Stony Brook University

A Search for First Generation Leptoquarks at the ATLAS detector

A Dissertation Presented

by

Regina Marie Caputo

to

The Graduate School

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Physics

Stony Brook University

August 2011

Stony Brook University

The Graduate School

Regina Marie Caputo

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of this dissertation.

John D. Hobbs – Dissertation Advisor
Professor, Department of Physics and Astronomy

Maria Victoria Fernandez-Serra – Chairperson of Defense
Assistant Professor, Department of Physics and Astronomy

Patrick Meade
Assistant Professor, Department of Physics and Astronomy

Hong Ma
Physicist, Physics Department
Brookhaven National Laboratory

This dissertation is accepted by the Graduate School.

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

A Search for First Generation Leptoquarks at the ATLAS detector

by

Regina Marie Caputo

Doctor of Philosophy

in

Physics

Stony Brook University

2011

Similarities between quarks and leptons, which are elementary particles, suggest an additional symmetry or communication between the two families. Leptoquarks are hypothetical particles that carry both lepton and baryon number and would represent this additional symmetry. They are proposed to exist in several extensions to the Standard Model such as Grand Unification Theories (GUTs) and technicolor models.

This work reports on the search for first generation scalar leptoquarks at the ATLAS detector using an integrated luminosity of 35 pb^{-1} collected during the 2010 LHC running. Leptoquarks are produced in pairs and each leptoquark decays into a lepton/quark pair. One resulting event topology is two high energy jets, one high energy electron and missing transverse energy arising from a neutrino. The background, predominantly from associated production of vector bosons with jets and top quarks, is estimated using Standard Model simulated data, normalized and checked against observations in control regions. Multijet (QCD) background is estimated using data driven methods, primarily the Matrix Method for shape determination and the Fitting Method for normalization. The number of events observed is in good agreement with background predictions.

First generation leptoquarks with a mass less than 319 GeV at excluded at a 95% CL for the branching fraction, β, of a leptoquark to an electron and quark of 0.5. Weaker limits are derrived for other branching fraction values.

To my parents.

Contents

List of Figures

List of Tables

List of Abbreviations

ATLAS - Just ATLAS, it's no longer an abbreviation for anything!

BCM - Beam Conditions Monitor

CERN - European Organization for Nuclear Research

CL - Confidence Level

CSC - Cathode Strip Chambers

E_T - The energy transverse to the beam line of a particle or system of particles

$\not{E}_T$ - The missing transverse energy defined by the negative sum of E_T

EWK - ElectroWeaK. The theory used to describe the combination of the electromagnetic force and the weak force

FSR - Final State Radiation

gg - gluon-gluon

IP - Interaction Point

ISR - Initial State Radiation

JER - Jet Energy Resolution

JES - Jet Energy Scale

LAr - Liquid Argon (usually referring to the Liquid Argon Calorimeters)

LHC - Large Hadron Collider

LLR - Log-Likelihood Ratio

LO - Leading Order (refers to the calculation of a cross section)

LQ - Leptoquark

LUCID - LUminosity measurement using a Cherenkov Integrating Detector

MBTS - Minimum Bias Trigger Scintillators

MC - Monte Carlo

MDT - Monitored Drift Tubes

NLO - Next-to-Leading Order (refers to the calculation of a cross section)

pp - proton-proton (usually referring to the type of hadrons in the collisions)

p_T - The momentum transverse to the beam line of a particle or system of particles

QCD - Quantum Chromo-Dynamics. The theory which is used to describe the strong force

QED - Quantum Electro-Dynamics. The theory which is used to describe the electromagnetic force

$q\bar{q}$ - quark-antiquark (usually referring to the type of fermions involved in an interaction)

RPC - Resistive Plate Chambers

SCT - SemiConductor Tracker

SM - Standard Model

TGC - Thin Gap Chambers

TRT - Transition Radiation Tracker

Acknowledgements

It's not possible to acknowledge everyone that helped me along the path while writing this, but this is my chance to be verbose. First I would like to thank my advisor, John, who is an excellent mentor. His guidance made working in a 3000 person collaboration seem manageable. He is the kind of scientist I aspire to be. I have the utmost respect for him, although maybe not his cheesy jokes (just kidding!). I'd also like to thank the Stony Brook/BNL leptoquark team: Carolina, Thomas, Burton, John, and David - you guys make an amazing team that I was lucky to be a part of. I must also thank the ATLAS collaboration - without whose hard work over the past 20 years this would not have been possible.

I'd also like to thank to my parents. They supported all of my idiosyncrasies, starting at a very early nerdy age. They always pushed me to do my best and take pride in whatever I did. I also have to thank my 4th grade teacher, Mrs. Hawkins, who told me that I should study science because I was curious and smart. She clearly knew how to speak to me.

Although high school was more about what I did outside of class than in it, one of the exceptions was chemistry from Mr. Mead. He was the reason I decided to become a scientist. It wasn't until later that I discovered what was most interesting about chemistry was actually physics, and here I am today.

Once in college, I found I had a real passion for physics. I was fortunate to have an excellent undergraduate education that completely prepared me for graduate school - with the help of my advisor, Dr. Collins. Mines was a bastion of unpretentious learning. There I met friends who will always be near and dear to my heart.

Once I started my graduate school career, I was again lucky to find amazing people not only at Stony Brook but also at CERN. Those homework sets seemed more bearable in the company of my fellow classmates. Looking back, working late into the night and on weekends didn't seem all that bad with the company I kept. Thanks to Claire who would always be up for tennis while chatting. She gave me advice that made 2^{nd} year enjoyable. At CERN, I was lucky to have an amazing friend and mentor, Stephanie, who also showed me how to survive in physics and maintain a balance in the rest of my life.

I'd also like to thank Sarah and Jon. During one of the most uncertain times of my life, I don't know what I would have done without their support. I'm honored to consider myself their friend. Thank you also goes out to Matt and Abbey, they are the best friends a person could ask for and will never be replaced.

Somewhere along the way I met Josh. We started out as friends, suffering through homework together, and after four years of wandering became much more. I can't believe we made it through. The love, affection, and respect I have for him will never be surpassed. To all these people, I owe a debt of gratitude, and definitely a thesis.

Chapter 1

Introduction

Particle physics is the study of the laws which govern the most fundamental and elementary constituents of matter. These laws are understood in terms of a Quantum Field Theory (QFT) called the Standard Model (SM) in which the particles observed in nature are represented by fields. This introduction gives an overview of the SM and motivation for the search described in this thesis.

The analysis proceeds in the following manner. The search described in this thesis was performed at ATLAS, a multi-purpose detector located at the Large Hadron Collider (LHC) particle-accelerator at the European Organization for Nuclear Research (CERN) during the first year of running. Chapter 2 of this thesis describes the ATLAS detector which makes this scientific endeavor possible. A group of thousands of physicists and engineers are tasked with the commissioning and operation of this massive endeavor. The start of the LHC physics program, despite some major setbacks, is a testament to the hard work of these people who made the newly built apparatus operate and to the quality of the science which it will produce.

The understanding of physics processes observed with ATLAS, however, requires more than understanding the detector. The interpretation of the detector signals requires sophisticated modeling of physical processes. A method of Monte Carlo simulations used to generate events and model the detector is described in Chapter 3.

Prior to performing the analysis, the collected data is analyzed and reconstructed as described in Chapter 4. Physical objects produced in the initial collision are reconstructed by combining the information from several sub-detector components to build an understanding of the particles arising from a collision.

Once the physical objects are reconstructed, the results can be interpreted as known physical processes and when deviations from these processes are found as new physics. The main component of this thesis describes a search for new physics. To carry out the search a preselected sample is defined which has a similar topology to the new physics final states (such as leptoquarks). This preselected sample is dominated by the backgrounds to the analysis and has a very low signal-to-background ratio. The background contributions are

estimated using simulated events. From the preselected sample, three control regions are defined to validate the background modeling in signal poor regions. Finally, an optimization procedure is applied to simulated events in the preselected sample to determine the final selection requirements used to obtain sufficient a signal-to-background ratio. The leptoquark selection is discussed in Chapter 5.

Systematic uncertainties are biases in measurements which result in the mean of many separate measurements differing from the actual value of the measured attribute. They can be from several sources such as an imperfect calibration of the detector or differences in the signal or background modeling. These systematic uncertainties are described in Chapter 6.

Having properly collected and modeled the data, the method used to assess the sensitivity of the results is described in Chapter 7. The statistical methods of interpreting the data are discussed. The final results and conclusions are discussed in Chapters 8 and 9 respectively.

1.1 An Introduction to the Standard Model

The Standard Model (SM) is a theory that describes the electromagnetic, weak and strong interactions which mediate the dynamics of elementary particles. It consists of three generations of spin-1/2 particles, known as fermions, and twelve spin-1 gauge bosons. A brief conceptual explanation will be introduced in this section, and a more detailed discussion of the mathematical formalism of the SM will be discussed in Section 1.2.

1.1.1 Quarks and Leptons

Experimental evidence has shown that ordinary matter is made up of a set of fundamental particles called *fermions*. They are referred to as fundamental because these particles are believed to have no further sub-structure. All fermions have half-integer angular momentum, or spin, and therefore obey Fermi-Dirac statistics. This implies that they have unique quantum states, so if two fermions have the same quantum numbers they may not occupy the same state. The fermions fall into three generations and two categories: quarks and leptons, which are shown in Table 1.1. Every fermion also has an associated anti-particle with opposite quantum numbers.

Leptons are fermions which are separated into two types: charged and neutral. The most familiar charged lepton is also the lightest: the electron, e. The two others are the muon, μ, and the tau, τ, which are heavier versions of the electron. The neutral leptons are called neutrinos, each denoted with the symbol ν. The different flavors of neutrinos correspond to their charged lepton counter-parts indicated by a subscript. For example, the first generation neutrino partner to the electron, is the electron neutrino, ν_e.

Quarks, unlike leptons, exist in bound states in nature. Due to the properties of the strong force (see Section 1.2.2), they must form pairs of quarks (mesons) or triplets (baryons). The most familiar combinations of quarks are protons and neutrons, both of which are baryons -

generically called hadrons. Protons are made up of a combination of first generation quarks: two up, (u), type quarks, and one down, (d), type yielding a charge of +1. Neutrons are made of the same first generation quarks, in different proportions: one u and two d quarks. This yields a particle with no net charge. The next two generations of up-type quarks are the charm, c, and the top, t - which is the heaviest quark. The analogous down-type quarks are the strange, s, and the bottom, b. The quantum numbers for the quarks and leptons can be seen in Table 1.1.

The fact that fermions of the same charge are held together in the nucleus shows the amount of energy provided by the strong force. This also means that the mass of hadrons is not solely the sum of the quark constituents, but also gluons - the carriers of the strong force.

Fermion	Charge	1^{st} Generation		2^{nd} Generation		3^{rd} Generation	
			mass (MeV/c^2)		mass (MeV/c^2)		mass (MeV/c^2)
Quarks	+2/3	u	1.5 to 3.0	c	1.25 GeV/c^2	t	174.2 GeV/c^2
	-1/3	d	3 to 7	s	70 to 120	b	4.20 GeV/c^2
Leptons	-1	e	0.51	μ	105.6	τ	1777
	0	ν_e	<2 eV/c^2	ν_μ	<0.19	ν_τ	<18.2

Table 1.1: Quarks and Leptons. Three generations of fermions (anti-particles not shown). Each family has identical gauge interactions, and differ only by mass and flavor quantum number [1]. The charge shown is relative to the negative of the electron charge.

1.1.2 Force Mediators

The SM is not only comprised of particles, but also describes their interactions. The particles interact via the exchange of integer-spin bosons. Since bosons have an integer-spin, they obey Bose-Einstein statistics, which implies that multiple bosons of the same quantum numbers can occupy the same state. The bosons which mediate the forces are summarized in Table 1.2. An important omission in this table is gravity. One of the shortcomings of the SM is that it not compatible with the gravitational force, and this will be discussed in Section 1.3.

The electromagnetic (EM) force is responsible for essentially all phenomena at scales between atoms and solar systems. It allows bound states of electrons with nuclei and for intermolecular forces in states of matter. These interactions are mediated by the exchange of photons, γ, which couple to all particles with electric charge. It is the second force by strength, but the best understood. Since photons are massless, the range of the EM force is infinite. It obeys the inverse square law, which means as two charged particles are separated, the force diminishes as the inverse of the separation squared.

Boson	Force	Mass (GeV/c^2)	Rel. Strength	Range
gluon (g)	Strong	0	1	10^{-15} m
photon (γ)	Electromagnetic	0	10^{-3}	∞
Z^0	Weak	91.1876	10^{-6}	10^{-18} m
$W^{\pm}$	Weak	80.403	10^{-6}	10^{-18} m
Higgs (H)	-	>114.4	-	-

Table 1.2: Force Mediators. These are the forces that are described by the Standard Model. They are mediated by bosons listed in the first column along with their mass and relative strength compared to the strong force, and their range [2]. From here, $\hbar = c = 1$.

The weak force governs nuclear β-decay, in which a neutron decays to a proton, electron, and anti-neutrino ($n \rightarrow p + e^- + \bar{\nu}^e$). Conversely, it allows the transmutation of protons to neutrons ($p \rightarrow n + e^+ + \nu^e$), which allows deuterium to form, and is the process which causes the sun to burn. These interactions are governed by three vector bosons: $W^{\pm}$ and the Z^0. A unique property of the weak force is that it allows the flavor between quarks and the flavor between leptons to change. The weak interaction is the only force in which all quarks and leptons participate.

The strong force is responsible for binding the nucleus of atoms together. This inter-quark force is mediated by a massless particle called a gluon, *g*. It binds protons and neutrons together in an atomic nucleus and quarks together in a proton. Gluons act on quarks, anti-quarks and other gluons and are the mediators of the strong force. Because the strength of the strong force grows with increasing distance quarks can not exist alone, an effect which is referred to as confinement.

1.2 Formalism of the Standard Model

The Standard Model is a Quantum Field Theory in which the particles are representations of the fields that have distinct values for a position in space and time. The SM uses Lagrangians, which are mathematical constructs that describe the evolution of a physical system. The power of using Lagrangians is that they can be constructed to be gauge invariant, which allows one way to describe a number of local continuous symmetries that are observed in nature. This means that the Lagrangian is invariant under a group of transformations and that the field is invariant under the transformation: $\psi(x) \rightarrow e^{i\alpha(x)}\psi(x)$ [3].

The true power of the SM lies in its gauge invariance, which lead to symmetries which in turn leads to conservation laws. Gauge fields are included in the Lagrangian to ensure its invariance under the local group transformations (called gauge invariance). Since the SM is quantized, the quanta of the gauge fields are called gauge bosons. The Standard Model's symmetry group ($SU(3) \times SU(2) \times U(1)$) gives rise to 12 bosons: photon (γ), $W^{\pm}$, Z^0, and

eight gluons (g).

1.2.1 Electromagnetic Force and Quantum Electrodynamics

The theoretical model for the electromagnetic (EM) force is Quantum Electrodynamics (QED). It is described by a $U(1)$ symmetry group. To describe the EM force, start with a free EM Lagrangian described in Equation 1.1.

$$\mathcal{L}_0 = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{1.1}$$

The Lagrangian must be invariant under local gauge transformations. To do this, consider first a modified derivative, D_μ, which transforms covariantly under phase transformations, and a vector field, A_μ as defined in Equation 1.2.

$$D_\mu \equiv \partial_\mu - ieA_\mu \tag{1.2}$$

A_μ is the gauge field corresponding to the photon and transforms as defined in Equation 1.3. If this field corresponds to a physical photon field, there must be a term which corresponds to its kinetic energy. Since it too must be invariant when transformed as the gauge field, it can only involve the gauge invariant field strength tensor, given by Equation 1.4.

$$A_\mu \rightarrow A_\mu + \frac{1}{e}\partial_\mu \alpha_\mu \tag{1.3}$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{1.4}$$

Substituting these into $\mathcal{L}_0$ gives the QED Lagrangian, given in Equation 1.5. It should be noted that an addition of a mass term would violate gauge invariance, therefore, the photon must be massless.

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu A_\mu \psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.5}$$

1.2.2 Strong Force and Quantum Chromodynamics

The aptly named strongest of the fundamental forces is the strong force. The theoretical model for the strong force is quantum chromodynamics (QCD), an $SU(3)$ gauge theory of color charged fermions. This symmetry requires an octet of gluons and three different color charges: red (r), blue (b), and green (g). The color combinations of the gluon octet is listed below:

- $(r\bar{b} + b\bar{r})/\sqrt{2}$, $-i(r\bar{b} - b\bar{r})/\sqrt{2}$
- $(r\bar{g} + g\bar{r})/\sqrt{2}$, $-i(r\bar{g} - g\bar{r})/\sqrt{2}$

- $(b\bar{g} + g\bar{b})/\sqrt{2}$, $-i(b\bar{g} - g\bar{b})/\sqrt{2}$
- $(r\bar{r} - b\bar{b})/\sqrt{2}$, $(r\bar{r} + b\bar{b} - 2g\bar{g})/\sqrt{6}$

QCD has the two unique properties of confinement and asymptotic freedom. Confinement is a property which means that the force between quarks does not decrease as they are separated. Instead, it would take an infinite amount of energy to completely separate two quarks. As a result, quarks when separated (in a particle collision for example) create quark/anti-quark pairs to reduce the energy of the system. This is what causes the creation of quark and gluon showers (or hadronization), which form jets. Asymptotic freedom is the property that means at higher energies the interactions between quarks and gluons become weaker. Conversely, at small energy scales (or large length scales) the interactions become stronger thus preventing the un-bonding of protons.

The free Lagrangian for the $SU(3)$ group for the quark color fields is given by Equation 1.6, where $j = 1, 2, 3$ denotes the three color fields.

$$\mathcal{L}_0 = \bar{q}_j(i\gamma^\mu\partial_\mu - m)q_j \tag{1.6}$$

To impose $SU(3)$ local gauge invariance on $\mathcal{L}_0$, apply the same covariant derivative as shown in Equation 1.2 but instead of the single gauge field A_μ and charge, $-e$, use G^a_μ and charge gT_a, where G^a_μ is a 3×3 matrix formed by the eight color gauge fields corresponding to the eight gluons and transforms as shown in Equation 1.7

$$G^a_\mu \rightarrow G^a_\mu - \frac{1}{g}\partial_\mu\alpha_a - f_{a,b,c}\alpha_b G^c_\mu \tag{1.7}$$

where $f_{a,b,c}$ are structure constants. Replacing ∂_μ by D_μ yields the final gauge invariant QCD Lagrangian in Equation 1.8, where the gluon field strength tensor, $G^a_{\mu\nu}$, is defined in Equation 1.9.

$$\mathcal{L}_{QCD} = \bar{q}_j(i\gamma^\mu\partial_\mu - m)q_j - g(\bar{q}\gamma^\mu T_a q)G^a_\mu - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a \tag{1.8}$$

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g f_{a,b,c} G^b_\mu G^c_\nu \tag{1.9}$$

Expanding these two equations yields terms that describe the free propagation of quarks (terms containing $q\bar{q}$), of gluons (G^2), and of the quark-gluon interaction ($gq\bar{q}G$). Terms with a cubic and quartic powers of G imply three and four gluon vertices which give gluon-gluon interactions.

1.2.3 Weak Force and Electroweak Symmetry

The weak force changes the flavor of quarks and leptons. There are two types of weak processes: charged (mediated by the $W^\pm$ bosons) and neutral (mediated by Z^0 boson). Since

all three of these bosons are massive, the strength of the weak force decreases rapidly with distance - and is hence the weakest of the fundamental forces (except gravity).

Although the EM and weak forces appear very different, unification of these forces occurs on the order of 100 GeV (unification energy) and is accomplished under an $SU(2) \times U(1)$ gauge group. The gauge bosons are the three W bosons of weak isospin from $SU(2)$ (W^1, W^2, and W^3), and the B^0 boson of weak hypercharge from $U(1)$ all of which are massless [4, 5].

As with QCD for strong interactions, there are a number of constraints which help build a formalism for EM and weak interactions. It has been observed that only left-handed fermions (right-handed anti-fermions) participate in flavor-changing weak interactions and that leptons appear in doublets containing a charged and neutral component. However, there is no right-handed contribution from neutrinos. These constraints lead us to the formalism of electroweak symmetry.

As with QCD and EM, the electroweak Lagrangian can be formed by applying a modified derivative to the free Lagrangian. This forms a Lagrangian which can be broken into two parts which is shown in Equation 1.10.

$$\mathcal{L}_{EWK} = \mathcal{L}_{gauge} + \mathcal{L}_{fermion} \tag{1.10}$$

The term, $\mathcal{L}_{gauge}$, describes the gauge interactions between the W and B particles. It takes a similar form as the kinetic energy term for the gauge field in the QED Lagrangian.

$$\mathcal{L}_{gauge} = -\frac{1}{4} W^i_{\mu\nu} W^{\mu\nu i} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \tag{1.11}$$

where,

$$\begin{aligned} W^i_{\mu\nu} &= \partial_\nu W^i_\mu - \partial_\mu W^i_\nu + g\epsilon^{ijk} W^j_\mu W^k_\nu \\ B_{\mu\nu} &= \partial_\nu B_\mu - \partial_\mu B_\nu \end{aligned}$$

The term, $\mathcal{L}_{fermion}$, has an analogous part in the QED Lagrangian and $i = 1, 2, 3$. It represents the the interactions of the gauge bosons with the fermions through the modified covariant derivatives. This is expanded in Equation 1.12

$$\mathcal{L}_{fermion} = \bar{\psi}_L i\gamma^\mu (D_\mu) \psi_L + \bar{\psi}_R i\gamma^\mu (D_{\mu,R}) \psi_R \tag{1.12}$$

where ψ_L term denotes the left-handed isospin doublet of the fermion $\begin{pmatrix} \nu_i \\ \ell_i \end{pmatrix}_L$ and since neutrinos do not appear to have a right-handed component, only the isospin singlet exists $(\ell_i)_R$. It should also be noted that the right-handed covariant derivative $(D_{\mu,R})$ has one fewer term. This is due to the right-handed fermions not coupling to isospin. The covariant derivatives are defined as:

$$
\begin{aligned}
D_\mu &= \partial_\mu + i\frac{g'}{2}B_\mu Y + i\frac{g}{2}\tau \cdot \mathbf{W}_\mu \\
D_{\mu,R} &= \partial_\mu + i\frac{g'}{2}B_\mu Y
\end{aligned}
$$

τ denotes the Pauli matrices, and g and g' are coupling constants [5]. The system defined by the two Lagrangians, $\mathcal{L}_{gauge}$ and $\mathcal{L}_{fermion}$, is consistent with a gauge theory of the weak isospin and weak hypercharge. However, nature is not so simple. The requirement of local gauge invariance means that the fields must be massless which conflicts with the experimental observation of massive W and Z bosons.

1.2.4 Spontaneous Symmetry Breaking

A solution to the massive gauge bosons is to let the masses be generated by a breaking of the vacuum state symmetry but retain the gauge symmetry of the full Lagrangian. This phenomenon is known as spontaneous symmetry breaking. In the case of electroweak theory, electroweak symmetry breaking is also referred to as the Higgs mechanism. To facilitate this type of symmetry breaking an additional Lagrangian must be added to $\mathcal{L}_{EWK}$ (the term $\mathcal{L}_{higgs}$) and is demonstrated in Equation 1.13.

$$\mathcal{L}_{EWK} = \mathcal{L}_{gauge} + \mathcal{L}_{fermion} + \mathcal{L}_{higgs} \tag{1.13}$$

Occam's Razor leads to the solution of adding a complex doublet of Higgs fields: $\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$ [3]. The contribution of the Higgs field to the Lagrangian is given by $\mathcal{L}_{higgs}$. It is given by Equation 1.14

$$\mathcal{L}_{higgs} = (D^\mu \Phi)^\dagger (D_\mu \Phi) - V(\Phi) \tag{1.14}$$

where, $V(\Phi)$ denotes the interaction of the Higgs field with itself and is represented by:

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda(\Phi^\dagger \Phi)^2$$

The only additional requirements are that $\mu^2, \lambda > 0$. These choices are mainly motivated by the requirement that while the electroweak symmetry is broken, the symmetry in electromagnetism remains. The constraint on μ^2 yields the following implications on the minimum:

$$\begin{aligned} \langle 0|\Phi|0\rangle &= \begin{pmatrix} 0 \\ \phi_0 \end{pmatrix} \\ \phi_0 &= \sqrt{\frac{-\mu^2}{2\lambda}} \\ V(\phi_0) &= \frac{\lambda}{4}v^4 \end{aligned}$$

where $v \equiv \sqrt{\frac{\mu^2}{\lambda}}$, is the vacuum expectation value and is found to be 246 GeV.

In this type of symmetry breaking, the Lagrangian of the system is invariant under a transformation but the vacuum state is not. This will lead to an assignment of a non-zero vacuum expectation value to the fields of the theory which in turn can be interpreted as a non-zero mass. This causes the W^3 and B^0 bosons ($W^{1,2}$ bosons) to mix to form the observed eigenstates of the Z and photon, A (also represented by γ), ($W^{\pm}$) by the weak mixing angle, θ_W. From here, the physical charged fields and two neutral fields are shown in Equation 1.15.

$$\begin{aligned} W_\mu^{\pm} &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp W_\mu^2) \\ Z_\mu &= -sin(\theta_W)B_\mu + cos(\theta_W)W_\mu^3 \\ A_\mu &= sin(\theta_W)W_\mu^3 + cos(\theta_W)B_\mu \end{aligned}$$

The bosons obtain their masses from the Higgs mechanism, and they are found to be:

$$\begin{aligned} M_W^2 &= \frac{1}{4}g^2g'^2 \\ M_Z^2 &= \frac{1}{4}(g^2+g'^2)v^2 \\ M_A^2 &= 0 \end{aligned}$$

This addition of an extra Higgs field to the gauge theory is the simplest implementation of the Higgs mechanism. The spontaneous symmetry breaking of the underlying local symmetry allows interactions of the Higgs field with other fields in the theory, so as to produce mass terms for the gauge bosons. This mechanism must also leave behind elementary scalar (spin-0) particles, known as Higgs bosons. Direct searches for the Higgs boson have been the subject of research for decades at LEP, the Tevatron, and now the LHC, and it has yet to be found.

The undiscovered Higgs boson, as well as other problems with the SM leaves the field

open for other more dynamic theories which explain electroweak symmetry breaking. These alternatives will be discussed in the Section 1.3 as well as the implications on potentially new families of particles.

1.3 Beyond the Standard Model

Despite the tremendous success of the Standard Model, there are a few shortcomings which have been puzzling particle physicists for decades. Loosely speaking, these fall into three categories, which are described below.

- Experimental Observations not Modeled: The SM has no theoretical framework for gravity, dark energy, and neutrino masses, to name a few.
- Theoretical Predictions not Observed: The SM has predicted the existence of the Higgs boson, which was discussed in Section 1.2.4. If the Higgs boson is not found, then EWSB must involve physics beyond the SM.
- Theoretical Oddities: Even if no physics is found beyond the SM, there are some odd features to the framework which have motivated other explanations. Some of these include: (1) the hierarchy problem (large corrections requiring a very finely tuned Higgs boson mass), (2) the large number of numerical parameters - the SM depends on 19 (determined from experiment), (3) the need for three generations of quarks and leptons.

All these challenges to the SM have motivated the search for new physics. The theoretical framework for these solutions can be categorized into two solutions: weakly coupled and strongly coupled.

Weakly coupled solutions use additional symmetries to explain the cancelation of quadratic divergences. The flagship of this type of theory is Supersymmetry (SUSY). SUSY requires that every elementary particle has a "super-partner" which differs by a half unit of spin. In other words, for every boson, there is a fermion partner with the same internal quantum numbers and vice versa. Strongly coupled solutions add a new asymptotically free gauge group to the SM. Technicolor is the most developed of these types of solutions. These solutions have their own difficulties in that they are almost incalculable. Additionally there are attempts to go beyond the electroweak symmetry and unite not only the EM and weak forces, but also the strong force. These theories are generically called Grand Unified Theories (GUTs). Some of the implications of these theories and different predictions will be discussed in the following sections.

1.4 Leptoquark Phenomenology

The remarkable symmetry between quarks and leptons in the standard model suggests that some more fundamental theory may exist which allows interactions between them. Such interactions are mediated by a new type of particle, a leptoquark (LQ). Leptoquarks appear in many extensions to the SM. They have a rich phenomenology, and share several general features. They are color-triplet bosons which carry both lepton and baryon number and can be either scalar (spin-0) or vector (spin-1). Scalar leptoquarks have one undetermined coupling at the $LQ-\ell-q$ vertex given by a Yukawa coupling, λ, whereas the vector version has two undetermined couplings. Leptoquarks have fractional electric charge which can be formed by adding the charges of one lepton and one quark [6, 7].

In this chapter, theoretical sources of leptoquark pair production mechanisms are discussed and the experimental signatures for such particles is presented. The current limits on leptoquark mass from other searches are also presented.

1.4.1 Leptoquarks in BSM Physics

To begin the discussion on leptoquarks, start with an effective Lagrangian which can be defined in the following terms in Equation 1.15 [8].

$$\mathcal{L}_{LQ} = \mathcal{L}^f_{|F|=0} + \mathcal{L}^f_{|F|=2} + \mathcal{L}^g_S + \mathcal{L}^g_V \tag{1.15}$$

F is the fermion number is defined by $F = L + 3B$, where L is the lepton number and B is the baryon number (+1/3 for quarks). S and V are the general forms for scalar and vector leptoquarks respectively. Each Lagrangian can be written in the following terms:

$$\begin{aligned}
\mathcal{L}^f_{|F|=0} &= (h_{2L}\bar{u}_R\ell_L + h_{2R}\bar{q}_L i\tau_2 e_R)S_{1/2} + \tilde{h}_{2L}\bar{d}_R\ell_L\tilde{V}^L_{1/2} \\
&+ (h_{1L}\bar{q}_L\ell_L + h_{2R}\bar{d}_R\gamma^\mu e_R)V_0 \\
&+ \tilde{h}_{1R}\bar{u}_R\gamma^\mu e_R\tilde{V}^R_0 + h_{3L}\bar{q}_L\boldsymbol{\tau}\gamma^\mu\ell_L V^L_1 + h.c.
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}^f_{|F|=2} &= (g_{1L}\bar{q}^c_R i\tau_2\ell_L + g_{1R}\bar{u}^c_R e_R)S_0 \\
&+ \tilde{g}_{R1}\bar{d}^c_R e_R\tilde{S}^R_0 + g_{3L}\bar{q}^c_L i\tau_2\boldsymbol{\tau}\ell_L S^L_1 \\
&+ (g_{2L}\bar{d}^c_R\gamma^\mu\ell_L + g_{2R}\bar{q}^c_L\gamma^\mu e_R)V_{1/2} \\
&+ \tilde{g}_2\bar{u}^c_R\gamma^\mu\ell_L\tilde{V}^L_{1/2} + h.c.
\end{aligned}$$

$$\mathcal{L}^g_S = \sum_{scalars}\left[(D^\mu_{ij}\Phi^j)^\dagger(D^{ik}_\mu\Phi_k) - M^2_S\Phi^{i\dagger}\Phi_i\right]$$

$$\mathcal{L}_V^g = \sum_{vectors} \left\{ \frac{-1}{2} G_{\mu\nu}^{i\dagger} G_i^{\mu\nu} + M_V^2 \Phi_\mu^{i\dagger} \Phi_i^\mu - i g_s \left[(1-\kappa_G) \Phi_\mu^{i\dagger} t_{ij}^a \Phi_\nu^j \mathcal{G}_a^{\mu\nu} + \frac{\lambda_G}{M_V^2} G_{\sigma\mu}^{i\dagger} t_{ij}^a G_\nu^{j\mu} \mathcal{G}_a^{\nu\sigma} \right] \right\}$$

where g_s is the strong coupling constant, t_a are the generators of $SU(3)_c$ and M_S and M_V are masses for scalar and vector leptoquarks respectively. A summary of the possible quantum numbers for leptoquarks is given in Table 1.3. Included in the table are the different possibilities for charge is given by Q (in units of proton charge), the lepton-quark Yukawa coupling given by λ, the decay branching fraction to a charged lepton given by β, and the fermion number given by F.

The parameters κ_G and λ_G are the anomalous couplings and assumed to be real. They are related to the anomalous magnetic moment, μ_V, and electric quadrupole moment, q_V, of the leptoquarks in the color fields as defined by the Equations 1.16.

$$\begin{aligned} \mu_{V,G} &= \frac{g_s}{2M_V}(2 - \kappa_G + \lambda_G) \\ q_{V,G} &= \frac{-g_s}{M_V^2}(1 - \kappa_G - \lambda_G) \end{aligned} \tag{1.16}$$

The field strength tensors, G and $\mathcal{G}$, and covariant derivative, D, are given by:

$$\begin{aligned} G_{\mu\nu}^a &= \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g_s f^{abc} A_{\mu b} A_{\nu c} \\ \mathcal{G}_{\mu\nu}^i &= D_\mu^{ik} \Phi_{\nu k} - D_\nu^{ik} \Phi_{\mu k} \\ D_\mu^{ij} &= \partial_\mu \delta^{ij} - i g_s t_a^{ij} A_\mu^a \end{aligned}$$

The following sections discuss the different many extensions to the SM which include leotpquarks.

Grand Unified Theories

Leptoquarks made their debut in the Pati-Salam $SU(4)$ color symmetry model. In this model, lepton number is treated as a fourth color: the four weak doublets of each generation are arranged as a four doublets of $SU(4)$ [6, 10]. This symmetry is then spontaneously broken such that the gluons remain massless, and the leptoquarks become massive. Leptoquarks in this model violate lepton family number, however, generational mixing is suppressed.

Another GUT that includes leptoquarks is the $SU(5)$ model. In this model, vector leptoquarks have unification scale masses, so they are not accessible by accelerators. Some scalar leptoquarks in this model are able to not only couple to a lepton and a quark, but also two quarks. If the two quark coupling is left unconstrained, then the leptoquark mass

Type	Q	Coupling	β	F
S_0^L	-1/3	$\lambda_L(e_L, u),\ -\lambda_L(\nu_e, d)$	1/2	2
S_0^R	-1/3	$\lambda_R(e_R, u)$	1	2
$\tilde{S}_0^R$	-4/3	$\lambda_R(e_R, d)$	1	2
S_1^L	-4/3	$-\sqrt{2}\lambda_L(e_L, d)$	1	2
	-1/3	$-\lambda_L(e_L, u),\ -\lambda_L(\nu_e, d)$	1/2	2
	+2/3	$\sqrt{2}\lambda_L(\nu_e, u)$	0	2
$V_{1/2}^L$	-4/3	$\lambda_L(e_L, d)$	1	2
	-1/3	$\lambda_L(\nu_e, d)$	0	2
$V_{1/2}^R$	-4/3	$\lambda_R(e_R, d)$	1	2
	-1/3	$\lambda_R(e_R, u)$	1	2
$V_{1/2}^R$	-1/3	$\lambda_L(e_L, u)$	1	2
	+2/3	$\lambda_L(\nu_L, u)$	0	2
$S_{1/2}^L$	-5/3	$\lambda_L(e_L, \bar{u})$	1	0
	-2/3	$\lambda_L(\nu_e, \bar{u})$	0	0
$S_{1/2}^R$	-5/3	$\lambda_R(e_R, \bar{u})$	1	0
	-2/3	$-\lambda_R(e_R, \bar{d})$	1	0
$\tilde{S}_{1/2}^L$	-2/3	$\lambda_L(e_L, \bar{d})$	1	0
	+1/3	$\lambda_L(\nu_e, \bar{d})$	0	0
V_0^L	-2/3	$\lambda_L(e_L, \bar{d}),\ -\lambda_L(\nu_e, \bar{u})$	1/2	0
V_0^R	-2/3	$\lambda_R(e_R, \bar{d})$	1	0
$\tilde{V}^R$	-5/3	$\lambda_R(e_R, \bar{u})$	1	0
V_1^L	-5/3	$\sqrt{2}\lambda_L(e_L, \bar{u})$	1	0
	-2/3	$-\lambda_L(e, \bar{d}),\ \lambda_L(\nu_e, \bar{u})$	1/2	0
	+1/3	$\sqrt{2}\lambda_L(\nu_L, \bar{d})$	0	0

Table 1.3: Leptoquark quantum numbers [9]. Charge is given by Q (in units of proton charge), the lepton-quark Yukawa coupling given by λ, the decay branching fraction to a charged lepton given by β, and the fermion number given by F.

must be large and again, inaccessible to accelerators. However, if the $LQ - q - q'$ coupling is set to zero, scalar $SU(5)$ leptoquarks can be light (on the order of 100 GeV) [11].

Leptoquarks also arise in superstring E_6 models [12, 13]. This opens up the addition of extra $U(1)$ symmetries. The low energy limit may contain leptoquarks as with other GUTs.

Supersymmetry

Supersymmetry also allows for leptoquark-like objects at the TeV mass scale if R-parity is not conserved and Yukawa terms are added to the superpotential. R-parity is denotes a

+1 for particles and -1 for supersymmetric partners. The assumption if R-parity is violated would be that particles and supersymmetric particles do not need to produced in equal amounts. The extra Yukawa term is what violates lepton number conservation and leads to scalar quarks produced singly as leptoquark-like objects [14].

Technicolor

Although the Higgs offers one solution to electroweak symmetry breaking, there are other options that can be modeled after chiral symmetry breaking in QCD, which does provide a mass for the W and Z bosons - only orders of magnitude too small. Strongly interacting theories like technicolor provide a different and dynamic solution to electroweak symmetry breaking. The initial inspiration of the technicolor force was a QCD-like strongly-interacting gauge theory [15].

Technicolor generates masses for W and Z bosons through new gauge interactions. The theory predicts new technicolor fermions, or technifermions, which form lowest bound states called technimesons. Although the initial models of technicolor have been heavily constrained by experiments on flavor changing neutral currents, extensions to technicolor are relatively unconstrained.

One feature of extended technicolor models increasing the group number symmetry allows large numbers of goldstone bosons. For example if the model is $SU(8)$ is broken 63 goldstone bosons will be produced. Three will be eaten by the electroweak gauge bosons. In the remaining 60, there will be "leptoquark mesons" composed of a techniquark and an anti-technilepton mediated by the technicolor force. These bosons will not only have a place in the technicolor sector. There are also analogous mesons to quarks and leptons, which would require a quark/anti-lepton mediator with the same SM quantum numbers.

1.4.2 Production at Hadron Colliders

The pair production of leptoquarks is via gg fusion and $q\bar{q}$ annihilation born (tree-level) diagrams illustrated in Figure 1.1. Because of the nature of pp colliders (as opposed to $p\bar{p}$ colliders like the Tevatron), gg fusion is the dominant process for scalar pair production at the LHC for low leptoquark masses ($< 1.5\ TeV$).

1.4.3 Couplings

The cross section for pair production of leptoquarks depends on gluon-leptoquark coupling terms. Gluon-leptoquark interactions are determined by the gauge symmetry of scalar QCD so that the predictions for pair production of scalar leptoquarks can neglect Yukawa terms [6, 16]. These productions involve only well-understood strong couplings, no undetermined Yukawa couplings. Single leptoquark production involves diagrams which depend

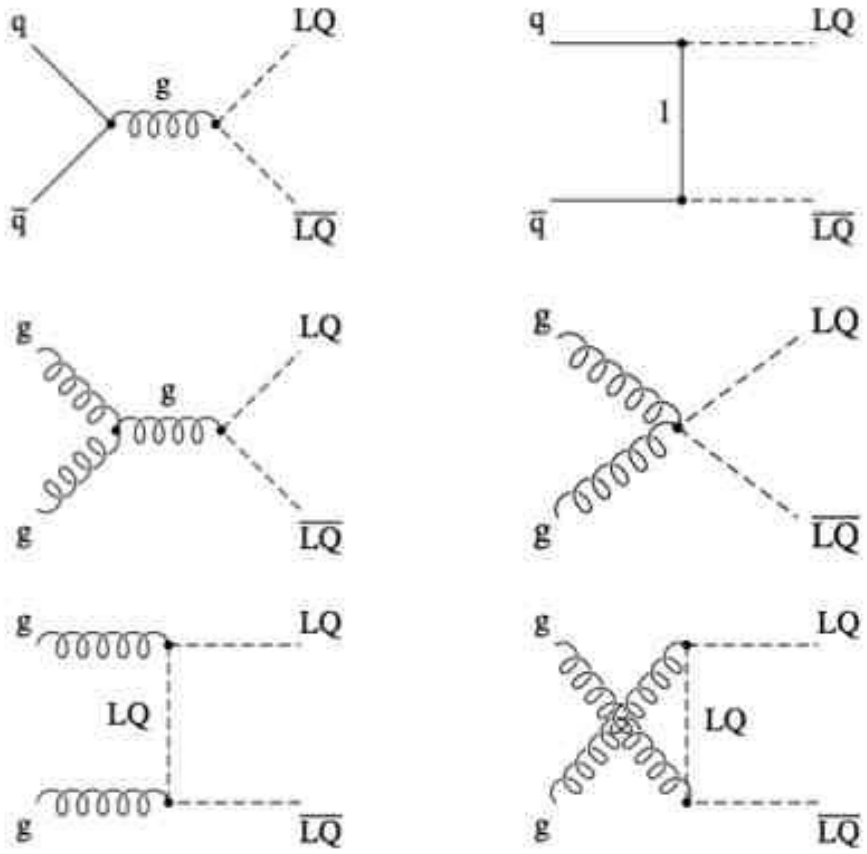


Figure 1.1: Leading Order Diagrams for leptoquark pair production. The dominant production mechanism at the LHC is gg for leptoquark masses < 1.5 TeV.

on the unknown new couplings based on the $LQ - \ell - q$ vertex. Because of this, single leptoquark production has more model dependence and is not considered in this thesis.

The decay rate of leptoquarks also depends on an unknown coupling, but the main impact on searches for pair production is that the coupling be large enough that the decay is well inside the detector volume. For leptoquark masses in the range considered in this thesis, 200 GeV$\leq M_{LQ} \leq$ 350 GeV, couplings greater than $e \times 10^{-6}$ with e the electron charge correspond to decay lengths less than roughly 1 mm. This sets the approximate sensitivity to the unknown coupling strength.

1.4.4 Leading and Next-to-Leading Cross Section

The cross sections for pair production of leptoquarks at the LHC are gluon-gluon fusion and quark-antiquark annihilation can be written as the Equations 1.17 [17].

$$\hat{\sigma}_{LO}[gg] = \frac{\alpha_s^2\pi}{96\hat{s}}\left[\beta(41-31\beta^2) + (18\beta^2-\beta^4-17)log\frac{1+\beta}{1-\beta}\right]$$

$$\hat{\sigma}_{LO}[q\bar{q}] = \frac{\alpha_s^2\pi}{\hat{s}}\frac{2}{27}\beta^3 \tag{1.17}$$

where $\sqrt{\hat{s}}$ is the invariant energy of the sub-process and $\beta = \sqrt{1-4M_{LQ}^2/\hat{s}}$. The renormalization and factorization scale, μ is set to M_{LQ}.

Since gg fusion is the dominant process at the LHC, QCD corrections strongly affect the parton cross sections near the production threshold. These corrections are positive for gluon initial states, therefore to accurately obtain the production cross section, higher order terms must be included. The perturbative expansion of the total parton cross section can be expressed in terms of scaling functions as shown in Equation 1.18.

$$\hat{\sigma}_{ij} = \frac{\alpha_s^2(M_{LQ}^2)}{M_{LQ}^2}\left[f_{ij}^{(B)}(\beta) + 4\pi\alpha_s(M_{LQ}^2)\left\{f_{ij}^{(V+S)}(\beta) + f_{ij}^{(H)}(\beta)\right\}\right] \tag{1.18}$$

where $i,j = g,q$ and the f_{ij} scaling functions can be expressed as follows [17], if $\beta << 1$:

$$f_{gg}^B = \frac{7\pi\beta}{384} \qquad f_{q\bar{q}}^B = \frac{\pi\beta^3}{54}$$

$$f_{gg}^{V+S}/f_{gg}^B = \frac{11}{336\beta} \qquad f_{q\bar{q}}^{V+S}/f_{q\bar{q}}^B = \frac{-1}{48\beta}$$

$$f_{gg}^H/f_{gg}^B = \frac{3}{2\pi^2}log^2(8\beta^2) - \frac{183}{28\pi^2}log(8\beta^2) \qquad f_{q\bar{q}}^H/f_{q\bar{q}}^B = \frac{2}{3\pi^2}log^2(8\beta^2) - \frac{107}{36\pi^2}log(8\beta^2)$$

When the NLO term is included, the stability of the cross section is checked by using two different parameterizations of the parton densities (CTEQ6 and MRST2002). The difference, although increases with leptoquark mass, does not exceed 10%. This means that the inclusion of the NLO cross section calculation stabilizes the theoretical prediction of the production of leptoquarks. The choice of factorization and renormalization scale, μ, near the leptoquark mass gives a conservative lower limit on the production cross section. The inclusion of higher order corrections increases the cross section over the Born calculations, performed for the same scale of QCD coupling parton density. These corrections shift the leptoquark mass spectrum limit upward by approximately 100 GeV, which can be easily probed at the LHC. Figure 1.2 shows the results of the cross section calculations performed for the design center-of-mass energy for the LHC.

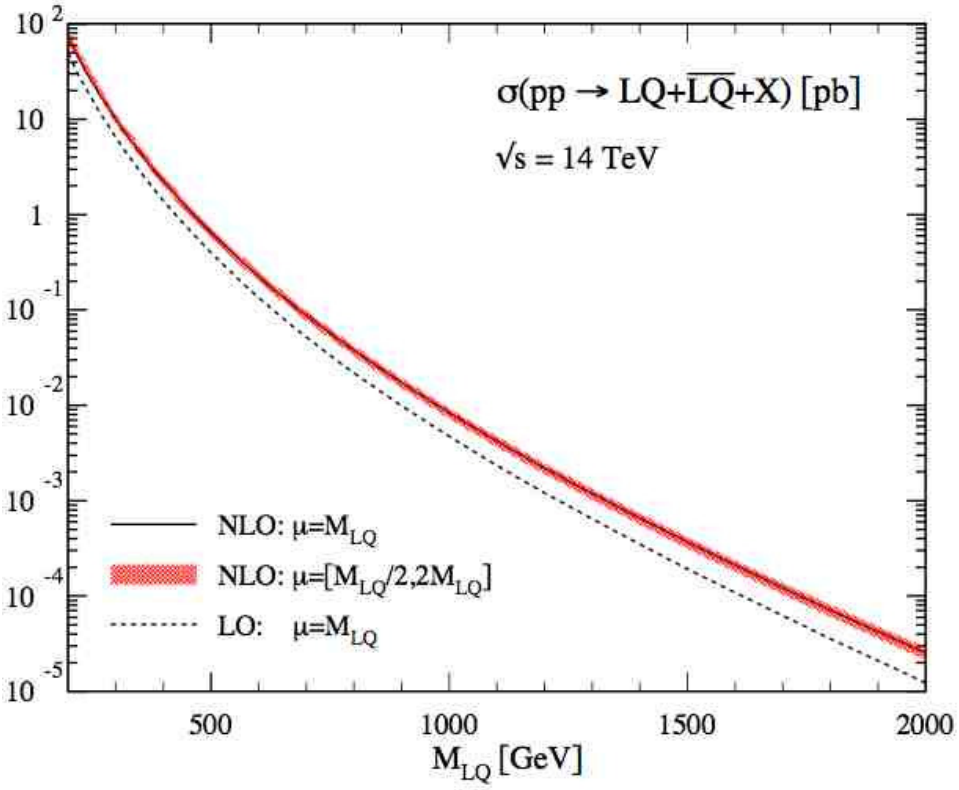


Figure 1.2: Cross Section for Leptoquarks for LHC Design Energy [17]. The cross section is also computed for the LHC $\sqrt{s} = 7$ TeV. Since this was the operational energy for the LHC for the analysis, the results are shown in Figure 5.1.

1.4.5 Experimental Signatures

The analysis of this thesis focuses on scalar pair production, $pp \rightarrow LQ\bar{LQ}$ using decay modes in which one leptoquark decays to a first generation charged lepton (electron) and a quark and the other decays to a neutrino and a quark. The leptoquark decays are required to involve leptons, neutrinos and quarks of the same generation, so the fundamental decay modes (ignore distinction of particle and antiparticle) are $LQ\bar{LQ} \rightarrow e\nu_e ud$ with the lepton/quark correspondence chosen to give the LQ and $\bar{LQ}$ opposite electric charge. Analysis using the decay modes in which both leptoquarks decay to a charged lepton and a quark are also performed [18].

1.4.6 Current Leptoquark Search Limits

Searches for leptoquarks have been performed at HERA (ep collisions), LEP (e^+e^-) and the Tevatron ($p\bar{p}$). Current leptoquark mass bounds come from searches carried out at the Tevatron which combine searches in the $\ell^+\ell^- qq'$ and $\ell\nu qq'$ final states. In addition to the ATLAS results, CMS similarly performed searches for leptoquarks in the $\ell^+\ell^- qq'$ and $\ell\nu qq'$ final state. It should be noted that the CMS $\ell\nu qq'$ results were submitted for publication after the results presented in this thesis.

The Tevatron 95% C.L. bounds at a leptoquark branching fraction to charged leptons (β) of 1/2, the point of maximum sensitivity in this $\ell\nu qq'$ final state, are $M_{LQ} > 284$ GeV for first generation. If only the $\ell\nu qq'$ final state is considered, the bounds are $M_{LQ} > 265$ GeV as can be seen in Figure 1.3.

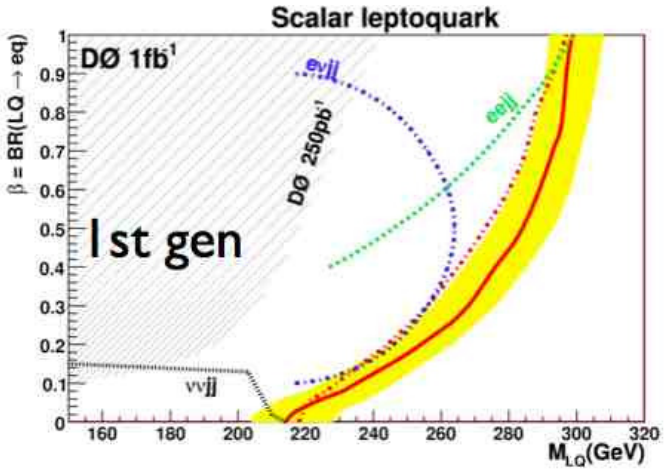


Figure 1.3: D0 Limit [19] for leptoquark production and decay into first generation quarks and leptons as a function of branching fraction, β, and leptoquark mass. The branching fraction is defined as the percentage of time a leptoquark decays into a charged lepton and a quark. The results for three different channels are displayed: decay to two electrons and two quarks is shown with a green dotted line, decay to an electron, a neutrino and two quarks is shown with a blue dotted line, and decay to two neutrinos and two quarks is shown with a black dotted line. The theoretical prediction combination of these three channels is shown with a dotted purple line. The observed limit is shown with a red line with $\pm 1\sigma$ in the yellow band. The previous D0 limits using 250 pb^{-1} of data are shown in gray hatching.

The analogous CMS bounds are approximately $M_{LQ} > 340$ GeV for $\beta = 0.5$ when the charged and neutral lepton channels are combined in the first generation. This exclusion plot is shown in Figure 1.4.

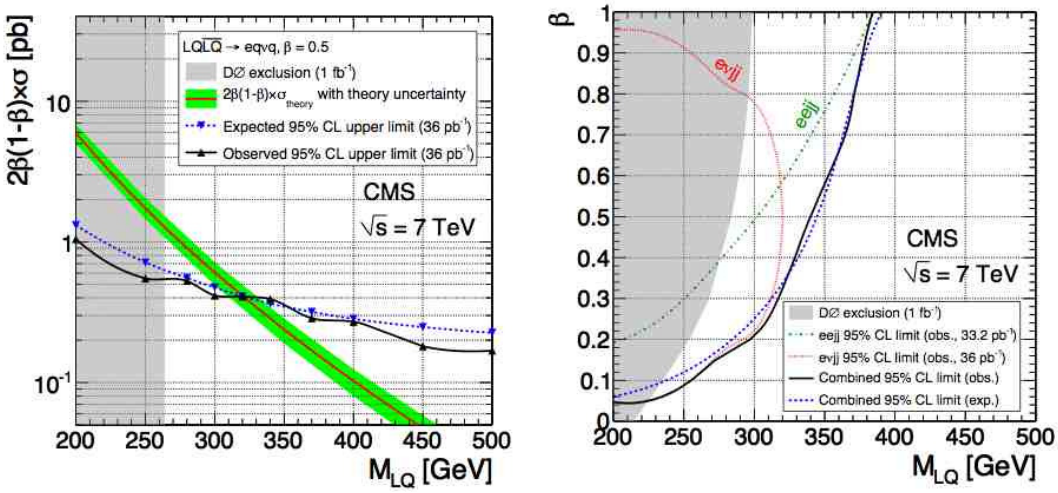


Figure 1.4: CMS limits on production of leptoquarks which decay into first generation quarks and leptons. The figure on the left shows the limits on the production cross section for branching fraction, β =0.5. The theoretical production is shown with a red line and the green band represents the uncertainty. The blue dotted line is the expected results and solid black is observed. The β vs. mass plane results for two channels are displayed on the right: decay to two electrons and two quarks is shown with a green dotted line, decay to an electron, and a neutrino and two quarks is shown with a solid pink line. The combination is shown with a black dotted line and the solid black line shows the data [20].

Chapter 2

Experimental Apparatus

The Large Hadron Collider (LHC) is a proton-proton collider run by the European Organization for Nuclear Research (CERN). It is located in the same tunnel that housed LEP, an electron-positron collider, crossing the boarder between Switzerland and France. The LHC was designed to run at a center-of-mass energy of $\sqrt{s} = 14$ TeV. However for the initial running period in 2010, it operated at $\sqrt{s} = 7$ TeV. With this run period, the LHC became the highest energy collider in the world. The data, in this thesis were collected during the period of running in 2010 by the ATLAS detector, one of two multi-purpose experiments located on the LHC. The layout can be seen in Figure 2.1. This chapter discusses the experimental apparatuses that make the analysis possible.

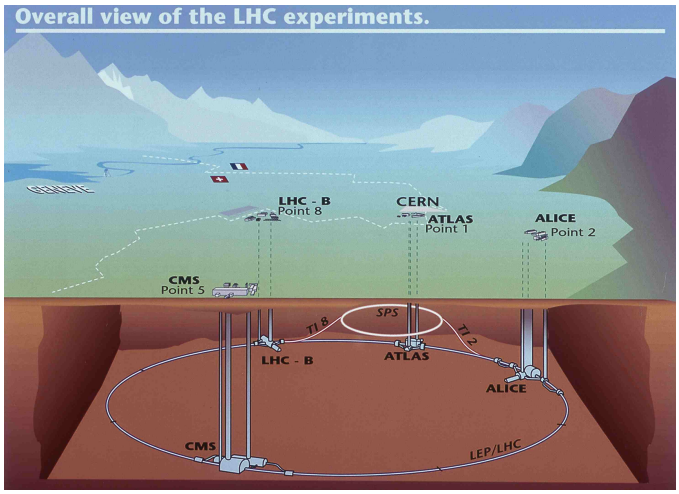


Figure 2.1: The LHC nestled between the Jura and Alps mountain ranges, straddling the border between France and Switzerland [21].

2.1 The Large Hadron Collider

The LHC is the final element in a chain of five accelerators used to create and accelerate the proton beams to their collision point. This section discusses these accelerators. A layout of the CERN accelerator complex is shown in Figure 2.2.

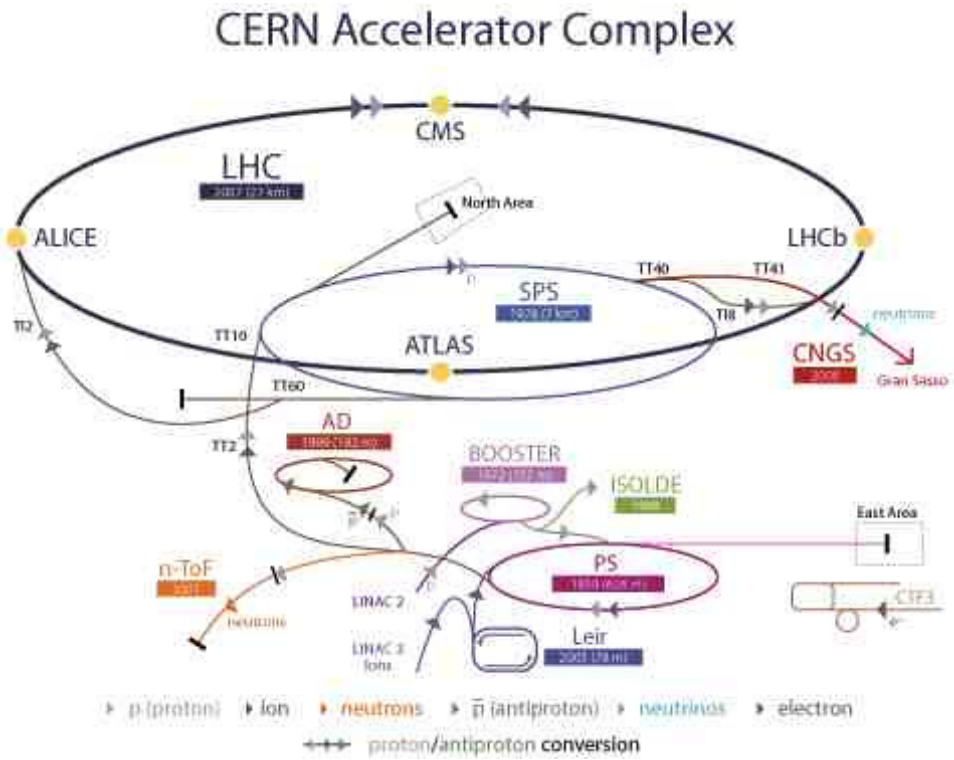


Figure 2.2: The LHC Accelerator Complex. The chain of five accelerators and associated experiments located at the CERN site [21].

2.1.1 The Proton Beam

The protons used in the collisions at the LHC are produced by ionizing Hydrogen gas in a duoplasmatron source. A schematic of this is shown in Figure 2.3 [22]. The process starts when a cathode filament emits electrons into a vacuum chamber. Hydrogen gas is introduced into the chamber, where it becomes ionized through interactions with the free electrons from the cathode, forming a plasma. The plasma is then accelerated through a series of charged grids, and becomes an ion beam (proton beam), moving at high speed from the aperture of the device.

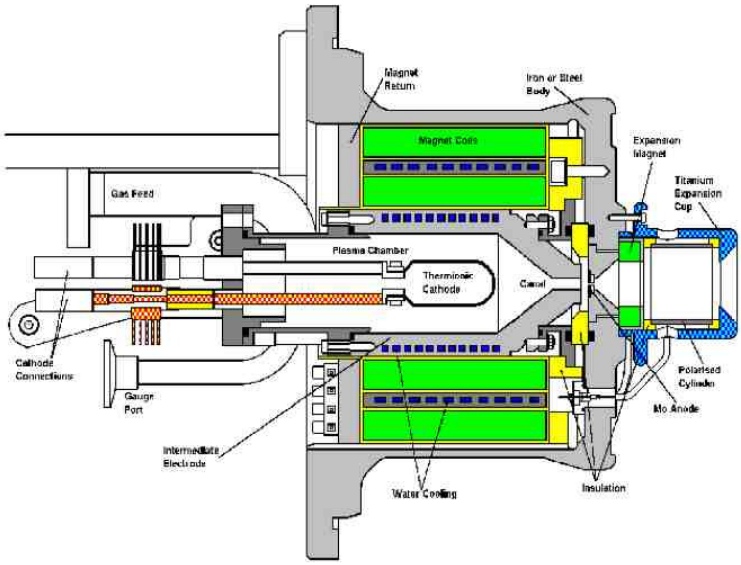


Figure 2.3: The Duoplasmatron. A cathode filament emits electrons into a vacuum chamber where it ionizes a hydrogen gas forming a plasma. The plasma is then accelerated through a series of charged grids, and becomes an ion beam (proton beam) [22].

The proton beam is then directed into a linear accelerator ("Linac2") which accelerates the beam to 50 MeV. This is the beginning of the accelerator chain shown in Figure 2.2. Once accelerated, these protons then are injected into the first of three synchrotrons: The Proton Synchrotron Booster (PSB, or "Booster"). The Booster both accelerates protons to 1.4 GeV and provides beam to the ISOLDE (Isotope Separator On-Line) experiment. From here, the Booster injects the proton beam into the Proton Synchrotron (PS) which increases the beam energy to 26 GeV. A physics complex called the East Area utilizes this beam when its not being sent elsewhere. Once the protons are up to energy, they are injected into the Super Proton Synchrotron (SPS). There are a number of facilities off of this ring: (1) CNGS, a fixed target experiment which sends neutrinos off to Gran Sasso, Italy, and (2) the North Area physics facility. Once the protons are accelerated to 450 GeV in the SPS, they are then injected into the main LHC ring.

2.1.2 The LHC

The proton beams were successfully circulated at the LHC for the first time in September 2008. The operations were interrupted due to an electrical short between magnets which caused a severe Helium leak and restarted in November 2009 at the injection energy of 450 GeV per beam. The first collision at a center-of-mass energy of $\sqrt{s} = 7$ TeV took place

on March 30, 2010 with an instantaneous luminosity ($\mathcal{L}_{inst}$) of 2×10^{27} cm^{-2}s^{-1} and by October 2010 the luminosity reached 10^{32} cm^{-2}s^{-1}. There were a maximum of 200 bunches simultaneously in the machine during this running, and approximately 10^{15} protons per bunch. These numbers will increase dramatically for the 2011/2012 runs. At the end of the 7 TeV experimental period (the end of 2012), the LHC will shut down for maintenance for up to two years so that it will be able to reach the design energy of 14 TeV.

There are six detectors constructed at the LHC. ATLAS and the Compact Muon Solenoid (CMS) are the two general-purpose experiments intended to operate at the peak energy and luminosity and search for new physics. The LHC machine also accelerates lead ion beams, which will collide at energies of up to 5.5 TeV per nucleon pair. A Large Ion Collider Experiment (ALICE) is the dedicated heavy-ion experiment. The focus of the Large Hadron Collider beauty (LHCb) experiment is b-quark physics in the forward region of the interaction point. The TOTal Elastic and diffractive cross section Measurement (TOTEM) detector is designed for the detection of protons from elastic scattering at small angles and is used to monitor accurately the LHC's luminosity. The Large Hadron Collider forward detector (LHCf) uses forward particles created inside the LHC as a source to simulate cosmic rays in laboratory conditions.

2.2 The ATLAS Detector

ATLAS is designed as a general-purpose experiment. A variety of different particles with a broad range of energies are produced by LHC collisions. Rather than focusing on a particular physical process, ATLAS is designed to detect and measure whatever form any new physical processes or particles might take. ATLAS follows this tradition made by previous experiments at other colliders, such as the Tevatron, which were designed based on this philosophy. However, because of the LHC's unprecedented energy and extremely high rate of collisions, ATLAS is required to be larger and more complex than any detector ever built.

2.2.1 Coordinate System

The ATLAS detector uses a right-handed coordinate system with the x-axis pointing to the center of the LHC ring, the y-axis pointing straight up, and the z-axis pointing along the beam direction. This is shown in Figure 2.4. There are two principle angles used in discussing event parameters and data: ϕ (the angle in the xy-plane measured with respect to the x-axis) and θ_{cm} (the angle in the xz-plane measured with respect to the x-axis). θ_{cm} is used to calculate the pseudorapidity (η). In the limit where the particle is traveling close to the speed of light, or in the approximation that the mass of the particle is nearly zero, pseudorapidity is numerically close to the definition of rapidity. η is defined in Equation 2.1.

$$\eta = -ln\left(tan\frac{\theta_{cm}}{2}\right) \tag{2.1}$$

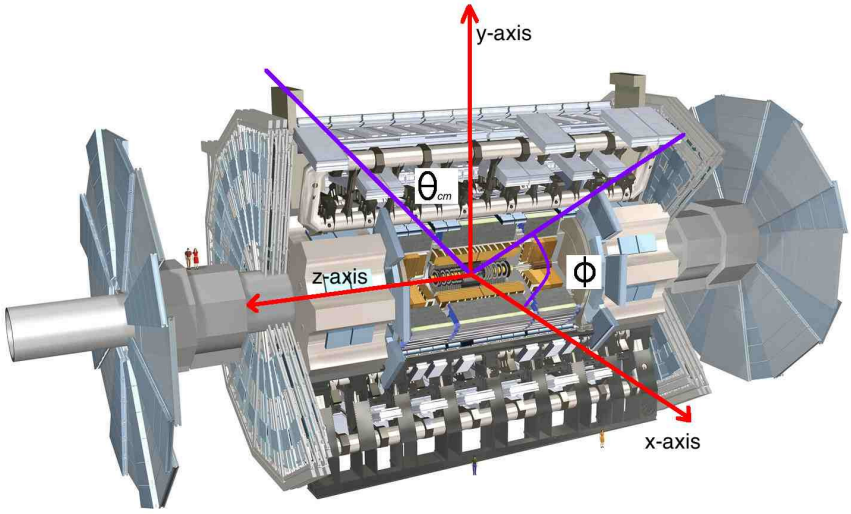


Figure 2.4: The ATLAS Coordinate System. The x-axis is pointed in the direction of the LHC center, the y-axis is up, and the z-axis is along the beam pipe. ϕ is the angle in the xy-plane (transverse plane). θ_{cm} is the angle in the xz-plane and is used to calculate the pseudorapidity (η) [21].

2.2.2 Luminosity

Luminosity ($\mathcal{L}$) is a measure of the number of inelastic pp collisions per units time and area and is defined in Equation 2.2.

$$\mathcal{L} = \frac{\mu n_b f_r}{\sigma_{incl}} = \frac{\mu_{vis} n_b f_r}{\epsilon \sigma_{incl}} = \frac{\mu_{vis} n_b f_r}{\sigma_{vis}} \tag{2.2}$$

μ is the average number of inelastic interactions per bunch crossing, n_b is the number of bunch crossings producing collisions per machine revolution, f_r is the machine revolution frequency, σ_{incl} is the cross-section for pp inelastic collisions, and σ_{vis} is the cross section seen by the detector. σ_{vis} is proportional to $\sum_x$ and $\sum_y$, which are measured by van der Meer (vdM) scans by each luminosity detector and are defined in Equation 2.3 [23]

$$\sum_x = \frac{1}{\sqrt{2\pi}} \frac{\int R_x(\delta) d\delta}{R_x(0)} \tag{2.3}$$

where $R_x(\delta)$ is the rate measured during a horizontal scan and the beams are separated by the distance δ ($\delta = 0$ represents the case of zero beam separation). In the case where $R_x(\delta)$ is Gaussian, $\sum_x$ corresponds to the standard deviation. The rate of a process ($\mathcal{R}$) can be defined in terms of the cross section (σ) and $\mathcal{L}$, shown in Equation 2.4.

$$\mathcal{R} = \sigma \times \mathcal{L} \tag{2.4}$$

For the 2010 run period the ATLAS detector had two dedicated detectors to determine the luminosity provided by the LHC in real-time (online), (LUCID and BCM) as well as another detector (MBTS) which was used to determine the luminosity after the run (offline) [24]. These detector systems can be seen in Figure 2.5. During data taking, the runs are broken into luminosity blocks (LB), which are the smallest time intervals for which the integrated, dead-time and prescale-corrected luminosity is reported. LBs for the 2010 run were two minutes in length.

LUCID

The LUminosity measurement using a Cherenkov Integrating Detector (LUCID) is a Cherenkov light detector primarily dedicated to online luminosity monitoring. LUCID calculates the relative luminosity by detecting inelastic *pp* scattering in the forward direction ($5.6 < |\eta| < 6.0$). By doing this it can both measure the integrated luminosity and provide online monitoring of the instantaneous luminosity and beam conditions [25]. It is designed to have a sufficient time resolution in order to identify individual bunch crossings. Of the luminosity detectors discussed, it is the farthest from the interaction point.

BCM

The Beam Conditions Monitor (BCM) consists of Chemical Vapor Deposition (CVD) diamond sensors arranged around the beam pipe in a cross pattern. The BCM is primarily designed to detect beam losses which can indicate a risk of damage to the ATLAS detectors. The fast readout of the BCM also provides a useful, low acceptance bunch-by-bunch luminosity signal at $|\eta| = 4.2$ [26]. The luminosity is determined from the average number of inelastic collisions per beam crossing ($\bar{N}_{LUCID}$ and $\bar{N}_{BCM}$). The number of collisions per crossing follows a Poisson distribution. From the fraction of beam crossings with no collisions $\bar{N}$ can be determined via the Poisson probability, $P(\sigma, \bar{N}) = e^{-\bar{N}}$.

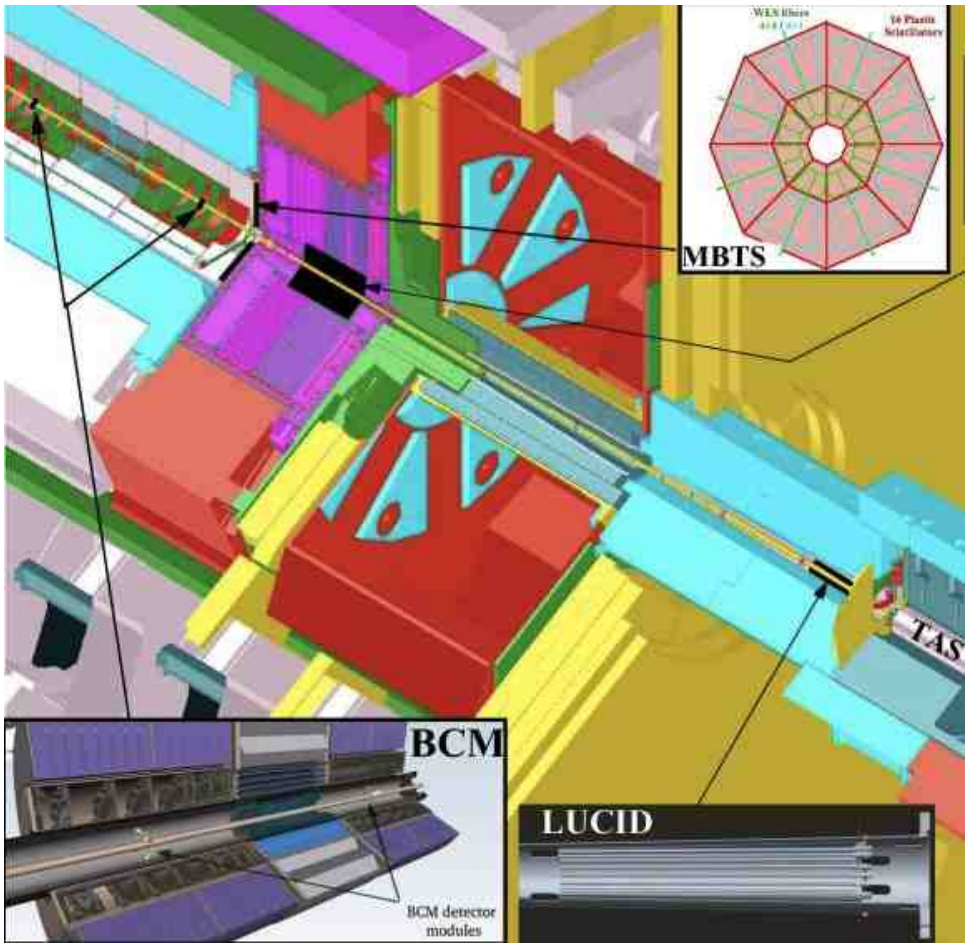


Figure 2.5: The Luminosity Detectors: LUminosity measurement using a Cherenkov Integrating Detector (LUCID), Beam Conditions Monitor (BCM), and Minimum Bias Trigger Scintillators (MBTS) shown in relation to the rest of the ATLAS detector [21].

MBTS

In addition to the online luminosity detectors, ATLAS has an offline algorithm to measure the luminosity which utilizes Minimum Bias Trigger Scintillators (MBTS) and the inner detector (ID). The MBTS consists of large-area scintillation counters placed before the endcap electromagnetic calorimeter. The MBTS data is triggered and read out through the standard ATLAS data acquisition system, which is different than the online algorithms used by LUCID and the BCM using dedicated hardware. Unfortunately there were not enough statistics in the 2010 run to be able to make an accurate offline luminosity measurement [25].

2.2.3 Tracking

The ATLAS tracking system is designed to track charged particles that originate from the interaction point (IP). It consists of three subdetectors, which specialize in different aspects of track reconstruction. In order of increasing radius are: (1) the silicon pixel detector, (2) the SemiConductor Tracker (SCT), and (3) finally the Transition Radiation Tracker (TRT). This system of detectors is housed within the two Tesla magnetic field of a superconducting solenoid magnet. The field curves the trajectory of the charged particles, allowing for precision measurements of both momentum resolution and charge in an environment of high track multiplicity to an $|\eta| < 2.5$.

The innermost pixel detector is finely granulated to measure vertices accurately. The SCT measures the particle momenta precisely and the TRT specializes in pattern recognition due to the large possible number of hits and contributes to electron identification. Figure 2.6 illustrates the layout of each detector which consists of barrel and endcap regions as met by particles coming from the interaction vertex [27].

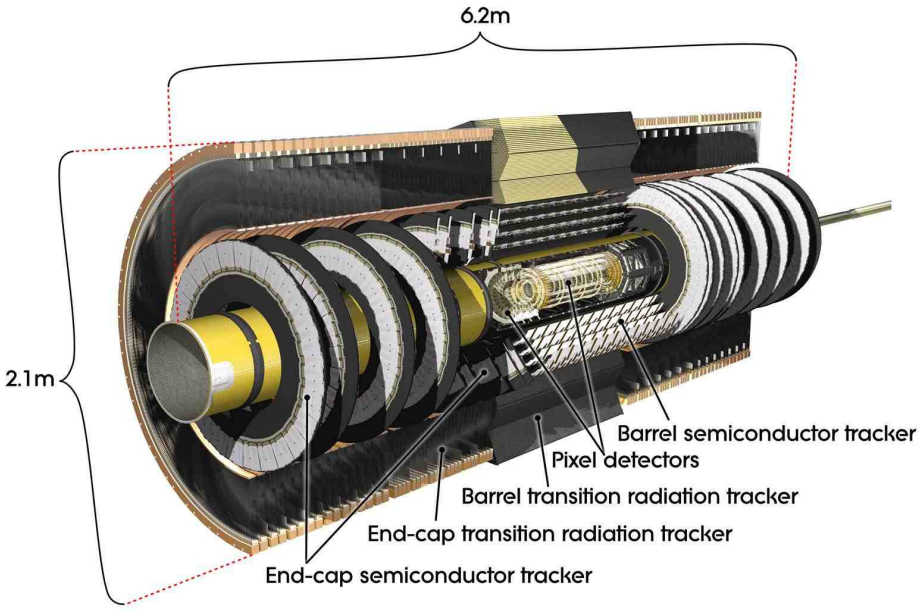


Figure 2.6: The Inner Detector. The innermost part of the tracking system is the pixel detector. It is also the most finely granulated. Outside of this is the SemiConductor Tracker, which is also made of silicon. The final part of the tracking system is the Transition Radiation Tracker [21].

The resolution of a track parameter X can be expressed as a function of p_T by: $\sigma_X(p_T) = \sigma_X(\infty)(1 \oplus p_X/p_T)$, where $\sigma_X(\infty)$is the asymptotic resolution expected at infinite momentum

and p_X is a constant representing the value of p_T for which the intrinsic and multiple-scattering terms in the equation are equal for the parameter X. The momentum resolution of the Inner Detector is σ/p_T=5×10$^{-4}p_T$/GeV $\oplus$ 0.01. This expression works well at high p_T (where the resolution is dominated by the intrinsic detector resolution) and at low p_T (where the resolution is dominated by multiple scattering). $\sigma_X(\infty)$ and p_X are implicitly functions of the pseudorapidity.

Pixel Detector

The pixel detector is located closest to the IP, a mere 50 mm away from the beam pipe. It consists of finely grained silicon chips designed to accurately reconstruct the primary vertex. The pixel detector is comprised of 1744 modules giving a total of nearly 80 million channels in a cylinder 1.4 m long, and 0.5 m in diameter. The barrel part of the pixel detector consists of the 3 cylindrical layers with the radial positions of 50 mm, 88 mm and 120 mm whose layout is shown in Figure 2.7. These three barrel layers are made of identical segments inclined with azimuthal angle of 20 degrees. There are 22, 38 and 52 segments in each of these layers respectively. Each segment is composed of 13 pixel modules. The sensors are made of a 250 μm thick high resistivity n-type silicon bulk, with p+ and n+ type implantation on opposite sides [28]. There are three disks on each side of the forward regions. One disk is made of 8 sectors, with 6 modules in each sector. Disk modules are nearly identical to the barrel modules.

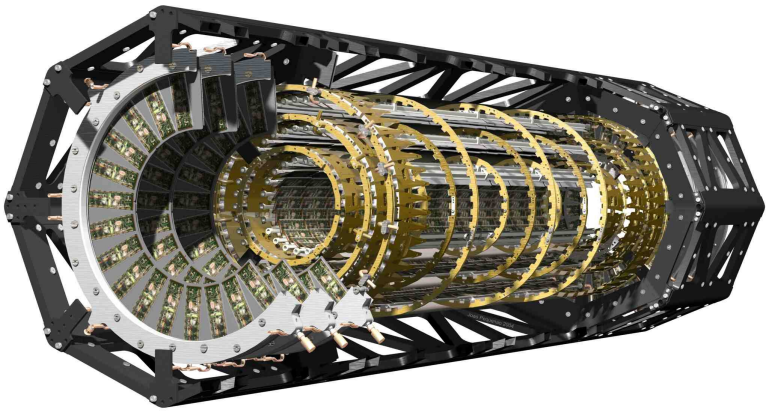


Figure 2.7: The Pixel Detector. It consists of three concentric barrels and three endcaps on each side [21].

A summary of the pixel detector's parameters can be found in Table 2.1. This geometry results in the following [27]:

- primary vertex resolution: 11 μm in the transverse plane and 45 μm in the z plane

- secondary vertex resolution: 50 μm in the transverse plane and 70 μm in the z plane
- impact parameter resolution: 10 μm in the transverse plane and 70 μm in the z plane

SemiConductor Tracker (SCT)

Moving outside of the pixel detector, the next layer of the tracking system is the Semi-Conductor Tracker (SCT). A slice of the SCT can be seen in relation to the rest of the inner detector in Figure 2.8. The layers in blue are the representation of the SCT, shown above the green pixel layers and the large tube which is the beam pipe. The job of the SCT is very similar to the pixel detector: to provide spacial measurements of charged tracks coming from the IP. The SCT gives 8 measurements per track and is comprised of 2,112 barrel and 968 endcap double sided modules with 6 million readout strips readout every 80 mm. The sensors are single sided p-on-n strips with an accuracy of 17 mm/layer in the direction transverse to the strips. In this way the information from the SCT can be added to the already existing pixel information and used to improve the tracking reconstruction and resolution [27, 29].

Just as for the pixel detector, the layers of the SCT are in the barrel region arranged in concentric cylinders around the beam axis while in the endcap regions they are arranged as disks perpendicular to the beam axis. This configuration can be seen in Figure 2.6. A summary of the parameters of the SCT can be found in Table 2.1.

Transition Radiation Tracker (TRT)

The final and outermost part of the tracking system in ATLAS is a combined straw tracker and transition radiation detector. This system is called the Transition Radiation Tracker (TRT) and can be seen in both Figure 2.6 and as the outermost slice in Figure 2.8. Each straw is filled with gas mixture (70% Xe, 27% CO_2 and 3% O_2) that becomes ionized when a charged particle passes through. The straws are held at a large negative voltage which drives the ionized particles to a fine wire in the center of the straw. The ionized electrons drift towards the central anode wire, and cause an avalanche of ion-electron pairs that amplifies the signal by a factor of 10,000. This creates a detectable electrical signal on the wire. It provides up to 36 two-dimensional measurements per track track.

The straws of the TRT are four mm in diameter with an inner wire made of gold plated Tungsten which is 0.03 mm in diameter. The signal from the wires leaves a path of hit straws which allow the trajectory to be reconstructed and based on the geometry of the TRT. The barrel contains 52,544 axial straws of 144 cm long at radii between 35 cm and 107 cm. The endcaps contain a total of 245,760 radial straws at radii between 64 cm and 107 cm [27, 30]. The TRT provides a spacial hit resolution of 0.150 mm for charged particle tracks with $|\eta| < 2.5$ and $p_T > 0.5$ GeV. A summary of these parameters can be found in Table 2.1. Since the amount of transition radiation is dependent on the speed of the particle, light particles leave strong signals. This allows the TRT to be particularly useful in electron reconstruction.

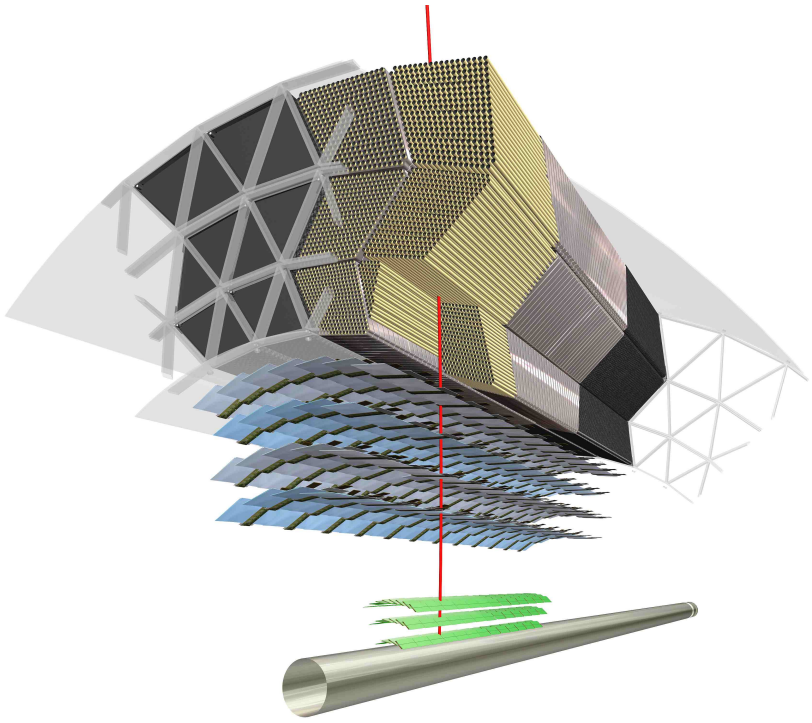


Figure 2.8: Inner Detector Slice. The three green layers closest to the IP are the three barrel pixel layers. Outside of the pixel layers in blue are the barrel parts of the SCT. The straws shown furthest from the beam pipe are the barrel part of the TRT [21].

Detector	Radial ext. (mm)	Length (mm)	Accuracy (μm)
Pixel B	$50.5 < R < 122.5$	$0 < \|z\| < 400.5$	10 (R-ϕ) 115 (z)
Pixel EC	$88.8 < R < 149.6$	$495 < \|z\| < 650$	10 (R-ϕ) 115 (z)
SCT B	$255 < R < 549$	$0 < \|z\| < 805$	17 (R-ϕ) 580 (z)
SCT EC	$251 < R < 610$	$810 < \|z\| < 2797$	17 (R-ϕ) 580 (z)
TRT B	$554 < R < 1082$	$0 < \|z\| < 780$	130
TRT EC	$617 < R < 1106$	$839 < \|z\| < 2744$	130

Table 2.1: Summary of parameters for Inner Detector. EC stands for endcaps and B is for barrel. The coverage of the inner detector extends to $|\eta| < 2.5$ [31].

2.2.4 Calorimeters

The system of calorimeters provides the capability to measure energy and position of electrons, photons and jets and to distinguish between electromagnetic (electrons) and hadronic (jets) objects. The purpose of these detectors is to have the desired particle deposit all

of its energy in the respective calorimeter, while allowing other particles to pass through with minimal energy loss. Additionally, the calorimeter needs to have good hermeticity to determine the missing transverse energy ($\not{E}_T$). There are two components to the ATLAS calorimeters: the electromagnetic calorimeters and the hadronic calorimeters. These two calorimeter systems are shown in Figure 2.9 and have energy resolution requirements listed in Table 2.2.

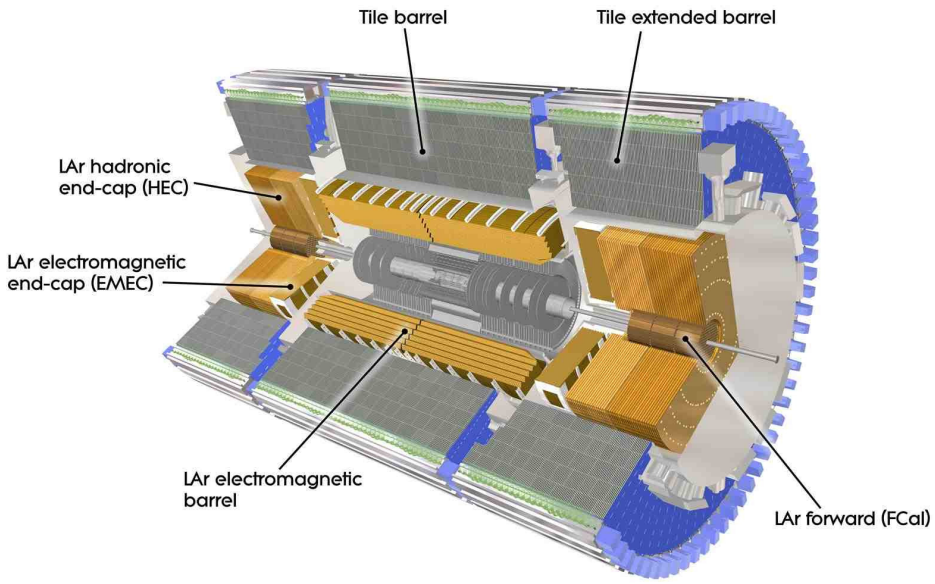


Figure 2.9: The electromagnetic and hadronic calorimeters. The electromagnetic calorimeter consists of a Liquid Argon (LAr) sampling calorimeter with accordion shaped electrodes. It is shown in gold. There are four parts to the LAr detector: the electromagnetic barrel and endcap, the hadronic endcap, and the LAr forward calorimeter (located at high $|\eta|$). The hadronic calorimeter also has a scintillating tile barrel and extended barrel [21].

Electron and photon measurement is based on showers produced by these particles when incident on matter. High-energy electrons predominantly lose energy in matter by bremsstrahlung, and photons by pair production. The characteristic amount of matter traversed for these related interactions is called the radiation length (X_0). For electrons, X_0 is the mean distance over which it loses all but 1/e of its energy by bremsstrahlung. For photons, X_0 is 7/9 of the mean free path for pair production. X_0 is also a convenient length scale by which to characterize the showering in the electromagnetic calorimeter. The electromagnetic calorimeter in ATLAS is a Liquid Argon (LAr) sampling calorimeter, which uses lead plates to cause electromagnetic particles to shower while allowing other particles such

as hadrons or muons to pass through with as little energy loss as possible [32].

Hadrons, which come from either quark and gluon radiation or are associated with the hard scatter, produce cascades of successive inelastic hadron-nuclear interactions. These showers are generally broader than electromagnetic ones. The depth of showering in a material is characterized by the interaction length (λ_0), which is defined as the mean path between hadronic interactions. The hadronic calorimeter on ATLAS is comprised of 3 different detectors: a scintillating tile calorimeter in the central region, LAr endcaps and a LAr forward calorimeter.

Detector	Resloution (σ_E/E)
Electromagnetic (LAr)	$10\%/\sqrt{E}\oplus$ 0.7%
Hadronic (Tile and LAr Had)	$50\%/\sqrt{E}\oplus$ 3%
Forward	$100\%/\sqrt{E}\oplus$ 10%

Table 2.2: Energy Resolution Requirements for electromagnetic, hadronic and forward calorimeters [32, 33].

Liquid Argon Calorimeters

The Liquid Argon (LAr) calorimeter plays a central role in ATLAS. This system of calorimeters is needed to reconstruct electromagnetic particles, jets, and $\not{E}_T$. The LAr calorimeter is a sampling calorimeter with accordion shaped absorbers and electrodes covering a pseudorapidity range $|\eta|$ <1.475 in the barrel region and $1.375 < |\eta| < 3.2$ in the endcap regions. This accordion shaped geometry provides continuous coverage in the azimuthal direction (ϕ) and can be seen in Figure 2.9. The calorimeter cells are mostly uniform in size in ϕ and η for each of the three sampling layers (especially in the central region of $|\eta| < 2.5$). Because the cells are uniform in η and ϕ, not linear dimension, the granularity is especially fine in the central region and more at larger η toward the forward region [34]. The granularity of the LAr EM calorimeter is summarized in Table 2.3.

The central barrel is comprised of two half barrels joined at η= 0. Each half barrel is made up of 1024 accordion-shaped lead absorbers. Read out kapton electrodes are positioned in the middle of the accordion gaps. The distribution of material in front of the LAr barrel and endcap calorimeters requires that a presampler be included in the calorimeter system to correct for energy lost in front of the calorimeter. The barrel and endcap presampler have 1 cm and 5 mm liquid argon active layers respectively, which is instrumented with electrodes roughly perpendicular to the beam axis in the barrel and parallel to it in the endcap.

The first sampling is situated closest to the IP and has the finest granularity in η. It is approximately six X_0 thick (including the presampler and dead material), which allows for precise position measurement and the potential for differentiation between π^0 and electron/photon shower shapes. The second sampling is the thickest layer, at least 22 X_0 in the

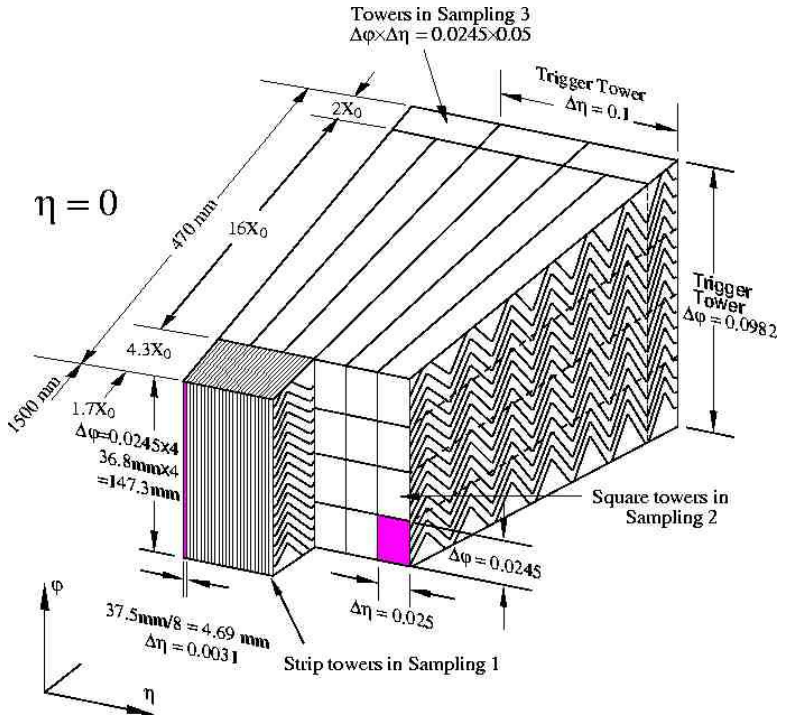


Figure 2.10: LAr Calorimeter. A segment of the LAr Calorimeter is shown. The first layer (sampling 1) is most finely grained in η. The second sampling is the thickest in X_0 and has square shaped cells equal in $\Delta\eta$ and $\Delta\phi$. The final layer (sampling 3) is shown at the very back with larger cells in $\Delta\eta$ than the previous layer. The presampler is not shown [32].

Detector	$0 < \|\eta\| < 1.8$	$1.8 < \|\eta\| < 2.0$	$2.0 < \|\eta\| < 2.5$	$2.5 < \|\eta\| < 3.2$
Presampler	0.025×0.1			
Sampling 1	0.003×0.1	0.004×0.1	0.006×0.1	0.1×0.1
Sampling 2	0.025×0.025	0.025×0.025	0.025×0.025	0.1×0.1
Sampling 3	0.050×0.025	0.050×0.025	0.050×0.025	
Trigger	0.1×0.1	0.1×0.1	0.1×0.1	0.2×0.2

Table 2.3: Granularity of LAr EM Calorimeter in $\Delta\eta \times \Delta\phi$ [32].

barrel and 26 X_0 in the endcaps. This geometry results in electrons and photons depositing most of their energy in this layer. The third and final sampling is designed to recover energy from particles energetic enough to pass beyond the second sampling. It is between two and twelve X_0 thick depending on η. Since hadronic showers are more likely to extend past the EM calorimeter into the hadronic calorimeter, this layer helps in the discrimination between electromagnetic and hadronic showers. The three sampling layers can be seen in Figure 2.10.

In addition to the electromagnetic LAr calorimeter system, there is also a hadronic LAr detector, with slightly different material and electronics [32, 35]. Each hadronic endcap calorimeter consists of two independent wheels each with 32 modules of outer radius 2.03 m. They cover a pseudorapidity between $1.5 < |\eta| < 3.2$. The wheels are made out of copper plates. A 8.5 mm gap between consecutive plates contains three parallel electrodes which split the gap into drift spaces of about 1.8 mm. The first wheel has two longitudinal segments which are eight and 16 layers in depth respectively and provide a depth of λ_0= 1.4 and 2.9 respectively. The second wheel has one segment of 16 layers which has a depth of λ_0= 5.7. These cells are fully pointing in ϕ, but only pseudo-pointing in η. To avoid a crack in material density at the transition between the endcap and the forward calorimeter ($|\eta|$= 3.1), the acceptance of the electromagnetic calorimeter goes down to $|\eta|$= 3.2. More detailed information about the granularity of the hadronic LAr endcap is given in Table 2.4.

Detector	First Wheel	Second Wheel
$\|\eta\| < 2.5$	0.1×0.1	0.1×0.1
$\|\eta\| < 3.1$	0.2×0.2	0.2×0.2
RO Channels	(768+736)$\times$ 2	704 $\times$ 2

Table 2.4: Granularity of LAr Hadronic Calorimeter in $\Delta\eta \times \Delta\phi$ [32].

The final LAr calorimeter is the forward calorimeter (FCAL). The FCAL is exposed to a higher level of radiation being closer to the beam ($3.0< |\eta| < 4.9$). Instead of an accordion geometry, the FCAL is a metal matrix with regularly spaced longitudinal channels filled with concentric rods and tubes which run parallel to the beam pipe resembling a stopper for the LAr endcap hole. A cross-section view is shown in Figure 2.11. It consists of 3 subsystems creatively differentiated: FCAL 1, 2 and 3. The innermost subsystem (FCAL 1) covers the acceptance $3.0< |\eta| < 4.9$. It is made of copper absorbers with 12,000 rods and tubes and has a depth of $\lambda_0 = 2.6$. FCAL 2 is the middle segment and is made of tungsten absorbers with 10,000 rods and tubes. It covers the acceptance $3.1< |\eta| < 4.9$ and has a depth of $\lambda_0 = 3.5$. Lastly, FCAL 3 is the farthest from the IP and is also made of tungsten absorbers with 10,000 rods and tubes. It covers the acceptance $3.2< |\eta| < 4.9$ and has a depth of $\lambda_0 = 3.4$. Thus, the FCAL detector provides a total depth of approximately 9.5 λ_0.

Uniformity of the LAr Calorimeters

During the 2008-2009 shutdown of the LHC, ATLAS was busily collecting data from cosmic rays. These cosmic rays were used to measure the energy response of the ATLAS liquid argon electromagnetic calorimeter to muons which are Minimizing Ionizing Particles (MIPs). Several hundreds of million cosmic events were collected during this period. Clusters of cells (1×3 in $\eta\times\phi$ in the 2^{nd} sampling and 2×1 in the 1^{st}) were used to measure the upper bound on the non-uniformity. The non-uniformity along the η direction of the calorimeter

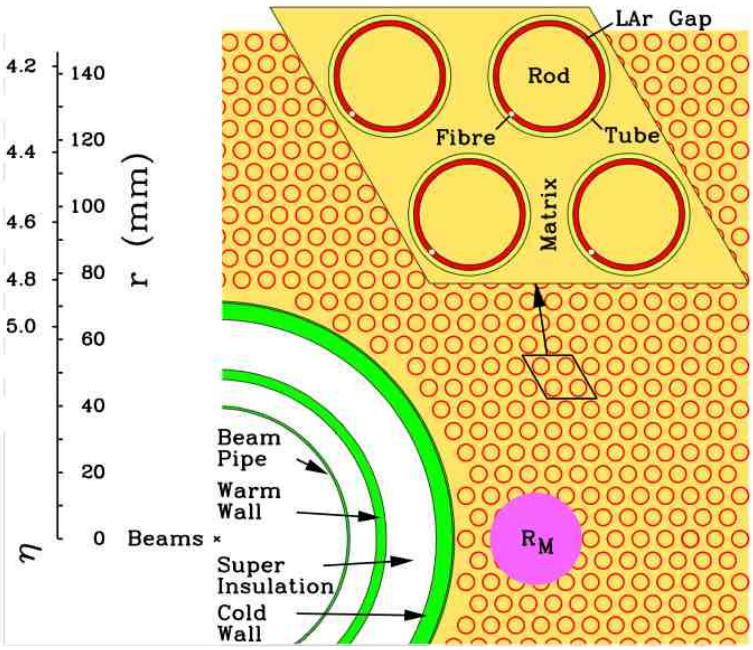


Figure 2.11: A cross section of the LAr Forward Calorimeters [32].

response in the first sampling layer was shown to be less than 1.7% in 2×1 cell clusters (corresponding to 0.00625×0.1 in $\Delta\eta \times \Delta\phi$). In the second sampling layer, the non-uniformity was shown to be less than 1% in 1×3 clusters (corresponding to 0.025×0.075 in $\Delta\eta \times \Delta\phi$). These results are shown in Figure 2.12. The region between $|\eta| < 0.7$ was mapped at the single cell level while a wider binning was needed between $0.7 < |\eta| < 1.4$ due to statistics. These numbers correspond to an average effective coverage of approximately 20% [34, 36].

This study was later done again with the full 2009 data set and once the LHC was running with collision data muons, which do not suffer from the fact that the muons do not come from the IP. When trying to understand in depth the sources of non-uniformity in the calorimeter response to electrons in ATLAS, an improved study of this kind could play an important role in separating the sources related to the calorimeter from those related to the upstream material.

Tile Calorimeter

The tile calorimeter is the barrel portion of the hadronic calorimeter [33]. It is located directly outside the LAr barrel calorimeter in the range $|\eta| < 1.0$, and has an extended barrel covering the range $0.8 < |\eta| < 1.7$. Its orientation with the LAr calorimeters can be seen in gray in Figure 2.9. The tile calorimeter is a sampling calorimeter using steel as the passive absorber and scintillating tiles as the active medium. Like the LAr EM calorimeter, there

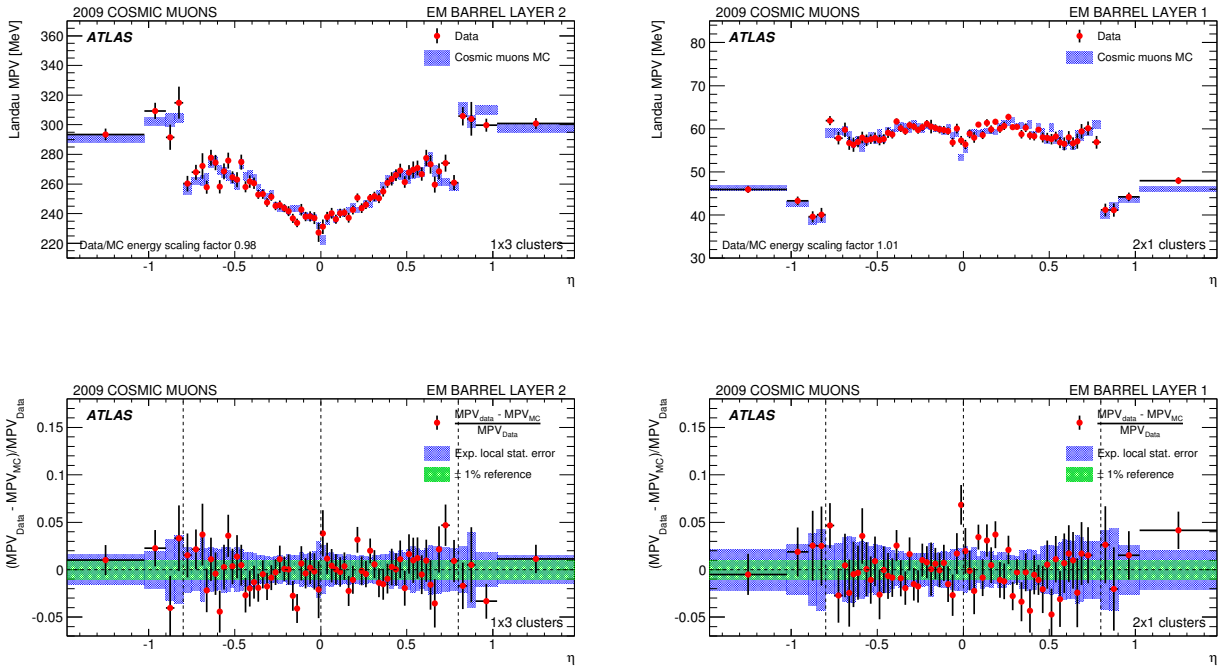


Figure 2.12: For the optimized clusters, evolution of MPV (upper plots) and MC (blue bands) as a function of η for middle clusters (left) and strip clusters (right), for data (dots with error bars). The measured local uniformity (lower plots) is plotted as a function of η (red dots). The expectation is indicated as the (blue) band and a ± 1 % reference strip is also shown (green) [36].

are three sampling layers. The tiles are 3 mm thick and point radially out from the beam line meaning that particles from the IP cross the tiles in a longitudinal direction at $\eta = 0$. This provides a finer readout segmentation in the z-plane which allows a more accurate determination of the shower shape. This is especially important for high p_T particles which form larger showers. To collect all the energy from hadronic particles, the tile is over ten interaction lengths (λ_0) thick. The energy deposited in the tile is read out by wavelength shifting (WLS) fibers attached to each end of the tile. These fibers are then coupled to a photomultiplier. The $\Delta\eta \times \Delta\phi$ granularity of the first two samplings is 0.1×0.1 and is 0.2×0.1 for the third [37].

2.2.5 Magnet System

The magnet system used at ATLAS had to be designed to provide the optimal conditions for particle identification and momentum measurement for the tracking and muon systems

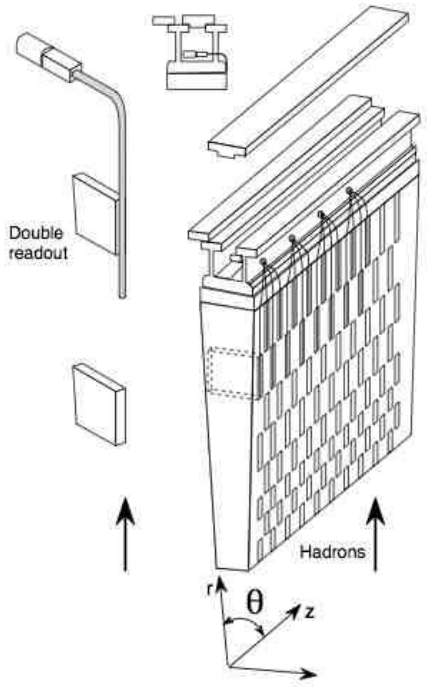


Figure 2.13: Tile Calorimeter. A slice of the tile calorimeter is shown demonstrating the orientation of the scintillating tiles. They are positioned in the longitudinal direction to give more segmentation in the z-direction [33].

separately. This required four independent magnets: the central solenoid and three toroid systems. The solenoid and torroid magnets can be seen Figure 2.14. The toroid magnets can also be seen in Figure 2.15 where they are integrated with the muon system.

Central Solenoid

The central solenoid encases the inner detector in a 2 T magnetic field. This field strength was chosen to maintain good tracking for lower p_T particles while still providing enough of a magnetic field to resolve highly energetic charged particles. This field is achieved by having 1154 turns around the central coil supplied with 7730 A. Another design consideration of the central solenoid was to minimize the amount of material before the electromagnetic calorimeter. The central solenoid is constructed from a single-layer coil that consists of high-strength Al-stabilised NbTi/Cu material. The axial length of the solenoid is 5.28 m and the radius is 1.27 m, resulting in a total thickness of 0.66 radiation lengths (X_0) [39, 40].

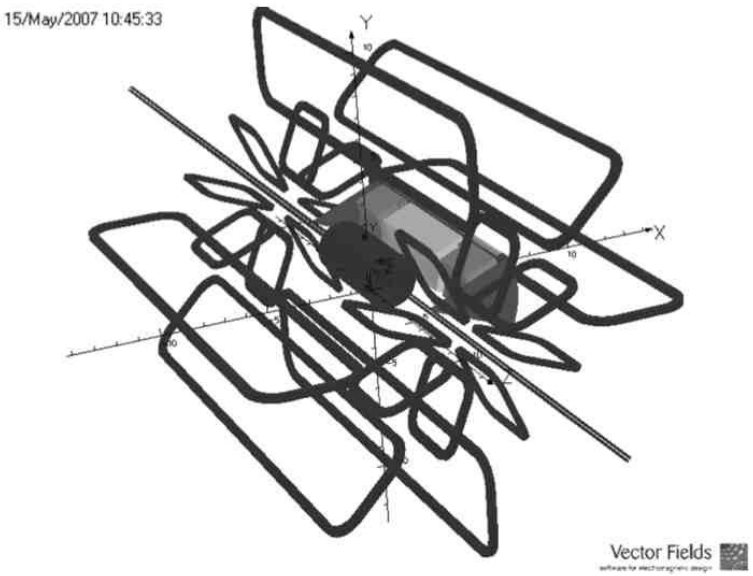


Figure 2.14: The Solenoid and Toroid Magnets. The solenoid magnet is located in the central region of the magnet system. It is barrel shaped and circles the beam pipe and inner detector. The large doughnut shaped magnets located around the solenoid and at the endcap regions are the toroid magnetic system [38].

Barrel and Endcap Toroids

The toroid system consists of three parts: a barrel part located outside the hadronic tile calorimeter and two endcaps located outside the liquid argon hadronic calorimeter. The 4 T barrel toroid magnet is the largest of the four superconducting magnets (25 m long, 20 m diameter). It provides the magnetic field for the ATLAS muon spectrometer. The barrel consists of eight identical racetrack-shaped coils encased in individual stainless steel cryostats. Each coil measures 25.5 m $\times$ 5.4 m and is made from an Al-stabilised Nb/Ti/Cu superconducting alloy. When all the eight modules are combined they produce the bending power for tracks in the central region ($|\eta| < 1$). For the region $1.4 < |\eta| < 2.7$, magnetic field is provided by the endcap toroids. The endcap toroids are similar to the barrel part in layout and choice of material. Only the dimensions are different and the fact that all endcap modules are encased in a single cryostat. In the endcap region, each module measures 4 m $\times$ 4.5 m [38, 41]. The layout of the toroid magnet systems can be seen in Figure 2.14.

2.2.6 Muon Systems

Muons are detected in ATLAS by measuring their energy loss by ionization and their trajectory in both the inner detector and a dedicated muon system [42, 43]. Since muons do

not interact strongly and have a relatively high mass, they can pass through the calorimeters depositing only a minimal amount of energy. Upon traversing the calorimeters, they enter the muon spectrometers which detect the charged particles. The muon spectrometers form the outer shell of the ATLAS detector. They lie on the outside of the calorimeter modules and cover a radius between 4.5 and 11 m extending approximately 23 m along the beam axis on both sides of the interaction point [31]. In the two endcap regions, muon chambers form large wheels, perpendicular to the z-axis and located at distances of $|z| = 7.4$ m, 10.8 m, 14 m, and 21.5 m from the interaction point. Layout of the muon system was designed to best measure the deflection of tracks by toroids. The arrangement of the muon chambers is shown in Figure 2.15. There are two precision chambers (Monitored Drift Tubes - MDT and Cathode Strip Chambers - CSC) and two trigger chambers (Resistive Plate Chambers - RPC and Thin Gap Chambers - TGC) that comprise the muon system.

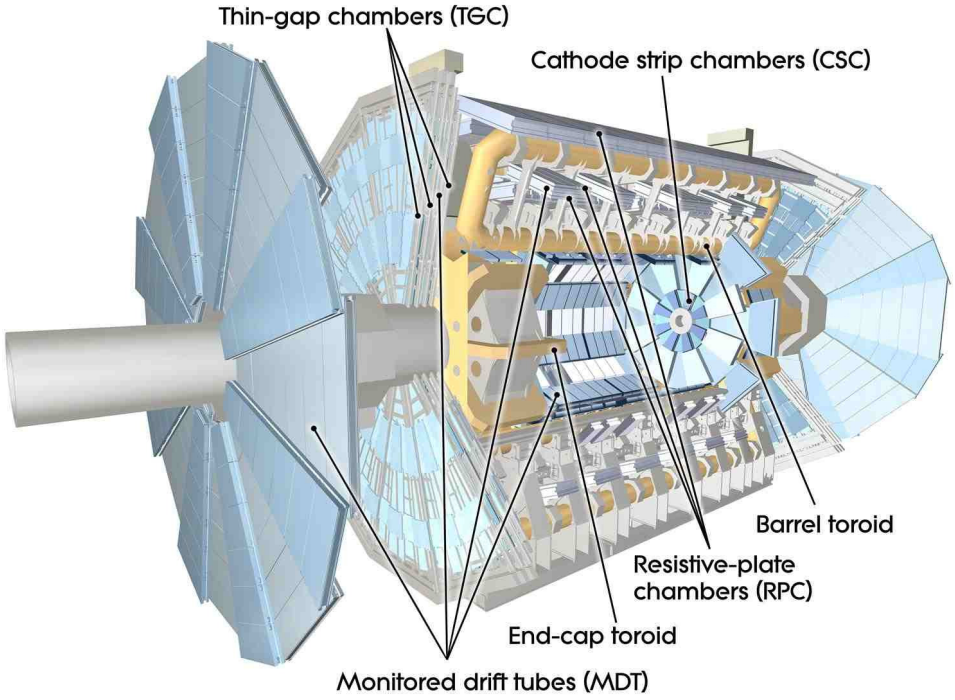


Figure 2.15: The Muon System consists of four different subsystems. The Monitored Drift Tubes (MDT) provide precision measurements for the direction of the tracks. Cathode Strip Chambers (CSC) provide directional measurements at large η. Resistive Plate Chambers (RPC) are the central trigger system and Thin-Gap Chambers (TGC) provide the trigger at large η [21].

The chambers are oriented such that particles which originate at the interaction point traverse three chambers around the beam axis. The design measures the curvature of the

tracks in three positions. In the endcaps, where the toroid cryostat prevents chambers from being placed inside the magnetic field, the muon momentum is measured from the difference in entry and exit angle of the magnet. The chambers are described in more detail below.

Monitored Drift Tubes (MDT)

The primary precision measurement in the bending direction for the muon system is done using a drift tube technology: the MDT chambers. It consists of three layers in the barrel and endcaps for $|\eta| < 2.0$, and two layers for $2.0 < |\eta| < 2.7$. These chambers consist of three to eight layers of drift tubes, operated at an absolute pressure of 3 bar, which achieve an average resolution of 80 μm per tube, or 35 μm per chamber.

Cathode Strip Chambers (CSC)

At large pseudorapidity ($2.0 < |\eta| < 2.7$), the third chamber of the system is the CSC. The CSCs with higher granularity are used in the innermost plane in the region. This detector technology is a multiwire proportional chamber with cathode strip readouts. There are precision coordinates at end of detector which contains 70,000 channels. The alignment is done by optically monitoring deviations from straight lines (large sections) and by tracks (small sections and cross checks) They provide 40 μm resolution per station in the bending plane (yz-plane) and 5 mm in the transverse plane.

Resistive Plate Chambers (RPC)

The muon trigger system covering the central region ($|\eta| < 1.05$) consists of RPCs. This technology utilizes bakelite plates with thin gas gap and has a 2^{nd} coordinate measurement in central region with 380,000 channels yielding 1.5 ns of time resolution.

Thin-Gap Chambers (TGC)

For larger pseudorapidities ($1.05 < |\eta| < 2.4$) TGCs provide the trigger, with a small overlap in the $|\eta|$ =1.05 region. This technology provides 2^{nd} coordinate measurement at ends of detector and has almost 440,000 channels. Both chamber types (RPCs and TGCs) deliver signals within 25 ns, thus providing the ability to identify an event with a beam-crossing. The trigger chambers measure both coordinates of the track, one in the bending plane and one in the non-bending plane.

2.3 Triggers and Data Acquisition

Because of the low cross sections of interesting physics processes, the LHC must run at high instantaneous luminosity to maximally produce interesting physics events. To achieve

this, the LHC the interaction rate will be on the order of one GHz. Most of these interactions will produce minimum bias QCD events which are not interesting from a physics perspective. It would be impractical - and impossible - to try to record all these events, therefore ATLAS must decide before the event is recorded which events it will keep and which it will discard. ATLAS deals with this through a system of triggers which reduce the interaction rate from 10^9 to a reasonable 10^2 read out rate.

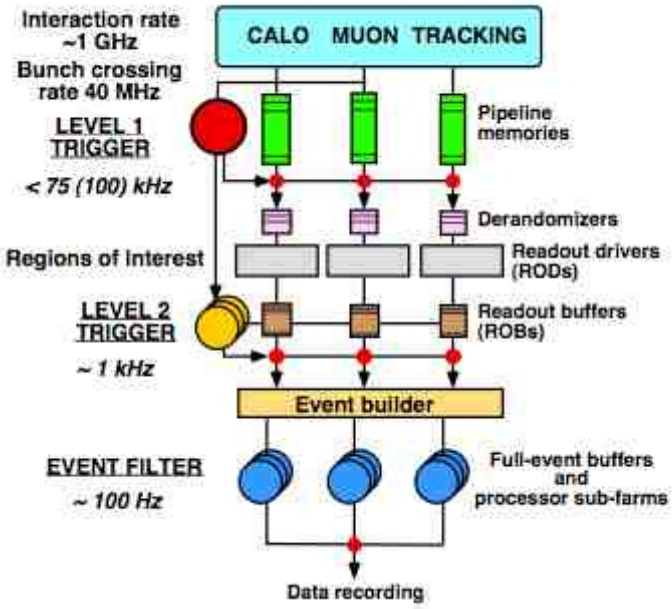


Figure 2.16: The ATLAS Trigger. The readout from the LVL1 triggers on the detectors starts the triggering sequence and is shown at the top. A decision is made to reduce the rate from a GHz to tens of kHz. After that, the LVL2 trigger further reduces the readout, being even more selective as to which events are kept. These events are stored in buffers so that a final decision can be made by the EF trigger, which reads out at a rate of ≈100 Hz [44].

At ATLAS, events of potential physics interest are selected by a three-level trigger system, with a final recording rate of about 200 Hz. The first level (LVL1) is implemented in customized hardware integrated throughout the detectors, and the two high level triggers (LVL2 and Event Filter) are software triggers. Events selected by LVL1 are read out from the front-end electronics systems of the detectors into readout buffers (ROBs). All of the data for the selected bunch crossing from all of the detectors are held in the ROBs either until the event is accepted or rejected by the second level trigger (LVL2). Accepted events have the data from the ROBs transferred by the Data Acquisition (DAQ) system to storage associated with the Event Filter, all of which is discussed in the upcoming subsections. There are over

500 trigger signatures defined for the physics program and not all the trigger signatures write all the data that passes. Each trigger has a defined "prescale" which is a number that denotes how many of the events that pass the trigger are recorded. For example a prescale of 10 records one out of every ten events that passes the requirements of that trigger. A map of ATLAS trigger decision process can be seen in Figure 2.16.

2.3.1 Level 1 Trigger

The first level (LVL1) is implemented in custom-made electronics, reducing the event rate to less than 100 kHz [44]. The initial selection is based on coarse granularity information from a subset of detectors. The calorimeters provide triggers for high p_T electrons, photons, jets, τ, and $\not{E}_T$. The RPCs and TGCs provide the first level trigger for muons. The LVL1 trigger decision is based high p_T objects or on combinations of objects required in coincidence or veto. It also uniquely identifies the bunch crossing of interest with a BCID. The LVL1 latency is a measurement from the time of the proton-proton collision until the trigger decision is available to the front-end electronics and is 2.5 μs. This information is combined in the central trigger processor (CTP) and a decision is made. Because of the initial low luminosity during the beginning of 2010, data was recorded based only on the decision of the LVL1 trigger.

2.3.2 High Level Trigger: Level 2 and Event Filter

The High Level Trigger (HLT) consists of the Level 2 (LVL2) and the Event Filter (EF) trigger [45]. HLT algorithms have access to data from all ATLAS sub-detectors. The LVL2 operates on the "regions of interest" as identified by LVL1. After a LVL2-accept is signaled, the event building collects all detector and trigger data and passes it to the Event Filter (EF) processes. The process of moving data from the ROBs to the EF is called event building. The event building infrastructure adapts the offline event reconstruction algorithms to an online environment. Whereas before event building each event is composed of many fragments, with one fragment in each ROB, after event building the full event is stored in a single memory accessible by an EF processor. Events accepted by the EF are written to mass storage. The LVL2 decision time is about 40 ms, and the EF, because event building must occur, is on the order of a few seconds.

The HLT activated when the LHC had reached a luminosity of 2×10^{30} cm^{-2}s^{-1}, at the end of July 2010. The trigger prescales for low threshold triggers must be adjusted several times over a run period to use the available recording bandwidth because the luminosity drops during the lifetime of an LHC fill. Three key features of the ATLAS HLT are vital to achieve the robustness and flexibility desired for the trigger operation:

- in-time performance monitoring and validation
- flexible configuration for easily adjustable prescales

- the error catching and recovery system for failures in the HLT execution

2.3.3 Data Acquisition

During the data taking period of 2010, ATLAS recorded 45.0 pb^{-1} of stable proton-proton collisions, with an average pp data recording efficiency of 93.6% [46]. From the start of running in March to the end of pp data taking in November the LHC peak luminosity went up by a factor of 100, and the definition of the LVL1 and HLT lines and prescales changed dramatically. After an event has been accepted by the LVL1 trigger, the DAQ system receives and buffers data from electronics located on the detector. These data are received by Readout Drivers (RODs), which act as a buffer for the the event before sending them on to Readout Subsystem (ROS) for a final trigger decision. If the data from an event passes the EF it is moved to the CERN computer center for permanent storage and distribution via the LHC Computing Grid (LCG).

The Detector Control System (DCS) permits operation of the ATLAS hardware with a homogeneous interface to all sub-detector components. These include high- and low- voltage power supplies, and environmental conditions such as temperature and humidity monitoring. Communication between the DAQ and DCS systems is also possible to coordinate data-taking depending on experimental conditions.

Chapter 3

Event Generation

Given the probabilistic nature of particle interactions, properties of a single event can not be uniquely predicted. Instead, a large statistical sample of events for a given process is built up using random numbers to reproduce quantum mechanical properties. The process of predicting the physics processes as well as the interactions with the detector is called Monte Carlo (MC) modeling. MC methods repeatedly model physical processes and vary the initial conditions in phase space within their statistical uncertainties.

Several different types of events are generated, combined, and normalized (using cross sections and luminosity) to make predictions about the normalization and the topology of the signal and backgrounds. These MC programs model physics interactions which produce an initial reaction and then decay and interact in the detector. The interactions include the following processes:

- Hard Scatter: the partons in the pp system collide directly and produce the desired physics process
- Particle Decay: particles produced in the hard scatter decay on time-scales comparable to that of QCD parton showers
- Initial and Final State Radiation (ISR/FSR): accelerated quarks and gluons radiate gluons both before and after the primary interaction via $q \rightarrow qg$ and $g \rightarrow gg$
- Hadronization: quarks via confinement combine and produce hadronic material in the form of jets
- Soft Interactions: initial state parton interactions which are not a part of the hard scatter

3.1 Event Generators

The MC event samples in this analysis were produced with either one of two full simulation generators used in tandem with additional packages that provide more accurate modeling for certain processes. Event generation begins with the hard process. The generation starts with the lowest order cross section using whichever parton distributions have been selected. Then, coherent parton shower algorithms are implemented. This is applies when there is only one scatter in the process, so for additional scattering (underlying events) additional generators are used. These different generators are described, and the specific ATLAS tunes for them are also discussed [47].

3.1.1 Pythia

Pythia is a general purpose program for the generation of high-energy physics events [48]. It provides a description of collisions between two incoming elementary particles at the Leading Order (LO). It returns particle information the at the four-vector level as opposed to distributions. Pythia generates events created from two incoming particles, and the user specifies the results from the interaction. For example two-to-two would be two incoming particles which result in two outgoing particles. The processes it models include hard and soft interactions, parton distributions, initial- and final-state parton showers, multiple interactions, fragmentation and decay. It uses the Parton Shower model to account for the QCD and QED radiations for a number of Standard Model and beyond the Standard Model physics processes. In cases where the physics final states are not understood well enough to give an exact description (such as for multihadronic final states), the program is based on a combination of analytical results and various QCD-based phenomenological models.

Showering evolves hadrons from the initial high momentum scale (Q) of the partons resulting from the collision down to a cut-off value (Q_0) to obtain the final state particles. Color confinement, which causes hadronization, is modeled using the String Model. The ATLAS MC produced with Pythia use the CTEQ6L1 [47] parton distribution functions (PDFs).

3.1.2 Herwig

Herwig, like Pythia, is a general-purpose MC event generator [49]. It also uses the Parton Shower approach for initial- and final-state QCD radiation, (ISR/FSR) including color coherence effects and azimuthal correlations both within and between jets. Some benefits of Herwig in particular are: (1) the matching of first-order matrix elements with parton showers, (2) a more correct treatment of heavy quark decays that Pythia, and (3) a wide range of Beyond the Standard Model processes.

Herwig is most commonly used with another generator called Jimmy. The ATLAS MC produced with Herwig and Jimmy use specific tunes for the MRST LO, CTEQ6L1 and

CTEQ6.6 PDFs. This is collectively titled "AUET1".

3.1.3 Jimmy

Jimmy is a library of routines linked with Herwig [50]. The high luminosity running conditions at the LHC result in several parton interactions during one bunch crossing (an effect known as multiple parton scattering). In addition to the most energetic (hard) two-to-two parton-parton scattering, sometimes there is a second semi-hard 2-to-2 parton-parton scattering that contributes particles to the underlying event. One of the weaknesses of Herwig alone, is that it can not produce these types of events well. Jimmy allows the generation of multiple parton scattering events in hadron-hadron, photon-photon or photon-hadron events. In particular, it is used as a generator of underlying events in events which contain high E_T jets and other hard processes in hadron-hadron collisions.

3.1.4 Alpgen

Alpgen is a matrix element generator is used to calculate leading order (LO) matrix elements for multiple particle final states [51]. The calculation of the LO matrix elements for the selected hard process is performed using the ALPHA algorithm and is extended to QCD interactions. It is interfaced with Herwig and Jimmy to compute the hadronization and multiple parton scattering and is used to describe the following final states: W+N partons and Z+N partons (with N$\leq$5).

Interfacing the Alpgen output to Herwig involves the risk that a theoretically equivalent parton is added during the shower evolution to one already generated in the matrix element calculation (i.e.: double counting). A solution implements a matching procedure for matrix elements and parton showers to remove an event which occurs in both generators. In case of Alpgen, the approach to remove double counted jet configurations is implemented using MLM matching, which is described below.

MLM Matching Scheme

Double counting is a major concern for any MC generation process. In the case of events produced with jets produced by multiple generators (i.e. Alpgenand Herwig), the MLM Matching Scheme is used to avoid double counting [52]. In multi-jet events where there is ambiguity in the multiplicity of the partons. For example, the following events should be considered separately: (1) three jets produced by three partons, and (2) two jets produced by two partons, with a third jet produced by gluon emission. Both events contain three jets, yet the different sources necessitate that they are treated separately.

The MLM Matching Scheme handles generated events in bins of exclusive parton multiplicities (nlp). When the events are given to Herwig, the event′s quarks and gluons (partons) are clustered into jets. The MLM matching is then done to find the jets associated with the

partons. Starting from the hardest parton, the jet which is closest to it in (η, ϕ) is selected. The distance to from the parton to the jet is defined as $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ If the distance between the parton and the jet center is smaller than a defined minimum, R_{min}, the parton and the jet match. The matched jet is removed from the list of jets, and matching for subsequent partons is performed. The event is fully matched if each parton has a matched jet. For events which satisfy matching, no extra jet may be present. If the event contains a jet cluster which is not matched to one of the *nlp* partons, it is rejected. Higher multiplicity bins (usually above 5) are produced inclusively. In this case, the algorithm is modified to require that each of the *nlp* partons is matched to a jet, making no such requirement on the additional jets in the event.

3.1.5 MC@NLO and Powheg

The drawback of the previously discussed generators is that they only compute the cross sections to LO. In many cases, exact next-to-leading-order (NLO) calculations are needed for a sufficiently accurate description of the physical process. One problem that arises from trying to merge NLO calculations with parton shower simulations (like Pythia and Herwig) is over counting because most shower based MC programs already implement an approximate NLO calculation. MC@NLO is a NLO event generator, which takes into account the potential over counting by subtracting out the approximation from the exact solution [53]. Additionally, it calculates the matrix element with additional single parton corrections to the hard process. These corrections give rise to weighted events, which are calculated into the expected yields. This weighting, however, does not have to be positive (since a subtraction occurs). To overcome the implication of an event with a negative cross section an additional generator, Powheg, is used in parallel.

Powheg (Positive Weight Hardest Emission Generator) is also a NLO method [54]. It overcomes the problem in MC@NLO of negative weights by ordering the generated events by p_T. Powheg can be interfaced to both shower and non-shower based generators, however, it should be used with generators that p_T order generated events. MC@NLO is used with Powheg to generate a number of background samples, notably $t\bar{t}$, and other diboson processes used in this analysis. It uses the PDF set CTEQ6.6 [47].

3.1.6 Summary

Table 3.1 shows a summary of the event generators discussed.

3.2 Detector Simulation

The events produced by the generators are then processed using detailed detector simulations. First, a record of the interactions of the particles in the detector are produced. The

	Physics	Precision	Hard Process	UE	ISR	FSR	Hadronization	Features
PYTHIA	SM, BSM	LO	X	color-connected	backward shower w/MPI	p_T ordered	string model	General purpose
HERWIG	SM, BSM	LO	X	from JIMMY	angle ordered		cluster	General purpose
JIMMY	UE	LO		MPI	from HERWIG			used for UE
ALPGEN	V+jets	LO	X	from HERWIG or PYTHIA				Final states with large multiplicities
MC@NLO	VV, $t\bar{t}$	NLO	X	from HERWIG				NLO ME yields neg weights
POWHEG	VV, $t\bar{t}$	NLO	paired with MC@NLO					corrects for neg weights in MC@NLO

Table 3.1: This provides a summary of the generator properties discussed in the previous section. UE stands for Underlying Event, ISR and FSR are the Initial and Final State Radiation respectively. LO and NLO represent the precision of the generator to either Leading Order and Next-to-Leading Order. V for vector bosons (W and Z), MPI stands for Multiple Parton Interactions, and finally ME stands for Matrix Element.

following interactions must be simulated:

- Interactions with the magnetic field
- Ionization of the generated particles
- Showering in the detector (mostly interactions of the particles with the detector through energy deposition)

The events are subsequently digitized and transformed into Raw Data Objects (RDOs). During the digitization process, the hits produced are translated into the output form similar to what is expected from the readout electronics in the actual experiment. The information obtained is used then to reconstruct the physics objects like tracks in the simulated inner detector and energy deposition in the simulated calorimeters. The final outputs are RDOs that resemble the real detector data.

The detector simulation is based on the GEANT4 toolkit which is used for simulating the passage of particles through matter [55]. It has been designed to handle the complex geometries of the ATLAS detector and includes tracking and calorimetry energy deposition. The detector description also includes misalignments in the inner detector and calorimeter and modeling of read-out electronics, cables, and other dead material. The physics processes offered cover a comprehensive energy range into the TeV scale.

3.3 Monte Carlo Corrections

Despite the best efforts, the MC remains an imperfect estimation of the data because of a number of effects which are not well modeled. Some of these include the underlying physics, imperfections in the detector, or multiple events in one read out. The corrections which are common in analyses are discussed in greater detail.

3.3.1 Trigger and ID efficiencies

The electron trigger and reconstruction efficiencies are measured using the standard tag-and-probe (described below) method in Z events. Background contamination in the Z event data is determined using a sideband method (also described below) and subtracted when calculating the ID efficiency. The trigger scale factor is defined in Equation 3.1 and ID scale factor is defined in Equation 3.2.

$$f_{trig} = \frac{\text{Probe electrons passing full selection and trigger}}{\text{Probe electrons passing full selection}} \tag{3.1}$$

$$f_{ID} = \frac{\text{Probe electrons passing full ID}}{\text{All preselected loose probe electrons}} \tag{3.2}$$

The results for f_{trig} are:

- Trigger: L1_EM14, $f_{trig} = 99.48\% \pm 0.52 \pm 1.0$
- Trigger: EF_e15_medium, $f_{trig} = 99.11\% \pm 0.10 \pm 1.0$

For f_{trig} the first uncertainty is the statistical uncertainty, and the second is the systematic uncertainty. Since both efficiencies are consistent with unity, no correction for the trigger efficiency is implemented. The ID efficiency, however, is η dependent and is listed in Table 3.2. The efficiencies are determined specifically for electrons used in this analysis, which are discussed in greater detail in Section 4.4.1.

The Tag-And-Probe Method for the Data vs. MC Corrections Process

Z bosons can decay decay into an electron and a positron (e^+e^-). This process produces a unique clean signal with electrons that can be reconstructed to reproduce an invariant mass of the parent particle. The method to find a signal (an electron) is defined as follows:

- Find a well reconstructed electron which satisfies a number of quality requirements - this is defined as the "tag"
- Look for another cluster in the electromagnetic calorimeter - this is defined as the "probe"

η range	Scale Factor (%)
(-2.47, -2.01)	93.5%
(-2.01, -1.52)	98.7%
(-1.37, -0.8)	97.0%
(-0.8, 0)	97.3%
(0, 0.8)	97.3%
(0.8, 1.37)	97.0%
(1.52, 2.01)	99.7%
(2.01, 2.47)	96.2%

Table 3.2: The extended ID efficiency MC scale factors. The requirements on the electron are Medium quality with a hit in the b-layer and an isolation requirement. These specifications are given in greater detail in Section 4.4.1. The uncertainty is 3.7%

The efficiency is then the ratio of the number of probe electrons that pass the quality requirements set for the electron under study (the tag) to the number of all the probe clusters found.

The Sideband Method

The sideband method defines a region using the invariant mass of the Z peak and the two regions on either side. The Z peak is at 91 GeV, so the range for example could be from 80 GeV to 105 GeV. A measurement of the background is determined by the regions outside of this range after interpolation into the signal region. It is then subtracted from the Z peak to increase the purity of electrons in this region.

3.3.2 Vertex Reweighting

In addition to the trigger and ID efficiency scaling, a weight based on the number of vertices in the event is also applied. The vertex multiplicity (or number of vertices) measured is dependent on the number of interactions that occur per bunch crossing. The number of vertices in MC events does not match the number of vertices in data events, which changes the shapes of some distributions. To account for this, a relative weight is applied to each event based on the vertex multiplicity. This is done to bring the data and simulated samples vertex multiplicity distributions into agreement. The weights in Table 3.3 show the vertex multiplicity scale factor. This is the ratio of the data to the MC vertex multiplicity. It is then applied as an event weight to the MC.

# Vertices	Weight (data/MC)
1	1.6415 ± 0.0071
2	1.1139 ± 0.0040
3	0.8496 ± 0.0035
4	0.7051 ± 0.0040
5	0.6249 ± 0.0056
6	0.6074 ± 0.0094
7	0.5715 ± 0.0169
8	0.6460 ± 0.0383
9	0.6968 ± 0.0838
10	0.8011 ± 0.2126
$\geq$11	0.2295 ± 0.2382

Table 3.3: The Vertex weighting scale factors for the samples described in Section 5. These are the vertex multiplicity ratios (data/MC) applied to MC depending on number of vertices in event. The number of vertices in the simulated events does not match the number of vertices in real data events, so an event weight is applied to bring these into agreement.

3.3.3 Pile-up

Pile-up is a general term describing situations that that can come from several different types of events. The ATLAS subdetectors are sensitive to hits several bunch crossings before and after the bunch crossing that contains the hard scatter event. Thus, multiple events are recorded in a single one. In addition to the effects of multiple bunch crossings, the different types of events that result in pile-up are:

- Minimum bias: the number of minimum bias interactions per bunch crossing depends on luminosity and bunch spacing. The hits of the minimum bias events are overlaid onto the hits from the hard scattering event.
- Cavern background: neutrons can propagate through the ATLAS cavern producing a neutron-photon gas. This gas produces a constant background, called cavern background, of low energy electrons and protons from spallation.
- Beam gas: this includes residual hydrogen, oxygen and carbon gases in the ATLAS beam pipe that may interact with the beam at any place along the beam pipe.
- Beam halo: this is the background resulting from interactions between the beam and up- stream accelerator elements.

Pile-up is simulated with the Athena-based pile-up application during the digitization step. All of the detector and electronic effects are taken into account during the pile-up event merging for the samples described in Section 5.

Chapter 4

Reconstruction and Identification

This chapter discusses how the data collected from the subdetectors are interpreted as physics objects (particles) using a series of algorithms. Recorded raw data are stored in Raw Data Output "RDO" files and are then processed to interpret a physical process in a physics analysis. This reconstruction is done using both online and offline algorithms. These algorithms are developed for the ATLAS Athena framework. Reconstructed objects discussed here are tracks, vertices, jets, electrons, muons and missing transverse energy ($\not\!\!E_T$).

After these physics objects are reconstructed, the output is saved in Event Summary Data (ESD) files. These files require the Athena framework to analyze. Further processing saves these ESD files as Analysis Object Data (AODs). End user analyses use Derived n-th Physics Data (DnPDs) which are flat ntuples produced from AOD files. ATLAS uses a modular Event Data Model (EDM) which is implemented offline in the Athena framework. The goal is to ensure maintainability during the long lifetime of the ATLAS experiment.

4.1 Tracks

The first step in the object identification process is reconstructing tracks. Tracks are reconstructed from the pixel, SemiConductor Tracker (SCT), and Transition Radiation Tracker (TRT) subdetectors (See Section 2). Charged particles deposit energy into cells and tubes of the inner detector and these depositions, along with the known geometry, are used to construct points in 2D and 3D corresponding to points along charged particle trajectories. The two track finding algorithms generally used are a track fitting global χ^2 and a Kalman Filter technique [56].

- GlobalChi2Fitter: a common and robust fitting technique which utilizes a χ^2 minimization where material effects are additional fitting parameters weighted by their expected variance due to their stochastic behavior. The fit parameters in the model are: the track parameters at the vertex (impact parameter, direction and momentum), the scattering angles (two for each scattering plane), and the energy losses.

- KalmanFilter (KF) Algorithm: a straight-forward implementation of Kalman filter technique which combines forward filtering, backward smoothing and an outlier rejection, and is also used for vertex reconstruction.

Before tracks can be found or fitted, the raw data from the pixel and SCT detectors are converted into clusters, and the TRT raw timing information are turned into calibrated drift circles. The default strategy to reconstruct tracks starts inside and works out. Track candidates are identified using space points in the three pixel layers and the first SCT layer. They are combined to make track seeds, which are extended throughout the SCT. Since the magnetic field is not perfectly homogenous and the track particles lose energy by traversing the detector material, a track is not an exact helix.

Normally, the track parametrization is chosen in a frame, where the z-axis is parallel to the magnetic field. There are five track parameters used in ATLAS:

- Charge over momentum magnitude: $\frac{q}{p}$
- Transverse impact parameter: d_0, which is the distance of the closest approach of the track projected to the interaction point into the xy-plane
- Longitudinal impact parameter: z_0, which is the z value of the point of closest approach determined above
- Azimuthal angle: ϕ_0 which is the angle of the momentum at the point of closest approach determined above
- Polar angle: θ is used to calculate the pseudorapidity, η

These parameters define the track candidates. These candidates are fitted using the algorithms described above. The outlier clusters are removed and fake tracks are rejected. Then, the selected tracks are extended into the TRT to associate drift circle information. Finally, the extended tracks are refitted with the full information of the pixel, SCT and TRT detectors. The quality of the refitted tracks is compared to the silicon-only track candidates. Hits on track extensions resulting in bad fits are labelled as outliers. Once the tracks are reconstructed, tracks that point to a common origin are then used to reconstruct vertices which are defined in the following section.

The reconstruction of tracks is closely coupled with the reconstruction of vertices. One technique which combines track and vertex finding is the Kalman Filter [57], which is described in greater detail in Section 4.2.

4.2 Vertices

When the reconstruction of tracks is finished, vertices, which represent the position where particles are produced or decay, can be found and fitted. The reconstruction of primary

vertices, where interactions between particles in the beam occur are very important for physics analyses. The EDM of ATLAS includes a core of algorithms which provide basic vertex reconstruction and extensions to them. All primary vertex finders use tracks which originate from the beam crossing area. The resolution of the primary vertex is 11μm in the rϕ-plane and 45 μm in the z-plane, and the secondary vertex resolution is 50 μm in the rϕ-plane and 70 μm in the z-plane.

The reconstruction of the primary vertex is subdivided into two stages: primary vertex finding (association of reconstructed tracks to a particular vertex candidate) and vertex fitting (reconstruction of actual vertex position and quality). These two stages can be either combined in a "finding-through-fitting" approach, or done separately in a "finding-after-fitting" approach.

In the "finding-after-fitting" approach, the reconstruction of vertices starts with the preselection of tracks compatible with the expected bunch crossing region. Tracks are then ordered according to the value of their z-impact parameter. Track clusters (tracks that point to a common origin) are then regarded as primary vertex candidates. Each candidate is then reconstructed using one of the vertex fitters and cleaned by iterating over outlying tracks. This process continues until there are no more outlying tracks.

In contrast, the "finding-through-fitting" method deals more cleanly with outlying tracks. Reconstruction starts again with a preselection of tracks originating from the bunch crossing. A single vertex seed is created out of this set of tracks and a vertex candidate is reconstructed. Outliers are then used to create a new vertex seed. With each iteration, a new vertex candidate is formed, and the fit of each vertex is done simultaneously with each track down-weighted with respect to the its vertex.

Kalman Filter

An example of a "finding-through-fitting" algorithm is a Kalman Filter technique. The purpose of the this technique is to use measurements of tracks observed over time, containing noise and other fluctuations, to produce values for vertices that are to be closer to the true values. It starts with an arbitrary initial vertex position and covariance matrix estimation. The result is a new vertex at the distance of closest approach. The highest p_T track is identified originating from the vertex. A second track is then added and a new vertex estimation is found. Finally, other tracks are included which originate from the final vertex position. The tracks are then smoothed using constraints that come from the vertex that they parameterize. This method provides information on the covariance matrix, χ^2, number of degrees of freedom from the vertex fit, and the tracks used for the fit. Compared to minimum bias events, signal events generally have higher track multiplicity and p_T. This distinguishes a signal vertex or "primary" vertex, from other vertices from minimum bias events. Proper determination of the primary vertex is crucial to the measurement of other objects in the reconstruction stream and the magnitude of $\not{E}_T$.

4.3 Jets

Quarks and gluons are the resulting particles from the strong interaction reactions. Color confinement dictates that they can not exist freely and must group together to form hadrons of neutral color. When a quark is separated from another hadron, new quark/anti-quark pairs appear and form mesons (typically pions). While this process, called hadronization, occurs and if the quarks and gluons are at a high p_T, "jets" of particles are formed. A jet is a narrow cone of hadrons originating from a quark or gluon. Measuring its properties can give insight as to the property of the original particle.

Particles in jets leave tracks in the inner detector and shower in the calorimeter, depositing most of the energy in the hadronic layer. Jets can be reconstructed using a variety of different algorithms, the most basic of which is a cone based algorithm. The algorithm starts with the energy deposition of a high p_T object in the hadronic calorimeter. It then defines the jet in terms of a "cone" in ΔR of energy deposition around the central cluster. There are two main problems with this basic definition of a jet: (1) adding of a soft or a collinear parton changes the jet object (not collinear safe) and (2) the output of the jet algorithm is not stable against addition of soft particles (not infrared safe). ATLAS implements a jet algorithm that is both infrared and collinear safe and is described below.

The jets are reconstructed using an anti-k$_T$ algorithm. The anti-k$_T$ algorithm uses the relationships listed in Equations 4.1 and 4.2

$$\Delta R_{i,j} = \sqrt{(\Delta\eta_{i,j})^2 + (\Delta\phi_{i,j})^2} \tag{4.1}$$

$$d_{i,j} = min\left(\frac{1}{k_{T,i}^2}\frac{1}{k_{T,j}^2}\right)\frac{\Delta R_{i,j}^2}{R^2}, \quad d_{B,i} = \frac{1}{k_{T,i}^2} \tag{4.2}$$

where i and j are two objects from the calorimeter. The algorithm reconstructs jets in the following way [58]:

- Recombination starts from highest p_T (hardest) objects
- Finds the smaller of $d_{i,j}$ and $d_{B,i}$ and combines them to form a jet
- If $d_{B,i}$, call i a jet and remove from list of objects
- Repeat until there are no objects left

This method of constructing jets has the consequence of soft particles (low p_T) tending to cluster with hard ones before they cluster among themselves just based on the hard cluster start. This also results in a regular boundary shape which means that the shape of the jets is largely unaffected by soft radiation. The result of using the anti-k$_T$ algorithm to form jets from energy depositions can be seen in Figure 4.1.

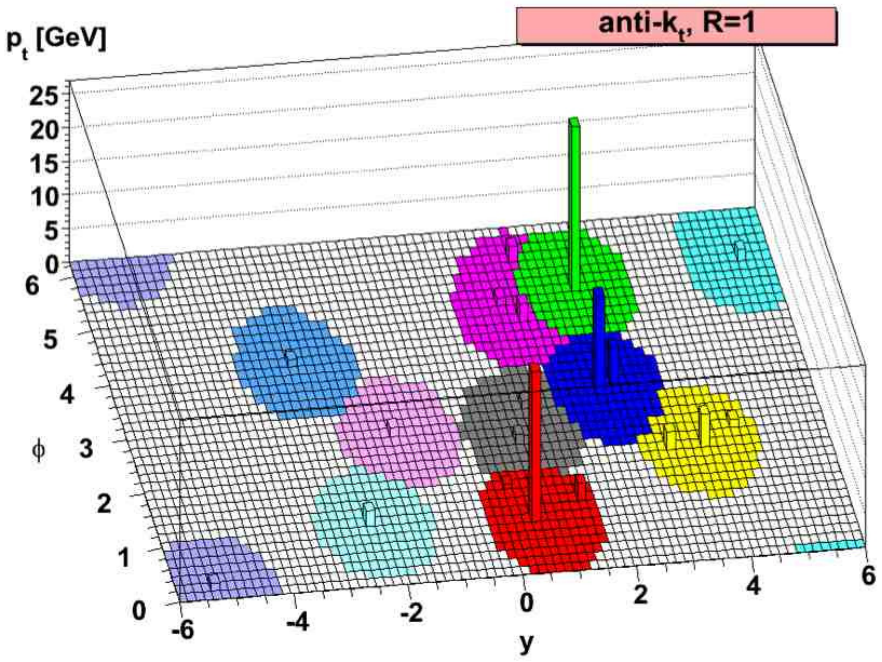


Figure 4.1: A simulation of the results of the anti-k$_T$ jet algorithm used to reconstruct jets. The spike in each cluster is the initial high p_T object on which the jet was constructed. Each color surrounding a spike represents a different jet found by the anti-k$_T$ algorithm. This specific example used an R parameter value of 1, which is larger than the values for ATLAS which are 0.4 or 0.6.

Reconstructed jets are calibrated to the electromagnetic (EM) scale. This is the energy scale which uses calibration constants for electrons and photons from the test beam, but is not correct for hadrons [59]. This energy scale also does not correct for detector effects including:

- Energy losses in inactive regions of the detector (dead material)
- Particles whose energy deposition is not contained by the calorimeter (leakage)
- Inefficiencies in calorimeter clustering and jet reconstruction

The goal of the jet energy scale calibration is to correct the energy and momentum of the jets measured in the calorimeter for these effects, using the kinematics of the corresponding Monte Carlo simulated jets. Jets calibrated in the EM energy scale are first reconstructed from using EM scale objects, then corrected to their true jet energy. The choice of jet energy scale calibration for the first ATLAS data is a jet-by-jet correction applied as a function of the jet transverse momentum and pseudorapidity, η.

4.4 Electrons and Photons

The electron and photon reconstruction and identification algorithms used in ATLAS are designed to achieve both a large background rejection and a high and uniform efficiency over the full acceptance of the inner detector and electromagnetic (EM) calorimeter for transverse energies (E_T) above 20 GeV. Electrons from the initial interaction should be isolated from other events in the detector and need to be discriminated from hadrons in QCD jets, from photon conversions, and from non-isolated electrons from heavy flavor decays. The isolation variable is defined as the energy in a cone around the central energy cluster, with the energy deposition removed. The reconstruction of isolation is shown in Figure 4.2. Photons also need to be distinguished between direct photon production and photons from hadronic decays.

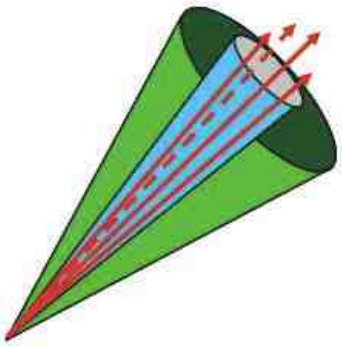


Figure 4.2: Construction of the Isolation variable. The isolation variable is defined as the energy in a cone (shown in green) around the central energy cluster (blue with red tracks). The central energy cluster is removed and the isolation energy can be calculated either as an absolute energy, or when taken as a ratio with the central cluster, a relative energy.

The ATLAS EM calorimeter has a fine segmentation in both the transverse and longitudinal directions of the showers. Finely segmented first and second EM calorimeter layers allows excellent e/γ and π_0 (hadronic) discrimination. In front of these two layers is a presampler layer to correct for energy loss in the material before the calorimeter and behind them is a back layer which provides corrections for highly energetic EM showers. There are separate reconstruction algorithms for photons and electrons and they will be discussed separately in Sections 4.4.1 and 4.4.2.

4.4.1 Electrons

Electron reconstruction begins with a seed cluster of energy deposited in a grid satisfying $E_T > 2.5$ GeV with the granularity in the second layer of the LAr EM calorimeter. This cluster is matched to tracks, which are extrapolated to the middle EM calorimeter layer and

have a p_T >0.5 GeV, requiring a match within a window of $\Delta\eta \times \Delta\phi = 0.05 \times 0.1$ centered on the cluster. The closest-matched track to this layer′s cluster barycenter is assigned to the electron candidate [60]. The electron author variable differentiates the the basis of reconstruction for the electromagnetic cluster. Electrons in this analysis are required to have an author of 1 or 3 which means the reconstruction is based on clustering algorithms. Electrons can also be defined by track based searching methods, however, those are not considered. There are fiducial requirements based on the acceptances of the detector. Electrons must satisfy: $|\eta|$ <2.47 but are vetoed in the crack region: 1.37< $|\eta|$ <1.52.

Electron reconstruction algorithm: isEM

The ATLAS reconstruction algorithm also has a 32-bit mask which is used to define the three different grades of purity of electrons, called **isEM**. The method requires a set of discriminating variables measured from the calorimeters and the inner detector. The response of the calorimeter is a function of pesudorapidity, η. The following η regions have been defined based on the topology of the calorimeter: $0 < \eta < 0.8$, $0.8 < \eta < 1.37$, $1.52 < \eta < 2.01$, and $2.01 < \eta < 2.47$ (the -η regions are considered separately from the +η regions). The calorimeter based measurements that are used to distinguish electrons from jets include:

- Hadronic leakage: the fraction of the electron's energy deposited in the first layer of the hadronic calorimeter. (Electrons deposit most of their energy in the EM calorimeter)
- Lateral shower shape: The ratio deposited in a cluster of 3×7 cells to 7×7 around the cell with the highest energy deposited. (Electrons have small clusters generally)
- Shower width: RMS of η of cells in a cluster weighted by the energy in the second layer of the EM calorimeter. (Electrons shower in a narrow region)

Jets with single or multiple hadrons pose as the main source of fake electrons. The first sampling layer of the electromagnetic calorimeter, with its very fine granularity in pseudorapidity, can be used to detect substructures within a shower and thus identify π_0 and pair-converted photons. To reject these fakes, **isEM** uses the following variables:

- Energy difference: between the highest energy cell and the next highest energy cell. (Electrons typically deposit most of their energy in one cell, whereas π_0 and photons deposit in two)
- Shower shape core: fraction of energy deposited in the core part of the cluster in the first sampling layer of the electromagnetic calorimeter.
- Weighted shower width in first layer: RMS of η of cells in a cluster weighted by the energy in the first layer of the EM calorimeter. (Electrons shower in a narrow region)

The inner detector is also used to discriminate electrons from photon conversions. The **isEM** variables associated with tracking are described below:

- Hits in the vertexing (b) layer: Discriminates against photon conversions and hadrons originating from c and b quarks.
- Hits in each layer of the inner detector: number of hits in pixel, SCT and TRT detectors
- Impact parameter: distance of the track perigee from the beam axis. This selection criterion discriminates against the particles from the secondary vertices.
- High Threshold TRT hits: the percentage of high-threshold hits in the TRT is used as a discriminating variable.
- Track/Cluster matching: the $\Delta\eta$ and $\Delta\phi$ of the tracks and the clusters must be small
- E_T/p_T: the ratio of cluster energy to track momentum.

These criteria are used to define three different classifications of electrons, which balance a need for purity with high efficiency. Not all electrons must pass all the requirements listed above. Since **isEM** is a bit-mask, each requirement returns either a pass or fail, so each electron grade has a different set of **isEM** variables it must pass, with higher grades requiring more selection criteria passed. These three grades of electrons with the **isEM** variables required are defined below:

- **Robust Loose:** the basic selection, uses as discriminant variables the EM shower shape (lateral shower containment and shower width) from the second layer of the LAr EM calorimeter and energy leakage into the hadronic calorimeters. This selection has the highest efficiency but the lowest background rejection.
- **Robust Medium:** uses, in addition to the requirements from Robust Loose, track quality variables (hits in silicon layers and transverse impact parameters), track-cluster matching ($\Delta\eta$ between cluster and track extrapolated to first layer), and additional hadronic rejection using shape information on energy deposition in first layer of EM calorimeter.
- **Robuster Tight:** additionally uses requirements on E/p, the number of hits in the TRT, and the ratio of high-threshold hits to the total number of hits in the TRT to discriminate against backgrounds. Electrons from conversions are rejected by requiring a hit in the first layer of the pixel detector (b-layer). Also, a conversion-flagging algorithm is used. The impact-parameter requirements in the medium selection are further tightened.

4.4.2 Photons

Photons and electrons both deposit there energy primarily in the electromagnetic calorimeter. Photons, however, are distinguished from electrons since there is no reconstructed track matched to the cluster. Photons use a reconstruction algorithm called **PhotonIDTool**, and is discussed below.

There are two types of photons that have to be considered in the detector: ones that convert into e^+e^-, and ones that do not. If there is no reconstructed track matched to an electromagnetic cluster, the object is classified as an unconverted photon candidate. However, if there are reconstructed tracks matched to the cluster that come from a displaced (conversion) vertex (as opposed to the interaction point), it then is classified as a converted photon candidate. This case is also discussed in greater detail below.

Photon reconstruction algorithm: PhotonIDTool

Photons have a 16 bit mask definition called **PhotonIDTool**, which is segmented into the same η regions as **isEM** in the EM calorimeter. The variables used in **PhotonIDTool** are essentially the same as **isEM** for electrons. The main difference between **PhotonIDTool** and **isEM** is that there are no tracking specific requirements in **PhotonIDTool**. It is used to make up the different classifications of photons (with only slight differences between the requirements for converted and unconverted photons). There are five different loose photons depending on the specific needs of the analysis and one tight definition. The tight definition uses energy deposition shape information in the first and second sampling of the calorimeter, which provides good differentiation between photons and π_0.

Converted Photons

Converted photon candidates are reconstructed if the conversion radius is below 800 mm (within the TRT). High efficiency in this range can only be achieved if double- and single-track conversions are taken into account. Single track conversions are those for which only one matching track is reconstructed. There are also no hits in the first pixel layer (b-layer). Photon conversions are assigned an author value of 16. A large fraction of single-track conversion candidates corresponds to a track reconstructed only in the TRT detector. Such tracks have no associated silicon detector information and have therefore poorer momentum resolution and η information.

Electron/Photon Ambiguity

There is inherent ambiguity between electrons and photons since they both appear as energy deposition in the EM calorimeter and are distinguished mainly by track-based information. There is an EM Ambiguity tool which addresses this issue. This tool allows for reevaluation of an object which was identified as a photon to be associated with what was

once classified as a bad track. This is then reclassified as an electron. The tool takes the calorimeter, the vertex, and the tracking information and defines the degree of "ambiguity" (or probability of misidentification) as loose or standard. The algorithm uses the following criteria:

- The presence of a TRT-only track that fails both standard and loose ambiguity
- The presence of a TRT-only track that fails standard but passes loose ambiguity
- Passes criteria that either has a bad or no track particle
- Passes standard conversion vertex ambiguity criteria

4.5 Muons

The muon detection system of the ATLAS detector is characterized by two high precision tracking systems, discussed in detail in Chapter 2, namely the Inner Detector and the Muon Spectrometer. The calorimeter systems ensure a hadron are filtered leaving muons with energy above 3 GeV.

In order to combine the muon tracks reconstructed in the Inner Detector and the Muon Spectrometer the muon identification software packages (*Muonboy*, STACO, MuID, MuTag) have been developed. The purpose of these packages is to associate segments and tracks found in the Muon Spectrometer with the corresponding Inner Detector track and calorimeter information in order to identify muons at their production vertex with optimum parameter resolution. A brief description of the principles of the different algorithms is listed below [61, 62]:

- *Muonboy* employs a standalone reconstruction. Muon tracks are reconstructed in the Muon Spectrometer only and the tracks are extrapolated to the beam line.
- STACO and MuID use both tracking systems to form a combined reconstruction. Muon tracks reconstructed in the Muon Spectrometer by *Muonboy*, or similar programs, are combined with the tracks reconstructed in the Inner Detector (ID) using a statistical method.
- MuTag focuses on tagged muons. Tagged muons are tracks reconstructed in the Inner Detector and extrapolated to the Muon Spectrometer by looking for nearby hits.

Muonboy, STACO, MuID, and MuTag are discussed in greater detail in the following sections.

Muonboy

The *Muonboy* algorithm uses pattern recognition to identify regions of activity in the muon system through the triggers (RPC and TGC). A region of activity is defined in η and ϕ where there is at least one RPC or TGC hit. The region is approximately $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$. It then reconstructs track segments in each muon chamber in the regions of activity. The track segments in each muon chamber are then combined to form three-dimensional muon track candidates in the magnetic field of the toroids. Then using the individual hit information, a global fit of the muon track candidates is done through the full muon system.

STACO and MuID

Both the MuID and STACO algorithms were developed to combine track information from the Muon Spectrometer and the Inner Detector to form what is referred to as a combined muon. They employ track "finding-and-fitting" algorithms independently and can act as cross checks to the ATLAS muon reconstruction.

The STACO algorithm performs a statistical combination of these two independent measurements from *Muonboy* and the ID tracks using the parameters of the reconstructed tracks and their covariance matrices. This combination improves the momentum resolution over a wide range of transverse momentum ($6 < p_T < 200$ GeV). It also allows for the rejection of muons from secondary interactions and the ones from π/K decays in flight. In general the ID dominates in the range of $p_T < 30$ GeV and the Muon Spectrometer in the region above 200 GeV.

The MuID algorithm is also a combined reconstruction algorithm. MuID identifies ID tracks as muons and provides the best possible muon parameter resolution in the vicinity of the production vertex. Muons reconstruction starts in the muon spectrometers using a reconstruction algorithm called MOORE (similar in principle to *Muonboy*). They are then propagated through the magnetic fields and the calorimeters and back to the vertex region. They are then matched to tracks using a χ^2 method. A combined fit is performed at the vertex for all combinations above a given χ^2 probability. Hits found in the two subdetectors with separate standalone algorithms are then used to combine tracks (MuID Combined mode). In case of a good combined fit, all matches to the ID are kept as identified muons [63].

MuTag

The MuTag algorithm has been developed to tag low p_T muons. The efficiency of the muon reconstruction in the Muon Spectrometers decreases very rapidly with decreasing p_T. One reason is that accurate tracking of low p_T muons in an inhomogeneous magnetic field is difficult. Another is that for muons with low p_T, the energy lost traversing the calorimeters becomes comparable to their energy, specially in the barrel region. The principle of MuTag is to start from with ID tracks, extrapolate them to the inner station of the muon system,

and try to match them with a segment reconstructed in these stations not yet associated with a combined track.

4.6 Missing Energy

The total transverse momentum (momentum in the xy-plane) is expected to be zero because at a first approximation there is no momentum in the transverse plane from the initial interaction. Since neutrinos can not be detected in the detector, their presence is inferred as an imbalance in the measured total $\boldsymbol{E_T}$ deposited. This missing energy can be determined by calculating the imbalance. The important quantities are:

- E_x - sum of energy deposited in the x-direction
- E_y - sum of energy deposited in the y-direction
- E_T - transverse energy in the xy plane. This is calculated by: $\sqrt{(E_x)^2 + (E_y)^2}$
- E_T^{sum} - Scalar Sum of transverse energy in calorimeters

At a first approximation the $\not{E}_T$ calculated first by summing the energies in the calorimeter cells given in Equation 4.3.

$$\boldsymbol{E_T} = \sum_{cells} \boldsymbol{E_T}^i \tag{4.3}$$

Once the energy sum is found in the calorimeter, taking the negative value of this gives us an expression for $\not{E}_T$ as defined in Equation 4.4.

$$\boldsymbol{E_T} = -\boldsymbol{\not{E}_T} \tag{4.4}$$

However other effects, such as contributions from muons and energy loss in the cryostat must be taken into account. The ATLAS $\not{E}_T$ reconstruction, for the purpose of the 2010 physics analyses, takes into account contributions from transverse energy deposits in the calorimeters, corrections for energy loss in the cryostat, and measured muons [64]:

$$\not{E}_{x,y} = \not{E}_{x,y}(calo) + \not{E}_{x,y}(cryo) + \not{E}_{x,y}(muon) \tag{4.5}$$

These three terms are discussed in greater detail in the following sections.

$\not{E}_T$(calo)

Using only information in the calorimeter, the $\not{E}_T$ can be reconstructed up to $|\eta| = 4.5$. The term for the $\not{E}_T$ in the calorimeter term is defined as follows:

$$
\begin{aligned}
\not{E}_x(calo) &= -\sum_{i=1}^{N_{cell}} E_i sin(\theta_i) cos(\phi_i) \\
\not{E}_y(calo) &= -\sum_{i=1}^{N_{cell}} E_i sin(\theta_i) sin(\phi_i) \\
\not{E}_T(calo) &= \sqrt{(\not{E}_x(calo))^2 + (\not{E}_y(calo))^2}
\end{aligned} \tag{4.6}
$$

Because of the high granularity of the calorimeter, noise in the cells must be suppressed. This is done by using an electronic noise suppression technique in the calorimeters which means only using cells in three-dimensional topological clusters (topoclusters). Topoclusters are only reconstructed if they follow a set of thresholds which are multiples (4/2/0) of σ of noise. This means that the seed cell must have a deposited energy of $|E_i| > 4\sigma^2_{noise}$, and neighboring cells are added if they have an energy of $|E_i| > 2\sigma^2_{noise}$. All the neighbors of these accumulated cells are then added as well.

The ATLAS calorimeters respond differently to hadrons and electrons. Therefore a special calibration scheme for hadronic energy deposits to correct for these differences. One way to do so is applying cell-level weights. This method is called a Global cell energy-density weighting calibration scheme (Global calibration or GCW). These weights boost low density signals, which are likely from hadronic activity. The other is to apply weights to the topoclusters individually. This is called Local cluster weighting calibration scheme (Local hadronic calibration or LCW). According to the topocluster topology, a topocluster is classified as electromagnetic or hadronic. Then corrections are applied to account for the differences in calorimeter response. Both of these methods determine the weights or corrections from Monte Carlo simulations and apply the changes by modifying Equation 4.6 [65].

$\not{E}_T$(muon)

The $\not{E}_T$ muon term is calculated from the momenta of muons measured in a smaller range of pseudorapidity than for the calorimeter ($|\eta| < 2.7$). In the region $|\eta| < 2.5$ only combined muons are considered. For large η regions only the muon spectrometer is required. $\not{E}_T(muon)$ is defined as follows:

$$
\not{E}_{x,y}(muon) = -\sum_{muons} E_{x,y}(muon)
$$

Since the acceptance of the Muon Spectrometer is not as large as for the calorimeters, the energy from these muons can be recovered and used to calculate $\not{E}_T$ by using the muon reconstruction algorithms which require only the ID and the calorimeters.

$\not{E}_T$(cryo)

In between the LAr barrel electromagnetic calorimeter and the Tile barrel hadronic calorimeter lies a cryostat about half of an interaction length thick. This layer can lead to significant energy losses in hadronic showers, which necessitates a correction in $\not{E}_T$. The $\not{E}_T$ reconstruction, when it is calibrated with GCW, recovers this loss of energy in the cryostat using the correlation of energies between the last layer of the LAr calorimeter and the first layer of the hadronic calorimeter. This correction is called the "cryostat term when used for jet energy correction. When the calorimeter term is calibrated with the LCW scheme, a corresponding correction is already done at topocluster level so no additional corrections are needed (i.e.,$\not{E}_T(cryo) = 0$).

Chapter 5

Leptoquark analysis

This chapter presents a search for pair production of scalar leptoquarks which decay into an electron/quark pair and a neutrino/quark pair. The analysis is performed in the following manner. A sample which has a similar topology to the LQ final state yet is dominated by backgrounds is defined. Data/MC comparison is done to validate the background modeling. From this selection, three additional control regions are defined to validate further the modeling by enhancing the dominant backgrounds: two W+jets enhanced regions and one $t\bar{t}$ (all of which have negligible signal contribution). Once the background modeling is validated, an optimization procedure (Random Grid Search) is applied to simulated events passing the preselected sample requirements to determine the final selection criteria. The region used in the analysis must have a sufficiently high signal-to-background ratio. These selection requirements are then applied to the data and the probability of seeing a signal+background hypothesis in an all background sample is measured.

5.1 Data and Monte Carlo Samples

5.1.1 Data

The data for this analysis were collected during the 2010 LHC running period. The total integrated luminosity recorded by ATLAS was 45.0±4.9 pb^{-1}. After data quality requirements the total was reduced to 35.2±3.8 pb^{-1}. Data quality information is used to remove runs and luminosity blocks which are marked as "bad" in the detector subsystems. A "bad" status indicates that a section of the subdetector was disabled or functioning abnormally. A combined Good Runs List (GRL) was established for both muon, and electron channels, and the requirements for the good runs are listed below with a description and the actual command to return the required list in parentheses:

- General Requirements: These cover the requirements that the data was from the 2010 7 TeV running and that ATLAS was included in the run (tag data10_7TeV and db

DATA and partition ATLAS)

- The LHC beam energy was at 3.5 TeV (lhc beamenergy 3400-3600)
- The LHC beams were stable (lhc stablebeams T)
- The silicon and muon systems have completed their start-up routine (ready 1)
- The current in the solenoid and toroid magnets is at operational levels - i.e. the magnets are on (mag s $>$ 6000 and mag t $>$ 18000)
- The status of all the detectors is good, which means that the data quality flags are green.

  (dq GLOBAL STATUS, CP TRACKING, CP EG ELECTRON BARREL, CP EG ELECTRON ENDCAP, CP MU MSTACO, CP JET JETB, CP JET JETEA, CP JET JETEC, CP MET METCALO, CP MET METMUON, LUMI, L1CAL, L1MUE, L1MUB, TRELE, TRMUO)
- Run period dependent luminosity block and summary flags equal to good status.

  (periods A and B: LBSUMM#DetStatus-v03-repro04-01 g.

  Other periods: LBSUMM#DetStatus-v03-pass1-analysis-2010X g with X the period (C-J)).

These requirements are applied to the run periods A-I2. The run number range, total integrated luminosity ($\mathcal{L}_{int}$), resulting integrated luminosity ($\mathcal{L}_{calc}$), and electron and muon triggers used are given in Table 5.1.

5.1.2 Monte Carlo

The background physics processes were modeled by MC produced using a combination of generators each with specific strengths for a particular background. W/Z+jet (also referred to as V+jet) events are generated using Alpgen interfaced to Herwig and Jimmy. Exclusive samples with zero to four additional partons and an inclusive samples with five or more partons are computed using Alpgen cross sections so that the sum of the parton samples equals the inclusive cross section calculated to NNLO [66, 67]. The MC generators described here are discussed in greater detail in Section 3. The inclusive NNLO cross section times branching fraction to a single lepton species are:

- $\sigma(W \to (l\nu)) = 10.46 \pm 0.42$ nb
- $\sigma(Z \to (ll)) = 1.069 \pm 0.04$ nb.

Data Period	Run Range	$\mathcal{L}_{int}$ (nb$^-$1)	$\mathcal{L}_{calc}$ (nb$^-$1)	Electron Trigger	Muon Trigger
A	152166-153200	0.4	0.4	L1_EM14	L1_MU10
B	153565-155160	9	8.1	L1_EM14	L1_MU10
C	155288-156682	9.5	8.4	L1_EM14	L1_MU10
D	158045-159224	320	280.	L1_EM14	L1_MU10
E1-E3	160387-160879	512	470.	L1_EM14	L1_MU10
E4-E7	160899-161948	606	530.	EF_e15_med	EF_mu10_MG
F	162347-162882	1980	1800.	EF_e15_med	EF_mu10_MG
G1	165591-165632	780	680.	EF_e15_med	EF_mu10_MG
G2-G6	165703-166383	8290	5200.	EF_e15_med	EF_mu10_MG
H	166466-166964	9300	7200.	EF_e15_med	EF_mu10_MG
I1	167575-167576	4200	3500.	EF_e15_med	EF_mu10_MG
I1-I2	167607-167844	20000	15600.	EF_e15_med	EF_mu10_MG_tight
TOTAL			35200.		

Table 5.1: The data sample used for this analysis shown divided in terms of data periods: A-I2. The run number range, total integrated luminosity ($\mathcal{L}_{int}$), resulting integrated luminosity ($\mathcal{L}_{calc}$), and electron and muon triggers are also listed.

Additional samples for cross checks are generated using Sherpa and Pythia. Simulated $t\bar{t}$ events are generated using MC@NLO and Powheg interfaced to Herwig for parton showering with the NLO cross sections including NNLL corrections [66, 68]. Single top events are generated using MC@NLO [69, 70]. The cross sections for the $t\bar{t}$ sample and single top all include branching fraction, B, which imposes a single lepton requirement in the decay. The cross sections used in this analysis are listed below:

- $\sigma(t\bar{t}) = 164.5^{+11.4}_{-15.7}$ pb
- $\sigma(t_{s-channel}) \times B = 3.9 \pm 0.4$ pb
- $\sigma(t_{t-channel}) \times B = 58. \pm 6$ pb
- $\sigma(t_{tW}) \times B = 13.1 \pm 0.13$ pb

Generated diboson events also include branching fraction, B, which requires at least a single lepton in the decay. They are produced using Herwig. The WZ production has the additional requirement that the invariant mass of the dilepton pair from the Z production: ($M_{ll} >$ 40 GeV). The ZZ production has the requirement that the invariant mass of the dilepton pair: ($M_{ll} > 60$ GeV). The NLO cross sections with the full systematic uncertainties are as follows [71]:

- $\sigma(WW) \times B = 44.9 \pm 2.2$ pb

- $\sigma(WZ) \times B = 18.0 \pm 1.3$ pb
- $\sigma(ZZ) \times B = 5.96 \pm 0.3$ pb

Signal events are generated by Pythia with cross sections and uncertainties determined using the CTEQ 6.6 parton distribution function (PDF) set [17]. Figure 5.1 shows the cross section as a function of LQ mass for $\sqrt{s} = 7$ TeV for both Pythia and CTEQ based calculations.

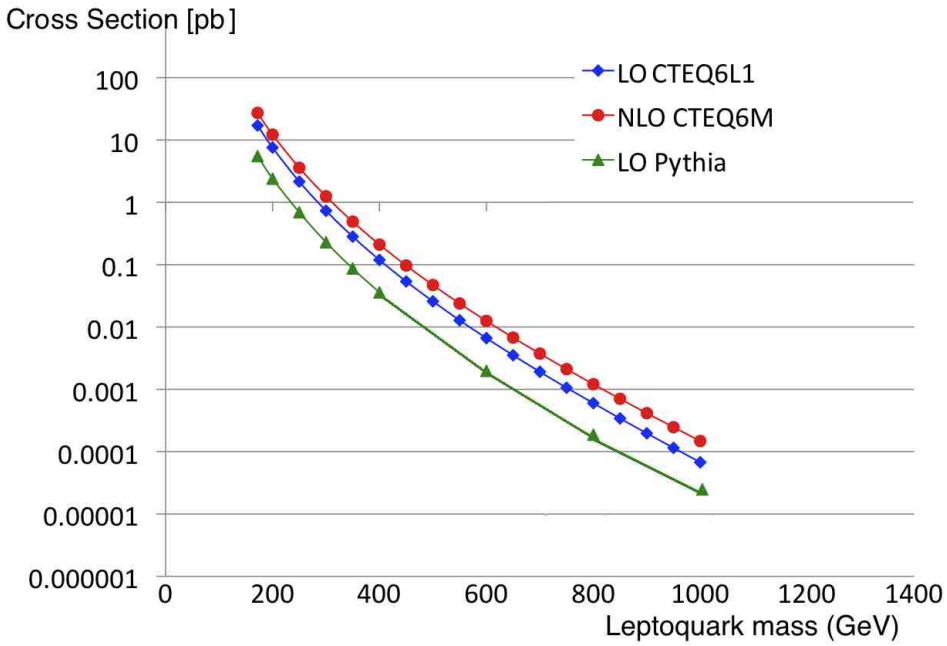


Figure 5.1: LQ Cross section as a function of leptoquark mass. In green is the resulting cross section from the leading order Pythia calculation. The blue dotted line is the LO calculation from CTEQ 6L1 PDF and the red dotted line shows the NLO calculation using CTEQ6M for $\sqrt{s} = 7$ TeV [17].

The unknown $LQ - \ell - q$ coupling (λ) is set to $0.1 \times \sqrt{4\pi\alpha_{EM}}$. This corresponds to an LQ full width of less than 1 MeV, and a negligible decay length. For pair production, changing λ does not affect the production cross section for $\lambda < 0.8$. This relationship is shown in Figure 5.2.

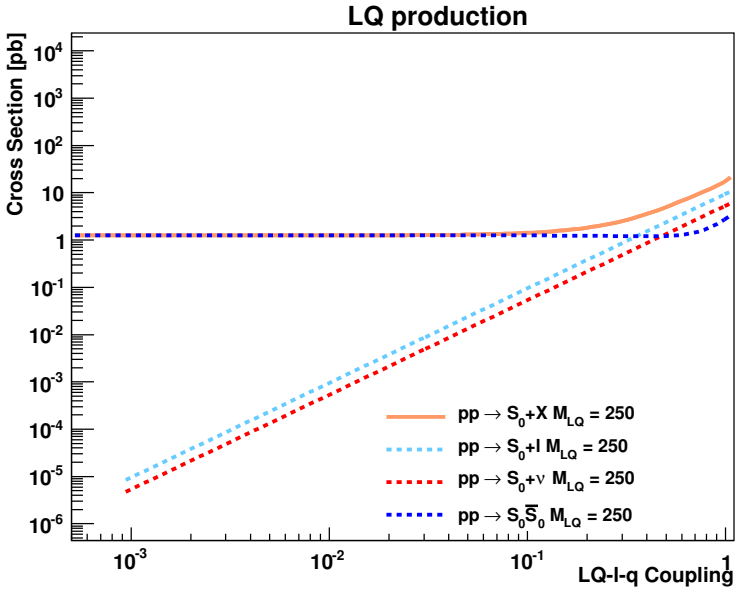


Figure 5.2: The leptoquark coupling as a function of cross section. In this plot S_0 represents the production of a leptoquark, and $S_0\bar{S}_0$ implies pair production, which is shown in dashed dark blue. All leptoquarks are produced with a mass of 250 GeV. This shows that for pair production, changing the coupling constant, λ, does not affect the production cross section for λ <0.8. This result was checked for leptoquark masses of 250 GeV (shown) and 500 GeV with the same results [17].

5.2 Object Selection

5.2.1 Kinematic Variable Definitions

Four event variables used in event selection are the reconstructed invariant mass of the leptoquark ($M(LQ)$), transverse mass of the leptoquark ($M_T(LQ)$), transverse mass of the W boson ($M_T(l, \not{E}_T)$), and S_T. The leptoquark masses are determined by pairing each lepton with each jet and calculating $M(LQ)$ and $M_T(LQ)$ from each pair. The combination of leptons and jets that provides the smallest mass difference determines which lepton is paired with which jet. The other two variables are defined as:

$$M_T(l, \not{E}_T) = \sqrt{2p_T^l \not{E}_T (1 - \cos\Delta\phi)} \tag{5.1}$$

$$S_T = p_T^l + p_T^{jet_1} + p_T^{jet_2} + \not{E}_T \tag{5.2}$$

5.2.2 Electrons

Electrons selected for the analysis satisfy the requirements for **RobustMedium** electrons described in Section 4.4.1. To better reject the mulitijet background, a relative isolation of E_T^{cone20}/E_T <0.2 is required. E_T^{cone20} is defined as the energy deposition in a ΔR cone of 0.2 from the center of the cluster, with the cluster itself removed. For additional photon conversion rejection, the matching track must have a hit in the first layer of the pixel detector (b-layer) when an active pixel module is traversed. The trigger used for the MC is L1_EM14. The triggers used in the data were separated by run period and listed in Table 5.1. Because parts of the calorimeter had dead or poorly functioning optical transmitters (OTx), these regions were removed from the data sets and masked in the MC.

5.2.3 Jets

Jets used in this analysis satisfy the requirements described in Section 4.3 for the anti-k$_T$ algorithm with a parameter $R = 0.4$. They have the additional requirement that there is at least one good jet in the event preselection and there are at least two good jets in the signal region.

A "good" jet is one that passes the following set of requirements:

- The majority of energy deposited is in the hadronic calorimeter (specific percentages vary with η and ϕ based on detector topology)
- The jets are in time with a beam crossing
- The jets that are reconstructed inside the fiducial volume of the inner detector have a matching track associated with them
- The jets originate in the hard-scatter interaction (jet-vertex-fraction, or JVF)

A "bad" jet fails at least one of the above requirements. An "ugly" jet is one that traverses there are dead or uncalibrated cells.

5.2.4 Missing Energy

$\not{E}_T$ used in this analysis is calibrated using the LCW (LocHadTopo) scheme defined in Section 4.6. If an ugly jet is present, the event is rejected because the $\not{E}_T$ is not properly reconstructed.

5.2.5 Muon Veto Requirement

Events in this analysis were vetoed if a muon in the event passed the combined STACO algorithm requirements listed in Section 4.5. Muons are specifically vetoed so that the muon channel and electron channel LQ searches are exclusive.

In addition to the event containing a combined STACO muon, if the ID track has the following requirements, the event is vetoed:

- At least one pixel hit
- Six SCT hits
- A required number of TRT hits (nTRTHits) compared to the number of hits that do not get reconstructed as a track (nTRTOutliers)
  Specifically the requirements are for $|\eta| < 1.9$:
  nTRTHits + nTRTOutliers > 5 and nTRTOutliers < 0.9 (nTRTHits + nTRTOutliers)
  And the requirements are for $|\eta| > 1.9$:
  if n = nTRTHits + nTRTOutliers > 5, then nTRTOutliers < 0.9n
- Match quality must satisfy $\chi^2 < 150$
- Impact parameters: $|d_0| < 0.1$ mm and $|z_0| < 5$ mm
- Combined muon $p_T > 20$ GeV
- Data periods A-F: $|\eta| < 2.0$; other data periods: $|\eta| < 2.4$
- Relative $p_T^{cone\ 20} < 0.25$, Illustrated in Figure 4.2

5.3 Background Yield Determination

All the background yields, except for multijet (MJ) or QCD events, and all signal yields are determined from simulated events after applying data/Monte Carlo scale factors described in Section 3.3. MJ backgrounds that come from jets which are misidentified as electrons are estimated using data driven methods. These methods are described in Section 5.3.3.

5.3.1 Simulated Background

The yield of the simulated background is determined by applying the defined selection criteria to determine an efficiency and an acceptance. These values are then multiplied by the cross section, and integrated luminosity of the data set. This is described in Equation 5.3.

$$N = \sigma \epsilon \alpha \mathcal{L} \tag{5.3}$$

5.3.2 Simulated Signal

The yield of the simulated signal is determined in much the same way as the simulated background. The efficiency and acceptance of the selection criteria are multiplied by the cross section and integrated luminosity of the data set. An interesting effect was observed due to the parton shower model implemented by Pythia. This required the regeneration of all the signal samples. This effect is discussed in greater detail in Appendix B.

5.3.3 Multijet Background

The shape of the multijet (MJ) background is determined by a "loose-tight" sample with the residual contamination from real electrons estimated to be 7%. Loose is defined by the **Robust Medium** requirement, which was the loosest sample which could be obtained with the trigger and tight is defined by electrons described in Section 4.4.1 (**Robust Medium**+b-layer+ isolation). The "loose-tight" sample is defined by events which pass the loose selection but fail the tight selection. The normalization is determined by fitting a sum of the $M_T(e, \not{E}_T)$ distributions of the simulated background and the MJ background to the data (See Equation 5.4).

$$\sum data = \sum MC \tag{5.4}$$

The total number of background events is constrained to the data yield. The fraction of MJ background, α, is the fit parameter. The negative log likelihood is then minimized to find the most likely fraction of MJ background. This fit with the resulting log likelihood plot is shown in Figure 5.3.

5.4 Event Selection

The event selection is divided into three parts. First a "preselection" region is defined which has a very low signal-to-background ratio. This is done to insure that there is agreement between the background rich modeled regions and the data. Once the preselected samples are in agreement, an the signal region is completely removed and the background modeled regions are again compared with data. This is called a Control Region. Finally, a signal rich region is defined where the majority of the background events are removed. This is defined as the signal region. These three regions will be discussed in greater detail in the following sections.

5.4.1 Preselection Region

The preselection criteria for events satisfies the following prescription:

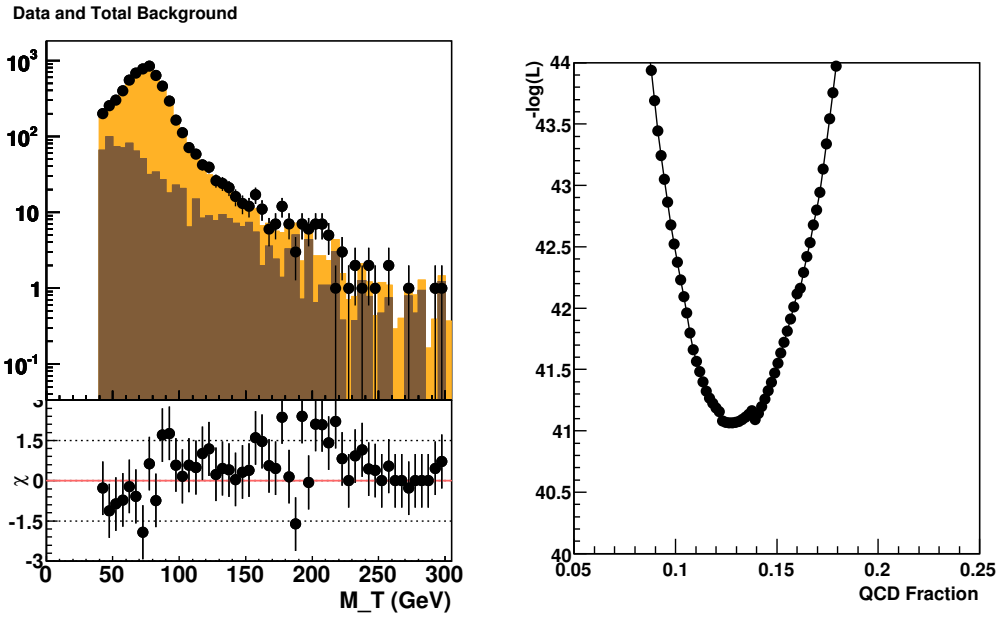


Figure 5.3: Fit Method for QCD estimation. The figure on the left shows the M_T distribution with the brown representing the multijet background, yellow all other backgrounds, and the black points data. The plot on the lower left shows the χ^2 fit for each bin. The total number of background events is constrained to the data yield and the fraction of MJ background, α, is the fit parameter. The plot on the right shows the negative log likelihood. The -log(L) is then minimized to find the most likely fraction of MJ background. In this plot it appears to be ≈12%.

- Pass the events listed in the GRL and have a good primary vertex as described in Section 4.2.
- Pass the appropriate trigger based on the run number.
- Exactly 1 good electron
- $\not{E}_T > 25$ GeV
- $M_T(e, \not{E}_T) > 40$ GeV
- At least 1 good jet

To remove residual events with badly measured jets, the event must fall outside of a pre-defined triangle of phase space. This is called a "triangle-cut". The region of phase space removed by the triangle is low $\not{E}_T$ with small $\Delta\phi(jet, \not{E}_T)$ and is is shown in Figure 5.4 for

the highest p_T good jet (upper) and the 2^{nd} highest p_T jet (lower row). The signal events are shown on the right and the QCD events on the left. Removing this region allows extra rejection of residual events with badly measured jets. Also if there is an event which has a good muon, the event is vetoed.

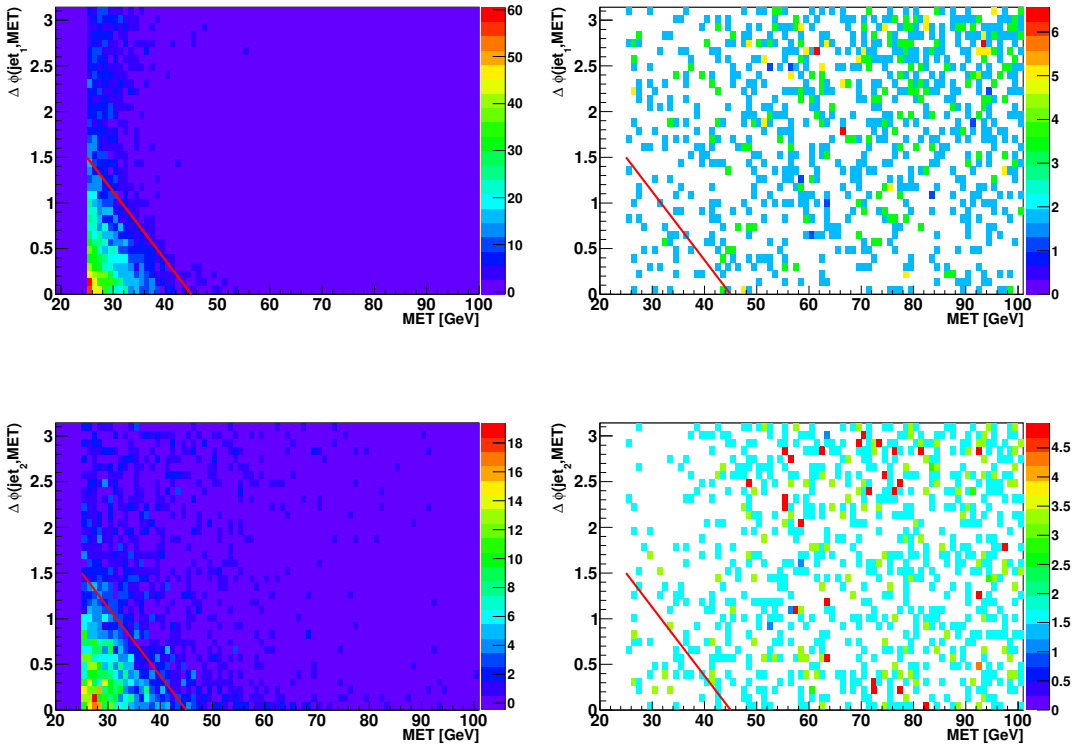


Figure 5.4: Triangle Cut. The region of phase space removed by the triangle is low $\not\!\!E_T$ with small $\Delta\phi(jet, \not\!\!E_T)$. The two triangles are defined with respect to the highest p_T good jet (upper row) and the 2^{nd} highest p_T jet (lower row), with the signal events shown on the right and the QCD events on the left.

Preselected Sample Yields

The results from each event selection criteria from the preselection are listed in Table 5.2. Each column applies the criteria from the column to the left. This style of representing results is called a cut flow. The results from the preselection yield 6090^{+990}_{-1130} predicted events where the uncertainties are systematic and statistical and 6088 observed events. The kinematic variable transverse momentum, p_T, distributions for: the electron, $\not\!\!E_T$, and the two highest p_T jets are shown in Figure 5.5. The event reconstruction variables: $M(LQ)$, $M_T(LQ)$, $M_T(l, \not\!\!E_T)$, and S_T are shown in Figure 5.6.

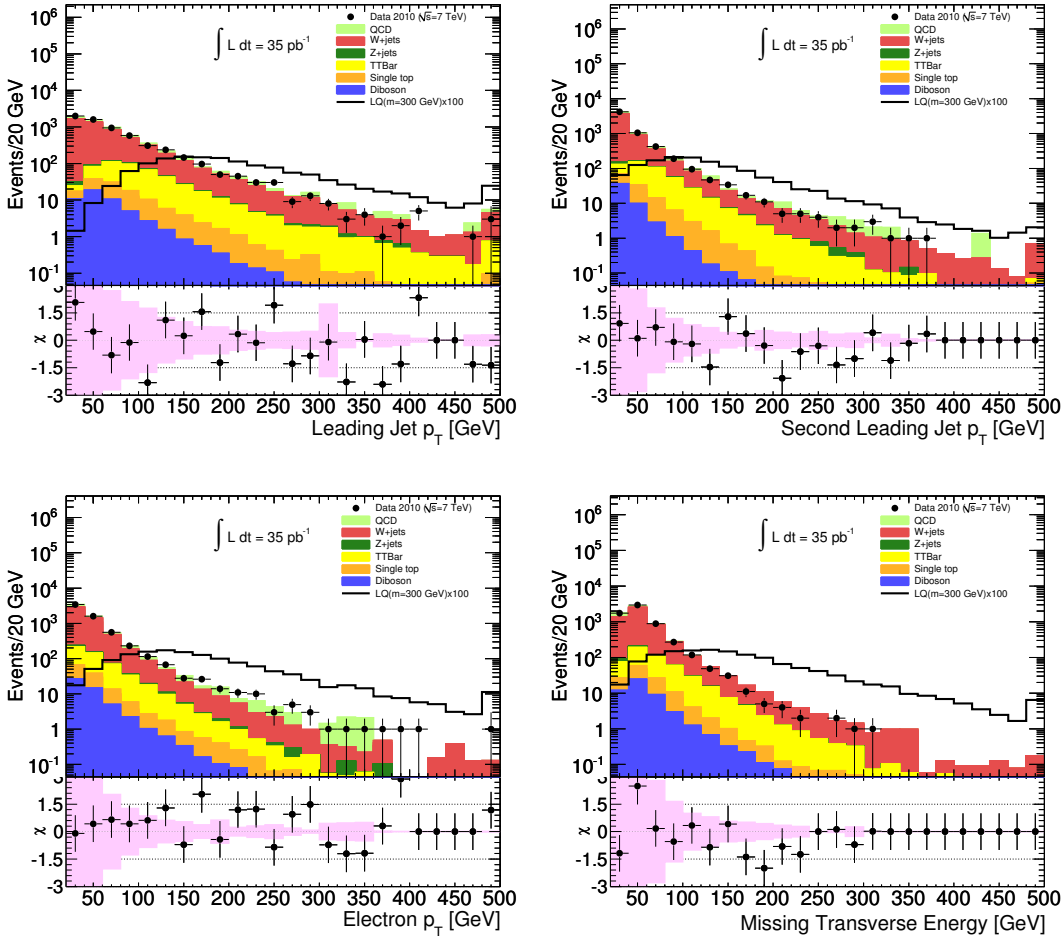


Figure 5.5: Preselection p_T distributions for the two highest p_T jets (upper left and right), the electron (lower left), and $\not\!\!E_T$ (lower right). The backgrounds included are listed in the legend and shown in the colored histograms. The signal displayed is a 300 GeV leptoquark with a cross section enhanced by $\times 100$ and is shown in the black line. The data shown has an integrated luminosity of 35 pb^{-1} and is in black dots. The plots below the p_T distributions are the results of Poisson probability calculation. This is used instead of a Gaussian because of the low statistics in high p_T regions. It is then translated into a χ^2 between the data and MC predictions with the systematic uncertainty plotted in pink.

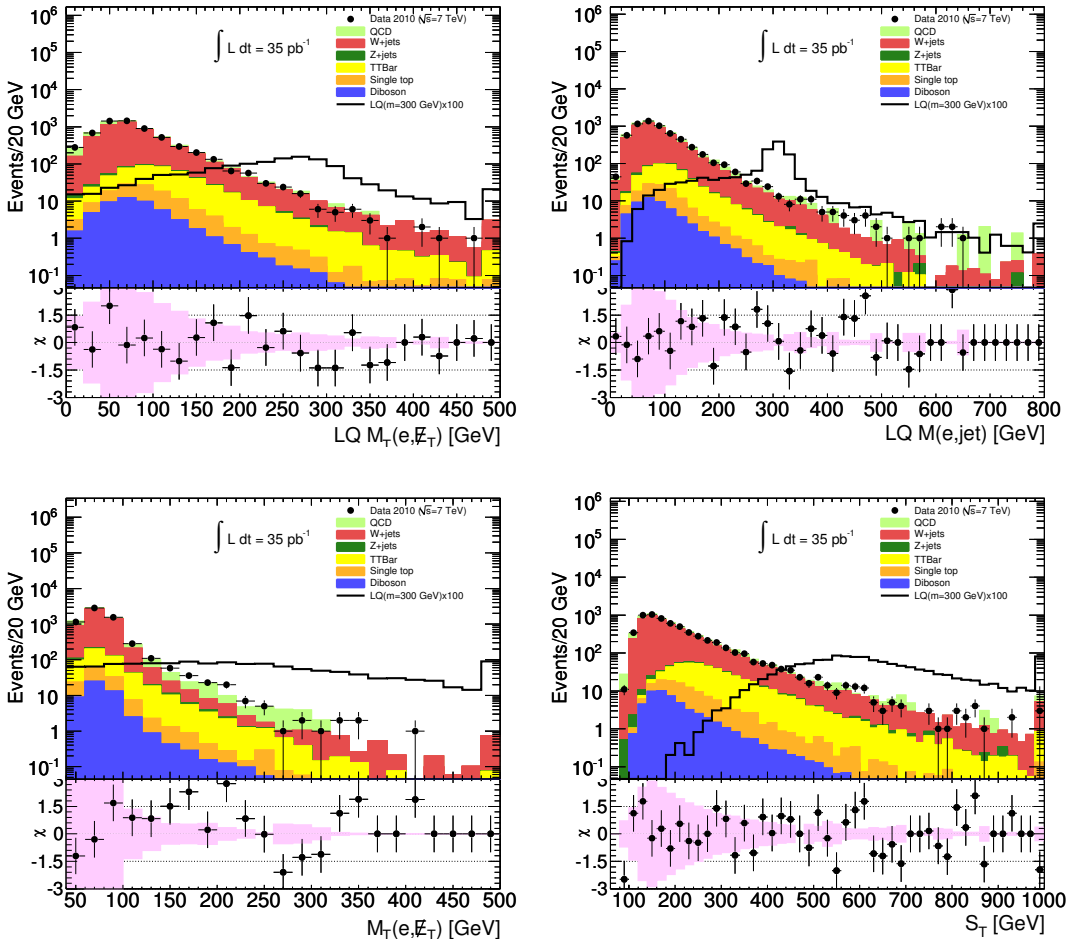


Figure 5.6: Event reconstruction variables: $M_T(LQ)$ (upper left), $M(LQ)$ (upper right), $M_T(l, \not{E}_T)$ (lower left), and S_T (lower right). The backgrounds included are listed in the legend and shown in the colored histograms. The signal displayed is a 300 GeV leptoquark with a cross section enhanced by $\times 100$ and is shown in the black line. The data shown has an integrated luminosity of 35 pb^{-1} and is in black dots. The plots below the p_T distributions are the results of Poisson probability calculation. It is then translated into a χ^2 between the data and MC predictions with the systematic uncertainty plotted in pink.

5.4.2 Control Regions

To insure that the background is modeled properly, a set of background enhanced regions has been defined in which the signal is essentially removed. There are three regions defined by the dominant backgrounds to the signal: two different W+jets enhanced regions and one $t\bar{t}$ region. All of the requirements of the control regions are in addition to the requirements already defined in the preselection. These control regions are described in the following subsections.

W+jets Control Region

There are two W+jets enhanced regions: a W+=2 jets region and a W+ >3 jets region. The W+=2 jets enhances the W mass peak and requires the associated production of exactly two jets. This emulates exactly the signal: one electron, $\not\!\!E_T$ and two jets. The region is defined by the following requirements:

- exactly 2 jets
- $M_T(e, \not\!\!E_T) < 150$ GeV
- S_T <175 GeV

There is also a high probability that leptoquarks will also be produced in association with additional jets, since the event is produced at high energies. Therefore, an additional W enhanced region is considered with additional jets. The W+ >3 jets is defined by:

- ≥ 3 jets
- $M_T(e, \not\!\!E_T) < 150$ GeV
- S_T <200 GeV

The resulting yields from the two W enhanced control regions are listed in Table 5.3 and distributions of event variables are plotted in Figures 5.7 and 5.8.

$t\bar{t}$ Control Region

$t\bar{t}$ production yields a topology with two W bosons and two b-flavored jets. Since there was no specific tagging of these b-flavored jets in this analysis, they appear as all other jets. This decay topology yields is similar to the leptoquark signal topology, but is has even more jets than the W+> 3 jets region already defined. The $t\bar{t}$ region is defined by:

- ≥ 4 jets

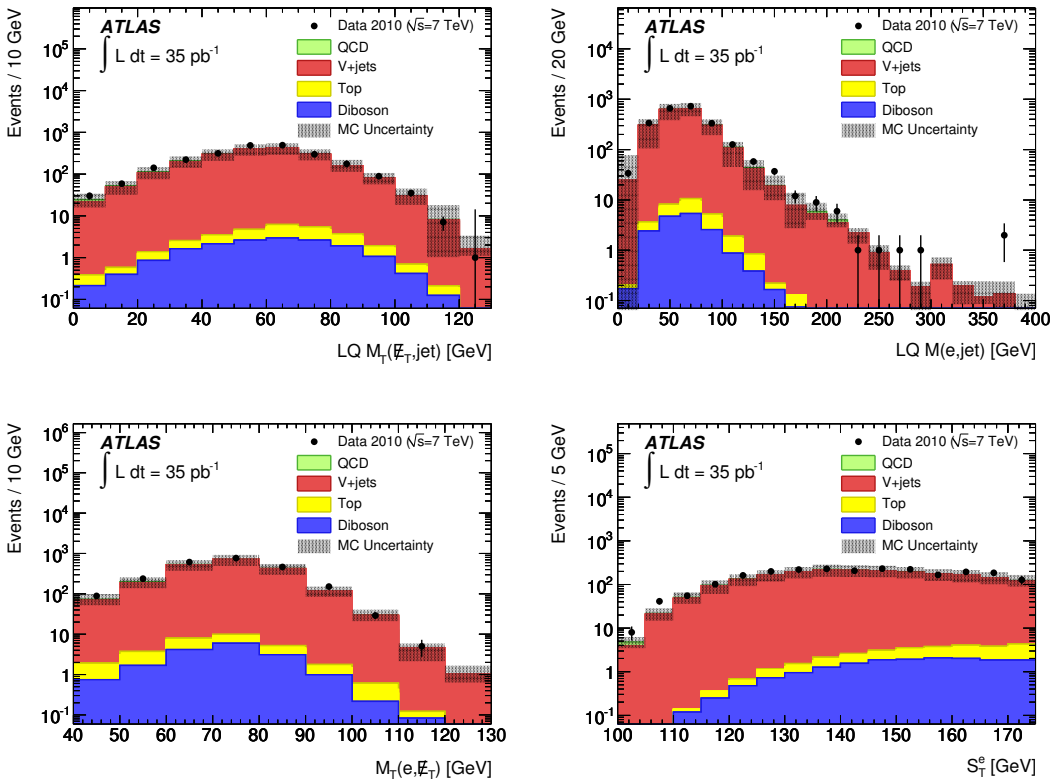


Figure 5.7: Event reconstruction variables from the W+=2 jets enhanced region: $M_T(LQ)$ (upper left), $M(LQ)$ (upper right), $M_T(l, \not{E}_T)$ (lower left), and S_T (lower right). The backgrounds included are listed in the legend and shown in the colored histograms. The systematic uncertainty for the MC is shown in the gray hatched regions. The data shown has an integrated luminosity of 35 pb^{-1} and is in black dots. It should be noted that the dominant background in these plots is the V+jets which is exactly what the event selection was designed to enhance.

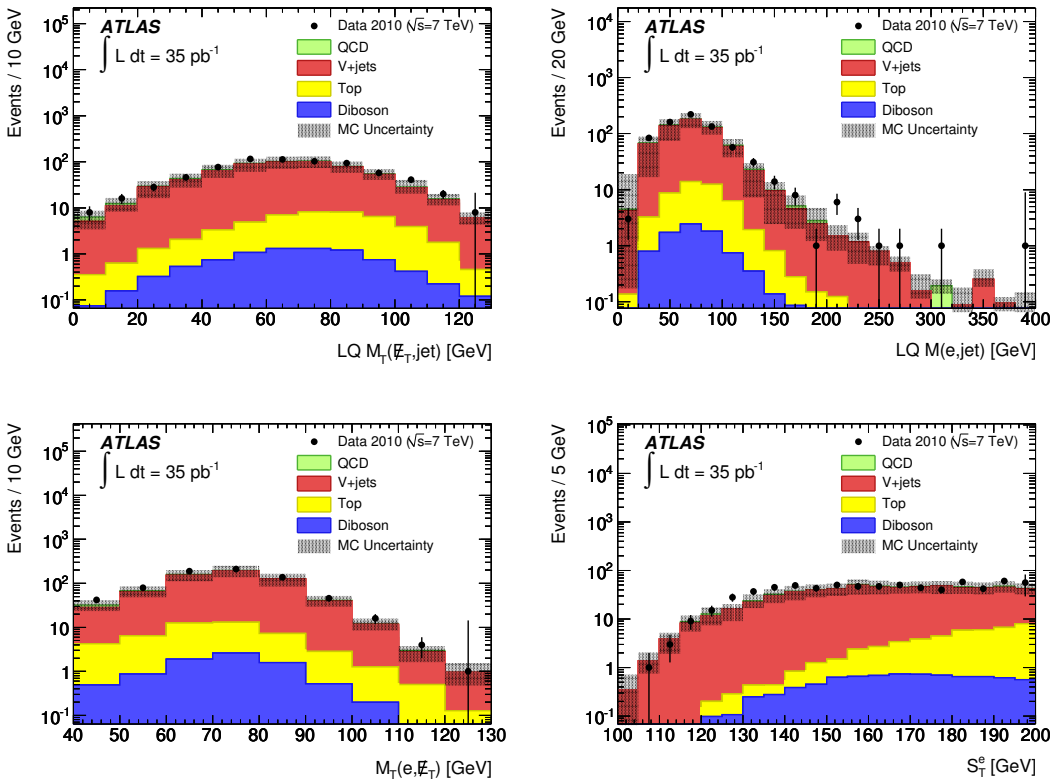


Figure 5.8: Event reconstruction variables from the W+>3 jets enhanced region: $M_T(LQ)$ (upper left), $M(LQ)$ (upper right), $M_T(l, \not{E}_T)$ (lower left), and S_T (lower right). The backgrounds included are listed in the legend and shown in the colored histograms. The systematic uncertainty for the MC is shown in the gray hatched regions. The data shown has an integrated luminosity of 35 pb^{-1} and is in black dots. In this series of histograms the dominant background is both V+jets with much more influence from $t\bar{t}$ than the W+=2 jet control region.

- $M_T(e, \not{E}_T) < 150$ GeV
- $p_T(jet_1)$ >50 GeV, $p_T(jet_2)$ >40 GeV, $p_T(jet_3)$ >30 GeV

The resulting yields of these additional requirements are listed in Table 5.3 and distributions of event variables are plotted in Figure 5.9.

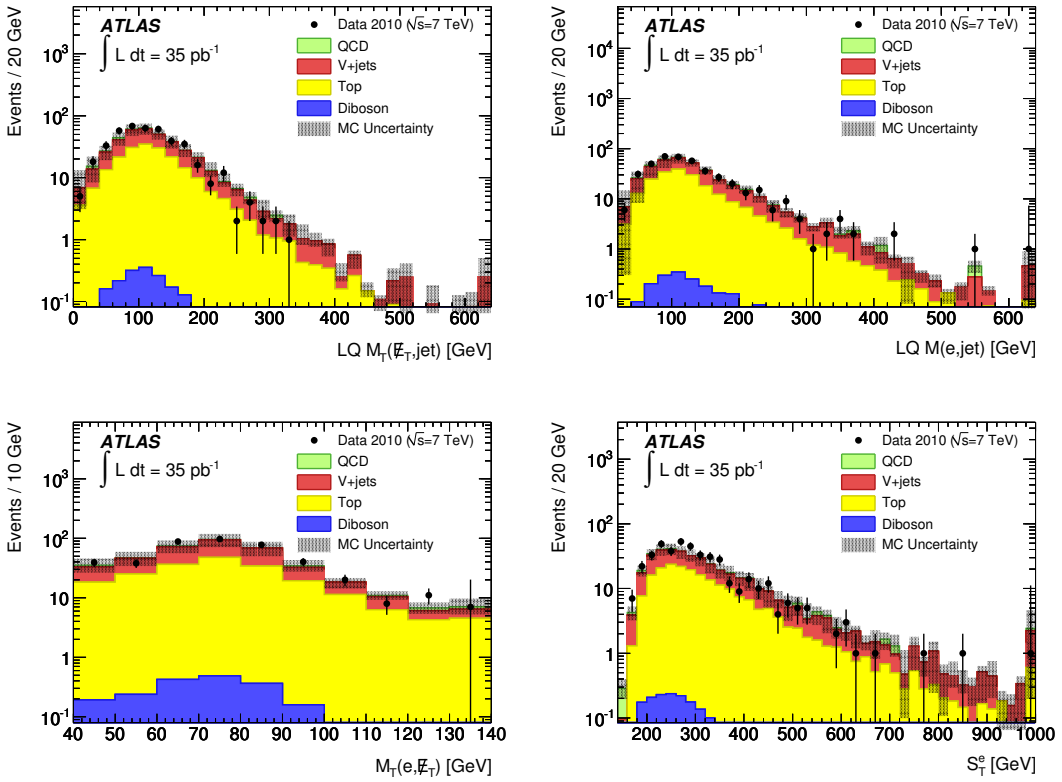


Figure 5.9: Event reconstruction variables from the $t\bar{t}$ jets enhanced region: $M_T(LQ)$ (upper left), $M(LQ)$ (upper right), $M_T(l, \not{E}_T)$ (lower left), and S_T (lower right). The backgrounds included are listed in the legend and shown in the colored histograms. The systematic uncertainty for the MC is shown in the gray hatched regions. The data shown has an integrated luminosity of 35 pb^{-1} and is in black dots. The dominant background in these histograms is clearly $t\bar{t}$.

Summarized Control Region Results

There is agreement within uncertainties in both the normalization and shape in all three defined control regions and the preselection region. This allows the analysis to proceed to the event selection stage where it focuses on signal enhanced regions.

5.5 Optimized LQ Selection

Once the background regions were validated, a method for determining a signal enhanced region must be established. A rigorous approach is to do a systematic search over a grid of points in an N-dimensional parameter space. A search over a regular grid, however, is inefficient since most of the phase space lies in regions where there is little or no signal. One way to more efficiently determine not only the optimal parameter space but also the optimal criteria is using a Random Grid Search (RGS) method as defined in Section 5.5.1.

5.5.1 Random Grid Search Optimization Method

The Random Grid Search (RGS) is a variant of traditional fixed-grid based optimizations. It works by choosing test points in an N-dimensional space corresponding to N selection variables specified by the user. Each test point defines the minimum corner of an N-dimensional box corresponding to the signal region. The optimal cut is the point which gives the highest signal significance. The unique feature of an RGS algorithm is the manner in which the test points are chosen. Each test point corresponds to the values of the N selection variables in one event in the simulated signal event sample. The RGS Method uses MC generated signal events as the values in N-dimentional parameter space [72]. The significance is calculated by minimizing the poisson probability that a background fluctuation is signal and background. This is discussed in greater detail in Chapter 7. The only difference is when determining the optimal values to use in the RGS method, only MC is used, no data.

The MC signal events used in the optimization were generated for leptoquarks with a mass of 300 GeV as defined in Section 5.1.2. A systematic uncertainty of 10% is assigned and is based on an approximation on the uncertainty from the luminosity. The distribution of the significance as a function of the background is show in Figure 5.10. This distribution is used to determine which selection of cuts will give the highest signal significance, this point is highlighted in red. This method is chosen over multivariate methods because simplicity is desirable in an early analysis. Multivariate methods are simply not needed to extend beyond previous searches. Also the number of variables has been limited to four and only variables dealing with the specific event topology and kinematics have been used.

5.5.2 Selection Criteria and Stability Tests

The point of highest significance corresponds to the results listed in Table 5.4. To be sure the points are not only giving the largest possible significance but also a stable point, a series of stability tests are performed. Since RGS uses as test points the values of every variable of the signal sample in every event, the major density of points will be in the bulk of the signal distributions. In other regions, like in the high/low mass tails, etc, the statistics are poorer and so the density of points will be smaller. It is imperative that when optimizing, a set of selection criteria is not the result of a statistical fluctuation. Figure 5.11 shows the

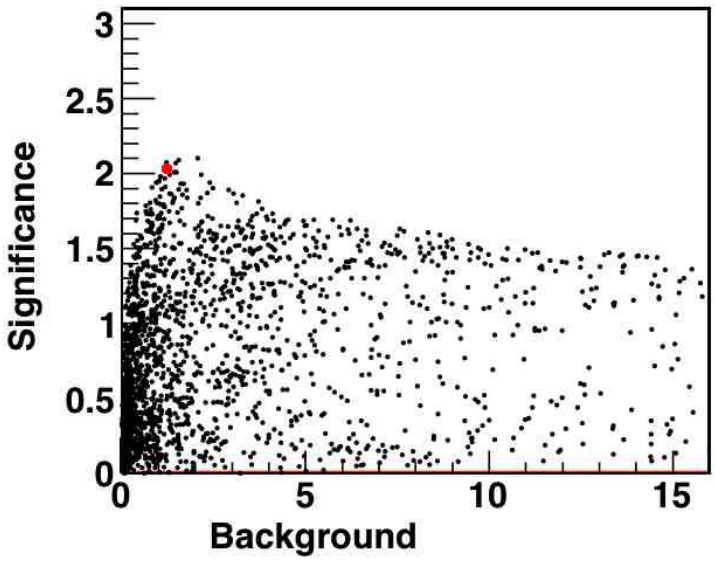


Figure 5.10: Significance plotted vs. background yield. Signal used was 300 GeV leptoquarks.

values of the selection cuts for the four variables used in this optimization for all cut points whose significance, s satisfies $s > 0.85 \times s_{max}$ with s_{max} the maximum significance.

With the RGS method an unstable point is easily identified and avoided. The final signal region selection requirements are determined by the RGS procedure and rounded to the nearest 10 GeV. The significance was checked both before and after rounding verifying that rounding did not affect the results.

5.5.3 Mass Dependent Optimization

Although only one MC signal (300 GeV leptoquarks) was used in the optimization, optimizations were performed for each of the following leptoquark masses, $M(LQ) = 250$, 300, and 350 GeV. The results from this optimization pass are shown in detail in Table 5.4. Table 5.5 shows a summary of the predicted significance values for each mass and each optimization applied to the different leptoquark masses considered. Since very limited mass dependence is seen in the results, the 300 GeV optimized selection was used for all masses.

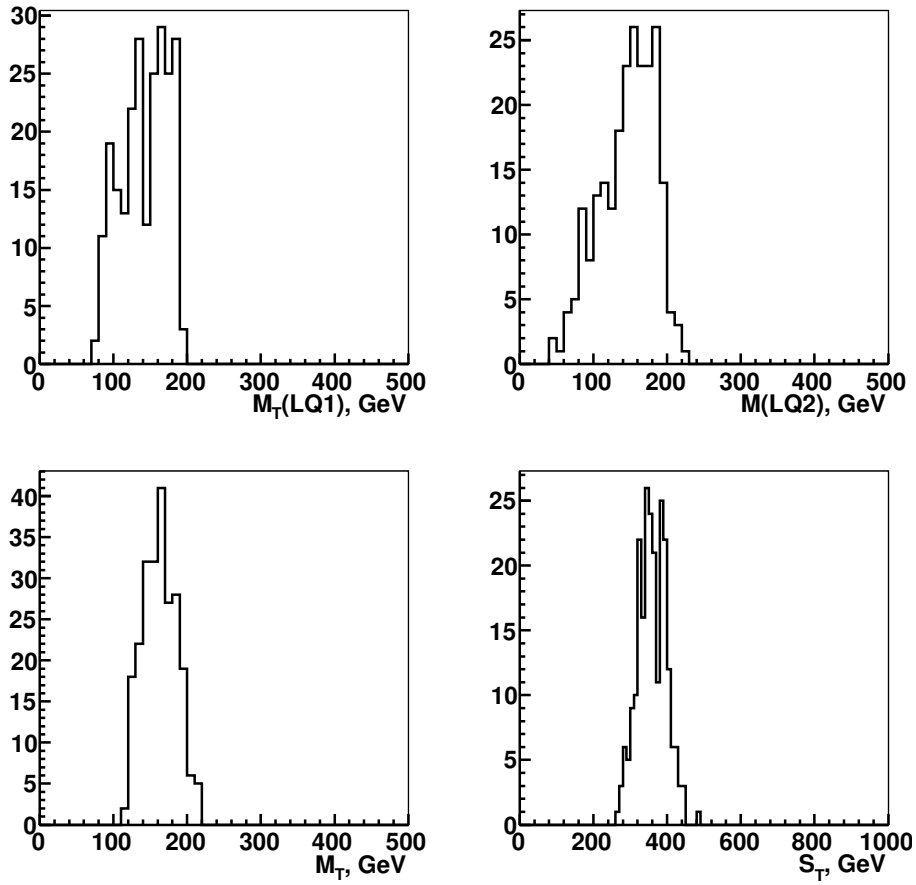


Figure 5.11: Values of the selection cuts for the four variables used in this optimization for all cut points whose significance, s satisfies $s > 0.85 \times s_{max}$ with s_{max} the maximum significance.

Source	Selection Yields							
	Good Event	Lepton	$\not\!\!E_T$	M_T	≥ 1 jet	Triangle	μ veto	$\geq 2jet$
$W \rightarrow e\nu$+jets	230196 ± 132	151209 ± 135	124166 ± 129	123051 ± 129	24687 ± 59	19869 ± 54	19868 ± 54	4645 ± 22
$W \rightarrow \tau\nu$+jets	52635 ± 91	5404 ± 33	2962 ± 24	2546 ± 23	555 ± 10	476.1 ± 9.2	475.0 ± 9.2	115.5 ± 3.6
$Z \rightarrow ee$+jets	30395 ± 25	13848 ± 29	419.4 ± 6.3	365.9 ± 5.9	264.4 ± 5.0	57.9 ± 2.3	57.7 ± 2.3	24.4 ± 1.5
$Z \rightarrow \tau\tau$+jets	10579 ± 27	1358 ± 11	351.7 ± 5.8	213.5 ± 4.6	72.3 ± 2.6	42.2 ± 2.0	35.4 ± 1.8	11.6 ± 1.0
$t\bar{t}$	2773.2 ± 2.3	709.3 ± 2.9	601.5 ± 2.8	496.6 ± 2.6	495.1 ± 2.6	435.3 ± 2.4	381.6 ± 2.3	372.9 ± 2.3
single top	604.0 ± 1.8	195.2 ± 2.0	160.5 ± 2.0	141.5 ± 1.9	135.8 ± 1.9	116.7 ± 1.8	112.1 ± 1.7	80.7 ± 1.6
Diboson	648.08 ± 0.66	257.18 ± 0.61	181.21 ± 0.55	165.38 ± 0.53	133.55 ± 0.49	107.71 ± 0.45	101.91 ± 0.44	60.56 ± 0.35
Multijet	—	—	—	—	—	—	1710 ± 14	779 ± 16
Total Bkg.	—	—	—	—	—	—	22740 ± 60	6090 ± 30
Data	—	—	—	—	—	—	23248	6088
LQ 250 GeV	58.91 ± 0.11	44.69 ± 0.27	43.08 ± 0.28	39.03 ± 0.29	39.01 ± 0.29	38.10 ± 0.30	38.10 ± 0.30	36.63 ± 0.30
LQ 300 GeV	20.58 ± 0.04	15.73 ± 0.09	15.28 ± 0.10	14.02 ± 0.10	14.01 ± 0.10	13.81 ± 0.10	13.81 ± 0.10	13.44 ± 0.10
LQ 350 GeV	8.11 ± 0.02	6.34 ± 0.04	6.18 ± 0.04	5.72 ± 0.04	5.71 ± 0.04	5.65 ± 0.04	5.65 ± 0.04	5.51 ± 0.04
LQ ($eqeq$) 250 GeV	119.52 ± 0.18	37.74 ± 0.56	10.80 ± 0.35	9.45 ± 0.33	9.45 ± 0.33	4.63 ± 0.23	4.62 ± 0.23	4.61 ± 0.23
LQ ($eqeq$) 300 GeV	41.69 ± 0.06	13.07 ± 0.20	4.24 ± 0.13	3.65 ± 0.12	3.65 ± 0.12	1.93 ± 0.09	1.93 ± 0.09	1.93 ± 0.09
LQ ($eqeq$) 350 GeV	16.45 ± 0.02	4.84 ± 0.08	1.76 ± 0.05	1.55 ± 0.05	1.55 ± 0.05	0.90 ± 0.04	0.90 ± 0.04	0.90 ± 0.04

Table 5.2: Preselected sample predicted and observed cut flow yields. All uncertainties except that for the MJ background are statistical only. The uncertainty for the MJ background is determined from the M_T fit described in Section 5.3.3. Table A.1 in Appendix A has full details for the subsamples.

Source	Control Region		
	$W+2$ jets	$W+\geq 3$ jet	$t\bar{t}$
V+jets	2100 ± 700	580 ± 190	180 ± 60
Top	21 ± 4	44 ± 9	210 ± 40
Diboson	17 ± 4	8.3 ± 1.9	2.1 ± 0.5
QCD	64 ± 14	68 ± 15	29 ± 7
Total Bkg	2200 ± 700	700 ± 200	420 ± 80
Data	2344	722	425

Table 5.3: The predicted and observed yields in the control samples. Top refers to both single top and $t\bar{t}$ events. Both statistical and systematic uncertainties are included. Table A.2 in Appendix A has the full event selection with the yields for every background individually and requirement.

LQ Mass	Selection Requirement (GeV)						
(GeV)	M_T^{LQ1}	M^{LQ2}	$M_T(e, \not{E}_T)$	S_T	Signal	Background	Significance
250	160	140	150	370	10.4	7.6	3.0
300	180	180	200	410	3.4	1.4	2.1
350	280	210	210	510	1.5	1.0	1.1

Table 5.4: Results of the significance optimization for the electron channel as a function of LQ mass. The selection requirements, predicted signal and background yields and the predicted significance are shown. A 10% overall background uncertainty is included in the optimization.

.

Input LQ	250 GeV Optimized Selection			300 GeV Optimized Selection			350 GeV Optimized Selection		
Mass (GeV)	Signal	Bkg	Signif.	Signal	Bkg	Signif.	Signal	Bkg	Signif.
250	10	6.3	3.0	5.6	1.6	3.1	3.3	1.0	2.3
300	5.0	6.3	1.6	3.4	1.6	2.1	2.7	1.0	1.9
350	2.4	6.3	0.8	1.7	1.6	1.0	1.5	1.0	1.1

Table 5.5: A comparison of the predicted yields and backgrounds resulting from the initial optimization as a function of the input leptoquark mass. Each row corresponds to the predictions for a specific input mass, and each (set of three) column(s) shows the result from a given optimized selection.

Chapter 6

Systematic Uncertainties

Systematic uncertainties are the best estimation of a systematic bias in our estimation of the signal and background efficiency or shape. They are derived for a variety of sources including:

- Data and Monte Carlo differences in efficiencies and resolutions for leptons, jets and $\not\!\!E_T$
- Instantaneous luminosity dependence and underlying event effects
- The methods used to determine multijet (MJ) backgrounds
- Contributions from theory

These uncertainties are the subject of this chapter.

6.1 Electron Trigger and Reconstruction Uncertainties

The electron reconstruction efficiency systematic uncertainties are determined by varying the fitting and sideband windows used to select the $Z \to ee$ peak and by varying the fitting function used to describe the Z line shape (described in Section 3.3.1). The trigger efficiency systematic uncertainty is given by the trigger group. The values are:

- Reconstruction systematic uncertainty: $\sigma_{sys}^{reco} = 3.7\%$
- Trigger systematic uncertainty: $\sigma_{sys}^{trig} = 1\%$

6.2 Electron Energy Scale and Resolution Uncertainties

The electron energy scale (EES) and electron energy resolutions (EER) mainly come from calorimeter calibrations. The e/γ and Monte Carlo Production (MCP) groups have derived official EES and EER uncertainties. In the case of the EES uncertainty, $\pm 1\sigma$ is added to the electron energy and propagated to the $\not\!\!E_T$. For the resolution, the uncertainty is in the form of a Gaussian smearing and a 1σ variation randomly chosen event-by-event is applied. The yields for each simulated background and the leptoquark signal are recomputed with the change converted to a fractional systematic uncertainty given by Equation 6.1, where $N_i^{\pm}$ is is the yield for the i^{th} sample for the plus or minus variation respectively and N_i is the yield for the default electron energy scale and resolution.

$$\sigma^{\pm} = \frac{N_i^{\pm} - N_i}{N_i} \tag{6.1}$$

The resulting systematic uncertainties found using the e/γ group variations are shown in Table 6.1. It should be noted that given the limited statistical precision for the $W \to \tau\nu$ and $Z \to \ell\ell$ samples, the $W \to e\nu$+jets systematic uncertainty are used for all V+jets backgrounds. Detailed tables are located in Table A.6 in Appendix A.

	Preselection			300 GeV Optimization		
	Scale			Scale		
Source	σ^-	σ^+	Resolution	σ^-	σ^+	Resolution
$W \to e\nu$+jets	-0.0113	0.0116	-0.0012	0.1098	—	—
$W \to \tau\nu$+jets	0.0261	-0.0334	-0.0082	—	—	—
$Z \to ee$+jets	-0.0884	0.1497	-0.0027	—	—	—
$Z \to \tau\tau$+jets	0.0274	0.0189	-0.0014	—	—	—
$t\bar{t}$	0.0024	-0.0020	-0.0008	0.0476	0.0238	—
single top	-0.0037	-0.0005	-0.0016	—	—	—
Diboson	-0.0091	0.0113	-0.0014	0.0447	-0.0494	—
LQ 250 GeV	-0.0010	0.0031	-0.0006	-0.0011	0.0044	0.0022
LQ 300 GeV	-0.0037	0.0015	—	-0.0050	-0.0088	0.0013
LQ 350 GeV	-0.0038	0.0016	—	-0.0043	0.0071	0.0019

Table 6.1: Systematic uncertainty from the electron energy scale and resolution uncertainty. More detailed tables are located in Table A.6 in Appendix A.

6.3 Jet Energy Scale and Resolution Uncertainties

The jet energy scale (JES) and resolution (JER) systematic uncertainties are determined using an official tool provided by the Jet/$\not\!\!E_T$ group called the JESUncertaintyProvider [73,

74]. The JES systematic analysis is carried out in a similar way to the electron energy scale systematic. The jet energies are modified by $\pm 1\sigma$ and then propagated to the event $\not\!\!E_T$. However, two additional systematic uncertainties are included. The first is a 5% uncertainty added in quadrature to account for differences in response of the calorimeters for quark and gluon jets. Gluon jets tend to have broader showers, and this difference is modeled using MC. The other is also a 5% uncertainty added in quadrature because the $\not\!\!E_T$ was calculated when the jet energy was determined at the electromagnetic (EM) scale. However, the JES uncertainties were propagated after the hadronic JES corrections have been applied.

Tables 6.2 and 6.3 show the final fractional uncertainties for the JES and JER respectively. Figure 6.1 shows the changed shape of the jet spectra distributions when the JES $\pm 1\sigma$ is applied. As with the electron systematics, statistics require that the W$\rightarrow$ eν+jets results are used for all V+jets sources. It should also be noted that the systematics for the optimization for the electron and muon (with $\not\!\!E_T$) channels have been combined because of limited Monte Carlo statistics. Table A.7 in Appendix A shows the full JES systematic uncertainty results for the preselection and optimization. Likewise, Table A.8 in Appendix A shows the full JER systematic uncertainty results.

	Electron		$e+\mu$ Combined			
	Preselection		Preselection		300 GeV Opt	
Sample	σ^+_{JES}	σ^-_{JES}	σ^+_{JES}	σ^-_{JES}	σ^+_{JES}	σ^-_{JES}
$W \rightarrow e\nu$+jets	0.109	-0.138	0.08	0.13	0.22	0.13
$W \rightarrow \tau\nu$+jets	0.130	-0.197	0.12	0.17	—	—
$Z \rightarrow ee$+jets	0.161	0.268	0.03	0.03	—	—
$Z \rightarrow \tau\tau$+jets	0.186	-0.134	0.10	0.06	—	—
$t\bar{t}$	0.021	-0.018	0.04	0.03	0.13	0.03
single top (s+t+tW)	0.037	-0.048	0.02	0.03	0.34	0.32
Diboson (WW, WZ, ZZ)	0.029	-0.106	0.08	0.14	0.17	0.24
LQ 250 GeV	0.001	-0.003	0.002	0.003	0.006	0.045
LQ 300 GeV	-0.001	-0.001	0.001	0.003	0.006	0.015
LQ 350 GeV	0.001	-0.002	0.002	0.004	0.007	0.009

Table 6.2: The jet energy scale fractional systematic uncertainties. Because of low statistics in the other channels, the $W \rightarrow e\nu$+jets systematic uncertainty are used for all V+jets backgrounds. The electron and muon uncertainties are combined for the optimized selection. Detailed tables are available in Appendix A in Table A.7.

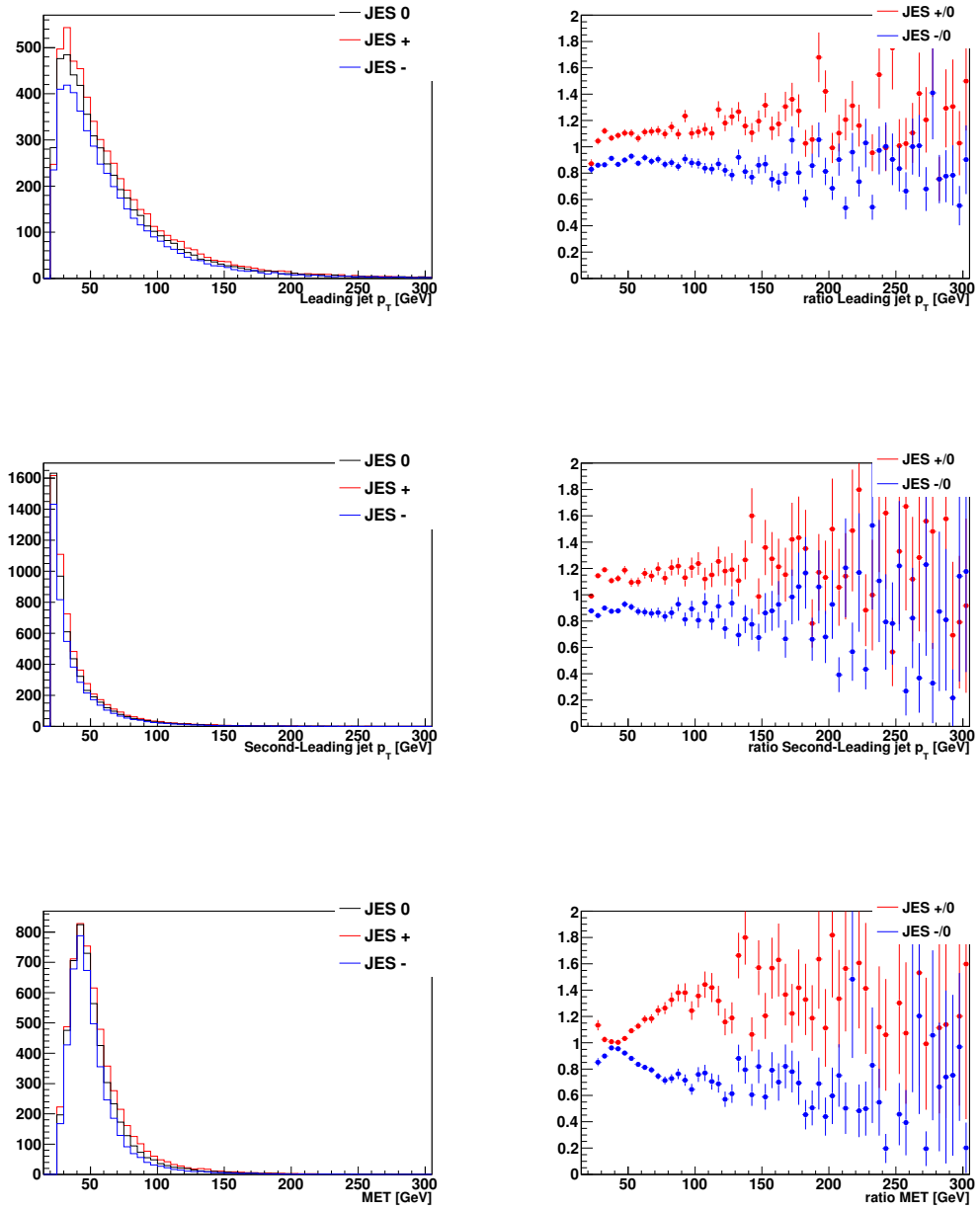


Figure 6.1: The jet energy and $\not\!\!E_T$ background spectra and their ratios with (JES $\pm$) and without (JES 0) the JES uncertainties for events in the preselected sample. The left column shows the spectra for the leading (top) and second (middle) jet and the $\not\!\!E_T$ (bottom) for the nominal JES and for the plus and minus one sigma variations. The right column shows the ratio of the shifted JES to the nominal JES as shown in the left column.

Source	Electron Chan. Preselection	$e+\mu$ Combined Preselection	$e+\mu$ Combined Optimized Sel.
$W \to e\nu$+jets	-0.102	0.117	0.09
$W \to \tau\nu$+jets	-0.099	0.089	—
$Z \to ee$+jets	0.808	0.092	—
$Z \to \tau\tau$+jets	0.095	0.074	—
$t\bar{t}$	0.015	0.034	0.26
single top (s+t+tW)	-0.018	0.026	—
Diboson (WW, WZ, ZZ)	-0.075	0.144	0.14
LQ 250 GeV	-0.003	0.003	0.020
LQ 300 GeV	-0.005	0.003	0.010
LQ 350 GeV	-0.003	0.004	0.011

Table 6.3: The fractional jet energy resolution systematic uncertainty. Because of low statistics in the other channels, the $W \to e\nu$+jets systematic uncertainty are used for all V+jets backgrounds and the electron and muon channels are combined for optimization. Detailed tables are available in Appendix A in Table A.8.

6.4 Event Pile-up Uncertainties

The systematic uncertainty from pile-up (multiple events in a single readout as an affect of changes in the instantaneous luminosity) is evaluated by comparing the results from simulated samples with and without minimum-bias events added. The fractional uncertainty is calculated in a similar way to both the electron and jet systematic fractional uncertainties using Equation 6.1. The results are shown in Table 6.4 both for the preselection and using the W $\to e\nu$+jets for all V+jets modes and the optimization after combining the electron and muon channel results. These combinations are made because of limited Monte Carlo statistics for events passing the optimized selection. The full event details are located in Table A.9 in Appendix A.

6.5 Multijet Background Uncertainties

Systematic uncertainties are determined on the MJ backgrounds by comparing results from alternate normalizations to those from the methods described in Section 5.3.3. There are two ways to evaluate this systematic: (1) change the fit parameter of the fit method or (2) chose a new method all together. Both of these options are used to determine the MJ background systematic uncertainty, and will be discussed in greater detail.

Source	Electron Chan. Preselection	$e+\mu$ Combined Preselection	$e+\mu$ Combined Optimized Sel.
$W \to e\nu$+jets	-0.071	-0.063	-0.05
$t\bar{t}$	-0.012	-0.012	-0.03
single top (s+t+tW)	-0.028	-0.026	-0.03
Diboson (WW, WZ, ZZ)	-0.035	-0.020	-0.06
LQ 250 GeV	-0.000	0.004	0.022
LQ 300 GeV	0.007	0.009	0.004
LQ 350 GeV	0.009	0.010	0.020

Table 6.4: The fractional pile-up systematic uncertainty. The results from the electron and muon channels are combined because of low Monte Carlo statistics remaining after the optimized selection. The $W \to e\nu$+jets uncertainties are used for all V+jets sources.

6.5.1 Alternate Fit Parameters

In addition to this alternative normalization, the fit-based normalization was calculated using the alternative variables: $\not\!\!E_T$ and p_T as the fit variables instead of M_T. These distributions are given in Figure 6.2 and can be compared to Figure 5.3 found in Section 5.3.3.

6.5.2 The Matrix Method

The matrix method is used as an alternative to the fit method. It starts by defining a "loose" region (electrons that only satisfy the **Robust Medium** requirement), a "tight" region (those that satisfied the requirements for the analysis), and two probabilities:

- ϵ_{fake}: the probability that a jet misidentified as an electron passes the enriched sample ID requirement also passes the full electron ID requirements
- ϵ_e: the probability for a real electron which passes the enriched sample electron ID selection to also pass the default selection

A MJ enriched sample ("loose") is defined by relaxing the electron ID requirements from those described in Section 4.4.1 to RobustMedium. The ratio of the $\not\!\!E_T$ distributions of the default selection to the MJ enriched is used to determine ϵ_{fake}. It is computed using the average value of the ratio in the range $0 < \not\!\!E_T < 15$ GeV. This is determined separately for electron candidates in the barrel and end-cap. ϵ_e is calculated from Z $\to$ ee data for $|M_{ee} - M_Z| < 10$ GeV. This probability is measured as a function of the electron-candidate p_T for candidates in the barrel and end-cap.

Using the number of MJ background fake electrons, N_{MJ}, and the number of real electrons, N_e, the number in the enriched sample (with the default $\not\!\!E_T > 25$ GeV requirement) can be found by solving the pair of equations:

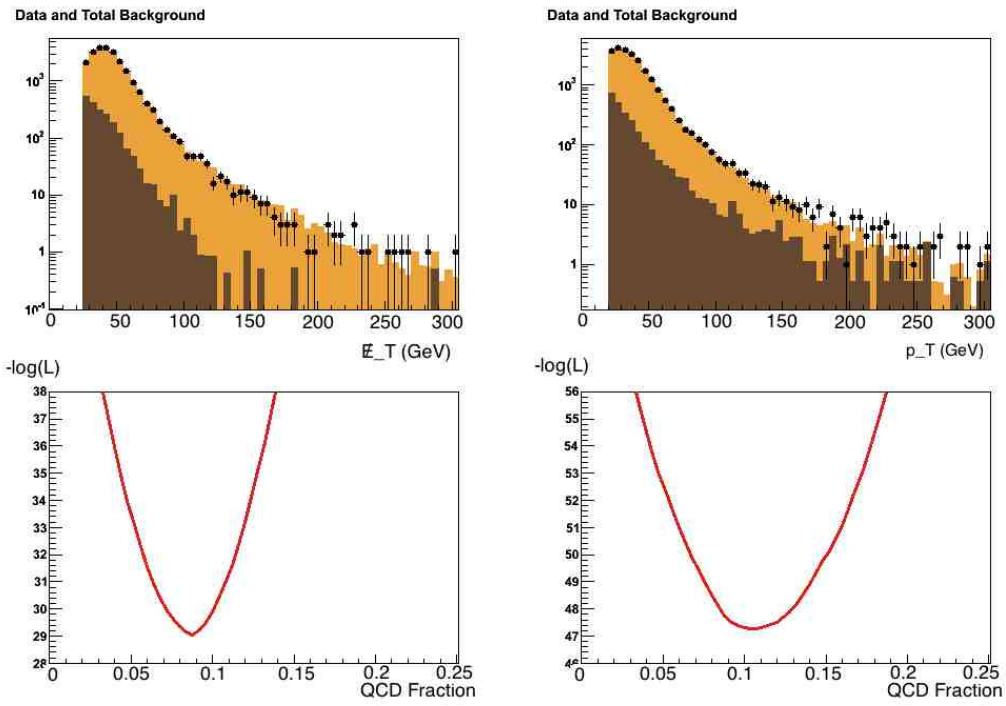


Figure 6.2: Alternative Variables for Fit Method. The $\not\!\!E_T$ distribution is shown on the upper left and the p_T distribution is shown on the upper right. The -log($\mathcal{L}$) for each is shown as a function of the fit fractions.

$$N_{enriched} = N_{MJ} + N_e \tag{6.2}$$

$$N = \epsilon_{fake} N_{MJ} + \epsilon_e N_e \tag{6.3}$$

where $N_{enriched}$ is the yield in the enriched sample with $\not{E}_T > 25$ GeV and N is the yield in the default preselected sample with the full electron ID. Solving this for the MJ background in the preselected sample gives:

$$N_{MJ} = \frac{\epsilon_e N_L - N_T}{\epsilon_e - \epsilon_{fake}} \tag{6.4}$$

The resulting value for N_{MJ} using the Matrix Method this is shown in Table 6.5.

6.5.3 Combined Multijet Systematic Uncertainty

The results from alternative MJ estimation methods are compared with that from the default method in Table 6.5. The maximum difference in yields is taken as the systematic uncertainty for the MJ background.

N_{MJ} (fit, default)	778
N_{MJ} (fit, $\not{E}_T$)	610
N_{MJ} (fit, p_T^e)	851
N_{MJ} (MM)	746
Fractional Syst.	22%

Table 6.5: The MJ background yields computed using the alternative variables in the fit method, and the matrix method, and the resulting systematic uncertainty. The systematic is taken as the largest (fractional) difference between the default method and all others.

6.6 Selection Check Using the Debug Stream

The analyses are also performed using the ATLAS debug stream as input. Sometimes processing the event output takes longer than the buffer time allotted. This time-out feature is built into the algorithms and kicks in before processing is complete. The data is dumped into a special "debug" stream, which is a safety store where potentially interesting, but problematic, data is filed for later reprocessing offline. One event passed the preselection and no events in the control regions or signal region.

6.7 Luminosity Uncertainties

The ATLAS standard uncertainty of 11% is applied to all background yields determined from simulated events [23]. This was later improved to 3.4%, but this improvement was not used for the analysis since luminosity was not a dominant systematic.

6.8 Cross Section Uncertainties

6.8.1 Signal

The systematic uncertainties on the production cross sections were evaluated by changing the scale up and down by a factor of two and by performing PDF re-weighting using the CTEQ uncertainty sets. For example, the signal M_{LQ} = 300 GeV the scale variation introduced an uncertainty of +13%, -14% and the PDF re-weighting gave a variation of 13%. The impact of these changes on the acceptance is negligible.

The production model uncertainties were derived by changing the generator level ISR, FSR and multiple interaction configurations. Initial studies based on simply turning off ISR, FSR and multiple interactions had negligible impact, except for the FSR case. This was further studied using the same ISR/FSR flag changes used by the W analysis, equivalent to the following Pythiasettings:

- parp(1)=0.192, paru(112)=0.192, mstu(112)=3, parp(61)=0.192

The result, an acceptance variation of 2% for the optimized selection, is taken as the production model systematic.

6.8.2 Background

The systematic uncertainties on the theory calculations used to normalize backgrounds are given in the text of Section 5.1.2.

6.9 W+jets Production Model Uncertainties

The systematic uncertainties for the production models of W+jets and Z+jets events are determined by making generator-level comparisons of the yields from samples with varied parameters to the yields from the nominal samples. This follows the procedures of the W+jets analysis group [66].

A default yield, N_0 is calculated using generator-level quantities corresponding to the reconstruction level variables used in the analysis from a standard Alpgen+Herwig MLM matched sample (See Section 3.1.4). The Alpgen input parameters are then varied individually by $\pm 1\sigma$ to produce samples which are used to calculate $\pm 1\sigma$ sigma yields, $N_{\pm}$.

	Preselection		Optimization	
Variation	σ^-	σ^+	σ^-	σ^+
dr	-0.125	0.008	-0.26	0.074
pt	0.02	-0.21	-0.31	-0.38
qfac	0.125	-0.12	-0.28	-0.13
ktfac	0.26	-0.21	0.23	-0.043
iqopt	0.032	—	0.05	—

Table 6.6: Differences in W+jets yields arising from varying the ALPGEN input parameters. The negative variation is arbitrarily defined by reducing the numerical value of the input parameter, and the positive variation is defined by increasing the numerical value of the input parameter. As with the standard sample, sample–specific k factors are derived by scaling the sum of the ALPGEN LO cross sections to the NLO value.

The systematic uncertainty is calculated using Equation 6.1, as with the other fractional systematic uncertainties. The following parameters are considered as a source of systematic uncertainties:

- Renormalization and factorization scale, Q: implemented in ALPGEN through a few choices:

  a real input parameter (**qfac**) that allows to vary overall scale of Q;

  the nominal value for **qfac** = 1 (variation: **qfac** = 0.5, 2);

  the nominal value for **iqopt** = 1 (variation: **iqopt** = 2,3);

  in estimating uncertainty only variation of **qfac** was used to avoid double counting of uncertainty due to renormalization and factorization scale.

- Minimal parton p_T: nominal p_T^{min} = 20 GeV, the variation p_T^{min} = 25 GeV applied symmetrically.

- Cone size $\Delta R_{i,j}$ for parton separation: nominal $\Delta R_{i,j} = 0.7$ and variation $\Delta R_{i,j} = 0.4$, 1.0

- Nodal reweighting parameter **ktfac**: additional partons/showers when MLM matching is used with nominal **ktfac** =1 and variation **ktfac** =0.5, 2.

The results are shown in Table 6.6. The largest absolute deviation is used for the systematic uncertainty giving $\sigma_{sys}^{W+jets} = \pm 40\%$ for the optimized selection criteria.

6.10 $t\bar{t}$ Production Model Uncertainties

Systematic uncertainties are evaluated for the $t\bar{t}$ production model [75]. A 4% uncertainty on the NLO production cross section, discussed in Section 5.1.2 is used. In addition, MC@NLO events are used as an alternative to the default Powheg+Pythia generation. Additional samples with top mass variations of 0.25 GeV, and variations in generator parameters have also been used. Table 6.7 shows the resulting systematic uncertainty.

Preselected	250 GeV Opt	300 GeV Opt	350 GeV Opt
13%	35%	35%	35%

Table 6.7: Systematic uncertainties arising from comparing the default $t\bar{t}$ Monte Carlo sample generated with Powheg+Pythia to a sample generated using MC@NLO. The same values are used for the electron and muon channels.

6.11 Summary

The preceding sections describe the individual sources of systematic uncertainty and how the uncertainties are determined and the values. Table 6.8 shows the summary for the preselected sample, and Table 6.9 shows the summary for the optimized selection.

Source	Production σ	Modeling	Electron Related	JES	JER	Luminosity	Pile-up
V+jets	4%	25%	$< 1\%$	+8%, -13%	12%	11%	6%
$t\bar{t}$	+7%, -10%	13%	$< 1\%$	4%	3.5%	11%	1.2%
Single top	10%	—	—	+2%, -3%	2.6%	11%	2.6%
Dibosons	5%	—	$< 1\%$	+8%, -14%	15%	11%	2%
LQ (300 GeV)	18%	—	$< 1\%$	$< 0.05\%$	$< 3\%$	11%	1%

Table 6.8: A summary of the systematic uncertainties for the preselection. The electron trigger, identification and energy scale and resolution uncertainties are small and grouped together. Some approximations have been made when grouping sources for this table, but full details are available in the text.

Source	Production σ	Modeling	Electron Related	JES	JER	Luminosity	Pile-up
V+jets	4%	40%	$< 5\%$	+22%, -13%	10%	11%	5%
$t\bar{t}$	+7%, -10%	35%	$< 2\%$	+13%, -3%	26%	11%	3%
Single top	10%	—	—	$\pm 33\%$	26%	11%	3%
Dibosons	5%	—	$< 1\%$	+17%, -24%	14%	11%	6%
LQ (300 GeV)	18%	2%	$< 1\%$	$< 3\%$	$< 3\%$	11%	2%

Table 6.9: A summary of the systematic uncertainties for the optimized selection. The electron trigger, identification and energy scale and resolution uncertainties are small and grouped together. Some approximations have been made when grouping sources for this table, but full details are available in the text.

Chapter 7

Limit Setting

Once a set of optimal cuts has been determined and these requirements have been applied to data and MC, final variable distributions are compared for the signal, the background and the observations from data. The final distributions become the input to a calculation of upper limits on the signal model search parameter (a cross section, for example). This chapter focuses on an overview of the statistical methods used to make this distinction.

There are two main statistical approaches that are accepted by the High Energy Physics community [1]:

- Frequentist approach: the result is interpreted by the frequency at which the observed outcome is likely to occur in a series of hypothetical repeatable experiments.
- Bayesian approach: that same result is evaluated to predict where the true value lies based on the best a-priori knowledge of the experimenter.

Either approach can be used to quantify the significance of a result, or to set limits on the production rate of the physical process. The analysis presented implements a "semi-Frequentist" method, used by the Collie [76] statistical package, and will be discussed in greater detail.

7.1 Overview

The final distribution is used to build two hypotheses: a test (H_1) and a null (H_0), which are described below [77]:

- Test Hypothesis, H_1: Also known as the signal+background hypothesis, the data can only be be described by signal+background.
- Null Hypothesis, H_0: Also known as the background only hypothesis, the data are described by by the background only.

The end goal of a search is to be able to reject H_0 in favor of H_1. To do this in an unbiased way, a test variable is used to define a region. This region sets the range of possible values of the test variable in which the null hypothesis would be rejected. The test of compatibility between the data and our hypothesis is quantified by a p-value. It is the probability that the hypothesis in question would fluctuate to the levels observed in data.

7.2 Log-Likelihood Ratio Test

The test statistic that gives the optimal method to distinguish the leptoquark signal from the standard model background is given by the likelihood ratio test [78]. It is utilized in the COLLIE statistical package and shown in Equation 7.1. The likelihood ratio is calculated for the data, H_0 (L(b—x)), and H_1 (L(s+b—x)). Since this is a random distribution rate, the likelihood function is a Poisson probability distribution, shown in Equation 7.2.

$$\Lambda(x) = \frac{\mathcal{L}(s+b|x)}{\mathcal{L}(b|x)} \tag{7.1}$$

$$\mathcal{L}(k|x) = \frac{(k)^x e^{-(k)}}{x!} \tag{7.2}$$

In this equation, x represents the data, b is the estimated background yield, s is the predicted signal yield, and $\mathcal{L}$ is the likelihood function. The combined likelihood over all the channels and all the bins can be obtained by taking the product of the individual likelihood functions. The multiple analysis channels and multiple data bins likelihood ratio test statistic is shown in Equation 7.3.

$$\Lambda(x) = \prod_{i}^{N_{channels}} \prod_{j}^{N_{bins}} \frac{(s_{ij}+b_{ij})^{x_{ij}} e^{-(s_{ij}+b_{ij})}}{x_{ij}!} / \frac{(b_{ij})^{x_{ij}} e^{-(b_{ij})}}{x_{ij}!}. \tag{7.3}$$

The most common way to represent this function is to take the negative logarithm and multiply by two. This yields a log-likelihood ratio test (LLR) and is shown in Equation 7.4.

$$\mathrm{LLR(x)} = -2\log(\Lambda(\mathrm{x})) = \sum_{\mathrm{i}}^{\mathrm{N_{channels}}} \sum_{\mathrm{j}}^{\mathrm{N_{bins}}} \left[\mathrm{s_{ij}} - \mathrm{x_{ij}} \ln\left(1 + \frac{\mathrm{s_{ij}}}{\mathrm{b_{ij}}}\right)\right] \tag{7.4}$$

7.3 Confidence Levels and p-Values

The p-value is the probability that the background only hypothesis fluctuates so as to at least the observed data. This can be calculated using the previously defined LLR variables and taking the integral of the pseudo-experiment LLR distribution above the value in data.

In contrast, the confidence level (CL) gives the guaranteed lower bound on the range of values which contain a measurement's true value [78]. A probability density function (pdf) can be parametrized against the LLR test statistic (x) and some unknown parameter (e.g. the signal cross section, (σ: P(x|σ)). The pdf can be used in Equation 7.5 where P the probability that the signal would yield our observed LLR for any given cross section. Pseudo-experiments are performed to determine the frequency of the LLR value for the measured data.

$$P(x_- < x < x_+|\sigma) = \int_{x_+}^{x_-} P(x|\sigma)dx = CL \tag{7.5}$$

In the case where the unknown parameter is bounded by 0 (like for the signal cross section), the CL and p-value have the same interpretation. Confidence levels can be determined from the LLR integral for the H_0 and H_1 hypotheses, which are called CL_B and CL_{S+B} and are written in Equations 7.6 and 7.7, respectively. They are the fraction of H_0 and H_1 pseudo-experiments with LLR values larger that the one observed in data.

$$\mathrm{CL_{S+B}} = \int_{\mathrm{LLR(s+b|x)}}^{\infty} \mathbf{P}(\mathrm{s+b|x'})\mathrm{d(LLR(s+b|x'))} \tag{7.6}$$

$$\mathrm{CL_{B}} = \int_{\mathrm{LLR(b|x)}}^{\infty} \mathbf{P}(\mathrm{b|x'})\mathrm{d(LLR(b|x'))} \tag{7.7}$$

In each experiment, the data is sampled from a Poisson distribution whose argument is either the expected background yield or the sum of the signal and background for the background-only and signal+background hypotheses, respectively. An example of these distributions are shown in Figure 7.1.

7.3.1 Treatment of Systematics

Neither the background nor the signal can be predicted precisely due to the systematic uncertainties in the analysis. In order to take these systematic uncertainties into account, the Bayesian approach is more natural than the Frequentist. Systematic uncertainties, which are discussed in Chpater 6, are introduced into the LLR calculation as nuisance parameters. This approach uses a Gaussian distribution for each prior (systematic uncertainty) unless an asymmetric uncertainty is given (e.g. the JES Systematic), in this case a bifurcated Gaussian is used. Each systematic is sampled within its $\pm 1\sigma$ uncertainty. The Bayesian posterior is the product of the likelihoods and is defined in Equation 7.8, where *sys* represents the systematics.

$$\mathbf{P}(s|x) = \int\int_{sys} \frac{\mathbf{P}(x|s,b,sys) \times \mathbf{P}(s) \times \mathbf{P}(sys)}{\mathbf{P}(x)} \tag{7.8}$$

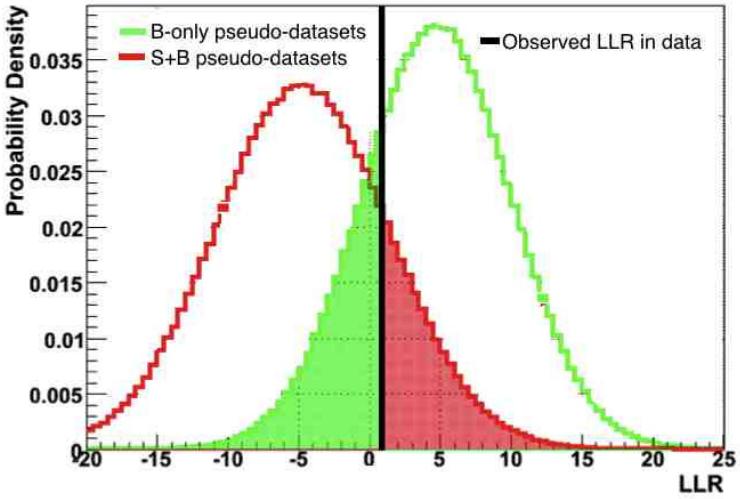


Figure 7.1: CL_{S+B} and CL_{B} [76]

Correlated systematics between backgrounds and the signal are commonly sampled (e.g. luminosity) whereas uncorrelated systematics (e.g. MJ normalization) are sampled independently. Confidence levels are defined by integrating the normalized probability distribution of LLR values from the observed LLR value to infinity.

7.4 Modified Frequentist approach

When using CL_{S+B} there is an inherent risk brought about by a dependence on the modeling of the background. For example, if fewer events are observed events than predicted by the background model, the limit gets better. The Modified-Frequentist approach offers a solution to this feature by defining a new confidence level, CL_S [77]. It is defined as the ratio of the confidence levels for the signal+background (CL_{S+B}) and background-only (CL_B) hypotheses as shown in Equation 7.9. This removes the dependence on background modeling.

$$\mathrm{CL_S} = \frac{\mathrm{CL_{S+B}}}{\mathrm{CL_B}} \tag{7.9}$$

7.5 Measuring a Cross Section

The ultimate purpose of the statistical analysis is to be able to state if a statistically significant observation of the signal has been made. If so, then a measurement of the cross

section can be calculated; if not, an upper-bound on the production cross section is determined.

When a excess is observed in the data, the LLR test statistic can be used to determine the cross section. To start, the signal cross section is incrementally shifted over a range of values. For each of these, a new set of H_1 pseudo-experiments are generated. The 1-CL_S values at each point forms a continuous distribution. The differential of this distribution is a Gaussian centered on the measured cross section. The H_0 hypothesis can then be excluded at the 5σ confidence level when 1-CL_B equals $1 - 2.8 \times 10^{-7}$. In the event that no data excess is observed, the production cross section is excluded at 95% confidence level when 1-CL_S equals 0.95.

Chapter 8

Results

The optimized selection criteria outlined in Section 5.5 are applied to the data and MC for the background and leptoquark signal with mass between 200 and 400 GeV (in 50 GeV increments). To provide the best signal vs. background discrimination several variables were investigated and tested as inputs to the significance calculation described in Section 7. The leptoquark invariant mass ($M_{LQ}(e, jet)$) distribution (shown for a 300 GeV leptoquark in Figure 8.1) was chosen as the input because of the clear peak in the leptoquark mass providing the most sensitivity, while the shape is flat for the background. The distributions of the remaining data, simulated signal and predicted background are discussed in the following sections.

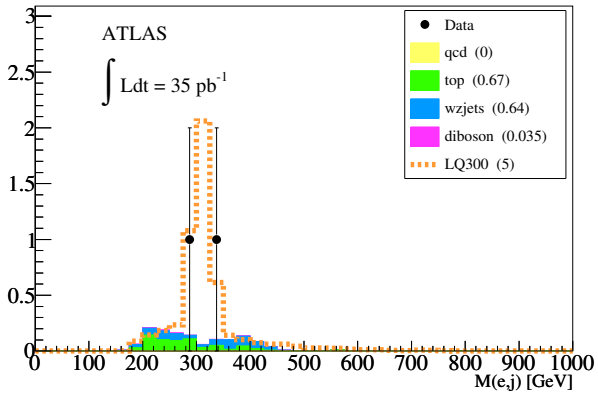


Figure 8.1: Input $M_{LQ}(e, jet)$ Distribution for limit calculation. This variable provides the most sensitivity for the leptoquark search.

8.1 Remaining Events and Event Displays

The results of each selection criteria (listed in Table 5.4) to data and MC prediction are shown in Table 8.1. The 300 GeV optimization is used for all leptoquark masses. Additional results for optimizations for 250 GeV and 350 GeV generated leptoquark events are shown in Tables A.3 and A.5, but not used.

	Selection Yields			
Source	M_T^{LQ1}	M^{LQ2}	$M_T(e, \not{E}_T)$	S_T
$W \to e\nu$+jets	160.6 ± 3.8	65.1 ± 2.4	0.65 ± 0.23	0.65 ± 0.23
$W \to \tau\nu$+jets	7.43 ± 0.85	3.64 ± 0.61	—	—
$Z \to ee$+jets	1.15 ± 0.32	0.75 ± 0.26	—	—
$Z \to \tau\tau$+jets	1.01 ± 0.30	0.14 ± 0.11	—	—
$t\bar{t}$	42.43 ± 0.81	17.83 ± 0.53	0.76 ± 0.11	0.61 ± 0.10
single top (s+t+tW)	4.10 ± 0.38	1.47 ± 0.23	0.06 ± 0.05	0.06 ± 0.05
Diboson	2.49 ± 0.07	1.12 ± 0.05	$(67.4 \pm 9.8) \times 10^{-3}$	$(63.4 \pm 9.6) \times 10^{-3}$
Multijet	6.52 ± 0.78	3.80 ± 0.46	—	—
Total Bkg.	225.7 ± 4.1	93.8 ± 2.6	1.54 ± 0.26	1.38 ± 0.26
Data	217	82	2	2
$LQ\,LQ \to eq\nu q'$ 250	22.52 ± 0.29	20.71 ± 0.29	9.71 ± 0.22	9.62 ± 0.22
$LQ\,LQ \to eq\nu q'$ 300	9.84 ± 0.11	9.27 ± 0.11	5.07 ± 0.09	5.05 ± 0.09
$LQ\,LQ \to eq\nu q'$ 350	4.42 ± 0.04	4.24 ± 0.04	2.56 ± 0.04	2.56 ± 0.04
$LQ\,LQ \to eqeq$ 250	1.83 ± 0.15	1.62 ± 0.14	0.24 ± 0.05	0.24 ± 0.05
$LQ\,LQ \to eqeq$ 300	0.98 ± 0.06	0.93 ± 0.06	0.24 ± 0.03	0.24 ± 0.03
$LQ\,LQ \to eqeq$ 350	0.50 ± 0.03	0.48 ± 0.03	0.15 ± 0.02	0.15 ± 0.02

Table 8.1: Optimized selection event selection yields for a 300 GeV LQ. The selection for a given column includes the requirements of columns to the left. The LQLQ signal samples with neutrinos assume a branching fraction, β, of 0.5, and the sample with no neutrinos assumes β=1.0. Table A.4 in Appendix A gives a detailed listing in terms of the subsamples merged to form this table.

The resulting distributions after the event selection are shown in Figure 8.2. The two remaining data events are shown in the black points, whereas the standard model background is shown in the color histograms. The black dotted, solid, and dashed lines are simulated leptoquark signal events generated at 250, 300 and 350 GeV respectively. The two events that pass selection have their kinematic properties listed in Table 8.2. The Atlantis-based event displays for these two events are shown in Figures 8.3 and 8.4 and were generated using Atlantis, an event display tool.

8.2 Combination with eejj Channel

The ATLAS efforts on the search for leptoquarks, however, did not end with the search for leptoquarks which decay into an electron/quark pair and a neutrino/quark pair (i.e.: the eνjj channel). One assumption made throughout the analysis is that leptoquarks conserve generation. This provided a natural way to combine different search channels to explore

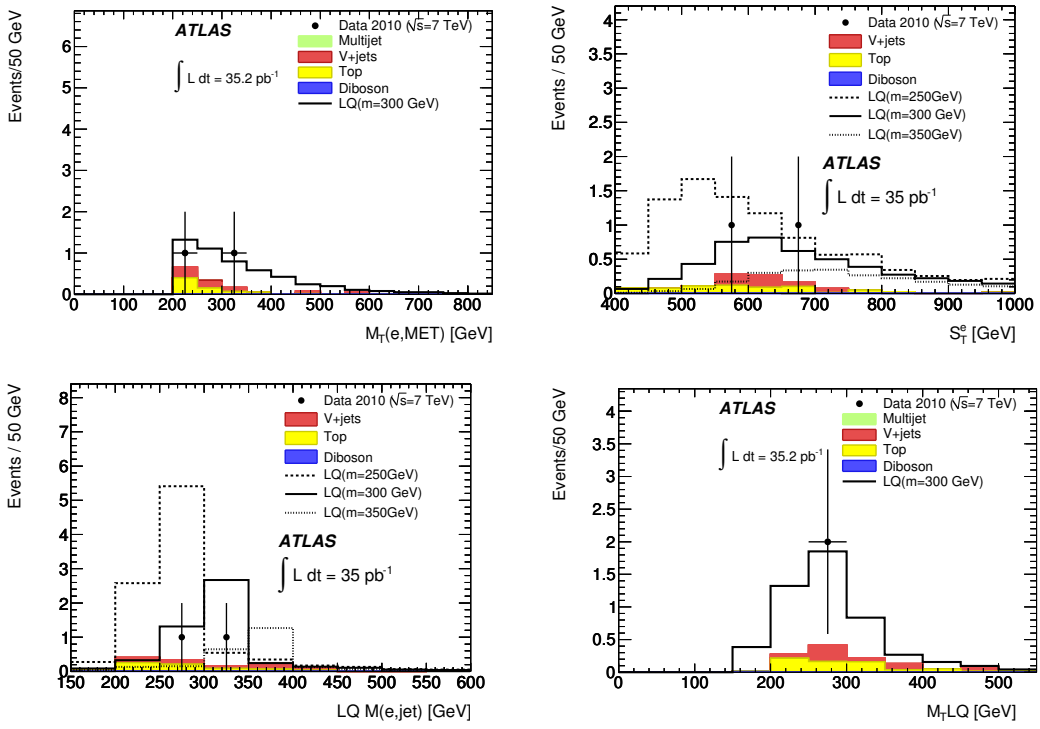


Figure 8.2: Signal Region Event Plots. The variables used in the optimization are plotted. M_T(e,$\not\!\!E_T$) is in the upper left and has a kinematic cut of 200 GeV. S_T is in the upper right and has a kinematic cut at 410 GeV. The invariant leptoquark mass is in the lower left and has a cut at 180 GeV and the transverse leptoquark mass is in the lower right and also has a cut at 180 GeV. The two remaining events are shown in the black points, whereas the standard model background is shown in the color histograms. The black dotted, solid, and dashed lines are simulated leptoquark signal events generated at 250, 300 and 350 GeV respectively.

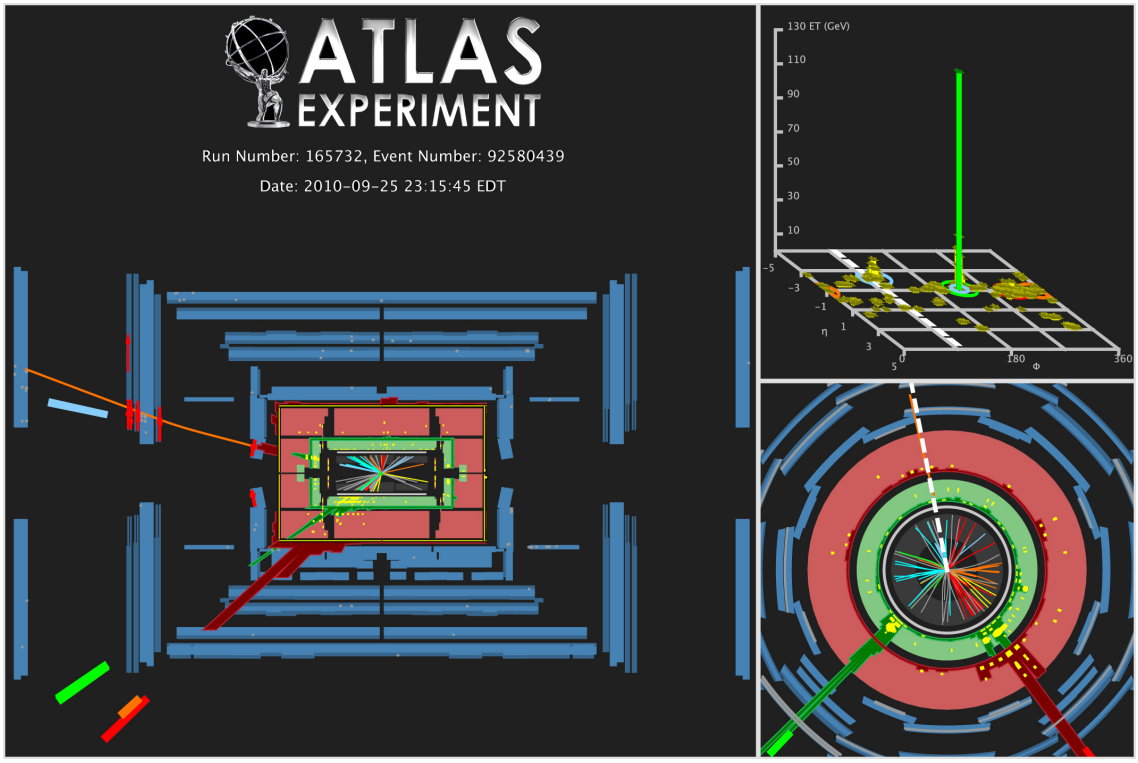


Figure 8.3: Event display for one of the two events passing the final selection. The run and event numbers are 165732 and 92580439.

in more detail β vs. mass space, where β is the branching fraction of leptoquarks into charged leptons (i.e. electrons). If the probability that leptoquarks decay into both neutrinos and electrons is equal then the eνjj channel has the highest probability of being observed. However, if they do not decay to both equally, other decay modes could be more sensitive. This opens up a search in which both leptoquarks decay into an electron and a quark, called the eejj channel.

Specifically the eejj channel results are combined with the analysis presented here (eνjj) and then used to set a cross section limit on leptoquark decay into first-generation fermions [18]. The signal significance for the cross section is calculated for each channel and each β. These cross section limits are then converted to: $\sigma_{LQ\bar{L}Q} \times \beta$ vs. mass plane. Two examples of $\sigma_{LQ\bar{L}Q} \times \beta$ vs. mass plots are shown in Figure 8.5 for β=0.5 (maximum for eνjj channel) and for β=1.0 (maximum for eejj channel). The figures show the theoretical cross section as a hatched blue line, the expected limit in a dashed blue with $\pm 1\sigma$ and $\pm 2\sigma$ in yellow and green respectively and the observed limit in solid red. The total integrated luminosity used was 35 pb^{-1}.

Several different β results are then combined to make the β vs. mass plane shown in

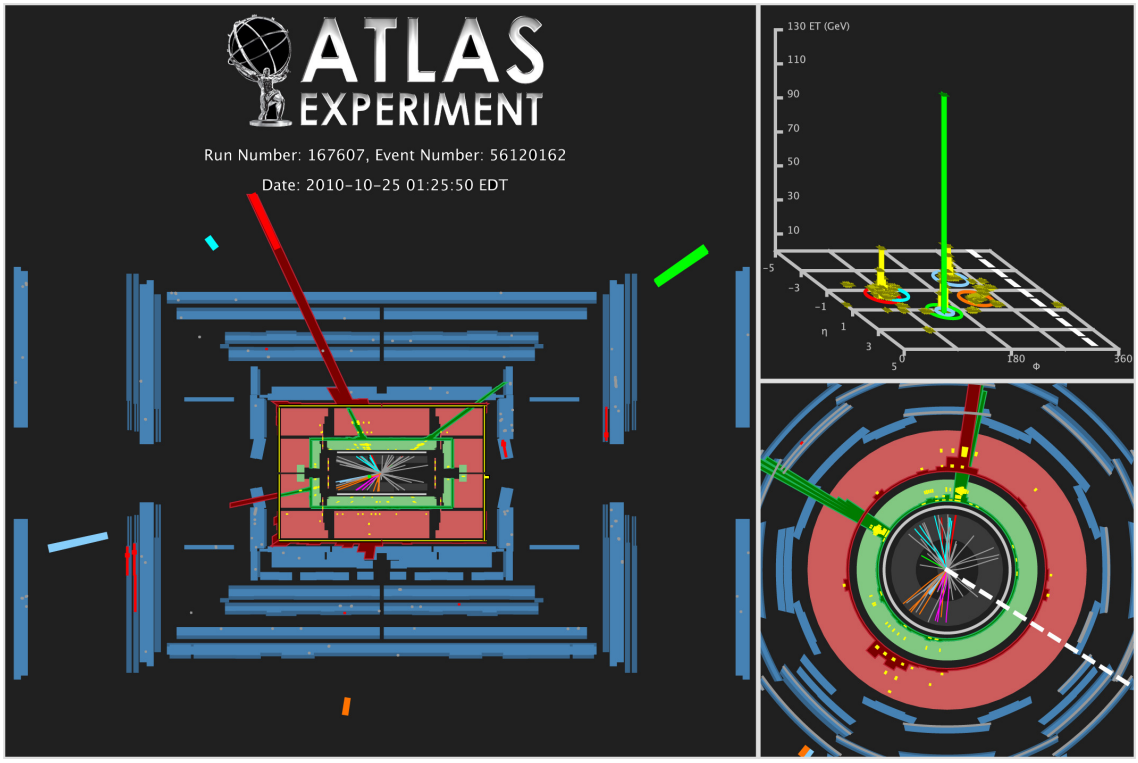


Figure 8.4: Event display for the one of the events passing the final selection. The run and event numbers are 167607 and 56120162.

Figure 8.6. The Figure shows the D0 result in gray, the eνjj in dashed pink and the eejj shown in dotted pink. The theoretical results with $\pm 1\sigma$ are in dashed blue and yellow bands and the observed limit is shown in solid red. The CMS results which include only eνjj channel because these were the only results ready at the time of publication are also shown in blue dotted line.

8.3 Summary

The results from the search for eνjj combined with the results from the eejj search for pair production of scalar leptoquarks have been reported, and no signal excess has been observed. This shows no evidence for leptoquark production, and 95% CL lower bounds on the LQ cross section are determined to be 319 GeV for the branching fraction, β, of a leptoquark to an electron and quark of $\beta = 0.5$. These results are published in [18].

Run (Event) Number	165732 (92580439)	167607 (56120162)
electron p_T	127 GeV	132 GeV
electron ϕ	-2.34	2.62
electron η	-1.15	1.11
$\not{E}_T$	104 GeV	230 GeV
$\not{E}_T\phi$	1.76	-0.58
Number of Jets	3	3
Jet 1, p_T	175 GeV	186 GeV
ϕ	-0.91	1.42
η	-0.91	-0.50
Jet 2, p_T	153 GeV	142 GeV
ϕ	1.9	-2.2
η	-2.2	-2.2
M_T^{LQ}	262 GeV	263 GeV
M^{LQ}	289 GeV	333 GeV
M_T	204 GeV	348 GeV
S_T	558 GeV	689 GeV

Table 8.2: Kinematic properties of events passing the optimized selection requirements.

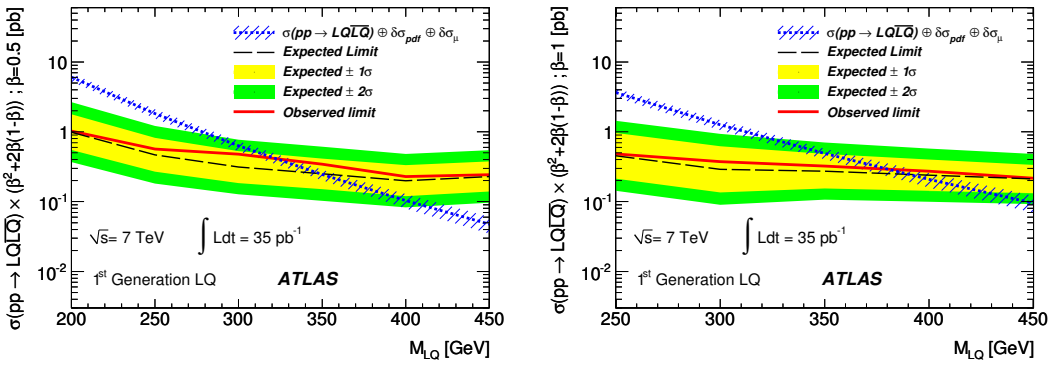


Figure 8.5: Limits for β=0.5, 1.0. β is defined as the percentage of time that leptoquarks decay into a charged lepton and a quark. The theoretical cross section as a hatched blue line, the expected limit in a dashed blue with $\pm1\sigma$ and $\pm2\sigma$ in yellow and green respectively with the observed limit in solid red. The total integrated luminosity used was 35 pb^{-1}.

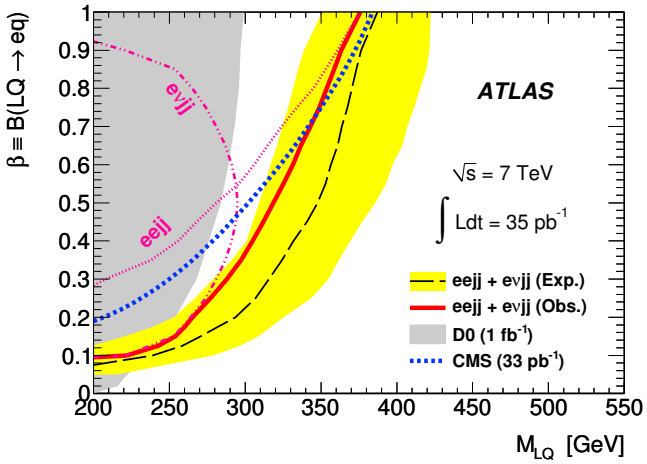


Figure 8.6: LQ combined limit. The previous limit from D0 is shown in gray, the electron+$\not\!\!E_T$ ($e\nu$jj) in dashed pink and the di-electron (eejj) in dotted pink. The theoretical results are shown in dashed blue, and with $\pm 1\sigma$ in the yellow bands and the observed limit is shown in solid red. The CMS results are also shown in blue dotted line.

Chapter 9

Conclusions

Results are presented from a search for pair production of scalar leptoquarks using the entire 2010 ATLAS data set. Searching for particular topologies allows an analysis to be sensitive to different models which is paramount to beyond the standard model searches. Relying too heavily on a particular model may cause us to overlook the region of phase space where new physics lies. The assumption of scalar production frees the search from unknown couplings leaving it to rely on the following two additional assumptions: leptoquarks conserve generation and are produced in pairs. The good agreement between between data and predicted background in predefined low signal regions demonstrates our ability to model, search and measure the cross sections of known Standard Model physics within the first year of operation of the LHC. Since this level of agreement was obtained, an optimized selection criteria was derived using MC simulated events and applied to data. After the selection was applied, agreement between the data and background-only prediction translates into bounds in the plane of branching fraction LQ→eq versus LQ mass. These results are published in [18].

The limiting factor of this analysis was controlling the systematics, specifically the W+jets and $t\bar{t}$ production model uncertainties and jet energy scale uncertainties. In future analyses, the modeling uncertainties can be reduced by additional MC production in the tails of the W+jets and $t\bar{t}$ distributions in the signal region. This will provide smaller fluctuations in regions of low statistics. The jet energy scale uncertainties have been reduced already after the publication of these results by better understanding of the calorimeter response and using more constraints by data. However, the most powerful tool available to push the bounds of the limits is the immense amount of data being collected by ATLAS since the start of the 2011 run. The reduction of these dominant systematics as well as inclusions of orders of magnitude more of data will allow more stringent limits and hopefully a convincing discovery of these particles.

The Standard Model of particle physics exhibits striking similarities between quarks and leptons which motivates the search for this additional symmetry known as leptoquarks. The equivalent of the entire ATLAS 2010 dataset (of 35 pb^{-1}) is collected in a single day of

consistent running and by the end of the 2012 run it is predicted that 5 fb^{-1} of data to be collected. This will allow 5σ discovery well into the TeV range. Additionally more time spent understanding the intricacies of the detector will allow searches in the additional channels of $\not{E}_T$ (neutrino-only) production (β near 0), and the third generation (decay into τ and top). The inclusion of the third generation will give insight to the charge of the leptoquarks since the bottom and top quarks can be distinguished. Searching in this extended region of phase space into the TeV range will surely unlock many secrets to unknowns beyond the Standard Model.

Bibliography

[1] K. Nakamura et al. (Particle Data Group), J. Phys. G **37**, 075021 (2010), URL `http://pdg.lbl.gov/index.html`.

[2] D. Perkins, *Introduction to High Energy Physics* (Cambridge University Press, 2000), ISBN 9780521621960.

[3] F. Halzen and A. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics* (John Wiley and Sons, New Jersey, 1984), ISBN 0-471-88741-2.

[4] E. D. Commins and P. H. Bucksbaum, *Weak Interactions of Leptons and Quarks* (Cambridge University Press, Cambridge, 1983), ISBN 0 521 23092 6.

[5] S. Dawson, in *Electroweak Symmetry Breaking* (1998), lectures given at the 1998 Summer School in High Energy Physics and Cosmology, `arXiv:hep-ph/9901280v1`.

[6] S. Davidson, D. Bailey, and B. A. Campbell, Zeitschrift Für Physik C **61**, 613 (1994).

[7] M. Krämer, T. Plehn, M. Spira, and P. M. Zerwas, Phys. Rev. D **71**, 057503 (2005), `hep-ph/0411038v2`.

[8] A. Belyaev, C. Leroy, R. Mehdiyev, and A. Pukhov, JHEP **09**, 005 (2005), `hep-ph/0502067`.

[9] D. E. Acosta and S. K. Blessing, Annual Review of Nuclear and Particle Science **49**, 389 (1999), URL `http://www.annualreviews.org/doi/abs/10.1146/annurev.nucl.49.1.389`.

[10] J. C. Pati and A. Salam, Phys. Rev. D **10**, 275 (1974).

[11] H. Georgi and S. L. Glashow, Phys. Rev. Lett. **32**, 438 (1974).

[12] E. Witten, Nucl. Phys. **B258**, 75 (1985).

[13] M. Dine et al., Nucl. Phys. **B259**, 549 (1985).

[14] J. L. Hewett and T. G. Rizzo, Phys. Rev. **D56**, 5709 (1997), `hep-ph/9703337`.

[15] K. Lane (2002), hep-ph/0202255.

[16] J. Blumlein, E. Boos, and A. Kryukov, Z. Phys. **C76**, 137 (1997), hep-ph/9610408.

[17] M. Krämer, T. Plehn, M. Spira, and P. M. Zerwas, Phys. Rev. Lett. **79**, 341 (1997), hep-ph/9704322.

[18] G. Aad et al. (ATLAS) (2011), 1104.4481.

[19] V. M. Abazov et al. (D0), Phys. Lett. **B681**, 224 (2009), 0907.1048.

[20] S. Chatrchyan et al. (CMS) (2011), 1105.5237.

[21] *Atlas experiment* (2011), http://www.atlas.ch/photos/index.html.

[22] C. Hill, *The cern hadron ion sources* (1999), c Hill for the Linac machine supervisors, http://linac2.home.cern.ch/linac2/sources/source.htm.

[23] G. Aad et al. (ATLAS), Eur. Phys. J. **C71**, 1630 (2011), 1101.2185.

[24] P. Jenni and M. Nessi, Tech. Rep. CERN-LHCC-2004-010. LHCC-I-014, CERN, Geneva (2004).

[25] G. Aad et al. (ATLAS), Tech. Rep. ATLAS-CONF-2011-011, CERN, Geneva (2011).

[26] D. Dobos and H. Pernegger, Tech. Rep. ATL-COM-INDET-2010-034, CERN, Geneva (2010).

[27] S. Haywood, L. Rossi, R. Nickerson, and A. Romaniouk, *ATLAS Inner Detector: Technical Design Report, 2*, Technical Design Report ATLAS (CERN, Geneva, 1997).

[28] F. Djama, Tech. Rep. ATL-COM-INDET-2010-132, CERN, Geneva (2010).

[29] N. Barlow, Tech. Rep. ATL-COM-INDET-2010-003, CERN, Geneva (2010).

[30] J. Degenhardt, Tech. Rep. ATL-COM-INDET-2010-149, CERN, Geneva (2010).

[31] G. Aad et al., *Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics* (CERN, Geneva, 2009).

[32] *ATLAS Liquid Argon Calorimeter: Technical Design Report*, Technical Design Report ATLAS (CERN, Geneva, 1996).

[33] *ATLAS Tile Calorimeter: Technical Design Report*, Technical Design Report ATLAS (CERN, Geneva, 1997).

[34] G. Aad et al. (ATLAS), Eur. Phys. J. **C70**, 723 (2010), 0912.2642.

[35] T. Barillari et al., Tech. Rep. ATL-LARG-PUB-2009-001-2. ATL-COM-LARG-2008-006. ATL-LARG-PUB-2009-001, CERN, Geneva (2008).

[36] M. Aurousseau, R. Caputo, J. Hobbs, M. Kado, S. Laplace, D. Tsybychev, J.-B. de Vivie de Regie, and A. Yurkewicz, Tech. Rep. ATL-COM-LARG-2009-052, CERN, Geneva (2009).

[37] P. Pralavorio, Tech. Rep. ATL-COM-PHYS-2010-722, CERN, Geneva (2010).

[38] A. Foussat et al., IEEE Transactions on Applied Superconductivity **18**, 367 (2008), ISSN 1051-8223.

[39] M. Aleksa et al., Journal of Instrumentation **3**, P04003 (2008), URL `http://stacks.iop.org/1748-0221/3/i=04/a=P04003`.

[40] *ATLAS Central Solenoid: Technical Design Report*, Technical Design Report ATLAS (CERN, Geneva, 1997).

[41] A. Dudarev et al., IEEE Transactions on Applied Superconductivity **19**, 1307 (2009), ISSN 1051-8223.

[42] W. Liebig and D. Orestano, Tech. Rep. ATLAS-COM-CONF-2010-065, CERN, Geneva (2010).

[43] R. Voss and J. Wotschack, *ATLAS Muon Spectrometer: Technical Design Report*, Technical Design Report ATLAS (CERN, Geneva, 1997).

[44] *ATLAS Level-1 Trigger: Technical Design Report*, Technical Design Report ATLAS (CERN, Geneva, 1998).

[45] *ATLAS High-Level Trigger, Data Acquisition and Controls: Technical Design Report*, Technical Design Report ATLAS (CERN, Geneva, 2003).

[46] J. Stelzer, Tech. Rep. ATL-COM-DAQ-2011-011, CERN, Geneva (2011).

[47] Tech. Rep. ATLAS-PHYS-PUB-2010-002, CERN, Geneva (2011).

[48] T. Sjostrand, S. Mrenna, and P. Z. Skands, JHEP **05**, 026 (2006), `hep-ph/0603175`.

[49] G. Corcella et al., Journal of High Energy Physics. **1**, 010 (2001), URL `http://dro.dur.ac.uk/1548/`.

[50] J. Butterworth, J. Forshaw, and M. Seymour, Zeitschrift für Physik C Particles and Fields **72**, 637 (1996), ISSN 0170-9739, 10.1007/s002880050286, URL `http://dx.doi.org/10.1007/s002880050286`.

[51] M. L. Mangano et al., JHEP **07**, 001 (2003), hep-ph/0206293.

[52] S. Hoeche et al. (2006), hep-ph/0602031.

[53] S. Frixione and B. R. Webber, JHEP **06**, 029 (2002), hep-ph/0204244.

[54] S. Frixione, P. Nason, and C. Oleari, JHEP **11**, 070 (2007), 0709.2092.

[55] S. Agostinelli et al. (GEANT4), Nucl. Instrum. Meth. **A506**, 250 (2003).

[56] T. Cornelissen et al., Tech. Rep. ATLAS-SOFT-PUB-2007-007, CERN, Geneva (2007).

[57] G. Piacquadio, K. Prokofiev, and A. Wildauer, Journal of Physics: Conference Series **119**, 032033 (2008), URL http://stacks.iop.org/1742-6596/119/i=3/a=032033.

[58] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 063 (2008), 0802.1189.

[59] Tech. Rep. ATLAS-CONF-2010-056, CERN, Geneva (2010).

[60] Tech. Rep. ATLAS-CONF-2010-005, CERN, Geneva (2011).

[61] R. Nicolaidou et al., Journal of Physics: Conference Series **219**, 032052 (2010), URL http://stacks.iop.org/1742-6596/219/i=3/a=032052.

[62] M. Bellomo et al., Tech. Rep. ATLAS-COM-CONF-2011-068, CERN, Geneva (2011).

[63] A. Ventura et al., in *Nuclear Science Symposium Conference Record, 2004 IEEE* (2004), vol. 3, pp. 1530 –1534 Vol. 3, ISSN 1082-3654.

[64] Tech. Rep. ATLAS-CONF-2010-038, CERN, Geneva (2010).

[65] Tech. Rep. ATLAS-CONF-2010-057, CERN, Geneva (2010).

[66] G. Aad et al. (ATLAS), Phys. Lett. **B698**, 325 (2011), 1012.5382.

[67] G. Aad et al., Journal of High Energy Physics **2010**, 1 (2010), URL http://dx.doi.org/10.1007/JHEP12(2010)060.

[68] S. Moch and P. Uwer, Phys. Rev. D **78**, 034003 (2008).

[69] J. M. Campbell, R. K. Ellis, and F. Tramontano, Phys. Rev. **D70**, 094012 (2004), hep-ph/0408158.

[70] N. Kidonakis, PoS **DIS2010**, 196 (2010), 1005.3330.

[71] J. Butterworth et al., Tech. Rep. ATL-COM-PHYS-2010-695, CERN, Geneva (2010).

[72] N. Amos et al., in *Proceedings of the International Conference on Computing in High Energy Physics* (1995), URL `http://www.hep.net/chep95/abstract/abs_155.htm`.

[73] Tech. Rep. ATLAS-CONF-2010-054, CERN, Geneva (2010).

[74] Tech. Rep. ATLAS-CONF-2011-007, CERN, Geneva (2011).

[75] G. Aad et al. (ATLAS), Eur. Phys. J. **C71**, 1577 (2011), `1012.1792`.

[76] W. Fisher (D0), Tech. Rep. FERMILAB-TM-2386-E, Fermilab (2006).

[77] T. Junk, Nucl. Instrum. Meth. **A434**, 435 (1999), `hep-ex/9902006`.

[78] R. Barlow, *Statistics: A Guide to The Use of Statistical Methods in the Physical Sciences*, Manchester Physics Series (Wiley, 1989), ISBN 9780471922957.

Appendix A

Expanded Tables

This appendix has alternative versions of yield and systematics tables shown in the text. The versions here show results for all subsamples prior to the merging used to produce the simplified tables in the body of the thesis

Source	Selection Yields							
	Good Event	Lepton	$\not\!\!E_T$	M_T	≥ 1 jet	Triangle	μ veto	$\geq 2jet$
$W \to e\nu$+0p	174974 ± 122	116650 ± 122	97991 ± 117	97958 ± 117	4321 ± 30	3475 ± 27	3475 ± 27	361.3 ± 8.7
$W \to e\nu$+1p	38369 ± 50	24568 ± 54	18553 ± 51	17995 ± 51	13572 ± 47	11152 ± 43	11152 ± 43	635 ± 12
$W \to e\nu$+2p	12084 ± 16	7309 ± 18	5567 ± 17	5243 ± 17	4958 ± 17	3817 ± 16	3817 ± 16	2370 ± 13
$W \to e\nu$+3p	3533.9 ± 7.4	2015 ± 10	1540.7 ± 9.2	1401.9 ± 9.0	1383.8 ± 9.0	1069.4 ± 8.3	1069.1 ± 8.3	933.9 ± 7.9
$W \to e\nu$+4p	937.0 ± 3.3	509.5 ± 4.8	392.2 ± 4.6	348.7 ± 4.4	347.9 ± 4.4	269.2 ± 4.1	269.1 ± 4.1	259.3 ± 4.1
$W \to e\nu$+5p	298.2 ± 1.6	157.9 ± 2.8	121.3 ± 2.7	105.3 ± 2.6	105.2 ± 2.6	86.4 ± 2.5	86.4 ± 2.5	85.4 ± 2.4
$W \to \tau\nu$+0p	28855 ± 75	3644 ± 28	1985 ± 21	1982 ± 21	94.4 ± 4.5	76.0 ± 4.0	76.0 ± 4.0	7.0 ± 1.2
$W \to \tau\nu$+1p	13804 ± 47	1178 ± 16	602 ± 11	389.3 ± 9.1	295.4 ± 8.0	265.2 ± 7.5	264.9 ± 7.5	15.1 ± 1.8
$W \to \tau\nu$+2p	6529 ± 18	413.2 ± 5.8	255.4 ± 4.6	127.1 ± 3.3	118.7 ± 3.2	96.6 ± 2.8	96.1 ± 2.8	59.0 ± 2.2
$W \to \tau\nu$+3p	2438 ± 10	126.4 ± 3.2	86.8 ± 2.7	33.9 ± 1.7	33.2 ± 1.7	26.9 ± 1.5	26.6 ± 1.5	23.0 ± 1.4
$W \to \tau\nu$+4p	748.9 ± 4.4	31.5 ± 1.6	24.3 ± 1.4	9.87 ± 0.90	9.87 ± 0.90	8.58 ± 0.84	8.58 ± 0.84	8.41 ± 0.84
$W \to \tau\nu$+5p	260.2 ± 2.1	10.55 ± 0.91	8.78 ± 0.84	3.56 ± 0.54	3.56 ± 0.54	2.96 ± 0.49	2.96 ± 0.49	2.96 ± 0.49
$Z \to ee$+0p	22919 ± 23	10346 ± 26	245.3 ± 4.9	213.9 ± 4.6	119.5 ± 3.4	20.5 ± 1.4	20.5 ± 1.4	1.93 ± 0.44
$Z \to ee$+1p	5206.8 ± 8.2	2464 ± 12	94.6 ± 3.0	83.0 ± 2.8	76.9 ± 2.7	18.8 ± 1.3	18.8 ± 1.3	6.69 ± 0.80
$Z \to ee$+2p	1641.4 ± 4.0	749.4 ± 6.4	45.7 ± 2.0	40.0 ± 1.9	39.2 ± 1.9	9.8 ± 1.0	9.8 ± 1.0	7.72 ± 0.85
$Z \to ee$+3p	469.4 ± 1.7	214.5 ± 3.4	22.6 ± 1.4	19.8 ± 1.3	19.5 ± 1.3	5.59 ± 0.71	5.37 ± 0.70	4.92 ± 0.67
$Z \to ee$+4p	126.07 ± 0.79	59.3 ± 1.7	8.08 ± 0.82	6.79 ± 0.76	6.79 ± 0.76	2.27 ± 0.45	2.27 ± 0.45	2.27 ± 0.45
$Z \to ee$+5p	32.85 ± 0.34	15.17 ± 0.77	3.09 ± 0.44	2.46 ± 0.40	2.46 ± 0.40	0.88 ± 0.24	0.88 ± 0.24	0.88 ± 0.24
$Z \to \tau\tau$+0p	6688 ± 23	921.3 ± 9.4	177.9 ± 4.2	158.8 ± 4.0	23.9 ± 1.5	6.68 ± 0.81	5.31 ± 0.73	0.30 ± 0.17
$Z \to \tau\tau$+1p	2420 ± 12	289.7 ± 5.1	97.3 ± 3.0	35.4 ± 1.8	29.5 ± 1.7	22.3 ± 1.5	18.6 ± 1.3	2.26 ± 0.46
$Z \to \tau\tau$+2p	1003.4 ± 6.5	104.0 ± 3.0	48.7 ± 2.1	13.4 ± 1.1	13.0 ± 1.1	9.05 ± 0.92	7.84 ± 0.86	5.69 ± 0.73
$Z \to \tau\tau$+3p	334.7 ± 3.2	32.0 ± 1.7	19.6 ± 1.3	4.27 ± 0.62	4.27 ± 0.62	3.16 ± 0.54	2.67 ± 0.49	2.42 ± 0.47
$Z \to \tau\tau$+4p	102.8 ± 1.4	8.82 ± 0.86	6.28 ± 0.73	1.27 ± 0.34	1.27 ± 0.34	0.76 ± 0.26	0.76 ± 0.26	0.76 ± 0.26
$Z \to \tau\tau$+5p	29.69 ± 0.54	2.59 ± 0.41	1.99 ± 0.36	0.40 ± 0.17	0.40 ± 0.17	0.25 ± 0.13	0.19 ± 0.12	0.19 ± 0.12
$t\bar{t}$	2773.2 ± 2.3	709.3 ± 2.9	601.5 ± 2.8	496.6 ± 2.6	495.1 ± 2.6	435.3 ± 2.4	381.6 ± 2.3	372.9 ± 2.3
single top (s)	198.7 ± 1.0	123.3 ± 1.5	101.7 ± 1.5	92.3 ± 1.5	87.7 ± 1.4	75.4 ± 1.4	75.3 ± 1.4	47.9 ± 1.2
single top (t)	13.43 ± 0.06	7.86 ± 0.08	6.47 ± 0.08	5.81 ± 0.08	5.62 ± 0.08	4.77 ± 0.08	4.74 ± 0.08	3.60 ± 0.07
single top (Wt)	391.8 ± 1.4	64.1 ± 1.4	52.3 ± 1.3	43.4 ± 1.2	42.6 ± 1.2	36.5 ± 1.1	32.1 ± 1.0	29.2 ± 1.0
WW	396.62 ± 0.61	176.23 ± 0.56	137.76 ± 0.52	127.35 ± 0.51	103.88 ± 0.47	84.05 ± 0.43	79.96 ± 0.42	47.87 ± 0.33
WZ	127.91 ± 0.19	33.18 ± 0.15	10.94 ± 0.09	8.74 ± 0.08	7.28 ± 0.07	5.53 ± 0.06	4.76 ± 0.06	2.07 ± 0.04
ZZ	123.55 ± 0.19	47.76 ± 0.17	32.51 ± 0.15	29.28 ± 0.14	22.39 ± 0.12	18.13 ± 0.11	17.19 ± 0.11	10.62 ± 0.09
Multijet	—	—	—	—	—	—	1710 ± 14	779 ± 16
Total Bkg.	—	—	—	—	—	—	22742 ± 57	6089 ± 27
Data	—	—	—	—	—	—	23248	6088
$LQ\,LQ \to eq\nu q'$ 250 ($\beta = 0.5$)	58.91 ± 0.11	44.69 ± 0.27	43.08 ± 0.28	39.03 ± 0.29	39.01 ± 0.29	38.10 ± 0.30	38.10 ± 0.30	36.63 ± 0.30
$LQ\,LQ \to eq\nu q'$ 300 ($\beta = 0.5$)	20.58 ± 0.04	15.73 ± 0.09	15.28 ± 0.10	14.02 ± 0.10	14.01 ± 0.10	13.81 ± 0.10	13.81 ± 0.10	13.44 ± 0.10
$LQ\,LQ \to eq\nu q'$ 350 ($\beta = 0.5$)	8.11 ± 0.02	6.34 ± 0.04	6.18 ± 0.04	5.72 ± 0.04	5.71 ± 0.04	5.65 ± 0.04	5.65 ± 0.04	5.51 ± 0.04
$LQ\,LQ \to eqeq$ 250 ($\beta = 1$)	119.52 ± 0.18	37.74 ± 0.56	10.80 ± 0.35	9.45 ± 0.33	9.45 ± 0.33	4.63 ± 0.23	4.62 ± 0.23	4.61 ± 0.23
$LQ\,LQ \to eqeq$ 300 ($\beta = 1$)	41.69 ± 0.06	13.07 ± 0.20	4.24 ± 0.13	3.65 ± 0.12	3.65 ± 0.12	1.93 ± 0.09	1.93 ± 0.09	1.93 ± 0.09
$LQ\,LQ \to eqeq$ 350 ($\beta = 1$)	16.45 ± 0.02	4.84 ± 0.08	1.76 ± 0.05	1.55 ± 0.05	1.55 ± 0.05	0.90 ± 0.04	0.90 ± 0.04	0.90 ± 0.04

Table A.1: Preselected sample predicted and observed cut flow yields in the electron channel. All uncertainties except that for the MJ background are statistical only. The uncertainty for the MJ background is determined from the M_T fit.

Source	Yields Region I	Region II	Region III
$W \to e\nu$+0p	300.5 ± 7.1	51.3 ± 3.0	0.60 ± 0.59
$W \to e\nu$+1p	408.0 ± 9.3	53.9 ± 3.2	0.66 ± 0.37
$W \to e\nu$+2p	1132.9 ± 9.4	142.1 ± 3.4	4.76 ± 0.63
$W \to e\nu$+3p	174.3 ± 3.8	234.5 ± 4.4	42.3 ± 1.9
$W \to e\nu$+4p	17.5 ± 1.2	68.9 ± 2.3	71.7 ± 2.4
$W \to e\nu$+5p	1.50 ± 0.38	14.7 ± 1.2	48.0 ± 2.0
$W \to \tau\nu$+0p	1.9 ± 1.2	0.13 ± 0.32	< 0.491
$W \to \tau\nu$+1p	7.9 ± 1.3	0.73 ± 0.40	< 0.496
$W \to \tau\nu$+2p	23.8 ± 1.4	2.79 ± 0.49	0.17 ± 0.12
$W \to \tau\nu$+3p	2.41 ± 0.46	4.71 ± 0.64	1.14 ± 0.31
$W \to \tau\nu$+4p	0.57 ± 0.22	1.92 ± 0.40	2.16 ± 0.42
$W \to \tau\nu$+5p	< 0.188	0.21 ± 0.13	1.77 ± 0.38
$Z \to ee$+0p	1.02 ± 0.32	0.39 ± 0.20	< 0.227
$Z \to ee$+1p	2.29 ± 0.47	0.27 ± 0.16	< 0.219
$Z \to ee$+2p	0.94 ± 0.30	1.20 ± 0.34	0.35 ± 0.18
$Z \to ee$+3p	0.10 ± 0.10	0.75 ± 0.26	1.61 ± 0.38
$Z \to ee$+4p	< 0.205	0.31 ± 0.17	1.06 ± 0.31
$Z \to ee$+5p	< 0.159	0.13 ± 0.10	0.53 ± 0.19
$Z \to \tau\tau$+0p	0.06 ± 0.08	0.08 ± 0.09	< 0.228
$Z \to \tau\tau$+1p	0.85 ± 0.29	0.31 ± 0.17	0.07 ± 0.08
$Z \to \tau\tau$+2p	2.52 ± 0.49	0.26 ± 0.16	0.26 ± 0.16
$Z \to \tau\tau$+3p	< 0.211	0.64 ± 0.24	0.44 ± 0.20
$Z \to \tau\tau$+4p	< 0.205	< 0.205	0.53 ± 0.22
$Z \to \tau\tau$+5p	< 0.160	< 0.160	0.19 ± 0.12
$t\bar{t}$	3.86 ± 0.25	31.69 ± 0.70	195.7 ± 1.7
single top (s channel)	14.66 ± 0.57	7.32 ± 0.41	4.20 ± 0.31
single top (t channel)	0.62 ± 0.03	0.48 ± 0.03	0.37 ± 0.02
single top (Wt)	1.40 ± 0.21	4.16 ± 0.36	7.78 ± 0.49
WW	15.96 ± 0.20	7.69 ± 0.14	1.80 ± 0.07
WZ	0.16 ± 0.01	0.13 ± 0.01	$(72.3 \pm 9.4) \times 10^{-3}$
ZZ	0.93 ± 0.05	0.49 ± 0.03	0.18 ± 0.02
Multijet	63.5 ± 1.3	67.8 ± 2.7	29.2 ± 1.5
Total Bkg.	2180 ± 16	700.1 ± 8.2	417.8 ± 4.6
Data	2344	722	425
$LQ\,LQ \to eq\nu q'$ 250 ($\beta = 0.5$)	< 0.014	0.02 ± 0.01	6.44 ± 0.19
$LQ\,LQ \to eq\nu q'$ 300 ($\beta = 0.5$)	< 0.005	$(2.1 \pm 2.1) \times 10^{-3}$	2.27 ± 0.10
$LQ\,LQ \to eq\nu q'$ 350 ($\beta = 0.5$)	< 0.002	< 0.002	0.80 ± 0.03
$LQ\,LQ \to eqeq$ 250 ($\beta = 1$)	< 0.028	< 0.028	1.59 ± 0.14
$LQ\,LQ \to eqeq$ 300 ($\beta = 1$)	< 0.010	< 0.010	0.55 ± 0.05
$LQ\,LQ \to eqeq$ 350 ($\beta = 1$)	< 0.004	< 0.004	0.22 ± 0.02

Table A.2: Control regions predicted and observed yields in the electron channel. Yields marked by $<$ are 95% CL upper bounds for cases in which no simulated events passed the seleection. Uncertainties are statistical only.

	Selection Yields			
Source	M_T^{LQ1}	M^{LQ2}	$M_T(e, \not{E}_T)$	S_T
$W \to e\nu$+0p	0.48 ± 0.32	0.48 ± 0.32	< 0.49	< 0.49
$W \to e\nu$+1p	12.53 ± 1.63	4.65 ± 0.99	< 0.49	< 0.49
$W \to e\nu$+2p	97.72 ± 2.86	48.64 ± 2.02	1.54 ± 0.36	1.41 ± 0.35
$W \to e\nu$+3p	78.00 ± 2.56	43.94 ± 1.93	0.70 ± 0.24	0.70 ± 0.24
$W \to e\nu$+4p	33.76 ± 1.65	21.98 ± 1.34	0.41 ± 0.18	0.41 ± 0.18
$W \to e\nu$+5p	17.89 ± 1.27	11.55 ± 1.03	0.26 ± 0.16	0.26 ± 0.16
$W \to \tau\nu$+0p	< 0.49	< 0.49	< 0.49	< 0.49
$W \to \tau\nu$+1p	0.68 ± 0.38	0.68 ± 0.38	< 0.50	< 0.50
$W \to \tau\nu$+2p	3.71 ± 0.56	1.93 ± 0.40	< 0.19	< 0.19
$W \to \tau\nu$+3p	3.48 ± 0.54	2.22 ± 0.44	< 0.20	< 0.20
$W \to \tau\nu$+4p	1.40 ± 0.34	0.91 ± 0.27	< 0.19	< 0.19
$W \to \tau\nu$+5p	0.98 ± 0.28	0.78 ± 0.25	< 0.19	< 0.19
$Z \to ee$+0p	< 0.23	< 0.23	< 0.23	< 0.23
$Z \to ee$+1p	0.08 ± 0.09	< 0.22	< 0.22	< 0.22
$Z \to ee$+2p	0.78 ± 0.27	0.44 ± 0.20	< 0.22	< 0.22
$Z \to ee$+3p	0.61 ± 0.24	0.44 ± 0.20	0.10 ± 0.10	0.10 ± 0.10
$Z \to ee$+4p	0.49 ± 0.21	0.43 ± 0.19	0.10 ± 0.09	0.10 ± 0.09
$Z \to ee$+5p	0.08 ± 0.07	0.08 ± 0.07	0.08 ± 0.07	0.08 ± 0.07
$Z \to \tau\tau$+0p	< 0.23	< 0.23	< 0.23	< 0.23
$Z \to \tau\tau$+1p	0.08 ± 0.09	< 0.22	< 0.22	< 0.22
$Z \to \tau\tau$+2p	0.63 ± 0.24	0.44 ± 0.20	< 0.22	< 0.22
$Z \to \tau\tau$+3p	0.45 ± 0.20	0.15 ± 0.12	< 0.21	< 0.21
$Z \to \tau\tau$+4p	0.24 ± 0.15	0.10 ± 0.09	< 0.20	< 0.20
$Z \to \tau\tau$+5p	0.19 ± 0.12	0.04 ± 0.05	< 0.16	< 0.16
$t\bar{t}$	64.43 ± 1.00	35.82 ± 0.75	2.66 ± 0.20	2.14 ± 0.18
single top (s)	3.01 ± 0.34	0.62 ± 0.15	0.04 ± 0.04	0.04 ± 0.04
single top (t)	0.28 ± 0.02	0.09 ± 0.01	0.00 ± 0.00	0.00 ± 0.00
single top (Wt)	3.64 ± 0.35	2.28 ± 0.28	0.26 ± 0.09	0.22 ± 0.09
WW	2.78 ± 0.08	1.66 ± 0.06	0.08 ± 0.01	0.07 ± 0.01
WZ	0.15 ± 0.01	0.07 ± 0.01	0.03 ± 0.00	0.03 ± 0.00
ZZ	0.86 ± 0.03	0.57 ± 0.02	0.08 ± 0.01	0.07 ± 0.01
Multijet	25.67 ± 0.51	18.82 ± 0.38	2.02 ± 0.04	2.02 ± 0.04
Total Bkg.	355.07 ± 5.00	199.79 ± 3.72	8.35 ± 1.36	7.63 ± 1.35
Data	350	189	11	10
LQ 250 GeV	20.43 ± 0.30	19.09 ± 0.29	10.86 ± 0.24	10.44 ± 0.24
LQ 300 GeV	8.47 ± 0.15	8.03 ± 0.15	5.04 ± 0.13	5.02 ± 0.13
LQ 350 GeV	3.68 ± 0.04	3.55 ± 0.04	2.44 ± 0.04	2.44 ± 0.04

Table A.3: Optimized selection cut flow yields summary in the electron channel for a 250 GeV LQ. Yields marked by $<$ are 95% CL upper bounds for cases in which no simulated events passed the selection.

Source	Selection Yields M_T^{LQ1}	M^{LQ2}	$M_T(e, \not{E}_T)$	S_T
$W \to e\nu$+0p	0.48 ± 0.32	0.48 ± 0.32	< 0.49	< 0.49
$W \to e\nu$+1p	8.29 ± 1.33	1.95 ± 0.64	< 0.49	< 0.49
$W \to e\nu$+2p	63.74 ± 2.31	22.21 ± 1.37	0.42 ± 0.19	0.42 ± 0.19
$W \to e\nu$+3p	50.58 ± 2.07	21.14 ± 1.34	0.09 ± 0.09	0.09 ± 0.09
$W \to e\nu$+4p	25.26 ± 1.44	13.50 ± 1.06	0.07 ± 0.08	0.07 ± 0.08
$W \to e\nu$+5p	12.24 ± 1.06	5.78 ± 0.73	0.06 ± 0.07	0.06 ± 0.07
$W \to \tau\nu$+0p	< 0.49	< 0.49	< 0.49	< 0.49
$W \to \tau\nu$+1p	0.68 ± 0.38	0.51 ± 0.33	< 0.50	< 0.50
$W \to \tau\nu$+2p	2.50 ± 0.46	0.62 ± 0.23	< 0.19	< 0.19
$W \to \tau\nu$+3p	2.50 ± 0.46	1.44 ± 0.35	< 0.20	< 0.20
$W \to \tau\nu$+4p	1.08 ± 0.30	0.62 ± 0.23	< 0.19	< 0.19
$W \to \tau\nu$+5p	0.66 ± 0.23	0.46 ± 0.19	< 0.19	< 0.19
$Z \to ee$+0p	< 0.23	< 0.23	< 0.23	< 0.23
$Z \to ee$+1p	0.08 ± 0.09	< 0.22	< 0.22	< 0.22
$Z \to ee$+2p	0.33 ± 0.18	0.26 ± 0.16	< 0.22	< 0.22
$Z \to ee$+3p	0.34 ± 0.18	0.16 ± 0.12	< 0.21	< 0.21
$Z \to ee$+4p	0.33 ± 0.17	0.25 ± 0.15	< 0.20	< 0.20
$Z \to ee$+5p	0.08 ± 0.07	0.08 ± 0.07	< 0.16	< 0.16
$Z \to \tau\tau$+0p	< 0.23	< 0.23	< 0.23	< 0.23
$Z \to \tau\tau$+1p	0.08 ± 0.09	< 0.22	< 0.22	< 0.22
$Z \to \tau\tau$+2p	0.44 ± 0.20	< 0.22	< 0.22	< 0.22
$Z \to \tau\tau$+3p	0.06 ± 0.07	< 0.21	< 0.21	< 0.21
$Z \to \tau\tau$+4p	0.24 ± 0.15	0.10 ± 0.09	< 0.20	< 0.20
$Z \to \tau\tau$+5p	0.19 ± 0.12	0.04 ± 0.05	< 0.16	< 0.16
$t\bar{t}$	42.43 ± 0.81	17.83 ± 0.53	0.76 ± 0.11	0.61 ± 0.10
single top (s)	1.62 ± 0.25	0.28 ± 0.10	< 0.05	< 0.05
single top (t)	0.18 ± 0.02	0.03 ± 0.01	< 0.00	< 0.00
single top (Wt)	2.31 ± 0.28	1.15 ± 0.20	0.06 ± 0.05	0.06 ± 0.05
WW	1.82 ± 0.07	0.81 ± 0.04	0.02 ± 0.01	0.02 ± 0.01
WZ	0.10 ± 0.01	0.04 ± 0.01	0.01 ± 0.00	0.01 ± 0.00
ZZ	0.56 ± 0.02	0.28 ± 0.01	0.03 ± 0.00	0.03 ± 0.00
Multijet	6.52 ± 0.78	3.80 ± 0.46	< 0.28	< 0.28
Total Bkg.	225.72 ± 17.34	93.81 ± 2.71	1.54 ± 1.33	1.38 ± 1.33
Data	217	82	2	2
$LQ\,LQ \to eq\nu q'$ 250 ($\beta = 0.5$)	22.52 ± 0.29	20.71 ± 0.29	9.71 ± 0.22	9.62 ± 0.22
$LQ\,LQ \to eq\nu q'$ 300 ($\beta = 0.5$)	9.84 ± 0.11	9.27 ± 0.11	5.07 ± 0.09	5.05 ± 0.09
$LQ\,LQ \to eq\nu q'$ 350 ($\beta = 0.5$)	4.42 ± 0.04	4.24 ± 0.04	2.56 ± 0.04	2.56 ± 0.04
$LQ\,LQ \to eqeq$ 250 ($\beta = 1$)	1.83 ± 0.15	1.62 ± 0.14	0.24 ± 0.05	0.24 ± 0.05
$LQ\,LQ \to eqeq$ 300 ($\beta = 1$)	0.98 ± 0.06	0.93 ± 0.06	0.24 ± 0.03	0.24 ± 0.03
$LQ\,LQ \to eqeq$ 350 ($\beta = 1$)	0.50 ± 0.03	0.48 ± 0.03	0.15 ± 0.02	0.15 ± 0.02

Table A.4: Optimized selection cut flow yields in the electron channel for a 300 GeV LQ. Yields marked by $<$ are 95% CL upper bounds for cases in which no simulated events passed the selection.

Source	Selection Yields			
	M_T^{LQ1}	M^{LQ2}	$M_T(e, \not{E}_T)$	S_T
$W \to e\nu$+0p	0.48 ± 0.32	< 0.49	< 0.49	< 0.49
$W \to e\nu$+1p	8.29 ± 1.33	1.60 ± 0.58	< 0.49	< 0.49
$W \to e\nu$+2p	63.74 ± 2.31	15.65 ± 1.15	0.42 ± 0.19	0.42 ± 0.19
$W \to e\nu$+3p	50.58 ± 2.07	14.80 ± 1.12	< 0.20	< 0.20
$W \to e\nu$+4p	25.26 ± 1.44	9.88 ± 0.90	0.07 ± 0.08	0.07 ± 0.08
$W \to e\nu$+5p	12.24 ± 1.06	4.00 ± 0.61	0.06 ± 0.07	0.06 ± 0.07
$W \to \tau\nu$+0p	< 0.49	< 0.49	< 0.49	< 0.49
$W \to \tau\nu$+1p	0.68 ± 0.38	0.15 ± 0.18	< 0.50	< 0.50
$W \to \tau\nu$+2p	2.50 ± 0.46	0.55 ± 0.22	< 0.19	< 0.19
$W \to \tau\nu$+3p	2.50 ± 0.46	1.25 ± 0.33	< 0.20	< 0.20
$W \to \tau\nu$+4p	1.08 ± 0.30	0.47 ± 0.20	< 0.19	< 0.19
$W \to \tau\nu$+5p	0.66 ± 0.23	0.46 ± 0.19	< 0.19	< 0.19
$Z \to ee$+0p	< 0.23	< 0.23	< 0.23	< 0.23
$Z \to ee$+1p	0.08 ± 0.09	< 0.22	< 0.22	< 0.22
$Z \to ee$+2p	0.33 ± 0.18	0.16 ± 0.12	< 0.22	< 0.22
$Z \to ee$+3p	0.34 ± 0.18	0.16 ± 0.12	< 0.21	< 0.21
$Z \to ee$+4p	0.33 ± 0.17	0.19 ± 0.13	< 0.20	< 0.20
$Z \to ee$+5p	0.08 ± 0.07	0.08 ± 0.07	< 0.16	< 0.16
$Z \to \tau\tau$+0p	< 0.23	< 0.23	< 0.23	< 0.23
$Z \to \tau\tau$+1p	0.08 ± 0.09	< 0.22	< 0.22	< 0.22
$Z \to \tau\tau$+2p	0.44 ± 0.20	< 0.22	< 0.22	< 0.22
$Z \to \tau\tau$+3p	0.06 ± 0.07	< 0.21	< 0.21	< 0.21
$Z \to \tau\tau$+4p	0.24 ± 0.15	0.10 ± 0.09	< 0.20	< 0.20
$Z \to \tau\tau$+5p	0.19 ± 0.12	< 0.16	< 0.16	< 0.16
$t\bar{t}$	42.43 ± 0.81	12.33 ± 0.44	0.54 ± 0.09	0.38 ± 0.08
single top (s)	1.62 ± 0.25	0.19 ± 0.08	< 0.05	< 0.05
single top (t)	0.18 ± 0.02	0.02 ± 0.01	< 0.00	< 0.00
single top (Wt)	2.31 ± 0.28	0.90 ± 0.18	0.06 ± 0.05	0.04 ± 0.04
WW	1.82 ± 0.07	0.54 ± 0.04	0.02 ± 0.01	0.01 ± 0.01
WZ	0.10 ± 0.01	0.03 ± 0.01	0.01 ± 0.00	0.01 ± 0.00
ZZ	0.56 ± 0.02	0.20 ± 0.01	0.02 ± 0.00	0.02 ± 0.00
Multijet	20.50 ± 2.46	9.83 ± 1.18	< 0.28	< 0.28
Total Bkg.	239.70 ± 4.77	73.53 ± 2.62	1.20 ± 1.34	1.02 ± 1.34
Data	217	60	1	1
LQ 250 GeV	17.07 ± 0.28	12.35 ± 0.25	4.56 ± 0.17	3.26 ± 0.14
LQ 300 GeV	7.61 ± 0.15	6.40 ± 0.14	3.03 ± 0.11	2.66 ± 0.10
LQ 350 GeV	3.46 ± 0.04	3.03 ± 0.04	1.59 ± 0.03	1.52 ± 0.03

Table A.5: Optimized selection cut flow yields summary in the electron channel for a 350 GeV LQ. Yields marked by $<$ are 95% CL upper bounds for cases in which no simulated events passed the seleection.

Source	Preselection Scale σ^-	Preselection Scale σ^+	Preselection Resolution	250 GeV Optimization Scale σ^-	250 GeV Optimization Scale σ^+	250 GeV Optimization Resolution	300 GeV Optimization Scale σ^-	300 GeV Optimization Scale σ^+	300 GeV Optimization Resolution	350 GeV Optimization Scale σ^-	350 GeV Optimization Scale σ^+	350 GeV Optimization Resolution
$W \to e\nu$+0p	−0.022	0.020	−0.001	—	—	—	—	—	—	—	—	—
$W \to e\nu$+1p	−0.008	0.008	−0.001	—	—	—	—	—	—	—	—	—
$W \to e\nu$+2p	−0.012	0.014	−0.001	0.000	0.000	0.000	0.167	0.000	0.000	0.000	0.000	0.000
$W \to e\nu$+3p	−0.011	0.009	−0.001	−0.125	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$W \to e\nu$+4p	0.006	0.004	−0.001	0.000	0.333	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$W \to e\nu$+5p	−0.014	−0.006	−0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$W \to \tau\nu$+0p	0.017	−0.030	−0.006	—	—	—	—	—	—	—	—	—
$W \to \tau\nu$+1p	0.054	−0.031	−0.009	—	—	—	—	—	—	—	—	—
$W \to \tau\nu$+2p	0.017	−0.038	−0.009	—	—	—	—	—	—	—	—	—
$W \to \tau\nu$+3p	0.024	−0.027	−0.006	—	—	—	—	—	—	—	—	—
$W \to \tau\nu$+4p	0.046	−0.009	0.000	—	—	—	—	—	—	—	—	—
$W \to \tau\nu$+5p	0.053	−0.079	−0.026	—	—	—	—	—	—	—	—	—
$Z \to ee$+0p	−0.076	0.080	0.000	—	—	—	—	—	—	—	—	—
$Z \to ee$+1p	−0.093	0.123	0.000	—	—	—	—	—	—	—	—	—
$Z \to ee$+2p	−0.093	0.169	−0.008	—	—	—	—	—	—	—	—	—
$Z \to ee$+3p	−0.109	0.156	0.000	0.000	0.000	0.000	—	—	—	—	—	—
$Z \to ee$+4p	−0.033	0.133	0.000	0.000	0.000	0.000	—	—	—	—	—	—
$Z \to ee$+5p	−0.067	0.333	0.000	−1.000	0.000	0.000	—	—	—	—	—	—
$Z \to \tau\tau$+0p	−0.049	−0.049	−0.016	—	—	—	—	—	—	—	—	—
$Z \to \tau\tau$+1p	−0.020	−0.020	−0.005	—	—	—	—	—	—	—	—	—
$Z \to \tau\tau$+2p	0.035	0.000	0.000	—	—	—	—	—	—	—	—	—
$Z \to \tau\tau$+3p	0.038	0.115	0.000	—	—	—	—	—	—	—	—	—
$Z \to \tau\tau$+4p	0.111	0.000	0.000	—	—	—	—	—	—	—	—	—
$Z \to \tau\tau$+5p	0.000	0.000	0.000	—	—	—	—	—	—	—	—	—
$t\bar{t}$	0.002	−0.002	−0.001	−0.008	0.000	0.008	0.048	0.024	0.000	0.000	−0.020	0.000
single top (s)	−0.003	−0.003	−0.003	0.000	0.000	0.000	—	—	—	0.000	−1.000	0.000
single top (t)	−0.005	−0.005	−0.001	0.000	0.000	0.000	—	—	—	—	—	—
single top (tW)	−0.005	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
WW	−0.008	0.009	−0.002	−0.034	0.034	0.000	0.000	−0.077	0.000	−0.063	0.000	0.000
WZ	−0.031	0.041	0.000	−0.100	0.000	0.050	−0.200	−0.100	0.000	−0.091	0.000	0.000
ZZ	−0.009	0.014	−0.001	0.080	−0.040	0.000	0.200	0.000	0.000	0.063	0.000	0.000
LQ 250 GeV	−0.001	0.003	−0.001	−0.013	0.003	0.001	−0.001	0.004	0.002	−0.005	0.011	0.003
LQ 300 GeV	−0.004	0.001	0.000	−0.011	0.004	0.001	−0.005	−0.009	0.001	−0.008	−0.011	−0.001
LQ 350 GeV	−0.004	0.002	0.000	−0.007	0.005	0.000	−0.004	0.007	0.002	−0.001	0.006	0.000

Table A.6: Systematic uncertainty arising from the electron energy scale and resolution uncertainty.

Sample	Preselection		250 GeV Opt		300 GeV Opt		350 GeV Opt	
	σ^{+}_{JES}	σ^{-}_{JES}	σ^{+}_{JES}	σ^{-}_{JES}	σ^{+}_{JES}	σ^{-}_{JES}	σ^{+}_{JES}	σ^{-}_{JES}
$W \to e\nu$+0p	0.269	−0.341	—	—	—	—	—	—
$W \to e\nu$+1p	0.312	−0.347	—	—	—	—	—	—
$W \to e\nu$+2p	0.073	−0.104	—	—	0.765	0.000	—	—
$W \to e\nu$+3p	0.033	−0.050	—	—	—	—	—	—
$W \to e\nu$+4p	0.027	−0.015	—	—	—	—	—	—
$W \to e\nu$+5p	0.026	−0.009	—	—	—	—	—	—
$W \to \tau\nu$+0p	0.244	−0.605	—	—	—	—	—	—
$W \to \tau\nu$+1p	0.396	−0.436	—	—	—	—	—	—
$W \to \tau\nu$+2p	0.083	−0.158	—	—	—	—	—	—
$W \to \tau\nu$+3p	0.076	−0.091	—	—	—	—	—	—
$W \to \tau\nu$+4p	0.043	−0.044	—	—	—	—	—	—
$W \to \tau\nu$+5p	0.104	−0.038	—	—	—	—	—	—
$Z \to ee$+0p	0.497	−0.224	—	—	—	—	—	—
$Z \to ee$+1p	0.176	0.133	—	—	—	—	—	—
$Z \to ee$+2p	0.190	0.289	—	—	—	—	—	—
$Z \to ee$+3p	0.076	0.434	—	—	—	—	—	—
$Z \to ee$+4p	0.035	0.570	—	—	—	—	—	—
$Z \to ee$+5p	−0.140	0.496	—	—	—	—	—	—
$Z \to \tau\tau$+0p	0.561	−0.527	—	—	—	—	—	—
$Z \to \tau\tau$+1p	0.445	−0.249	—	—	—	—	—	—
$Z \to \tau\tau$+2p	0.109	−0.130	—	—	—	—	—	—
$Z \to \tau\tau$+3p	0.092	−0.000	—	—	—	—	—	—
$Z \to \tau\tau$+4p	0.091	−0.127	—	—	—	—	—	—
$Z \to \tau\tau$+5p	0.400	0.000	—	—	—	—	—	—
$t\bar{t}$	0.021	−0.018	—	—	0.292	0.020	—	—
single top (s)	0.046	−0.057	—	—	—	—	—	—
single top (t)	0.040	−0.067	—	—	—	—	—	—
single top (tW)	0.022	−0.031	—	—	—	—	—	—
WW	0.021	−0.120	—	—	0.134	−0.014	—	—
WZ	0.062	0.015	—	—	−0.083	−0.083	—	—
ZZ	0.062	−0.070	—	—	0.177	−0.245	—	—
LQ 200 GeV	0.001	−0.003	—	—	−0.004	−0.066	—	—
LQ 250 GeV	−0.001	−0.001	—	—	−0.005	−0.024	—	—
LQ 300 GeV	0.001	−0.002	—	—	−0.003	−0.008	—	—

Table A.7: The jet energy scale systematic uncertainties on the yields for each source for the electron channel. Blanks indicate cases for which no MC events passed the selections. Given the limited statistical precision for the $W \to \tau\nu$ and $Z \to \ell\ell$ samples, the $W \to e\nu$+jets systematic uncertainty will be used for all V+jets backgrounds.

Sample	Preselection	250 GeV Opt	300 GeV Opt	350 GeV Opt
$W \to e\nu$+0p	−0.256	—	—	—
$W \to e\nu$+1p	−0.242	—	—	—
$W \to e\nu$+2p	−0.086	—	0.163	—
$W \to e\nu$+3p	−0.032	—	—	—
$W \to e\nu$+4p	0.027	—	—	—
$W \to e\nu$+5p	0.009	—	—	—
$W \to \tau\nu$+0p	−0.305	—	—	—
$W \to \tau\nu$+1p	−0.286	—	—	—
$W \to \tau\nu$+2p	−0.091	—	—	—
$W \to \tau\nu$+3p	0.020	—	—	—
$W \to \tau\nu$+4p	−0.029	—	—	—
$W \to \tau\nu$+5p	0.080	—	—	—
$Z \to ee$+0p	0.224	—	—	—
$Z \to ee$+1p	0.746	—	—	—
$Z \to ee$+2p	0.923	—	—	—
$Z \to ee$+3p	0.754	—	—	—
$Z \to ee$+4p	0.979	—	—	—
$Z \to ee$+5p	1.412	—	—	—
$Z \to \tau\tau$+0p	0.250	—	—	—
$Z \to \tau\tau$+1p	0.140	—	—	—
$Z \to \tau\tau$+2p	−0.031	—	—	—
$Z \to \tau\tau$+3p	0.245	—	—	—
$Z \to \tau\tau$+4p	0.194	—	—	—
$Z \to \tau\tau$+5p	0.789	—	—	—
$t\bar{t}$	0.015	—	0.312	—
single top (s)	−0.041	—	—	—
single top (t)	−0.019	—	—	—
single top (tW)	0.021	—	—	—
WW	−0.090	—	−0.016	—
WZ	0.084	—	−0.242	—
ZZ	−0.036	—	0.006	—
LQ 250 GeV	−0.003	—	−0.023	—
LQ 300 GeV	−0.005	—	−0.004	—
LQ 350 GeV	−0.003	—	0.007	—

Table A.8: The jet energy resolution systematic uncertainties on the yields for each source for the electron channel. Blanks indicate cases for which no MC events passed the selections. Given the limited statistical precision for the $W \to \tau\nu$ and $Z \to \ell\ell$ samples, the $W \to e\nu$+jets systematic uncertainty will be used for all V+jets backgrounds.

Source	Electron Channel			
	Preselection	250 GeV Opt.	300 GeV Opt	350 GeV Opt
$W \to e\nu$+0p	−0.336	—	—	—
$W \to e\nu$+1p	−0.213	—	—	—
$W \to e\nu$+2p	−0.020	−0.125	0.167	−0.167
$W \to e\nu$+3p	−0.026	−0.257	−0.010	—
$W \to e\nu$+4p	−0.010	0.000	0.000	0.000
$W \to e\nu$+5p	−0.022	0.000	0.000	0.000
$W \to \tau\nu$+0p	−0.599	—	—	—
$W \to \tau\nu$+1p	−0.176	—	—	—
$W \to \tau\nu$+2p	−0.077	—	—	—
$W \to \tau\nu$+3p	−0.024	—	—	—
$W \to \tau\nu$+4p	−0.037	—	—	—
$W \to \tau\nu$+5p	0.079	—	—	—
$Z \to ee$+0p	−0.808	—	—	—
$Z \to ee$+1p	−0.428	—	—	—
$Z \to ee$+2p	−0.368	—	—	—
$Z \to ee$+3p	−0.186	0.000	—	—
$Z \to ee$+4p	−0.167	1.000	—	—
$Z \to ee$+5p	−0.133	—	—	—
$Z \to \tau\tau$+0p	−0.334	—	—	—
$Z \to \tau\tau$+1p	−0.412	—	—	—
$Z \to \tau\tau$+2p	0.070	—	—	—
$Z \to \tau\tau$+3p	−0.042	—	—	—
$Z \to \tau\tau$+4p	−0.111	—	—	—
$Z \to \tau\tau$+5p	−0.333	—	—	—
$t\bar{t}$	−0.012	−0.039	0.000	0.130
single top (s)	−0.016	0.000	—	—
single top (t)	−0.011	0.000	—	—
single top (tW)	−0.048	0.110	−0.029	−0.029
WW	−0.022	−0.086	−0.137	−0.235
WZ	−0.033	−0.033	0.362	0.022
ZZ	−0.095	−0.181	−0.217	−0.004
LQ 250 GeV	−0.000	0.003	0.018	0.024
LQ 300 GeV	0.007	0.005	−0.021	−0.043
LQ 350 GeV	0.009	0.018	0.016	0.010

Table A.9: Fractional systematic uncertainty arising from event pile up for the electron channel. The uncertainty is defined as the fractional change in the yields of samples without and with pile up and is shown for the preselection and all three mass–dependent optimizations.

Appendix B

Original vs. Updated Signal MC Event Comparison

The simulated samples originally used in this analysis shown in Table B.1 were generated using a recently available Pythia shower model which has p_T ordered showering. Studies show that this gave a very different hadronic energy around the electron than seen in other ATLAS samples, for examples the $t\bar{t}$ samples used in this analysis. The samples were regenerated (also in Table B.1) using the older Pythia virtuality ordered parton shower model. This also necessitated an underlying event tune change, but the impact of the tune change is small.

New Monte Carlo Samples Mass (GeV)	Request ID	Processing Tags
200	119142	e699, s765, s767, r1302, r1306
250	119143	e699, s765, s767, r1302, r1306
300	119144	e699, s765, s767, r1302, r1306
350	119145	e699, s765, s767, r1302, r1306
400	119146	e699, s765, s767, r1302, r1306
Old Monte Carlo Samples Mass (GeV)	Request ID	Processing Tags
200	115166	e579, s765, s767, r1302, r1306
250	115167	e579, s765, s767, r1302, r1306
300	115168	e579, s765, s767, r1302, r1306
350	115169	e579, s765, s767, r1302, r1306
400	115170	e579, s765, s767, r1302, r1306

Table B.1: All samples are `mc09` production at $\sqrt{s} = 7$ TeV created from AODs. The New MC samples: The samples were regenerated using the older Pythia virtuality ordered parton shower model. The Old MC samples: The generation information for the original simulated signal samples. These samples are no longer used because of a (suspected) bug in LQ productions with the Pythia p_T ordered showering algorithm.

Table B.2 shows the signal cut flow for the original and new simulated samples. The initial estimates indicated a 10% increase in electron ID efficiency resulting from the change. However, a striking increase in the signal yield for the optimized selections is seen. Figures B.1 and B.2 show comparison distributions of important kinematic variables between the original and new simulations. The decreased hadronic energy results in an higher (transverse) energies in the events giving rise to the improved acceptance.

Selection	Old MC	New MC
lepton selection	13.91 ± 0.15	15.73 ± 0.09
$\not{E}_T > 20$ GeV	13.37 ± 0.15	15.28 ± 0.10
$M_T > 40$ GeV	12.02 ± 0.16	14.02 ± 0.10
$N_J \geq 1$	12.02 ± 0.16	14.01 ± 0.10
triangle	11.74 ± 0.16	13.81 ± 0.10
$N_J \geq 2$, lepton veto	11.56 ± 0.16	13.44 ± 0.10
M_T^{LQ}	7.61 ± 0.15	9.84 ± 0.11
M^{LQ}	6.90 ± 0.15	9.27 ± 0.11
M_T	3.37 ± 0.11	5.07 ± 0.09
S_T	3.36 ± 0.11	5.05 ± 0.09

Table B.2: Cut flow yields comparing the old and new signal MC. The 300 GeV signal samples are shown.

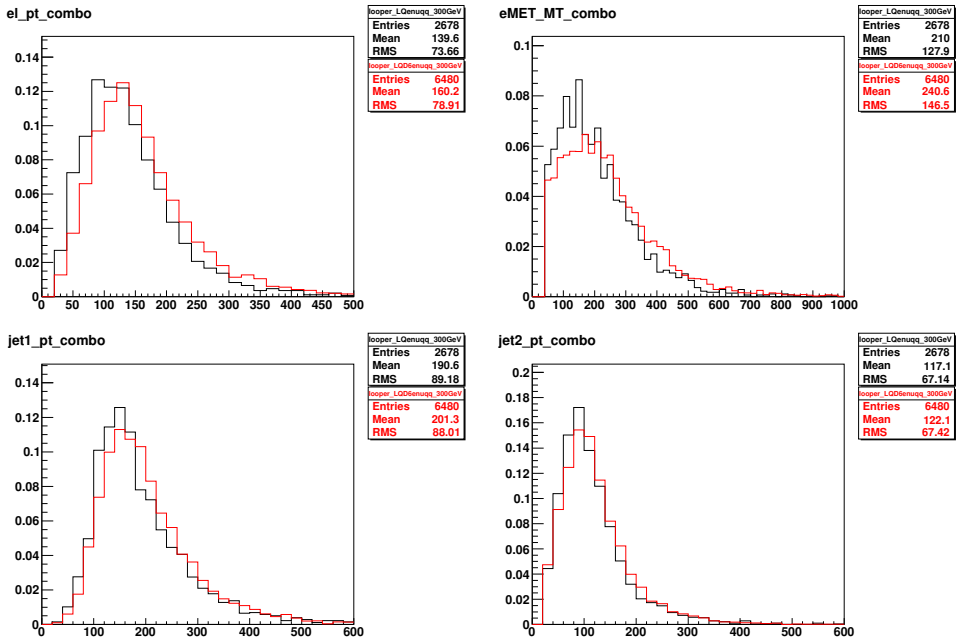


Figure B.1: Electron p_T, $M_{LQ} = 300$ GeV after the loose preselection (top left), and $\not{E}_T$, $M_{LQ} = 300$ GeV after the loose preselection (top right). The red histogram is the new MC sample, and black is the original MC sample. Leading jet p_T, $M_{LQ} = 300$ GeV after the loose preselection (lower left), and 2nd jet p_T, $M_{LQ} = 300$ GeV after the loose preselection (lower right). The red histogram is the new MC sample, and black is the original MC sample.

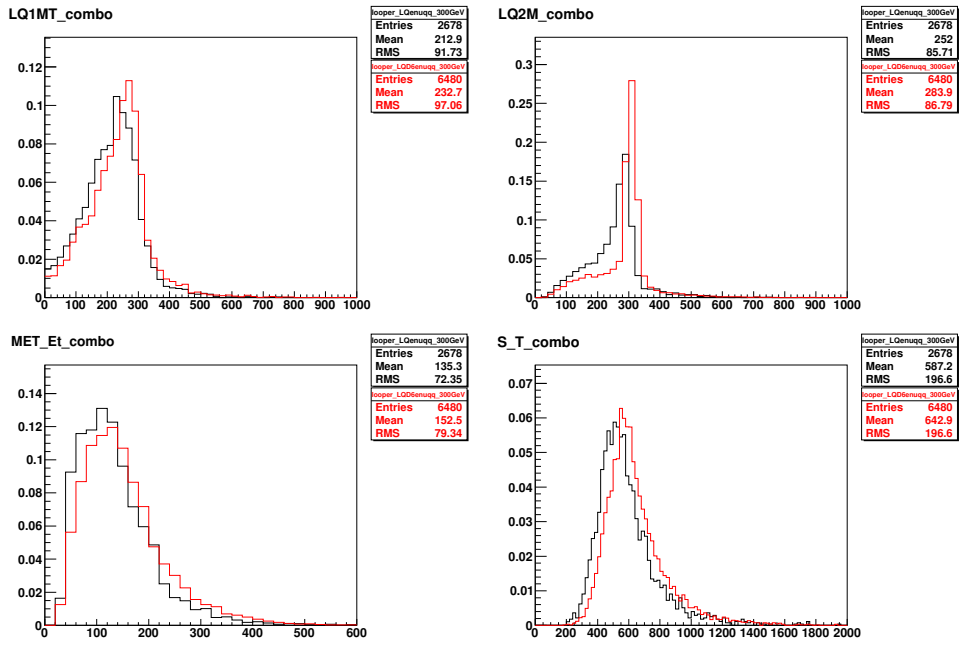


Figure B.2: M_T^{LQ}, $M_{LQ} = 300$ GeV after the loose preselection (top left), and M^{LQ}, $M_{LQ} = 300$ GeV after the loose preselection (top right). The red histogram is the new MC sample, and black is the original MC sample. M_T, $M_{LQ} = 300$ GeV after the loose preselection (lower left), and S_T, $M_{LQ} = 300$ GeV after the loose preselection (lower right). The red histogram is the new MC sample, and black is the original MC sample.

Appendix C

E/p from photon conversions in photon+jet samples

C.1 Introduction

The ATLAS electromagnetic (EM) calorimeter is part of the larger Liquid Argon (LAr) calorimeter system. To calibrate the EM calorimeter, particles of a known mass can be reconstructed from events which decay to dielectrons such as the J/ψ, the Υ, and the Z. All three of these can be cleanly reconstructed and their known mass provides a standard for calibration. However, the resulting electron energy spectra have large gaps for which the calibration is untested. It is important to understand the detector response in different energy ranges, specifically in the regions between mass of reconstructed dielectrons from the J/ψ, the Υ and the Z but also extending beyond these energies, because this is where new physics is likely to lie.

The purpose of this study is to use high cross section, high transverse energy (E_T) events to investigate the detector response and subsequently check the detector calibration over a wide energy range and compare to the results to known sources. However, since there is no invariant mass to reconstruct the calibration must rely on inter-calibration between the ID and the EM calorimeter using electrons only. Two potential sources with high cross sections at high E_T are photon+jet and dijet samples. Photon conversions from these two data sets are reconstructed using both the inner detector (ID) and the EM calorimeter. The E_T/p_T is calculated to compare the response from the ID and the EM calorimeter in into p_T and η bins. The p_T and η binned values of E_T/p_T (from here referred to as E/p) are then fit to determine the most probable value (MPV). Assuming the inner detector is calibrated, the correct calibration of the EM calorimeter over a complete energy range can then be determined.

C.2 E/p calibration using Photon + Jet samples

The data set investigated is a photon+jet MC sample chosen because the primary photon is produced over a large energy range. The converted photon produces two isolated electrons. The process has a high cross section, so the data set could be utilized after a short period of running.

The major background to the photon+jet sample is multiijet (QCD) events. They differ from photon+jet samples in that the photons do not come from the primary interaction. Instead the photons result dominantly from the decays of π^0s, and are typically embedded in jets. Therefore, the photons are not isolated and are at a lower energy than the primary interaction making it harder to use these for calibration.

C.2.1 Photon + Jet Monte Carlo

The MC data sets used were D3PDs from the standard ATLAS production Egamma ntuples (NTUP_EGAMMA version 15.6.10). The D3PDs were made in the production chain from the bulk produced MC09 data sets at $\sqrt{s} = 7$ TeV. The MC information is listed in Table C.1. Each data set contains events produced within a $\hat{p}_\perp$ range of the hard scatter in the rest frame of the center of mass. The different ranges are referred to as *jet number* during the rest of this Appendix. The sample names, $\hat{p}_\perp$ range of the hard scatter, and number of produced events in 35 pb^{-1} using a cross section calculated by Pythia are as follows:

Jet Number (Range [GeV])	Request ID	Number of Events	Equivalent $\mathcal{L}$ [pb $^{-1}$]	Events in $\mathcal{L} = 35$ pb^{-1}
J1 ($17 \leq \hat{p}_\perp \leq 35$)	108087	4994464	22.1	7915950
J2 ($35 \leq \hat{p}_\perp \leq 70$)	108081	1998486	115.5	605675
J3 ($70 \leq \hat{p}_\perp \leq 140$)	108082	999662	657.7	53200
J4 ($140 \leq \hat{p}_\perp \leq 280$)	108083	998676	11954.5	2924
Total				8577749

Table C.1: Photon + Jet Monte Carlo Information. Processing tag: e505, s765, s767, r1305, r1306, p179. This yields a total of 8.6×10^6 events over all hard scatter p_T ranges. These were produced from Egamma ntuples: NTUP_EGAMMA version 15.6.10.

C.2.2 Event Selection

This study utilizes photon conversions giving two isolated electrons in the ID. The EM cluster object has both E_T and isolation requirements to insure the remaining events are of high quality for the E/p study. Since the conversion occurs in the ID, electrons from the

conversion also have track requirements as is necessary to have a high quality calculation of E/p.

Good Runs List The good runs list is defined by the e/γ working group for run periods A-I2 during the 2010 run. It is only implemented when running on data. See Table C.2.

Period	Number of runs	Run Range	$\mathcal{L}$ [pb^{-1}]
A	17	[152166, 153159]	0.000
B	5	[153565, 155160]	0.007
C	8	[155228, 156682]	0.008
D	20	[158045, 159224]	0.281
E	22	[160387, 161948]	0.996
F	10	[162347, 162882]	1.761
G	16	[165591, 166383]	6.859
H	8	[166466, 166964]	7.150
I	7	[167575, 167844]	21.268
Total	113		38.332

Table C.2: GRL for data in periods A-I during 2010 run.

Vertex The highest p_T vertex (primary vertex) of the interaction must have the following:

- At least three tracks from the vertex
- Maximum displacement z $\leq$ 150 mm from the interaction point

This requirement is implemented on both data and MC.

Triggers The trigger used for both the 2010 data and the photon+jet samples is the g40_loose trigger. This is a photon trigger that requires EM clusters to satisfy $E_T \geq 40$ GeV at Event Filter level and is not prescaled. It is a primary e/γ trigger.

Photon Quality The photons reconstructed passed the tight requirement of the **PhotonIDTool**, see Section 4.4.2 for more information. Photons were required to be within $|\eta| = 2.47$ and were rejected in the crack region ($1.37 < |\eta| < 1.52$) as were photons in regions with bad OTx (as with electrons, see Section 4.4.1). The photons were required to have an author of 16, which means that they are conversion photons, as determined by the photon reconstruction algorithms which look for displaced vertices from electron pairs.

The E_T of the photons was required to be 25 GeV and pass a relative isolation of 10% in a ΔR cone of 20. In retrospect a photon E_T of 40 GeV would have been a more appropriate choice given the trigger requirements. The results from the photon selection are shown in Table C.3.

Sample	ave. γ per event	OTx	Author	η	E_T	E_T isolation	Tight γ	Total
J1	2.6	45%	32%	92%	100%	63%	76%	16.5%
J2	3.1	45%	35%	92%	100%	57%	89%	22.5%
J3	3.4	45%	37%	92%	100%	50%	92%	23.5%
J4	3.7	45%	37%	93%	100%	43%	93%	23.0%

Table C.3: Results of event selection in photon+jet sample.

Conversion Reconstruction The reconstruction of the photon conversion was done by track cluster matching in the Egamma D3PD maker. The Photon Recovery Tool finds electrons from converted photons and reconstructs the photon. Each track and cluster object are matched via the requirement $\Delta R < 0.05$. One EM cluster matches two tracks because the two electrons resulting from the photon conversion are too close together to be resolved in the EM calorimeter. Additionally, an EM cluster object in each event had to have a vertex candidate match two tracks in the photon container.

Track Selection The requirements on the tracks assigned to the photon conversion are that the vertex candidate has two tracks. TRT only conversion tracks were not considered, and a fiducial cut on the conversion radius ($R_{conv} > 600$ mm) was applied. There was also a $\chi^2 < 10$ requirement to ensure that the reconstructed tracks were of good quality. See Figure C.1.

Conversion Selection The location of the truth vertex after the selection requirements in the detector is shown in Figure C.2. Figure C.3 shows the silicon regions of the ID with truth photons on the left, the reconstructed photon conversion vertex located on the right, and the data on the bottom.

C.2.3 Event Yields

Photon + jet signal

The efficiencies for each stage of the event selection and the total efficiency are listed after the event selection in Table C.4. Figure C.4(a) shows the photon E_T spectrum for the com-

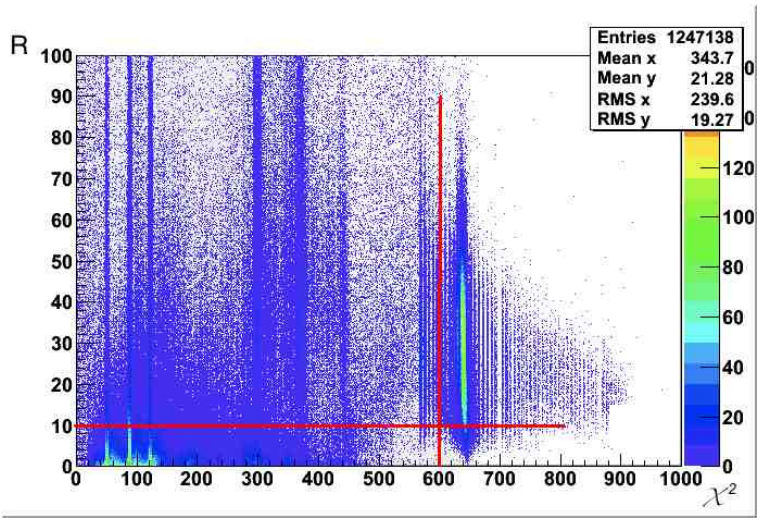


Figure C.1: Conversion radius (R) vs. Vertex χ^2

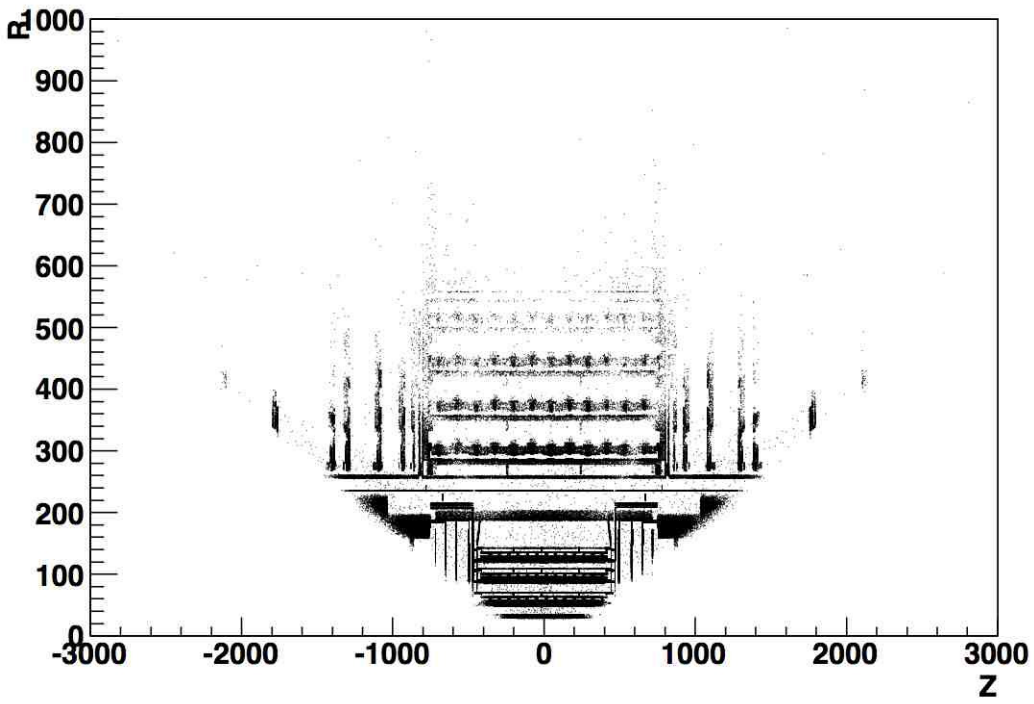
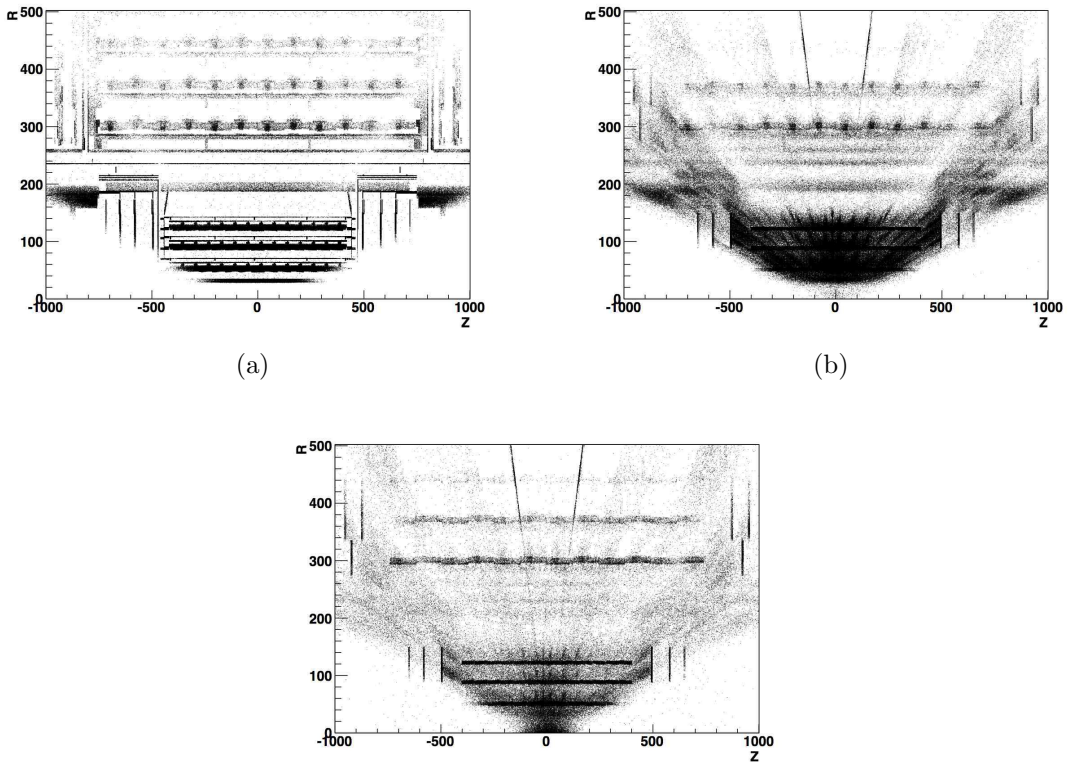


Figure C.2: Location of truth conversion vertex in the RZ-plane.



(a) (b)

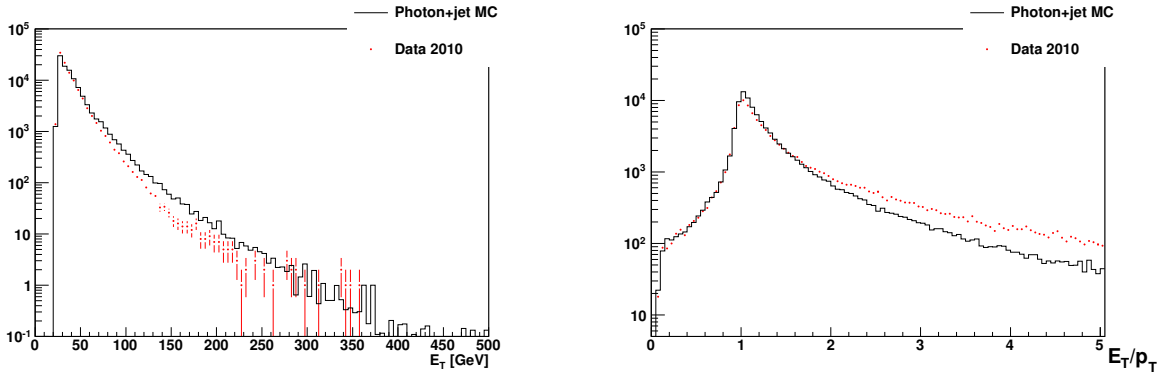
(c)

Figure C.3: Location of conversion vertex in the RZ-plane (a) true (b) reco (c) data

bined J1-J4 photon+jet samples plotted with the E_T distribution for the data. Figure C.4(b) shows the E/p. The MC is scaled to the data yield.

Sample	Vertex	Trigger	Photon	E/p	Total
J1	98.7%	84.2%	2.97%	100%	2.47%
J2	98.7%	89.5%	9.22%	100%	8.15%
J3	98.7%	91.1%	9.46%	100%	8.51%
J4	98.7%	94.0%	8.22%	100%	7.62%
Combined					2.91%

Table C.4: Photon+jet Event Selection cut flow.

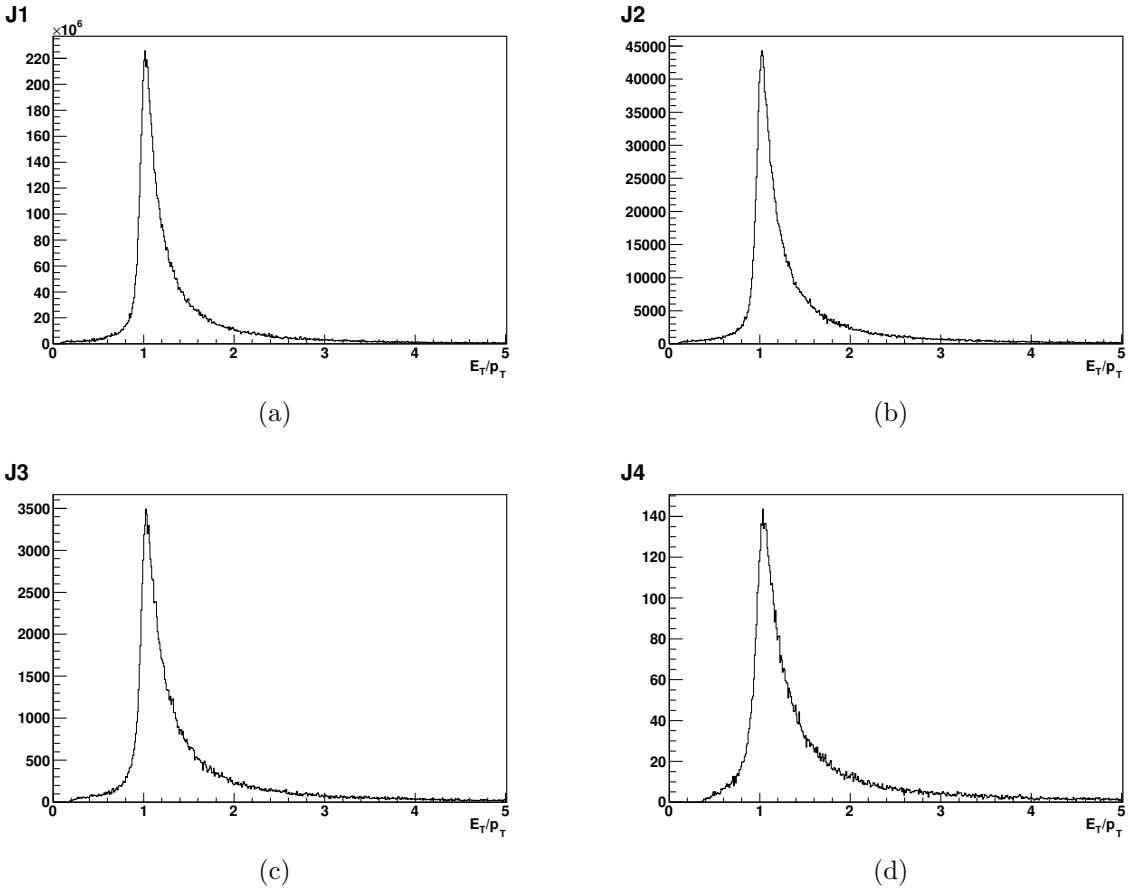


(a) (b)

Figure C.4: (a) Photon E_T spectrum with both data and MC. These events pass the event selection. (b) E/p distribution with data and MC. The MC is scaled to the data yield. There is a discrepancy in the tails of both distributions. This is most likely because there is no accounting for the dijet background which would most likely be in the tails of the E/p distribution.

In initial analysis of the E/p spectrum, there was a tail that extended to E/p > 3. These events occurred from two dominant causes: a threshold bias from the generator and TRT only reconstructed tracks. Sometimes the photon track p_T is significantly less than the p_T of the photon from the Jn sample that is associated with it. Generally these events were located near the generator boundaries (near 17 GeV for the J1 sample, for example). The generator thresholds introduce a resolution bias near the generator cut off. The events located near the generator boundary are removed to eliminate this bias. Additionally TRT only conversions were rejected with events that had high values of χ^2. This removed the majority of events that had an E/p > 3.The resulting distributions for the individual samples are shown in

Figure C.5 and the combined results with the data are shown in Figure C.4(b). There is a discrepancy in the tails of both distributions. This is most likely due to the fact that there is no accounting for the dijet background which would most likely be in the tails of the E/p distribution.



(a) (b) (c) (d)

Figure C.5: Results of event selection in photon+jet sample. (a) J1 (b) J2 (c) J3 (d) J4

C.3 Fit Model

The E/p distributions were made in bins of the combined p_T of the conversion electrons ($p_{T,e1} + p_{T,e2}$). The bin size was determined by statistics; each bin required at least 10000 conversions in MC and 3000 in data (see Table C.5). The E/p distributions were then used as described in the Section C.3.1 and C.3.2 to check the calorimeter response and calibration. The E/p values were also binned by the η of the EM cluster object to observe the response

as a function of η. The p_T binned values of E/p distribution were fit and the parameters of the fit were plotted also as described in Sections C.3.1 and C.3.2.

Sample	p_T bin ranges (GeV)
J1	22-26, 26-28, 28-30, 30-33, 33-35
J2	40-43, 43-46, 46-50, 50-55, 55-63, 63-70
J3	75-83, 83-95, 95-120, 120-140
J4	145-160, 160-188, 188-264, 264-280
Jall	35-38, 38-41, 41-45, 45-50, 50-57, 57-70, 70-102, 102-280
data	35-37, 37-39, 39-42, 42-45, 45-49, 49-55, 55-66, 66-102, 102-280

Table C.5: A list of the results of the p_T binning. J1-J4 all separated into 4 p_T bins. The J all sample is the cross section weighted combination of the J1-J4 samples. Each bin for MC has approximately 10000 events. The data bins are also included, however, the bins have only 3000 events so that there are approximately the same number of bins for the Jall and the data samples.

C.3.1 Reconstructed Monte Carlo

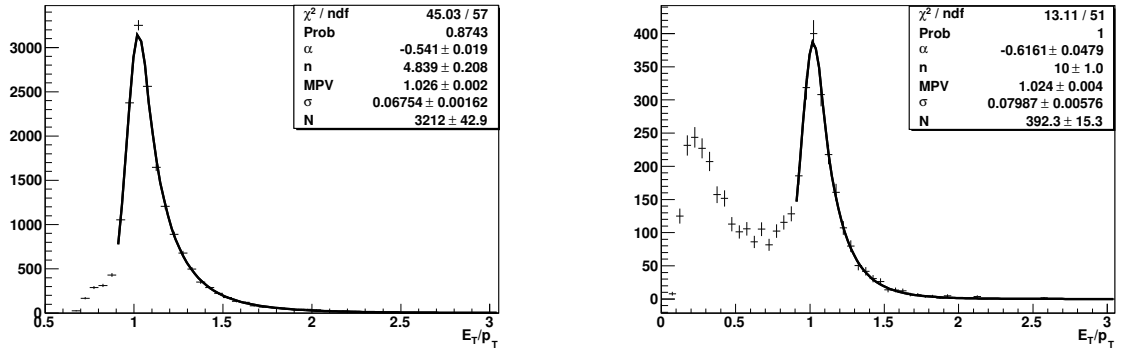
$$f(x;\alpha,n,\bar{x},\sigma) = N \cdot \begin{cases} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}, & \text{if } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \cdot \left(B - \frac{x-\bar{x}}{\sigma}\right)^{-n}, & \text{if } \frac{x-xb}{\sigma} \leq -\alpha \end{cases} \tag{C.1}$$

$$A = \left(\frac{n}{\mid \alpha \mid}\right)^n \cdot e^{-\frac{|\alpha|^2}{2}}$$

$$B = \frac{n}{\mid \alpha \mid} - \mid \alpha \mid$$

The reconstructed MC and data E/p distributions are fit using a Crystal Ball function shown in Equation C.1. The Crystal Ball function defines a probability density function used to describe lossy processes and consists of a Gaussian core and a power-law low end tail. The parameters are: α, n, $\bar{x}$, and σ. $\bar{x}$ is the most probable value (MPV) of the peak of the Gaussian (which would be correspond to the peak of E/p) and σ is the Gaussian width. α is the transition point between the Gaussian and power-law and n is the power. In this study, the tail was at the high end of the spectrum, so the sign of α was switched to accommodate the high end lossy track p_T tail. An example of this fit to the data in one p_T

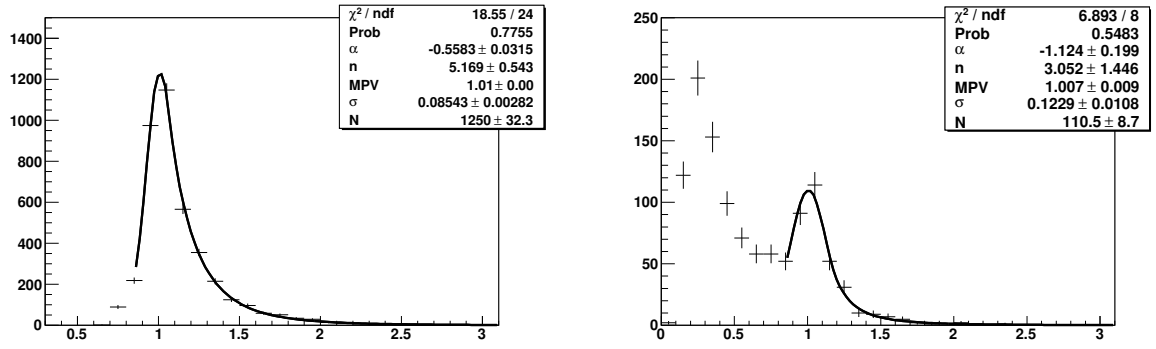
bin is shown in Figure C.6. The Figure on the left is a E/p distribution for low p_T photons. The Crystal ball function describes the data well in these regions, however when the p_T increases a secondary peak starts to grow in low E/p regions (generally $E_T/p_T < 0.6$). The second peak comes from poorly reconstructed or overlapping tracks in the ID. These tracks generally have higher χ^2, convert in the outer regions of the Silicon detectors, and overlap with each other, which is what causes the second peak. At very high p_T, charged particles in the ID curve less under the influence of the magnetic field. This makes p_T resolution worse in the ID in these ranges.



(a) (b)

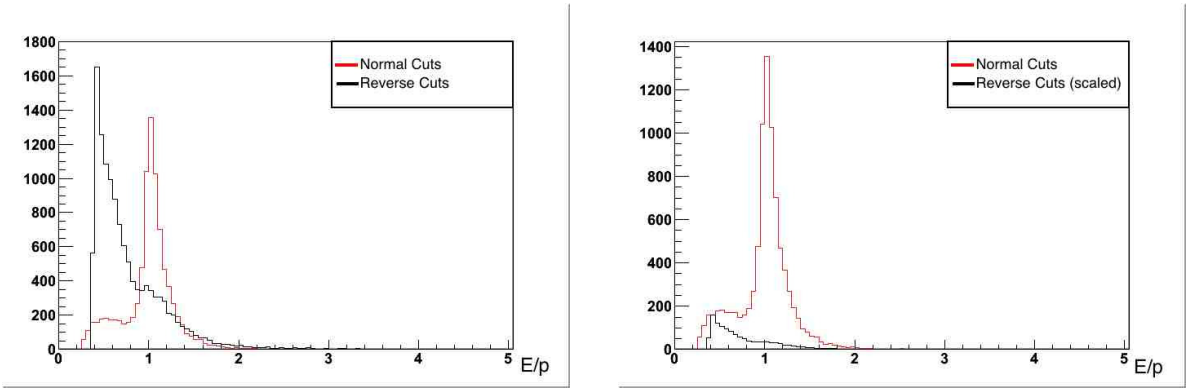
Figure C.6: Example of fitting the reconstructed MC in the bins (a) $35 < p_T < 38$ GeV (b)$102 < p_T < 280$ GeV using the Crystal Ball function for the combined Jall photon + jet sample. The Crystal Ball function describes the E/p distribution well at low p_T. At very high p_T, charged particles in the ID curve less under the influence of the magnetic field. This makes p_T resolution worse in the ID in these ranges.

Figure C.7 shows the results from fitting the Crystal Ball function to the data. A similar effect is seen in both the reconstructed photon+jet MC and the data. The Crystal Ball function describes the E/p distributions in data well. Again when the p_T increases ,a secondary peak starts to grow in low E/p regions. In the data, however, the peak is twice the size of the secondary peak modeled in the MC. This is most probably due to dijet (QCD) contamination in the data. There was an attempt to model this dijet contamination by reversing the χ^2 and TRT fiducial cuts in the data. If the extra contamination is modeled by reversing these cuts, then the additional shape can be added into the MC photon+jet sample. The results of the reverse cuts are shown in Figure C.8. Although it is quite clear from the plots that reversing the cuts does in fact enhance the low E/p peak, the shape is quite different than the secondary peak. Therefore this method alone can not be used to model the dijet contamination.



(a) (b)

Figure C.7: Crystal Ball Fits to Data (a) $35 < p_T < 37$ GeV (b)$102 < p_T < 280$ GeV

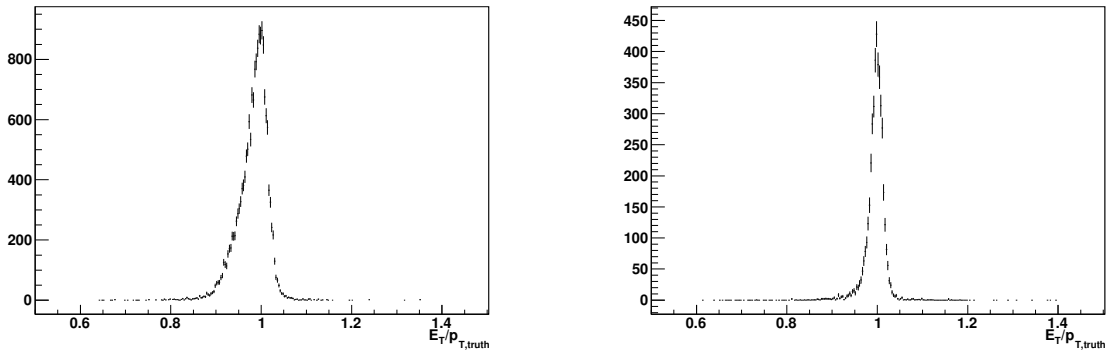


(a) (b)

Figure C.8: Reverse Cuts to Model QCD

C.3.2 E/p using true p_T: Tracking Systems Impact

The E/p spectrum computed using true particle p_T (truth MC) is shown in Figure C.9. The resulting spectrum is not a single value exactly one because only the p_T is from the truth particle container of the photon. E_T is the reconstructed value from the photon EM object as in Section C.3.1. This is done to observe how the value of E/p depends on the tracking systems. The truth MC peaks much more sharply than the reconstructed MC making fitting to a function difficult as shown in Figure C.9. Instead of fitting to a Crystal Ball function, the mean and error on the mean were used to compare plots. This is justified because the of sharpness of the peak and because the focus is on the peak, not the other fit parameters.



(a) (b)

Figure C.9: Example of the truth MC in the bins (a) $35 < p_T < 38$ GeV (b)$102 < p_T < 280$ GeV for the combined Jall photon + jet sample. This shows clearly that the reconstructed calorimeter E_T is better at high E_T - as opposed to low E_T

C.4 Results

This section presents the results of the methods described in the previous sections. The impact of the resolution on E/p is presented. The photon + jet samples are first studied individually by jet sample, Jn. Then the samples are combined once by cross section and once by giving equal weight to each MC event (this is to accurately represent the number of events in relation to each other and to represent the number of events achieved after a period of data taking). Finally the results for detector response are presented as functions of p_T and η.

C.4.1 Resolution and Generator Threshold Bias

To determine whether the sharp rise at low p_T is caused by the tracking reconstruction or the calorimeter reconstruction, the resolution of both the $1/p_T$ (tracking variable) and E_T (calorimeter variable) is shown in Figure C.10. This study was done with an older MC set that the rest of the Appendix. This used MC08 photon+jet samples generated at $\sqrt{s} = 10$ TeV and should be redone with the new MC09 production. The resolution is defined as one standard deviation value (σ) of a Gaussian fit to the Offset given by Equation C.2 where (a) is the variable which is being investigated. The resolution plots for $1/p_T$ and E_T should result in a Gaussian centered at zero. The MC reconstructed plots show that the $1/p_T$ is causing the E/p to spike. Furthermore, the MC truth plots do not show this effect. The overall offset also implies that the η integrated p_T is usually reconstructed too low and in particular at low p_T, the the value of $p_{T,reco}$ is disproportionately small.

$$Offset(a) = \frac{a_{reco} - a_{true}}{a_{true}} \tag{C.2}$$

As stated in Section C.2.1, the photon+jet samples start at a true photon energy of 17 GeV. There are tracks which have a p_T significantly below the $\hat{p}_\perp$ of the hard scatter and are excluded from the study. The second point in the reconstructed plot shown in Figure C.10 is about 20% higher than the expected value from the truth E/p. Since this point is near the generator cut off there is a bias introduced because there are missing events which would be normally be present that are less than than 17 GeV. From preliminary studies it appears that removing events within 5 GeV of the generator threshold for the J1 sample would remove the missing event bias. This effect is less significant in the higher energy jet samples due to overlap from the previous jet sample and the lower cross section. The remaining results can be seen in Figure C.14(b)

C.4.2 Fixing Parameters

The Crystal ball parameters as a function of p_T resulting from fits to their E/p distributions for the reconstructed MC are shown in Figure C.11. For Figure C.11(b) (the MPV), the parameters plotted from the MC truth are the mean of the $E_T/p_{T,truth}$ distributions.

Figure C.11(c) shows that the parameter, n, appears to be constant. The values of n were fit to a constant excluding the outlier and found to be $n_0 = 4.8 \pm 0.2$. The results are shown in Figure C.12. The MPV plots were fit to a constant, to determine the E/p over the p_T range. The overall results (shown in Figure C.12) do not show significant improvement on the constant fit to the E/p MPV plots (shown in the upper right of Figure C.11(b) and Figure C.12(b)). Instead the same MPV was found for both circumstances.

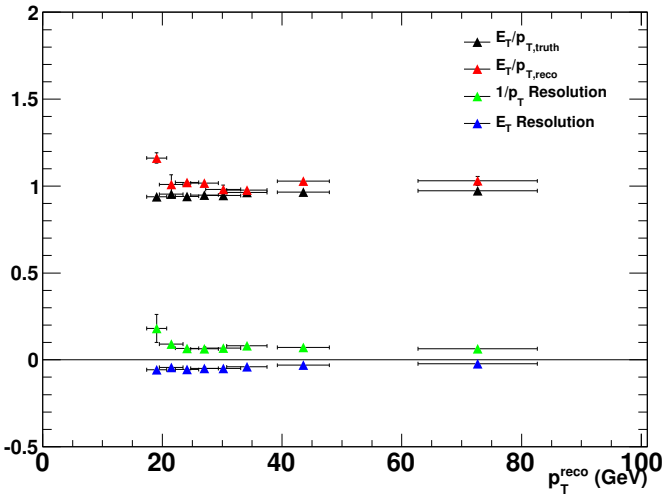


Figure C.10: The peak positions of $E_T/p_{T,reco}$, $E_T/p_{T,truth}$, and E_T and $1/p_T$ resolutions vs. p_T.

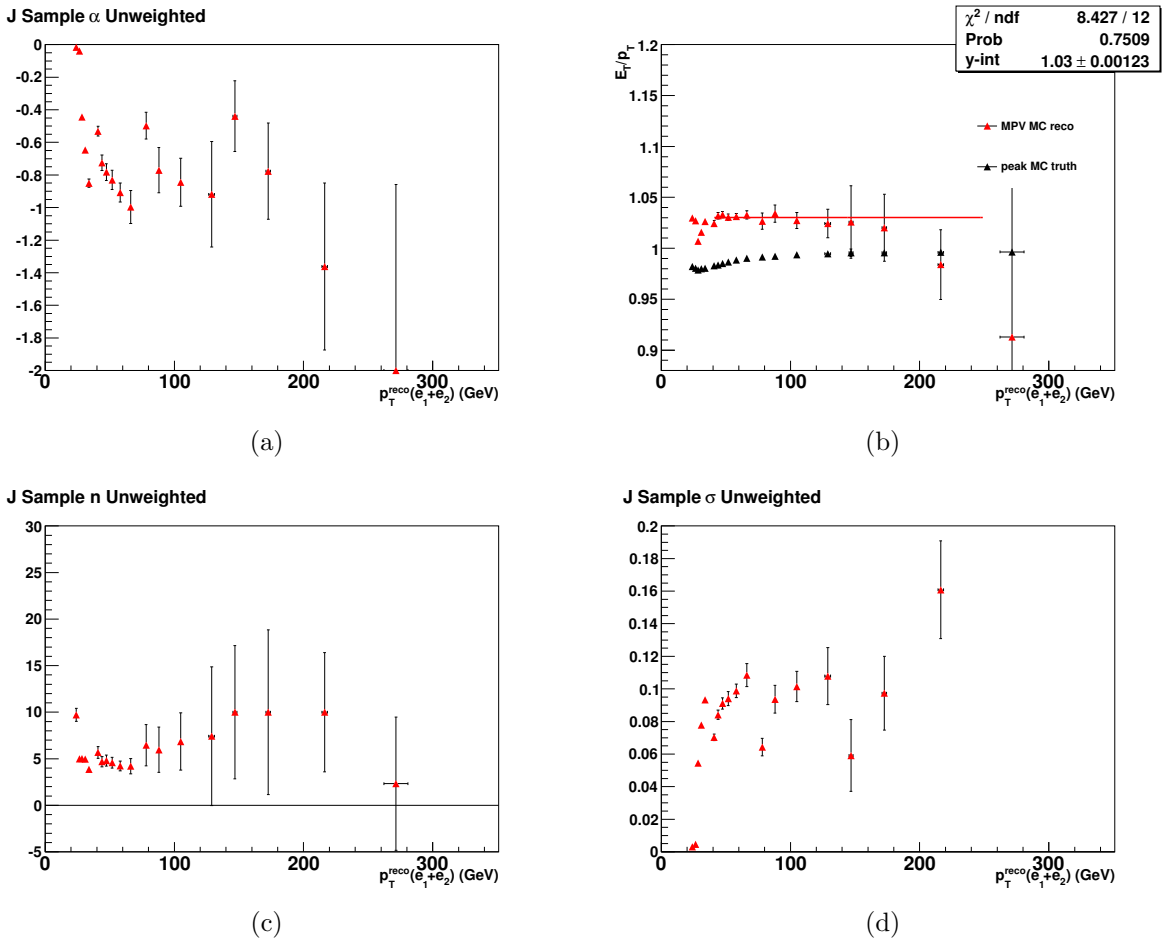


Figure C.11: The parameters of the Crystal Ball function. The plots show results from the J1-J4 samples, overlaid. (a) α, (b) MPV from fits to the E/p as functions of p_T, (c) n and, (d) σ.

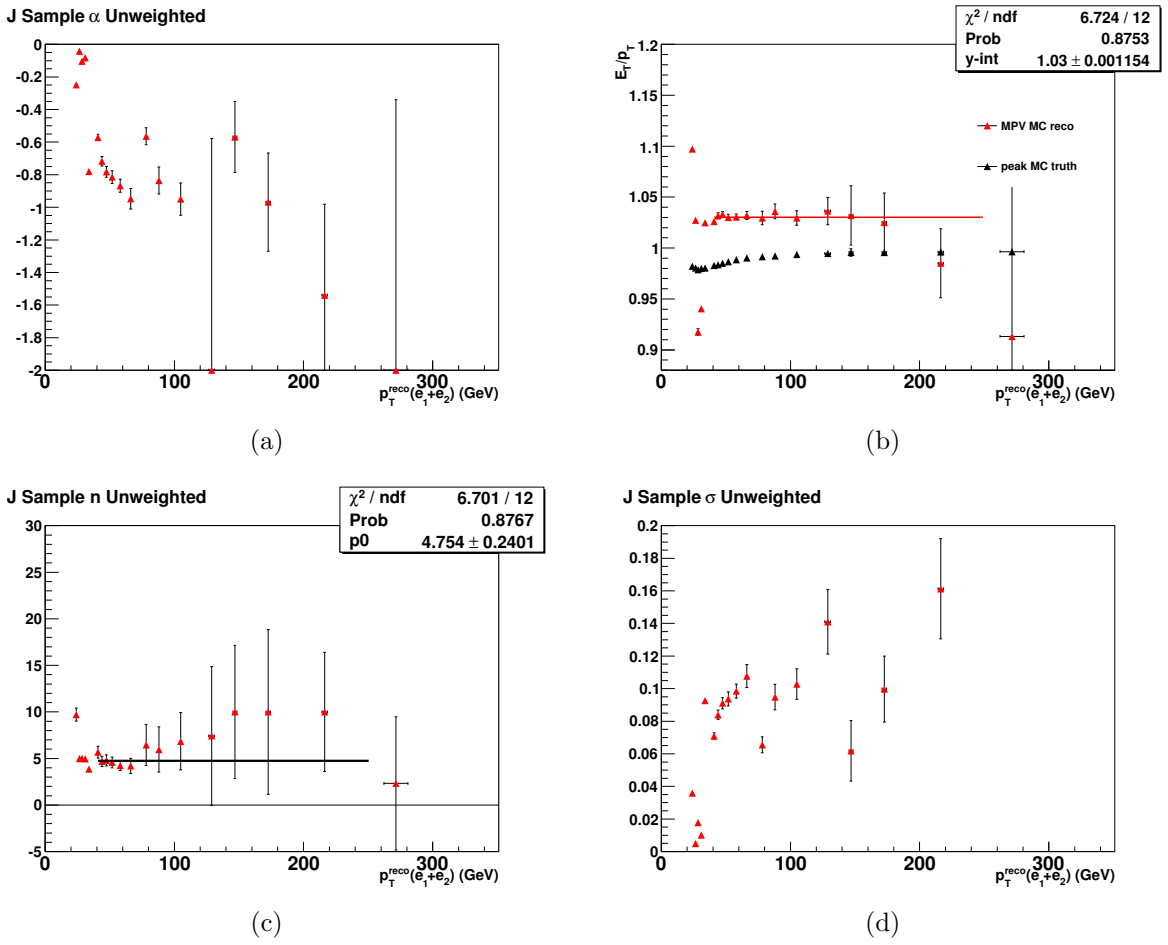


Figure C.12: The parameters of the Crystal Ball function with n fixed to $n_0 = 4.8 \pm 0.1$. The plots show a combination of the J1-J4 samples, unweighted. This shows that fixing n improves the 4th point, however overall results do not show significant improvement. (a) α, (b) MPV (c) n fixed to 4.8±0.2 (d) and σ.

C.4.3 Response

Since the MC had fewer statistics than would be expected in 35 pb^{-1} of data in the lower Jn samples, the first results look at the individual J samples ignoring relative cross section difference. Figure C.13 shows the response for the individual J samples. The response vs. p_T was fit to a constant and the results are included. The highest p_T sample, J4, shows a trailing off in the resolution because the tracks are harder to reconstruct at high p_T. Figure C.14 shows the response for all J samples. Figure C.14(a) shows the results from Figure C.13 overlaid and unweighted. The resulting reconstructed J samples are fit to a constant in the range 40 - 250 GeV. The results of this are displayed in the upper right corner of the plot. Figure C.14(b) shows the result of the cross-section combined MC and the data. Both of these are again fit to constants and the results are shown in the upper right. The two fits are not consistent with each other within errors, however the errors are calculated only using statistics no systematics. A systematic study should be done to complete the analysis.

C.4.4 η Dependence

Each p_T bin from Figures C.13 and C.14 were separated into four η regions. The four regions studied are defined by the EM barrel calorimeter: $\eta < -0.8$, $-0.8 < \eta < 0$, $0 < \eta < 0.8$, and $\eta > 0.8$. The results are shown in Figure C.15. The η value plotted on the x axis was determined by taking the average of the η in each of the regions. In each of the η bins there are $\sim$2000 events. The J2 region is the most consistent over η. This region in p_T is well calibrated because it is the range Z calibration. There is structure in both J3 and J4 at large $|\eta|$. In general for the points of lower p_T and large $|\eta|$ the values for E/p were greater than one, whereas at higher p_T and large $|\eta|$ the values for E/p were less than one. In the central η regions, the values for E/p were closer to one.

Figure C.16 shows the results from the cross section weighted photon+jet samples on the left and the data on the right. There was one poorly fit point in the MC which caused large statistical errors in the $-0.8 < \eta < 0$ bin. This was investigated, but since the rest of the fits were good, and the value for the response was reasonable, another fit was not attempted.

C.5 Summary and Future Work

This study reports use of E/p from photon conversions in photon+jet samples to investigate the detector response and calibration over a wide range of energies. The J4 sample shows that we can explore energy ranges to 300 GeV. Each MC data set used was comprised of at least 1,000,000 events. To get this number of events at the highest energy would require an integrated luminosity of 1 fb^{-1} of data. However, 35 pb^{-1} is enough data to utilize photon conversions to almost the J3 range and lower ($\sim$ 150 GeV). This strategy looks promising for extrapolating between J/ψ, Υ and Z in photon+jet sample immediately now that data

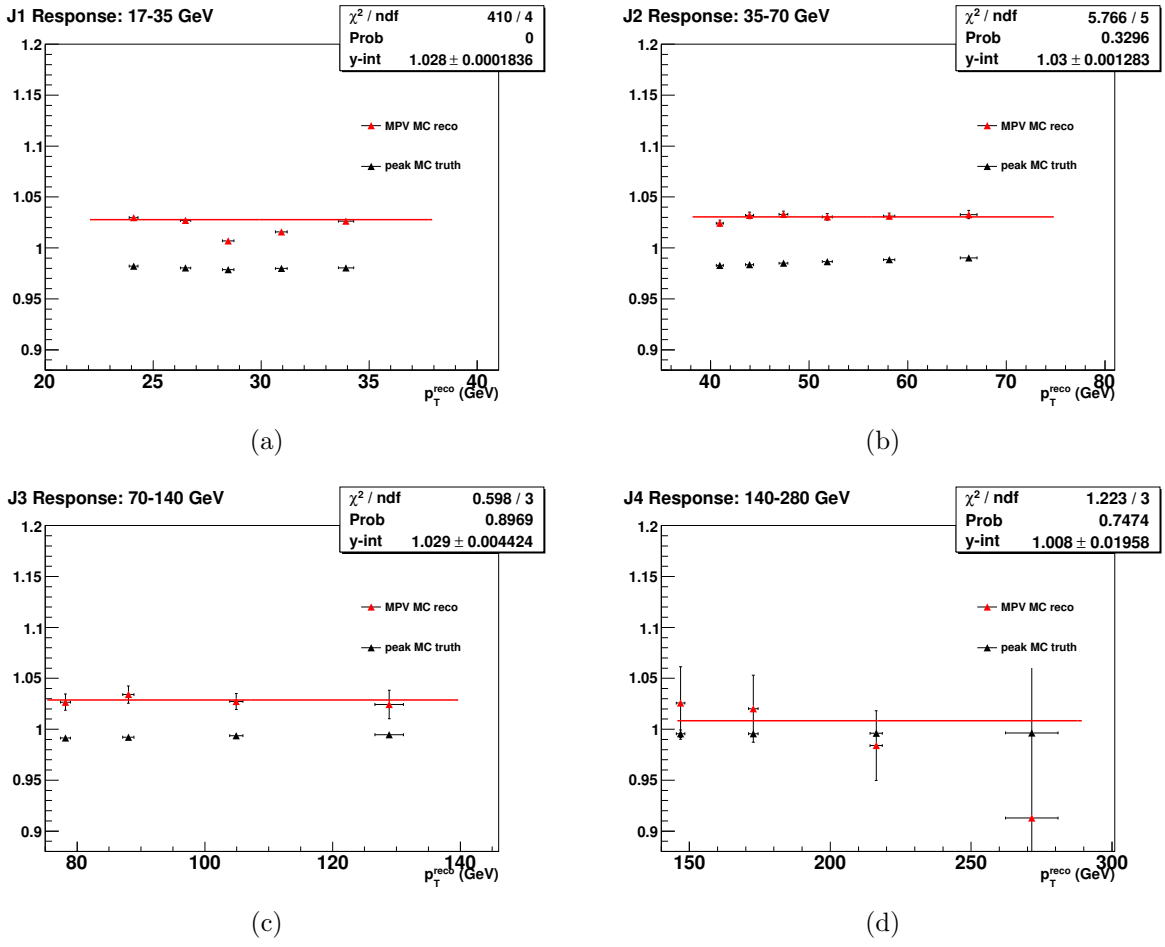
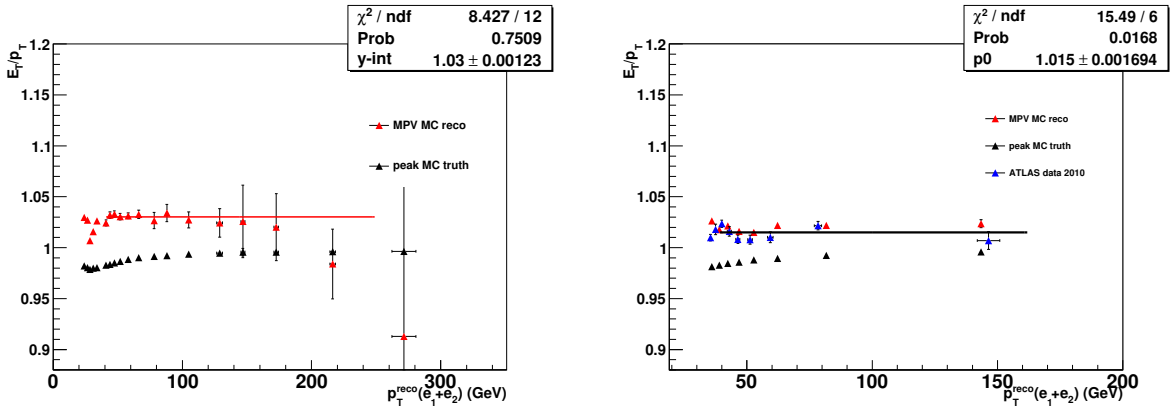


Figure C.13: The results of fitting the individual samples. (a) J1 (b) J2 (c) J3 (d) J4



(a) (b)

Figure C.14: (a) The results of the MPV of the Crystal Ball fit for which represents the E/p for unweighted J1-J4 samples. (b) The results of the MPV of the Crystal Ball fit which represents the E/p for the cross section weighted samples. The results are also fit to a constant which is shown in the plot.

has begun to be collected. Especially for the J1-J3 data samples, a constant provides a reasonable fit and was found to be 1.015±0.002 (statistical errors only, no systematics). In the J4 Response plot, all but the last point was consistent with a constant. The last point is also near of the generator limits (280 GeV) of the hard scatter, and therefore may be bias.

The combined cross section weighted samples also show that the E/p response is nearly constant over the energy range 25-140 GeV. The total MC photon+jet sample had ∼9,000,000 events. The unweighted sample combination also shows E/p is consistent to about 180 GeV and arcs down from 180-280 GeV. From this study, E/p is 1.030±0.001 for $|\eta| < 2.5$. When looking at the central region, $|\eta| < 1$, E/p was found to be 1.005±0.002 . Initial studies were done to find a way to remove these events at high η or understand why the E/p is more poorly reconstructed. However, further studies are needed to draw conclusions.

This study focused on the intrinsic ability to check the calorimeter calibration in energy ranges not probed by J/ψ, Υ, and Z decays. Additional work is needed to account for systematic effects in the track and calorimeter reconstruction. Also, the study should be extended with tests using MC data sets with a miscalibrated EM calorimeter. Dijet MC and QCD enhanced data samples are needed to derive conclusive results for dijet E/p distribution shapes. Additionally cross checks in the calculation of E/p with the J/ψ and Z samples should be done to cross check the method.

Special thanks to J.F. Marchand, M. Begel, H. Ma and S. Snyder for their discussions concerning photons and python. Thanks to Andrew Hamilton for producing the photon skims from the e/γ data streams.

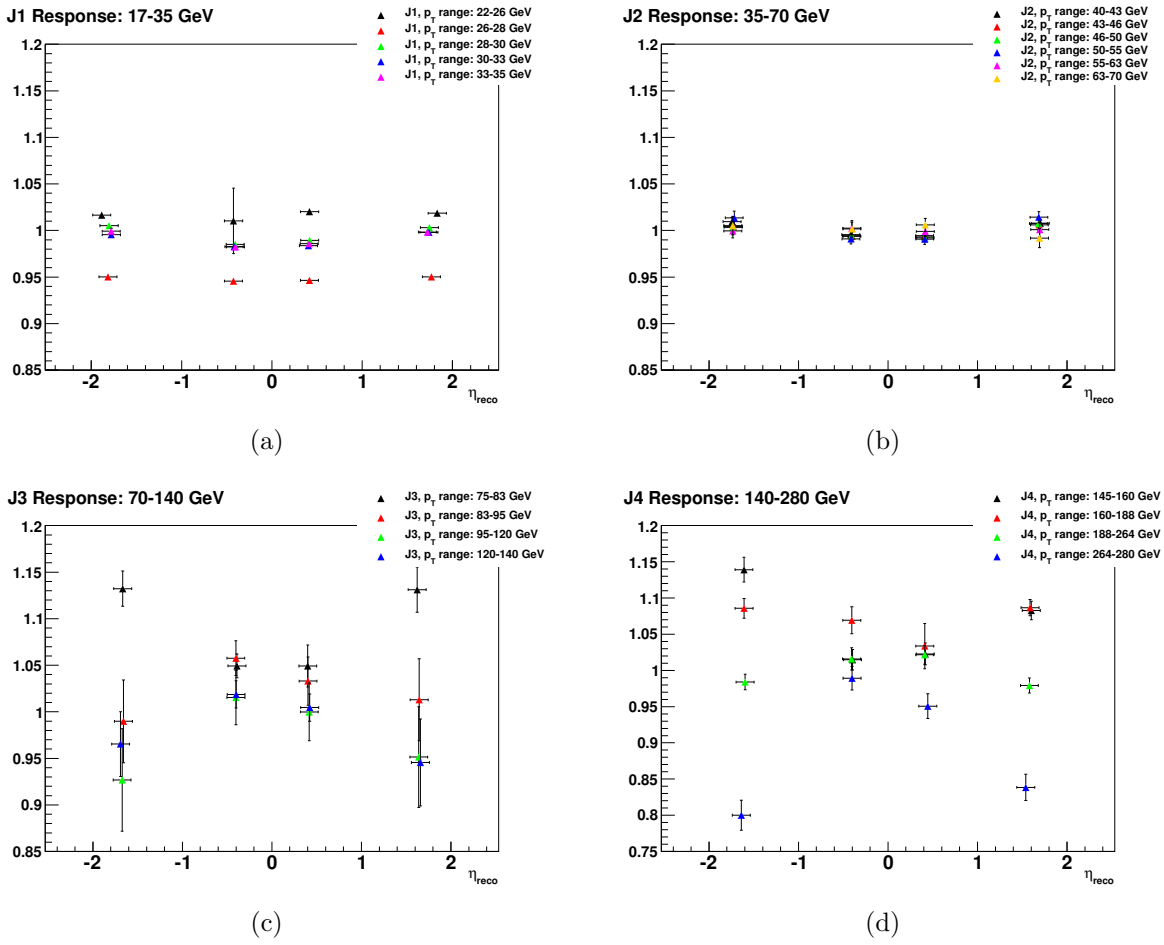
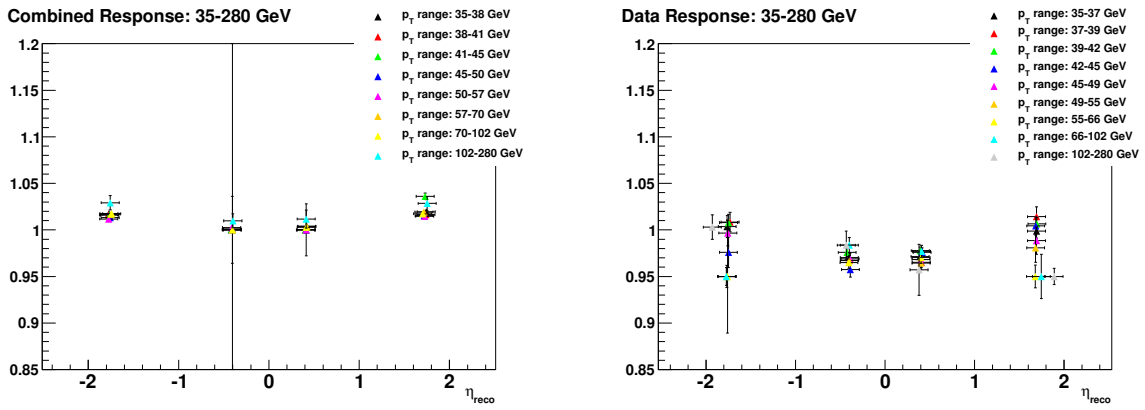


Figure C.15: η binned samples (a) J1 (b) J2 (c) J3 (d) J4



(a) (b)

Figure C.16: η binned Response: Combined and Data