Masters Report
Electron Identification with the ATLAS Detector

Jaspreet Sidhu
Supervisor: Professor Richard Teuscher

August 29, 2007

Introduction

The ATLAS experiment is a general-purpose particle detector being constructed at the Large Hadron Collider (LHC) at CERN. The collider is contained in a 27km circumference tunnel located underground at a depth ranging from 50 to 150m. The LHC will be the highest energy particle collider ever built with proton beams colliding at a center-of-mass energy of 14 TeV and a design luminosity of 10^{34} cm^{-2}s^{-1}. The experiment will be seven times the energy scale and 100 times the luminosity of the Tevatron at Fermilab, which is currently the highest energy particle collider in the world. The ATLAS detector, one of five being built in the LHC tunnel, is to be 46 meters long and 25 meters in diameter, and will weigh about 7,000 tones, which will make it the largest collider detector ever built.

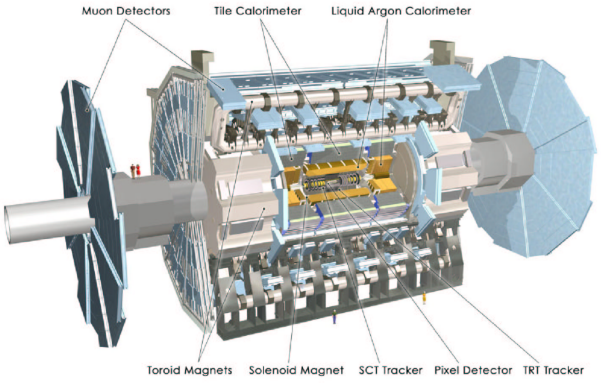

Figure 1: The ATLAS experiment.

The ATLAS experiment is a multi-purpose detector. Its inner detector is composed of pixel detectors, silicon strip detectors and transition radiation drift tubes covering a pseudo-rapidity of $|\eta| < 2.5$. This is surrounded by a lead/liquid argon electromagnetic calorimeter, a steel/scintillating tile hadronic calorimeter, and then finally a large air-core muon spectrometer

The ATLAS detector intends to investigate physics in the TeV energy scale for the first time. One of the main physics goals for the ATLAS experiment is to understand electroweak symmetry breaking and hence search for the Higgs boson. ATLAS also hopes to investigate physics beyond the Standard model such as Super Symmetry and the search for extra dimensions. The ATLAS experiment will no doubt be a very influential experiment in changing our understanding of high-energy physics in energy regimes never before probed.

Analysis

The aim of this study is to understand and analyze the methods of electron identification with ATLAS. This is achieved by using a $Z \rightarrow ee$ signal sample. The goal is to understand basic quantities and distributions and to determine reconstruction and identification efficiencies. A quantitative description of QCD background is determined, specifically electron fake rates, as a means to minimize mis-identified jets. Trigger performance is also determined.

Physics Motivation

Events with electrons in the final state are important signatures for many physics analyses in ATLAS. Isolated high p_T electrons will not be an easy task to identify at the LHC due to the large QCD backgrounds that are present due to high p_T jets. A rejection factor of 10^5 against jets will be required to efficiently identify electrons. Thus good electron identification is required over a broad energy range.

The process $Z \rightarrow ee$ will be an important sample for calibration of the electromagnetic calorimeter. The ATLAS detector will see first beam from the LHC in the middle of 2008. The data collected in the first few months will be used for calibrating the different sub detectors that are part of ATLAS. This will be crucial because any measurements and discoveries will depend on the calibration being precise. This channel is essentially background-free and can be used as a stand-alone calibration. With a production rate of 1 kHz at a luminosity of 10^{33} cm^{-2}s^{-1} (10% nominal) a sufficient intercalibration can be done within a few days of data taking.

Datasets

In this study several dataset samples are used. Official CSC Monte Carlo data samples were obtained using the ATLAS grid computing system. Specifically for the signal sample $Z \rightarrow ee$, the sample $trig1\_misal\_csc11.005144.PythiaZee.recon.AOD.v12000603$ was used. This sample contains 7250 $Z \rightarrow ee$ events reconstructed in ATHENA version 12.0.6 with trigger information included. Events are generated with csc11 data made with geometry ATLAS-DC3-02 (generation is independent of the geometry). For detector simulation and digits, data is processed with geometry ATLAS-CSC-01-02-00, which includes the most recent measurements of the magnetic field, a misaligned geometry to simulate an imperfect detector calibration, and material distortions. More information about this can be found on the Inner Detector Software Twiki webpage[1].

[1] https://twiki.cern.ch/twiki/bin/view/Atlas/InnerDetectorSoftware

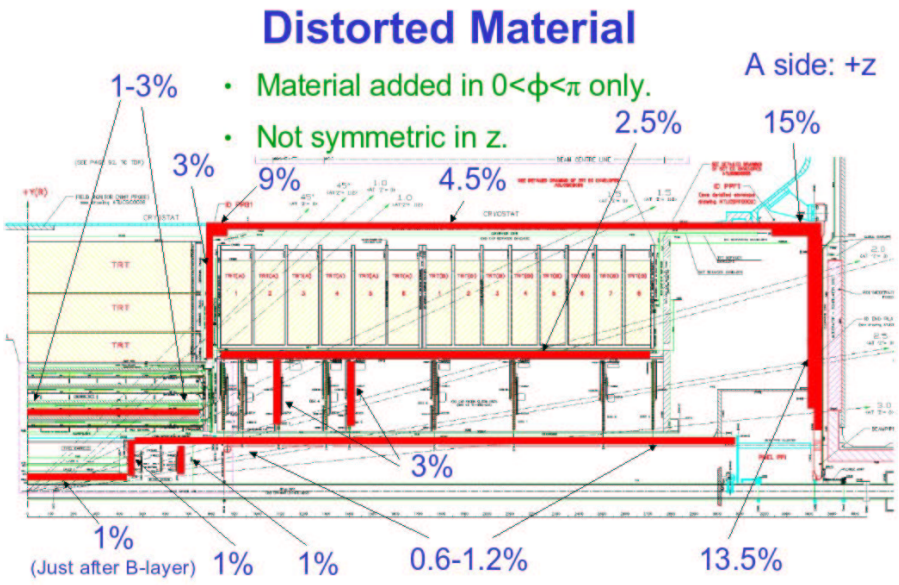

Figure 2: Illustration of the geometry ATLAS-CSC-01-02-00 which includes new magnetic field, misaligned geometry with material distortions.

To determine QCD jet fake rates, official CSC jet samples produced in ATHENA version 12.0.6 are used. These jet samples represent jets in different p_T ranges. These are

- J0 Dijet pt range 8-17 GeV
- J1 Dijet pt range 17-35 GeV
- J2 Dijet pt range 35-70 GeV
- J3 Dijet pt range 70-140 GeV
- J4 Dijet pt range 140-280 GeV
- J5 Dijet pt range 280-560 GeV
- J6 Dijet pt range 560-1120 GeV

Specifically the samples J0, J2, J3, J4 and J6 were obtained from the grid and used in the analyses. At the time of this report J1 and J5 samples were unavailable. The number of events obtained are :

- J0: 83,250 events
- J2: 19,00 events
- J3: 10,00 events
- J4: 33,250 events
- J6: 2,750 events

Electron reconstruction

Electron analysis is performed on the electron "container", which is a collection of electron objects in ATHENA. The reconstruction of an electron candidate placed in the electron container in ATLAS is done by the egammaRec algorithm, which is created at the time of AOD [2] building. The egammaRec algorithm uses information coming from the Electromagnetic Calorimeter and Inner Detector. The egammaRec algorithm ensures that there be a matched track to a calorimeter cluster determined by the CaloClusterMaker and that the ratio of the hadronic energy of the total energy be less than 20%. Particles in the electron container at this stage contain loose candidates and are not necessarily electrons.

Electron Identification

Several electron identification methods are used in ATLAS. These include Cut, Neural Net and Likelihood methods. A cut based method is studied to determine good electron objects. The description of the cuts used is now given.

- **Cut 1** Good track and $p_T > 15GeV$
- **Cut 2** Egamma Object
- **Cut 3** isEM

Cut 1 The first step in the process is the elimination of electrons within a lower p_T bound. This restriction is set at 15GeV and is implemented because we are not concerned with "soft" (low-energy) electrons. As well as this, we ask for the existence of a good quality track. A cluster in the calorimeter is matched to a track in a two-step process. First the η [3] and the ϕ [4] of the

[2] AOD is Analysis Object Data, the standard data file for analysis.

[3] $\eta = -\ln tan\frac{\theta}{2}$, where θ is the polar angle from the beam axis.

[4] ϕ : azimuthal angle around the beam axis, where $tan\phi = \frac{p_x}{p_y}$.

track (at its origin) are compared to that of the cluster. If there is agreement to within 0.05 in η, and 0.1 in ϕ, step two proceeds. Here, the track is extrapolated towards the cluster. $\Delta\eta$ and $\Delta\phi$ are calculated at each layer, to ensure closeness with the track.

Cut 2 Egamma object

We require that the electron object be an "egamma object". The electron candidates in the container are grouped in two collection, "softe" and "egamma". The first is track-seeded, while the second is cluster-seeded. Since the objects are grouped into two collections set by the method author, we ask that the author of the object be egamma. This ensures the objects are not low energy electrons. Softe objects are used to identify soft electrons of p_T around $5 - 10$ GeV with no isolation. These objects are important for B-physics.

Cut 3 isEM()

The isEM method is composed of a bit field of several flags, each of which corresponds to a specific algorithm, or set of algorithms. These algorithms use a combination of electron shower shape properties, calorimeter information and inner detector information including the Transition Radiation Tracker (TRT) to identify good electrons. The algorithms consist of series of cuts; if a cut is not passed, then a bit is set in the isEM flag.

Electron identification based on the calorimeter relies on the layout of energy deposits. It is expected that electrons will shower predominantly in the electromagnetic calorimeter, with small, if any, deposits in the hadronic calorimeter; typically the energy deposits within the hadronic calorimeter are less that 2% for an electron. Because of this, the measure of 'hadronic leakage' proves useful when discriminating good electrons. Hadronic leakage considers the ratio of the transverse energy reconstructed in the hadronic calorimeter to the transverse energy reconstructed in the electromagnetic calorimeter, $\frac{Et(hadronic)}{Et(electromagnetic)}$. isEM places an upper bound of 5% on this ratio.

isEM also considers identification based on the electromagnetic calorimeter alone, through the shower shape formed by energy deposits. It is expected that electromagnetic showers deposit most of their energy in the second sampling of the electromagnetic calorimeter. The shape constructed typically occurs within a certain range of unit cells, this range is: $\Delta\eta = 3$ and $\Delta\phi = 7$. This means the electromagnetic shower is contained within these 3×7 unit cells. Typically, jets do not scatter in such a confined manner. The process of isEM considers the extension of the electron shower space. Rather than considering only the 3×7 cells, $\Delta\eta$ is extended to 7, so a 7×7 unit region is evaluated. For electrons, we expect the ratio of energy within the smaller space to the larger space to be close to one. That is, the lateral, (along η), leakage for electrons is small. This process helps eliminate fake electrons from jets.

So far, we have discussed aspects of isEM that deal with jets of high energy (through hadronic leakage), or wide shower distributions (by electromagnetic shower shape). The remaining fake

electrons can be examined by making use of the fine granularity in rapidity within the first sample of the electromagnetic calorimeter. Here, discriminations can be made based on the species composing a jet. For example, jets with π^0's have two peaks in (η, ϕ) in their decay to two photons; the fine granularity within the first electromagnetic sample helps identify and discard such fake electron candidates.

All the aspects above correspond to the workings of isEM using calorimeter information, it also makes use of the inner detector. isEM further examines the tracks which are matched to an electromagnetic cluster. This closer examination ensures that there are at least nine precision hits along an inner detector track line. As well as this, the energy measured in the electromagnetic calorimeter is compared to the momentum measured in the inner detector. For an electron, the momentum is expected to match the energy. All the 'bits' used in isEM are given below.

The bits encoded in isEM based on the egamma Cluster are :

- ClusterEtaRange
- ClusterHadronicLeakage
- ClusterMiddleSampling
- ClusterFirstSampling

The bits encoded based on the egamma Track are

- TrackEtaRange
- TrackHitsA0
- TrackMatchAndEoP
- TrackTRT

Thus if isEM == 0, i.e. all cuts are passed then a good electron candidate is found. One can use a bit mask to use only certain parts of the ID criteria, i.e. (isEM $\&$ bitmask)==0. Depending on the bit mask used one can vary the stringency of the isEM cut, and hence not require that all cuts be passed. As such three levels of isEM cuts have been defined, isEM loose, isEM medium and isEM tight. Depending on the analysis and requirements one may be more useful than others. In version 13 of ATHENA release these three isEM methods have been set without the need to specify the bit mask. However the version used in this study is 12.0.6 and the appropriate bits were used to match what will be defined in version 13.

The definition of the three isEM methods are:

- **Cut3a: isEM loose** (isEM $\& \ 0 \times 7$)==0 , Identification based only on information from the calorimeter
- **Cut3b: isEM medium** (isEM $\& \ 0 \times 3FF$)==0, Identification based on all cuts except TRT
- **Cut3c: isEM tight** (isEM)==0, Identification based on all cuts

The specific isEM cut used will depend on one's analysis.

Below are figures illustrating the process of electron identification for $Z \rightarrow ee$ events. Distributions of the p_T and η of the reconstructed electrons after each cut are shown. Reconstructed electrons are also matched to truth [5] electrons in the truth container within a $\Delta R < 0.1$. ΔR is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. The matched electrons are also shown in red.

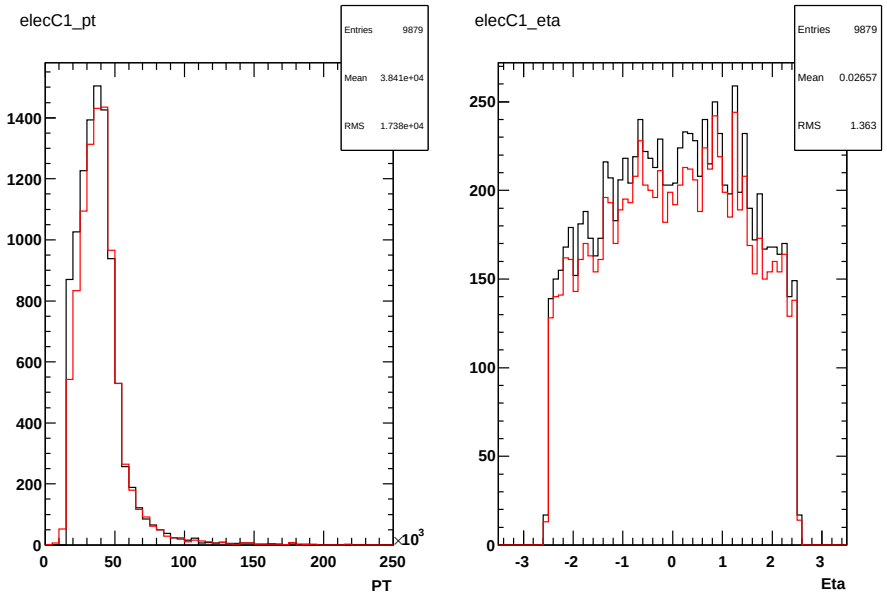

Figure 3: Distribution of p_T and η after Cut 1: Good track and $p_T > 15GeV$; Reconstructed electrons in black and Matched electrons to Truth in red.

Efficiency of Reconstruction

The efficiency of reconstruction is defined as

$$\epsilon_{Rec} = \frac{Number\ Reconstructed\ and\ Matched}{Number\ of\ Truth\ Electrons}$$

[5] Here 'truth' refers to Monte Carlo level quantities.

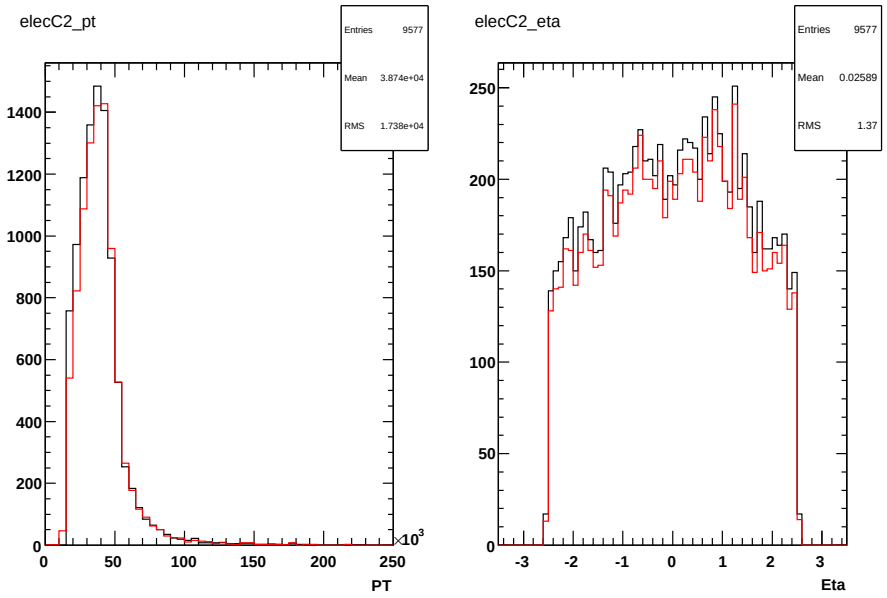

Figure 4: Distribution of p_T and η after Cut 2: Egamma Object; Reconstructed electrons in black and Matched electrons to Truth in red.

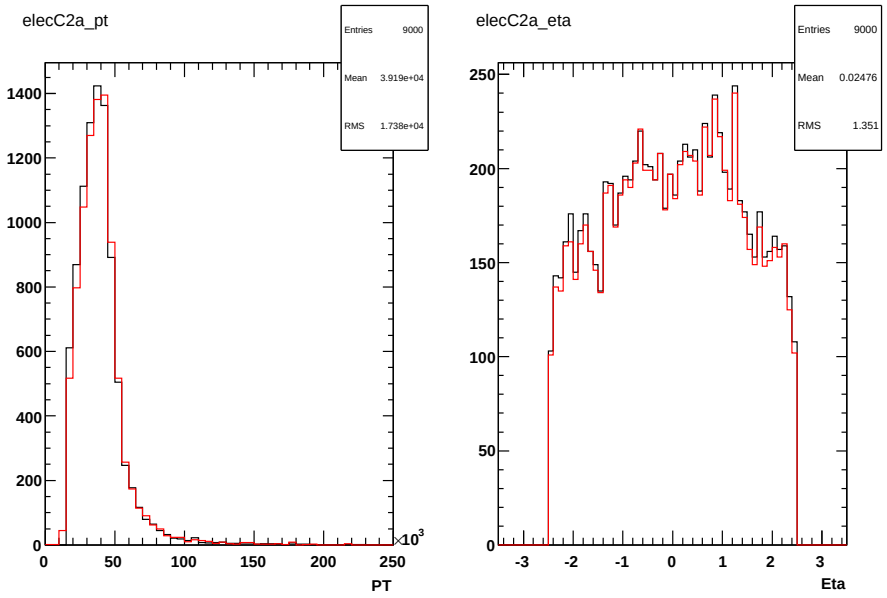

Figure 5: Distribution of p_T and η after Cut 3a: isEM loose; Reconstructed electrons in black and Matched electrons to Truth in red.

In the signal sample used there are 7250 $Z \to e^+e^-$ events yielding 14500 truth electrons.

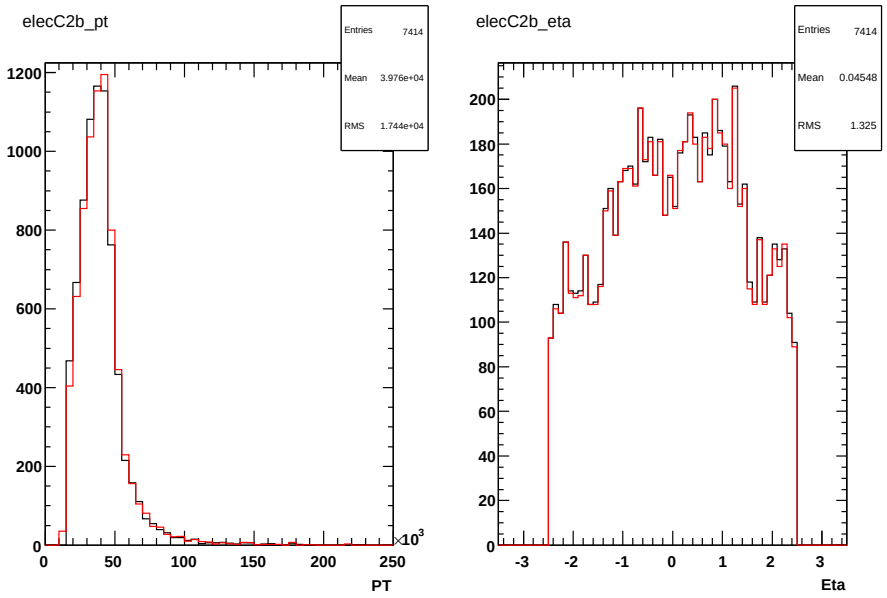

Figure 6: Distribution of p_T and η after Cut 3b: isEM medium; Reconstructed electrons in black and Matched electrons to Truth in red.

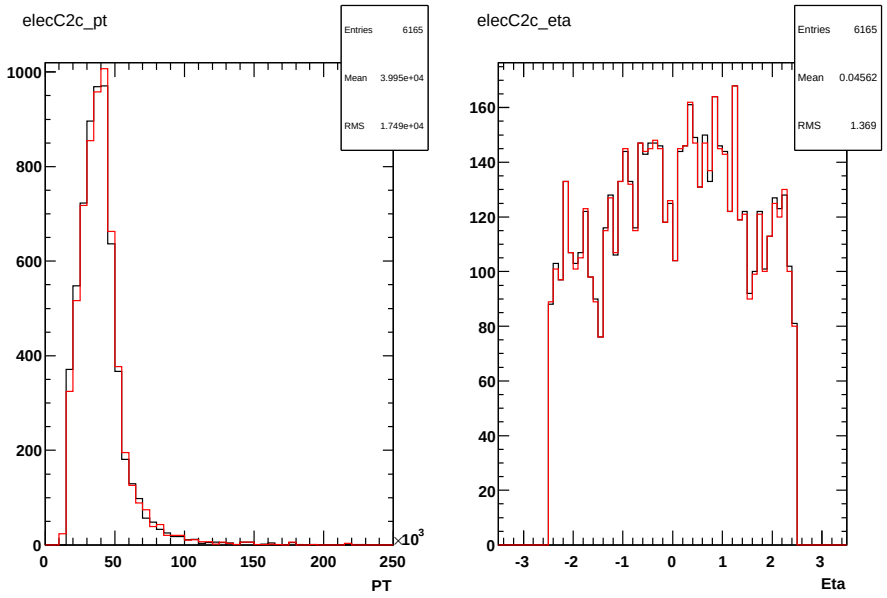

Figure 7: Distribution of p_T and η after Cut 3c: isEM tight; Reconstructed electrons in black and Matched electrons to Truth in red.

The Purity of reconstruction is defined as

$$Purity = \frac{Number\ of\ Reconstructed\ Electrons\ Matched}{Number\ of Reconstructed\ Electron\ Candidates}$$

	Number of Electrons	Number Matched	ϵ_{Rec} (%)	Purity (%)
No Cuts	26132			
Cut 1	9879	9158	68.1	92.7
Cut 2	9577	9084	66.0	94.9
Cut 3a	8331	8217	57.4	98.6
Cut 3b	7054	7029	48.6	99.6
Cut 3c	6165	6145	42.5	99.7

Table 1: Efficiency of Reconstruction and Purity of Electrons After Cuts where 14500 truth electrons are generated.

As we can see from Table 1, the efficiency of reconstruction decreases as more stringent cuts are applied, while the purity of the sample increases.

The figure below displays the reconstruction efficiency as a function of p_T and η for the three isEM methods. The efficiency decreases within the crack region of $1.3 < |\eta| < 1.5$ and generally increases as a function of p_T.

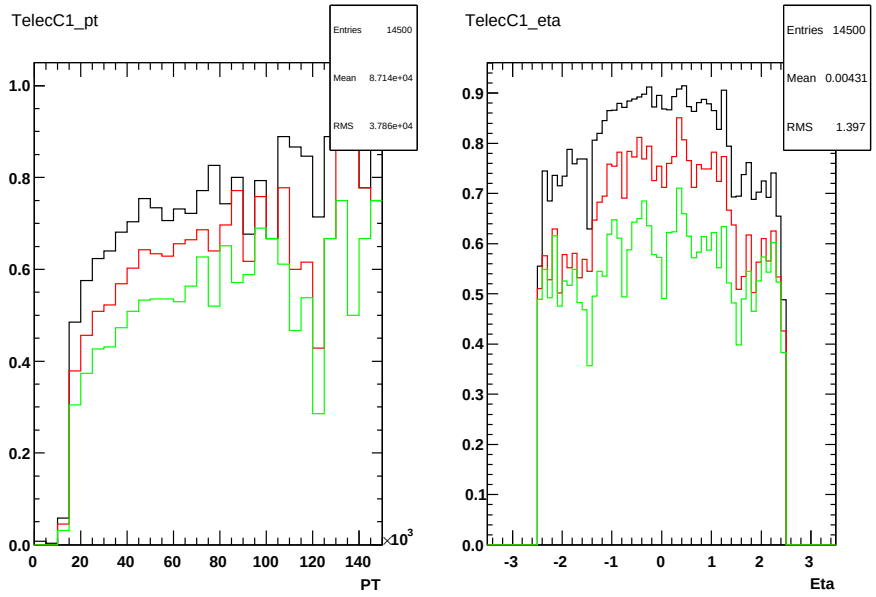

Figure 8: Reconstruction Efficiency as a a function of p_T and η for the three isEM methods; isEM loose (Black), isEM medium (Red), isEM tight (Green).

It is also interesting to compare the resolution of energy for the three isEM methods. The energy resolution is defined as

$$Resolution = \frac{Reconstructed\ Transverse\ Energy}{Truth\ Transverse\ Energy}$$

Below are the resolution distributions for isEM loose, medium and tight.

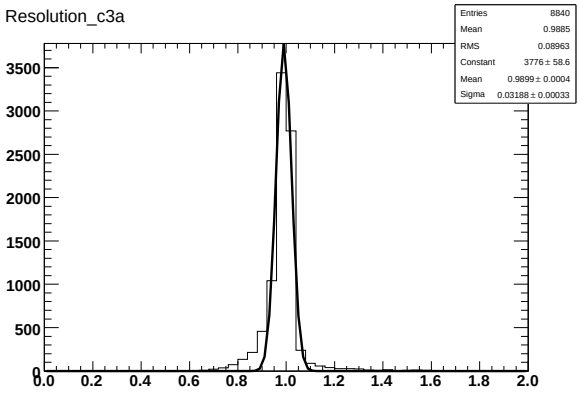

Figure 9: Energy Resolution for isEM loose.

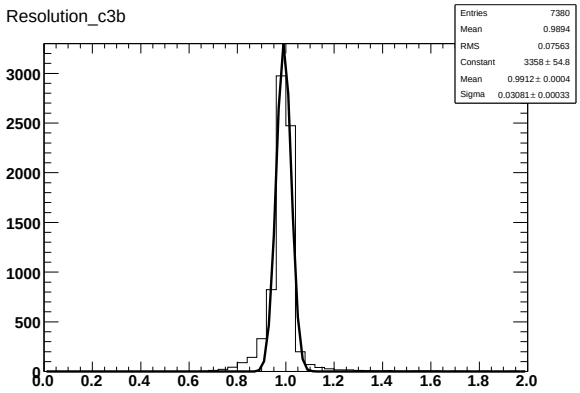

Figure 10: Energy Resolution for isEM medium.

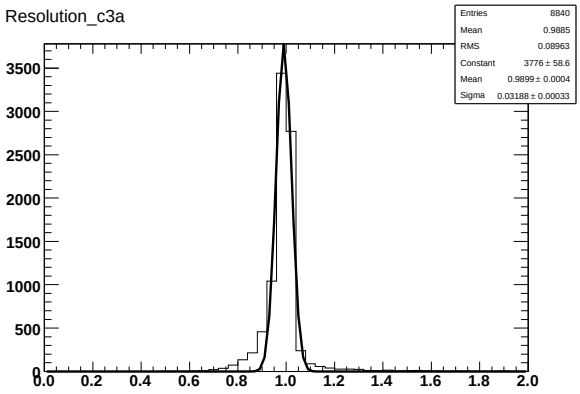

Figure 11: Energy Resolution for isEM tight.

As we can see from Table 2 the resolution of the energy is smaller as isEM is more stringent.

	Mean (GeV)	σ (GeV)
Cut 3a isEM loose	0.989	0.379
Cut 3b isEM medium	0.990	0.03025
Cut 3c isEM tight	0.991	0.03016

Table 2: Resolution of Energy of the three isEM methods.

Electron 1 isEM cut	Electron 2 isEM cut	Mean (GeV)	σ (GeV)
None	None	88.53	4.78
Loose	Loose	88.90	4.26
Loose	Medium	89.08	4.22
Loose	Tight	83.13	4.18
Medium	Medium	89.19	4.18
Medium	Tight	89.23	4.09
Tight	Tight	89.25	4.08

Table 3: Resolution of Z mass using the various isEM methods applied to the reconstructed electrons.

Signal Reconstruction

Having identified good electron objects, the invariant mass of the system is then calculated to reconstruct the Z boson. The invariant mass of the system will depend on which isEM method is used to reconstruct the electrons. Table 3 details this dependency.

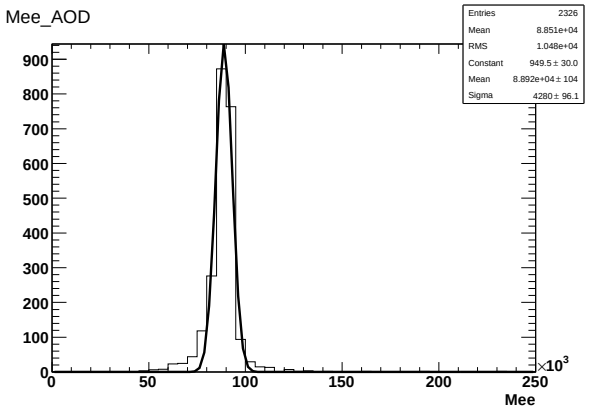

Figure 12: Reconstructed Z boson mass from both electrons passing isEM tight cuts.

From Table 3 we can see that as the electron identification method becomes stricter the resolution of the Z mass is tighter and approaches the accepted measured value of 91.19 GeV as obtained from the Particle Data Group (PDG).

Trigger Statistics

The ATLAS trigger system is structured in three levels for online event selection. These are identified as LVL1, LVL2 and LVL3 or EF (Event Filter). Events are selected according to a Trigger menu. Simply stated this is just a selection criterion. At each level of the trigger system, decisions are made based upon selection criteria that are improved upon and refined through subsequent levels.

For each event, a TriggerDecision object is produced by an Algorithm derived class, TriggerDecisionMaker and is stored at AOD [6] is building. The TriggerDecision provides the decision of the trigger on each event. It lists which trigger signatures were running at each trigger level, and which were satisfied by the event. It also returns the overall result, accept or fail, from each trigger level and for the whole trigger. It is currently set up for a luminosity of $L = 10^{33} cm^{-2} s^{-1}$.

Specifically we are interested in Electron Object triggers, to understand which events are identified as having good electrons that will pass the trigger criteria. Electron candidates are selected at Level-1 using the calorimeter information. LVL2 reconstruction uses information on the transverse energy and the direction of the electromagnetic cluster selected by the LVL1 trigger. The LVL2 calorimeter trigger refines the LVL1 information using full-granularity information from the calorimeters. If LVL2 cuts are passed, then the event is passed to the Event Filter (EF). The EF selection cuts from the Trigger menu that we are interested are e25i (an isolated electron of 25 GeV) and 2e15i (2 isolated electrons of 15 GeV) which are the main single and double object electron physics triggers respectively.

For e25i the default cuts for η in bins [0..0.75, 0.75..1.5,1.5..1.8,1.8..2.0,2.0..2.5] are

- E_T (cluster) threshold = 20 GeV
- E_T of leakage into hadronic calorimeter $<$ [3.8 GeV, 3.8 GeV, 3.8 GeV, 3.8 GeV, 3.8 GeV], no leakage cut applied if $E_T(cluster) > 90GeV$
- Rcore [7] threshold $>$ [0.895,0.895,0.895,0.895,0.895]
- Eratio [8] $>$ [0.730,0.730,0.730,0.730,0.730]

For 2e15i the default cuts for η in bins [0..0.75, 0.75..1.5,1.5..1.8,1.8..2.0,2.0..2.5] are

- E_T (cluster) threshold = 12 GeV

[6] AOD is Analysis Object Data, the standard data file for analysis.

[7] Rcore = $\frac{E37}{E77}$, which is the ratio of energy contained in a $\Delta\eta \times \Delta\phi = 3 \times 7$ window to that in a 7×7 window in the second sampling of the EM Calorimeter.

[8] $Eratio = \frac{E1-E2}{E1+E2}$, where E1 and E2 are the first and second maximum strips of energy in the first sampling of the EM Calorimeter.

	No Trigger	e25i (ϵ %)	2e15i (ϵ %)
isEM Loose	6057	4915 (81.1)	2177 (34.9)
isEM Medium	5369	4545 (84.7)	2071 (38.6)
isEM Tight	4773	4134 (86.6)	1964 (41.1)

Table 4: Number of Events passing the EF electron trigger selections

- E_T of leakage into hadronic calorimeter $<$ [1.0 GeV, 1.0 GeV, 4.0 GeV, 1.5 GeV, 1.0 GeV,], no leakage cut applied if $E_T(cluster) > 90GeV$
- Rcore threshold $>$ [0.90, 0.89, 0.89, 0.90, 0.89]
- Eratio $>$ [0.60, 0.70, 0.75, 0.85, 0.90]

Table 4 lists the number of events passing the stated EF trigger for the three isEM methods as well as its efficiency. The efficiency of the trigger selection is

$$\epsilon(\%) = \frac{Number\ of\ Events\ after\ Trigger\ Selection}{Number\ of\ events\ with\ no\ Trigger}$$

The number of events with no Trigger are events that pass the electron identification scheme without considering trigger information. Thus the efficiency of the trigger explains which events that pass our identification method will also pass the electron physics triggers.

The trigger efficiency increases as the isEM method is more stringent. However, the e25i electron trigger is far more efficient at selecting events of interest for this study than the 2e15i electron trigger.

QCD data

The electron reconstruction process is now applied to our QCD jet simulated data. The samples used are described in the section Datasets and are the J0, J2, J3, J4 and J6 samples. The distributions of p_T and η of identified electrons of the jet samples using the isEM tight method are shown in Fig.13-17. From Fig. 13 we can see that only 1 candidate from 83250 events passed the tight selection. A summary of the electron identification scheme applied to the jet samples are given in Table 5.

Every electron reconstructed in Table 5 is a fake electron since there are no truth electrons in our jet samples. J4 and J6 samples have the most significant portion of reconstructed electrons, contributing to a sizable fake rate. One concludes that the probability of a jet faking an electron increases with p_T.

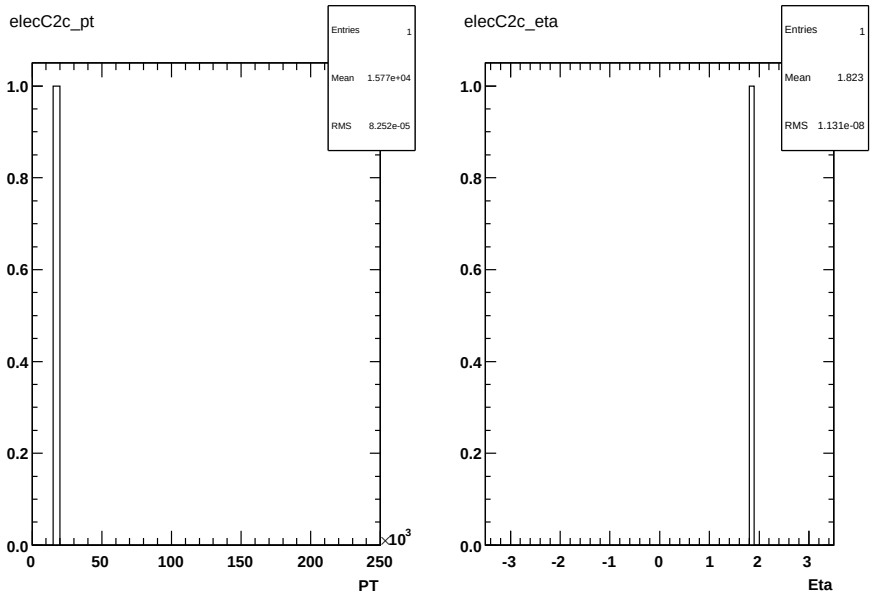

Figure 13: Distribution of p_T and η after Cut3c: isEM tight for the J0 jet sample

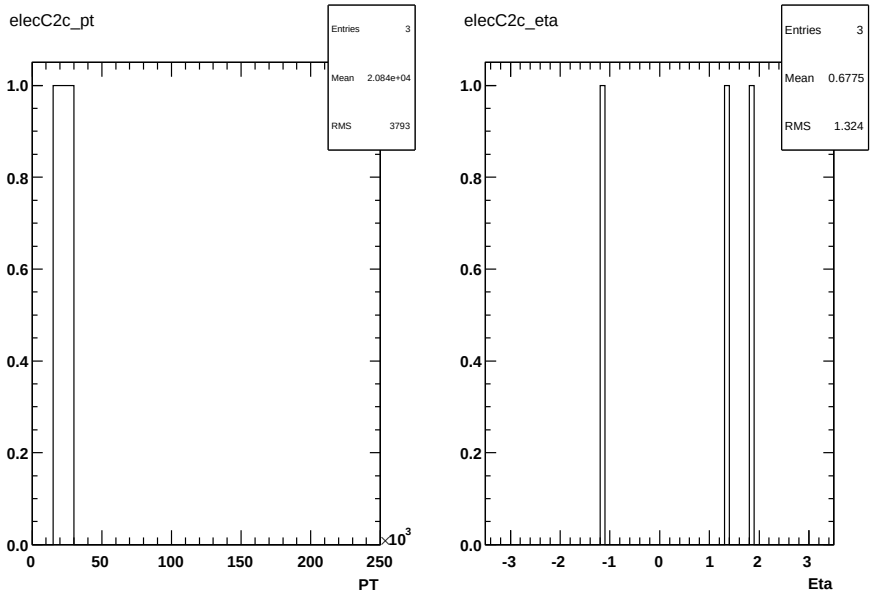

Figure 14: Distribution of p_T and η after Cut3c: isEM tight for the J2 jet sample

Considering trigger information, Table 6 summarizes the details of the electron trigger selections. The electron trigger selections are highly efficient at removing jet sample events that are likely to reconstruct fake electrons. The isEM tight cut is by far the most efficient by nearly removing all events that can contain jets faking electrons.

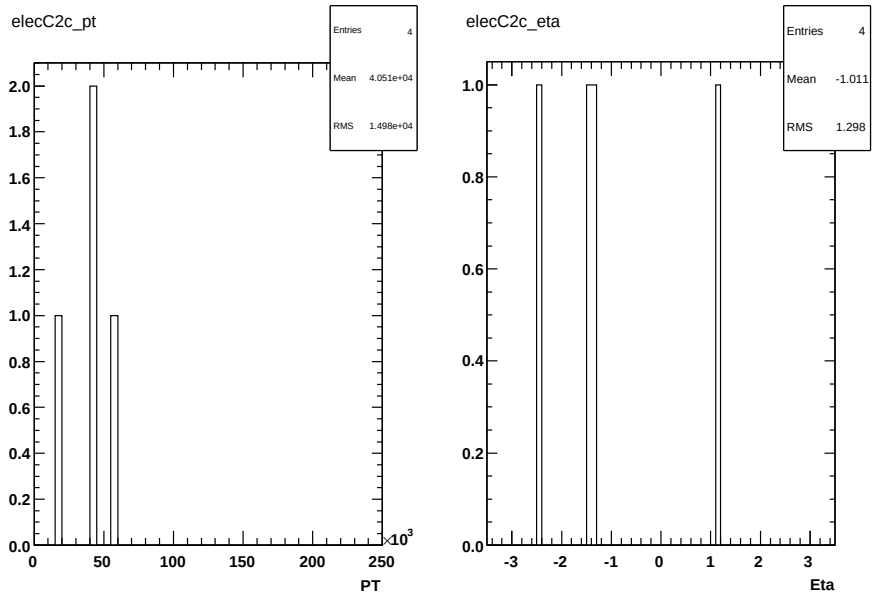

Figure 15: Distribution of p_T and η after Cut3c: isEM tight for the J3 jet sample

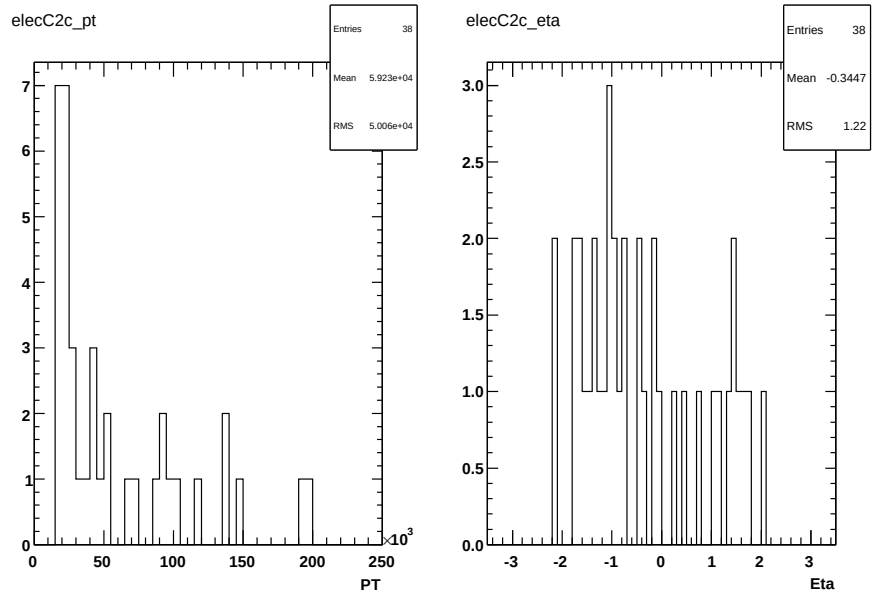

Figure 16: Distribution of p_T and η after Cut3c: isEM tight for the J4 jet sample

Conclusion

In this study, an electron identification scheme has been outlined for a comparative look at three different isEM methods, Loose, Medium and Tight, including reconstruction efficiencies. This is achieved by using a $Z \rightarrow ee$ signal sample. A quantitative description of QCD background is

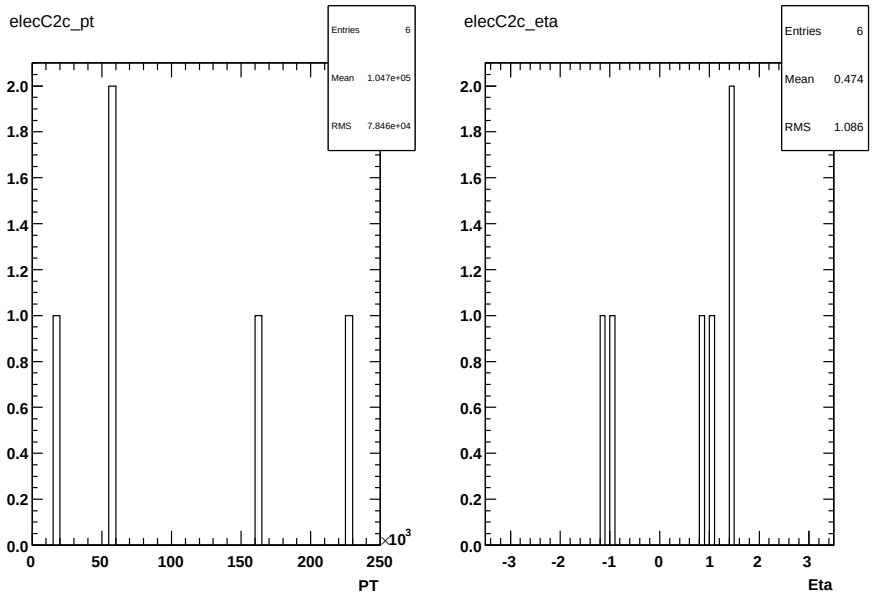

Figure 17: Distribution of p_T and η after Cut3c: isEM tight for the J6 jet sample

	J0	J2	J3	J4	J6
Total Number of Events	83250	19000	10000	33250	2750
isEM Loose	7	239	339	1602	160
isEM Medium	1	43	77	229	25
isEM Tight	1	3	4	36	6

Table 5: Reconstructed electrons using identification scheme for the three isEM methods

		J0	J2	J3	J4	J6
isEM Loose	No Trigger	7	235	332	1552	175
	e25i	2	5	4	11	0
	2e15i	0	0	0	0	0
isEM Medium	No Trigger	1	42	77	228	24
	e25i	0	2	2	4	0
	2e15i	0	0	0	0	0
isEM Tight	No Trigger	1	3	4	38	6
	e25i	0	0	0	3	0
	2e15i	0	0	0	0	0

Table 6: Number of Events passing the electron triggers for the three isEM methods

determined and a trigger analysis of electron physics triggers has also been included.