

---

# Measurement of the electron identification efficiency for $Z \rightarrow ee$ using ATLAS 2015 data

*DESY Summer Student Programme, 2017*

Kathryn Coldham

*Queen Mary University of London, United Kingdom*

Supervisor

Dr. Thorsten Kuhl

PhD students

Baishali Dutta and Akanksha Vishwakarma



September 4, 2017

## Abstract

Working with the DESY ATLAS Group, the data and Monte Carlo electron identification efficiencies were determined for  $Z \rightarrow ee$ . This was for 2015 ATLAS data for pp collisions at a centre of mass energy of  $\sqrt{s} = 13$  TeV. The efficiencies were used to calculate the scale factor for the Monte Carlo (MC) simulation so that the MC simulation can be matched to the data and be used to model physics processes.

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The Large Hadron Collider and the CERN Injector Chain</b>	<b>1</b>
<b>3</b>	<b>The ATLAS Detector</b>	<b>2</b>
<b>4</b>	<b>ATLAS Coordinate System</b>	<b>3</b>
<b>5</b>	<b>Electron Selection</b>	<b>4</b>
<b>6</b>	<b>Electron Identification Efficiency and Monte Carlo Scale Factor</b>	<b>5</b>
<b>7</b>	<b>Results</b>	<b>6</b>
7.1	Z Boson: Data and Monte Carlo Comparison Plots . . . . .	6
7.2	Electron Candidates: Data and Monte Carlo Comparison Plots . . . . .	9
7.3	Efficiency and Scale Factor Plots . . . . .	12
<b>8</b>	<b>Conclusions</b>	<b>14</b>

## 1 Introduction

For this study, the data samples used were part of the 13 TeV ATLAS data that was collected in 2015. These were associated with a bunch crossing of 25ns.

The electron candidates produced by the  $Z \rightarrow ee$  process were investigated. Electrons are a type of lepton that leave a clean signal in the ATLAS detector due to having little background. They are present in many physics processes so determining the electron identification efficiencies for Monte Carlo (MC) and data are important. To determine these efficiencies, loose, medium and tight identification cuts were applied to the reconstructed electrons for  $Z \rightarrow ee$ . The efficiencies were plotted as functions of the pseudorapidity  $\eta$  and the transverse momentum  $P_T$  for each identification cut. The  $\eta$  and  $P_T$  dependent scale factors for each identification cut were then found by calculating the ratio between the data identification efficiency and the MC identification efficiency. The scale factors can be used to match data and MC so that the MC simulation can be used to model other physics processes.

## 2 The Large Hadron Collider and the CERN Injector Chain

The Large Hadron Collider (LHC) is a 27-kilometre-long, circular accelerator and is part of European Organisation for Nuclear Research (CERN). It is situated astride the French-Swiss border and collides two counter-rotating beams of proton bunches, with each bunch containing  $\sim 10^{11}$  protons. The collisions were designed to have a centre of mass energy of 14 TeV and the LHC has a design luminosity of  $10^{34} \text{cm}^{-2} \text{s}^{-1}$  [1].

The LHC project was approved by the CERN council in 1994 and was built by re-using the linear accelerators, the Proton Synchrotron Booster, the Proton Synchrotron and Super Proton Synchrotron of its predecessor, the Large Electron-Positron Collider (LEP). These form part of the accelerator chain for the protons before they reach the LHC ring. Protons originate from a bottle of hydrogen gas found at one end of the Linac 2 linear accelerator. After passing through Linac 2, the protons experience an energy increase of 50MeV. The protons then pass through the Proton Synchrotron Booster, which consists of four superimposed synchrotron rings and accelerates protons to 1.4 GeV. They are then injected into the Proton Synchrotron, which accelerates them up to 25GeV. The proton beam then reaches the Super Proton Synchrotron, where they are accelerated up to 450GeV and are injected into the LHC ring. Once inside the LHC ring, the proton bunches go on to collide at four main points, at each an experiment has been built. These are: A Large Ion Collider Experiment (ALICE), A Toroidal LHC Apparatus (ATLAS), Compact Muon Solenoid (CMS) and Large Hadron Collider beauty experiment (LHCb) [2].

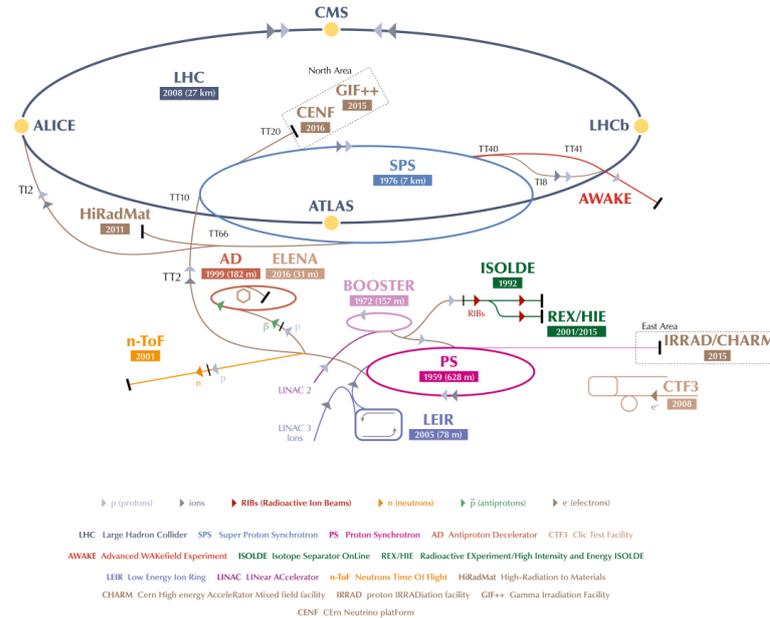


Figure 1: A schematic of the CERN Accelerator Complex [3].

### 3 The ATLAS Detector

The ATLAS (A Toroidal LHC ApparatuS) detector is a general-purpose experiment that is situated on the LHC ring and is designed to study a wide range of physics processes. At 46-metres-long and with a 25-metre-wide diameter, this cylindrically-shaped detector weighs around 7,000 tonnes. It consists of concentric layers that surround a central beam pipe, inside which protons collide. The products of these collisions are then detected using the detector technologies utilised in each of the concentric layers, which are: the Inner Detector, the Electromagnetic Calorimeter, the Hadronic Calorimeter and the Muon Spectrometer [4].

The Inner Detector (ID) is found closest to the beam axis and it is divided into three subsections: the pixel tracker, the semiconductor tracker (SCT) and the transition radiation tracker (TRT). The ID is used to obtain information about the direction of travel, momentum and charge of charged particles, such as the electron. The paths of the charged particles bend under the Lorentz force due to the action of a 2 T solenoidal magnetic field around the ID [4].

Surrounding the Inner Detector is the Electromagnetic Calorimeter (ECAL) and this is where particles that undergo the electromagnetic interaction, such as electrons, are detected; the energy deposit made is called a cluster. Lead was chosen as the absorber material due to its high density and this is interleaved with liquid argon in an accordion structure. Next is the Hadronic Calorimeter (HCAL), a tile calorimeter that is used to absorb particles that are subjected to the strong interaction, such as hadrons. Furthest

away from the beam axis is the Muon Spectrometer which is used to obtain signals when muons pass through. A 4 T toroidal magnetic field surrounds the Muon Spectrometer to also bend the muons' paths by the Lorentz force [4].

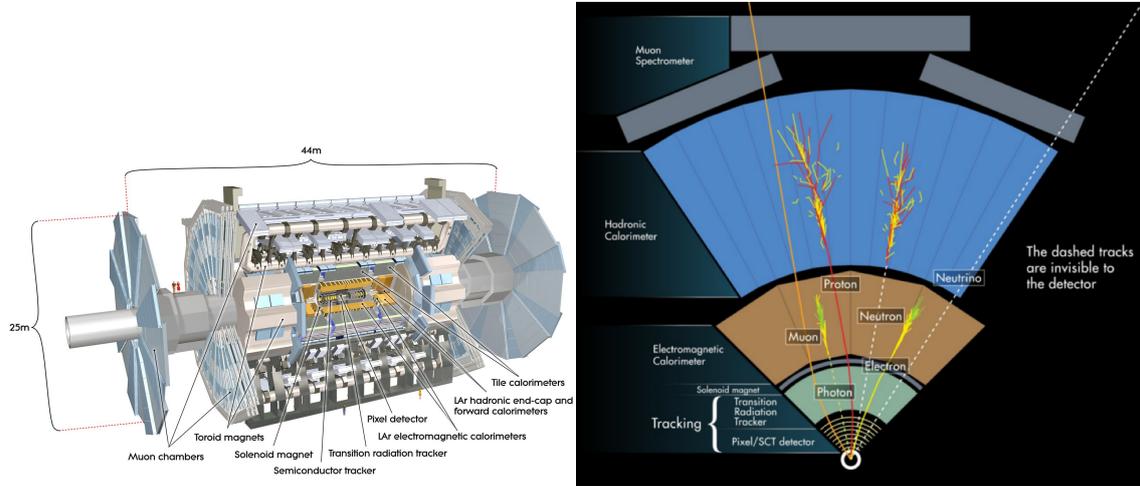


Figure 2: Left: The ATLAS detector [6]; Right: A schematic view of the ATLAS detector subsystems [5].

## 4 ATLAS Coordinate System

Since the ATLAS detector is cylindrically symmetric, a cylindrical coordinate system is used. This is shown by figure 3.

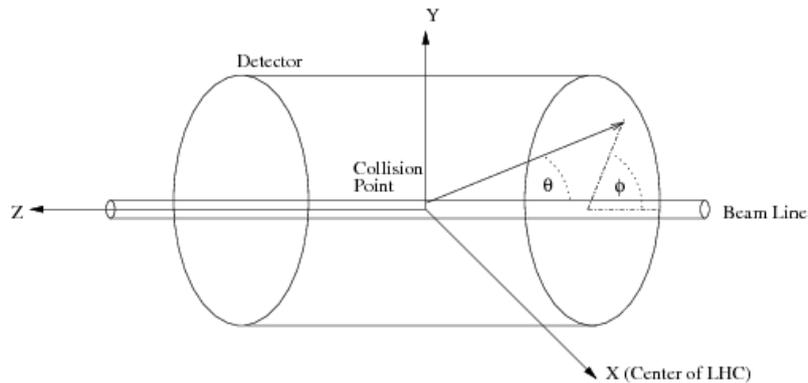


Figure 3: The cylindrical coordinate system [6].

The proton-proton collisions occur inside the beam pipe, at the centre of the detector, which is taken as the origin of the coordinate system. From this, there are three axes:  $x$ ,  $y$  and  $z$ . The proton beams travel along the  $z$ -axis, which is the longitudinal direction, and the  $x$ -axis points towards the centre of the LHC ring. The  $x$ - $y$  plane is perpendicular to the beam axis and is called the transverse plane. In this plane, measurable quantities such

as the transverse momentum,  $P_T$ , are measured. The azimuthal angle,  $\phi$ , surrounds the beam axis whilst the polar angle,  $\theta$ , surrounds the x-axis. By using equation 1, the polar angle can be used to obtain the variable pseudorapidity, which is represented by  $\eta$  and is invariant under the Lorentz boost in the longitudinal direction. For example, the variation in  $\eta$  for the different sections of the ID is shown in figure 4. The pseudorapidity can then be used with the azimuthal angle,  $\phi$ , to calculate the radial distance of the physics objects in the detector,  $\Delta R$ , as shown by equation 2 [7].

$$\eta = -\ln(\tan(\theta/2)) \quad (1)$$

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad (2)$$

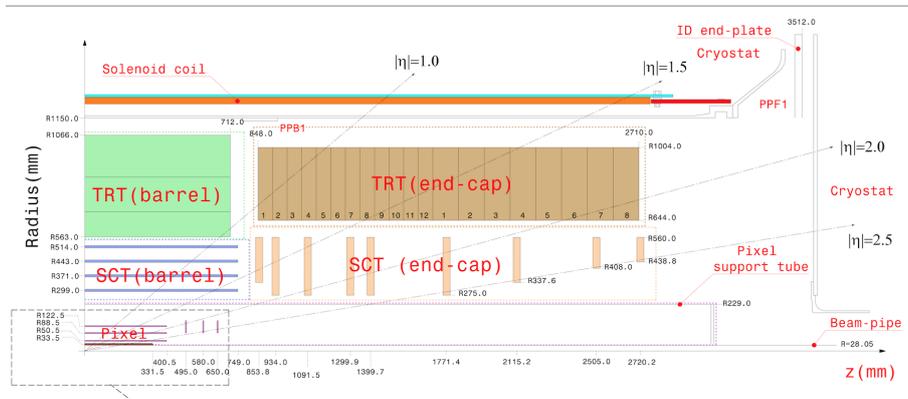


Figure 4: The variation in  $\eta$  values for the different ID sections [8].

## 5 Electron Selection

The  $Z \rightarrow ee$  events that are produced in pp collisions are recorded by the ATLAS detector. This Z boson decays into an electron and a positron and, since they are charged particles, they leave a track in the inner detector. They then reach the electromagnetic calorimeter and each leave an energy deposit, which is known as a cluster. This cluster is then extrapolated backwards to match the ID track to give the electron trajectory. The electrons are reconstructed from the track and cluster information using a set of selection criteria [9, 10].

The identification stage follows this, which enables for the rejection of some of the background processes from the data. These processes include  $J/\psi \rightarrow ee$ ,  $Z \rightarrow ee\gamma$  and other physics processes with the same ee final state. The work for this project is for the identification stage, where the loose, medium and tight identification criteria were used on the kinematic variables for the Z boson and electrons. These criteria have different strictness levels, with tight being the strictest and loose being the least strict [9, 10].

In addition to the identification criteria, trigger selections are applied and the electrons are then isolated from nearby objects to select the nearby objects as the final selection step for the electron candidates [9, 10].

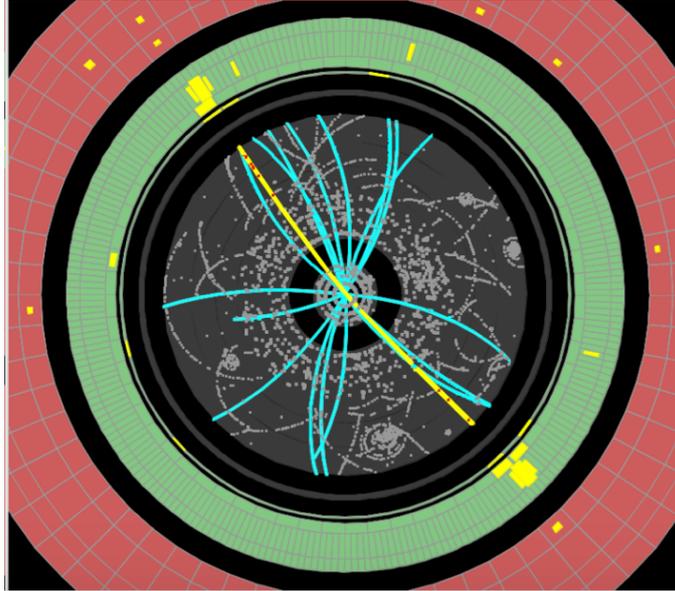


Figure 5: The electron tracks in the inner detector and the electron clusters in the ECAL. The inner detector is represented by the grey region and the ECAL is represented by the green region. The electron tracks are represented by straight, yellow lines in the inner detector and the tracks are positioned back-to-back. The clusters are represented by yellow blobs in the ECAL [11].

## 6 Electron Identification Efficiency and Monte Carlo Scale Factor

The total electron efficiency,  $\epsilon_{total}$ , can be obtained by using equation 3, where the efficiencies for reconstruction, identification, trigger and isolation are represented by  $\epsilon_{reconstruction}$ ,  $\epsilon_{identification}$ ,  $\epsilon_{trigger}$  and  $\epsilon_{isolation}$ , respectively [9].

$$\epsilon_{total} = \epsilon_{reconstruction} \times \epsilon_{identification} \times \epsilon_{trigger} \times \epsilon_{isolation} \quad (3)$$

However, this study only considered the electron identification for the  $Z \rightarrow ee$  process, for which electrons and positrons were paired using the tag and probe method. In this method, strict selections are applied to one of the electrons, called the "tag" electron, in order to reject background processes. After this, the second electron of the electron-positron pair, which is known as the "probe" electron, is used to determine the electron identification efficiencies for the Monte Carlo (MC) simulation and 2015 ATLAS data. The efficiency on the probe electrons for each selection cut can be calculated by using equation 4 [9, 10].

$$\epsilon = \frac{\text{number of probe electrons that pass the selection criteria}}{\text{total number of probe electrons}} \quad (4)$$

The MC simulation simulates the  $Z \rightarrow ee$  process by using a random sampling technique [10]. From this, distributions were obtained for the invariant mass of the Z boson and the  $\eta$ ,  $\phi$  and  $P_T$  values for the Z boson, tag electron and probe electron. These distributions were then compared to the corresponding distributions for the data. From these distributions, the data and MC efficiencies were calculated. Next, the scale factors were calculated by dividing the data identification efficiency by the MC identification efficiency. The determined scale factors can be applied to the MC simulation to match the data. The equation for the MC scale factor for a particular identification criteria is shown by equation 5 [10].

$$\text{Scale Factor} = \frac{\text{Data identification efficiency}}{\text{MC identification efficiency}} \quad (5)$$

## 7 Results

### 7.1 Z Boson: Data and Monte Carlo Comparison Plots

The invariant mass,  $\eta$ ,  $\phi$  and  $P_T$  variables of the Z boson candidate for the MC simulation were compared to the corresponding variables in the data. These comparison plots are shown in figures 6-15. A 70-110 GeV mass window cut was applied to reject the high and low energy tails of the Z boson invariant mass distribution, which might be dominated by background processes. Electron candidates were only selected with  $P_T > 20$  because, below this value, background processes would dominate. Figures 10-12 show the MC and data comparison plots for the tag electrons (electron candidate 1) and figures 13-15 show the MC and data comparison plots for the probe electrons (electron candidate 2). The data and MC distributions are normalised to one with respect to the total number of events in data and MC simulation. The results from figures 10-15 were then used in subsection 7.2 to determine the electron identification efficiency and MC scale factor as a function of  $\eta$  and  $P_T$ .

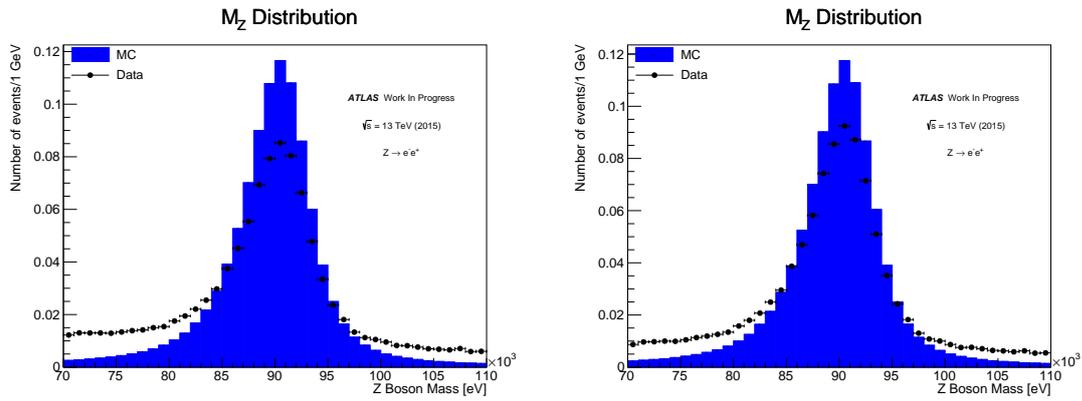


Figure 6: Left: Z boson mass distribution with no identification cuts applied on both electron candidates; Right: Z boson mass distribution with a tight identification cut applied to electron candidate 1 and a medium identification cut applied to electron candidate 2.

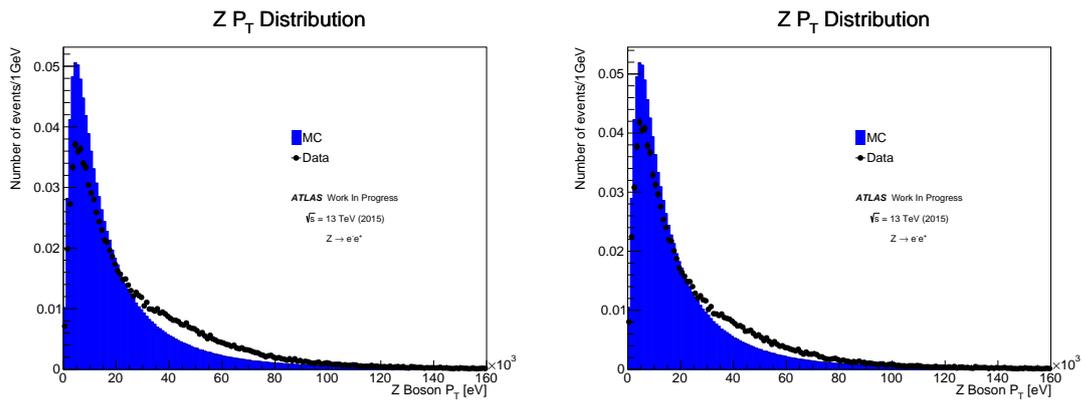


Figure 7: Left: Z boson transverse momentum distribution with no identification cuts applied on both electron candidates; Right: Z boson transverse momentum distribution with a tight identification cut applied to electron candidate 1 and a medium identification cut applied to electron candidate 2.

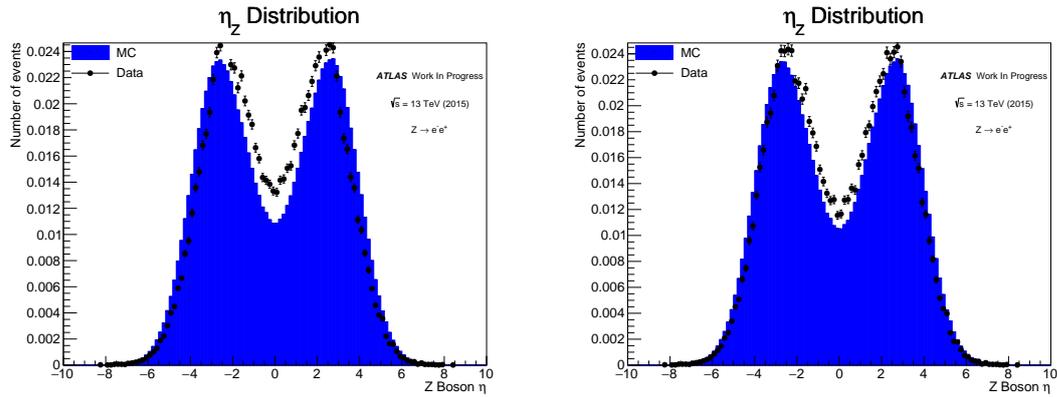


Figure 8: Left:  $Z$  boson  $\eta$  distribution with no identification cuts applied on both electron candidates; Right:  $Z$  boson  $\eta$  distribution with a tight identification cut applied to electron candidate 1 and a medium identification cut applied to electron candidate 2.

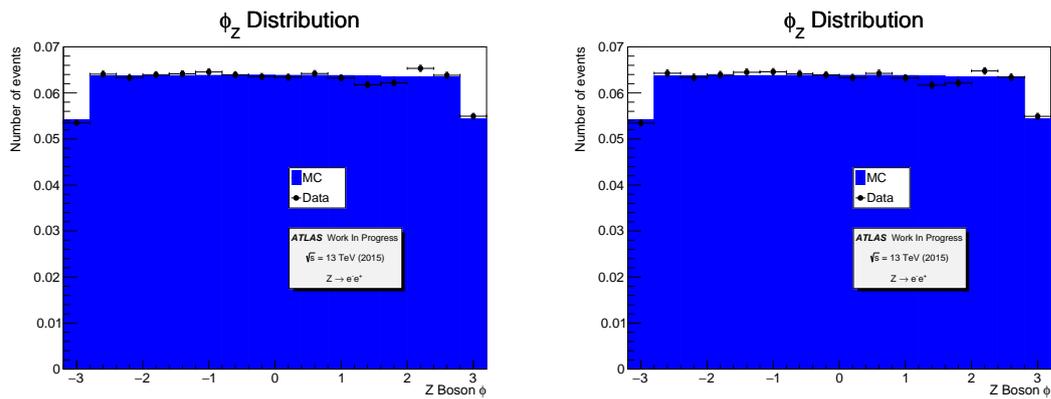


Figure 9: Left:  $Z$  boson  $\phi$  distribution with no identification cuts applied on both electron candidates; Right:  $Z$  boson  $\phi$  distribution with a tight identification cut applied to electron candidate 1 and a medium identification cut applied to electron candidate 2.

## 7.2 Electron Candidates: Data and Monte Carlo Comparison Plots

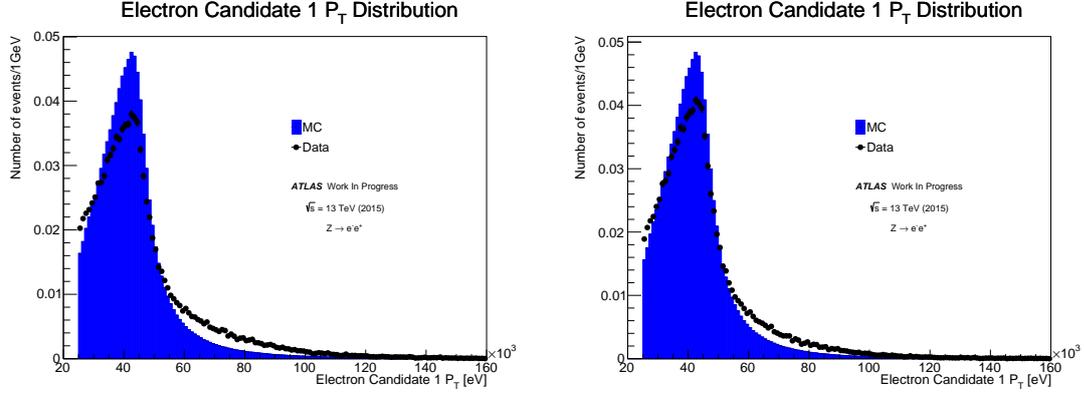


Figure 10: Left: Electron candidate 1 transverse momentum distribution with tight selection cut applied to candidate 1; Right: Electron candidate 1 transverse momentum distribution with a tight selection cut applied to electron candidate 1 and a medium selection cut applied to electron candidate 2.

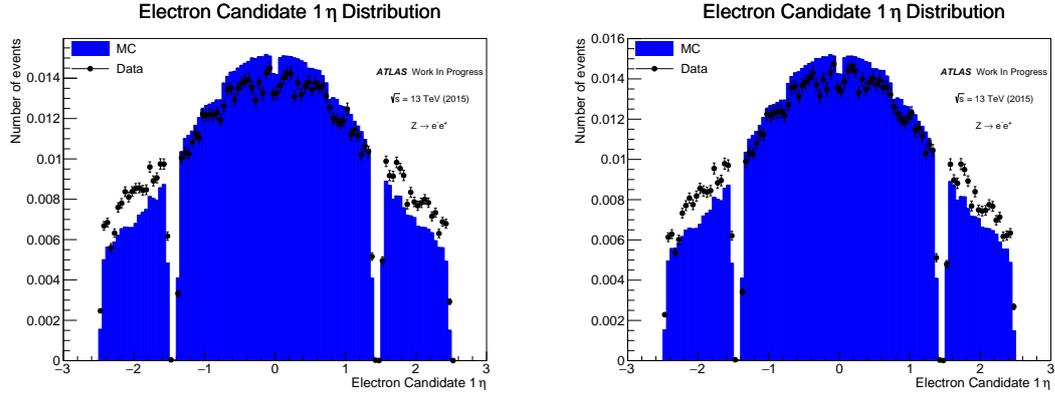


Figure 11: Left: Electron candidate 1  $\eta$  distribution with tight selection cut applied to candidate 1; Right: Electron candidate 1  $\eta$  distribution with a tight selection cut applied to electron candidate 1 and a medium selection cut applied to electron candidate 2.

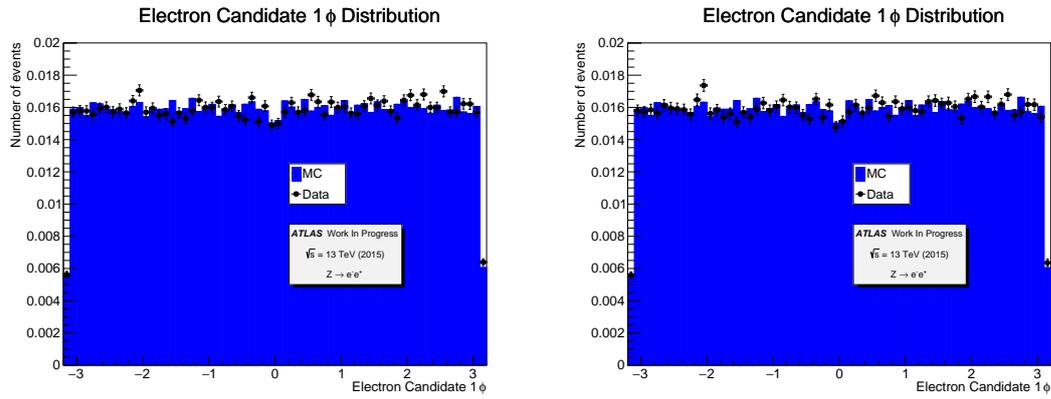


Figure 12: Left: Electron candidate 1  $\phi$  distribution with tight selection cut applied to candidate 1; Right: Electron candidate 1  $\phi$  distribution with a tight selection cut applied to electron candidate 1 and a medium selection cut applied to electron candidate 2.

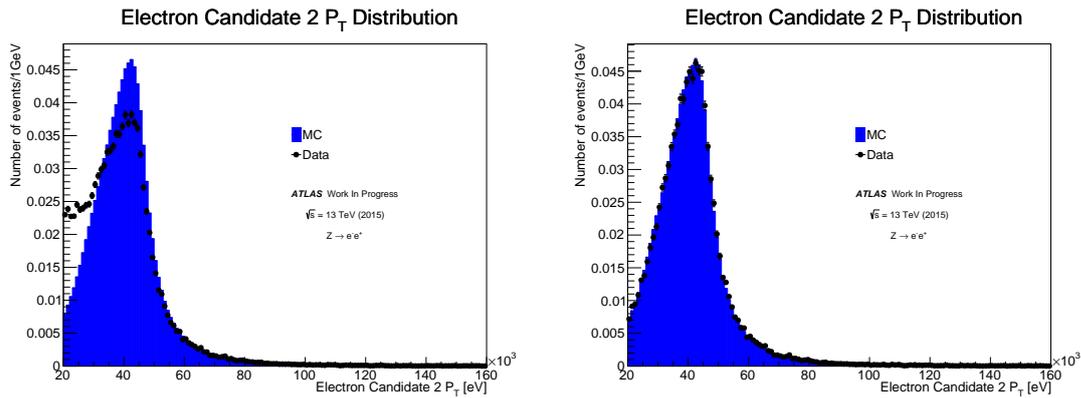


Figure 13: Left: Electron candidate 2 transverse momentum distribution with a tight selection cut applied to electron candidate 1 and a medium selection cut applied to electron candidate 2; Right: Electron candidate 2 transverse momentum distribution with tight selection cut applied to candidate 1.

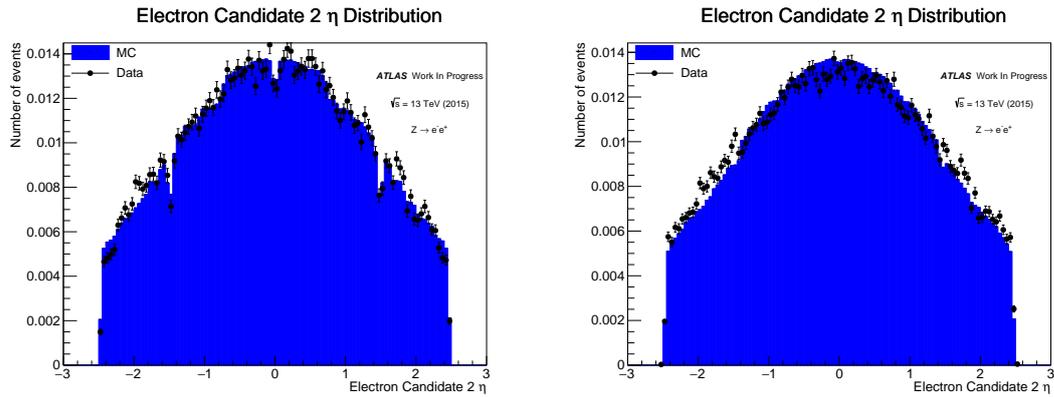


Figure 14: Left: Electron candidate 2  $\eta$  distribution with tight selection cut applied to candidate 1; Right: Electron candidate 2  $\eta$  distribution with a tight selection cut applied to electron candidate 1 and a medium selection cut applied to electron candidate 2.

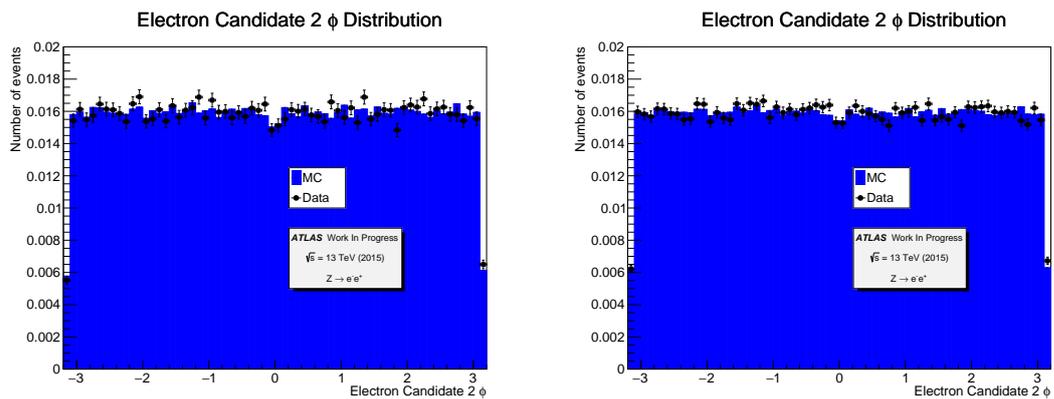


Figure 15: Left: Electron candidate 2  $\phi$  distribution with tight selection cut applied to candidate 1; Right: Electron candidate 2  $\phi$  distribution with a tight selection cut applied to electron candidate 1 and a medium selection cut applied to electron candidate 2.

### 7.3 Efficiency and Scale Factor Plots

To calculate the identification efficiencies on the probe electrons, a base selection of a 70-110 GeV mass window cut, electrons with  $P_T > 20$  GeV and a tight identification cut on tag electron were applied. Finally, the distributions for both data and MC that were either with or without a loose identification cut on the probe electrons were compared to calculate the loose identification efficiency as function of  $\eta$  and  $P_T$ . The loose identification efficiencies are shown by figure 16. This was repeated for the medium and tight identification efficiencies, as shown by figures 17 and 18, respectively. By calculating the ratio between the data identification efficiency and the MC identification efficiency for each identification cut, the scale factors for the loose, medium and tight efficiencies were calculated for  $\eta$  and  $P_T$ , as shown in figure 19.

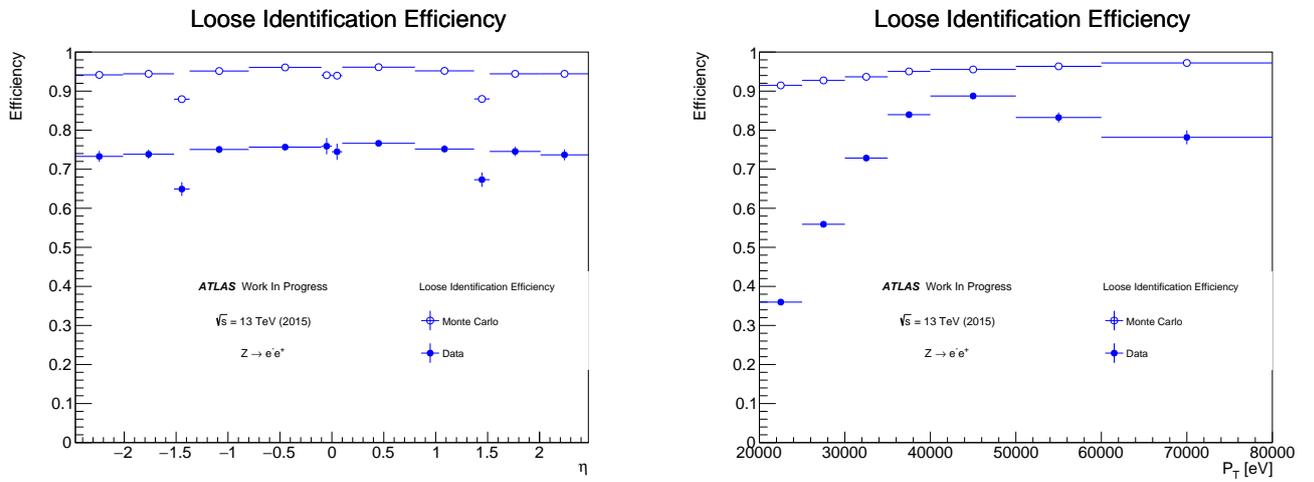


Figure 16: Left: The loose efficiency as a function of the pseudorapidity for data and Monte Carlo; Right: The loose efficiency as a function of  $P_T$  for data and Monte Carlo.

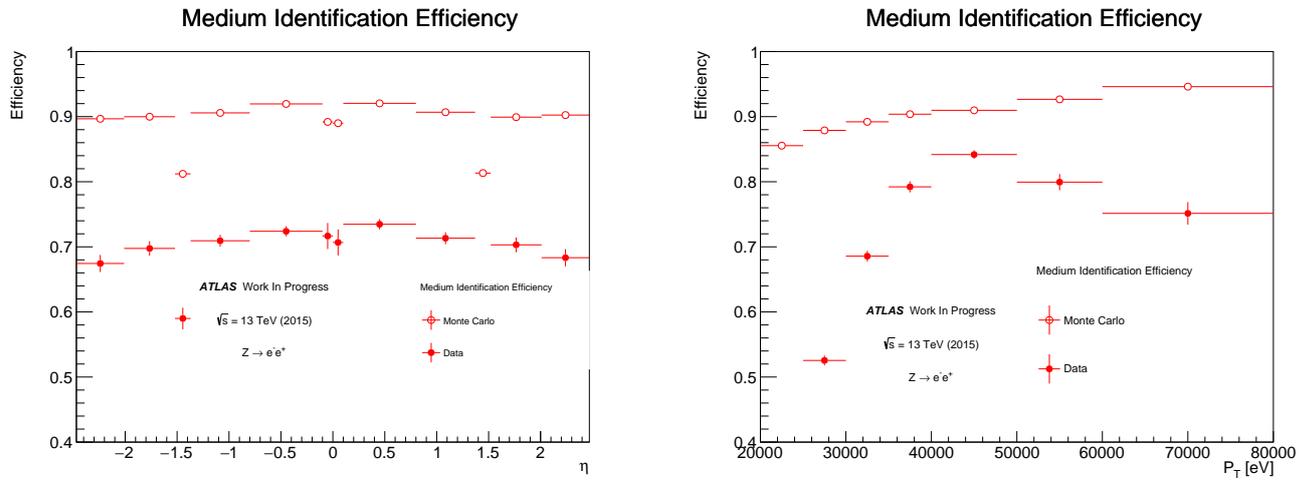


Figure 17: The medium efficiency as a function of the pseudorapidity for data and Monte Carlo; Right: The medium efficiency as a function of  $P_T$  for data and Monte Carlo.

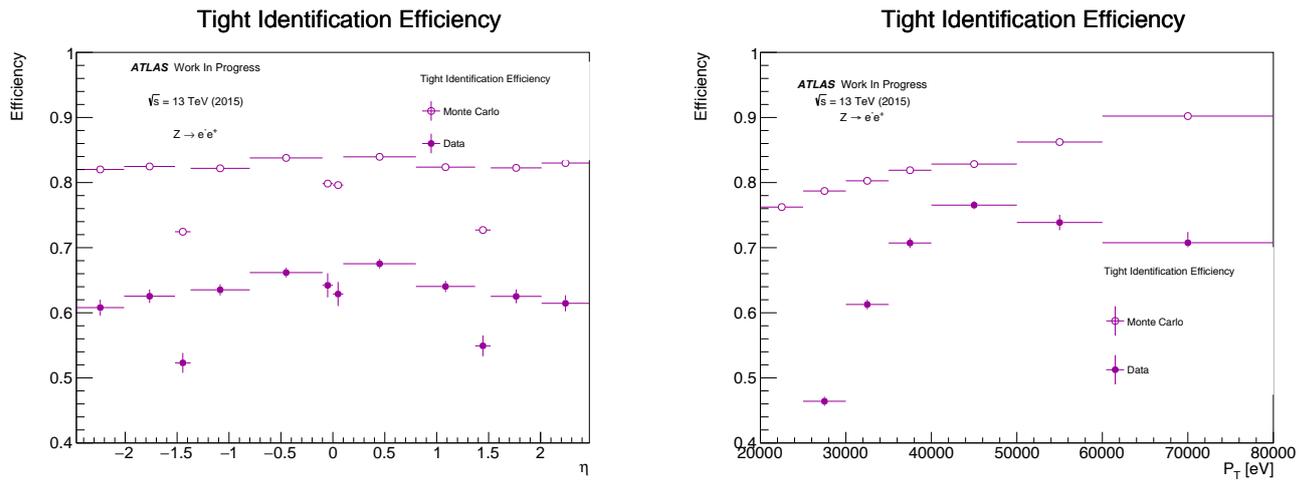


Figure 18: The tight efficiency as a function of the pseudorapidity for data and Monte Carlo; Right: The tight efficiency as a function of  $P_T$  for data and Monte Carlo.

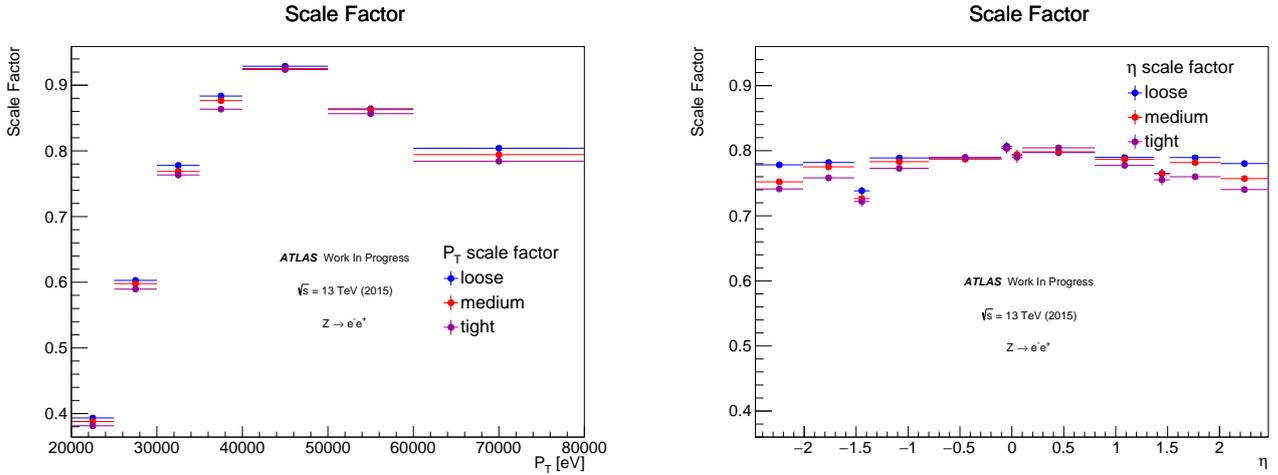


Figure 19: Left:  $P_T$  scale factors for the loose, medium and tight identification cuts; Right:  $\eta$  scale factors for the loose, medium and tight identification cuts

## 8 Conclusions

The MC and data comparison plots for the Z boson are shown in figures 6-9, and the MC and data comparison plots for the electron candidates are shown in figures 10-15. In these, the data distributions are broader compared to the MC distributions as the data is contaminated by different background processes. The identification cuts reduced some of the background so the difference between the MC and data distributions became lower after applying the cuts. The difference between the data and MC distributions was lower for the tight identification cut than the medium identification cut due to higher background reduction.

The efficiency plots are shown in figures 16-18. When tighter cuts were applied, less electrons passed the tight cut and hence the numerator of equation 5 decreased but the denominator, which is the total number of electrons, remained the same. This is why the loose identification efficiency is the highest and the efficiency became lower with a medium ID selection and was lowest with tight ID cuts. The identification efficiencies were also lower in the data than in MC due to the presence of background in the data. The  $P_T$  dependent efficiency has lower values at very low  $P_T$ , then increases and again decreases. For the  $P_T$  efficiency distributions, the highest data efficiency occurred in the 40-50 GeV bin. This agrees with expectations because the invariant mass of the Z boson, which is around 91 GeV, is shared between the electron and positron. At both the low and high energy tails, the backgrounds are from  $J/\psi \rightarrow ee$  and  $Z \rightarrow ee\gamma$ . Top and other physics processes with an electron-positron pair in the final states would contribute. These background processes in turn reduce the identification efficiencies at low and high energies by not being able to pass the identification cuts. The  $\eta$  efficiency distributions were almost uniform but displayed a drop in the region  $1.37 < |\eta| < 1.52$ , which is the transition region between the ECAL barrel and endcap regions. This region contains a large amount of non-detecting material so electrons are not well described. The effect of background processes is more sensitive in  $P_T$  than  $\eta$ .

The scale factor plots are shown in figure 19. The peak in the scale factor for  $P_T$

occurred in the 40-50 GeV bin and the scale factor was the largest for the loose identification cut and the smallest for the tight identification cut. With the exception of the  $1.37 < |\eta| < 1.52$  region, the  $\eta$  scale factor distribution remained almost uniform and had scale factor values between 0.7 and 0.8. The scale factor for  $\eta$  was also the largest for the loose identification cut and the smallest for the tight identification cut. The measured scale factors can now be applied to the MC simulation to mimic the data so the simulation can be used for different physics processes in data.

## Acknowledgements

I would like to thank my project supervisor, Dr Thorsten Kuhl, and PhD students Baishali Dutta and Akanksha Vishwakarma for providing assistance whenever I encountered difficulty. I would especially like to thank Baishali for giving an in-depth explanation for each stage of the project, the physics processes involved and for proof-reading and suggesting edits for my report. Thank you very much to Dr Gernot Maier and Katrin Varschen for the smooth organisation of the DESY Summer School. Finally, I would like to express my gratitude to DESY for offering me this incredible opportunity.

## References

- [1] Bruning, O *et al.*, *The LHC Main Ring*, LHC Design Report, **1**, 263 (2004).
- [2] The LHC Study Group, *The Large Hadron Collider Conceptual Design Report*, European Organisation for Nuclear Research, **1**, 7 (1995).
- [3] Mobs, Esma Anais, *The CERN accelerator complex*, CERN Document Server, (2016). Available: <https://cds.cern.ch/record/2225847?ln=en>. Last accessed 22nd Aug 2017.
- [4] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report 1*, European Organisation for Nuclear Research, **1**, (1999).
- [5] Unknown author, *Detecting Particles*, University of Oxford, **1**, (2016). Available: <http://collider.physics.ox.ac.uk/detecting.html>. Last accessed 30th Aug 2017.
- [6] Schott, Matthias *et al.*, *Review of single vector boson production in pp collisions at  $\sqrt{s} = 7$  TeV*, inspirehep.net, (2014). Available: <https://inspirehep.net/record/1294662/plots>. Last accessed 1st Sept 2017.
- [7] John Allison, *The Road to Discovery: Detector Alignment, Electron Identification, Particle Misidentification, WW Physics, and the Discovery of the Higgs Boson*, Pennsylvania U., **1**, (2012).
- [8] ATLAS Collaboration, *A study of the material in the ATLAS inner detector using secondary hadronic interactions*, arXiv, **1**, (2012).
- [9] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, European Organisation for Nuclear Research, **1**, (2012).
- [10] In conversation with Baishali Dutta

- 
- [11] ATLAS Collaboration, *Z* → *ee* candidate in 7 TeV collisions. twiki.cern.ch. (2010).  
Available: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic>.