A Feasibility Study on Measuring a Strange Sea Asymmetry in the Proton at ATLAS

Laura Gilbert

Subdepartment of Particle Physics

University of Oxford

l.gilbert1@physics.ox.ac.uk

10th October 2008

Abstract

The findings of a study into the feasibility of placing experimental constraints on a hypothetical asymmetry in the strange component of the proton sea, and on the strange components as a whole, at ATLAS are presented here.

The strange and antistrange components of the proton sea (collectively known as the "strange sea") are generally assumed to be charge-symmetric, however the possibility of an asymmetric strange sea is indicated in theory and experiment, making this an interesting area for study in data. The strange component of the hadron sea at low values of x is not well constrained in data, and a measurement placing stronger limits on this will be vital in enabling Standard Model and beyond-Standard Model physics to be distinguished at the LHC. This Monte Carlo study addresses the question of whether the current limits on the strange sea could be improved using the ATLAS experiment, and whether if an asymmetry exists in the strange sea it is likely to be detectable at ATLAS.

This study was performed using approximately 1 fb^{-1} of simulated data. It was found that it will not be possible to measure properties of the strange sea in 1 fb^{-1} of data. It is indicated that new limits on the strange component of the sea, and a possible measurement of a strange sea asymmetry, are feasible with around 300 fb^{-1} of data provided that the levels of certain background channels with large cross-sections (namely $c\bar{c}$ and $b\bar{b}$ backgrounds) could be convincingly shown to be controlled in further, higher-volume studies with data and Monte Carlo.

Acknowledgments

I would like to thank my advisor, Jeff Tseng, whom I admire greatly both as a brilliant physicist and inspirational man. He is endlessly patient, never steered me wrong, always took my side, instilled in me a deep appreciation of a well-formed statistic, and never let me get away with anything.

I offer my gratitude to all of my colleagues at Oxford in particular Amanda Cooper-Sarkar and Tony Weidberg, and Maarten Boonekamp at CERN, for their very generous help and advice. I would also like to thank Nick and Alys Brett, Sam Harper, Muge Karagozm, Chris Dennis and Ian Preston for their friendship and for making work fun, the latter especially for offering to meerkat-sit whilst I got some work done, and Juliet Biggs and Dave Bonfield for being wonderful housemates.

I am also deeply appreciative for the help of Pat Roche who supported me when I started teaching and provided valuable career advice along the way.

I am very grateful to DELL for their support of my graduate studies, and in particular to Mark Cobban, and my co-workers and friends at DELL in Austin, Texas, in particular Monica Kashyapp, Garima Kochhar, Saeed Iqbal, Ron Pepper, and Jenwei Hsieh.

Most importantly, I would like to thank my family and close friends for their tireless love, support and encouragement, and for filling my life with joy. In particular I am forever grateful to my father Mike, who showed me the romance and power of science from an early age, and to my mother Anita, who instilled in me the conviction that nothing is unachievable if only one tries hard enough. Without their unwavering support in every possible aspect of my life, and without the example they have set me throughout, I could not have achieved any of the things that I have been proud of. This thesis is dedicated to them. I thank my sister Cassandra for her humour, for years of moral support and for her admirable, if only occasionally fruitful, efforts to prevent me dressing like a graduate student.

Special mention goes to Tom Dawson, who has been my one-man cheer team and my hero. I learn something new from him every day, and am fortunate to have him in my life.

Finally I would like to thank Nicki Cooper-Sarkar and Eva Moore who have been there for me unfailingly every time I've needed them.

Contents

1 Introduction

ATLAS [1] is a particle physics experiment being built at CERN in Geneva. It is due to start taking data in 2009 with the primary aim of searching for new physics beyond that which is known today. The detector will measure physics events provided by collisions between protons at the Large Hadron Collider (LHC). The LHC has a high discovery potential for new physics because it has a greater center of mass energy of collisions and higher design luminosity than any previous particle physics experiment.

The proton consists of two valence up quarks and one valence down quark in a sea of virtual quarks and gluons, bound together by the strong interaction. In the accepted theory of the strong interaction at high energies the proton sea is expected to have quark-antiquark symmetry. However a flavour asymmetry in the lightest sea quarks has been discovered in previous experiments, leading naturally to the question of whether the more exotic sea components may also exhibit some asymmetry. The first place to search for this will be in the behaviour of the third lightest quark, the strange component of the nucleon sea.

Deviations from the current predictions of the distributions of the strange sea could significantly affect the measurements of certain observables at ATLAS. For example, the ratio of the LHC Z^0 to total $W^{\pm}$ production cross-sections, $R_{ZW} = \sigma_{Z^0}/(\sigma_{W^+} + \sigma_{W^-})$ is likely to be sensitive to uncertainties in the strange and antistrange distributions [2]. The strange and antistrange quarks contribute to Z and W production via $s\bar{s} \rightarrow Z$ and $sc \rightarrow W$.

Much of the focus at ATLAS is the search for new physics. Many of the models which will be tested are extensions to the Standard Model (SM), which contain new gauge bosons. These will be identified by deviations from the SM predictions in dilepton events, the major background to which will be the decay of SM W and Z bosons. Uncertainties in predictions of the production cross-sections of these gauge bosons due to poor constraints on the strange sea must be reduced in order to distinguish between the SM and new physics.

A measurement of an asymmetry in the strange-antistrange distribution would also go some way towards explaining the origins of the NuTeV Anomaly [3]. The NuTeV collaboration measured the square of the sine of the weak mixing angle ($sin^2\theta_W$, where θ_W is the ratio of the masses of the W to Z bosons) to be approximately three standard deviations above the accepted Standard Model prediction, which was determined by fits to other electroweak measurements [4]. The NuTeV analysis assumed that the strange and antistrange distributions were symmetric, and a positive measure of the second moment of the strange-antistrange distribution $\int_0^1 x(s(x) - \bar{s}(x)) >> 0$ could reduce this effect to around 1.35 standard deviations from the SM fit [5][6][7].

To probe the strange quark components of the proton sea the channels $s + g \rightarrow c + W^-$ and $\bar{s}+g \rightarrow \bar{c}+W^+$ were studied, because in these processes differences in the momentum distributions of the charm and W boson could only arise from differences in $s/\bar{s}$ momenta. This makes these channels an ideal choice of process in which to study asymmetries in the strange sea. The W boson was identified by its characteristic decay into an electron and a neutrino, since these particles will be clearly identifiable at ATLAS, and the presence of a charmed quark was inferred by the reconstruction of a D^* meson (see Section 2.3 for details).

2 Theoretical Background

2.1 The Standard Model

The Standard Model of Particle Physics is a theory encompassing current knowledge of matter and its behaviour in terms of of the interactions of fundamental point-like particles. These carry a property of internal angular momentum, known as *spin.* The components of the theory which make up matter are the particles with half-integral spin (*fermions*) while those with integral spin (*bosons*) usually mediate forces between particles. Force-carrying bosons are known as *gauge* bosons.

The SM incorporates three forces tied into two major sub-theories, the Electroweak (EW) theory and Quantum Chromodynamics (QCD). The electromagnetic force is described by the theory of Quantum Electrodynamics (QED). This introduces massless photons (γ) as the gauge bosons which mediate the force. The weak force, in which massive W and Z bosons mediate the weak interactions, is combined with QED to form the EW theory. Quantum Chromodynamics (QCD) is the theory describing the strong force interactions, which are mediated by gauge bosons known as gluons.

The SM does not include any theory of gravitation, but the effects of gravity on fundamental particles interacting at high energies are negligible so this omission does not adversely affect the power of the Standard Model to predict particle interactions in the high energy collisions.

The set of fundamental fermions can be divided into two types: those which are able to interact via the strong force, and those which are not. The fermions which are sensitive to the strong force are known as quarks, those which are not are leptons. Quarks and leptons are both able to interact via the weak force. All quarks and half of the leptons carry electromagnetic charge (positive or negative) and these are able to interact via the electromagnetic force (in which particles with charges in the same sense repel, those with oppositely signed charges attract). Charged leptons carry charges of ± 1 (in units of Coulombs), whereas quarks carry charges of $\pm\,\frac{1}{3}$ or $\pm\,\frac{2}{3}$. The photons which mediate the Electromagnetic force are uncharged, as are the Z bosons, but the weak force mediating W boson carries a charge of ± 1. The uncharged leptons are known as neutrinos, which are thought to be virtually massless and thus interact exclusively via the weak force. The strong force interactions involve three possible charges, the *colour* charges, labelled red, green and blue. Individual types of quark are able to carry any one of these, and gluons carry combinations of the colour charges (with the restriction that they are not permitted to be colourless overall). QED and QCD are described in more detail in Section 2.1.1, and the weak force and EW theory in Section 2.1.2.

Each of the quarks and leptons has an *antiparticle* partner with an identical mass and spin but opposite electromagnetic and colour charges. (They are named after their particle partners such that the antiparticle of the up quark is called the antiup, excepting the electron for which the antiparticle is the positron.) They also hold opposite quantum numbers relating to particle “flavour”, for example the strange quark has the property of “strangeness” $S = -1$ and the antistrange quark $S = +1$, and in the case of quarks they hold an inverse colour charge, anticolour.

There are three so-called “generations” of each of the quarks and leptons. Each generation is made up of a pair of complementary quarks, and one of leptons. The pairs are known as “families”. The sets of quark or lepton families act identically under the three SM forces but have different masses. Ordinary matter is made up of only the particles from the lightest family, since when the heavier particles are created they are able to undergo interactions which result in transformations of these particles into the lightest counterpart, via the exchange of gauge bosons. The most important properties of the fundamental fermions of the SM are listed in Table 1 (the antiparticle partners of the fermions are omitted as their properties can be inferred). The particles are divided into their three generations, and their measured masses and widths are taken from the Particle Data Group (PDG) [8] numbers.

Gen.	**QUARKS**				**LEPTONS**			
	Name	Symbol	Charge	Mass	Name	Symbol	Charge	Mass
1	Up	u	$+\frac{2}{3}$	1.5-3.0 MeV	Electron	e	-1	0.511 MeV
	Down	d	-$\frac{1}{3}$	3-7 MeV	e Neutrino	ν_e	0	~0
2	Charm	c	$+\frac{2}{3}$	1.25 ± 0.09 GeV	Muon	μ	-1	105.7 MeV
	Strange	s	-$\frac{1}{3}$	95 ± 25 MeV	μ Neutrino	ν_μ	0	~0
3	Top	t	$+\frac{2}{3}$	172.5 ± 2.7 GeV	Tau	τ	-1	1777 MeV
	Bottom	b	-$\frac{1}{3}$	4.20 ± 0.07 GeV	τ Neutrino	ν_τ	0	~0

Table 1: Properties of spin $\pm\frac{1}{2}$ fermions found in the standard model

The properties of the gauge bosons are listed in Table 2, along with the forces they mediate.

BOSONS					
Name	Symbol	Force Mediated	Charge	Mass	Spin
Photon	γ	Electromagnetism	0	0	1
W boson	W	Weak	± 1	80.4 GeV	1
Z boson	Z	Weak	0	91.2 GeV	1
Gluon	g	Strong	0	0	1

Table 2: Properties of integral spin bosons found in the standard model

2.1.1 QED and QCD

QED is the theory of the electromagnetic interaction, describing the interactions between electrically charged particles. The gauge bosons which propagate the electromagnetic force (photons) are not themselves charged. The potential of the QED field (V_{QED}) can be derived from the well-known Coulomb force between two charged particles of charges q_1e and q_2e (where e is the electronic charge), and is given by:

$$V_{QED}(q_1, q_2, \mathbf{r}) = -q_1 q_2 \frac{\alpha}{|\mathbf{r}|} \tag{1}$$

(in natural units, i.e. $\epsilon_0 = c = h = 1$) where the *fine-structure constant* α measures the strength of the electromagnetic interaction. It is defined by:

$$\alpha = \frac{e^2}{4\pi} \tag{2}$$

where α depends on Q^{2*}, which is a measure of momentum transferred between the particles. As a result the particle charge e must also depend on Q^2, so $e(Q^2)$ is the "effective charge" of one particle as seen from the viewpoint of the other. It turns out that the charge when there is no interaction, i.e. $Q^2 = 0$, is very small (at this point it becomes the well-known electronic charge $e = 1.60218$ x $10^{-19}C$) and α does not become large on the scales current experiments are able to reach.

The higher the energy of the collision the greater the momentum transfer will be and so the greater the charge of each seem to be to the other. Therefore as the charges get closer together the greater the effective charge they feel and the stronger the QED field is, and as they move far apart the interaction between them lessens in strength, i.e.

$$\frac{d}{dQ^2} e^2(Q^2) > 0 \tag{3}$$

---

[*] Q is the kinematic variable of momentum transfer in particle scattering, i.e. if particles 1 and 2 scatter to produce particles 3 and 4 with momenta p_1, p_2, p_3 and p_4 respectively then $Q = p_3 - p_1 = p_4 - p_2$

or

$$\frac{d}{dQ^2}\alpha(Q^2) > 0 \tag{4}$$

QCD is the theory of the strong interaction, describing all interactions of quarks and gluons, where gluons are the guage bosons which mediate the strong force. The gluon carries colour charge itself, unlike QED in which the photon carries no electromagnetic charge. Also unlike the single photon of QED there is more than one type of gluon. As a result gluons undergo "self-interaction", they can interact directly with one another and so the simple Coulomb law force of QED no longer applies. Furthermore, because gluons mediate interactions between two quarks, each of which has one of the three colour charges red (r), green (g) or blue (b) (or the equivalent anticolours) the gluon itself will carry the sum of two colour charges, depending on the interaction it mediates. The possible gluon states are given as linear combinations of sets of colour-anticolour pairs (for example one possible combination might be $r\bar{b} - b\bar{r}$). In this representation it is possible to create nine linearly independent states - however one of these states would be a "colour singlet" (i.e. $r\bar{r} + g\bar{g} + b\bar{b}$). This gluon would be non-interacting and so is not allowed in QCD, leaving eight types of gluon overall.

By analogy with QED a fine structure constant α_s can be defined for QCD:

$$\alpha_s = \frac{g^2}{4\pi} \tag{5}$$

where g is the effective QCD charge (colour). At fairly short distances the potential due to the strong force field can be approximated to of the sum of a Coulomb term and a potential which varies linearly with r [9].

$$V_{QCD} \propto \alpha_s(-\frac{1}{r} + \sigma r) \tag{6}$$

The reason for this linear potential is that the strength of the force between two quarks depends on the density of gluons between them. As energy is put in to increase the distance between them the gluons holding them together are able to interact and produce more gluons, increasing the strength of the force. As the separation between two quarks increases a narrow tube of self-interacting gluons is produced, increasing the strength of the force with the distance (unlike QED in which the force lessens with distance). As a result of this the behaviour of α_s with increasing Q^2 is in the opposite sense to the QED α (Equation 4):

$$\frac{d}{dQ^2}\alpha_s(Q^2) < 0 \tag{7}$$

This leads to a phenomenon known as *confinement*: because the force increases as two quarks are separated it would take an infinite amount of energy to separate them, and they are thus unable to exist as single entities (free quarks have never been observed).

Because of confinement particles carrying colour charge (quarks or gluons) which are produced in a scattering interaction must be bound into composite bodies known as *hadrons*. The presence of one free coloured particle excites quark/antiquark pairs and gluons from the vacuum, which in turn excite more in a *fragmentation* process. These group together, eventually creating a narrow cone or *jet* of hadrons leaving the interaction point along the path of the original particle in a process called *hadronisation*.

Hadrons are required to have a total colour charge of zero, and thus they are divided into two different types. The baryons contain three quarks (or three antiquarks), one of each with colour charge red, green and blue (or antired, antigreen and antiblue) which equates to a total colour of zero. The mesons contain a quark and an antiquark, for example a quark with red colour charge and an antiquark with antired, which again results in a composite particle which is colourless overall. (Hadronic structure is discussed more fully in Section 2.2.)

Because of confinement the coupling between quarks decreases as quarks move closer together, giving the strong force field the property of *asymptotic freedom*. This means that the strong force is relatively weak within the hadron itself and the quarks can there be considered to be able to move independently of one another. The higher the energy of the hadron, the higher the momenta of the individual quarks and gluons which make it up, and the weaker their strong interactions appear. The strength of the strong interaction is described by the strong coupling constant α_s, which therefore depends on the energy of the hadronic system.

It transpires that the amplitude of the interaction between quarks from different hadrons (a physical quantity which can be measured) can be expanded as a polynomial equation in terms of α_s [10]. At short distances and high energies α_s becomes small enough that this amplitude can be considered to be a perturbation of a simple potential. In this regime perturbation theory can be applied to calculate physical observables such as the interaction strengths and cross-sections (interaction probabilities) of different types of quark interactions, the theory in this case is known as perturbative QCD (pQCD).

At longer distances or lower energies the potential becomes highly non-linear (the non-perturbative region) and so analytic or perturbative solutions to the theory are difficult, if not impossible, to find.

2.1.2 The Weak Force and Electroweak Theory

The weak interaction is mediated by three massive gauge bosons: the W^+ and W^- bosons, which are an electrically charged particle-antiparticle pair, and the Z^0 which is neutral. These gauge bosons are very massive (around a hundred times more massive than a proton), the mass of the W being 80.425 ± 0.038 GeV and of the Z 91.187 ± 0.002 GeV, meaning that they are in general very short-lived. As a result the range of the weak force is very limited, and so weak interactions proceed much more slowly than other SM interactions, and over shorter ranges.

The weak interaction has some unusual characteristics. It is the only SM force which interacts with neutrinos. It is also the only force which is capable of changing the flavour of the particles whose interactions it mediates.

The strong interaction conserves particle flavour, as does the electromagnetic, but the weak force is able to change the flavour of leptons and quarks (note that it is not sensitive to quark colour). For example in beta decay a neutron decays to a proton via the weak interaction, emitting an electron and a neutrino. On the quark level this corresponds to a d quark from the neutrino decaying into a u quark and emitting a negatively charged W boson which then decays into an electron and neutrino. A Feynman diagram for this process is shown in Figure 1.

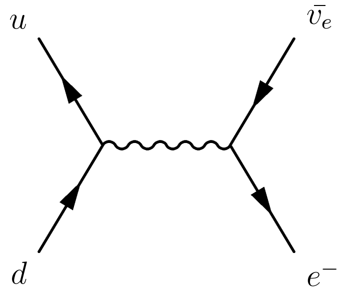


Figure 1: Feynman diagram for beta decay

It is most common for quark flavour changes to occur within a generation, but this is not always the case. In the weak kaon decay $K^- \rightarrow \pi^0 e^- \bar{\nu_e}$ the strange quark from the kaon is allowed to decay into a u quark, Figure 2. The property of strangeness is lost in this interaction, which would be forbidden in any other type of interaction. These cross-family interactions are less common than those within a family, and interactions between the first and third families are very unusual. However lepton interactions are contained within families, so a W is permitted to decay into an electron and an electron neutrino, but not a muon and an electron neutrino for example (this is the conservation of lepton number).

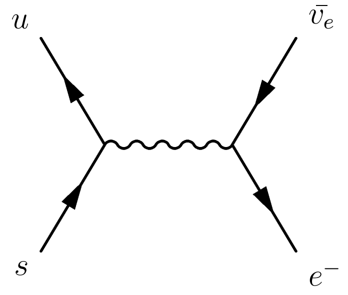


Figure 2: Feynman diagram for $K^- \rightarrow \pi^0 e^- \bar{\nu_e}$ decay

The reason for this difference in the behaviours of the strong and weak forces is that the quarks are effectively eigenstates of the strong and electromagnetic forces, but not of the weak. If the strong interaction eigenstates are the u, d, s, c, b and t quarks, which could be written

$$\begin{pmatrix} u \end{pmatrix}, \begin{pmatrix} d \end{pmatrix}, \begin{pmatrix} c \end{pmatrix}, \begin{pmatrix} s \end{pmatrix}, \begin{pmatrix} t \end{pmatrix}, \begin{pmatrix} b \end{pmatrix}$$

then it turns out that the weak eigenstates are the three families of lepton, and also linear combinations of the strong force eigenstates. These could be represented as follows:

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}, \begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \begin{pmatrix} t \\ b' \end{pmatrix}$$

where d', s' and b' are related to the strong eigenstates by

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \mathcal{M} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

$\mathcal{M}$ is known as the Cabibbo-Kabayashi-Maskawa (CKM) matrix. The elements of the matrix have been measured experimentally, and the limits on each are given here within 90% confidence limits. The values along the leading diagonal are close to one, meaning that most weak interactions are kept within the same quark family. It can be seen that the mixed states include small amounts of the adjacent families, and that there is very little mixing between the first and third families.

$$\mathcal{M} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.9745 - 0.9760 & 0.217 - 0.224 & 0.0018 - 0.0045 \\ 0.217 - 0.224 & 0.9737 - 0.9753 & 0.036 - 0.042 \\ 0.004 - 0.013 & 0.035 - 0.042 & 0.9991 - 0.9994 \end{pmatrix}$$

There are several commonly used parametrisations of the CKM matrix, one of the most common (the Wolfenstein parametrisation) is:

$$\mathcal{M} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^{\triangle})$$

where λ, A, ρ and η are four independent variables. The first three are real, and η is complex. These numbers have been measured experimentally and are of the order unity. λ is an important physical constant which appears in all of the weak mixing, the value of which has been measured in a variety of weak interaction experiments (such a the weak decay of B of K mesons in e^+e^-collision experiments such as CLEO [11] and BaBar [13], hadron-hadron collision experiments such as NA48 [12], and neutrino-nucleon collision experiments such as NuTeV [3]) to be approximately 0.2.

Another unusual property of the weak force is its observed preference for interactions with left-handed particles (and right-handed antiparticles). In the limit that the particles in question are effectively massless (true for neutrinos, and for sufficiently energetic electrons) the charged weak current (i.e. those interactions involving the W boson) interacts exclusively with particles with these characteristics. This is built into the theory of the weak force by describing the W boson propagator as a “V-A”-type propagator (vector minus axial vector). Because of this dependence on axial properties of the particles it interacts with, the weak force violates the parity (P) and charge-parity (CP) symmetries. This bias leads to a slight inherent bias towards the creation of matter over antimatter in the universe.

In 1979 Glashow, Salam and Weinberg won a Nobel Prize for their work on unifying the weak force with the electromagnetic force into one mathematically functional field theory with four gauge bosons, two neutral and two with opposite charges. They showed that it was possible (and convenient) to model the electromagnetic and weak forces as manifestations of the same unified electroweak force. The three lepton families are themselves eigenstates of the electroweak force, explaining why they are not mixed whilst the quarks are. The theory is highly successful in predicting standard model interactions, but it depends heavily upon the existence of a mechanism which allows symmetry breaking within this theory, to permit the gauge bosons to have different masses (the photon is massless, the W and Z bosons are very massive). This theoretical mechanism is known as the Higgs mechanism, and it replies upon another massive boson which has not yet been detected, but is famously known as the Higgs boson. Discovery of the Higgs boson is one of the primary goals of the LHC.

The mathematical workings of the Electroweak theory and Higgs mechanism are outside the scope of this thesis.

2.2 The Structure of the Proton

Protons are hadrons, classically consisting of two up (u) quarks and one down (d) quark (known as the *valence quarks*) which are held together by a *sea* of gluons. The gluons in the sea are able to split into short-lived quark-antiquark pairs, which can recombine back into gluons, meaning that the proton contains not only the valence quarks but also pairs of *sea quarks*. The sea particles are *virtual*, obeying Heisenberg's Uncertainty Principle (i.e. the more energetic they are the shorter time they can exist for). Virtual particles obey conservation laws in the same way that *real* particles do, but unlike real particles they are not required to obey the relativistic relationship

$$E^2 = p^2 + m^2 \tag{8}$$

Particles which do not follow this relationship are known as *off-mass-shell*. The further off-mass-shell a virtual particle is the less probable its existence, and for this reason the heavy quark components of the sea are statistically much less dense than the light quark contributions. The valence quarks, sea quarks and gluons which make up hadrons are collectively known as *partons*.

Most of the current knowledge of the properties and densities of the different types of parton within the proton has been extracted from Deep Inelastic Scattering (DIS) experiments, such as those at the HERA [14] collider. In these, electrons and protons were collided together at high energies to probe the sub-structure of the protons, to gain information about the densities of the individual types of parton within the proton and how its momentum is shared out amongst them.

High-energy collisions between two protons can be approximated as collisions between effectively free partons, as was discussed in Section 2.1.1. Therefore hadron-hadron colliders can also used to gain further knowledge of hadronic structure. ATLAS is one such experiment.

2.2.1 Parton Distribution Functions

Parton distribution functions (PDFs) give the probability of finding a parton in a hadron as a function of $x^\dagger$ and Q^2 [15]. They are determined by the results of DIS experiments.

Consider an interaction between two hadrons, hadron A and hadron B, colliding with equal and opposite momenta. A parton from each hadron participates in the interaction, producing products c (a parton) and X (anything). This interaction $A + B \rightarrow X + jet$ is shown in Figure 3. c will fragment to produce a jet of hadrons with high *transverse momentum* $p_T{}^\ddagger$.

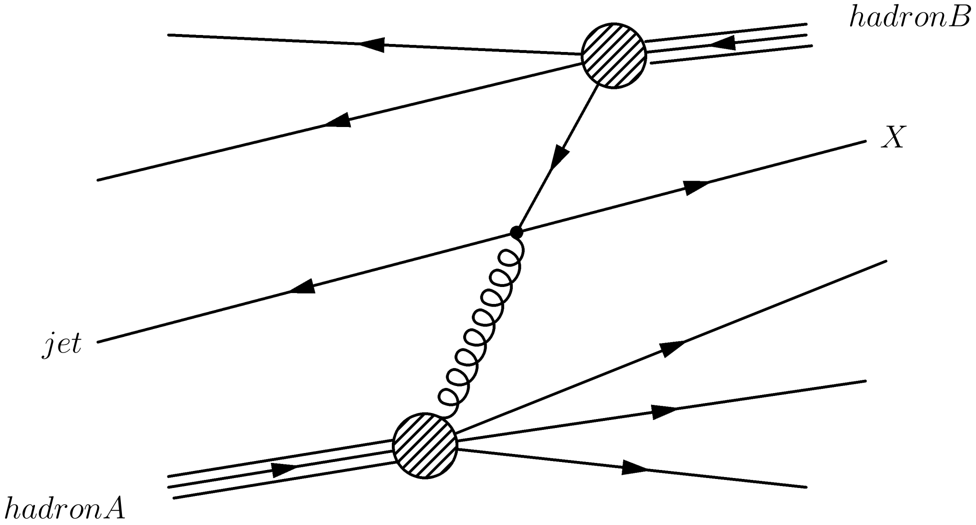


Figure 3: Diagram showing the interaction between two hadrons to produce a jet and another product (X), including spectator partons.

The interaction of partons a and b is known as the *hard process* (forming the physically

---

[†] The "Bjorken x" measures the fraction of the momentum of the incident proton carried by a specified parton, such that $0 < x < 1$

[‡] The transverse momentum of a particle is its momentum resolved into the plane perpendicular to the momentum of the incident colliding hadrons

interesting part of the event), and the partons from the proton which do not interact in the hard process are known as the *spectator* partons. Soft interactions between spectator partons form the *underlying event* which produce many low-p_T particles in the detector.

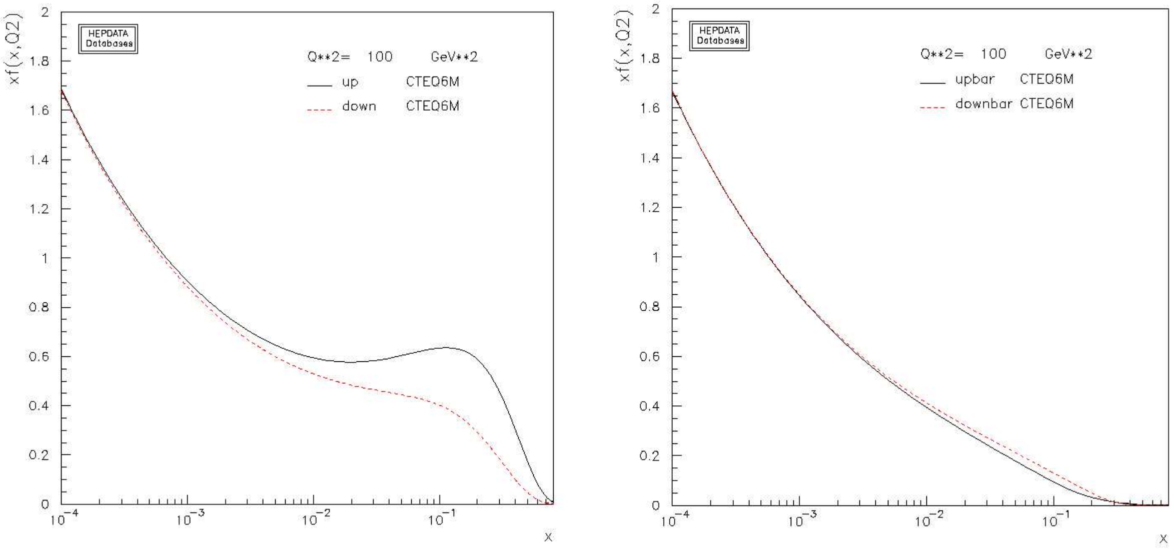
Let the incident parton from hadron A be a parton of type a, and from B be of type b such that the hard process can be described by $a + b \rightarrow c + X$. a and b carry momentum fractions x_a and x_b. Then the probability of parton a being found in hadron A and interacting at a certain value of Q^2 is given by $f_{a/A}(x_a, Q^2)$, and similarly for parton b, where the functions $f_{n/N}(x_n, Q^2)$ are known as the parton number distributions for partons of type n in a hadron of type N. The parton momentum distribution is given by $x_n f_{n/N}(x_n)$, this is what is meant by the "parton distribution function". The cross-section $d\hat{\sigma}$ of the short-scale interaction between partons a and b which produces $X + c$ can be calculated using perturbative QCD in terms of the PDFs of partons of type a and b.

Different experiments allow partons with different ranges in x to be probed, and fits have been performed on the available data from individual experiments which have been combined into global analyses of the likely forms of the parton distribution functions. Fitting of PDFs to data is a highly complex process, the methods and results of which are discussed in detail in Chapter 6 of [16]. Figure 4 shows the PDFs of partons in the proton as given by the CTEQ6m [17] analysis for $Q^2 = 10^4$ (corresponding to the production of particles of mass $\sim$ 100 GeV). The CTEQ6m PDF assumes symmetric distributions for the strange/antistrange, charm/anticharm and bottom/antibottom sea contributions (at $Q = 100$ GeV the top quark component of the proton sea is negligible). The excess of u over d valence quarks can easily be seen in Figure 4(a).

2.2.2 Sea Quark/Antiquark Asymmetries and the Meson Cloud Model

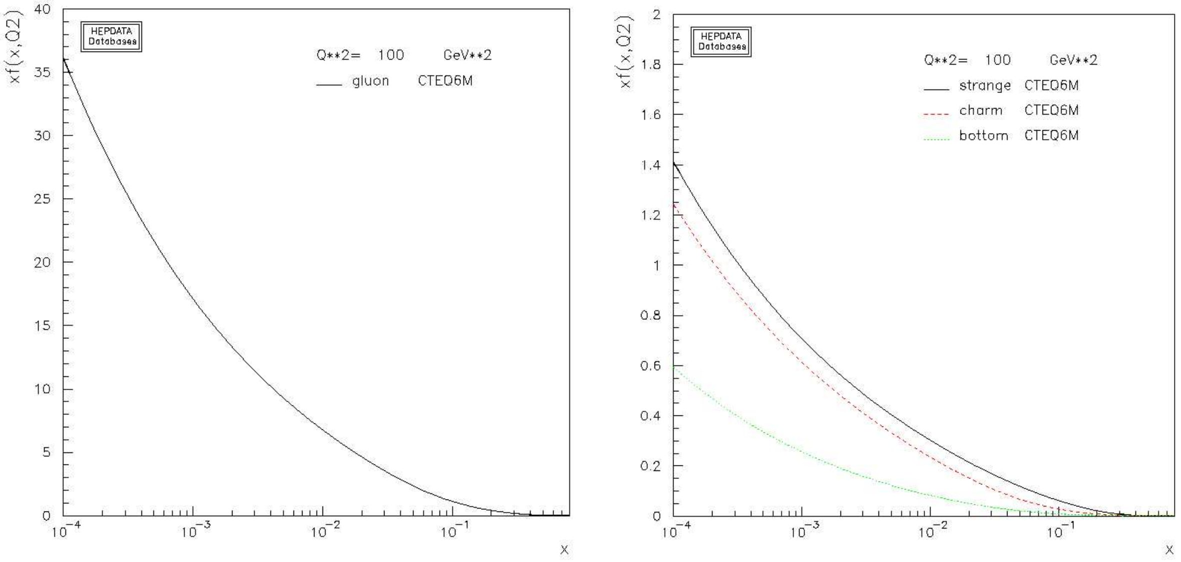
pQCD predicts that the distributions of $\bar{u}$ and $\bar{d}$ quarks in the proton should be almost flavour symmetric. However the New Muon Collaboration (NMC) measured a significant violation of the Gottfried sum rule [§] [18], finding a flavour asymmetry in the light sea quarks $\int_0^1 (\bar{d}(x) - \bar{u}(x))dx$ of around 30% [19], [20]. This observation was subsequently confirmed by the NA51 [21], E866 [22], [23] and HERMES [24] collaborations. This implies that the distribution of $\bar{d}$ significantly dominates that of $\bar{u}$ in the proton (shown in 4(b)). $\bar{d}(x) - \bar{u}(x)$ is also found to have a strong x dependence [25]. Various models have been proposed to explain this flavour asymmetry, which must depend upon non-perturbative processes acting upon the sea components.

[§] The Gottfried Sum Rule is given by $GSR = \int_0^1 \frac{dx}{x}(F_2^{ep}(x) - F_2^{en}(x))$ where $F_2^{ep}(x)$ and $F_2^{en}(x)$ are structure functions calculated from electron-proton or electron-neutron scattering experiments respectively ($F_2(x) = \Sigma e_i^2 x^2 f_i(x)$ where the sum is over hadronic partons i with charge e_i, and $f_i(x)$ are the parton number distributions as before). In the context of the quark-parton model this gives the difference between the numbers of each type of quark in the proton and in the neutron. The quark-parton model assumes that the proton and neutron are equivalent and opposite in terms of u and d quark components, and otherwise identical, such that $u_n(x) = d_p(x)$, $u_p(x) = d_n(x)$, $s_n(x) = s_p(x)$ and so forth, so the sum becomes $GSR = \frac{1}{3}\int_0^1 dx(u_p(x) - d_p(x) + \bar{u}_p(x) - \bar{d}_p(x)) = \frac{1}{3} + \frac{2}{3}\int_0^1 dx(\bar{u}_p(x) - \bar{d}_p(x))$. Since $\bar{u}$ and $\bar{d}$ have very similar masses they are expected within pQCD to have very similar distributions, leading to $GSR \approx \frac{1}{3}$.



(a) PDFs of u (black) and d (red) quarks in the proton

(b) PDFs of $\bar{u}$ (black) and $\bar{d}$ (red) quarks in the proton



(c) PDFs of gluons (black)

(d) PDFs of $s/\bar{s}$ (black), $c/\bar{c}$ (red) and $b/\bar{b}$ (green) quarks in the proton (these quark/antiquark distributions are assumed symmetric).

Figure 4: Parton distribution functions of partons in the proton according to the CTEQ6m PDF set, at $Q^2 = 100$ GeV^2 (note different scales). The plots are taken from [26].

As explained in Section 2.2 hadrons are considered to be made up of three valence quarks in a fluctuating sea of gluons, quarks and antiquarks. As the hadronic momentum scale increases perturbative processes are increasingly expected to occur which generate $q\bar{q}$ pairs with the same probability and momentum distributions as one another. However even at LHC energies a small fraction of the sea quarks may be formed in non-perturbative processes, which could lead to the generation of slightly asymmetric q and $\bar{q}$ sea components.

This is not entirely surprising since application of the Pauli exclusion principle at quark level would also lead to an excess of $\bar{d}$ over $\bar{u}$ (this theory is encompassed in the "Bag Model" [27]). Put simply, if there were the same number of valence us as ds in the proton there would be no asymmetry, but since there are two valence us and one d in the proton the structure of the vacuum inside the proton is different to that outside. To an external probe the change in the structure of the vacuum manifests as an intrinsic nonperturbative sea of $q\bar{q}$ pairs [28], leading to a measured total asymmetry in $\int_0^1 \frac{\bar{d}(x)}{\bar{u}(x)} dx$ of the order $\frac{5}{4}$ [29], [30], [31], although this is insufficient to match the experimental results.

The most significant framework available in which the generation of a non-perturbative, asymmetric nucleon sea is inherent is the Meson Cloud Model (MCM) [32]. This model assumes that the nucleon can fluctuate with a low probability into a bound state consisting of a meson and a baryon. Depending on the hadron in question this leads to different environments in which a sea quark or its equivalent antiquark can exist - clearly should a proton fluctuate into any baryon-meson bound state the antiquark must be confined (albeit temporarily) into the mesonic state and the quark into the baryon. This will potentially lead to different behaviours for different types of sea quarks which will depend on the specific hadronic states produced. It seems likely that the observed asymmetry in $\bar{u}$ and $\bar{d}$ is a combination of both the effects of Pauli exclusion and with MCM predictions.

There are a number of mathematical approaches which have been used to infer the consequences of various MCMs, including the parametrisation of the nucleon-meson-baryon vertex using form factors [33], and the use of two-body wave functions to describe the states [34]. Reference [35] uses a simple approach to studying the predictions of the MCM for the strange components of the sea. It assumes that $q\bar{q}$ pairs are initially produced perturbatively. These then combine with the valence quarks to produce a hadronic bound state which fluctuates into a meson-baryon pair. This approach allows the description of the non-perturbative processes of the MCM in terms of the well-known hadronic scheme of the SM.

Symmetric s and $\bar{s}$ parton distribution functions are often assumed (as was shown in Figure 4), but this is neither established or rejected experimentally, and a non-zero strange sea asymmetry is indicated within the MCM [35], [36].

Figure 5 is taken from [35]. It shows the non-perturbative strange sea distributions for s and $\bar{s}$ and the asymmetry $(s(x) - \bar{s}(x))$ calculated within the MCM. It is calculated at a low value of Q^2, 0.8 GeV^2, a significantly different regime to that at ATLAS (Q^2 $\sim$ 100 GeV^2). At ATLAS, therefore, the number of perturbatively produced (and presumably charge-symmetric) strange and antistrange quarks is expected to greatly dominate over the potentially asymmetric non-perturbatively produced strange and antistrange quarks, but the actual numbers of these non-perturbative quarks may be quite similar to the low Q^2 case. This model predicts a definite asymmetry (remembering that the calculations are based on some assumptions about the form of the $s\bar{s}$ distributions in strange hadrons which are not well known) in which an excess of s quarks is seen at most values of x. However $\bar{s}$ dominates in the low x region, independent of the exact form of the s, $\bar{s}$ distributions used, suggesting that if there is a strange sea asymmetry found at ATLAS it is likely to be negative.

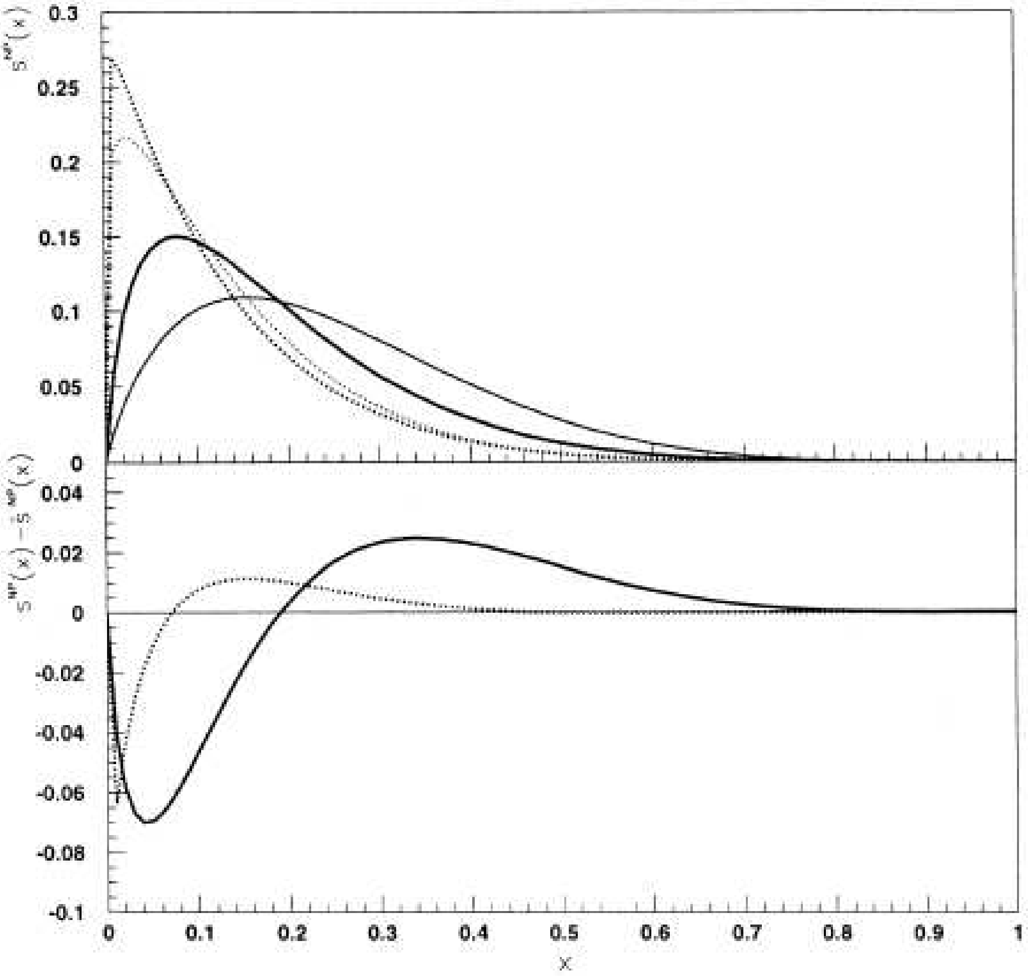


Figure 5: Upper: The non-perturbative strange distributions $xs(x)$ and $x\bar{s}(x)$ in the nucleon sea: two theoretically feasible distributions (solid and dashed) which fit experimental constraints are shown. The thinner lines represent s and the thicker $\bar{s}$. Lower: the asymmetry in the nucleon strange sea as calculated from the two above distributions ($Q^2 = 0.8$ GeV^2). See [35] for details.

Figure 6 shows the results of a global analysis of the strangeness number and momentum asymmetry functions which are consistent with experimental data (2007) gathered from the HERA and fixed-target experiments, shown for $Q^2 = 10$ GeV^2. These were produced by the CTEQ collaboration (as were the distributions shown in Figure 4). They also indicate a likely excess of $\bar{s}$ over s at low x.

Existing experimental constraints on strange PDFs are relatively weak, and in particular below $x \sim 10^{-2}$ there is very little data available on the flavour composition of the sea [2], which can be seen in Figure 6.

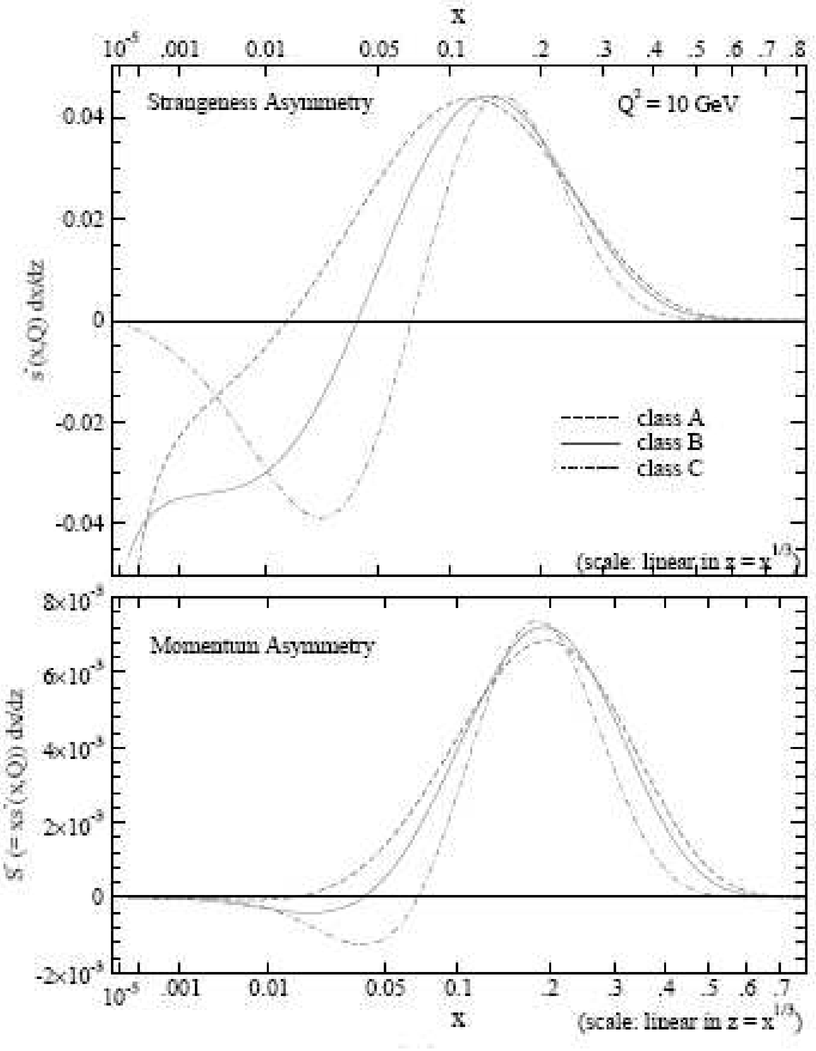
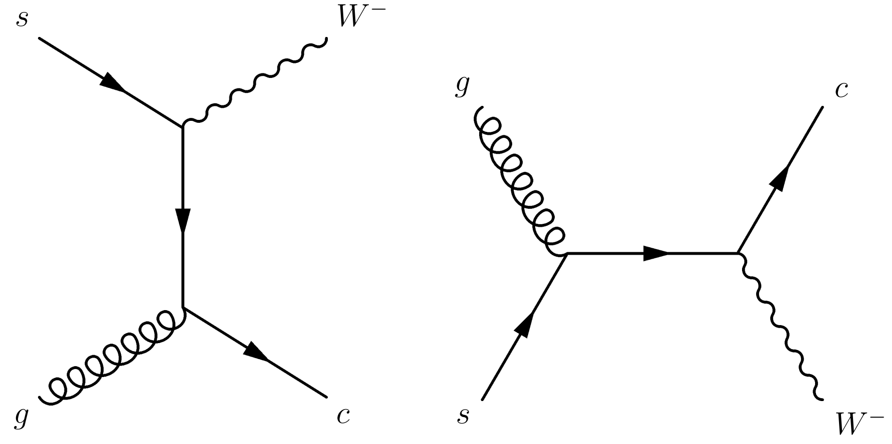


Figure 6: Upper: number asymmetry functions ($s(x) - \bar{s}(x)$) which are consistent with experimental data. Lower: equivalent strangeness momentum asymmetry functions $x(s(x) - \bar{s}(x))$. The solid line is the central fit, the other two are alternative fits. The functions are multiplied by a Jacobian factor dx/dz so that the area under the curve is the corresponding integral over x ($Q^2 = 10$ GeV^2). See [37] for details.

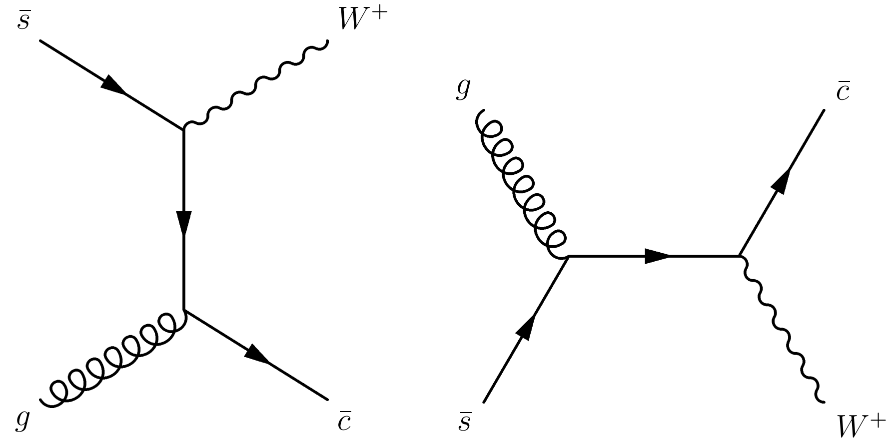
Figure 6 suggests that the number asymmetry of $\bar{s}$ over s could reach a maximum of up to 0.04 at some values of x within the range of W boson production at ATLAS, indicating an excess of $W^+ + D^{*-}$ events over $W^- + D^{*+}$ events of up to around 7% at some values of x.

Detailed study is needed of new experimental data in order to constrain the nucleonic strange PDFs further, the results of which could either exclude or increase the likelihood of a symmetric strange sea distribution. These would enable better prediction of hadron interaction cross-sections in future, and bring an increased understanding of non-perturbative parton interactions.

2.3 The Signal Channel $\mathbf{s} + g \rightarrow W + c$



(a) $s + g \rightarrow c + W^-$: t-channel (b) $s + g \rightarrow c + W^-$: s-channel



(c) $\bar{s} + g \rightarrow \bar{c} + W^+$: t-channel (d) $\bar{s} + g \rightarrow \bar{c} + W^+$: s-channel

Figure 7: t- and s-channel Feynman diagrams for the signal channel. a) and b) are different time orderings for the process $s + g \rightarrow c + W^-$, c) and d) represent $\bar{s} + g \rightarrow \bar{c} + W^+$

Feynman diagrams for the signal channel $s + g \rightarrow c + W$ are shown in Figure 7 for the cases of an incident strange and antistrange quark. Any differences in the properties of the s and $\bar{s}$ will manifest in differences in the c quark and W boson distributions.

The decays of the W boson to either an electron with an electron neutrino ($W \rightarrow e\nu_e$) or a muon with a muon neutrino ($W \rightarrow \mu\nu_\mu$) are highly distinctive processes, accounting for approximately 10% of W boson decays each. These can be identified in the detector by the presence of a single energetic electron or muon (with energies of the order of half the mass of the W boson) and an equivalent amount of *missing energy*, corresponding to the presence of a neutrino. Neutrinos are minimally-interacting, chargeless and nearly massless particles which escape detection in the detector. Their presence is inferred from a momentum balance in the detector, known as missing energy (this will be discussed further in Section 3.5.6). When data-taking begins, on-line hardware and software triggers will identify events with this dual signature as likely to contain a W boson and assign these events to a collection of probable W events for physics analysis, so Monte Carlo samples in which W bosons decay into $e/\mu+\nu$ are a useful basis for a study into the production of W bosons with associated charm. The generation of these samples is discussed in Section 4. W bosons can also decay hadronically, but these events will be more difficult to identify since large numbers of multi-jet events will be produced at ATLAS via many different production mechanisms.

For simplicity, only the electronic decay of the W boson was used in this study. To boost statistics in data, the decay to muons could also be used. The main disadvantages in using electrons are that they are subject to Bremsstrahlung effects (as discussed in section 6.2.2), and that electrons can be produced by the interaction of other particles with the detector as well as in the proton-proton interactions (the effects of which will be considered in section 9.2.9). The reconstruction efficiency and spatial and energy resolutions of electrons will be better than those of muons in the real detector, however, for the energy ranges accessible for leptons produced by the W boson decays.

The decay $W \rightarrow \tau\nu_\tau$ in which $\tau \rightarrow e\nu_e\nu_\tau$ could look very similar to $W \rightarrow e\nu$, which channel will be discussed in Section 8.

The charm quark will be bound into a hadron found within a jet, which must be reconstructed in order to infer the presence of the original charm. By their nature hadrons containing charm are short-lived and so a hadron with an easily identifiable decay chain was chosen, the D^* meson. This meson was chosen to tag the presence of a charm because it is very unstable and can follow a distinctive decay path. In this analysis the channel is chosen in which the D^* decays almost instantaneously to a D^0 meson and a charged pion (henceforth referred to as the *bachelor pion* π_B). The D^0 then decays to a charged kaon K and an oppositely charged pion π. The D^0 has a proper decay length of 123 μm, and factoring in a Lorentz boost from its motion away from the interaction point, the point at which it decays will generally be noticeably displaced from the primary vertex. The D^0 decay point is referred to as the *secondary vertex.* The decay chain is described in Figure 8. Selection of D^*s will be discussed in more detail in Section 7.1.

2.4 Measuring a Strange Sea Asymmetry

In order to measure any asymmetry in the protonic strange sea the numbers of events produced in the interactions $s + g \rightarrow W^- + c$ and $\bar{s} + g \rightarrow W^+ + \bar{c}$ were counted, as discussed in Section 2.3. The theoretical asymmetry measured between these channels will be given by:

$$\mathcal{A} = \frac{\mathcal{N}(s) - \mathcal{N}(\bar{s})}{\mathcal{N}} \tag{9}$$

where $\mathcal{N}(s)$ will be the total number of strange quarks events produced in this interaction, and $\mathcal{N}(\bar{s})$ the number of antistrange. $\mathcal{N}$ is the total number of strange and antistrange events measured:

$$\mathcal{N} = \mathcal{N}(s) + \mathcal{N}(\bar{s}) \tag{10}$$

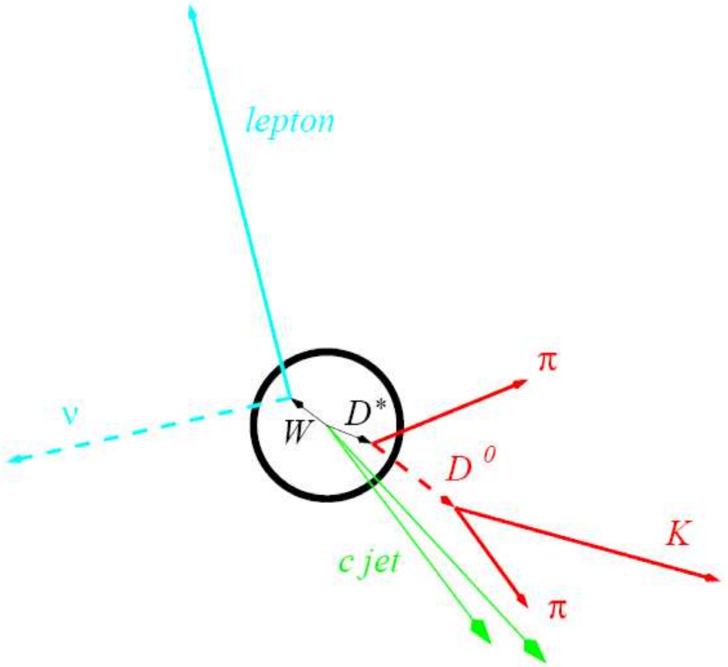


Figure 8: Kinematic view of the signal decay chains used in this analysis. Decays found within the circle will be indistinguishable from instantaneous decays at the primary vertex, those on the outside will usually appear displaced from the primary vertex.

However the actual events counted were those in which the W boson decayed to $e\nu$ and the charm was bound into a D^*, which then decayed into a charged pion and a D^0 meson. The D^0 decayed further into a charged kaon and pion. The particles also must fall into the relevant acceptance regions of the detector in order to be counted, and there will be an additional efficiency factor which gives the probability of a particle being successfully reconstructed provided that it falls in the acceptance region. Therefore the number of the strange quark events that will be measured is given by:

$$\begin{aligned}\mathcal{N}(s) = \mathcal{L}\sigma_{(sg\to W^- c)} f(c\to D^{*+}) B(W^- \to e^-\bar{\nu})\epsilon_{e^-\bar{\nu}}A_{e^-\bar{\nu}} B(D^{*+}\to D^0\pi^+_B)\epsilon_{\pi^+_B}A_{\pi^+_B} \\ \times B(D^0\to K^-\pi^+)\epsilon_{K^-\pi^+}A_{K^-\pi^+} \quad (11)\end{aligned}$$

where $\mathcal{L}$ is the luminosity of the experiment; $f(c\to D^{*+})$ is the hadronisation fraction of the charm into a D^{*+}; B is the branching fraction of the decay indicated; ϵ is the efficiency of reconstruction of the type of particle(s) indicated, and A the acceptance. Similarly the number of antistrange quark events measured is given by:

$$\begin{aligned}\mathcal{N}(\bar{s}) = \mathcal{L}\sigma_{(\bar{s}g\to W^+\bar{c})}f(\bar{c}\to D^{*-})B(W^+\to e^+\nu)\epsilon_{e^+\nu}A_{e^+\nu}B(D^{*-}\to \bar{D}^0\pi_B^-)\epsilon_{\pi_B^-}A_{\pi_B^-}\\ \times B(\bar{D}^0\to K^+\pi^-)\epsilon_{K^+\pi^-}A_{K^+\pi^-}\end{aligned} \tag{12}$$

Assuming that the branching and hadronisation fractions are charge symmetric (eg. $B(D^0 \to K^-\pi^+) = B(\bar{D}^0 \to K^+\pi^-)$ and so forth), as they are both measured and predicted to be, and defining the product of the efficiency and acceptance of each set of particles as $\beta = \epsilon A$ this leads to an asymmetry of the form:

$$\mathcal{A} = \frac{[\sigma_{(sg\to W^-c)}\beta_{e^-\bar{\nu}}\beta_{\pi_B^+}\beta_{K^-\pi^+} - \sigma_{(\bar{s}g\to W^+\bar{c})}\beta_{e^+\nu}\beta_{\pi_B^-}\beta_{K^+\pi^-}]}{[\sigma_{(sg\to W^-c)}\beta_{e^-\bar{\nu}}\beta_{\pi_B^+}\beta_{K^-\pi^+} + \sigma_{(\bar{s}g\to W^+\bar{c})}\beta_{e^+\nu}\beta_{\pi_B^-}\beta_{K^+\pi^-}]} \tag{13}$$

The measurement efficiencies and acceptances for the different types of particle relate directly to the numbers of associated W bosons and D^* mesons measured, such that

$$\sigma_{(sg\to W^-c)}\beta_{e^-\bar{\nu}}\beta_{\pi_B^+}\beta_{K^-\pi^+} \propto N(W^-D^{*+}) \tag{14}$$

and

$$\sigma_{(\bar{s}g\to W^+\bar{c})}\beta_{e^+\nu}\beta_{\pi_B^-}\beta_{K^+\pi^-} \propto N(W^+D^{*-}) \tag{15}$$

where $N(W^-D^{*+})$ is the number of measured events in which an associated W^- and D^{*+} meson are found together, and similarly for $N(W^+D^{*-})$. The measured asymmetry then becomes:

$$\mathcal{A} = \frac{N(W^-D^{*+}) - N(W^+D^{*-})}{N(W^-D^{*+}) + N(W^+D^{*-})} \tag{16}$$

More details of this measurement can be found in Section 7.3.

This calculation assumes that there is no background to this signal, whereas in reality this cannot be true. There are other channels which produce a W boson and D^* meson of opposite signs (or can appear to do so), some of which will not be charge-symmetric. The effects of these can be estimated in Monte Carlo, as will be discussed in Section 9.2.1. The majority of backgrounds in data, however, will result from random coincidences of W bosons and D^* mesons, either true or falsely reconstructed. These backgrounds will be charge-symmetric, and thus the same number of them will be found in the WD^* pairs with the correct sign correlation for signal reconstruction, and also the other "wrong" sign correlation (in which the W and D^* have the same sign). Taking this into account, the numbers of pairs of WD^*s with the wrong sign correlation can be used as an estimate of the level of background to the signal. The effects of the asymmetric backgrounds should then be incorporated as an uncertainty on the background level.

As a result, a better measure of the asymmetry is given by

$$\mathcal{A} = \frac{[N(W^-D^{*+}) - N(W^-D^{*-})] - [N(W^+D^{*-}) - N(W^+D^{*+})]}{[N(W^-D^{*+}) - N(W^-D^{*-})] + [N(W^+D^{*-}) - N(W^+D^{*+})]} \tag{17}$$

in which $N(x, y)$ is the number of measured events in the signal peak from either the correctly or wrongly sign-correlated WD^* events as indicated.

It should be stressed that this approach will be useful in data, but it will not be used here, since the Monte Carlo dataset used in signal reconstruction (Section 4) contains much of the asymmetric background. In this dataset there is very little chance of the production of random, uncorrelated pairs of Ws and D^*s.

3 The ATLAS Experiment

ATLAS is a particle physics experiment based at the Large Hadron Collider at CERN, which is due to start running in 2009. It will search for new physics in proton-proton collisions of higher energies than produced in any previous particle physics experiment, and also seeks to increase understanding of the interactions and properties of known particles.

ATLAS is one of the largest collaborative efforts ever attempted in the physical sciences. There are around 2100 scientists and engineers participating from more than 165 institutions in 35 countries.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a ring accelerator (a synchrocyclotron), 27 km in circumference, built approximately 100 m below ground at the European Organization for Nuclear Research (CERN) near Geneva on the French-Swiss border.

Narrow beams of very high energy protons are produced in a series of smaller accelerators, culminating in the Super Proton Synchotron (SPS), and these are then injected into the LHC. The collider consists of two interleaved synchotron rings enclosed within superconducting dipole magnets operating at up to 9 Tesla which are cooled to 1.9K by superfluid helium. The LHC contains two beams of protons travelling in opposite directions around the two rings, and additional quadrupole magnets are used to focus the beams. The proton beams will be allowed to collide at four separate points along the ring. The protons are grouped into bunches of particles travelling together, rather than moving as continuous beams. There will be around 2800 bunches of $\sim 10^{11}$ protons per bunch. These will be allowed to collide every 25 ns for data-taking.

Around the collision points the detectors for the four major LHC experiments are being built: ATLAS, CMS, LHCb and ALICE. The first two of these are multi-purpose detectors, LHCb is dedicated to B-physics and ALICE is optimised to study collisions between heavy ions in order to study the properties and behaviour of the quark-gluon plasma that will be formed momentarily at the collision point.

Figure 9 shows the overall layout of the LHC and the experiments positioned around it.

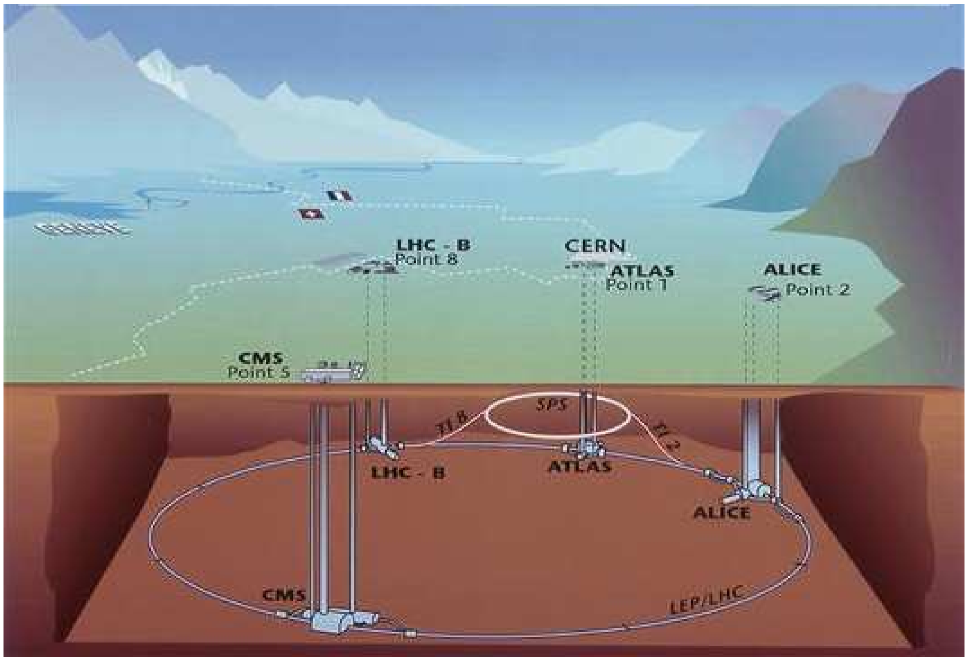


Figure 9: Schematic view of the LHC, showing the LHC and SPS rings, the positions of the detectors of the major experiments and the CERN facility. The Swiss-French border is marked. [38]

The design luminosity of the LHC is 10^{34} $\text{cm}^{-2}\text{s}^{-1}$. This equates to an integrated luminosity $\mathcal{L}$ of 100 fb^{-1} per year. At full design luminosity ("high luminosity") there will be an average of 23 proton-proton interactions per crossing. The centre-of-mass energy of the proton-proton collisions is expected to reach 14 TeV, producing energy densities similar to those present in the universe less than a billionth of a second after the Big Bang. For an initial period of around three years the LHC will run at "low luminosity", with an instantaneous luminosity of 10^{33} cm^{-2}s^{-1} and integrated luminosity 10 fb^{-1} per year.

3.2 The Atlas Detector

ATLAS (*A Toroidal LHC ApparatuS*) is a general-purpose detector currently in the final assembly phase at the LHC. It has been optimised to measure the broadest possible range of particles that may be produced in the very high energy proton-proton collisions at the LHC, and is thus larger and more complex than any detector built before. The ATLAS detector is 46 metres long and 25 metres in diameter, weighing around 7000 tonnes. It is cylindrically symmetric, consisting of a central barrel region and end-caps. Beams of protons enter the detector from each side, travelling down the central beam pipe, and collide at its centre.

The ATLAS co-ordinate system is defined such that the z axis is along the beam pipe, and corresponds with the direction of motion of the incoming protons. The x, y plane is normal to the z axis, and ϕ is the azimuthal angle (measured from the horizontal x axis) in this plane. θ measures the polar angle from the collision point on the z axis. Pseudorapidity, η [¶], is often used to define the position of detector components.

[¶] Pseudorapidity η is defined as $-ln[tan\frac{\theta}{2}]$. It is used to define the position of detector components (which are essentially cylindrically symmetric), where θ is the angle between the beam pipe and detector component measured at the collision point. Pseudorapidity can also be calculated for the path of a particle moving away from the collision in the detector, where θ is the angle between the particle momentum vector and the beam line, and defines the region of the detector that a particle can be detected in.
For comparison, rapidity y is an intrinsic property of a particle, defined as $\frac{1}{2}ln\frac{E+p_L}{E-p_L}$ where E is the energy of the particle and p_L its longitudinal momentum. In the massless limit this can be reduced to the same definition as that for pseudorapidity, i.e. $\eta = y$ for massless particles.
Pseudorapidity is a useful variable in hadron colliders because, in the massless particle approximation the difference in pseudorapidity between two particles, $\Delta\eta$, is invariant under a Lorentz boost of the centre of mass of the system along the z axis. In a hadron-hadron collider the momenta of the incident hadrons is known, but it is the partons that make up the hadronic sub-structure that interact, and they carry variable amounts of the hadrons momentum. As a result the exact momentum of the interacting particles in the z direction is unknown for any given collision.

Figure 10 details the basic layout of the detector. The Inner Detector, consisting of the Semiconductor Pixel Detector, the Silicon Tracker and the Transition Radiation Tracker, sits in the central region within a solenoid coil which immerses it in a 2 Tesla axial magnetic field. This provides detailed tracking and vertexing information as well as momentum resolution for charged particle tracks up to $|\eta| = 2.5$.

Beyond the coil sit the calorimeters, which extend out to $|\eta| \approx 5$. They are divided into Electromagnetic and Hadronic Calorimetry. The goal of the calorimetry is to stop, measure the energy of, and identify electrons and photons (mainly the electromagnetic calorimeters) and jets (mainly hadronic). The majority of the electromagnetic calorimeters are liquid argon (LAr) calorimeters, in which layers of lead or stainless steel are interspersed with narrow layers of liquid argon. Interactions of particles with the lead cause the creation of secondary electrons (see Section 3.5.2 later), which in turn cause ionisation in the liquid argon. This enables a current to flow between two copper electrodes in the argon, producing a signal in proportion to the amount of ionisation. The calorimetry has fine granularity, and is cylindrically symmetric, which enables it to provide good missing transverse energy resolution. Light electromagnetically interacting particles (mainly electrons and photons) will be contained entirely in the Electromagnetic Calorimeter, whereas hadrons will lose some of their energy there but will continue on to the Hadronic Calorimeter where most of the rest of it will be absorbed. The only Standard Model particles that should escape the calorimetry are neutrinos and high-energy muons.

The final layer of the detector consists of several layers of Muon Spectrometers. These surround the hadronic calorimeter and provide high precision measurements of the momenta of the muons which escape the calorimetry.

The individual detector components are discussed in more detail in the following sections.

The detector specifications are found in [39].

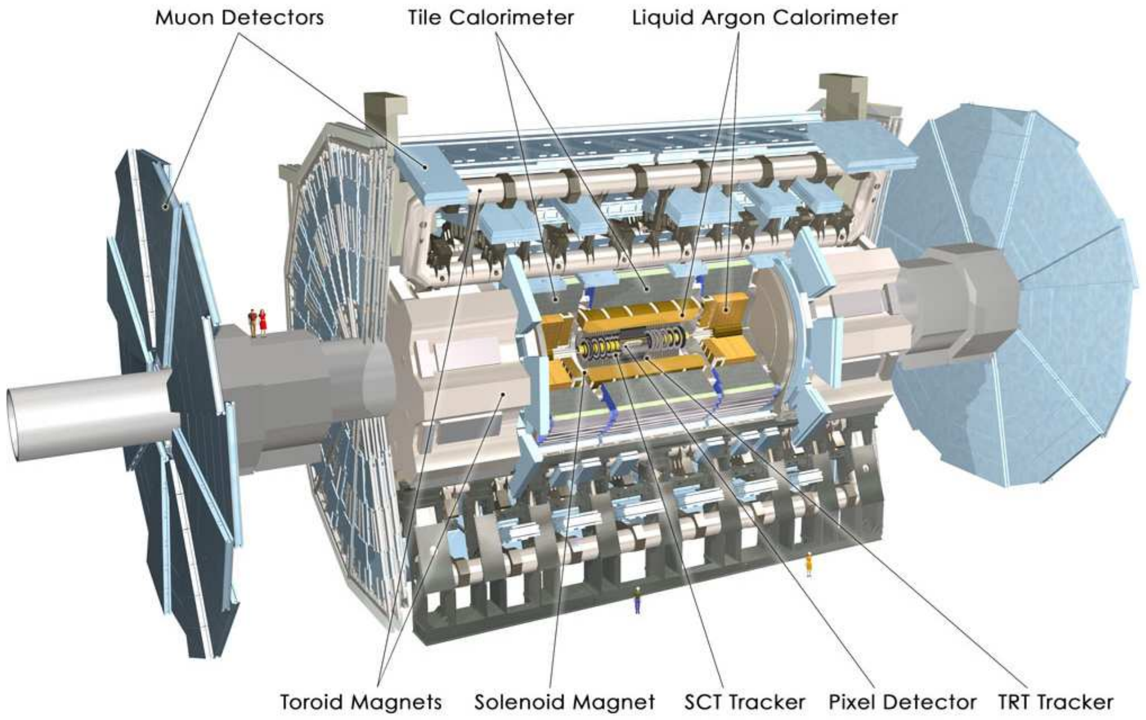


Figure 10: The layout of the ATLAS detector [38]

3.2.1 Inner Detector

The types of physics process that are expected to occur at the LHC will result in large numbers of charged particles with very low momenta (∼1000 per bunch crossing *event*). In order to separate and identify these, very high precision measurements are needed with fine-granularity tracking detectors.

The Inner Detector (ID) is contained within a cylinder of length 7 m and radius 1.15 m and sits inside the central solenoid. It sits around the interaction point and is designed to be able to measure the transverse momenta of charged particles accurately. It enables the accurate reconstruction of charged particle tracks, which allows reconstruction of the position of the primary interaction point (the *primary vertex*) and any *secondary vertices* which arise from particle decays. It consists of three layers of subdetectors: a silicon pixel detector is the innermost, followed by a silicon strip detector (the SemiConductor Tracker, SCT) and then a transition radiation detector (the Transition Radiation Tracker, TRT). Figure 11 shows the layout of the ID.

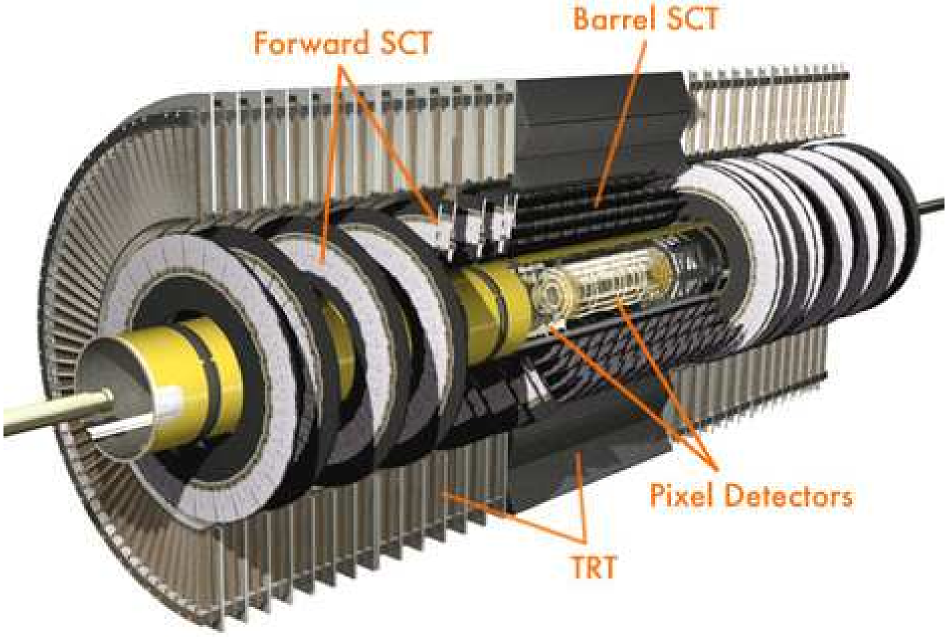


Figure 11: View of the ATLAS Inner Detector

The individual pixel detectors of the pixel detector are arranged in concentric cylinders around the interaction point and are of very fine granularity (50 × 400 μm) to provide accurate position measurements for the paths of the charged particles. The pixel detector consists of a barrel region and endcaps. The barrel region has of three layers, thus contributing in general three space points to the measurement of a charged particle path, the first of which is at a 4 cm radius from the interaction point (it is not expected to have a long operating life due to the intensity of radiation incident on it). There are two further cylinders at 10 and 13 cm. There are five endcap disks on each end. The system is tiled with identical silicon pixel modules which are 63.4 mm long and 21.4 mm wide, and can achieve a position resolution of 12 μm in $[r, \phi]$ space and around 70 μm in z.

The pixel detector is followed by the layers of the SemiConductor Tracker, which provides precision measurements of momentum. It consists of four concentric cylinders in the barrel region (thus generally contributing around four space points per particle to the track) ranging from radii of 30 - 52 cm from the interaction point, and nine disks in the forward region on each side, covered with silicon strip modules. The barrel region covers the pseudorapidity range $|\eta| < 2.5$, and this is used for precision tracking. The size of the silicon modules is 6.36 × 12.8 cm, and the spatial resolution of the system is around 16 μm in $[r, \phi]$ space and 580 μm in z.

The Transition Radiation Tracker (TRT) [40] is found at the outer radii of the ID, between 62 and 108 cm. The TRT is a "straw tube" tracker, consisting of 370 000 cylindrical drift tubes operating at a high rate, and it provides typically about 36 tracking points per charged particle. Its purposes are both to provide continuous tracking and to ease discrimination of electrons from hadrons (in particular pions) using transition radiation[‖]. The "straw tubes" are 4 mm diameter drift tubes filled with a Xenon gas mixture (designed to maximise the transition radiation) which run in the z direction in the barrel ($|\eta| < 0.9$) and radially in the end caps (18 disks on each side, extending the pseudorapidity coverage to $|\eta| < 2.5$).

The relative precisions of the various inner detector measurements are well matched, so that no single measurement dominates the momentum resolution.

The ID measurements provided are combined into *tracks*, in which the most likely momenta and paths of particles are interpolated from the set of individual space points.

[‖] Transition radiation is produced by the motion of a relativistic charged particle across the boundary between media with different dielectric constants. The particle radiates soft X-rays with energies that increase rapidly with the Lorentz γ factor of the incident particle (γ increases as particle mass decreases). Electrons will therefore produce far more transition radiation than hadrons: for electron identification efficiency of 90% in the TRT the pion contamination is expected to be $< \sim 1\%$

3.2.2 The Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECal) [41] is a lead/liquid-argon detector with accordion geometry. It is divided into a barrel and two end caps and preceded in $|\eta| < 1.8$ by a "presampler", a detector which is used to correct for the energy lost in the material upstream of the calorimeter (mainly the solenoid magnet). This is a 11mm thick active layer of liquid argon, in the barrel region only. It provides a sampling of any pre-existing showering in front of the ECal. It measures only the energy of pre-showers and has granularity $\eta \times \phi = 0.025 \times 0.1$.

The barrel region, covering approximately $|\eta| < 1.5$ consists of three sampling layers with $\eta \times \phi$ granularities of 0.025×0.025, 0.003×0.1 and 0.05×0.025 (from the innermost out). Three layers of end caps cover the $1.4 < |\eta| < 3.2$ region with very similar granularities to the three barrel layers, but there are "cracks" (areas of reduced coverage) at the boundaries between the barrel and endcaps, around $1.4 < |\eta| < 1.6$, where there is dead material in front of the calorimeters which absorb particle energy. This decreases the energy resolution in these regions. The ECal is symmetric in ϕ (there are no azimuthal cracks). The positioning of the ECal relative to the inner detector is shown in Figure 12.

The primary purpose of the Electromagnetic Calorimetry systems is the detection and measurement of the energy of electrons and photons. It has a fast response time and good containment so that it can be used for triggering, and has fine granularity and good energy resolution for accurate energy measurements. It provides measurements up to $|\eta| < 3.2$, but its precision region is limited to $|\eta| < 2.5$ to correspond with the region of ID coverage.

The energy resolution of the ECal is expected to vary with particle energy, and also with rapidity y (due to geometric effects). The energy dependence of the energy resolution in the Electromagnetic Calorimeter is expected to include components which take the

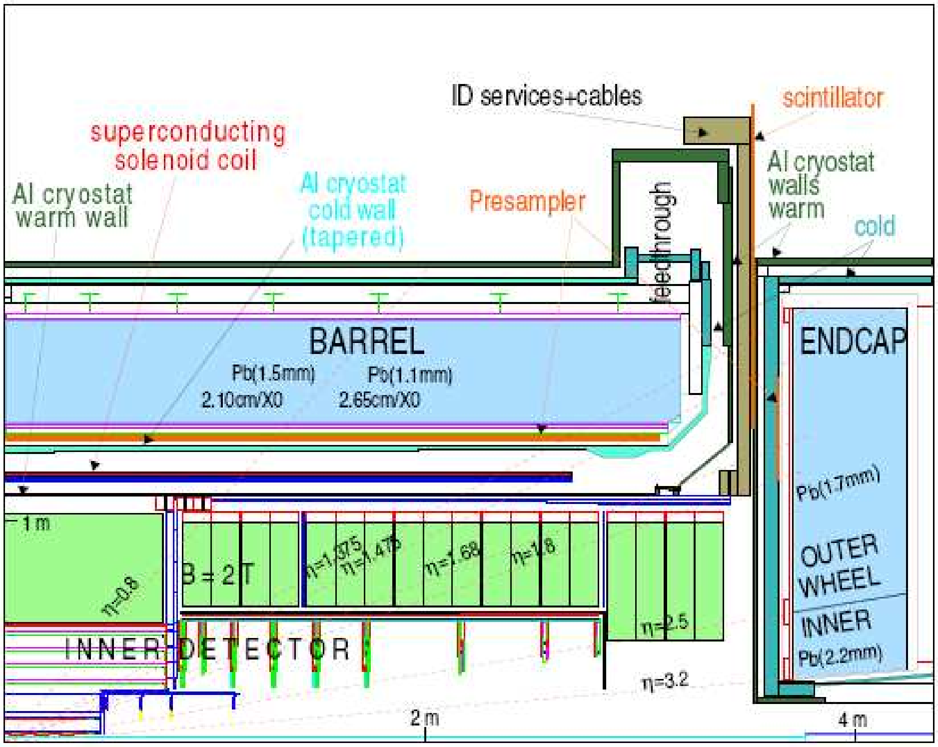


Figure 12: Lateral view of the Inner Detector and the Electromagnetic Calorimeter of the ATLAS detector

following forms:

- $\frac{\sigma}{E}_{fluct} \propto \frac{a}{\sqrt{E}}$: Intrinsic energy resolution of an ideal calorimeter, due to inherently statistical processes, eg. the physical development of the shower and its natural fluctuations, and to instrumental and calibration limits. This term usually dominates.
- $\frac{\sigma}{E}_{syst} \propto$ b: Constant term, due to instrumental effects from detector non-uniformities, non linearities and calibration errors (not energy dependent).
- $\frac{\sigma}{E}_{instr} \propto \frac{c}{E}$: Noise term, due to readout chain electronic noise. This depends on the readout circuit, the detector is designed to minimise this term.

To meet the LHC physics requirements for accurate Higgs mass reconstruction and Z' searches the resolution is designed to have a sampling term $a < 10~\%\sqrt{GeV}$ and constant term $b < 1\%$ ([42]). c is to be be kept as small as possible. These sum in quadrature to give a total fractional resolution $\frac{\sigma(E)}{E}$ which varies with particle energy E as

$$\frac{\sigma(E)}{E} = \frac{a^2}{E} \oplus b^2 \oplus \frac{c^2}{E^2} \tag{18}$$

For example, the transverse energy resolution of a jet with $p_T sim$ 10 GeV is around 5%, and of a 100 GeV jet is around 1-2%.

3.2.3 The Hadronic Calorimeter

The main goals of the ATLAS hadronic calorimetry are to reconstruct hadronic jets, to measure the total *missing transverse energy* (MET) (see Section 3.5.6) and to enhance the particle identification capability provided by the electromagnetic calorimetry. In order to do this accurately the hadronic calorimetry needs to cover a wide rapidity range and to have good granularity.

There are a range of different types of Hadronic Calorimeters in different detector regions, covering a pseudorapidity range of $|\eta| < 4.9$ in total. It consists of three main devices. In the barrel region $|\eta| < 1.7$ is the Tile Calorimeter, made of plastic scintillator tiles sandwiched between layers of iron. It is followed by the Hadronic Endcap Calorimeter, a liquid argon calorimeter, up to $|\eta| = 3.2$. There is then the high density Forward Calorimeter, covering the pseudorapidity range $3.1 < |\eta| < 4.9$. For $|\eta| < 2.5$ the granularity of the Hadronic Calorimetry is around 0.1×0.1 in $\eta \times \phi$. Beyond this range it ingreases to 0.2 × 0.2. For more details of the hadronic and electromagnetic calorimetry see [43].

The energy resolution of the hadronic calorimeters follows a similar form to that of the electromagnetic, but is generally worse. It is around 10% for $p_T = 50$ GeV jets, and drops to $\sim 6\%$ around 400 GeV.

3.2.4 The Muon Spectrometer

The Muon Spectrometer defines the overall dimensions of the ATLAS detector, surrounding the calorimetry, and it accounts for most of the volume of the detector. It is made up of drift tubes and is immersed in an air-core toroidal magnetic field of $\sim$4 T. It contains three stations of high precision tracking chambers and trigger chambers with very fast response time, extending out to $|\eta| < 3.2$ overall. The Muon Spectrometer can perform independent measurements of muon momenta or can be used in combination with the ID. It is necessary to measure muon p_T to high accuracy since muons are promising clean signatures of many new physics signals. The muon spectrometer measures muon transverse momenta with a resolution of 2-3% for muons with $p_T \sim 100$ GeV, and increases to around 10% for 1 TeV muons. See [44] for more details.

3.3 The ATLAS Software

All of the software used in ATLAS is encompassed in a common framework, ATHENA [45], which allows the many different applications which are used to communicate with one another. Athena is derived from the GAUDI framework [46] which was developed for the LHCb experiment. It manages the order in which algorithms are run, chaining them together, and provides access to common services such as message logging, the writing of histograms and ntuples and data access. The POOL persistency scheme is used to write data to disk, and data can also be stored in a common memory store called the transition event store. The framework allows later algorithms in the chain to read from either of these places.

The main components of the ATLAS software are:

- **The production of simulated events.** The stages involved in this are the generation of physically realistic simulated events from the proton-proton collisions, simulation of the effects of the physical detector on the particles as they pass through it, and finally a digitisation stage which mimics the physical and electonic response of the detector components to the passage of the charged particles, producing an output of the same form as data from the real detector. This process is discussed in detail in Section 3.4.

- **Reconstruction of particles**, either from physical data from the detector or simulated data. A number of algorithms are included which perform pattern recognition, track fitting, vertex determination, energy measurement and corrections for particle energy loss and so forth. The result is a set of reconstructed objects (particle tracks, jets, and *clusters* - areas of energy deposited in the calorimeter when a particle strikes it), many of which are grouped into containers according to the types of particles they are likely to be associated with, with appropriate measures of likelihood. *Truth* information can also be made available for comparison of the reconstructed with simulated data, which represents how the data would look if reconstruction was perfect. Reconstruction will be discussed in Section 3.5.

- **Physics analysis.** Tools for the analysis of the reconstructed data are included in ATHENA, and users can add their own analysis code.

- **Event Display** programs are available for visualisation of events. Currently these are HEPVIS [47], which provides a highly detailed three-dimensional viewer, and Atlantis [48] which shows a two-dimensional projection. They produce a graphical representation of the detector geometry, the event data (in the form of *hits* at the points where a particle interacts with the detector), and the reconstructed paths of the charged particle. They can also display the detector truth information, which lists where a particle would have been seen to hit the detector if it were perfectly reconstructed.

3.4 Event Generation, Simulation and Digitisation

Realistic simulations of physics processes in particle interactions and the response of the detector to the passage of particles are crucial to the success of particle physics experiments. Simulations covering the entire kinematic range reachable by the experiment are used to improve understanding of the detector response and to enable optimisation of the detector performance before data-taking starts. Simulations of specific physics processes of interest are used to develop analysis techniques in advance of data-taking, and samples including known background processes to potential new theories will be needed for comparison with real data to test whether new physics has been discovered.

The production of simulated datasets for physics analysis in ATLAS is performed in three stages: event generation, detector simulation and digitisation.

3.4.1 Event Generation

The event generation phase models the proton-proton interaction using Monte Carlo principles. Only the primary interaction, the initial products of the proton-proton collisions and their subsequent decay and hadronisation are simulated, according to physical models inspired by QCD and QED [49]. Appropriate parton distribution functions must be passed into the event generator to enable it to model the effects of proton structure on interactions between the partons in the incident protons.

Event generators work in three stages:

1. Hard Scattering: The interaction between the two incoming partons is modelled. This can be calculated using perturbative QCD.

2. Parton Showering: The partons produced in the interaction split to form showers of more partons. These can also be calculated in perturbative QCD, but the computations required for exact results are vast, so approximations are made in order to perform the event generation on a sensible time-scale.

3. Hadronisation: Models the formation of hadrons from outgoing partons. This includes the description of the underlying event (the result of soft interactions between the beam remnants).

The result of event generation is a sampling of all of the possible final states for specific types of process (the physical accuracy of which depends upon the physical models built into the event generator), and the output is a set of 4-vectors describing the stable final state particles (those with lifetimes $> \sim 10^{-8}$s). This information is then passed into a detector simulation.

There are a number of event generation programs available in ATLAS. These include:

- HERWIG ([50]) is a general-purpose event generator, primarily calculating physics processes to leading order only (i.e. the simplest production mechanisms, $2 \rightarrow 2$ or $2 \rightarrow 3$ body interactions at most). As a result, the matrix elements diverge for the production of particles with low momenta, and it is unreliable in the very soft-p_T region. Estimates are included for the effects of higher order diagrams on cross-sections, but not for the differences in outgoing particle kinematics, so some areas of phase space that should be populated if all relevant production mechanisms were included will be empty.

  HERWIG uses a cluster model for hadronisation. After parton showering, all of the outgoing gluons are split into light quark-antiquark pairs (or sometimes two of each). The quarks are then combined with their neighbours to form colour-singlet clusters, which are fragmented into hadrons (the clusters are split into different numbers of hadrons depending on their mass and properties). The spectator partons are also included in clusters and treated in a similar way, providing the underlying event.

In ATLAS, HERWIG is accessed the JIMMY interface [51], which provides a library of routines that enable generation of multiple parton scattering events when linked to HERWIG. HERWIG itself then deals with the hadronisation process. Several other basic event generators can be linked in to use the HERWIG hadronisation routines.

- Pythia ([52]) is another general-purpose event generator and shares many features with HERWIG. Again, it performs calculations to the lowest order, but includes regularisation to remove the singularities that HERWIG suffers from and thus describes low-p_T data more accurately. The most significant difference between Pythia and HERWIG is the hadronisation model used: Pythia uses the Lund (string fragmentation) model [53]. The Lund model is based on the assumption that confinement is linear, i.e. that the force between a quark and antiquark increases in proportion to the separation between them. As two quarks move apart after parton showering a uniform tube of colour is structed uniformly between them, giving rise to a picture of confinement with a potential that rises linearly. As the energy stored increases the tube may break, producing a quark-antiquark pair. Each will attach to one of the loose ends of the broken tube, leaving two pairs each with a colour tube stretched between them and the process can repeat. Whilst the invariant mass of a string remains high enough to produce more pairs further breakings will, eventually leaving only on-mass-shell hadrons.

  Pythia allows the user to chose between combinations of particle decay modes for each type of particle and so is useful in generating large numbers of rare events.

- MC@NLO ([54], [55]) can be used to generate events for several important processes including the production of W, Z and Higgs bosons and top quarks up to next-to-leading order (NLO). It deals with particle interactions up to first order in α_s and can thus produce events of the type $X + jet$ where X is the particle of interest. It provides phase space and kinematic information on the initial interaction and decay products and is interfaced to HERWIG, which provides the hadronisation. Because it uses true NLO calculations of matrix elements, it is able to provide a more realistic view of phase space than leading order generators, and so should be able to accurately reproduce the kinematics of high-p_T jets in $X + jet$ production. (In leading order generators extra jets can only come from showering the incoming particles, meaning they are usually unphysically soft and not in the correct kinematic region for the study of jets coming from the hard process.)

  Because of the models used in MC@NLO it is possible for two kinematically identical events to be produced, one from gluon emission in the hard scattering phase, and the other when lower order production is followed by a parton shower that produces a gluon, which can result in the generation of essentially the same Feynman diagram. In order to avoid this double-counting, all observables are calculated as a sum of a positive and negative term. As a result, some events are generated in MC@NLO with negative weights, meaning that their contribution should be subtracted from rather than added to the total cross-section and phase space in order to produce a physical cross-section for the process in question.

- CompHep ([56], [57], [58]) is a standalone program that can be used as a full event generator but is most useful for quickly calculating leading order cross-sections and distributions of particle scattering in various QFT models. It can take into account all QCD and Electroweak diagrams, the masses of fermions and bosons, and the widths on unstable particles.

It is especially useful in the estimation of the rates of background processes to physics signals.

- ALPGEN ([59]) is used to approximate the kinematics of events in which several hard, well separated jets are expected, originating from from hard QCD processes or from the decay of heavy particles such as gauge bosons or heavy quarks. These types of events are expected to be common in the LHC. The underlying events are generated in leading order 2→2 processes, and additional high-p_T jets are produced during the shower evolution. It therefore produces processes which appear higher-order but is not truly a higher-order generator. ALPGEN can be interfaced with either Pythia or Herwig, which perform the necessary hadronisation.

3.4.2 Full Detector Simulation (GEANT4)

Once in the detector, the particles that were originally produced in the collision will alter. Some with short lifetimes decay into other particles, which might then decay further. Most will interact physically with the matter that makes up the detector, both the active detecting elements and the support structures. The job of the detector simulation code is to describe this process as closely as possible. Parameters such as the calibration and alignment can also be varied in the simulation to allow the study of how these systematic effects will alter the results of an analysis.

ATLAS uses the highly detailed GEANT4 [60] detector simulation program, which models detector components to the microscopic scale. It simulates all of the important physics processes in the real detector including multiple scattering, continuous energy loss, Bremsstrahlung of electrons, conversions of photons and nuclear interactions with hadrons. The GEANT4 detector simulation requires many billions of calculations to be performed which severely limits the number of complete events that can be produced on a reasonable timescale.

In a separate digitisation phase, GEANT4 models the electronic response of each part of the detector, including the effect of digitisation of data and the simulation of noisy and dead readout channels. This follows the main GEANT4 simulation in a separate step and is usually fast, producing data which is of a very similar form to that expected from the real readout electronics. Overall GEANT4 produces of the rough order of one to ten events per hour on a standard worker node, depending on the complexity of the generated event and computer speed.

After digitisation, reconstruction software is applied to the fully simulated events (as it would be to real data) in order to extract physical information about the event.

3.4.3 Fast Detector Simulation (ATLFAST)

The ATLAS Fast detector simulation, ATLFAST [61], parameterises the response of the detector in order to simulate detector events quickly and simply. It takes into account the only basic geometry of the detector: rather than using GeoModel it uses a brief list of the basic properties of the vital detector components, for example the $|\eta|$ coverage of the ID precision physics regions, the size of the barrel and end-cap calorimeters and the granularity of the hadronic calorimeter. Rather than truly reconstructing events it takes the stable particles from the event generators and directly records them as reconstructed particles, smearing some of their properties (e.g. energy and momentum) by a normally distributed random number to imitate the real detector resolutions. It does not include particle interactions with the detector (e.g. scattering) and has 100% particle reconstruction rates for particles such as electrons, photons and muons provided that they pass a minimum transverse momentum cut and are found within the acceptance window for the relevant (precision) parts of the detector that are designed to pick them up. ATLFAST provides a good overview of event shapes and generates $\sim$ 100000 events /hr^{-1}, making it indispensable for fast estimates of signal and background rates in specific physics channels.

3.5 Reconstruction

The ATLAS reconstruction software takes information from the readout of the individual detector components (or from the digitised full simulation data), and builds it into a picture of where particles passed through the detector components. It then reconstructs particle energies and momenta to identify different types of particle and determine their kinematics.

Reconstruction takes place in three basic phases. Firstly, data is reconstructed within each detector component, and combined into tracks (resulting from the passage of charged particles through the tracking detectors) or clusters (representing energy deposits in the calorimeters). Secondly the information from all of the subdetectors is combined and used to identify key objects such as electrons, muons, photons and hadronic jets. The third stage is the application of analysis-specific algorithms for identification of events of physical interest. The following sections give a brief overview of reconstruction in the ATLAS experiment.

3.5.1 Track Finding and Vertexing in the Inner Detector

When a charged particle passes through the tracking sub-detectors it generates a meaurable hit in each detector component it passes through. Combining all detector hits within an event enables the reconstruction of the path and momentum of the particle.

Figure 13 shows a simulation of the production of a Higgs boson which decays to a bottom and anti-bottom quark, as viewed in the transverse plane in the barrel region of the ID. The hits in the different regions of the ID can clearly be seen: the most central three rings of hits represent the layers of the pixel detector, which are followed by the four layers of the SCT, surrounded by the densely-populated TRT layers. Therefore a track in the ID is generally identified from three well-separated consecutive hits in the layers of the pixel detector and four in the SCT, followed by around 36 closely-spaced points in the TRT. In events with high detector occupancy such as this distinguishing the tracks of individual particles is a challenge and is the task of the pattern recognition software.

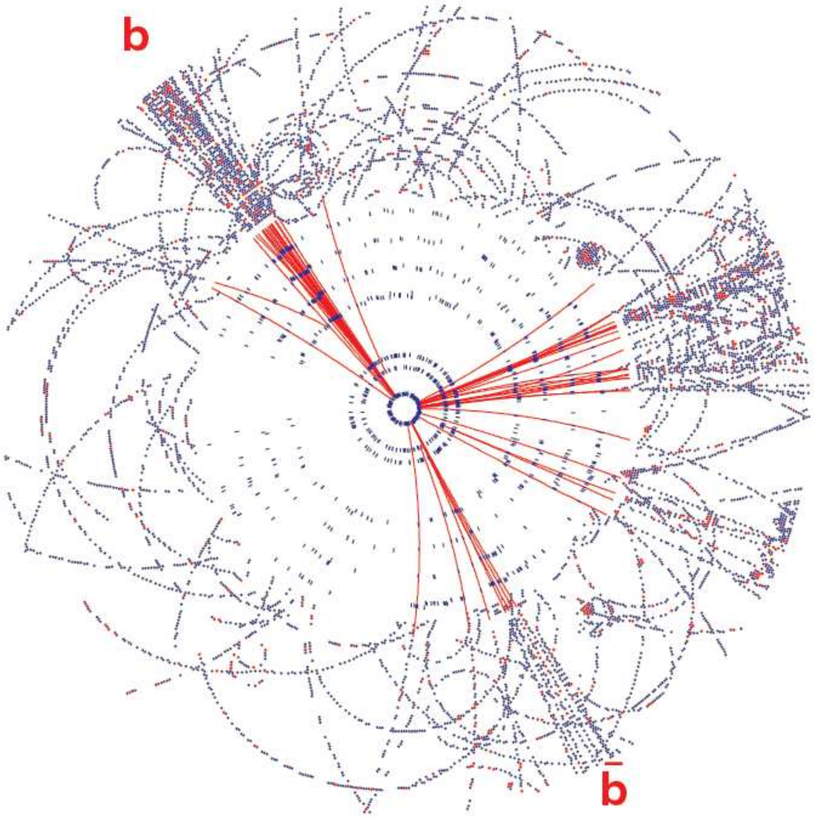


Figure 13: Example of a simulated event in the Inner Detector ($H \rightarrow b\bar{b}$) [38]

In general, to reconstruct tracks, pattern recognition software [62] is used to determine which inner detector hits belong to which tracks and provide an estimate of the track parameters. This information is then passed to a track fitting algorithm which attempts to calculate a realistic track trajectory for each charged particle. Due to the soleniodal field the tracks in the ID follow a helical path whose radius decreases as the particle loses momentum in passing through the detector. The five parameters of the helix will depend on the particle momentum and the magnetic field strength. The fit is significantly complicated by the fact that the magnetic field is not completely uniform, by the energy loss of the particles and by interactions with the detector components, such as Bremsstrahlung of electrons, photon conversions into electron/positron pairs and multiple scattering in which the direction of the particle's trajectory is deflected by interactions with the consecutive detector layers.

Quality cuts are applied to the reconstructed tracks. The newTracking algorithm requires that the tracks have a minimum p_T of 100 MeV, pass within 1mm of the nominal interaction point in the transverse plane and 10cm in z. It also requires that the track contains (at least) seven hits in the combined pixel and SCT layers (note that there are more than seven layers in the endcap regions) and that there is a hit in the innermost pixel layer that circles the interaction point.

Five parameters are used to descibe the properties of a track. The track parameters used at ATLAS are:

- q/p, the product of the inverse of the value of the momentum of the track with the charge of the particle (extrapolated from the direction of its curvature in the magnetic field)
- d_0 is the *impact parameter*, which is the distance of the track from the z axis at the point of closest approach to the z axis
- ϕ_0 is the angle of the track in the (x, y) plane at this point

- θ_0 is the angle of the track with the nominal x of the track at this point
- z_0 is the z-coordinate at this point

For illustration, Figure 14 (taken from [62]) shows a two dimensional projection of a track passing close to the interaction point at the axis with the relevant parameters labelled.

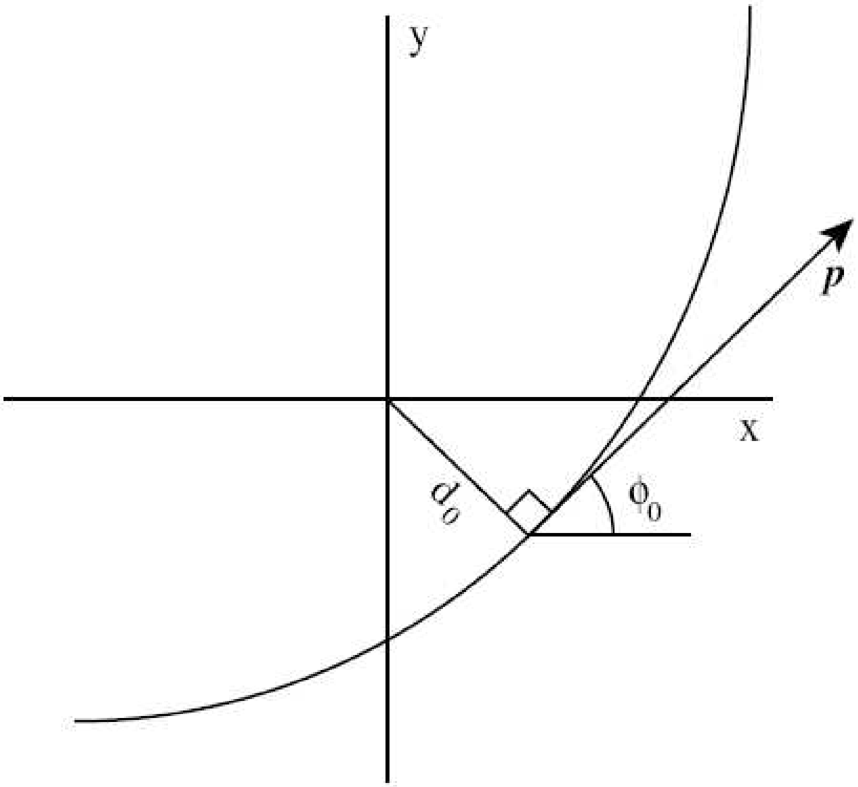


Figure 14: Scematic description of track parameters

Once the tracks are reconstructed they are projected towards the interaction point. The primary vertex position in the (x, y) is defined at the centre of the beam pipe, $(0, 0)$. To find the z co-ordinate a sliding window of 0.7 cm in z is moved along the interaction region of the beam pipe. The window with the largest number of tracks projected into it is considered to contain the primary vertex, the position of which is given by the mean z

of all the tracks in the window. All tracks which are identified with the primary vertex are removed and the sliding window proceedure is repeated to find other vertices.

The resolution of the transverse momentum of charged particle tracks is discussed in more detail in Section 7, but as a guide the transverse momentum resolution of a track with $p_T = 20$ GeV is around 0.2%, for $p_T = 400$ GeV it is around 2%, and for 1 TeV 10% is expected.

3.5.2 Clustering in the Calorimeters

The primary processes by which electrons and photons interact with matter are respectively Bremsstrahlung and pair production. In Bremsstrahlung the electron's (or positron's) momentum is affected by an electromagnetic interaction with a nucleus, at which point it radiates an energetic photon so that momentum is conserved. Fast moving photons when passing through matter split into electron and positron pairs, which in turn will be slowed by Bremsstrahlung. The process repeats, forming a *shower* of electromagnetic particles. The ECal is designed to bring these particles to a stop by forcing them to lose their energy through repeated particle production in quick succession (it is made of very dense material to increase the interaction rate), until the energy of the resulting particles is too low for further multiplication. After this the remaining energy is dissipated in Compton scattering of photons and ionisation losses for the electrons. The energy of the original particle (electron or photon) is thus passed into the ECal cells, and a cluster is reconstructed where the shower is found.

Collisions between hadrons and nuclei can result in strong interactions which produce quarks or gluons, which will themselves hadronise, and so showers of hadrons are formed in this way. The probability of hadrons passing close enough to nuclei to interact via the very short-range strong force is much smaller than the probabiliy of electrons interacting via the much longer-range electromagnetic force, and so the amount of showering in the ECal is quite low. The lightest hadrons are neutral pions, π_0, many of which will be produced in most hadronic showers. Their dominant decay mode is $\pi_0 \rightarrow \gamma\gamma$ (with a branching fraction of almost 99%), and these photons produce electromagnetic showers, meaning that hadronic showers which begin in the ECal are often detected. Whilst hadrons frequently start to shower in the ECal they are not usually stopped by it, so hadronic showers will generally stretch over both calorimeters. The role of the HCal is to collect the remainder of the hadronic energy.

The geometric region in which a passing particle deposits its energy in the calorimeters is called a cluster. These are searched for with one of two main algorithms, a sliding window algorithm [63] [64], or topological clustering [65]. The default sliding window clustering algorithm is used here. It proceeds in three steps: tower building, seed finding, and cluster filling.

In tower building the calorimeter is divided into segments in $(\eta \times \phi)$ space of (0.025×0.025) for the ECal and (0.01×0.01) for the HCal. The energies deposited in all of the calorimeter cells in this angular slice of the detector are summed to give the tower energy.

In seed finding a window of fixed size is created where $N_\eta \times N_\phi = 5 \times 5$ are the numbers of towers in η and ϕ. To reconstruct clusters, the energies of the towers within every possible window in the calorimeter are summed, and where the total transverse energy of the window is > 3 GeV a cluster "seed" is considered to have been found. A more precice measure of the position of the seed is then calculated using a smaller window of 3×3 around the area and choosing the point with the largest total energy to be the centre. If other seeds are found within 2×2 of the seed the one with the largest transverse energy is kept and the others rejected.

The positions of the seeds are used by the reconstruction software to fill clusters. A cluster includes all cells which are located inside a rectangle of variable size around the seed, usually $N_\eta \times N_\phi = 5 \times 5$, 3×5 or 3×7. The size of the window chosen depends on the type of particle to be reconstructed. For example, 3×5 windows are generally preferred for electron reconstruction (electromagnetic showers are generally narrow compared with hadronic ones), but a larger windows would be used if a photon conversion is reconstructed since this would produce an electron and positron pair, causing a spatially larger deposit in the calorimeter than a single electron or photon. For this reason the final size of the cluster is chosen after preliminary particle idenification.

In the real detector energy deposits in the calorimeter cells will be read out by the electronics in terms of a current. Conversion factors are needed to express this current in terms of the physical energy of the particles that produced it. These conversion factors were estimated for the detector simulation using Monte-Carlo simulations of $Z \to e^+e^-$ events to match simulated and reconstructed particle energies and masses. When the ATLAS detector starts to run they must be extracted from studies in data. This correction is known as the *EM scale*, and it is applied to all calorimeter towers by default.

3.5.3 Identification and Reconstruction of Electrons and Photons

To identify electrons a track in the inner detector is matched to the position of a jet in the Electromagnetic Calorimeter. The track is required to have $\frac{E}{p} \sim 1$, since the electronic mass is very small compared with its momentum at LHC energies. Large amounts of transition radiation in the TRT further indicate likely electron candidates.

Charged hadrons will leave some of their energy in the ECal in ionisation or, particularly in the case of pions, in decays to electromagnetic particles (the dominant decay of the π^0 is into two photons, and charged pions often lose their charge in “charge exchange” interactions with nucleons via $\pi^- p \to \pi^0 n$). Discrimination between electrons and pions is performed based on the shapes of the showers produced, the primary differences being that the electromagnetic particles interact immediately with the ECal whereas hadronic showers usually begin further downstream for the reasons stated in 3.5.2, and that electrons or photons will form narrower clusters in (η, ϕ) than those created in hadronic decays.

3.5.4 Identification and Reconstruction of Muons

Muons, being significantly heavier than electrons, do not find their trajectories noticably altered by the dense material of the calorimeters and leave only ionisation traces in these. They form tracks in the inner detector, and further tracks in the muon spectrometer.

The muon spectrometer provides standalone muon identification and measurement, typically using three different segments of muon tracking detectors which each reconstruct a segment of the muon trajectory. These are combined into a muon path, which is traced back through the detector to the ID (correcting for its estimated energy loss in the calorimeter) and combined with an ID track in order to provide an accurate measure of muon momentum.

3.5.5 Reconstruction and Calibration of Hadronic Jets

The presence of a hadronic jet implies the formation of a quark or gluon in the initial collision (or subsequent interactions). Hadronic jets consist of groups of nearly collinear hadrons, many of which form individual tracks in the ID, and the summed energy of the entire jet measures the energy of the initial parton. Jets are reconstructed in ATLAS using a number of different jet-finding algorithms (as described in [39] Section 9.1.2). The default jet reconstruction algorithm in ATLAS is the cone algorithm. The basic approach of the cone algorithm is to start with a cluster, the formation of which was described in Section 3.5.2, take the tower within it with the highest energy as as "seed" for the jet and then form a cone around that point with a specified radius in (η, ϕ) space of $\Delta R = \sqrt(\eta^2 + \phi^2)$ where usually $\Delta R = 0.4$ or 0.7. The centre of the cone is then iterated around every tower in the cluster to find the point at which the maximum energy density is contained within the cone, and the jet is considered to exist there.

The jet clustering algorithms produce jets which are not calibrated (or rather, are calibrated to the EM scale). The calorimetry will respond very differently to hadrons than electrons and so jet calibration algorithms must be applied in order to reconstruct jets with physically accurate energies. A weighting to transform from the EM to the *jet energy scale* is used in which each cell is weighted by a factor which depends upon the energy density within it, and the detector region. This is applied to each cell of the hadronic calorimeter. The weighting factors were extrapolated for the ATLAS detector simulations using Monte Carlo samples of di-jet events (those in which the primary interaction produces two QCD jets back-to-back). They must be calculated again in data when ATLAS starts, probably using γ + jet events. In these events the photon and jet will be produced with equal and opposite momenta and so the known EM energy scale will allow extraction of the jet energy scale.

The energy of the hadronic jet is then calculated by the following basic method:

$$E_{jet}^{reco} = \sum_{icells\in} w_i(\rho_i)E_i \tag{19}$$

where E_i is the energy deposited in the ith calorimeter cell, ρ_i is the energy density in the cell (energy/cell volume), and w_i is a weighting function that depends on ρ_i. The weights are calculated by minimising the quantity

$$\chi^2 = \sum_{jets} (\frac{E_{jet}^{reco} - E_{jet}^{truth}}{E_{jet}^{truth}})^2 \tag{20}$$

for the di-jet sample. The process is described in detail in [39] Section 9.1.1.

Many detector effects must be take into consideration when reconstructing jets, and complex corrections are made in the reconstruction software for a wide range of effects including the different responses of the calorimeter cells to charged or neutral hadrons, non-linear detector responses, fluctuations in the magnetic field, "leakage", in which the full particle energy is not contained within the reconstructed clusters, the finite granularity of the calorimeter cells and the effects of electronic noise. Algorithms are also implemented to deal with the case in which two or more jets overlap in the calorimeters. [66] gives a more detailed explanation of how jets are reconstructed in ATLAS, and [67] for a general discussion of jet reconstruction algorithms.

3.5.6 Reconstruction of Missing Transverse Energy

Conservation of momentum indicates that the total transverse momentum of each event should equal zero. The total transverse momentum of the particles which escape detection (or missing transverse energy, MET) is thus taken to equal the negative of the vector sum of the transverse momenta of all the particles detected. The MET variables used for physics analysis are reconstructed by summing the energy deposits in all calorimeter cells. Corrections are applied to take into account energy lost in the material between the electromagnetic and hadronic calorimeters, and also the differing calibrations for hadrons, electrons and photons. Muon corrections are also applied in which the energy deposited in the calorimeter by the muon is removed and the muon transverse momentum reconstructed from the muon spectrometers is added instead (since muon energy is not accurately measured in calorimeters). The missing transverse energy variables are again calibrated in a similar fashion to the jet energy corrections in Section 3.5.5 above.

3.6 Event Display

Athena contains a module called GeoModel which contains all of the information about the physical detector, such as properties of the materials of which each component is made, and the detailed position of each part of the detector. This information is stored in a shared database so that each part of the software can work with the same database to produce consistent results.

Figure 15 shows a set of screenshots taken from HEPVIS. The event shown is an example of the production of a W boson with a jet where the W boson decays immediately to an electron and neutrino. This event is taken from the (fully simulated and reconstructed) sample used for this analysis, the generation of which will be described in Section 4. The various detector components are displayed according to the GeoModel information.

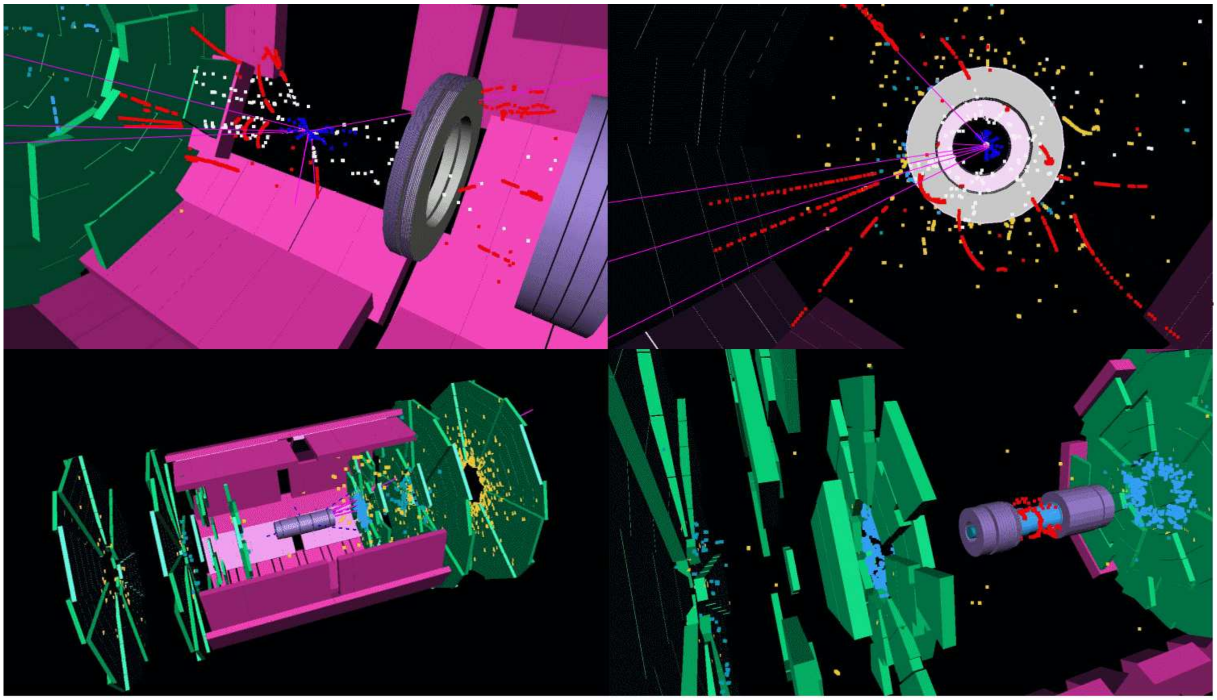


Figure 15: HEPVIS representations of a $(W \rightarrow e\nu) + jet$ event in the ATLAS detector

The top left picture in 15 shows a close-up of the interaction point from just outside the ID with a few of the TRT endcaps visible (grey), the electron can be seen moving towards the right of the picture and the particles which form the jet are moving to the left. The hits in the pixel detector are shown in dark blue, the SCT in white and the TRT in red. The pink lines emanating from the interaction point describe the initial momenta of the charged particles leaving the interaction point. The barrel regions of the muon spectrometer are shown in pink and the endcaps in green. The top right picture shows an end-on view of the same event, looking down the beam pipe. The bottom pictures show the same event from different angles from farther out in the detector, where the light blue and yellow markers represent hits in different regions of the muon chambers.

4 Generation of the Signal Dataset

The generic leading order (LO) Feynman diagram for W boson production is shown in Figure 16, where q_a, $\bar{q}_b$ are different types of quark with unit charge difference (for example $u + \bar{d} \rightarrow W^+$, $s + \bar{c} \rightarrow W^-$, etc.).

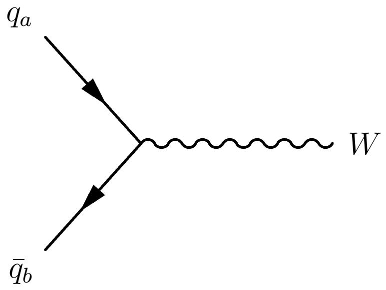


Figure 16: First order Feynman diagram for W boson production

The Feynman diagrams for the signal channel that were shown in Figure 7 are examples of second order W production, the full set of which are shown in Figure 17. The central and right diagrams would become the signal channel in the case $q_a = s/\bar{s}$, $q_b = c/\bar{c}$. Therefore a dataset of second order-produced W bosons was created for this study, as it includes the signal channel, and also many of the major W boson-containing backgrounds (these will be discussed in Section 9).

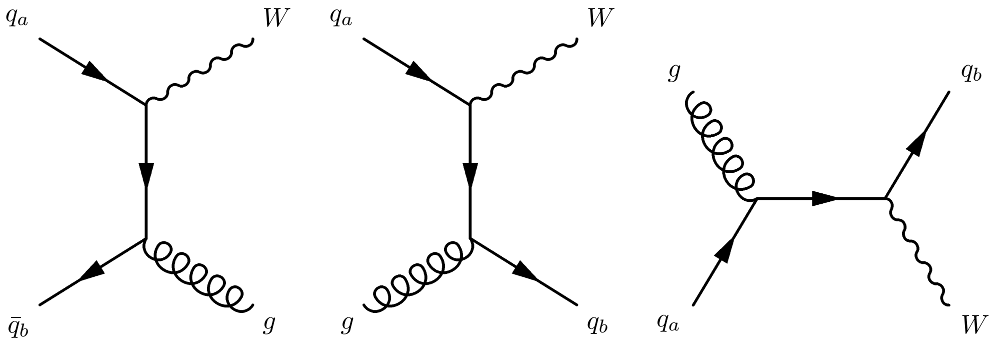


Figure 17: Secon order Feynman diagrams for W boson production

The MC@NLO event generator (version 3.1 [68]) was used to create a set of events for this analysis in which W bosons were produced to next-to-leading-order, and these were required to subsequently decay into electrons/positrons and neutrinos. (MC@NLO is a relatively slow and cumbersome generator, but it provides complete NLO calculations for the production of and leptonic decay of W bosons, as was discussed in Section 3.4.1.) The NLO CTEQ6m [17] PDF set was used, which assumes a symmetric strange and antistrange quark distribution, and similarly for the charm and bottom sea components.

The phase space information for the generation of three million events of each of $W^- \rightarrow e^-\bar{\nu}$ and $W^+ \rightarrow e^+\nu$ was prepared in MC@NLO (See [69] for a detailed description of the generation process: note that the W boson mass was set to 80.425 GeV with a width of 2.124 GeV), and this was passed to HERWIG for hadronisation. The minimum p_T of the quarks and gluons created in the hard process was set to 10 GeV, and the minimum p_T of those created in secondary scatters, the interactions between spectator partons, was set to 3.85 GeV in HERWIG. These cuts were included to decrease the time spent by the generator in calculating soft processes which would not produce particles which would be reconstructable in the detector, whilst rejecting a negligible number of events that could be reconstructed into detector objects. Two additional modules were called from HERWIG to complete the event generation: Photos [70] was used to deal with final state QED radiative corrections, and Tauola [71] was added in order to deal with τ-lepton decay. The samples were then run through the ATLFAST fast detector simulation package (in ATLAS software version 12.0.3).

The cross-sections, σ, for $W^+ \rightarrow e^+\nu$ and $W^- \rightarrow e^-\bar{\nu}$ generated in this manner were calculated after HERWIG hadronisation to be:

$$\sigma_{W^+\rightarrow e^+\nu\ NLO} \simeq 2.18 \pm 0.06 \text{ nb}$$

$$\sigma_{W^-\rightarrow e^-\bar{\nu}\ NLO} \simeq 1.76 \pm 0.07 \text{ nb}$$

After subtracting negative weights the total number of physical W^+ events remaining in the sample was 2.46×10^6, and there were 2.44×10^6 W^- events (4.90×10^6 in total). The sample corresponds to approximately 1 fb^{-1} of data (N.B. the experiment is expected to accumulate about 1 fb^{-1} in around a month of running in the low luminosity phase.) The data were re-weighted by a factor $\mathcal{W}$, using the cross-sections to give a total number of events that would be expected per fb^{-1} (i.e. 2.18 million W^+ and 1.76 million W^- events). The calculation used was

$$\mathcal{W} = \frac{\sigma \mathcal{L}}{N_w} \tag{21}$$

where σ is the cross-section of the channel, $\mathcal{L}$ is the integrated luminosity required (one fb^{-1}) and N_w is the number of physical events in the dataset.

W bosons created in higher-than-NLO production mechanisms were not be included in the study since production of W bosons to higher orders in α_s than NLO is both highly suppressed and very complex to reproduce in Monte Carlo. This leads to a systematic uncertainty in the production cross-section of the W bosons: NNLO corrections are thought to result in a 2-4% uncertainty in the $W \to e\nu$ production cross-section at ATLAS [72] (this assumes selection of electrons with transverse momenta > 20 GeV and $|\eta| < 2.5$, and MET > 20 GeV, which are similar to the cuts applied to select W bosons in this study, Section 5.5. See [73] also). Therefore an additional uncertainty of 4% will be added here, so the relevant W boson production cross-sections become

$$\begin{aligned} \sigma_{W^+ \to e^+\nu\ NLO} &\simeq 2.18 \pm 0.15 \text{ nb} \\ \sigma_{W^- \to e^-\bar{\nu}\ NLO} &\simeq 1.76 \pm 0.14 \text{ nb} \end{aligned}$$

It was observed that ~76.5% of the sample was produced from leading order W generation diagrams, and the remaining ~23.5% given by next to leading order diagrams. Gluons are produced in the final state in ~ 10% of events (Figure 17).

5 Kinematics of W Boson Production

The kinematics of the events produced in Section 2.4 were studied before they were passed into the detector simulation programs. The plots shown in this section are created from the generator level information only (from the MC@NLO generator) and include no detector effects. The samples used here were one million of each of the $W^+ \to e^+\nu_e$ and $W^- \to e^-\bar{\nu}_e$ events. Plots of rapidity, pseudorapidity and longitudinal momentum are weighted by production cross-section such that the W^- plots correspond to the same luminosity as the one million W^+ event sample. These are marked with **. The plots of transverse energy and mass are not weighted and show raw numbers of events, for easy comparison of their shapes.

5.1 W Boson Production at ATLAS

As shown in Figures 16 and 17 W bosons are produced from incident partons in combinations of $q\bar{q}$, $g\bar{q}$ and qg. In a proton-proton collider such as ATLAS the gluons and antiquarks must always come from the proton sea, whereas the quarks can be either valence or sea. There are twice as many valence u as d quarks available in the proton (u quarks contribute to W^+ production and d quarks to W^-) so the cross-section for W^+ boson production at the LHC will be significantly greater than that of W^- bosons.

Figure 18 (taken from [74] Figure 15-2) shows the kinematic constraints on the momentum fraction x of the pairs of incident partons (one from each proton) that can produce of particles of mass M at rapidity y at the LHC. It indicates that W bosons found within the pseudorapidity window for accurate reconstruction of electrons in the ATLAS detector ($\eta < 2.5$) are generally produced from partons with $10^{-4} < x < \sim 10^{-1}$. As a rule most of the proton momentum is carried by the valence quarks ($x \sim 0.1$) and individual sea partons tend to carry small momentum fractions, existing in the low x region ($x \ll 0.1$). Bosons at the W mass are produced from partons a and b when $x(a)x(b) \approx 0.0001$.

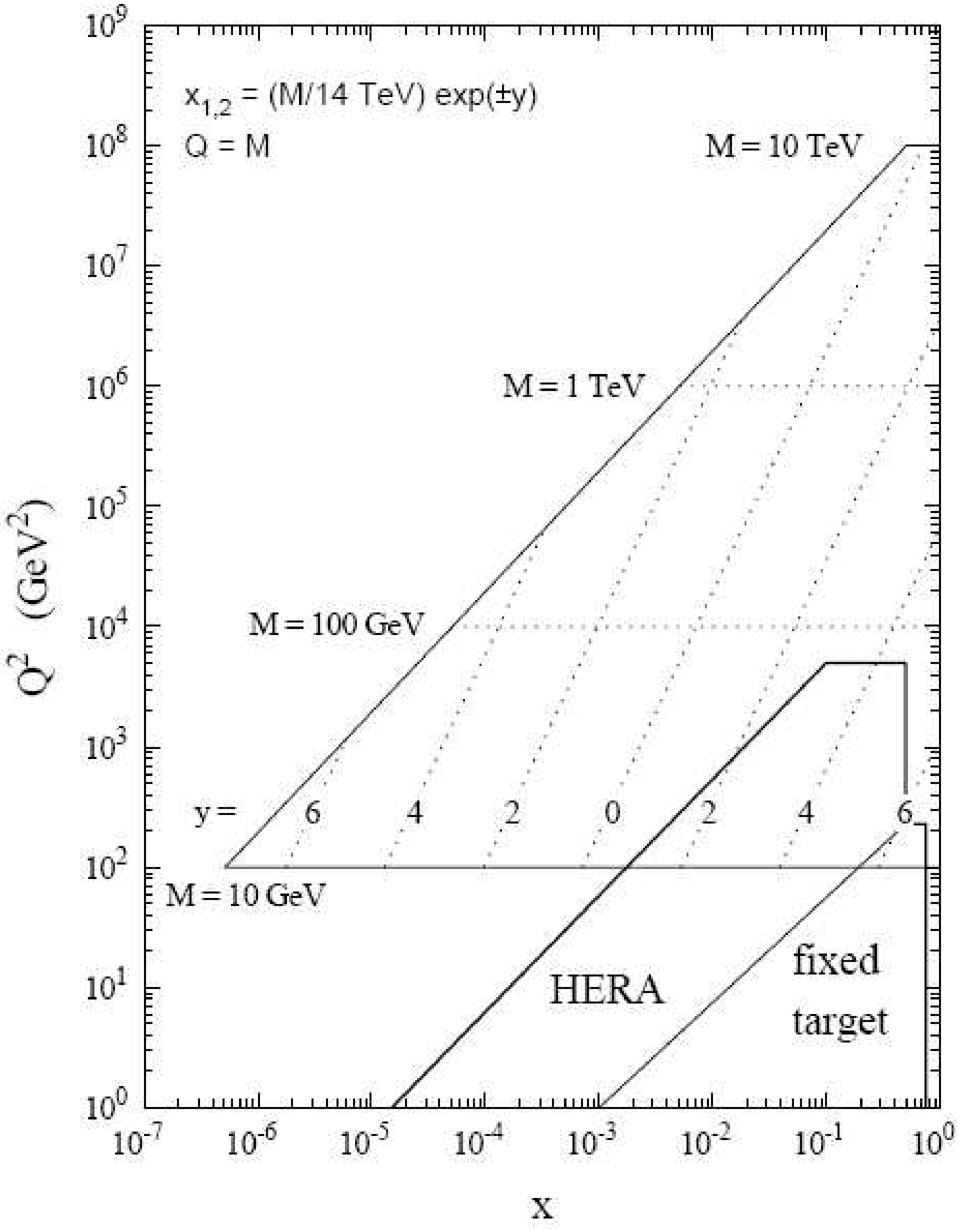


Figure 18: Parton kinematics at the LHC in the (x, Q^2) kinematic plane for the production of a particle of mass M at rapidity y (dotted lines).

5.2 Kinematic Properties of W bosons

The MC@NLO sample described in Section 4 was divided into two parts: those in which the W bosons were generated in first order production mechanisms (accounting for around 76.5% of the sample), and those which were generated in second (the remaining 23.5%). The kinematic properties of each were studied separately.

5.2.1 W bosons formed in first order production mechanisms

In leading order generation of W bosons the W can only be produced travelling longitudinally down the beam pipe. This is a result of conservation of momentum, since the interacting partons will have negligible transverse momenta. Therefore the W boson transverse energy (E_T) is zero. However the mass, rapidity (y) and longitudinal momentum (p_z) for each W boson can be reconstructed, and these are shown in Figures 19, 20 and 21 respectively.

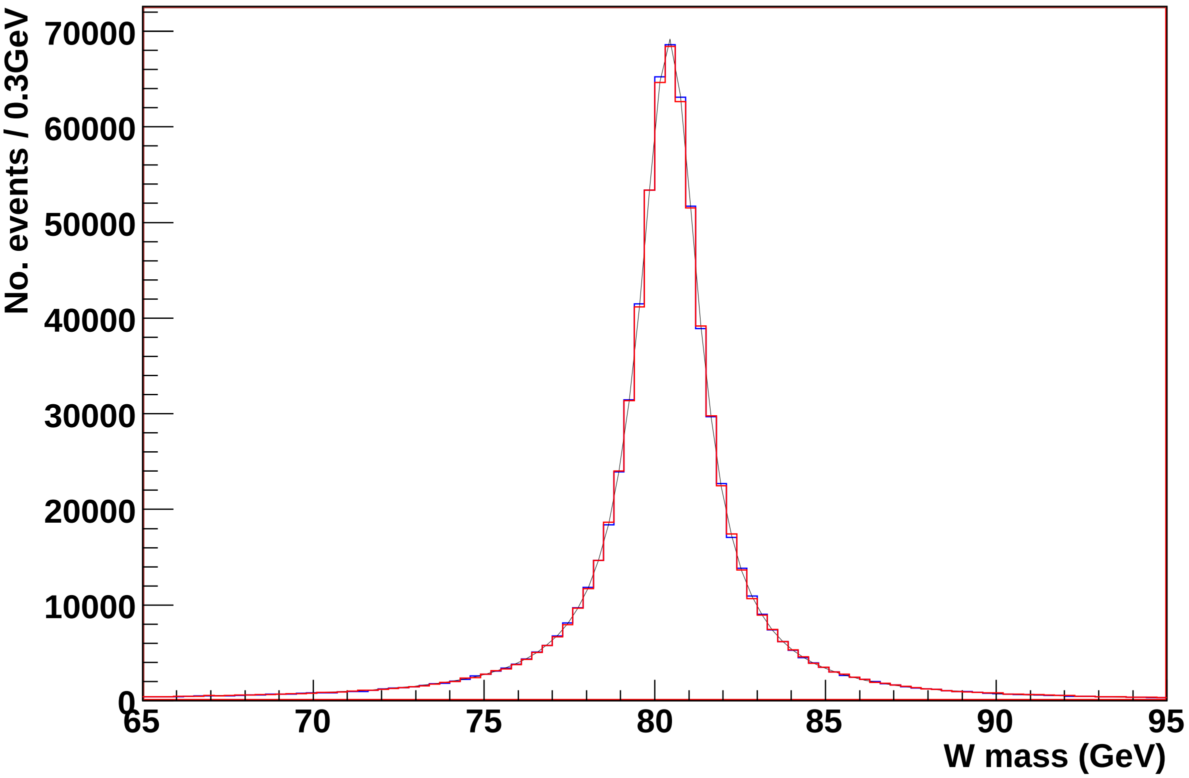


Figure 19: W boson masses, W^- (blue) and W^+ (red), for W bosons produced in first order diagrams. The W^- distribution is fitted with Breit-Wigner distribution (black, dashed).

A Breit-Wigner distribution was fitted to the W^+ and W^- mass plots (Figure 19) as per the following formula:

$$f_{BW}(M) = N \frac{M^2\Gamma^2}{(M^2 - M_{res}^2)^2 + M_{res}^2\Gamma^2} \tag{22}$$

The fitted mass of the W^+ is 80.430 ± 0.007 GeV with a width of 2.132 ± 0.005 GeV (χ^2 = 124.8 with 97 degrees of freedom), and the W^- has mass 80.428 ± 0.007 GeV and width 2.125 ± 0.005 GeV (χ^2 = 123.7 with 97 degrees of freedom). These correlate well with the W mass and width specified at generation (Section 4).

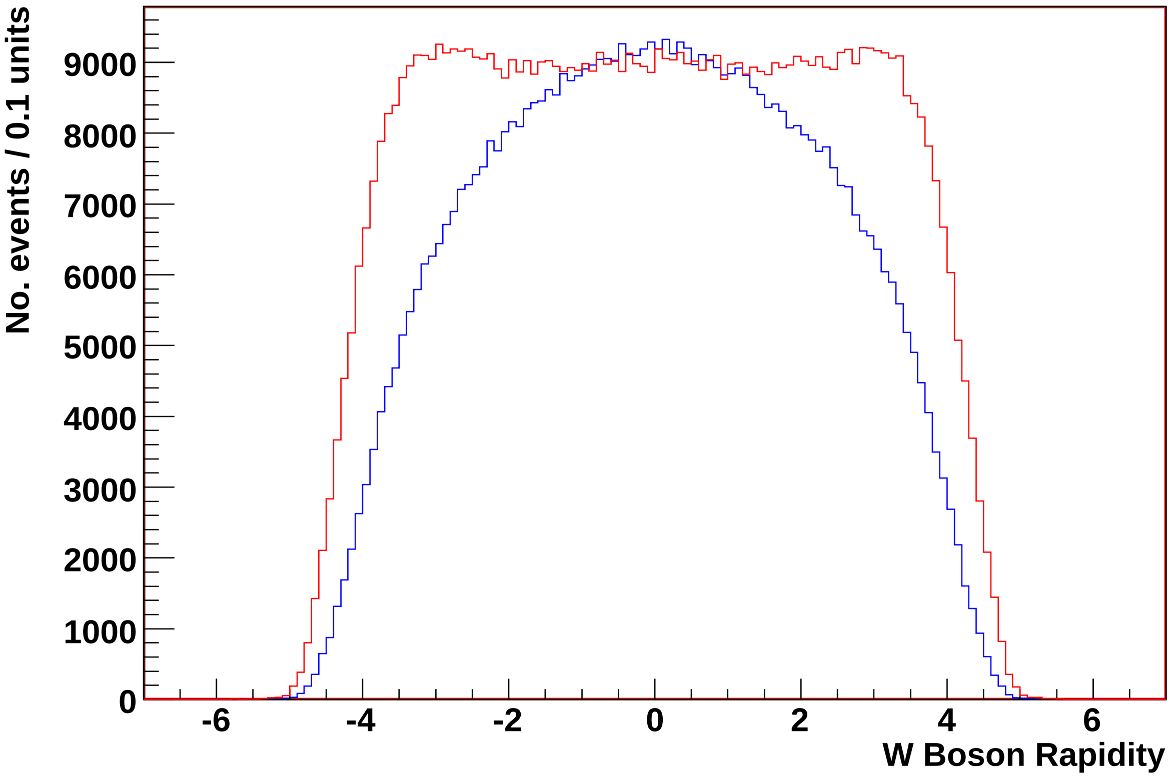


Figure 20: Rapidity of W bosons produced in first order diagrams, W^- (blue) and W^+ (red). **

The rapidity distributions of the W^+ and W^- bosons are shown in Figure 20. The difference in their shapes results from an underlying difference in the parton momentum fractions during production.

If a W boson is produced at first order (Figure 16) from the interaction of two partons with roughly the same momenta, i.e. when $x(a) \approx x(b) \approx 0.01$, then by conservation of momentum the W will have $p_Z \approx 0$, placing it in the central region of the detector at $y \approx 0$. If there is a large difference in parton momenta, e.g. $x(a) \approx 0.1$ and $x(a) \approx 0.001$, then the W boson will be produced at high rapidity, in the forward regions of the detector. Therefore the high rapidity W bosons are likely to have been produced from one valence and one sea parton, whereas the W bosons produced in the central region are likely to have come from sea-sea interactions. As a result the number asymmetry due to the excess of u over d valence quarks will be manifest away from the central detector region; at $y \approx 0$ the numbers of W^+ and W^- should be approximately equal.

As was discussed in 2.2.2, results from the E866 collaboration [22] have shown that u valence quarks carry a higher momentum fraction than d valence quarks (i.e. $x(u) > x(d)$ on average), and both of these to be higher than the sea quarks, for which $x(\bar{d})$ is measured to be $\gtrsim x(\bar{u})$. (A full description of known parton behaviours over the range of low-x values and their contribution to W production at the LHC can be found in [25]). As a result the W^+ bosons created from the interaction between $u + \bar{d}$ will generally be produced with a higher longitudinal momentum than the W^- from $d + \bar{u}$, since there will be a larger average difference in the momenta of the interacting quarks in the $u + \bar{d}$ case. This effect pushes the r.m.s. rapidity of the W^+ boson to a higher value than that of the W^-, which explains the differeing shapes of the W^+ and W^- rapidities seen in Figure 20. Since $y = 0$ corresponds to $p_z = 0$ this also leads to a greater fraction of W^- than W^+ bosons being found at low longitudinal momenta, which is seen in Figure 21.

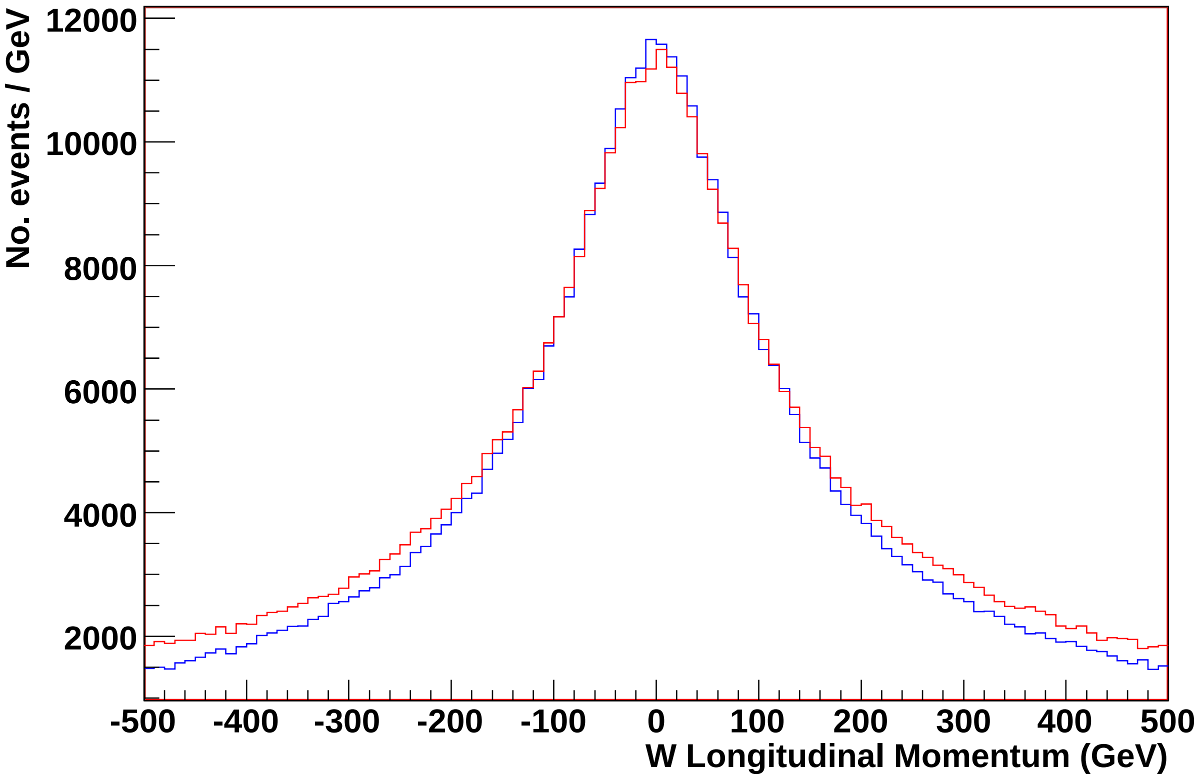
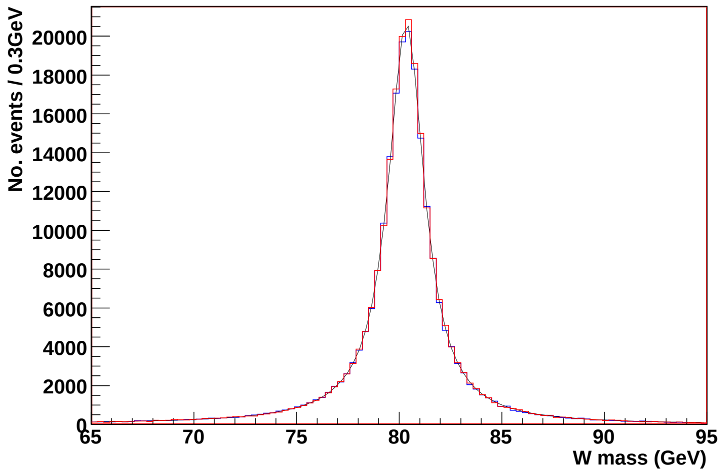


Figure 21: Longitudinal momentum of W bosons produced in first order diagrams, W^- (blue) and W^+ (red). **

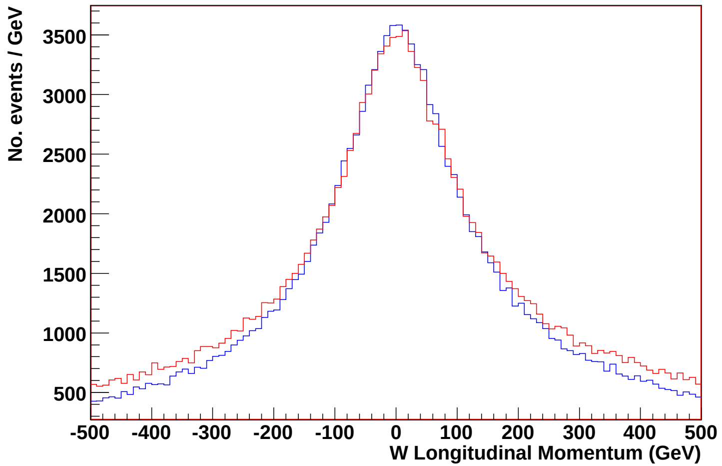
The asymmetries inherent in the production of W^+ and W^- from differences in the numbers of and momentum fractions carried by up and down type quarks in the proton must be considered when attempting to measure asymmetries due to differences in the strange and antistrange sea quark contributions as they can produce charge-asymmetric backgrounds.

5.2.2 W bosons formed in second order production mechanisms

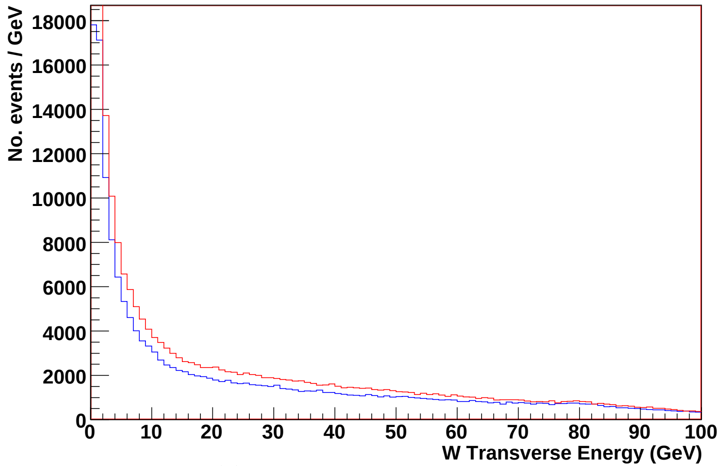
Figure 22 shows properties (mass, p_z, E_T) of the W bosons generated in the second order production mechanisms (Figure 16). There is very little difference from the leading order case (Section 5.2.1), except that now that the W bosons are produced in conjunction with a *partner* quark or gluon they have non-zero transverse momentum.



(a) W Masses, W^- fitted with Breit-Wigner distribution



(b) W longitudinal momenta **



(c) W transverse energies

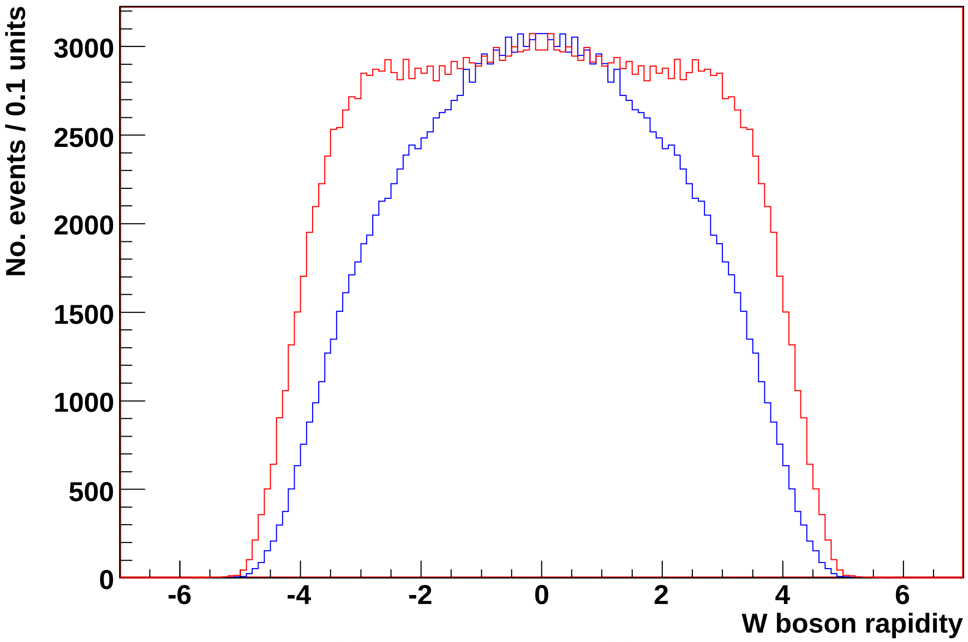
Figure 22: W^- (blue) and W^+ (red) rapidities, longitudinal momenta and masses for W bosons generated in second order diagrams. W^- mass is fitted with a Breit-Wigner distribution (black, dashed)

Figure 22(c) demonstrates that this is usually small.

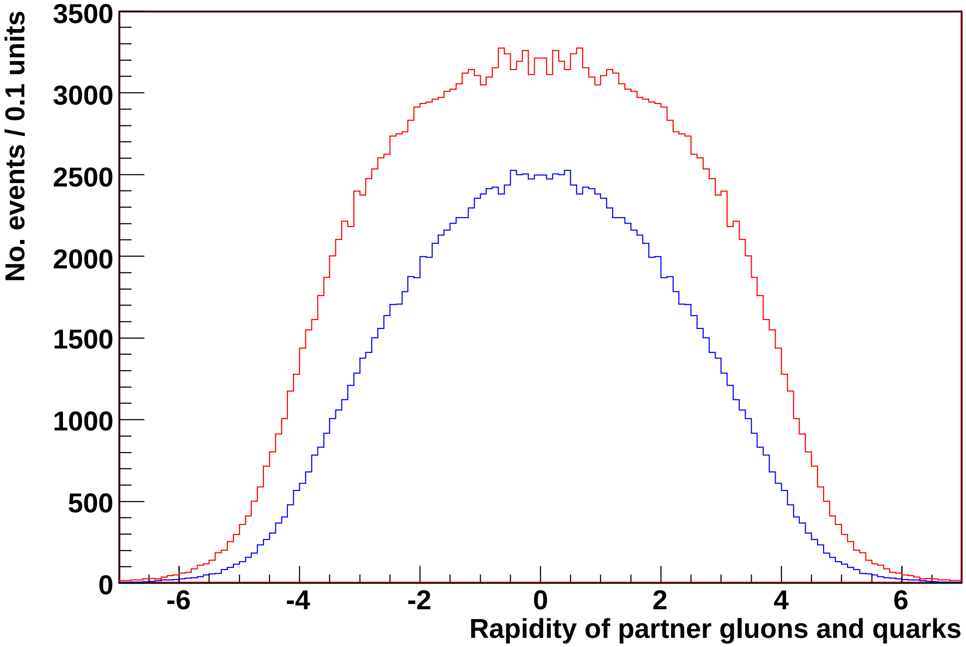
Again, fitting a Breit-Wigner to the mass distributions shown in Figure 22(a), the mass of the W^+ is found to be 80.355 ± 0.003 GeV with a width of 2.145 ± 0.007 GeV (χ^2 = 150.2 with 97 degrees of freedom), the W$^-$ has mass 80.346 ± 0.003 GeV and width 2.162 ± 0.007 GeV (χ^2 = 116.2 with 97 degrees of freedom).

Figure 23 shows the rapidity distribution of the second order-produced W bosons (Figure 23(a)) and also the rapidity of the gluon or quark partner (Figure 23(b)) produced with it. The W boson rapidity distributions are narrower than in the first order case (Figure 20) because the W now recoils against another parton which takes some of the total longitudinal momentum from the quarks entering the interaction. The rapidity distributions of the partner particles are narrower than those of the equivalent W bosons, implying that they tend to be produced with lower longitudinal momenta than the W. This is not surprising, since their masses will almost invariably be very much lower than that of a W boson.

An error in the generation software originally caused the second order rapidities to exhibit a small forward-backward asymmetry, with the W boson being emitted preferentially at positive rapidities and the partner at negative. This bug was reported, and subsequently for all particles produced from the second order diagrams plots of $\pm\frac{|y|}{2}$ or $\pm\frac{|\eta|}{2}$ are used in the place of y or η to give symmetric distributions.



(a) W boson rapidities **



(b) Partner rapidities **

Figure 23: W boson and partner quark/gluon rapidities for W^- (blue) and W^+ (red) produced at second order.

Figure 24 shows the partner rapidity distributions in the cases in which the partner is a charm or anticharm quark. The shapes of the two plots are very similar as would be expected, since the dominant generation mechanism for W+c production is the interaction of a strange or anti-strange quark with a gluon, and the PDFs for the strange and anti-strange sea quarks used were equal.

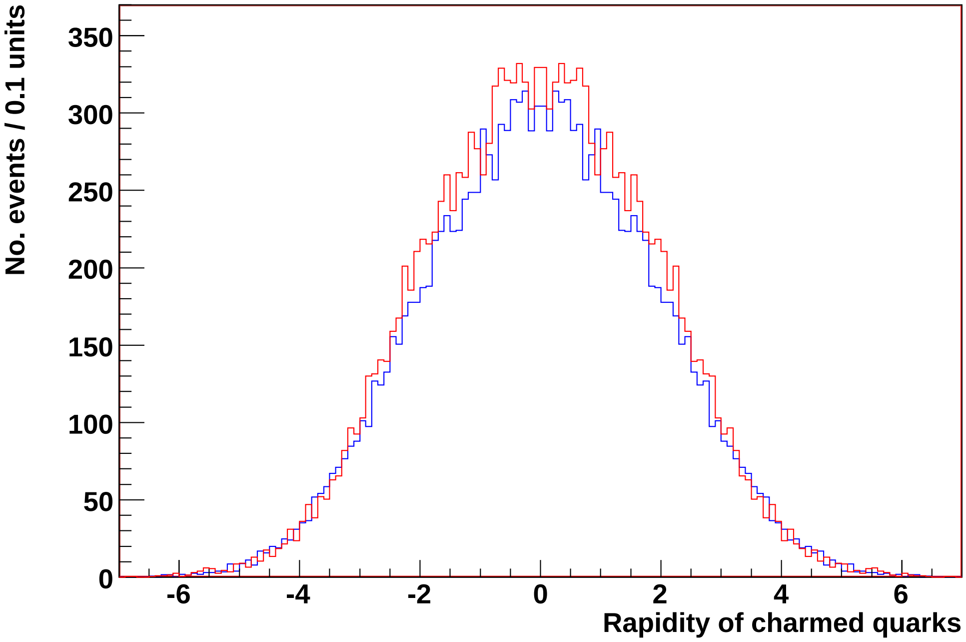


Figure 24: Rapidity distributions for the charm quarks produced with W^- bosons at second order (blue) and anticharm quarks with W^+ (red). **

5.3 Kinematic Properties of Leptons Produced from W boson Decay

5.3.1 Leptons from W bosons formed in first order production mechanisms

Figure 25 shows plots of the η, E_T and p_z of the electrons, positrons, neutrinos and antineutrinos that result from the decay of the W bosons produce in first order diagrams.

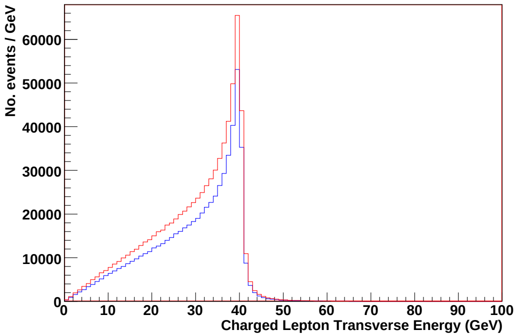
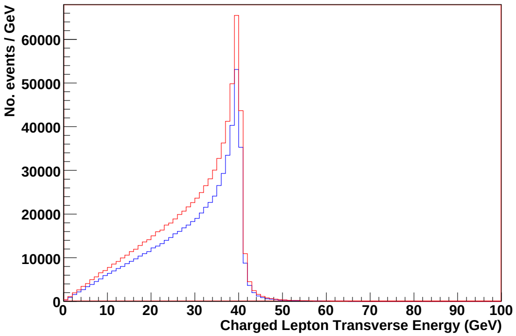
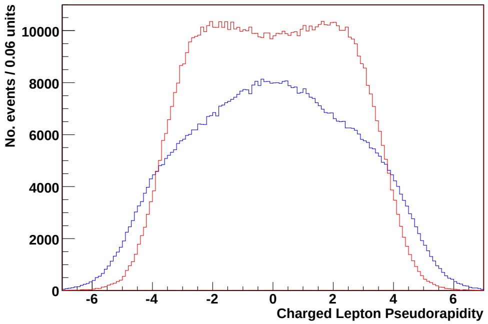
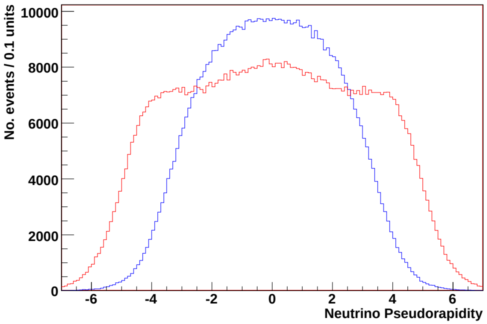
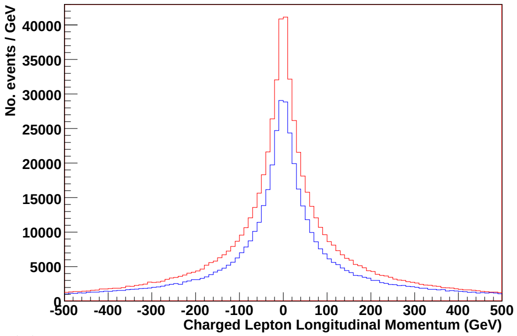
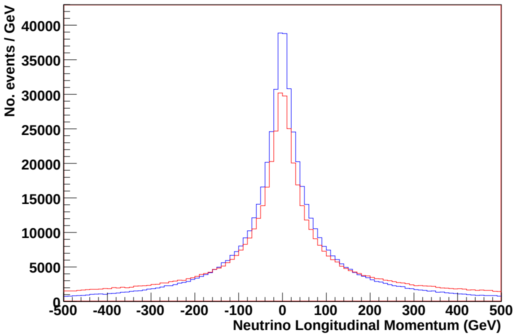
	e^- (blue) and e^+ (red)	**$\bar{\nu}$ (blue) and ν (red)**
$\mathbf{E}_T$	 (a) charged lepton transverse energy	 (b) neutral lepton transverse energy
η	 (c) charged lepton pseudorapidity	 (d) neutral lepton pseudorapidity
$\mathbf{p}_z$	 (e) charged lepton longitudinal momentum	 (f) neutral lepton longitudinal momentum

Figure 25: e^- and $\bar{\nu}_e$ (blue), and e^+ and ν_e (red) transverse energies, rapidities and longitudinal momenta, produced from the decay of W bosons generated by first order diagrams.

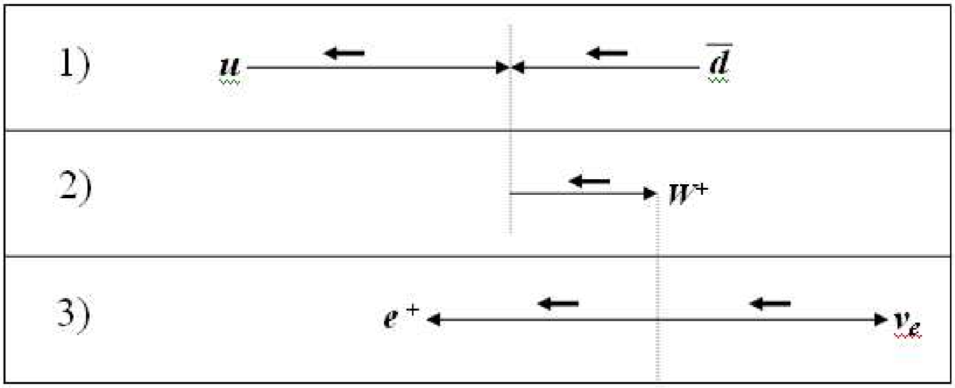
The transverse energies shown in Figures 25(a) and 25(b) cut off sharply at half the mass of the W boson as would be expected (the maximum lepton transverse energy being half of the W boson mass, since the W boson is produced at rest in the transverse plane in this case).

The W boson charge asymmetry will obviously follow through to the rapidity distributions of the leptons produced from the W boson decay. However these distributions are altered by helicity[**] considerations, which cause the positrons to be preferentially emitted back inwards along the line of flight of the W, and the neutrinos to be emitted forward in the case of the W^+. This is because the weak force interacts preferentially with left-handed particles and right-handed antiparticles. The W boson only interacts with particles of negative helicity and antiparticles of positive helicity in the case that for the interacting particles $\frac{E}{p}$ ∼1 (this is known as the "massless limit"). The effects of this on the momentum of leptons produced in W decay are demonstrated in Figure 26. For simplicity only the most common production modes for W bosons are shown ($u + \bar{d} \rightarrow W^+$ and $d + \bar{u} \rightarrow W^-$), and the massless limit is assumed for the quarks and leptons. Similar effects would be seen for other first and second order production channels.

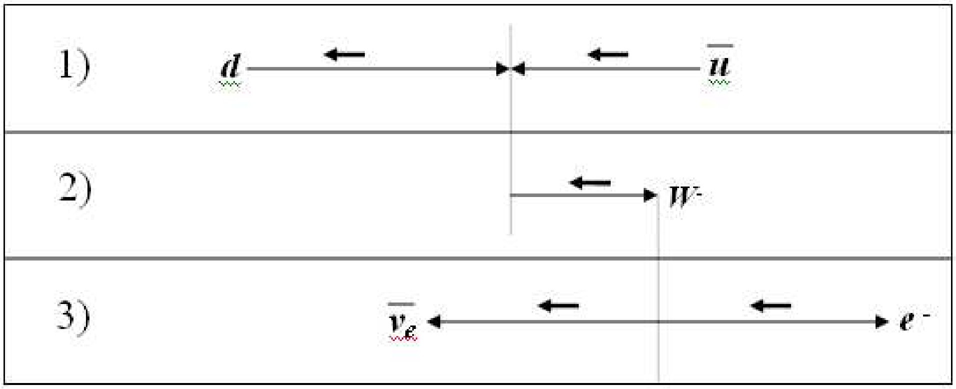
This effect results in positron pseudorapidity distributions that peak at lower rapidities than the W^+, with the neutrinos thrown out to higher rapidities. In the case of the W^- the electrons will be emitted preferentially forward along the W boson line of flight, and the antineutrino emitted backwards. Since the W^- peaks at central rapidity this simply means the electron rapidity distribution should be wider than that of the W^-, and the antineutrino narrower. These effects can be seen in Figures 25(c) and 25(d).

The fact that the positrons/neutrinos tend to be produced at higher rapidities than the electrons/antineutrinos explains why they also tend to higher longitudinal momenta (see Figures 25(e),25(f)).

[**] Helicity H is the projection of the vector spin of a particle onto its momentum vector, $H = \frac{\mathbf{s.p}}{|\mathbf{s.p}|}$. It is positive when the projection of the particle's spin onto the direction of its motion is positive, it is negative when the spin projection points in the opposite direction to the particle momentum



(a) 1) A W^+ boson is created from the interaction of a u of negative and a $\bar{d}$ of positive helicity. 2) Generally $x(u) > x(\bar{d})$ and so the W^+ momentum will be in the direction of the original momentum of the u. To conserve spin the helicity of the W^+ will be negative. 3) The W^+ decays into a positron and a neutrino. The neutrino must have negative helicity and will continue approximately along the W^+ boson line of flight, whereas the positron will have positive helicity and will move in roughly the opposite direction.



(b) 1) A W^- boson is created from the interaction of a d of negative and a $\bar{u}$ of positive helicity. 2) Generally $x(d) > x(\bar{u})$ and so the W^- momentum will be in the direction of the original momentum of the d. To conserve spin the helicity of the W^- will be negative. 3) The W^- decays into an electron and an antineutrino. The electron must have negative helicity and will continue approximately along the W^- boson line of flight, whereas the antineutrino will have positive helicity and will move in roughly the opposite direction.

Figure 26: Description of the effects of helicity conservation on the production and leptonic decay of W bosons.

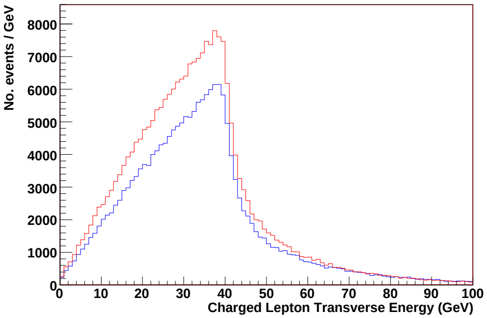
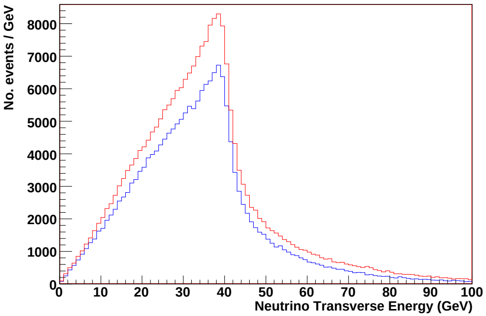
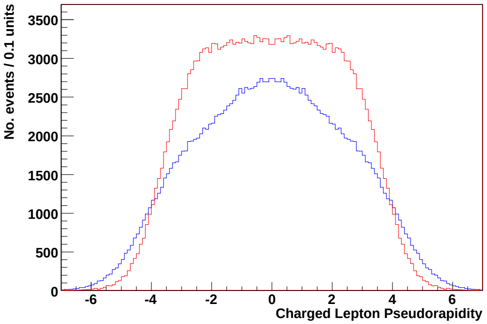
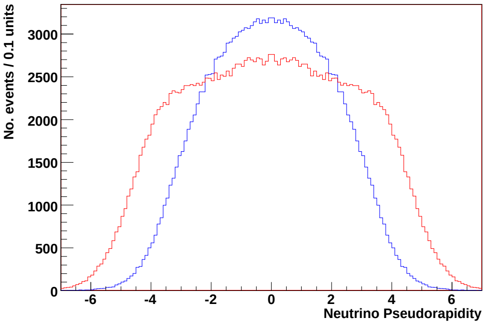
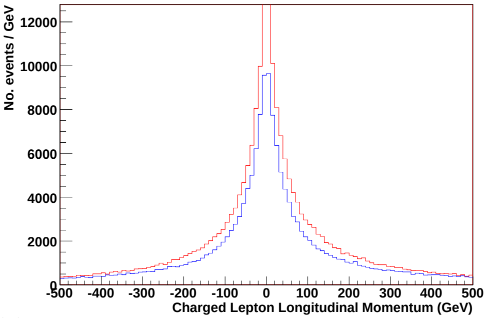
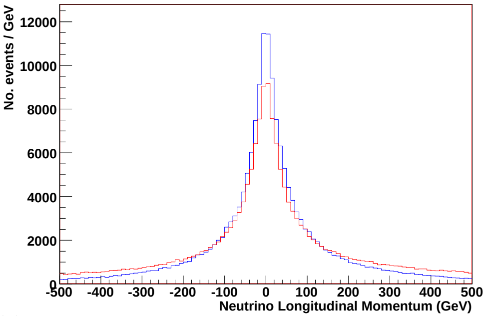
	e$^-$ (blue) and e$^+$ (red)	**$\bar{\nu}$ (blue) and ν (red)**
$\mathbf{E}_T$	 (a) Charged lepton transverse energy	 (b) Neutral lepton transverse energy
η	 (c) Charged lepton pseudorapidity	 (d) Neutral lepton pseudorapidity
$\mathbf{p}_z$	 (e) Charged lepton longitudinal momentum	 (f) Neutral lepton longitudinal momentum

Figure 27: e^- and $\bar{\nu}$ (blue), and e^+ and ν (red) rapidities, transverse energies and longitudinal momenta, produced from the decay of W bosons generated by second order diagrams.

5.3.2 Leptons from W bosons formed in second order production mechanisms

Figure 27 shows plots of the η, E_t and p_L of the electrons, positrons, neutrinos and antineutrinos that result from the decay of the second order generated Ws. The η distributions (Figures 27(c) and 27(d)) follow a similar shape to the leading order case (Figures 25(c) and 25(d)), as do the p_z distributions (Figures 27(e) and 27(f), compared with Figures 25(e) and 25(f)). The transverse energy distributions (Figures 27(a) and 27(b)) are also fairly similar to their first order counterparts (Figures 25(a) and 25(b)), but are much broader and less sharply peaked because in this case the W bosons from which they are created are produced with non-zero E_T.

5.4 Combined generator level plots

The mass distribution for all W^+ and W^- bosons generated in the combined first and second order production mechanisms is shown in Figure 28. The mean W mass is 80.410 ± 0.006 GeV with width 2.133 ± 0.003 GeV (χ^2 = 126.9 with 97 degrees of freedom).

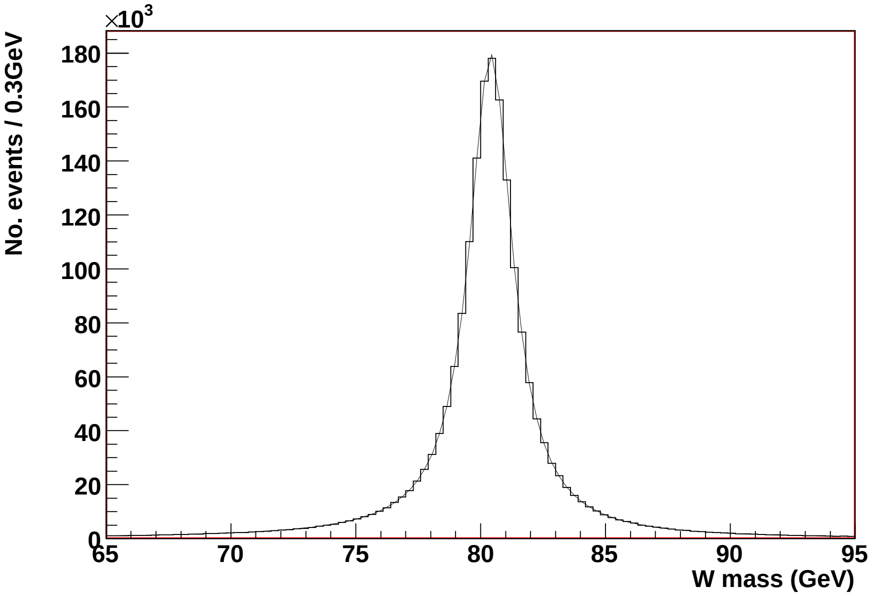


Figure 28: Total NLO W$^-$ (blue) and W$^+$ (red) masses

In data it is not possible to reconstruct the W mass directly, since the neutrino cannot be detected and so its longitudinal momentum is unknown; only transverse properties of the neutrino can be extrapolated from energy and momenum conservation. We define the property of "transverse mass" for the W:

$$\mathrm{M}_T = \sqrt{2 e_{ET} \nu_{ET} (1 - cos(e_\phi - \nu_\phi))}$$

(where e_{ET} and ν_{ET} represent the electron and neutrino transverse energies, the latter corresponding to MET in the detector, and e_ϕ and ν_ϕ their azimuthal angles) and measure this instead (as discussed in Section 3.5.6).

Figure 29 shows the transverse mass distribution calculated for the W^+ and W^- samples in combination.

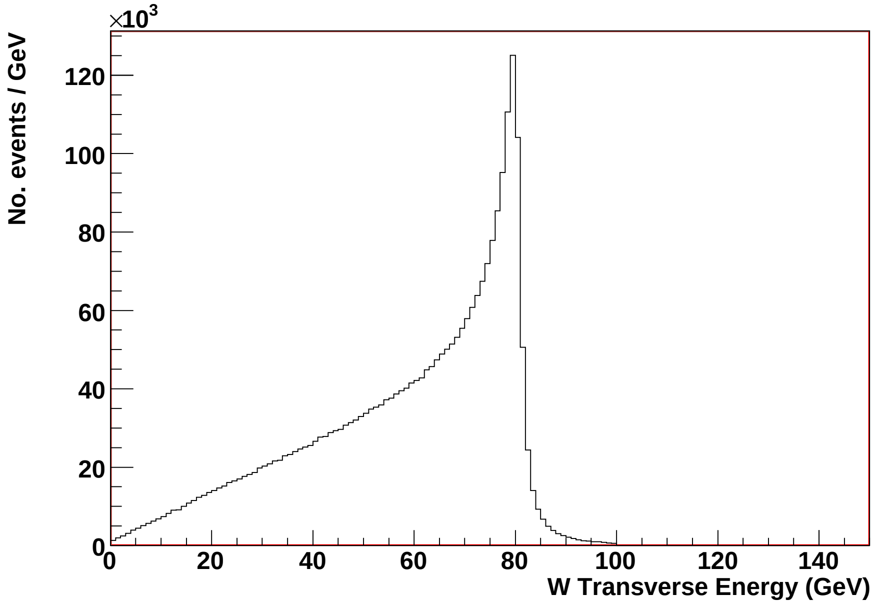
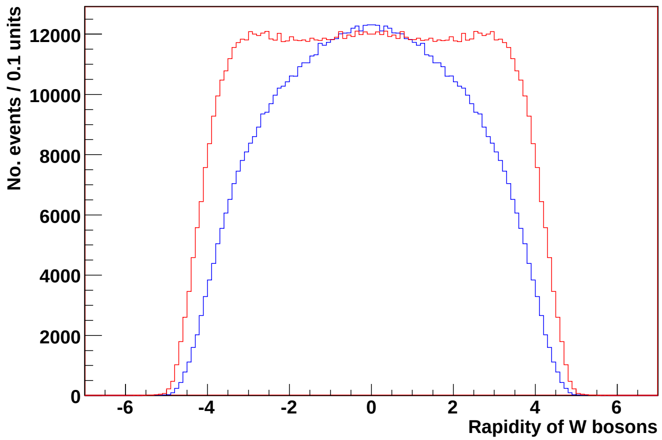
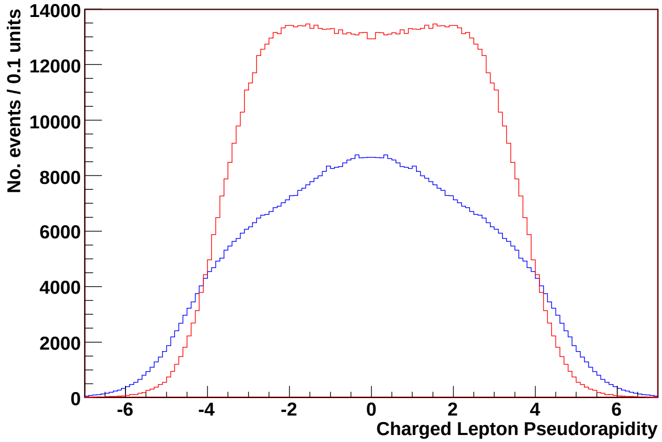


Figure 29: Total NLO W transverse mass for combined W^- and W^+ samples

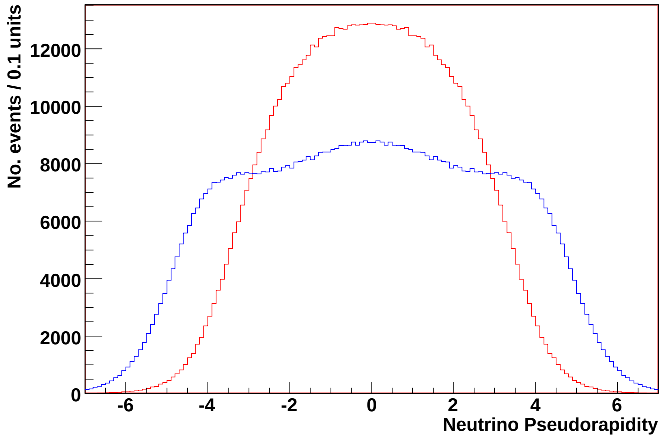
For reference the rapidities of the W^+ and W^- for the combined first and second order (NLO) channels and the total electron, positron, neutrino and antineutrino pseudorapidities are plotted in Figure 30.



(a) Total W^- (blue) and W^+ (red) rapidities



(b) Total electron (blue) and positron (red) pseudorapidities



(c) Total antineutrino (blue) and neutrino (red) pseudorapidities

Figure 30: Rapidity distributions of W bosons from the complete $W^- \rightarrow e^-\bar{\nu}_e$ (blue) and $W^+ \rightarrow e^+\nu_e$ (red) MC@NLO samples, and the pseudorapidity distributions of their leptonic decay products.

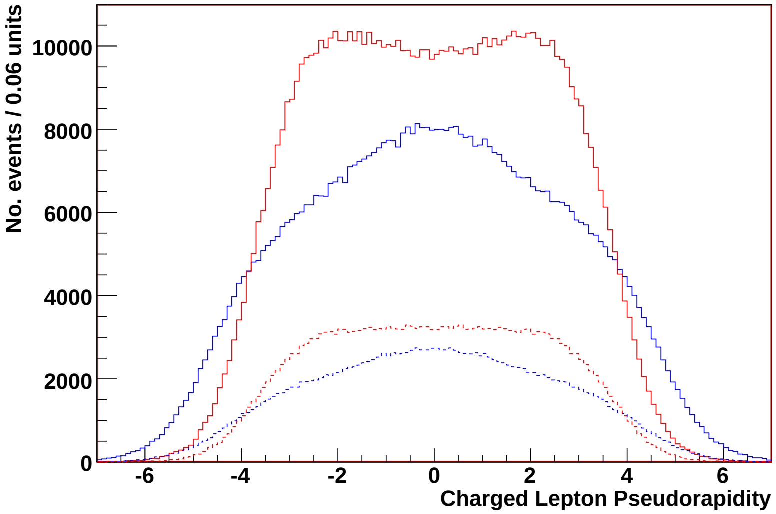
5.5 Effects of W boson charge asymmetry on selection of W Bosons

The ATLAS Technical Design Report suggests the following selection cuts for the reconstruction of W bosons ([74] Section 16.1.2):

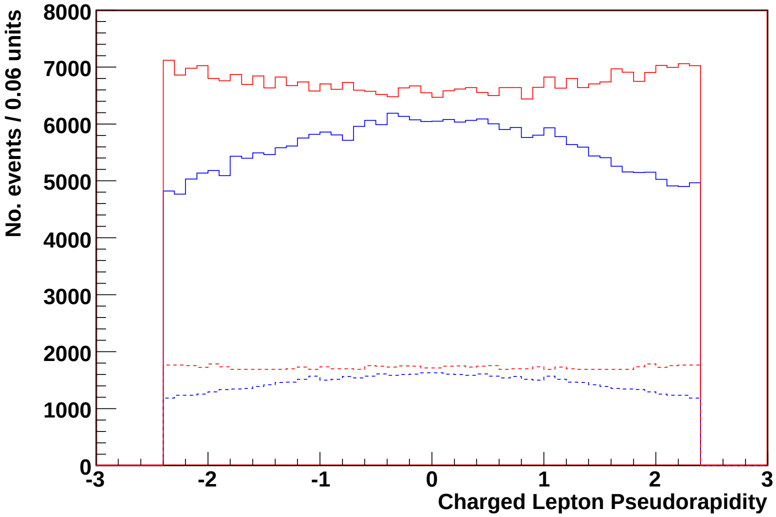
- An isolated electron with $p_T > 25$ GeV inside the region devoted to precision physics, $|\eta| < 2.4$.
- Missing transverse energy $MET > 25$ GeV.

Figure 31(a) compares the pseudorapidity distributions for the electrons and positrons from first and second order-produced W boson decay. Cuts of $|\eta| < 2.4$, $E_T > 25$ GeV were applied to the electron and positron, and a cut of $E_T > 25$ GeV to the ν or $\bar{\nu}$ produced with it (equivalent to a cut on MET of > 25 GeV). The same distributions were then plotted after the cuts, as shown in Figure 31(b).

Table 3 shows the number of events from each of first and second order production both before and after the W selection cuts for each of the W^+ and W^- samples. These are also represented as a percentage of the (remaining) sample. The percentage of each type of event that passes the W boson selection cuts is given. It can be seen that the first order-produced events are around 7% more likely to pass selection cuts than the second order. This is because the leptons produced in the decay of the first order-generated W bosons have a higher mean transverse energy than in the second order case. Electrons from the W^- boson decay are also on average 1.5% more likely to pass selection cuts than positrons from W^+ decay because they are generally produced at more central rapidities than the positrons.



(a) Generator level plots of electron (blue) and positron (red) pseudorapidity distributions before W boson selection cuts



(b) Generator level plots of electron (blue) and positron (red) pseudorapidity distributions after W boson selection cuts

Figure 31: Electron (blue) and positron (red) pseudorapidity distributions before and after W selection cuts applied at generator level (from samples of one million of each W^+ and W^-). The solid lines are leptons from first order-produced W boson decay, and the dotted lines from second-order W bosons.

Sample	**1st/2nd order**	**No. Events Before** W **Selection Cuts (%)**	**No Events After** W **Selection Cuts (%)**	**Fraction Passing Selection Cuts (%)**
W^+	1st	761709 (76.4)	323111 (79.6)	42.4 ± 0.1
W^+	2nd	234709 (23.6)	82787 (20.4)	35.3 ± 0.3
W^-	1st	762572 (76.6)	332488 (79.4)	43.6 ± 0.1
W^-	2nd	233019 (23.4)	86210 (20.6)	37.0 ± 0.3

Table 3: Numbers and percentages of $W^- \rightarrow e^-\bar{\nu}$ and $W^+ \rightarrow e^+\nu$ events (from samples of one million events each) which pass the basic W boson selection cuts listed above at generator level.

6 Studies on W Bosons in Full Simulation

The $W \rightarrow e\nu$ samples detailed in section 4 were produced at the request of the ATLAS Computing Systems Commissioning (CSC) group, which is responsible for testing the ATLAS computing model. These tests include the running of "Data Challenges", which involve the generation and then full simulation and reconstruction of large numbers of events of certain physics processes, representative of the types of physics that it should be possible and useful to study early on in the running of ATLAS. These MC@NLO samples, and a similar set of $W \rightarrow \mu\nu$ events, were provided for "Data Challenge 3". The CSC group ran a subset of these events through first HERWIG to complete the event generation step, and then the GEANT4 full detector simulation package, followed by digitisation and reconstruction. It was necessary to validate the samples to show that they were behaving as expected before they could be released to the physics community for analysis work. This section contains a small part of the validation performed on these fully-simulated samples.

The fully simulated samples provided by the CSC group were not large enough to perform the analysis on the strange sea that will be described in Section 7, but the results can be used to estimate the effects of the real detector on the final results of that study. A more complete validation study of this sample, both at generator level and in full simulation, and including the muon samples, is found in [75].

6.1 Sample Details

Prior to full simulation a generator level filter was applied to the electron as to select those with both $p_T > 10$ GeV and $|\eta| < 2.7$. The fully simulated samples studied here are designated as follows:

W$^-$→e$^-\nu$ sample: *csc11.005250.McAtNloWenu.recon.CBNT.v11004201* (49 Working Files)

W$^+$→e$^+\nu$ sample: *csc11.005254.McAtNloWplusenu.recon.CBNT.v11004205* (48 Working Files)

Table 6.1 details the number of events, cross-section and filter efficiencies for each data set. The centre of mass energy was held constant at 14 TeV. The version of the ATLAS software which was used for reconstruction in each case is also noted. There are often notable differences in datasets reconstructed in different major software versions since the reconstruction software is still heavily under development, but within a specific release branch (11.0.42 in this case) the differences in version between 11.0.42_1 and 11.0.42_5 are not large.

Process	**Reconstruction Software version**	**Filter Efficiency (%)**	**# Filtered Events**	**Cross-Section (nb)**
W$^+$→e$^+\nu$	11.0.42_5	63.2 ± 1.6	47800	1.76 ± 0.14
W$^-$→e$^-\nu$	11.0.42_1	65.8 ± 1.6	45850	2.18 ± 0.15

Table 4: Details of the fully simulated $W \rightarrow e^-\bar{\nu}$ and $W \rightarrow e^+\nu$ datasets

6.2 Electron Identification and Resolution

6.2.1 Selection of Electron Candidates

Table 5 lists the cuts, with their efficiencies, which were applied to select "good" electrons and positrons from the samples from which to reconstruct W bosons. Efficiencies are given for each cut for both the events which are simulated and reconstructed with the GEANT4 full detector simulation, and also for "detector truth", which shows what would be expected if the detector acted ideally and reconstruction was perfect. (N.B. In the plots which follow comparing the simulated and truth variables the truth plots are normalised to the same number of events as the fully simulated plots.)

	SIMULATED			**TRUTH**		
	Cut	**Efficiency (%)**		**Cut**	**Efficiency (%)**	
		W^+	W^-		W^+	W^-
1.	$\geq$ one identified e/γ candidate	95.84 $\pm$ 0.09	96.52 $\pm$ 0.09	$\geq$ one identified e/γ candidate	92.78 $\pm$ 0.12	93.80 $\pm$ 0.11
2.	$ET_e > 25$ GeV	66.06 $\pm$ 0.22	74.07 $\pm$ 0.20	$ET_e > 25$ GeV	65.14 $\pm$ 0.22	72.93 $\pm$ 0.21
3.	$\lvert\eta_e\rvert < 2.4$	63.42 $\pm$ 0.22	71.66 $\pm$ 0.21	$\lvert\eta_e\rvert < 2.4$	62.39 $\pm$ 0.22	70.52 $\pm$ 0.21
4.	IsEM=0 $\geq$ one assoc. track	45.76 $\pm$ 0.23	51.89 $\pm$ 0.23	Truth electron $\geq$ one true electron track	62.21 $\pm$ 0.22	69.96 $\pm$ 0.21
5.	$\Delta R_{[e,jet]} <0.7$	45.49 $\pm$ 0.23	51.64 $\pm$ 0.23	$\Delta R_{[e,jet]} <0.7$	61.47 $\pm$ 0.22	69.65 $\pm$ 0.21
6.	$E_e - E_{jet} < 5GeV$	45.02 $\pm$ 0.23	51.14 $\pm$ 0.23	$E_e - E_{jet} < 5GeV$	61.47 $\pm$ 0.22	69.65 $\pm$ 0.21

Table 5: Electron selection cut efficiencies: full simulation and detector truth

The cuts are explained in detail below:

1. There is at least one electron (e) or photon (γ) candidate reconstructed by the software. The Egamma reconstruction software uses information from the Electron Calorimeter to identify likely electron and photon candidates (see [76] for details.)

2. The transverse energy of this candidate must be at least 25 GeV for good reconstruction.

3. The candidate must lie in the rapidity window of the inner detector coverage, $|\eta| < 2.5$, so a pseudorapidity cut of $|\eta| < 2.4$ is applied here to enable the accurate reconstruction of tracks whilst avoiding edge effects.

4. In reconstructed data a flag, the IsEM flag, is set to 0 if a candidate is a "good" electron or photon. Potential e/γ candidates are selected by a set discrimination cuts based on ECal information as follows:

- Electromagnetic particles typically deposit a small amount of energy in the HCal. A property of "hadronic leakage" is defined as the ratio of the transverse energy recontructed in first compartment of the HCal within a window $\Delta\eta$ x $\Delta\phi = 0.2$ x 0.2 to the total transverse energy reconstructed in the Ecal. This is required to be $< 2\%$.
- There are a variety of cuts based upon the expected shape of the electromagnetic shower which are used to reject hadrons, including the decay of $\pi^0 \rightarrow \gamma\gamma$. Electrons and photons tend to deposit the majority of their energy in the second ECal layer. Various variables are reconstructed from information in this layer, the most significant of which is the "lateral shower shape". This is the ratio of the energy reconstructed in a 3 x 7 cluster to the energy in a 7 x 7 cluster around the seed (cluster reconstruction was explained in Section 3.5.2). Because electrons form narrow clusters this ratio should be close to 1 for electrons. The exact position of this and the other shower shape cuts depend upon the rapidity of the electron candidate. These are explained in detail in [77].

To remove photons there are a further set of cuts, based on ID information, which are applied.

- It is required that for each electron candidate identified by the software there should be at least one Inner Detector track associated with it.
- For a track to be considered a possible electron candidate the minimum requirements are that it has transverse momentum > 5 GeV and that it is composed of at least nine hits in the precision region of the ID (pixel detector and SCT), at least two of which are in the pixel detector, one of these being in the innermost layer, and that the impact parameter of the track is < 0.1 cm.
- The egamma reconstruction software then matches the electron candidate clus-

ter found in the ECal with any track within $\Delta\eta_{e,track} < 0.05$, $\Delta\phi_{e,track} < 0.1$ and $0.5 < \mathrm{E}_e/\mathrm{p}_{track} < 4$ of an energy deposit associated with an e/γ candidate in the calorimeter (where the granulatiry is $\Delta\eta_{e,track} = \eta_e - \eta_{track}$ and $\Delta\phi_{e,track} = \phi_e - \phi_{track}$ [††], see [39] section 7.2.2.4 for further details).

- It then assigns the electron candidate to the best matched track available within these constraints. This process is described in [76] (notice that for software versions 13 upward these default cuts will be looser).

See [78] for a more detailed explanation of electron selection. Note that for reconstruction software earlier than version 12 the cut based on TRT information should be ignored due to a bug in the software (version 11 is used here, therefore the TRT cut is not included).

The requirements are placed on $\Delta\eta$ and $\Delta\phi$ on the truth information for statistical comparison, where the both the electron candidate and the associated track is required to have come from a true electron.

5. This specifies that there should be a reconstructed jet (Section 3.5.5) in the same cone in $[\phi,\eta]$ space as the electron candidate such that

$$\Delta R_{[e,jet]} = \sqrt{[\eta_e - \eta_{jet}]^2 + [\phi_e - \phi_{jet}]^2} < 0.7 \qquad (23)$$

(using the “cone algorithm”) and requiring E_{Tseed} (energy of the seed) > 2 GeV, $E_{Ttotalcone}$ (energy of the cone) > 10 GeV (which will drop to 1 GeV and 7 GeV respectively in reconstruction software versions past 12.0.4). η_{jet} and ϕ_{jet} correspond

---

[††] where η_e is the pseudorapidity of the electron candidate calculated in the first sampling of the electromagnetic calorimeter where the granularity is very fine ($\Delta\phi$ x $\Delta\eta$ = 0.1 x 0.003), ϕ_e is the azimuthal angle of the electron candidate calculated in the second compartment (where $\Delta\phi$ x $\Delta\eta$ = 0.025 x 0.025). η_{track} is the pseudorapidity of the track calculated in the Inner Detector and extrapolated to the calorimeter, similarly for the azimuthal angle ϕ_{track}. E_e is the energy of the candidate in the Electromagnetic Calorimeter and p_{track} the momentum of the track calculated in the Inner Detector.

to the transverse-energy-weighted mean pseudorapidity and azimuth of all calorimeter cells belonging to the jet measured with respect to the reconstructed vertex. In the central rapidity region ($|\eta| < 2.5$) the hadronic calorimeter granularity is $\Delta\eta$ x $\Delta\phi = 0.1$ x 0.1.

This $\Delta R_{[e,jet]}$ distribution for all combinations of jets and the remaining electron candidates is shown in Figure 32.

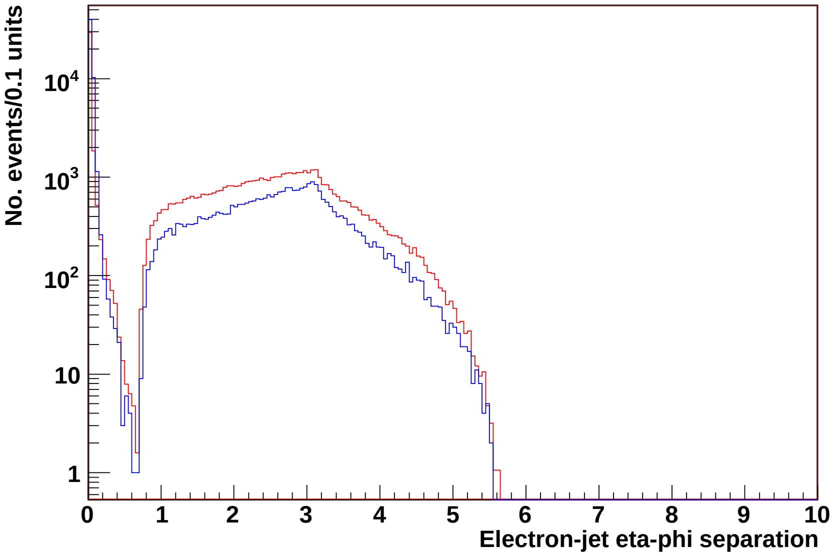
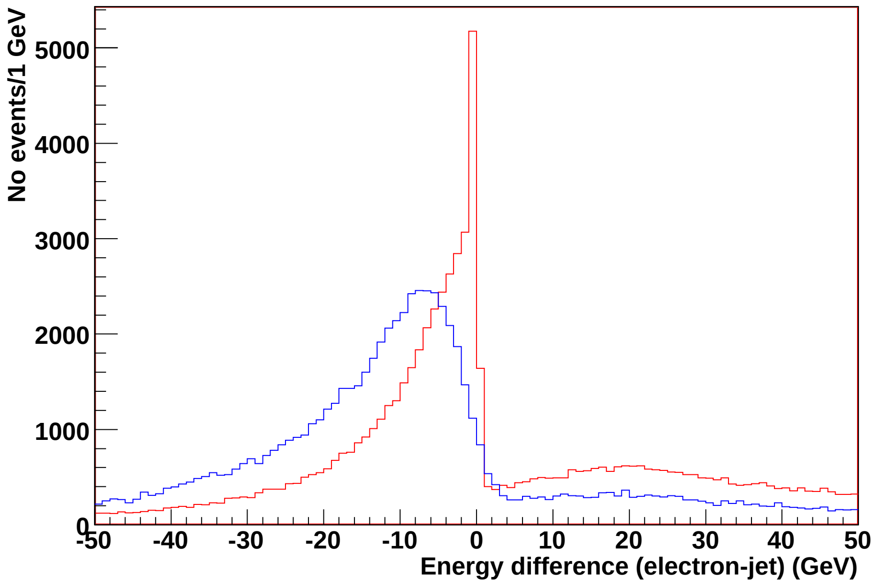
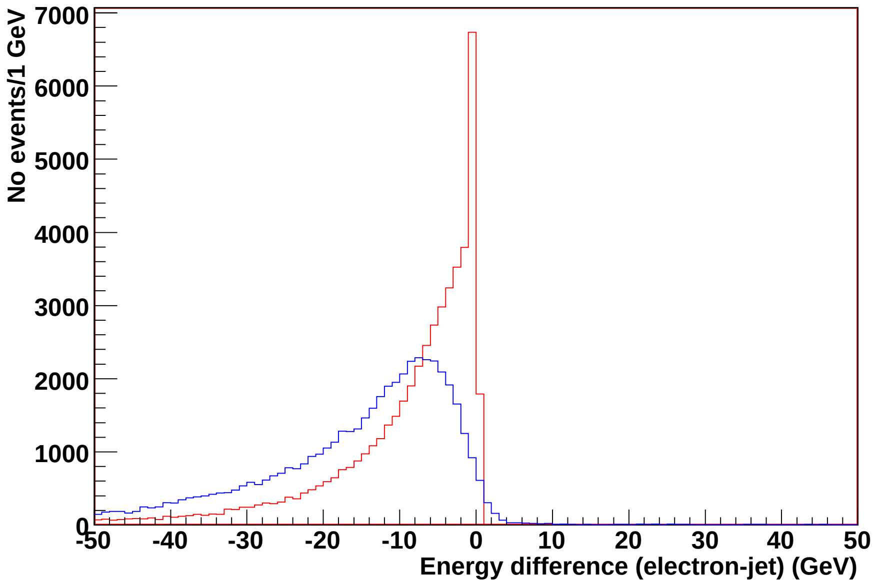


Figure 32: $\Delta R_{[e,jet]}$ for reconstruction (blue) and truth (red), for all combinations of an electron candidates and a jet.

The energy difference between the electron candidate and jet is also plotted in Figure 33. Figure 33(a) shows this distribution before the $\Delta R_{[e,jet]} < 0.7$ cut is made and Figure 33(b) after.



(a) $E_e - E_{jet}$ before $\Delta \mathrm{R}_{[e,jet]}$ cut



(b) $E_e - E_{jet}$ after $\Delta \mathrm{R}_{[e,jet]} < 0.7$ cut

Figure 33: $E_e - E_{jet}$ for reconstruction (blue) and truth (red), for all combinations of an electron candidates and a jet.

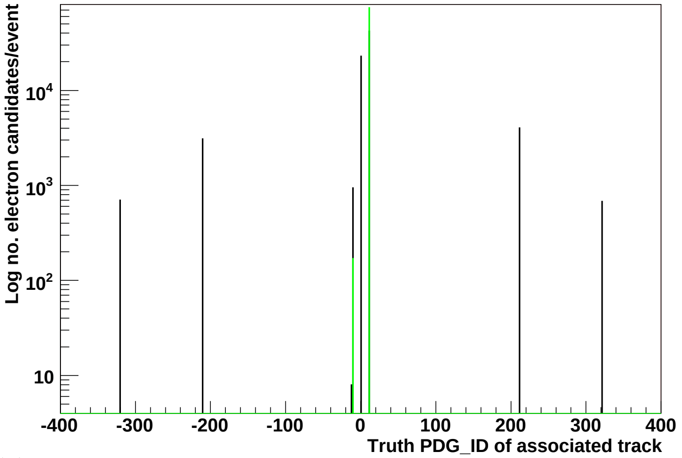
6. Figure 33 suggests that the jet matching cut can be refined further by placing a requirement on the energy difference between the electron and jet. Generally in the detector the energy of the reconstructed electron is less than that of the associated

jet, due to the nature of the corrections that are applied to the jet energy to account for energy having been lost from particles on the journey through the detector to the calorimeters. This is calibrated for hadrons and tends to be slightly over-corrected for electrons. Therefore a cut requiring the energy of the electron candidate to be no more than 5 GeV higher than a jet that represents the same electron is included to remove any electron/jet pairs that coincidentally match in ΔR.

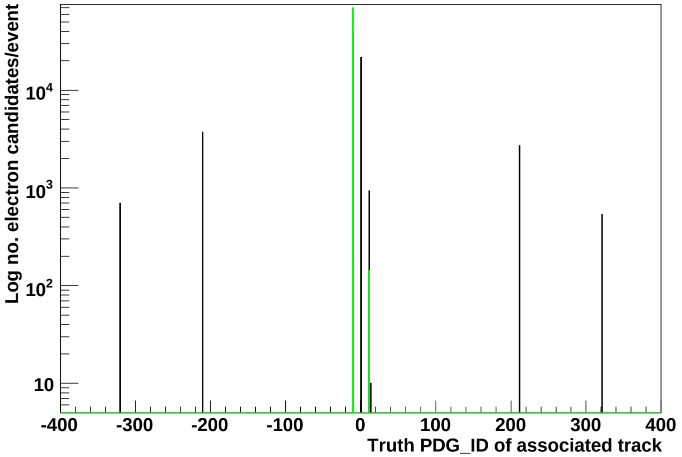
Table 5 shows that a higher percentage of electrons from W^- boson decays pass the selection cuts than positrons from the W^+ bosons. This is to be expected because the W^+ decays produce positrons at higher rapidity on average from the electrons from the W^- decays (as was discussed in Section 5.1).

Of the events passing the cuts from both samples 36 contain two electron candidates, although just one electron is seen in the truth. These are likely to result from hard Bremsstrahlung of the original electron from the decay of the W boson. In this process the electron decelerates rapidly in the detector, radiating a high energy photon which might then split into an electron-positron pair.

"Truth" information is available for each of the reconstructed tracks, giving a particle ID code for the type of generated particle the track is most closely associated with. Figure 34(a) shows the ID codes for the reconstructed tracks that are identified with the simulated electrons both before (black) and after (green) all of the electron selection cuts. Figure 34(b) shows the same distribution for the positron sample. These demonstrate that the selection code has removed all particles which are neither electrons nor positrons.



(a) tracks associated with electron or positron candidates in the W^- sample



(b) tracks associated with electron or positron candidates in the W^+ sample

Figure 34: Particle identification numbers of tracks identified with electron or positron candidates both before (black) and after (green) all electron/positron selection cuts. Particle IDs 11 are electron tracks, -11 are positrons, 12 are muons and particles with IDs over 100 are hadrons. ID 0 suggests that there is no true particle associated with the reconstructed track.

6.2.2 Electron Charge Identification

The direction of curvature of the track associated with a particle is used to extract its charge. For the W^- and W^+ samples, however, there are in total 105 "wrong-charge" electron candidates in the samples after the selection cuts listed in Table 5, i.e. the direction of curvature of some of the reconstructed tracks is opposite to what is expected for the sample type. The truth information associated with each reconstructed electron candidate is shown in Figure 34, demonstrating the positively-charged particles passing electron selection cuts are truly associated with positrons, rather than with electrons which have been charge-misidentified in the simulated detector. The events in which there were two passing electron candidates (36 events) account for some, but not all, of these.

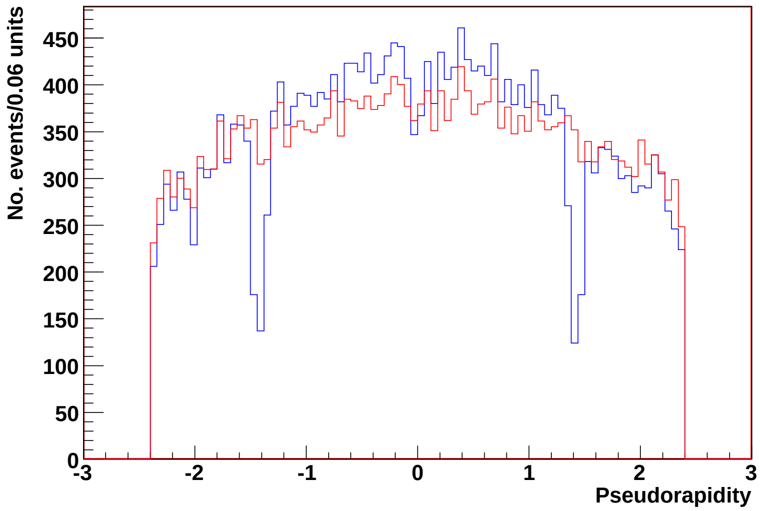
There are no positrons at the generator level or in the detector truth for the W^- sample, and no electrons in the W^+ so wrong-charge electron candidates must arise from detector interactions. The effect is mirrored to the same extent in the W^+ sample as the W^- which suggests that it is again likely to be a result of hard Bremsstrahlung. Considering the W^- sample, in this case one of the original electrons would emit a photon, which could take most of the energy of the original electron. If this then split asymmetrically into an electron-positron pair in which the positron was produced at a much higher energy than the electron then only the positron would pass the selection cuts. Many more electrons and positrons are likely to pass the selection cuts from asymmetric than symmetric Bremsstrahlung, since the transverse energy requirement is quite high at 25 GeV compared with the most likely transverse energy of the original electron (which is around 40 GeV, see Figure 36(b) later).

There are 105 events in which at least one wrong-charge electron or positron passed the selection cuts, out of a total of $\sim$45000. This amounts to the incorrect identification of the sign of electrons and positrons produced from the decay of the W bosons generated in this manner in $2.35 \pm 0.10\%$ of cases.

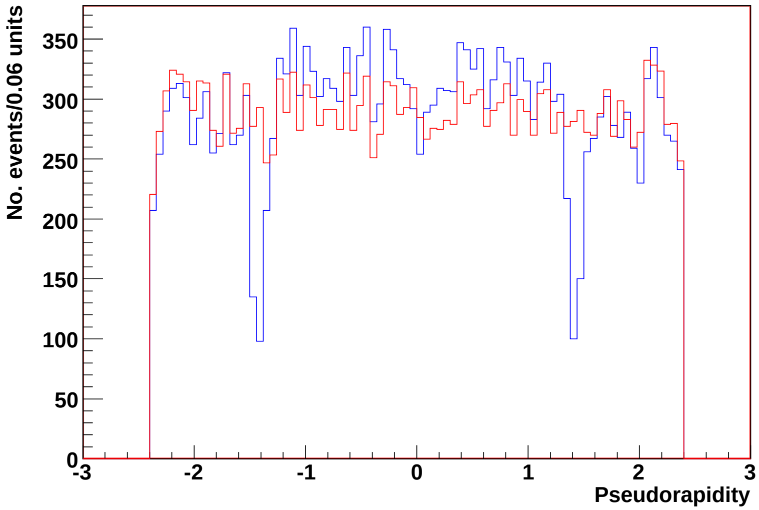
6.2.3 Electron η, ϕ, $\mathbf{E}_T$ and $\mathbf{p}_z$ plots after selection cuts

Figure 35 shows the pseudorapidities (η) of the electron candidates remaining after all the selection cuts were applied for each of the samples. The plots follow similar shapes to the electron/positron pseudorapidity plots shown at generator level in Figure 30(c). The detector cracks in the Electromagnetic Calorimeter at $|\eta| \approx 1.5$ can be seen in the reconstructed lepton rapidities. (Notice that because the truth plots are normalised to the same area under the graph as the simulated plots, the truth plots appear lower overall. This is because the detector cracks have cut out some simulated events that have not been removed in truth.)

Figure 36 shows the electron azimuthal angles (ϕ), transverse energies (E_T) and longitudinal momenta (p_z). As would be expected these plots look very similar for e^+ and e^- and so they are combined in this figure. The W boson decays isotropically in the transverse plane so the ϕ distribution shown in Figure 36(a) is flat. The E_T distribution seen in Figure 36(b) peaks just under 40 GeV which is again to be expected since the majority of the W bosons would be produced close to rest in the transverse plane and would thus split into an electron and neutrino with approximately equal transverse momenta, equating to roughly half of the W rest mass. The p_z distribution (Figure 36(c)) peaks at 0, reflecting the fact that the larger the component of electron momentum is in the transverse plane, the more likely the electron is to pass the p_T cut. Electrons travelling entirely longitudinally (i.e. down the beam pipe with total momentum $= p_z$) cannot be detected at all. The detector cracks at $\eta \approx 1.5$ are again visible as slight dips in longitudinal momenta roughly 60 GeV.

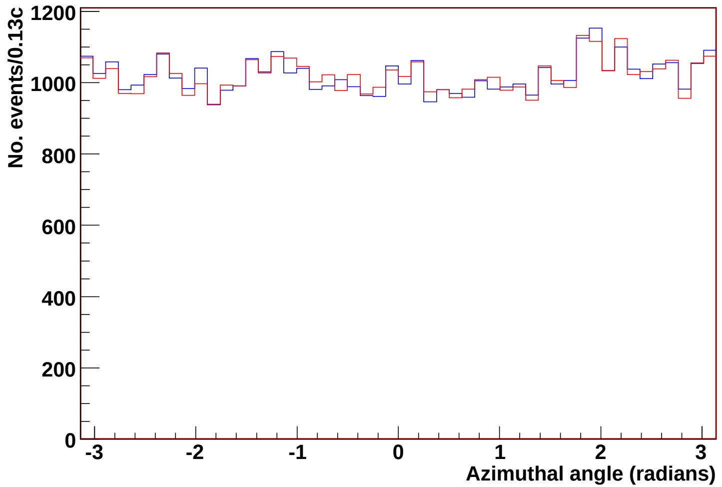


(a) e$^-$ pseudorapidity

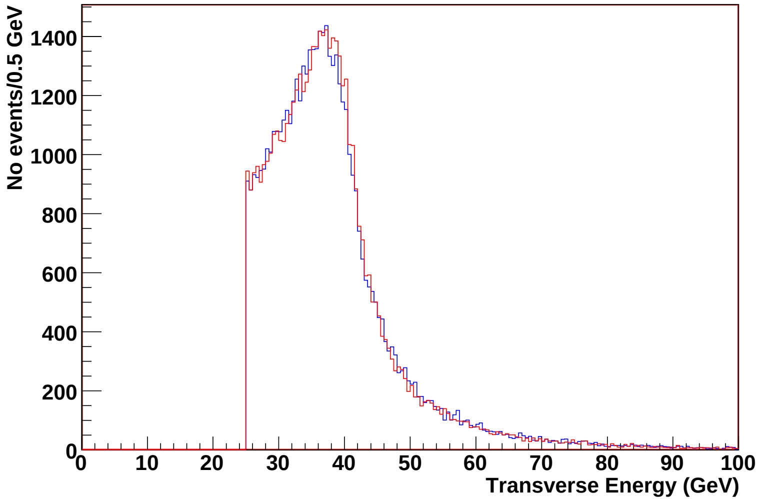


(b) e$^+$ pseudorapidity

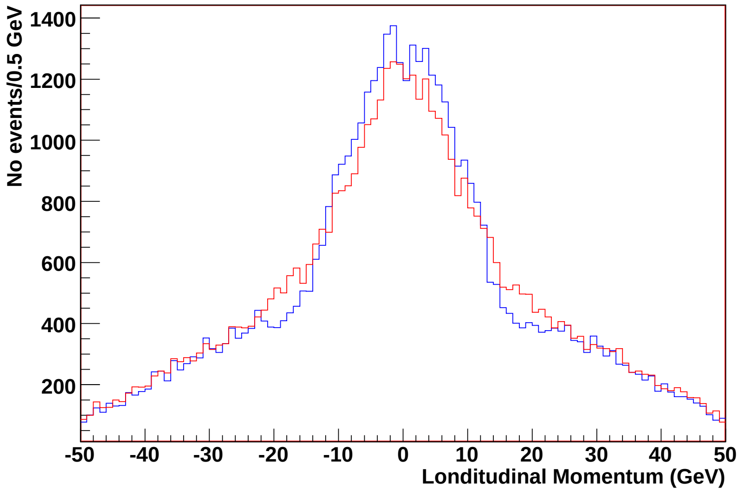
Figure 35: Electron pseudorapidity for reconstruction (blue) and truth (red) in the W^- (upper) and W^+ (lower) samples



(a) electron/positron azimuthal angle



(b) electron/positron transverse energy



(c) electron/positron longitudinal momentum

Figure 36: Electron kinematic variables from the combined W^+ and W^- samples plotted for reconstruction (blue) and truth (red)

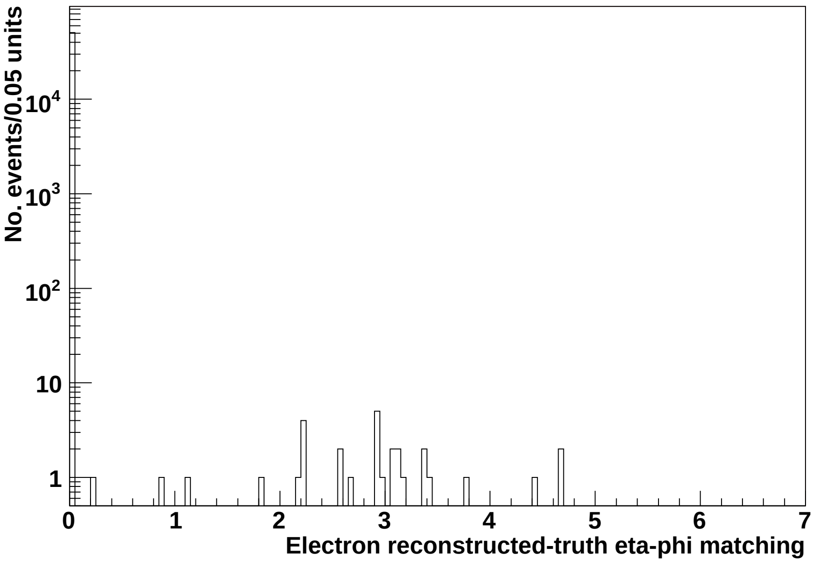
6.2.4 Spatial Resolution

In order to investigate spatial resolutions of the electrons a match is required between the reconstructed and true electrons for each event. For every reconstructed charged electron that passed selection its proximity in $[\eta, \phi]$ space to each truth electron with an energy of at least 20GeV found in the event record was calculated:

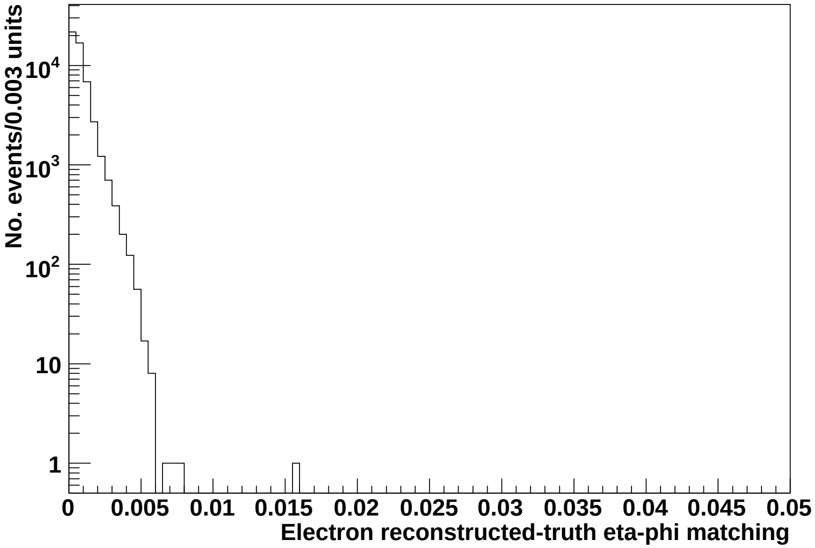
$$\Delta R_{[reconstructed,truth]} = \sqrt{[\eta_{reconstructed} - \eta_{truth}]^2 + [\phi_{reconstructed} - \phi_{truth}]^2} \qquad (24)$$

Electrons falling into the crack regions of the detector are excluded from all resolution calculations. These regions are taken to be $1.4 < |\eta| < 1.6$, which are areas of reduced coverage in the Electromagnetic Calorimeter. From this point on the W^+ and W^- samples are combined and “electron” is taken to mean electron or positron.

$\Delta\mathrm{R}_{[reconstructed,truth]}$ is plotted for electrons (on two different scales) in Figure 37.



(a) electron/positron $\Delta\mathrm{R}_{[reconstructed,truth]}$



(b) electron/positron $\Delta\mathrm{R}_{[reconstructed,truth]}$, detail

Figure 37: Electron spatial fractional resolution

The outliers to the electron spatial resolution distribution (i.e. those with ie. $\Delta R \geq\sim 0.01$ shown in Figure 37(a)) come from the events in which more than one reconstructed electron was found in the same event. Rejecting these events from the sample removed all of the outlying unmatched pairs. The events with only one wrong-charge electron candidate were retained. Since these were produced from hard Bremsstrahlung followed by a very asymmetric photon conversion, momentum considerations indicate that the wrong-sign electron would follow a very similar path to the original electron produced in the W decay.

The spatial resulution can be seen to be better than $\delta R \simeq 0.005$, which is comparable with the calorimeter granularity given in Section 6.2.1.

6.2.5 Electron Transverse Energy Resolution

Figure 36(b) shows the transverse energy distributions of the electrons. The transverse energy fractional resolution σ is plotted in Figure 38 by comparing the transverse energy of the reconstructed electron as calculated from the Electronic Calorimenter (ET_{reco}) with that of its truth counterpart (ET_{truth}) as follows:

$$\sigma = (\frac{(ET_{reco} - ET_{truth})}{ET_{truth}}) \tag{25}$$

A Gaussian was fitted to the distribution in Figure 38. There is a negative non-gaussian tail on this distribution from electrons which lost a fraction of their energy by Bremsstrahlung when passing through the detector, this tail is not included in the Gaussian fit. The fit is improved by also removing a small amount of the positive tail where the statistics are low. The fitted region is from -0.025 to 0.1%. From this the electron transverse energy resolution is found to be 1.8% and it peaks very close to 0, at -0.0013 ± 0.0001, implying that the simulated Electromagnetic Calorimeter is well calibrated.

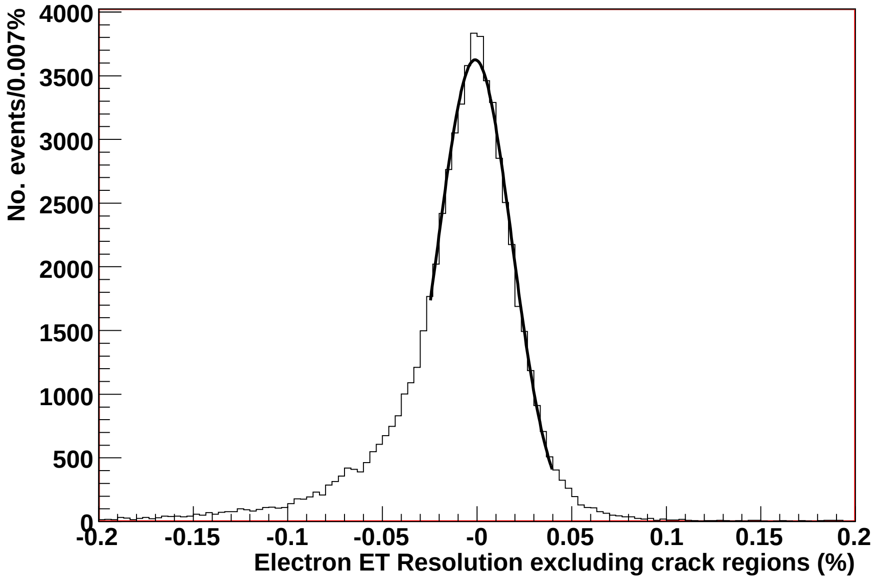


Figure 38: Electron/positron ET resolution: (sim-truth)/truth

The energy resolution of the Electromagnetic Calorimeter is expected to vary with particle energy according to

$$\frac{\sigma}{E} = \frac{a^2}{E} \oplus b^2 \oplus \frac{c^2}{E^2}$$

as discussed in Section 3.2.2 (the fit to transverse energy should be of the same form), and also with rapidity. For illustration, Figure 39 shows the resolution in two different windows of electron ET. It can be seen that the resolution is better at higher electron energies.

The electron candidates were divided into groups in 5 GeV windows of transverse energy, and the transverse energy resolution for each group was calculated as before, by fitting a Gaussian to the resolution curve for each group. Figure 40 shows the transverse energy resolution for electrons selected within 5 GeV windows of ET as a function of the median ET of the window. For this resolution study all electrons of 10 GeV and over and that pass all the other selection cuts listed in Table 5 are included (although elsewhere only electrons with transverse energies of 25 GeV and above are used).

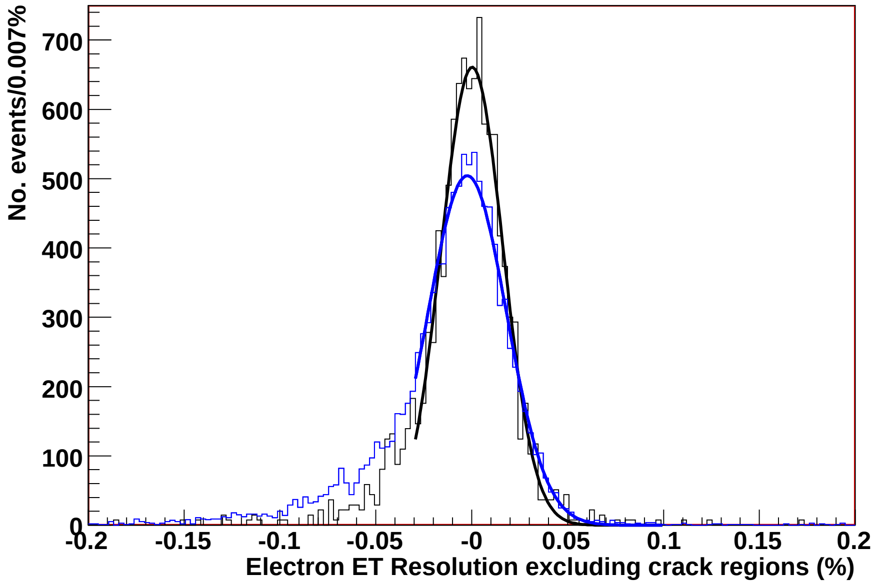


Figure 39: Electron ET resolution in the ET range 50-55 GeV (black) and 30-35 GeV (blue, normalised to same area as black plot).

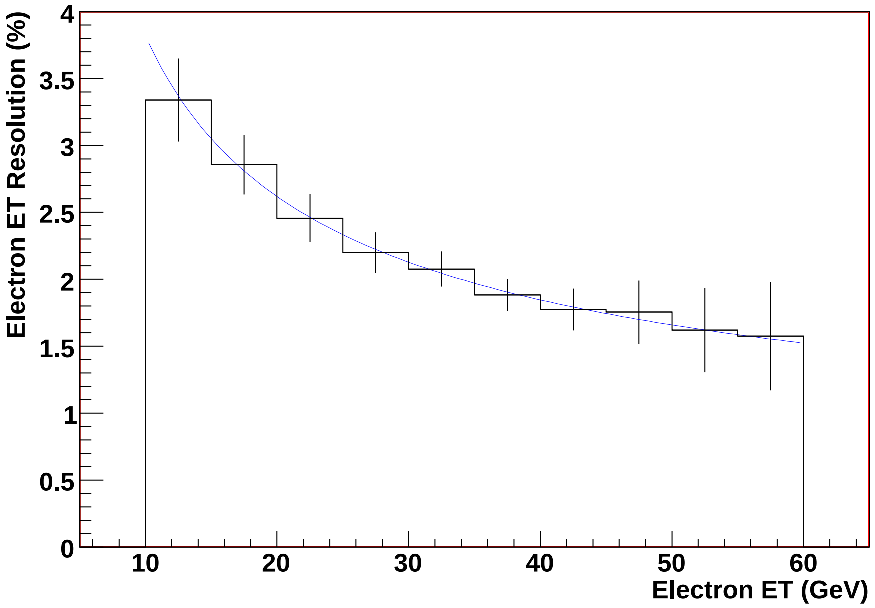


Figure 40: Electron ET resolution as a function of energy resolution (black). ET distribution is shown for shape comparison (dotted) and the fitted function (blue)

Fitting the above function to the available data for transverse energy the constants are found to be a=11.0 ± 4.3%$\sqrt{GeV}$, b=0.5 ± 1.8%, c=15 ± 38%GeV. For the number of events available this cannot be measured with greater accuracy (error bars are plotted where the error in σ is given by $\sqrt{\frac{\sigma(1-\sigma)}{N}}$, where N = number of electrons in the group). Comparison with previous studies (for example [79, 80, 81, 82, 83]) suggests these numbers are within the expected limits, and they fit within the stated ATLAS design requirements noted in Section 3.2.2.

This suffices as a proof of principle for validation purposes. For more accurate results larger samples would be needed, the mean rather than median ET of each range in transverse energy would be used for the data points, and the fit should be performed at different specific rapidities. [79] discusses in detail the effects of rapidity and materials on the measurements.

6.3 Reconstruction of Missing Energy

Figure 41 shows the MET spectrum for all events that pass the W selection cuts. It can be seen from these plots that MET is not well calibrated in this version of the reconstruction software. The reconstructed distributions vary in shape significantly from the truth.

6.3.1 Missing Transverse Energy Resolution

The resolution of the missing transverse energy is given by:

$$Resolution = \sigma(MET_{reco} - MET_{truth}) \tag{26}$$

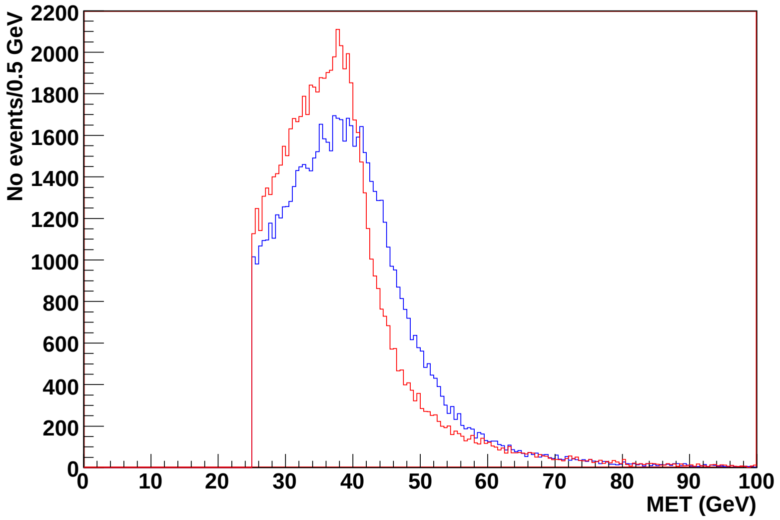


Figure 41: Missing transverse energy distributions for reconstruction (blue) and truth (red)

It is found by plotting the difference between the values of the reconstructed and true MET for every event and fitting a Gaussian curve to the distribution and finding the width. This distribution is shown in Figure 42, with a Gaussian fitted to the central region of the distribution, between -50 and 85 GeV. From this the resolution of the missing transverse energy in this sample is found to be 5.32 ± 0.06 GeV. The Gaussian is centred at 1.69 ± 0.04 GeV ($\chi^2 = 61.12$ with 38 degrees of freedom), demonstrating that the reconstructed MET is not well calibrated to the true value in this version of the detector simulation software, as was seen in Figure 6.3.1.

The Gaussian fit does not match the resolution curve well away from the central region. This is because of the different responses of the calorimetry to low and high energy particles. More accurately, the MET resolution of the detector is expected to vary with the recoil transverse energy of the event, where the recoil transverse energy (ΣE_T) in this case is the sum of the energies of all the particles in the event except for the electron from the W boson decay (taken to be the electron with the highest transverse energy).

Electrons leave energy in the form of jets in the electromagnetic calorimeters, so to find the recoil in this case the energy of the jet which is matched in $[\eta, \phi]$ space to the electron ($\Delta R_{[e,jet]} < 0.7$, as in Section 6.2.1) must be subtracted from the total calorimeter energy.

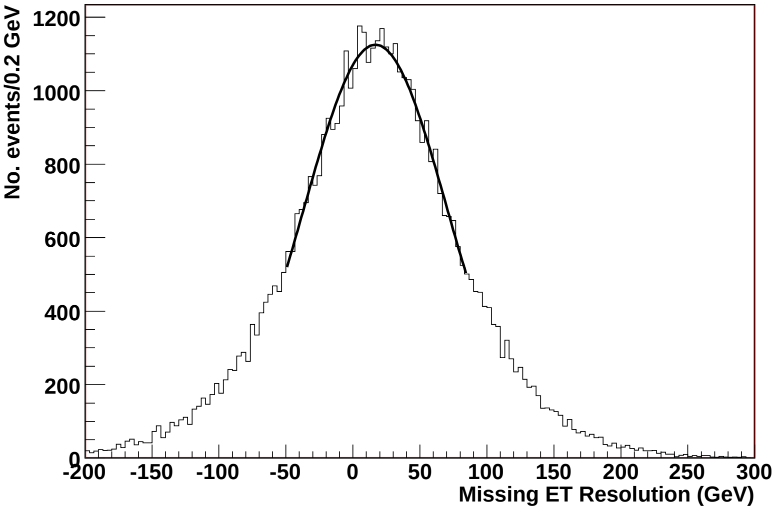


Figure 42: MET resolution

The samples were then divided into sets with recoil transverse energies in 10 GeV windows (the first being 0-10 GeV, the second 10-20 GeV etc.), and the MET resolution was calculated for each window of events. Figure 43 shows the resolution in two different windows of ΣE_T.

The resolution for each window of ΣE_T was plotted against the median value of ΣE_T for the window in each case as shown in Figure 44. The shape of the total recoil tranverse energy is also plotted for comparison (not to scale).

Missing transverse energy resolution is expected to follow the form:

$$\sigma_{MET}(\Sigma E_T) = p_0 \sqrt{\Sigma E_T} \tag{27}$$

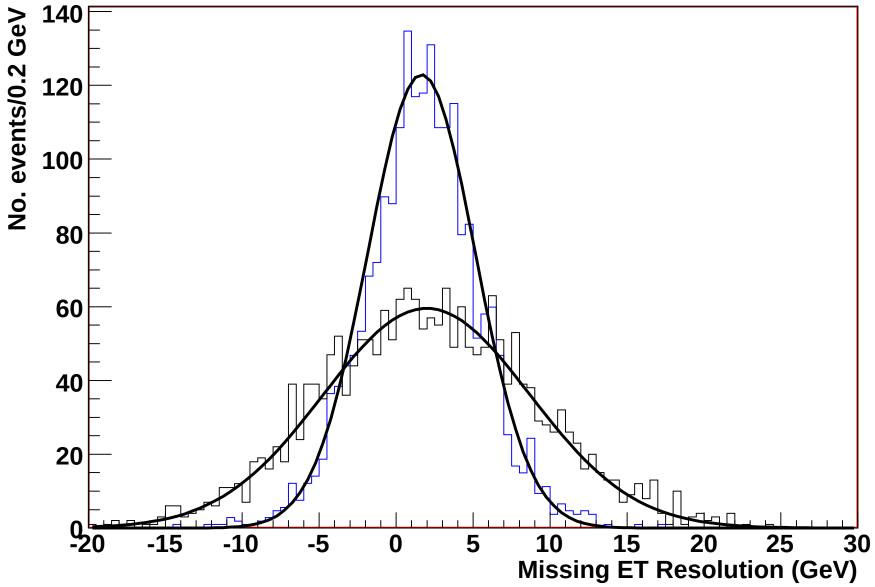


Figure 43: MET resolution in the recoil transverse energy range 120-130 GeV (black) and 10-20 GeV (blue, normalised to same area as black plot). For the 120-130 GeV plot the MET resolution is 6.72 ± 0.11 GeV, the Gaussian fit is centred at 1.98 ± 0.15 GeV ($\chi^2 = 77.3$ with 80 degrees of freedom). For the 10-20 GeV the MET resolution is 3.50 ± 0.06 GeV, the Gaussian fit is centred at 1.65 ± 0.76 GeV ($\chi^2 = 66.3$ with 52 degrees of freedom).

Where p_0 is an arbitrary parameter ([84]) and σ_{MET} is the resolution of the missing transverse energy. In order to test this a second order polynomial was fitted to the MET resolution plots in Figure 44 such that:

$$\sigma_{MET}(\Sigma E_T) = A + B(\Sigma E_T)^C \tag{28}$$

where the constant term is added to allow for the MET mis-calibration. The best fit polynomial for the electron sample was found to be given by:

$$\sigma_{MET,e}(\Sigma E_T) = (2.000 \pm 0.230) + (0.013 \pm 0.011)(\Sigma E_T)^{0.5 \pm 0.1} GeV \tag{29}$$

which is in the form of a square root, as expected ($\chi^2 = 36.5$ with 13 degrees of freedom).

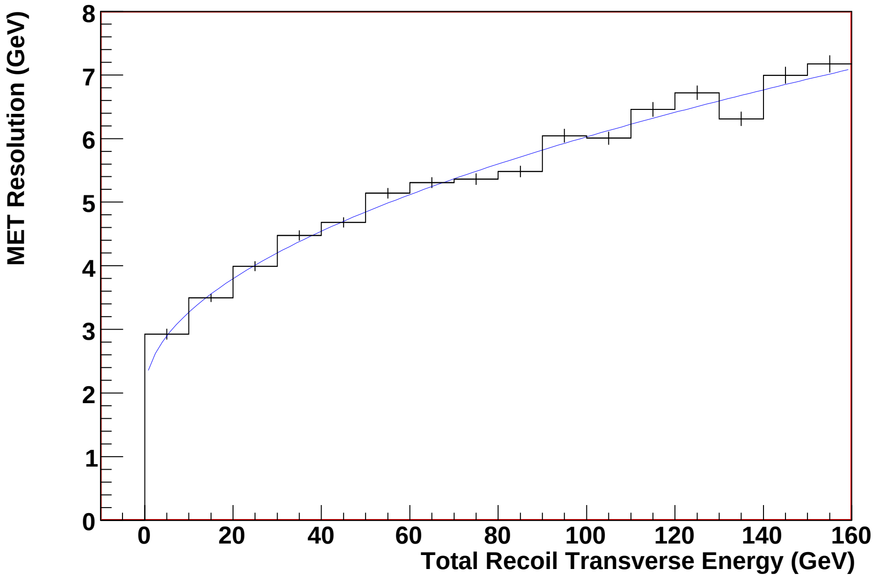


Figure 44: MET resolution as a function of total recoil transverse energy ΣE_T in the electron/positron samples. A second order polynomial is fitted to the MET resolution distribution. ΣE_T is plotted in the background (not to scale) as a dotted line.

6.3.2 Correlation between electron and missing energy momenta

One aspect of MET calibration in the case of $W \rightarrow e\nu$ reconstruction is the consideration of how well the MET vector direction correlates with the direction of the electron transverse momentum vector. The properties of the MET when its vector is resolved first in the direction perpendicular to the lepton line of flight and then parallel to it are studied in this section.

Resolving MET perpendicular to the direction of electron line of flight: In directions perpendicular to the momentum of the electron we expect the component of missing transverse energy to be zero, because at first order the neutrino will be emitted with equal and opposite transverse momentum to the electron so there will be no missing energy perpendicular to the lepton flight path (at generator level). Even in the higher order cases helicity considerations lead the lepton and neutrino to be emitted preferentially in opposite directions and so a sharp peak around zero is expected from the second order diagrams, as seen in Figure 22(c).

These distributions will experience Gaussian smearing due to the imperfect detector resolution, so a narrow Gaussian peak is expected around zero from the first order events, and a wider one from the second order events. This overall shape can be seen in Figure 45, which shows the component of missing transverse energy resolved in the direction perpendicular to the electron line of flight. Note that the distribution of W boson transverse momenta in the second order case is not itself Gaussian (it falls away more rapidly from the narrow central peak rapidly in a way suggestive of a $\frac{1}{x}$ curve, see Figure 22(c)) and so in the second order case this will be very approximate. Again the truth (red) distribution is significantly different in shape to the reconstructed (blue), as the inaccurate reconstruction that was seen in Figure 41 has somewhat washed out the features of the distribution.

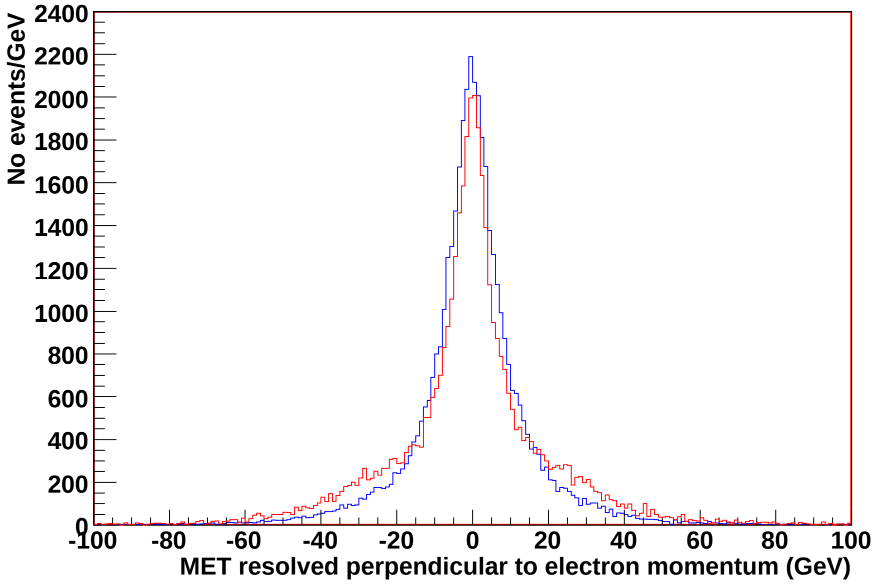


Figure 45: Missing transverse energy resolved perpendicular to the transverse momentum of the electron for reconstruction (blue) and truth (red).

As explained above, the plots shown in Figure 45 could be expected to follow the very approximate form of the sum of two Gaussians (the wider from second order W production, and the narrower from first order). Figure 46 shows a double Gaussian fitted to the MET perpendicular to the electron line of flight for the reconstructed events. The fit appears fairly good, suggesting that this variable could be used to differentiate between leading order and higher order W production.

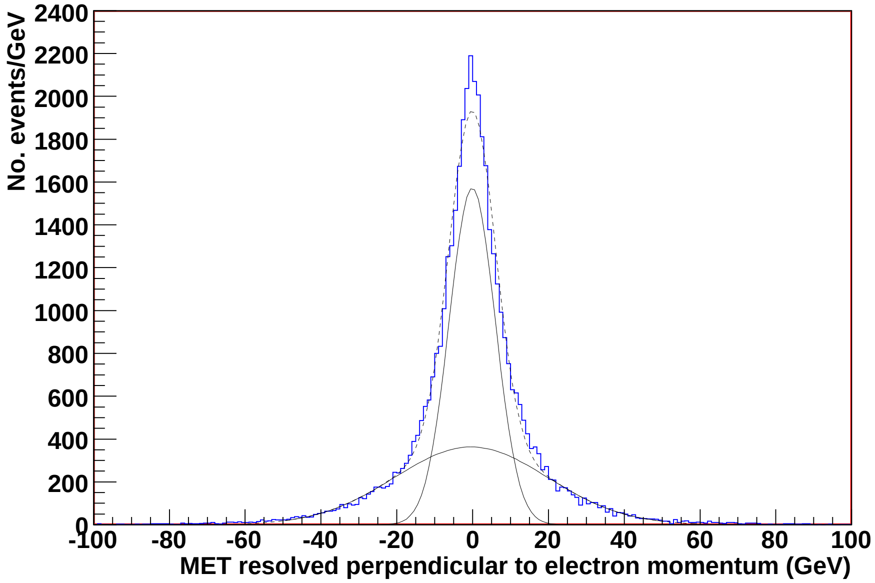


Figure 46: Missing transverse energy perpendicular to the transverse momentum of the electron. The fit shown (dotted) is the sum of two Gaussians, which are also plotted individually (continuous)

The double Gaussian fit is as follows, where the first of the two Gaussians corresponds to the larger first order peak:

$$Ae^{\frac{-(x-B)^2}{2C^2}} + De^{\frac{-(x-E)^2}{2F^2}} \tag{30}$$

Where A=1588 ± 16 events/0.02 units; B=-86 ± 53 MeV; C=6269 ± 78 MeV; D=329 ± 10 events/0.02 units; E=-495 ± 170 MeV; F=21210 ± 240 MeV. These are both centred close to 0 as would be expected. The area under a Gaussian curve centred at 0, Ke^{-ax^2}, is given by

$$\int_{-\infty}^{\infty} dx K e^{-ax^2} = K\sqrt{\frac{\pi}{a}} \tag{31}$$

The areas under the first and second order Gaussian curves shown in Figure 46 were calculated (in the former case $K = A$, $a = \frac{1}{2C^2}$ and in the latter $K = D$, $a = \frac{1}{2F^2}$) and the ratio of the area under the first and second order curves was found to be ~ 1.4.

Table 3 in Section 5.5 shows the fraction of events in the W samples that passed the basic W boson selection cuts (electron $|\eta| < 2.4$ and $p_T > 25$ GeV, MET > 25 GeV) at generator level. It indicates that approximately four times as many events should pass the W selection cuts from the first order generation of W bosons than second order for both electrons and positrons (with some variation between W^+ and W^-). The cuts applied at generator level are comparable to the basic criteria for events passing W selection in full simulation, and so the ratio between the numbers of first and second order events that should remain after the W selection criteria in full simulation should be similar.

There is therefore a large discrepancy between the ratio of the number of electrons falling into each of the fitted Gaussian peaks in Figure 46 and the ratio of the numbers of electrons passing the selection cuts that were generated in each of first and second order production of W bosons. This is not unexpected since the resolution of the electron transverse energies (Figure 38) has a significant non-Gaussian tail which depends on the energy of the electrons, and the MET resolution is also not purely Gaussian, but varies with the energy of the event (Figure 42).

A more comprehensive study on the various detector effects involved would be needed to produce a more accurate fit to the distribution of the missing transverse energy in the plane perpendicular to the lepton line of flight for each of first and second order W

production, should it be needed to reject backgrounds of first order-produced W (or Z) bosons in data, but for the following analysis it will be seen not to be required.

Resolving MET parallel to the direction of electron line of flight: Similarly it would be expected that, resolved in the direction parallel to the charged lepton momentum, the missing transverse energy should be on average equal and opposite to the electron transverse energy. Analagously to the perpendicular case discussed above, this would be exactly true in the first order case at generator level since the W would be produced with no transverse energy. If the lepton p_T and missing p_T resolved in the direction parallel to the lepton line of flight were summed a Gaussian distribution sharply peaked at 0 in the first order case and less sharply in the second order case should result. Figure 47 shows the sum of the lepton transverse energy and the missing transverse energy measured in the direction parallel to the transverse momentum of the electron (the positive direction is taken as the direction in which the electron is travelling).

In this case, the reconstructed distributions peak a few hundred MeV below zero due to the mis-calibration in missing energy reconstruction (seen in 6.3.1), although the truth peaks around 0 as expected. Both the true and reconstructed versions also have larger positive than negative tails which implies that the lepton has a slightly higher transverse energy than the missing energy in general. This is again due to over-correction of the lower energy electrons for the effects of Bremsstrahlung in the detector during reconstruction.

A similar double Gaussian fit could be performed here as above but due to the miscalibration the Gaussians would be centred far from zero and would thus be a less convincing measure of the type of production process than the distributions shown in Figure 45.

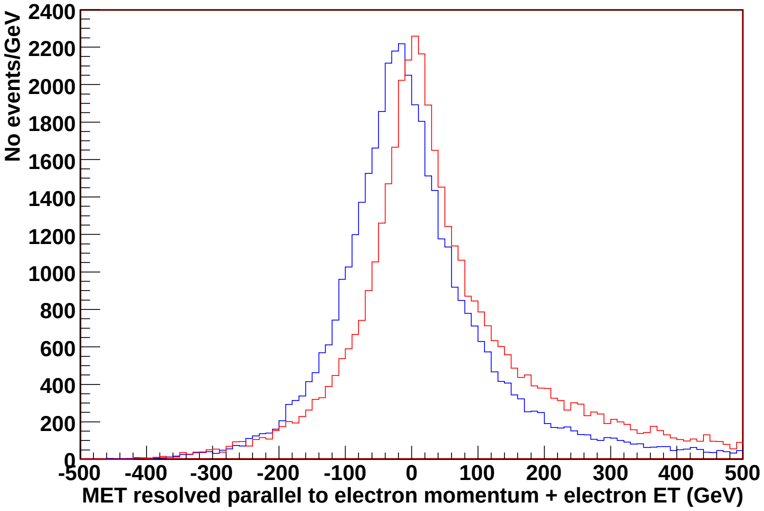


Figure 47: The sum of the electron transverse energy and missing transverse energy in the direction parallel to electron momentum for reconstruction (blue) and truth (red)

7 Signal Events in Fast Simulation

The channels $s + g \rightarrow W^- + c$ and $\bar{s} + g \rightarrow W^+ + \bar{c}$ were used to study the strange sea (as discusssed in Section 2.3). This section details the method used to reconstruct the W bosons and charm quark produced in these channels.

7.1 Preliminary $W + D^*$ Selection

The W boson was identified by its decay into an electron, which was reconstructed in the detector as in Section 3.5.3, and a neutrino, the presence of which was inferred by the existence of a significant amount of missing energy (Section 6.3) in the detector.

As was discussed in Section 2.3 the charm quark was identified in the case that it was bound into a D^* meson during hadronisation. This was followed by a decay into a D^0 meson and a bachelor pion, and the subsequent decay of the D^0 via the channel $D^0 \rightarrow K\pi$.

Momentum considerations indicate that the W boson, and thus the resulting electron, are likely to be found in a different region of the detector to the jet produced by the charm quark. As a result, an isolation criterion can be applied to the electron, stating that it should not be found close to or within any hadronic jets. Also there would usually be only one energetic electron and one high-energy jet in such events.

Table 6 lists the decay chains of the W boson and charm quark together with the (charge-symmetric) branching fractions for each decay, which will be used to search for the signal events in the cases of an intial strange quark or antistrange quark interacting with a gluon. (All branching fractions are taken from [8].)

Initial Interaction	$s + g \rightarrow c + W^-$	$\bar{s} + g \rightarrow \bar{c} + W^+$	**Branching Fraction** (%)
W decay mode	$W^- \rightarrow e^- \ \bar{\nu}$	$W^+ \rightarrow e^+ \ \nu$	10.7 ± 0.1
Charm decay chain	$c \rightarrow D^{*+}$	$\bar{c} \rightarrow D^{*-}$	25.5 ± 1.7
	$D^{*+} \rightarrow D^0 \ \pi^+$	$D^{*-} \rightarrow \bar{D}^0 \ \pi^-$	67.7 ± 0.5
	$D^0 \rightarrow K^- \ \pi^+$	$\bar{D}^0 \rightarrow K^+ \ \pi^-$	3.80 ± 0.09

Table 6: Branching fraction for the decay channels used for signal reconstruction

To identify the events containing the signal channel in the dataset the following method is used to select the W bosons and charm quarks:

W Boson Selection:

1. One electron in the event, produced with transverse momentum > 25 GeV.

2. Missing energy in the event such that MET > 25 GeV

3. The electron was found in the pseudorapidity range $|\eta| < 2.4$. The inner detector tracking range is $|\eta| < 2.5$, so this cut ensures that the electron is in the tracking region and away from edge effects which could reduce the accuracy of the measurement of its momentum.

4. The electron was *isolated*, i.e. not found in the same immediate region of the detector as a jet. This is tested in ATLFAST (by default) in the following way: if there is an electromagnetic cluster within a radius of $\Delta\text{R}<0.15$ of the projected track, where

$$\Delta\text{R}_{[e,cluster]} = \sqrt{[\eta_e - \eta_{cluster}]^2 + [\phi_e - \phi_{cluster}]^2}$$

   then this cluster is assumed to be associated with the electron. If there is another cluster within $0.15 < \Delta R < 0.4$ of the reconstructed electron it is not isolated and fails the cut. (This cut will be very important in the rejection of physics backgrounds to the signal involving semileptonic decays of hadrons within quark or gluon jets.)

Of the W^+ sample 43.24 ± 0.03% of the generated events passed these selection cuts, and 45.19 ± 0.03% of the W^-.

Charm Selection:

1. Two tracks were selected from the event. It was assumed that one corresponded to a kaon and one a pion with the appropriate masses. Their track momenta and these masses were used to assign each track an energy (using Equation 8), and then the track momenta and assumed energies were combined to reconstruct a D^0 candidate. The mass of the D^0 candidate would be close to the true D^0 mass if the two tracks selected were indeed associated with a kaon and pion from a real D^0 decay.

   This process was repeated for every pair of tracks in the event (discounting the tracks already associated with electrons), producing two D^0 candidates for each pair of tracks (one is produced where the first track is assumed to be produced by a kaon and the second a pion, then another when the second track is assigned to represent the pion and the first the kaon).

   In order to reduce the number of calculations needed, a very loose requirement was placed on the mass of each D^0 candidate such that it must be within 200 MeV/c^2 of the true D^0 mass (1864.6 MeV/c^2) before continuing.

2. A third track was now taken to represent the π_B and combined with the D^0 candidate in the same way in order to produce a D^* candidate. The mass of the D^* candidate was calculated. Again this process was repeated, combining the properties calculated for every D^0 candidate with every other track in the event, leading to a very high multiplicity of D^* candidates in each event.

3. At the LHC there will be very large numbers of low p_T particles present for each event. Figure 48 shows the transverse momentum distribution plotted for all recon-

structed tracks. $p_T = 0.5$ GeV is the standard transverse momentum threshold for reconstruction of charged particle tracks at ATLAS. This is because below this the effects of multiple interactions mean the uncertainties in measuring the track momentum become too large for the tracks to be useful, and also the track curvature becomes too large for the tracking algorithm to identify. Therefore in ATLFAST only charged particles with (true) $p_T > 0.5$ GeV are reconstructed in the inner detector.

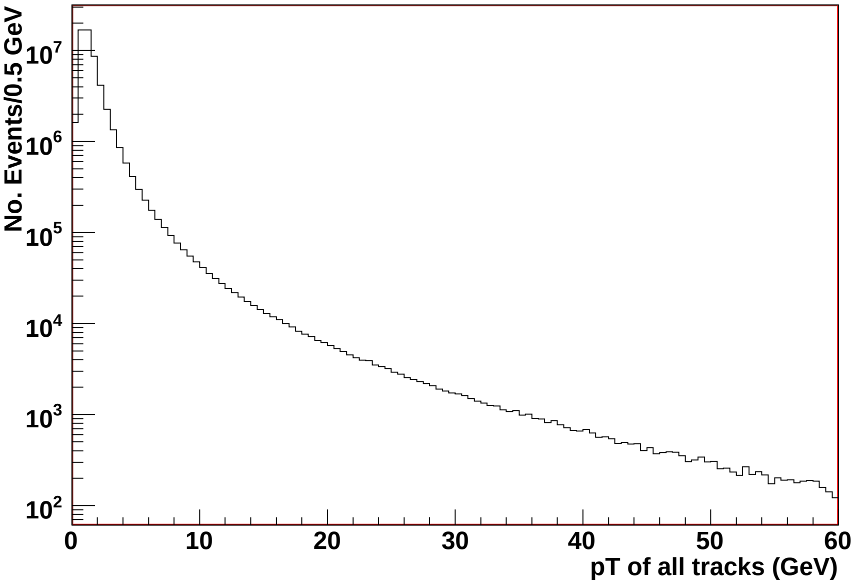
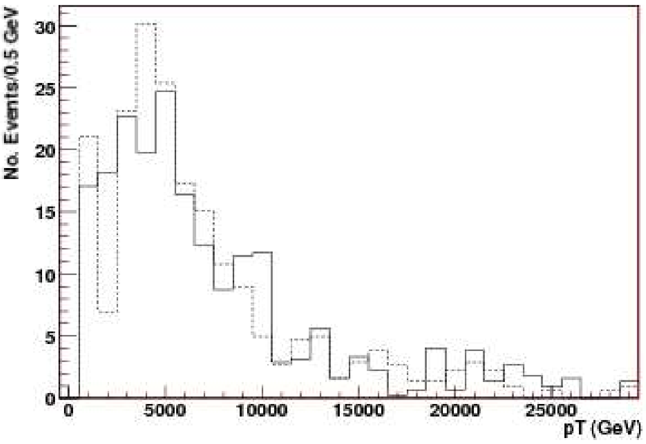
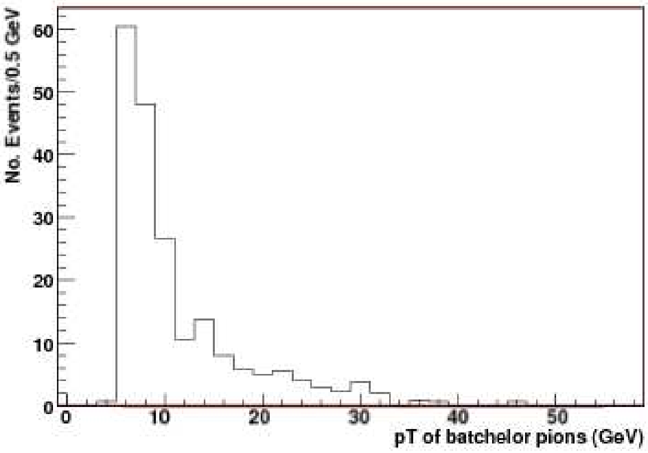


Figure 48: Transverse momentum of all reconstructed tracks, excluding those associated with electrons.

It is possible to search through the record of the event generator to identify which events of the sample contained true D^*s that both passed the W selection cuts and contained a charm quark that decayed through the channel of interest. For this subset of events the p_T distributions of each of the kaon, pion and bachelor pion tracks that resulted from the D^* decay were plotted, and the results are shown in Figure 49. Only events with tracks which passed the low p_T threshhold of 0.5 GeV are plotted, there are approximately 220 such events.



(a) kaon p_T (continuous) and pion p_T (dotted) (MeV)



(b) bachelor pion p_T (MeV)

Figure 49: Tranverse momentum distributions of kaons and pions resulting from D^* decay after W boson selection cuts

In order to reduce the computing time, the lowest p_T kaon and pion candidates are rejected by requiring that their transverse momenta are constrained to > 1 GeV. The tracks associated with bachelor pions were only constrained to $p_T > 0.5$ GeV. From Figure 49 it can be seen that the kaon and pion transverse momenta peak around 5 GeV and so only a significant fraction of the signal will be removed by the p_T cuts on the kaon and pion tracks (fewer than 10% of the reconstructed tracks are removed). The bachelor pion is produced at a much lower momentum on average than the kaon and pion from the D^0 decay, so a large fraction of the bachelor pions will not be reconstructed by the software, causing a large loss of signal. (Specialised algorithms are under development for the reconstruction of lower p_T tracks which may improve the acceptance of this analysis in future: see [85] for details).

4. There are several sign correlations inherent in this signal (as can be seen from Table 6).
   (a) Firstly the electron has the opposite sign to that of the bachelor pion, i.e. charge(e) × charge(π_B) < 1
   (b) Secondly the kaon and pion making up the D^0 have opposite signs, i.e. charge(K) × charge(π) < 1
   (c) Thirdly the kaon and bachelor pion have opposite signs, i.e. charge(K) × charge(π_B) < 1
5. After these criteria were applied, the difference between the masses of the D^* and D^0 candidates reconstructed for each set of three tracks remaining were plotted. For true D^* and D^0 mesons (whose masses are 2010.0 ± 0.5 MeV/c^2 and 1864.6 ± 0.5 MeV/c^2 respectively) the mass difference will be:

$$\delta m = m(D^*) - m(D^0) = 145.4 MeV \tag{32}$$

(similar to the mass of a pion).

The reconstructed masses of the D^* and D^0 themselves tend to be wide due to uncertainties in the measurement of p_T of the tracks, which decrease as the track p_T decreases. The uncertainty of the track inverse transverse momentum, $\frac{1}{p_T}$, is given in [39] Section 3.3.1.6 as

$$\sigma(\frac{1}{p_T}) \simeq 0.36 \oplus \frac{13}{p_T\sqrt{sin\theta}}(TeV^{-1}c) \tag{33}$$

which indicates that for a bachelor pion of $p_T \sim 0.5$ GeV the uncertainty in the track p_T is of the order 7 MeV, for a kaon of $p_T \sim$1 GeV it will be around 15 MeV, and for a pion of $p_T \sim 1.5$ GeV it will be of order 23 MeV. Combining the fractional uncertainties in p_T resolutions of the kaon and pion tracks indicates a minimum uncertainty on the D^0 mass of around 60 MeV, and including the bachelor pion the minimum uncertainty on reconstructing the D^* mass will be around ~100 MeV in this analysis. However the fractional uncertainty in the mass difference δm will be merely that of the (usually low p_T) bachelor pion track so the measurement uncertainty on δm could be as low as 7 MeV.

The intrinsic widths of the masses of the D^* and D^0 are smaller still (180 keV and 170 keV respectively) so the mass difference peak produced should be sharp and narrow. Furthermore, because of the method used to reconstruct this signal a mass difference variable calculated from three tracks chosen at random can never be lower than the charged pion mass (139.6 MeV) meaning that this signal peak is likely to be found in a region of relatively low combinatoric background (as will be seen in Figure 50).

The difference between the masses of the D^* and D^0 candidates was plotted for every combination of tracks that passed the cuts up to this point. The results of this are shown in Figure 50, where the blue line includes every track combination which passes the cuts up to this point after the detector simulation; the pink line shows all the events that contained a true W and a D^* that passed the cuts, representing the signal, and the black dotted line represents the detector truth information, (i.e. the result if reconstruction was perfect for the full sample, without detector smearing, after the cuts). The signal peak is visible around 145.4 MeV/c^2 as expected. The events from the W^+ and W^- samples are weighted by their cross-sections to give numbers of events per inverse femtobarn of data in all subsequent plots.

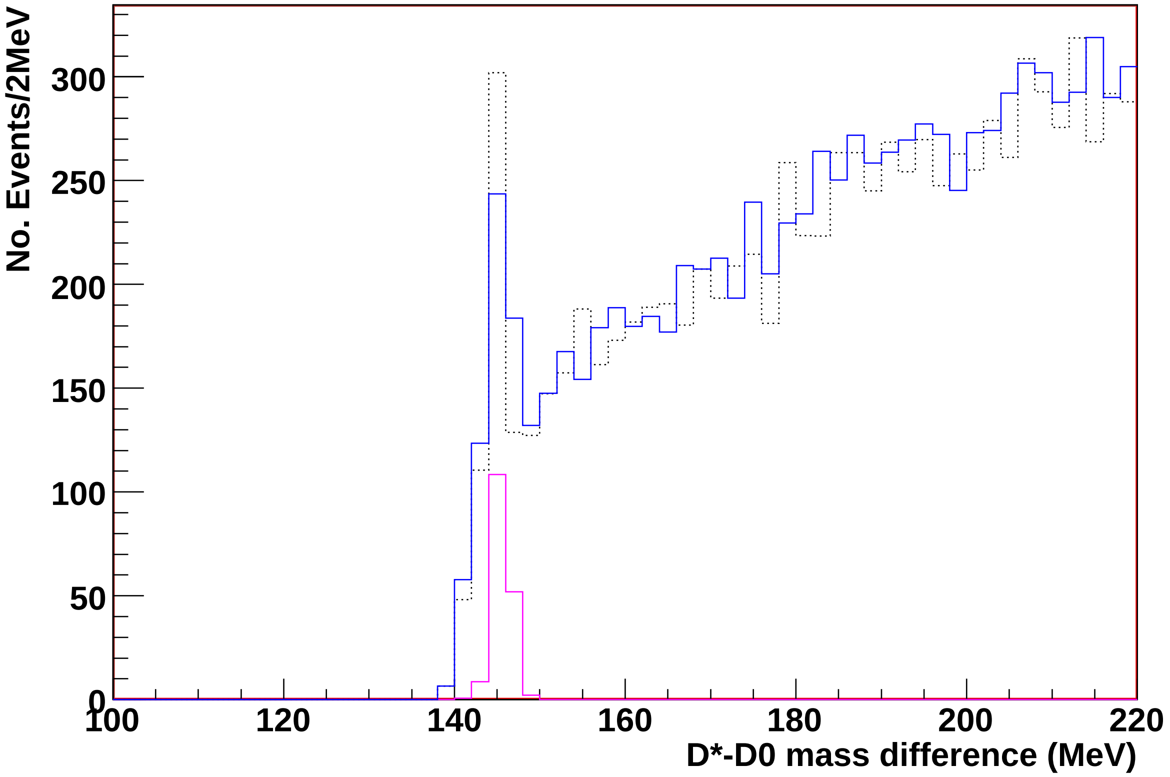


Figure 50: D^*-D^0 mass difference (MeV) calculated for each D^* candidate passing preliminary selection cuts. The blue line represents the full sample, black the detector truth for the full sample, and pink the true (smeared) events in which a D^* was produced with a W at generator level. Each event is weighted by the cross-section such that the plot shows the number of events that would be found in 1fb^{-1} of data.

7.2 Optimised D^* Meson Selection Cuts

To amplify the signal further selection cuts were then applied as described below. To reject as many of the background events as possible without removing too much of the signal a standard $\frac{s^2}{s+b}$ optimisation procedure was applied to the following cuts in order to find the best position for the cut to maximise signal significance. s here represents the number of true $W + D^*$ events that pass the cut (i.e. those plotted in pink in Figure 50) and $s + b$ is the number of combined signal and background events for the sample (blue) in each case. The background reduced here is the combinatoric background within the current $W \to e\nu$ sample (as seen in Figure 50), and only the events with a mass difference of < 170 MeV were considered in the optimisation as they represent the most important region in which to reduce the background levels. The statistical uncertainty in measuring the sum of signal and background events is then $\sqrt{s+b}$ (from Poisson statistics), and the statistical significance of a measurement, i.e. the number of standard deviations the signal is away from the null hypothesis, will be given by $\frac{s}{\sqrt{s+b}}$. This variable is squared for simplicity, so the value of $\frac{s^2}{s+b}$ was measured for different cut positions in each case and the optimal position for the cut was taken where this variable was optimised.

To avoid biasing the signal in our test sample the optimisation procedure for cuts was performed on a new set of samples which were created in the same format to the ones used for the signal analysis, but containing different events (this was done simply by re-running the Monte Carlo generation process with different random number seeds). Each cut in turn was applied over a range of cut positions and the value at which $\frac{s^2}{s+b}$ was maximised was chosen as the best place for the cut. The cuts were optimised in the order given. Any events which failed an optimised cut were rejected before the optimisation proceedure was started on the next cut.

In the following plots the distributions for the true signal events (events in which a D^* was produced with a W at generator level, which were reconstructed by the software and passed all cuts so far) are plotted in blue. The distributions for the entire sample (background and signal) are normalised to the same area under the graph and plotted in black for shape comparison. The position of the optimised cut point is marked in pink.

The final optimised cut values were then applied to the original test sample and the results are given in section 7.3.

1. $|m(D^0_{reconstructed}) - m(D^0_{true})| <$ **40 MeV**

   This cut, the position of which is shown in Figure 51, constrains the mass of the reconstructed D^0 candidate to be within 40 MeV of the true D^0 mass.

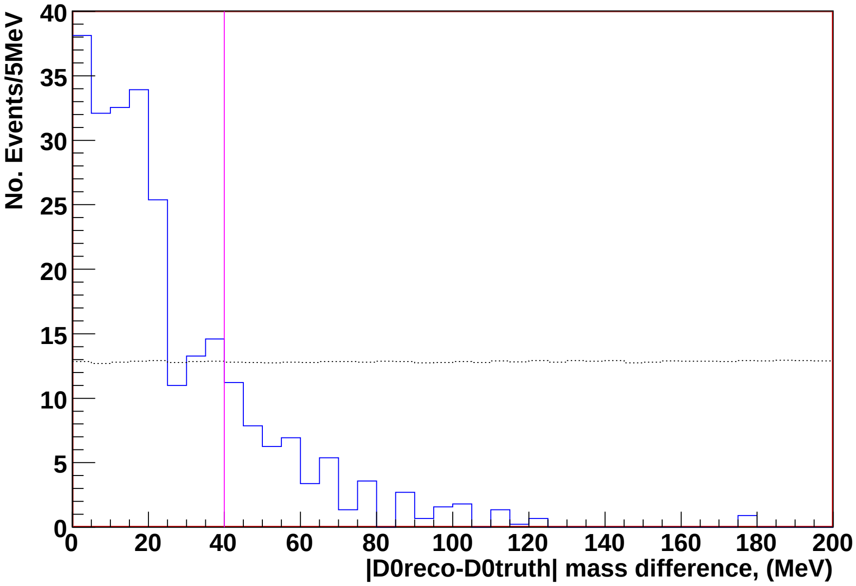


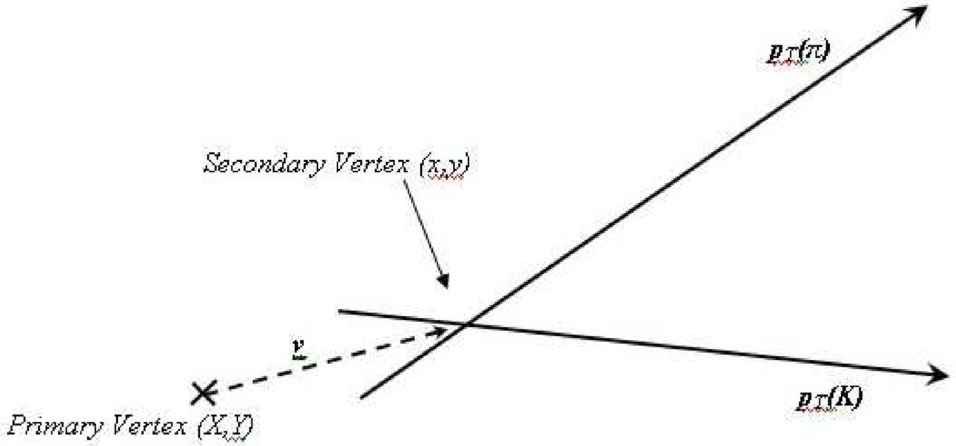
Figure 51: $m(D^0_{reconstructed})$ - $m(D^0_{true})$: mass difference (MeV) between reconstructed D^0 candidate and true D^0 mass. (The same distribution for the entire sample - background and signal - are normalised to the same area under the graph and plotted in black for comparison.)

2. **Signed** $\mathbf{L}_{xy} >$ **0.35 cm**

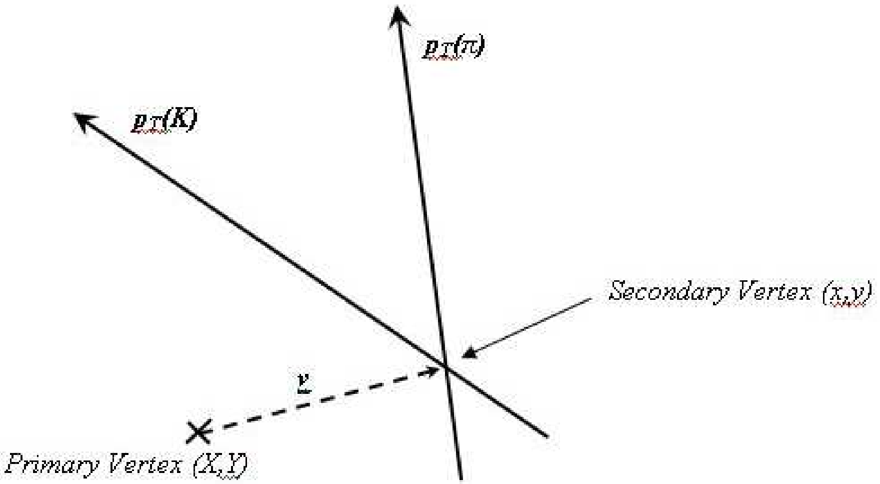
   The D^0 meson has a finite lifetime and should travel on average several millimetres from the interaction point before it decays. Therefore if the K and π tracks originated from the decay of a D^0 meson, then they will have been formed at a secondary vertex point *(x,y)*. The point at which the track p_T vectors cross is defined as the secondary vertex, assuming a straight line approximation for the particle tracks. The position of the primary interaction in the transverse plane is defined as *(X,Y)*, as shown in Figure 52, and is approximately at the centre of the beampipe. The variable L_{xy} is then simply the length of the vector ($\underline{v}$) which represents the direction of the vector of D^0 transverse momentum (the dotted lines). If the K and π tracks come from a real D^0 decay they will be pointing away from the primary vertex as shown in Figure 52(a).

If on the other hand the tracks selected were not produced from a D^0 decay the tracks could be pointing in any direction (Figure 52(b)).

If the vector sum of the tracks' transverse momenta points back to the primary vertex they are, by conservation of momentum, very unlikely to have been created by a D^0 travelling from the primary vertex and then decaying. The L_{xy} variable is signed to take this into account such that if $\underline{v}.(\underline{p_T}(K) + \underline{p_T}(\pi)) > 0$ it is positive, otherwise it will be negative. Random combinations of tracks should produce a distribution that peaks around 0.



(a) Reconstruction of secondary vertex in the case where two tracks come from decay of a neutral particle formed at the primary vertex



(b) Reconstruction of secondary vertex in the case where two tracks intersect at random

Figure 52: Reconstruction of secondary vertices

The shape of the signed L_{xy} distributions for the true signal events and full sample are shown in Figure 53.

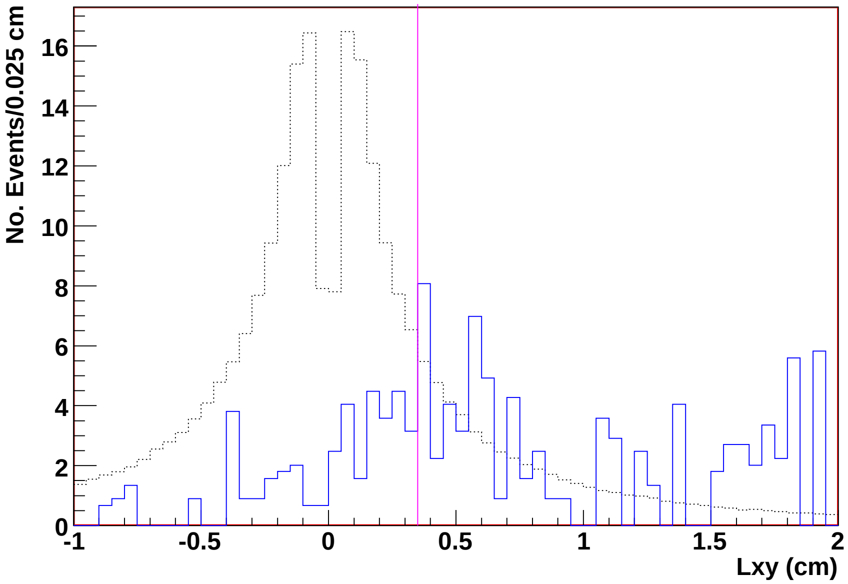


Figure 53: Reconstructed signed L_{xy} of secondary vertex (cm) (the same distribution for the entire sample - background and signal - are normalised to the same area under the graph and plotted in black for comparison).

3. $|\frac{d_0(\pi_B)}{\sigma(d_0(\pi_B))}| < \mathbf{3}$

The D^* has a negligible lifetime and so both the D^0 itself, and the bachelor pion produced with it, should be *prompt* (i.e. produced at the primary vertex). The impact parameter d_0 of a track measures its distance of closest approach to the primary vertex in the transverse plane (again assuming the straight line approximation) as shown in Figure 54.

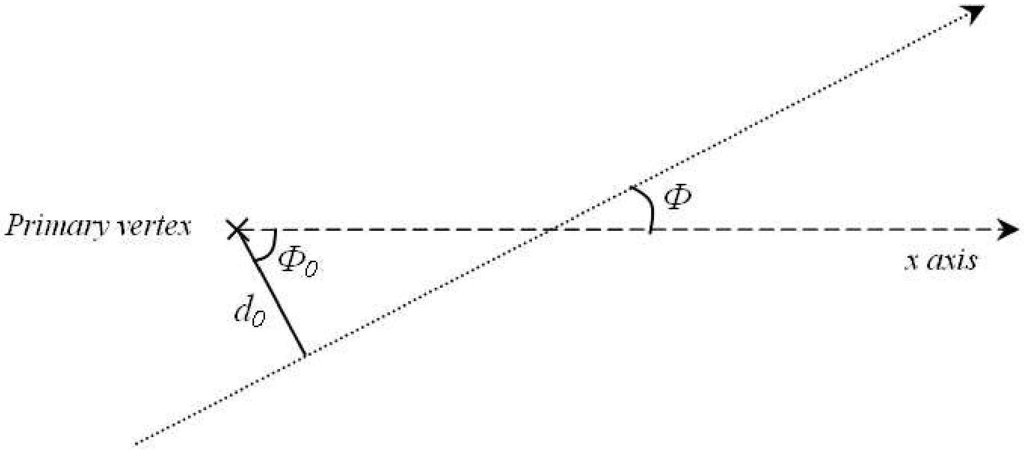


Figure 54: Definition of the impact parameter of a track

In ATLAS impact parameters are signed according to the following convention: if

$$\Phi_0 - \Phi = \tfrac{\pi}{2} + 2\pi n$$

then the track d_0 is signed positive. Otherwise it is negative.

The variable $|\frac{d_0(\pi_B)}{\sigma(d_0(\pi_B))}|$, where $d_0(\pi_B)$ is the impact parameter of the bachelor pion and $\sigma(d_0(\pi_B))$ its uncertainty, is plotted in Figure 55 for the signal and for the full sample.

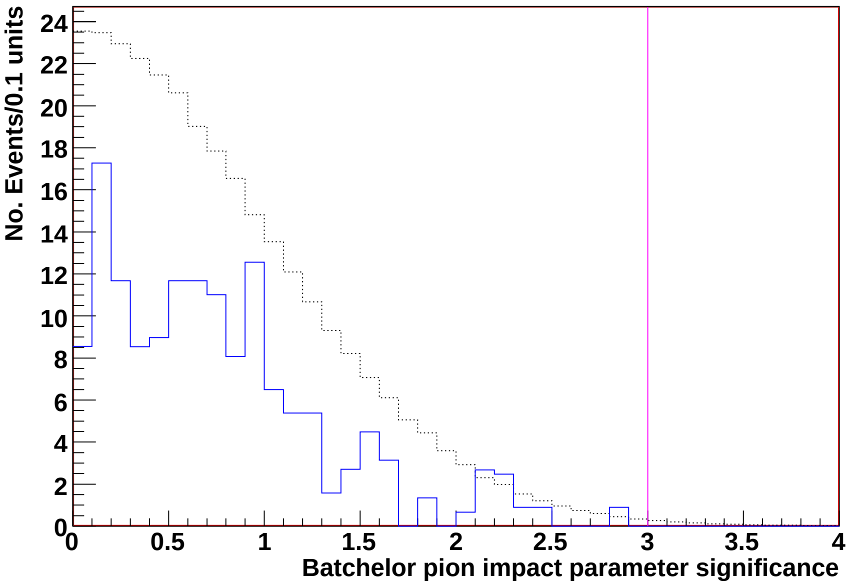


Figure 55: The impact parameter significance of the bachelor pion (the same distribution for the entire sample - background and signal - are normalised to the same area under the graph and plotted in black for comparison).

The optimisation proceedure showed that there was little benefit in cutting on this variable in Monte Carlo. A three standard-deviation cut was applied which will remove those tracks still identified with bachelor pions which would have been very unlikely to have come from the primary vertex. This removes very little of the background in this sample and does not significantly improve the signal to background ratio. However this cut should become useful in rejecting backgrounds involving B mesons (formed in other production mechanisms than those found in this sample), which have relatively long lifetimes and can decay to produce D^*s. D^*s formed from b jets will be produced far from the primary vertex, and in that case the impact parameters of their bachelor pions will be significantly larger on average than those formed in prompt c jets. The expected impact parameter significance of bachelor pions produced from the decay of D^*s formed in b jets is compared qualitatively with those from charmed jets in Figure 56.

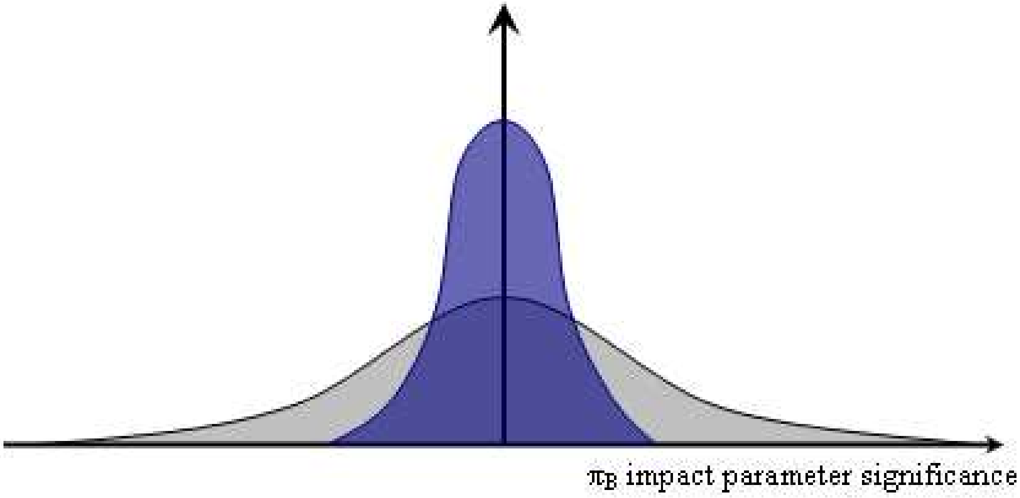


Figure 56: Qualitative view of the expected shapes of the impact parameter significance of the bachelor pion produced from the decay of D^* mesons formed in c (navy) and b (grey) jets.

It should be possible to tune this cut in data if needed.

4. $\mathbf{d}_0(K) \times d_0(\pi) < 0cm^2$

   In the signal channel a D^0 meson is produced at the primary vertex, it travels a distance L_{xy} in the transverse plane and then decays into a kaon and pion. Therefore, assuming straight tracks, if the kaon and pion tracks were extrapolated back towards the primary vertex they should pass the primary vertex on opposite sides as demonstrated in Figure 57. This implies that their impact parameters should be oppositely signed (which would not be the case for around half of the *combinatoric background*, the background formed by the random combinations of unassociated tracks).

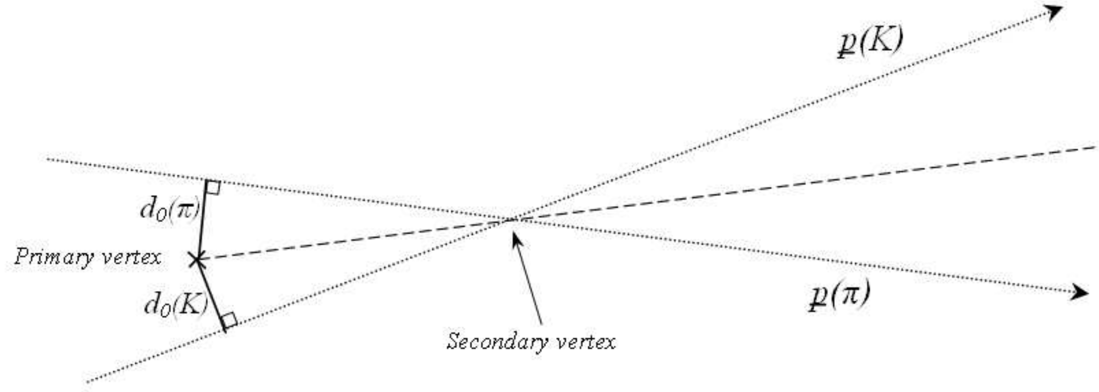


Figure 57: The kaon and pion tracks from D^0 decay will pass on opposite sides of the primary vertex, and are thus oppositely charged

Figure 58 shows the product of the impact parameters of the assigned kaon and pion tracks of the events still passing the selection cuts up to this point. In the signal case this product is usually negative, and the optimal place for the cut was found to be at $d_0(K) \times d_0(\pi) < 0$ cm^2.

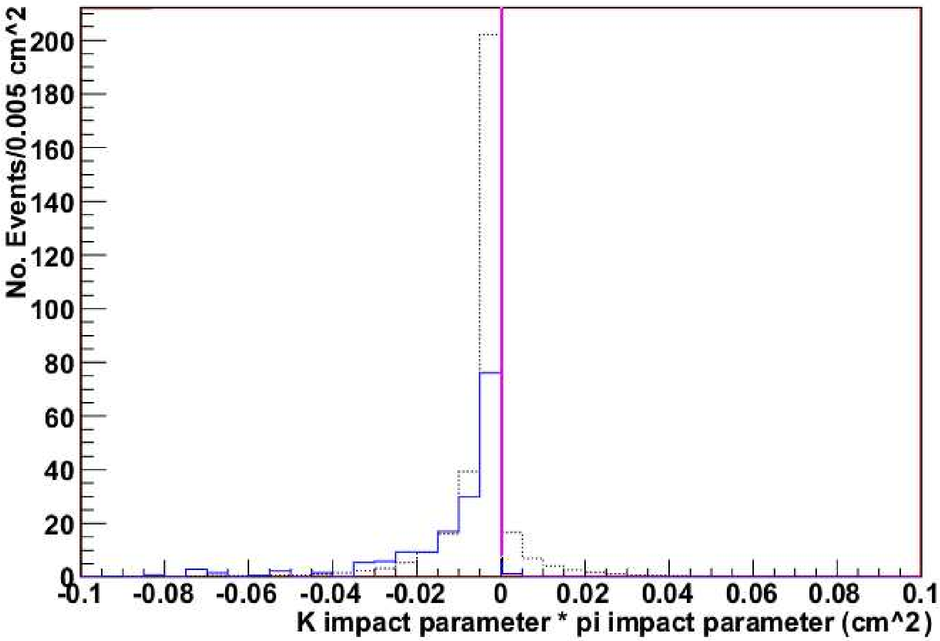


Figure 58: Product of the impact parameters of the kaon and pion (the same distribution for the entire sample - background and signal - are normalised to the same area under the graph and plotted in black for comparison).

5. $|d_0(D^0)| < \mathbf{0.2cm}$

   A transverse momentum vector was reconstructed for the D^0 candidate by combining the transverse momenta of the tracks assigned to the kaon and pion, and from this the magnitude of the impact parameter of the D^0 was extrapolated. Since the D^* meson of our signal channel will decay instantaneously at the primary vertex the D^0 impact parameter should be small. Figure 59 shows the calculated impact parameter for the reconstructed true D^0 and for the remaining sample. The signal and background levels are quite low by this point meaning that the optimisation process is more prone to uncertainty, but a cut on $|d_0(D^0)| < 0.2$ cm is indicated.

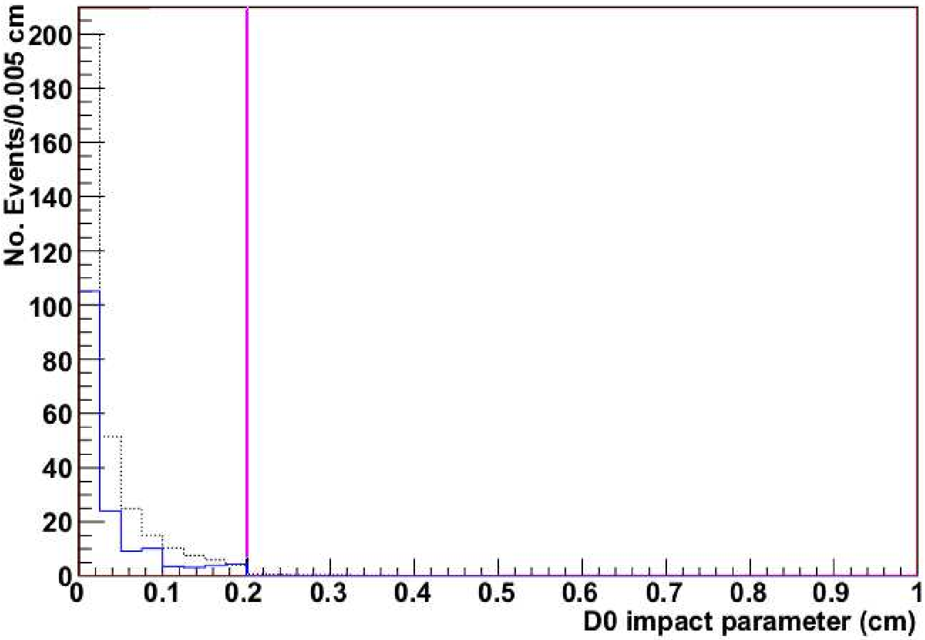


Figure 59: Impact parameter of the reconstructed D^0 (unsigned) (the same distribution for the entire sample - background and signal - are normalised to the same area under the graph and plotted in black for comparison).

The positions of these optimised cuts are listed in Table 7.

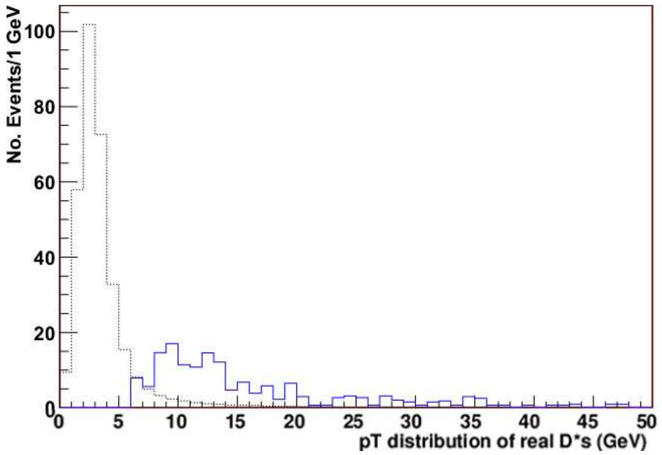
	Optimised Cut
1)	$\lvert m(D^0_{reconstructed}) - m(D^0_{true})\rvert < 40$ MeV
2)	Signed $L_{xy} > 0.35$ cm
3)	$\lvert\frac{d_0(\pi_B)}{\sigma(d_0(\pi_B))}\rvert < 3$
4)	$d_0(K) \times d_0(\pi) < 0$ cm^2
5)	$\lvert d_0(D^0)\rvert < 0.2$cm

Table 7: Summary of the optimised D^* selection cuts

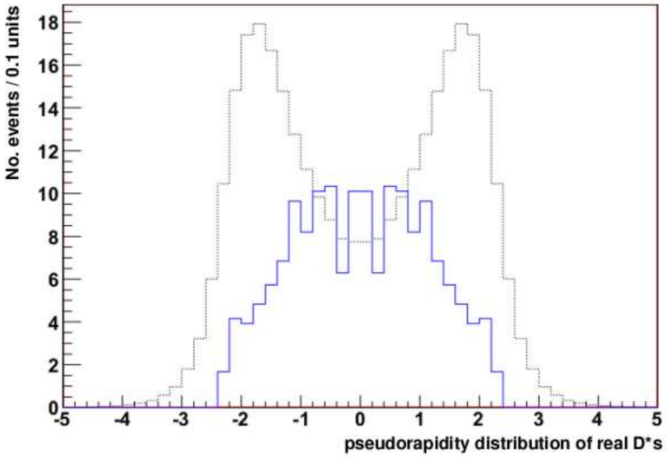
6. **Transverse Momentum and Pseudorapidity of D^* Meson**:

The transverse momentum and pseudorapidity of the reconstructed D^* were then plotted for the real D^*s and the remaining combinatoric background in the test sample, as shown in Figure 60.

(N.B. the pseudorapidity of the D^* is not a meaningful approximation to its rapidity since it is a massive particle, but it can be used as a discriminant to reduce the combinatoric background.)



(a) Transverse momenta of signal D^*s compared with the background (MeV)



(b) Pseudorapidity of signal D^*s compared with the background

Figure 60: Transverse momenta and pseudorapidities of the signal D^*s (blue) remaining after optimised selection cuts compared with the shape of the remaining background (black, dotted)

Cuts on $p_T(D^*) > 6$ GeV and $|\eta(D^*)| < 2.4$ were applied to reduce the background without reducing the signal.

7. **Angle between D^* and W boson in the Transverse Plane**:

In the signal channel the W boson and c quark are expected to be produced approximately back-to-back in the transverse plane. The transverse momentum vector of the W boson was constructed from the electron and missing energy transverse momentum vectors, and the angle between the reconstructed D^* meson and W boson was then calculated. This is shown in Figure 61 for the signal events and remaining combinatoric background.

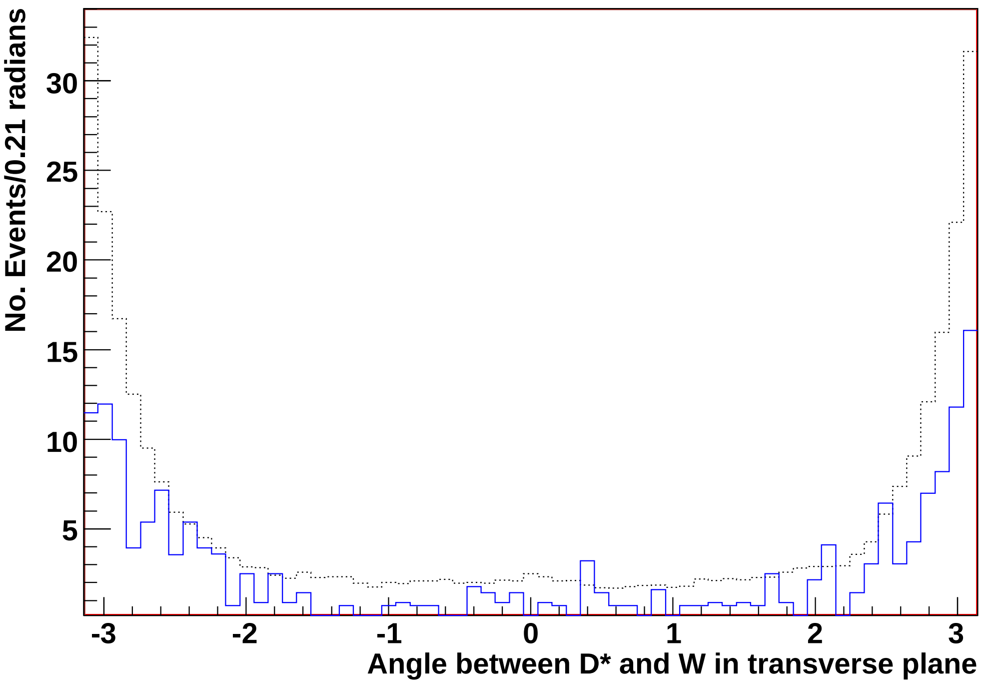


Figure 61: Angle between the electron and reconstructed D^* in the case that the event contains a signal D^*s (blue) compared with the shape of the remaining background (black, dotted)

The real $W + D^*$ events and combinatoric background follow similar distributions and so this variable cannot be used to search for D^*s here. This is unsurprising as the sample generated included either only a W boson decaying to $e\nu$ (first order W production), or a W with a jet (second order). A large proportion of the first order-produced events will have been removed during the D^* selection, and any jet produced in second-order W production will tend to leave the interaction in the opposite hemisphere to the electron. This means that most of the tracks that will form the combinatoric background to the signal will still be found at a large angle from the electron. This cut should become useful when considering, for example, backgrounds involving multiple jets, or b jets in which both the electron and D^* can be formed in the same jet.

8. **MET Perpendicular to Electron Line of Flight**

   In section 6.3.2 the possibility of selecting $W \rightarrow e\nu$ events generated from second order diagrams by considering the width of the distribution of the component of missing transverse energy resolved perpendicular to the electron line of flight was discussed. This variable is plotted in Figure 62 for both the events which contain a true D^* and for the combinatoric background, after all of the previous selection cuts were applied.

   Since the majority of the first order-produced W bosons will have already been rejected, the real $W+D^*$ events and the combinatoric background again have similar shapes by this stage in the selection process and so this variable cannot be used to search for D^*s in this sample (cutting on this variable does not improve the signal significance). However it may become useful for rejection of other (mostly non-W producing) backgrounds.

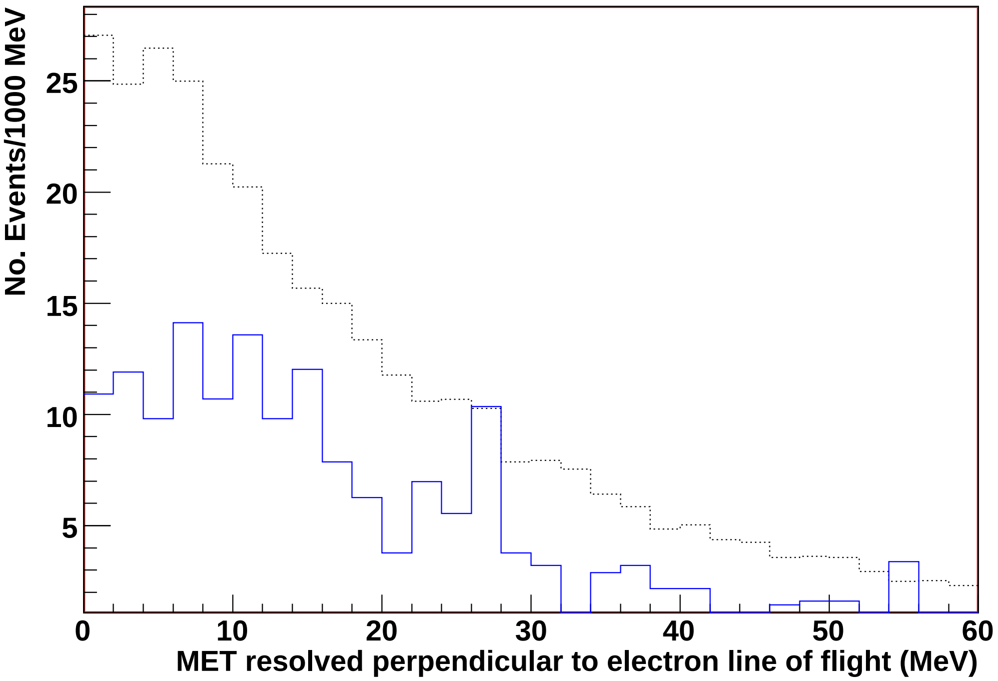


Figure 62: Component of missing transverse energy resolved perpendicular to electron line of flight in the case that the event contains a signal D^*s (blue), compared with the shape of the remaining background (black, dotted)

7.3 Results of the D^* Search

After applying the selection cuts the D^*-D^0 mass difference δm for the combined W^+ and W^- boson samples was plotted again, as shown in Figure 63. The events are weighted by their cross-sections to give a measure of the number of events per fb^{-1} of data. The signal peak is found primarily between 144 and 148 MeV in mass difference, so this is chosen as the region in which to consider the signal and (later) background levels.

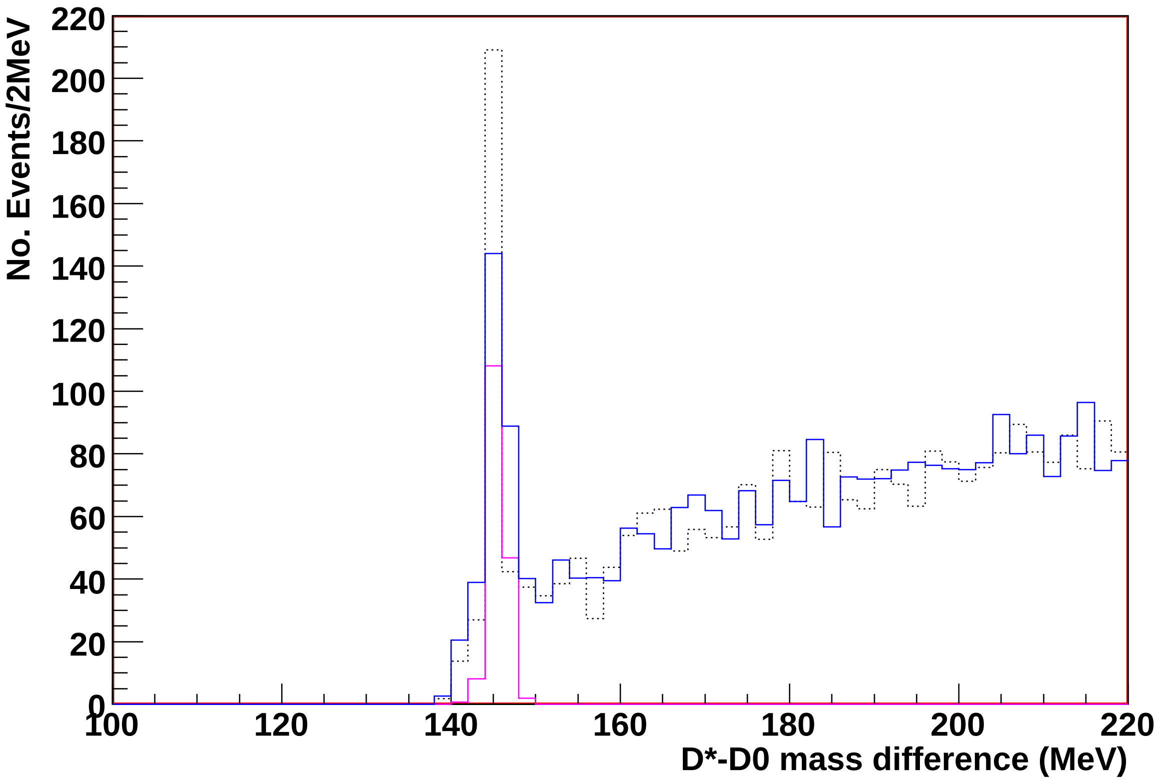
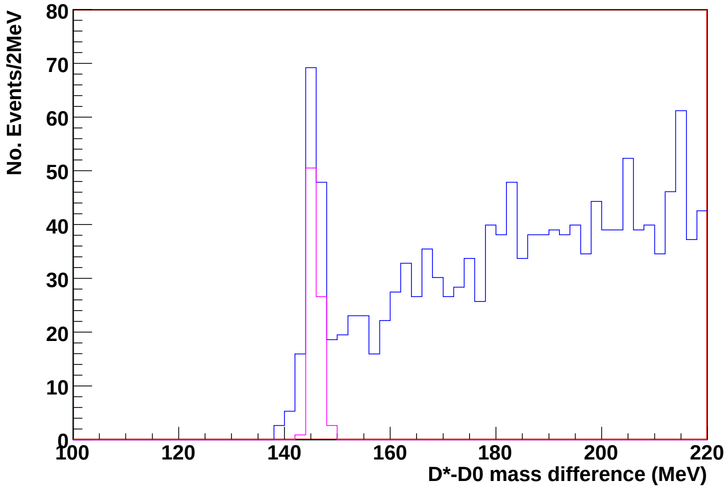
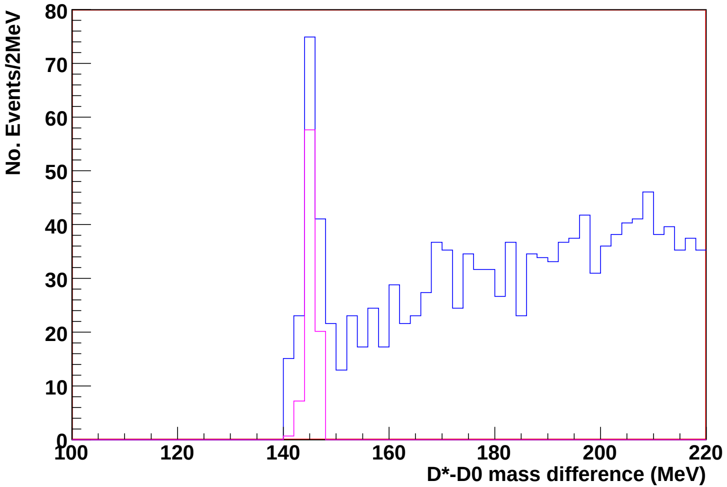


Figure 63: D^*-D^0 mass difference δm (MeV) calculated for each D^* candidate passing all selection cuts. The blue line represents the full sample, black the detector truth for the full sample, and pink the true events in which a D^* was produced with a W at generator level (detector-smeared). The plots are normalised to represent the number of events that would be found in inverse femotobarn of data.

The δm distribution for the individual W^+ and W^- boson samples are shown in Figure 64. The plots are again weighted by their cross-sections to fb^{-1} of data. The level of the combinatoric background in the W^+ sample is seen to be slightly higher than that in the W^- sample since the former has a larger number of events per fb^{-1}.



(a) D^{*}-D^{0} mass difference δm (MeV) calculated for the W^{+} sample



(b) D^{*}-D^{0} mass difference δm (MeV) calculated for the W^{-} sample

Figure 64: D^{*}-D^{0} mass difference δm (MeV) calculated for each D^{*} candidate passing all selection cuts for the W^{+} and W^{-} samples individually. The blue line represents the full sample and pink the true events in which a D^{*} was produced with a W at generator level. The plots are normalised to represent the number of events that would be found in inverse femotobarn of data.

The number of signal events above the background in the mass difference range 144 MeV $< \delta m <$ 148 MeV was estimated from Figure 63 for the entire sample. The number of the true $W + D^{*}$ events passing all cuts and falling into this mass region (i.e. the pink events in Figure 63) were also counted and then weighted by the cross-section of each channel to give a number of events per fb^{-1}. The same process was applied to the plots shown in Figure 64 for the W^+ and W^- samples individually, and the results are given in Table 8.

Sample	Measured no. events/fb^{-1}	True no. events/fb^{-1}
W^+	82 ± 25	77
W^-	83 ± 25	78
$W^+ + W^-$	166 ± 37	155

Table 8: Number of events per fb^{-1} from each of the $W \rightarrow e\nu$ datasets which pass the W boson selection cuts (to the nearest event). Because the statistics are low the measured numbers of events and the uncertainties on those are estimated by manually fitting a straight line to background under the peak, counting the numbers of events by hand and combining uncertainties from the straight line fit and poisson errors on the measurement of each bin of the peak itself (a Gaussian fit would be preferable with larger statistics)

Approximately the same number of W^+ events pass the selection events as W^-, as would naively be expected (the measurement uncertainty on the size of the peaks in each case is around 30% of the signal for 1 fb^{-1} of data). However, as was mentioned in Section 5.1, the production mechanisms of all types of events that could form a real W boson and D^* meson together are not necessarily all charge-symmetric.

Table 9 shows the type of production mechanisms that produced the true events in which a D^* meson was produced with a W boson, both of which passed the selection cuts detailed up to this point. The numbers of passing events produced from the $\bar{s} + g \rightarrow W^+ + \bar{c}$ and $s + g \rightarrow W^- + c$ signal channels per fb^{-1} are expected to be equal, and this is seen to be possible within a the quoted uncertainties. A number asymmetry of true $s + g \rightarrow W^- + c$ events over $\bar{s} + g \rightarrow W^+ + \bar{c}$ events of 3.4 ± 4.2 is observed, which is consistent with the expected value of 0 (since the strange and antistrange quark PDFs used were equal).

Production channels other than $s + g$ are backgrounds to the true signal. There is a *flavour excitation* mode which contributes to the $s + g$ signal events, the mechanism for this and its and significance are explained below.

Sample		$W^+ \to e^+\nu$	$W^- \to e^-\bar{\nu}$
Total Events (million)		2.18 ± 0.15	1.76 ± 0.14
		2.46 (2.99) (0.53)	*2.44 (2.99) (0.55)*
Passing Events		77.1 ± 5.3	77.8 ± 6.2
		87 (93) (6)	*108 (125) (17)*
Signal Channel	$\mathbf{s} + g \to W + c$	66.5 ± 4.6	61.3 ± 4.9
		75 (79) (4)	*85 (93) (8)*
Flavour Excitation	$g \to c\bar{c}$	44.3 ± 3.1	36.1 ± 2.9
	$s + c \to W$	*46 (50) (4)*	*50 (54) (4)*
c Producing	$\mathbf{d} + g \to W + c$	2.7 ± 0.2	7.9 ± 0.6
Channels		*3 (3) (0)*	*11 (13) (2)*
Flavour Excitation	$g \to c\bar{c}$	0.8 ± 0.1	5.0 ± 0.4
	$d + c \to W$	*1 (1) (0)*	*7 (7) (0)*
	$\mathbf{b} + g \to W + c$	0	0
		0 (0) (0)	*0 (0) (0)*
b Producing	$\mathbf{u} + g \to W + b$	0.8 ± 0.1	0
Channels		*1 (1) (0)*	*0 (0) (0)*
	$\mathbf{c} + g \to W + b$	1.8 ± 0.1	2.9 ± 0.2
		2 (2) (0)	*4 (6) (2)*
g Producing	$\mathbf{u} + d \to W + g$	4.4 ± 0.3	5.0 ± 0.4
Channels		*5 (7) (2)*	*7 (12) (5)*
	$\mathbf{u} + s \to W + g$	0	0.7 ± 0.1
		0 (0) (0)	*1 (1) (0)*
	$\mathbf{c} + d \to W + g$	0	0
		0 (0) (0)	*0 (0) (0)*
	$\mathbf{c} + s \to W + g$	0	0
		0 (0) (0)	*0 (0) (0)*

Table 9: Production channels for the generation of the events in which the true W bosons and D^* mesons which passed the selection cuts were produced. In normal font are the number of passing events when the samples are weighted by cross-section. The actual number of passing events (unweighted) from each sample are indicated in italics below, and the first set of italicised brackets indicates the number of positively weighted events and the second the negative. The uncertainties given are taken from the uncertainty on the measurements of the cross-sections.

Returning to the Feynman diagrams for $s+g \rightarrow W^- + c$ (Figures 7(a) and 7(c)), included in these there is a physical time-ordering in which one proton contributes a gluon, which then splits collinearly into a charm-anticharm pair. This process is known as flavour excitation. Following this one of the charmed quarks interacts with a strange quark from the other proton, producing a W boson (for illustration of this see Figure 65); the other is free to form a jet. Because gluon splitting is the primary process which gives rise to the $q/\bar{q}$ sea this time ordering could be considered to be an interaction with a sea quark, rather than a gluon.

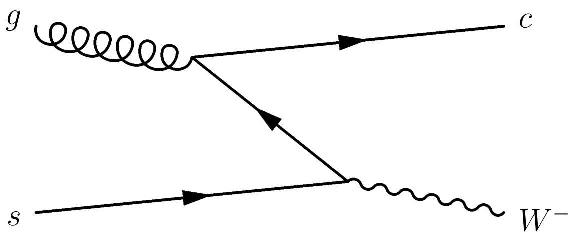


Figure 65: Diagram showing flavour excitation production of $W+c$

As a result any asymmetry in the $c/\bar{c}$ component of the proton sea might affect the production of $W+c$. The process of gluon splitting is itself pertubative and should lead to a symmetric $c/\bar{c}$ distribution at the point at which the gluon splits, after which non-perturbative processes may alter the distributions. The rate of gluon splitting into $c/\bar{c}$ is significantly lower than $s/\bar{s}$ so this is unlikely to be a large effect. It will be assumed here that the flavour excitation mode induces no new asymmetry but this should be experimentally verified in the future.

As shown in Table 9 it transpires that, in the representation of this event generator at least, the flavour excitation mode is a significant source of signal and also some background events. Whilst it is sensitive to the protonic strange quark distribution it might also depend to some extent on the charm PDFs, which are assumed but not known to be symmetric. The effect of a possible asymmetry in the flavour excitation mode on the measurement of the strange sea asymmetry in this channel will be assumed to be negligible here, but would bear further theoretical scrutiny before a study of this type was performed in data. It should be noted that this is a feature of the MC@NLO event generation and is likely not to be physically representative of the fraction of events which might be affected by the charmed sea in data.

The low statistics in some background channels (seen in Table 15) means that this study is most useful as a guide to the types of events which pass the selection cuts; it will not provide a statistically convincing measure of the background contributions to the reconstructed signal (generating the amount of data that would be needed to improve this to a useful level were not achievable on the time-scale of this project).

The channels which produce a charm quark with a W boson without the presence of an incident strange quark ($d/b + g \rightarrow W + c$) form irreducible backgrounds to the signal (i.e. they form the same products with almost the same kinematics). These will be discussed in Section 9.2.1.

Other channels in which the W is produced with a b quark ($u/c + g \rightarrow W + b$) or gluon ($u/c + d/s/b \rightarrow W + g$) also form backgrounds which are difficult to reduce further. b quarks form jets containing some type of B meson, which might decay into a D^* and which can pass the selection cuts if the D^* is produced with a small impact parameter, and gluons can split into $c\bar{c}$ or $b\bar{b}$ pairs, again forming a jet in which a D^* can be formed. These will also be dicussed in Section 9.2.1.

7.4 Summary of $W + D^*$ Selection Cuts

Table 10 lists the numbers of signal events (a W boson produced with a D^* meson) that pass each selection cut, along with percentage acceptances corresponding to each cut.

Cut No.	**Description of cut**	**Signal Events passing cuts /fb^{-1} from W^+ sample (No. (%))**	**Signal Events passing cuts /fb^{-1} from W^- sample (No. (%))**
0	No cuts	8540 (100)	8540 (100)
1	W Boson Selection: $\lvert\eta_e\rvert$¡2.4, $p_{Te} > 25GeV$, $MET > 25GeV$	1630 (19.1)	1630 (19.1)
2	select 3 tracks, all with $\lvert\eta_{track}\rvert < 2.4$, $p_T >\sim 0.5GeV$ (minimum for track reconstruction)	89.6 (1.05)	81.1 (0.95)
3	charge(K) × charge(π) < 1; charge(K) × charge(π_B) < 1; charge(e) × charge(π_B) < 1	79.8 (0.93)	76.6 (0.90)
4	preliminary cut on D^0 mass: mass of reconstructed D^0 is within 200 MeV of the true D^0 mass	78.1 (0.92)	75.9 (0.89)
5	$p_{TK} > 1GeV$; $p_{T\pi} > 1.5GeV$; $p_{T\pi_B} > 0.5GeV$	67.4 (0.79)	75.9 (0.89)
6	$\lvert\eta_{D^*}\rvert$¡2.4; $p_{TD^*} > 6GeV$	67.4 (0.77)	66.3 (0.78)
7	D^0 is within 40 MeV of the true D^0 mass	67.4 (0.79)	66.2 (0.78)
8	L_{xy} of secondary vertex > 0.35	67.4 (0.79)	62.7 (0.73)
9	impact parameter significance of bachelor pion $\lvert\frac{d_0(\pi_B)}{\sigma(d_0(\pi_B))}\rvert < 3$	66.5 (0.78)	61.9 (0.73)
10	product of impact parameters of kaon and pion -ve $d_0(K) \times d_0(\pi) < 0$ cm	66.5 (0.78)	61.3 (0.72)
11	impact parameter of D^0 small $\lvert d_0(D^0)\rvert < 0.2$cm	66.5 (0.78)	61.3 (0.72)

Table 10: Acceptance of consecutive selection cuts for true $W^+ + D^{*-}$ and $W^- + D^{*+}$ events

The acceptance for both channels is similar, as expected. The overall acceptance of the selection process is found to be on average 0.75%. The most significant reductions in acceptance arise from the fact that the D^* mesons are generally produced at low transverse momenta and so the kaon and pion tracks that result from the D^*s are frequently not reconstructed.

7.5 Application of Electron Reconstruction Efficiencies from Full Simulation Studies to Signal Reconstruction

The most significant difference between the fast detector simulation and a physical detector in this study will be the efficiency of electron reconstruction. ATLFAST assumes that all electrons produced within the inner detector tracking region ($|\eta| < 2.5$) will be reconstructed provided that they pass the relevant cuts on transverse momentum and the isolation criteria. In the real detector this will not be the case and a significant fraction of these electrons would not be sucessfully reconstructed. An approximate measure of the difference between ATLFAST and the real detector can be gained by comparing the reconstruction efficiencies of electrons in ATLFAST and GEANT4. The electron reconstruction efficiencies given in GEANT4 in this section for the NLO $W \rightarrow e\nu$ samples were calculated in Section 6.2.1 (the results were given in Table 5) and the efficiencies for GEANT4 are taken from this.

To recap, in Section 6.2 a series of cuts were applied to identify isolated electrons with $p_T > 25$ GeV and $|\eta| < 2.4$ in full simulation. The generator level samples used in Section 6 were a subset of the signal samples used in ATLFAST, but in GEANT4 a filter was applied at generator level such that only events containing a true electron/positron with $p_T >$10 GeV and $|\eta| < 2.7$ were passed into the full simulation (to reduce computation time).

To compare the electron reconstruction efficiencies in the full and fast simulations, the original signal sample was used, the same generator level filter was applied as in GEANT4, and then it was then passed through ATLFAST as usual. Cuts were then applied to select electrons with $p_T > 25$ GeV and $|\eta| < 2.4$ again, and the ATLFAST isolation criteria detailed in Section 7.1 were included. The results are compared with the results for the full simulation in Table 11.

Sample	Probability of electrons passing selection cuts in **Fast Simulation** (%)	Probability of electrons passing selection cuts in **Full Simulation** (%)	Multiplication factor relating fast → full sim
$W^+ \rightarrow e^+\nu$	68.0 ± 0.03	45.0 ± 0.2	0.662 ± 0.003
$W^- \rightarrow e^-\nu$	75.6 ± 0.02	51.1 ± 0.2	0.676 ± 0.003

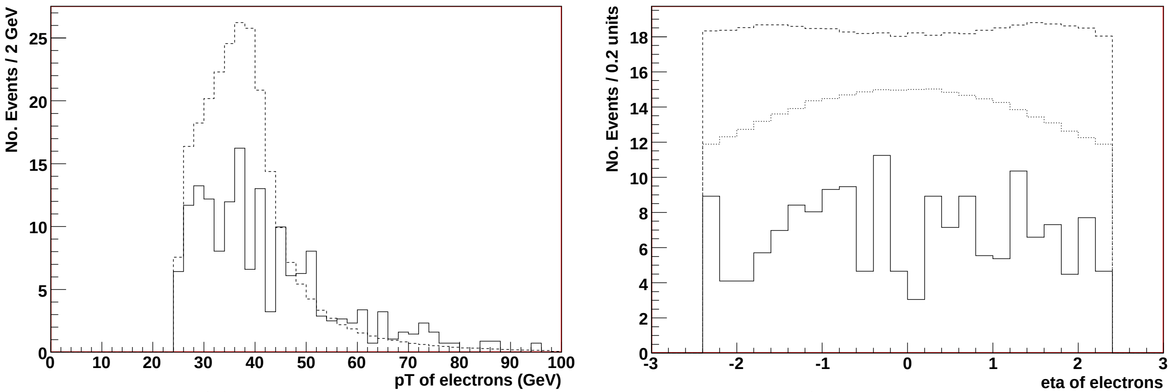
Table 11: Comparison of electron reconstruction efficiencies in full and fast simulation for $W \rightarrow e\nu$ samples

From these numbers the approximate fractions of electron and positrons from W boson decay (with $p_T > 25$ GeV and $|\eta| < 2.4$) that would pass the fast simulation and also the full simulation W boson selection cuts was deduced: 66.2% of the electrons from the W^+ sample which were reconstructed in the fast simulation would also be reconstructed in the realistic detector simulation, and 67.6% from the W^- (the difference in reconstruction efficiencies is due to the difference in the η distributions of the W^+ and W^- bosons produced as was discussed in Section 5.1).

In general, the electron and positron reconstruction efficiencies will depend strongly on their transverse momentum and pseudorapidity distributions as discussed in Section 6.2.5. Figure 66(a) shows the p_T distribution of the electrons and positrons from the true events which pass the selection cuts compared with the original electron p_T distribution for all of the events before selection (only the events with electron $p_T > 25$GeV are shown). It suggests that the cuts do not have a very strong effect on the shape of the electron transverse energy distribution, although the very low statistics remaining at this point make it impossible to be sure.

Figure 66(b) shows the η distribution of the passing electrons and positrons combined, over the range $|\eta| < 2.4$ only. If separated into electrons and positrons individually the distributions appear very similar to one another, as would be expected as the majority of the mechanisms that produced them are charge-symmetric, so combining them into one distribution is reasonable. In the background are shown, not to scale, the original shapes of both the electrons (dotted) and positrons (dashed) produced from W decay, which were not produced in a charge-symmetric way. The original positron distribution tends to increase towards the higher values of η available, whereas the electron decreases. The overall distribution of electrons and positrons after the cuts appears fairly flat (although with the low statistics available this is impossible to state with any certainty), and so represents a middle-ground between the η distributions of the original uncut electrons and positrons.

Because of this, as a working approximation it will be assumed that the reconstruction efficiency for electrons in the fully simulated sample for events passing all of the cuts will be approximately mid-way between those quoted in Table 11 for the W^+ and W^-, i.e. around 66.9%, $\simeq$ 67%.



(a) Electron p_T (b) Electron η

Figure 66: Combined p_T and η distributions of electrons and positrons from true events which pass all selection cuts (solid line). For shape comparison, the same distribution for all events before most cuts are shown, not to scale (dotted lines). These distributions are shown after cuts of $p_T > 25$ GeV and $|\eta| < 2.4$ on the electron/positron only. In the case of the η distributions the electron and positron distributions are initially quite different, and so they are both shown. The upper dotted line is the shape of the electron distribution before the cuts, and the lower the positron.

Table 12 shows the approximate numbers of events from the $W \rightarrow e\nu$ datasets that would be expected to pass the selection cuts with more realistic electron/positron reconstruction efficiencies of 67%. (The numbers with ALTFAST reconstruction efficiencies only were shown in Table 8). The errors were calculated by assuming that the entire event shape is reduced by 33%. It should be noted that the combinatoric background may be higher in full simulation due to the reconstruction of fake tracks, as well as the scattering of secondary particles in the detector.

Sample	Reconstruction Efficiency	Measured no. Events/fb^{-1}	True no. Events/fb^{-1}
W^+	$\sim 67\%$	55 ± 19	52
W^-	$\sim 67\%$	55 ± 19	52
$W^+ + W^-$	$\sim 67\%$	111 ± 28	104

Table 12: Approximate number of events per fb^{-1} from each of the $W \to e\nu$ datasets which pass the W boson selection cuts following application of realistic electron reconstruction efficiencies. The numbers and uncertainties given here were calculated in the same way as those in Table 8: the distributions seen in Figures 67 and 68 were multiplied by the efficiency factor 0.67 and then measured again.

For simplicity, in the consideration of backgrounds in Section 9, 67% will be used as an estimate of the electron reconstruction efficiency in the real detector for backgrounds in which the electrons and positrons are produced with similar transverse energy distributions to the signal. Differences in electron and positron reconstruction rate will not generally be considered. For reference when considering these backgrounds, the numbers of each type of event passing signal selection cuts as seen in Table 9 are given again in Table 13 after the inclusion of a 67% electron reconstruction efficiency.

Sample		$W^+ \to e^+\nu$	$W^- \to e^-\bar{\nu}$
Total Events (million)		2.18 ± 0.15	1.76 ± 0.14
Passing Events		51.7 ± 3.6	52.1 ± 4.2
Signal Channel	$\mathbf{s} + g \to W + c$	44.6 ± 0.7	41.1 ± 3.3
Flavour Excitation	$g \to c\bar{c}$	29.7 ± 2.1	24.2 ± 1.9
c Producing	$\mathbf{d} + g \to W + c$	1.8 ± 0.1	5.3 ± 0.4
Flavour Excitation	$g \to c\bar{c}$	0.5 ± 0.1	3.3 ± 0.3
	$\mathbf{b} + g \to W + c$	0	0
b Producing	$\mathbf{u} + g \to W + b$	0.5 ± 0.1	0
	$\mathbf{c} + g \to W + b$	1.2 ± 0.1	1.9 ± 0.1
g Producing	$\mathbf{u} + d \to W + g$	2.9 ± 0.2	3.3 ± 0.3
	$\mathbf{u} + s \to W + g$	0	0.5 ± 0.1
	$\mathbf{c} + d \to W + g$	0	0
	$\mathbf{c} + s \to W + g$	0	0

Table 13: Production channels for the generation of the events in which the true W bosons and D^* mesons which passed the selection cuts were produced, following application of a 67% electron/positron reconstruction efficiency. In normal font are the number of passing events when the samples are weighted by cross-section. The actual number of passing events (unweighted) from each sample are indicated in italics below, and the first set of italicised brackets indicates the number of positively weighted events and the second the negative. The uncertainties given are taken from the uncertainty on the measurements of the cross-sections.

7.6 Summary of Results of Signal Search

The signal channel used in this analysis has several distinguishing features. It has a clear signature containing just one electron of relatively high energy, an equivalent amount of missing energy, and a charmed jet which is spatially well separated from the electron. The isolated electron can be reconstructed by the Inner Detector and the Electromagnetic Calorimeter, and information from the ATLAS calorimetery and Muon Spectrometers can be combined to reconstruct missing transverse energy to a level that will be acceptable for identification of W bosons. The ATLAS detector has also been designed to offer excellent tracking resolution for charged particles with momenta as low as 0.5 GeV, which should enable the reconstruction of even relatively low energy charmed mesons in the jet from their decays into other particles. The D^* meson was chosen to tag the presence of an outgoing charm quark because its decay into a prompt charged π and neutral D^0, and the subsequent decay of the D^0 into a charged kaon and pion, provide a clear signature of only three tracks, which keeps the level of the combinatoric background manageable. This signature includes three track charge correlations and, because the D^0 will often travel for a significant distance before decaying, a secondary vertex can be reconstructed. These features enable the rapid reduction of the combinatoric background. Finally, although both the D^* and D^0 masses would be significantly smeared and thus difficult to identify amongst the background, plotting the difference in their masses for each event reconstructed produces a peak which is narrow (a few MeV wide) and in a region of low combinatoric background, making this analysis method a good choice for reconstruction of the $W + c$ signal.

This analysis was performed on approximately 2.5 million of each of $W^+ \to e^+\nu$ and $W^-+ \to e^-\bar{\nu}$ events, normalised to an inverse femtobarn of simulated data. The events were produced in MC@NLO, to up to next-to-leading-order and passed through the ATLFAST detector simulation. The resulting signal peaks in the sample of W^+ bosons were of measured size 82 ± 25 events, corresponding to 77 true events in which a W boson and D^* meson were formed and followed the stated decay chain, and 83 ± 25 events measured in the W^- sample, corresponding to 78 truth events (Section 7.3, all here numbers are fb^{-1} simulated data). The reconstruction efficiency of true signal events is found to be 0.78% for the W^+ sample and 0.72% for the W^- (Table 10). The measurement uncertainty on the individual W^+ and W^- signal peaks is around 30%.

Comparison of ATLFAST with GEANT4 for samples of this type (Section 7.5) suggests that approximately 33% of the electrons and positrons in the events which pass the signal selection cuts in ATLFAST would not be reconstructed in GEANT4 or, by inference, in the real detector. This brings the sizes of the measured signal peaks down to approximately 55 ± 19 in each case (52 true events). The size of the measurement error on this number is difficult to estimate since the levels of combinatoric background could be significantly higher in data than in ATLFAST; on the other hand the signal is in an area of naturally low combinatoric background.

The results of studies in full simulation (Chapter 6) suggest that there is an approximately 2.4% electronic charge misidentification rate, suggesting that 2.4% of events which should pass the cuts with the correct sign correlation between the electron and the D^* meson will not be reconstructed in data. This will add a small extra uncertainty of less than one event per fb^{-1}.

The measured number asymmetry from these numbers $\mathcal{A}$ (Equation 16) is in the range 0.4 ± 7.8%. Clearly in order to extract meaningful information about a strange sea asymmetry at ATLAS several orders of magnitude more data would be needed (this will be discussed in Section 10).

8 $W \rightarrow \tau\nu$ as a source of signal and background events

The decay of $W \rightarrow \tau\nu_\tau$ has almost the same likelihood of occurring as $W \rightarrow e\nu_e$ with a branching fraction of 10.74% (again taken from [8]), compared to 10.72% for the electronic decay. If the τ then decays via $\tau \rightarrow e\nu_e\nu_\tau$ (for which the branching fraction is 17.84%) then the event can pass the same selection cuts and will be sensitive to the strange quark distribution in the same way as the $W \rightarrow e\nu$ signal events. However the electrons will be produced at a lower transverse momentum than the original signal electrons and so the reconstruction efficiency should be considerably lower.

8.1 Results of Searching for the Signal in the $W \rightarrow \tau\nu$ Datasets

Approximately 2.5 million of each $W^+ \rightarrow \tau^+\nu_\tau$ and $W^- \rightarrow \tau^-\bar{\nu}_\tau$ were created in MC@NLO and passed through the signal selection process (the τs were free to decay via all permitted SM channels). Of the W^+ sample 1.25% of the generated events passed the W boson selection cuts, and 1.49% of the W^-. These numbers are compared with those for the $W \rightarrow e\nu_e$ samples in Table 14. The ratio between the number of e and τ events passing the selection cuts is also shown. These are of order 3%, which is considerably smaller than the branching fraction of the decay channel $\tau \rightarrow e\nu_e\nu_\tau$, as expected.

Sample	Percentage passing W selection cuts (%)	Sample	Percentage passing W selection cuts (%)	Multiplication factor $e \rightarrow \tau$ of passing events
$W^+ \rightarrow e^+\nu_e$	43.24 ± 0.03	$W^+ \rightarrow \tau^-\bar{\nu}_e$	1.254 ± 0.007	0.0290 ± 0.0002
$W^- \rightarrow e^-\bar{\nu}_\tau$	45.19 ± 0.03	$W^- \rightarrow \tau^-\nu_\tau$	1.488 ± 0.008	0.0329 ± 0.0001

Table 14: Percentages of events from the $W \rightarrow e\nu$ samples which pass W boson selection cuts compared with $W \rightarrow \tau\nu$

The D^*-D^0 mass difference for these samples in combination (again weighted by the cross-section of each channel is shown in Figure 67. The signal peak is not as clearly separated from the background than in the $W \rightarrow e\nu$ case, and levels of both signal and combinatoric background are significantly lower.

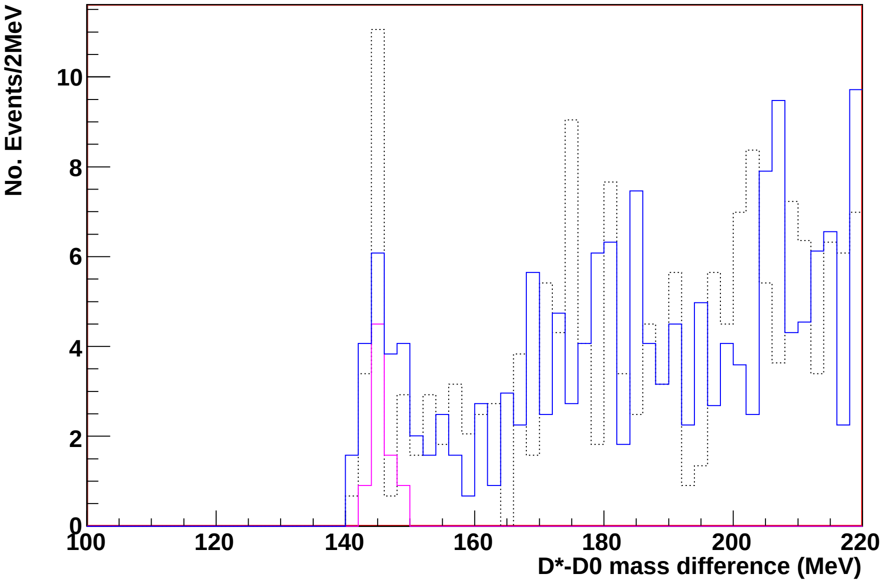


Figure 67: D^*-D^0 mass difference δm (MeV) for the combined $W \rightarrow \tau\nu$ samples, calculated for each D^* candidate passing all selection cuts. The blue line represents the full sample, black the detector truth for the full sample, and pink the true events in which a D^* was produced with a W boson at generator level (detector-smeared). Each event is weighted by the cross-section such that the plot shows the number of events that would be found in one fb^{-1} of data.

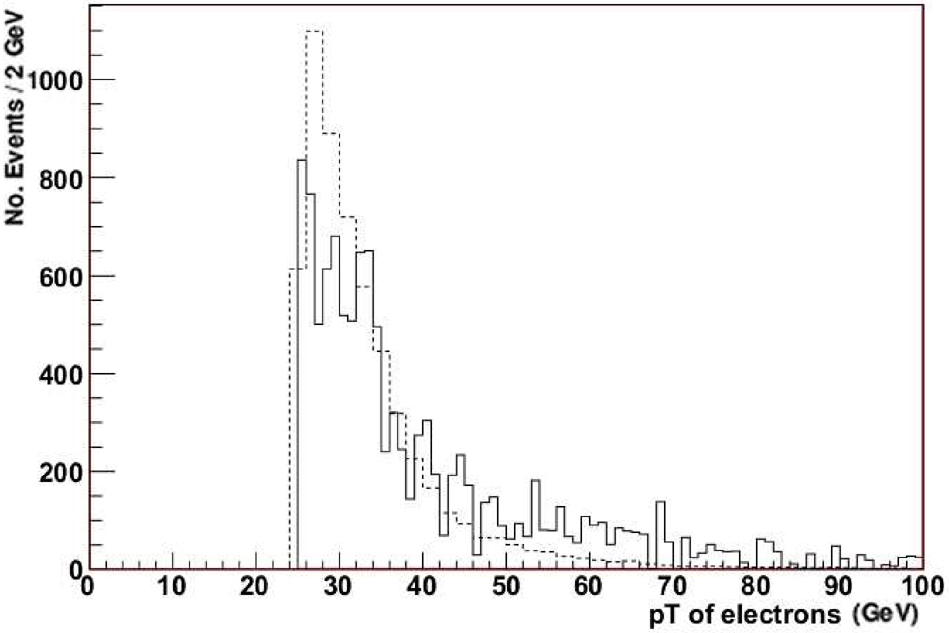


Figure 68: p_T distribution of electrons and positrons from true $W \rightarrow \tau\nu$ events which pass all selection cuts (solid line) compared with the same distribution for all events before cuts (dotted lines, not to scale, only events with $p_T > 25$ GeV are shown)

Two truth events (i.e. $<\sim$ 7 events at 95% CL) from the (un-normalised) $W^+ \rightarrow \tau^+\nu_\tau$ and five from the $W^- \rightarrow \tau^-\bar{\nu}_\tau$ (i.e. $<\sim$ 11 events at 95% CL) pass the selection cuts and are found in the signal mass difference range of 144 MeV $< \delta m <$ 148 MeV. When normalised to one fb^{-1} this equates to a contribution to the signal peak of $<\sim$ 14 events per fb^{-1} in total at 95% CL, or $<\sim$ 9 events fb^{-1} with a more realistic electron reconstruction efficiency of 67% (Section 7.5). Table 15 shows the generation mechanisms of the true events which pass all cuts. As in Table 9 two different uncertainties are given, the symmetric error arises from the uncertainty in the measurement of the cross-section of each channel, and the asymmetric represents the 90% confidence interval.

Sample		$W^+ \to e^+\nu$	$W^- \to e^-\bar{\nu}$
Total Events (million)		2.18 ± 0.15	1.76 ± 0.14
		2.46 (2.99) (0.53)	*2.44 (2.99) (0.55)*
Passing Events		1.8 ± 0.1	3.6 ± 0.2
		2 (3) (1)	*5 (7) (2)*
Signal Channel	$\mathbf{s} + g \to W + c$	0.9 ± 0.1	0
		0 (1) (1)	*-1 (1) (2)*
Flavour Excitation	$g \to c\bar{c}$	0.9 ± 0.1	0
	$s + c \to W$	*1 (1) (0)*	*-1 (0) (1)*
c Producing	$\mathbf{d} + g \to W + c$	0	1.4 ± 0.1
Channels		*0 (0) (0)*	*2 (2) (0)*
Flavour Excitation	$g \to c\bar{c}$	0.9 ± 0.1	0
	$d + c \to W$	*1 (1) (0)*	*0 (0) (0)*
	$\mathbf{b} + g \to W + c$	0	0
		0 (0) (0)	*0 (0) (0)*
b Producing	$\mathbf{u} + g \to W + b$	0	0.7 ± 0.1
		0 (0) (0)	*1 (1) (0)*
	$\mathbf{c} + g \to W + b$	0	0.7 ± 0.1
		0 (0) (0)	*1 (1) (0)*
g Producing	$\mathbf{u} + d \to W + g$	0.9 ± 0.1	1.4 ± 0.1
Channels		*1 (1) (0)*	*2 (2) (0)*
	$\mathbf{u} + s \to W + g$	0	0
		0 (0) (0)	*0 (0) (0)*
	$\mathbf{d} + c \to W + g$	0	0
		0 (0) (0)	*0 (0) (0)*
	$\mathbf{s} + c \to W + g$	0	0
		0 (0) (0)	*0 (0) (0)*

Table 15: Production channels for the generation of the $W \to \tau\nu$ events in which the true W bosons and D^* mesons which passed the selection cuts were produced. In normal font are the number of passing events when the samples are weighted by cross-section. The actual number of passing events (unweighted) from each sample are indicated in italics below, and the first set of italicised brackets indicates the number of positively weighted events and the second the negative. The uncertainties given are taken from the uncertainty on the measurements of the cross-sections.

Table 15 indicates that the events which pass the signal selection cuts from the $W \to \tau\nu$ samples are not primarily produced from the interaction of a strange quark with a gluon. This is not unexpected as W bosons which are produced from the interaction of valence quarks will be produced with higher energies than those from sea quarks and are thus more likely to pass the W selection cuts. This suggests that $W \to \tau\nu$ events should be treated as background rather than as a source of signal events. A much larger study would be needed, however, to quantify the level of background from $W \to \tau\nu$ decays more accurately. Because the numbers of passing events of each separate generation process seen in Table 15 are very small, they suffer from very large uncertainties, so only the total background level in this sample will be used.

9 Rejection of Background Channels

9.1 Overview of Backgrounds to the Signal

There are many channels that can mimic the signal channel by producing either real W bosons and/or D^*s or "fake" ones. Fake W bosons are found when isolated electrons occurr which did not result from the decay of a real W, found with missing energy, and fake D^* candidates can be constructed from random track combinations that happen to match the properties of a real D^*. They fall into the following four broad types (labelled $\mathcal{A}$, $\mathcal{B}$, $\mathcal{D}$ or $\mathcal{D}$):

- $\mathcal{A}$: real W boson, real D^* meson. This group includes the signal, and also
    - Associated production of WD^* pairs in the primary interaction from channels other than the signal channel
    - Unassociated production of a W and D^*, for example in jets produced in the primary interaction
- $\mathcal{B}$: real W boson, fake D^* meson, including
    - Combinatoric background to signal (already considered in Section 7.3)
- $\mathcal{C}$: fake W boson, real D^* meson, including
    - cases involving $Z \to e^+e^-$ in which one electron is not reconstructed correctly in the detected and appears as missing energy
- $\mathcal{D}$: fake W boson, fake D^* meson, including
    - Combinatoric background to cases $\mathcal{C}$

The backgrounds listed below are considered in order in Section 9.2. The dominant type of background is indicated in brackets, bearing in mind that any case involving a real D^* will also include a possibility of reconstructing fake D^*s in the combinatoric background.

1. $W + jet, W \to e\nu$ production, including the irreducible background and those in which the non-charm jet produces a D^* ($\mathcal{A}$).
2. $W + jet, W \to \tau\nu$ production, where $\tau \to e\nu_e\nu_\tau$ ($\mathcal{A}$).
3. $Z + jet, Z \to e^+e^-$ in which one e is lost ($\mathcal{C}$).
4. $Z + jet, Z \to \tau^+\tau^-$ in which one τ is lost and the other decays via $\tau \to e\nu_e\nu_\tau$ ($\mathcal{C}$).
5. Electroweak diboson production: WW, WZ or ZZ where one boson decays leptonically and one decays to $q\bar{q}$ ($\mathcal{A}$ or $\mathcal{C}$).
6. $q_i\bar{q}_i$ in which one quark decays semi-leptonically to produce an electron and a neutrino, and a D^* is formed in one of the jets ($\mathcal{C}$).
7. $W \to q_i\bar{q}_j$ (where q_k are heavy quarks of flavour $k = c$ or b) in which one q_{hk} decays semileptonically and a D^* is found in one of the jets ($\mathcal{C}$).
8. $Z \to q_i\bar{q}_i$ in which one q_k decays semileptonically and a D^* is found in one of the jets ($\mathcal{C}$).
9. Events containing a real D^* formed in a heavy-quark jet and a fake electron, which in this case is an electron that is produced from particle interactions with the detector rather than in the primary interaction ($\mathcal{C}$).

ATLFAST is sufficient for estimating levels of backgrounds in all of these cases except the last, in which GEANT4 is required to simulate the effect of particle interactions in the real detector. As discussed in Section 7.5, appropriate electron reconstruction efficiencies of 67% will be applied for electrons and positrons produced in true W and Z boson decays, which will have p_T distributions similar to the signal.

Backgrounds in which single W bosons are produced will usually involve a number asymmetry from the dual facts that there are more W^+ bosons than W^- at ATLAS, and that the W^- bosons are more often produced in the inner detector tracking region than the W^+, as was discussed in Section 5.1.

9.2 Simulation and Reduction of Backgrounds

In the generation of the samples detailed below, the minimum p_T of partons created in the hard process was 10 GeV, and of those created in secondary scatters 3.85 GeV, as discussed in Section 4. No other significant constraints are applied except where detailed.

For reference the branching fractions of relevant decay channels are listed in Table 16 (including hadronisation where relevant). Again, these are taken from [8] and are charge-symmetric.

Particle	**Decay Channel**	**Branching Fractions (%)**
W	$\rightarrow e\nu$	10.72 ± 0.16
	$\rightarrow \tau\nu$	10.74 ± 0.27
	$\rightarrow cX$	33.6 ± 2.7
Z	$\rightarrow e^+e^-$	3.363 ± 0.004
	$\rightarrow \tau^+\tau^-$	3.370 ± 0.008
	$\rightarrow c\bar{c}$	11.81 ± 0.33
	$\rightarrow b\bar{b}$	15.13 ± 0.05
c	$\rightarrow D^*X$	25.5 ± 0.017
c	$\rightarrow eX$	9.6 ± 0.4
b	$\rightarrow D^*$	17.3 ± 0.2
b	$\rightarrow e\nu X$	10.68 ± 0.22
D^*	$\rightarrow D^0\pi$	67.7 ± 0.5
D^0	$\rightarrow K\pi$	3.80 ± 0.09
τ	$\rightarrow e\bar{\nu}_e\nu_\tau$	17.84 ± 0.06

Table 16: Branching fractions for standard model decay channels

9.2.1 Non-signal W $+ jet, W \rightarrow e\nu$ Production

The Irreducible Background: Down and bottom quarks can interact with gluons to produce a W boson and charmed quark in the same way as the strange quarks do in the signal channel (t-channel examples of this interaction are given in Figure 69; there are also equivalent s-channel diagrams [‡‡]). These contributions cannot be reduced because their kinematics will be nearly identical to those of the signal (the only difference arising from the x distributions of the different partons), but their effects can be studied in Monte Carlo.

[‡‡] t-channel refers to Feynman diagrams (in 2→2 production) in which the two incoming particles fuse to create a virtual intermediate which then splits into the two product; s-channel diagrams are those in which one incoming particle emits an intermediate to form one of the products, the other absorbs the intermediate to form the other.

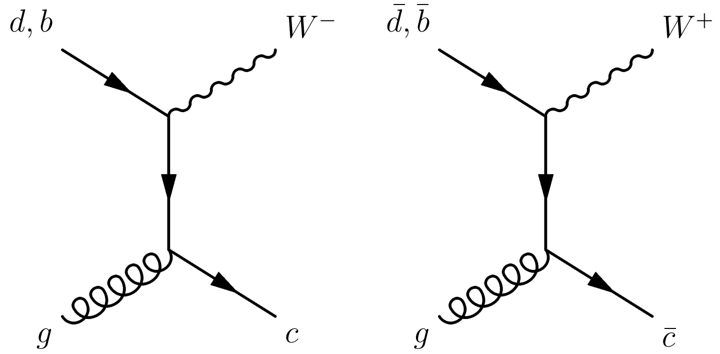


Figure 69: Example Feynman diagrams for the irreducible backgrounds to the signal channel

These channels are significantly less common than the equivalent strange quark generated diagrams which make up the signal, since they involve a change in quark flavour which is discouraged in electroweak interactions. The cross-section of the events generated from the down quark contribution will be supressed by a factor $\sim \lambda^2$ compared with the signal (where λ is a constant, $\lambda \sim 0.2$, as discussed in Section 2.1.2) and those from the bottom quark diagrams will be supressed by a factor $\sim \lambda^4$, so they are not expected to make large contributions to the total $W + c$ cross-section. It must, however, be taken into account that whilst the b and $\bar{b}$ quarks are rare in the proton and can probably be assumed to carry very similar average momentum fractions, this is very much not the case for the d and $\bar{d}$ quarks. As explained in detail in Section 5, since there are valence down quarks in the proton there will be significantly more ds than $\bar{d}$s in total, meaning that more W^- D^{*+} pairs will be produced than W^+ D^{*-} from these diagrams. Furthermore, valence quarks carry more momentum than sea quarks on average, so the W^- and D^{*+} produced from the valence d quarks are likely to be produced at higher rapidities and with greater energies than those from the sea d quarks or the W^+ and D^{*-} produced from the sea $\bar{d}$ quarks. There will thus be both an inherent number asymmetry and a momentum asymmetry in both the W bosons and D^* meson due to the down-containing diagrams.

Considering the truth information of the events which passed all signal cuts (Table 9), it was found that 3 of the W^+ (corresponding to < 7.4 at 90% CL) and 11 of the W^- events (¡ 17.7 at 90% CL) were generated from $d+g$ diagrams. Weighting by cross-section and including a 67% electron reconstruction efficiency, this corresponds to contributions of $<\sim 5$ and $<\sim 9$ of the W^+ and W^- respectively at 90% CL to the signal peak.

These numbers indicate that around 4% of the passing events from the W^- and 10% from the W^+ samples are likely to have been produced by d quark diagrams, which leads to a number asymmetry of W^- and D^{*+} over W^+ and D^{*-} events of approximately $\mathcal{A}$ = -3.4 ± 0.6% (from numbers given in Table 13), where in this case the number asymmetry $\mathcal{A}$ is given by

$$\mathcal{A} = \frac{\mathcal{N}(\bar{d}) - \mathcal{N}(d)}{\mathcal{N}} \tag{34}$$

where $\mathcal{N}(d)$ is the number of $d+g \rightarrow W^- + c$ events which passed the cuts, and $\mathcal{N}(\bar{d})$ the number from $\bar{d}+g \rightarrow W^+ + \bar{c}$. $\mathcal{N}$ is the total number of (correctly sign-correlated) WD^* events measured. As expected the asymmetry is negative, but the numbers involved (~ 7 passing events) are too small for a this to be a very accurate measure of the asymmetry produced by this channel.

None of the equivalent b quark diagrams are found to contribute to the signal peak. Calculation of the channels $b+g \rightarrow W+c$ in COMPHEP shows that its cross-section is less than 0.1% the size of the $s+g \rightarrow W+c$ cross-section, so this background can be neglected.

W produced with b jet: NLO production of W bosons when the associated jet originates from a b quark will form a background in the case that a D^* is formed in the jet (from the decay of a B meson). These D^*s will in general be formed away from the primary vertex (see Figure 56). A cut was placed on the impact parameter of the bachelor pion in the signal selection to reduce this background (and other later backgrounds involving b quark jets). The t-channel Feynman digrams for these channels are shown in Figure 70.

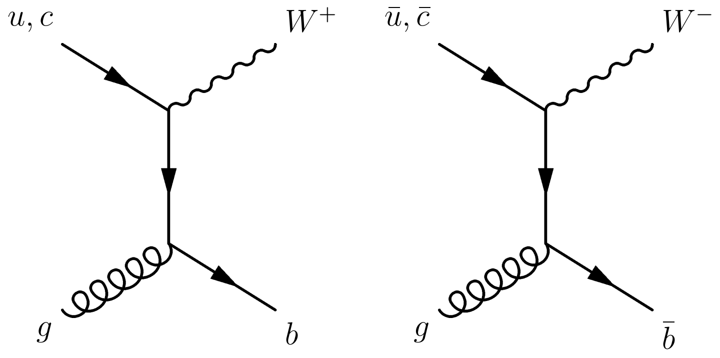


Figure 70: Example Feynman diagrams for b-jet producing backgrounds to the signal channel

The numbers of true events from contributing channels are again shown in Table 13, with a total of 3.6 events per fb^{-1} passing selection cuts. This corresponds to fewer than $\sim$ 7 events per fb^{-1} at 90% CL (including a 67% electron reconstruction efficiency). This background again would tend to be asymmetric, more bs being produced than $\bar{b}$s from the channels $u + g \rightarrow W + b$ (corresponding to a negative asymmetry). An asymmetry of -0.2 ± 0.3% is indicated.

W produced with g followed by gluon splitting: As mentioned in Section 4, around $\sim$ 10% of Ws produced (as measured in the signal sample) are produced in conjunction with gluons. If one of these gluons splits colinearly into a $c\bar{c}$ pair then a charmed jet will be produced so the event could pass selection cuts in the same way as the signal (see Figure 71). However, the Feynman diagrams for production of $W + g$ are not specifically sensitive to the proton strange quark distribution, making this a background to the signal if a D^* of an opposite charge to the W is produced in the jet (a similar effect could be seen if the gluon splits to $b\bar{b}$). D^*s produced in this manner are likely to have lower transverse momenta than in the signal, making them less likely to be selected.

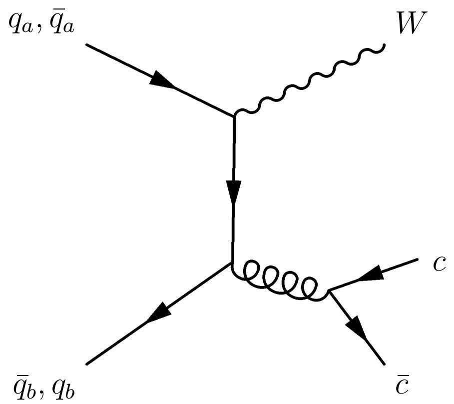


Figure 71: Feynman Diagram for the channel $q_a + q_b \rightarrow g + W$

The true $D^* + W$ events which passed the selection cuts were analysed in Table 13, and of them around 7 per fb^{-1} were found to have originated from gluon splitting diagrams in conjunction with either a W^+ or a W^-. This corresponds to a background level of $\lesssim$ 10 events at 90% CL.

There is again an asymmetry inherent in the diagram shown in Figure 71. More W^+ bosons will be produced in general in this way because there is a higher number of valence u than d quarks in the proton, meaning that overall there should be more gluons produced with a W^+ boson. It can be seen from Table 15 that more events of this type pass the selection cuts in the W^- case, however, leading to a small negative asymmetry, $\mathcal{A}$ = -0.9 ± 1.5 %. This suffers from large uncertainties, so much higher statistics would be needed to quantify this effect further.

Summary of backgrounds included in signal sample: To get an idea of the likely contributions to the signal peak and asymmetries for each of these channels, the numbers truth events to pass selection cuts in each type of background event are shown (taken from Table 13 and the number asymmetries $\mathcal{A}$ are calculated for each channel in Table 17. The same numbers for the true signal channel, and for the overall numbers of $W + D^*$ events that pass the selection cuts are given for comparison. The total combined asymmetry of these backgrounds is also shown in 17.

Channel	**No. Passing Events (truth) per fb^{-1}**		**Asymmetry (%)**
	W^+	W^-	
all $W + D^*$ events passing selection cuts	51.7 ± 3.5	52.1 ± 4.2	-0.4 ± 7.8
Signal $s + g \rightarrow W + c$	44.6 ± 0.7	41.1 ± 3.3	3.4 ± 4.2
Irreducible backgrounds $d/b + g \rightarrow W + c$	1.8 ± 0.1	5.3 ± 0.4	-3.4 ± 0.6
b jet producing backgrounds $u/c + g \rightarrow W + b$	1.7 ± 0.2	1.9 ± 0.1	-0.2 ± 0.3
gluon splitting backgrounds $q_a + \bar{q}_b \rightarrow W + g$	2.9 ± 0.2	3.8 ± 0.4	-0.9 ± 0.6
Background Total	6.4 ± 0.5	11.0 ± 0.9	-4.5 ± 1.5

Table 17: The measured numbers of events contributing to signal peak in $W^+ \rightarrow e^+\nu$ and $W^-+ \rightarrow e^-\bar{\nu}$ samples from the three different types of background channel included in the sample. Uncertainties shown come from uncertainties in the measurement of the cross sections of the channels. Also given are the measured number asymmetries (positive when W^+ contribution is larger)

Table 17 indicates that the overall number asymmetry included in the signal peaks given in Section 7.3 due to asymmetric background channels is -4.5 ± 2.4% (there is an overall excess of W^-D^*+ events over W^-D^*+ background events as would be theoretically expected). These figures are subject to fairly large uncertainties (around 50%) due to the low cross-sections and acceptances of the signal selection process. This asymmetry is small in comparison with the 30% measurement uncertainty given in Section 7.3 for an inverse femtobarn of data. In order to extract any meaningful asymmetry, several orders of magnitude of data greater than 1 fb^{-1} would be required, as will be discussed in Section 10.

9.2.2 Non-signal $\mathbf{W} + jet, W \rightarrow \tau\nu$ Production

This channel was discussed in detail in Section 8. As was shown in Table 15, the total contribution to the signal peak is around 5.4 ± 0.3 events per fb^{-1}, or 3.6 ± 0.2 events if a 67% electron reconstruction efficiency is included.

There are more W^- events passing the selection cuts than W^+, which contributes to the asymmetry of the signal peak of -1.7 ± 0.4%.

9.2.3 $\mathbf{Z} + jet, Z \rightarrow e^+e^-$

NLO production of Z bosons in the channel $c + g \rightarrow Z + c$ (Figure 72) could mimic the signal channel if the Z boson decays into e^+e^- (as could potentially the same diagrams with b quarks instead of c, where the outgoing bottom quark then goes on to decay into a charm). The W boson selection cuts require the presence of one electron with $E_T >$ 25 GeV and $|\eta|$¡2.4, and missing transverse energy > 25 GeV, which could only occur if one of the electrons produced in the Z decay was not reconstructed in the detector ("lost", for example in an area of the detector with poor coverage). In this case the event could appear to contain one electron with $E_T > 25$ GeV and a similar amount of missing transverse energy (equal to the transverse energy of the missing electron).

Furthermore, because of the strong sign correlations used to select the signal, the remaining electron would need to be charged oppositely to the charm quark charge which will be true in approximately half of cases. Because the Z boson has a mass $\sim$ 10 GeV greater than the W mass, the electrons produced in the Z decay will have around a 5 GeV higher average energy than in the W case and thus be more likely to pass selection cuts. However the cross-section for Z boson production is lower than W at ATLAS so the background from this channel is still expected to be low.

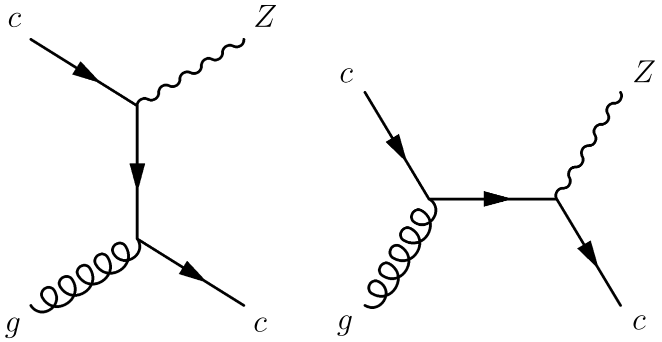


Figure 72: Feynman Diagram for the background channel $c + g \rightarrow c + Z$

To quantify the level of background expected from this channel ∼1.5 million $Z \rightarrow e^+e^-$ events were generated using MC@NLO (similarly to the W sample, this sample includes both leading order Z production and next-to-leading order $Z + jet$ production). The cross-section was calculated in HERWIG to be

$$\sigma_{Z \rightarrow e^+e^-\ NLO} \simeq 0.80 \text{ nb}$$

When the signal selection cuts listed in Table 10 were applied directly to this sample no events passed all of the selection cuts.

However, ATLFAST assumes a 100% electron reconstruction efficiency within the Inner Detector acceptance region, and as discussed in Section 7.5 this is unrealistic. To estimate this effect, the signal cuts were altered by removing the MET cut, and requiring two electrons in the event, both with transverse energy > 25 GeV and at least one with $|\eta| < 2.4$. It was assumed that the electron which was not constrained in η could potentially be lost, creating MET in the detector.

Four truth events passed these selection criteria (Figure 73); two e^+D^{*-} and two e^-D^{*+} pairs were found. This equates to < 10 events at 90% CL, or < 5 events per fb^{-1} at 90%

CL. In reconstruction, these appear virtually indistinguishable from the combinatoric background.

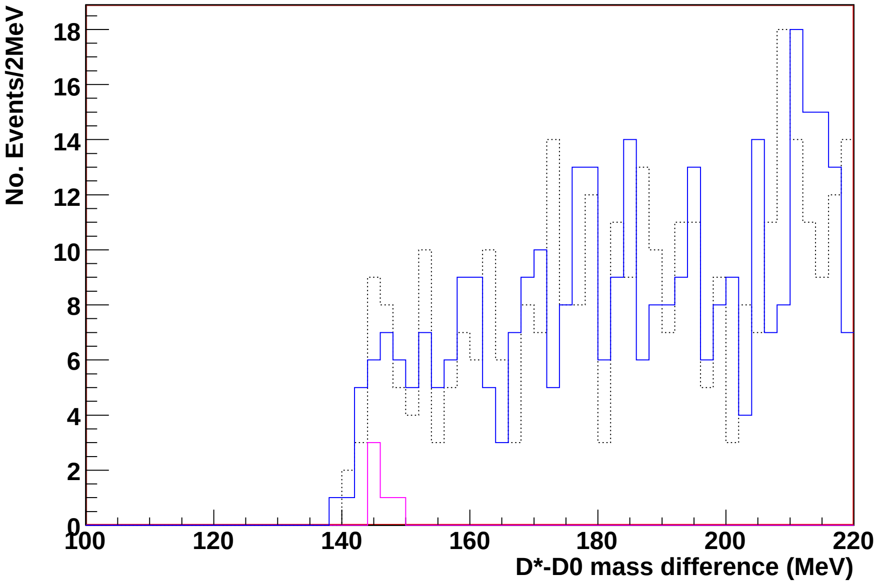


Figure 73: D^*-D^0 mass difference δm (MeV) for the combined $Z \rightarrow e^+e^-$ samples, calculated for each D^* candidate passing all selection cuts. The blue line represents the full sample, black the detector truth for the full sample, and pink the true events in which a D^* was produced with a W boson at generator level (detector-smeared). These distributions is not weighted by cross-section.

Assuming a ∼67% probability of each of the two electrons being reconstruced and requiring that only one is reconstructed, only 44% of these events would be likely to pass the full selection cuts including an MET cut in a realistic detector, bringing the background limit from this channel to < 2.2 events per fb^{-1} at 90% CL. (N.B. There are $Z + c$, $Z + g$ and $Z + b$ backgrounds which would mimic the signal in the same sort of way as the equivalent W background. These are included in this sample and will not be separately considered.)

Z boson production is inherently symmetric, so there should be no charge asymmetry due to this channel. The contribution to the overall charge asymmetry of the signal peak was measured in this case to be $\mathcal{A} = 0.00 \pm 0.02\%$, where the uncertainty was calculated using 90% confidence limits on the two events of each sign pairing that were found to pass selection above: The upper limit on 2 events is 5.9 to 90% confidence and the lower is 0.5. (Multiplying these numbers by the 44% reconstruction rate suggested above and dividing by the overall height of the signal peak given in Section 7.5 gives a maximum asymmetry on this channel of 0.02.)

9.2.4 $\mathbf{Z} + jet, Z \rightarrow \tau^{+}\tau^{-}$

If instead of decaying to two electrons the Z boson decayed to two τ leptons (the branching fractions of $Z \rightarrow e^{+}e^{-}$ and $Z \rightarrow \tau^{+}\tau^{-}$ being almost identical), one of which then decayed to an electron and neutrinos (which occurs $< 20\%$ of the time), then this channel could also mimic the signal. The τ decay produces two neutrinos, so there will be some missing energy, but from helicity considerations they would be expected to be produced approximately back-to-back. Since MET is calculated from momentum asymmetries in the detector, unless there is a large difference in the energies of the two neutrinos their momenta will cancel out to a large extent, suggesting that the MET generated in this way will be much lower than the signal MET. On the other hand, it is much more likely that one of these events will pass the missing energy selection cut if the second τ is lost in the detector.

The electrons produced in this process will also be much lower in energy than those of the signal channel since they are further down the decay chain. This effect was seen in the difference in strength between the $W \rightarrow e\nu$ and $W \rightarrow \tau\nu$ signal channels, in which the latter was shown to provide fewer than 10% of the number of events that passed all signal selection cuts of the former. Therefore the background contribution from this $Z \rightarrow \tau^{+}\tau^{-}$ channel will be $< 10\%$ of the background from NLO $Z \rightarrow e^{+}e^{-}$, implying a contribution to the background of $\ll 1$ event per fb^{-1}. The contribution to the number asymmetry should also be negligible.

9.2.5 Electroweak Diboson Production (WW, WZ, ZZ)

The production of pairs of W and Z bosons forms backgrounds to the signal when one of the gauge bosons decays into a quark-antiquark pair, one of which is heavy and produces a D^* meson in its jet, the other decaying semileptonically. The gauge bosons are likely to be produced approximately back-to-back, which could mimic the signal quite closely; however their production cross-sections are very low.

The decay channels most likely to mimic the signal are as follows (where X is any particle or set of particles):

- WW: one $W \to e\nu$, the other $W \to cX$ (primarily $c\bar{s}$)
- ZZ: one $Z \to e^+e^-$ (one lost), the other $Z \to c\bar{c}$ or $Z \to b\bar{b}$
- WZ: $W \to e\nu$ and $Z \to c\bar{c}$ or $Z \to b\bar{b}$, or $Z \to e^+e^-$ and $W \to cX$

Events in which $Z \to b\bar{b}$ could pass the selection cuts if a B meson formed in a b jet decayed into a D^* meson. The decay of Z bosons to $b\bar{b}$ is more common than $c\bar{c}$, but the hadronisation/decay fraction of $b \to D^*X$ is a smaller than $c \to D^*X$ (see Table 16). The $Z \to c\bar{c}$ and $Z \to b\bar{b}$ are then of roughly similar significances. On the other hand, W decays into b quarks are very rare, the branching ratio is around $B(W \to c\bar{b}) = 0.06\%$ (taken from the Pythia decay table, there is no PDG data available for this decay) so this channel is neglected here.

Table 18 shows the cross-section for each of the listed channel, calculated with HERWIG. This is multiplied by the total branching fraction calculated for the decay channels listed above, and the subsequent hadronisation of a charm into a D^* followed by its eventual decay into π_B, $K\pi$ as used for signal selection (represented by $c \rightarrow K\pi\pi$ here, $B =$ 0.656%). This gives a fractional cross-section for the likelihood of an electron and a D^* which decays via the signal route being present in an event.

	HERWIG cross section σ (pb)	Description of Decay Channel	Branching Fraction B	Fractional Cross Section σB (pb)	No. Events /fb^{-1}
WW	70	$2\ x\ (W\rightarrow e\nu\ x\ W\rightarrow cX)\ x\ c\rightarrow K\pi\pi\ )$	4.73 x 10^{-4}	3.3 x 10^{-2}	3.3
ZZ	11	$2\ x\ (Z\rightarrow e^+e^-\ x\ (Z\rightarrow c\bar{c} + Z\rightarrow b\bar{b})\ )\ x\ c\rightarrow K\pi\pi\ )$	1.19 x 10^{-4}	1.3 x 10^{-3}	0.13
WZ	27	$[(W\rightarrow cX\ x\ Z\rightarrow e^+e^-) + (W\rightarrow e\nu\ x\ (Z\rightarrow c\bar{c} + Z\rightarrow b\bar{b})]\ x\ c\rightarrow K\pi\pi\ )$	2.63 x 10^{-4}	7.1 x 10^{-3}	0.71

Table 18: Number of events per fb^{-1} in electroweak diboson decay that decay in modes used for signal selection

Together these backgrounds contribute a total of around 4 events per fb^{-1} that might mimic the signal without taking into account a reduction due to requirements on the sign correlation between the passing electron and D^*, or the selection cuts, which should remove a very large percentage of these events. Only around 0.7% of the true, correctly sign-correlated, signal events survived reconstruction and all selection cuts, section 7.3 and a similar or lower $W + D^*$ reconstruction efficiency should occur here. The combined contribution of these channels to the background can therefore be assumed to be negligible, as can the contribution to any number asymmetry.

9.2.6 q$\bar{q}$ Production

The rates of production of $q\bar{q}$ pairs are very high at hadron colliders such as the LHC. Pairs of heavy quarks ($c\bar{c}$, $b\bar{b}$ and $t\bar{t}$) are therefore able to produce D^*s and electrons in large numbers. Their cross-sections for production at ATLAS were calculated as

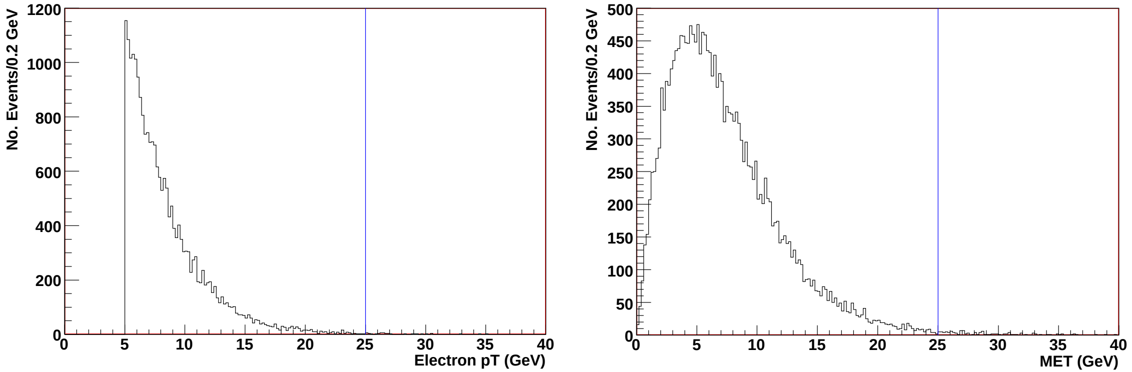
1. $\sigma(c\bar{c}) = 14.5$ mb (to leading order, Pythia)
2. $\sigma(b\bar{b}) = 3.3$ mb, as calculated to next-to-leading order by MC@NLO. Pythia gives the leading order cross section to be 0.4 mb. Another semi-NLO generator, ALPGEN, predicts the NLO $b\bar{b}$ cross section to be around 2mb. This suggests that the MC@NLO generator may be overestimating the $b\bar{b}$ cross section, but the cross-section is not well known at this point. The value 3.3 mb will be used here, which will likely lead to a conservative estimate of the background level.
3. $\sigma(t\bar{t}) = 0.5$ nb (to leading order, Pythia)

MC@NLO is able to produce $b\bar{b}$ and $t\bar{t}$ events, but is not currently able to calculate $c\bar{c}$ to NLO. Pythia was used for $t\bar{t}$ since an order of magnitude estimate was sufficient here. Notice that the LO and NLO cross sections calculated in Pythia and MC@NLO for the $b\bar{b}$ sample are a factor of 10 different. This implies that the cross sections should be used for order of magnitude estimation only, especially in the case of $c\bar{c}$ which cannot easily be generated at NLO at present. These channels are considered in detail below (the NLO sample is used for $b\bar{b}$).

1) The $c\bar{c}$ channel can mimic the signal if one of the charm quarks decays through the D^* channel as usual; and the other decays semileptonically to an electron and other particles. In the vast majority of cases a neutrino is produced with it). These processes are supressed relative to the signal channel for the following reasons:

- The electron is produced in a charmed jet and will usually fail the isolation cuts.
- Both the electron and missing p_T will tend to be lower than those produced from direct W boson decay in the signal channel.

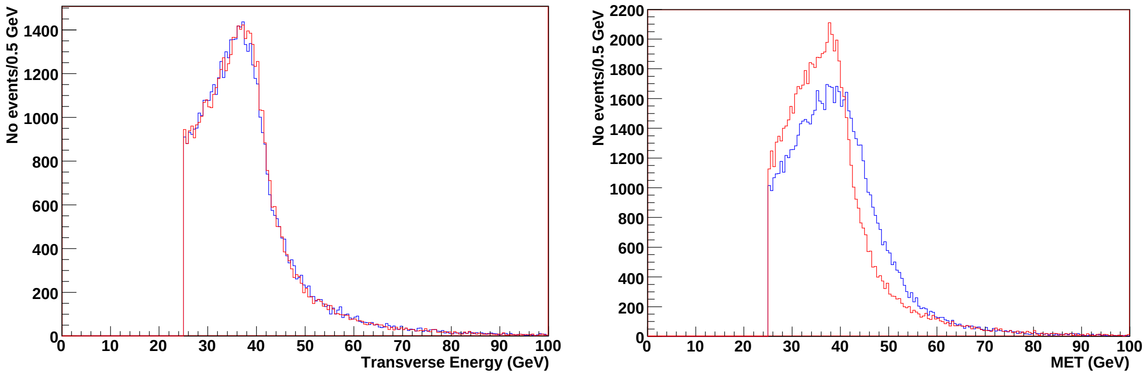
Figure 74 shows the distributions of the electron transverse momentum and the missing transverse energy for $c\bar{c}$ events which decay semileptonically (electrons with $p_T < 5$GeV are not reconstructed in ATLFAST) prior to the application of selection cuts. It is clear that very few events will pass the W boson selection cuts (electron p_T and MET both $>$ 25 GeV, the positions of the cuts are marked on the plots).



(a) Electron p_T (b) MET

Figure 74: Electron p_T and MET distributions for electrons and neutrinos produced in charm decay of $c\bar{c}$ events

Comparison of these distributions with those of the signal sample (Figures 41 and 36(b)) shows that both the electron p_T and MET are significantly lower in $c\bar{c}$ decay. The relevant signal plots (full simulation) are reproduced in Figure 75 for easy comparison (N.B. for electrons $E_T \approx p_T$).



(a) Electron E_T

(b) MET

Figure 75: Electron E_T and MET distributions for electrons and neutrinos produced in signal sample

The electrons in produced in this channel are found almost exclusively with $p_T < 30$GeV. Because of their very different energy spectra to the signal electrons a much lower electron reconstruction efficiency is likely. To investigate this , the relative reconstruction efficiencies in the fast and full simulation were calculated in the same way as in Section 7.5, but for only electrons and positrons with $p_T < 30$ GeV. The relative reconstruction efficiencies in this case are reduced to around 31.2% for positrons and 27.6% for electrons(Table 19). An average electron reconstruction efficiency of 29% will be assumed here, and for other background channels which produce similarly low-p_T electrons.

Sample	Probability of es passing selection cuts in **Fast Simulation** (%)	Probability of es passing selection cuts in **Full Simulation** (%)	Multiplication factor relating fast → full sim
$W^+ \to e^+\nu$	32.1 ± 0.06	10.0 ± 0.1	0.312 ± 0.006
$W^- \to e^-\nu$	37.0 ± 0.07	10.2 ± 0.1	0.276 ± 0.005

Table 19: Comparison of electron reconstruction efficiencies in full and fast simulation for $W \to e\nu$ samples for electrons with transverse momenta below 30 GeV

Because of its large cross-section, in order to be able to reject the $c\bar{c}$ backgrounds to < 10 events per fb^{-1} using the signal selection criteria directly, approximately 5 x 10^{11} would be needed. Generating a sample this large was not feasible in terms of computing time or disk space. Instead 1.75×10^6 $c\bar{c}$ events were generated in Pythia. They were then analysed separately for events which pass first the W boson selection cuts, and then events in which a D^* was sucessfully reconstructed. An estimate of the level of backgrounds was made by assuming that the two conditions are independent and considering the likely W and D^* sign correlations.

52 events passed the W selection cuts, and 18 events passed the D^* selection cuts, 9 D^{*+} and 9 D^{*-} (considering the true D^* events only for simplicity). In combination this implies a probability of one event passing both cuts of 6.1×10^{-10} ($=\frac{52}{1.75\times10^6} \times \frac{18}{1.75\times10^6}$). Then applying the lower p_T estimated electron reconstruction efficiency of 29% this is reduced to 1.8×10^{-10}. There will be 1.42×10^9 $c\bar{c}$ events per fb (from the production cross-section), so the number that would pass the selection cuts if the D^* and electron were produced from different c quarks is ~ 0.25 events per fb $^{-1}$.

Because a leading-order generator was used it is generally accepted that the cross-section could be inaccurate for the LHC by a factor of up to ~ 2, meaning that this limit might be more realistically $\sim< 0.5$ events / fb^{-1}. It is therefore very likely that the total background contribution from $c\bar{c}$ is < 1 event / fb^{-1}.

For reference the mass difference plot for the $c\bar{c}$ sample after the application of D^* selection cuts is shown in Figure 76.

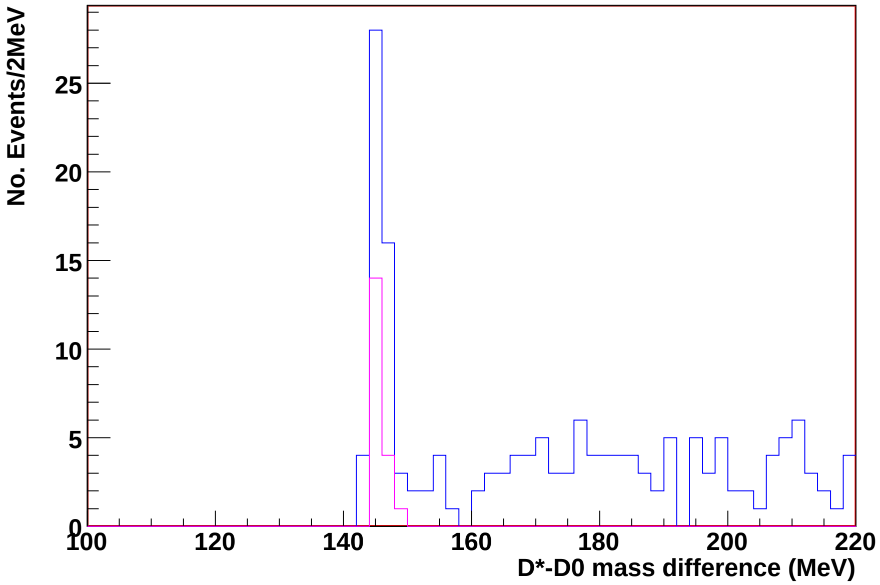


Figure 76: D^*-D^0 mass difference (MeV) calculated for each D^* candidate passing the D^* selection cuts only (no W boson selection cuts) in the $c\bar{c}$, un-normalised. The blue line represents the full sample and pink the true events in which a D^* was produced at generator level.

The background contribution from this channel will be charge-symmetric parton distribution functions for c and $\bar{c}$ in the proton sea are assumed. The limit on the uncertainty to the charge asymmetry contributed by this channel to the signal peak was calculated from the fact that nine D^{*+} events passed the original D^* selection cuts (or between 4.3 and 15.3 at 90% CL) and nine D^{*-} events. The same arguments were used to reduce these numbers to between 0.06 and 0.21 D^{*+} events fb^{-1} at 90% CL. This leads to a number asymmetry in the region of the signal peak of $\mathcal{A} = 0.0 \pm 0.1\%$.

2) The $b\bar{b}$ channel is able to mimic the signal in two ways:

1. If one $b \rightarrow e\nu X$ and the other $b \rightarrow D^*X$ followed by a D^* decay to $K\pi\pi$ as usual. The total probability for this decay chain occurring is 0.097%, or a fractional cross-section of $\sim 3.2\ \mu$b.

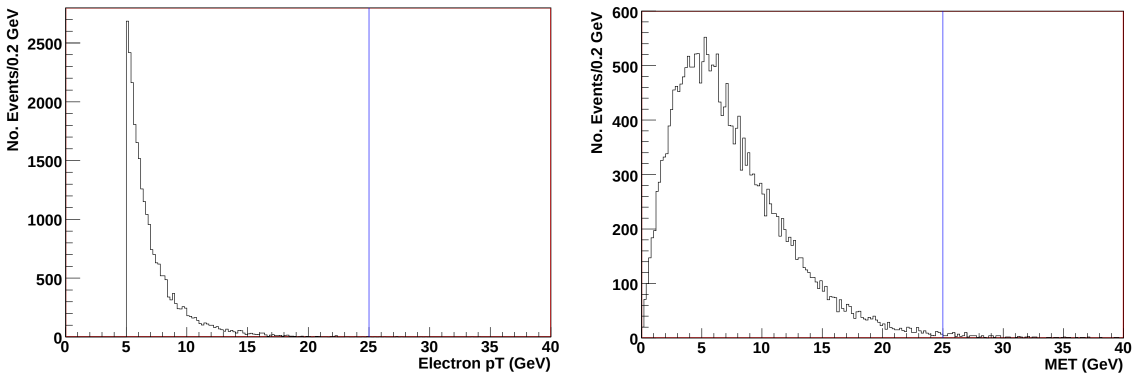
2. If one $b \rightarrow D^{*} e \nu X$ followed by the D^{*} decay to $K\pi\pi$. The other b is free to decay in any fashion. The total probability of this decay chain is then 0.142%, or ~ 4.6 μb.

If the two types cases above have similar probabilities of passing the selection cuts (neglecting any sign correlations) then around 60% of the events that pass selection would be of the second type.

These events are likely to be very strongly rejected because:

- The electron is produced in a b jet and will usually fail the isolation cuts
- Both the electron and missing p_T will tend to be lower than those produced from direct W boson decay in the signal channel, similarly to the $c\bar{c}$ case above. Figure 77 shows the distributions of the electron transverse momentum and the missing transverse energy for $c\bar{c}$ events prior to any selection cuts. Again 29% will be taken as an estimate of the electron reconstruction efficiency.
- B hadrons tend to have long lifetimes and so the impact parameter of the bachelor pion will on average have a much larger impact parameter. than in the signal case where the D^{*} is produced at the primary vertex (see Figure 56).
- In case 1 above the electron will carry the same sign as the D^{*} and the resulting bachelor pion, and thus these events will not usually pass the sign correlation cuts unless the sign of the electron is incorrectly reconstructed. Studies in full simulation suggest that the electron charge is misidentified just 2.35% of the time in the signal sample (see Section 6.2.2) so this will not be common; applying this value will lower the fractional cross-section of case 1 to ~ 75 nb.
- In case 2 the same sign correlations will apply as in the signal. However the D^{*} and electron are produced from the same jet, meaning that they should be found

close together in the detector. An additional cut on the angle between the D^* and electron could be effective in reducing this background to a manageable level. (This was discussed in Section 7.2 but is not currently applied.)



(a) Electron p_T

(b) MET

Figure 77: Electron p_T and MET distributions for electrons and neutrinos produced in charm decay of $b\bar{b}$ events

In this case, to be able to reject the $b\bar{b}$ background to < 10 events per fb^{-1} using the signal selection 1.6 x 10^{12} events respectively would need to be generated (= 10 events × luminosity / $b\bar{b}$ cross section). Again this is an unfeasibly high number of events. 5×10^6 $b\bar{b}$ events (MC@NLO) were created, the events were analysed separately for W and D^* selection, as in the $c\bar{c}$ sample.

22 events passed the W selection cuts and 7 true D^* events passed the D^* selection cuts (4 D^{*+}, 3 D^{*-}), or < 14 events at 90% CL. Together these imply a probability of one event passing both cuts of $2 \times \frac{22}{5\times10^6} \times \frac{14}{5\times10^6} = 2.7 \times 10^{-11}$. Applying an electron reconstruction efficiency of 29% this is reduced to 7.3×10^{-12}. The cross-section of 3.3 mb implies that there are 3.3×10^{12} $b\bar{b}$ events per fb, so the number that would pass the selection cuts if the D^* and electron were produced from different b quarks is ~ 24 events / fb^{-1} at 90% CL.

If $< 60\%$ of the sample passing the events are of the second kind of decay, in which the W and D^* are oppositely signed, < 14 events / fb^{-1} at 90% CL could be expected to pass. However these < 14 events will have come almost exclusively from the same original b quark in order to pass the selection cuts. As a result, it is highly likely that these can be rejected in data by the application of a cut on the angular difference between the D^* and electron (as discussed in Section 7.2) with very little loss to the signal. With such small numbers of events it should be possible to exclude almost all of them. Assuming a fairly conservative 90% rejection rate suggests a remaining background level from these events of < 1.4 events fb^{-1} at 90% CL.

Assuming that these events were sucessfully rejected, the remainder would be the ~ 10 events in which the W and D^* were produced in different jets, and the sign of one of the electron was mis-identified (the sign of the D^* will be well constructed from low-p_T tracks). A mis-identification rate of 2.35% would reject this part of the background to $<$ 0.2 events / fb^{-1}, at 90% CL so it can be neglected.

In total it is then likely that the total background contribution from $\mathbf{b}\bar{b}$ is < 2 events / $\mathbf{f}b^{-1}$.

In order to reject this background with real confidence, studies in data of how the D^* and electron distributions from $b\bar{b}$ decay relate would be required. It is likely that they are reducible down to the $\sim 1\%$ level by judicious selection, but were this not to be the case the background is still less than $\sim 10\%$ of the signal, and the background will be charge-symmetric if symmetric parton distribution functions for c and $\bar{c}$, and b and $\bar{b}$ in the proton sea are assumed.

The mass difference plot for the $b\bar{b}$ samples after the application of D^* selection cuts are shown in Figure 78. The background contribution from this channel should again be charge-symmetric. A number asymmetry was calculated from the fact that four D^{*+}

events passed the original D^* selection cuts (or between 1.5 and 8.6 at 90% CL) and three D^{*-} events (or between 1.1 and 7.4 at 90% CL). The same arguments were used to reduce these numbers to between 0.2 and 0.9 D^{*+} events fb^{-1} at 90% CL, and 0.1 and 0.7 D^{*+} events fb^{-1} at 90% CL. This leads to a number asymmetry in the region of the signal peak of $\mathcal{A} = 0.09\ ^{+0.56}_{-0.65}\%$. As the background is expected to be charge symmetric and the statistics are very low it is reasonable to simplify this to $\mathcal{A} = 0.0 \pm 0.6\%$.

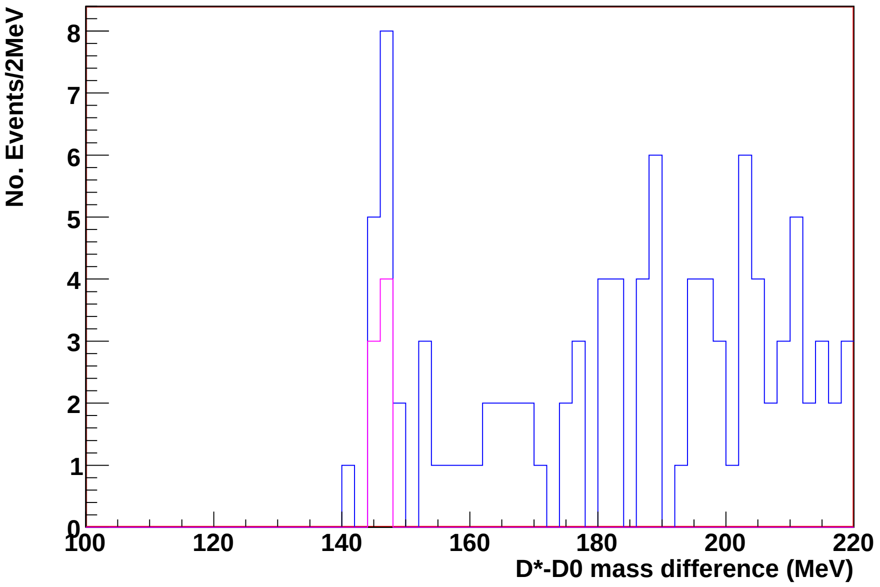


Figure 78: D^*-D^0 mass difference (MeV) calculated for each D^* candidate passing the D^* selection cuts only (no W boson selection cuts) in the $b\bar{b}$ sample, un-normalised. The blue line represents the full sample and pink the true events in which a D^* was produced at generator level.

3) Top quarks invariably decay into W+b, so the $t\bar{t}$ background rejection will follow the same lines as the $b\bar{b}$, excepting that the π_B and D^0 will tend to have large impact parameters, which will cause the events to be rejected. Since the $t\bar{t}$ cross-section is lower than the $b\bar{b}$ by several orders of magnitude, if the $b\bar{b}$ background is tightly constrained it is likely that this background can be discounted without further calculation.

In summary, the total background contribution from $q\bar{q}$ production is estimated to be < 3 events / fb^{-1}. There will be no significant charge asymmetry due to these channels ($\mathcal{A} \approx 0.0 \pm 0.7\%$).

9.2.7 W $\rightarrow q_i\bar{q}_j$

$W \rightarrow q_i\bar{q}_j$ is most likely to form a background to the signal when $W \rightarrow s\bar{c}/c\bar{s}$ or $W \rightarrow c\bar{b}/c\bar{b}$ (the latter decay being Cabibbo supressed).

The sc events are very unlikely to pass the selection cuts because the strange quark jets rarely produce energetic electrons. The main mechanisms for electron production in the s jet is the decay of pions. Charged pions can produce electrons directly, whereas the dominant decay of neutral pions is to a pair of photons which can split into e^+e^- pairs. In both cases the electrons are likely to be formed at quite low energies, and are unlikely to pass either the p_T or the isolation cuts.

The cb events are the most likely of the available pairings to pass the selection cuts, in the case that either of the c or b jets produces a D^* and the other an electron. This is still quite unlikely to be a significant background because the charges of the D^* and electron in this case will be the same as the W, and so will be rejected unless the electronic charge is misidentified. There is also a possibility of the b jet producing both the electron and D^*. This was discussed in the consideration of the background from the direct production of $b\bar{b}$ pairs in Section 9.2.6 above and was found to be unlikely to contribute significantly.

Pythia was used to generate this data set (leading order W production followed by hadronic decay). The cross-section of all combined possible quark-antiquark decay channels is $\sim$ 108 nb, but to reduce computation time, the W boson decay channels were constrained to include only events in which at least one of the quarks was a c, b or t. The cross-section for this reduced set was calculated by Pythia to be 53.9 nb, so this channel produces around 54 million events per fb^{-1}, too many to generate (again, the Pythia cross-section is expected to be within a factor of 2 of the true physical value). Instead 1.75 $\times$ 10^6 events were generated and as in Section 9.2.6 the events were analysed separately for W bosons and D^* mesons. $W \rightarrow s\bar{c}/c\bar{s}$ is the dominant mode in this dataset, and less than 0.3% of the sample is $W \rightarrow c\bar{b}/c\bar{b}$.

12 events passed the W selection cuts and 5 real D^* events passed the D^* selection cuts (3 D^{*+} and 2 D^{*-}), or < 11 at 90% CL. This implies a probability of any one event passing both sets of cuts of 3.9×10^{-11} ($=\frac{12}{1.75\times10^6} \times \frac{11}{1.75\times10^6}$). Applying the lower p_T estimated electron reconstruction efficiency of 29% this is reduced to 1.1×10^{-11}. The cross-section of 53.9 nb implies that there are 5.4×10^8 of these events produced per fb, so the number that would pass the selection cuts if the D^* and electron were produced from different quarks is ~ 0.006 events / fb^{-1}.

The background contribution from this channel is then $\ll 1$ event / fb^{-1}.

This background will be charge-asymmetric since there will be more W^+ than W^- bosons available. Using 90% confidence intervals again, the asymmetry measured here is found to be negligible, $\mathcal{A} \approx 0.00003 \pm 0.00002\%$. The asymmetry is found by subtracting the number of D^{*-} (3) from the number of D^{*+}s (2) leaving 1, then re-weighting this by the cross-section to 0.006 events / fb$^(- 1)$. Dividing this by the size of the signal peak at fb$^(- 1)$, around 155 events (see Section 7.6) gives the number asymmetry contributed to the signal peak in this channel. The maximum asymmetry allowed to 90% CL is found by noting that 2 events is > 0.5 events at 90% CL and 3 events is < 6.7 at 90% CL and recalculating the asymmetry with these numbers.

9.2.8 Z $\rightarrow q_i\bar{q}_i$

The decay of $c\bar{c}$ and $b\bar{b}$ pairs in ways that may mimic the signal channel has already been discussed in 9.2.6. The background from $Z \rightarrow q_i\bar{q}_i$ follows very similar arguments, although they are expected to display slighty different kinematic properties since the $q\bar{q}$ pairs here will be produced by on-mass-shell guage bosons, rather than by virtual intermediates. The MET and electron p_T distributions turn out to follow very similar shapes to the $q\bar{q}$, and so the lower p_T reconstruction efficiency is assumed to be appropriate.

2.17×10^6 $Z \rightarrow q_i\bar{q}_i$ events were generated in Pythia (i.e. LO Z production). The Z boson decay channels were constrained to include only $c\bar{c}$, $b\bar{b}$ and $t\bar{t}$. The cross-section for this was 12.8 nb.

Again, the events were analysed separately for W bosons and D^*s, since it is very much less likely that an event could pass all selection cuts when the e and D^* come from the same jet.

- 644 events passed the W selection cuts
- 40 events passed the D^* selection cuts (again considering the real D^* events for simplicity).
- Together these imply a probability of one event passing both cuts of 1.1×10^{-8} $(=\frac{644}{2.17\times10^6} \times \frac{40}{2.17\times10^6})$
- Applying the lower p_T estimated electron reconstruction efficiency of 29% this is reduced to 3.2×10^{-9}
- The cross-section of 12.8 nb implies that there are 1.3×10^7 of these events produced per fb, so the number that would pass the selection cuts if the D^* and electron were produced from different quarks is $\sim$ **0.04 events / f**b^{-1}

The estimated contribution from this channel is then $\ll$ **1 event / f**b^{-1}. The generation mechanism in this case is charge symmetric and the effect upon the asymmetry measurement will again be negligible.

9.2.9 D* + "fake" electron

The vast majority of events in ATLAS will contain at least one high energy hadronic jet so there will be a high overall rate of D^* production. Detector effects can result in the reconstruction of isolated electrons and missing energy, although will be rare for one such event to contain both a sufficiently energetic electron and also enough MET to pass signal selection cuts. In order to quantify the background from these events a pre-existing fully-simulated sample (CSC sample id. 5802 [86], 600k events) was employed.

Sample 5802 is a fully-simulated sample of mainly of dijets (i.e. including all LO processes which form two hadronic jets) generated with Pythia. It also contains a small number of other processes (LO W, γ, $t\bar{t}$ production, each weighted according to their cross-section). The quarks involved in the hard scattering are each required to have $p_T > 15$ GeV, and at least one jet must be produced with $p_T > 17$ GeV and $|\eta| < 2.7$. The mechanisms included in this sample are given in Table 20 (f indicates a fermion). The calculated cross-section (2.28 mb) is theoretically expected to be accurate to within a factor ~ 2 of the physical value.

Process	**Cross-section (mb)**
$pp \to f\bar{f}, gg, gq, q\bar{q}$	2.28
$pp \to q\gamma, g\gamma$	4.17×10^{-4}
$pp \to f^*\gamma^*$	4.53×10^{-4}
$pp \to W^{\pm}$	1.41×10^{-4}
$pp \to t\bar{t}$	4.91×10^{-7}

Table 20: Interactions included in Sample 5802 and their cross-sections

Because these events are fully simulated the electron selection cuts detailed in 6.2 were applied, with an MET cut of $MET > 25$ GeV, in order to select W bosons. No events passed the W selection cuts, i.e. < 3 events at 90% CL.

Meanwhile, charmed jets were produced in 23% of the events, and it is reasonable to assume that in the majority of cases the presence of a fake electron in the detector will be independent of charm production. Applying the relevant branching fractions (Table 6) suggests that D^*s will be produced in roughly 0.15% of events (this will be an order of magnitude estimate, not considering those from b jets etc.). Applying this probability to the maximum of three fake electrons likely to be found per fb^{-1} drops the likely level of background from this sample to < 0.01 events at 90% CL, i.e. negligible, even without considering the low acceptance for D^* reconstruction. Any contribution to the number asymmetry is then also likely to be neglible.

9.3 Summary of Background Processes

Table 21 indicates the numbers of events per fb^{-1} likely to pass selection cuts in each of the backgrounds discussed above, with a brief description of the channel. Expected charge asymmetries are quoted where relevant. The asymmetry in each case is calculated from Equation 16 and represents the expected percentage excess of passing W^+ events over W^- (neg. is indicated where the asymmetry is negligible). There are 21.1 ± 1.6 events fb^{-1} in the most significant backgrounds, discussed in Sections 9.2.1 and 9.2.2, and a further contribution of < 5.2 events at 90% CL from the backgrounds discussed in Sections 9.2.3-9, leading to a limit on the backgrounds of < 27.9 events fb^{-1} at 90% CL from the combined backgrounds. The total number asymmetry from the backgrounds listed is -6.2 ± 1.9 at 90% CL, found by summing the numbers of each of W^+D^{*-} and W^-D^{*+} events per fb^{-1} given by each channel and then calculating the asymmetry as normal.

Sect.	**Background Description**	**Conditions for mimicking signal**	**Generator**	**No. events/fb^{-1} passing cuts (to 90% CL)**	**Number Asym.** $\mathcal{A}$(%)
7.3	NLO $W \to e\ \nu$ $s+g \to W+D^*$	True Signal	MC@NLO	W^+: 44.6 ± 0.7 W^-: 41.1 ± 3.3	3.4 ± 4.2
9.2.1	NLO $W \to e\nu$ from d, b diagrams	Same products and kinematics as signal but no strange quark	MC@NLO (included in signal above)	W^+: 6.4 ± 0.5 W^-: 11.1 ± 0.9	-4.5 ± 1.5
	$q_a\ q_b \to W\ g$ $g \to c\bar{c}\ (b\bar{b})$	g splits → charm jet. (D* p_T lower than signal.)			
	$q\ g \to W\ b$	b → cX, c will have lower p_T than in signal.			
9.2.2	NLO $W \to \tau\nu$ $\tau \to e\nu\nu$	Contains true signal, but backgrounds dominate	MC@NLO	3.6 ± 0.2	-1.7 ± 0.4
9.2.3	NLO $Z \to e^+e^-$	One electron lost, mis-idendified as MET.	MC@NLO	< 2.2	0.00 ± 0.02
9.2.4	NLO $Z \to \tau^+\tau^-$	One $\tau \to e\nu\nu$ other not detected	inferred from above	$\ll$1	*neg.*
9.2.5	WW	One $W \to e\nu$ other $W \to cX$	HERWIG	$\ll$1	*neg.*
	ZZ	One $Z \to e^+e^-$, other $Z \to c\bar{c}$ one electron lost			
	WZ	$W \to e\nu$ with $Z \to c\bar{c}$ or $W \to cX$ with $Z \to e^+e^-$			
9.2.6	$c\bar{c}$	One $c \to e\nu X$ Other $c \to D^*X$	Pythia	<1	0.0 ± 0.7
	$b\bar{b}$	One $b \to e\nu X$ other $b \to D^*X$ or one $b \to D^*e\nu X$	MC@NLO	<2	
	$t\bar{t}$	$t \to Wb$ 100% of the time (similar to $b\bar{b}$)	Pythia	$\ll$1	
9.2.7	LO $W \to c\bar{s}/s\bar{c}$	$c(\bar{c}) \to D^*X$, electron produced in a quark jet	Pythia	$\ll$1	*neg.*
	LO $W \to c\bar{b}/b\bar{c}$	$q \to e\nu X$ other $q \to D^*X$ or one $b \to D^*e\nu X$			
9.2.8	LO $Z \to c\bar{c}/b\bar{b}$	One $q \to e\nu X$ other $q \to D^*X$ or $b \to D^*e\nu X$	Pythia	$\ll$1	*neg.*
9.2.9	D^* + fake W	D^* from q/g jets, e/MET from detector effects	HERWIG Full Sim	$\ll$1	*neg.*
	Total			<27.9	-6.2 ± 1.9

Table 21: List and description of background processes to the signal

10 Discussion and Conclusions

The total size of the signal peak, including both sign combinations and using realistic electron reconstruction efficiencies, is 111 ± 28 events per fb^{-1}, with 55 ± 19 events measured in each of the W^+D^{*-} and W^-D^{*+} cases. Using Equation 16 this gives a measured number asymmetry in the range 0.0 ± 34.5%.

Clearly for one fb^{-1} of data the measurement uncertainties will dwarf any measurements of the strange sea. If a measurement of the strange sea is to be attempted, then significantly more than 1 fb^{-1} of both real data and Monte Carlo simulation of the channels $W \rightarrow e\nu$ and $W \rightarrow \tau\nu$ would be needed. A simple estimate suggests that 10 fb $^{-1}$ would bring the measurement uncertainty on the individual W^+ and W^- signal peaks to around 10% and 100 fb^{-1} to 3% (around a year of data taking at high luminosity). With 10 fb^{-1} or more of data a Gaussian fit could be used to measure the signal peak, so improving reducing the uncertainty further. If 300 fb^{-1} of data was used, and the statistics were improved by the incorporation of the channel $W \rightarrow \mu\nu$ (assuming this could be reconstructed with a similar efficiency), the measurement uncertainty could be reduced to around 1%. Of course, this study was based upon W bosons produced at NLO, and as mentioned in Section 4 NNLO corrections are expected to alter the W production cross-sections by up to ∼3%. In order to reduce the measurement uncertainty to around 1%, careful measurements of the W boson cross-section at ATLAS would be needed.

The total number of background events likely to minic the signal and pass selection cuts is fewer than 27.9 events per fb^{-1} at 90% CL. The number asymmetry from the backgrounds is -6.2 ± 1.9 at 90% CL, which is well below the level of the measurement uncertainty on the peak in 1 fb^{-1} of data.

The most significant backgrounds to the signal are those which most closely mimic the signal; the NLO production of $W \rightarrow e\nu$ processes described in 9.2.1 in which a real W boson and D^* are produced but are not created from the interaction of a strange quark and gluon. These backgrounds are all inherently charge asymmetric. The current Monte Carlo statisics are too small to very accurately estimate the levels of backgrounds from these individual channels, but together they account for 17.5 ± 1.4 events per fb^{-1} at 90% CL and contribute a number asymmetry of -4.5 ± 1.5%.

The next most significant background channel is NLO production of W bosons followed by their decay to $\tau\nu_\tau$. This contributed a small number of events which pass the selection cuts, 3.6 ± 0.2 $D^* + W$ events in total per fb^{-1}, but for kinematic reasons the events which pass signal selection are mostly generated from the same background channels as above rather than the signal $s + g$ channel. The set of passing events from the $W \rightarrow \tau\nu_\tau$ sample is also likely to be charge asymmetric. This channel will contribute a number asymmetry of -1.7 ± 0.4% to the reconstructed peak.

Other backgrounds which are likely to contribute measurably to the reconstructed peak if the statistics are increased by around two orders of magnitude are the production of a $Z + jet$ followed by the decay of the Z to e^+e^- and the loss of one of these electrons in the detector (< 2.2 events per fb^{-1}), and production of pairs of $c\bar{c}$ and $b\bar{b}$ pairs (< 1 event per fb^{-1} and < 2 events per fb^{-1} respectively). None of these channels are expected to contribute measurably to the number asymmetry. The rejection of these $q\bar{q}$ backgrounds is based upon the assumption that an event is much more likely to pass if the electron and D^* were produced in different jets, which assumption must be confirmed in studies on ATLAS data, and larger-scale simulation studies, in future in order to confirm that these backgrounds can be controlled sufficiently to measure the signal channel to an acceptable level of accuracy.

In simulation, a ten-fold increase in statistics would provide a better estimate of the fraction of the peak in the $W \rightarrow e\nu$ sample which was produced from background channels to the signal, and an increase to 100 fb^{-1} would enable the accurate measurement of the charge asymmetries that will arise from the differences in kinematics of the background W^+ and W^- producing channels, which would be necessary to extract a reliable measurement of any extra asymmetry from the strange sea. This amount of simulated data would take several months to generate, but were this measurement to be attempted using physical data it would be vital to do so. Equivalently larger simulated samples should also be generated for the NLO $W \rightarrow \tau\nu$ channel to measure its likely contribution to the number asymmetry at higher statistics, and, as mentioned above, large samples of ATLAS dijet data would be needed to test whether the $c\bar{c}$ and $b\bar{b}$ backgrounds can be sufficiently constrained.

As was mentioned in Section 2.4, the approach to extracting the background information in data would be different, since the background levels from each channel will be combined. There is also likely to be a significant increase in the levels of random occurances of uncorrelated reconstructed Ws and D^*s in the real detector (as opposed to the fast simulation), forming a charge-symmetric background to the signal. The numbers of same-sign WD^* events which pass all selection cuts should be plotted, and there will be a significant peak at in the $D^* - D^0$ mass difference plot. This can be seen in the $W \rightarrow e\nu$ datasets used here, Figure 79, but as was shown in Section 9.2.1 the backgrounds included in this sample were not predominantly symmetric and so this method was not used in this Monte Carlo study. (N.B. the equivalent plot for the opposite sign events was shown in Figure 63.)

ATLFAST does not accurately represent the reconstruction efficiency of electrons in the real detector. Results of studies in the full detector simulation, GEANT4 (Section 6) have been used to estimate the efficiency of electron reconstruction (69% for the signal sample, 29% for the electrons produced in several background channels with transverse momenta $\lesssim$ 30 GeV) and the rate of electron charge misidentification ($\approx$ 2.35%), but it would be necessary to measure these more accurately in data, including their dependence on electron η and p_T if measurements of the strange sea were attempted using this method. Other necessary measurements include the rate of electronic charge misidentification, the reconstruction rate of low p_T tracks and missing energy resolution. The levels of combinatoric background are also likely to be significantly higher in data than ATLFAST due to the higher multiplicity of track candidates likely to be identified in the physical detector, which can be studied in data.

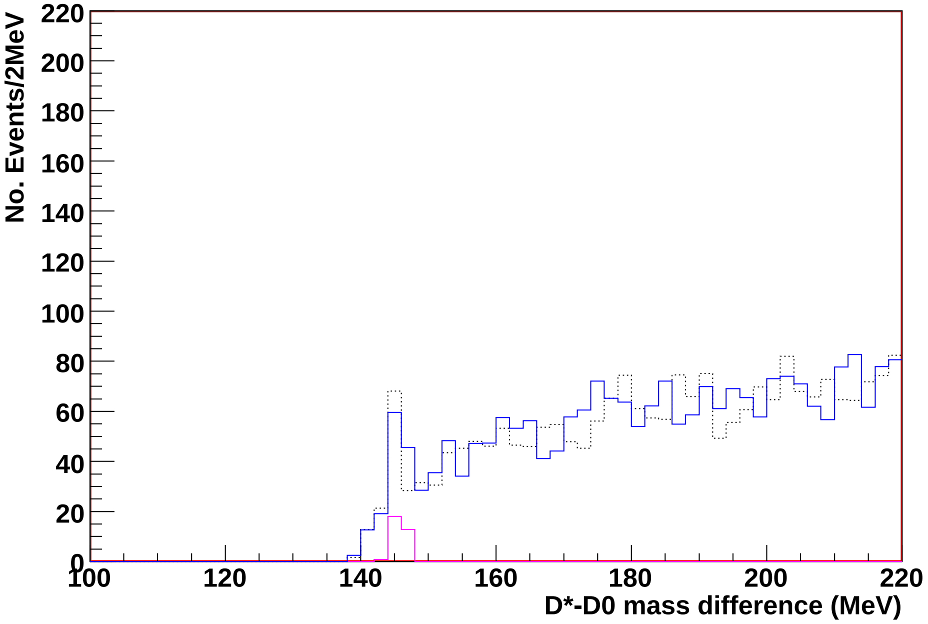


Figure 79: $D^* - D^0$ mass difference δm (MeV) calculated for each D^* candidate passing all selection cuts, except that the reconstructed D^* and W have the same sign. The blue line represents the full sample, black the detector truth for the full sample, and pink the true events in which a D^* was produced with a W at generator level (detector-smeared). The plots are normalised to represent the number of events that would be found in inverse femotobarn of data.

It would be a significant challenge to look for an asymmetry in the strange sea components in this channel due to its small cross-section and the low efficiency of the D^* selection process, but it remains likely to be the most promising channel for this study in ATLAS. Bearing in mind that the sea components of the proton carry momentum fractions on average of $x <\sim 0.01$, and that W bosons formed in the acceptance region of ATLAS are produced from partons with $10^{-4} < x <\sim 10^{-1}$ (as described in Section 5.1), Figures 5 and 6 suggest an overall dominance of $\bar{s}(x)$ over $s(x)$ is likely (which should correspond to a negative number asymmetry).

Figure 6 suggests that the maximum possible number asymmetry of $\bar{s}$ over s is around 7% at some specific values of x. Should an asymmetry exist it is likely that it could be significantly smaller than this. Even at 300 fb^{-1} the measurement uncertainty on the signal peak will be 1% or more, suggesting that this method would be insensitive to asymmetries of $< 1\%$ at the expected LHC luminosity. Measuring this channel is therefore likely to be most useful in improving current constraints on a strange sea asymmetry, but could possibly detect a positive strange sea asymmetry.

The asymmetries quoted here are upon the total number of $W + D^*$ events found within the ATLAS acceptance region for accurate track reconstruction ($|\eta| < 2.4$). Since any strange sea asymmetry which exists will have a dependency upon the value of x at which the quarks are produced, which itself relates to the rapidities at which the W and D^* are formed, it would be useful to measure this asymmetry as a function of the rapidities of the outgoing particles. This could be matched with various sets of Monte Carlo data generated from different asymmetric strange PDFs (for example those used in [37]) to give an idea of the most likely protonic strange sea distribution functions. The statistics in this Monte Carlo study are not be high enough to make differential asymmetry measurements over a range of rapidities possible.

11 Summary

Monte Carlo simulations were used investigate the possibility of placing tighter limits on the strange sea parton distribution functions at ATLAS, and to test whether if an asymmetry exists in the strange proton sea it is likely to be detectable at ATLAS. The strange proton sea is not currently well constrained at low-x, and a non-zero asymmetry between the strange and antistrange components of the sea is indicated in theory and in some experimental data. Several important measurements of SM cross-sections will depend upon better knowledge of the strange sea.

The mechanism ($s + g \rightarrow W + c$) was selected to study the strange quark sea of the proton, because it is sensitive to the strange and antistrange sea components (it relies upon the same gluon distribution for both charge conjugates). The decays of the W boson into an electron and neutrino and the c quark into a D^* were chosen as their reconstruction is relatively free of background. W boson production at the LHC is sensitive to the low x regions of the proton sea, and so this channel will probe the most useful parts of the strange proton PDFs.

Approximately 2.5 million events of each charge conjugate of $W \rightarrow e\nu$ were generated at NLO, representing around 1 fb^{-1} of data. The ATLAS fast simulation ATLFAST was used to approximate the effects of the detector on the sample. Reduced electron reconstruction efficiencies were included as a more realistic estimate of the effects of the detector on electrons. Clear peaks were constructed in each of these channels, representing the combined reconstruction of a W boson and D^* meson of opposite signs via a set of selection cuts. The size of the resulting signal peak was 111 ± 28 events per fb^{-1}, 55 ± 19 events per fb^{-1} measured in each of the W^+ and W^- cases, leading to a number asymmetry in the range 0.0 ± 34.5% at 90% CL.

The likely level of additional background from channels not already included in the $W \to e\nu$ sample is estimated from Monte Carlo simulation studies to be < 27.9 events per fb^{-1} at 90% CL. The number asymmetry contribution indicated from the background studies is -6.2 ± 1.9% at 90% CL. However, extensive studies of the backgrounds would be needed in real data, in particular the very high cross-section $c\bar{c}$ and $b\bar{b}$ channels, would be needed before it could be concluded that the uncertainty from the backgrounds could be controlled to a manageable level.

For one fb^{-1} of data the measurement uncertainties will dwarf any attempted measurements of the strange sea and certainly no information on strange sea asymmetries could be gained. 300 fb^{-1} would be required to extract useful information on the properties of the strange sea. At this level, should the largest asymmetry allowed in the meson cloud model and in fits to previous data occur it should also be possible to measure some number differences between $s(x)$ and $\bar{s}$. Since the strange sea asymmetry varies with x, the most information could be gained from extracting the number asymmetry as a function of rapidity. Very little is known about the true strange sea distributions at low-x values, so whether such an asymmetry is seen or not this would be a valuable channel to study in terms of placing tighter constraints on the what is known of the nucleonic strange sea parton distribution functions.

References

[1] The ATLAS Homepage: *http://atlas.web.ch/Atlas/Welcome.html*

[2] P. Nadolsky *Strange, charm and bottom flavors in CTEQ global analysis* arXiv:0809.0945 [hep-ph] (2008)

[3] The NuTeV Collaboration *A Precise Determination of Electroweak Parameters in Neutrino-Nucleon Scattering* Phys.Rev.Lett. 88 (2002) 091802; Erratum-ibid. 90 (2003) 239902(2003)

[4] The LEP Collaborations *A Combination of Preliminary Electroweak Measurements and Constraints on the Standard Model* CERN-EP/2001-98, hep-ex/0112021.

[5] Y. Ding and B. Ma *Contribution of asymmetric strangeantistrange sea to the PaschosWolfenstein relation* Phys. Lett. B, 590, 216 (2004)

[6] F. Wei and B. Zou *The $s - \bar{s}$ asymmetry in the nucleon and "NuTeV" anomaly* Phys. Lett. B 5 (2008) 501

[7] B. Portheault *Strange Sea Asymmetry from Global QCD Fits* hep-ph/0406226v1 (2003)

[8] S. Eidelman et al. *Review of Particle Physics* Physics Letters B 592, 1 (2004)

[9] Y. Sumino, *"Coulomb + linear" form of the static QCD potential in operatorproduct expansion* Phys. Lett. B 595 (2004) 387392

[10] The CTEQ Collaboration *Handbook of perturbative QCD* http://www.phys.psu.edu/ cteq/handbook/v1.1/handbook.pdf

[11] CLEO *http://www.lns.cornell.edu/Research/EPP/CLEO/*

[12] NA48 *http://www.etap.physik.uni-mainz.de/na48_eng.php*

[13] BaBar *http://www-public.slac.stanford.edu/babar/*

[14] HERA *http://adweb.desy.de/mpy/hera/*

[15] D. Soper, *Parton Distribution Functions* hep-lat/9609018v1 (1996)

[16] Robin Devenish and Amanda Cooper-Sarkar, *Deep Inelastic Scattering*, Oxford University Press

[17] CTEQ Collaboration: Pumplin, J., et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis* JHEP 0207, 012 (2002), [hep-ph/0201195]

[18] K. Gottfried, Phys. Rev. Lett. 18, 1174 (1967).

[19] P. Amaudruz et al., *Gottfried sum from the ratio* F_2^n/F_2^p , Phys. Rev. Lett. 66 2712 (1991)

[20] M. Arneodo et al., Phys. Rev. D 50, R1 (1994)

[21] A. Baldit et al., *Study of the isospin symmetry breaking in the light quark sea of the nucleon from the Drell-Yan process*, Phys. Lett. B322, 244 (1994)

[22] E.A. Hawker et al., *Measurement of the Light Antiquark Flavor Asymmetry in the Nucleon Sea*, Phys. Rev. Lett. 80, 3715 (1998)

[23] R. S. Towell et al., *Improved measurement of the* $\bar{d}/\bar{u}$ *asymmetry in the nucleon sea*, Phys. Rev. D 64, 052002 (2001)

[24] K. Ackerstaff et al., *Flavor Asymmetry of the Light Quark Sea from Semi-inclusive Deep-Inelastic Scattering* , Phys. Rev. Lett. 80, 3715 (1998)

[25] A. M. Cooper-Sarkar *Low-x physics at and W and Z production at the LHC* [hep-ph/0512228]

[26] Durham University On-line Plotting and Calculation, http://durpdg.dur.ac.uk/hepdata/pdf3.html

[27] J.W. Rohlf, *Modern Physics from A to Z0* Chapter 18, Wiley, 1994

[28] A. I. Signal and A. W. Thomas, *Calculation of Quark Distribution Functions Using Bag-model Wave Functions*, Phys. Rev. D40, 2832 (1989)

[29] S. Kumano, *Flavor Asymmetry of Antiquark Distributions in the Nucleon* Phys. Rep. 303, 183 (1998)

[30] J. Speth and A. W. Thomas, Adv. Nucl. Phys 24, 83 (1998)

[31] R. Vogt, *Physic of the Nucleon Sea Quark Distributions* Prog. Part. Nucl. Phys 45, S105 (2000)

[32] A. W. Thomas, *A limit on the pionic component of the nucleon through SU(3) flavour breaking in the sea*, Phys. Lett. B126, 97 (1983)

[33] A. Signal and A. W. Thomas, *Possible Strength Of The Nonperturbative Strange Sea Of The Nucleon* Phys. Lett. B191, 205 (1987)

[34] S. J. Brodsky and B. Q. Ma, *The Spin and Flavor Content of Intrinsic Sea Quarks* Phys. Lett. B381, 317 (1996) hep-ph/9707408

[35] H. R. Christiansen and J. Magnin *Strange/Anti-strange Asymmetry in the Nucleon Sea* Phys. Lett. B 445 (1998), hep-ph/9801283

[36] F. G. Cao and A. I. Signal *The quark-antiquark asymmetry of the strange sea of the nucleon* Phys. Lett. B, 559, 229 (2003)

[37] F. Olness et al., *Neutrino Dimuon Production and the Strangeness Asymmetry of the Nucleon* hep-ph/0312323v2 (2004)

[38] CERN public images server: http://cdsweb.cern.ch/collection/Photos

[39] ATLAS Collaboration, *Technical Design Report Volume 1*, ATLAS TDR 14, CERN/LHCC 99-14 (1999)

[40] V. Mitsou *The ATLAS Transition Radiation Tracker* hep-ex/0311058v1 (2003)

[41] P. Krieger *The ATLAS Liquid Argon Calorimeter* 2005 IEEE Nuclear Science Symposium Conference Record N21-5

[42] ATLAS Collaboration, *Calorimeter Performance Technical Design Report*, CERN/LHCC 96-40 (1996).

[43] ATLAS Collaboration *Calorimeter Performance Technical Design Report* CERN/LHCC/96-40, ATLAS TDR 1 (1996)

[44] ATLAS Muon Collaboration *Muon Spectrometer Technical Design Report* CERN/LHCC/9722 (1997)

[45] The ATHENA Homepage, *http://atlas.web.cern.ch/Atlas/GROUPS/ SOFTWARE/OO/architecture/*

[46] G. Barrand et al., *GAUDI - A software architecture and framework for building HEP data processing applications*, Comp. Phys. Comm. 140 (2001) 45

[47] J. Boudreau and V. Tsulaia, *The HEPVis/v-atlas FAQ*, *http://boudreau.home.cern.ch/boudreau/v-atlas-hepvis.htm*

[48] The Atlantis Homepage, *http://atlantis.web.cern.ch/atlantis/*

[49] M. Dobbs et al., Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics hep-ph/0403045 (2004)

[50] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M.H. Seymour and B.R. Webber, *HERWIG 6.5* JHEP 0101 (2001) 010 [hep-ph/0011363] [hep-ph/0210213]

[51] J. M. Butterworth, J. R. Forshaw and M. H. Seymour, *Multiparton Interactions in Photoproduction at HERA* CERN-TH/95-82, MC-TH-96/05, UCL-HEP 96-02, Zeit. fr Phys. C72 (1996) 637-646 [hep-ph/9601371]

[52] T. Sjstrand et al., *High-Energy-Physics Event Generation with PYTHIA 6.1* [hep-ph/0010017]

[53] T. Sjostrand, *Jet Fragmentation of Multiparton Configurations in a String Framework* Nucl. Phys. 248, 469 (1984)

[54] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP 0206 (2002) 029 [hep-ph/0204244]

[55] S. Frixione, P. Nason and B.R. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, JHEP 0308 (2003) 007 [hep-ph/0305252]

[56] E.Boos et al. CompHEP Collaboration, *CompHEP 4.4: Automatic computations from Lagrangians to events*, Nucl. Instrum. Meth. A534(2004), p250 [hep-ph/0403113].

[57] A.Pukhov and et al., *CompHEP - a package for evaluation of Feynman diagrams and integration over multi-particle phase space. User's manual for version 3.3* [hep-ph/9908288]

[58] Comphep home page: *http://comphep.sinp.msu.ru*

[59] M.L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP 0307:001,2003, hep-ph/0206293

[60] S. Agostinelli et al., *GEANT4 - A Simulation Toolkit*, Nuclear Instruments and Methods A 506 (2003) 250-303

[61] E. Richter-Was, D. Froidevaux, L. Poggioli, *ATLFAST 2.0 - a fast simulation package for ATLAS*, ATLAS Internal Note, ATLAS-PHYS-98-131 (1998)

[62] T. Cornelissen, Doctoral Thesis *Track Fitting in the ATLAS Experiment*, CERN-THESIS-2006-072, (2006)

[63] D. Rousseau, *ATLAS reconstruction software*, Eur.Phys.J C33 (2004)

[64] The ATLAS Collaboration *https://twiki.cern.ch/twiki/bin/view/Atlas/SlidingWindowClustering*

[65] The ATLAS Collaboration *https://twiki.cern.ch/twiki/bin/view/Atlas/TopologicalClustering*

[66] *ATLAS WIKI: Jet Reconstruction in ATLAS* , https://twiki.cern.ch/twiki/bin/view/Atlas/JetS

[67] G. Blazey et al., *Run II Jet Physics*, hep-ex/0005012v2 (2000)

[68] S. Frixione and B. Webber *The MC@NLO 3.1 Event Generator* CavendishHEP05/09 GEFTH6/2005

[69] L. Gilbert *Generation using MC@NLO for DC3* http://www-pnp.physics.ox.ac.uk/∼gilbert/McAtNlo.html

[70] Golonka, Piotr and Was, Zbigniew *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, Eur. Phys. J. C45 (2006) 97-107, hep-ph/0506025

[71] S. Jadach, Z. Was and J.H. Kuehn *TAUOLA - A library of Monte Carlo programs to simulate decays of polarized leptons* Comp. Phys. Commun. 64 (1991) 275

[72] F. Petriello and K. Melnikov, *The W boson production cross section at the LHC through* $O(\alpha_s^2)$, Phys. Rev. Lett. 96, 231803, (2006)

[73] S. Alekhin *The NNLO predictions for the rates of the W/Z production in* $p(\bar{p})p$ *collisions*, hep-ph/0307219 (2003)

[74] ATLAS Collaboration, *Technical Design Report Volume 2*, ATLAS TDR 15, CERN/LHCC 99-15 (1999)

[75] L.Gilbert *DC3 Samples* $W\rightarrow(e/\mu)/nu$ *at Next to Leading Order* http://www-pnp.physics.ox.ac.uk/~gilbert/MCatNLOWnote.pdf

[76] *Egamma ATLAS Reconstruction and Performance*, https://twiki.cern.ch/twiki/bin/view/Atlas/EgammaATLASRecoPerformance

[77] S. Mattig, Doctoral Thesis *Determination of Electron Trigger Efficiencies using* $Z^0 \rightarrow e^+e^-$ *Events at the LHC*

[78] *ATLAS WIKI: ElectronIDIsEM*, https://twiki.cern.ch/twiki/bin/view/Main/ElectronIDIsEM

[79] *Performances of the Calorimeter Modules for Electrons and Pions* http://atlas.web.cern.ch/Atlas/SUB_DETECTORS/TILE/testbeam/summary_paper/particles/electr

[80] B. Lund-Jensen, on behalf of the ATLAS Liquid Argon Electromagnetic Calorimeter Group, *The Status of the Construction and the Performance of the ATLAS Liquid Argon Electromagnetic Calorimeters*, IEEE Transactions On Nuclear Science, Vol. 51, No. 5, October 2004 2181

[81] Werner Witzeling for the ATLAS Collaboration *The ATLAS Detector - a Status Report* Presented at the XIVth International Workshop on High Energy Physics and Quantum Field Theory, Moscow, Russia, 27 May - 2 June 1999. http://theory.sinp.msu.ru/ qfthep04/99_/pro99/Witzeling.pdf

[82] Christophe Clément, On behalf of the ATLAS Liquid Argon Community *ATLAS Electromagnetic Calorimeter Testbeam Results* http://lappweb.in2p3.fr/Calor2000/Contributions/Ionization/clement.pdf

[83] Ron Madaras, *ATLAS Electron ID*, LBL ATLAS Group Physics Meeting, http://www-atlas.lbl.gov/physics/Ron_Atlas_EMID.pdf (May 27 2005)

[84] Ambreesh Gupta, *Missing Transverse Energy Reconstruction in the ATLAS Calorimeter* AIP Conf. Proc. 867 (2006) 453-459

[85] Michael Leyton, *Measuring the pseudorapidity density and transverse momentum spectrum of charged particles at 14 TeV* Atl-Co-Phys-2007-051 (2007)

[86] T. Hollins *SCT Hybrid Testing and the Production of Direct Photons in the ATLAS Experiment at the LHC* (http://www.ep.ph.bham.ac.uk/publications/thesis/tih_thesis.pdf)

[87] N. Carrer and A. Dainese *Charm and beauty production at LHC* http://arxiv.org/PS_cache/hep-ph/pdf/0311/0311225v1.pdf