Studies of Jets, Subjets and Higgs Searches with The ATLAS Detector

Sarah J. Baker

UCL

University College London

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

Abstract

Measurements are made on data collected at a centre-of-mass energy of 7 TeV using the ATLAS detector at the LHC. Jets from non-collision sources for the inclusive jet cross-section are evaluated for jet transverse momentum values from 20 GeV to 1.5 TeV and for rapidity values up to $|y| < 4.4$. Measurements are made using anti-k_T jets with the distance parameters R = 0.4 and R = 0.6, corresponding to 37 pb^{-1} of data. Analysis of the sub-structure of QCD jets at high transverse momentum is performed using anti-k_T R = 1.0 and Cambridge-Aachen R = 1.2, using 35 pb^{-1} of data. The first measurement of jet mass is presented.

Acknowledgements

Contents

List of Figures

9

List of Tables

Chapter 1

The Standard Model

1.1 Introduction

Particles and the interactions that govern them are understood in terms of the Standard Model. The theory was developed from the premise that fundamental particles may be represented in terms of quantisable fields. By imposing invariance under certain symmetries, a renormalisable Lagrangian can be constructed that encapsulates the dynamic and kinematic properties of the component fields. As an interesting consequence of Noether's theorem [1], the restrictions imposed by symmetry gives rise to associated conservation laws. These laws (such as conservation of energy, momentum and charge) are directly observed in nature.

The Standard Model is formally represented as a gauge theory of the SU(3) x SU(2) x U(1) symmetry groups. These correspond to the strong, weak and electromagnetic mixing sectors respectively. The electromagnetic interaction binds together atoms and molecules, and dictates the behaviour observed in electromagnetic fields. The weak interaction underpins processes involved in some radioactive decays of nuclei, such as beta decay, and the decays of nucleons and heavier quarks. The strong force binds quarks and gluons into protons and neutrons, and is responsible for the multitude of hadronic particles that have been observed in high energy processes. Gauge invariance corresponds to invariance under local transformations; all known particles can be understood in terms of degrees of freedom of fields which obey one (or more) of these symmetries. The fourth known force, gravity, is not included in the Standard Model.

Figure 1.1: The Particle Content of the Standard Model.

No quantum field theory has yet been developed which successfully incorporates gravity into a quantum field theory in a satisfactory manner. Whist this omission highlights the need for deeper theoretical insight, the comparatively low strength of gravity with respect to the other forces to some extent justifies this approach - gravity becomes largely insignificant when probing very high energy scales.

The Standard Model describes 12 fermions of spin 1/2, grouped into three leptons and quark families (as shown in Figure 1.1). Force is mediated by the gauge bosons; gluons, $W^{\pm}$, Z and photons. The non-Abelian (non-commuting) properties of SU(3) x SU(2) give rise to gluon, W^+ and W^- self-interactions. Mass is introduced to the Z and W bosons through electroweak symmetry breaking. This process requires the introduction of an additional scalar field and ultimately predicts the existence of an additional particle, the Higgs Boson.

There are several limitations apparent in the Standard Model. It does not, for example, explain the values of various parameters that must be determined from experiment, such as particle mass and charge. However, its success is based on testable predictions and grants the Standard Model recognition as one of the finest achievements in modern

science. Validation of the Standard Model and confirmation of the underlying process responsible for electroweak symmetry breaking are some of the most important tasks for the Large Hadron Collider (LHC), based at CERN, Switzerland. The existence of a particle consistent with a Standard Model Higgs Boson with a statistical significance of 5 sigma[1] was announced by CERN on July 4th 2012 [2, 3].

This thesis outlines contributions to some of the first Standard Model measurements undertaken using the ATLAS detector, at the LHC. An introduction to the framework of Quantum Chromodynamics (QCD) is provided in Chapter 1, with focus on the theoretical ideas behind jets given in Chapter 2. Chapter 3 outlines the operational running and setup of the ATLAS detector over the data taking periods of 2009 to 2012. The technical considerations of jet reconstruction in ATLAS are discussed in Chapter 4. Measurements of the inclusive jet cross-section are described in Chapter 5, with extension into jet Sub-structure measurements in Chapter 6.

1.2   The Theory of the Strong Interaction

Quantum Chromodynamics is the mathematical framework describing strong interactions, namely those between quarks and their associated force carriers, the gluons.

1.2.1   Development of QCD

By the late 1950's bubble chamber experiments had given rise to a plethora of seemingly 'fundamental' particles, which were organised using a technical framework called the 8-Fold Way [4]. Quantum Chromodynamics arose from the need to understand the wide spectrum of observed phenomena using a more meaningful conceptual framework. The evolution of QCD is an interesting facet of particle physics; however it is informative to consider the results with respect to the broader concepts to which they allude, rather than in terms of the path of historical development. Thus, here chronology is largely ignored in a discussion of some of the important experimental and theoretical

---

[1]5 sigma corresponds to the chance of random occurrence being less than 99.99995, and is taken as the bench-mark value for the discovery of a new particle or phenomenon.

results obtained in the last century.

The QCD particle content of the Standard Model comprises of spin 1/2 fermions and spin 1 bosons, the quarks and gluons respectively. There is no experimental evidence to date to suggest that these particles are comprised from more fundamental constituents. In the quark picture the proton is described as a bound state of two up quarks and a down quark and the neutron as a bound state of two down quarks and an up quark. The existence of the particles such as Δ^{++}, with the $3^{\rm rd}$ component of angular momentum $J_3 = 3/2$ and a charge of $+2$ requires three up quarks with identical spin quantum numbers. Such a combination of fermions violates the Pauli Exclusion Principle and necessitates the introduction of the strong charge, or 'colour'. The colour charge is labelled red, green and blue such that (anti)quarks can take three colours: (anti)red, (anti)green and (anti)blue.

Experimental support for the postulated quark-like structure of baryonic matter was found in deep inelastic scattering measurements, which indicated that particles (primarily leptons) scattered against point-like, spin 1/2 constituents of the nucleons [5]. The presence of spin 1 gluons was inferred from consideration of the total momentum shared by the quarks. Within the Standard Model gluons form an octet under SU(3) and exist in a superposition of colour and anticolour states.

1.2.2   Renormalisation, Asymptotic freedom and Confinement

For calculations of QCD cross-sections above the first order of perturbation theory (leading order), loop diagrams are introduced. Calculations then require integration over the loop momentum and these integrals diverge in the limit of infinite momentum, yielding an infinite result. To give meaningful results which can be compared directly with experiment a process called renormalisation is applied. This process absorbs the infinities into physical quantities; however, it also introduces a renormalisation parameter u_r, dependent on the renormalisation scale. Thus the process introduces an uncertainty, which must be factored in to the theoretical calculation. Ultimately, this method permits the construction of an explicit form of the coupling constant α_s, which now

scales with energy. The gluon self-interaction means that at low momentum transfer Q^2 this coupling becomes very large, a property referred to as 'confinement'. Furthermore, strong coupling weakens with increasing Q^2 resulting in 'asymptotic freedom'. A direct experimental consequence of confinement is that quarks and gluons are never observed in isolation, and are instead observed as collimated showers of particles. This leads to a complex and interesting phenomenology, described more fully in Chapter 2. At low energies quarks combine as hadrons, bound-state colourless quark triplets and doublets known as baryons and mesons respectively. However, at high energies and with sufficient momentum transfer quarks can be freed from hadrons. Increased separation between quarks causes the potential in the gluon field connecting the quarks to rise and once this potential is strong enough more quarks, anti-quarks and gluons are created in a process known as 'fragmentation'. The new particles act to screen the initial quarks from each other, and reduce the potential between the initial quarks. Eventually the partons combine into hadrons via 'hadronisation'.

1.2.3   Parton Density Functions and Factorisation

Asymptotic freedom means that at high energies quarks can be modelled by point-like constituents in the parton model [6]. Deep inelastic scattering experiments provided a means of understanding the quark composition of hadrons. Nucleons are now known to be composite objects, comprising of the valence quarks (which determine the quantum numbers of the hadrons), gluons, and a sea of quarks that are created and annihilated via quantum fluctuations. The momentum transfer, Q^2, between the probe particle and the target hadron is representative of the distance scales resolvable by the probe particle. The total momentum of the proton is shared among these constituent partons where the distribution of the momentum fraction x carried by each parton is expressed as a probability that a particular parton has a given x. A structure function is used to encapsulate the lack of knowledge of the nucleon structure and kinematics of the collision. The momentum distribution for a quark or gluon is called parton distribution function, or PDF. Experimentally x depends on Q^2 (Figure 1.2), which can be regarded as a result of the increased resolving power or the number of partons that can influence the probe, either from gluons splitting into $q\bar{q}$ pairs or radiation from
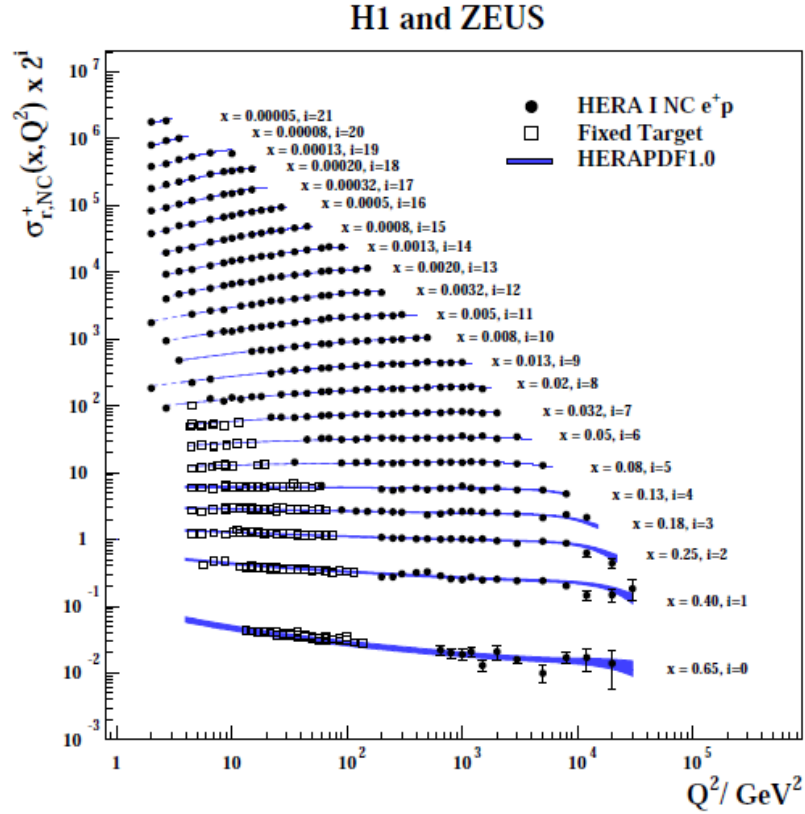
Figure 1.2: HERA e+p reduced cross-section and fixed-target data as a function of Q^2. The error bars indicate the total experimental uncertainty [7].

quarks. The `DGLAP` formalism [8] uses splitting functions to perturbatively calculate the renormalised parton densities with the Q^2 dependence. The `DGLAP` formalism gives information on the evolution of the PDF, however the shape of the PDF must be modelled using both experimental data on the structure functions and parameterisations from consideration of the properties of the nucleons they describe. Whilst PDFs are not themselves observable, they can be extracted from combinations of measurements for independent structure functions.

It is possible to consider the perturbative evolution of quarks and gluons independently of the soft non-perturbative effects such as hadronisation, fragmentation and PDFs through a process of factorisation. Factorisation absorbs infinities, as in the case of renormalisation, and introduces a scale uncertainty u_f^2 which must be incorporated

into the final theoretical calculation. Factorisation has been mathematically proven for many processes, and the separation of the hard underlying process and the soft contribution is a powerful analytical tool. In short, perturbation theory models the hard process, however PDFs must be incorporated to compare to the full theoretical prediction.

Chapter 2

The Theory of Particle Jets

2.1 Introduction

At high energies Quantum Chromodynamic processes produce quarks and gluons, often in large numbers. These particles are never observed directly; a quark, or gluon will fragment and then hadronise into a collimated spray of particles, as described in Chapter 1. This collection of particles can itself be considered as an object, referred to as a 'jet' of particles. Figure 2.1 is a dijet event recorded by ATLAS and shown on the ATLAS event display, ATLANTIS [9]. Visualisations such as this demonstrate intuitively what is meant by the term jet; however, a more rigorous approach is required in order recover precise information about the underlying hard process. Defining what, exactly, constitutes a jet is one of the main theoretical challenges of high energy particle physics.

This chapter discusses several methods of defining jets that are commonly used by the ATLAS experiment and for the work in this thesis. It also outlines the ideas behind 'sub-structure analysis', a method which utilises the internal structure of a jet in order to extract further useful information. Relevant to this thesis are the jet sub-structure variables, which provide discrimination between different decay topologies and particle types.
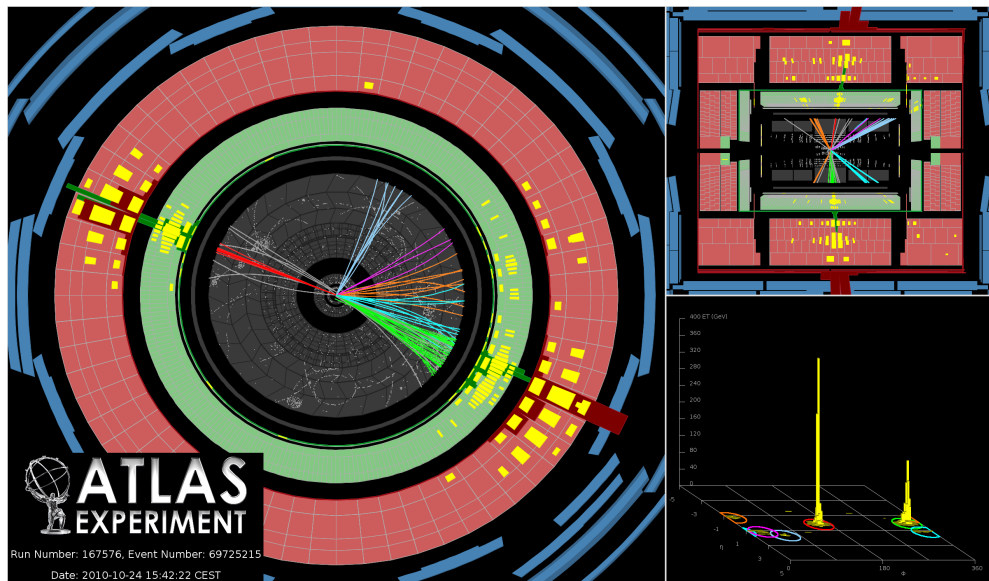
Figure 2.1: The highest mass central dijet event and the highest p_T jet collected by the end of October 2010. The two central jets have an invariant mass of 2.6 TeV and the highest p_T jet is 1.3 TeV.

2.2 Defining Jets

To construct jets out of a group of constituent particles there must be a set of rules to determine which particles are considered to belong to the same jet. Furthermore, there needs to be some specification to define how the four-momentum of the input particles are combined to give the four-momentum of the final jet object.

The Les Houches accords [10] specify a jet definition to contain the following:

- A jet algorithm to define how to map the constituent four-momenta into jets.

- Full specification of the jet algorithm parameters.

- A recombination scheme to determine the final four-momentum of a jet from its constituents.

Conceptually, a jet is an object produced using a particular set of these specifications.

2.3   Further Requirements for Jets

An effective jet definition must also fulfil other nominal requirements. It must be simple to implement in an experimental analysis, or theoretical calculation. It must be defined at any order of perturbation theory, and yield finite cross-sections at any order. These cross-sections must not be largely influenced by differences in modelling of the hadronisation process. Additional infinitely soft particles radiated by the primary partons should not modify the result of jet finding (infra-red safety), and jet finding should be insensitive to particles radiated at very small angle with respect to the initial parton (collinear safety). Furthermore, the input jet algorithm should reproduce the same physics for the final state irrespective of the inputs to the jet, whether they are partons, particles or calorimeter-level objects.

2.4   Types of Jet Algorithm

Jet algorithms are in general divided into two distinct types: cone algorithms and successive recombination algorithms.

Cone algorithms reconstruct jets in accordance the direction of energy flow. For example, the ATLAS cone algorithm takes a constituent above a specified energy cut-off (seed) inside a cone in η, ϕ space with radius R. The four-momenta of all particles that fall within R are summed to obtain an initial (proto-jet) direction. The cone is then redrawn using the proto-jet direction as the starting point, and the four-momentum of the constituents within a radius R are used to obtain a new jet direction. This process is iterated until the jet direction does not change significantly in successive iterations; the jet is then considered stable. Finally a 'split-merge' process is applied. A number of constituents may be overlapping; these constituents must be either assigned to one of the two jets. If they are not assigned to an individual jet, the two jets are merged together. Soft parton emissions may affect this process, thus the ATLAS Cone algorithm is not infra-red safe. In many ATLAS analyses cone algorithms are considered obsolete and the approach of successive recombination is preferred.

Successive recombination algorithms iteratively merge constituents according some

distance parameter, such as distance in space or p_T. Equations 2.1 and 2.2 show the most common type used in ATLAS:

$$d_{ib} = p_{Ti}^{2\alpha} \tag{2.1}$$

$$d_{ij} = min(p_{Ti}^{2\alpha}, p_{Tj}^{2\alpha})\frac{\delta R_{ij}^2}{R^2} \tag{2.2}$$

where $\delta R_{ij}^2 = \sqrt{(\delta\phi_{ij}^2 + \delta y_{ij}^2)}$. The values p_{Ti}, y and ϕ are the transverse momentum, rapidity and azimuth of the ith four-vector respectively. R is the algorithm distance parameter. Initially, d_i and d_{ij} are calculated for all constituents i and j. If the minimum for all values of d_{ij} and d_i is the value of d_{ij} then objects i and j will be combined by summing their four-momentum. If d_{ib} is the minimum value then i is declared to be a final jet and removed. This process is iterated until no particles remain.

Setting $\alpha = 1$ gives the k_t algorithm [11], whereas $\alpha = $ -1 is the anti-k_T algorithm [12]. Choosing $\alpha = 0$ removes the p_T dependence to give the angular ordered Cambridge-Aachen algorithm [13]. All three cases fulfil the requirements set out in Section 2.3. The anti-k_T algorithm tends to cluster lower p_T (softer) particles around a higher p_T (harder) core. The final jets tend to be conical in shape, which can be useful when correcting the jets for experimental effects (such as correcting for energy losses found in jets that are found in poorly instrumented regions of the detector).

2.5   Sub-structure in Jets

The LHC operates at energies significantly above the electroweak scale and thus interesting physics channels often produce particles with electroweak scale masses and a significant Lorentz boost. The decay products of boosted particles tend to be highly collimated and therefore it can be useful to reconstruct them within a single jet. The internal structure of this jet may then contain information about physics which took place above the hadronisation scale but below the p_T scale of the hard process. This information can be recovered by applying sophisticated jet reconstruction techniques,

such as sub-structure analysis.

The use of jet sub-structure is motivated by the need to improve jet mass resolution in increasingly high-energy and luminosity environments, where effects such as pile-up (multiple proton-proton collisions, as discussed in Chapter 3) tend to degrade the hadronic signal. The method dates back to Seymour, who considered a Higgs boson decaying to WW [14]. Two methods were utilised. In the first, jets were made with the inclusive k_t algorithm, with R = 1. A wide radius is used to ensure the decay products of the event are fully captured. The final step of the clustering procedure is reversed to recover the two 'sub-jets' from the W decay products, which were combined to form the final jet. The difference in η, ϕ space (ΔR) between these sub-jets was then used to define a smaller R. Clustering was then performed on the jet constituents using the smaller R. The second method used a cone algorithm with a (small) value R = 0.25 to resolve the two sub-jets of the decay seperately during the initial jet clustering stage. However, with the latter technique, the gluon radiation from $W \rightarrow q\bar{q}$ was lost resulting in poorer mass resolution. Seymour observed that the first method of unclustering a large R jet and re-clustering the constituents with a smaller R parameter meant that any radiation from the decay could be preserved in the form of a third sub-jet. Further studies by Butterworth, Cox and Forshaw [15] showed that in WW scattering, the k_t distance $\sqrt{d_{12}}$ (described in Section 2.5.2) between the two W sub-jets was close to the W mass in W decays, but tended to have lower values in QCD jets. This variable could therefore offer a means of discrimination between signal and background in this channel.

In 2008, the Cambridge-Aachen algorithm was considered as a means to identify low mass Higgs decays [16]. A low mass Higgs particle decays preferentially to b-quarks. The Cambridge-Aachen algorithm uses sequential recombination with respect to increasing angular distance. The angular separation of two b-quarks produced in a Lorentz boosted Higgs decay is largely dependent on the mass of the Higgs particle. Higgs production via an associated vector boson is both kinematically challenging and subject to high backgrounds at the LHC. Until recently this production mode had been largely disregarded as a viable search channel for the Higgs for ATLAS. However, if a boosted Higgs is considered, the backgrounds rates are comparatively reduced and the

efficiency of identifying b hadrons is improved. By reconstructing the $H \to b\bar{b}$ system as a single jet, sub-structure techniques can then be used to recover the signal.

2.5.1   Splitting and Filtering Tools

A software tool was developed as part of this thesis to perform sub-structure analysis within the ATLAS framework. The procedure involves reversing the Cambridge-Aachen clustering process to identify the point at which two relatively low mass objects were combined to make a significantly more massive object. This point corresponds to the decay of a heavy particle. The mass resolution of the decay products can then be improved by re-clustering using a smaller R parameter such that the number of jet constituents originating from underlying event and pile-up (defined in Chapter 3) are reduced.

The process takes a Cambridge Aachen jet j, and reverses the final clustering step when two objects (the sub-jets) j_1 and j_2 were combined to make j. The mass of j_1 is defined as the greater of j_1 and j_2. If j cannot be unclustered (i.e. if it is a single particle or the distance δR between j_1 and j_2 ($\delta R_{j1,j2}$) is less than 0.3) then it is not a suitable candidate, so j is disregarded and the process is stopped. The cut-off $\delta R_{j1,j2}$ corresponds to the angle for which resolving the sub-jets becomes difficult due to the angular resolution of the calorimeter. If the difference in the mass of j_1 to the overall mass of the jet j passes a defined cut-off and the k_t difference between j_1 and j passes a defined cut-off then the process is stopped. If not j is redefined as j_1 and the process is repeated from the beginning until these criteria are fulfilled. Finally, the constituents of j are re-clustered using the Cambridge-Aachen algorithm with an R-parameter of $R_{filt} = \min(0.3, \delta R_{j1,j2}/2)$ to produce n sub-jets $s_1, s_2 \ldots s_n$ ordered in descending p_T. The three highest p_T subjets are kept:

$$j = \sum_{i=1}^{\min(n,3)} s_i.\qquad\qquad(2.3)$$

The mass cut-off, and k_t cut-off are tuned to low mass Higgs searches to be 0.67 and

0.09 respectively. When the mass cut-off is reached, it signifies a significant mass drop and is identified as the two-body decay of a heavy object. The k_t symmetry requirement between the two sub-jets acts to suppress splittings from quarks and gluons, which tend to decay more asymmetrically in k_t than those of a Higgs decaying to b-quarks.

Studies by Butterworth et. al established the sub-structure method as a competitive means of identifying Higgs decays in ATLAS. Figure 2.2 shows the mass distribution of sub-jets from a 126 GeV $H \rightarrow b\bar{b}$ (black and red) at 8 TeV and 14 TeV using the same technique. The Higgs is produced with an associated vector boson Z or W (denoted V). The V decays leptonically as either a) $V \rightarrow l\bar{l}$, b) $V \rightarrow l\nu$ or c) $V \rightarrow \nu\bar{\nu}$. Since neutrinos are not reconstructed directly in ATLAS the third channel is always measured experimentally as missing transverse energy E_T (Chapter 3 provides further detail on neutrino reconstruction). Considering jets with $p_T > 200$ GeV reduces the cross-section of this channel by 95%. However, the remaining high p_T events have several advantages for mass reconstruction. The decay system is heavier and therefore more likely to be found within the central detector; identifying b-quarks is more effective in this region due to the coverage provided by the tracker. At such high transverse momentum the decay products from the Higgs are often highly collimated and can be reconstructed more effectively using a single jet. Sub-structure techniques can then be applied to improve the mass resolution of the signal.

Studies are performed using the Rivet software [17] and with Monte-Carlo samples produced using the HERWIG event generator, as described in Chapter 3. Jets are created from all particles (excluding neutrinos) using Cambridge-Aachen with R = 1.2. A candidate lepton is defined as one that has $p_T > 20$ GeV and $|\eta| < 2.5$. These values are based on experimental trigger thresholds and efficiencies during 8 TeV data collection. Channels a), b) and c) are equated to 2, 1 and 0 charged leptons respectively. For channel a) it is required that the leptons are oppositely charged, with a total lepton $p_T > 200$ GeV and and invariant mass close to that of the Z (between 80 GeV to 100 GeV). For channel b) it is required that there is missing transverse energy consistent with the presence of a W Boson with $p_T > 200$ GeV. For channel c) it is required that there is missing energy of greater than 200 GeV. The $t\bar{t}$ background is high in b) and
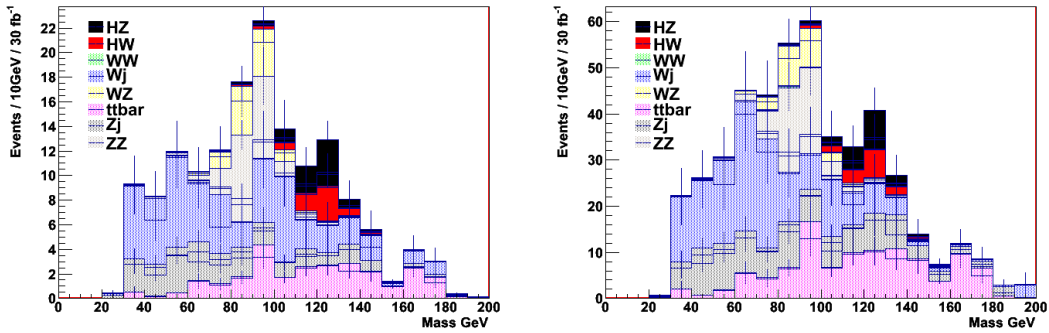
Figure 2.2: Mass of the hardest three subjets in HERWIG Monte-Carlo events for $H \rightarrow b\bar{b}$ with associated Z or W decaying leptonically and backgrounds at a centre-of-mass energy of 8 TeV (left) and 14 TeV (right). Plots are normalised to 30 fb^{-1}.

c); a further veto is applied in these two channels to remove all events with anti-k$_T$ 0.4 jets in the event which have a p_T above 30 GeV and do not overlap with the lepton or higgs candidates. This veto has a large cost in terms of statistics; however, signal to background ratios are improved significantly. Further improvements can be obtained through fine-tuning and other optimisations, such as increasing the Higgs p_T cut, and reducing the value of the Cambridge-Aachen R parameter.

Sub-structure analysis is applied to the candidate Higgs jet. The two hardest sub-jets are identified as originating from b-quarks by a geometrical match with parton level b-hadrons in the event. These jets are weighted according to an estimated efficiency that a b-hadron is correctly identified (b-tagged) or that another particle incorrectly identified as a b-hadron (mis-tagged) in the detector. This conservatively estimated as 60% b-tag rate and 0.02% mis-tag rate. Figure 2.2 shows the mass distribution of the three hardest sub-jets.

2.5.2 k_t Splitting Scales

The k_t splitting scales are defined by re-clustering the constituents of the jet with the k_t recombination algorithm.

$$\sqrt{d_{12}} = min(p_{Tj1}, p_{Tj1}) \times \delta R_{j1j2}, \tag{2.4}$$

where j_1 and j_2 are the two jets before the final clustering step. $\sqrt{d_{23}}$ is calculated in a similar fashion but for the penultimate clustering step. The variable $\sqrt{d_{12}}$ can used to distinguish the heavy particle decays (which tend to be symmetric) from those produced by asymmetric QCD splittings. The expected values are $m/2$ for heavy particles and $\lesssim 20$ GeV for QCD, although QCD splittings tend to have a tail which can extend to high values.

2.5.3   N-Subjettiness

For a jet k, the N-subjettiness variables [18] τ_N are defined as:

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} min\{\Delta R_{1,k}, \Delta R_{2,k}, ..., \Delta R_{N,k}\} \tag{2.5}$$

$$d_0 = \sum_k p_{T,k} R \tag{2.6}$$

where $\Delta R_{i,k} = \sqrt{(\delta\eta)^2 + (\delta\phi)^2}$ is the distance from sub-jet i to constituent k.

Jet constituents are clustered using using the exclusive k_t algorithm into N sub-jets. The quantity τ_N is effectively is a measure of how localised the constituents are to the axes of each sub-jet N.

Figure 2.3 shows a schematic of typical QCD and W jet splittings and the typical calorimeter cell energies in an η, ϕ plane for the W jets and QCD jets. Unlike the QCD jets, the W jets tend to be composed of two distinct regions of energy. Cells are coloured according to how the k_t algorithm divides the cells into two sub-jets. The marker size for each calorimeter cell is proportional to the logarithm of the particle energies in the cell. The open square shows the overall jet direction and the open circles show the two sub-jet directions. The ratio τ_2/τ_1 (τ_{21}) can then be considered as a measure of the alignment of the jet energy to the open square compared to that of the open circles. its is therefore a direct indication of how well the jet can be decomposed into two component sub-jets. Similarly, the ratio τ_3/τ_2 (τ_{32}) is used to describe how well a jet can be decomposed into three sub-jets. Figure 2.4 shows distributions of τ_1, τ_2 and τ_{21} for W and QCD jets. It can be seen that the ratio of the two variables gives
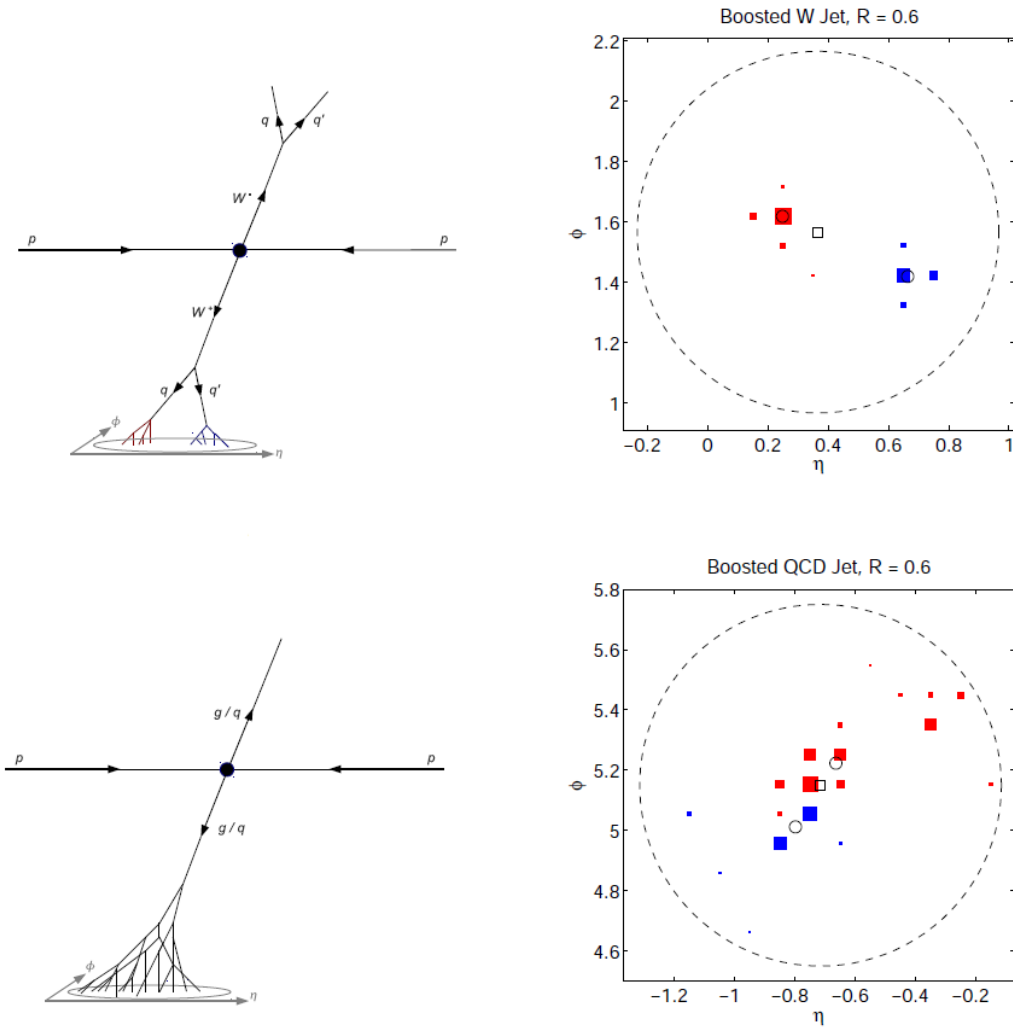
Figure 2.3: Schematic of the fully hadronic decay splittings in W jets (top left) and dijet (bottom left) events and typical event displays showing calorimeter energy distribution for anti-k$_T$ jets with R = 0.6 for W jets (top right) and QCD jets (bottom right) [18].

greater separation, and therefore greater discriminating power between these two types of event.
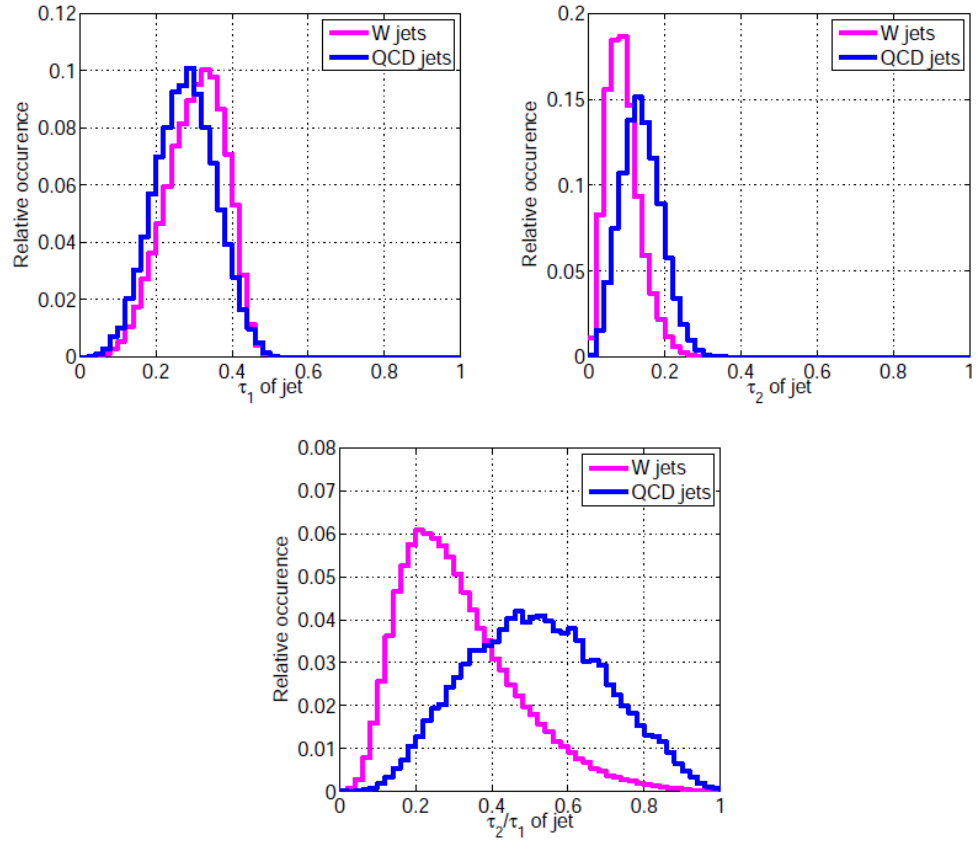
Figure 2.4: Distributions of τ_1 (left) and τ_2 (centre) and τ_2/τ_1 (right) for boosted W and QCD anti-k$_T$ jets with R $= 0.6$. Jets are taken in an invariant mass window of 65 GeV $<$ jet$_{\text{mass}}$ $<$ 95 GeV on jet with $p_T > 300$ GeV and $\eta < 1.3$ [18].

Chapter 3

The ATLAS Experiment

3.1   Introduction

The Large Hadron Collider (LHC) is designed to accelerate and collide protons at energies up to the Terascale. The properties of fundamental particles produced in these proton-proton collisions drive the design of particle detectors.

This chapter outlines the design of the ATLAS detector, focusing on the components that are most important for jet reconstruction. The Inner Detector (ID) (Section 3.3.1) performs vertex reconstruction and tracking for charged particles. The calorimeters (Section 3.3.2) are designed for energy measurements. Events are selected using a purpose-built trigger system (Section 3.4) and reconstructed using the ATLAS offline software (Section 3.5).

3.2   The LHC

CERN is the most advanced particle physics centre in the world and hosts the LHC, the highest-energy particle accelerator ever built. In 2009 it collided protons at a centre-of-mass energy of 900 GeV. In March 2010 it reached the record breaking centre-of-mass energy of 7 TeV, and in April 2012 the LHC ran at a centre-of-mass energy of 8 TeV. At design specifications, the LHC is expected to accelerate beams of protons to a centre-of-mass energy of 14 TeV and to reach luminosities of 10^{34}cm^{-2}s^{-1}. Under these conditions beam crossings occur at 25 ns intervals, with O(20) interactions per

bunch crossing.

Beams of protons are circulated in opposite directions within two separate vacuum tubes. The electromagnetic charge of the protons enables bunches of protons to be bent using a series of superconducting dipole magnets, and then focused using quadrupole magnets. The quadrupole magnets act to collide the two beams at four interaction points. The LHC hosts four detectors which are situated at one of the four interaction points. ATLAS is one of two general purpose detectors located on the LHC ring. At the points where the beam crosses, proton-proton interactions can occur. The main interaction of interest occurring between partons within the proton is described as the 'hard' process. However, along with the particles produced in the hard process are additional particles from the proton remnants, as well as radiation emitted from incoming and outgoing particles. These two contributions are known collectively as 'underlying event'.

The number of interactions expected per proton-proton bunch pair is calculated from the cross-section σ and instantaneous luminosity $\mathcal{L}$. The cross-section is a measure of the effective scattering area, given in units of barns (b) 10^{-24}cm^2 and instantaneous luminosity is measured in units of b^{-1}s^{-1}. The rate R is given by:

$$R = \frac{dn}{dt} = \mathcal{L} \times \sigma \qquad (3.1)$$

Instantaneous luminosity is given by:

$$\mathcal{L} = \frac{n_b f_r n_1 n_2}{2\pi \sigma_x \sigma_y} \qquad (3.2)$$

where n_b is the number of bunches per beam, f_r is the rotation frequency of the beam, n_1 and n_2 are the numbers of particles in the two colliding bunches, and σ_x and σ_y characterise the widths of the beam profile in the horizontal and vertical directions respectively.
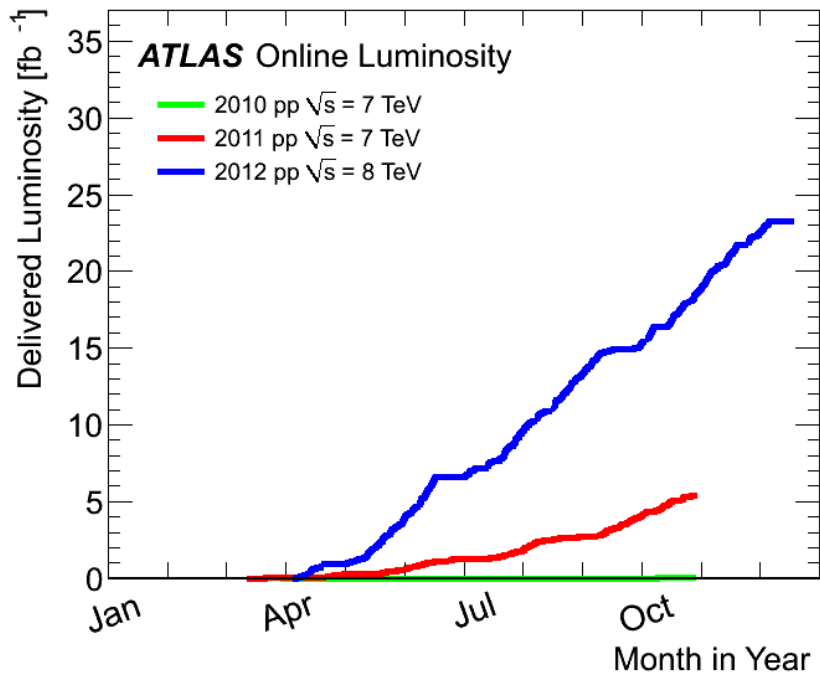
Figure 3.1: The cumulative online luminosity versus day delivered to ATLAS during stable beams and for proton-proton collisions. This is shown for 2010 (green), 2011 (red) and 2012 (blue) running.

An increase in the instantaneous luminosity can be obtained by reducing the area of the beam, or increasing the number of protons per bunch, or the number of bunches in the beam. Improvements in luminosity tend to increase a background effect known as 'pile-up', caused by multiple interactions happening in the same event as the process of interest. Reducing the beam area or 'squeezing' the beam creates additional interactions per beam crossing, an effect known as 'in-time' pile-up. An increase in the number of protons per bunch or number of bunches leads to a decrease in the time between each interaction, and can lead to contamination from consecutive bunches, or 'out-of-time' pile-up. Additional out-of-time pile-up can occur due to overlap between multiple subsequent bunches, known as 'bunch trains'. During initial running conditions pile-up was negligible. In May 2010, the beam size was decreased, which led to an increase in the number of events with more than a single proton-proton event by 10%.

Figure 3.1 shows the integrated luminosity for 2010, 2011, and 2012. The integrated
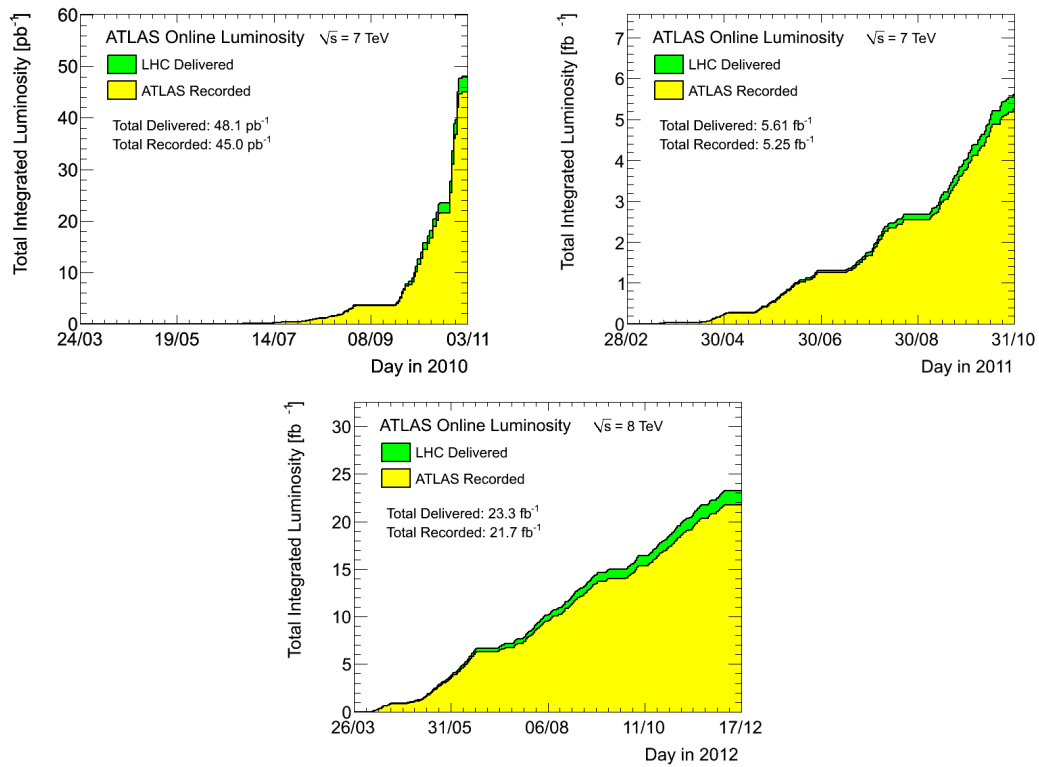
Figure 3.2: The cumulative online luminosity versus day delivered to ATLAS during stable beams and for proton-proton collisions. This is shown separately for 2010 (top left), 2011 (top right) and 2012 (bottom).
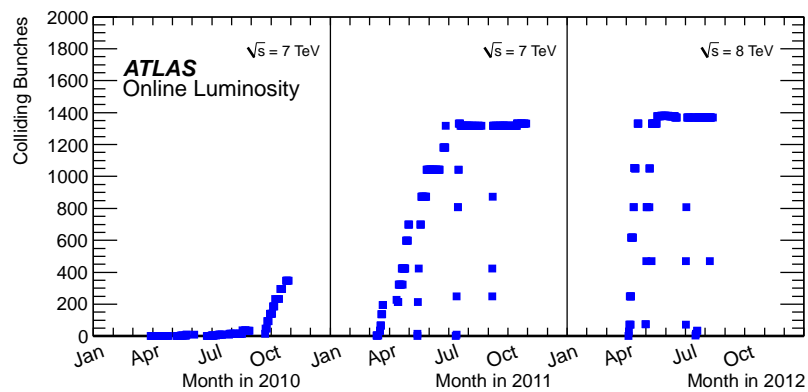
Figure 3.3: The number of colliding bunches in ATLAS versus time during the proton-proton runs of 2010, 2011 and 2012.
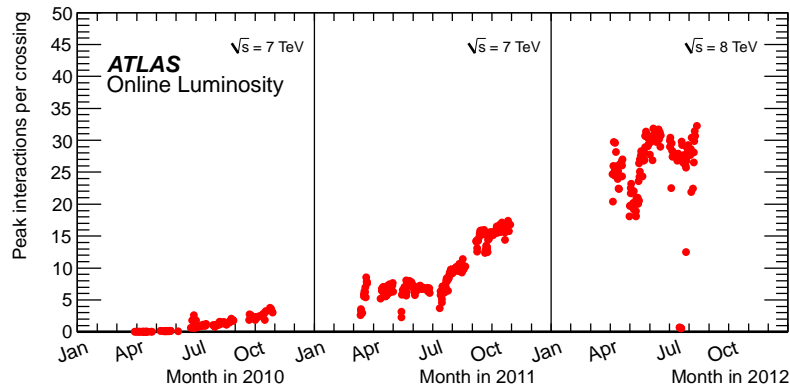
Figure 3.4: Maximum pile-up levels in ATLAS versus time during the proton-proton runs of 2010, 2011 and 2012.

luminosity for these data taking periods are shown individually in Figure 3.2. The delivered luminosity is the luminosity recorded from the start of 'stable' beams in the LHC, until ATLAS is turned off to allow for beam studies or repair. The luminosity is determined from counting rates measured by the luminosity detectors. The body of work included in this thesis covers the data taking periods on 2010 and 2011 only; however, it is nonetheless interesting to observe the trends found in data to 2012. Figure 3.3 shows the number of increasing bunches being collided across data taking periods and 3.4 shows the number of pile-up interactions per week (the maximum value when both beams are stable). Figures 3.1 to 3.4 are taken from the ATLAS public results webpage. Many processes that are to be measured at the LHC have cross-sections that are relatively small. Therefore a high integrated luminosity is required in order to produce enough events to give a statistically significant result. The LHC will run over several years in order to achieve the required total luminosity needed to study rarer processes. At optimal luminosity conditions, the high numbers of particles produced in collisions impose a formidable challenge to the commissioning and design of ATLAS. The calorimeter and Inner Detector sub-systems must be radiation hard, with high granularity.
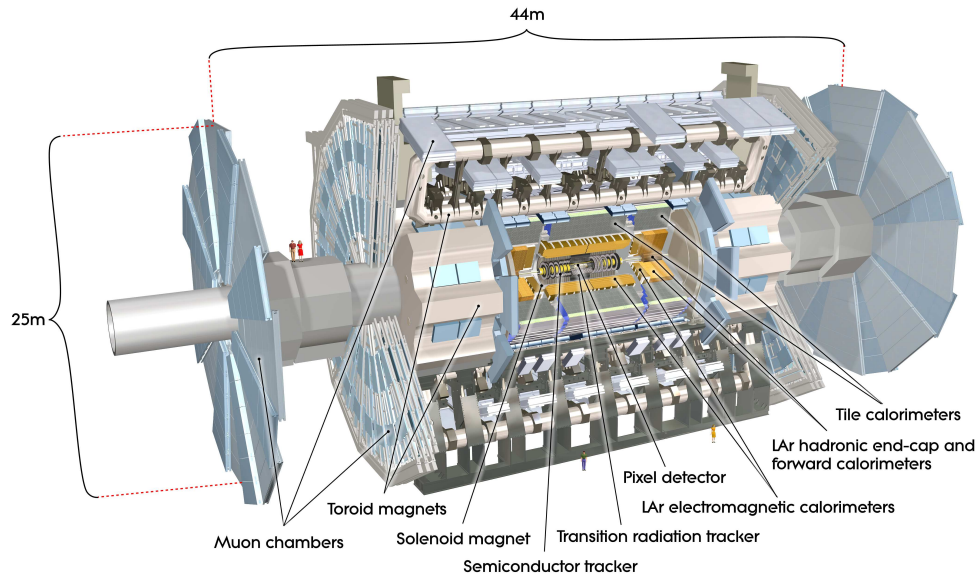
Figure 3.5: An illustration of the ATLAS detector, showing the layout of the Inner Detector, calorimeter systems and muon system

3.3   ATLAS

ATLAS is described using a cartesian co-ordinate system, where the x-axis points to the middle of the LHC ring, and the y-axis points vertically upwards and the z-axis is aligned with the beam direction. The azimuthal angle ϕ and the polar angle θ are taken with respect to the beam. The r coordinate describes the radial distance with respect to the interaction point.

Rapidity is defined by:

$$y = \frac{1}{2} ln \left(\frac{E + p_z}{E - p_z} \right)$$ (3.3)

where E and p_z are the particle energy and momentum in the z-axis. At high energy this can be approximated by pseudorapidity η, which is related to the polar angle θ:

$$\eta = -ln \left[tan \left(\frac{\theta}{2} \right) \right]$$ (3.4)

In the relativistic limit, rapidity and pseudorapidity converge. Momentum transverse to the direction of the beam p_T is defined as:

$$p_T = \sqrt{p_x^2 + p_y^2} \tag{3.5}$$

Initial partons have virtually no momentum in the transverse direction. Therefore, outgoing particles with large amounts of p_T are most likely the result of a hard interaction.

The purpose of the ATLAS detector is to identify the type, momentum and energy of the particles produced in each proton-proton collision. To this end, ATLAS adopts a traditional construction, with tracking systems, calorimeters and muon detector layered radially from the beam pipe (Figure 3.5). All stable standard model particles produced in a proton-proton collision can be identified by a combination of these three ATLAS subsystems. The position and momenta of charged particles are measured by the Inner Detector, situated in a magnetic field provided by a 2 T solenoidal magnet. Surrounding the Inner Detector are the electromagnetic and hadronic calorimeters, which measure the energy deposition of both electrically charged and neutral particles. Beyond these, the muon spectrometer and the toroidal magnet system are used to measure the momentum and position of muons. Neutrinos do not interact directly with the detector and instead are identified by 'missing energy' transverse to the beam direction. Conservation of momentum means the total of transverse energy should be virtually zero. For a fully hermetic calorimeter the presence of missing transverse energy is indicative of a particle which did not interact with the detector.

3.3.1   The Inner Detector

Tracking is designed to measure the momentum of charged particles and the position of the primary interaction point (primary vertex) of a collision. Precise tracking also permits the identification of long lived particles such as b-hadrons. Their comparatively
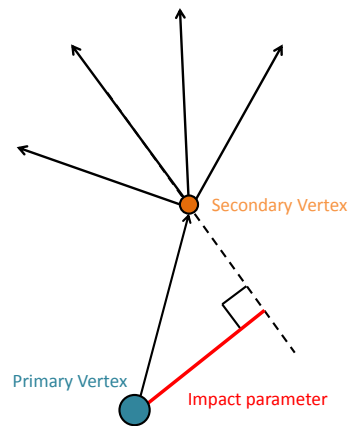
Figure 3.6: Schematic showing the impact parameter (distance of closest approach), for a decaying b-hadron.
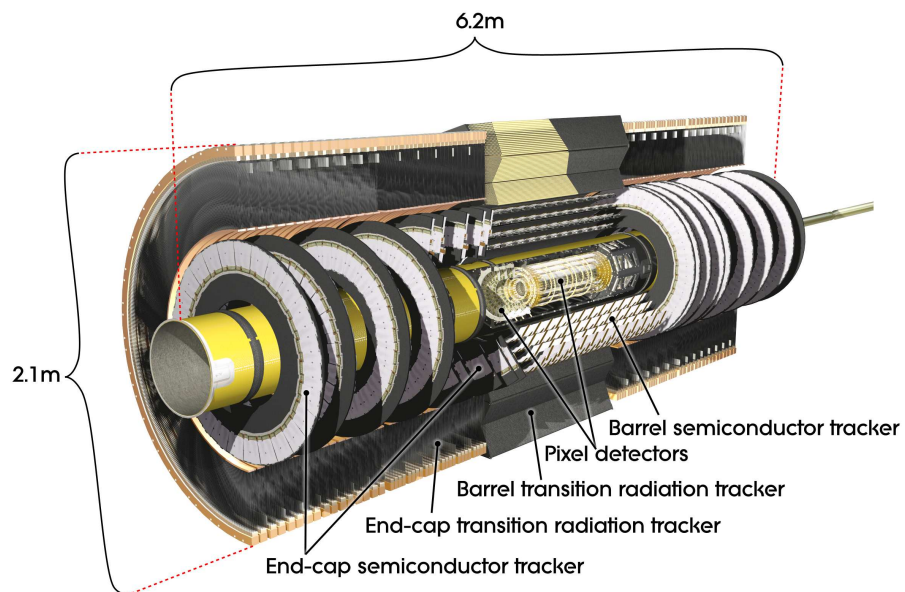
Figure 3.7: Illustration of the ATLAS Inner Detector, showing the pixel, strip silicon detectors and transition radiation detectors layered radially from the beam pipe and end-caps.

long lifetime (about 1.5 ps) allows identification from measuring the displacement of the b-hadron decay vertex (secondary vertex) with respect to the primary interaction vertex. The long lifetime and relatively large mass of the *b*-hadrons also permits identification via measurement of their large impact parameters, shown schematically in Figure 3.6.

The ATLAS Inner Detector (Figure 3.7) is comprised of three sub-systems arranged around the beam-axis in concentric cylinders, with disks in the end-caps. It extends 7 m in z and has a radius of 1.15 m, providing full coverage in ϕ and coverage up to $|\eta| < 2.5$.

Closest to the beam are the high resolution silicon detectors: the Pixel detector and Semiconductor Tracker (SCT). Surrounding these is the Transition Radiation Tracker (TRT). The entire Inner Detector is immersed in a 2 T magnetic field which is aligned with the beam axis, produced by a central solenoid. Electrically charged particles curve under the influence of this magnetic field in accordance to their charge sign and mass. They interact with material in the ID sub-detectors through ionisation, and are detected as an electrical signal or pulse which is recorded as a 'hit'. By combining the hits in the three sub-detectors, the particle curvature (and therefore particle momentum) can be assessed. Pattern recognition software is used to reconstruct the individual hits into particle 'tracks' [19].

The Silicon Detectors The silicon Pixel detector is designed for high granularity around the collision region. Each pixel layer is segmented in $R - \phi$ and z. The primary advantage of a pixel design is in the reduced hit-rate per channel in comparison to strip detectors. This allows for the fast electronic read-out time needed in the high particle density environment found close to the beam. The SCT consists of four pairs of cylindrical layers in the central region and nine disks at either end. The silicon micro-strips measure $R - \phi$, and permit z measurements by orienting the layers of strips. The SCT is critical for effective tracking perpendicular to the beam, since it measures particles over a much larger radial distance than the Pixel detector.

The Transition Radiation Tracker The TRT is comprised of drift tubes containing

a Xenon-based gas mixture. The drift tubes are aligned with the beam axis to measure $R - \phi$, but provide no z information. At the ends, the tubes are arranged in a fan layout to give information in $\phi - z$. The gas in the tubes ionises when traversed by a charge particle and negative ions drift to a charged central wire to produce a current pulse. Between the tubes, different media with different indices of refraction are placed. Relativistic charged particles that transition between the different media can emit 'transition radiation' in the form of photons, the quantity of which is depends on the Lorentz boost of the particle. A high degree of transition radiation can therefore be used to identify light particles such as electrons. The radiated photons contribute to an electronic pulse with a high threshold, which is distinguishable from the lower threshold pulse produced by particles which did not produce transition radiation. The TRT also contributes significantly to the momentum measurement since the lower precision is offset by the larger number of measurements (discussed in detail in Chapter 6) and due to longer track length measurements that it provides.

3.3.2 Calorimeter

Surrounding the Inner Detector is the calorimeter system (Figure 3.8). The calorimeters are designed to measure the energy of particles. This is especially important for neutral particles, since they are not detected by the tracking systems. Energy must be measured for both hadronic and electromagnetic particles, which 'shower' when they interact with material to produce a cascade of secondary particles. The depth of the ATLAS calorimeter must be sufficient to contain these showers. Hadrons penetrate further and shower more broadly than electromagnetic particles and so the positioning of the hadronic calorimeter is designed to allow for this and reduce the effect of 'punch-through'. Punch-through is caused by secondary particles exiting the calorimeter and faking muon signals in the muon detector. ATLAS calorimetry employs a method known as 'sampling' to enable a more compact design. Alternating layers of active material and passive material are used to measure energy and to absorb the shower respectively. The ATLAS calorimeters are non-compensating; some of the energy of a hadronic collision is lost to nuclear recoils and breakdowns or by undetected secondary particles such as muons escaping the calorimeter. The effect is accounted for in jet reconstruction by applying a jet energy correction. This is discussed in Chapter 4.
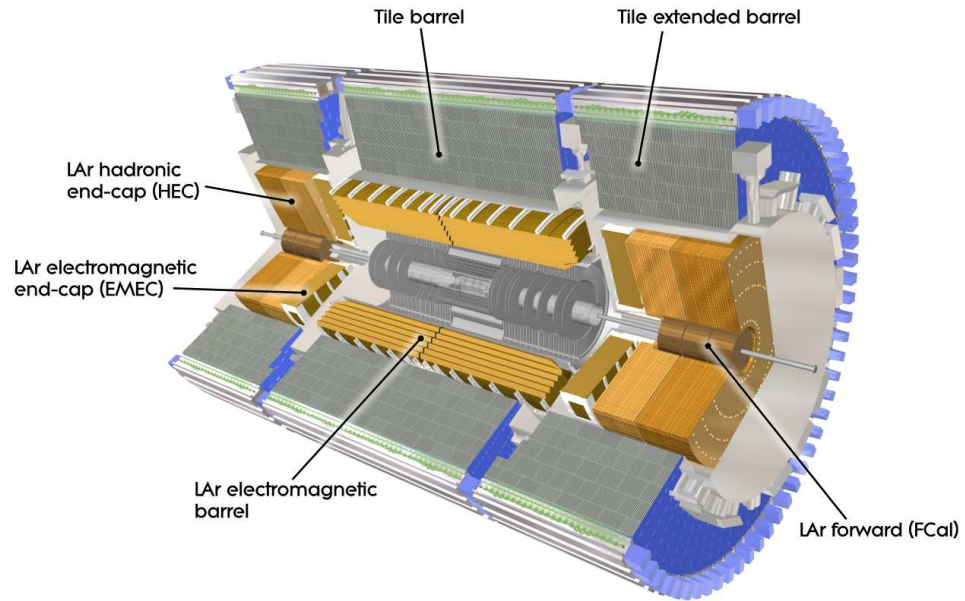
Figure 3.8: The ATLAS calorimeters, showing the Liquid-Argon calorimeters in orange and the tile calorimeters in grey.
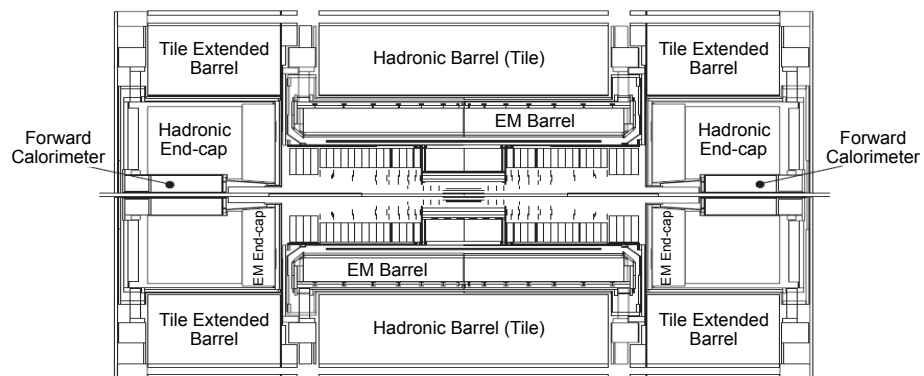
Figure 3.9: The transverse $(r - z)$ positioning of the of the ATLAS calorimeters.

A schematic of the ATLAS calorimeters is provided in Figure 3.9. Calorimetry is provided for $|\eta| < 4.9$. High granularity electromagnetic (EM) calorimetry is provided by the liquid-argon (LAr) sampling calorimeters, which give coverage up to $\eta < 3.2$. Beyond this, the forward calorimeter (FCAL) provides electromagnetic calorimetry up to $|\eta| < 4.9$. In the region where $|\eta| < 1.7$, hadronic calorimetry is performed by a scintillator-tile calorimeter. This is divided into a main central barrel and two smaller extended barrels situated either side. For $|\eta| > 1.5$ hadronic calorimetry is performed using the LAr technology in the Hadronic End-Caps (HEC). For very forward measurements up to $|\eta| < 4.9$, the FCAL also provides hadronic measurement capability.

Electromagnetic Calorimetry The EM calorimeter uses liquid-argon as an active material and lead as an absorber arranged in an 'accordion' structure. This design acts to reduce any azimuthal cracks, while allowing full ϕ symmetry. As charged particles pass though the calorimeter they ionise the liquid-argon. The electrons produced drift toward copper electrodes in the read-out cells under the presence of a magnetic field. Two half barrels span to $|\eta| < 1.475$, and the end-caps (divided into two coaxial wheels) cover the region $1.375 < |\eta| < 3.2$. The crack region $1.375 < |\eta| < 1.52$ contains material needed to cool the detector. For $|\eta| < 2.5$ the EM calorimeter is finely grained to allow precise measurements of electrons and photons, and is segmented into three longitudinal layers radially from the beam pipe. Beyond $|\eta| < 2.5$ the EM calorimeter is divided into two sections in depth, and more coarsely grained.

Hadronic Calorimetry Hadronic calorimetry is shared by the Tile, the HEC and the FCal calorimeters. The Tile calorimeter is made from alternating steel as an absorber with scintillating plates as the active material. It is situated outside the EM calorimeter and is radially segmented into three separate Layers. The HEC consists of two LAr end-caps with copper absorber plates, placed outside the EM end-cap calorimeters and sharing the same cryostat. The FCAL is divided into two parts situated at either end of ATLAS, covering $3.2 < |\eta| < 4.9$. It is comprised of concentric tubes situated parallel to the beam axis, each containing a metal rod surrounded by liquid-argon. The two end parts are each divided into three modules: the inner module contains copper for EM

measurements, and the two deeper modules use tungsten for hadronic measurements.

3.3.3   Muon System

Muon momentum is measured using a purpose-built muon system (Figure 3.10). A magnetic field is provided by three barrel and two in the end-cap air toroids. The air core minimises the deflections due to multiple scattering, and the field provides a bending moment to the muons. High precision tracking in $|\eta| < 2.7$ is provided using the Monitored Drift Tubes (MDT) and Cathode Strip Chambers (CSC). Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC) are used for selecting interesting events containing muon candidates.

3.4   Trigger Systems

The high rate delivered by the LHC far exceeds the computing resources available to ATLAS to record and store all the data. Bunch crossings occur at a rate of up to 40 GHz; this data must be compressed to 200 MHz for storage. A method of selecting interesting events, known as 'triggering' is implemented. The trigger system should ideally reject as much background as possible without biasing against events of interest. The ATLAS trigger is a 'three-tier' system, comprised of both hardware-based and software-based triggering systems. Initially, events are selected using a hardware (Level-1) trigger. Level-1 selects regions of interest from the calorimeter and muon systems, which are passed to the intermediate software based trigger (Level-2). At Level-2 the full detector information is available, and harder selection criteria are applied to each event. Events that pass Level-2 are passed to another software trigger (the Event Filter or EF). The EF performs an analysis similar to the full event reconstruction, and combined with Level-2 is known collectively as the 'high level' (HL) trigger. Events that pass this trigger are stored for offline processing. Limitations in bandwidth means that in some cases, only a fraction of the events that pass the trigger are kept. Such a trigger is described as 'prescaled', where a prescale of 1 means that all events that pass the trigger are kept and the fraction of events that passes the trigger is given by 1/prescale. Once data has passed through the trigger systems and the rate reduced to manageable levels it is written to tape and stored.
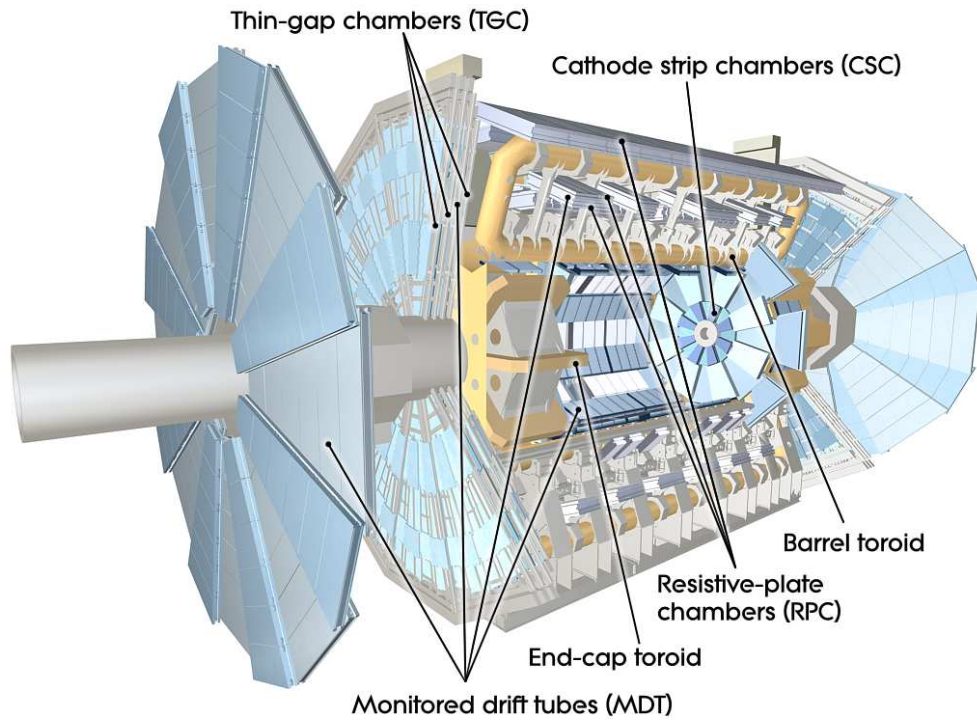
Figure 3.10: The ATLAS muon system.

3.5 Software

Monte-Carlo event generators permit theoretical predictions for Standard Model processes to be combined with information concerning the proton structure, underlying event, fragmentation and hadronisation as discussed in Chapter 1. In general they model the hard scatter in a $2 \to 2$ or $2 \to n$ process, convoluted with an appropriately chosen PDF. The probability that each parton will split into further particles is simulated by a parton shower model. Once a prescribed energy cut-off scale is reached, a hadronisation model is implemented.

The event generators used in the analyses discussed in this thesis are HERWIG [20], PYTHIA [21] and ALPGEN [22]. PYTHIA implements a $2 \to 2$ matrix element at leading order. The matrix element represents the perturbative calculation of the hard process and leading order refers to the first order in the perturbative expansion. The parton

showers are modelled in a leading-logarithmic approximation using p_T ordering [23]. The hadronisation stage is implemented as the Lund string model [24]. Multiple parton interactions are also simulated. The PDF set applied in PYTHIA is the modified leading order set MRST LO* [25]. ALPGEN models the matrix element at leading order for multi-parton processes, where additional partons are included as (up to 6) extra legs of the diagram. It is interfaced to the HERWIG generator to model a leading-logarithmic, angular-ordered parton shower. HERWIG uses a cluster model where gluons split in $q\bar{q}$ pairs which then regrouped into hadrons. The multiple parton interactions are modelled using JIMMY [26]. The PDF used for the ALPGEN samples is CTEQ6L1 [27]. HERWIG++ [28] is an extension of HERWIG, it uses an angular ordered parton shower and implements custom hadronisation and multi-parton interactions models [29]. The PDF used is also MRST LO*. Underlying event models in the PYTHIA and HERWIG are tuned from minimum-bias events (all events, with no bias from restricted trigger conditions) from ATLAS data [30, 31]. For next-to-leading order NLO calculations (calculated to the second order of perturbation theory), NLOJET++ 4.1.2 [32] is implemented and convoluted with the parton distribution functions via the APPLGRID software [33]. The ATLAS detector 'full simulation' is based on GEANT4 [34], and consists of a complete simulation of detector material and geometry, with extensive calorimeter modelling. This modelling takes 30-60 minutes per event.

Chapter 4

Jets in ATLAS

4.1 Introduction

Jets in ATLAS are made using the standard jet clustering algorithms (discussed in Chapter 2), implemented via the FASTJET package [35]. The choice of jet algorithm depends on the physics requirements of the analysis: measurements of inclusive jets in Chapter 5 use anti-k_T jets with R = 0.4 or 0.6 and sub-structure jets in Chapter 6 use anti-k_T jets with R = 1.0 and Cambridge-Aachen jets with R = 1.2. For the purposes of this discussion, jets from these two separate analyses will be referred to as the 'inclusive jets', and the 'sub-structure jets' respectively.

This chapter outlines the procedure applied to ATLAS calorimeter cells in order to reconstruct jets (Section 4.2). This process requires the application of jet calibration (Section 4.4), quality control (Section 4.6) and a method to compare the resulting jet distributions with theoretical predictions, referred to as 'unfolding' (Section 4.7).

4.2 Inputs to Jet reconstruction

4.2.1 Calorimeter Cells

The simplest configuration of calorimeter cells is a 'tower'. Towers are constructed using a geometrical projection through η and ϕ, typically 0.1×0.1 in size. Noise and pile-up can cause large energy fluctuations in a tower, which in some situations will result in negative overall tower energy. Since the jet finders require a four-vector input

with positive energy, noisy towers with negative energy are combined with neighbour-
ing positive towers to create an object with an overall positive energy.

'Topoclusters' are more complex calorimeter objects, consisting of groups of calorime-
ter cells which are clustered into three-dimensional energy deposits, designed to repro-
duce the shower development in the calorimeter. They are seeded by a cell with cell
energy $|E_{cell}| > 4\sigma_{cell}$ where σ_{cell} is the root-mean-squared (RMS) noise of the cell, in-
dicating the average magnitude of the error due to electronic noise and pile-up. Neigh-
bouring cells which satisfy $|E_{cell}| > 2\sigma_{cell}$ are combined with the seed cell. Finally,
any surrounding cells satisfying $|E_{cell}| > 0$ are added to the cluster. Inactive cells and
cells with high electronic noise are excluded from this process. Both towers objects and
topoclusters objects are assigned a mass value of 0 in reconstruction and the object
energy is taken as the total sum of the constituent cell energies. The object direction
is calculated to point back to the geometric centre of the calorimeter.

4.2.2 Tracks

Jets constructed from tracks (track jets) have sources of systematic uncertainty that
tend to be uncorrelated to those of the calorimeter thus track jets provide an invalu-
able cross-check of the systematic uncertainty of jets seeded from calorimeter objects
(calorimeter jets). This comparison is exploited for the calculation of the systematic
uncertainties for jet energy resolution and energy scale in ATLAS jets (Section 4.3).
The tracks are reconstructed using the Inner Detector. Track mass is generally taken to
be that of the pion. The track quality is determined using indicators such as the total
number of hits in the various layers of the Inner Detector, the impact parameter or
track momentum. The precise selection criteria used for track jets studies is discussed
in Chapter 6.

4.3 Jet Energy Resolution

The jet energy scale (JES) and jet energy resolution (JER) are related to the mean and width respectively of the measured energy of a sample of jets at a given energy. For inclusive and sub-structure jets the scale and resolution are derived in Monte-Carlo, and then validated using in-situ methods.

The measured energy of a jet can fluctuate from its actual value for a variety of reasons including the modelling of physics in Monte-Carlo and calorimeter effects. Differences can also occur due to effects such as radiation losses in the jet reconstruction process and statistical fluctuations of the shower development. Determining how well the Monte-Carlo models these effects is particularly important when comparing the final jet measurements to theoretical predictions.

The JER is calculated in Monte-Carlo by taking the width of the jet energy response R_E and dividing by the mean:

$$R_E = E_{calo}/E_{truth} \tag{4.1}$$

E_{truth} corresponds to the energy of jets reconstructed at the hadronic or 'truth' level (excluding muons and neutrinos) and E_{calo} corresponds to jets reconstructed with full calorimeter simulation implemented.

For the inclusive jets, the JER is primarily validated using in-situ methods [36]. The two main techniques that are used both agree with Monte-Carlo to within 10% as demonstrated in Figure 4.1. The first method (red) exploits the energy conservation in a dijet system in the transverse plane. Dijets are chosen which fall in similar regions of detector (and thus vulnerable to similar resolution effects), with back-to back topology ($\Delta\phi > 2.8$). A transverse momentum cut is made on any third jet to limit the impact of gluon radiation. The width of the asymmetry in the p_T of each jet (asymmetry is defined as the difference in the p_T of each jet, divided by their sum) is used constrain
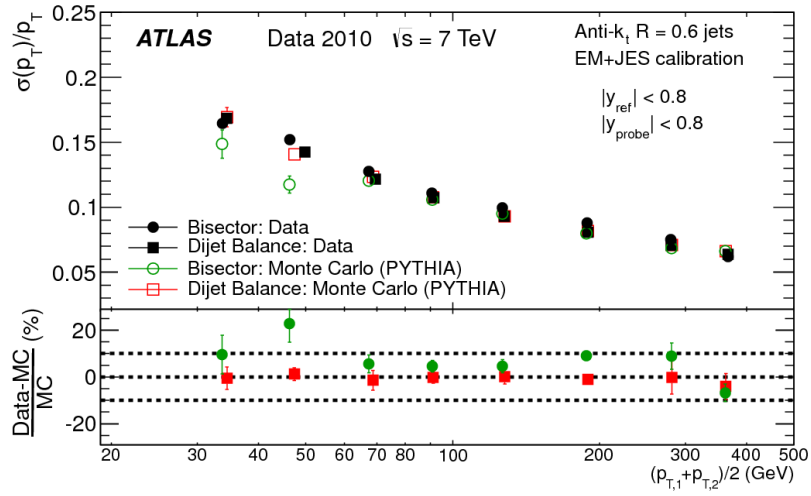
Figure 4.1: The JER correction determined in ATLAS for dijet balance and bisector methods in data and Monte-Carlo [36].
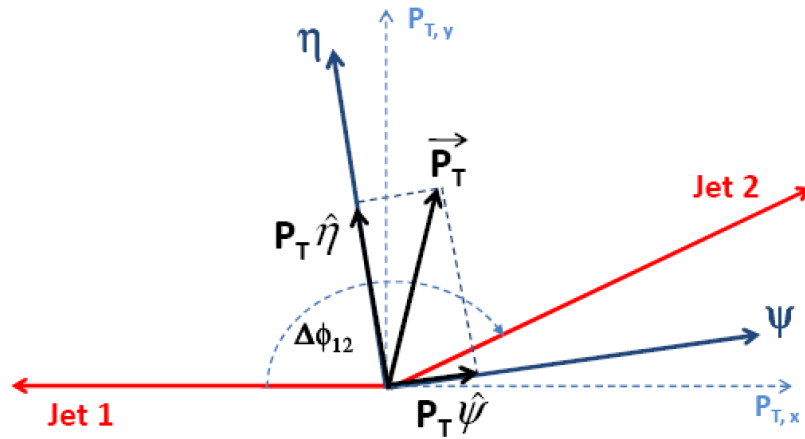
Figure 4.2: The bisector method applied to a dijet system. The η-axis corresponds to the azimuthal bi-sector of the dijet system in the plane transverse to the beam axis. The ψ-axis is defined as orthogonal to the η-axis [36].

the overall resolution. The second method (green) uses a balance vector P_T, defined as the vector sum of the p_T of a dijet system. This is illustrated in Figure 4.2. For a perfectly balance dijet $P_T = 0$. This vector is projected on to the axes that bisect the dijet system, ψ and η. There are a number of sources that give rise to fluctuations in the projections and therefore non-zero values in the variance the ψ and η components. However, at particle-level the magnitudes of these fluctuations are expected to be the same, except for the differences which arise due to contributions from calorimeter effects.

The main uncertainties for the JER come from different physics modelling in the Monte-Carlo and from the uncertainties of these two situ-methods. The uncertainty of dijet balancing is 3% to 4% for $20 - 500$ GeV jets and the bisector method gives an uncertainty of 4% to 6% for $20 - 500$ GeV jets.

For the sub-structure jets, mass reconstruction is also important and so the jet mass response R_m is defined:

$$R_m = \text{mass}_{\text{calo}}/\text{mass}_{\text{truth}} \tag{4.2}$$

The jet mass resolution (JMR) and JMR uncertainty in ATLAS were not measured prior to the sub-structure measurement described in this thesis. Furthermore, JER was only evaluated for anti-k_T $R = 0.4$ and 0.6 jets and therefore needed to be re-evaluated for larger sub-structure jets. The resolution uncertainty measurements are taken from Monte-Carlo only. Uncertainties are derived by studying the variations in resolution created by varying the detector the geometry and varying the modelling of hadronic physics in GEANT and in different Monte-Carlo generators. Resolution uncertainties of 20% are found for all the variables considered in the substructure analysis, except the n-subjettiness uncertainties, which are found to be 10%.

4.4   Jet Energy Scale

The ATLAS calorimeter is non-compensating (the response to hadrons is lower than the response to electrons and photons of the same energy). This is primarily due to energy deposited in the calorimeter that is undetected, such as that lost in nuclear excitations, nuclear breakdowns and escape of 'invisible' particles such as neutrinos and muons. Calorimeter cells are initially calibrated to the electromagnetic (EM) scale, which correctly measures the energy deposited by electromagnetic showers. The ATLAS EM scale is established from test-beam measurements using electrons and muons [37]. A 'jet energy scale correction' is applied to calibrate hadronic objects from the EM scale to the correct energy. This correction also accounts for effects such as energy losses due to dead material, and for inefficiencies which arise during calorimeter clustering and jet reconstruction.

4.4.1   Jet Energy Scale Correction

Jets constructed from calorimeter objects are corrected from the electromagnetic scale to the hadronic scale. Inclusive jets use a jet-level correction, where the calibration is applied to each reconstructed jet. The correction described here for inclusive jets is usually referred to as 'EM+JES'. Conversely, sub-structure jets are corrected for non-compensation using an energy correction at the cluster level before jet reconstruction. Since it is applied before jet finding the correction helps improve knowledge of the particle-level structure of the jet, which can be advantageous when considering internal structure. After jet reconstruction other effects which act to alter the energy response (such as dead material and reconstruction inefficiencies) are then corrected for using a jet-level correction.

4.4.2   Pre-Calibration

Before calibration a 'jet origin' correction is applied to the jet inputs. The inputs initially point to the geometric centre of the detector; improvements in the angular resolution and p_T can be achieved by recalculating the inputs to point to the primary (interaction) vertex.

For the sub-structure jets, pile-up is removed by selecting events with only one primary vertex. For inclusive jets, a correction for pile-up is applied prior to JES calibration. Each additional vertex in an event corresponds to approximately 0.5 GeV of additional energy per jet. An average energy correction is subtracted from the EM scale jet to account for any extra energy coming from the additional interaction, calculated as a function of the number of primary vertices and pseudorapidity of the jet.

4.4.3   Jet Energy Scale Calibration

The jet energy scale is established in Monte-Carlo by matching reconstructed calorimeter jets to Monte-Carlo particle truth jets (excluding muons and neutrinos). The distribution of the energy response between calorimeter and matched particle jets is used to determine an average jet energy response. The correction is obtained by dividing the particle jet energy by the EM-scale energy of the calorimeter. This correction is energy and pseudorapidity and dependent, and is shown as a function of these variables in Figure 4.3. The effects of the changing detector environment with η can be clearly observed.

For forward jets, the transverse momentum balance between central and forward jets in dijet events is used to derive the JES uncertainty. The method is commonly referred to as the 'η-intercalibration'. The systematic uncertainties on central jets are lower due to the high tracking precision available from the ID. The technique is similar to the dijet balance method described in Section 4.3, but using differing regions of the detector. Selection cuts are applied to improve the dijet topology and remove third jet radiation as far as possible [40]. The region $|\eta| < 0.8$ is taken as a reference and used to evaluate the uncertainties of jets in more forward regions. Due to momentum conservation, these jets are expected to have equal p_T.

The inclusive JES is further validated by balancing the jet with another physics object with well measured energy. Such methods include calculating the p_T balance between the sum of the transverse track momenta associated to the jet, using a system
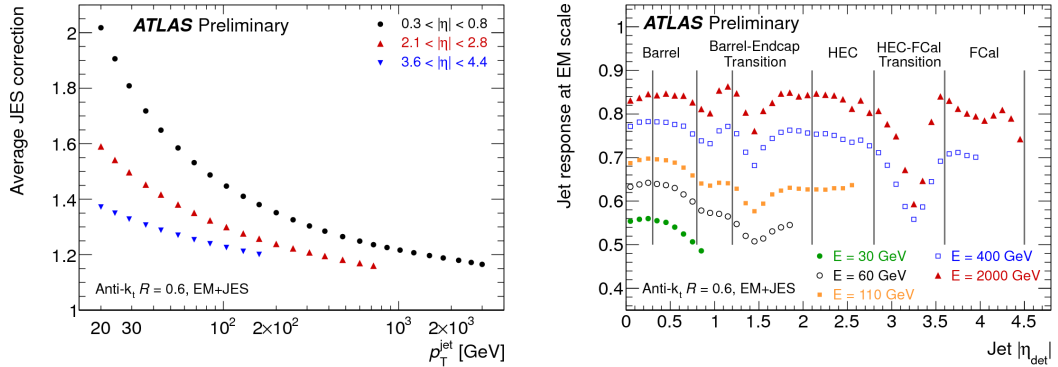
Figure 4.3: The JES scale corrections for inclusive jet measurements as a function on η and p_T [40].

of low p_T jets or a photon. The track-based method gives high statistical precision and extends across the p_T range of the inclusive measurement. However, it is subject to high systematic uncertainty of 6% for very low p_T jets and about 3-4% for $40 < p_T < 800$ GeV. Photon balancing has a 1% systematic uncertainty, but is statistically limited and only valid for $p_T > 300$ GeV. Balancing high p_T jets against the recoil system of low p_T jets enables validation to within 5% for jets up to 1 TeV, with statistical an uncertainty comparable to (or less) than the systematic uncertainty.

For sub-structure the method of deriving and applying the final jet level correction is similar to the inclusive method, although there are three important differences. Firstly, only central jets are measured and so no η-intercalibration is required. Secondly, sub-structure measurements use calorimeter cells which are calibrated to correct the energy response before jet finding. Hadronic showers are broader and penetrate deeper than electromagnetic showers and so the clusters categorised as being either electromagnetic or hadronic, in accordance to their shape, energy density and depth in the calorimeters. The hadronic clusters are corrected to the hadronic scale by applying calibration constants before jet finding [38]. Further corrections for effects such as dead material and reconstruction inefficiencies are then applied in a jet-level correction, after reconstruction. Finally, sub-structure calibration also acts to restore the mass response of the calorimeter. Mass correction factors are calculated from Monte-Carlo as a function of η and applied in a similar fashion to p_T correction factors. The final calibration is there-

fore dependent on η, p_T and mass [39]. The jet mass scale uncertainty is determined by considering that the calorimeter and inner detector are expected to have uncorrelated systematic uncertainties. A comparison between track jets and calorimeter jets can be used to determine the uncertainty on the calorimeter, provided that track reconstruction for large high p_T jets is well modelled by simulation and that the uncertainties associated with tracking are small. Establishing the uncertainties for tracking forms a large part of this thesis and is discussed in detail in Chapter 6. The final uncertainty from the scale uncertainties are between 3% and 5%.

4.4.4   Post-Calibration

Following calibration, a small η-dependent correction is applied to account for energy losses in regions of the detector with inactive material, such as the transition regions between calorimeters. In general the correction is small, but can reach up to 5-7% in the most poorly instrumented regions.

4.5   Uncertainties of the Jet Energy Scale

In the central detector the jet energy can be measured with a precision of about 2% to 3% over a wide transverse momentum range. The main sources of systematic uncertainty on the JES for inclusive jets are shown in Figure 4.4. These include: non-closure (deviation from unity) when the JES correction is applied to reconstructed Monte-Carlo; uncertainty on the single hadron response of the calorimeter; detector noise modelling in Monte-Carlo; the uncertainty associated with physics modelling by Monte-Carlo generators and the uncertainties arising from the η-intercalibration between central and forward jets.

Single Particle Calorimeter Response The response of the calorimeters to isolated (charged) hadrons is evaluated by considering the ratio of their energy E which is deposited in the calorimeter, to the momentum p as calculated from the associated track in the Inner Detector. Figure 4.5 shows the ratio of $\langle E/p \rangle$ in Monte-Carlo compared to data. The dotted lines indicate unity $\pm 5\%$.
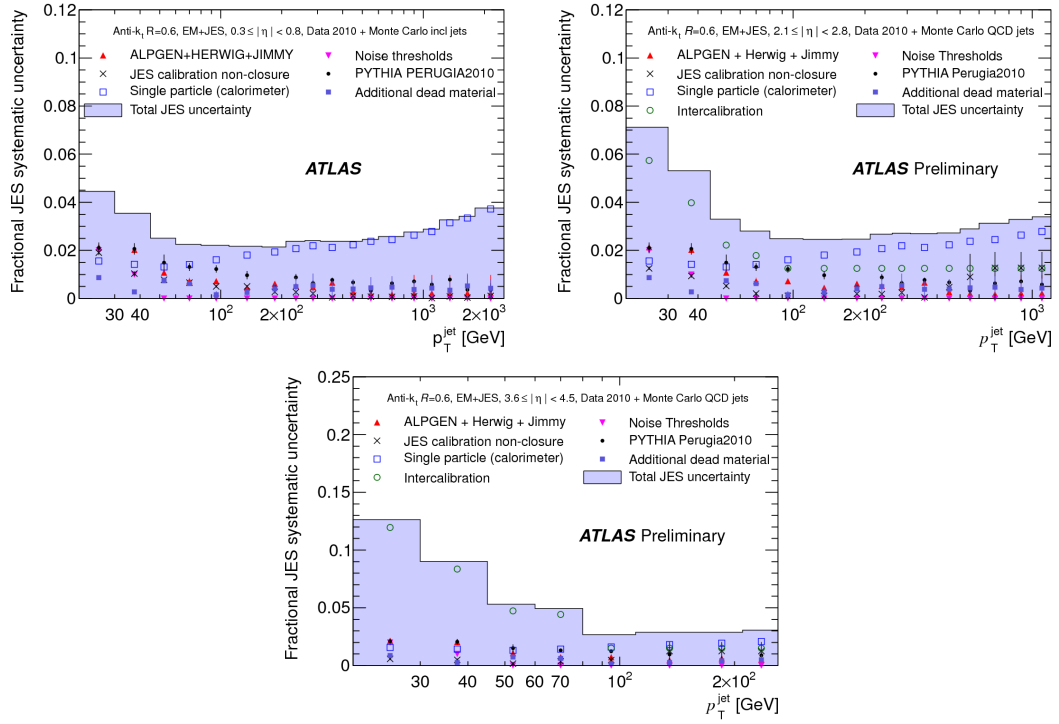
Figure 4.4: The fractional contribution to JES systematic uncertainty of anti-k$_T$ jets with R = 0.6 in three different η regions [40].
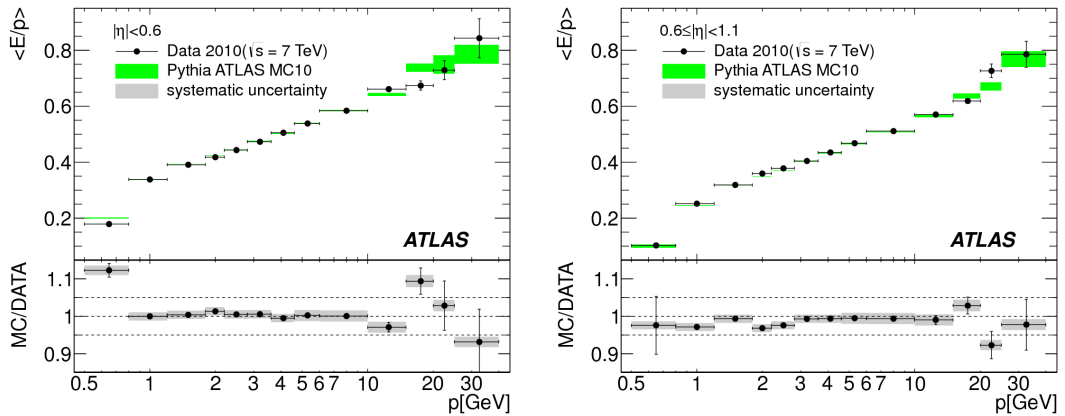
Figure 4.5: $\langle$ E/p $\rangle$ at $\sqrt{s} = 7$ TeV as a function of the track momentum for $|\eta| < 0.6$ (left) and $0.6 \leq |\eta| < 1.1$ The dotted lines indicate unity $\pm 5\%$ (right) [41]

Non-Closure Uncertainties arise from non-closure in the jet p_T and energy response when the JES is reapplied to original Monte-Carlo from which the correction factors were derived. This deviation from unity indicates that the p_T of the calibrated calorimeter jet is not properly restored by the correction factors. This is a result of using the energy correction to restore the transverse momentum; if the masses of the calorimeter and the truth jet are different there will be a bias in the p_T. The systematic uncertainty due to the non-closure is taken as the largest deviation of the response from unity between energy and p_T.

Modelling of Dead Material The effect of dead material is checked by varying the material budget modelling in Monte-Carlo by adding more or less material to the material in the detector, within the material budget uncertainty. The modelling of the detector in Monte-Carlo simulation is well described. No additional uncertainty for punch-through effects is assigned for high p_T jets.

Noise The uncertainty from discrepancies between noise descriptions (measured as the RMS of the cell energy when no collision have taken place) in Monte-Carlo and the data is estimated by lowering and raising the signal-to-noise thresholds for clusters from the nominal values by 10% and 7%.

η-intercalibration The systematic uncertainties on the JES in the central region $0.3 < |\eta| < 0.8$ are well constrained and therefore this region is used as the baseline with which the other regions are calibrated using the dijet balance method. The uncertainty due to η-intercalibration is taken as the RMS deviation of the Monte-Carlo to data.

4.6   Quality Control

4.6.1   Luminosity

Luminosity is measured independently by several detectors, the results of which are combined to produce the overall ATLAS luminosity measurement and uncertainty. The

absolute calibration is derived using van-der-Meer scans which determine the observed event rate by scanning the beams across each other in horizontally and vertically. ATLAS operates quality monitoring during data taking to ensure detector conditions for specific periods of luminosity are optimal. These include checks on if there was a stable beam present in the detector and tests of whether the luminosity measurement and trigger systems are functioning within expected parameters. All data used in this thesis is required to pass these quality checks.

4.6.2   Quality Requirement for Jets

Selection criteria act to remove fake or mis-measured jets; the process by which they are removed is referred to as jet 'cleaning'. These types of jets arise from various sources. These include noise 'spikes' in calorimeter cells in the HEC, coherent noise in calorimeter cells in the EM calorimeter and jets arising from photons produced by cosmic muons and other non-collision backgrounds. Jets are typically distinguished between being fake and being badly measured (due to falling into less well instrumented areas of the detector) according to the cleaning criteria they pass. The variables used to clean the jets are defined as follows:

f_{HEC}: The fraction of the jet energy in the Hadronic End-cap (HEC).

Q_{LAr}: Fraction of LAr cells with $Q-$factor > 4000 (Q-factor is a quantity which quantifies the difference between the measured pulse shape and the predicted pulse shape in a given LAr cell).

Q_{HEC}: As Q_{LAr} but for cells in the HEC.

f_{EM}: The fraction of the jet energy in the EM calorimeters.

t: The mean difference in time between the cells in the jet and the event time.

f_{charged}: The ratio of the total p_T of track associated to the jets divided by the p_T of the jet.

f_{Max}: The maximum energy fraction in any one calorimeter layer.

The sub-structure and inclusive measurement each use a combination of selection cuts on these variables. Requirements are placed on the minimum number of tracks pointing to a primary vertex candidate. This is a particularly effective method for

removing jets that are reconstructed from cosmic muons which have passed through the ATLAS cavern into the detector. Events are selected with a primary vertex that is consistent with the luminous region (beam-spot) position with at least five associated tracks, each with transverse momentum > 0.5 GeV.

4.7 Unfolding

Unfolding is a process that corrects distributions reconstructed at the detector-level to the equivalent distributions if all detector effects (smearing and inefficiencies) are mitigated - the perfect detector. These 'ideal' distributions allow direct comparisons to theory and other experiments, without a priori knowledge of the detector. Distributions for the final-state particle (hadron level) are compared to those after the detector simulation is applied (detector level). The two important methods of unfolding used in inclusive and sub-structure are bin-by-bin and Iterative Dynamically Stabilised (IDS).

4.7.1 Bin-by-bin

The bin-by-bin method uses the ratio between the hadron level and detector level distributions; the ratio is then applied as a correction factor to the measured data. Typically distributions are binned in p_T and η. The two primary issues for this method are the fraction of events which stay in a comparable bin between hadron level and detector level (efficiency) and the fraction of events in detector level also found in the same bin at the hadron level (purity). If bin migrations are large compared to bin width, wide bins are necessary.

4.7.2 Iterative Dynamically Stabilised

Iterative Dynamically Stabilised unfolding uses a two-dimensional transfer matrix A_{ij} to encode the effect of the detector. Calorimeter jets and true jets are matched by a geometric selection, and a transfer matrix is constructed to represents the probability for an event reconstructed in bin i to originate from bin j. The data is multiplied by the transfer matrix. The result is optimised using an iterative procedure by scaling rows of the matrix to match the corrected result. At each step the hadron level Monte-Carlo is reweighted to the shape of the corrected data in order to prevent statistical fluctuations

from being enhanced by successive iterations. A_{ij} does not include unmatched jets, so an efficiency factor is applied to data to account for the unmatched jets before unfolding, and then a corresponding factor is applied once again at the end of the unfolding to restore the number of jets.

Chapter 5

The Inclusive Jet Cross-Section

5.1 Introduction

The inclusive jet cross-section is a benchmark analysis for the LHC. Jet production is the most abundant high p_T process at the LHC and the measurement relies on the effective commissioning and operational running of both the machine and detector, as well as a firm understanding of the systematic uncertainties of jet reconstruction. Furthermore, a solid understanding of jet reconstruction in ATLAS provides the foundation for many other physics analyses, including top production, Higgs searches and physics beyond the Standard Model. The inclusive jet cross-section is sensitive to the combination of the QCD matrix element and parton densities, yielding precise measurements of the strong coupling constant α_S and providing information on the structure of the proton. The energetic reach of the LHC extends beyond that of any previous collider experiment, permitting QCD to be probed up to unprecedented energy scales.

In this chapter the ATLAS analysis of the inclusive single jet cross-section is discussed. Measurements are made using anti-k$_T$ jets with R = 0.4 and R = 0.6. The initial measurement used a total integrated luminosity of 17 nb^{-1} of data collected at a centre-of-mass energy of 7 TeV. This measurement is published in `EPJC`. This is extended to include low p_T (down to $p_T > 20$ GeV) and very forward (regions spanning beyond the central calorimeter up to $|\eta| = 4.4$) jets. The final measurement corresponds to a total integrated luminosity of 37 pb^{-1} [42].

5.2 Inclusive Jet Measurements

5.3 Jet Reconstruction

Jets are made using the anti-k_T jet clustering algorithm (described in Chapter 2), with R = 0.4 and R = 0.6. They are built from all stable particles (particles with a lifetime of 10 ps or greater). In the case of Monte-Carlo this definition includes neutrinos and muons. Jets are constructed and calibrated according to the method described in Chapter 4. Jets are selected to satisfy $p_T > 20$ GeV and $|y| < 4.4$

5.4 Trigger Selection

In this measurement trigger efficiency is defined as the ratio between the number of jets in an event where at least one of the jets fired the trigger to the number of all jets in all events. A trigger is considered 'fully efficient' at the p_T when this efficiency > 99%. The trigger strategy requires a specific trigger per jet for a particular p_T and $|y|$ bin, where the $|y|$ bins are chosen to reflect the geometry of the detector. For jets with $20 < p_T < 60$ GeV triggers that select on minimum bias data are used. Central jet triggers are used all other p_T for $|y| < 3.6$. Forward jet triggers are used for jets with jets $3.2 < |y| < 4.9$.

Background Source	Removal Cuts
HEC Spikes	$f_{\mathrm{HEC}} > 1 - |Q_{\mathrm{HEC}}|$
EM Noise	$f_{\mathrm{EM}} > 0.9$ **AND** $|Q_{\mathrm{LAr}}| > 0.8$ **AND** $|\eta_{\mathrm{EM}}| < 2.8$
Cosmics	$|t| > 10$ ns
Cosmics	$f_{\mathrm{EM}} < 0.05$ **AND** $f_{\mathrm{charged}} < 0.1$ **AND** $|\eta_{EM}| < 2$
Cosmics	$f_{\mathrm{EM}} > 0.95$ **AND** $f_{\mathrm{charged}} < 0.05$ **AND** $|\eta_{EM}| < 2$

Table 5.1: Cleaning variables specific to the inclusive jet analysis.

5.5 Jet Quality

The selection cuts in Table 5.1 are applied to remove backgrounds as described in Chapter 4. This combination of cuts is referred to as 'medium' to reflect the stringency of the requirement.

Events are required to contain at least one primary vertex consistent with the position of the LHC beam-spot (luminous region), with five or more tracks with $p_T > 150$ MeV. The vertex distribution in data is found to be very well modelled by a Gaussian fit (Figure 5.1), shown here for 1792 nb^{-1} of data. However, when viewed on a logarithmic scale, non-Gaussian contributions become apparent in the tails. These vertices tend to be the result of non-collision sources (Section 5.7) and so can be used validate and improve the cleaning cuts applied to remove fake jets.

5.6 Vertex Reconstruction in the Inclusive Jet Cross-section

For the first 17 fb^{-1} of data, jet variables were calculated with respect to the position of beam-spot rather than to the event primary vertex. This approach was taken due to a technical issue in the production of the dataset; however, this choice was considered justified provided that no significant physics impact was found.

A comparison was made using the detector geometry information to recover the primary vertex information, in order to redefine the jet position back to the primary vertex. A simple average was calculated over all calorimeter cells in a given layer to obtain the primary vertex r, and z positions. These were used to recalculate variables such as p_T and η taking the primary vertex position as the new origin of the jet. These variables were plotted for anti-k$_T$ = 0.4 jets for data and Monte-Carlo taken from PYTHIA. Figure 5.2 demonstrates the effect on p_T for three different η bins covering the central and crack regions of the detector. The broadening of the data distribution in more central bins show that this effect is greatest on central jets, however the effect is just a few % of the overall jet p_T.

As a fraction of the overall data, events with $|z| > 200$ mm is 0.06%. The difference
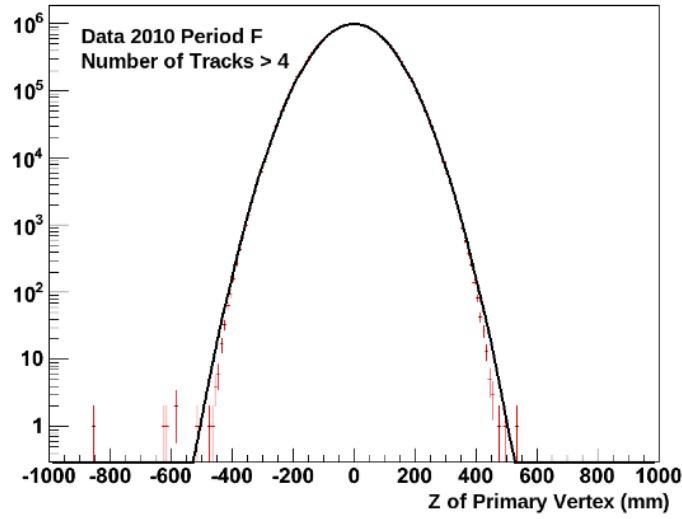
Figure 5.1: Vertex position in data, shown on a logarithmic scale, fitted with a Gaussian distribution corresponding to approximately 1792 nb^{-1} of data

in the p_T for such events compared to events with $|z| < 100$ mm is small. Consequently, the uncertainty from mis-modelling of the vertex position is taken to be negligible in the final result.

5.7 Removal of Non-Collision Backgrounds in the Inclusive Jet Measurement

Non-collision background refers to reconstructed jets that were not produced from normal collisions of the LHC beams. Primarily they consist of either beam induced backgrounds (BIBs) or cosmic ray showers.

5.7.1 Beam Induced Backgrounds

BIBs originate from proton losses upstream of the interaction point, which induce secondary cascades of particles. In general, the rate of such particles is proportional to the beam current and depends on the operational conditions of the LHC.

- **Tertiary halo** caused by protons lost on limiting apertures.
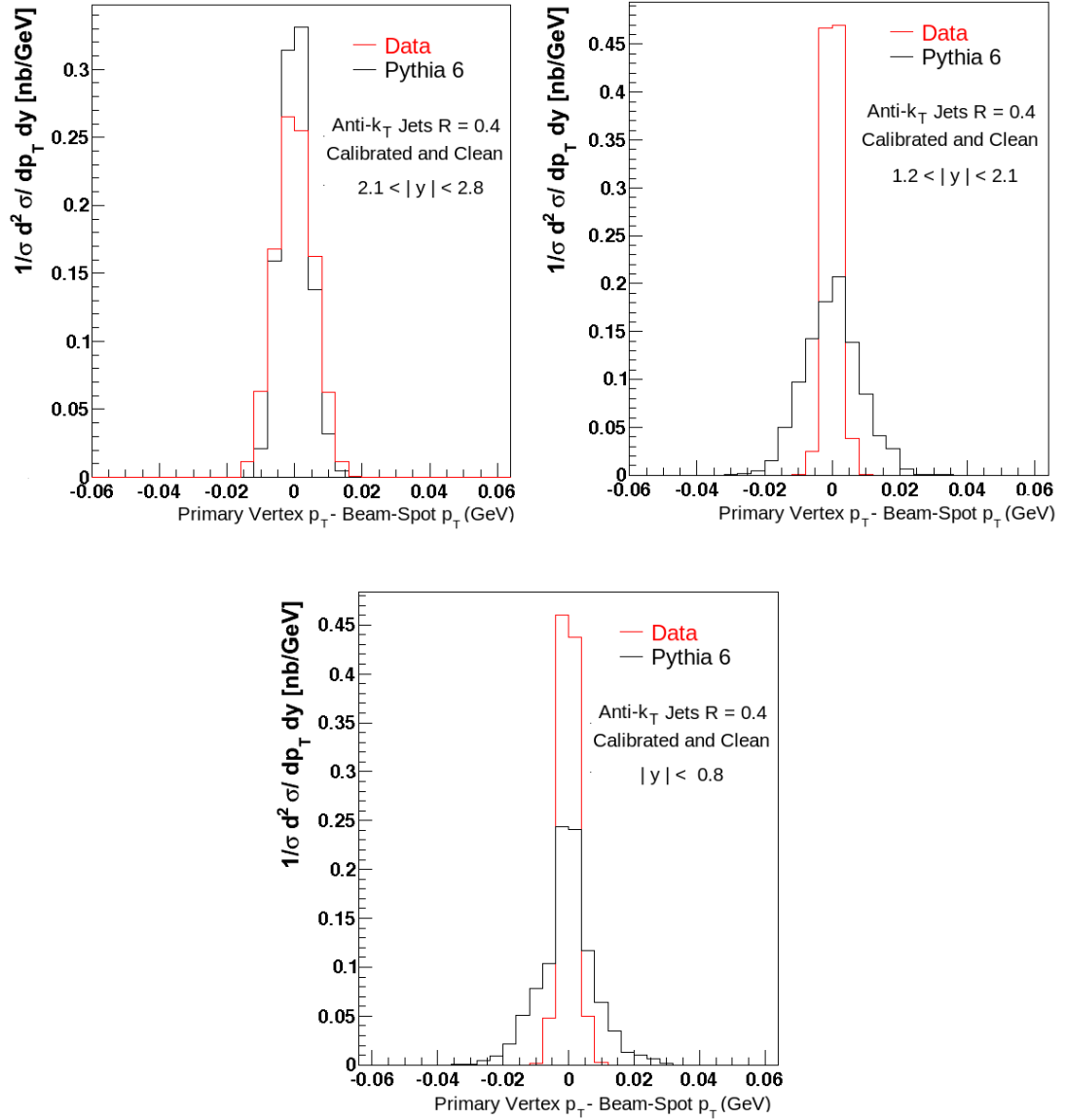
Figure 5.2: Number of events (normalised to unity) verses the difference between the p_T calculated for the primary vertex to that of the beam-spot as a fraction of beam-spot p_T. Jets are made with anti-k_T R = 0.4, with 'medium' cleaning (defined in Table 5.1). Results are shown for three different rapidity regions.

- **Inelastic beam-gas** caused via inelastic collisions of protons with gas inside the beam-pipe.

- **Elastic beam-gas** caused by elastic scattering of beam gas resulting in small deflections in the trajectory of the protons.

- **Parasitic and satellite collisions** refer to collisions of normal bunches, offset from the nominal ATLAS interaction point.

Sometimes un-captured beam is re-captured in a bunch, forming a secondary or 'ghost' bunch. Tertiary halo and elastic beam-gas protons tend to result in a relatively collimated flux of secondaries. High energy muons from these types of BIB are generally found within a radius of 2 m at the level of the interaction point. The inelastic BIB tends to result in a long tail of high energy muons in the horizontal plane.

5.7.2   Cosmic Ray Showers

Cosmic-ray showers are produced in the atmosphere, from extra-solar high energy particles colliding with atmospheric particles to produce a cascade of secondaries. After traversing the ATLAS cavern, this background tends to be comprised of primarily muons. They may pass a jet trigger if sufficient energy is deposited in the calorimeter, or else by joining a real signature to cause a lower threshold jet to pass a trigger which otherwise would not have been fired. When tracks are produced from cosmic-ray events their impact parameter with respect to the interaction point is typically large. This geometric feature has been found to be very efficient at rejecting cosmic-ray background. In Table 5.2, the cut flow for a sample of data triggered on cosmic events is shown. The primary vertex requirements placed on the jets remove all but 13 of the events, with only one of these remaining events containing a jet with sufficient p_T to enter the final analysis.

5.7.3   Triggering on Non-Collision Background

The bunch crossing ID (BCID) is used to characterise proton bunches. The bunches are categorised by the following criteria and are used for triggering on non-collision events:

Selection Cut	Number of events Passing
Initial Number of Events	5054990
Events with Optimal Detector Conditions	2588622
Events with Primary Vertex	13
Events with Jet $p_T > 20$ GeV	1
Events with Jet $\eta < 4.4$	1
Events after Jet Cleaning	1

Table 5.2: Cut Flow Table of Events which Pass Cosmic Trigger Requirements for 1.97 fb^{-1} of Data

Selection Cut	Number of events/jets Passing
Initial Number of Events	10018864
Events with Optimal Detector Conditions	8470799
Events with Primary Vertex	8272153
Events with Jet $p_T > 60$ GeV, $\eta < 4.4$	2667707
Jets after Cleaning	3950343
Jets passing Unpaired trigger for 10 GeV jets	242

Table 5.3: Cut Flow of Jet Events for 1.97 fb^{-1} of Data

- **Paired** a bunch in both LHC beams in the same bunch crossing.

- **Unpaired isolated** a bunch in only one LHC beam with no bunch in the other beam (within 3 BCIDs).

- **Unpaired non-isolated**, a bunch in only one LHC beam with a nearby (within 3 BCIDs) bunch in the other beam.

- **Empty** a bunch crossing containing no proton bunches.

It is required that cleaning sufficiently remove backgrounds produced from non-collision sources, else this must be accounted for in the systematic uncertainty in the measurement. Contributions from cosmic sources were tested by considering events passing the cosmic event triggers and 'Empty' triggers (defined above). Beam background is checked using data passing jet triggers and also the 'Unpaired' triggers, as
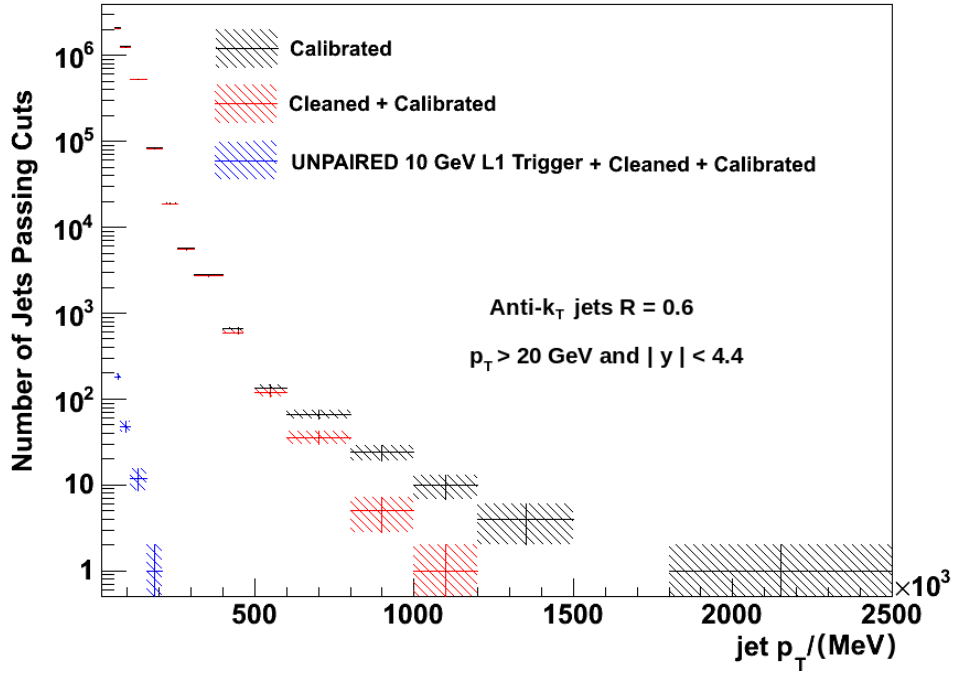
Figure 5.3: Non-collision jets passing events selection and cleaning criteria and passing Level-1 Unpaired Triggers.

defined above. The full analysis chain is applied to a sample of data (corresponding to 1.97 fb^{-1} in luminosity) and shown in Figure 5.3. The number of inclusive jets caused by BIBs that pass the analysis cleaning criteria is $< 0.007\%$ of the number of clean jets. Just 1 cosmic event passes after cleaning is applied.

The experimental uncertainties that contribute to this measurement (except those from statistical and the luminosity uncertainty of 3.4%) are shown in Figure 5.4. The jet energy resolution uncertainties are shown in white and jet energy scale uncertainties shown in yellow. All other uncertainties shown in green. The negligible contribution of jets that arise from non-collision sources to the cross-section after jet cleaning contribute 0% to the uncertainty (green). Figures 5.5 and 5.6 show the same experimental uncertainties applied to the inclusive and dijet cross-sections. The grey band is the quadratic sum of the experimental systematic uncertainties, to which BIBs and cosmic muons contribute an overall uncertainty of 0%. The error-bars indicate the statistical

uncertainty. The uncertainty due to the luminosity measurement is not included. This plot also includes the theory uncertainty, shown as the orange, hatched band. The sources of experimental and theoretical uncertainty on the measurement are discussed in Section 5.8 and Section 5.9.

5.8 Experimental Uncertainties

The experimental uncertainties (Figure 5.4) of the inclusive jet measurement are dominated by uncertainties from the JES calibration. Due to the steeply falling p_T spectra of the measurement, a small shift in transverse momentum uncertainty corresponds to a large change in the jet p_T spectrum.

JES uncertainty The JES for inclusive jet measurements is determined from Monte-Carlo and validated in-situ. Details are given in Chapter 4. The main sources of systematic uncertainty on the JES for inclusive jets include: deviation from unity of the response when the JES correction is applied to reconstructed Monte-Carlo; uncertainty on the single hadron response of the calorimeter; modelling of dead material and noise threshold in Monte-Carlo and the uncertainties arising from the η-intercalibration between central and forward jets.

JER uncertainty The main uncertainties in the inclusive jet JER arise from the different physics modelling in Monte-Carlo and from the uncertainties from dijet balancing and the bisector analyses (Details are given in Chapter 4).

Pile-Up The uncertainty after applying the pile-up correction is estimated as a function of the number of primary vertices. For two primary vertices per event with $p_T > 20$ GeV and $0.3 > |y| > 0.8$ the uncertainty is 1%. With $2.1 > |y| > 2.8$ it is 2%. For jets with $p_T > 200$ GeV, the uncertainty due to pile-up becomes negligible.

Reconstruction Efficiency The jet reconstruction efficiency is derived using the Monte-Carlo simulation and the systematic uncertainty evaluated by comparing to track jets. The jet reconstruction efficiency is well-described by the Monte-Carlo sim-
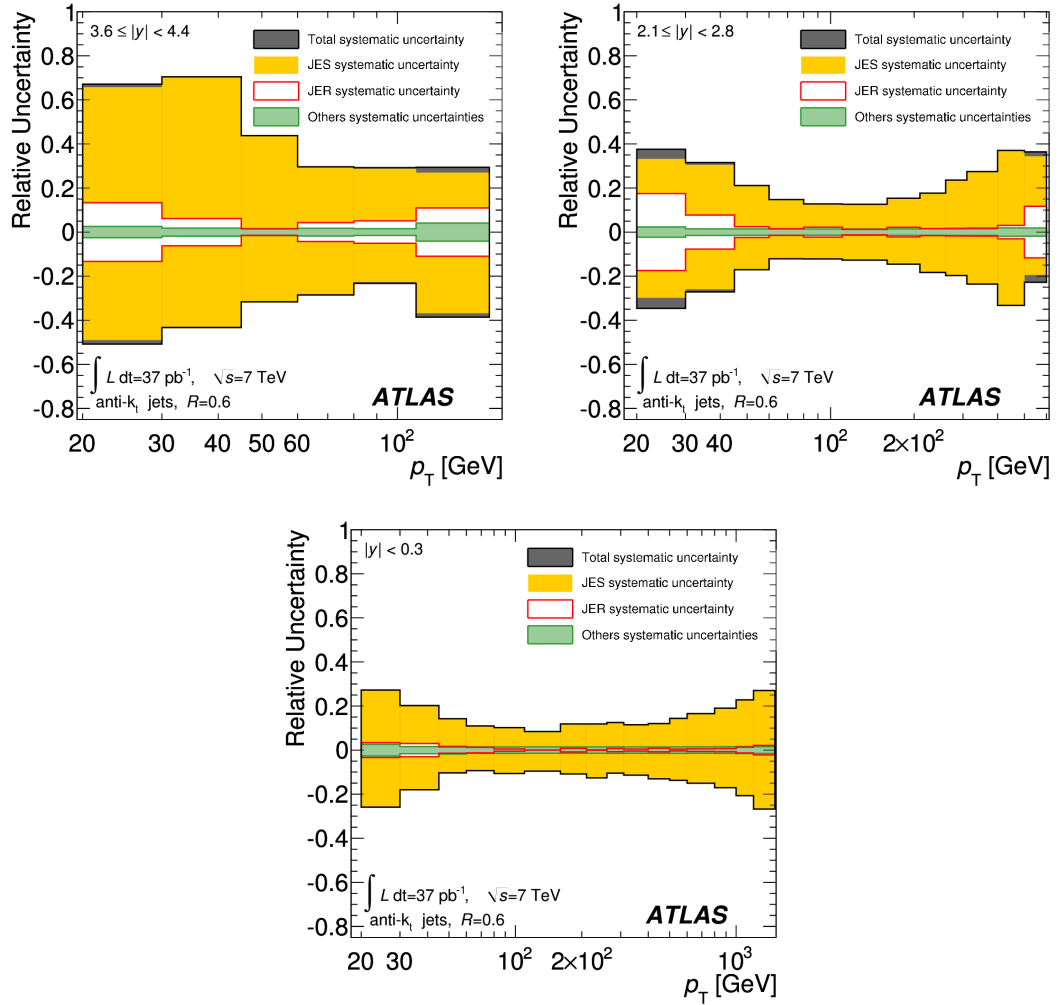
Figure 5.4: The total systematic uncertainty on the inclusive jet cross-section measurement for anti-k_T jets with R = 0.6 in three rapidity bins. Statistical uncertainty and the 3.4% uncertainty of the integrated luminosity are not shown.

ulation. The associated systematic uncertainty is below 2% for jets with $p_T < 30$ GeV and negligible for higher p_T.

Luminosity This is derived independently by several detectors, the results of which are combined to produce the overall ATLAS luminosity measurement and uncertainty. The uncertainty for the inclusive cross-section from luminosity during this data taking period is 3.4%.

Corrections for trigger and detector efficiencies are performed during unfolding from true (including muons and neutrinos) to reconstructed jets from Monte-Carlo. The first 17 nb^{-1} of the inclusive measurement uses the bin-by-bin method of unfolding and the full 2010 measurement shown in Figures 5.5 and 5.6 uses IDS (as discussed in Chapter 4).

5.9   Comparison to theory

The inclusive jet measurement is compared to theoretical predictions for next-to-leading order (NLO) perturbative QCD (pQCD) using NLOJET++ 4.1.2. The renormalisation and factorisation scales were set to the p_T of leading jet of the event and the uncertainty determined by varying the renormalisation scale from half, to double this value. The NLO calculations predict partonic cross-sections; however, these are not measurable in practice. Therefore NLO pQCD calculations were corrected for non-perturbative effects by applying a 'soft correction', derived using PYTHIA Monte-Carlo. This is obtained by calculating the ratio of the cross-section of inclusive jets with the underlying event (or hadronisation) implemented, to the cross-section of the Monte-Carlo without hadronisation and underlying event applied. The two separate corrections are multiplied together and then applied to each bin of p_T. The soft QCD corrections depend on the size of the jets. Wide jets cluster more additional particles from the underlying event. This effect is more significant at low p_T, where the contribution from underlying event is fractionally larger compared to the p_T of the hard process. Narrower jets are more susceptible to losing particles outside the jet clustering radius from the process of hadronisation. At low p_T, the jet constituents become less collimated which causes the impact of this effect to increase.

5.10   Theoretical Uncertainties

The theoretical uncertainty arises primarily from systematic uncertainties on the PDFs, the renormalisation and factorisation scales and the value taken α_s. They are derived using the APPLGRID software, combined with NLOJET++. These are shown, along with the experimental errors in Figures 5.5 and 5.6.

Renormalisation and Factorisation scales The theory prediction is recalculated while varying the renormalisation and factorisation scales independently to double and half their nominal values. The envelope of uncertainty is then taken as the total uncertainty. Overall, it contributes less than 5%, with the largest values found at low p_T.

PDF uncertainty PDFs are parameterised into a number of eigenvectors, the number of which depends on the PDF used. These parameters are shifted according to systematic uncertainty of the measurement, and the uncertainty is propagated through the PDF fits.

α_s **uncertainty** The cross-section is recalculated using special PDF sets with α_s changed within its known uncertainties. The uncertainties due to varying α_s contribute 2-3% for low p_T and increases to 4-5% for higher transverse momentum jets.

Non-perturbative corrections The theoretical prediction is calculated using different tunes of PYTHIA and HERWIG++. The overall uncertainty is taken as the maximum spread. This reaches up to 2-6% for anti-k$_T$ R = 0.6 and up to 10% for R = 0.4.

5.11   Inclusive Jet Cross-Section Summary and Discussion

The inclusive jet cross-section is measured from 20 GeV out to 1.5 TeV, for anti-k$_T$ R = 0.4 and R = 0.6 jets. The measurement is extended to the forward regions, up to $|y| <$ 4.4. The inclusive jet cross-section is well described by theoretical predictions within experimental uncertainty. The contribution from this thesis involved understanding the extent and impact of non-collision backgrounds on the measurement. These were shown to have minimal impact on the overall uncertainty of the measurement, once jet cleaning is applied. The inclusive jet spectrum is purely a QCD $2 \rightarrow 2$ process, and thus provides direct information on the initial partons. Furthermore, the cross-section is used in the measurement of α_s, and therefore contributes to an important uncertainty on the calculation of PDFs.
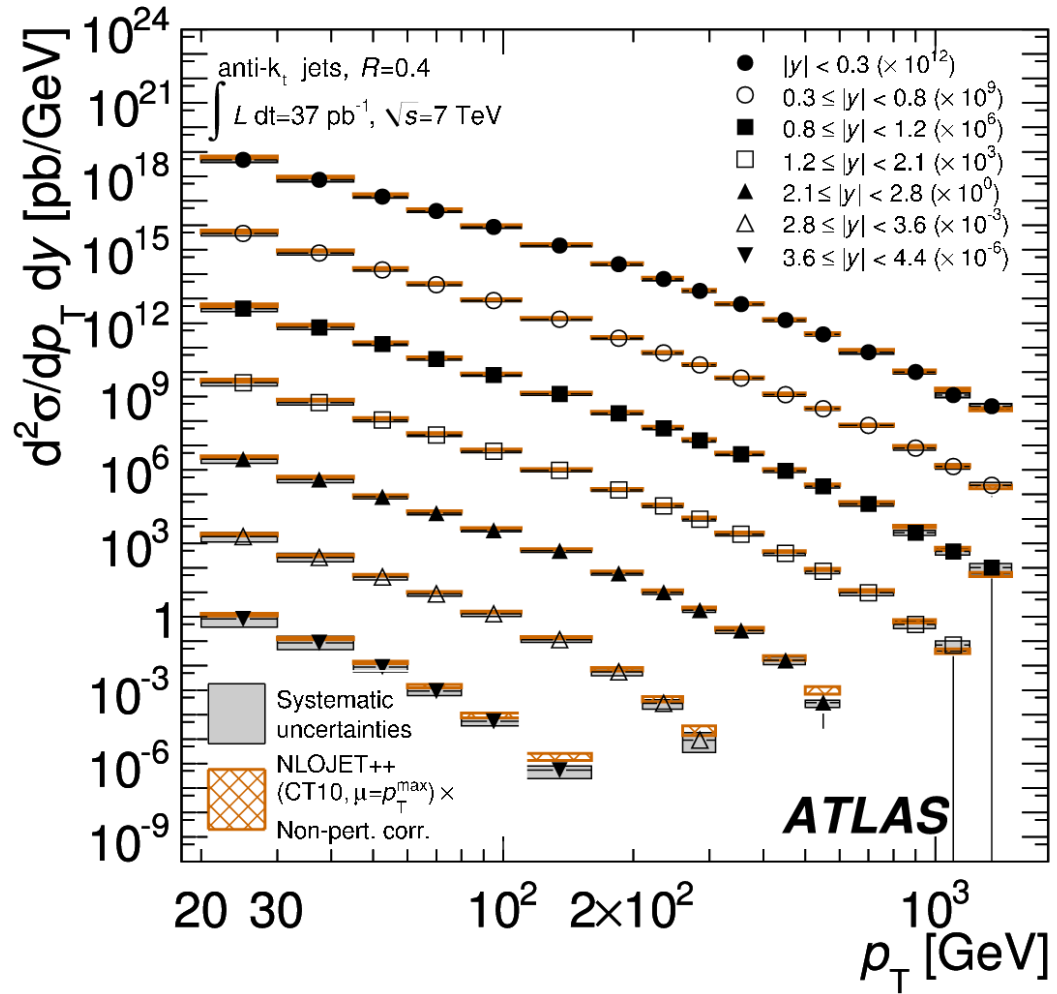
Figure 5.5: The inclusive jet double-differential cross-section as a function of jet p_T in different regions of $|y|$ for anti-k$_T$ R = 0.4 [42].
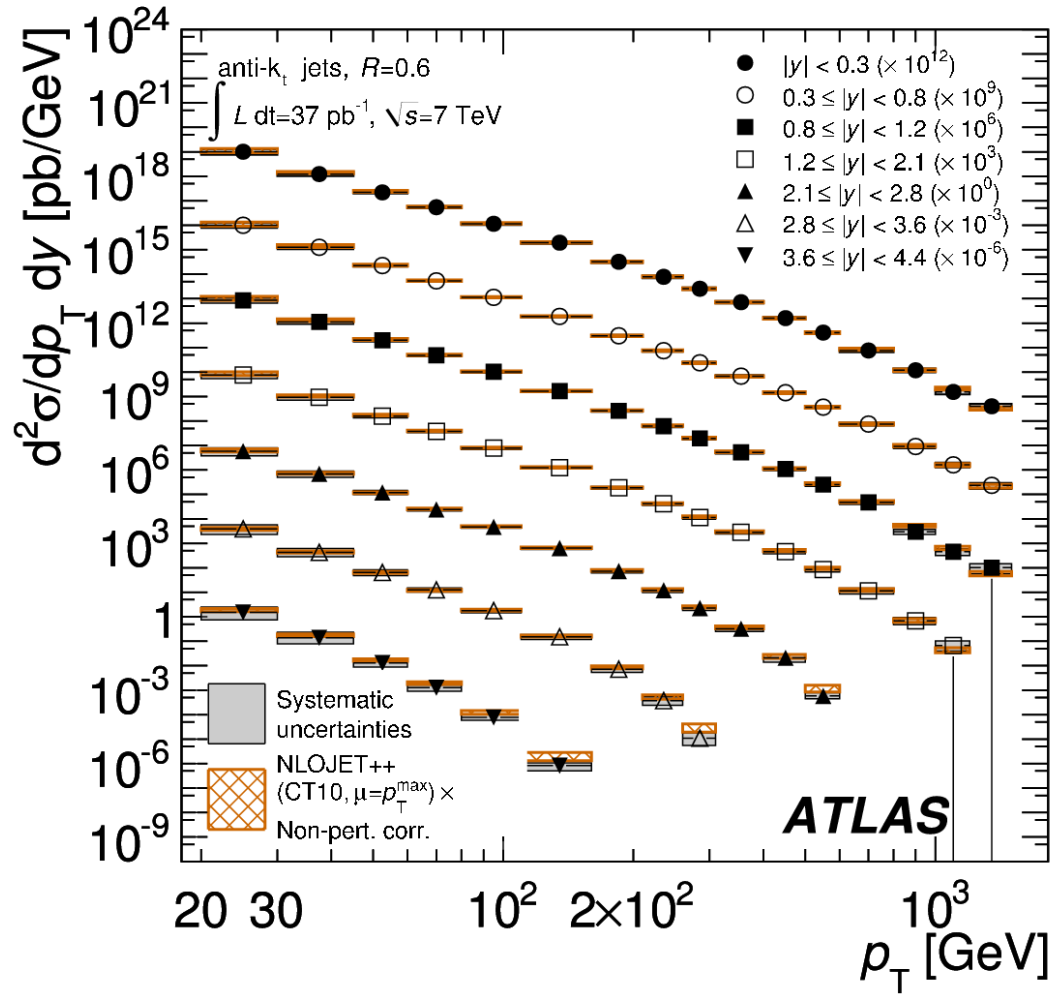
Figure 5.6: The inclusive jet double-differential cross-section as a function of jet p_T in different regions of $|y|$ for anti-k_T R = 0.6 [42].

Chapter 6

Sub-structure Measurements in ATLAS

6.1 Introduction

The sub-structure of jets has been proposed as a means to distinguish a signal of inter-esting new physics from QCD backgrounds. Jets are composite objects with inherent internal structure, which can be exploited to reduce effects such as pile-up and under-lying event. The technique is being used in several Standard Model measurements, as well as in searches for new physics [43]. However, prior to making such measure-ments, confidence in the modelling of the sub-structure of jets must first be established.

The following chapter outlines the first published jet mass measurement from the ATLAS detector [39]. Measurements of jet mass, k_T splitting scales and n-subjettiness ratios (as defined in Chapter 2) are performed for anti-k_T jets with R = 1.0 and Cambridge-Aachen jets with R = 1.2. Jet mass after splitting and filtering is mea-sured for Cambridge-Aachen jets with R = 1.2. The measurement has been made with an inclusive sample of jets, collected at a centre-of-mass energy of 7 TeV. Events are required to pass the selection criteria specified below, resulting in a total integrated luminosity of 35 pb^{-1}.

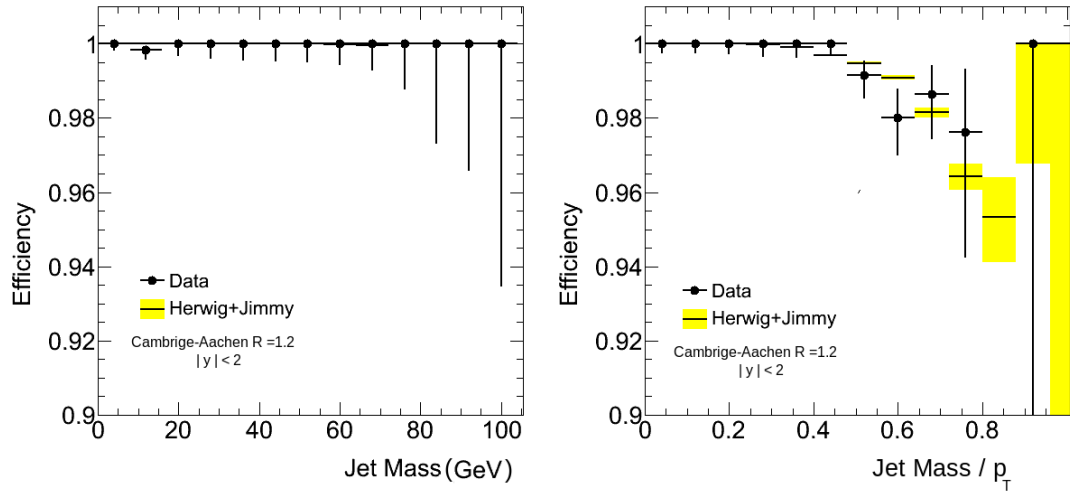Figure 6.1: Trigger efficiencies for mass (left) and mass/p_T (right), using the Level-1 95 GeV jet trigger, for Cambridge-Aachen jets with $400 < p_T < 500$ GeV
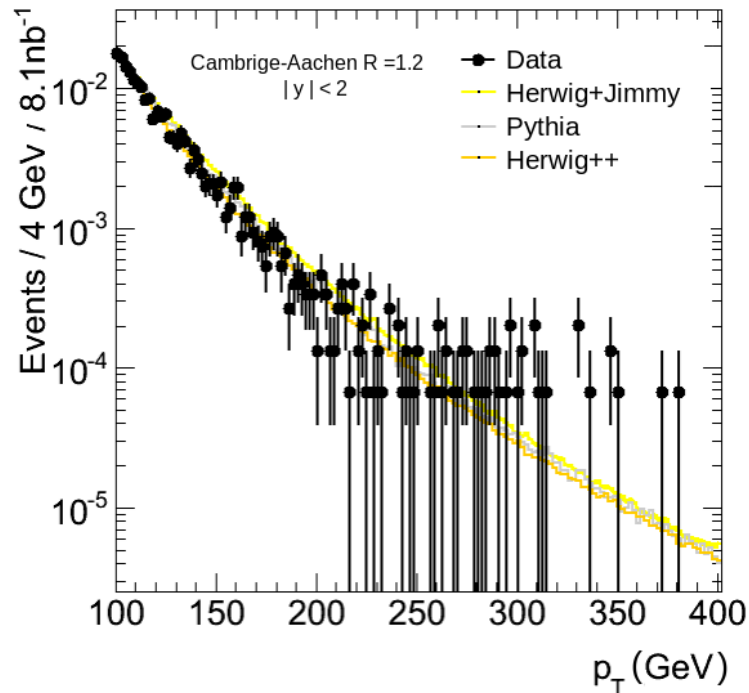
.

Figure 6.2: The p_T distribution of the three hardest sub-jets for Cambridge-Aachen R = 1.2 jets with $|\eta| < 2.0$. Figures have been normalised to unity.
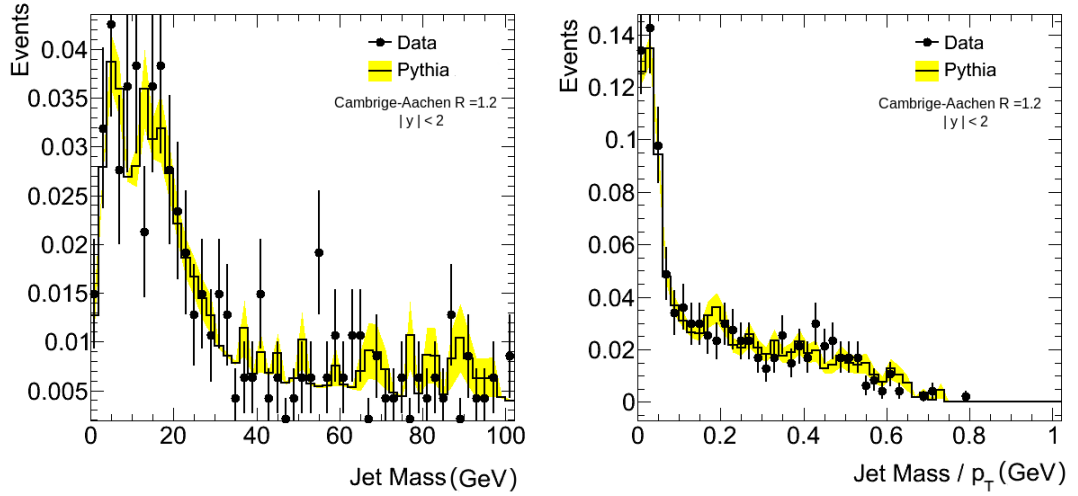
Figure 6.3: Mass (left) and mass/p_T (right) of the three hardest sub-jets. Figures have been normalised to unity.

6.2 Sub-structure Studies In Early Data

A sample of events recorded at a centre-of-mass energy of 7 TeV and representing the first 8 nb^{-1} of data are studied. Events are triggered using the Level-1 95 GeV jet trigger. Jets are made with the Cambridge-Aachen jet finder with R = 1.2 using 'locally calibrated' clusters as inputs. Local calibration refers to the calibration applied to restore the response of hadronic clusters and is discussed in detail in Chapter 4.

Level-1 95 GeV jet trigger efficiencies for Cambridge-Aachen jets with $400 < p_T < 500$ and $|\eta| < 2.0$, are shown for mass and mass/p_T in Figure 6.1. Trigger efficiency is here defined as the ratio between the number of jets events where at least one of the jets fired the trigger to the number of all jets in all events. For the range $400 < p_T < 500$, efficiency with respect to p_T is above 99%. The efficiency of the trigger against mass is shown to be well modelled by Monte-Carlo. The distribution of mass/p_T is directly related to the angular scale between sub-jets. This variable is well modelled by Monte-Carlo.

After trigger selection, quality selection variables (as defined in Chapter 4) are applied according to the inclusive jet specification shown in Table 5.1 on every anti-k$_T$ R

= 0.6 jets in the event. An event with an anti-k$_T$ R = 0.6 jet which fails to pass the selection criteria is disregarded. Jets are chosen with $400 < p_T < 500$ GeV, $|\eta| < 2.0$, and with the number of tracks pointing to the vertex < 4, and with a vertex position that is consistent with the luminous region (beam-spot). Sub-structure analysis (as described in Section 2.5) is performed on the Cambridge-Aachen jets, and the p_T spectrum (Figure 6.2), mass and mass/p_T (Figure 6.3) of the three hardest sub-jets are compared with Monte-Carlo. Distributions are initially well described by the Monte-Carlo, after sub-structure analysis had been applied.

6.3   Sub-structure Measurements on the 2010 Dataset

6.4   Jet Reconstruction

Jets are reconstructed using the anti-k$_T$ algorithm with distance parameter R = 1.0, the Cambridge-Aachen algorithm with distance parameter R = 1.2. The Cambridge-Aachen are also shown after splitting and filtering (using the method described in Chapter 2). Using these two types of jet definition allows an assessment to be made of the impact of the choice of jet algorithm on the measurement. Jets with wide radii are preferred for boosted object searches in order to increase the efficiency with which the boosted decay products are captured. The jets are calibrated according to the specification described in Chapter 4 for sub-structure jets.

6.5   Trigger Selection

The ATLAS trigger system is described in Chapter 3. Events are required to pass the Level-1 hardware triggers, with no offline High Level Trigger (HLT) selection applied. In the lowest p_T bin (200-300 GeV), jets at the electromagnetic scale (see Chapter 2) must pass a 30 GeV jet trigger. The 30 GeV trigger is used until the point when this trigger becomes prescaled. All other events are triggered using a 95 GeV jet trigger which has no prescale applied over the entire data taking period.

6.6 Jet Quality

Jets are required to be within the range $|\eta| < 2$, in four 100 GeV p_T bins spanning 200 to 600 GeV. In the p_T range selected the jets are on the trigger high efficiency plateau (efficiency > 99%). The $|\eta|$ restrained ensures they are contained entirely within the barrel and end-cap sub-detectors.

Background Source	Removal Cuts
HEC Spikes	$f_{\text{HEC}} > 0.5$ **AND** $|Q_{\text{HEC}}| > 0.5$
HEC Spikes	$|E < 0| > 60$ GeV
EM Noise	$f_{\text{EM}} > 0.485$ **AND** $|Q_{\text{LAr}}| > 0.8$ **AND** $|\eta_{\text{EM}}| < 2.8$
Cosmics	$|t| < 10$ ns
Cosmics	$f_{\text{EM}} < 0.05$ **AND** $f_{\text{charged}} < 0.05$ **AND** $|\eta_{EM}| < 2$
Cosmics	$f_{\text{EM}} < 0.05$ **AND** $|\eta_{EM}| \geq 2.$
Cosmics	$f_{\text{Max}} > 0.489$ **AND** $|\eta_{EM}| < 2$

Table 6.1: Cleaning variables specific to the sub-structure analysis.

Selection cuts are applied to remove non-collision backgrounds and other sources of noise. The 'standard' jet cleaning criteria in ATLAS are optimised for anti-k_T R = 0.6 jets. Therefore, as with early data studies (Section 6.3), the cleaning is performed using jets reconstructed using anti-k_T R = 0.6. Events with an anti-k_T R = 0.6 jet which exceeds a threshold of 30 GeV and fails the selection criteria are disregarded. This selection corresponds to a reduction in data of 3%.

Events are required to contain a primary vertex consistent with the position of the LHC beam-spot, with five or more tracks with $p_T > 150$ MeV. The 2010 dataset was subject to increasing levels of pile-up contamination with increasing luminosity. Pile-up has been shown to have a strong impact on jet mass reconstruction and sub-structure [44]. In order to mitigate the problems associated with pile-up, events are selected with exactly one primary vertex. This selects 22% of the total dataset. Vertex position is well modelled by Monte-Carlo [39].

The Monte-Carlo samples are subject to the same selection criteria as data, ex-

cept the cleaning derived for the removal of noise and non-collision backgrounds. The cleaning selection cuts have a negligible impact on Monte-Carlo.

6.7   Detector-level Distributions

Detector-level distributions for Cambridge-Aachen 1.2 are shown for jet p_T, η, mass in Figure 6.4 and the sub-structure variables $\sqrt{d_{12}}$, $\sqrt{d_{23}}$, τ_{21} and τ_{32} are shown in Figure 6.5. The simulated distributions are normalised to the data separately in each plot. The Monte-Carlo describes the shape of the data for these variables well before systematic uncertainties are included. This provides confidence in the effectiveness underlying modelling of these variables in QCD for both HERWIG and PYTHIA. Better levels of agreement are found after splitting filtering is applied (as discussed in Chapter 2). Some of the jets fail the selection after the splitting and filtering procedure, and so there are approximately four times fewer jets in the distribution [39].

6.8   Experimental Uncertainties

As with the inclusive jet cross-section, the most important systematic uncertainty on this measurement is due to the modelling of the response of the calorimeter. However, the methods used to evaluate the uncertainties are previously performed only for anti-k_T with R = 0.4 and R = 0.6. The jet energy scale (JES) and jet energy resolution (JER) need to be evaluated for the larger Cambridge-Aachen R = 1.2 and anti-k_T with R = 1.0 jets used in this analysis. Furthermore, for sub-structure measurements, the mass of jets must also be calibrated. Thus, the jet mass scale (JMS) uncertainty and jet mass resolution (JMR) uncertainty must also be calculated and included in the final result. The calculation of jet mass resolution is performed in Monte-Carlo, as described in (Chapter 4). It is possible to reapply the methods used to validate the JES uncertainty for inclusive jets to the sub-structure jets; however, no corresponding evaluation of JMS is found for jets in ATLAS. One solution is to use jets created from tracks in the Inner Detector (track jets) to constrain the uncertainty on jets made from the calorimeter (calorimeter jets). Track jets and calorimeter jets have different sources of systematic uncertainty, and therefore comparisons between the two can be used to constrain the uncertainty of the calorimeter. This method, known as the 'double ratio'

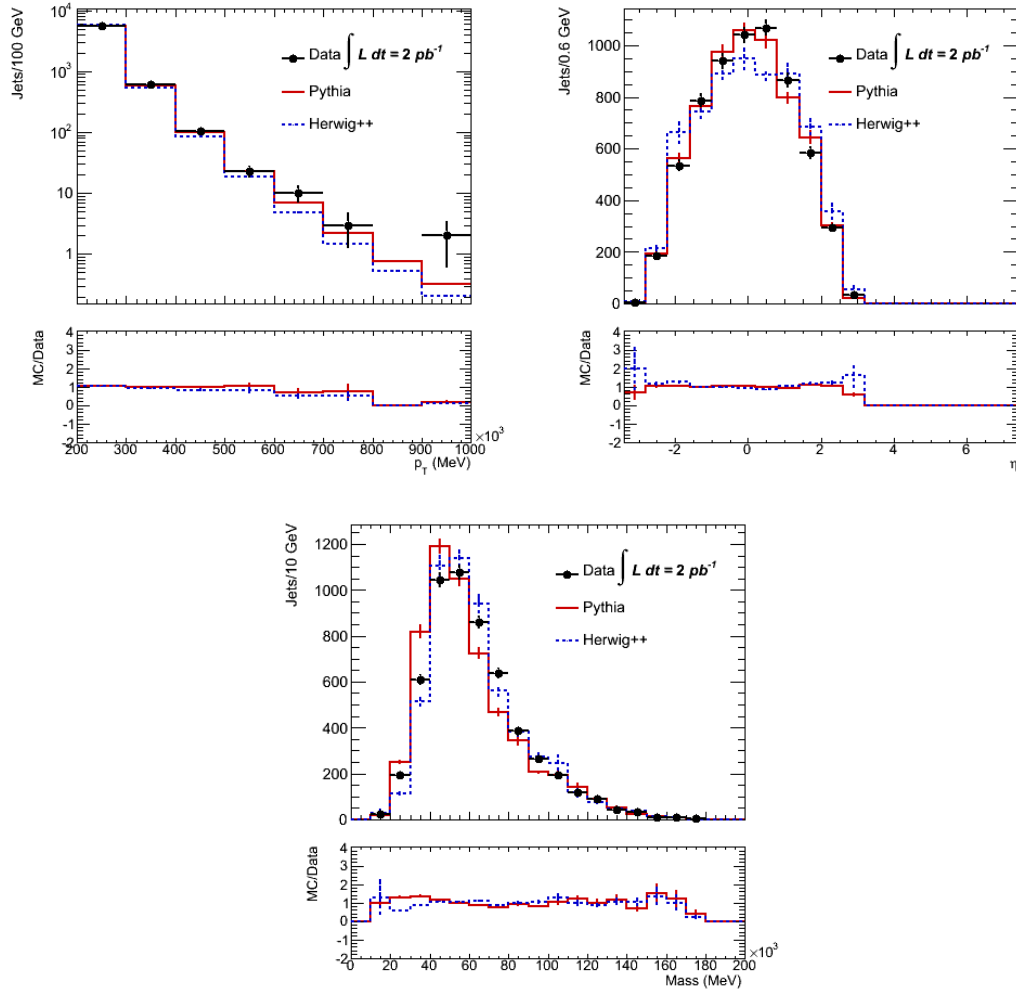Figure 6.4: p_T (top left), η (top right) and mass (bottom) distributions for Cambridge-Aachen Jets for $p_T > 200$ GeV

method, necessitates that the uncertainty on tracking is relatively small. The extension of this method from evaluating energy, to mass and other variables is straightforward. Therefore, all scale uncertainties for the variables in this analysis are calculated using this method of comparison.
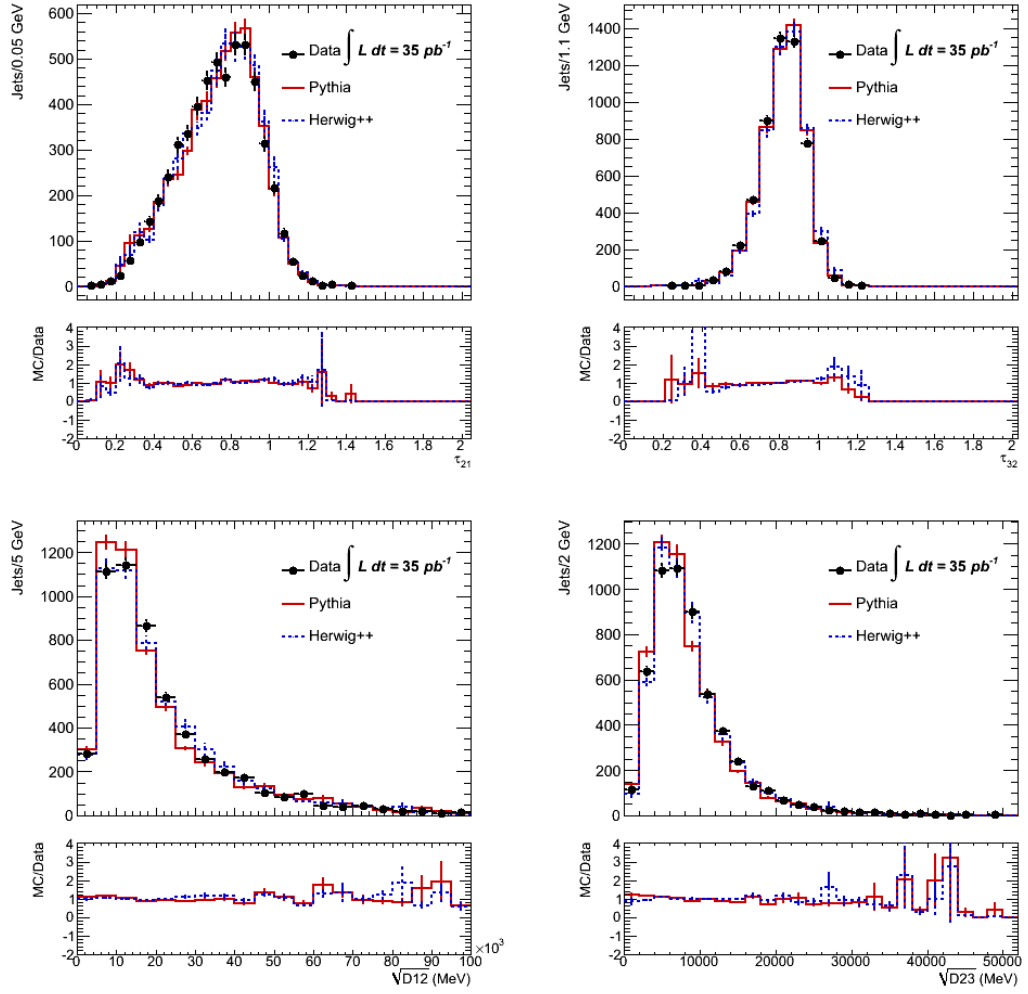
Figure 6.5: τ_{21} (top left), τ_{32} (top right) $\sqrt{d_{12}}$ (bottom right) and $\sqrt{d_{23}}$ (bottom left) distributions for Cambridge-Aachen Jets for $p_T > 200$ GeV

6.9 The Double Ratio

Track jets are matched to calorimeter jets by means of a geometrical selection using ΔR between the track jet and calorimeter jet with $\Delta R < 0.3$. The uncertainty is evaluated for each variable X with $X = p_T$, η, $mass$, $\sqrt{d_{12}}$, $\sqrt{d_{23}}$, τ_{21} and τ_{32} by computing the ratio r^X (defined below) between the distributions of X for calorimeter jets and track jets.

$$r^X = \frac{X_{calorimeter-jet}}{X_{track-jet}} \tag{6.1}$$

At low p_T, low momentum charged particles can fall below the minimum cut-off used for selecting charged tracks. This may cause biases in the ratio for the case where charged tracks are lost. However, if the Monte-Carlo correctly models the data for this effect then the double ratios ρ^X (as defined by equation 6.2) are expected to be consistent with unity.

$$\rho^X = \frac{r^X{}_{data}}{r^X{}_{Monte-Carlo}} \tag{6.2}$$

The value of ρ^X relies on the performance of the Inner Detector and the modelling of the charged to total momentum of jets by simulation. Jets are restricted to $\eta > 1.0$ so that they are fully within the tracking acceptance. The double ratio plots for data, PYTHIA and HERWIG are shown in Figures 6.6 to 6.7.

The uncertainty for each Monte-Carlo is calculated as the average deviation of the double ratio from unity. The final uncertainty is taken as the maximum of all the Monte-Carlo. The uncertainty values used in the final unfolding also included SHERPA - a variety of Monte-Carlo samples are preferred to account for uncertainties from different fragmentation and hadronisation models. This final value is measured at 3% to 5%.

6.10   Uncertainties on Tracks

The double ratio method relies on the assumption that the uncertainties associated with tracking are small in comparison to those of the calorimeter. Previous studies for anti-k$_T$ R = 0.6 demonstrated that the dominant tracking uncertainty arises from inefficiencies in track reconstruction [45].

The tracking efficiencies and resolution for jets can differ from those of isolated tracks. Within a jet the tracks may be close together spatially leading to multiple
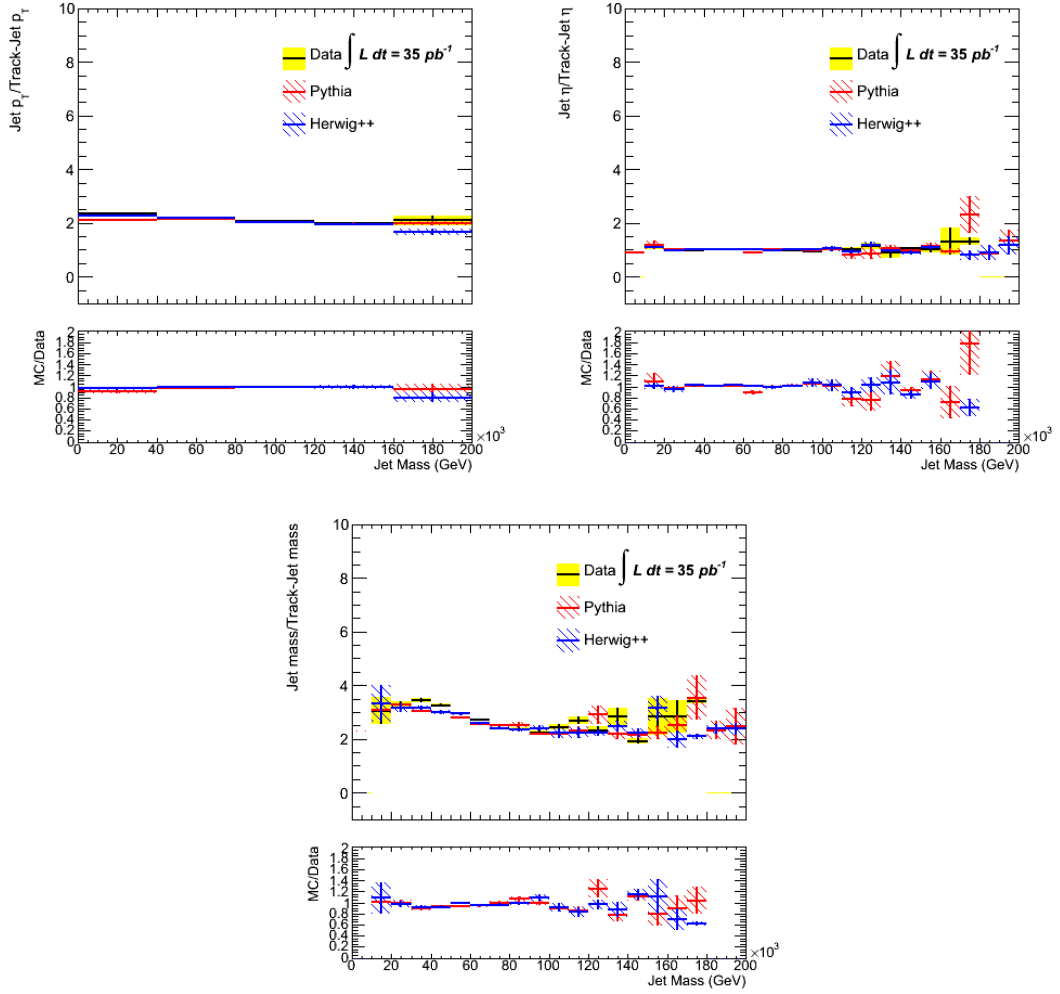
Figure 6.6: r^X of jet p_T (top left) and η (top right) mass (bottom) of the calorimeter to that determined by tracks versus the calorimeter jet. The bottom frame shows the double ratio ρ^X.

hits in a particular layer of the ID. The reconstructed position of the track may be affected by the presence of another track. These 'shared' hits increase the probability of incorrectly assigning a hit to a track or not associating the track to a corresponding TRT hit. At high densities, 'fake' tracks may be reconstructed from the combination of hits from multiple truth particles. Furthermore, high p_T tracks that are close together geometrically can share hits over several ID layers. In this case, the two tracks may be incorrectly reconstructed as a single track, leading to a loss of efficiency as a function of increasing p_T. The impact of these effects for non-standard, high p_T jets needed to

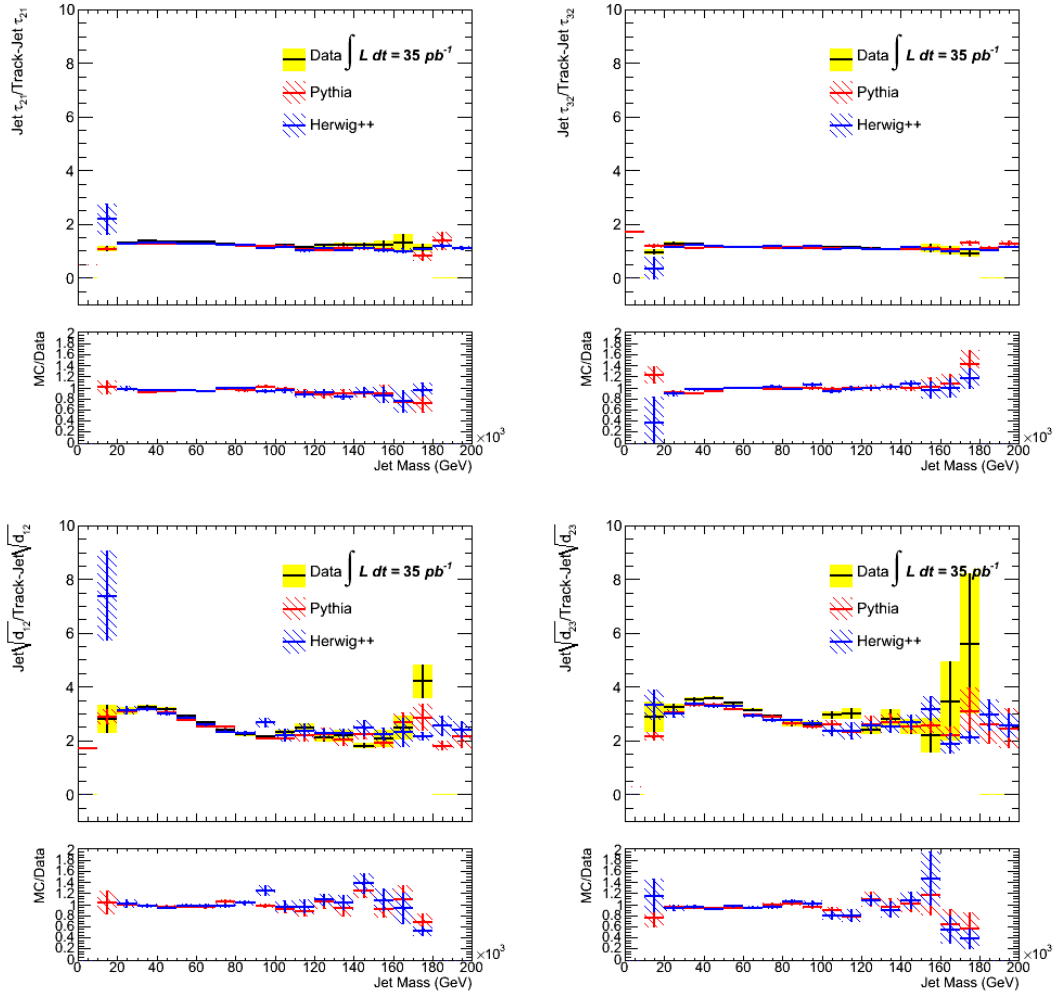Figure 6.7: r^X of jet τ_{21} (top left), τ_{32} of (top right), $\sqrt{d_{12}}$ (bottom left) and $\sqrt{d_{23}}$ (bottom right) of the calorimeter to that determined by tracks versus the calorimeter jet mass for jets with 300 - 400 GeV in p_T. Shown are the data and a variety of Monte-Carlo models. The bottom frame shows the ratio of the Monte-Carlo models to data ρ^X.

be evaluated. The robustness of the sub-structure variables $\sqrt{d_{12}}$, $\sqrt{d_{23}}$, τ_{21} and τ_{32} against tracking inefficiencies must also be understood.

6.11   Degradation of Resolution

The distribution of the average number of hits associated to a track and the average number of shared jet hits for the inner detector sub-systems is compared to PYTHIA, HERWIG++ and HERWIG+JIMMY Monte-Carlo event generators. Tracks are measured with respect to the primary vertex. Quality is ensured with the section that $p_T > 0.5$ GeV, $N_{pixel} \geq 1$, $N_{SCT} \geq 6$, $d_0 < 1.5$ mm and $z_0 . \sin(\theta) < 1.5$ mm. N_{pixel} and N_{SCT} are the number of hits in the pixel and SCT detectors, respectively. The variables d_0 and z_0 are the transverse and longitudinal impact parameters, and θ is the polar angle. The requirement that track $p_T > 0.5$ GeV is used to remove tracks arising from softer interactions. Tracks are associated to a calorimeter jet if ΔR between the track and the jet is less than the distance parameter used in the jet reconstruction. Geometrical matching of this type is more effective for the anti-k_T algorithm, which tends to produce cone-like jets. For this reason, only results for anti-k_T jets are given. However, distributions were found to be broadly consistent between both jet definitions.

The density of tracks is expected to be higher in the core of the jet and for jets with a higher p_T. Therefore, distributions are shown with respect to the calorimeter jet transverse momentum and as a function of ΔR between the track and the center of the jet. Figures 6.8 to 6.10) show the average number of hits for the Pixel layer, SCT and TRT, and the fraction of shared hits for both the Pixel layer and SCT. For the SCT and Pixel layer the of number of hits decreases slightly with increasing p_T. The fraction of shared hits increases with rising p_T by a factor of 5, and increases with respect to core of the jet by a factor of 2. The SCT and Pixel detector are consistently better modelled by PYTHIA Monte-Carlo. However, for the highest p_T jets considered (900 - 1000 GeV) the average number of TRT hits in PYTHIA can be up to 1.5 hits more than TRT hits in data with PYTHIA performing the least well of all the generators. However, in the maximum jet p_T bin (500 - 600 GeV) used in the sub-structure analysis, the difference between Monte-Carlo and data remains less than 1 hit in the TRT for all generators.

This is considered sufficiently small that all generators may be used in further tracking studies.

PYTHIA in general performs best in modelling the ID sub-systems and thus is used as the baseline for the inefficiency studies shown in section 6.11.1. PYTHIA is also used as the primary Monte-Carlo sample for deriving the tracking systematic uncertainty for the final measurement, with the other generators providing a comparison to determine the effect on the systematic uncertainty arising from different fragmentation and soft physics models.

6.11.1   Fake Rates and Tracking Efficiencies

Monte-Carlo based truth studies were undertaken to determine the contribution to overall systematic uncertainty from tracking inefficiency. Efficiency is studied for anti-k_T R = 1.0 and Cambridge-Aachen R = 1.2, as a function of p_T, $mass$, $\sqrt{d_{12}}$, $\sqrt{d_{23}}$, τ_{21} and τ_{32}. Fake rates and badly measured tracks are further constrained by considering the cases where the p_T of the track is larger than p_T jet. This situation will only occur in the case when the either the track or the jet is mis-measured.

Reconstructed tracks in the Monte-Carlo sample are matched to truth using a hit level truth-matching method. Each hit on track is assigned a weight (10 for pixel, 5 for SCT, 1 for TRT)[1] and the 'probability' P_{match} is defined the sum of the weights of hits coming from a truth particle, divided by the sum of all weights. By default the track is matched to the truth particle with the highest probability.

$$P_{match} = \frac{10 \times N^{common}_{pixel} + 5 \times N^{common}_{SCT} + 1 \times N^{common}_{TRT}}{10 \times N^{track}_{pixel} + 5 \times N^{track}_{SCT} + 1 \times N^{track}_{TRT}} \quad (6.3)$$

'N^{common}' denotes the number of hits common between the track and truth particle and 'N^{track}' denotes the total number of hits on the track. P_{match} can be considered a measure how many hits on the track are also associated with truth particles. A value of zero for P_{match} indicates no hits that come from truth particles or a 'fake' track. This

---

[1]The weights reflect the detector design in terms of expected number of hits, as shown in section 6.11.
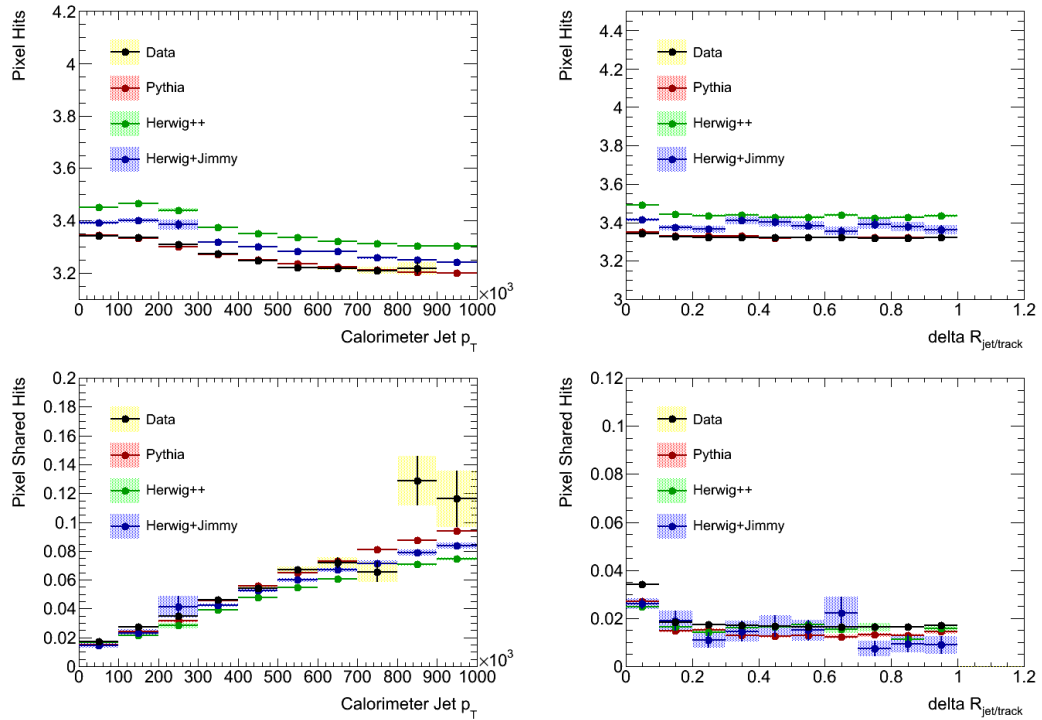
Figure 6.8: Pixel hits as a function of p_T (top left) and Δ R (top right) and Pixel Shared Hits as a function of p_T (bottom left) and Δ R (bottom right) for anti-k$_T$ 1.0 jets

method of hit matching does not account for tracks that are formed by hits from the decay products of a primary particle. A more rigorous approach requires the inclusion of additional terms in 6.3 which encompass more fully the topologies of such decays, in order to conserve the primary particle momentum. Also, the number of hits is a discrete variable, and therefore P_{match} becomes difficult to interpret for low momentum tracks, with a small number of hits. Consequently, alternative methods have since been developed [46]. However, this particular variable has the advantage of being independent of the detector resolutions, and provides a suitable indicator of quality for the high transverse momentum jets that are considered in boosted analyses. The distribution of P_{match} is shown in Figure 6.11. Zero value or fake P_{match} values contribute less than 0.02% over the full kinematic range. Although the track match cut-off is somewhat arbitrarily chosen, a very high cut on truth-match parameter may remove tracks that are incorrectly assigned hits from another track, particularly in the high density core
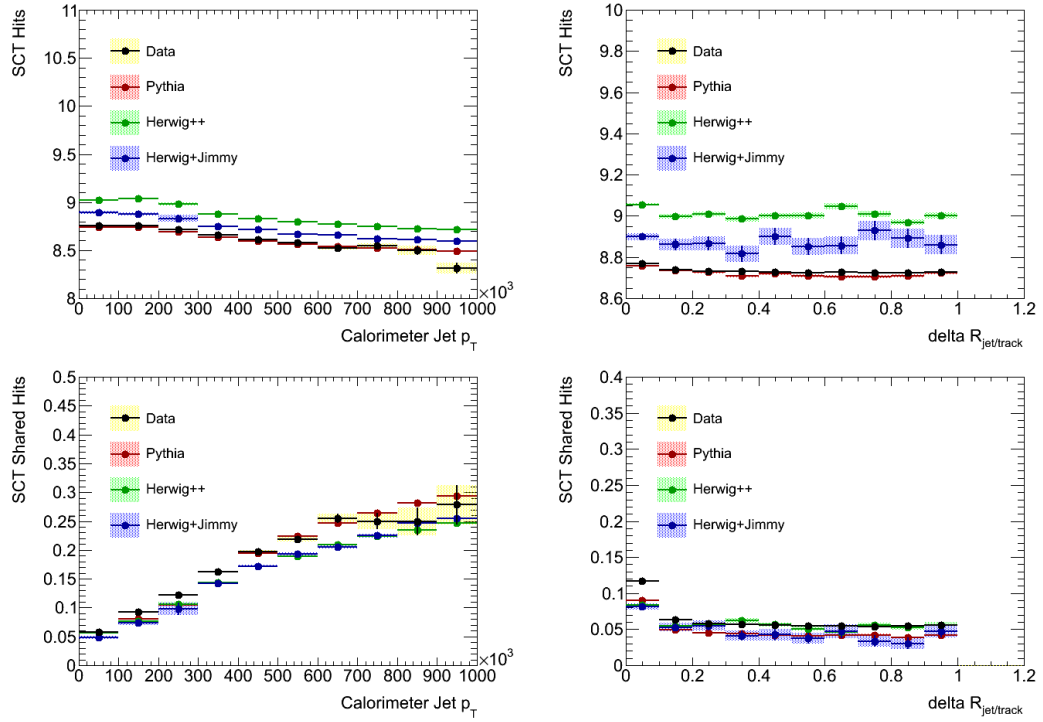
Figure 6.9: SCT hits as a function of p_T (top left) and Δ R (top right) and SCT shared hits as a function of p_T (bottom left) and Δ R (bottom right) for anti-k_T 1.0 jets
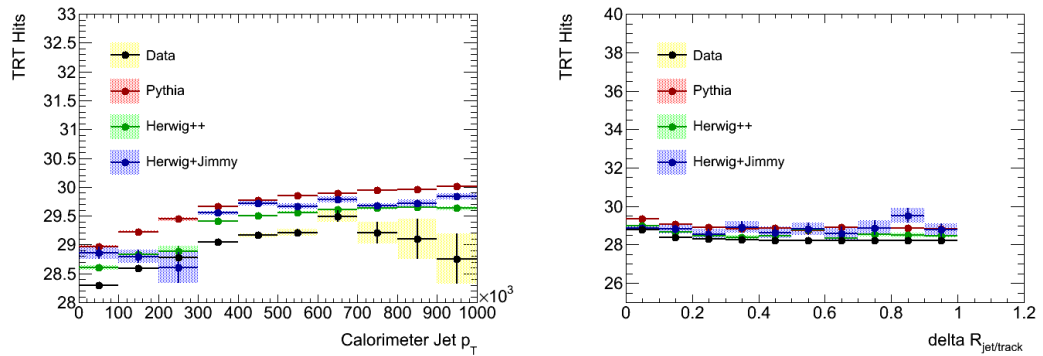

Figure 6.10: TRT hits as a function of p_T (left) and Δ R (right) for anti-k_T 1.0 jets
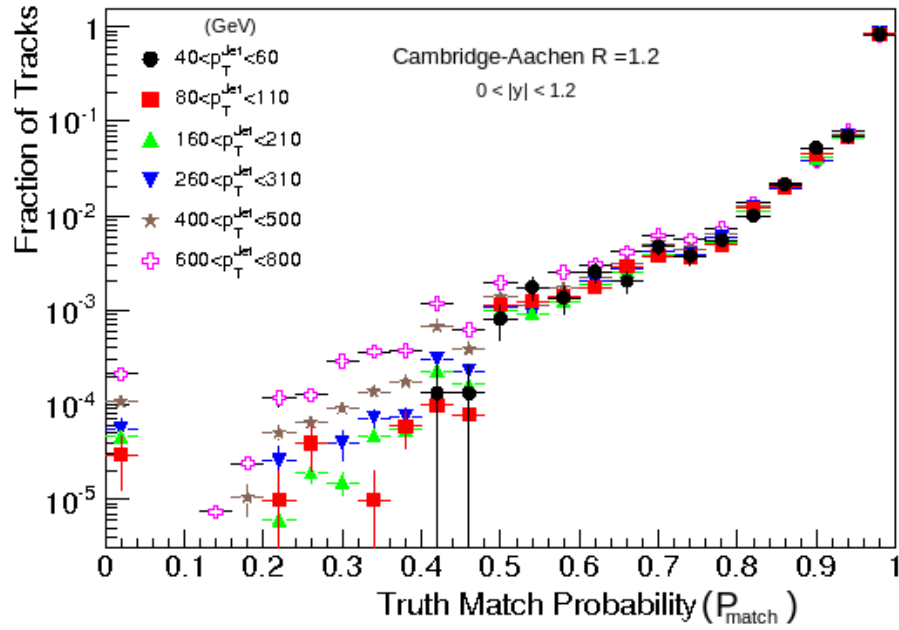
Figure 6.11: Probability of a true track being consistent with a reconstructed track in Monte-Carlo.

of the jet. This situation is avoided by choosing a value of P_{match} such that genuine track candidates are not rejected.

In order to understand the impact of shared on the tracking, the following ratio R_{match} is considered:

$$R_{match} = \frac{track_{match}}{truth_{all}} \tag{6.4}$$

Where $track_{match}$ is calculated for each variable X as the sum of tracks in a track jet associated to the calorimeter jet where $X = p_T$, η, $mass$, $\sqrt{d_{12}}$, $\sqrt{d_{23}}$, τ_{21} and τ_{32} with a truth match passing the match value cut-off. $truth - all$ are all truth particles within a geometrical match to a truth jet. A minimum p_T cut is assigned to the truth particles and tracks of 0.5 GeV. No statistical errors are shown due to the strong correlations between the numerator and denominator and also migrations between bins

making a simple interpretation difficult. This is shown as function of p_T for several values of P_{match} in figure 6.12. R_{match} never reaches 100% due to losses from hadronic interactions. There is a large p_T dependence, with high p_T tracks are being preferentially lost. It was demonstrated in section 6.11 that the numbers of recorded Pixel and SCT hits on a track are minimally dependent on the transverse momentum of the calorimeter jet. More importantly, the number of shared hits increases as a function of p_T. Inefficiencies arise in the case where two very high p_T tracks are close together, such that only one of the two tracks is properly reconstructed. For the sub-structure jets are considered up to $p_T < 600$ GeV the p_T distribution shown in figure 6.4 demonstrates that only a tiny fraction of jets have a transverse momentum above this cut-off. Overall effects of around 3% are seen for p_T and mass. For sub-structure variables these distributions have a minimal dependence on increasing p_T (Figure 6.13), and the uncertainty is conservatively estimated at 1.5%.

A constraint is provided on the fraction of fake, or badly measured tracks by considering the ratio of the p_T of the track to that of the associated calorimeter jet. Although the shape of this distribution is dependent on the fragmentation model used in the Monte-Carlo, this ratio should only be greater than unity in the case of mis-measurement. Data is found to have a higher number of mis-measured jets than Monte-Carlo, an effect that becomes more apparent with increasing p_T. The number of mis-measured tracks is below 0.1% across all p_T bins. Further indication of badly measured tracks could be observed as tails in the distribution of the impact parameter distribution d_0 (figure 6.14). These are not observed.

Final scale uncertainties are determined by adding in quadrature the uncertainty on the ID tracking with the deviation from unity from the double ratios. The resulting scale uncertainties on p_T, η, $mass$, $\sqrt{d_{12}}$, $\sqrt{d_{23}}$, τ_{21} and τ_{32} are 3-6%. At very high p_T statistical effects on the double ratio tend to dominate the systematic uncertainties on the final measurement.
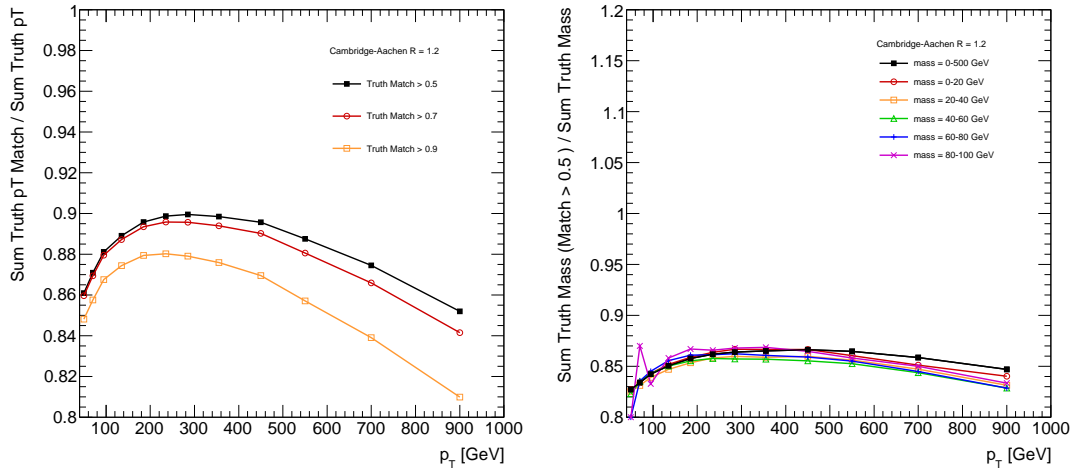
Figure 6.12: R_{match} for p_T (left) using three different values of P_{match} and mass (right) with $P_{match} = 0.5$ for anti-k_T R = 1 jets.

6.11.2   Constraining the Scale and Resolution using the W peak

One of the primary motivations for sub-structure techniques is to discover new particle resonances. However, this method can also be used to reconstruct the mass of known particles, such as the Z or W bosons. By measuring the difference in the mean shift and difference in width of the distribution between data and Monte-Carlo for a known mass resonance an independent constraint can be placed on the JMS and JMR for sub-structure jets in ATLAS. Figure 6.15 shows a mass distribution for events consistent with $WH \to \ell\nu b\bar{b}$ [47]. Events in the initial study were split and then filtered using the method applied in the Monte-Carlo substructure studies performed in Chapter 2. Events are selected with a W consistent with having $p_T > 200$ GeV for split and filtered Cambridge-Aachen R = 1.2 jets with $p_T > 180$ GeV. For this thesis the plot is then fitted using the combination of a 5th order polynomial and Gaussian, as shown in Figure 6.16. These represent the first study in ATLAS using a known particle resonance to constrain the scale and resolution for sub-jets. The difference in mean and width in Monte-Carlo to data found a 4% uncertainty in jet mass scale, consistent with the results found double ratio method and a resolution uncertainty of 20% to 23%. These values are in excellent agreement with the values derived for the JMR and JMS by other methods in this analysis. A full treatment of the systematic uncertainties is beyond the scope of
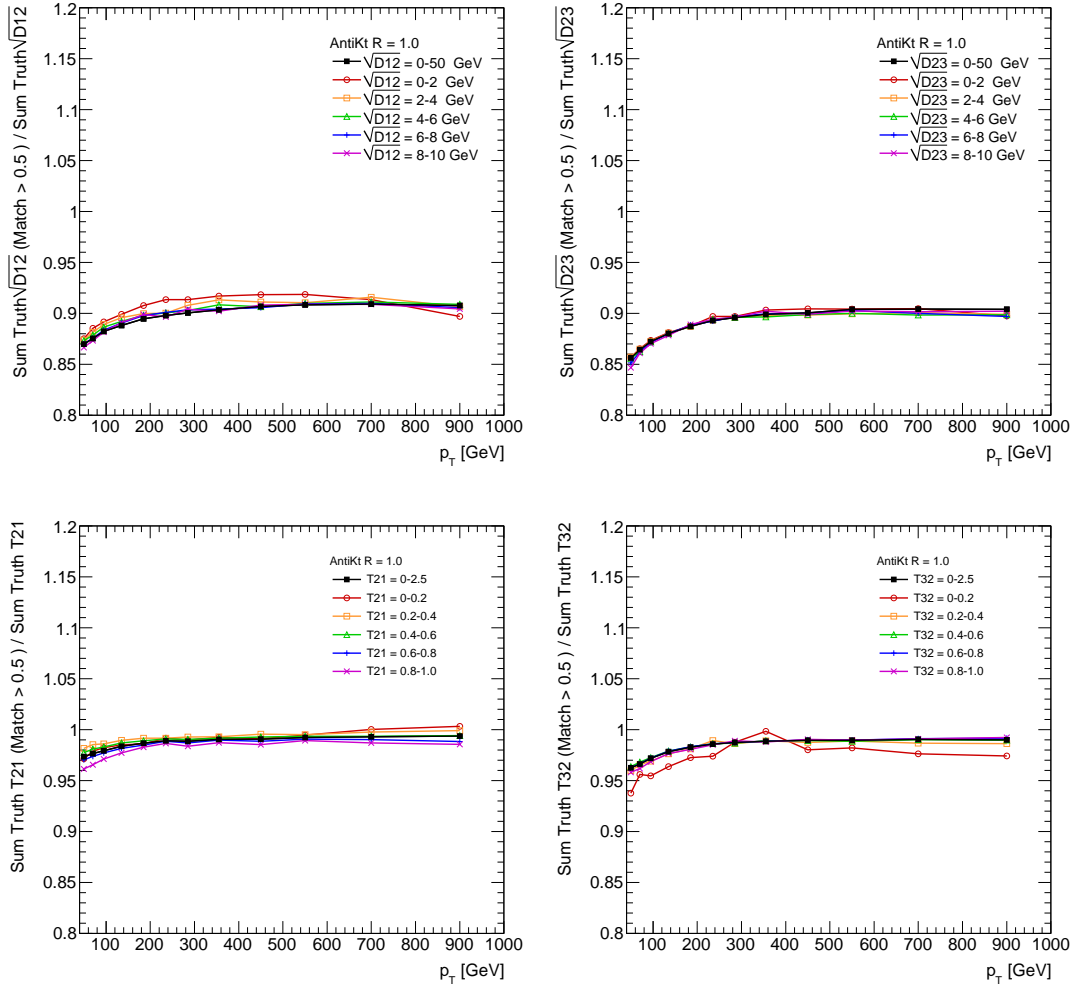
Figure 6.13: R_{match} for $\sqrt{d_{12}}$ (top left) and $\sqrt{d_{23}}$ (top right) and ratio of τ_{21} (left) and τ_{32} (right) with $P_{match} = 0.5$ for anti-k_T R = 1 jets.

this thesis. However, with a full treatment of background shape and model uncertainty this is potentially a very effective method to constrain systematic uncertainty on future sub-structure measurements in ATLAS and establishing the mass scale for a boosted Higgs resonance.

6.11.3   Unfolding

Final distributions are corrected for detector resolution and acceptance and compared to theoretical models. Iterative Dynamically Stabilised (IDS) unfolding is applied (Section 4.7) using a central value is provided by PYTHIA due to the better performance
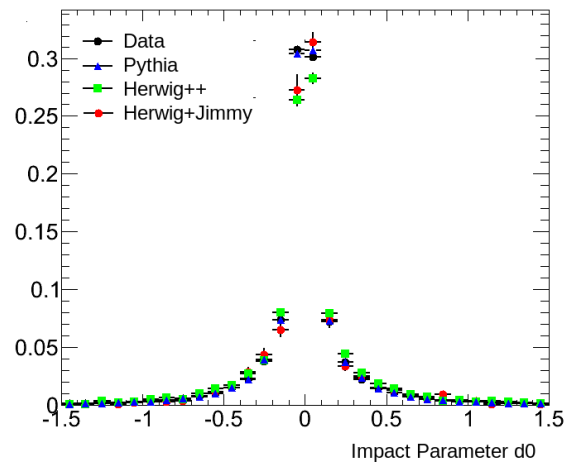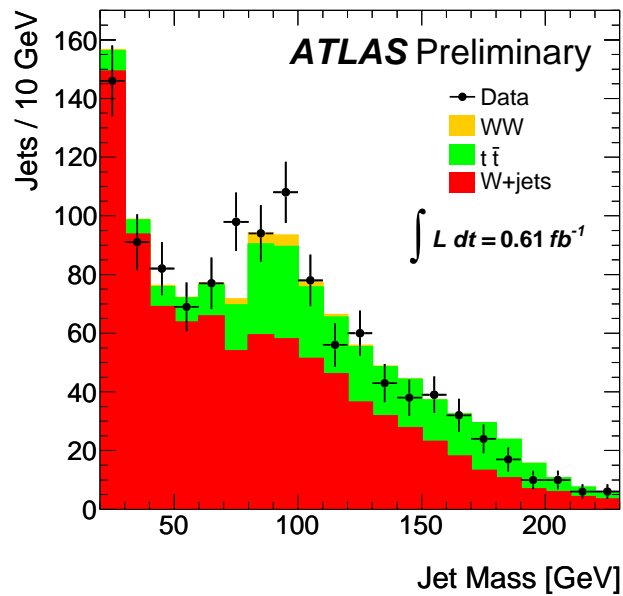
Figure 6.14: Distribution of impact parameter d_0.

Figure 6.15: The jet mass distribution of sub-jets with p_T 180 GeV in events consistent with a $WH \to \ell\nu b\bar{b}$ boson decay with $p_T > 200$ GeV. The distribution is compared to the uncorrected MC simulation prediction for $t\bar{t}$, W+jets and WW processes [47].
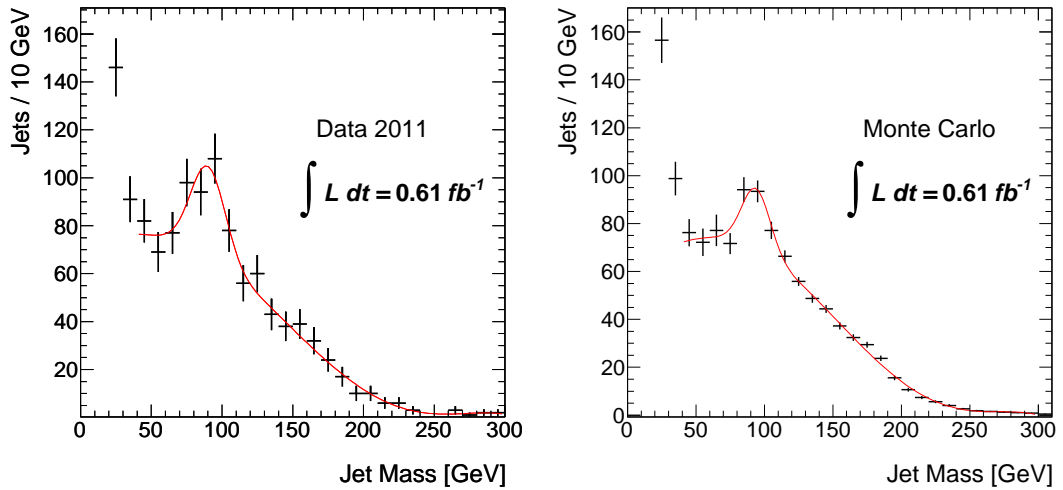
Figure 6.16: Fits are performed on the same sample as Figure 6.15 for data and Monte-Carlo using a 5th order polynomial and a Gaussian distribution.

in modelling the soft physics when compared to HERWIG, and the higher statistics available. The systematic uncertainty for the unfolding is assessed by repeating the procedure using different Monte-Carlo samples [39]. The cross-sections for filtered jet mass are shown in Figures 6.17 to 6.19. The sub-structure variables are shown for anti-k_T jets for Figures 6.20 to 6.23. For the ratio plots, the statistical uncertainty on Monte-Carlo predictions does not include the data statistical uncertainty. In general, the modelling by PYTHIA is softer than that seen in data. HERWIG++ tends to reproduce a harder mass spectrum. The jet mass is best described after splitting and filtering, and shows agreement between Monte-Carlo and data within statistical uncertainties. The sub-structure variables exhibit generally good agreement with both PYTHIA and HERWIG++.

6.12 Sub-Structure in QCD Jets Summary and Discussion

The high granularity of the ATLAS detector provides excellent resolution at small scales, which can be utilised to recover information about the underlying physics in a jet. In order use this technique in searches for new particle resonances, the impact of systematic uncertainties on sub-structure analysis must be assessed. This study is performed in the 2010 data, corresponding to an integrated luminosity of 35 pb^{-1}.

Jet mass, p_T, η and sub-structure variables are measured in QCD jets for Cambridge-Aachen R = 1.2 and anti-k_T R = 1.0. Data is subject to selection criteria designed to remove events for which the detector conditions are considered unreliable or those contaminated by detector noise and non-collision background. The systematic uncertainties on the calorimeter are assessed using jets made from the Inner Detector. The tracking systematic uncertainties were shown to be small (1% to 2%) and are folded into the final uncertainties. The results show a marked improvement for the modelling of jet mass for the jets after the splitting and filtering process has been applied. This result for QCD helps gives confidence in the sub-structure method, and established a method of deriving the uncertainties for sub-structure jets. This is essential if this method is to be used in future physics analyses in ATLAS.
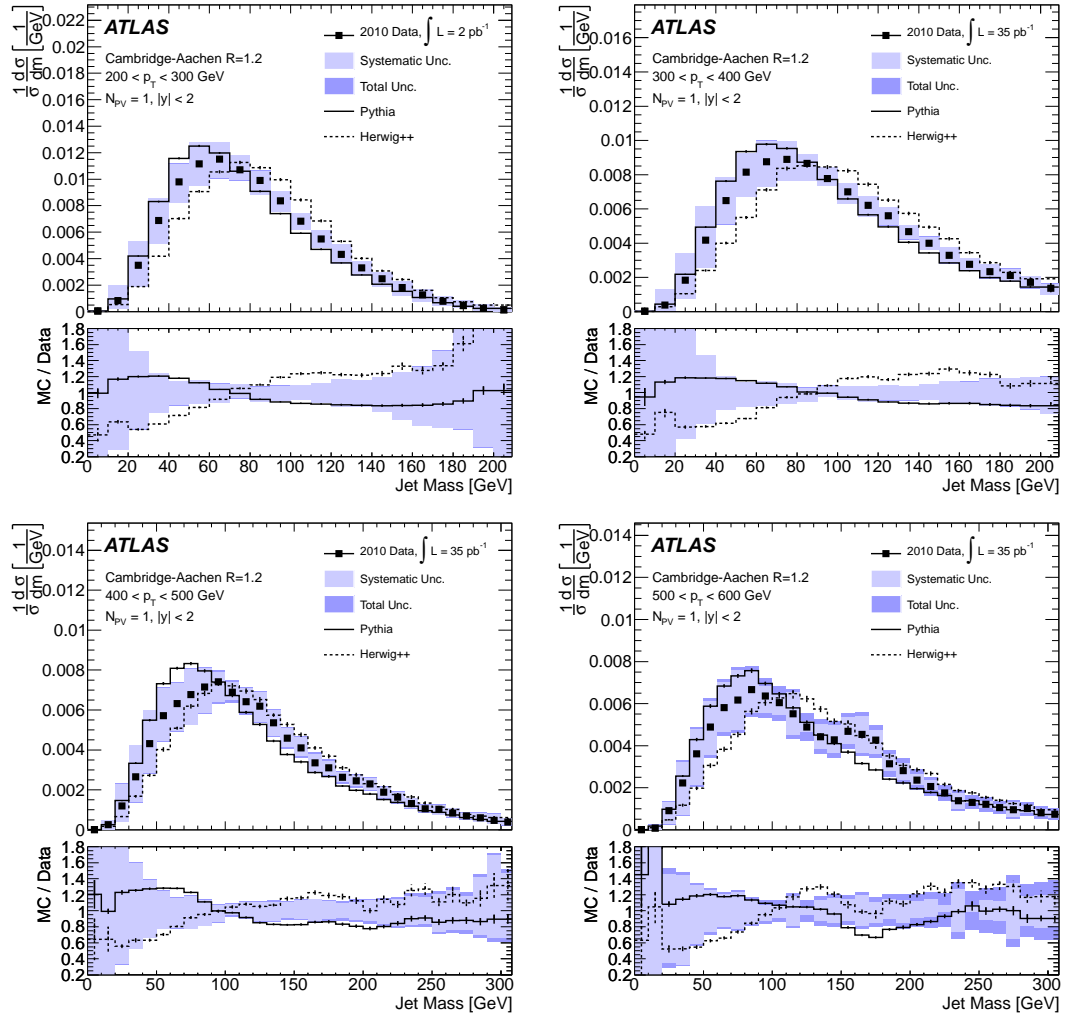
Figure 6.17: Normalised cross-sections as a function of mass of Cambridge-Aachen jets with $R = 1.2$ in four different p_T bins [39].
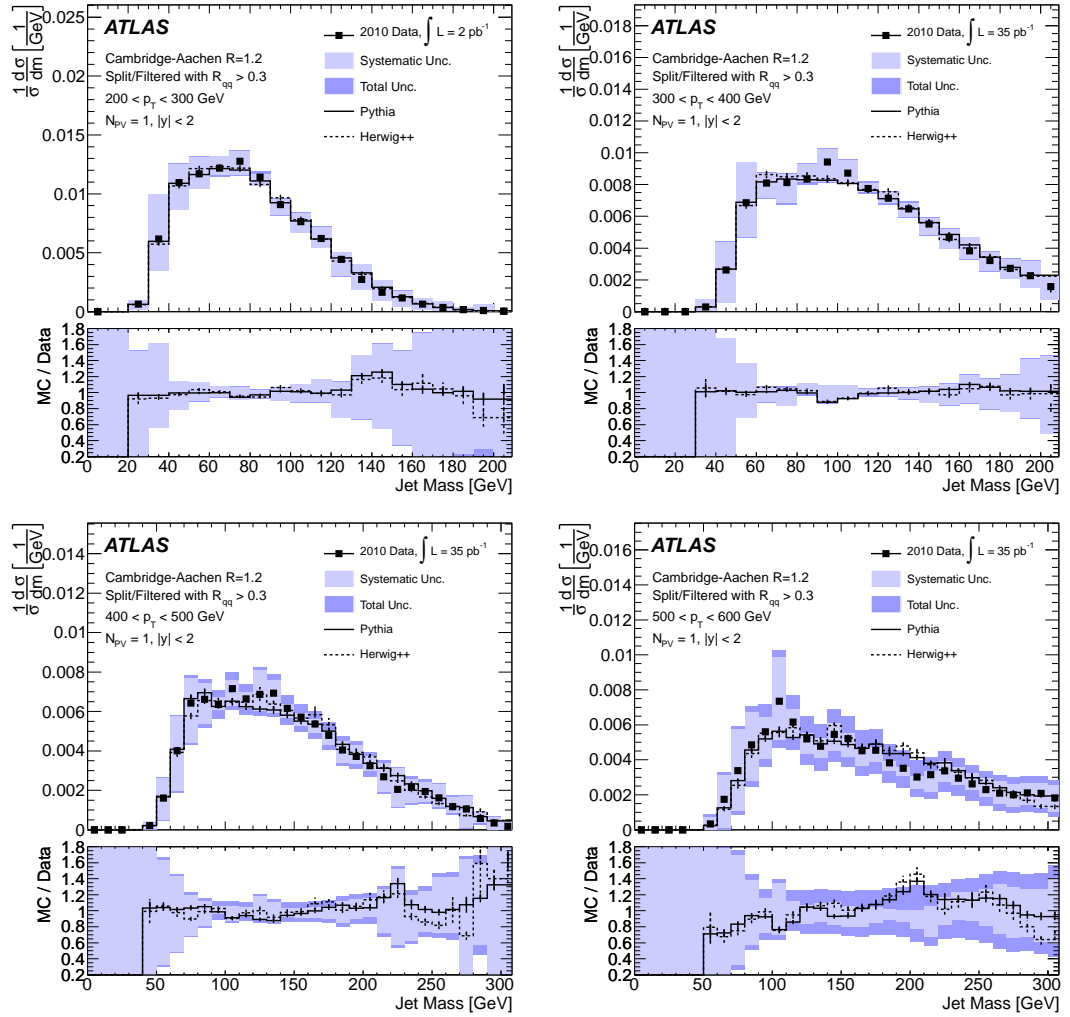
Figure 6.18: Normalised cross-sections as a function of mass of Cambridge-Aachen jets with $R = 1.2$ after splitting and filtering in four different p_T bins [39].
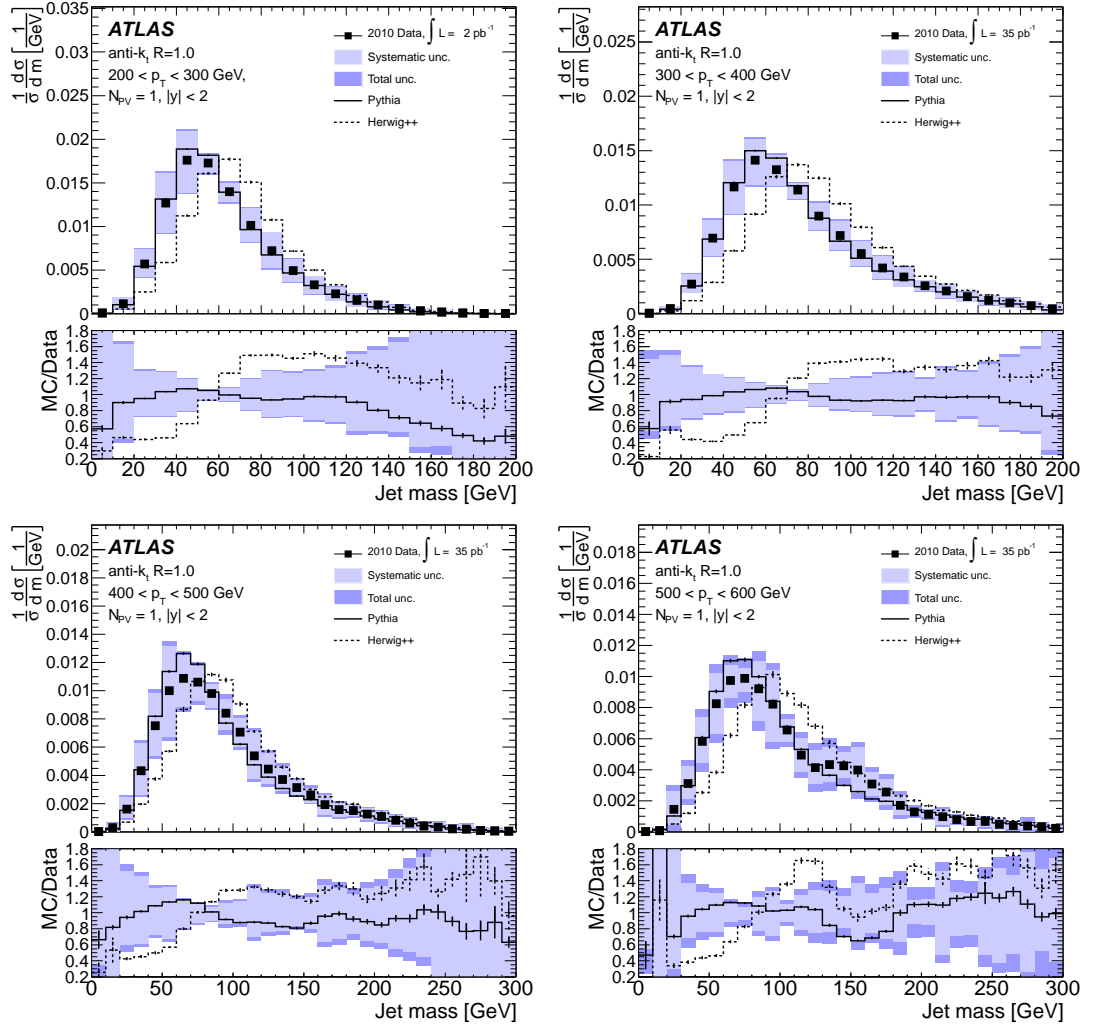
Figure 6.19: Normalised cross-sections as a function of mass of anti-k_T jets with $R = 1.0$ in four different p_T bins [39].

Figure 6.20: Normalised cross-sections as a function of $\sqrt{d_{12}}$ of anti-k$_T$ jets with $R = 1.0$ in four different p_T bins [39].
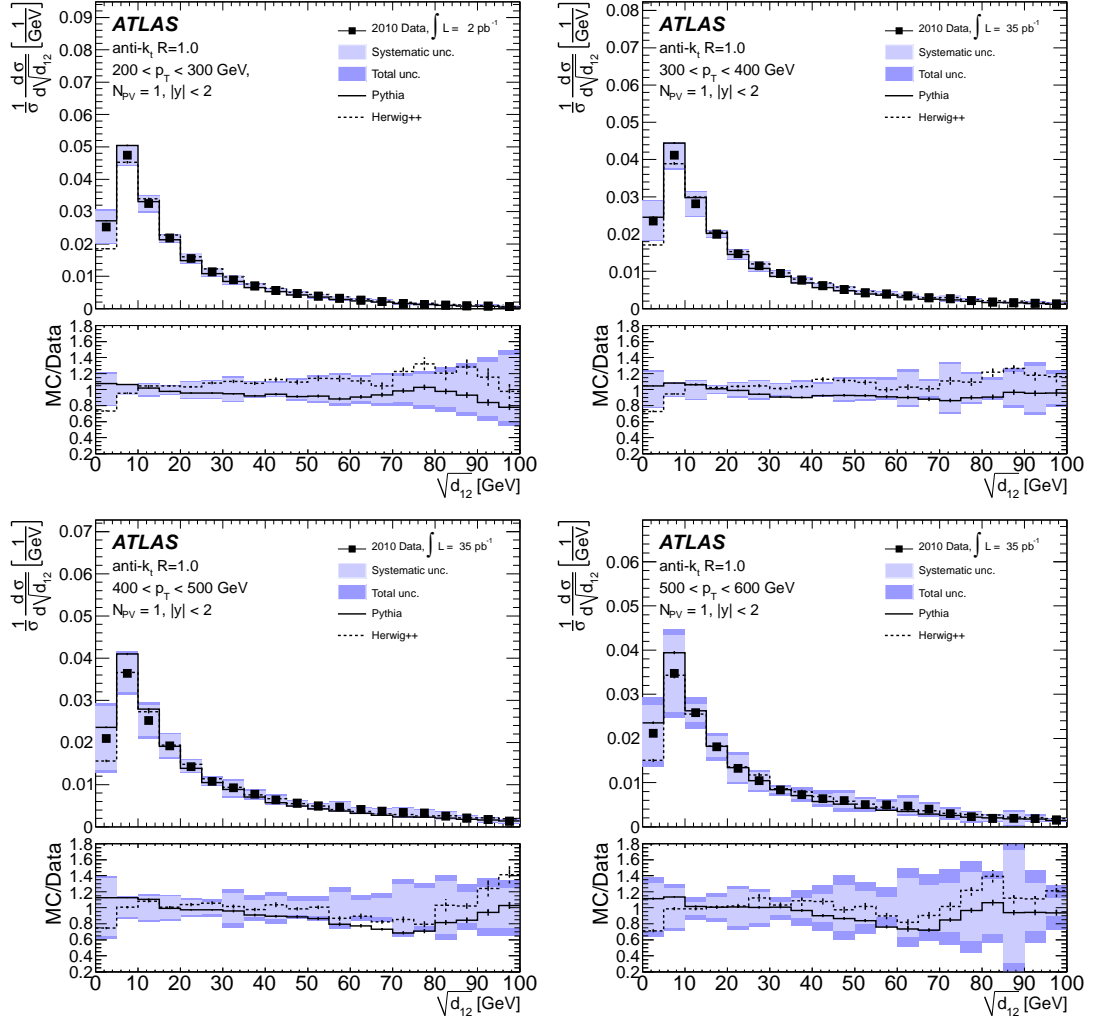
Figure 6.21: Normalised cross-sections as a function of $\sqrt{d_{23}}$ of anti-k_T jets with $R = 1.0$ in four different p_T bins [39].

Figure 6.22: Normalised cross-sections as a function of τ_{21} of anti-k$_T$ jets with $R = 1.0$ in four different p_T bins [39].
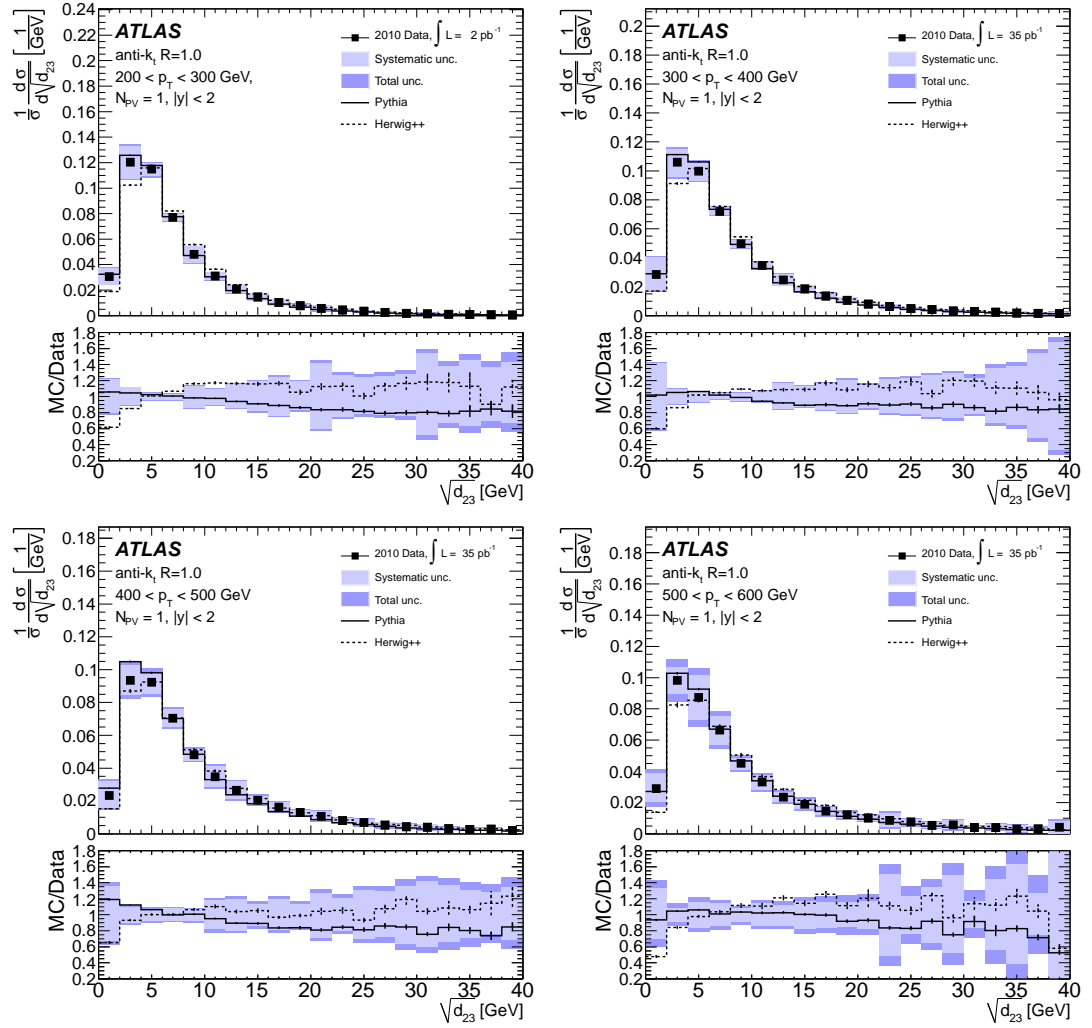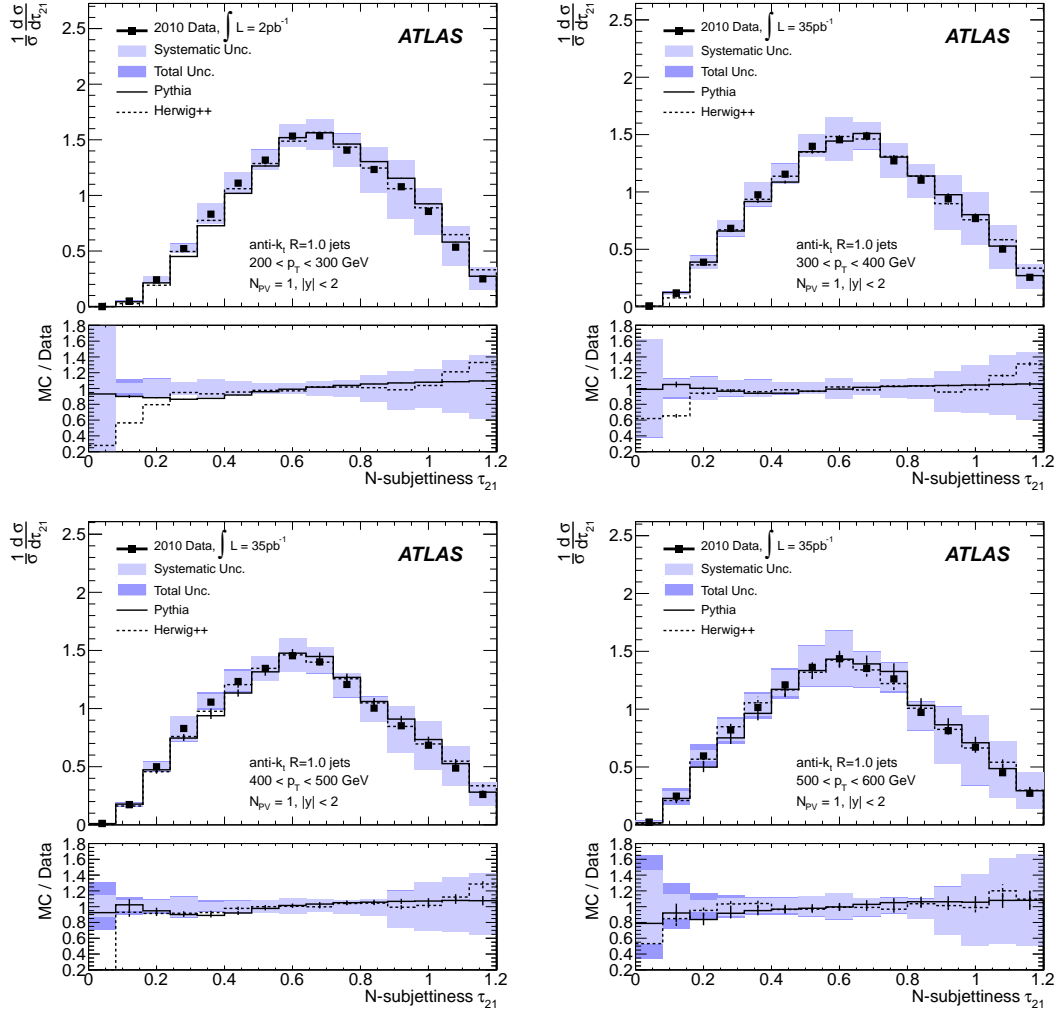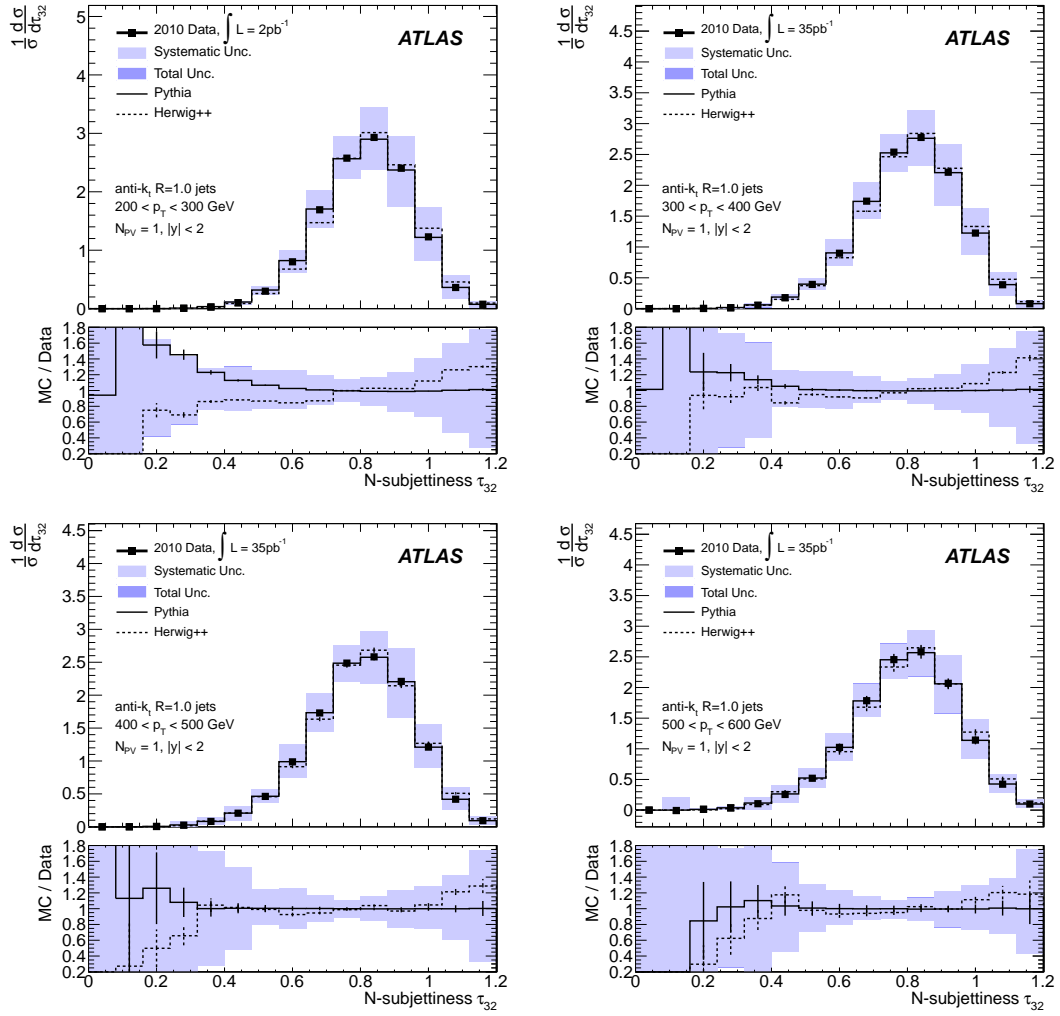
Figure 6.23: Normalised cross-sections as a function of τ_{32} of anti-k$_T$ jets with $R = 1.0$ in four different p_T bins [39].

Chapter 7

Conclusions

This thesis represents work towards the measurement of the systematic uncertainty for several key jet analyses performed in ATLAS, using jets reconstructed at a centre-of-mass energy of 7 TeV. The differential cross-section of inclusive jets are measured in the central and forward detectors, from 20 GeV to 1.5 TeV. This measurement spans two orders of magnitude in transverse momentum, beyond the energetic reach of any previous experiment. Background contributions from non-collision sources are measured. Jet sub-structure is studied in data, and jet mass measurements are performed in ATLAS for the first time. Mass scales are constrained for sub-structure jets using a comparison between the calorimeter and the tracking detector and by fitting the reconstructed W boson mass peak. Constraints are placed on the systematic uncertainty of the jets reconstructed using the tracking systems through a detailed study of the performance of track reconstruction at high p_T.

Bibliography

[1] Noether E, *Invariante Variationsprobleme*, Nachr. D. Knig. Gesellsch. D. Wiss. Zu Gttingen, Math-phys. Klasse (1918) 235257.

[2] The ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys.Lett. B **716** (2012) 1-29.

[3] The CMS Collaboration,*Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett. B **716**(2012) 306.

[4] M. Gell-Mann and Y. Ne'eman, *The eightfold way: with new contributions from the authors*, Advanced book classics. Perseus Pub, (2000).

[5] R. Cahn and G. Goldhaber, *The experimental foundations of particle physics*, Cambridge University Press, (2009).

[6] R. Feynman, *Photon-hadron interactions*, Advanced book classics. Addison-Wesley, (1998).

[7] H1 and ZEUS Collab., F.D. Aaron et al., Published in JHEP01 (2010) 109.

[8] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nuclear Physics B **126** (1977) 2, 298 318.

[9] E. Jensen, *Supersymmetry search using Z^0 bosons produced in neutralino decays at the ATLAS detector*, CERN-THESIS-2011-027.

[10] C. Buttar et al., *Standard Model Handles and Candles Working Group: Tools and Jets Summary Report*, arXiv:0803.0678 [hep-ph].

[11] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, *Longitudinally invariant k_t clustering algorithms for hadron-hadron collisions*, Nucl. Phys. B **406** (1993) 187224.

[12] M. Cacciari, G. Salam, and G Soyez, *The anti-kt jet clustering algorithm*, JHEP **0804** (2008) 063.

[13] Y. Dokshitzer, G. Leder, S. Moretti, and B. Webber, *Better jet clustering algorithms*, JHEP **08** (1997) 001.

[14] M. H. Seymour, *Searches for new particles using cone and cluster jet algorithms: A Comparative study*, Z. Phys. C **62** (1994) 127.

[15] J. M. Butterworth, B. E. Cox and J. R. Forshaw, *WW scattering at the LHC*, Phys. Rev. D **65** (2002) 096014.

[16] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, *Jet substructure as a new search channel for the Higgs at the LHC*, Phys. Rev. Lett. **100**, 242001 (2008).

[17] A. Buckley et al., *Rivet user manual*, arXiv:1003.0694 [hep-ph].

[18] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, JHEP 1103 015, (2011).

[19] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, and A. Salzburger *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, ATL-SOFT-PUB-2007-007, CERN, Geneva, Mar, 2007.

[20] G. Corcella et al., *HERWIG 6.5 release note*, arXiv:0210213 [hep-ph].

[21] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026.

[22] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001.

[23] T. Sjostrand and P. Z. Skands, *Transverse-momentum-ordered showers and interleaved multiple interactions*, Eur.Phys.J. **C39** (2005) 129154.

[24] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton fragmentation and string dynamics*, Phys. Rep. **97** (1983) 31145.

[25] A. Sherstnev and R. S. Thorne, *Parton distributions for LO generators* Eur. Phys. J. C **55** (2008) 553575.

[26] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. C **72** (1996) 637646.

[27] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012.

[28] M. Bahr, S. Gieseke, M. Gigg, D. Grellscheid, K. Hamilton, et al., *Herwig++ Physics and Manual*, Eur.Phys.J. **C58** (2008) 639707.

[29] M. Bahr, S. Gieseke, and M. H. Seymour, *Simulation of multiple partonic interactions in Herwig++*, JHEP **08 07** (2008) 076.

[30] ATLAS Collaboration, *First tuning of HERWIG/JIMMY to ATLAS data*, ATL-PHYS-PUB-2010-014, CERN, Geneva, October, 2010.

[31] ATLAS Collaboration, *Response of the ATLAS calorimeter to single isolated hadrons produced in proton proton collisions a centre-of-mass energy of $\sqrt{s} = 900GeV$*, ATLAS-CONF-2010-017, CERN, Geneva, March, 2010.

[32] Z. Nagy, *Next-to-leading order calculation of three jet observables in hadron hadron collision*, Phys. Rev. **D68** (2003) 094002.

[33] T. Carli et al., *A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project*, Eur. Phys. J. **C66** (2010) 503524.

[34] J. Allison et al., *Geant4 Developments and Applications* IEEE Transactions on Nuclear Science **53** (2006) 270-278.

[35] G. P. Salam, M. Cacciari, and G. Soyez http://www.lpthe.jussieu.fr/ salam/fastjet/

[36] ATLAS Collaboration, *Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s} = 7$ TeV*, Tech. Rep. ATLAS-CONF-2010-054, CERN, Geneva, 2010.

[37] A. Bazan, *The ATLAS liquid argon calorimeter read-out system*, IEEE Trans. Nucl. Sci. **53** (2006) 735740.

[38] ATLAS Collaboration, *Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at sqrt(s) = 7 TeV*, ATLAS Note ATLAS-CONF-2010-053 (2010).

[39] ATLAS Collaboration, *Jet mass and substructure of inclusive jets in sqrt(s) = 7 TeV pp collisions with the ATLAS experiment*, JHEP 1205 (2012) 128.

[40] ATLAS Collaboration, *Jet energy scale and its systematic uncertainty for jets produced in proton-proton collisions at $\sqrt{s} = 7$ TeV and measured with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2010-056, CERN, Geneva, 2010.

[41] ATLAS Collaboration, *Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC* Tech. Rep. CERN-PH-EP-2012-005, CERN, Geneva, 2012.

[42] ATLAS Collaboration, *Measurement of inclusive jet and dijet cross sections in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS detector*, Eur. Phys. J. **C71** (2011).

[43] J. M. Butterworth, J. R. Ellis, A. R. Raklev, and G. P. Salam, *Discovering baryon-number violating neutralino decays at the LHC*, Phys. Rev. Lett. **103** (2009) 241803.

[44] A. Abdessalam et al., *Boosted objects as a probe of beyond the standard model physics*, J. Phys. G. (2010).

[45] ATLAS Collaboration, *ATLAS jet energy scale uncertainties using tracks in proton proton collisions at sqrt(s) = 7 TeV* ATLAS-CONF-2011-067 (2011).

[46] ATLAS Collaboration, *Track Reconstruction Efficiency in sqrt(s) = 7 Data for tracks with $p_T > 100$ MeVs* ATL-COM-PHYS-2010-682 (2011).

[47] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a b-quark pair with the ATLAS detector at the LHC* ATLAS-CONF-2011-103 (2011).