Searches for Non-Resonant New Physics in the High Energy Di-Electron Spectrum with ATLAS at the LHC

Liam Duguid

Department of Physics,

Royal Holloway, University of London,

Egham, Surrey, UK, TW20 0EX.

`lduguid@cern.ch`

Supervisor: Dr. Tracey Berry

A thesis submitted to the University of London for the

Degree of Doctor of Philosophy

January 24, 2014

DECLARATION

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the document.

Liam Duguid

Write thanks here

Abstract

Write abstract here

Preface

- **Chapter 1: Theory**

  This chapter covers an overview of the Standard Model (SM) of particle physics and then continue on to Beyond the Standard model (BSM) phenomena. The main focus is on the idea of Non resonant excesses in the dilepton Drell-Yan (DY) spectrum of which two examples are discussed. The first example is Contact Interactions, a model which describes many BSM phenomena that can show as four fermion contact interaction that exhibit a divergence from the SM DY spectrum. The example shown is that of a quark-lepton composite model where at a certain energy level quarks and leptons can form composite particles. The second example given is the Arkani-Hamed, Dimopoulos, and Dvali (ADD) model. This is a graviton theory with the addition of large extra spacial dimensions to dilute gravity. These large extra spacial dimensions create Kaluza-Klein resonances of the graviton very close to each other and so exhibit signs of Non-resonance behaviour. A look at past results for similar searches is also discused here.

- **Chapter 2: Experiment**

  This chapter covers an overview of the ATLAS experiment and the LHC. A particular focus will be given to the inner tracking detector and energy Calorimeters of ATLAS as these systems are the parts used in the detection of di-electron events used in this analysis. Although parts of the detector will be discussed in some respect.

- **Chapter 3: Trigger**

  This chapter focuses on the triggering system for selecting data events in the ATLAS detector. An overview of the whole system will be given but a focus made on the "egamma" part which selects electron and photon events. A slight detour will be made discussing the effect of increases in the luminosity of the LHC beam through the 2011-2012 data taking period and efforts taken to reduce high rates of data acquisition this entailed in the "egamma" chain.

- **Chapter 4: Reconstruction**

  This chapter details the algorithms used in reconstructing electrons and photons from the detector

output. It also contains a discussion on ATLAS assignments of *tight*, *medium* and *loose* electrons.

- **Chapter 5: Event Selection**

  This chapter covers the main event selection of di-electron events for the non-resonance analysis on the 20 fb^{-1} recorded in 2012. There will also be a discussion of and need for corrections applied to energy measurements.

- **Chapter 6: Background Estimate**

  This chapter discuses the estimate made of the background processes to the non-resonant signal. It covers the Monte Carlo (MC) generated to estimate these backgrounds as well as corrections applied to match MC to the data collection conditions used and corrections to account for next to next to leading order calculation effects.

- **Chapter 7: Signal Search**

  This chapter shows the search for new physics in the data collected in the 2012 data taking period. This includes a description of the MC used to predict the signals as well as comparison between the Data and the MC prediction of the background. Also looked at are the significance or p-value of any divergences from the SM background prediction.

- **Chapter 8: Statistical Analysis and Conclusion**

  This final chapter discuses a statistical treatment of the results. First discussed is possible sources of systematic error in the analysis as well as levels of statistical error. Next a Bayesian approach is taken to setting lower limits on the scale of new physics predicted by this analysis. Last is a discussion of how this result can be interpreted and a comparison to past searches for similar phenomena and current ones. A_μ

Trigger work:

- optimisation of 2012 electron triggers at LV1 up to HLT for higher luminoscity conditions and the required rate reductions.

Z Prime Analysis:

- Ran analysis of several MC sample for people. (including Black holes, pythia DY and samples for

Non-resonant Analysis:

- Full analysis coded and run by liam.

- production of reverse ID jets sample for Non-resonant anlysis.

- optimisation of new isolation cut.

- Study of new opposite sign cut and effect on reverse ID jets sample.

- Study of cosTheta* variable data MC comparison in control region.

- optimisation of binning for statistical analysis.

- Limit setting via Baysian analysis toolkit.

Contents

Chapter 1

Theory

1.1 Standard Model

-current status of SM. quick overview

1.2 Non-resonant new Physics

Beyond the Standard Model (BSM) or new physics models is a staple of the physics programs of the LHC detectors. Any theoretical models not contained within the Standard Model (SM) can fall in to this category and LHC experiments aim to search for as many of these models as are feasible within scope (Proton-Proton collisions and within the energy range of the LHC). Within the detection channel of two lepton decays (di-lepton), one evidence of new physics is non-resonant signals. This physics would show as a divergence from the SM prediction in the di-lepton mass spectrum unlike the resonant signals of particles such as the Z boson particle which shows as a peak in the di-lepton mass spectrum.

Non-resonant signals could be the results of many BSM theoretical models but two main theorys are presented here and their searches compose the rest of this thesis.

1.2.1 Contact Interaction Theory

The Standard Model (SM) assumes quarks and leptons to be fundamental particles in nature. This assumption is not without compelling argument but like the proton beforehand there is no reason quarks and leptons should not be composite structures or bound states of more fundamental particles, often referred to as Preons[?], at an energy scale Λ we have yet to reach.

One way quark-lepton compositness would exhibit itself as in 4-fermion contact interactions between two quarks from the incoming protons producing two final state leptons. This is how compositness is

searched for at the ATLAS detector.

Figure 1.1: Feynman diagrams of a contact interaction (left) and the predominant background Drell-Yan production (right).

Experimentally this interaction would be seen as a deviation from the Standard Model (SM) Drell-Yan (DY)($q\bar{q} \rightarrow \gamma/Z \rightarrow \ell^+\ell^-$) dilepton mass spectrum (seen in Fig. 1.1).

Without knowing the intermediate process one can write a Lagrangian describing the new Interaction:

$$\mathcal{L} = \frac{g^2}{2\Lambda^2}[\eta_{LL}(\bar{\psi}_L\gamma_\mu\psi_L)(\bar{\psi}_L\gamma^\mu\psi_L) + \eta_{RR}(\bar{\psi}_R\gamma_\mu\psi_R)(\bar{\psi}_R\gamma^\mu\psi_R) + 2\eta_{LR}(\bar{\psi}_L\gamma_\mu\psi_L)(\bar{\psi}_R\gamma^\mu\psi_R)]. \qquad (1.1)$$

Where g is the coupling constant and ψ_L and ψ_R are the left and right handed fermionic fields respectively. The sign of η defines whether the new interaction interferes constructively ($\eta = -1$) or destructively ($\eta = +1$) with the DY and is always unity. For previous analyses [**?**, **?**, **?**] a benchmark model of just the Left-Left (LL) component has been used and defined by $\eta_{LL} = \pm 1$ and $\eta_{RR} = \eta_{LR} = 0$. Here however an investigation of each of the three parameters is done individually. Both the LL and RR cases are expected to behave similarly however the LR case exhibits a different angular dependence than either of the other formalisms or the DY background. A differential cross section for this new interaction is given by

$$\frac{d\sigma}{dm_{\ell\ell}} = \frac{d\sigma_{DY}}{dm_{\ell\ell}} - \eta\frac{F_I}{\Lambda^2} + \frac{F_C}{\Lambda^4}, \qquad (1.2)$$

where $m_{\ell\ell}$ is the dilepton mass and Λ is the scale of the new physics. In the case of quark/lepton compositeness Λ refers to the point at which fermions stop being bound as SM quarks and leptons. F_I and F_C define the interference DY-CI term and the pure CI term respectively. The experimental signature of this effect is a broad increase on the SM background at high invariant mass for all models while the LR formalism also sees a different forward backwards asymmetry.

– add angular difference – add past limits

1.2.2 ADD Theory

ADD is a theory used to solve the hierarchy problem via large extra spacial dimensions first described by Arkani-Hamed, Dimopoulos, and Dvali (ADD) [**?**]. The ADD model predicts a graviton that propagates the extra dimensions and acquires Kaluza-Klein (KK) modes that show as a broad increase in the SM background. This is the reason the search for ADD is done in conjunction with CI.

1.3   Past Searches

Chapter 2

Experiment

This chapter will explore the ATLAS experiment in order to explain data specific to this analysis is obtained. First however is a discussion of the Large Hadron Collider which supplies the ATLAS experiment with proton collisions.

2.1   The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest and most powerful particle collider in the world with a circumference of 27 km and and design centre of mass collision energy of 14 TeV. During the 2012 run the accelerator was run at a centre of mass energy of 8 TeV while providing an integrated luminosity of just above 20 fb^{-1} throughout the year to its two general purpose experiments, CMS and ATLAS, that latter of which provided data for this analysis. However the LHC can not run in isolation to provide beams for its 4 main experiments, instead it is the last and newest accelerator in a chain of accelerators which extract protons from a hydrogen canister with little to no momentum and inject them in to the LHC as a 450 GeV beam. The proton source is a device called a Duoplasmatron which injects hydrogen gas in to a strong electric field striping electrons from their nuclei. The remaining protons are injected in to Linac 2, a linear accelerator which accelerates them to an energy of 50 MeV. The BOOSTER or Proton Synchrotron Booster (PBS) comes next in the chain and accelerates protons from 50 MeV to 1.4 GeV to be injected in to the main Proton Synchrotron (PS). The PS accelerates protons up to an energy of 25 GeV and again injects them in to another accelerator, the Super Proton Synchrotron (SPS). The SPS is the final stage before injection in to the LHC ring and pushes protons to an energy of 450 GeV. Protons from the SPS then get injected in to the LHC in both counter revolving directions and accelerated to their final collision energy. For the data used in the analysis that follows (the 2012 data run) the final proton beam energy is 4 TeV giving a final centre of mass collision energy of 8 TeV.

The LHC itself is built in the same tunnel used by the Large Lepton-Positron (LEP) collider. Based at CERN (Centre of European Nuclear Research) the 27 *km* tunnel is between 50 to 175 *m* underground and like CERN itself crosses the French-Swiss border just outside Geneva. Construction of the LHC started in 2001 after the LEP collider was decommissioned dismantled with excavation of the caverns for the LHC's four main experiments starting slightly before in 1998. The LHC is a synchrotron machine requiring 1,232 super-conducting Niobium-Titanium dipole magnets each providing an 8.33 *T* magnetic field to direct the proton beams around its loop and an additional 392 quadrupole magnets of the same type to focus the beams for the collision points. The super conducting magnets operate at 1.9 *K* with the whole accelerator requiring 96 tonnes of liquid helium to remain cooled.

-protons per bunch

-number of bunches

-luminosity (design and running for all of these)

Four collision points exist around its circumference providing events to the four experiments; ATLAS (A Toroidal LHC Apparatus), CMS (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment) and LHCb (Large Hadron Collider beauty). ATLAS and CMS are both general purpose experiments designed to look for a variety of physics. ALICE is designed specifically to study quark-gluon plasma in heavy ion collisions scheduled for the end of each LHC run period while LHCb looks for beauty mesons in a search for CP-violation. There are also three additional LHC detectors in various stages of deployment without their own collision points; TOTEM (Total Elastic and diffractive cross section Measurement), LHCf (LHC forward) and MoEDAL (Monopole and Exotics Detector at the LHC) which measure separate beam properties. TOTEM shares CMS's collision point aiming to measure the proton cross-section very accurately while LHCf shares ATLAS's collision point measuring the very forward region of collision with the hope of investigating the source of ultra-high-energy cosmic rays. MoEDAL shares a cavern with LHCb and targeted to search for magnetic monopoles and other highly ionising stable massive particles.

2.2 ATLAS - A Toroidal LHC Apparatus

- image of whole detector

The ATLAS detector sits 100 m underground just over the road from the main CERN site and at 45 m long, 25 m in diameter and weighing over 7,000 tons is one of largest and most complex particle physics experiments in the world. The Detector it's self can be divided in to four main subsystems and from the interaction point out they are; the Inner Detector or tracking detector, the Calorimeters both Electro-Magnetic

(EM) and Hadronic, the Magnet system and the Muon Spectrometer. There is also a small set of forward detectors for accurately measuring the integrated luminosity provided to ATLAS by the LHC.

$$\eta = -\ln[\tan(\frac{\theta}{2})] \tag{2.1}$$

- detector coordinates system.

- sub detector coverage and diagram.

Broadly the detector is also divided in to the barrel region (cylinder surrounding the interaction point) and endcap region (circles covering the ends of the barrel region) which use slightly different configurations and technology in order to cover a full range in η. Following is a description of each main subsystem while focusing particularly on both the Inner Detector and EM Calorimeter as these are the important systems in identification of electrons used for this analysis.

- calibration?

2.2.1   Inner Detector

-good diagram

The Inner Detector is ATLAS's main tracking detector fitted as the closest part to the interaction point. A tracking detector is needed to trace charged particles from the interaction point out to the calorimetry system and give two bits of information; a charged particles position to match with the calorimeters (or Muon Spectrometer in the case of muons) and when a magnetic field is present an estimate of a particles momentum to compare with the calorimeter obtained from the radius of its curve. It is composed of three different tracking technologies in order going out from the collision point; the Pixel Detector (PD), the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). The Inner Detector was designed to precisely measure charged tracks in the energy range 0.5 GeV - 150 GeV while complimenting the energy measurements of the Calorimetry system. Covering a range of $|\eta| < 2.5$ and full range in ϕ the Inner Detector with the help of the 2 T magnetic field imposed by the solenoid magnet (discussed bellow) boasts a momentum resolution of $\sigma_{p_T}/p_T = 0.05\% \, p_T \oplus 1\%$ to charged tracks. In its design it was also important for the Inner Detector to be able to distinguish between multiple primary vertices at the collision point, referred to as pile-up, as well as secondary vertices from sources such as the hadronisation of b quarks. (-stats as to precision?)

Pixel Detector

The Pixel Detector is the first layer and closest to the beam line consisting of three layers of silicon pixels. Because of its proximity to the beam line the pixel detector had to be designed to be heavily radiation hard and understood so its performance could be predicted over an extended period of radiation exposure. The Pixel Detector is made of a barrel and two endcaps composed of 1744 modules all together.

(- keep cool to avoid leakage current.) (-read out channels?)

Semiconductor Tracker

The SCT consists of the same technology as the PD but organised in to four layers in the barrel region and 9 layers in each endcap. Due to the amount of electronics cooling is rather important to this layer and so the sensors in each module are glued to each side of a thermally conductive spine that gives the SCT structure and transports heat out via the mounting point of each module keeping them at their operating temperature of $-7\,°C$. (-read out channels?)

Transition Radiation Tracker

The TRT uses a completely different tracking technology to the rest of Inner detector it uses a straw detectors composed of 4 mm diameter polymide tubes each with a $31\,\mu m$ diameter gold plated Tungsten-Rhenium wire. Due to the small diameter of straws the TRT can obtain the high read-out rate needed for experiments at the LHC. The barrel region consists of 50,000 of these straws with a readout at each end providing 100,000 readout channels. The endcaps contain another 320,000 straws only read out a one end giving the TRT a total of 420,000 channels. Each channel measures drift time giving a resolution of $170\,\mu m$ in each straw. The straws are filled with a high Xenon concentration ($Xe(70\%)CO_2(27\%)O_2(3\%)$) of gas in order to detect electrons via radiated photons as they traverse the material between straws (-what is this material?). This is achieved by giving each straw two timing thresholds, the lower to discriminate tracking hits (direct hits) while the higher threshold discriminates transition radiation hits. (-eta ranges of barrel end cap.) (-number of straw hits in each region)

2.2.2 Calorimeters

-good diagram

While the Inner Detector only measures charged particles, the Calorimeters measure both neutral and charged particles and is split in to two different for particles with differing properties. The inner Electromagnetic Calorimeter is designed to measure primarily electrons and photons as well as pions while the Hadronic Calorimeter found outside it looks for hadrons such as neutrons and protons

Electromagnetic Calorimeter

Hadronic Calorimeter

2.2.3   Magnet System

- Solenoid magnet - Toroid magnet

2.2.4   Muon Spectrometer

-good diagram

2.2.5   Forward Detectors

- performance and faults?

Chapter 3

The Trigger & Data Acquisition

3.1 Level-1 Trigger

3.2 Level-2 Trigger

3.3 Event Filter

3.4 Data Acquisition

3.5 Trigger Menu and Rates

3.5.1 The "e/γ" Trigger Menu

3.5.2 Trigger Rates in High Luminosity Regime

The ATLAS trigger system comprises a hardware-based Level-1 (L1) trigger and a software-based High Level Trigger (HLT), subdivided into the Level-2 (L2) and Event Filter (EF). Due to the bandwidth limitations of the trigger each level is restricted to a certain output rate. During 2011 the L1 output rate was kept below 60 kHz, L2 below 5 kHz and the EF output rate at around 400 Hz averaged over the LHC fills. The bandwidth allocated to the e/γ triggers was approximately 30% of the total EF output rate. Electron and photon identification is accomplished by a set of η- and E_T-dependent rectangular cuts variables [**?**, **?**]. Throughout 2011 data taking at ATLAS the luminosity continued to increase putting pressure on the trigger's ability to control the output rate. Several methods were employed to reduce the trigger rate and in the e/γ trigger a variable threshold and hadronic core isolation was investigated to reduce the rate of the Level-1 trigger. In order to keep within time constraints only a coarse granularity is available Level-1 trigger in regions of 0.4 η. Threshold requirements were therefore investigated varying every 0.4 η. The effect of

a hadronic core isolation was also investigated on the selection of electrons which defines a region in the hadronic calorimeter behind the e/γ candidate in which you require and minimum amount of energy to be deposited to distinguish between jets and e/γ objects.

Figure 3.1: Performance of the first level of the ATLAS e/γ trigger before(EM16) and after(EM16VH) variable thresholds and hadronic core isolation are applied.

These attempts where successful and rate reduced to compensate for the high luminosity environment. Fig. 3.1 shows the performance of the trigger after these changes had been made. It can be seen that a minimal impact of these new requirements is felt.

There was also a contribution to the maintenance of the e/γ trigger software run in the ATLAS detector, both these tasks forming the authorship qualification. The authorship task culminated in presentation of a poster on behalf of the e/γ trigger group at the Computing and High Energy Physics (CHEP) conference held New York in May 2012. The contribution is included two pages previously and details the performance of the e/γ trigger in the 2011 run period [**?**].

Chapter 4

Reconstruction

- Reconstruction of electron objects and selection of "good" candidates.

4.1 Data Formats and Software

4.2 Electron Reconstruction

Chapter 5

Event Selection

The main event selection for this analysis is based on a standard cut-flow selection used within ATLAS to select high energy di-electron events. Following will be the basic outline of each requirement an event must satisfy followed by a small discussion of optimisations done to some cuts for this analysis. We then discuss corrections implemented at this stage in addition to the reconstruction mentioned in the previous chapter.

5.1   Analysis Selection

The following selection is made data after being reconstructed by the ATLAS software.

Event Selection

- Each event is required to contain at least one reconstructed primary vertex with more than 2 charged tracks traceable to it.

- Event is required have passed the chosen unprescaled electron trigger (EF_g20_loose).

Electron Selection

- Each electron is required to have a transverse momentum (p_T) greater than 30 GeV.

- Electron $|\eta| < 2.47$ and not lie within the detector crack region $1.37 \leq |\eta| \leq 1.52$ due to a decreased energy resolution.

- Electrons are required to pass identification criteria on the transverse shower shape, the longitudinal leakage into the hadronic calorimeter, and the association to an inner detector track, defined together as a medium electron identification.

- If expected electron is required to have signal in the inter most level of the tracking detector (B-layer). Used to suppress background from photon conversions.

Dielectron Selection

- Selection of two highest p_T electrons left in event.

- Lead Isolation (A cone around the candidate in the calorimeter is required to have $< 0.007 \times E_T + 5.0\, GeV$ deposited in it) of the highest p_T electron in the event is used to suppress jet's background.

- Lead Isolation (A cone around the candidate in the calorimeter is required to have $< 0.0022 \times E_T + 6.0\, GeV$ deposited in it) of the highest p_T electron in the event is used to suppress more jet's background.

- Dielectron invariant mass (m_{ee}) is required to be greater than or equal 80 GeV.

- Same sign requirement. Require both electrons to be identified with the same sign.

5.1.1   Isolation Requirement

It was decided a re-investigation of the isolation requirement was needed, updating the cut from its previous iteration in the di-electron analysis on 2011 ATLAS data. The previous threshold was a flat, less than 7 GeV, cut on the p_T corrected (define pT corrected) electro-magnetic calorimeter cluster isolation (definition probably required) of the highest p_T electron in the selected pair. The first investigation was to see how this cut performed in the selection of MC signal at 8 TeV centre of mass energy. Due to the better statistics found in the *DY→ee* MC sample and this being an irreducible background and therefore indistinguishable from the signal this was used in the following investigation.

It can be seen in fig. # that this flat cut of 7 GeV causes an increasing efficiency loss at high energy and was deemed unsuitably high for this iteration of the analysis due to the higher reach in energy expected from the higher centre of mass energy. Two methods that could be used in conjunction were proposed to solve this issue as well as the possibility of introducing a requirement on the second highest energy electron in our selected pair.

The first alternative was a different algorithm of calculating the isolation variable, topo isolation (definition required). This was deemed unsuitable after a short investigation due the similarity of distribution of final result and problems with the implementation of this algorithm in ATLAS reconstruction code. (remember and find out the exact reason)

The second possibility was an isolation requirement varying with energy. Fig. 5.1 shows the distribution of DY MC events in E_T and cluster isolation. It can be seen that electrons become less isolated under this definition of isolation as energy of the electron increases. This is to be expected as higher energy electrons produce larger showers in the EM calorimeter and so are less well restrained to only a few colorimeter cells in the electron shower.
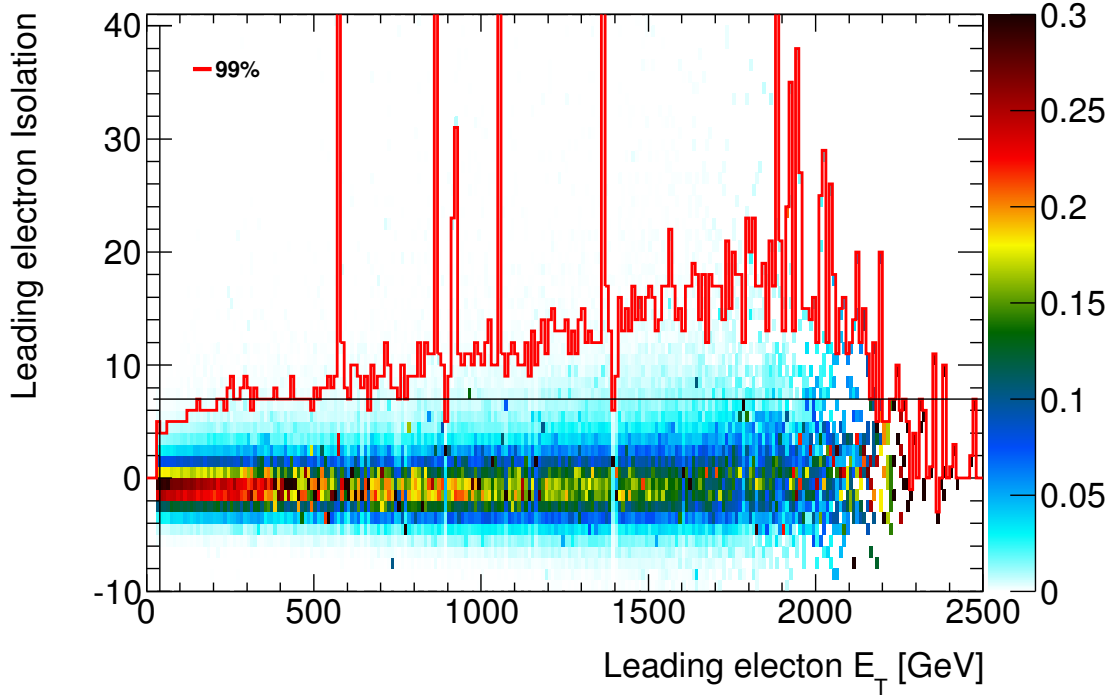
Figure 5.1: Distribution of DY MC in E_T and cluster isolation for the highest energy electon. Colour density shows the fraction of electons from that E_T coulum found in cell. The red line shows the 99% acceptance point of electrons in the E_T coloum. While the Black vertical and horizontal lines show the p_T and old isolation cut respectivly.

In order to define a requirement varying in pT the 99% acceptance point for each pT column was looked at and a 1st order polynomial fit to these points by eye was done. The 99% acceptance points can be seen in fig. 5.2 as well as the estimated fit which would form the cut. The same thing was looked at for the second highest pT electron and can be seen in fig. 5.3.

The two first order polynomials shown here correspond to isolation requirements of;

$$Lead\ Isolation < 0.007 \times E_T + 5.0\ GeV$$

$$Subleading\ Isolation < 0.022 \times E_T + 6.0\ GeV$$

for the highest and second highest energy electrons respectively.

An analysis of the efficiency of these cuts on signal can be seen in fig. 5.4 and fig. 5.5 where it can be seen they maintain a flat behaviour as E_T increases.

The main source of background the isolation cut is imposed to reduce is jets that fake an electron signal in the detector. It is therefore important to measure the effect of this new requirement. Jets background is estimated via a reverse ID method on data (see Section #) and so lacks statistics at high energy. For this reason it is impossible to optimise the isolation requirement against rejection of high energy fakes as seen in fig. #. Fig. # shows the efficiency of acceptance of jet fakes against pT and shows a sizeable increase in efficacy over the flat cut (fig. #) used previously.
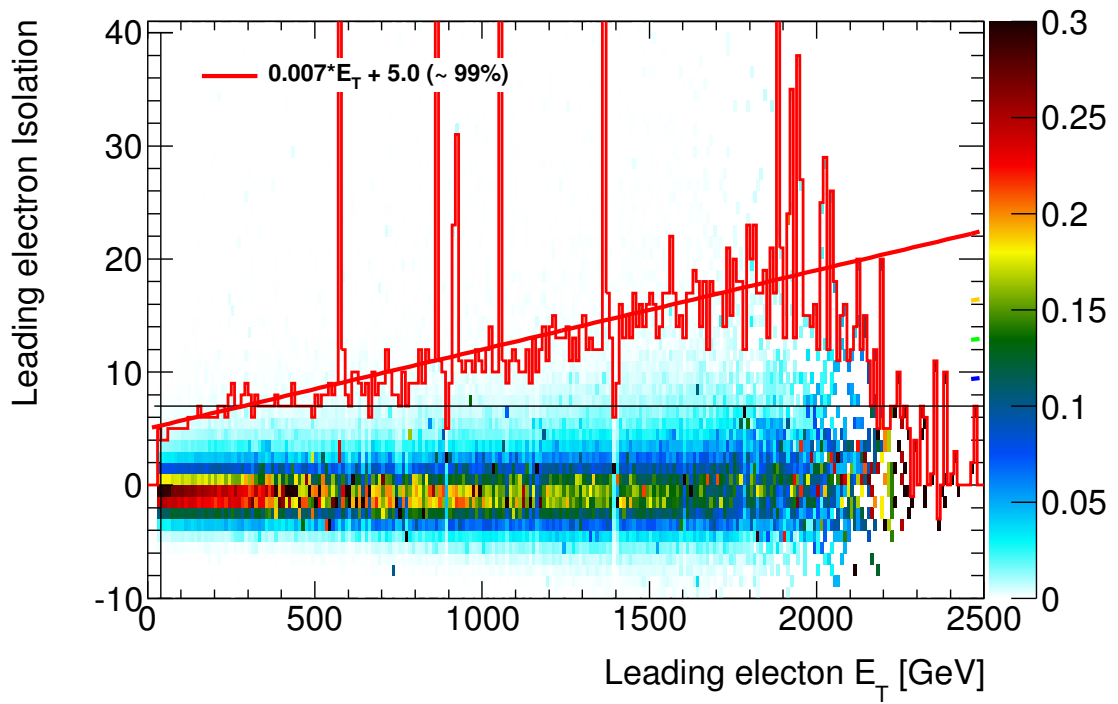
Figure 5.2: Similar plot to Fig. 5.1 but with a fit 99% as a possible issolation requirment.

5.1.2 Same Sign requirement

5.2 Corrections

5.2.1 Energy Resolution Correction
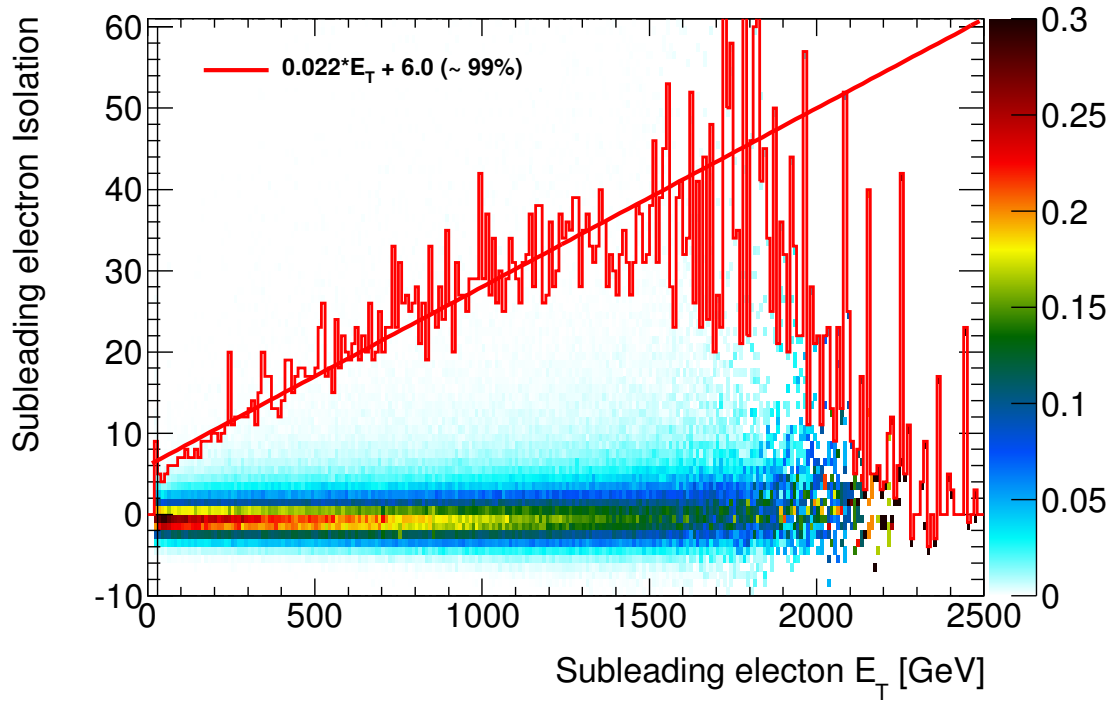
Figure 5.3: Similar plot to Fig. 5.2 but for second highest energy electrons after the requirment proposed in Fig. 5.2 is applied.
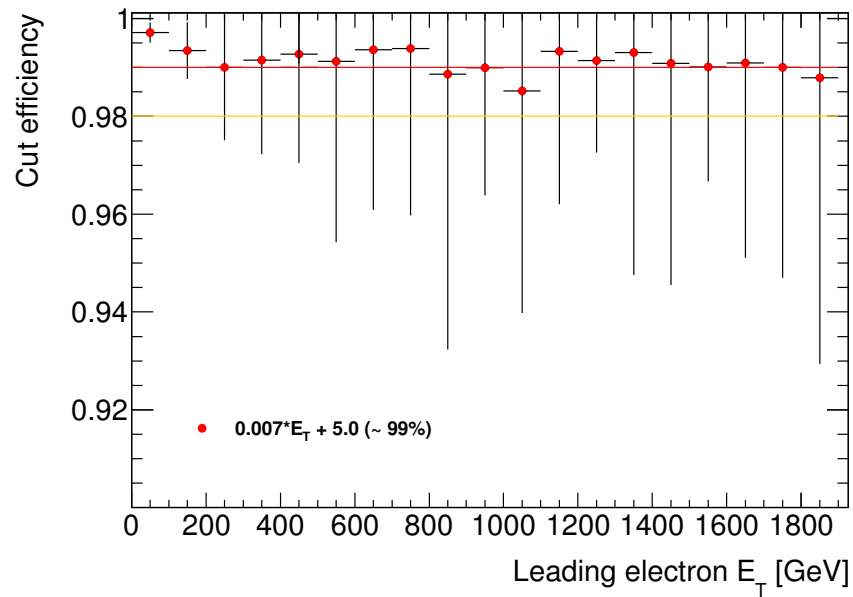
Figure 5.4: Efficency of new leading electon isolation cut on selection of signal MC. Red and orange lines indicate the 99% and 98% efficency levels respectivly.

Figure 5.5: Efficency of new subleading electon isolation cut oxn selection of signal MC. Red and orange lines indicate the 99% and 98% efficency levels respectivly.

Chapter 6

Background Estimate

6.1 Monte Carlo samples

- MC's used, PDF's, k-factors

6.1.1 Detector Simulation and Object Reconstruction

-GEANT 4 and ATHENA

6.1.2 PDF Choice and NNLO Corrections

6.1.3 MC Corrections

6.2   Fake Factor Multi-Jet Estimate

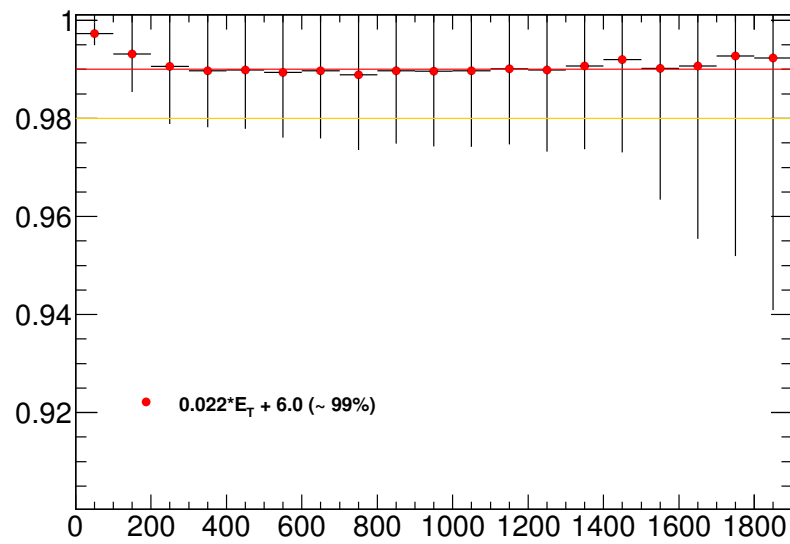One of the major sources of background to di-electron signals are di-jets or electron+jets (mainly W+jets) events where one or both selected leptons are jets faking electron signatures. The method for estimating this background, described here, is a "fake factor" or "matrix-method". This is a data-driven method where electrons are selected by a tight (N_{tight}) and loose (N_{loose}) selection. The tight selection is the standard electron selection used in this analysis while the loose selection has no isolation requirement and must only pass a loose++ egamma definition (see Chapter 4) with no track matching criteria. N_{tight} is therefore by design a subset of N_{loose}. Two more hidden values are also assigned *real* and *fake* referring to true source of each electron. This gives us two coefficients to determine from data.

$$f = \frac{N_{tight}^{fake}}{N_{loose}^{fake}} \qquad r = \frac{N_{tight}^{real}}{N_{loose}^{real}} \qquad (6.1)$$

The fake rate f denotes the probability that a *fake* electron which passes the loose requirement also passes tight while r refers to the probability that a *real* electron which passes the loose requirement also passes the tight. Reconstructed events are split in to two distinct groups, tight(T), and loose while failing tight(L), where *Tight* is now no longer a subset of *Loose*. This allows us to relate our reconstructed events to the underling truth events via a matrix of fake rates shown in Eq. 6.2.

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = \begin{pmatrix} r_1 r_2 & r_1 f_2 & f_1 r_2 & f_1 f_2 \\ r_1(1-r_2) & r_1(1-f_2) & f_1(1-r_2) & f_1(1-f_2) \\ (1-r_1)r_2 & (1-r_1)f_2 & (1-f_1)r_2 & (1-f_1)f_2 \\ (1-r_1)(1-r_2) & (1-r_1)(1-f_2) & (1-f_1)(1-r_2) & (1-f_1)(1-f_2) \end{pmatrix} \begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix} \qquad (6.2)$$

The first index in Eq. 6.2 refers to the highest p_T electron while the second index refers to the second highest p_T electron. So N_{LT} indicates the reconstructed events with highest p_T electron only passing the *Loose* selection while the second highest p_T electron passes *Tight* selection. The indices 1 and 2 refer to fake rates (f) and efficiencies (r) on leading and sub-leading electrons respectively.

The interesting part for this study is the contribution to N_{TT} coming from sources other than N_{RR}, these can be seen in Eq. 6.3.

$$N_{TT}^{\ell+jets} = r_1 f_2 N_{RF} + f_1 r_2 N_{FR}$$

$$N_{TT}^{di-jets} = f_1 f_2 N_{FF}$$

$$N_{TT}^{\ell+jets \; \& \; di-jets} = r_1 f_2 N_{RF} + f_1 r_2 N_{FR} + f_1 f_2 N_{FF} \tag{6.3}$$

This function however contains hidden variables and so Eq. 6.2 is inverted to derive a better formalism.

$$\begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix} = \alpha \begin{pmatrix} (f_1-1)(f_2-1) & (f_1-1)f_2 & f_1(f_2-1) & f_1 f_2 \\ (f_1-1)(1-r_2) & (1-f_1)r_2 & f_1(1-r_2) & -f_1 r_2 \\ (r_1-1)(1-f_2) & (1-r_1)f_2 & r_1(1-f_2) & -r_1 f_2 \\ (1-r_1)(1-r_2) & (r_1-1)r_2 & r_1(r_2-1) & r_1 r_2 \end{pmatrix} \begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} \tag{6.4}$$

where,

$$\alpha = \frac{1}{(r_1-f_1)(r_2-f_2)} \tag{6.5}$$

The fraction of selected events with at least one fake is then given by Eq. 6.2.

$$\begin{aligned} N_{TT}^{\ell+jets \; \& \; di-jets} = \;\; & \alpha r_1 f_2 [(f_1-1)(1-r_2)N_{TT} + (1-f_1)r_2 N_{TL} + f_1(1-r_2)N_{LT} - f_1 r_2 N_{LL}] \\ + \;\; & \alpha f_1 r_2 [(r_1-1)(1-f_2)N_{TT} + (1-r_1)f_2 N_{TL} + r_1(1-f_2)N_{LT} - r_1 f_2 N_{LL}] \quad (6.6) \\ + \;\; & \alpha f_1 f_2 [(1-r_1)(1-r_2)N_{TT} + (r_1-1)r_2 N_{TL} + r_1(r_2-1)N_{LT} + r_1 r_2 N_{LL}] \end{aligned}$$

$$\begin{aligned} = \;\; & \alpha[r_1 f_2(f_1-1)(1-r_2) + f_1 r_2(r_1-1)(1-f_2) + f_1 f_2(1-r_1)(1-r_2)]N_{TT} \\ & + \alpha f_2 r_2[r_1(1-f_1) + f_1(1-r_1) + f_1(r_1-1)]N_{TL} \\ & + \alpha f_1 r_1[f_2(1-r_2) + r_2(1-f_2) + f_2(r_2-1)]N_{LT} \\ & - \alpha f_1 f_2 r_1 r_2 N_{LL} \end{aligned} \tag{6.7}$$

Equation 6.7 shows the derived formula relating the multi-jet background to fake rates, efficiencies and four independent samples selected from data. Detailed here is this method used on the full 20 fb^{-1} of integrated luminosity from ATLAS's 2012 run.

6.2.1 Real electron efficiency estimation

The real electron efficiency is defined as Eq. 6.1 $r = N_{tight}^{real}/N_{loose}^{real}$. This is determined from MC using a mass binned Drell-Yan sample. The efficiencies are found for both the leading and sub-leading electrons and binned in 8 p_T and three eta bins of $|\eta| < 1.37$ (barrel), $1.52 < |\eta| < 2.01$ and $2.01 < |\eta| < 2.47$ (endcap). The efficiency is distributed between 90 - 96% as can be seen in Fig. 6.1.
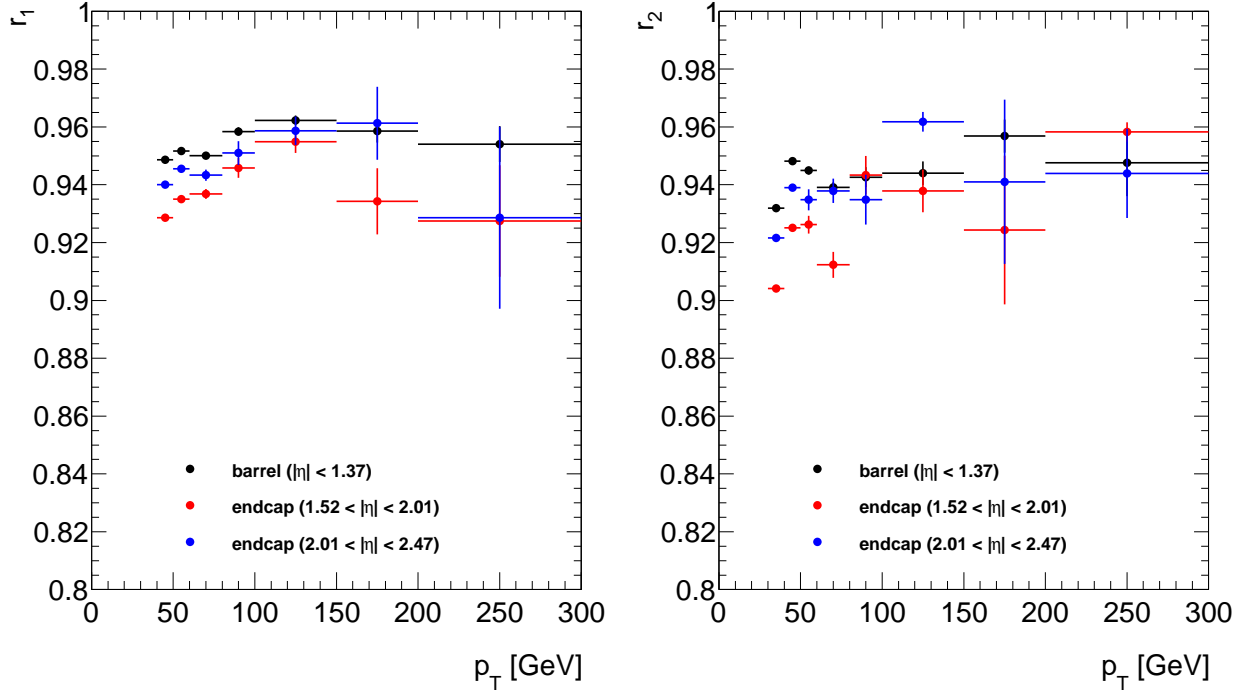

Figure 6.1: Real electron efficiencies obtained from Drell-Yan MC and binned in p_T and three coarse η bins covering the barrel and two endcap regions. Efficiencies for leading electrons are shown on the left while those for subleading electron are on the right.

6.2.2 Fake electron rate estimation

The default method selected for analysing the fake rates is a single object method selection on the jet stream data. This gives the main advantage of more statistics and a higher energy reach compared to methods such as using tag and probe on the egamma stream data. An array of triggers are used for selecting suitable events based on the single jet trigger EF_jX_a4chad (where X = 25, 35, 45, 55, 80, 110, 145, 180, 220, 280, 360). Events are associated to groups with the lowest trigger threshold they pass as each trigger has a different prescale. Objects are selected with the AntiKt4TopoEMJets algorithm and then matched to objects in the egamma stream with a $\Delta R < 0.1$. Objects also have to pass the medium jet-cleaning criteria (define this). Two further steps are taken to suppress real electrons from W decays and real Drell-Yan events. A veto of

$E_{Tmiss} > 25\ GeV$ is introduced to combat the former while events with two medium++ or loose++ electrons with $|m_{tag\ \&\ probe} - 91\ GeV| < 20\ GeV$ are vetoed to counter the real Drell-Yan.

The fake rate is then defined as Eq. 6.1 $f = N_{tight}^{fake}/N_{loose}^{fake}$ with distributions selected using the standard event selection on the matched egamma objects. Due to the different prescales of each trigger a separate set of fake rates are calculated for each trigger, these are then combined as a weighted average of all fake rates. Fig. 6.2 shows the distribution of fake rates for leading and subleading fakes which are distributed between 3 - 20%.
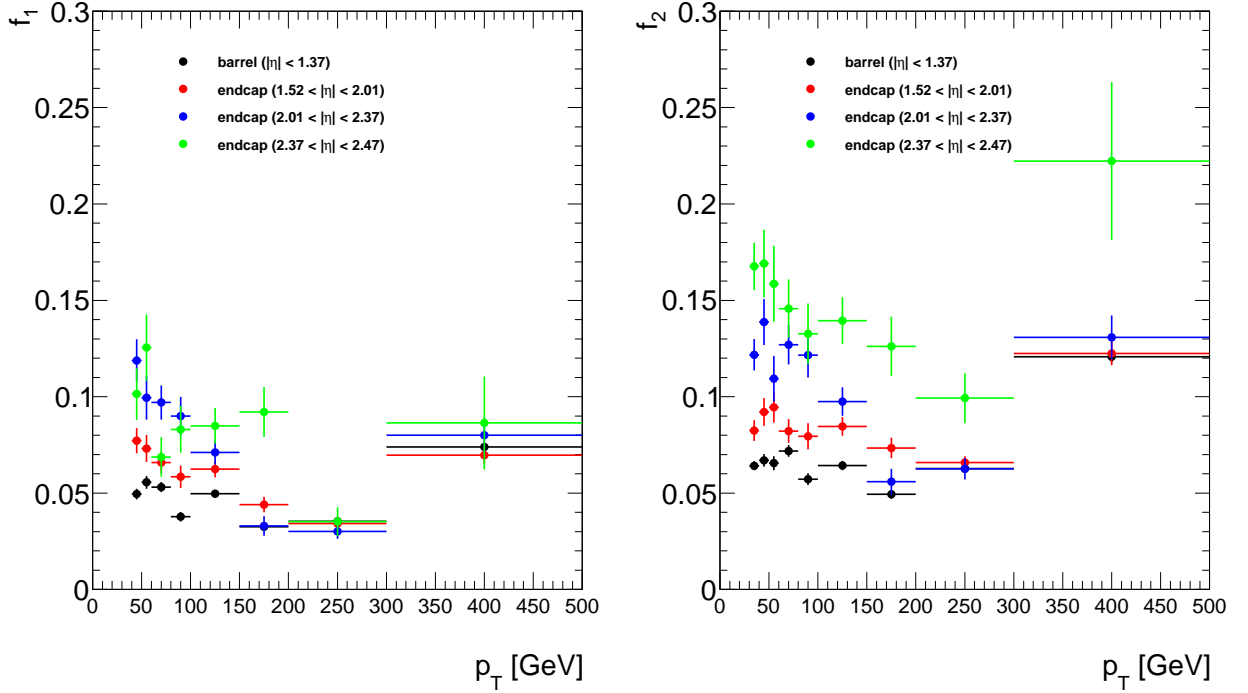
Figure 6.2: Fake rates obtained from data and binned in p_T and four coarse η bins covering the barrel and three endcap regions. Fake rates for leading electrons are shown on the left while those for subleading electron are on the right.

6.2.3 Properties of Multi-Jet Background

In order to compose the final sample events are organised by the distributions N_{TT}, N_{TL}, N_{LT} or N_{LL} and weights are applied according to each electrons p_T, η with respect to Eq. 6.7 and the corresponding efficiencies and fake rates. Fig. 6.3 shows these distributions before the efficiencies and fake rates are applied to weight to the final background prediction. In addition to these steps an extra fit is then applied at low invariant mass due to contamination due to the Z boson peak. This method is not suited to predicting the Multi-Jet background in the Z boson peak region and so a fit is obtained between 120 GeV and 400 GeV and stitched from 110 GeV and bellow. This gives a good estimate to the integral in this region for use in

scaling MC's to luminosity but is not predicted to be good at predicting other variables in this region.
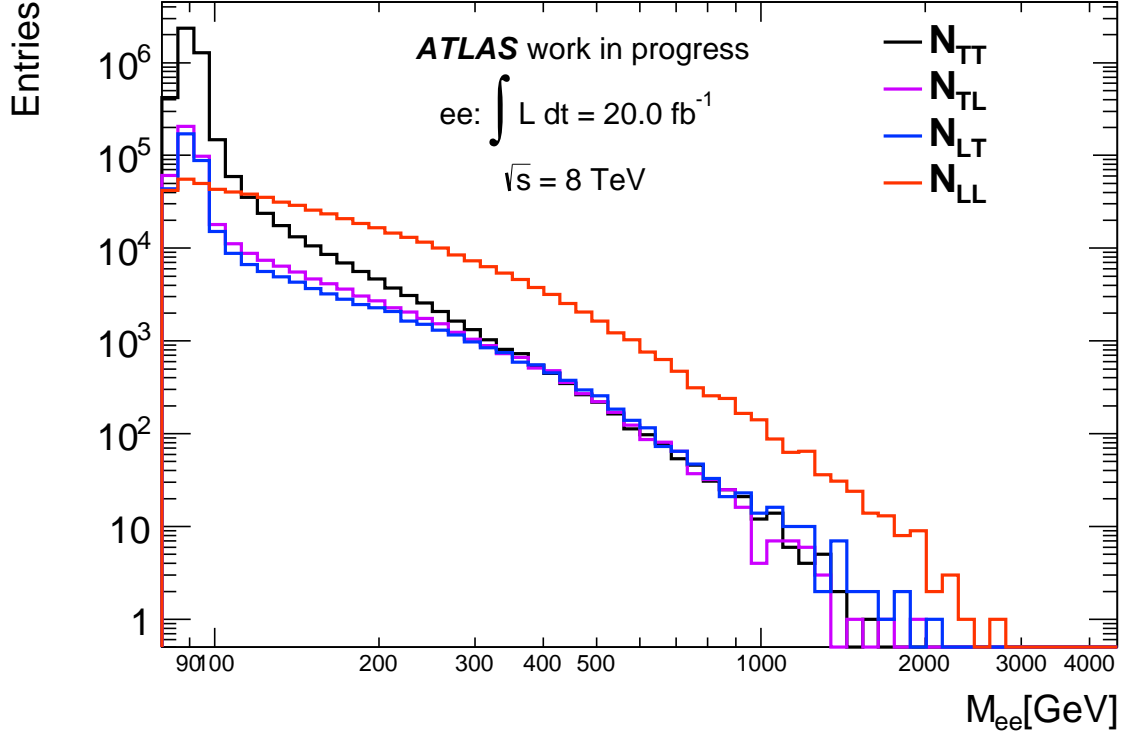
Figure 6.3: Distribution of N_{TT}, N_{TL}, N_{LT} and N_{LL} from data with no weightings applied.

6.2.4   Other methods and estimation of Error

Two other methods and variations upon them were used to test the validity of this method as well as estimate the systematic error of this background estimates procedure. These two methods are both tag and probe measurements on either the jet stream of data, or the egamma stream where the method is more an "inverse" tag and probe with the selection of a tag with high probability of being a jet. Variations are also made on the method by assuming r_1 and $r_2 = 1.0$ in all cases as well as changing the definition of loose but fail tight. These variations simplify the equations slightly but the method remains the same. Fig. 6.4 shows all of these variations compared to the default method used to obtain the estimation. This figure then gives us a good estimate to the systematic uncertainty of the multi-jet estimate which has been chosen to be 20%. (ref zprime support note)
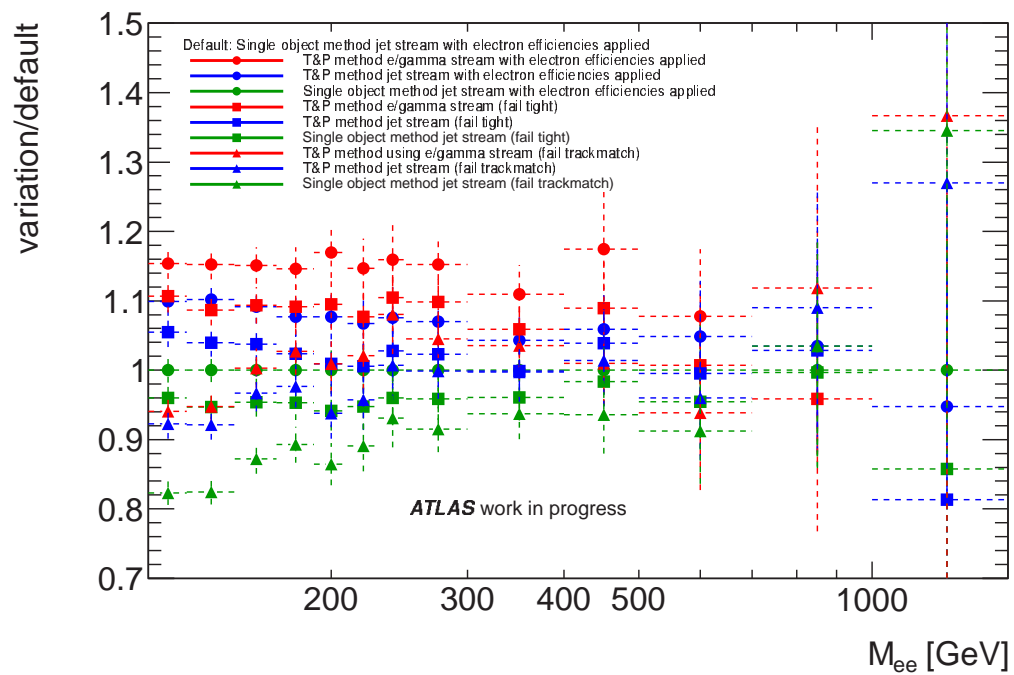
Figure 6.4: Ratio of the final background estimate of all method variations to the default method. The ratio starts at 116 GeV and ends at 1500 GeV. Source: resonant support note, will not be able to use this.

Chapter 7

Signal and Results

7.1 Signal Monte Carlo

- MC, PDF's, k-factors and parameters

7.1.1 PDF and Corrections

7.2 Results

- Data MC comparisons

Chapter 8

Statistical Analysis

8.1   Systematics

8.2   Signal Search & P-Values

8.3   Setting Limits

8.3.1   Contact Interation Limits

8.3.2   ADD Limits

Chapter 9

Conclusion

9.1 Interpretation

9.2 Looking Forward

Appendices

Appendix A

Non-Resonance Analysis 2011

A.1 Data and Background Processes

A.1.1 Data

All data used in the CI analysis is taken from the LHC 2011 $\sqrt{s} = 7TeV$ proton-proton collision data of which ATLAS recorded $4.9fb^{-1}$ of electron candidate data. Data was collected with stable LHC beams and a fully operational inner detector and calorimeter each being important in the identification of good electron candidates.

A.1.2 Background

The main background the CI signal in the electron channel is Drell-Yan (DY) $\rightarrow ee$ production mediated by a photon or the Z boson. However there are other small contributions from $t\bar{t}$, diboson, W + jets and QCD production. $t\bar{t}$ background consists of events with $t\bar{t}$ production decaying to amongst other things two electrons. Diboson background can involve either an event producing two W bosons, two Z boson or one W and one Z boson in which two electrons also exist in the final decay state. Both these productions can result in two electrons with a large combined invariant mass which could mimic CI production. W + jets production can result in an electron and a jet faking an electron left in the final state, while QCD refers to events where two jets fake electrons. These all combine to form the background sample.

A.1.3 Background Production

Contributions from SM processes were primarily simulated via leading-order (LO) $PYTHIA$ [?] Monte Carlo (MC) event generation and using $Geant4$ [?] to simulate the ATLAS detector. This method was used to generate a $Z \rightarrow ee$ sample for the low dielectron invariant mass region ($m_{ee} < 120GeV$) and a

$DY \rightarrow ee$ mass binned sample for the invariant high mass ($m_{ee} > 120 GeV$) to keep high statistics at high invariant mass. Four other samples are included to produce the background estimate, these are $t\bar{t}$ (produced with *MC@NLO*), diboson(WW, WZ and ZZ decays produced with *HERWIG*), W + jets (produced with *ALPGEN*, *JIMMY* and *HERWIG*) and QCD (produced using a data driven method [**?**]).

A.1.4 Signal

Five benchmarks for the value of Λ where chosen for the CI signal samples for both constructive and destructive interference. Like the DY these were also produced with LO *PYTHIA* containing both the pure DY contribution as well as the interference and pure CI components. Samples where produced above dilepton invariant mass of 120 GeV to increase statistics above the Z^0 peak where new physics would appear.

m_{ee} [GeV]	120-300	300-600	≥ 600
$\Lambda^- = 3$ TeV	9.8583	0.96765	0.77563
$\Lambda^- = 4$ TeV	9.3400	0.50510	0.26647
$\Lambda^- = 5$ TeV	9.0960	0.36935	0.12733
$\Lambda^- = 7$ TeV	8.9686	0.28401	0.048406
$\Lambda^- = 12$ TeV	8.9037	0.24774	0.021028
$\Lambda^+ = 3$ TeV	8.9018	0.65247	0.66034
$\Lambda^+ = 4$ TeV	8.8756	0.33164	0.20668
$\Lambda^+ = 5$ TeV	8.7362	0.25646	0.085934
$\Lambda^+ = 7$ TeV	8.8101	0.22656	0.028772
$\Lambda^+ = 12$ TeV	8.9045	0.22774	0.014418

Table A.1: Table of CI sample cross sections [pb^{-1}].

Corrections are applied to the MC samples. A correction handling the number of proton-proton collisions seen in each event is added to all MC samples due to unknown ATLAS run conditions for the full data set. QCD and Electroweak K-factor corrections are applied as a function of invariant mass to the signal samples and SM DY samples.

A.2 Electron Identification and Selection

The selection of electron candidates for the CI analysis can be split in three main parts, selection of a good event, selection of a set of good electrons and selection of a good dielectron pair.

Event Selection

- Each event is required to contain at least one reconstructed primary vertex with more than 2 charged

tracks traceable to it.

- Event is required have passed the chosen unprescaled electron trigger (EF_g20_loose).

Electron Selection

- Each electron is required to have a transverse momentum (p_T) greater than 25 GeV.

- Electron $|\eta| < 2.47$ and not lie within the detector crack region $1.37 \leq |\eta| \leq 1.52$ due to a decreased energy resolution.

- Electrons are required to pass identification criteria on the transverse shower shape, the longitudinal leakage into the hadronic calorimeter, and the association to an inner detector track, defined together as a medium electron identification.

- If expected electron is required to have signal in the inter most level of the tracking detector (B-layer). Used to suppress background from photon conversions.

Dielectron Selection

- Selection of two highest p_T electrons left in event.

- Isolation (A cone around the candidate in the calorimeter is required to have $< 7 GeV$ deposited in it) of the highest p_T electron in the event is required to suppress QCD jet background.

- Dielectron invariant mass (m$_{ee}$) is required to be greater than or equal 70 GeV.

These remaining candidates are then the results. The signal region is defined as $m_{ee} > 150 GeV$ while the $70 \leq m_{ee} \leq 110 GeV$ region is the control region.

m_{ee} [GeV]	70-110	110-200	200-400
DY	1231053.7 ± 1109.5	26756.7 ± 163.6	2964.0 ± 54.4
$t\bar{t}$	879.6 ± 29.7	1008.8 ± 31.8	315.8 ± 17.8
Dibosons	1827.1 ± 42.7	415.4 ± 20.4	146.6 ± 12.1
QCD + W+jets	2885.7 ± 53.7	1892.0 ± 43.5	510.5 ± 22.6
Total	1236646.0 ± 1112.0	30072.9 ± 173.4	3936.9 ± 62.7
Data	1236646	29816	4026

400-800	800-1200	1200-3000
266.0 ± 16.3	12.2 ± 3.5	1.5 ± 1.2
20.5 ± 4.5	0.3 ± 0.6	0.0 ± 0.2
16.5 ± 4.1	0.9 ± 0.9	0.1 ± 0.3
49.5 ± 7.0	2.0 ± 1.4	0.3 ± 0.5
352.4 ± 18.8	15.4 ± 3.9	1.9 ± 1.4
358	17	3

Table A.2: Table of data yeild compared to background MC scaled to luminosity of data. Errors shown are statistical only.

A.2.1 Data and Background Comparison

Table A.2 shows the number of data events remaining after selection of dielectron candidates compared to all sources of MC background after the same candidate selection and scaled to the data luminosity. The simulated MC samples also undergo a scaling factor to scale within the Z boson peak in the control region. As can be seen the background prediction matches very closely to data within the statistical errors shown.

The QCD background shown is not predicted using MC simulation but instead by via a reverse identification criteria on data [**?**]. This reducible background is due to QCD multijet production where jets are mis-measured as electron candidates.

Figure A.1: p_T distribution of the leading (left) and subleading (right) electrons showing data, MC background and example CI signal samples compared to data.

Figure A.2: η distribution of the leading (left) and subleading (right) electrons showing data, MC background compared to data.

Control plots were produced to display that the distributions were behaving as predicted such as the p_T (Fig. A.1) and the η (Fig. A.2) distributions.

m_{ee} [GeV]	110-200	200-400	400-800	800-1200	1200-3000
$\Lambda^- = 3$ TeV	18790.8 ± 137.1	5022.4 ± 70.9	2766.3 ± 52.6	1089.2 ± 33.0	673.3 ± 25.9
$\Lambda^- = 4$ TeV	18212.5 ± 135.0	3707.1 ± 60.9	1102.5 ± 33.2	356.9 ± 18.9	214.3 ± 14.6
$\Lambda^- = 5$ TeV	17821.5 ± 133.5	3310.5 ± 57.5	653.1 ± 25.6	160.6 ± 12.7	97.7 ± 9.9
$\Lambda^- = 7$ TeV	17711.1 ± 133.1	3018.8 ± 54.9	385.0 ± 19.6	56.1 ± 7.5	26.5 ± 5.1
$\Lambda^- = 12$ TeV	17693.4 ± 133.0	2992.7 ± 54.7	296.5 ± 17.2	20.4 ± 4.5	5.6 ± 2.4
$\Lambda^+ = 3$ TeV	18106.6 ± 134.6	4063.8 ± 63.7	2103.3 ± 45.9	918.1 ± 30.3	621.4 ± 24.9
$\Lambda^+ = 4$ TeV	17958.1 ± 134.0	3178.6 ± 56.4	765.6 ± 27.7	288.0 ± 17.0	194.9 ± 14.0
$\Lambda^+ = 5$ TeV	18026.6 ± 134.3	2895.6 ± 53.8	432.1 ± 20.8	111.4 ± 10.6	78.8 ± 8.9
$\Lambda^+ = 7$ TeV	17926.4 ± 133.9	2857.5 ± 53.5	278.2 ± 16.7	34.3 ± 5.9	19.1 ± 4.4

Table A.3: Table of CI signal yields for 4.9 fb^{-1}.

m_{ee} [GeV]	≥ 1300
$M_S = 1500$ GeV (GRW)	94.8 ± 9.7
$M_S = 2000$ GeV (GRW)	42.7 ± 6.5
$M_S = 2500$ GeV (GRW)	11.3 ± 3.4
$M_S = 3000$ GeV (GRW)	3.2 ± 1.8

Table A.4: Table of ADD analysis region yields for 4.9 fb^{-1}.

A.2.2 New Physics Signal Expectation

Tables A.3 and A.4 show the yeild from the CI and ADD MC signals used after scaling to data luminosity. The ADD yield is only shown in a single bin above 1300 GeV as the ADD statistical analysis uses only a one bin approach to set a limit of a general increase over SM background. Table A.5 shows the same one bin approach to the data MC comparison table.

m_{ee} [GeV]	≥ 1300
DY	1.1 ± 1.1
$t\bar{t}$	0.0 ± 0.1
Dibosons	0.1 ± 0.3
QCD + W+jets	0.2 ± 0.4
Total	1.4 ± 1.2
Data	2.0

Table A.5: Table of data and MC yields for ADD analysis region.

Figure A.3: Dielectron invariant mass distribution for data and Monte Carlo simulation. Lines show expected distributions for the pressence of Contact Interactions.

Figures A.3 (A.5) show the dielectron invariant mass distribution comparing data to background MC while showing the effect CI (ADD) would have on this spectrum. Figures A.4 (A.6) then show the same spectrum but with an intergrated invariant mass distribution instead which indicates better general increases in the dielectron spectrum.

A.3 Statistical Analysis

If no evidence for new physics is found then a Bayesian counting method can be used to set a limit on the scale (Λ) of new physics. A comparison between the observed events yields and expected yield for a range of different CI benchmarks is done using

$$\mu = n_{DY+CI}(\theta, \bar{\nu}) + n_{non-DYbg}(\bar{\nu}), \tag{A.1}$$

where μ is the number of expected events in each mass bin and n_{DY+CI} and n_{non-DY} are the number of events predicted by a particular benchmark signal sample and the number of predicted non DY background events respectively. ν is a set of Gaussian nuisance parameters that account for systematic uncertainties in the analysis while θ corresponds to the energy scale Λ. The likelihood function for observing a set of $\bar{n}$ events in N mass bins is therefore given by:

$$\mathcal{L}(\bar{n} \mid \theta, \bar{\nu}) = \prod_{k=1}^{N} \frac{\mu_k^{n_k} e^{-\mu_k}}{n_k!} \tag{A.2}$$

as a product of Poission probabilities for each mass bin k. Using Bayes' theorem this gives posterior probability

$$\mathcal{P}(\theta \mid \bar{n}) = \frac{1}{\mathcal{Z}} \mathcal{L}_M(\bar{n} \mid \theta) P(\theta) \tag{A.3}$$

where Z is a normalisation constant and $\mathcal{L}_{\mathcal{M}}$ is the marginalised likelihood after all nuisance parameters have been integrated out. A prior probability for $P(\theta)$ is chosen to be flat in $1/\Lambda^2$, motivated by the CI differential cross-section (Eq. 1.2).

A 95% confidence level (CL) limit is found by finding Λ_{lim} that satisfies $\int_0^{\theta_{lim}} P(\theta \mid \bar{n}) d\theta = 0.95$ with $\theta = 1/\Lambda^2$. For this analysis the Bayesian Analysis Toolkit (BAT) [?] was used to do this calculation.

A.4 Results

Results for the 2011 analysis are currently undergoing approval from ATLAS. Therefore the results presented here are provisional limits.

Channel	ee	$\mu\mu$	ee+$\mu\mu$
Expected CI constructive	13.73 TeV	12.24 TeV	15.10 TeV
Expected CI destructive	10.41 TeV	10.23 TeV	11.42 TeV
Observed CI constructive	11.60 TeV	12.07 TeV	12.70 TeV
Observed CI destructive	8.76 TeV	9.17 TeV	9.63 TeV
Expected ADD	2.84 TeV	2.71 TeV	2.94 TeV
Observed ADD	2.71 TeV	2.72 TeV	2.94 TeV

Table A.6: Table of 95% confidence level limits found in the CI and ADD analyses

An expected limit was obtained for the data set by generating 1000 Standard Model like pseudoexperiments. The Baysian limit setting method is then applied to each of these 1000 pseudoexperiments to get a distribution of 95% CL limits. The median of these distributions is then taken as the expected limit for each channel. These results can be seen in table A.6 in which the expected limits are generally found to be higher because they dictate a scenario where no deviation from the Standard model background is found. In reality slight statistical fluctuations are found in the data set and so limits seen are lower than expected.

These limits are found to be higher than the limits in the last iteration of this analysis seen in the paper provided and the first limits set by ATLAS on ADD.

Figure A.4: Dielectron intergrated invariant mass distribution for data and total background Monte Carlo simulation. Lines show expected distributions for the presence of Contact Interactions.

Figure A.5: Dielectron invariant mass distribution for data and Monte Carlo simulation. Lines show expected distributions for the pressence of ADD.

Figure A.6: Dielectron intergrated invariant mass distribution for data and total backpacking Monte Carlo simulation. Lines show expected distributions for the presence of ADD.

Bibliography