{-0.0003}$	$0.9739^{+0.0005}_{-0.0006}$	$0.982^{+0.001}_{-0.001}$
0.8 to 1.442	$0.9364^{+0.0007}_{-0.0007}$	$0.9597^{+0.0004}_{-0.0004}$	$0.9668^{+0.0008}_{-0.0008}$	$0.979^{+0.002}_{-0.002}$
1.442 to 1.556	$0.779^{+0.004}_{-0.005}$	$0.841^{+0.003}_{-0.003}$	$0.842^{+0.006}_{-0.006}$	$0.86^{+0.02}_{-0.02}$
1.556 to 2	$0.749^{+0.002}_{-0.002}$	$0.786^{+0.002}_{-0.002}$	$0.798^{+0.003}_{-0.003}$	$0.810^{+0.008}_{-0.008}$
2 to 2.1	$0.737^{+0.004}_{-0.004}$	$0.769^{+0.004}_{-0.004}$	$0.779^{+0.007}_{-0.008}$	$0.82^{+0.02}_{-0.02}$

Table 5.5: The electron trigger efficiency in MadGraph MC.

object with which to perform T&P to measure it in data. Failure to form superclusters is generally due to dead crystals in ECAL, which are accounted for in the detector response simulation. Scale factors for matching a track given that a supercluster has already been found were measured centrally by the CMS collaboration using T&P [85]. A summary of their method follows.

The events used to measure the reconstruction scale factors are selected with the dedicated electron T&P Trigger: `HLT_Ele20_CaloIdVT_CaloIsoVT_TrkIdT_TrkIsoVT_SC4_Mass50`. This trigger requires one electron with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ which must also pass very tight isolation and ID requirements while requiring only a low energy ($E_{\mathrm{T}} > 4\,\mathrm{GeV}$) supercluster as the other leg. The trigger rate is kept down by requiring that the invariant mass of these two objects is greater than 50 GeV.

The events selected by the trigger are further required to pass a set of selection requirements. The tag electron is required to pass `Tight`, have $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, and $|\eta| < 2.5$. Electrons are rejected if they fall in the seam between EB and EE ($1.4442 < |\eta| < 1.566$). The tag must also be matched to the tight leg of the T&P trigger. The

probe supercluster has minimal requirements applied; it is required to have tracker isolation < 0.15. For the MC sample, the tag is required to be matched to a generator level electron with $\Delta R < 0.2$. Additionally, the event is required to have low particle flow missing energy ($E_{\mathrm{T}}^{\mathrm{PF\ miss}} < 20\,\mathrm{GeV}$) in order to reject poorly reconstructed events.

The events were binned in terms of probe's p_{T} and η as well as whether the probe passed or failed. In each bin, the m_{ee} distribution was constructed and a template consisting of the sum of a Gaussian smeared $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ MC sample and an exponential background was fitted. The number of events predicted by the signal fit on the passing sample, failing sample, and sum of the two samples was used to get the efficiency. A similar process was performed on MC, although instead of a fit a simple counting of passing events was performed (as there is no background in MC). The resulting scale factors are given in table 5.6.

$\lvert\eta\rvert$	20–30 GeV	30–40 GeV	40–50 GeV	> 50 GeV
0.0 to 0.8	$0.982^{\pm 0.003}_{\pm 0.012}$	$0.988^{\pm 0.001}_{\pm 0.008}$	$0.990^{\pm 0.001}_{\pm 0.004}$	$0.990^{\pm 0.001}_{\pm 0.004}$
0.8 to 1.4442	$0.993^{\pm 0.002}_{\pm 0.012}$	$0.993^{\pm 0.001}_{\pm 0.008}$	$0.993^{\pm 0.001}_{\pm 0.004}$	$0.991^{\pm 0.001}_{\pm 0.004}$
1.4442 to 1.566	$1.016^{\pm 0.012}_{\pm 0.020}$	$0.985^{\pm 0.004}_{\pm 0.009}$	$0.987^{\pm 0.004}_{\pm 0.004}$	$0.974^{\pm 0.009}_{\pm 0.006}$
1.566 to 2.0	$0.988^{\pm 0.003}_{\pm 0.012}$	$0.993^{\pm 0.002}_{\pm 0.008}$	$0.992^{\pm 0.001}_{\pm 0.004}$	$0.990^{\pm 0.003}_{\pm 0.004}$
2.0 to 2.5	$1.002^{\pm 0.004}_{\pm 0.012}$	$1.004^{\pm 0.002}_{\pm 0.008}$	$1.005^{\pm 0.002}_{\pm 0.004}$	$0.998^{\pm 0.004}_{\pm 0.004}$

Table 5.6: Scale factors for GSF electron reconstruction. The upper uncertainty listed is statistical, the lower is systematic.

5.3.4 Electron Identification

Not all electrons which are reconstructed pass the ID criteria used in this analysis, specifically `Medium` and `Tight`, the details of which are covered in section 4.3. The efficiency of going from a reconstructed electron to one which passes the identification criteria is measured centrally by the CMS collaboration using T&P [86]. A summary of their method follows.

The events used for this measurement were selected using two triggers:

HLT_Ele20_CaloIdVT_CaloIsoVT_TrkIdT_TrkIsoVT_SC4_Mass50, which is described above, and HLT_Ele17_CaloIdVT_CaloIsoVT_TrkIdT_TrkIsoVT_Ele8_Mass50, which requires one electron with $p_{\mathrm{T}} > 17\,\mathrm{GeV}$ and tight isolation and ID requirements while also requiring a reconstructed second electron (as opposed to a supercluster as required by the first trigger) with $p_{\mathrm{T}} > 8\,\mathrm{GeV}$. It further requires a dielectron invariant mass of $m_{\mathrm{ee}} > 50\,\mathrm{GeV}$.

The tag electrons are required to pass Tight, have $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, and $|\eta| < 2.5$; they are rejected if they fall in the seem between EB and EE ($1.4442 < |\eta| < 1.566$). The tag is not required to match the trigger. Probe electrons have the same η requirements as tags, but are only required to have $p_{\mathrm{T}} > 10\,\mathrm{GeV}$. Passing probes pass the ID criteria under investigation, failing probes fail the ID criteria. The invariant mass of the tag and probe pair is required to be near the Z mass peak ($60\,\mathrm{GeV} < m_{\mathrm{ee}} < 120\,\mathrm{GeV}$). The electrons are required to have charges of opposite sign. In MC, the probe is only required to be matched to a generator electron with $\Delta R < 0.2$.

The efficiencies are then calculated by fitting the m_{ee} distributions using a template constructed with a $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ MC sample and an exponential background. The three categories (passing probes, failing probes, and all probes) are then simultaneously fit with this template and the number of fitted signal events is used to derive an efficiency. A simple count of events is used for the MC efficiency instead of a fit. The resulting scale factors are given in tables 5.7 and 5.8.

$\lvert\eta\rvert$	20–30 GeV	30–40 GeV	40–50 GeV	50–200 GeV
0.0 to 0.8	$0.960^{+0.003}_{-0.003}$	$0.978^{+0.001}_{-0.001}$	$0.981^{+0.001}_{-0.001}$	$0.982^{+0.002}_{-0.002}$
0.8 to 1.4442	$0.936^{+0.004}_{-0.004}$	$0.958^{+0.002}_{-0.002}$	$0.969^{+0.001}_{-0.001}$	$0.969^{+0.002}_{-0.002}$
1.4442 to 1.566	$0.933^{+0.015}_{-0.017}$	$0.907^{+0.008}_{-0.008}$	$0.904^{+0.004}_{-0.004}$	$0.926^{+0.011}_{-0.011}$
1.566 to 2.0	$0.879^{+0.007}_{-0.007}$	$0.909^{+0.003}_{-0.003}$	$0.942^{+0.002}_{-0.002}$	$0.957^{+0.004}_{-0.004}$
2.0 to 2.5	$0.974^{+0.004}_{-0.004}$	$0.987^{+0.004}_{-0.004}$	$0.991^{+0.003}_{-0.003}$	$0.999^{+0.005}_{-0.005}$

Table 5.7: Scale factors for Tight electron ID.

$\lvert\eta\rvert$	20–30 GeV	30–40 GeV	40–50 GeV	50–200 GeV
0.0 to 0.8	$0.986^{+0.002}_{-0.001}$	$1.002^{+0.001}_{-0.001}$	$1.005^{+0.001}_{-0.001}$	$1.004^{+0.001}_{-0.001}$
0.8 to 1.4442	$0.959^{+0.003}_{-0.003}$	$0.980^{+0.001}_{-0.001}$	$0.988^{+0.001}_{-0.001}$	$0.988^{+0.002}_{-0.002}$
1.4442 to 1.566	$0.967^{+0.007}_{-0.013}$	$0.950^{+0.006}_{-0.007}$	$0.958^{+0.005}_{-0.005}$	$0.966^{+0.009}_{-0.009}$
1.566 to 2.0	$0.941^{+0.005}_{-0.005}$	$0.967^{+0.003}_{-0.003}$	$0.992^{+0.002}_{-0.002}$	$1.000^{+0.003}_{-0.003}$
2.0 to 2.5	$1.020^{+0.003}_{-0.003}$	$1.021^{+0.003}_{-0.003}$	$1.019^{+0.002}_{-0.002}$	$1.022^{+0.004}_{-0.004}$

Table 5.8: Scale factors for `Medium` electron ID.

Chapter 6

Event Selection

6.1 Acceptance

The acceptance region is a definition of what events, assuming that there are no limitations due to the detector design, are included in the analysis. The acceptance region is what the final result of this analysis will be corrected to, ensuring that other measurements and theoretical predictions can be compared to our results without having to account for the detector response. The acceptance region defines what sort of physics results we can make statements about, and also determines the value of the effective cross section of the Z.

Our acceptance is defined by the kinematics of the two electrons and the mass of the Z boson. One of the electrons, called the central electron, is required to have $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and to be within the central region of the detector (hence the name) with $|\eta| < 2.1$. The other electron, called the extended electron, has looser requirements; it must have $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and is not required to be as central with $|\eta| < 2.4$. The requirements on the central electron were selected in conjunction with the $\mathrm{Z} \to \mu^+\mu^-$ measurement of ϕ^* at CMS so that that measurement and this one could be easily combined into a joint measurement. The pseudorapidity limit was selected to match the most efficient region of CMS's single muon trigger, while the transverse momentum threshold was dictated by threshold on single electron trigger. The pseudorapidity limit on the extended electron was chosen to keep all of the muons within the region covered by the muon system, while the transverse momentum threshold was selected because

for $p_{\mathrm{T}} < 20\,\mathrm{GeV}$ the rate of fake electrons increases.

There are other particles, like the J/ψ, that decay to e^+e^- pairs as shown in figure 6.1. Fortunately, none of these other particles are near the Z in mass, and so we can eliminate them from our acceptance by requiring a mass near the Z peak. We therefore define our mass window around the nominal Z mass of 91 GeV, extending from 60 GeV to 120 GeV.

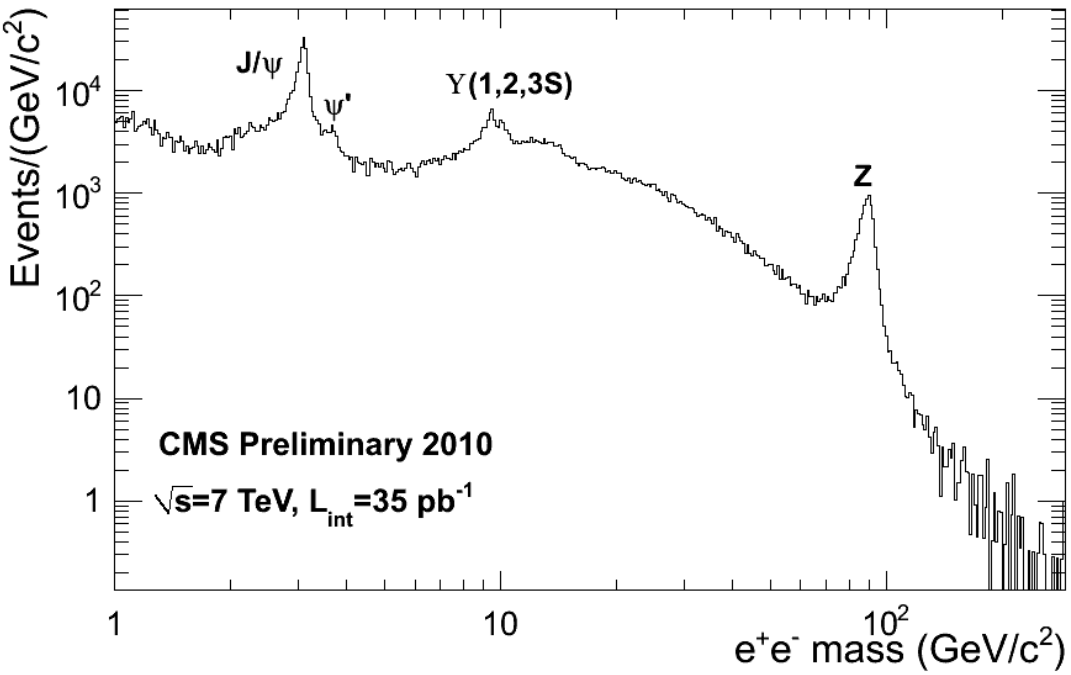


Figure 6.1: The spectrum of e^+e^- events as measured by CMS in 2010.

6.2 Object Selection

The events in the data that will be used in the analysis are selected by looking at the physics objects contained within them. A selection is made on both the individual electrons, and on the Z boson they are assumed to come from.

6.2.1 Electron Selection

The requirements used to select electrons are chosen to be tighter than the requirements used at the trigger level in order to make calculating the various efficiencies easier. For an event to be considered it must have at least two electrons that pass the acceptance requirements for extended electrons: $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.4$. If three or more electrons pass this initial requirement, only the two with the highest p_{T} are considered. One of these two electrons must be within $|\eta| < 2.1$ and it must also have $p_{\mathrm{T}} > 30\,\mathrm{GeV}$.

A central electron in the event is required to pass `Tight` and to be matched to one of the electrons that passed `HLT_Ele27_WP80` with $\Delta R < 0.3$. The other electron must pass `Medium`. The definitions of `Medium` and `Tight` are discussed in section 4.3. If both electrons are central electrons, then only one of them need pass `Tight`, but this same electron must also match the trigger; the other central electron need only pass `Medium`.

No charge requirement is applied because the number of signal events removed by this selection is larger than the number of background events suppressed.

The distributions of electron p_{T} and η for the highest p_{T} electron and the second highest p_{T} electron after all selection requirements are applied are shown in figures 6.2 and 6.3.

6.2.2 Z Selection

The Z boson decays too quickly to leave any direct signal in the detector, so it is reconstructed from its decay products: the two electrons whose selection is described in section 6.2.1. In events where the two electrons pass the selection, the Z is constructed by taking the sum of the electron four-vectors. The resulting invariant mass of the Z must be within the region set by the acceptance ($60\,\mathrm{GeV} < m_{\mathrm{ee}} < 120\,\mathrm{GeV}$).

The distributions of m_{ee}, Z Q_{T}, and Z Y after these selection requirements are applied are shown with MadGraph signal MC in figures 6.4, 6.6, and 6.8, and with powheg + Pythia6 (`Z2star`) signal MC in figures 6.5, 6.7, and 6.9. The MC samples have had all of the corrections discussed in section 4.2 applied, as well as the scale factors discussed in section 5.3. Even so, the MadGraph sample does not reproduce the distributions seen in data.

The effect of these disagreements is considered in section 7.1.10. The disagreement

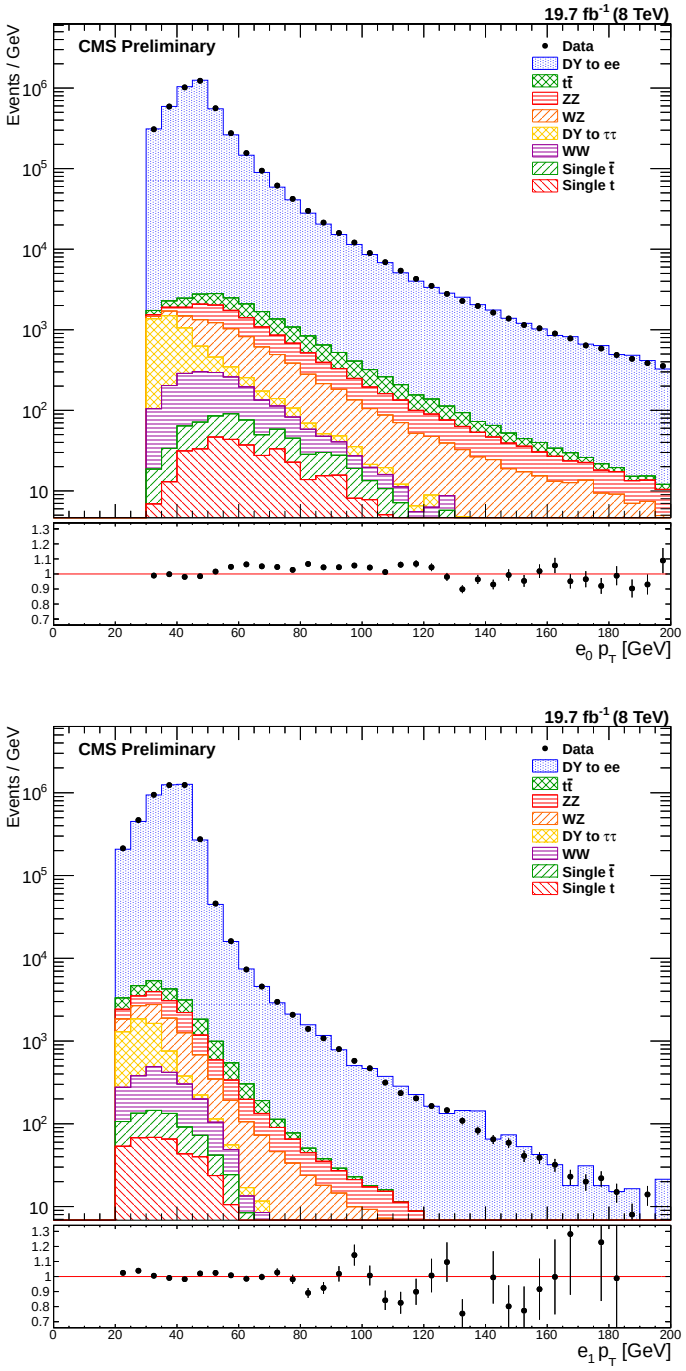


Figure 6.2: The p_{T} distribution of the higher (top) and lower (bottom) p_{T} electrons in data (points) and MadGraph signal MC and the background MC samples (histograms) for events passing the full analysis selection.

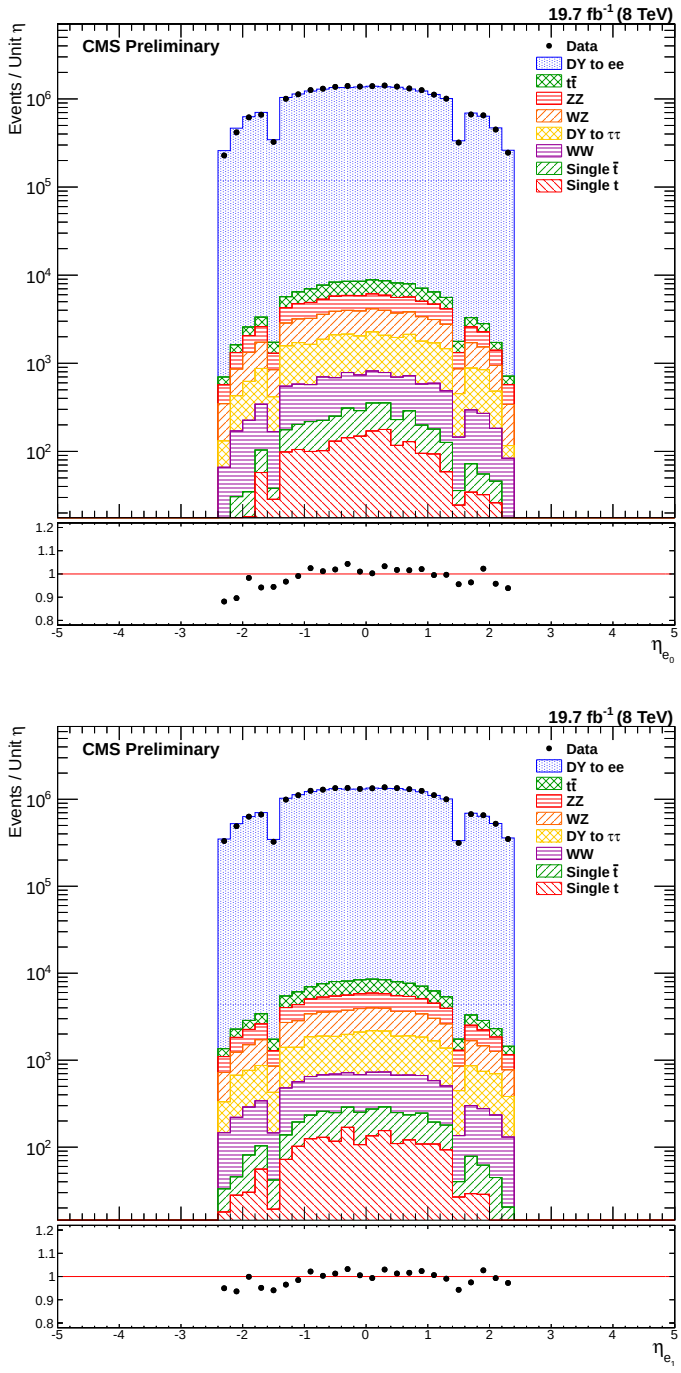


Figure 6.3: The η distribution of the higher (top) and lower (bottom) p_{T} electrons in data (points) and MadGraph signal MC and the background MC samples (histograms) for events passing the full analysis selection.

in the Z Q_{T} distribution is not considered as a source of uncertainty as ϕ^* is related to boson Q_{T}, and hence the disagreement is motivation for the measurement, not something to be corrected for.

6.3 Background Estimation

The distribution of background events is estimated using the MC samples discussed in section 5.2.2, with the exception of the QCD multi-jet and W + jets backgrounds, which are computed using data. The various MC samples are reweighted so that they have equivalent luminosity to the data. The MC datasets also have various corrections applied which are discussed in section 5.3. These scale factors are applied as weights to each MC event in order to make them better match the data.

After the reweighting, the selection requirements are applied to to the MC. The number of events that survive the selection are taken as the number in our data, and are subtracted off from the data events.

Overall, the selection requirements discussed in the previous sections leave a very pure sample of Z boson events: 99.44 % of events are signal. The dominant backgrounds are the diboson backgrounds and $\mathrm{t\bar{t}}$. We define Z bosons produced in association with another weak boson as a background because of their different production mechanism as opposed to single Z boson events. The fraction of events from each of the considered backgrounds is listed in table 6.1.

6.3.1 Electron–Muon Control Sample

Many of the backgrounds—specifically $\mathrm{t\bar{t}}$, WW, DY $\rightarrow \tau^+\tau^-$, tW, and $\bar{\mathrm{t}}$W—produce leptons via independent decay chains. Therefore, the flavor of the two leptons is not constrained to be the same. An electron–muon (e–μ) background dominated control sample is used to test how well the MC reproduce the various backgrounds.

Events for the e–μ sample are selected from data taken with the `HLT_IsoMu24_eta2p1`. This trigger required a single muon with $p_{\mathrm{T}} > 24\,\mathrm{GeV}$ and $|\eta| < 2.1$. Events selected from the dataset provided by this trigger are required to have one muon with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.1$ and one electron with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.4$. In order to further suppress Z decays, events are rejected if there is a third lepton of either flavor

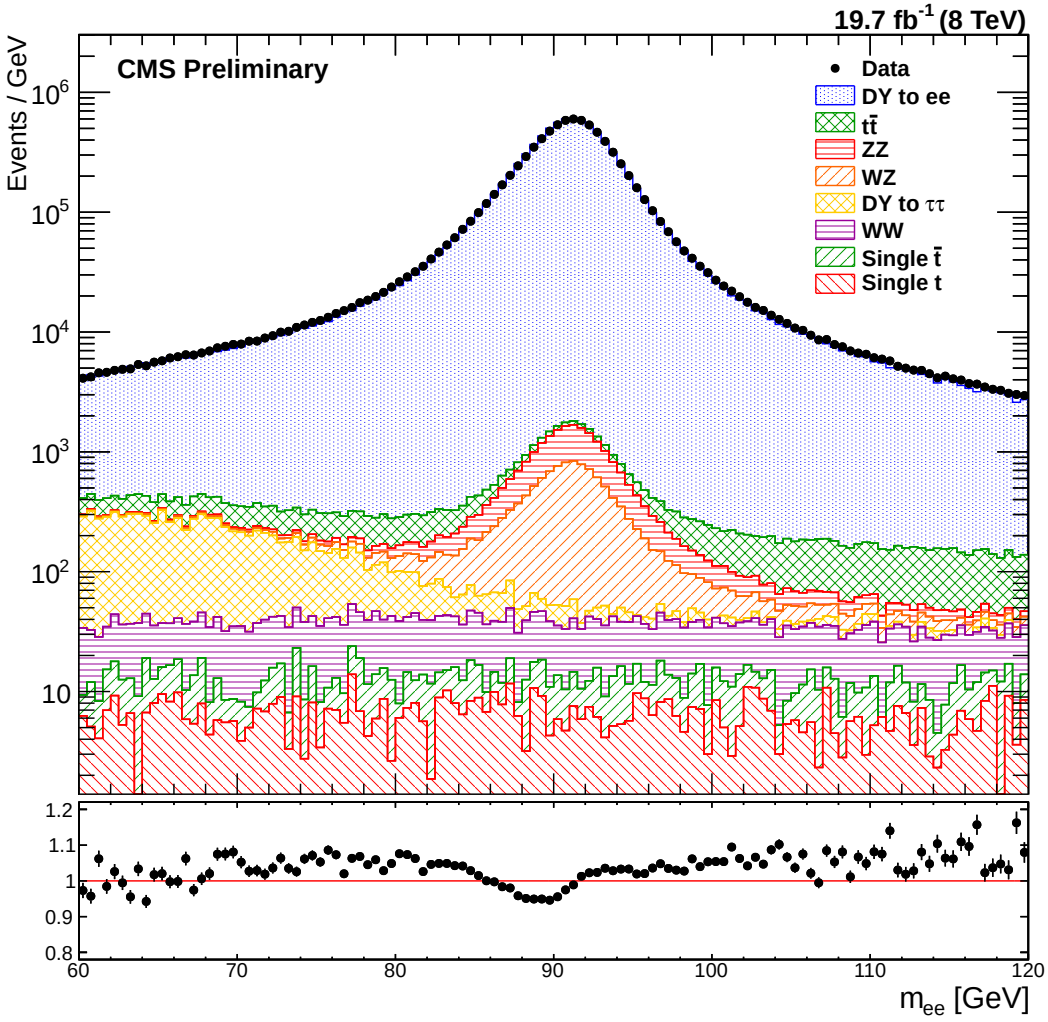


Figure 6.4: The m_{ee} distribution of all events passing the final selection in data (points) and MadGraph signal MC and the background MC samples (histograms).

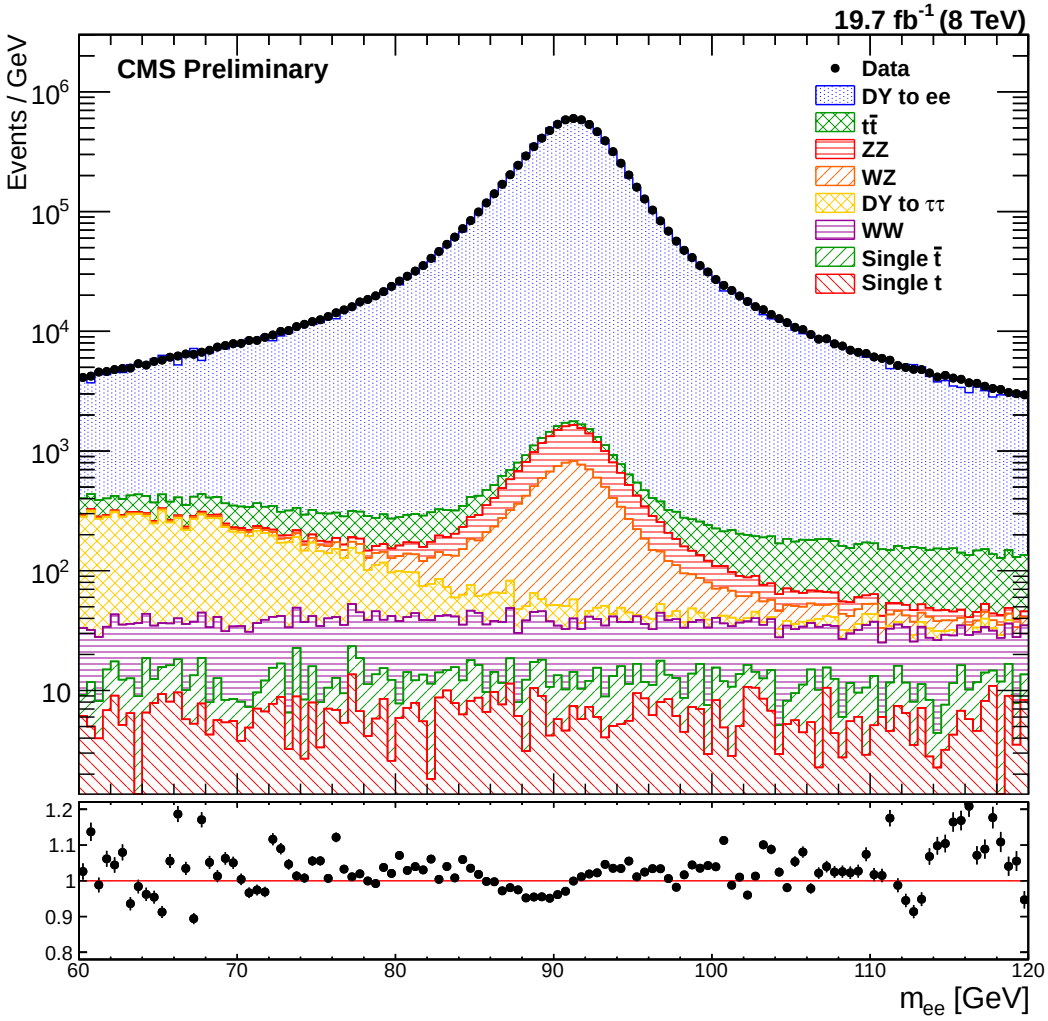


Figure 6.5: The m_{ee} distribution of all events passing the final selection in data (points) and POWHEG + PYTHIA6 (`Z2star`) signal MC and the background MC samples (histograms).

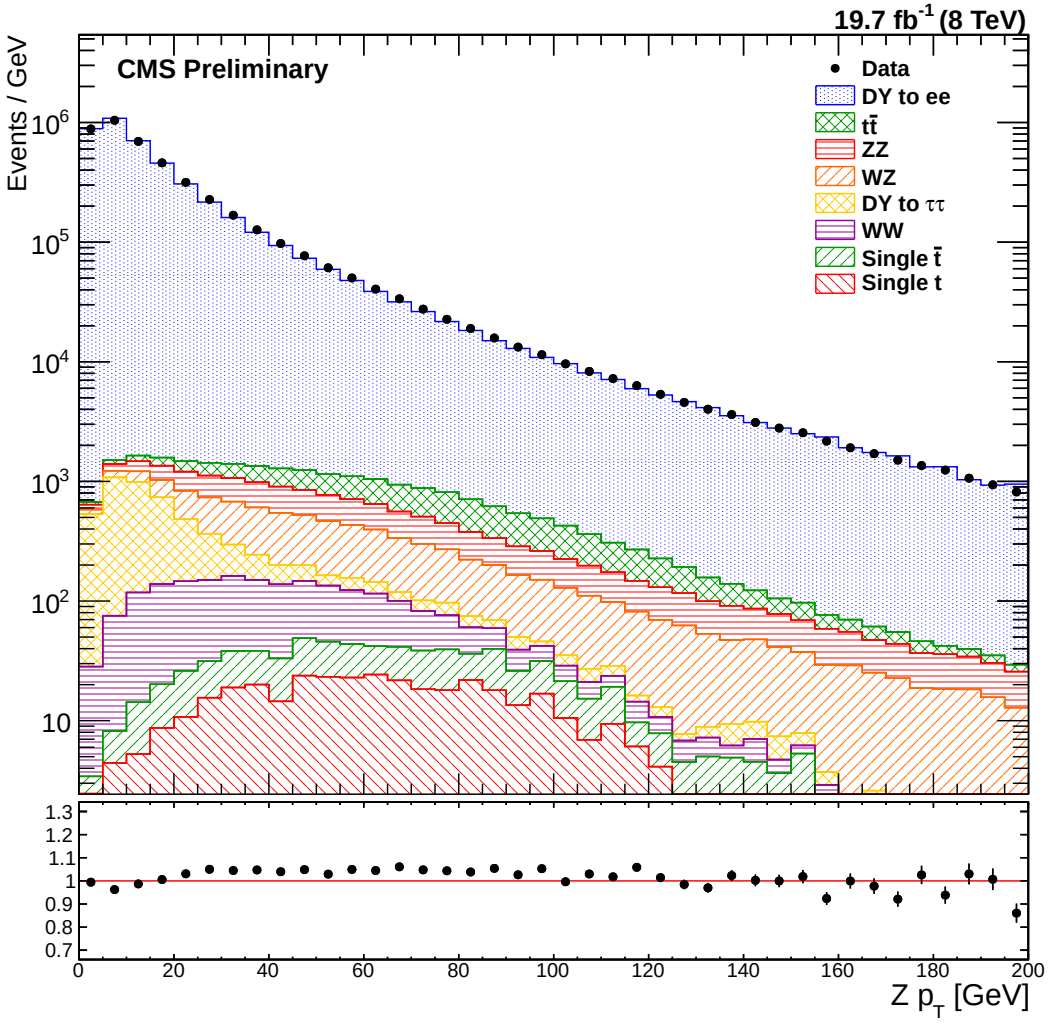


Figure 6.6: The Q_{T} distribution of Z bosons for all events passing the final selection in data (points) and MadGraph signal MC and the background MC samples (histograms).

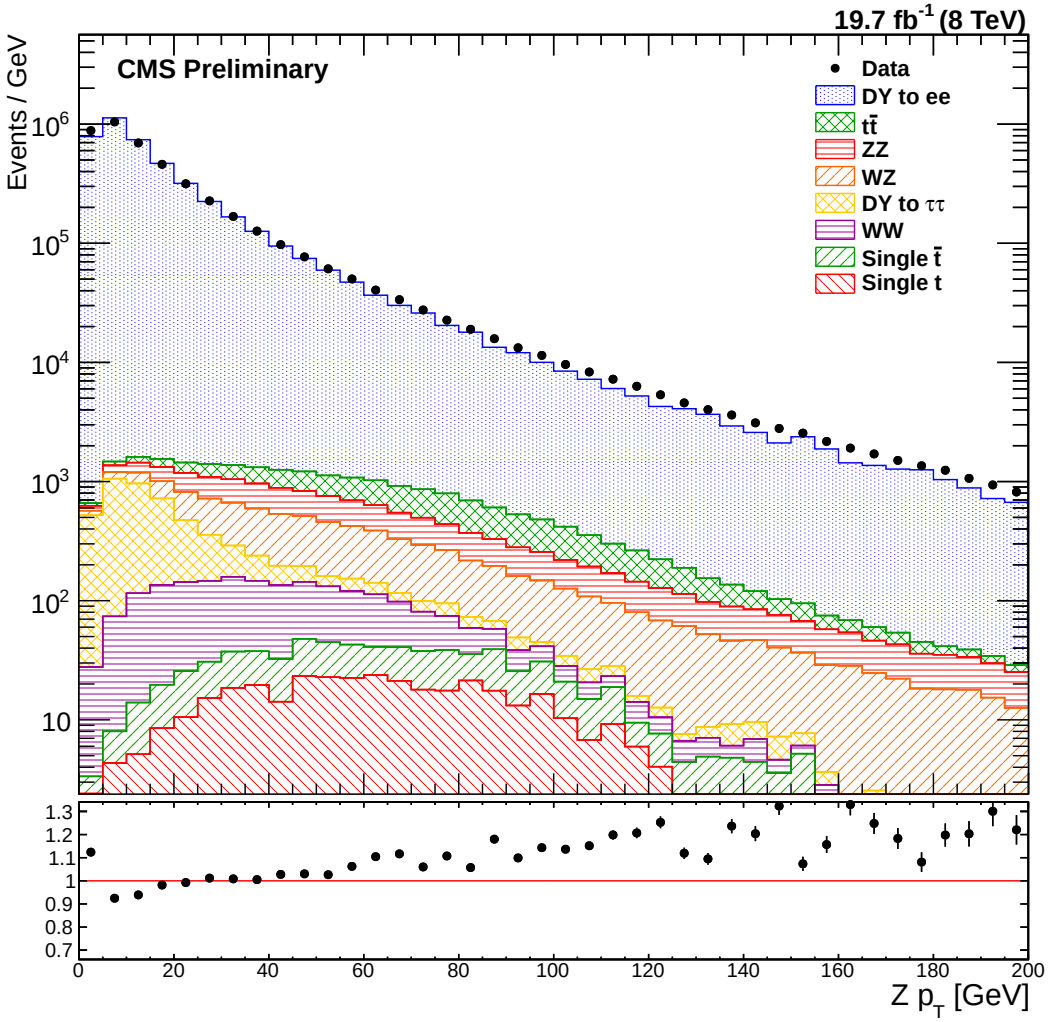


Figure 6.7: The Q_{T} distribution of Z bosons for all events passing the final selection in data (points) and POWHEG + PYTHIA6 (`Z2star`) signal MC and the background MC samples (histograms).

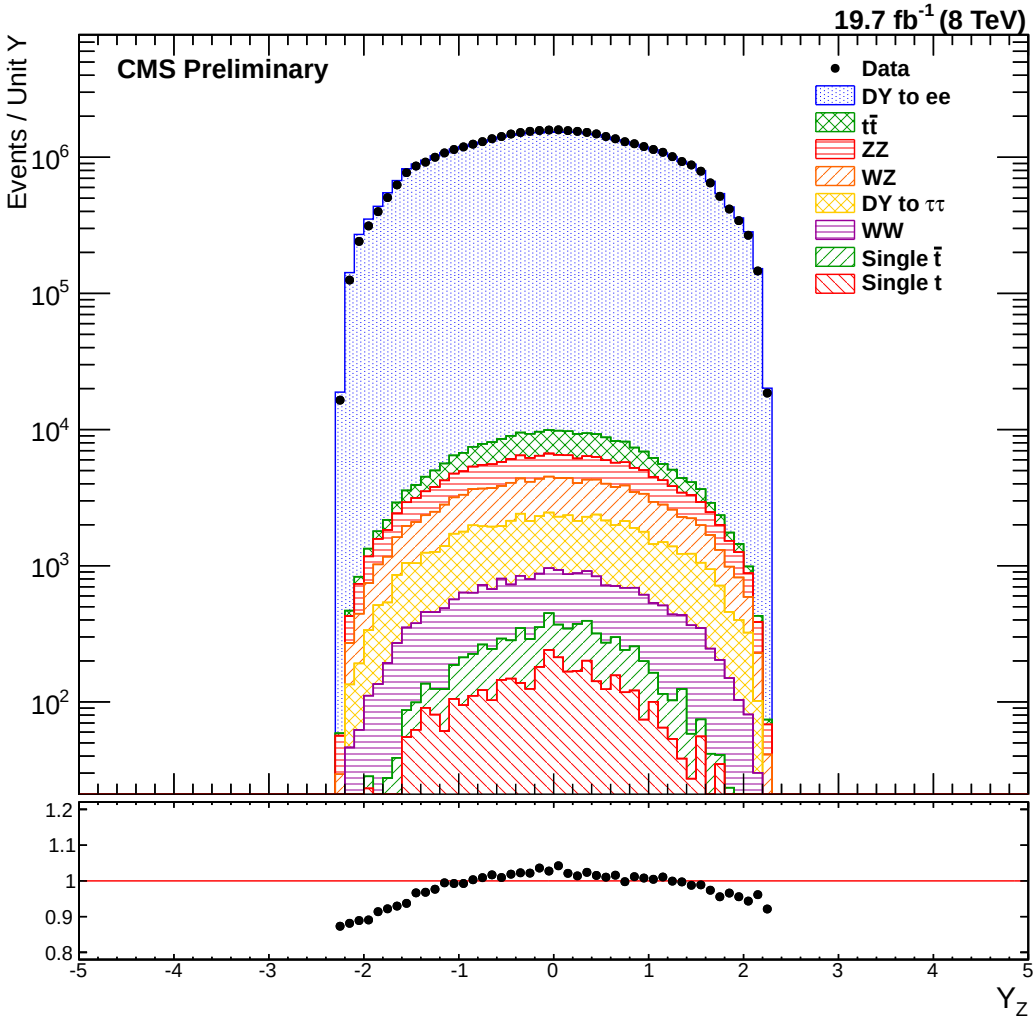


Figure 6.8: The Y distribution of Z bosons for all events passing the final selection in data (points) and MadGraph signal MC and the background MC samples (histograms).

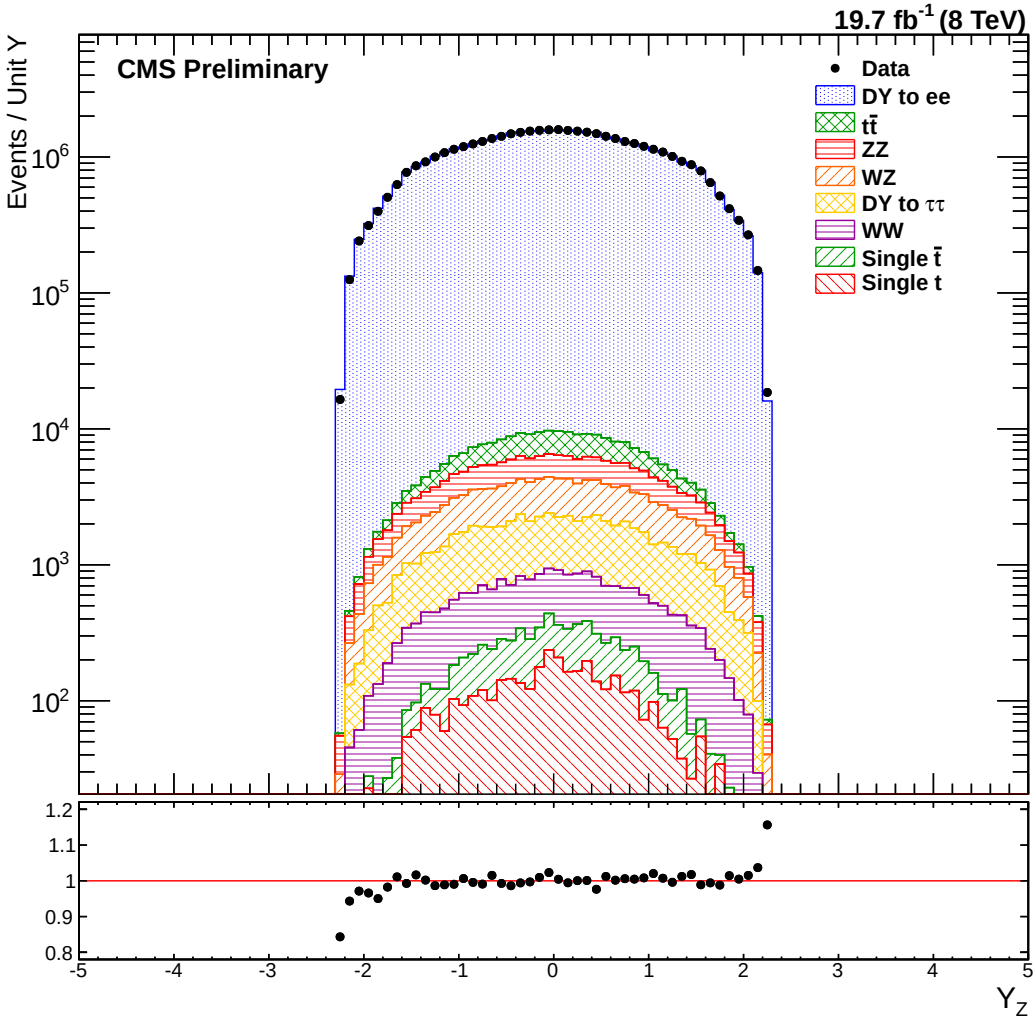


Figure 6.9: The Y distribution of Z bosons for all events passing the final selection in data (points) and POWHEG + PYTHIA6 (`Z2star`) signal MC and the background MC samples (histograms).

Process	of total	of background
Signal: DY $\rightarrow e^+e^-$	99.44 %	N.A.
$t\bar{t}$	0.15 %	26.6 %
ZZ	0.12 %	21.6 %
WZ	0.12 %	21.4 %
DY $\rightarrow \tau^+\tau^-$	0.09 %	16.2 %
WW	0.03 %	5.6 %
QCD multi-jet and W + jets	0.03 %	5.6 %
tW and $\bar{t}$W	0.02 %	2.9 %

Table 6.1: The estimated background contamination in the dataset as a percentage of all events, and the fraction of all background events from each background.

with $p_T > 20\,\text{GeV}$ and $|\eta| < 2.4$. Figure 6.10 shows a comparison of the MC and data in the control region. Although no deviations from 1 are seen, this ratio is taken as a scale factor to correct the previously listed backgrounds in each bin. A comparison of the $Z \rightarrow \mu^+\mu^-$ MC samples before and after this scale factor is applied is shown in figure 6.11.

This method could not be used to calibrate the WZ and ZZ background MC samples as they contain actual Z bosons. Instead a 20 % uncertainty is taken on their theoretical cross-section.

6.3.2 QCD multi-jet and W + jets Background Estimation

The centrally-produced QCD multi-jet and W + jets MC samples do not have enough events to make accurate estimations of their respective background in this analysis so instead a data-driven method is employed. The same requirements as discussed in section 6.2.1 are applied to both the data and the MC samples with the additional requirement that both electrons must have the same charge. This additional selection requirement removes most of the signal while removing only half of the expected QCD multi-jet and W + jets backgrounds because these processes are independent and hence there is no correlation between the charge of the two leptons.

The data and MC are divided into subsamples with each subsample corresponding to

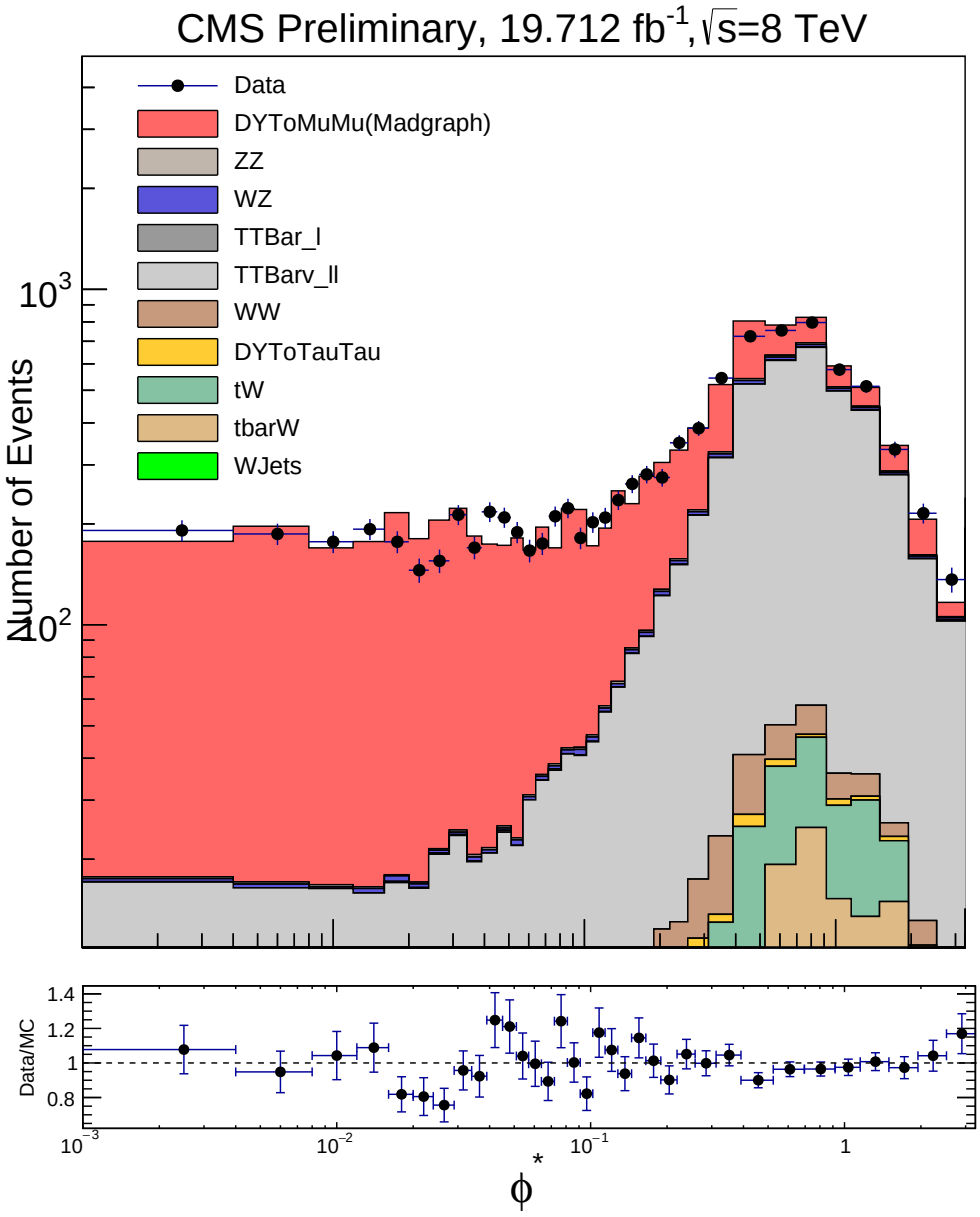


Figure 6.10: The ϕ^* distribution of events from the e–μ control sample. The data (points) match the MC (histograms) expectation well. The ratio, shown in the bottom plot, is taken as a scale factors and used to correct the backgrounds in each bin.

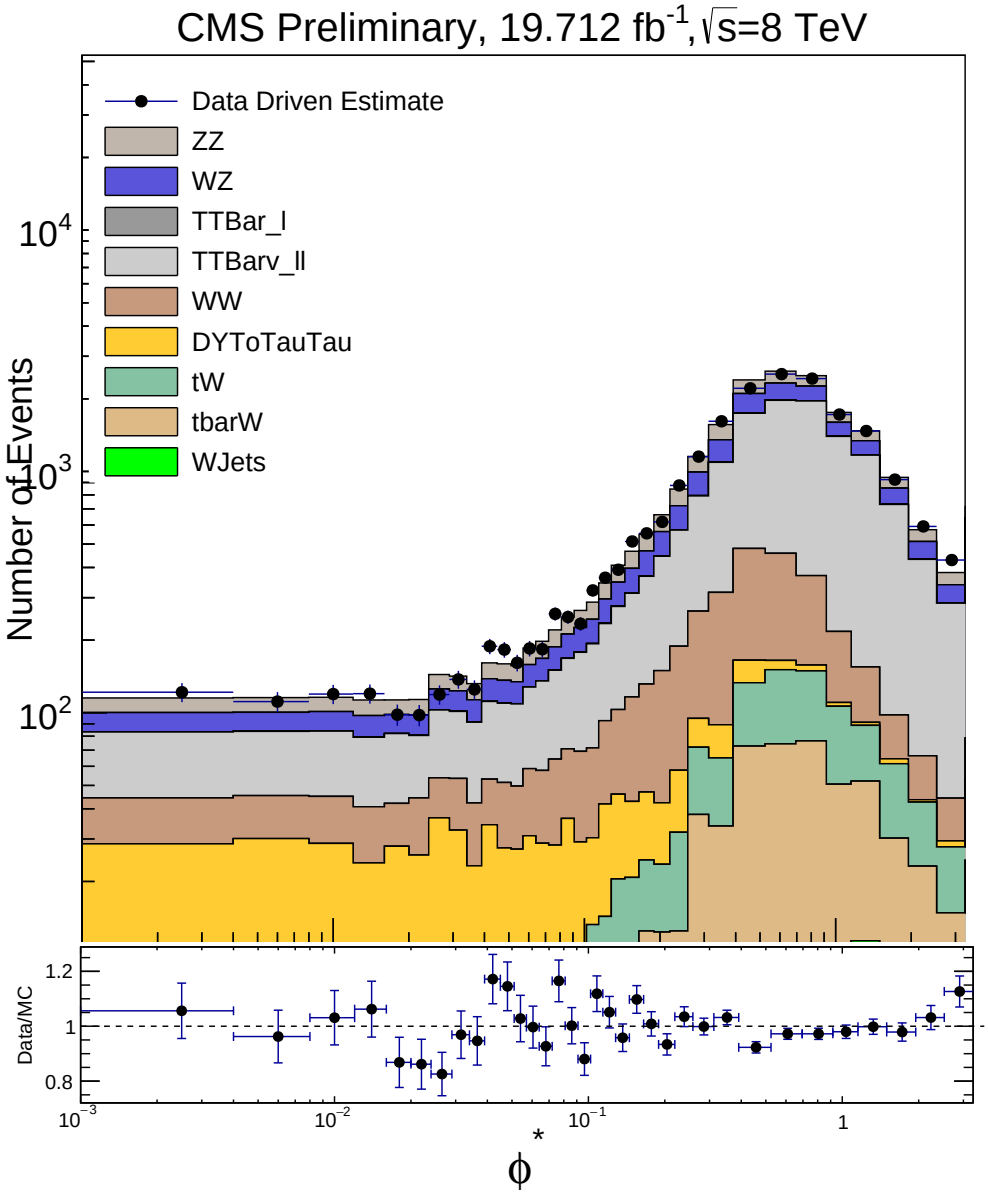


Figure 6.11: The ϕ^* distribution of $Z \to \mu^+\mu^-$ events in MC before and after the e–μ correction. The samples without the e–μ derived scale factors applied are shown in the histograms, while the points show the sum of all the MC samples after the scale factors are applied.

a bin in the ϕ^* distribution. The m_{ee} distribution of the events in each data subsample is fit with a combination of an MC template and an analytic background function. The MC templates used in the fits are created by reweighting each MC sample to have the same luminosity as the data and then summing the samples together. The shape of the resulting distribution is the template with its amplitude fit as a free parameter.

The analytic background function, F_{BG}, is the same function used by Haupt to model the QCD multi-jet and $\mathrm{W}+\mathrm{jets}$ backgrounds in his $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ shape measurement [87]. It is composed of a falling exponential—which fits the general shape of the QCD multi-jet and $\mathrm{W}+\mathrm{jets}$ background distribution—multiplied by a complementary error function—which cuts off the exponential at low mass. The exact form of this function is given by equation (6.1). The sum of the template and the background function is given by equation (6.2).

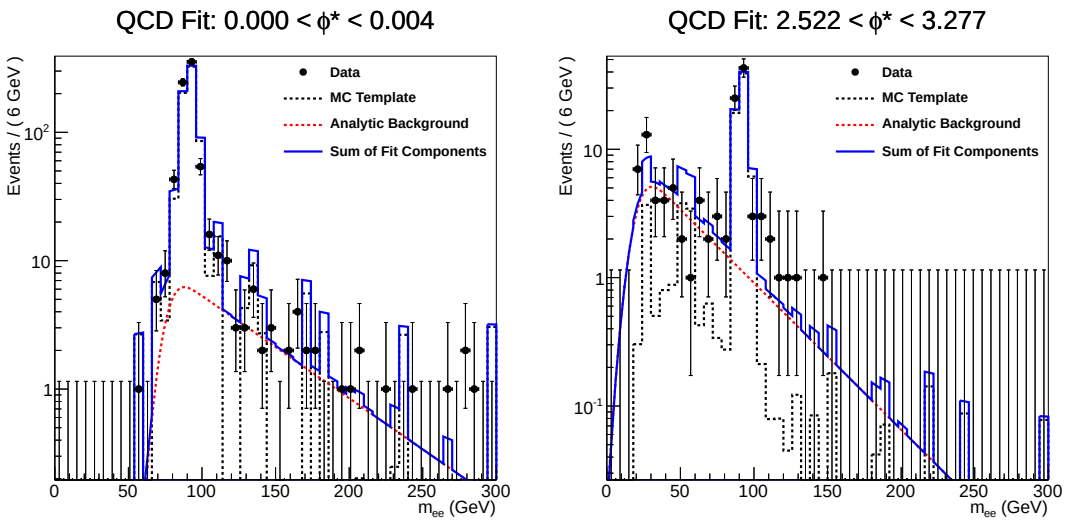
$$F_{\mathrm{BG}}(x;\gamma,\delta,\varepsilon) = e^{-\gamma x}\,\mathrm{erfc}\left(\frac{\varepsilon - x}{\delta}\right) \tag{6.1}$$

$$\alpha\, T_{\mathrm{MC}} + \beta\, F_{\mathrm{BG}}(x;\gamma,\delta,\varepsilon) \tag{6.2}$$

The background due to QCD multi-jet and $\mathrm{W}+\mathrm{jets}$ in each ϕ^* bin is taken to be twice the integral of the analytic background component from $60\,\mathrm{GeV} < m_{\mathrm{ee}} < 120\,\mathrm{GeV}$. The factor of two comes from the fact that the estimate was performed using only events containing same-charge electrons while the full analysis accepts both same- and opposite-charge events which are equally likely in the background. Two example fits are shown in figure 6.12. All of the fits are shown in appendix C. The estimated QCD multi-jet and $\mathrm{W}+\mathrm{jets}$ backgrounds in each ϕ^* bin are showing in figure 6.13, and presented in table C.1.

6.4 Unfolding

Particle detectors are incredibly sophisticated machines designed to make precise measurements of the various decay products created during the collisions. However, there are physical limitations that prevent a perfect measurement from being made. The finite



(a) (b)

Figure 6.12: Examples of the QCD multi-jet and W + jets data-driven background fits for the first and last ϕ^* bins. The data are shown as points with error bars, MC template as a dashed histogram, the analytic background function as the dashed line, and the sum of the template and function as a solid histogram. The full set of plots are shown in appendix C.

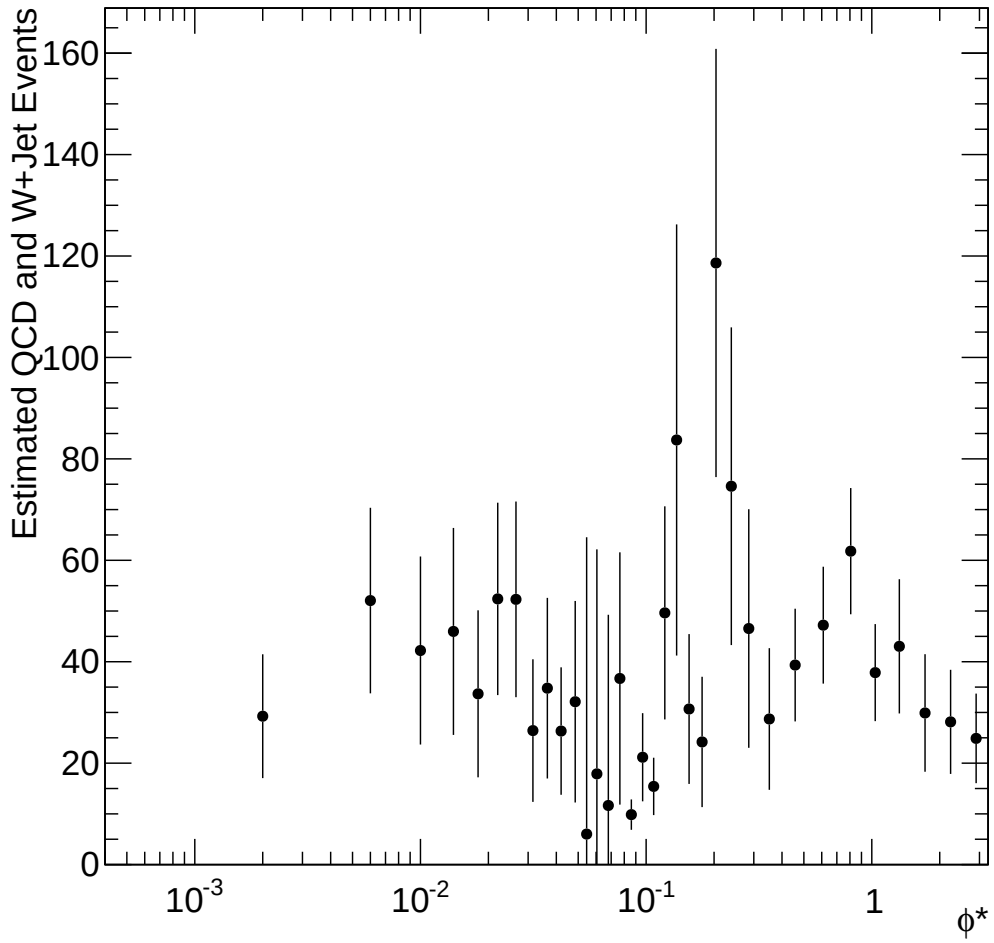


Figure 6.13: Estimate of the QCD multi-jet and W + jets background in each ϕ^* bin. Table C.1 presents the same information in numeric form.

energy, momentum, and position resolution of the various subdetectors impose limitations on the final measurements. In order to allow our measurement to be compared to theoretical predictions, we correct for these detector effects with a process known as unfolding.

We unfold in two steps: first the data are unfolded to correct for bin migration, second we correct for the imperfect efficiency of the trigger and reconstruction. The data are unfolded against the dressed definition of generator level electrons as discussed in section 2.3.3.

The unfolding was tested with the MadGraph DY $\to \ell^+\ell^-$ MC sample, and the powheg + Pythia6 (`Z2star`) DY $\to e^+e^-$ MC sample. These samples will be referred to as MadGraph and powheg respectively.

6.4.1 Bin Migration

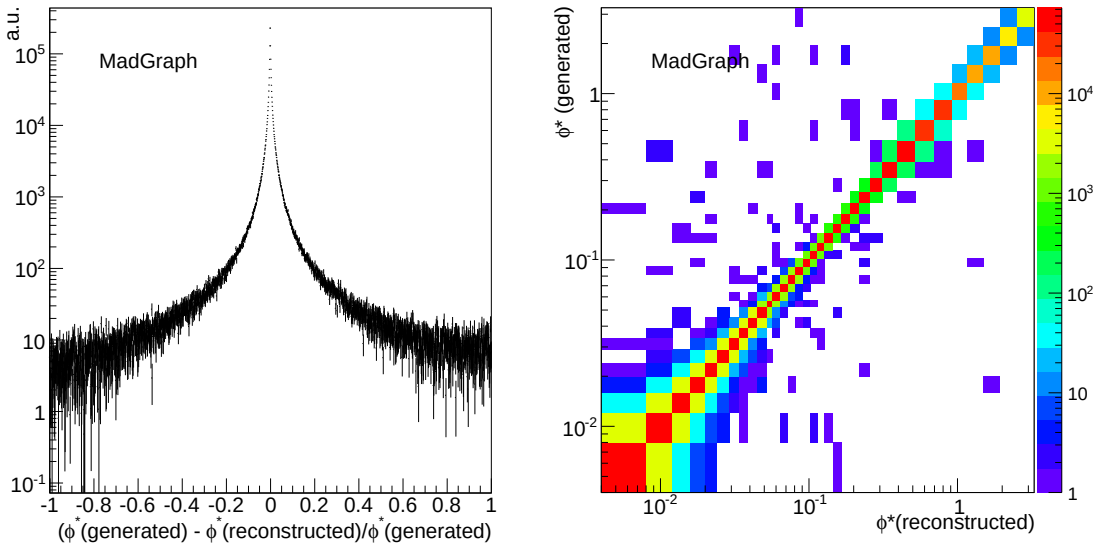
The finite resolution of CMS's angular position measurements leads to a finite resolution of the reconstructed ϕ^* distribution. Events which at the generator level would have ended up in a certain ϕ^* bin may instead migrate to one of the neighboring bins. The ϕ^* reconstruction resolution, defined as $\left(\phi^*_{\text{Gen}} - \phi^*_{\text{Reco}}\right)/\phi^*_{\text{Gen}}$, is shown in figure 6.14a.

MC events are used to create a response matrix that describes how generator level events are reconstructed in CMS. This matrix is shown in figure 6.14b. The amount of bin migration, as measured by the off-diagonal elements, is 5.5 %. Through unfolding this matrix is inverted, allowing us to transform the data to remove the undesired bin migration.

RooUnfold is used to perform the unfolding [88]. It implements Bayes' theorem as described in [89], and iteratively applies it to invert the response matrix. RooUnfold uses a limited number of iterations in order to terminate the algorithm before finding the true (but highly unstable) inverse. Four iterations are used in this analysis.

Closure Tests

This unfolding procedure was tested in several ways using the two signal MC samples discussed in section 5.2.2. First, the reconstructed ϕ^* distributions from the MadGraph and powheg sample were unfolded using their own generator level quantities



(a) (b)

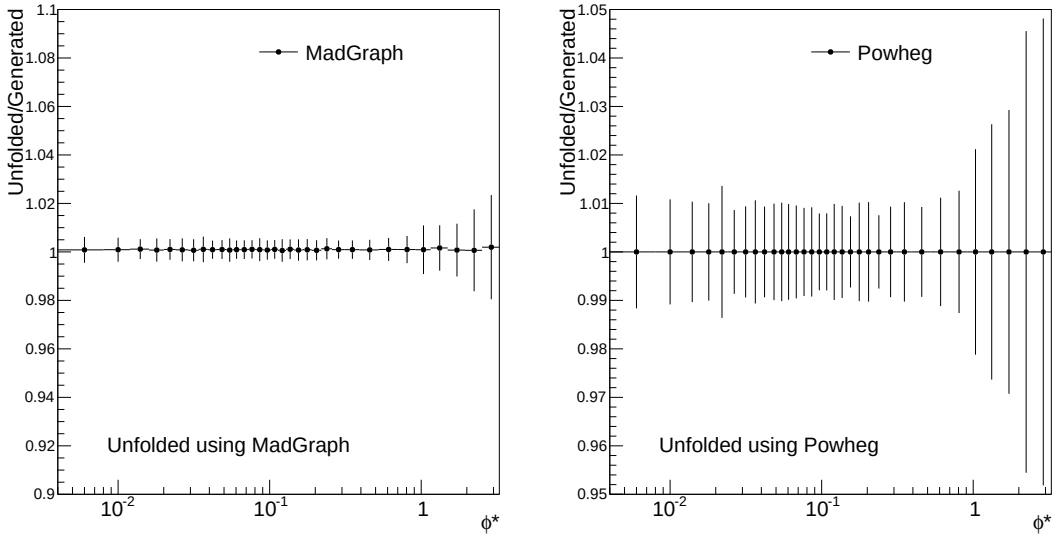
Figure 6.14: The ϕ^* reconstruction resolution (left) and bin migration (right) for events generated with MadGraph passing our event selection.

and compared to their generator level ϕ^* distributions. The result of this test for MadGraph is shown in figure 6.15a, while the result for powheg is shown in figure 6.15b.

Second, each MC signal sample was also divided into two independent halves. The reconstructed ϕ^* distribution in each half was unfolded using the generator level quantities from the other half. The two halves were then compared to each other. The result of this test for MadGraph is shown in figure 6.16a, while the result for powheg is shown in figure 6.16b.

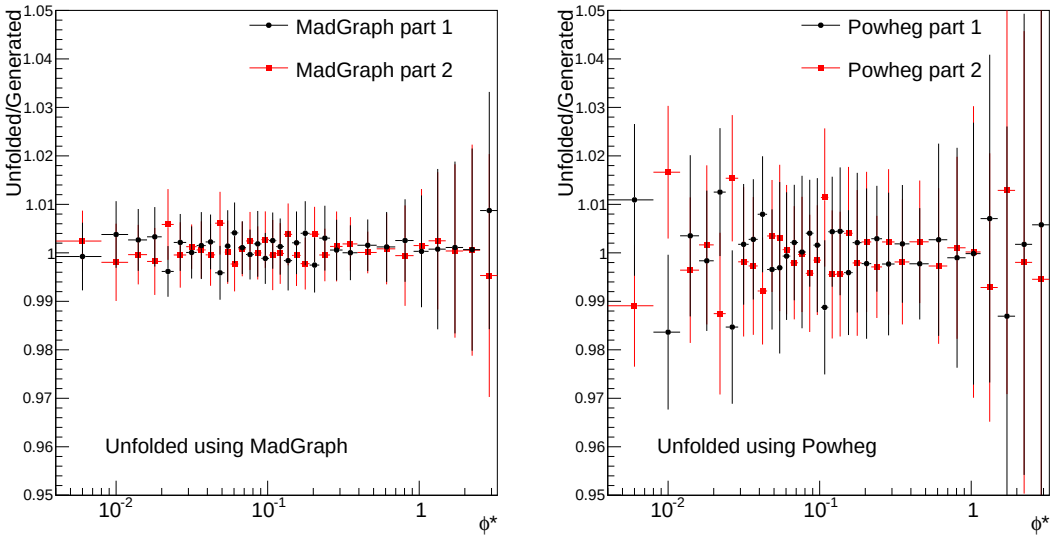
Third and finally, the reconstructed ϕ^* distribution from the MadGraph sample was unfolded with the generator level quantities from the powheg sample and vice versa. The result of unfolding powheg with MadGraph is shown in figure 6.17a, while the result of unfolding MadGraph with powheg is shown in figure 6.17b.

In most cases the results are consistent within the assigned statistical error bars. However, in the case of the MadGraph sample unfolded with powheg, there is disagreement between in the low ϕ^* bins. This is due to RooUnfold under estimating the uncertainties, as discussed in section 6.4.1.



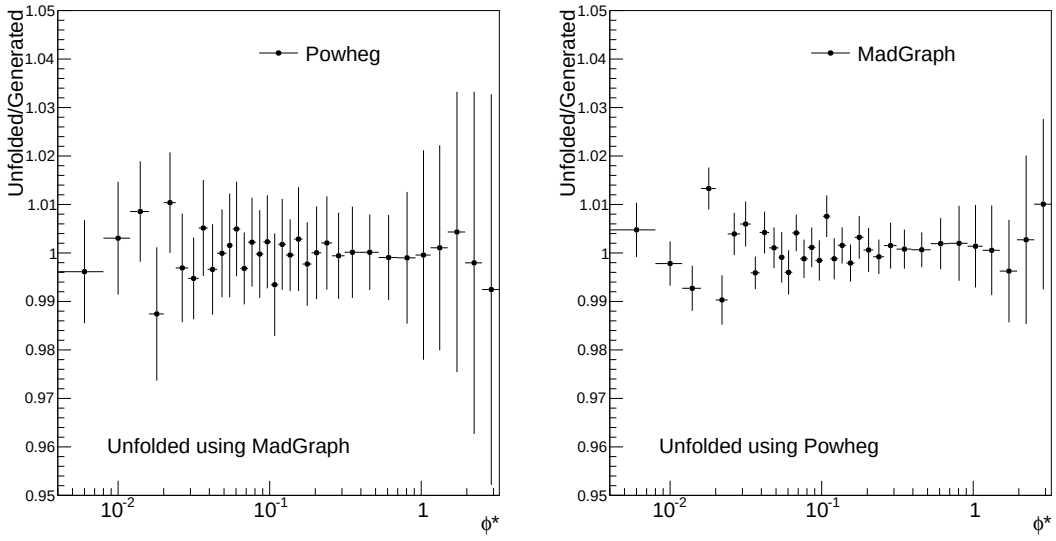
(a) (b)

Figure 6.15: The ratio between the unfolded reconstructed and generated ϕ^* distributions using the MadGraph sample (left) and powheg sample (right) for events passing our event selection. The unfolding was done with the same MadGraph (left) and powheg (right) sample.



(a) (b)

Figure 6.16: The ratio between the unfolded reconstructed and generated ϕ^* distributions in two (circular and square points) statistically independent samples events generated with MadGraph (left) and powheg (right) for events passing our event selection. Each reconstructed ϕ^* distribution was unfolded using the statistically independent sample produced by the same generator.



(a) (b)

Figure 6.17: On the left is shown the ratio between the unfolded reconstructed and generated ϕ^* distributions in the POWHEG sample unfolded with a flat MADGRAPH sample. On the right is show the same ratio but for MADGRAPH unfolded with POWHEG.

Statistical Uncertainties

The uncertainties shown in figures 6.15 to 6.17 are the statistical uncertainties returned by RooUnfold. These uncertainties only take the number of events in the reconstructed distribution into account. The uncertainties are correlated by the off-diagonal elements. While RooUnfold can return the full covariance matrix, we instead use a simpler approximation provided by RooUnfold where the uncertainties are based only on the diagonal elements. We tested the effects of this simplification on the uncertainties using two tests discussed below.

In the first test, each bin in the POWHEG ϕ^* distribution was regenerated from a Gaussian with the value of the original bin as the mean and the value of the statistical uncertainty on the bin from the original distribution as the standard deviation. Using this process, 500 ϕ^* distributions were generated, giving us 501 total distributions as the original distribution was also used. These distributions were then unfolded using the full MadGraph sample. The results are shown in figure 6.18 where the square points are the median value of the 501 distributions, and the error bars show the extent of the value in the central 68.2 % of the distributions. These are compared to the errors provided by RooUnfold. Both methods lead to identical results.

In the second test, 5000 and 50,000 randomly selected events from the POWHEG sample were used instead of the data, and 501 distributions were constructed in the same manner as above. The results of the test with 5000 events is shown in figure 6.19a, while the result of the test with 50,000 events is shown in figure 6.19b. As expected, when fewer reconstructed events are used in the unfolding, the uncertainties reported by RooUnfold are larger. In both tests, the simplified error handling produces the same results and so is used.

While RooUnfold properly handles the statistical uncertainty due to the number of events in the data being unfolded, it does not account for statistical uncertainty due to the number of generator level events used to create the bin migration matrix. This leads to an underestimation of the total uncertainty which is especially pronounced in cases where the number of events in the MC sample is smaller than the number of data events, as is the case in this analysis.

A toy MC method is used to propagate the uncertainty from the bin migration matrix to the unfolded distribution. In this method, 500 new bin migration matrices

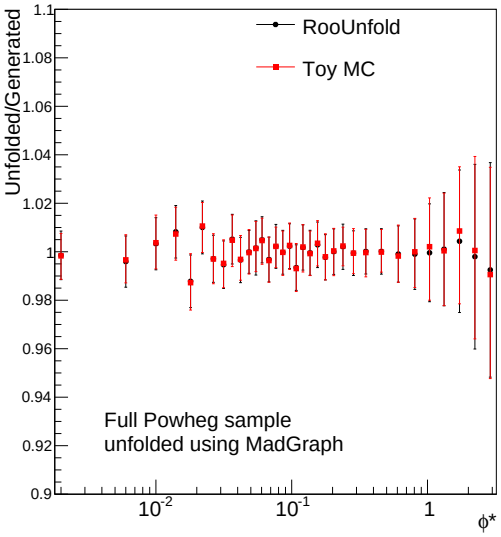
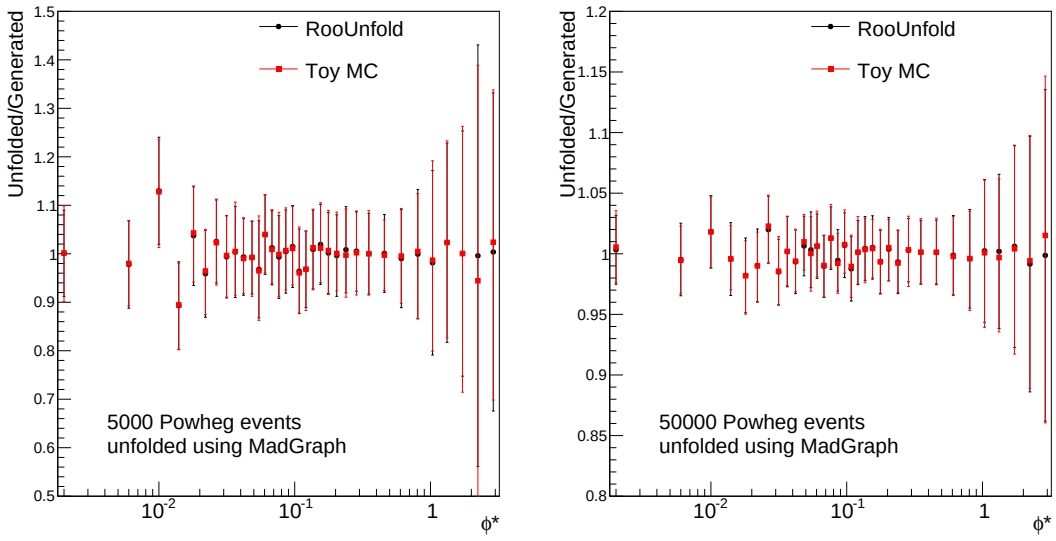


Figure 6.18: A comparison of the simplified RooUnfold uncertainties and toy MC based uncertainties. The ratio between the unfolded reconstructed and generated ϕ^* distributions for the powheg sample unfolded with MadGraph. The circular points show the results using the uncertainty from the RooUnfold package, while the square points show the results using 500 toy MC variations.



(a) (b)

Figure 6.19: A comparison of the behavior of RooUnfold uncertainties and toy MC based uncertainties as the number of events in the distribution to unfold changes. The ratio between the unfolded reconstructed and generated ϕ^* distributions for 5000 and 50,000 powheg events are shown on the left and right, respectively, unfolded with MadGraph. The circular points show the results using the uncertainty from the RooUnfold package, while the square points show the results using 500 toy MC variations.

are generated by randomly fluctuating each bin of the original matrix generated using 5000 events from the POWHEG sample. Bins with 0 events are left at 0. Bins with a small number of events—where the number of events in the bin divided by the statistical uncertainty is less than 5—are fluctuated using a Poisson distribution with the number of events in the original bin as the most probable value. All other bins are fluctuated using a Gaussian with mean equal to the number of events in the original bin and standard deviation equal to the statistical uncertainty on the original bin. The reconstructed ϕ^* distribution from the full MADGRAPH sample is then unfolded with each of the 501 bin migration matrices. The results are shown in figure 6.20. Figure 6.20a shows the uncertainties from ROOUNFOLD as compared to the uncertainties derived using the toy MC method. Figure 6.20c shows the extent of the ϕ^* bin values generated by the ensemble of toy MC bin migration matrices. Figure 6.20b shows the pull of the bins in figure 6.20a. Figure 6.20d shows the original bin migration matrix.

At large ϕ^* values the uncertainty calculated with the toy MC method goes to 0. This is because there are no off-diagonal bins in this region of the original bin migration matrix and so no off-diagonal bins are allowed to appear due to our fluctuations. This problem is greatly reduced as we use more generator level events to construct the matrix, as can be seen in figures 6.21 and 6.22.

The total statistical uncertainty due to the bin migration unfolding is the sum of the uncertainty reported by ROOUNFOLD due to the number of events in the data and the uncertainty calculated using the toy MC method in quadrature.

Systematic Uncertainties

The unfolding to correct for bin migration is dependent on the way in which the bin migration is simulated in MC. Differences in the ϕ^* distribution between MC and data can lead to systematic uncertainties on the final result. Two such potential differences between the MC and the data are considered. The first potential difference is in the shape of the ϕ^* distribution, and the second is in the resolution of the ϕ^* distribution.

One of the advantages of the Bayesian unfolding method used in this analysis is that it, unlike a simpler bin-by-bin correction, is theoretically insensitive to the distribution of ϕ^* at the generator level. In fact, D'Agostini recommends using a flat generated distribution as one easy method of ensuring that there are enough generated events in

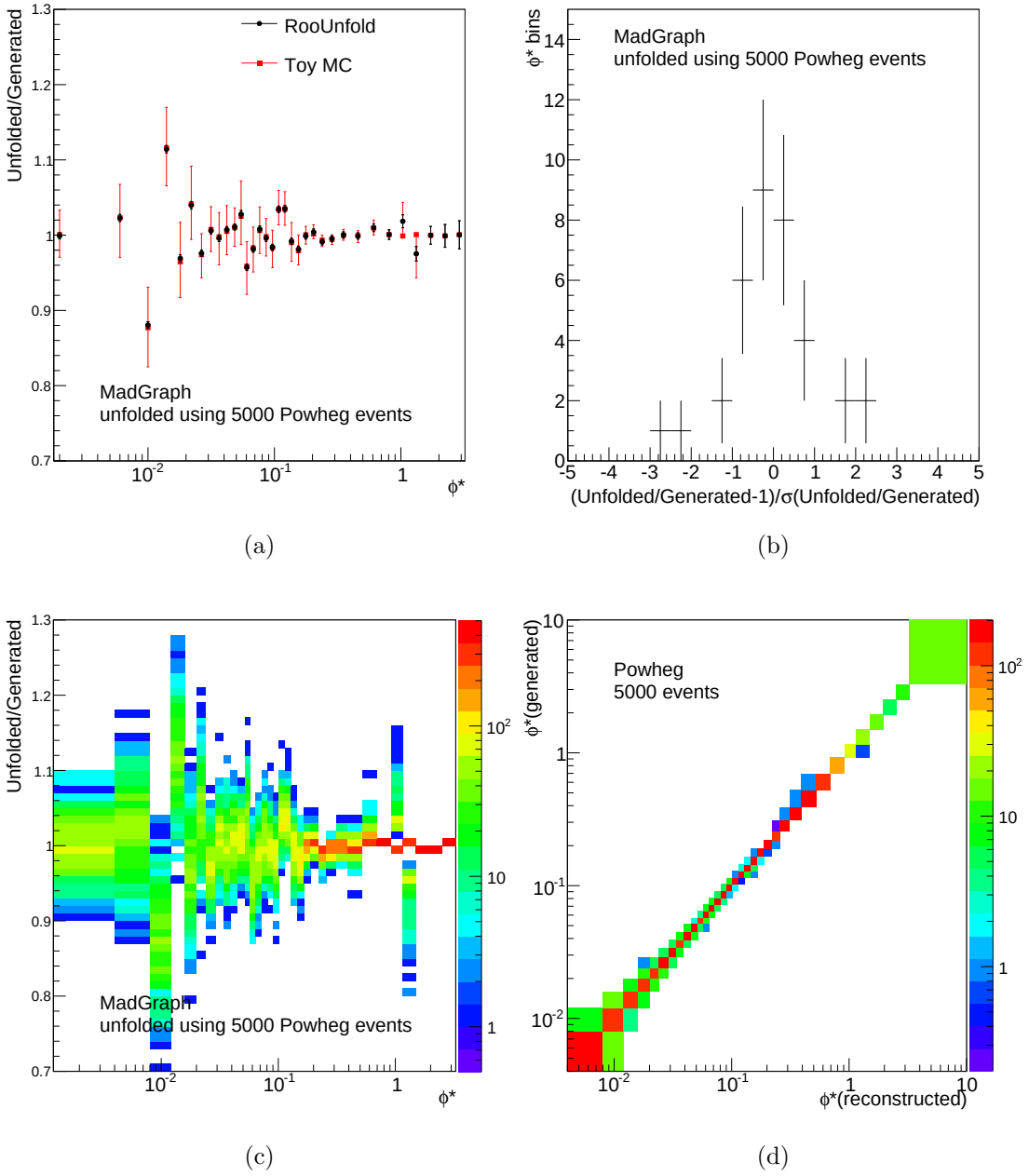


Figure 6.20: (a) The ratio between the unfolded reconstructed and generated ϕ^* distribution in MadGraph, unfolded with 50,000 powheg events. The circular points show the uncertainty reported by RooUnfold, and the square points show the result of 500 toys. (b) The pull of the bins of (a). (c) The 500 unfolded toy MC variations and the nominal distribution used to produce (a). (d) The bin migration matrix created with the 5000 powheg events.

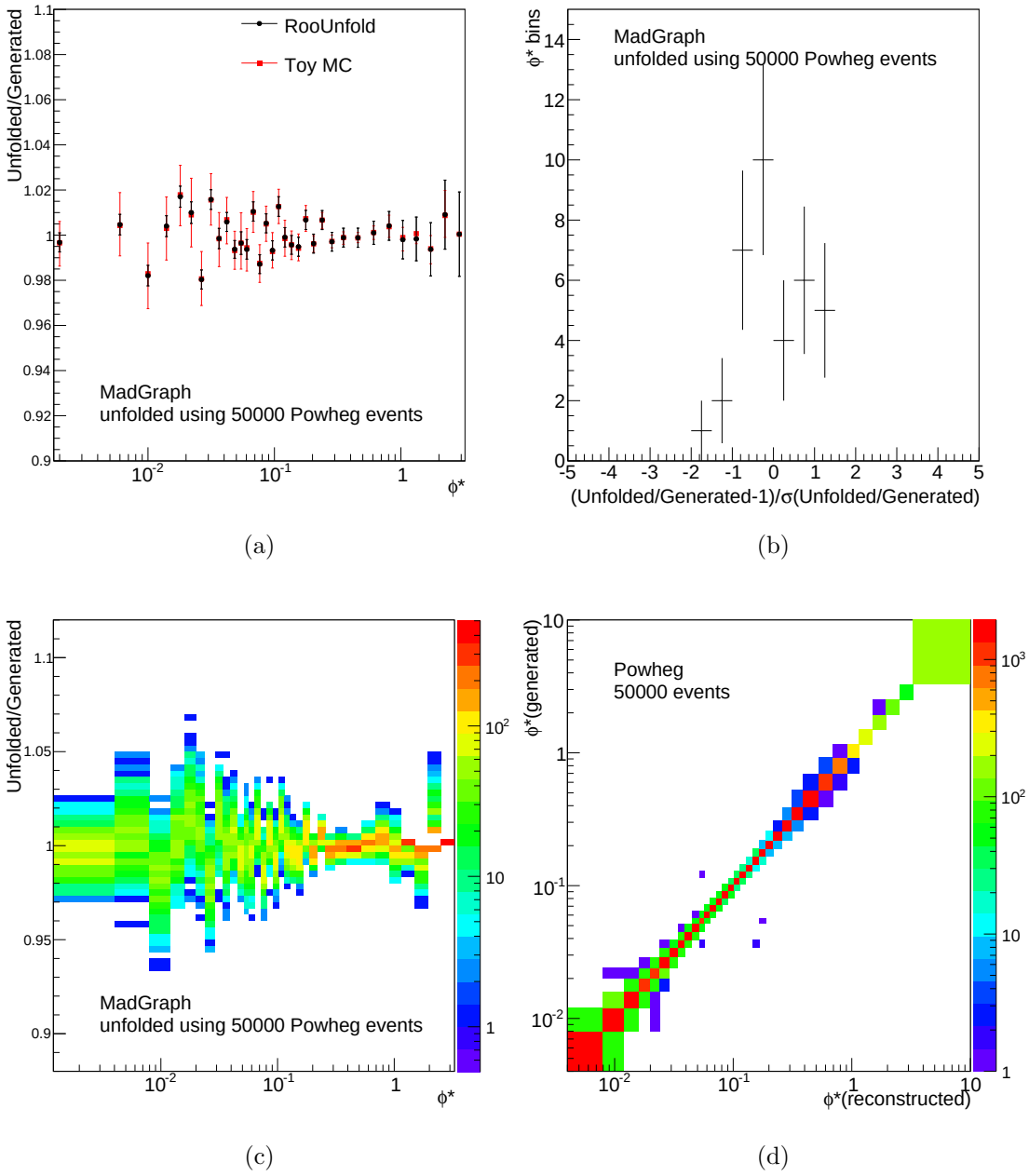


Figure 6.21: (a) The ratio between the unfolded reconstructed and generated ϕ^* distribution in MadGraph, unfolded with 50,000 powheg events. The circular points show the uncertainty reported by RooUnfold, and the square points show the result of 500 toys. (b) The pull of the bins of (a). (c) The 500 unfolded toy MC variations and the nominal distribution used to produce (a). (d) The bin migration matrix created with the 50,000 powheg events.

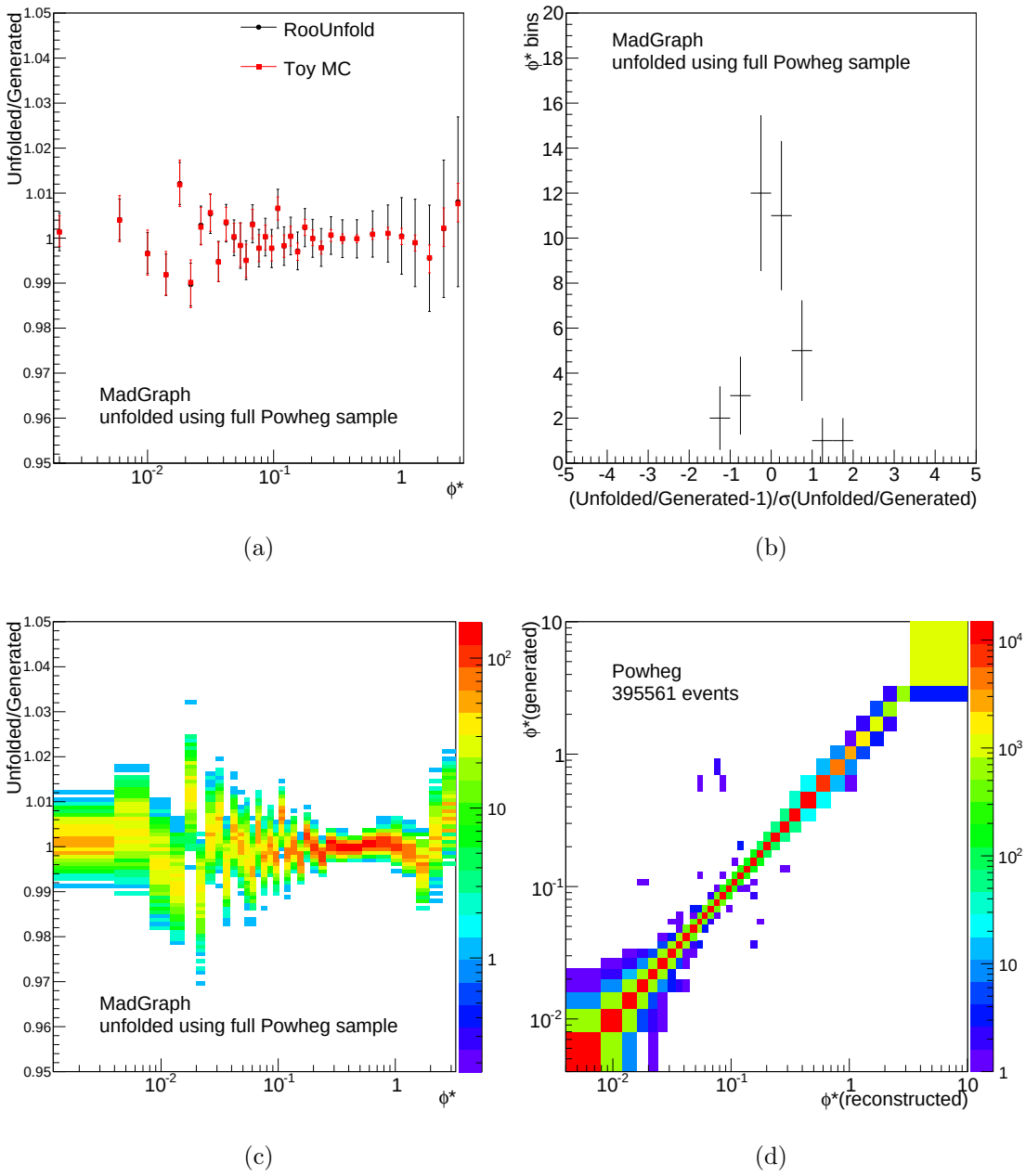
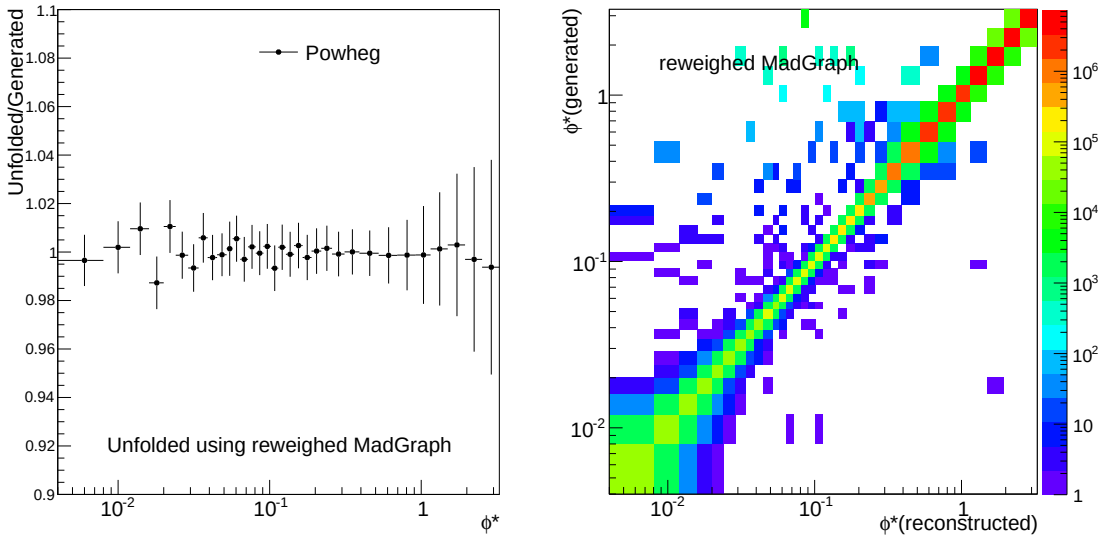


Figure 6.22: (a) The ratio between the unfolded reconstructed and generated ϕ^* distribution in MadGraph, unfolded with all powheg events. The circular points show the uncertainty reported by RooUnfold, and the square points show the result of 500 toys. (b) The pull of the bins of (a). (c) The 500 unfolded toy MC variations and the nominal distribution used to produce (a). (d) The bin migration matrix created with all the powheg events.

each bin [89]. We use this recommendation as the basis for a test of the systematic uncertainty. The ϕ^* distribution at the generator level in the MADGRAPH sample is inserted into a histogram with bin widths of $\phi^* = 0.011$. Each bin is weighted with a weight equal to the inverse number of events in the bin so that the distribution is flattened. The full POWHEG sample is then unfolded using this MADGRAPH distribution. The response matrix from this modified MADGRAPH distribution is show in figure 6.23b. The ratio of the unfolded POWHEG distribution over the generated distribution is shown in figure 6.23a. This ratio uses only the uncertainties provided by ROOUNFOLD and so underestimates the total uncertainty. No deviations from 1 are seen, and so no systematic uncertainty is assigned for the shape of the generator level ϕ^* distribution.



(a) (b)

Figure 6.23: The ratio between the unfolded reconstructed and generated ϕ^* distributions in the POWHEG sample unfolded with a flat MADGRAPH sample. The bin migration matrix for the reweighted MADGRAPH sample is shown on the right.

Differences in the resolution can arise due to differences in how the detector is simulated and how the detector responds in reality. Various corrections are applied to the MC in order to make it more closely match the data; these are discussed in section 5.3. The uncertainties from these corrections are discussed in section 7.1.

6.4.2 Efficiency Correction

Not every event which should be detected by CMS is. Some events are lost at each stage of event selection and reconstruction. These lost events must be corrected for in order to accurately compare our results to theory. We therefore apply corrections for the reconstruction efficiency, the electron ID efficiency, and the trigger efficiency.

The efficiency corrections are applied after the bin migration corrections. The correction factors are derived for each ϕ^* bin using the $Z \to e^+e^-$ MadGraph sample discussed in section 5.2.2. Two sets of events are selected. The first set, the "acceptance set", is created by applying the acceptance definition discussed in section 6.1 to the generator level quantities in the MC. The second set, the "final selection set", is created by selecting MC events by applying the full analysis selection to the reconstructed level quantities in the MC. The efficiency in each ϕ^* bin is calculated by counting the number of events in the "final selection set" in a bin, and dividing by the number of events in the same bin in the "acceptance set". This gives us an average efficiency composed of all the efficiencies of the events in the bin. The main advantage of using this average efficiency instead of correcting each event's efficiency individually is that any correlations between the various efficiencies are automatically taken into account in this process. The average efficiencies in each bin as calculated with both MadGraph and powheg are shown in figure 6.24.

The efficiency as a function of ϕ^* shows some interesting features. The first feature is that that the efficiency is approximately flat from $0 < \phi^* < 0.1$. This indicates that, in this region, any differences seen between the various generator ϕ^* distributions and the data are not efficiency-induced. The second feature is the increase in the efficiency at high ϕ^*. This is due to the fact that those Z bosons are boosted and so the electrons have additional energy leading to higher p_T. High p_T electrons are more efficiently triggered, reconstructed, and selected with ID and isolation cuts. The third feature is the dip in efficiency at $\phi^* \approx 0.2$. This dip is due to the fact that this ϕ^* region marks the transition between unboosted and boosted Z bosons. Electrons from these Z bosons can decay away from the direction of the boost and hence have lower momentum in the rest frame of the detector, giving them lower p_T. This puts them into the least efficiency bins in terms of trigger efficiency, reconstruction efficiency, and identification efficiency.

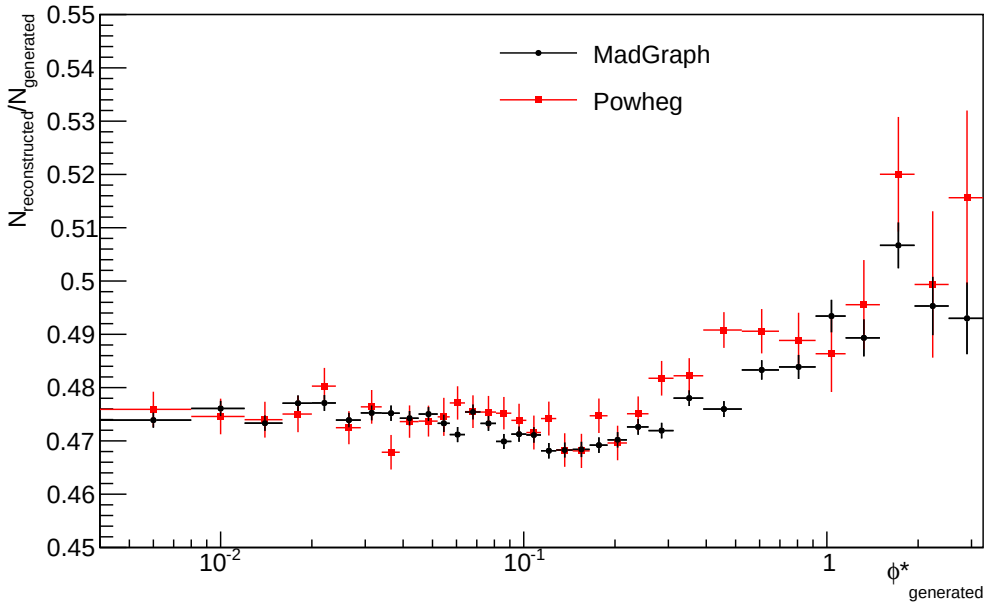


Figure 6.24: The ratio of $Z \to e^+e^-$ events generated with MadGraph (circles) and powheg (squares) which pass our event selection over the number of events in our fiducial region as a function of the generated ϕ^*.

Chapter 7

Analysis

7.1 Uncertainties

A measurement is only useful if it is known how well it was made. The following sections describe the sources of uncertainty in this measurement and estimate their impact on the final result.

7.1.1 Statistical Uncertainties

The statistical uncertainties due to the number of data events are propagated through the unfolding method using toy MC, as discussed in section 6.4.1. For the normalized cross section measurement, the statistical uncertainty is corrected for the uncertainty of the total cross section as follows:

$$\left(\varepsilon_i^{\text{Norm.}}\right)^2 = \left[\varepsilon_i^{\text{Abs.}} \cdot \left(\frac{1}{N} - \frac{n_i}{N^2}\right)\right]^2 + \sum_{i \neq j} \left(\varepsilon_j^{\text{Abs.}} \cdot \frac{n_i}{N^2}\right)^2 \tag{7.1}$$

where $\varepsilon_i^{\text{Norm.}}$ is the statistical uncertainty for normalized bin i, $\varepsilon_j^{\text{Abs.}}$ is the statistical uncertainty for bin j in the absolute distribution, n_i is the number of events in bin i in the absolute distribution, and N is the number of events given by the integral of the absolute distribution. The correction yielded by this formula is small.

The statistical uncertainty ranges from 0.26 % to 1.21 % for both the absolute and normalized cross section measurement. It is one of the dominant uncertainties for the

normalized cross section measurement.

7.1.2 Statistical Uncertainties from the Monte Carlo Samples

The MadGraph and powheg signal MC samples have fewer events which pass our final selection than there are events in the data. These samples are used to unfold the data and so their statistical uncertainty affects the final measurement. The effect of the statistical uncertainty on the bin migration matrix is propagated through to the final result via the use of toy MC as detailed in section 6.4.1. The uncertainty from using this method with the MadGraph sample is 0.1 % to 0.2 % for both the absolute and normalized cross section measurements.

In addition to affecting the unfolding, the low number of events in the MC affects the efficiency correction discussed in section 6.4.2 and shown in figure 6.24. These uncertainties vary from 0.3 % to 1.3 % for both the absolute and normalized cross section measurements. These uncertainties are calculated in the same manner for powheg, but the effect is larger as there are fewer events in that sample.

These two sources of uncertainty are measured separately. The final uncertainty due to using a specific signal MC sample is the sum in quadrature of both of these sources. This combined uncertainty is the dominant uncertainty for the normalized cross section measurement, having a slightly larger effect in each ϕ^* bin than the statistical uncertainty due to the data.

7.1.3 Luminosity Uncertainty

The integrated luminosity is measured at CMS using the occupancy in the pixel detector during minimum-bias events [90]. This luminosity measurement is calibrated by using van der Meer scans—a method to measure the beam size in which the two beams are offset and then "swept" across each other as that offset is reduced [91].

The integrated luminosity for the run period considered in this analysis is known to 2.6 %. This uncertainty is taken to be fully correlated bin-by-bin in ϕ^* for the absolute cross section measurement, where it is by far the dominant uncertainty. The luminosity cancels in the normalized cross section measurement and so the uncertainty only affects

the background subtraction. This effect is negligible compared to the uncertainty already present due to the background subtraction, which is discussed in section 7.1.7. The large uncertainty on the luminosity is the primary motivation behind making a normalized cross section measurement.

7.1.4 Pileup Uncertainty

As discussed in section 5.2.2, the high beam intensity at the LHC leads to multiple proton-proton interactions at each bunch crossing. This is modeled in MC by overlaying multiple simulated minimum-bias events on top of each simulated event. The distribution of pileup in MC is reweighted to match the data distribution based on the calculated instantaneous luminosity and the inelastic proton-proton cross section. The uncertainty due to this reweighting process is calculated by varying the inelastic cross section by plus and minus 5 %, recalculating the data distribution of pileup, and reweighting the MC samples to match this new distribution. The full analysis is then performed with these MC samples and the differences between the ϕ^* distributions is taken as a systematic uncertainty. The pileup uncertainty for the absolute cross section measurement ranges from 0.21 % to 0.58 %, while the uncertainty for the normalized cross section measurement varies from < 0.01 % to 0.64 %.

7.1.5 Trigger, Reconstruction, and Identification Scale Factors Uncertainty

Differences between the MC and data are corrected for using scale factors. Three different sets of scale factors are used to reweight the MC samples: trigger scale factors (discussed in section 5.3.2), reconstruction scale factors (discussed in section 5.3.3), and identification scale factors (discussed in section 5.3.4).

In all three cases, the uncertainties of the scale factors are propagated through to the final measurement using 500 toy MC variations. In this method, every toy is constructed by drawing each scale factor from a Gaussian probability distribution with its mean set to nominal value of the scale factor and its width set to the quadrature sum of the uncertainties of the scale factor. Each toy is then used to weight the MC samples used in this analysis, and the full analysis is performed with that newly weighted samples.

The uncertainty in the final result due to one of the three types of scale factors is taken to be defined by the central 68.2 % of results from the toys.

This procedure for propagating the uncertainty is performed independently for each of the three types of scale factors. The total uncertainty due to the scale factors is the sum in quadrature of the three results. For the absolute cross section measurement this uncertainty is approximately 0.4 %, while for the normalized cross section measurement it ranges from 0.02 % to 0.35 %.

7.1.6 p_{T} Scale Uncertainty

One of the advantages of the ϕ^* variable is that it is not computed using the momentum of the electrons and instead uses only the angles between them, which are generally better measured. This makes ϕ^* less sensitive to any potential problems with the p_{T} measurement of electrons.

However, the measurement of the p_{T} of the electrons is used to determine which events are included in our sample. Therefore, a shift in the p_{T} scale of the detector will either add or remove events that have electrons near the p_{T} selection requirement boundaries. To determine the uncertainty due to the p_{T} scale, we vary the p_{T} values of all of the electrons up and down by 0.3 %, which is a conservative estimate of the uncertainty of the p_{T} scale. The largest difference in each ϕ^* bin between the nominal result and the results with the modified p_{T} scale is taken as the uncertainty in that bin. The uncertainty due to the p_{T} scale for the absolute cross section measurement is 0.07 % to 0.17 %, while the uncertainty for the normalized cross section measurement varies from < 0.01 % to 0.10 %.

7.1.7 Background Subtraction Uncertainty

The background subtraction, which is discussed in section 6.3, deals with three separate categories of backgrounds. The uncertainty in each category is determined with a different method, and these uncertainties are added in quadrature to determine the total uncertainty due to the background subtraction.

The first category consists of the various backgrounds with two independent decay chains each of which can produce a lepton: $\mathrm{t\bar{t}}$, WW, DY $\rightarrow \tau^+\tau^-$, tW, and $\bar{\mathrm{t}}$W. The

contributions from these backgrounds are estimated by using an e–μ control sample as discussed in section 6.3.1. The uncertainty from the scale factors derived using this method are propagated through to the final result using 500 toy MC variations. For each variation, the scale factors are randomly drawn from a Gaussian distribution with mean equal to the nominal value of that scale factor and width equal to the uncertainty of the scale factor. Each toy is then used to weight the background MC samples which are then used to perform the background subtraction. The full analysis is then run with the newly background-subtracted data samples. The uncertainty due to the subtraction of this category of background for each bin in ϕ^* is defined by the spread of the central 68.2 % values obtained by the toys.

The second category consists of the backgrounds with a real Z boson: ZZ and WZ. For these samples we can not use the e–μ control sample so the uncertainty is calculated by taking a correlated 20 % uncertainty on the theoretical cross section and propagating this through to the final result.

The third and final category consists of the QCD multi-jet and W+jets backgrounds. The method of estimating this background is discussed in section 6.3.2. Instead of taking the uncertainties from the fit, which would not account for any systematics in the method used, a conservative 100 % uncertainty is assigned to this category.

The uncertainty due to the background subtraction for the absolute cross section measurement varies from 0.02 % to 0.64 %, and from 0.03 % to 0.59 % in the normalized cross section measurement.

7.1.8 PDF and Cross Section Uncertainties

As discussed in section 2.2.2, the kinematics of the Z boson depend on the internal composition and kinematics of the protons as they collide. The reconstructed ϕ^* distribution is therefore dependent on the PDFs used to generate the signal MC sample.

The uncertainty due to this choice of PDF is calculated following the recommendation of the PDF4LHC working group for the POWHEG MC signal sample [92]. PDFWeightProducer is used to reweight the POWHEG sample using the `CT10` PDF set. A total of 26 different pairs of weights are used by the tool to fully account for the uncertainty inherent in the PDF set; these weights are provided by the `CT10` collaboration specifically for this purpose. Each pair of weights consists of a variation of a PDF parameter,

with one weight corresponding to adjusting the parameter up and the other weight to down. Each of these weights are used to reweight the POWHEG sample, and the analysis is performed with this newly weighted sample. The uncertainty due to each weight is taken to be the difference with the nominal ϕ^* distribution. Two uncertainties are calculated: the one due to all of the upward parameter adjustments, and one to all the downward parameter adjustments, where in each case the uncertainties are added in quadrature. The largest of these two uncertainties for each ϕ^* bin is taken as the total uncertainty.

For the MADGRAPH MC signal sample, which is an LO sample generated with a LO PDF, the uncertainties can not be calculated in this manner. Instead, the PDF includes an uncertainty of the cross section which is used to scale the sample as calculated by FEWZ. Propagation of this cross section uncertainty through the analysis is achieved by both adding it to and subtracting it from the theoretical cross section and subsequently reweighting the MADGRAPH sample. The difference in the final ϕ^* distribution is taken as the uncertainty due to the FEWZ cross section. This uncertainty is the dominant uncertainty from the MADGRAPH sample.

7.1.9 Final State Radiation Uncertainties

The theory behind FSR, in which an electron radiates a photon, is discussed in section 2.3.3. These photons can affect the reconstruction of the Z, but this is taken into account during the unfolding as discussed in section 6.4. Hence, uncertainties in the modeling of FSR affect the unfolding and the final measurement.

The uncertainty is calculated using the FSRWEIGHTPRODUCER, which augments the PYTHIA QED calculation with exact $O(\alpha)$ and $\alpha(p_T^2)$ couplings and reweights the MC sample as if it had been produced with these calculations from the start. The effect of this reweighting on the final ϕ^* distribution is $\leq 0.34\,\%$ for the absolute cross section measurement and $\leq 0.03\,\%$ for the normalized cross section measurement.

7.1.10 Uncertainty from Four Vector Corrections

The m_{ee} distribution in the MADGRAPH signal MC sample does not precisely match the distribution in data, as seen in figure 6.4. This discrepancy remains even after applying

the various energy and momentum corrections to the electrons discussed in section 4.2. In order to determine the effect this has on the final measurement, the MadGraph signal MC sample is reweighted to remove this difference. The ratio between the nominal ϕ^* value and the value derived after this reweighting is shown in figure 7.1. The circular points show the ratio of the reconstructed ϕ^* distributions, while the square points show the ratio of the generated ϕ^* distributions. The errors are binomial. Most of the points are consistent with 1, and so no systematic uncertainty is assigned for this disagreement.

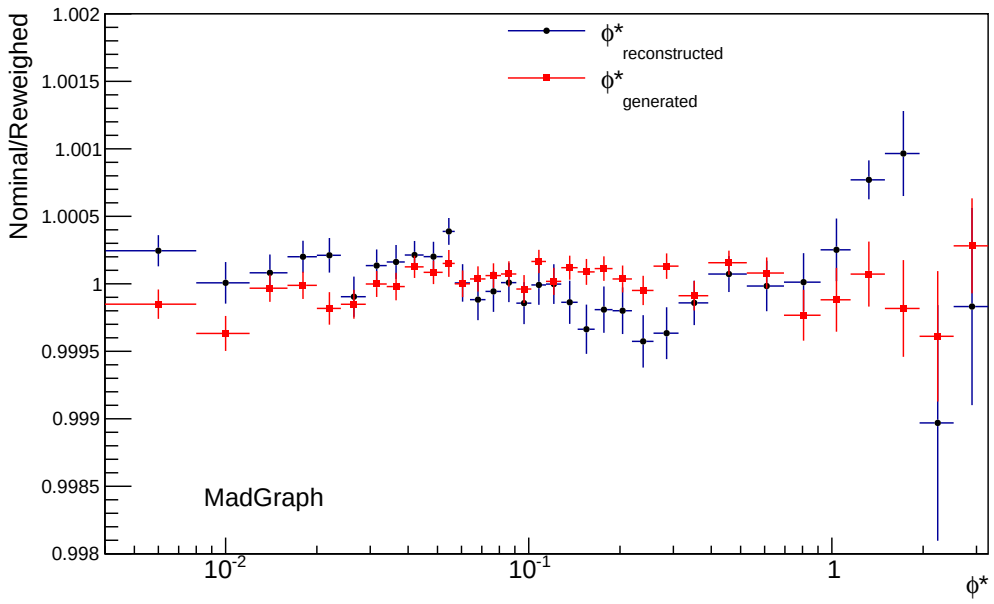


Figure 7.1: The ratio of ϕ^* in MadGraph before and after reweighting to remove the difference in the m_{ee} distribution between MC and data seen in figure 6.4. The circular points are the ratio in the reconstructed quantity, while the square points are the ratio in the generated quantity. The uncertainties are binomial.

Likewise, the Z Y distribution in MadGraph does not match the distribution in data, as seen in figure 6.8. The same reweighting procedure performed for the m_{ee} case above is also used here to force the distributions to agree. The ratio between the nominal ϕ^* value and the value derived after this reweighting is shown in figure 7.2. The circular points show the ratio of the reconstructed ϕ^* distributions, while the square points show the ratio of the generated ϕ^* distributions. The disagreement is on the

order of 0.05 % for most of the distribution, increasing to 0.5 % in the highest ϕ^* bins. Although this is larger than the effect seen in the m_{ee} reweighting, it is much smaller than the statistical uncertainties and so is not included in the final results.

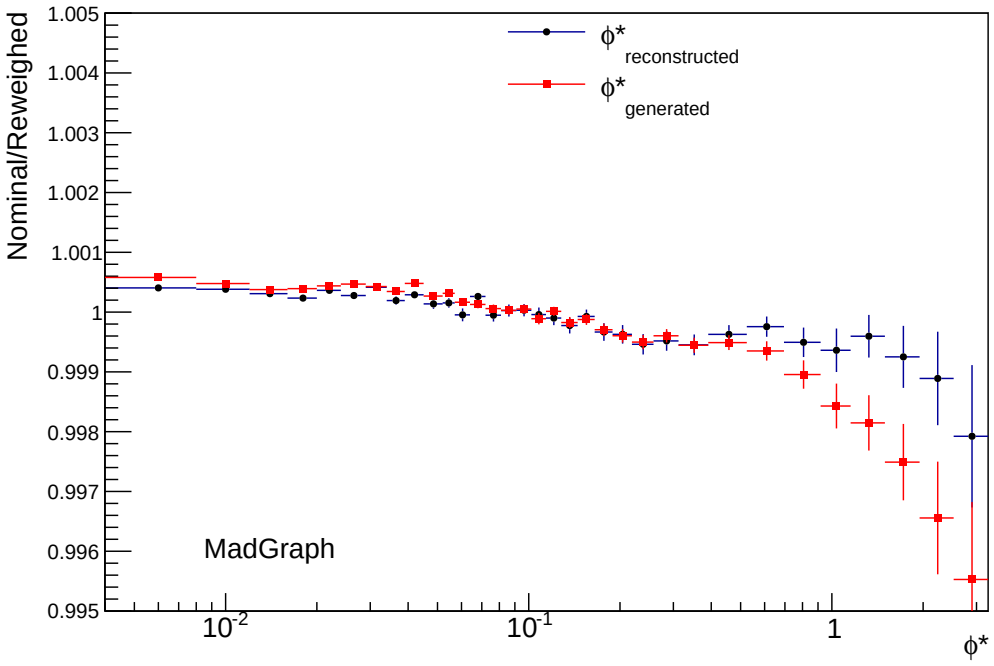


Figure 7.2: The ratio of ϕ^* in MadGraph before and after reweighting to remove the difference in the Y distribution between MC and data seen in figure 6.8. The circular points are the ratio in the reconstructed quantity, while the square points are the ratio in the generated quantity. The uncertainties are binomial.

7.2 Uncertainty Figures

The values of the various uncertainties in each ϕ^* bin are presented in the figures that follow. Figure 7.3 shows the uncertainties in the normalized ϕ^* cross section in data unfolded with MadGraph while figure 7.4 shows the uncertainties in the data unfolded with powheg+Pythia6 (`Z2star`). The uncertainties of the normalized ϕ^* cross section in MadGraph and powheg are shown in figures 7.5 and 7.6, respectively. All of these

values are presented in tables in appendix B. Figures showing the uncertainties for the absolute ϕ^* cross sections are presented in appendix A.

7.3 Results

Presented below are our normalized cross section measurements of the differential ϕ^* cross section for Z bosons decaying to an electron pair in the detector region defined in section 6.1. Absolute cross sections are presented in appendix A. Two sets of data are used to make the measurement, one unfolded with MadGraph and one unfolded with powheg + Pythia6 (`Z2star`); other than the different methods of unfolding, these datasets are identical. The data distributions of ϕ^* are compared to distributions from the MadGraph MC signal sample and the powheg + Pythia6 (`Z2star`) MC signal sample, as well as the two custom powheg samples discussed in section 5.2.3. The uncertainties of the two custom powheg samples are the same as for the centrally produced sample.

7.3.1 Normalized Differential Cross Section

The normalized differential cross section measurement using data unfolded with MadGraph is shown in figure 7.7 and given in tabular form in table 7.1. The lower plot in figure 7.7 is shown in more detail in figure 7.8. As the luminosity has canceled out in the normalization, the primary uncertainty of the data is now the statistical uncertainty from the MadGraph sample used to unfold it. The limit number of events in the MC samples is also the primary uncertainty on the MadGraph distribution and the three powheg distributions.

The normalized differential cross section measurement using data unfolded with powheg + Pythia6 (`Z2star`) is shown in figure 7.9 and given in tabular form in table 7.2. The lower plot in figure 7.9 is shown in more detail in figure 7.10. The primary uncertainty of the data is now the statistical uncertainty from the powheg sample used to unfold it.

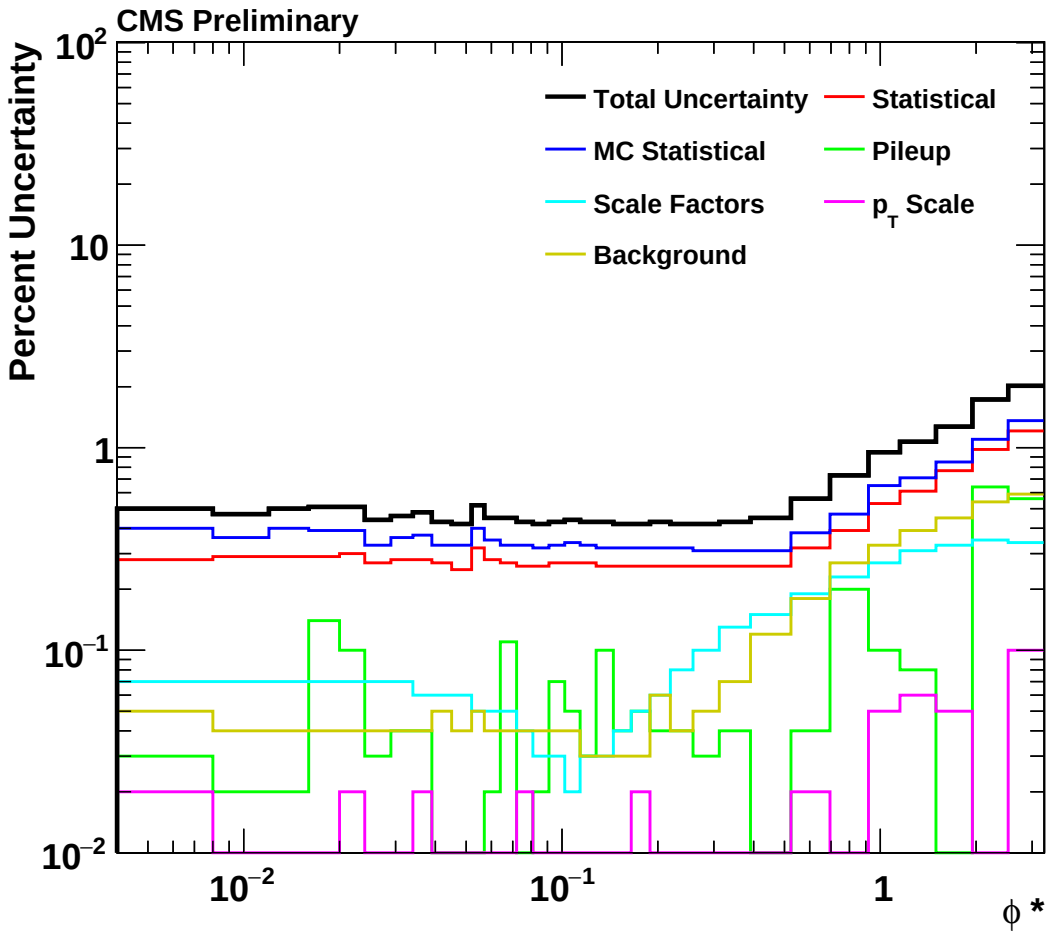


Figure 7.3: The uncertainties (in %) for the normalized cross section measurement made with data unfolded with MadGraph. The total value is the sum in quadrature of all the other values. These uncertainties are also presented in tabular form in table B.1.

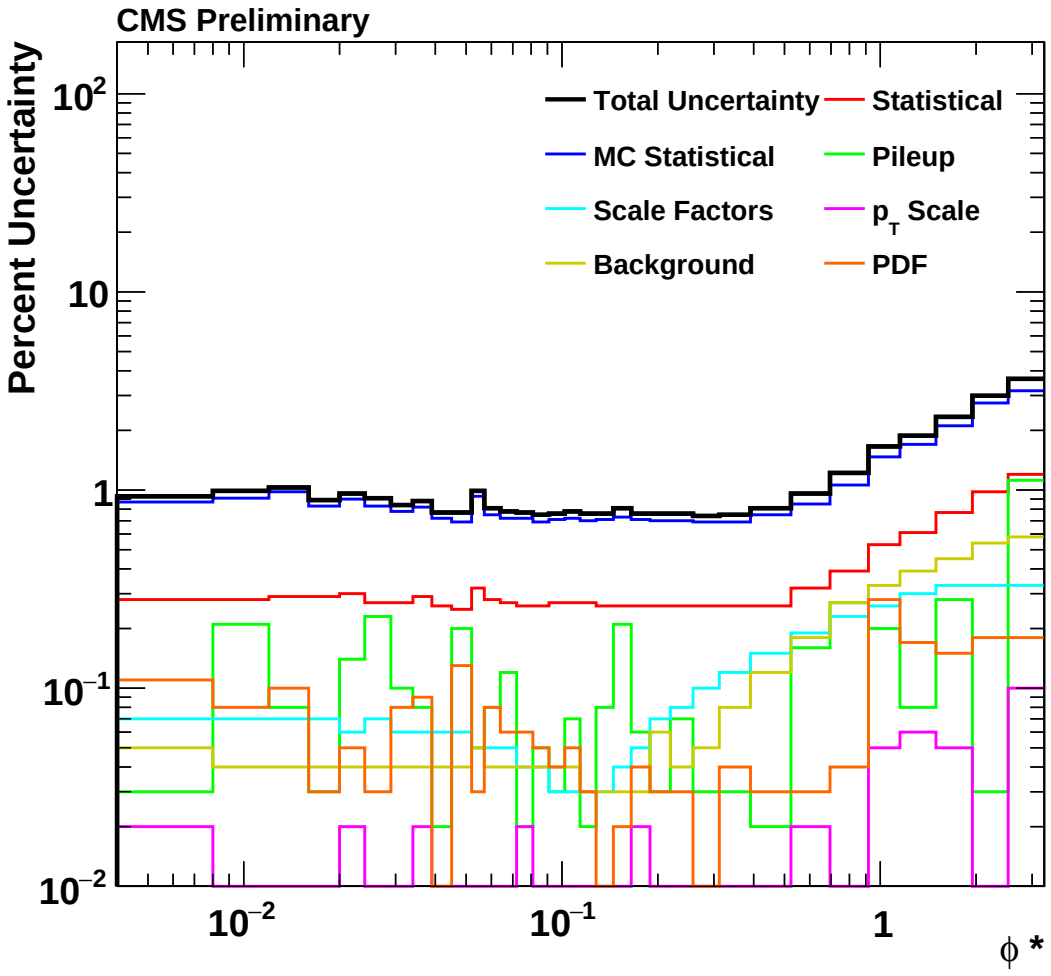


Figure 7.4: The uncertainties (in %) for the normalized cross section measurement made with data unfolded with POWHEG + PYTHIA6 (`Z2star`). The total value is the sum in quadrature of all the other values. These uncertainties are also presented in tabular form in table B.2.

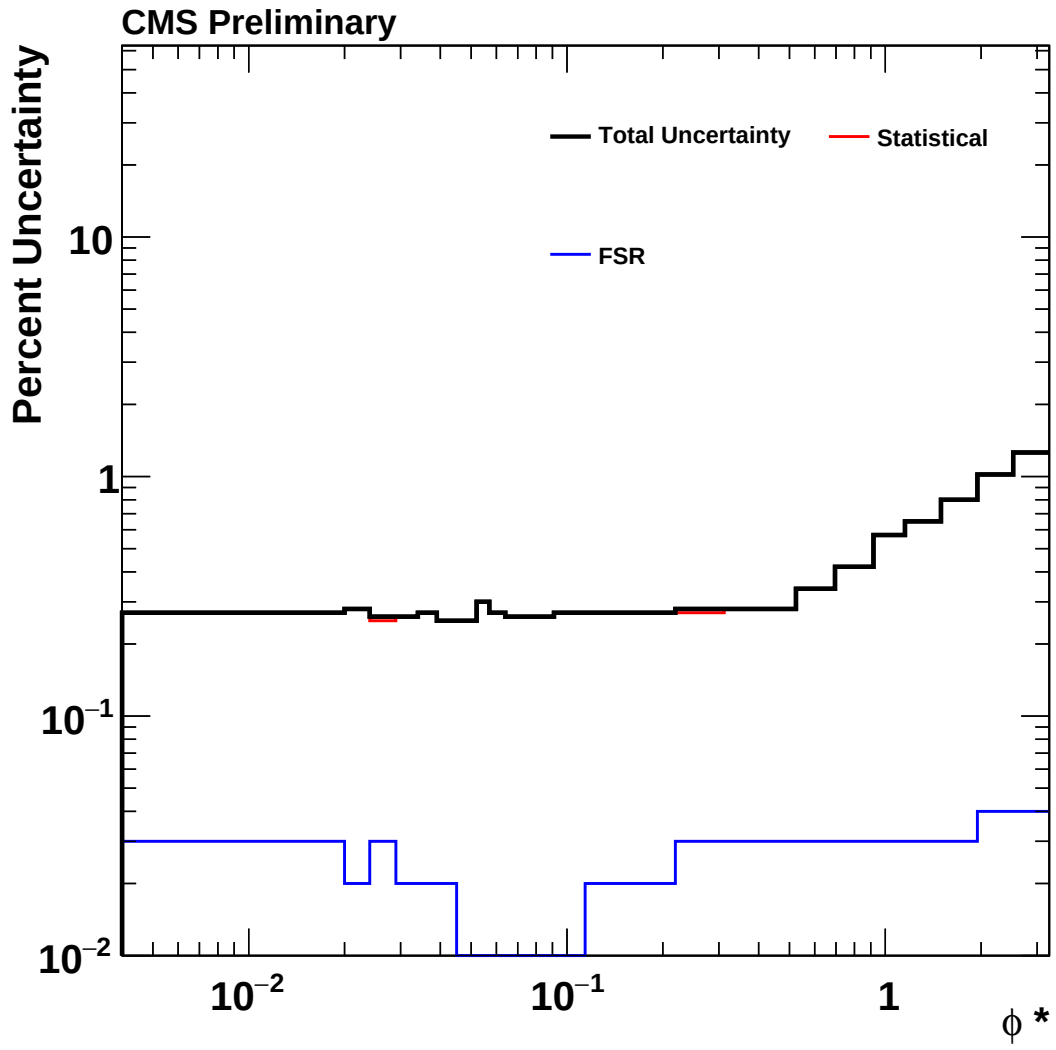


Figure 7.5: The uncertainties (in %) for the normalized cross section from the MadGraph MC sample. These uncertainties are also presented in tabular form in table B.3.

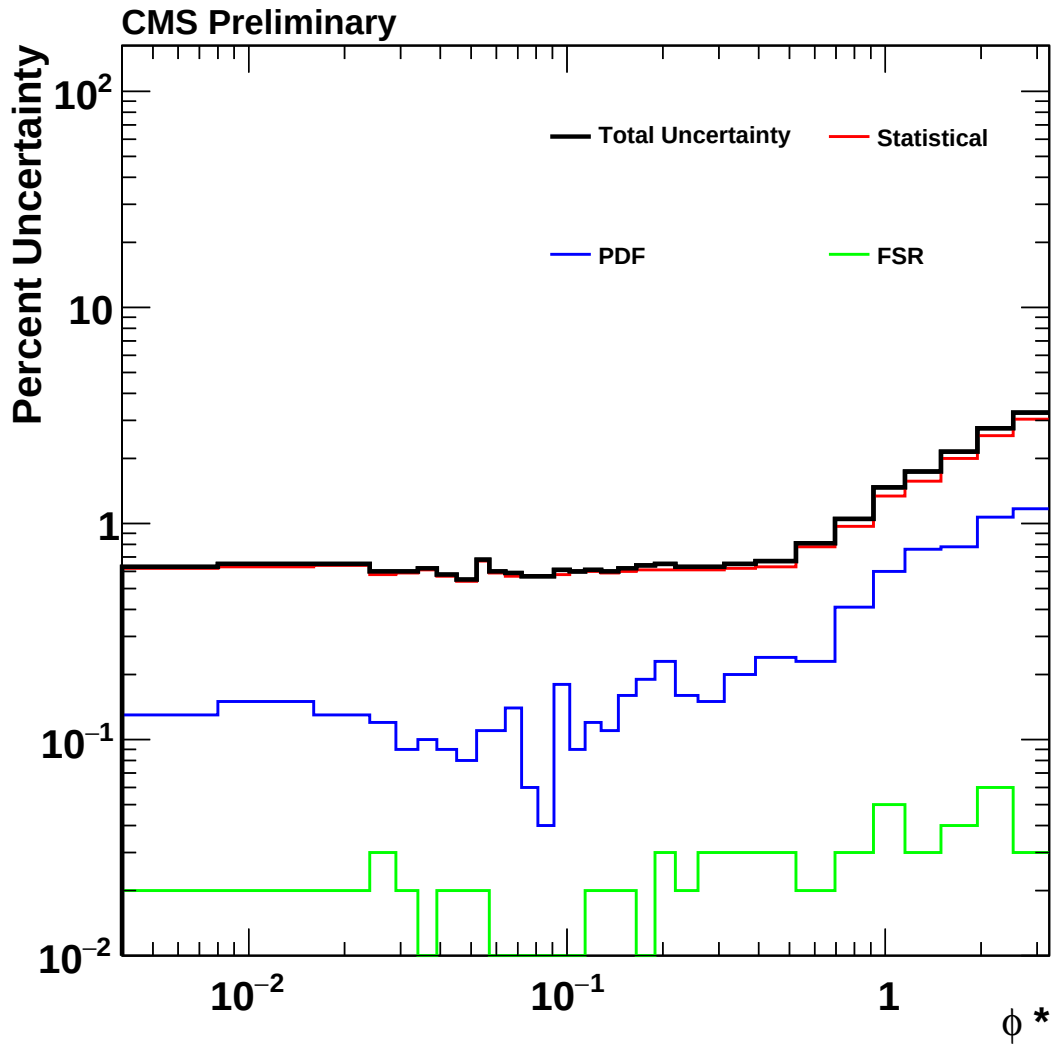


Figure 7.6: The uncertainties (in %) for the normalized cross section from the POWHEG MC sample. These uncertainties are also presented in tabular form in table B.4.

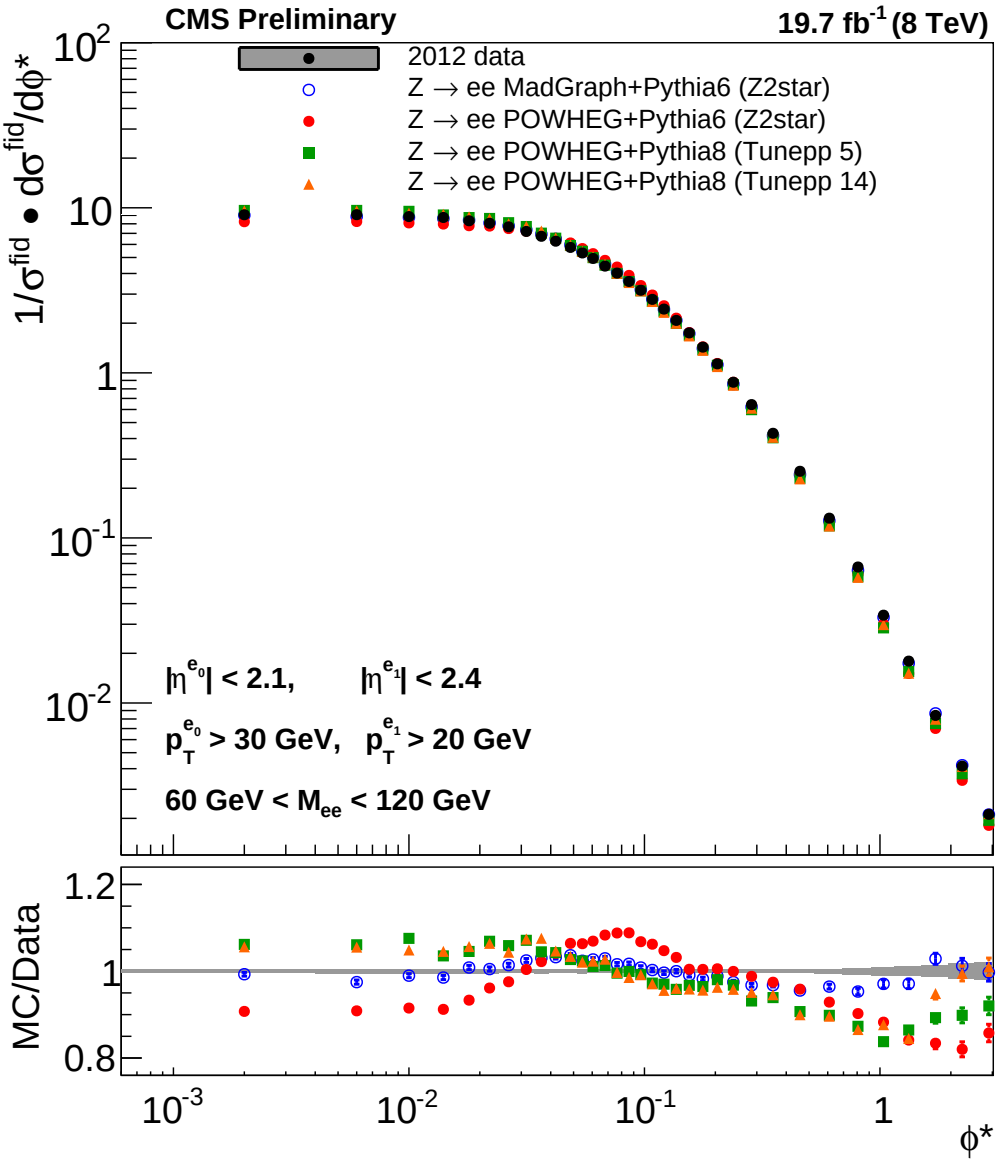


Figure 7.7: The normalized differential cross section with respects to ϕ^* for $Z \to e^+e^-$ events in our fiducial region from data unfolded with MadGraph, and the same distributions in MadGraph and three versions of powheg. A close up of the lower plot is shown in figure 7.8.

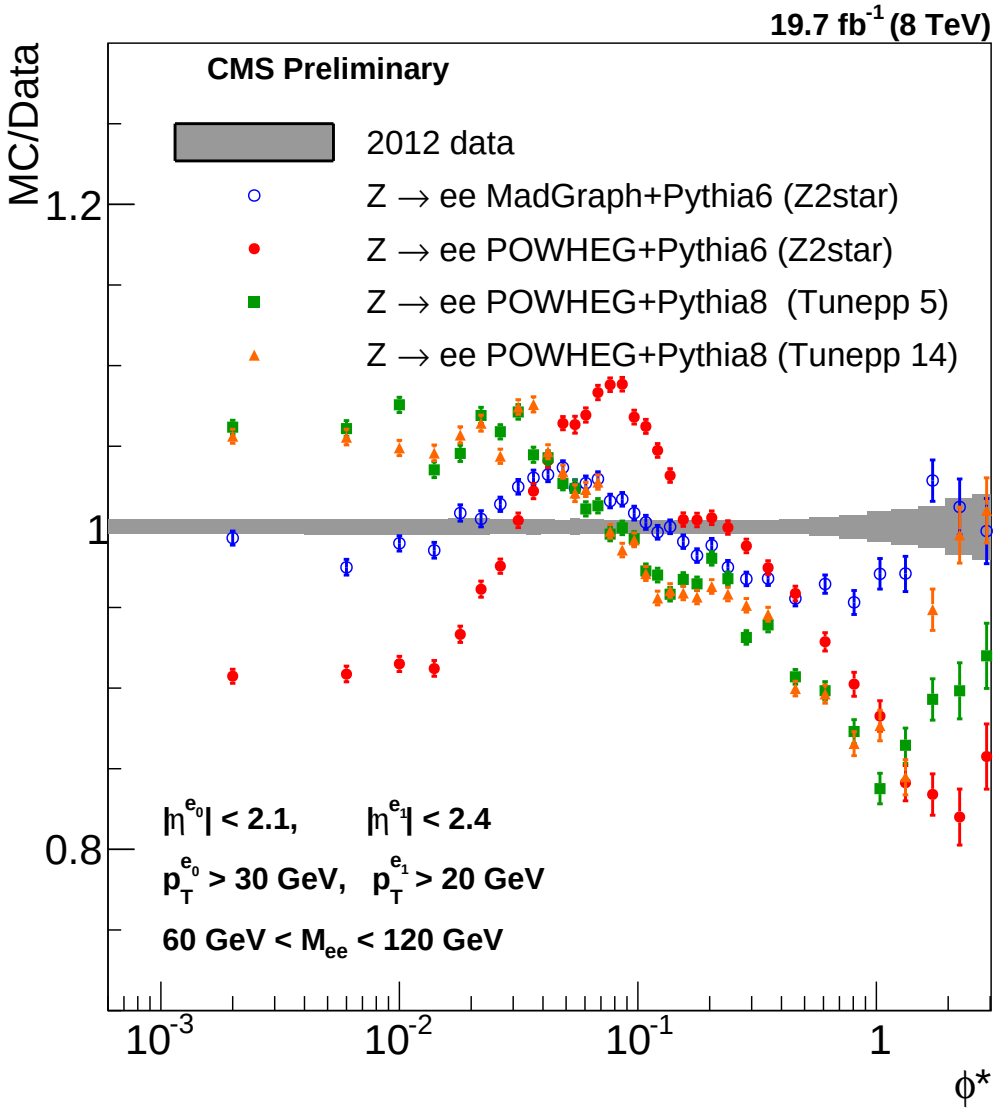


Figure 7.8: Close up of the ratio plot from figure 7.7 for the normalized cross section measurement unfolded with MadGraph. The error band indicates the uncertainty in the data, while the open circles show the ratio of MadGraph over data, the filled circles show the ratio of powheg + Pythia6 (`Z2star`) over data, the filled squares show the ratio of powheg + Pythia8 (`Tunepp5`) over data, and the filled triangles show the ratio of powheg + Pythia8 (`Tunepp14`).

ϕ^* range	Data	MadGraph	Z2star	Tunepp5	Tunepp14
0.000–0.004	9.08 ± 0.04	9.01 ± 0.02	8.24 ± 0.05	9.64 ± 0.06	9.59 ± 0.06
0.004–0.008	9.09 ± 0.05	8.86 ± 0.02	8.26 ± 0.05	9.64 ± 0.06	9.59 ± 0.06
0.008–0.012	8.85 ± 0.04	8.76 ± 0.02	8.10 ± 0.05	9.52 ± 0.06	9.28 ± 0.06
0.012–0.016	8.74 ± 0.04	8.61 ± 0.02	7.97 ± 0.05	9.05 ± 0.06	9.14 ± 0.06
0.016–0.020	8.34 ± 0.04	8.42 ± 0.02	7.79 ± 0.05	8.72 ± 0.06	8.82 ± 0.06
0.020–0.024	8.06 ± 0.04	8.10 ± 0.02	7.75 ± 0.05	8.62 ± 0.06	8.58 ± 0.06
0.024–0.029	7.67 ± 0.03	7.78 ± 0.02	7.48 ± 0.04	8.12 ± 0.05	8.01 ± 0.05
0.029–0.034	7.19 ± 0.03	7.37 ± 0.02	7.22 ± 0.04	7.70 ± 0.05	7.72 ± 0.05
0.034–0.039	6.72 ± 0.03	6.93 ± 0.02	6.87 ± 0.04	7.02 ± 0.04	7.23 ± 0.04
0.039–0.045	6.28 ± 0.03	6.48 ± 0.02	6.53 ± 0.04	6.55 ± 0.04	6.57 ± 0.04
0.045–0.052	5.75 ± 0.02	5.97 ± 0.01	6.12 ± 0.03	5.91 ± 0.03	5.95 ± 0.03
0.052–0.057	5.33 ± 0.03	5.45 ± 0.02	5.66 ± 0.04	5.45 ± 0.04	5.44 ± 0.04
0.057–0.064	4.93 ± 0.02	5.07 ± 0.01	5.28 ± 0.03	4.99 ± 0.03	5.05 ± 0.03
0.064–0.072	4.43 ± 0.02	4.57 ± 0.01	4.80 ± 0.03	4.49 ± 0.03	4.56 ± 0.03
0.072–0.081	4.02 ± 0.02	4.09 ± 0.01	4.38 ± 0.02	4.00 ± 0.02	4.01 ± 0.02
0.081–0.091	3.58 ± 0.02	3.640 ± 0.010	3.90 ± 0.02	3.58 ± 0.02	3.53 ± 0.02
0.091–0.102	3.17 ± 0.01	3.192 ± 0.009	3.38 ± 0.02	3.14 ± 0.02	3.14 ± 0.02
0.102–0.114	2.79 ± 0.01	2.795 ± 0.008	2.96 ± 0.02	2.71 ± 0.02	2.71 ± 0.02
0.114–0.128	2.43 ± 0.01	2.426 ± 0.007	2.55 ± 0.02	2.36 ± 0.01	2.33 ± 0.01
0.128–0.145	2.079 ± 0.009	2.079 ± 0.006	2.14 ± 0.01	1.99 ± 0.01	2.00 ± 0.01
0.145–0.165	1.746 ± 0.007	1.730 ± 0.005	1.75 ± 0.01	1.69 ± 0.01	1.67 ± 0.01
0.165–0.189	1.432 ± 0.006	1.406 ± 0.004	1.438 ± 0.009	1.381 ± 0.009	1.369 ± 0.009
0.189–0.219	1.134 ± 0.005	1.121 ± 0.003	1.140 ± 0.007	1.112 ± 0.007	1.092 ± 0.007
0.219–0.258	0.877 ± 0.004	0.855 ± 0.002	0.877 ± 0.006	0.849 ± 0.005	0.841 ± 0.005
0.258–0.312	0.642 ± 0.003	0.622 ± 0.002	0.635 ± 0.004	0.598 ± 0.004	0.611 ± 0.004
0.312–0.391	0.430 ± 0.002	0.417 ± 0.001	0.420 ± 0.003	0.404 ± 0.003	0.407 ± 0.003
0.391–0.524	0.253 ± 0.001	0.2421 ± 0.0007	0.243 ± 0.002	0.230 ± 0.002	0.228 ± 0.002
0.524–0.695	0.1317 ± 0.0007	0.1271 ± 0.0004	0.1224 ± 0.0010	0.1184 ± 0.0010	0.1181 ± 0.0010
0.695–0.918	0.0667 ± 0.0005	0.0635 ± 0.0003	0.0602 ± 0.0006	0.0582 ± 0.0006	0.0577 ± 0.0006
0.918–1.153	0.0340 ± 0.0003	0.0330 ± 0.0002	0.0300 ± 0.0004	0.0285 ± 0.0004	0.0298 ± 0.0004
1.153–1.496	0.0179 ± 0.0002	0.0174 ± 0.0001	0.0151 ± 0.0003	0.0155 ± 0.0003	0.0152 ± 0.0003
1.496–1.947	0.0084 ± 0.0001	0.00865 ± 0.00007	0.0070 ± 0.0002	0.0075 ± 0.0002	0.0080 ± 0.0002
1.947–2.522	0.00415 ± 0.00007	0.00420 ± 0.00004	0.00340 ± 0.00009	0.0037 ± 0.0001	0.0041 ± 0.0001
2.522–3.277	0.00212 ± 0.00004	0.00211 ± 0.00003	0.00181 ± 0.00006	0.00195 ± 0.00006	0.00214 ± 0.00007

Table 7.1: The normalized differential cross section in pb with respects to ϕ^* for $Z \to e^+e^-$ events in our fiducial region from data unfolded with MadGraph, and the same distributions in MadGraph and three versions of powheg. The column `Z2star` is the distribution from powheg + Pythia6 (`Z2star`), `Tunepp5` is powheg + Pythia8 (`Tunepp5`), and `Tunepp14` is powheg + Pythia8 (`Tunepp14`). These results are shown graphically in figure 7.7.

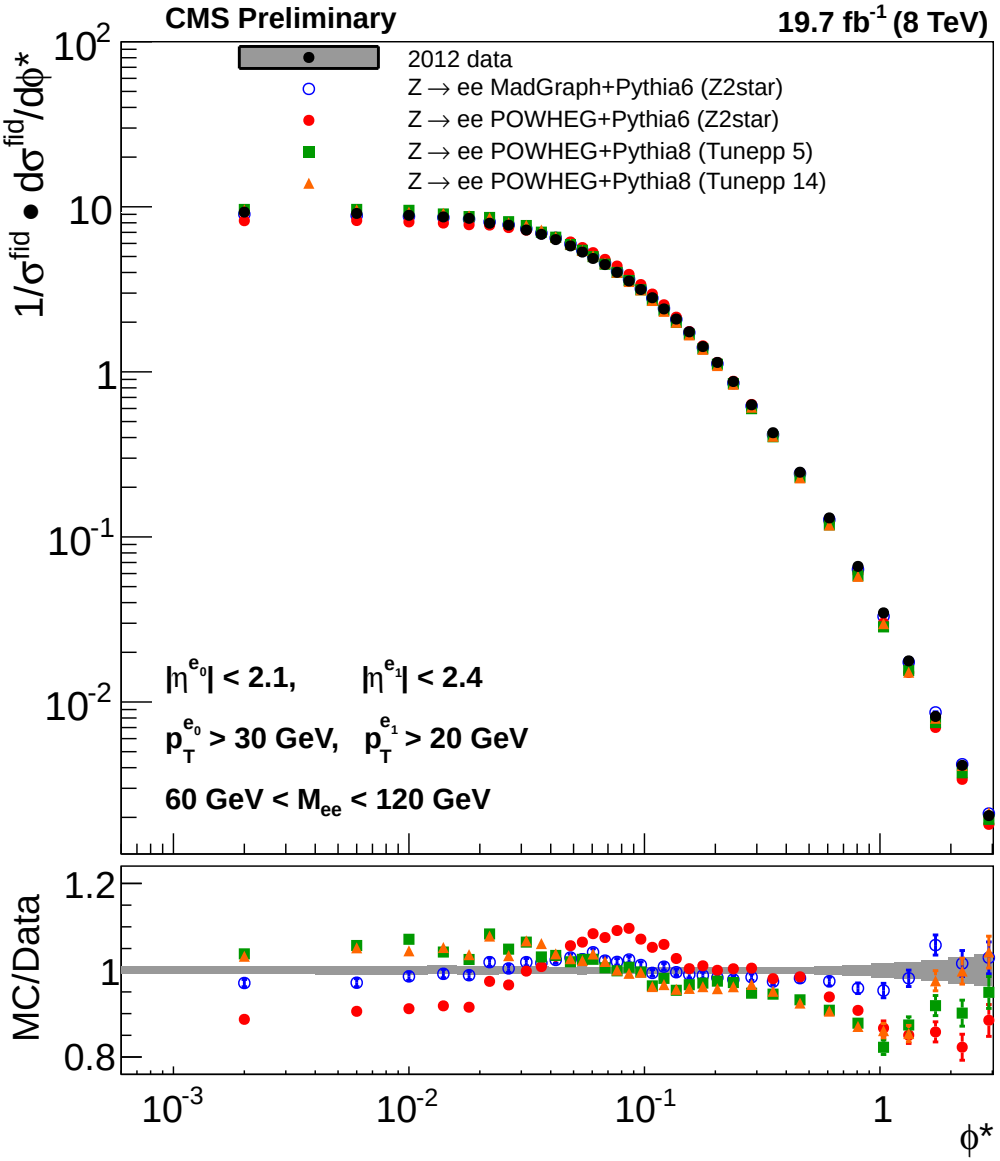


Figure 7.9: The normalized differential cross section with respects to ϕ^* for $Z \to e^+e^-$ events in our fiducial region from data unfolded with POWHEG + PYTHIA6 (`Z2star`), and the same distributions in MADGRAPH and three versions of POWHEG. A close up of the lower plot is shown in figure 7.10.

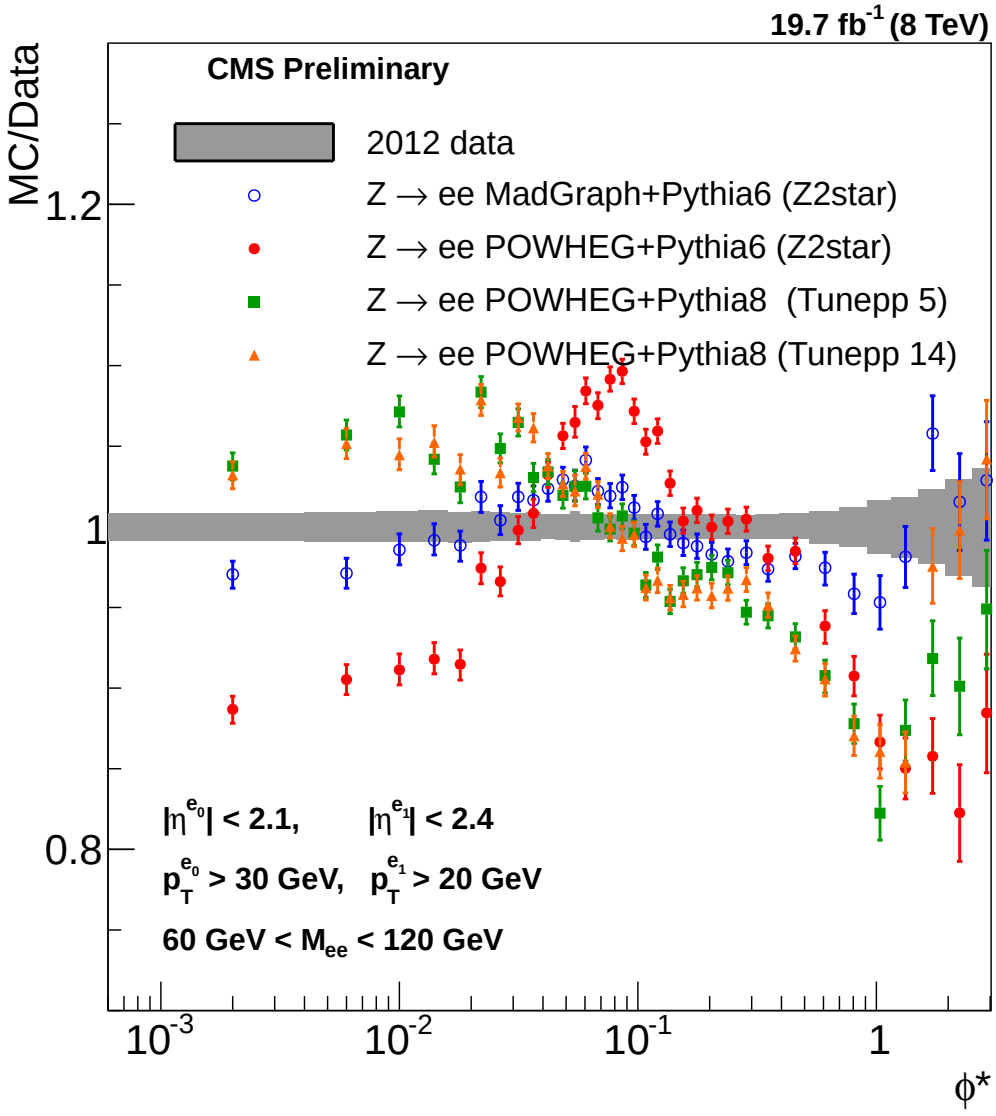


Figure 7.10: Close up of the ratio plot from figure 7.7 for the normalized cross section measurement unfolded with POWHEG + PYTHIA6 (`Z2star`). The error band indicates the uncertainty in the data, while the open circles show the ratio of MADGRAPH over data, the filled circles show the ratio of POWHEG + PYTHIA6 (`Z2star`) over data, the filled squares show the ratio of POWHEG + PYTHIA8 (`Tunepp5`) over data, and the filled triangles show the ratio of POWHEG + PYTHIA8 (`Tunepp14`).

ϕ^* range	Data	MadGraph	Z2star	Tunepp5	Tunepp14
0.000–0.004	9.29 ± 0.08	9.01 ± 0.02	8.24 ± 0.05	9.64 ± 0.06	9.59 ± 0.06
0.004–0.008	9.12 ± 0.08	8.86 ± 0.02	8.26 ± 0.05	9.64 ± 0.06	9.59 ± 0.06
0.008–0.012	8.89 ± 0.09	8.76 ± 0.02	8.10 ± 0.05	9.52 ± 0.06	9.28 ± 0.06
0.012–0.016	8.68 ± 0.09	8.61 ± 0.02	7.97 ± 0.05	9.05 ± 0.06	9.14 ± 0.06
0.016–0.020	8.51 ± 0.08	8.42 ± 0.02	7.79 ± 0.05	8.72 ± 0.06	8.82 ± 0.06
0.020–0.024	7.95 ± 0.08	8.10 ± 0.02	7.75 ± 0.05	8.62 ± 0.06	8.58 ± 0.06
0.024–0.029	7.75 ± 0.07	7.78 ± 0.02	7.48 ± 0.04	8.12 ± 0.05	8.01 ± 0.05
0.029–0.034	7.23 ± 0.06	7.37 ± 0.02	7.22 ± 0.04	7.70 ± 0.05	7.72 ± 0.05
0.034–0.039	6.81 ± 0.06	6.93 ± 0.02	6.87 ± 0.04	7.02 ± 0.04	7.23 ± 0.04
0.039–0.045	6.33 ± 0.05	6.48 ± 0.02	6.53 ± 0.04	6.55 ± 0.04	6.57 ± 0.04
0.045–0.052	5.80 ± 0.04	5.97 ± 0.01	6.12 ± 0.03	5.91 ± 0.03	5.95 ± 0.03
0.052–0.057	5.32 ± 0.05	5.45 ± 0.02	5.66 ± 0.04	5.45 ± 0.04	5.44 ± 0.04
0.057–0.064	4.87 ± 0.04	5.07 ± 0.01	5.28 ± 0.03	4.99 ± 0.03	5.05 ± 0.03
0.064–0.072	4.47 ± 0.04	4.57 ± 0.01	4.80 ± 0.03	4.49 ± 0.03	4.56 ± 0.03
0.072–0.081	4.01 ± 0.03	4.09 ± 0.01	4.38 ± 0.02	4.00 ± 0.02	4.01 ± 0.02
0.081–0.091	3.55 ± 0.03	3.640 ± 0.010	3.90 ± 0.02	3.58 ± 0.02	3.53 ± 0.02
0.091–0.102	3.15 ± 0.02	3.192 ± 0.009	3.38 ± 0.02	3.14 ± 0.02	3.14 ± 0.02
0.102–0.114	2.81 ± 0.02	2.795 ± 0.008	2.96 ± 0.02	2.71 ± 0.02	2.71 ± 0.02
0.114–0.128	2.41 ± 0.02	2.426 ± 0.007	2.55 ± 0.02	2.36 ± 0.01	2.33 ± 0.01
0.128–0.145	2.09 ± 0.02	2.079 ± 0.006	2.14 ± 0.01	1.99 ± 0.01	2.00 ± 0.01
0.145–0.165	1.75 ± 0.01	1.730 ± 0.005	1.75 ± 0.01	1.69 ± 0.01	1.67 ± 0.01
0.165–0.189	1.42 ± 0.01	1.406 ± 0.004	1.438 ± 0.009	1.381 ± 0.009	1.369 ± 0.009
0.189–0.219	1.141 ± 0.009	1.121 ± 0.003	1.140 ± 0.007	1.112 ± 0.007	1.092 ± 0.007
0.219–0.258	0.874 ± 0.007	0.855 ± 0.002	0.877 ± 0.006	0.849 ± 0.005	0.841 ± 0.005
0.258–0.312	0.632 ± 0.005	0.622 ± 0.002	0.635 ± 0.004	0.598 ± 0.004	0.611 ± 0.004
0.312–0.391	0.428 ± 0.003	0.417 ± 0.001	0.420 ± 0.003	0.404 ± 0.003	0.407 ± 0.003
0.391–0.524	0.247 ± 0.002	0.2421 ± 0.0007	0.243 ± 0.002	0.230 ± 0.002	0.228 ± 0.002
0.524–0.695	0.130 ± 0.001	0.1271 ± 0.0004	0.1224 ± 0.0010	0.1184 ± 0.0010	0.1181 ± 0.0010
0.695–0.918	0.0663 ± 0.0008	0.0635 ± 0.0003	0.0602 ± 0.0006	0.0582 ± 0.0006	0.0577 ± 0.0006
0.918–1.153	0.0347 ± 0.0006	0.0330 ± 0.0002	0.0300 ± 0.0004	0.0285 ± 0.0004	0.0298 ± 0.0004
1.153–1.496	0.0177 ± 0.0003	0.0174 ± 0.0001	0.0151 ± 0.0003	0.0155 ± 0.0003	0.0152 ± 0.0003
1.496–1.947	0.0082 ± 0.0002	0.00865 ± 0.00007	0.0070 ± 0.0002	0.0075 ± 0.0002	0.0080 ± 0.0002
1.947–2.522	0.0041 ± 0.0001	0.00420 ± 0.00004	0.00340 ± 0.00009	0.0037 ± 0.0001	0.0041 ± 0.0001
2.522–3.277	0.00205 ± 0.00007	0.00211 ± 0.00003	0.00181 ± 0.00006	0.00195 ± 0.00006	0.00214 ± 0.00007

Table 7.2: The normalized differential cross section in pb with respects to ϕ^* for $Z \to e^+e^-$ events in our fiducial region from data unfolded with powheg+Pythia6 (`Z2star`), and the same distributions in MadGraph and three versions of powheg. The column `Z2star` is the distribution from powheg + Pythia6 (`Z2star`), `Tunepp5` is powheg + Pythia8 (`Tunepp5`), and `Tunepp14` is powheg + Pythia8 (`Tunepp14`). These results are shown graphically in figure 7.9.

7.4 Discussion and Conclusion

A measurement of the normalized differential cross section of the Z boson decaying to electron pairs in terms of ϕ^* was performed using $19.7\,\text{fb}^{-1}$ of $\sqrt{s} = 8\,\text{TeV}$ data collected with the CMS detector. This measurement was compared to the predictions from MadGraph + Pythia6 (`Z2star`), powheg + Pythia6 (`Z2star`), powheg + Pythia8 (`Tunepp5`), and powheg + Pythia8 (`Tunepp14`).

The prediction from MadGraph is everywhere within 5 % of the distribution measured in data unfolded with MadGraph. The generator underestimates the cross section in both the low and high ϕ^* bins, but agrees in the very lowest bin and the three highest bins. In the intermediate region, MadGraph switches from underestimating to overestimating the cross section, but does so smoothly. Surprisingly, the prediction from MadGraph agrees better with the data unfolded with powheg + Pythia6 (`Z2star`), and not just because the uncertainty is larger. The central value of the prediction from MadGraph is closer to the data even disregarding the uncertainty.

The prediction from the powheg + Pythia6 (`Z2star`) sample is significantly worse than the prediction from MadGraph. Compared to the MadGraph unfolded data, powheg's prediction underestimate the cross section by 10 % in the low ϕ^* region, 15 % in the high ϕ^* region, and overestimates the cross section by 10 % in the intermediate region. In fact powheg is only consistent with the data in five bins, although four of these are grouped together around $\phi^* \approx 0.15$. powheg + Pythia6 (`Z2star`) does equally poorly when compared with the data unfolded with powheg + Pythia6 (`Z2star`).

The prediction from the powheg + Pythia8 (`Tunepp5`) and powheg + Pythia8 (`Tunepp14`) samples are as poor as the prediction from the powheg+Pythia6 (`Z2star`) sample, although in a different manner. Unlike the previous two samples, these two overestimate the cross section by 10 % in the lowest ϕ^* bins and then underestimate by 15 % in the highest ϕ^* bins. The two distributions are roughly compatible in the low ϕ^* bins, but deviate from each other significantly in the highest bins.

While none of the simulated samples matches the data well, it is interesting to note how varied the predictions are. The samples created with MadGraph + Pythia6 (`Z2star`) and powheg + Pythia6 (`Z2star`) use different generators but the same

hadronizer and tune. However, their shape is still very different, indicating that the generator is responsible for much of the disagreement. The samples created with POWHEG+PYTHIA6 (`Z2star`) and with POWHEG + PYTHIA8 (`Tunepp5`) and POWHEG + PYTHIA8 (`Tunepp14`) use the same generator but different hadronizers and tunes. Again, the distributions are very different, even more so then when comparing the MADGRAPH distribution to POWHEG one. This indicates the hadronizer and tune play a very large role in determining the ϕ^* distribution. Finally, the samples created with POWHEG + PYTHIA8 (`Tunepp5`) and POWHEG + PYTHIA8 (`Tunepp14`) use the same generator and the same hadronizer but different tunes. While the agreement between these two samples is the closest, many bins still disagree indicating that the tune alone can have an impact on the final distribution.

In the future, the measurement presented in this thesis can be used to optimize MC generation by adjusting the generator, hadronizer, and tune to achieve the best agreement possible. The MC samples created from this optimized simulation could then be used to make the templates used in the M_{W} measurement.

The ϕ^* variable, as expected, had very limited dependence on the p_{T} scale uncertainty of the detector, which was the smallest of the systematic uncertainties considered. Still, the measurement was limited by the systematic uncertainties, particularly the small number of MC events used for the unfolding. Fortunately, reducing this uncertainty is easy (although computationally expensive) as all that is required is to generate additional MC events. The statistical uncertainties were the next largest uncertainty. As the remaining uncertainties were all vanishingly small, future versions of this analysis will be able to improve on the measurement without needing to develop new analysis techniques.

References

[1] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579 – 588, 1961, http://dx.doi.org/10.1016/0029-5582(61)90469-2.

[2] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, November 1967, http://dx.doi.org/10.1103/PhysRevLett.19.1264.

[3] A Salam. Elementary particle physics: Relativistic groups and analyticity. In *Proceedings to the Eighth Nobel Symposium, Almquvist and Wiksell*, page 367, May 1968, http://dx.doi.org/10.1126/science.168.3936.1196-a.

[4] J.C. Maxwell. *A Treatise on Electricity and Magnetism.* Number v. 1-2 in A Treatise on Electricity and Magnetism. The Clarendon press for Macmillan and Company, 1873.

[5] J. C. Maxwell. A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 155(0):459–512, January 1865, http://dx.doi.org/10.1098/rstl.1865.0008.

[6] Hendrik Lorentz. Simplified Theory of Electrical and Optical Phenomena in Moving Systems. *Proceedings of the Royal Netherlands Academy of Arts and Sciences*, 1:427442, 1899.

[7] Albert Einstein. On the electrodynamics of moving bodies. *Annalen Phys.*, 17:891–921, 1905, http://dx.doi.org/10.1002/andp.200590006.

[8] P. A. M. Dirac. The quantum theory of the emission and absorption of radiation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 114(767):243–265, March 1927, http://dx.doi.org/10.1098/rspa.1927.0039.

[9] P. A. M. Dirac. The quantum theory of the electron. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 117(778):610–624, February 1928, http://dx.doi.org/10.1098/rspa.1928.0023.

[10] P. A. M. Dirac. A theory of electrons and protons. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 126(801):360–365, January 1930, http://dx.doi.org/10.1098/rspa.1930.0013.

[11] Carl D. Anderson. The positive electron. *Phys. Rev.*, 43:491–494, March 1933, http://dx.doi.org/10.1103/PhysRev.43.491.

[12] Willis Lamb and Robert Rutherford. Fine Structure of the Hydrogen Atom by a Microwave Method. *Physical Review*, 72(3):241–243, August 1947, http://dx.doi.org/10.1103/physrev.72.241.

[13] H. A. Bethe. The Electromagnetic Shift of Energy Levels. *Physical Review Online Archive (Prola)*, 72(4):339–341, August 1947, http://dx.doi.org/10.1103/physrev.72.339.

[14] S. Tomonaga. On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields. *Progress of Theoretical Physics*, 1(2):27–42, 1946, http://dx.doi.org/10.1143/ptp.1.27.

[15] Julian Schwinger. On Quantum-Electrodynamics and the Magnetic Moment of the Electron. *Physical Review Online Archive (Prola)*, 73(4):416–417, February 1948, http://dx.doi.org/10.1103/physrev.73.416.

[16] R. P. Feynman. Space-Time Approach to Quantum Electrodynamics. *Physical Review Online Archive (Prola)*, 76(6):769–789, September 1949, http://dx.doi.org/10.1103/physrev.76.769.

[17] F. Dyson. The Radiation Theories of Tomonaga, Schwinger, and Feynman. *Physical Review*, 75(3):486–502, February 1949, http://dx.doi.org/10.1103/physrev.75.486.

[18] James Chadwick. Intensitätsverteilung im magnetischen spektren der β-strahlen von radium B + C. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, 14:383–391, 1914.

[19] Wolfgang Pauli. Open letter to the group of radioactive people at the Gauverein meeting in Tübingen. Letter, December 1930.

[20] Enrico Fermi. Versuch einer theorie der β-strahlen. i. *Zeitschrift für Physik*, 88(3-4):161–177, 1934, http://dx.doi.org/10.1007/BF01351864.

[21] G. D. Rochester and C. C. Butler. Evidence for the existence of new unstable elementary particles. *Nature*, 160(4077):855–857, December 1947, http://dx.doi.org/10.1038/160855a0.

[22] R. Brown, U. Camerini, P. H. Fowler, H. Muirhead, C. F. Powell, and D. M. Ritson. Observations with electron-sensitive plates exposed to cosmic radiation. *Nature*, 163(4133):82–87, January 1949, http://dx.doi.org/10.1038/163082a0.

[23] T. D. Lee and C. N. Yang. Question of parity conservation in weak interactions. *Phys. Rev.*, 104:254–258, October 1956, http://dx.doi.org/10.1103/PhysRev.104.254.

[24] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, February 1957, http://dx.doi.org/10.1103/PhysRev.105.1413.

[25] Richard L. Garwin, Leon M. Lederman, and Marcel Weinrich. Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon. *Phys. Rev.*, 105:1415–1417, February 1957, http://dx.doi.org/10.1103/PhysRev.105.1415.

[26] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96:191–195, October 1954, http://dx.doi.org/10.1103/PhysRev.96.191.

[27] E. C. G. Sudarshan and R. E. Marshak. Chirality invariance and the universal fermi interaction. *Phys. Rev.*, 109:1860–1862, March 1958, http://dx.doi.org/10.1103/PhysRev.109.1860.2.

[28] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, August 1964, http://dx.doi.org/10.1103/PhysRevLett.13.321.

[29] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, October 1964, http://dx.doi.org/10.1103/PhysRevLett.13.508.

[30] E. Rutherford. LXXIX. The scattering of α and β particles by matter and the structure of the atom. *Philosophical Magazine Series 6*, 21(125):669–688, May 1911, http://dx.doi.org/10.1080/14786440508637080.

[31] E. Rutherford. LIV. Collision of α particles with light atoms . IV. An anomalous effect in nitrogen. *Philosophical Magazine Series 6*, 37(222):581–587, June 1919, http://dx.doi.org/10.1080/14786440608635919.

[32] James Chadwick. Possible existence of a neutron. *Nature*, 129(3252):312, February 1932, http://dx.doi.org/10.1038/129312a0.

[33] H Yukawa. On the Interaction of Elementary Particles. Part I. *Proc. Phys. Math. Soc. Jpn.*, 17:48–57, 1935, http://ptps.oxfordjournals.org/content/1/1.1.short.

[34] C. M. G. Lattes, G. P. S. Occhialini, and C. F. Powell. Observations on the tracks of slow mesons in photographic emulsions. *Nature*, 160(4066):453–456, October 1947, http://dx.doi.org/10.1038/160453a0.

[35] Tadao Nakano and Kazuhiko Nishijima. Charge independence for v-particles. *Prog. Theor. Phys.*, 10(5):581–582, November 1953, http://dx.doi.org/10.1143/ptp.10.581.

[36] Kazuhiko Nishijima. Charge independence theory of v-particles. *Prog. Theor. Phys.*, 13(3):285–304, March 1955, http://dx.doi.org/10.1143/ptp.13.285.

[37] M. Gell-Mann. The interpretation of the new particles as displaced charge multiplets. *Nuovo Cim*, 4(S2):848–866, April 1956, http://dx.doi.org/10.1007/bf02748000.

[38] M. Gell-Mann. A schematic model of baryons and mesons. *Physics Letters*, 8(3):214–215, February 1964, http://dx.doi.org/10.1016/s0031-9163(64)92001-3.

[39] G Zweig. An SU_3 model for strong interaction symmetry and its breaking; Version 1. Technical Report CERN-TH-401, CERN, Geneva, January 1964, https://cds.cern.ch/record/352337.

[40] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, 30:1343–1346, June 1973, http://dx.doi.org/10.1103/PhysRevLett.30.1343.

[41] H. David Politzer. Reliable perturbative results for strong interactions? *Phys. Rev. Lett.*, 30:1346–1349, June 1973, http://dx.doi.org/10.1103/PhysRevLett.30.134s.

[42] F.J. Hasert et. al. Search for elastic muon-neutrino electron scattering. *Physics Letters B*, 46(1):121–124, 1973, http://dx.doi.org/10.1016/0370-2693(73)90494-2.

[43] J. J. Aubert et. al. Experimental observation of a heavy particle J. *Phys. Rev. Lett.*, 33:1404–1406, December 1974, http://dx.doi.org/10.1103/PhysRevLett.33.1404.

[44] J. E. Augustin et. al. Discovery of a narrow resonance in e+e- annihilation. *Phys. Rev. Lett.*, 33:1406–1408, December 1974, http://dx.doi.org/10.1103/PhysRevLett.33.1406.

[45] S. W. Herb et. al. Observation of a dimuon resonance at 9.5 GeV in 400-GeV proton-nucleus collisions. *Phys. Rev. Lett.*, 39:252–255, August 1977, http://dx.doi.org/10.1103/PhysRevLett.39.252.

[46] CDF Collaboration. Observation of top quark production in $\overline{p}p$ collisions with the collider detector at fermilab. *Phys. Rev. Lett.*, 74:2626–2631, April 1995, http://dx.doi.org/10.1103/PhysRevLett.74.2626.

[47] D0 Collaboration. Search for high mass top quark production in $p\overline{p}$ collisions at $\sqrt{s} = 1.8$ tev. *Phys. Rev. Lett.*, 74:2422–2426, March 1995, http://dx.doi.org/10.1103/PhysRevLett.74.2422.

[48] UA1 Collaboration. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV. *Physics Letters B*, 122(1):103–116, 1983, http://dx.doi.org/10.1016/0370-2693(83)91177-2.

[49] UA2 Collaboration. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider. *Physics Letters B*, 122(56):476–485, 1983, http://dx.doi.org/10.1016/0370-2693(83)91605-2.

[50] UA1 Collaboration. Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider. *Physics Letters B*, 126(5):398–410, 1983, http://dx.doi.org/10.1016/0370-2693(83)90188-0.

[51] UA2 Collaboration. Evidence for Z0→e+e- at the CERN pp collider. *Physics Letters B*, 129(12):130–140, 1983, http://dx.doi.org/10.1016/0370-2693(83)90744-X.

[52] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, 2012, http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[53] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, 2012, http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[54] K.A. Olive et al. Review of Particle Physics. *Chin.Phys.*, C38:090001, 2014, http://dx.doi.org/10.1088/1674-1137/38/9/090001.

[55] Giuseppe Bozzi, Stefano Catani, Giancarlo Ferrera, Daniel de Florian, and Massimiliano Grazzini. Production of DrellYan lepton pairs in hadron collisions: Transverse-momentum resummation at next-to-next-to-leading logarithmic accuracy. *Physics Letters B*, 696(3):207 – 213, January 2011, http://dx.doi.org/10.1016/j.physletb.2010.12.024.

[56] Sonny Mantry and Frank Petriello. Transverse momentum distributions in the nonperturbative region. *Phys. Rev. D*, 84:014030, July 2011, http://link.aps.org/doi/10.1103/PhysRevD.84.014030.

[57] Thomas Becher and Matthias Neubert. Drell–Yan production at small q_t, transverse parton distributions and the collinear anomaly. *The European Physical Journal C*, 71(6), July 2011, http://dx.doi.org/10.1140/epjc/s10052-011-1665-7.

[58] M. Baak, J. Cúth, J. Haller, A. Hoecker, R. Kogler, K. Mönig, M. Schott, and J. Stelzer. The global electroweak fit at NNLO and prospects for the LHC and ILC. *The European Physical Journal C*, 74(9), July 2014, http://dx.doi.org/10.1140/epjc/s10052-014-3046-5.

[59] A.D. Martin, W.J. Stirling, R.S. Thorne, and G. Watt. Parton distributions for the LHC. *The European Physical Journal C*, 63(2):189–285, 2009, http://dx.doi.org/10.1140/epjc/s10052-009-1072-5.

[60] Sidney D. Drell and Tung-Mow Yan. Massive lepton-pair production in hadron-hadron collisions at high energies. *Phys. Rev. Lett.*, 25:316–320, August 1970, http://dx.doi.org/10.1103/PhysRevLett.25.316.

[61] Sidney D. Drell and Tung-Mow Yan. Massive lepton-pair production in hadron-hadron collisions at high energies. *Phys. Rev. Lett.*, 25:902–902, September 1970, http://dx.doi.org/10.1103/PhysRevLett.25.902.2.

[62] A. Banfi, S. Redford, M. Vesterinen, P. Waller, and T.R. Wyatt. Optimisation of variables for studying dilepton transverse momentum distributions at hadron colliders. *The European Physical Journal C*, 71(3), 2011, http://dx.doi.org/10.1140/epjc/s10052-011-1600-y.

[63] ATLAS Collaboration. Measurement of angular correlations in Drell–Yan lepton pairs to probe boson transverse momentum at with the ATLAS detector. *Physics Letters B*, 720(1-3):32–51, March 2013.

[64] D0 Collaboration. Precise study of the z/γ^* boson transverse momentum distribution in $p\overline{p}$ collisions using a novel technique. *Phys. Rev. Lett.*, 106:122001, March 2011, http://dx.doi.org/10.1103/PhysRevLett.106.122001.

[65] D0 Collaboration. Measurement of the ϕ^*_η distribution of muon pairs with masses between 30 and 500 GeV in 10.4fb^{-1} of $p\overline{p}$ collisions. *Phys. Rev. D*, 91:072002, April 2015, http://dx.doi.org/10.1103/PhysRevD.91.072002.

[66] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*. CERN, Geneva, 2004, http://dx.doi.org/10.5170/CERN-2004-003-V-1.

[67] CMS Collaboration. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006, http://cds.cern.ch/record/922757.

[68] CMS Collaboration. CMS Physics: Technical Design Report Volume 2: Physics Performance. *J. Phys. G*, 34(CERN-LHCC-2006-021. CMS-TDR-8-2):995–1579. 669 p, 2007, http://dx.doi.org/10.1088/0954-3899/34/6/S01.

[69] CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation*, 9(10):P10009, 2014, http://dx.doi.org/10.1088/1748-0221/9/10/P10009.

[70] CMS Collaboration. Electron reconstruction and identification at sqrt(s) = 7 TeV. Technical Report CMS-PAS-EGM-10-004, CERN, Geneva, 2010, http://cds.cern.ch/record/1299116.

[71] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, et al. Electron reconstruction in CMS. *Eur.Phys.J.*, C49:1099–1116, 2007, http://dx.doi.org/10.1140/epjc/s10052-006-0175-5.

[72] W Adam, R Frühwirth, A Strandlie, and T Todorov. Reconstruction of electrons with the gaussian-sum filter in the CMS tracker at the LHC. *Journal of Physics G: Nuclear and Particle Physics*, 31(9):N9, 2005, http://stacks.iop.org/0954-3899/31/i=9/a=N01.

[73] Y. Chen, E. DiMarco, M. Spiropulu, S. Xie, B. Mangano, G. Petrucciani, V. Sharma, and A. Vartak. Electron energy reconstruction using a multivariate regression. CMS Note 2012/327, CERN, 2012.

[74] CMS Collaboration. Measurements of the new Higgs-like boson at 125 GeV in the two photon decay channel. CMS Note 2013/253, CERN, 2013.

[75] M. Oreglia. A study of the reactions $\psi' \to \gamma\gamma\psi$. PhD Thesis from Stanford University, December 1980, http://www.slac.stanford.edu/cgi-wrap/getdoc/slac-r-236.pdf.

[76] CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical Report CMS-PAS-PFT-09-001, CERN, Geneva, April 2009, https://cds.cern.ch/record/1194487.

[77] CMS Collaboration. Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV. Technical Report CMS-PAS-PFT-10-002, CERN, Geneva, March 2010, https://cds.cern.ch/record/1247373.

[78] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5: Going Beyond. *Journal of High Energy Physics*, 2011(6), 2011, http://dx.doi.org/10.1007/JHEP06(2011)128.

[79] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *Journal of High Energy Physics*, 2004(11):40, 2004, http://stacks.iop.org/1126-6708/2004/i=11/a=040.

[80] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower monte carlo programs: the POWHEG BOX. *Journal of High Energy Physics*, 2010(6), 2010, http://dx.doi.org/10.1007/JHEP06(2010)043.

[81] Emanuele Re. Single-top Wt-channel production matched with parton showers using the POWHEG method. *The European Physical Journal C*, 71(2), 2011, http://dx.doi.org/10.1140/epjc/s10052-011-1547-z.

[82] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. Pythia 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026, 2006, http://dx.doi.org/10.1088/1126-6708/2006/05/026.

[83] Z. Wąs. Precision simulations with TAUOLA and PHOTOS. *Nuclear Physics B - Proceedings Supplements*, 169(0):16 – 21, 2007, http://dx.doi.org/10.1016/j.nuclphysbps.2007.02.113. Proceedings of the Ninth International Workshop on Tau Lepton Physics TAU06.

[84] S. Agostinelli et al. Geant4–a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and*

Associated Equipment, 506(3):250–303, 2003, http://dx.doi.org/10.1016/S0168-9002(03)01368-8.

[85] I. Kravchenko, R. Kamalieddin, K. Sung, A. Juodagalvis, and E. Olaiya. Electron reconstruction efficiency scale factors for 2012 data, June 2013, https://indico.cern.ch/event/257547/material/slides/0?contribId=2. Talk presented to CMS Joint ECAL DPG/EGM POG meeting.

[86] L. K. Saini, I. Kravchenko, and Y. Maravin. A study of efficiencies and scale factors for cut-based electron identification at CMS experiment using data from proton-proton collisions at $\sqrt{s} = 8$ tev. CMS Note 2014/055, CERN, 2014.

[87] Jason A. Haupt. The differential shape of the Z to ee cross-section as a function of Z rapidity. PhD Thesis from the University of Minnesota, May 2011, http://conservancy.umn.edu/handle/11299/107836.

[88] Tim Adye. Unfolding algorithms and tests using RooUnfold. Proceedings of the PHYSTAT 2011 Workshop, CERN, Geneva, Switzerland, January 2011, CERN-2011-006, pp 313-318, 2011, arXiv:1105.1160.

[89] G. D'Agostini. A multidimensional unfolding method based on Bayes' theorem. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 362(23):487 – 498, 1995, http://dx.doi.org/10.1016/0168-9002(95)00274-X.

[90] CMS Collaboration. CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update. *CERN*, 2013, https://cds.cern.ch/record/1598864.

[91] S. van der Meer. Calibration of the effective beam height in the ISR. Technical Report CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, June 1968, https://cds.cern.ch/record/296752.

[92] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, S. Forte, A. Glazov, J. Huston, R. McNulty, T. Sjöstrand, and R. Thorne. The PDF4LHC Working Group Interim Recommendations. *ArXiv e-prints*, January 2011, arXiv:1101.0538.

Appendix A

Other Measurements Using Dressed Electrons

A.1 Uncertainty Figures

The values of the various uncertainties in each ϕ^* bin are presented in the figures that follow. Figure 7.3 shows the uncertainties in the normalized ϕ^* cross section in data unfolded with MadGraph while figure 7.4 shows the uncertainties in the data unfolded with powheg + Pythia6 (`Z2star`). The uncertainties on the normalized ϕ^* cross section in MadGraph and powheg are shown in figures 7.5 and 7.6, respectively. All of these values are presented in tables in appendix B.

A.2 Absolute Differential Cross Section

The absolute differential cross section measurement using data unfolded with MadGraph is shown in figure A.5 and given in tabular form in table A.1. The lower plot in figure A.5 is shown in more detail in figure A.6. As previously discussed in this chapter, the primary uncertainty on the data distribution is from the integrated luminosity. The primary uncertainty for the MadGraph sample is the fewz calculated overall cross section used to scale the distribution, while the primary uncertainty for the powheg samples is the uncertainty calculated by varying the `CT10` PDF weights.

The absolute differential cross section measurement using data unfolded with powheg+

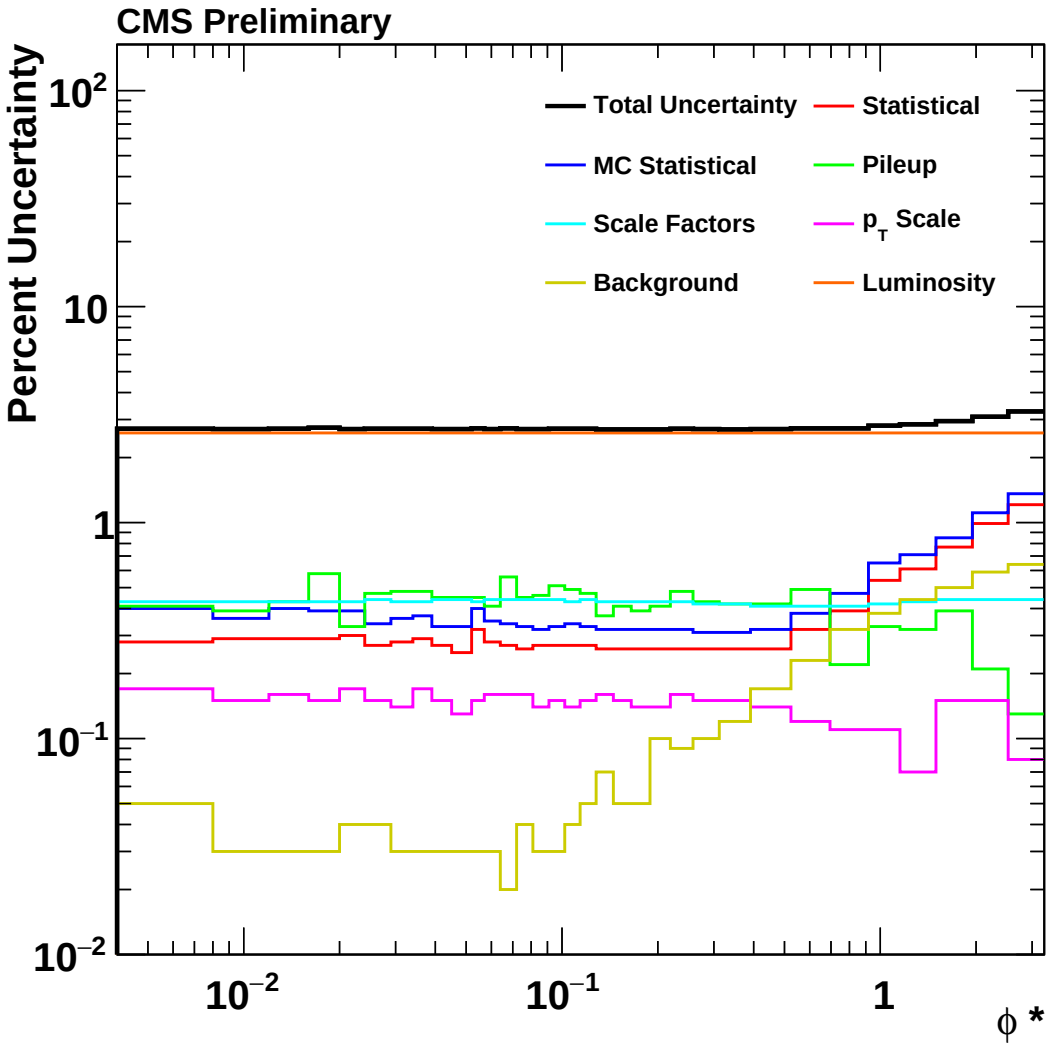


Figure A.1: The uncertainties (in %) for the absolute cross section measurement made with data unfolded with MadGraph. The total value is the sum in quadrature of all the other values. These uncertainties are also presented in tabular form in table B.5.

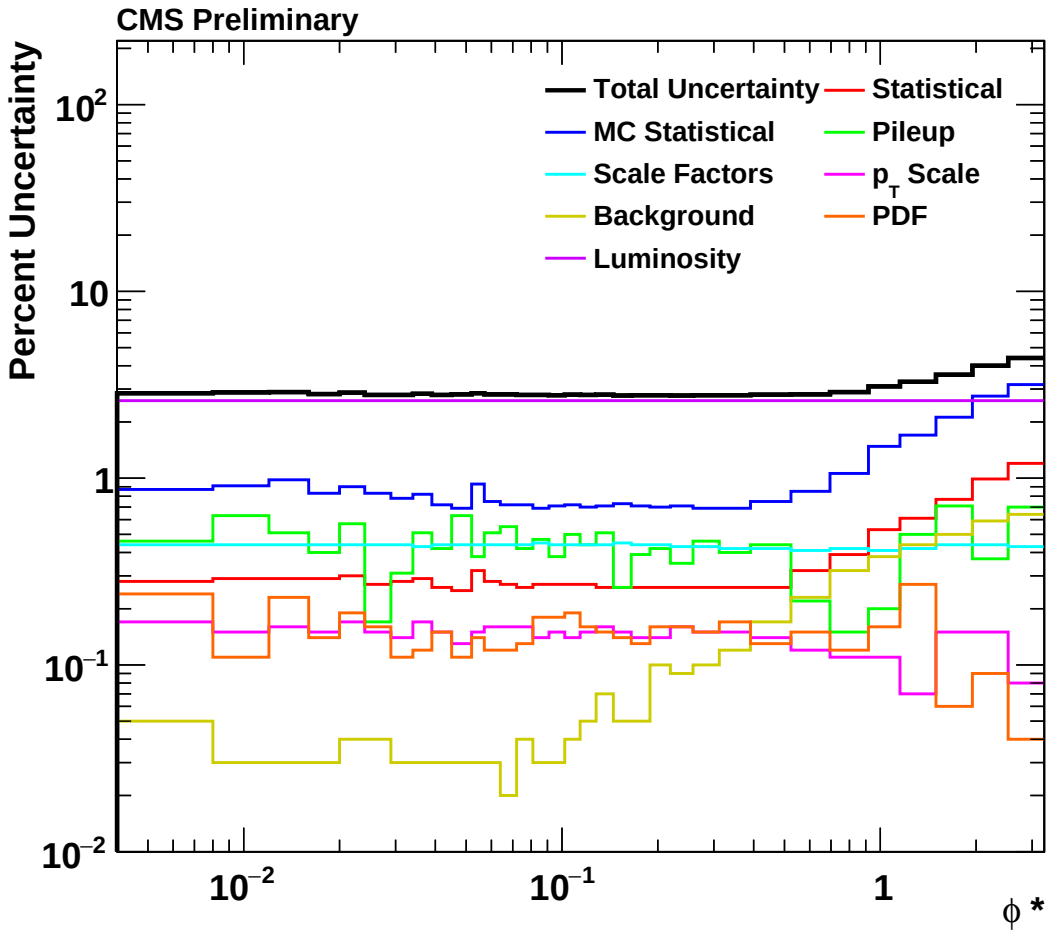


Figure A.2: The uncertainties (in %) for the absolute cross section measurement made with data unfolded with POWHEG + PYTHIA6 (`Z2star`). The total value is the sum in quadrature of all the other values. These uncertainties are also presented in tabular form in table B.6.

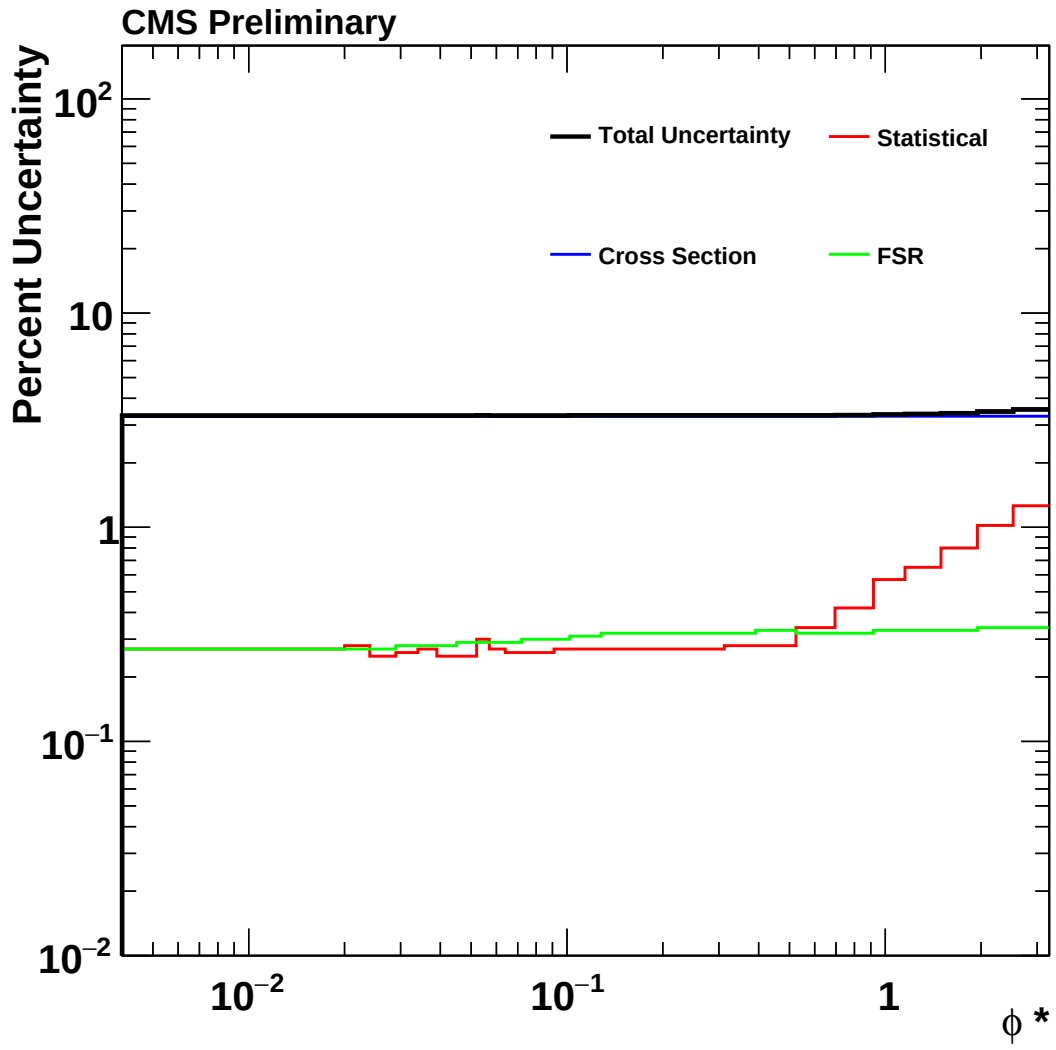


Figure A.3: The uncertainties (in %) for the absolute cross section from the MadGraph MC sample. These uncertainties are also presented in tabular form in table B.7.

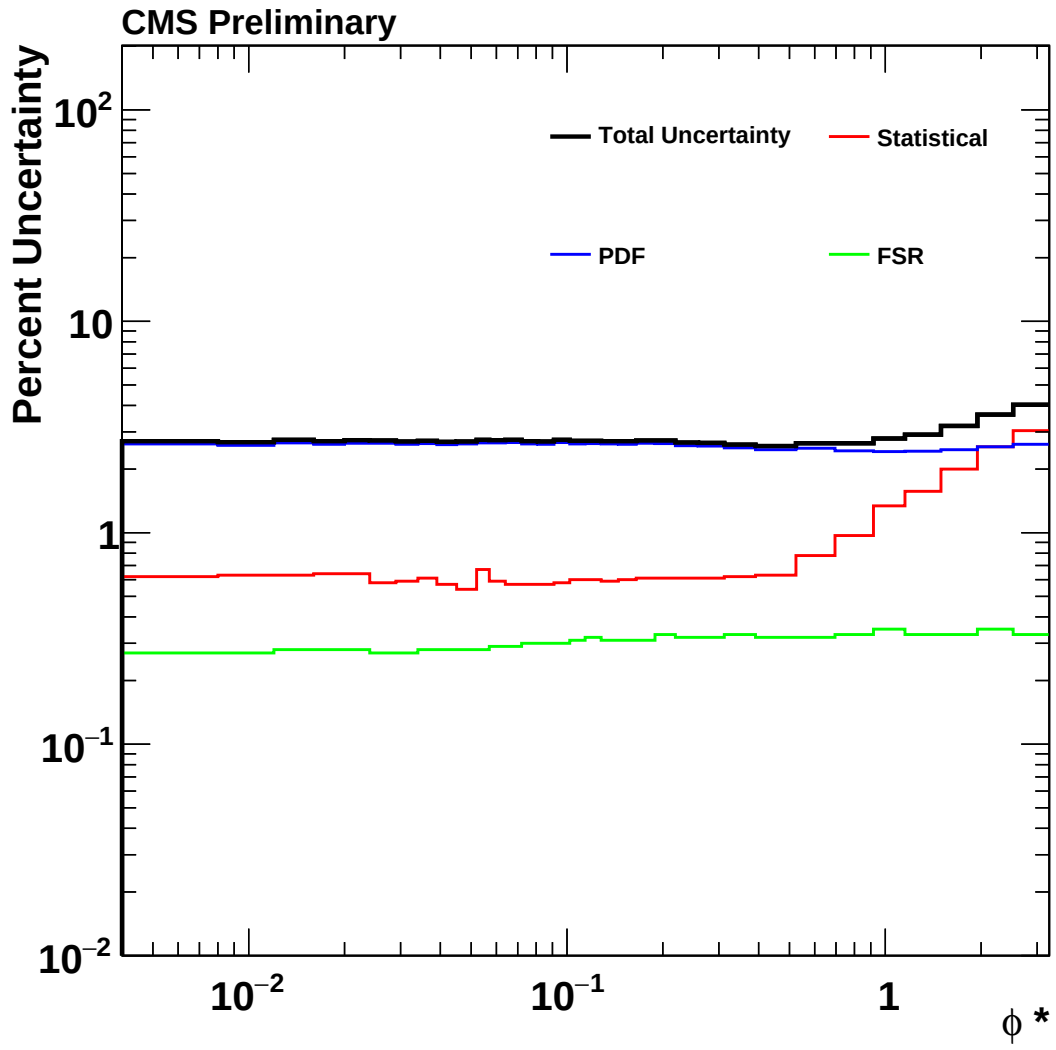


Figure A.4: The uncertainties (in %) for the absolute cross section from the POWHEG MC sample. These uncertainties are also presented in tabular form in table B.8.

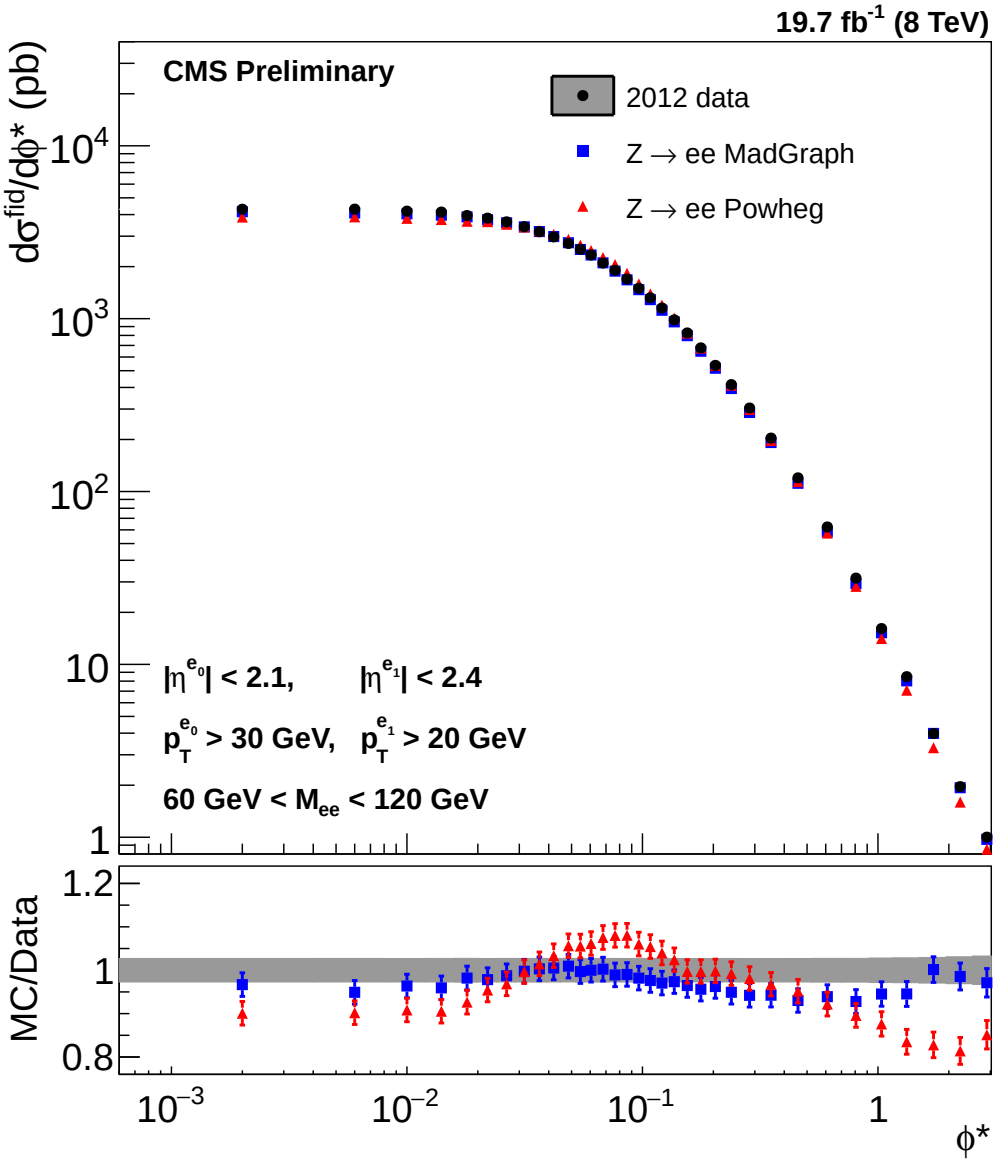


Figure A.5: The absolute differential cross section with respects to ϕ^* for $Z \to e^+e^-$ events in our fiducial region from data unfolded with MadGraph, and the same distributions in MadGraph and powheg + Pythia6 (`Z2star`). A close up of the lower plot is shown in figure A.6.

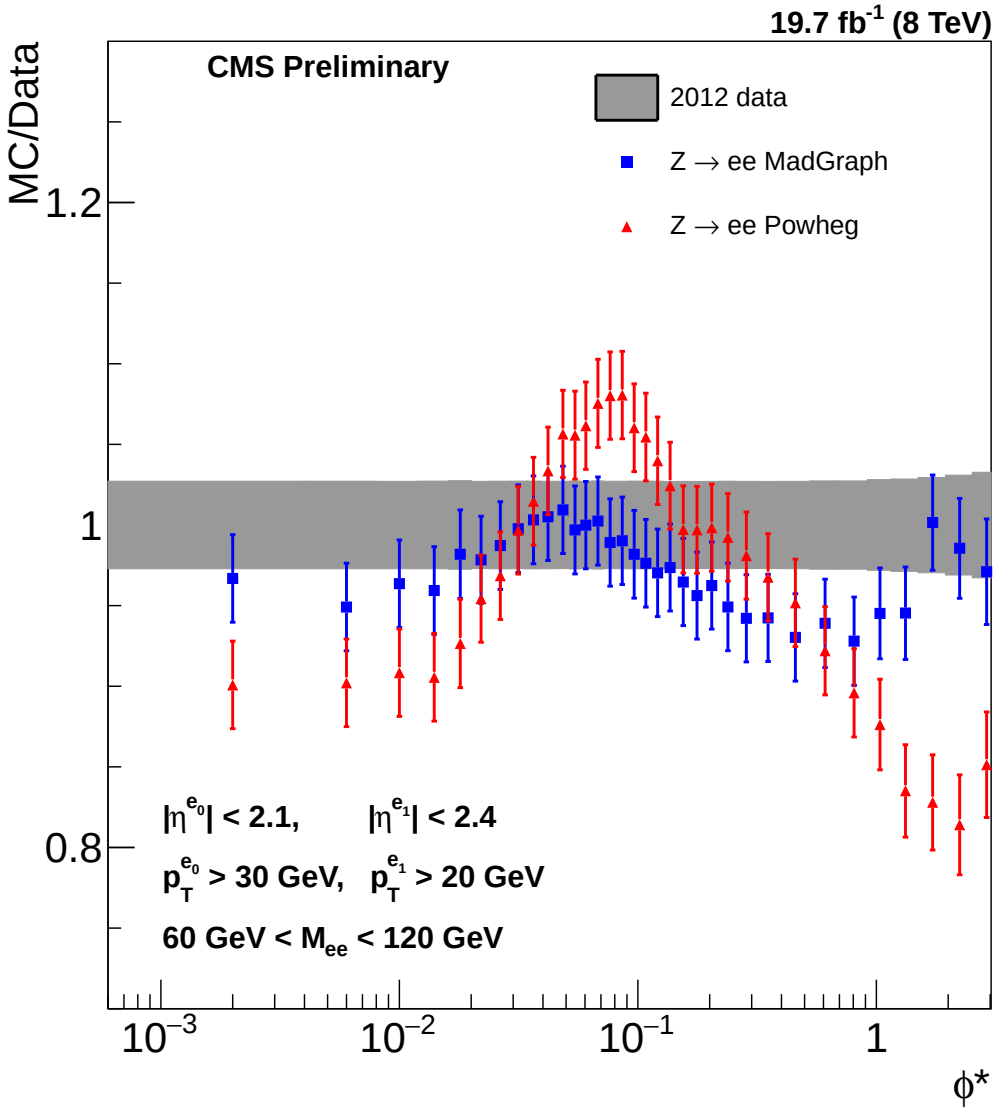


Figure A.6: Close up of the ratio plot from figure A.5 for the absolute cross section measurement unfolded with MadGraph. The error band indicates the uncertainty in the data, while the square points show the ratio of MadGraph over data, and the triangle points show the ratio of powheg + Pythia6 (`Z2star`) over data.

ϕ^* range	Data (pb)	MadGraph (pb)	powheg (±)
0.000–0.004	4297 ± 117	4155 ± 138	3871 ± 105
0.004–0.008	4302 ± 117	4083 ± 136	3881 ± 105
0.008–0.012	4190 ± 113	4037 ± 134	3806 ± 102
0.012–0.016	4137 ± 113	3969 ± 132	3746 ± 103
0.016–0.020	3951 ± 109	3879 ± 129	3660 ± 99
0.020–0.024	3816 ± 103	3734 ± 124	3642 ± 100
0.024–0.029	3632 ± 99	3586 ± 119	3518 ± 96
0.029–0.034	3404 ± 93	3396 ± 113	3393 ± 92
0.034–0.039	3182 ± 87	3192 ± 106	3229 ± 88
0.039–0.045	2971 ± 81	2987 ± 99	3071 ± 83
0.045–0.052	2724 ± 74	2750 ± 91	2878 ± 78
0.052–0.057	2521 ± 69	2514 ± 84	2662 ± 73
0.057–0.064	2335 ± 63	2335 ± 78	2479 ± 68
0.064–0.072	2099 ± 57	2104 ± 70	2257 ± 62
0.072–0.081	1904 ± 52	1883 ± 63	2057 ± 56
0.081–0.091	1694 ± 46	1678 ± 56	1831 ± 49
0.091–0.102	1498 ± 41	1471 ± 49	1589 ± 44
0.102–0.114	1320 ± 36	1288 ± 43	1392 ± 38
0.114–0.128	1152 ± 31	1118 ± 37	1198 ± 33
0.128–0.145	984 ± 27	958 ± 32	1008 ± 27
0.145–0.165	827 ± 22	797 ± 27	824 ± 22
0.165–0.189	678 ± 18	648 ± 22	676 ± 18
0.189–0.219	537 ± 15	517 ± 17	536 ± 15
0.219–0.258	415 ± 11	394 ± 13	412 ± 11
0.258–0.312	304 ± 8	287 ± 10	298 ± 8
0.312–0.391	204 ± 6	192 ± 6	197 ± 5
0.391–0.524	120 ± 3	112 ± 4	114 ± 3
0.524–0.695	62 ± 2	59 ± 2	58 ± 2
0.695–0.918	31.6 ± 0.9	29.3 ± 1.0	28.3 ± 0.7
0.918–1.153	16.1 ± 0.5	15.2 ± 0.5	14.1 ± 0.4
1.153–1.496	8.5 ± 0.2	8.0 ± 0.3	7.1 ± 0.2
1.496–1.947	4.0 ± 0.1	4.0 ± 0.1	3.3 ± 0.1
1.947–2.522	1.96 ± 0.06	1.93 ± 0.07	1.60 ± 0.06
2.522–3.277	1.00 ± 0.03	0.97 ± 0.03	0.85 ± 0.03

Table A.1: The absolute differential cross section in pb with respects to ϕ^* for $Z \to e^+e^-$ events in our fiducial region from data unfolded with MadGraph, and the same distributions in MadGraph and powheg + Pythia6 (`Z2star`). These results are shown graphically in figure A.5.

PYTHIA6 (`Z2star`) is shown in figure A.7 and given in tabular form in table A.2. The lower plot in figure A.7 is shown in more detail in figure A.8. The primary uncertainty on the data distribution is still from the integrated luminosity, although the MC statistical uncertainty is larger in the highest ϕ^* bins.

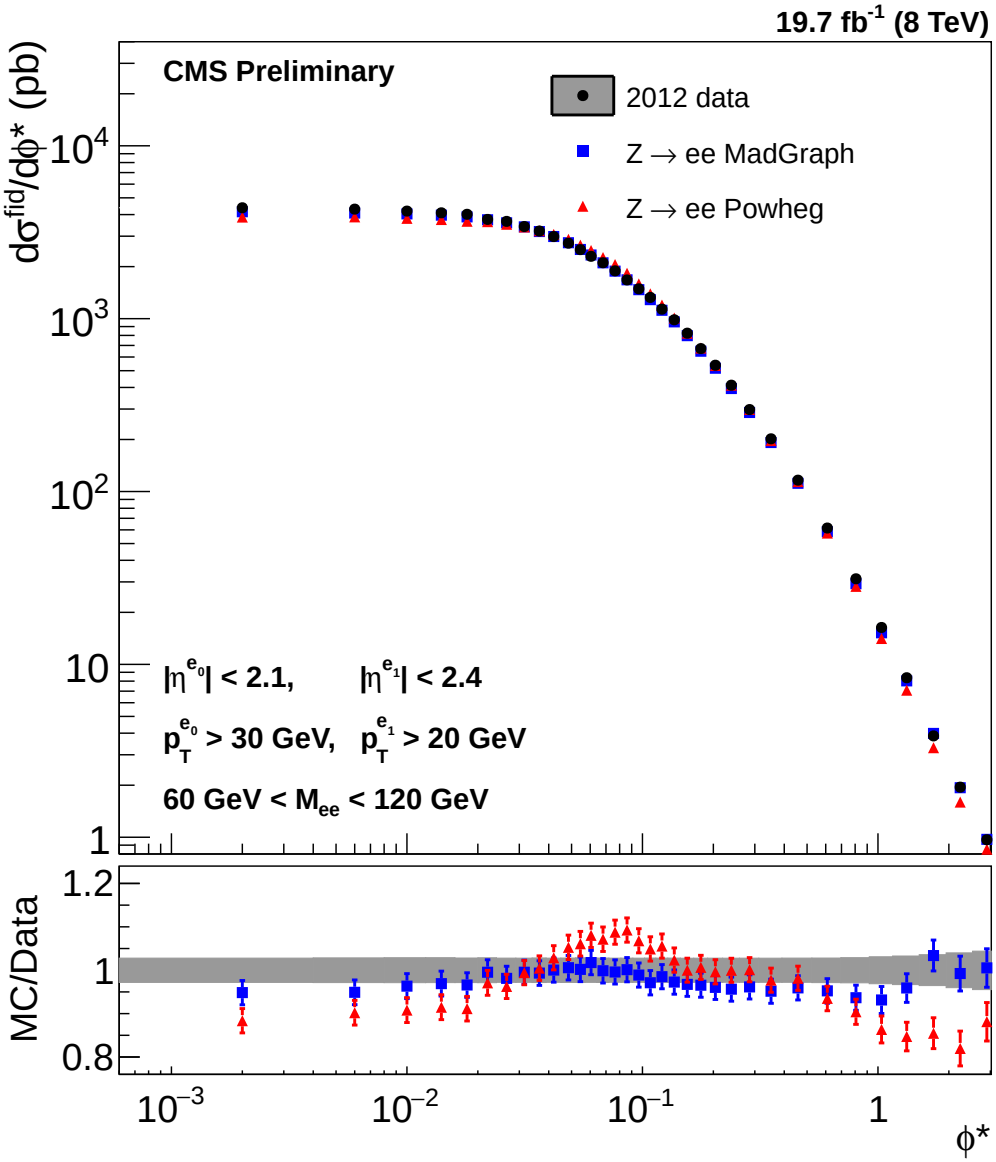


Figure A.7: The absolute differential cross section with respects to ϕ^* for $Z \to e^+e^-$ events in our fiducial region from data unfolded with POWHEG + PYTHIA6 (`Z2star`), and the same distributions in MADGRAPH and POWHEG + PYTHIA6 (`Z2star`). A close up of the lower plot is shown in figure A.8.

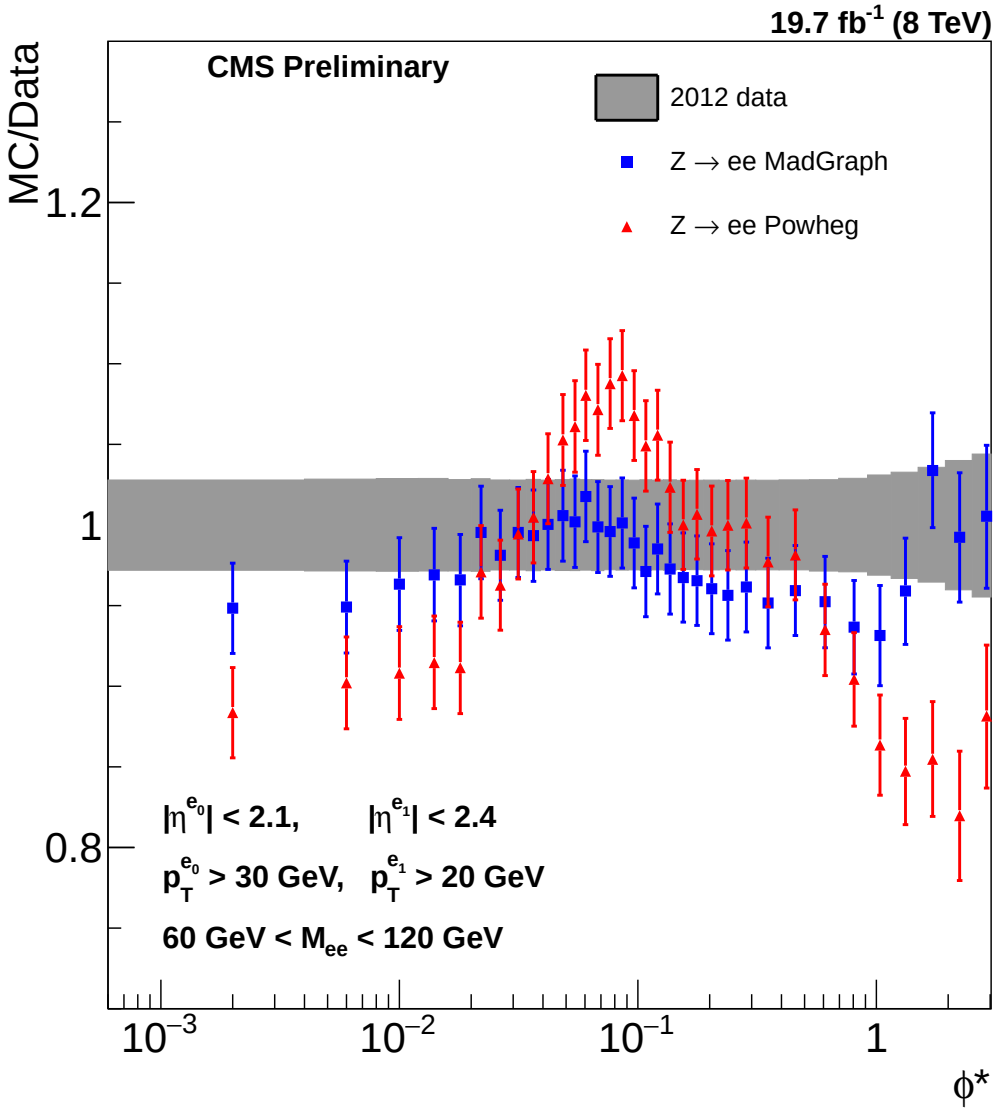


Figure A.8: Close up of the ratio plot from figure A.7 for the absolute cross section measurement unfolded with POWHEG + PYTHIA6 (`Z2star`). The error band indicates the uncertainty in the data, while the square points show the ratio of MADGRAPH over data, and the triangle points show the ratio of POWHEG + PYTHIA6 (`Z2star`) over data.

ϕ^* range	Data (pb)	MadGraph (pb)	powheg ($\pm$)
0.000–0.004	4381 ± 122	4155 ± 138	3871 ± 105
0.004–0.008	4302 ± 122	4083 ± 136	3881 ± 105
0.008–0.012	4191 ± 121	4037 ± 134	3806 ± 102
0.012–0.016	4096 ± 118	3969 ± 132	3746 ± 103
0.016–0.020	4016 ± 113	3879 ± 129	3660 ± 99
0.020–0.024	3751 ± 108	3734 ± 124	3642 ± 100
0.024–0.029	3654 ± 102	3586 ± 119	3518 ± 96
0.029–0.034	3412 ± 95	3396 ± 113	3393 ± 92
0.034–0.039	3213 ± 91	3192 ± 106	3229 ± 88
0.039–0.045	2985 ± 83	2987 ± 99	3071 ± 83
0.045–0.052	2734 ± 77	2750 ± 91	2878 ± 78
0.052–0.057	2509 ± 71	2514 ± 84	2662 ± 73
0.057–0.064	2295 ± 65	2335 ± 78	2479 ± 68
0.064–0.072	2107 ± 59	2104 ± 70	2257 ± 62
0.072–0.081	1891 ± 53	1883 ± 63	2057 ± 56
0.081–0.091	1676 ± 47	1678 ± 56	1831 ± 49
0.091–0.102	1488 ± 41	1471 ± 49	1589 ± 44
0.102–0.114	1327 ± 37	1288 ± 43	1392 ± 38
0.114–0.128	1135 ± 32	1118 ± 37	1198 ± 33
0.128–0.145	985 ± 28	958 ± 32	1008 ± 27
0.145–0.165	824 ± 23	797 ± 27	824 ± 22
0.165–0.189	671 ± 19	648 ± 22	676 ± 18
0.189–0.219	538 ± 15	517 ± 17	536 ± 15
0.219–0.258	412 ± 11	394 ± 13	412 ± 11
0.258–0.312	298 ± 8	287 ± 10	298 ± 8
0.312–0.391	202 ± 6	192 ± 6	197 ± 5
0.391–0.524	116 ± 3	112 ± 4	114 ± 3
0.524–0.695	61 ± 2	59 ± 2	58 ± 2
0.695–0.918	31.3 ± 0.9	29.3 ± 1.0	28.3 ± 0.7
0.918–1.153	16.4 ± 0.5	15.2 ± 0.5	14.1 ± 0.4
1.153–1.496	8.4 ± 0.3	8.0 ± 0.3	7.1 ± 0.2
1.496–1.947	3.9 ± 0.1	4.0 ± 0.1	3.3 ± 0.1
1.947–2.522	1.95 ± 0.08	1.93 ± 0.07	1.60 ± 0.06
2.522–3.277	0.97 ± 0.04	0.97 ± 0.03	0.85 ± 0.03

Table A.2: The absolute differential cross section in pb with respects to ϕ^* for $Z \to e^+e^-$ events in our fiducial region from data unfolded with powheg + Pythia6 (`Z2star`), and the same distributions in MadGraph and powheg + Pythia6 (`Z2star`). These results are shown graphically in figure A.7.

Appendix B

Uncertainty Tables

The uncertainties of the normalized cross section measurement are shown in figures 7.3 to 7.6, and for the absolute cross section measurement are shown in figures A.1 to A.4. In this appendix, the values of these uncertainties are presented in tabular form.

B.1 Explanation of the Columns

The tables with information about the uncertainty of the data, tables B.1, B.2, B.5, and B.6, contain the following columns:

ϕ^* Range:
: The range of ϕ^* values included in the bin.

Total Uncertainty (Total):
: The sum in quadrature of the all of the uncertainties.

Statistical Uncertainty (Stat.):
: The uncertainty due to the limited number of events in the data, as discussed in section 7.1.1.

Total Systematic Uncertainty (Total Syst.):
: The sum in quadrature of all of the systematic uncertainties including, for the absolute distribution, the luminosity uncertainty of 2.6 %.

Monte Carlo Statistical Uncertainty (MC Stat.):
: The uncertainty due to the limited number of events in the MC samples, as discussed in section 7.1.2.

Pileup Uncertainty (Pileup):

The uncertainty due to the pileup reweighting, as discussed in section 7.1.4.

Scale Factor Uncertainty (SF):

The uncertainty due to the scale factors, as discussed in section 7.1.5.

p_{T} Scale Uncertainty (p_{T} Scale):

The uncertainty due to p_{T} scale, as discussed in section 7.1.6.

Background Subtraction Uncertainty (Bkg.):

The uncertainty due to background subtraction, as discussed in section 7.1.7.

The tables with information related to the uncertainty of the MC distributions, tables B.3, B.4, B.7, and B.8, have the following columns:

ϕ^* Range:

The range of ϕ^* values included in the bin.

Total Uncertainty (Total):

The sum in quadrature of the all of the uncertainties.

Statistical Uncertainty (Stat.):

The uncertainty due to the limited number of events in the MC sample.

Parton Density Function (PDF):

The uncertainty due to choice of PDF used to generate the POWHEG MC, as discussed in section 7.1.8.

Theoretical Cross Section Uncertainty (Cross Section):

The uncertainty in the theoretical cross section of the MADGRAPH MC, as discussed in section 7.1.8.

Final State Radiation Uncertainty (FSR):

The uncertainty due to the modeling of FSR, as discussed in section 7.1.9.

B.2 Tables

The values of the various uncertainties in each ϕ^* bin are presented in the tables that follow. Tables B.1 and B.2 present the uncertainties in the normalized ϕ^* cross section in data unfolded with MADGRAPH and with POWHEG + PYTHIA6 (`Z2star`). Tables B.3 and B.4 present the uncertainties of the normalized ϕ^* cross section in MADGRAPH and

POWHEG MC, respectively. Tables B.5 and B.6 presents the uncertainties in the absolute ϕ^* cross section in data unfolded with MADGRAPH and with POWHEG + PYTHIA6 (`Z2star`). Tables B.7 and B.8 present the uncertainties on the absolute ϕ^* cross section in MADGRAPH and POWHEG MC, respectively.

ϕ^* Range	Total	Stat.	Total Syst.	MC Stat.	Pileup	SF	p_{T} Scale	Bkg.
0.000–0.004	0.43	0.26	0.34	0.33	0.04	0.07	0.01	0.05
0.004–0.008	0.50	0.28	0.41	0.40	0.03	0.07	0.02	0.05
0.008–0.012	0.47	0.29	0.37	0.36	0.02	0.07	0.00	0.04
0.012–0.016	0.50	0.29	0.41	0.40	0.02	0.07	0.01	0.04
0.016–0.020	0.51	0.29	0.42	0.39	0.14	0.07	0.00	0.04
0.020–0.024	0.51	0.30	0.41	0.39	0.10	0.07	0.02	0.04
0.024–0.029	0.44	0.27	0.35	0.33	0.03	0.07	0.01	0.04
0.029–0.034	0.46	0.28	0.37	0.36	0.04	0.07	0.01	0.04
0.034–0.039	0.48	0.28	0.38	0.37	0.04	0.06	0.02	0.04
0.039–0.045	0.43	0.27	0.34	0.33	0.01	0.06	0.00	0.05
0.045–0.052	0.42	0.25	0.34	0.33	0.01	0.06	0.00	0.04
0.052–0.057	0.52	0.32	0.40	0.40	0.01	0.05	0.01	0.05
0.057–0.064	0.45	0.28	0.35	0.35	0.02	0.05	0.01	0.04
0.064–0.072	0.45	0.27	0.36	0.33	0.11	0.05	0.01	0.04
0.072–0.081	0.43	0.26	0.34	0.33	0.01	0.04	0.02	0.04
0.081–0.091	0.42	0.26	0.33	0.32	0.02	0.03	0.00	0.04
0.091–0.102	0.43	0.27	0.34	0.33	0.07	0.03	0.00	0.04
0.102–0.114	0.44	0.27	0.35	0.34	0.05	0.02	0.00	0.04
0.114–0.128	0.43	0.27	0.33	0.33	0.03	0.03	0.01	0.03
0.128–0.145	0.43	0.26	0.34	0.32	0.10	0.03	0.01	0.03
0.145–0.165	0.42	0.26	0.33	0.32	0.04	0.04	0.00	0.03
0.165–0.189	0.42	0.26	0.33	0.32	0.05	0.05	0.02	0.03
0.189–0.219	0.43	0.26	0.33	0.32	0.04	0.06	0.01	0.06
0.219–0.258	0.42	0.26	0.33	0.32	0.04	0.08	0.01	0.04
0.258–0.312	0.42	0.26	0.33	0.31	0.03	0.10	0.01	0.05
0.312–0.391	0.43	0.26	0.35	0.31	0.04	0.13	0.00	0.07
0.391–0.524	0.45	0.26	0.37	0.31	0.00	0.15	0.00	0.12
0.524–0.695	0.56	0.32	0.46	0.38	0.04	0.19	0.02	0.18
0.695–0.918	0.73	0.39	0.62	0.47	0.20	0.23	0.01	0.27
0.918–1.153	0.95	0.53	0.79	0.65	0.10	0.27	0.05	0.33
1.153–1.496	1.07	0.61	0.87	0.71	0.08	0.31	0.06	0.39
1.496–1.947	1.27	0.77	1.02	0.85	0.00	0.33	0.05	0.45
1.947–2.522	1.73	0.98	1.43	1.10	0.64	0.35	0.00	0.54
2.522–3.277	2.02	1.21	1.62	1.36	0.56	0.34	0.10	0.59

Table B.1: The uncertainties (in %) for the normalized cross section measurement made with data unfolded with MadGraph.

ϕ^* Range	Total	Stat.	Total Syst.	MC Stat.	Pileup	SF	p_{T} Scale	Bkg.	PDF
0.000–0.004	0.82	0.26	0.77	0.77	0.05	0.07	0.01	0.05	0.01
0.004–0.008	0.93	0.28	0.88	0.87	0.03	0.07	0.02	0.05	0.11
0.008–0.012	0.99	0.28	0.94	0.91	0.21	0.07	0.00	0.04	0.08
0.012–0.016	1.03	0.29	0.99	0.98	0.08	0.07	0.01	0.04	0.10
0.016–0.020	0.89	0.29	0.84	0.83	0.03	0.07	0.00	0.04	0.03
0.020–0.024	0.96	0.30	0.91	0.90	0.14	0.06	0.02	0.04	0.05
0.024–0.029	0.91	0.27	0.87	0.83	0.23	0.07	0.01	0.04	0.03
0.029–0.034	0.84	0.27	0.80	0.78	0.10	0.06	0.01	0.04	0.08
0.034–0.039	0.88	0.29	0.83	0.82	0.08	0.06	0.02	0.04	0.09
0.039–0.045	0.77	0.26	0.72	0.72	0.02	0.06	0.00	0.04	0.01
0.045–0.052	0.77	0.25	0.73	0.69	0.20	0.06	0.00	0.04	0.13
0.052–0.057	0.99	0.32	0.94	0.93	0.05	0.05	0.01	0.05	0.03
0.057–0.064	0.81	0.28	0.77	0.75	0.08	0.05	0.01	0.04	0.08
0.064–0.072	0.78	0.27	0.74	0.72	0.12	0.05	0.01	0.04	0.06
0.072–0.081	0.77	0.26	0.73	0.72	0.02	0.04	0.02	0.04	0.06
0.081–0.091	0.75	0.26	0.70	0.69	0.05	0.04	0.00	0.04	0.05
0.091–0.102	0.76	0.27	0.71	0.71	0.03	0.03	0.00	0.04	0.04
0.102–0.114	0.78	0.27	0.73	0.72	0.07	0.03	0.00	0.04	0.05
0.114–0.128	0.76	0.27	0.71	0.70	0.02	0.03	0.01	0.03	0.03
0.128–0.145	0.76	0.26	0.71	0.71	0.08	0.03	0.01	0.03	0.01
0.145–0.165	0.81	0.26	0.76	0.73	0.21	0.04	0.00	0.03	0.02
0.165–0.189	0.76	0.26	0.72	0.71	0.06	0.05	0.02	0.03	0.04
0.189–0.219	0.76	0.26	0.71	0.70	0.03	0.07	0.01	0.06	0.03
0.219–0.258	0.76	0.26	0.71	0.70	0.07	0.08	0.01	0.04	0.03
0.258–0.312	0.74	0.26	0.70	0.69	0.03	0.10	0.01	0.05	0.01
0.312–0.391	0.75	0.26	0.71	0.69	0.03	0.12	0.00	0.08	0.04
0.391–0.524	0.81	0.26	0.77	0.75	0.02	0.15	0.00	0.12	0.03
0.524–0.695	0.96	0.32	0.90	0.85	0.16	0.19	0.02	0.18	0.03
0.695–0.918	1.22	0.39	1.15	1.06	0.27	0.23	0.01	0.27	0.04
0.918–1.153	1.66	0.53	1.57	1.47	0.20	0.26	0.05	0.33	0.28
1.153–1.496	1.88	0.61	1.78	1.70	0.08	0.30	0.06	0.39	0.17
1.496–1.947	2.34	0.77	2.21	2.11	0.28	0.33	0.05	0.45	0.15
1.947–2.522	2.99	0.98	2.82	2.75	0.03	0.33	0.00	0.54	0.18
2.522–3.277	3.64	1.20	3.43	3.17	1.12	0.33	0.10	0.58	0.18

Table B.2: The uncertainties (in %) for the normalized cross section measurement made with data unfolded with POWHEG + PYTHIA6 (`Z2star`).

ϕ^* Range	Total	Stat.	FSR
0.000–0.004	0.27	0.26	0.03
0.004–0.008	0.27	0.27	0.03
0.008–0.012	0.27	0.27	0.03
0.012–0.016	0.27	0.27	0.03
0.016–0.020	0.27	0.27	0.03
0.020–0.024	0.28	0.28	0.02
0.024–0.029	0.26	0.25	0.03
0.029–0.034	0.26	0.26	0.02
0.034–0.039	0.27	0.27	0.02
0.039–0.045	0.25	0.25	0.02
0.045–0.052	0.25	0.25	0.01
0.052–0.057	0.30	0.30	0.01
0.057–0.064	0.27	0.27	0.01
0.064–0.072	0.26	0.26	0.00
0.072–0.081	0.26	0.26	0.00
0.081–0.091	0.26	0.26	0.01
0.091–0.102	0.27	0.27	0.01
0.102–0.114	0.27	0.27	0.01
0.114–0.128	0.27	0.27	0.02
0.128–0.145	0.27	0.27	0.02
0.145–0.165	0.27	0.27	0.02
0.165–0.189	0.27	0.27	0.02
0.189–0.219	0.27	0.27	0.02
0.219–0.258	0.28	0.27	0.03
0.258–0.312	0.28	0.27	0.03
0.312–0.391	0.28	0.28	0.03
0.391–0.524	0.28	0.28	0.03
0.524–0.695	0.34	0.34	0.03
0.695–0.918	0.42	0.42	0.03
0.918–1.153	0.57	0.57	0.03
1.153–1.496	0.65	0.65	0.03
1.496–1.947	0.80	0.80	0.03
1.947–2.522	1.02	1.02	0.04
2.522–3.277	1.26	1.26	0.04

Table B.3: The uncertainties (in %) for the normalized cross section from the MadGraph MC sample.

ϕ^* Range	Total	Stat.	PDF	FSR
0.000–0.004	0.63	0.62	0.13	0.02
0.004–0.008	0.63	0.62	0.13	0.02
0.008–0.012	0.65	0.63	0.15	0.02
0.012–0.016	0.65	0.63	0.15	0.02
0.016–0.020	0.65	0.64	0.13	0.02
0.020–0.024	0.65	0.64	0.13	0.02
0.024–0.029	0.60	0.58	0.12	0.03
0.029–0.034	0.60	0.59	0.09	0.02
0.034–0.039	0.62	0.61	0.10	0.01
0.039–0.045	0.58	0.57	0.09	0.02
0.045–0.052	0.55	0.54	0.08	0.02
0.052–0.057	0.68	0.67	0.11	0.02
0.057–0.064	0.60	0.59	0.11	0.01
0.064–0.072	0.59	0.57	0.14	0.01
0.072–0.081	0.57	0.57	0.06	0.00
0.081–0.091	0.57	0.57	0.04	0.00
0.091–0.102	0.61	0.58	0.18	0.00
0.102–0.114	0.60	0.60	0.09	0.01
0.114–0.128	0.61	0.60	0.12	0.02
0.128–0.145	0.60	0.59	0.11	0.02
0.145–0.165	0.62	0.60	0.16	0.02
0.165–0.189	0.64	0.61	0.19	0.01
0.189–0.219	0.65	0.61	0.23	0.03
0.219–0.258	0.63	0.61	0.16	0.02
0.258–0.312	0.63	0.61	0.15	0.03
0.312–0.391	0.65	0.62	0.20	0.03
0.391–0.524	0.67	0.63	0.24	0.03
0.524–0.695	0.81	0.78	0.23	0.02
0.695–0.918	1.05	0.97	0.41	0.03
0.918–1.153	1.47	1.34	0.60	0.05
1.153–1.496	1.74	1.57	0.76	0.03
1.496–1.947	2.15	2.00	0.78	0.04
1.947–2.522	2.76	2.55	1.07	0.06
2.522–3.277	3.26	3.04	1.17	0.03

Table B.4: The uncertainties (in %) for the normalized cross section from the POWHEG MC sample.

ϕ^* Range	Total	Stat.	Total Syst.	MC Stat.	Pileup	SF	p_T Scale	Bkg.
0.000–0.004	2.72	0.26	2.70	0.33	0.48	0.43	0.16	0.02
0.004–0.008	2.72	0.28	2.70	0.40	0.41	0.43	0.17	0.05
0.008–0.012	2.71	0.29	2.69	0.36	0.39	0.43	0.15	0.03
0.012–0.016	2.72	0.29	2.71	0.40	0.43	0.43	0.16	0.03
0.016–0.020	2.75	0.29	2.73	0.39	0.58	0.43	0.15	0.03
0.020–0.024	2.71	0.30	2.69	0.39	0.33	0.43	0.17	0.04
0.024–0.029	2.72	0.27	2.70	0.34	0.47	0.44	0.15	0.04
0.029–0.034	2.72	0.28	2.71	0.36	0.48	0.43	0.14	0.03
0.034–0.039	2.72	0.29	2.71	0.37	0.48	0.43	0.17	0.03
0.039–0.045	2.71	0.27	2.70	0.33	0.45	0.44	0.15	0.03
0.045–0.052	2.71	0.25	2.70	0.33	0.45	0.44	0.13	0.03
0.052–0.057	2.73	0.32	2.71	0.40	0.45	0.43	0.15	0.03
0.057–0.064	2.71	0.28	2.70	0.35	0.41	0.44	0.16	0.03
0.064–0.072	2.73	0.27	2.72	0.34	0.56	0.44	0.16	0.02
0.072–0.081	2.71	0.26	2.70	0.33	0.45	0.44	0.16	0.04
0.081–0.091	2.71	0.27	2.70	0.32	0.46	0.44	0.14	0.03
0.091–0.102	2.72	0.27	2.71	0.33	0.51	0.44	0.15	0.03
0.102–0.114	2.72	0.27	2.71	0.34	0.49	0.43	0.14	0.04
0.114–0.128	2.72	0.27	2.70	0.33	0.47	0.44	0.15	0.05
0.128–0.145	2.70	0.26	2.69	0.32	0.37	0.43	0.16	0.07
0.145–0.165	2.70	0.26	2.69	0.32	0.41	0.43	0.15	0.05
0.165–0.189	2.70	0.26	2.69	0.32	0.39	0.43	0.14	0.05
0.189–0.219	2.70	0.26	2.69	0.32	0.41	0.43	0.14	0.10
0.219–0.258	2.72	0.26	2.70	0.32	0.48	0.43	0.16	0.09
0.258–0.312	2.71	0.26	2.69	0.31	0.43	0.42	0.15	0.10
0.312–0.391	2.70	0.26	2.69	0.31	0.42	0.42	0.15	0.12
0.391–0.524	2.71	0.26	2.69	0.32	0.42	0.41	0.14	0.17
0.524–0.695	2.73	0.32	2.72	0.38	0.49	0.41	0.12	0.23
0.695–0.918	2.73	0.39	2.70	0.47	0.22	0.41	0.11	0.32
0.918–1.153	2.81	0.54	2.76	0.65	0.33	0.42	0.11	0.38
1.153–1.496	2.85	0.61	2.79	0.71	0.32	0.43	0.07	0.44
1.496–1.947	2.95	0.77	2.84	0.85	0.39	0.44	0.15	0.50
1.947–2.522	3.09	0.99	2.93	1.11	0.21	0.44	0.15	0.59
2.522–3.277	3.27	1.21	3.04	1.36	0.13	0.44	0.08	0.64

Table B.5: The uncertainties (in %) for the absolute cross section measurement made with data unfolded with MadGraph. The total value and the total systematic value includes the uncertainty of 2.6 % due to the luminosity.

ϕ^* Range	Total	Stat.	Total Syst.	MC Stat.	Pileup	SF	p_{T} Scale	Bkg.	PDF
0.000–0.004	2.79	0.26	2.78	0.77	0.36	0.44	0.16	0.03	0.16
0.004–0.008	2.85	0.28	2.83	0.87	0.46	0.44	0.17	0.05	0.24
0.008–0.012	2.88	0.29	2.87	0.91	0.63	0.44	0.15	0.03	0.11
0.012–0.016	2.89	0.29	2.87	0.98	0.51	0.44	0.16	0.03	0.23
0.016–0.020	2.82	0.29	2.80	0.83	0.40	0.44	0.15	0.03	0.14
0.020–0.024	2.87	0.30	2.85	0.90	0.57	0.44	0.17	0.04	0.19
0.024–0.029	2.79	0.27	2.78	0.83	0.17	0.44	0.15	0.04	0.16
0.029–0.034	2.79	0.28	2.77	0.78	0.31	0.44	0.14	0.03	0.11
0.034–0.039	2.83	0.29	2.81	0.82	0.51	0.43	0.17	0.03	0.12
0.039–0.045	2.79	0.26	2.77	0.72	0.42	0.44	0.15	0.03	0.15
0.045–0.052	2.81	0.25	2.80	0.69	0.63	0.44	0.13	0.03	0.11
0.052–0.057	2.85	0.32	2.83	0.93	0.38	0.44	0.15	0.03	0.14
0.057–0.064	2.81	0.28	2.80	0.75	0.51	0.44	0.16	0.03	0.12
0.064–0.072	2.81	0.27	2.80	0.72	0.55	0.44	0.16	0.02	0.12
0.072–0.081	2.79	0.26	2.77	0.72	0.42	0.44	0.16	0.04	0.13
0.081–0.091	2.79	0.27	2.78	0.69	0.47	0.45	0.14	0.03	0.18
0.091–0.102	2.78	0.27	2.77	0.71	0.38	0.44	0.15	0.03	0.18
0.102–0.114	2.80	0.27	2.79	0.72	0.50	0.44	0.14	0.04	0.19
0.114–0.128	2.79	0.27	2.77	0.70	0.44	0.44	0.15	0.05	0.16
0.128–0.145	2.80	0.26	2.79	0.71	0.51	0.44	0.16	0.07	0.15
0.145–0.165	2.77	0.26	2.76	0.73	0.26	0.45	0.15	0.05	0.14
0.165–0.189	2.78	0.26	2.77	0.71	0.39	0.44	0.14	0.05	0.13
0.189–0.219	2.78	0.26	2.77	0.70	0.42	0.44	0.14	0.10	0.16
0.219–0.258	2.77	0.26	2.76	0.71	0.35	0.43	0.16	0.09	0.16
0.258–0.312	2.78	0.26	2.77	0.69	0.46	0.43	0.15	0.10	0.15
0.312–0.391	2.78	0.26	2.77	0.69	0.40	0.42	0.15	0.12	0.17
0.391–0.524	2.80	0.26	2.78	0.75	0.44	0.42	0.14	0.17	0.13
0.524–0.695	2.81	0.32	2.79	0.85	0.22	0.41	0.12	0.23	0.15
0.695–0.918	2.89	0.39	2.87	1.06	0.15	0.42	0.11	0.32	0.12
0.918–1.153	3.10	0.53	3.06	1.48	0.20	0.41	0.11	0.38	0.16
1.153–1.496	3.28	0.61	3.22	1.70	0.50	0.42	0.07	0.44	0.27
1.496–1.947	3.58	0.77	3.49	2.12	0.71	0.44	0.15	0.50	0.06
1.947–2.522	4.00	0.99	3.88	2.75	0.37	0.44	0.15	0.59	0.09
2.522–3.277	4.40	1.20	4.23	3.17	0.70	0.43	0.08	0.64	0.04

Table B.6: The uncertainties (in %) for the absolute cross section measurement made with data unfolded with POWHEG + PYTHIA6 (`Z2star`). The total value and the total systematic value includes the uncertainty of 2.6 % due to the luminosity.

ϕ^* Range	Total	Stat.	Cross Section	FSR
0.000–0.004	3.32	0.26	3.30	0.27
0.004–0.008	3.32	0.27	3.30	0.27
0.008–0.012	3.32	0.27	3.30	0.27
0.012–0.016	3.32	0.27	3.30	0.27
0.016–0.020	3.32	0.27	3.30	0.27
0.020–0.024	3.32	0.28	3.30	0.27
0.024–0.029	3.32	0.25	3.30	0.27
0.029–0.034	3.32	0.26	3.30	0.28
0.034–0.039	3.32	0.27	3.30	0.28
0.039–0.045	3.32	0.25	3.30	0.28
0.045–0.052	3.32	0.25	3.30	0.29
0.052–0.057	3.33	0.30	3.30	0.29
0.057–0.064	3.32	0.27	3.30	0.29
0.064–0.072	3.32	0.26	3.30	0.29
0.072–0.081	3.32	0.26	3.30	0.30
0.081–0.091	3.32	0.26	3.30	0.30
0.091–0.102	3.32	0.27	3.30	0.30
0.102–0.114	3.33	0.27	3.30	0.31
0.114–0.128	3.33	0.27	3.30	0.31
0.128–0.145	3.33	0.27	3.30	0.32
0.145–0.165	3.33	0.27	3.30	0.32
0.165–0.189	3.33	0.27	3.30	0.32
0.189–0.219	3.33	0.27	3.30	0.32
0.219–0.258	3.33	0.27	3.30	0.32
0.258–0.312	3.33	0.27	3.30	0.32
0.312–0.391	3.33	0.28	3.30	0.32
0.391–0.524	3.33	0.28	3.30	0.33
0.524–0.695	3.33	0.34	3.30	0.32
0.695–0.918	3.34	0.42	3.30	0.32
0.918–1.153	3.36	0.57	3.30	0.33
1.153–1.496	3.38	0.65	3.30	0.33
1.496–1.947	3.41	0.80	3.30	0.33
1.947–2.522	3.47	1.02	3.30	0.34
2.522–3.277	3.55	1.26	3.30	0.34

Table B.7: The uncertainties (in %) for the absolute cross section from the MadGraph MC sample.

ϕ^* Range	Total	Stat.	PDF	FSR
0.000–0.004	2.72	0.62	2.63	0.27
0.004–0.008	2.71	0.62	2.63	0.27
0.008–0.012	2.68	0.63	2.59	0.27
0.012–0.016	2.75	0.63	2.66	0.28
0.016–0.020	2.71	0.64	2.62	0.28
0.020–0.024	2.74	0.64	2.65	0.28
0.024–0.029	2.73	0.58	2.65	0.27
0.029–0.034	2.70	0.59	2.62	0.27
0.034–0.039	2.72	0.61	2.64	0.28
0.039–0.045	2.69	0.57	2.61	0.28
0.045–0.052	2.70	0.54	2.63	0.28
0.052–0.057	2.75	0.67	2.66	0.28
0.057–0.064	2.74	0.59	2.66	0.29
0.064–0.072	2.75	0.57	2.67	0.29
0.072–0.081	2.71	0.57	2.63	0.30
0.081–0.091	2.70	0.57	2.62	0.30
0.091–0.102	2.75	0.58	2.67	0.30
0.102–0.114	2.72	0.60	2.63	0.31
0.114–0.128	2.72	0.60	2.64	0.32
0.128–0.145	2.71	0.59	2.63	0.31
0.145–0.165	2.71	0.60	2.62	0.31
0.165–0.189	2.73	0.61	2.65	0.31
0.189–0.219	2.73	0.61	2.64	0.33
0.219–0.258	2.67	0.61	2.58	0.32
0.258–0.312	2.66	0.61	2.57	0.32
0.312–0.391	2.61	0.62	2.52	0.33
0.391–0.524	2.57	0.63	2.47	0.32
0.524–0.695	2.65	0.78	2.51	0.32
0.695–0.918	2.65	0.97	2.44	0.33
0.918–1.153	2.79	1.34	2.42	0.35
1.153–1.496	2.91	1.57	2.43	0.33
1.496–1.947	3.20	2.00	2.47	0.33
1.947–2.522	3.62	2.55	2.55	0.35
2.522–3.277	4.03	3.04	2.62	0.33

Table B.8: The uncertainties (in %) for the absolute cross section from the POWHEG MC sample.

Appendix C

QCD multi-jet and W + jets Background Fits

The results of the fits in each ϕ^* bin to determine the QCD multi-jet and W + jets background contributions, as discussed in section 6.3.2, are presented below. Table C.1 presents the background in each ϕ^* bin due to QCD multi-jet and W + jets as estimated from these fits.

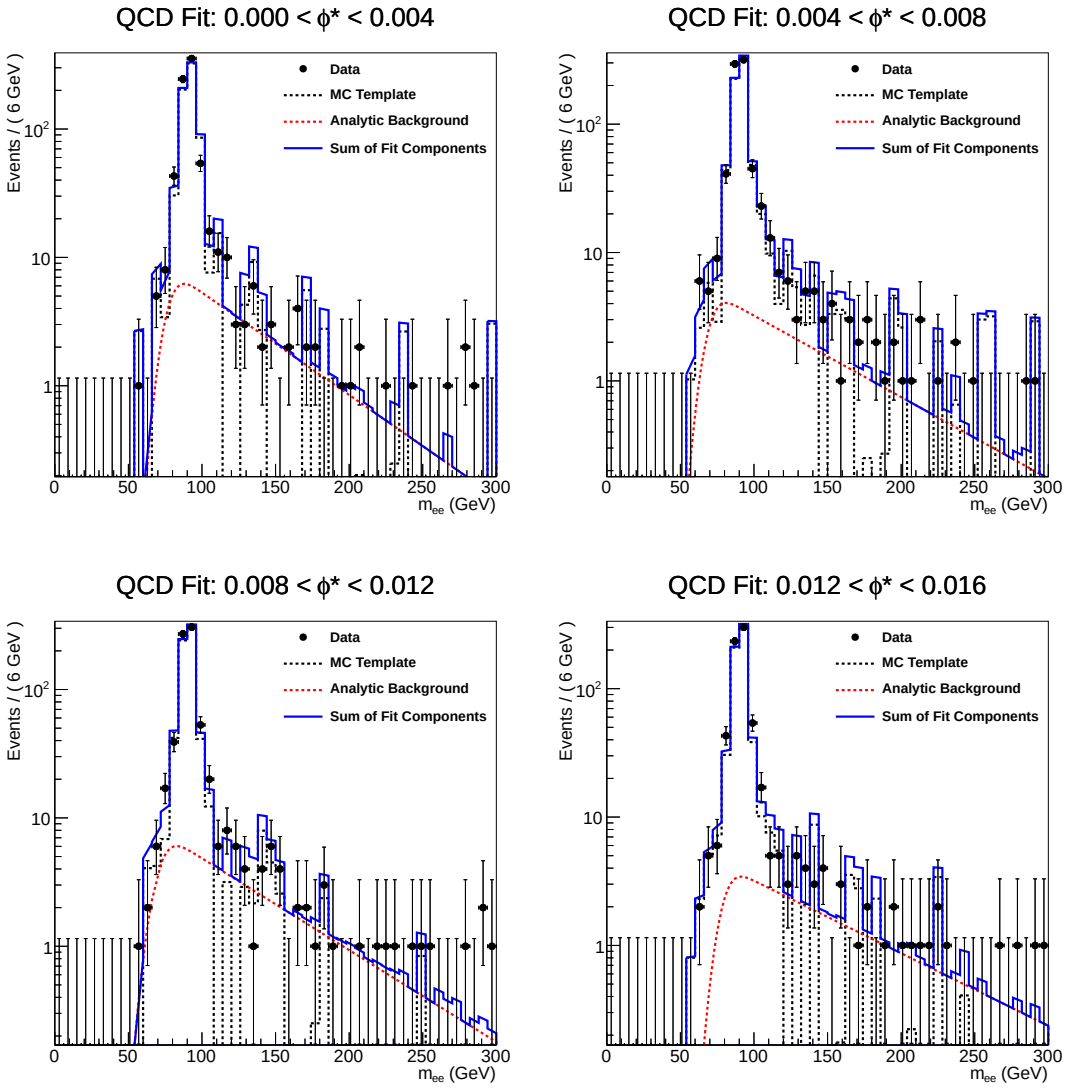


Figure C.1: The QCD multi-jet and W + jets data-driven background fits for the first set of four ϕ^* bins. The data are shown as points with error bars, MC template as a dashed histogram, the analytic background function as the dashed line, and the sum of the template and function as a solid histogram.

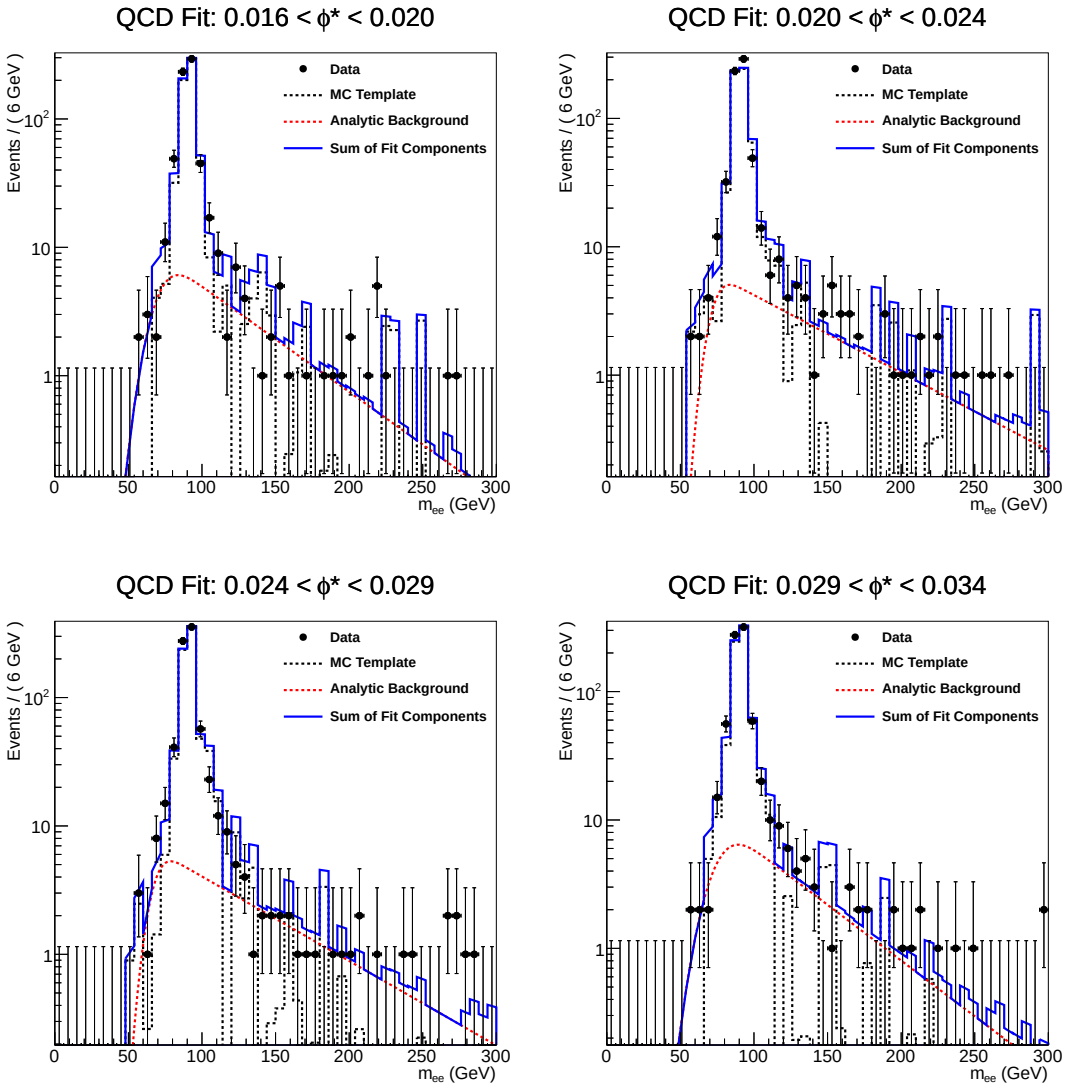


Figure C.2: The QCD multi-jet and W+jets data-driven background fits for the second set of four ϕ^* bins. The data are shown as points with error bars, MC template as a dashed histogram, the analytic background function as the dashed line, and the sum of the template and function as a solid histogram.

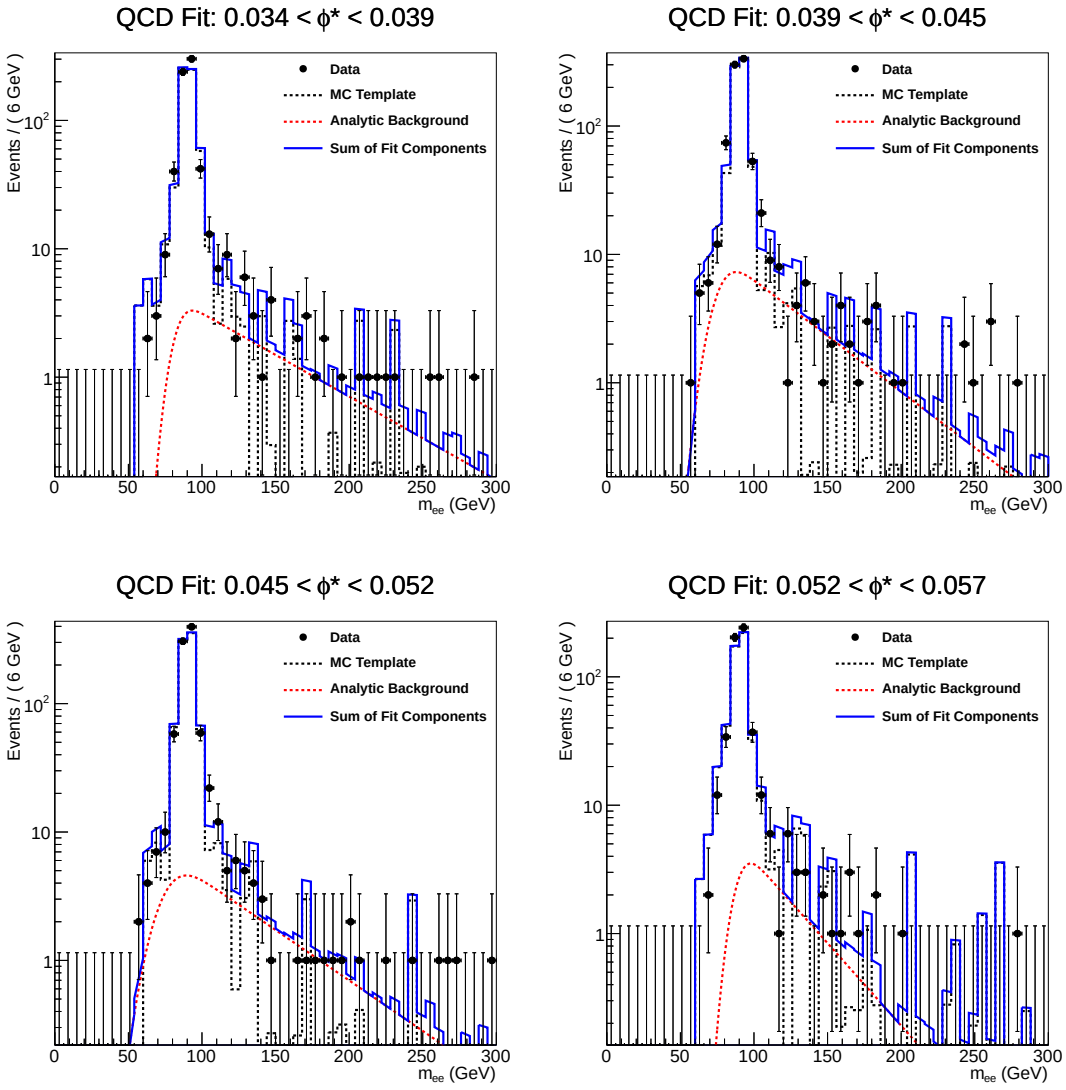


Figure C.3: The QCD multi-jet and W + jets data-driven background fits for the third set of four ϕ^* bins. The data are shown as points with error bars, MC template as a dashed histogram, the analytic background function as the dashed line, and the sum of the template and function as a solid histogram.

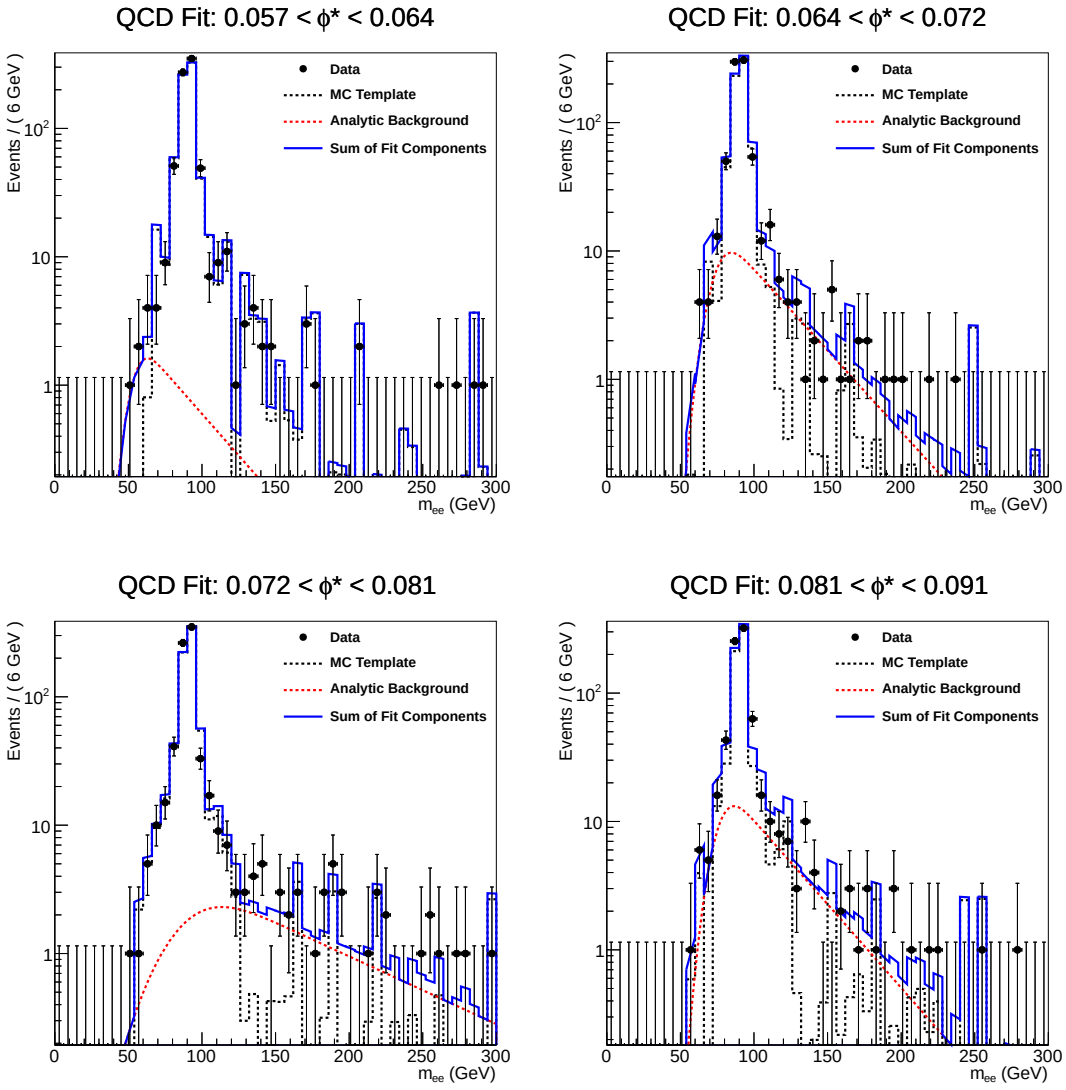


Figure C.4: The QCD multi-jet and W + jets data-driven background fits for the forth set of four ϕ^* bins. The data are shown as points with error bars, MC template as a dashed histogram, the analytic background function as the dashed line, and the sum of the template and function as a solid histogram.

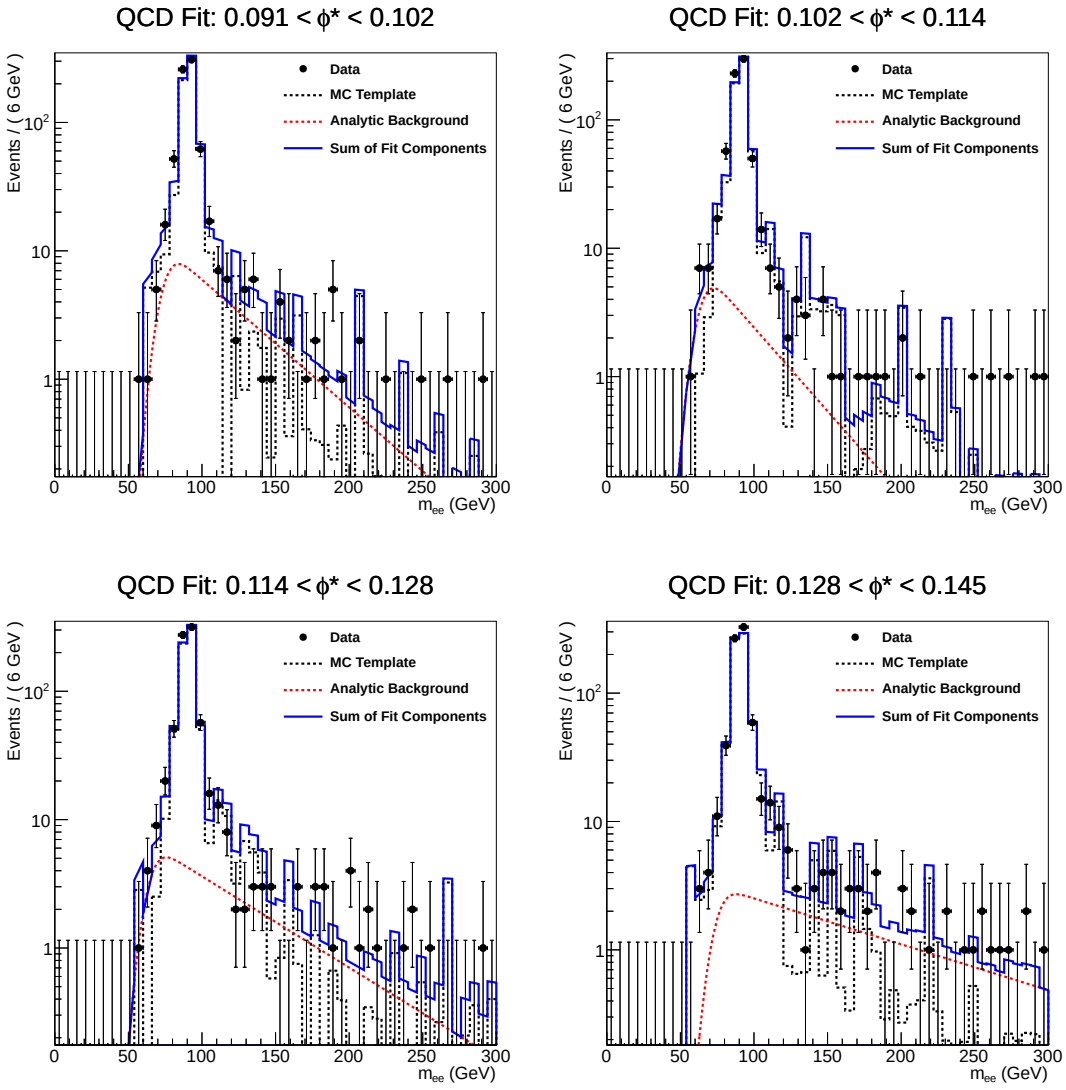


Figure C.5: The QCD multi-jet and W + jets data-driven background fits for the fifth set of four ϕ^* bins. The data are shown as points with error bars, MC template as a dashed histogram, the analytic background function as the dashed line, and the sum of the template and function as a solid histogram.

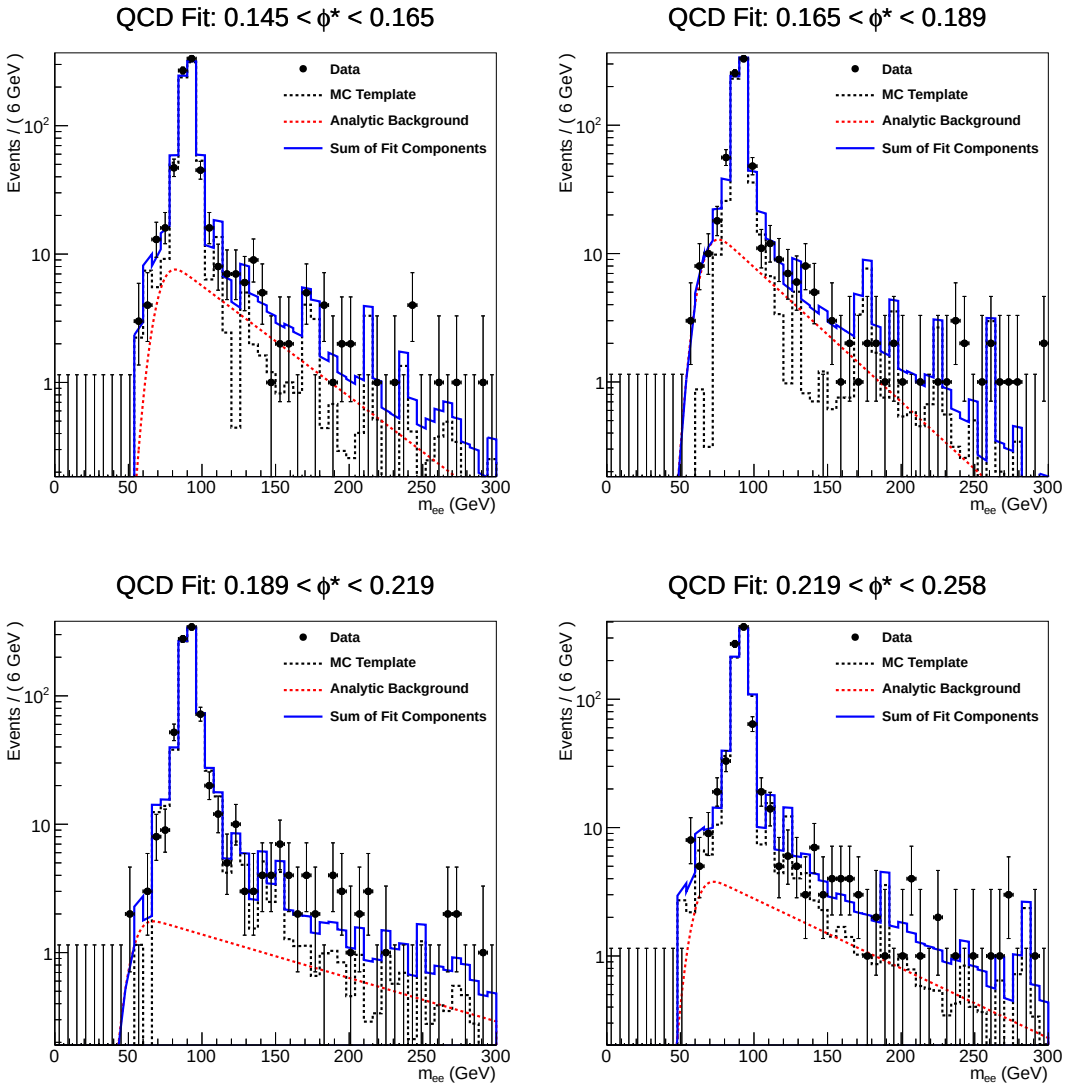


Figure C.6: The QCD multi-jet and W + jets data-driven background fits for the sixth set of four ϕ^* bins. The data are shown as points with error bars, MC template as a dashed histogram, the analytic background function as the dashed line, and the sum of the template and function as a solid histogram.

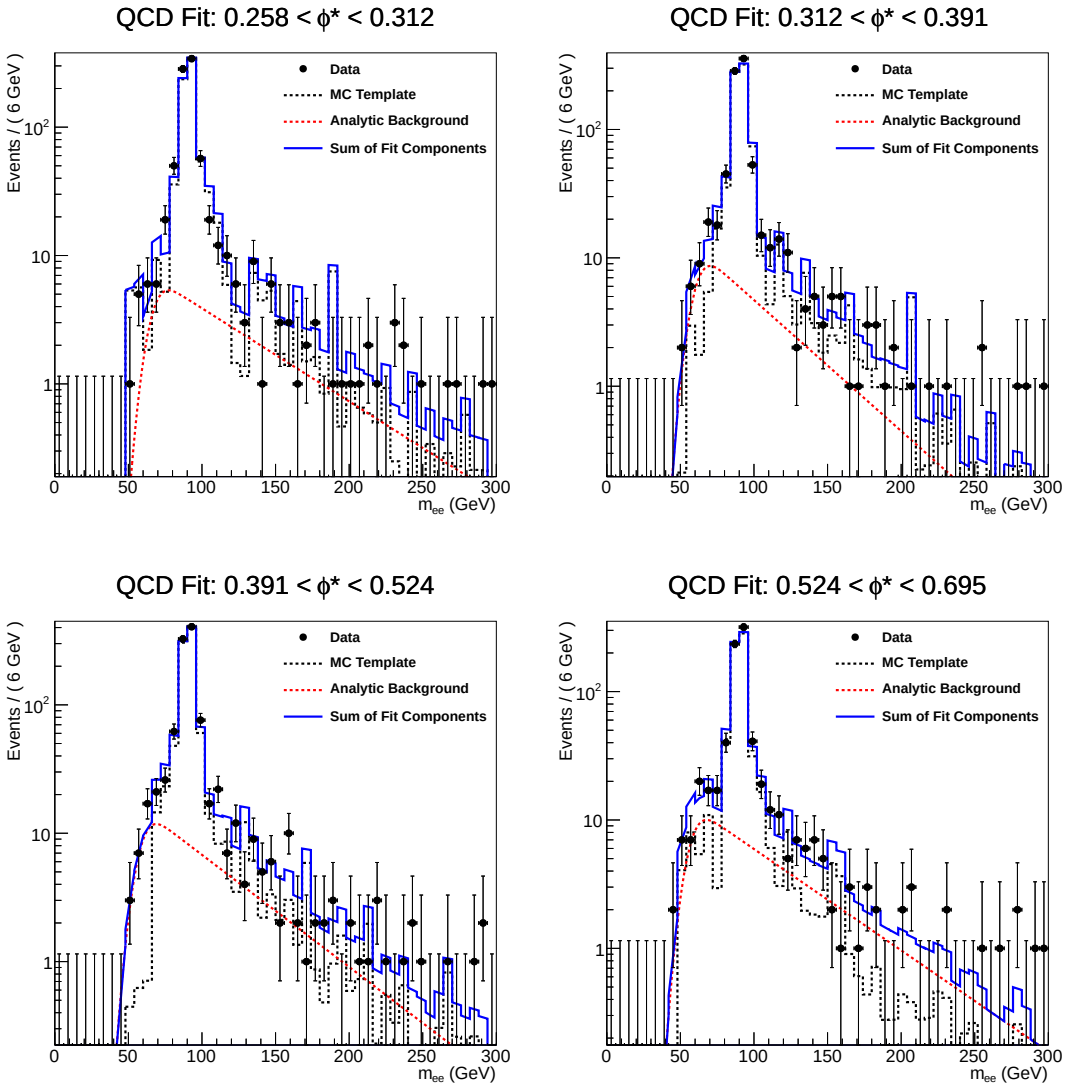


Figure C.7: The QCD multi-jet and W+jets data-driven background fits for the seventh set of four ϕ^* bins. The data are shown as points with error bars, MC template as a dashed histogram, the analytic background function as the dashed line, and the sum of the template and function as a solid histogram.

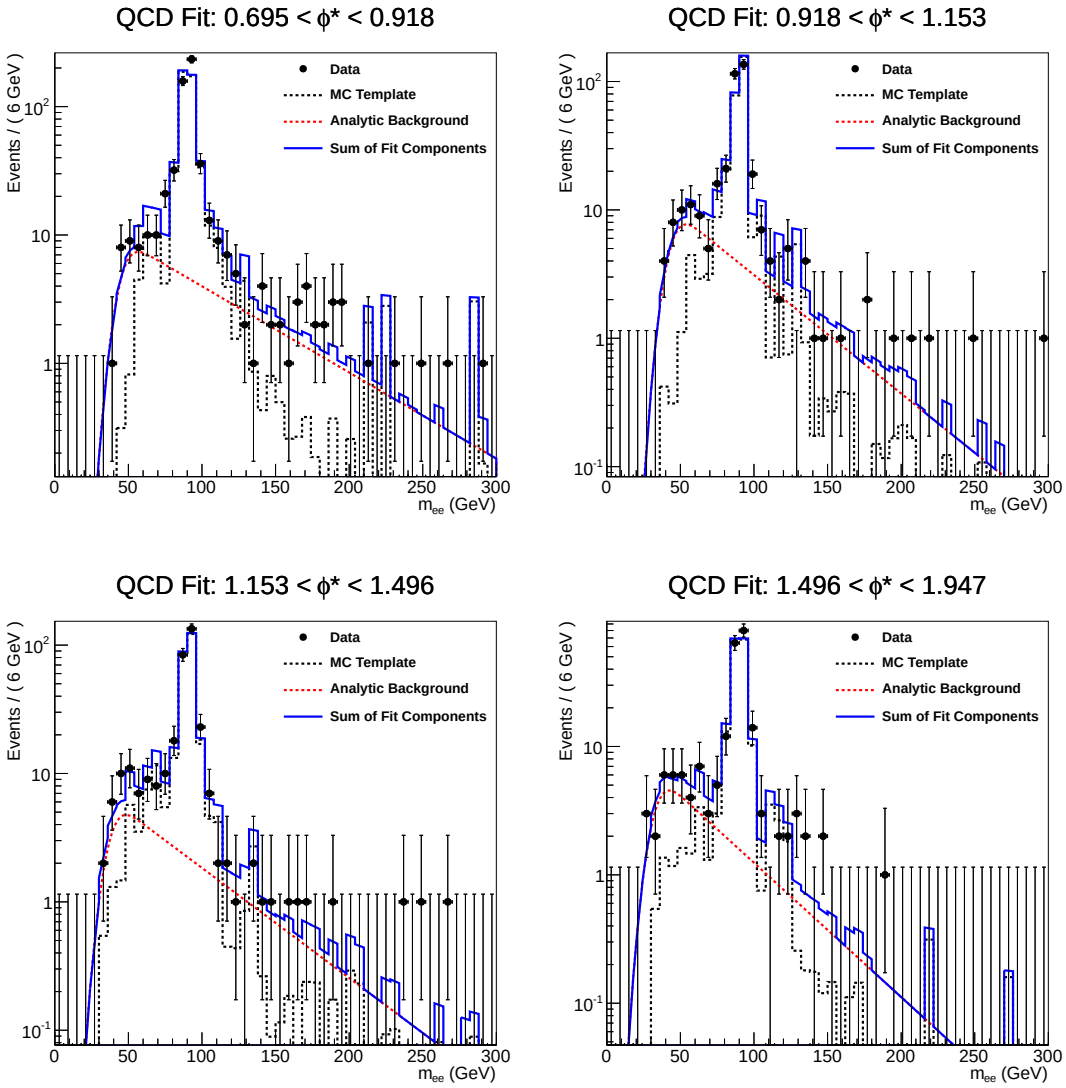


Figure C.8: The QCD multi-jet and W + jets data-driven background fits for the eighth set of four ϕ^* bins. The data are shown as points with error bars, MC template as a dashed histogram, the analytic background function as the dashed line, and the sum of the template and function as a solid histogram.

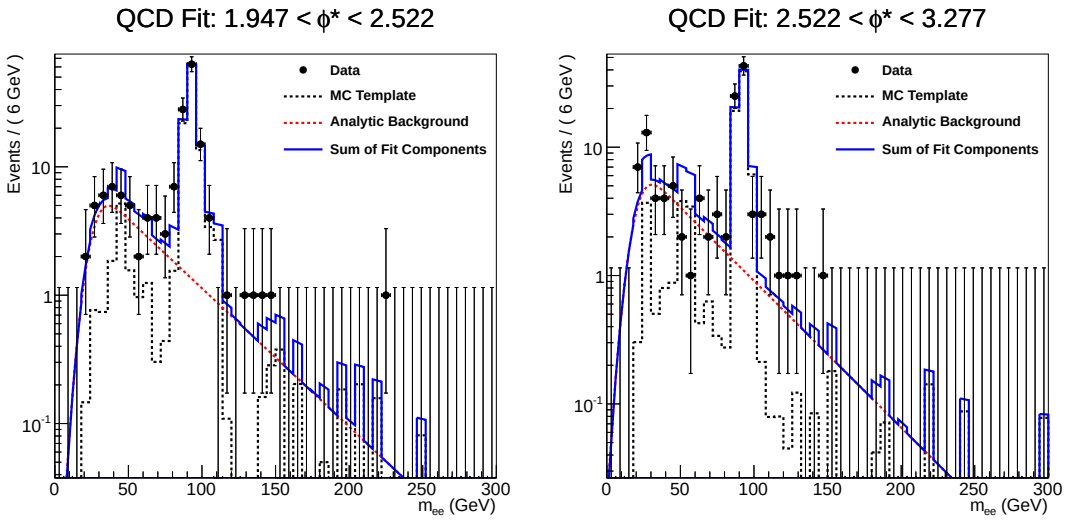


Figure C.9: The QCD multi-jet and W + jets data-driven background fits for the last two ϕ^* bins. The data are shown as points with error bars, MC template as a dashed histogram, the analytic background function as the dashed line, and the sum of the template and function as a solid histogram.

ϕ^* Range	Event Count
0.000–0.004	29 ± 12
0.004–0.008	52 ± 18
0.008–0.012	42 ± 19
0.012–0.016	46 ± 20
0.016–0.020	34 ± 16
0.020–0.024	52 ± 19
0.024–0.029	52 ± 19
0.029–0.034	26 ± 14
0.034–0.039	35 ± 18
0.039–0.045	26 ± 13
0.045–0.052	32 ± 20
0.052–0.057	6 ± 59
0.057–0.064	18 ± 44
0.064–0.072	12 ± 38
0.072–0.081	37 ± 25
0.081–0.091	10 ± 3
0.091–0.102	21 ± 9
0.102–0.114	15 ± 6
0.114–0.128	50 ± 21
0.128–0.145	84 ± 43
0.145–0.165	31 ± 15
0.165–0.189	24 ± 13
0.189–0.219	119 ± 42
0.219–0.258	75 ± 31
0.258–0.312	47 ± 24
0.312–0.391	29 ± 14
0.391–0.524	39 ± 11
0.524–0.695	47 ± 12
0.695–0.918	62 ± 12
0.918–1.153	38 ± 10
1.153–1.496	43 ± 13
1.496–1.947	30 ± 12
1.947–2.522	28 ± 10
2.522–3.277	25 ± 9

Table C.1: The estimated number of QCD multi-jet and $\mathrm{W}+\mathrm{jets}$ events in each ϕ^* bin from the fits.

Appendix D

Glossary and Acronyms

Every occupation develops its own jargon, and while this leads to concise communication between members of the group, it often hinders comprehension by the uninitiated. Where possible, jargon has been minimized, but as with any technical publication, some is unavoidable. To aid the understanding of the reader, commonly used terms and acronyms have been defined below.

D.1 Glossary

Background sample: The MC samples simulating the various backgrounds to the analysis, for example, $t\bar{t}$.

Barrel: The central region in η of the detector.

Compact Muon Solenoid (CMS): The detector that collected the data used in this thesis. See section 3.2.

Endcap: The portion of each subdetector that is flat and covers the high $|\eta|$ regions.

Generator Level: The information about an MC event as determined by the MC generator, before the event is passed through a detector simulation.

Hadronization: The process by which color charge is hidden from observation by producing colorless hadrons.

Hadron: Color neutral combinations of three quarks, for example protons and neutrons.

Interaction Point: The region at the center of the detector where proton-proton collisions occur.

Jet: A spray of high energy particles that originate from a colored object as it tries to maintain its colorless state.

Large Hadron Collider (LHC): The collider used to produce the data used in this thesis. See section 3.1.

Molière Radius: The radius in which 90 % of the energy of an electromagnetic shower is contained within for a given material.

Monte Carlo (MC): Simulated data. See section 5.2.

Particle Flow: An algorithm for reconstruction particles using information from multiple subdetectors. See section 4.4.3.

Parton: The individual constituents of a proton including the valence quarks, gluons, and sea quarks. See section 2.2.2.

Pileup: Additional proton-proton interactions which occur during an event.

Prescaled: To reduce the rate of a trigger by randomly throwing out events the trigger accepted.

Primary Vertex: The reconstructed location of the proton-proton interaction.

Reconstructed Level: The information about an event determined from the detector (for data) or the simulation of the detector response (in MC).

Reconstruction: The process of taking raw data from the detector and creating objects useful for physics. See chapter 4.

Sea Quarks: Pairs of quarks and antiquarks from gluon splitting that exist within each hadron.

Signal Sample: The MC samples simulating the signal ($DY \to e^+e^-$) that this analysis looks for, as opposed to background samples.

Simulation: See *Monte Carlo*.

Toy Monte Carlo: A highly simplified Monte Carlo sample.

Transverse Momentum: Momentum transverse to the beamline. In general denoted p_T, but Q_T is used specifically to mean the transverse momentum of the Z or W boson.

Trigger: A system that analyses events as they are happening and decides which ones to keep. See section 3.2.6.

Truth Level: See *Generator Level*.

Valence Quark: The quarks which give rise to the quantum numbers of the proton, specifically the two up quarks and the down quark.

D.2 Acronyms

Table D.1: Acronyms

Acronym	Meaning
CB	Crystal ball
CERN	Originally from *Conseil Européen pour la Recherche Nucléaire*, now the European Organization for Nuclear Research
CMS	Compact Muon Solenoid
CSC	Cathode stripe chambers
DAQ	Data acquisition
DT	Drift tubes
DY	Drell–Yan
EB	Electromagnetic calorimeter barrel
ECAL	Electromagnetic calorimeter

Continued on next page

Table D.1 – Continued from previous page

Acronym	Meaning
EE	Electromagnetic calorimeter endcap
ES	Electromagnetic calorimeter preshower
FNAL	Fermi National Accelerator Laboratory
FSR	Final state radiation
GSF	Gaussian-sum filter
HB	Hadronic calorimeter barrel
HB	Hadronic calorimeter endcap
HCAL	Hadronic calorimeter
HF	Forward hadronic calorimeter
HLT	High-level trigger
HO	Hadronic calorimeter outer
ID	Electron identification
ISR	Initial state radiation
L1	Level-1 trigger
LHC	Large Hadron Collider
LO	Leading order
MC	Monte Carlo
NLO	Next-to-leading order
NNLO	Next-to-next-to-leading order
PDF	Parton distribution function
PSB	Proton synchrotron booster
PS	Proton synchrotron
QCD	Quantum Chromodynamics
QED	Quantum Electrodynamics
RPC	Resistive place chambers
SLAC	Stanford Linear Accelerator Center
SPS	Super Proton Synchrotron
T&P	Tag and probe

Continued on next page

Table D.1 – Continued from previous page

Acronym	Meaning
TEC	Tracker endcap
TIB	Tracker inner barrel
TID	Tracker inner disk
TOB	Tracker outer barrel