Jet physics with the ATLAS experiment at the LHC

James Edward Martyn Robinson
University College London

Submitted to University College London in fulfilment
of the requirements for the award of the
degree of **Doctor of Philosophy**

March 28th, 2012

Declaration

I, James Edward Martyn Robinson 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.

James Robinson _____

Abstract

A variety of jet measurements are made using data collected during the first year of 7 TeV proton-proton collisions from the general-purpose ATLAS experiment at the LHC. The data are compared to leading-order and next-to-leading order Monte Carlo predictions, which have been interfaced with a parton shower, as well as to next-to-leading order perturbative QCD calculations, which have been corrected for soft effects. In each case, state-of-the-art jet algorithms are used, allowing for a better comparison between data and theory.

Two distinct types of analysis are presented in this thesis, measurements of jet cross-sections and investigations of QCD emissions in dijet systems. Double differential jet cross-sections, as a function of jet transverse momentum and rapidity or dijet mass and rapidity separation, provide an exacting test of QCD across several orders of magnitude. The study of QCD radiation in dijet systems is performed by vetoing on any QCD activity above a veto scale, Q_0 much greater than Λ_{QCD}, permitting the study of a wide range of perturbative QCD phenomena. Less well understood areas of phase space, in which standard event generators have large theoretical uncertainties, can be probed in both widely separated dijet systems and in those with average transverse momenta much greater than the veto scale.

Acknowledgements

Contemporary particle physics experiments are, in general, conducted by large collaborations. Many thousands of people work on the experiments based at the LHC and, as such, it is necessary to make clear which parts of this thesis are my own work.

Over the last century, a great deal of research has been performed into the fundamental structure of matter, much of which is condensed into the Standard Model of particle physics; some of the relevant parts of this theoretical framework are briefly discussed in Chapter 1. Equally, the ATLAS and LHC collaborations have worked over the last twenty years to design and build the physical machinery without which the measurements described in this thesis could not have been made. A short description of the layout and workings of some parts of the ATLAS detector is made in Chapter 2 although, by necessity, this cannot do justice to the huge complexity of this machine. The ATLAS collaboration also provides a data transfer system and software framework, within which context all of the following results were obtained.

The contents of Chapter 4 are strongly based on technical work, performed in collaboration with Dag Gillberg and Andrew Pilkington among others, which was used as one of the inputs to the determination of the ATLAS jet energy scale uncertainty. A novel soft radiation correction is also presented here; accordingly, all of the plots and numbers here are my own work and, although consistent with, are not identical to the official results.

In Chapter 5, my contribution was primarily to the determination of the cross-section in the forward region, $2.8 \leq |y| < 4.4$, particularly in developing the trigger strategy used here, together with Dag Gillberg and Justin Keung. In Chapter 6, I developed and validated the two-jet trigger strategy, in conjunction with Eric Feng and Christopher Meyer, based on an idea from Mario Campanelli. All the final cross-sections shown in both of these chapters represent official ATLAS results.

The contents of Chapter 7 and Chapter 8 represent an amalgam of published and ongoing research into QCD radiation phenomena. In the first instance, my contribution was primarily in determining background contributions and the stability of the final distributions under various changes to the selection criteria. Final plots shown here are again, official ATLAS results. In the later work, I transferred and revalidated the two-jet trigger strategy for a new set of variables, while also providing the bin-by-bin unfolded data. All plots shown in this chapter represent my own work. In both cases, I am heavily indebted to the other contributors to these analyses, including Andrew Pilkington, Pauline Bernat and Gareth Brown.

The work presented in this thesis would not have been possible without the assistance and encouragement of the High Energy Physics group at University College London, particularly Adam Davison and James Monk, who have shown a lot of patience despite my frequent physics and technical questions. Obviously, the ATLAS Collaboration, and the Jet and Missing-E_{T} group in particular, have been instrumental in allowing me to complete this work; I owe particular thanks to Andrew Pilkington and Dag Gillberg in this regard. It should, of course, be understood that all of the work in this thesis was performed with the help and guidance of my supervisor, Mario Campanelli.

Finally, thanks are due to my parents and my sister for the love and support that they have given me, not just over the past three years but throughout my life.

Preface

The Large Hadron Collider (LHC) is the largest and highest energy particle accelerator in the world, designed to collide protons with an unprecedented centre-of-mass energy of 14 TeV and instantaneous luminosity of 10^{34} cm^{-2} s^{-1}. In addition, the LHC has a heavy ion collision programme, aiming to collide lead nuclei with a centre-of-mass energy of 5.5 TeV. In the early phase of operation, the proton-proton programme at the LHC has been operating with reduced centre-of-mass energies of up to 7 TeV; these nevertheless represent the highest energy collisions that have yet been attained in a particle accelerator.

The accelerator was built by the European Organisation for Nuclear Research (CERN) and is situated in a 27 km long circular tunnel spanning the Swiss-French border near Geneva. A superconducting helium-cooled dipole magnet system, operating at 8.3 T is used to guide protons or lead nuclei around this ring.

The LHC was built with the aim of testing the current understanding of high energy physics, with its programme ranging from precise measurements of Standard Model parameters through to searches for new physics phenomena and investigations of the properties of strongly interacting matter at extreme energy densities. ATLAS (A Toroidal LHC ApparatuS) is, together with CMS (Compact Muon Solenoid), one of two general purpose detector experiments which have been collecting data at the LHC. Ideally suited to explore the TeV energy domain, ATLAS will play an important role in the possible resolution of these fundamental questions in particle physics.

In this thesis, a number of separate analyses are presented, each aiming to probe our understanding of QCD in this new energy regime, with $\sqrt{s} = 7$ TeV. Differential cross-sections of inclusive jets and dijets are performed across two orders of magnitude in jet transverse momentum and dijet mass and are compared to next-to-leading

order theoretical predictions. Radiation between dijets is examined as a possible means of discriminating between DGLAP and BFKL-like parton evolution schemes.

Finally, a more technical contribution to this thesis is a technique for ascertaining the uncertainty on the jet energy scale through intercalibration between jets in different regions of the detector.

Contents

List of Figures

List of Tables

"God does not play dice with the universe; He plays an ineffable game of His own devising, which might be compared, from the perspective of any of the other players [i.e. everybody], to being involved in an obscure and complex variant of poker in a pitch-dark room, with blank cards, for infinite stakes, with a Dealer who won't tell you the rules, and who smiles all the time."

— Terry Pratchett, *Good Omens*

Chapter 1.

Theoretical Framework

"I am now convinced that theoretical physics is actually philosophy."
— Max Born

1.1. Quantum Chromodynamics

Quantum chromodynamics (QCD) is the non-Abelian $SU(3)$ gauge theory of the strong interaction. Initially, it appears similar to QED, with the single electric charge replaced by three conserved "colour" charges; there are, however, important differences between the two.

In the QED Lagrangian, Equation (1.1), the electron carries one unit of charge, $-e$, the positron carries one unit of anti-charge $+e$ and the force is mediated by a massless photon.

$$\mathcal{L}_{QED} = \bar{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi - e\bar{\psi}\gamma^{\mu}A_{\mu}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.1}$$

where for a photon field A_{μ},

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

Similarly, in the QCD Lagrangian, Equation (1.2), the quarks carry colour charge, r, g, b, anti-quarks carry anti-charge, $\bar{r}, \bar{g}, \bar{b}$ and the force is mediated by massless gluons. As this is an exact $SU(3)$ symmetry, the strong interaction is invariant under rotations in colour space.

$$\mathcal{L}_{QCD} = \bar{\psi}_{q,a} \left(i\gamma^{\mu} \partial_{\mu} \delta_{ab} - m\delta_{ab} \right) \psi_{q,b} - g_s G^A_{\mu} \left(\bar{\psi}_{q,a} \gamma^{\mu} T^{ab}_A \psi_{q,b} \right) - \frac{1}{4} F^A_{\mu\nu} F^{\mu\nu}_A \tag{1.2}$$

where for a gluon field G^A_{μ},

$$F^A_{\mu\nu} = \partial_{\mu} G^A_{\nu} - \partial_{\nu} G^A_{\mu} - g_s f_{ABC} G^B_{\mu} G^C_{\nu}$$

In Equation (1.2), the γ^{μ} are the Dirac γ-matrices and repeated indices are summed over, following the Einstein summation convention. The ψ are quark-field spinors for a quark of flavour q and mass m_q, with a colour-index a that runs from $a = 1$ to $N_c = 3$; quarks form the fundamental representation of the $SU(3)$ colour group. The G^A correspond to the gluon fields, with A running from 1 to $N_c^2 - 1 = 8$; there are eight types of gluon, forming the adjoint representation of the $SU(3)$ colour group. The T^{ab}_A are eight 3×3 matrices, which form the generators of the $SU(3)$ group, corresponding to the fact that the interaction between a gluon and a quark rotates the quark in colour space. Finally, g_s is the QCD coupling constant and $F^A_{\mu\nu}$ the field tensor, where f_{ABC} are the structure functions of the $SU(3)$.

In QED, photons are electrically neutral and hence do not carry the charge of the EM interaction. In contrast, gluons do carry colour charge, one unit of colour and one of anti-colour, so gluon self-interactions are to be expected. Considering only the quark and gluon interaction terms the following simplified model of the QCD Lagrangian can be obtained:

$$\mathcal{L}_{QCD} = \text{“}q\bar{q}\text{”} + \text{“}G^2\text{”} + g_s \text{“}q\bar{q}G\text{”} + g_s \text{“}G^3\text{”} + {g_s}^2 \text{“}G^4\text{”} \tag{1.3}$$

Here, the q correspond to quarks and the G to gluons, with each term representing an allowed interaction or propagator within QCD; the G^3 and G^4 terms represent

triple and quartic gluon vertices. The non-abelian $G^A_{\mu\nu}$ term in the Lagrangian, which gives rise to these terms, is thus the major reason for the greater complexity of QCD in comparison to QED. As a result of gluon self-interactions, a QCD process contains more diagrams at a given order than would be true for QED.

1.2. Asymptotic Freedom

The interactions permitted by the QCD Lagrangian mean that a quark or a gluon can emit a gluon, while gluons can make quark or gluon loops. Analogously to QED, these interactions are governed by a coupling constant, with $\alpha_{\mathrm{S}} = \frac{{g_s}^2}{4\pi}$.

Just as is the case with electric charge screening in QED, the sum of the infinite series of higher order diagrams is equivalent to a single diagram with a "running" coupling constant. However, in contrast to QED, α_{S} *decreases* with increasing Q^2, as can be seen in Equation (1.4), resulting in "anti-screening" of colour charge.

$$\alpha_{\mathrm{S}}\left(Q^2\right) \simeq \frac{\alpha_{\mathrm{S}}\left(Q_0^2\right)}{1 + B\alpha_{\mathrm{S}}\left(Q_0^2\right)\ln\left(\frac{Q^2}{Q_0^2}\right)} \tag{1.4}$$

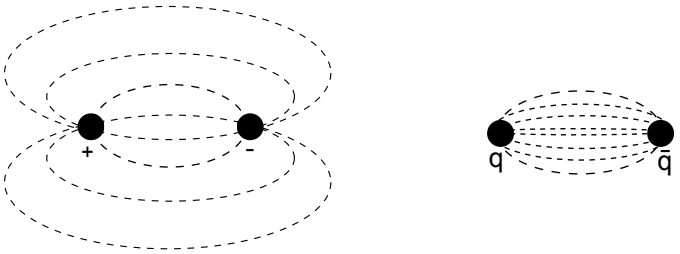
where for N_c colours and N_f quark flavours, the constant, B, is given by

$$B = \frac{11N_c - 2N_f}{12\pi}$$

At high Q^2, or equivalently at small distances, the effect of these higher order interactions can be conceptualised as spreading out the colour charge of an object into a "colour cloud". A test charge inside the colour cloud will experience a smaller force than if it were a large distance away. At small distances, therefore, quarks interact through colour fields of reduced strength and asymptotically behave like free particles. This phenomenon is known as asymptotic freedom and allows perturbation theory to be used in this region, permitting high precision tests, similar to those possible in QED [11].

1.3. Colour Confinement

It is believed, although not yet proven, that all observable free particles must be colourless. This is known as the colour confinement hypothesis: only colour singlet states can exist as free particles. Consequently quarks are never observed in isolation but only in bound states as colourless hadrons. Any energy injected into a hadron does not separate the quarks, but goes into creating $q\overline{q}$ pairs and hence further hadrons.



(a) Field lines arising from a QED dipole

(b) Field lines arising from a QCD dipole

Figure 1.1.: A comparison of the fields between QED (electric) and QCD (colour) dipoles: (a) the electric field surrounding two opposite electric charges and (b) the colour field surrounding a $q\overline{q}$ pair are shown. In comparison to the QED case, the lines of force in QCD are compressed. As the $q\overline{q}$ separation, r, is increased, the cross-sectional area of the QCD "flux tube" remains constant.

Gluon self-interactions provide a plausible means by which colour confinement could arise. When two coloured objects separate, the exchange of virtual gluons "squeezes" the lines of force between them so that they are closer together than in the QED case and a "flux tube" of interacting gluons is formed (see Figure 1.1). As the distance, r, between them increases, the cross-section area of this tube remains approximately constant. Since the number of field lines depends only on the colour of the sources, the field strength in the tube must also remain constant, while the field energy grows in proportion to the volume of the tube. This means that the energy of the $q\overline{q}$ system increases linearly with separation and hence that it would require an infinite amount of energy to fully separate two coloured objects [12].

1.4. Evolution Equations

1.4.1. Deep Inelastic Scattering and the Quark-Parton Model

The process in which the lepton from an $e^{\pm}$p collision interacts with a quark from the proton is known as deep inelastic scattering (DIS). If the interacting quark carries a fraction, x of the proton's 4-momentum, p, then the cross-section for this process is given by

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}x\,\mathrm{d}Q^2} = \frac{4\pi\alpha^2}{Q^4}\left[\left(1-y-\frac{m_p^2y^2}{Q^2}\right)\frac{F_2\left(x,Q^2\right)}{x}+y^2F_1\left(x,Q^2\right)\right] \tag{1.5}$$

or, in the high energy limit, where $Q^2 \gg m_p^2y^2$

$$= \frac{4\pi\alpha^2}{Q^4}\left[(1-y)\frac{F_2\left(x,Q^2\right)}{x}+y^2F_1\left(x,Q^2\right)\right] \tag{1.6}$$

where $Q^2 = -q^2$ is related to the momentum transfer between the quark and the electron, $y = \frac{p_p \cdot q}{p_p \cdot p_e}$ is the fraction of the electron's energy transferred to the proton in the proton's rest frame, m_p is the proton mass and $F_1\left(x,Q^2\right)$ and $F_2\left(x,Q^2\right)$ are, respectively, the pure magnetic and the electromagnetic proton structure functions, which describe the momentum distribution of the quarks within the proton.

In 1969, Bjorken proposed that the $F_1\left(x,Q^2\right)$ and $F_2\left(x,Q^2\right)$ structure functions should exhibit scaling behaviour in the deep inelastic limit, $Q^2 \to \infty$ while x stays finite [13]. Initially, this prediction, that the structure functions should be approximately independent of Q^2, was experimentally verified; more detailed investigation eventually showed the choice of x with which the measurement was made to be serendipitous: at low x, $F_2\left(x,Q^2\right)$ rises with increasing Q^2, while at high x, $F_2\left(x,Q^2\right)$ falls with increasing Q^2. It is nevertheless true that the structure functions depend more strongly on x, a kinematic quantity, than on Q^2, related to the energy of the collision. In addition, it is observed that $F_1\left(x,Q^2\right)$ and $F_2\left(x,Q^2\right)$ are not independent, but satisfy the Callan-Gross relation:

$$F_2(x) = 2xF_1(x) \tag{1.7}$$

By analogy with $e^- \mu^- \to e^- \mu^-$ scattering, the differential cross-section for elastic scattering between an electron and a quark with electric charge, $q_i = \pm 2/3, \pm 1/3$, and carrying a fraction, x, of the proton's 4-momentum, can be expressed as

$$\begin{aligned} \frac{\mathrm{d}\sigma}{\mathrm{d}Q^2} &= \frac{2\pi\alpha^2 q_i^2}{Q^4}\left[1 + (1-y)^2\right] \\ &= \frac{4\pi\alpha^2 q_i^2}{Q^4}\left[(1-y) + \frac{y^2}{2}\right] \end{aligned} \tag{1.8}$$

In the proton, the fraction of quarks having x in the range between x_0 and $x_0 + \delta x$ is $f_q(x_0, Q^2)\delta x$ where f_q is the parton distribution function (PDF) for the quark concerned. Taking the distribution of quark momenta in this way, the cross-section for scattering from a particular quark type within the proton is

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}Q^2}\right|_{x \to x+\delta x} = \frac{4\pi\alpha^2}{Q^4}\left[(1-y) + \frac{y^2}{2}\right] q_i^2 f_{q_i}(x, Q^2)\delta x \tag{1.9}$$

summing over all quarks and anti-quarks in the proton, since the gauge boson interacts equally with both, gives the expression

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}Q^2\,\mathrm{d}x} = \frac{4\pi\alpha^2}{Q^4}\left[(1-y) + \frac{y^2}{2}\right] \sum_i q_i^2 f_{q_i}(x, Q^2) \tag{1.10}$$

Comparing Equation (1.10) and Equation (1.6) yields the parton model prediction for the proton structure functions:

$$F_2\left(x,Q^2\right) = 2xF_1\left(x,Q^2\right) = \sum_i \overbrace{xf_{q_i}(x,Q^2)}^{quark} + \overbrace{x\overline{f_{q_i}}(x,Q^2)}^{anti-quark} \tag{1.11}$$

In the proton, quarks, anti-quarks and gluons each have non-zero PDFs. At present the f_q cannot be analytically calculated within QCD: perturbation theory cannot be used due to the large coupling constant. The evolution of the PDFs with x and Q^2, however, can be determined and tested experimentally.

1.4.2. QCD Compton Scattering and Boson-Gluon Fusion

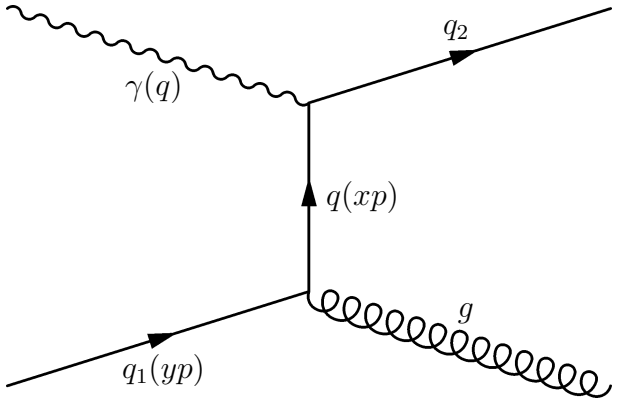


Figure 1.2.: QCD Compton scattering ($\gamma\, q \to \mathrm{g}\, q$).

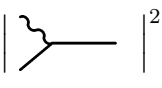
In the QCD Compton process (Figure 1.2), it can be shown that:

$$\hat{\sigma}(\gamma q \to gq) \;=\; \int_{\mu^2}^{\hat{s}/4} \mathrm{d}p_{\mathrm{T}}^2 \,\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^2} \;=\; e_i^2\sigma_0\frac{\alpha_{\mathrm{S}}}{2\pi}\int_{\mu^2}^{\hat{s}/4}\frac{\mathrm{d}p_{\mathrm{T}}^2}{p_{\mathrm{T}}^2}\,P_{qq}(z) \tag{1.12}$$

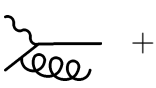
where $\sigma_0 = \frac{4\pi^2\alpha}{\hat{s}}$ is the $\gamma^* p$ total cross-section, p_{T} is the transverse momentum of the outgoing quark (see Section 1.8), μ^2 is an infrared cut-off to prevent divergences as $p_{\mathrm{T}}^2 \to 0$ and $P_{qq}(z) = \frac{4}{3}\left(\frac{1+z^2}{1-z}\right)$ is the probability of a quark emitting a gluon and so

becoming a quark with momentum reduced by a fraction z. Considering pure QED and leading order QCD diagrams gives:

$$\left.\frac{F_2\left(x,Q^2\right)}{x}\right|_{\gamma^* q\to q(g)} = \Big|$$



$$\Big|^2 + \Big|$$

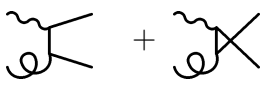


$$\Big|^2$$

$$= \sum_q e_q^2 \int_x^1 \frac{\mathrm{d}y}{y} f_q(y,Q^2)\left[\delta\left(1-\frac{x}{y}\right) + \frac{\alpha_\mathrm{S}}{2\pi} P_{qq}\left(\frac{x}{y}\right)\ln\left(\frac{Q^2}{\mu^2}\right)\right] \tag{1.13}$$

At this order in QCD, boson-gluon fusion must also be considered, with a similar expansion:

$$\left.\frac{F_2\left(x,Q^2\right)}{x}\right|_{\gamma^* g\to q\bar{q}} = \Big|$$



$$\Big|^2$$

$$= \sum_q e_q^2 \int_x^1 \frac{\mathrm{d}y}{y} f_g(y,Q^2)\frac{\alpha_\mathrm{S}}{2\pi} P_{gq}\left(\frac{x}{y}\right)\ln\left(\frac{Q^2}{\mu^2}\right) \tag{1.14}$$

where f_g is the gluon PDF in the proton and $P_{qg}(z) = \frac{1}{2}\left(z^2+(1-z)^2\right)$ is the probability that a gluon splits into a $q\bar{q}$ pair. Two further splitting functions are possible: $P_{gq}(z) = \frac{4}{3}\left(\frac{1+(1-z)^2}{z}\right)$ is the probability that a quark emits a gluon which carries away a fraction z of its momentum, and $P_{gg}(z) = 6\left(\frac{1-z}{z} + \frac{z}{1-z} + z(1-z)\right)$ is the probability that a gluon splits into two gluons; in each case these are unregulated probabilities, which contain singularities at $z=0$ or $z=1$. Combining these results allows equations governing the evolution of the structure functions to be obtained.

$$\frac{\mathrm{d}f_q(x,Q^2)}{\mathrm{d}\ln Q^2} = \frac{\alpha_\mathrm{S}}{2\pi}\int_x^1 \frac{\mathrm{d}y}{y}\left(f_q(y,Q^2)P_{qq}\left(\frac{x}{y}\right) + f_g(y,Q^2)P_{qg}\left(\frac{x}{y}\right)\right) \tag{1.15}$$

$$\frac{\mathrm{d}f_g(x,Q^2)}{\mathrm{d}\ln Q^2} = \frac{\alpha_\mathrm{S}}{2\pi}\int_x^1 \frac{\mathrm{d}y}{y}\left(\sum_i f_{q_i}(y,Q^2)P_{gq}\left(\frac{x}{y}\right) + f_g(y,Q^2)P_{gg}\left(\frac{x}{y}\right)\right) \tag{1.16}$$

Equations (1.15) and (1.16) comprise the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations of parton evolution [14–17]. As shown here they are leading order in α_S but they have long been used at next-to-leading order and have recently been calculated at next-to-next-to-leading order [18,19]. Although derived by considering deep inelastic scattering, these equations describe the partonic structure of the proton and are applicable more generally.

1.5. QCD Factorisation

At higher orders, the DGLAP equations become a summation over a series of integrals, each one similar in form to Equations 1.15 and 1.16. Within the context of DIS, initial state radiation means that the momentum of the parton at the point when it interacts with the photon can differ from its momentum when it was extracted from the proton. Mostly, such momentum modifying emissions are collinear with the parton and are often considered as altering the structure of the proton rather than forming part of the calculation of the parton-photon interaction. The separation between these two categories is defined using a factorisation scale, μ_F. Emissions with p_T above μ_F are included in the calculation, while emissions with lower p_T are accounted for in the proton PDFs. This process, known as QCD factorisation, separates the long-distance components (PDFs), which are universal, from the short-distance hard scattering, which is process dependent.

Additionally, due to the running of α_S as shown in Equation (1.4), a choice must be made about the value of Q at which α_S is evaluated. This value, known as the renormalisation scale, μ_R, is usually chosen to be of the order of a typical momentum transfer in the process being considered.

$$\sigma_{\text{jet}} = \sum_a \sum_b \overbrace{f_a\left(x_1, \mu_F^2\right) f_b\left(x_2, \mu_F^2\right)}^{PDFs} \overbrace{\hat{\sigma}_{a,b}\left(p_{\text{p1}}, p_{\text{p2}}, \alpha_S\left(\mu_R^2\right), \frac{Q^2}{\mu_F^2}, \frac{Q^2}{\mu_R^2}\right)}^{Hard\ scatter} \tag{1.17}$$

Equation (1.17) demonstrates how QCD factorisation can be applied to proton-proton collisions. Here the factorisation scale, μ_F, is used in the evolution of the

PDFs and fragmentation while the renormalisation scale, μ_R, is connected to the momentum transfer at which the integration of QCD equations stops. Q^2 is a hard scale that characterises the parton-parton interaction while f_a and f_b are the PDFs of the interacting protons: typically jet measurements use $\mu_F = \mu_R = Q$. Since the perturbation series is an asymptotic expansion, there is a limit to the precision with which any theoretical quantity can be calculated. Independently varying the renormalisation and factorisation scales away from their chosen values, usually by factors of two, provides an indication of the uncertainties on theoretical predictions.

1.6. Hadronisation and Jets

Perturbative QCD calculations may have coloured partons in the final state, but, due to colour confinement, only the colourless hadrons deriving from them are observed experimentally.

Consider a quark and anti-quark produced in, for example, an e^- e^+ annihilation. Initially, as the quarks separate, the interaction between them can be modelled by imagining a colour flux tube formed between them. As the separation increases, so does the energy stored in this flux tube and hence the probability of QCD radiation, which will be predominantly shallow-angled with respect to the originating parton. Thus, one parton can radiate gluons, which will in turn radiate $q\overline{q}$ pairs and so on, with each new parton nearly collinear with its parent. This process, known as parton showering, produces partons of successively lower energy until, at some point, the interactions exit the region in which perturbative QCD is valid.

Additionally, the coloured partons produced during this showering process must combine in bound states to form colourless hadrons. This is known as hadronisation and results in collimated sprays of colourless particles, known as jets; the first evidence of jets arising from quarks was obtained in e^+ $e^- \rightarrow q\overline{q}$ events at the SPEAR collider in 1975 [20]. Both hadronisation and the parton shower are inherently non-perturbative and must, therefore, be described using phenomenological models.

1.7. Jets at Hadron Colliders

Jets are unavoidable in hadron colliders, where the QCD $2 \rightarrow 2$ scattering of partons is the dominant hard process. This simple picture is complicated by higher order QCD interactions, particularly hard gluon radiation, as well as by soft QCD effects that have to be described using empirical models.

It is fundamentally impossible to examine the properties of partons: they are not physical objects but propagators, and their representation may vary according to, for instance, the Monte Carlo generator used. Jets, on the other hand, are well-defined objects that, although not the same as partons, provide us with a window through which to investigate them.

Through looking at jets, QCD can be studied by measuring parameters such as α_S or the top-quark mass. QCD calculations and phenomenological Monte Carlo models can be tested by measuring jet cross-sections, providing constraints for future PDF fits. The QCD evolution equations can be tested by looking at rapidity gaps in dijet systems, while jet substructure techniques provide a possible means of identifying new physics. Reconstructing decaying massive particles and constraining the structure of the proton both rely on accurate jet measurements.

1.8. Co-ordinates at Hadron Colliders

In a hadron collider the centre-of-mass frame of the hadrons is not usually the same as the centre-of-mass frame of the interacting partons. Energy and angular separations are not invariant under Lorentz boosts and, for a detector constructed in the hadronic centre-of-mass frame, particles will appear more collimated or dispersed depending on their boost.

It is therefore important, particularly when dealing with jets which are produced with a range of different boosts, to choose variables which are longitudinally Lorentz invariant with which to classify events. Rapidity, y, is a spatial coordinate describing the angle of a particle relative to the beam axis. At a hadron collider, particle production is roughly constant as a function of y.

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{1.18}$$

Pseudorapidity, η, is a well-defined function of the polar angle, θ which also provides a close approximation to rapidity; they are identical in the limit of massless particles.

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right) = \frac{1}{2} \ln \frac{p + p_z}{p - p_z} \tag{1.19}$$

For angles approximately perpendicular to the beam axis, $\theta = \pi/2 + \delta\theta$, it can be seen that distance in η is equivalent to distance in θ:

$$\begin{aligned} \Delta\eta &= -\ln \left(\tan \frac{\theta_1}{2} \right) + \ln \left(\tan \frac{\theta_2}{2} \right) \\ &= -\ln \left(1 + 2\frac{\delta\theta_1}{2} \right) + \ln \left(1 + 2\frac{\delta\theta_2}{2} \right) \\ &= -\ln \left(1 + \delta\theta_1 \right) + \ln \left(1 + \delta\theta_2 \right) \\ &= \delta\theta_2 - \delta\theta_1 = \Delta\theta \end{aligned} \tag{1.20}$$

Physical momenta are usually measured in terms of transverse momentum, p_{T}

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2} \tag{1.21}$$

Polar angle in the transverse plane, ϕ, is left unchanged by longitudinal boosts.

1.9. Luminosity at Hadron Colliders

The luminosity of a proton-proton collider can be expressed as

$$\mathcal{L} = \frac{R_{inel}}{\sigma_{inel}} \tag{1.22}$$

where R_{inel} is the rate at which inelastic collisions occur and σ_{inel} the corresponding cross-section. For a collider operating at a revolution frequency, f_r with n_b bunches crossing per revolution, this can be expressed as

$$\mathcal{L} = \frac{\mu n_b f_r}{\sigma_{inel}} = \frac{\mu^{vis} n_b f_r}{\epsilon \sigma_{inel}} = \frac{\mu^{vis} n_b f_r}{\sigma_{vis}} \tag{1.23}$$

where μ is the average number of inelastic interactions per bunch crossing, ϵ the efficiency for one inelastic collision to satisfy the event selection criteria for a visible process of interest and, for the chosen process, σ^{vis} and μ^{vis} are the cross-section and number of inelastic interactions per bunch crossing respectively.

The absolute luminosity, $\mathcal{L}$, can be inferred from measured accelerator parameters, using the equation

$$\mathcal{L} = \frac{n_b f_r n_1 n_2}{2\pi \Sigma_x \Sigma_y} \tag{1.24}$$

where n_1 and n_2 are the numbers of particles in the two colliding bunches, while Σ_x and Σ_y characterise the widths of the beam profile in the horizontal and vertical directions respectively. Σ_x and Σ_y are usually measured using van der Meer scans, in which the observed event rate is recorded while scanning the opposing beams across each other in the horizontal and vertical directions.

The luminosity determined using Equation (1.24), together with the measured value of μ_{vis} can then be used to calculate σ_{vis} using Equation (1.23), without relying on knowledge of the total inelastic cross-section or measurements of detector inefficiencies.

1.10. Jet Algorithms

To first order defining a jet is simple: it relies on identifying a coherent stream of particles originating at the interaction point. However, it is neither feasible nor repeatable to do this manually on an event-by-event basis: a well-defined jet algorithm is needed.

A jet algorithm is a fully specified set of rules for projecting information from a large number of hadron-like objects onto a small number of parton-like objects. These rules should work at all levels in order to allow fair and straightforward comparisons between data and theory. A jet algorithm should be able to take final-state particles from Monte Carlo, partons produced in a fixed order pQCD calculation or detector level objects such as calorimeter towers or tracks as input, while its output should be order independent and minimally dependent on hadronisation and detector effects that are often only understood empirically.

Jets need to be theoretically well-behaved, in other words, they must be invariant under minor modifications of their constituents. Specifically, adding a soft parton should not change the jet clustering results, a property known as infrared safety, while the identified jets should similarly remain unchanged if one parton is replaced by a collinear pair of partons, usually termed collinear safety. Without these requirements, real-virtual cancellations in next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) QCD calculations are not possible, producing divergent results. This leads to large uncertainties, removing much of the benefit obtained from calculating NLO quantities in the first place. From an experimental perspective, since detectors can resolve neither the full collinear nor the full infrared structure of an event, it is important that jet definitions are invariant in these limits.

The construction of a jet is unavoidably ambiguous in at least two ways. The first of these is the decision over which particles will be combined to form a jet: in other words the choice of which jet algorithm and associated parameters to use. The second is the question of how to recombine the momenta of these particles into a single four-vector to be assigned to the identified jet.

When taken together, these different elements determine the particular jets that will be identified in an event, although, in general, physical results, such as particle discovery, masses or couplings, should be independent of the choice of

jet definition. There are two main classes of jet algorithm: cone algorithms and sequential recombination algorithms.

1.10.1. Cone Algorithms

Cone algorithms have historically been important in hadron collider experiments, due to their apparent conceptual simplicity and fast execution time. The basic idea underlying algorithms of this type is to cluster objects together based on their proximity in $y - \phi$ space. Given a cone with centroid, C, and radius, R, a typical cone algorithm will take all objects i satisfying

$$\sqrt{\left(y^i - y^C\right)^2 + \left(\phi^i - \phi^C\right)^2} \leq R \tag{1.25}$$

and use them to determine the properties of the cone:

$$\begin{aligned} p^C = \left(E^C, \mathbf{p^C}\right) &= \sum_i \left(E^i, p_x^i, p_y^i, p_z^i\right), \\ \overline{y^C} = \frac{1}{2} \ln \frac{E^C + p_z^C}{E^C - p_z^C}, &\qquad \overline{\phi^C} = \tan^{-1}\left(\frac{p_y^C}{p_x^C}\right) \end{aligned} \tag{1.26}$$

If $y^C = \overline{y^C}$ and $\phi^C = \overline{\phi^C}$ then the sum of the momenta of all particles inside the cone points in the same direction as the cone centroid and the cone is identified as "stable". If the cone is not stable, the clustering step is repeated using the new centroid. This process is repeated iteratively until all identified cones are stable.

Cone algorithms of this sort usually rely on an initial step in which a subset of the available objects are identified as "seeds": the initial centroids used to start the clustering step. Usually, all objects above some p_{T} threshold are selected as seeds: something which is inherently infrared unsafe, as the particular distribution of soft objects or electronic noise will affect which objects are above this threshold.

Additionally, cone algorithms need a procedure to deal with the case in which two stable cones overlap: usually this comes in the form of a parameter which determines

whether to merge the two cones or to split them along a plane perpendicular to the line joining their centroids. This split/merge parameter is commonly expressed as a function of the percentage of p_{T} overlap between the constituents of the two cones.

Except in cases of overlap, cone algorithms produce regular, circular jets. This regularity of shape facilitates the comparison and calibration of jets. Unfortunately, with the exception of the SISCone algorithm [21], cone algorithms are, in general, infrared and collinear unsafe, making them a poor choice for comparisons between data and theory.

1.10.2. Recombination Algorithms

As an approximation, the development of a jet can be thought of as a consequence of repeated $1 \rightarrow 2$ branching of quarks and gluons within QCD. Sequential recombination algorithms aim to work their way backwards through these branches, repeatedly combining pairs of particles into a single one.

Clearly, an algorithm using this approach must define which pair of particles to combine at each step and how to determine when the end of the process has been reached. Obviously at each step, the best particles to combine are those which are in some way "closest" to one another; a metric to determine the distance between each pair of particles must therefore be defined:

$$d_{ij}^2 = \min\left(p_{\mathrm{T},i}^{2p},\, p_{\mathrm{T},j}^{2p}\right)\frac{\Delta R_{ij}^2}{R^2}, \qquad d_{iB} = p_{\mathrm{T},i}^{2p} \tag{1.27}$$

where $\Delta R_{ij}^2 = \left(y^i - y^j\right)^2 + \left(\phi^i - \phi^j\right)^2$, d_{ij} represents the distance between particles i and j and d_{iB} the distance between particle i and the beam remnant. The different types of recombination algorithm can be distinguished by the value of p that they use

$$p = \begin{cases} 1 & k_\perp \text{ algorithm [22, 23]} \\ 0 & \text{Cambridge/Aachen algorithm [24, 25]} \\ -1 & \text{anti-}k_t \text{ algorithm [26]} \end{cases} \tag{1.28}$$

In each case, given a minimal interjet separation, R, the algorithm proceeds by calculating $d_{i(j,B)}$ between each pair of objects and finds the smallest of these. If this is from the set d_{ij} then objects i and j are combined; if it is from the set d_{iB} then i is identified as a jet and removed from the list of objects. This process is repeated iteratively until there are no objects left. Historically, this has been computationally intensive, but recently an efficient implementation has been developed and is available through the FastJet software package [1, 2].

Sequential recombination algorithms of this type guarantee that all jets will be separated by a distance of at least R on a $y - \phi$ cylinder. The absolute number of jets is not infrared safe, as soft jets can be identified near the beam remnant, but, if a p_T threshold is instituted, the number of jets above the threshold represents an infrared safe quantity.

The Cambridge/Aachen algorithm is the simplest recombination algorithm, relying only on distance weighting. Because of this, it is sensitive to the distribution of soft objects, often producing irregularly shaped jets. Because the Cambridge/Aachen algorithm has a clustering hierarchy which is dependent on angle, it is possible to view a specific jet on a range of angular scales, a feature which is commonly used in studies of jet substructure.

The $k_\perp$ algorithm forms clusters from pairs of low p_T objects. By combining objects in this way, it is inherently collinear safe, proactively including soft QCD radiation. This means that $k_\perp$ jets can have irregular shapes which are complicated to deal with experimentally, particularly when attempting to correct for the effects of underlying event or pileup.

The anti-k_t algorithm forms clusters from pairs of high-p_T particles, disfavouring clustering between pairs of soft objects. Unlike $k_\perp$ and Cambridge/Aachen, reversing the order of clustering in anti-k_t cannot be usefully related to sequential branching in QCD, essentially anti-k_t proceeds by agglomerating soft material into an existing

hard subjet, rather than by first constructing the soft subjet and then recombining it with the hard subjet. Most pairwise clusterings will involve at least one hard object and such an object will therefore tend to accumulate all softer objects within R of its centroid. This results in anti-k_t jets having a roughly circular shape with radius $\simeq R$, allowing for easy experimental calibration. This feature has led to anti-k_t being adopted as the default jet algorithm used by the ATLAS and CMS experiments.

A comparison between different jet algorithms can be seen in Figure 1.3. Here the jets produced by $k_\perp$, Cambridge/Aachen, anti-k_t and SISCone are shown, using the same input distribution of particles and the same R-parameter in each case. The two highest p_T jets in the event, here shown in green and red, are essentially identical in each case, although the precise details of which particles are combined are different in each case. Additionally, there is some variation in the ordering of the lower p_T jets. The circular shapes of the anti-k_t jets and the irregular outline of the $k_\perp$ jets are particularly notable.

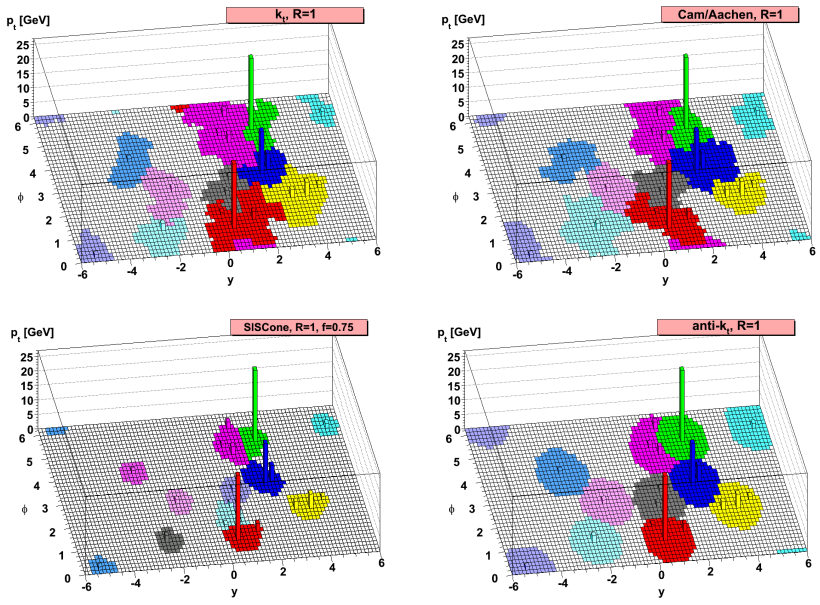


Figure 1.3.: A sample HERWIG (see Section 1.11.2) generated parton level event, overlaid with random soft particles, clustered with four different jet algorithms, illustrating the shapes and sizes of the resulting hard jets. For $k_\perp$ and Cambridge/Aachen the detailed shapes are partially determined by the specific set of soft particles used, and would change if these were to be modified [1, 2].

1.11. Monte Carlo Generators

The Monte Carlo method refers to any procedure that makes use of random numbers and probabilistic statistics to solve problems. Monte Carlo methods are used extensively in numerical analysis and simulation of natural phenomena. In the context of particle physics, Monte Carlo generators are used to produce theoretical simulations of real events. Often a variety of different programs are used: feasibly one program could generate a hard process while another evolves it through a parton shower algorithm and a third hadronises the coloured products of the shower. Different Monte Carlo generators often simulate different physics models, using different matrix elements, PDFs, evolution equations, parton showers or hadronisation models. Because of this, comparing a variety of Monte Carlo models to data provides a test of the compatibility of different theories with experimental results. The simulation of physics events in ATLAS is carried out in two steps:

- **Event generation**: Theoretical and phenomenological models are used to simulate the physics processes which result from proton-proton collisions, producing particle level information as an output.
- **Detector simulation**: These final-state particles are passed through a full simulation of the ATLAS detector [27] that is based on GEANT4 [28] and aims to reproduce the behaviour of particles as seen in test-beam studies.

After a given Monte Carlo generator has produced particle level predictions which have been passed through the ATLAS simulation, the events are treated in the same way as data: in particular, jets are reconstructed and calibrated using the reconstruction chain discussed in Section 3.2.

1.11.1. PYTHIA

The PYTHIA [29] Monte Carlo generator implements leading-order matrix elements from perturbative QCD for $2 \to 2$ processes, followed by p_{T}-ordered parton showers, calculated in the leading-logarithm approximation and finally the Lund string model for hadronisation. The underlying event in PYTHIA consists mainly of multiple-parton interactions interleaved with the initial state parton shower. Samples used in ATLAS are generated using the ATLAS Minimum Bias Tune 1 (AMBT1) set of

parameters [30], in which the non-diffractive model is tuned to ATLAS measurements of charged particle production at $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV. The AMBT1 tune uses the Martin-Roberts-Stirling-Thorne (MRST) LO* PDFs [31, 32].

PYTHIA 6.423 events produced with the Perugia 2011 [33] set of tuned parameters are also used. These parameters have been tuned to reproduce the jet shape and hadronic event shape spectra seen in LEP and Tevatron data.

1.11.2. HERWIG

The HERWIG [34] generator uses identical leading order matrix elements to PYTHIA, but applies an angular-ordered parton shower and a clustering hadronisation model. For the underlying event, HERWIG 6 is linked to JIMMY [35] to provide multiple partonic interactions. HERWIG++ [36], the latest version of HERWIG, directly implements a JIMMY-like approach to the underlying event. The HERWIG++ event samples are generated using the MRST LO* PDF set with the LHC Underlying Event 7-2 (LHC-UE7-2) tune for the underlying event [37].

1.11.3. ALPGEN

The ALPGEN [38] Monte Carlo generator provides leading order matrix elements with up to six partons in the final state. The ALPGEN samples are generated using the CTEQ6L1 PDF set, from the Coordinated Theoretical-Experimental Project [39] and are then passed through HERWIG and JIMMY (see Section 1.11.2) to provide parton showering, hadronisation and multiple partonic interactions with the ATLAS Underlying Event Tune 1 (AUET1) set of parameters [40].

1.11.4. POWHEG

POWHEG [41–43] is a generator capable of simulating next-to-leading order inclusive jet and dijet production. POWHEG allows the use of either PYTHIA or HERWIG to shower the partons, hadronise them, and model the underlying event. The advantage over the standard $2 \to 2$ matrix elements that are provided in PYTHIA and HERWIG, is that the emission of a third hard parton is calculated at matrix element level,

allowing observables that are dependent on the third jet to be calculated more accurately.

In the POWHEG algorithm, the genesis of each event comes from a QCD $2 \to 2$ partonic scatter. The renormalisation and factorisation scales are then both set to be equal to the transverse momentum of the outgoing partons, before the hardest partonic emission in the event is generated using the Martin-Stirling-Thorne-Watt (MSTW) 2008 NLO PDFs. Once the hardest partonic emission is simulated, the events can be evolved to the hadron level. The fact that POWHEG has a full parton shower interface ensures that, unlike in the case of fixed order NLO calculations, accurate predictions can be produced for multijet final states without the need to independently estimate soft corrections and their associated uncertainties.

1.11.5. NLOJET++

NLOJET++ is not a Monte Carlo generator, but rather a C++ program for explicitly determining leading and next-to-leading order cross-sections. It uses a modified form of the Catani-Seymour dipole subtraction method to calculate a variety of QCD processes in proton-proton collisions. Currently, NLOJET++ is able to compute n-jet cross-sections at next-to-leading order for $n \leq 3$ as well as four-jet cross-sections at leading order.

In order to provide results which are comparable with data, final distributions produced by NLOJET++ have to be corrected for the effects of hadronisation. Usually this is done on a bin-by-bin basis: final distributions are produced using NLOJET++ and separately using a leading-order Monte Carlo such as PYTHIA. The PYTHIA distributions are produced both at parton level, taking only the final state partons in the event, and at hadron level, once these partons have passed through the relevant hadronisation model. The ratio between these two distributions is then applied to the NLOJET++ distribution as a correction for the effects of hadronisation. Several different leading order Monte Carlos are used in order to allow an estimation of the uncertainty in physics modelling arising from this procedure.

1.11.6. HEJ

High Energy Jets (HEJ) is a Monte Carlo generator aiming to simulate production of multiple jets at a hadron collider. It is based on the BFKL kernel, implementing an all-order resummation of the perturbative terms which dominate the production of well-separated multijet events. This calculation is made in the limit of infinite rapidity separation between all partons produced in the event and is therefore most suited to events in which the most forward and most backward jets are widely separated.

The $2 \to n$ matrix elements are studied for $n \geq 2$ in the Multi-Regge kinematic limit in which scattering amplitudes are dominated by t-channel gluon exchange [44] and each of the successively emitted, rapidity-ordered gluons has a similar momentum to the previous ones. In this limit, the perturbative terms which describe multiple gluon emissions can be factorised, with the final scattering amplitude depending only on the transverse momenta of the emitted gluons.

This method approximates real emissions at all orders, also including virtual emissions in a way that ensures that any soft divergences from the virtual corrections cancel with those from gluon emissions. This results in a regularised matrix element at all orders in α_S, with contributions coming from any number, greater than two, of hard jets. The jet rates have been fully matched to tree-level accuracy up to a total of four jets.

As the centre-of-mass energy increases, the hard radiative corrections that HEJ was developed to describe become increasingly important; particularly when considering dijet systems with a large invariant mass or a large rapidity difference between the jets.

At present, HEJ is only capable of producing parton level predictions as it does not yet have proper parton shower matching. However, since it is not a fixed order calculation, the number of partons produced is variable, and relatively large in comparison to standard fixed order generators. Accordingly, in order to provide meaningful results, the partons produced as output by HEJ must be clustered into jets and corrected for soft effects in the same manner as described for NLOJET++ in Section 1.11.5.

Chapter 2.

The Large Hadron Collider and the ATLAS Detector

"I'm not concerned about all hell breaking loose, but that a PART of hell will break loose... it'll be much harder to detect."

— George Carlin

2.1. Overview

The Large Hadron Collider (LHC) at CERN has extended the frontiers of particle physics through its unprecedented energy and luminosity. In 2010, the LHC collided proton bunches, each containing more than 10^{11} particles, 20 million times per second, providing 7 TeV proton-proton collisions at instantaneous luminosities of up to 2.1×10^{32} cm^{-2} s^{-1}. A diagrammatic view of the CERN accelerator complex is shown in Figure 2.1

The high interaction rates, radiation doses, particle multiplicities and energies, when combined with the requirements for precision measurements, have set new standards for the design of particle detectors. ATLAS [5] is one of two general purpose detectors at the LHC that have been built to probe proton-proton and heavy ion collisions.

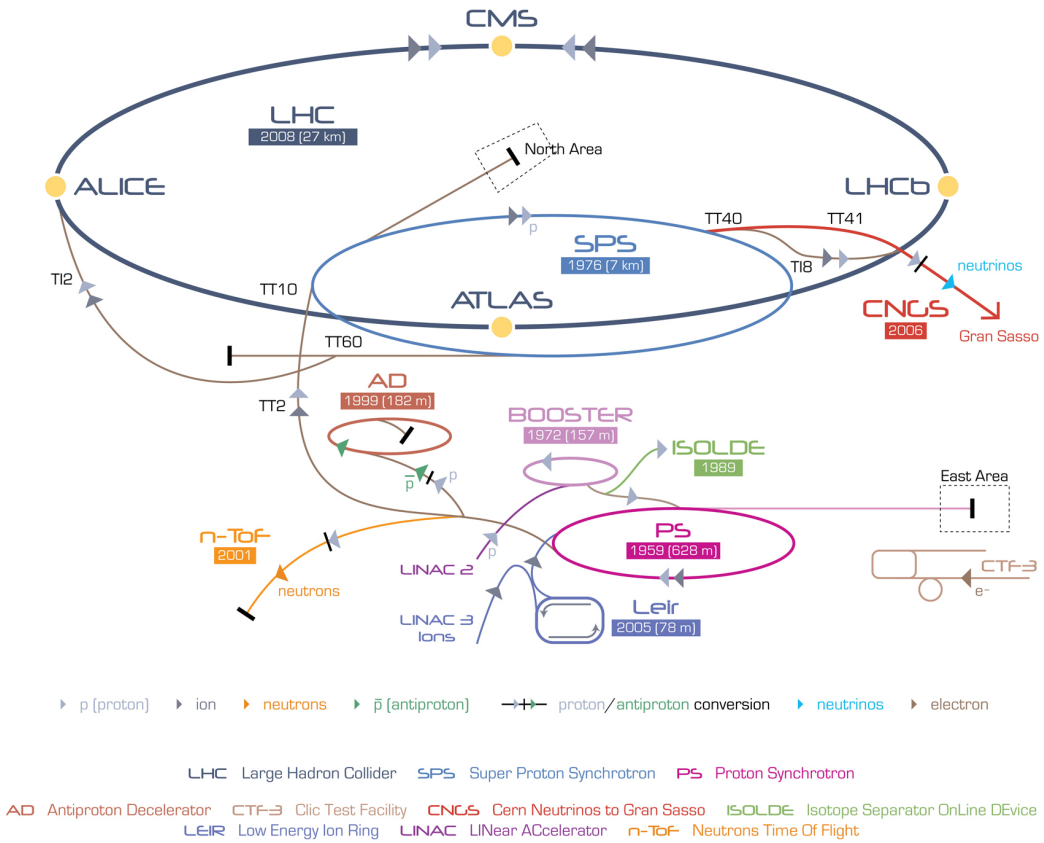


Figure 2.1.: Overview of the CERN accelerator complex [3]: a succession of particle accelerators that are used to reach sequentially higher energies. Each accelerator boosts a beam of particles, before injecting it into the next one in the sequence. Protons, obtained through ionising hydrogen atoms, are injected from the linear accelerator (LINAC2) into the PS Booster, then the Proton Synchrotron (PS), followed by the Super Proton Synchrotron (SPS), before finally reaching the Large Hadron Collider (LHC) ring. Protons circulate in the LHC ring for 20 minutes before reaching their maximum energy.

2.2. The ATLAS Detector

The ATLAS detector, shown in Figure 2.2, is divided into different detector subsystems, among them the tracker, the calorimeter system, the minimum bias trigger scintillators and the muon system. This thesis deals mainly with the calorimeter and will therefore only briefly mention other detector components.

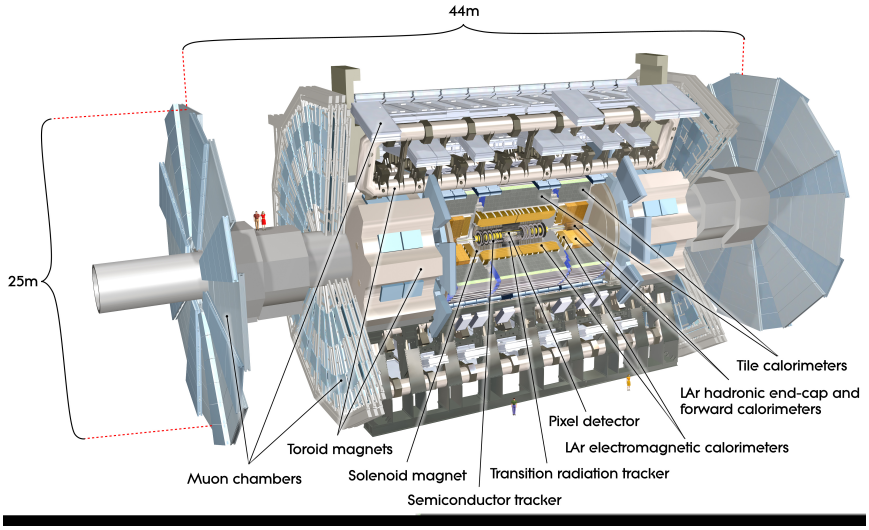


Figure 2.2.: Overview of the full ATLAS detector showing the four major subsystems, with human figures to scale [4]. The inner detector, consisting of the SCT and TRT, measures the momentum of charged particles. The calorimeter systems, the LAr EM, endcap and forward calorimeters, together with the tile calorimeters, measure the energies of interacting particles. The muon chambers, on the outer edges of the detector, identify muons and measure their momenta. The solenoid and toroid magnet systems bend charged particles, allowing for greater precision in momentum measurement.

The inner (tracking) detector has complete azimuthal coverage and spans the pseudorapidity[1] range $|\eta| < 2.51$. In this region, where the track density is large, the desired high-precision measurements require excellent resolution of momenta and of vertex positions. In order to make these high-granularity measurements, a variety of

[1] ATLAS uses a right-handed coordinate system with the origin taken at the nominal interaction point (IP) in the centre of the detector, the positive x-axis pointing inwards towards the centre of the LHC ring and the positive y-axis pointing upwards. The side of the detector with positive z is termed side-A while the negative z side is known as side-C. Cylindrical coordinates (r, θ) are used in the transverse plane, with ϕ the azimuthal angle around the beam axis. The pseudorapidity and rapidity are defined as discussed in Section 1.8. Distances, ΔR, in the pseudorapidity-azimuthal angular space are defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

tracking technologies are used. Layers of silicon pixel detectors, silicon microstrip detectors (SCT) and straw tube transition radiation tracking detectors (TRT) are all are surrounded by a solenoid magnet that provides a uniform magnetic field of 2 T. This allows tracks with $p_{\mathrm{T}} \leq 500$ MeV to achieve resolutions of between 0.4% and 1% depending on location in η. Additionally, tracks with $p_{\mathrm{T}} < 500$ MeV have resolutions of between 0.1% and 2%, again depending on their η [45].

The ATLAS muon spectrometer, covering the pseudorapidity range $|\eta| < 2.7$, can identify and reconstruct muons without input from any other subdetectors. As muons are the only interacting particles that can reliably pass through the calorimeter systems without stopping, they can be cleanly detected in the surrounding spectrometer. A series of drift tubes are used to detect the magnetic deflection of muon tracks while and cathode strip chambers provide precision position measurements. Overall, this system aims to identify muons with momenta between 3 GeV and 1 TeV.

The Minimum Bias Trigger Scintillators (MBTS) aim to select soft collisions between two interacting protons. The scintillators are mounted on the inner surface of the liquid argon (LAr) end-cap cryostats and cover a pseudorapidity range of $2.12 \leq |\eta| < 3.85$. The MBTS system is constructed from 2 cm polystyrene-based scintillator counters, 16 on each side of the detector, which, between them, cover the full azimuthal range.

2.2.1. Calorimeter Overview

Figure 2.3 shows an overview of the sampling calorimeters at ATLAS. The calorimeter systems cover the full range, $|\eta| < 4.9$, although different techniques must be used in different η regions, due to the varying radiation environment and the requirements of the physics processes of interest.

High granularity liquid-argon (LAr) electromagnetic (EM) sampling calorimeters cover the range $|\eta| < 3.2$. In the range $|\eta| < 1.7$, the hadronic calorimetry is provided by a scintillator-tile calorimeter. This is separated into a large central barrel and two smaller extended barrel cylinders, one on either side of the main barrel. In the end-caps, $|\eta| > 1.5$, LAr technology is again used for the hadronic calorimeters, matching the outer $|\eta|$ limits of the end-cap electromagnetic calorimeters. The LAr

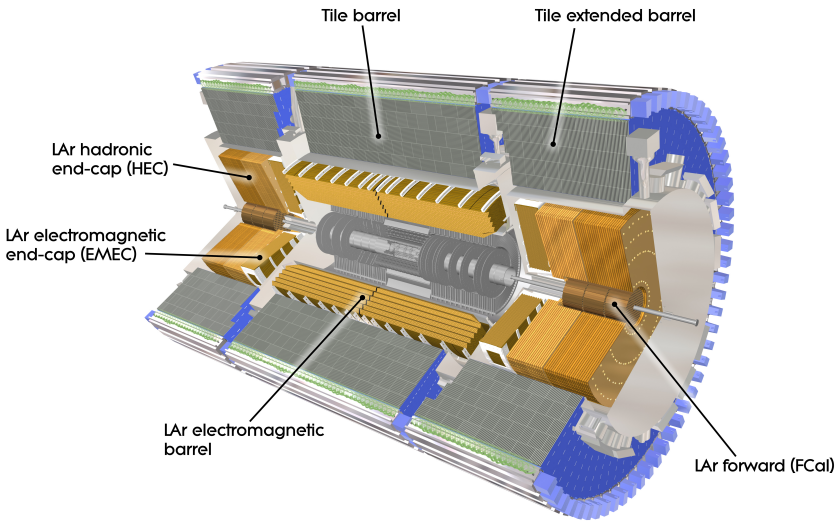


Figure 2.3.: Cross-sectional overview of the ATLAS calorimeter systems [5]. The LAr-based electromagnetic, forward and hadronic end-cap calorimeters are shown in brown, while the hadronic tile calorimeters are shown in gray and green.

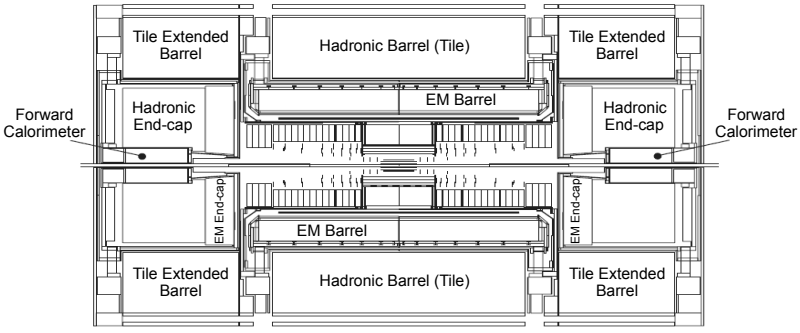


Figure 2.4.: Schematic transverse view (r-z view) of the calorimeters in the ATLAS detector [5]. A cylindrical coordinate system is used, with the z-axis along the proton-beam.

forward calorimeters extend out to $|\eta| = 4.9$, providing both electromagnetic and hadronic energy measurements.

For the inner detector, $|\eta| < 2.5$, the EM calorimeter is fine-grained to allow precision measurements of electrons and photons to be made. The other calorimeters in this region are coarser grained, although still possess sufficient resolution to satisfy the physics requirements for jet reconstruction and measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$.

Calorimeter depth is an important design consideration: the calorimeters must contain electromagnetic and hadronic showers and limit punch-through into the muon system. Electromagnetic showers are characterised by their narrow lateral profiles and are longitudinally parameterised by their radiation length X_0. Hadronic showers usually have a larger transverse spread and their nuclear interaction length, λ, is typically an order of magnitude greater than X_0, although this is material-dependent.

The barrel of the EM calorimeter is more than 22 X_0 thick, while the thickness of the end-cap is more than 24 X_0. For high energy jets, the active calorimeter comprises 9.7 λ in the barrel and 10 λ in the end-caps; providing adequate resolution. Figure 2.5 shows the total amount of material, including non-instrumented sections, in units of λ, as a function of $|\eta|$. Taking all of this into account, the total thickness, as demonstrated using test beams, is sufficient to reduce punch-through below the level of prompt or decay muons. Together with the large η coverage, this also ensures a good $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement, which is a particularly important signature for those analyses looking for evidence of supersymmetry.

2.2.2. Liquid Argon Electromagnetic Calorimeter

The electromagnetic calorimeter barrel, $|\eta| < 1.475$, and end-caps, $1.375 < |\eta| < 3.2$, are each housed in their own cryostat. The central solenoid, located in front of the electromagnetic calorimeter, is a significant source of dead material - approximately 1 λ at its greatest extent, restricting the maximal achievable calorimeter performance. Placing the solenoid and the liquid argon calorimeter inside a single vacuum vessel mitigates this effect by eliminating two vacuum walls which would otherwise be required. The barrel calorimeter consists of two identical half-barrels, separated by a 4 mm gap at $\eta = 0$. The two end-cap calorimeters are each subdivided into two

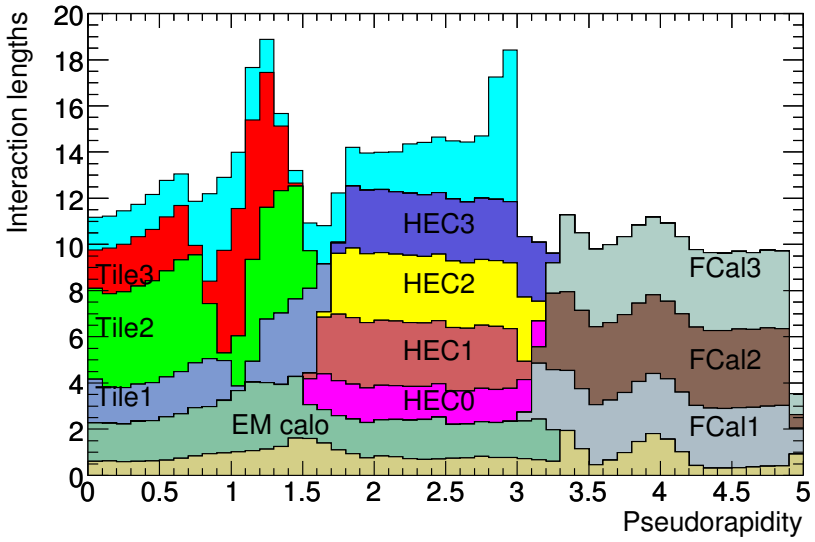


Figure 2.5.: Cumulative amount of material in the ATLAS detector, in units of interaction length, λ, as a function of $|\eta|$. The coverage from each individual calorimeter component is shown separately, while the sections closest to the interaction point, which are not instrumented for calorimetry, are shown in brown [5].

coaxial wheels: the outer wheel covering the region $1.375 \leq |\eta| < 2.5$ and the inner wheel $2.5 \leq |\eta| < 3.2$.

The electromagnetic calorimeter is a lead-LAr detector with lead absorber plates along its full coverage and readout provided by accordion shaped electrodes. This accordion geometry avoids the necessity for azimuthal cracks, providing complete symmetry in ϕ. For $|\eta| < 2.5$, the region instrumented for precision physics, the calorimeter is segmented in depth into three longitudinal layers, numbered from 1 to 3 outwards from the beam axis. Outside this region, the end-cap inner wheel is divided into only two sections in depth and also has a coarser lateral granularity.

A presampling detector is installed for $|\eta| < 1.8$ in order to allow the energy lost by electrons and photons upstream of the calorimeter to be measured and corrected for. This consists of an active LAr layer which is 1.1 cm thick in the barrel region, and 0.5 cm thick in the end-cap.

2.2.3. Hadronic Calorimeters

Tile Calorimeter

As can be seen in Figure 2.4, the tile calorimeter is located outside the electromagnetic calorimeter, extending radially from 2.28 m to 4.25 m. It is a sampling calorimeter, consisting of steel absorbing plates interleaved with scintillating tiles, which provide the active material. The hadronic barrel covers $|\eta| < 1.0$ and is complemented by two extended barrels covering $0.8 \leq |\eta| < 1.7$; each of these subdetectors is divided azimuthally into 64 modules in ϕ.

Similarly to the electromagnetic calorimeter, the tile is segmented into three layers: approximately 1.5, 4.1 and 1.8 λ thick for the barrel and 1.5, 2.6 and 3.3 λ thick in the extended barrel. Two sides of the scintillating tiles are read out by wavelength shifting fibres into two separate photomultiplier tubes. The readout cells, built by grouping fibres together in the photomultipliers, are pseudo-projective towards the interaction region.

Liquid Argon Hadronic End-cap Calorimeter

The Hadronic End-cap Calorimeter (HEC) is a copper detector consisting of two independent wheels per end-cap, located directly behind the end-cap electromagnetic calorimeter (Figure 2.4) and sharing the same LAr cryostats. The HEC covers the region $1.5 \leq |\eta| < 3.2$, overlapping with both the tile calorimeter and the forward calorimeter.

Each of the HEC wheels, which span $0.475 \leq r < 2.03$ m, is divided into 32 identical wedges in ϕ and two segments in depth, giving four layers per end-cap in total. The wheels closest to the interaction point are built from 25 mm parallel copper plates, while those further away use 50 mm copper plates: in each case the first of these plates is half-thickness. As the HEC is a sampling calorimeter, the copper plates are interleaved with 8.5 mm gaps filled with liquid argon, which provides the active medium.

2.2.4. Liquid Argon Forward Calorimeter

The Forward Calorimeter (FCAL) consists of one module at either end of the detector, approximately 10 λ deep, providing coverage for the region $3.2 \leq |\eta| < 4.9$. It is integrated into the end-cap cryostats although recessed by about 1.2 m from the electromagnetic calorimeter; because of the limited available depth a high-density design is required. The two FCALs each contain three modules: a copper calorimeter, optimised for electromagnetic measurements, and two tungsten calorimeters, predominantly aiming to perform hadronic measurements. Each module has the form of a matrix, with regularly spaced longitudinal channels filled with concentric tubes, parallel to the beam axis. Each tube contains a rod, and is otherwise filled with liquid argon which acts as the sensitive medium.

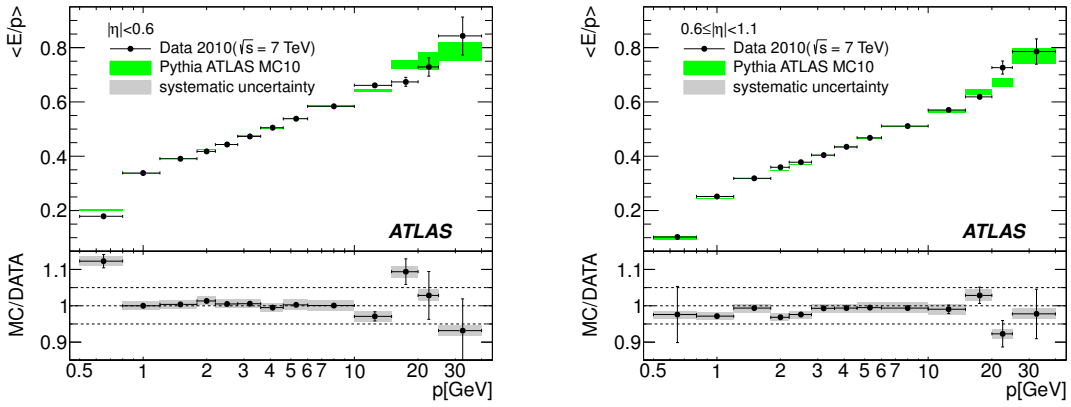
2.3. Detector Performance

2.3.1. Single Hadron Detector Response

The response of the calorimeters to isolated charged hadrons can be examined by considering the ratio of the energy, E, deposited in the calorimeter to the momentum, p, of the associated track. Comparing the $\langle E/p \rangle$ distribution in data and Monte Carlo allows an evaluation to be made of the modelling uncertainty on the calorimeter response in Monte Carlo. Figure 2.6 shows $\langle E/p \rangle$ as a function of track momentum for two different bins in $|\eta|$, after background subtraction. The agreement between data and Monte Carlo simulation is within $\sim 2\%$ for particles with momenta in the 1-10 GeV range, increasing to around 5% for momenta in the 10-30 GeV range.

2.3.2. Jet Energy Resolution

Fully calibrated jets, see Section 3.2 for details, reconstructed using the anti-k_t algorithm with $R = 0.6$ are used to study the jet resolution. Two different methods are used here: the first coming from dijet p_{T} balancing and the second from constructing quantities parallel and transverse to the dijet system, known as the bisector technique.



(a) $|\eta| < 0.6$ **(b)** $0.6 \leq |\eta| < 1.1$

Figure 2.6.: $\langle E/p \rangle$ as a function of the track momentum for (a) central and (b) more forward tracks. The black markers represent $\sqrt{s} = 7$ TeV collision data, while the green rectangles show the Monte Carlo prediction, with the vertical width indicating the associated statistical uncertainty. The lower sections show the ratio of the Monte Carlo simulation prediction to collision data. The grey band indicates the size of the systematic uncertainty on the measurement [6].

For jets with $|y| < 2.8$, the results from these two methods are found to be consistent in both data and Monte Carlo.

Figure 2.7 shows the fractional jet energy resolution as a function of the average p_{T} of the dijet system. The calorimeter jet reconstruction and selection efficiency relative to track jets is determined in data and Monte Carlo using a tag and probe technique and found to be in good agreement, within systematic uncertainties, for track jets with p_{T} values between 5 and 40 GeV [7].

2.4. Trigger System

When the design bunch spacing is reached, with collisions occurring every 25 ns, the event rate of 40 MHz is five orders of magnitude greater than the rate at which data can be recorded, which is limited to 400 Hz. This means that the trigger systems must achieve an overall rejection factor of 10^5 while avoiding biases and retaining the highest possible proportion of interesting events. During 2010, the bunch spacing was kept at 50 ns, slightly easing this pressure.

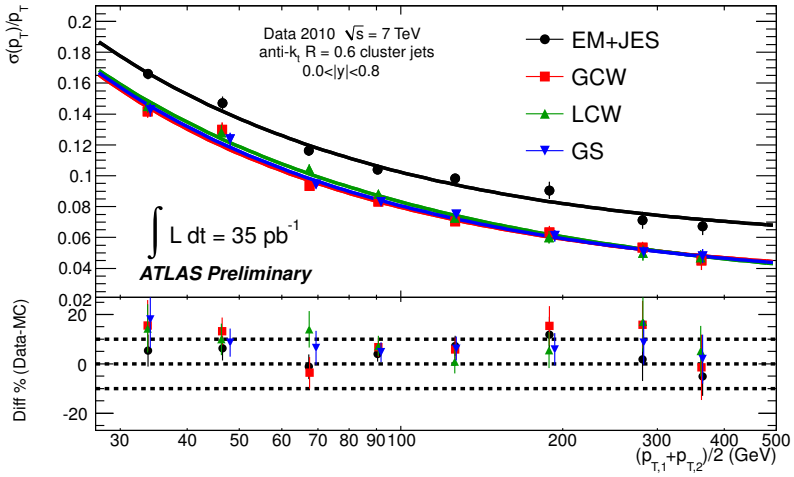


Figure 2.7.: Fractional jet energy resolution as a function of average dijet p_{T} for four different jet calibration schemes: the EM+JES, Global Cell Weighting (GCW), Local Cluster Weighting (LCW) and Global Sequential (GS) calibrations; only EM+JES has been fully validated for use in ATLAS, although other calibration schemes may be used in future data taking. Lines of best fit are shown in each case. The lower section shows the relative difference between Monte Carlo and the results from data. The black dotted lines indicate a relative uncertainty of $\pm 10\%$ [7].

The ATLAS trigger system has three distinct levels: Level-1 (L1), Level-2 (L2), and the event filter (EF). Each level of the trigger refines the decision made by previous levels by applying additional, stricter selection criteria. Initially, event information is accepted from the readout electronics and buffered; the hardware-based L1 trigger system then uses a subset of the available detector information to reject events: budgeting 2.5 ms of processing time per event, it is able to reduce the overall rate to 75 kHz. Events which pass the L1 trigger are then transferred to the L2 and EF trigger systems, collectively known as the high-level trigger (HLT). Each HLT trigger is seeded from a specific lower level trigger and is able to examine relevant features of the event in greater detail in order to make an overall trigger decision. Events passing one or more triggers after the final, "Physics" decision are passed to the data acquisition system.

The central trigger processor implements a 'trigger menu', comprising a list of trigger selections together with their associated prescales; only triggers specified by the processor are allowed to run in each event. The majority of trigger menu items are prescaled: only recording a set proportion of otherwise acceptable events.

This allows optimal use of available bandwidth in the face of changing luminosity conditions.

2.4.1. Level-1 Trigger

Due to the limited time available, the L1 trigger only uses a small amount of the available detector information: namely the calorimeter and muon systems. The aim of the L1 system is to identify high p_{T} leptons, photons and jets as well as events with large $E_{\mathrm{T}}^{\mathrm{miss}}$ or $E_{\mathrm{T}}^{\mathrm{total}}$. High p_{T} muons are identified using trigger chambers in the barrel and end-cap regions of the muon spectrometer, while the identification of other interesting event features is based on reduced-granularity information from all parts of the calorimeter system. Events which pass the L1 trigger selection are transferred to the L2 processing system and, if accepted by the HLT, onwards to the data acquisition systems via point-to-point links.

In each event, the L1 trigger also defines one or more Regions-of-Interest (RoIs): the geographical coordinates in $\eta - \phi$ space of those regions within the detector where its selection process has identified interesting features. The RoI data includes information on the type of feature identified and the criteria passed, usually some sort of threshold. This information is passed to the HLT for later use.

2.4.2. High-Level Trigger

The RoI information from L1 is used to seed the L2 trigger. L2 triggers are able to use all of the available detector data in the region covered by the RoIs at full granularity and precision; this corresponds to approximately 2% of the total event data. An average event processing time of 40 ms is allowed at this stage, resulting in an output event rate of approximately 3.5 kHz. The event filter performs the final stage of event selection, implementing analysis procedures similar to those performed offline to reduce the final event rate to roughly 400 Hz. In particular, EF jet triggers implement the anti-k_t algorithm, although this was not used to reject events during 2010.

2.4.3. Jet Triggers

Although the spectrum of jets produced is steeply falling, it would be preferable from the point of view of physics analyses, particularly differential cross-section measurements, to have a roughly uniform rate across the full jet E_{T} spectrum. Collecting sufficient statistics across the spectrum is also important for the measurement of detector and trigger efficiencies. Accordingly, a series of inclusive jet triggers are used, each with a higher threshold than the previous one. By changing the prescales of these triggers, the jet trigger menu can be optimised to achieve a roughly flat event rate across the E_{T} spectrum despite rising luminosity. The jet-trigger thresholds can also be changed, but this is a less frequent operation.

The jet trigger is split into logically independent systems: one covering the central region, $0 \leq |\eta| < 3.2$ and one covering the forward region, $3.2 \leq |\eta| < 4.9$. At Level-1, these systems use information from different calorimeter subsystems: central trigger jets rely on information from the electromagnetic barrel, tile and end-cap calorimeters, while forward trigger jets use information from the forward calorimeters only. At Level-2, however, information from multiple trigger systems may be considered, depending on the position of the L1 RoI, while in the Event Filter, information from the whole detector is combined [46].

Additionally, the Minimum Bias Trigger Scintillators (MBTS), located in front of the end-cap cryostats and covering $2.09 \leq |\eta| < 3.84$, are often used in jet-based analyses to provide fully efficient triggering for low p_{T} jets. The trigger L1_MBTS_1, requiring at least one hit in the minimum bias scintillators, is the primary trigger used to select minimum-bias events in ATLAS.

2.4.4. Measurements of Trigger Efficiency from Data

Understanding the efficiencies of relevant triggers is an important part of any physics analysis. As little weight as possible needs to be given to techniques relying solely on performance in Monte Carlo models; in-situ methods are much preferred.

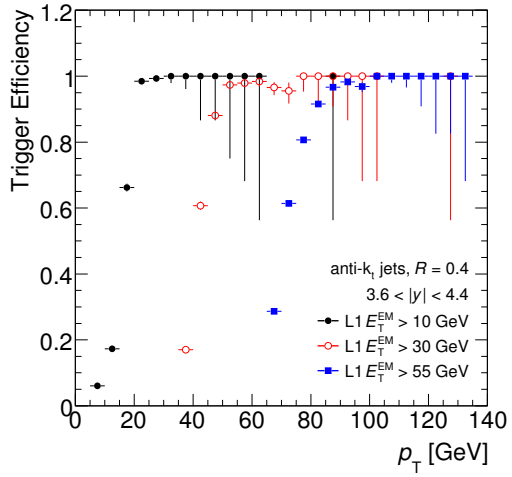
The two major in-situ methods are the "orthogonality" and "bootstrap" methods. In each case, these rely on constructing a sample of events which can then be examined to see what proportion of them pass the trigger of interest. Orthogonality

relies on taking events that pass a trigger which is known to be uncorrelated to the trigger under consideration. This can, however, introduce physics biases: muon triggers, for instance, could increase the proportion of heavy flavour jets. Jet trigger efficiencies are therefore usually determined using the bootstrap method in which events passing a lower threshold jet trigger, which is known to be on plateau in the p_{T} region of interest, are used. Firstly, events which pass minimum-bias triggers are used to measure the efficiency of those triggers with the lowest p_{T} thresholds, then the low threshold jet triggers are used as the baseline to select events with which to examine higher p_{T}-threshold triggers.

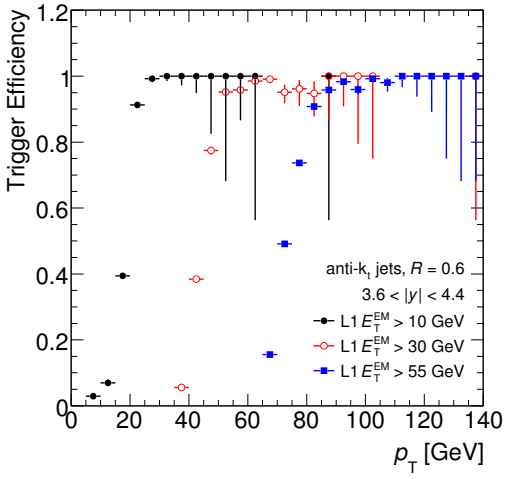
Efficiencies can be determined either on a *per-jet* or, more commonly, on a *per-event* basis. Per-event efficiencies are the fraction of events in which at least one jet passes the appropriate trigger among all events with at least one jet at the given p_{T}. Per-jet efficiencies are the fraction of jets passing the appropriate trigger among all jets at the given p_{T}. Per-jet efficiencies are more complex, due to their reliance on matching between offline jets and trigger objects and are rarely used.

Figure 2.8 shows per-event trigger efficiencies for six different triggers in the region $3.6 \leq |y| < 4.4$. Here the characteristic serpentine shape of a typical efficiency curve can be seen, rising from zero efficiency to reach a plateau. In Figure 2.8a and 2.8b the efficiencies for L1 triggers are shown. The likelihood that a given L1 trigger will fire, before prescale, is shown as a function of offline jet p_{T}, using anti-k_t jets, with $R = 0.4$ on the left and $R = 0.6$ on the right. Three different L1 thresholds are shown, each requiring a greater level of L1 $E_{\mathrm{T}}^{\mathrm{EM}}$ in order to trigger. The efficiency for the lowest threshold, L1 $E_{\mathrm{T}}^{\mathrm{EM}} > 10$ GeV, must be determined using MBTS triggers, but higher thresholds can be bootstrapped from previous jet triggers.

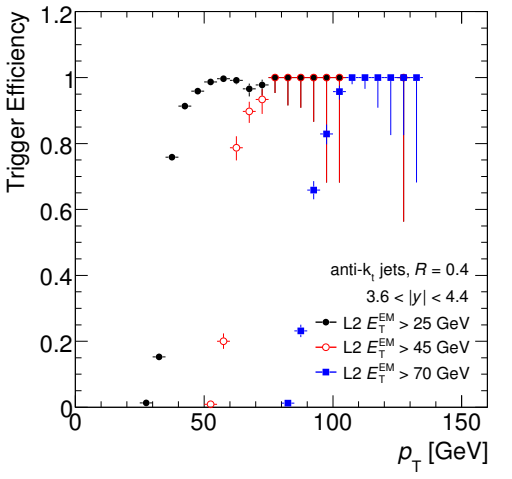
In Figure 2.8c and 2.8d, the efficiency curves for L2 triggers are shown, again as a function of offline jet p_{T}. The correlation between the $E_{\mathrm{T}}^{\mathrm{EM}}$ threshold and the offline p_{T} at which plateau is first reached is evident, as is the distinctive shape with which efficiency rises as a function of offline p_{T}. Specific details relevant to the efficiency curves shown here are discussed further in Section 5.4.2.



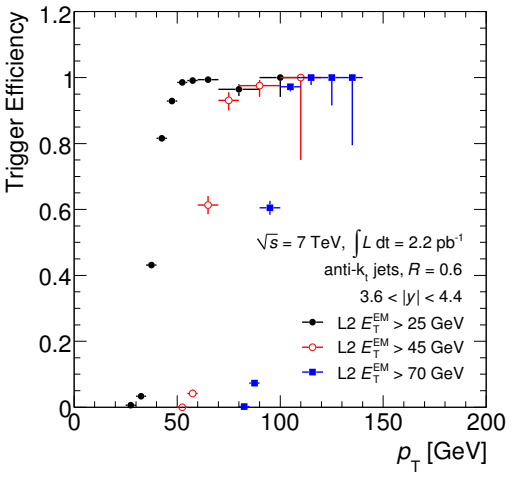
(a) L1 efficiency, anti-k_t $R = 0.4$ jets



(b) L1 efficiency, anti-k_t $R = 0.6$ jets



(c) L1+L2 efficiency, anti-k_t $R = 0.4$ jets



(d) L1+L2 efficiency, anti-k_t $R = 0.6$ jets

Figure 2.8.: Jet trigger efficiency at L1 (top) and combined L1+L2 (bottom) as a function of reconstructed jet p_{T} for anti-k_t jets with $R = 0.4$ (left) and $R = 0.6$ (right) in the forward region $3.6 \leq |y| < 4.4$, shown for three different trigger thresholds in each case. The trigger thresholds are at the electromagnetic scale, while the jet p_{T} is at the calibrated scale (see Section 3.2). Due to the presence of a dead FCAL trigger tower, which spans 0.9% of the $\eta - \phi$ acceptance, the efficiency is not expected to reach 100%.

Chapter 3.

Analysis Tools

"An architect's most useful tools are an eraser at the drafting board, and a wrecking bar at the site."
— Frank Lloyd Wright

3.1. Introduction

Progressing from raw detector output to final results and distributions, which can be compared to theoretical predictions, involves a series of sequential steps. The initial steps are generic to any analysis: the rejection of events for data quality reasons, the reconstruction of jets from calorimeter signals and their subsequent calibration are, in general, performed identically for each analysis. Additionally, unfolding for detector effects, in other words correcting distributions made at detector level back to the final-state particle level, is often necessary.

3.2. Jet Reconstruction and Calibration

The default jet-clustering algorithm in ATLAS is anti-k_t (see Section 1.10.2), implemented using FastJet [1, 2], with two different R-parameters: narrow jets with $R = 0.4$ and wide jets with $R = 0.6$.

Jets are reconstructed by applying a jet-clustering algorithm to calorimeter signals, and subsequently performing a calibration step, to correct for known detector effects. Two different inputs from the calorimeter can be used for jet-finding, towers and topological clusters.

Towers are formed by collecting cells into bins of a regular $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ grid, depending on their location, and summing up their signals, or a fraction of their signal corresponding to the overlap area fraction between the tower bin and the relevant cell. This summing stage is non-discriminatory, in other words all calorimeter cells are used in the towers. Towers with negative signals are recombined with nearby positive signal towers until the net signal is positive, thus all resulting towers have a valid physical four-vector and can directly be used by the jet finders. This approach can be understood as an overall noise cancellation rather than suppression, since noisy cells will still contribute to the jets.

Topological cell clusters [47] are an attempt to reconstruct three-dimensional energy depositions in the calorimeter [48, 49]. Firstly, any cell which satisfies $|E_{cell}| > 4\sigma_{cell}$ is identified as a "seed cell", where σ is the RMS noise of the cell due to electronic effects and pile-up. Any cell neighbouring a seed cell, which itself satisfies $|E_{cell}| > 2\sigma_{cell}$, is incorporated into the topocluster, and this process is repeated iteratively until there are no longer any cells adjacent to the topocluster with $|E_{cell}| > 2\sigma_{cell}$. Finally, an outer layer of cells is added, here accepting any surrounding cells which satisfy $E_{cell} > 0$ is added. Known hot cells or dead cells are excluded from this process. In contrast to using signal towers, using this 4-2-0 clustering scheme inherently incorporates noise suppression and results in fewer cells being included in the jet clustering step. In jet reconstruction, each topocluster is considered to be a massless particle with energy $E = \sum E_{cell}$ and position given by the energy-weighted centroid of cells in the cluster, with the direction pointing back towards the geometric centre of the detector.

Jets produced in this way are reconstructed at the electromagnetic (EM) scale, which is the basic signal scale for the ATLAS calorimeters. It accounts correctly for the energy deposited in the calorimeter by electromagnetic showers - validated using test-beam measurements with electrons and muons. It does not, however, correct for the lower hadron response, and a series of calibration steps is therefore needed to bring an uncalibrated, EM-scale jet to the hadronic scale energy scale [9, 10].

Jets which fall below the reconstruction threshold of 7 GeV are discarded before calibration.

3.2.1. Pile-up Correction

Jets calibrated at the EM-scale are affected by energy deposits arising from multiple proton-proton interactions within the same bunch crossing, known as pile-up. Pile-up can be either out-of-time, in other words, occurring slightly before or after the hard interaction or in-time, occurring contemporaneously. The ATLAS calorimeter response is such that the time integral of the signal corresponding to a single particle is zero: in other words a sharp positive peak is followed by a longer but lower amplitude trough. This ensures that the calorimeter is not saturated when large numbers of particles arrive in close succession but also means that, while in-time pile-up provides a positive contribution to the calorimeter signal, out-of-time pile-up will provide negative contribution. A correction to remove the average effects of these additional proton-proton interactions, derived using minimum bias data, is applied at the electromagnetic scale: the average additional E_{T} per calorimeter tower, measured as a function of η and the number of reconstructed primary vertices N_{PV}, is subtracted from each jet.

3.2.2. Jet Origin Correction

The calorimeter clusters used for jet reconstruction are assumed to originate from the geometrical centre of ATLAS. The jet origin correction first corrects each calorimeter cluster to instead point back to the primary vertex with the highest $\sum {p_{\mathrm{T}}}^{track}$ of the event; the beam spot is used if there is no primary vertex.

The kinematics of each calorimeter cluster are recalculated using the direction from the primary vertex to the centroid of the cluster. The raw jet four momentum is then redefined as the four vector sum of the clusters. This correction improves the angular resolution while the jet energy is unaffected. A small improvement in jet p_{T} resolution is introduced due to the changing jet direction, although this is rarely larger than 1%. Most of the effect of the correction comes from the z-position of the primary vertex.

3.2.3. Final Jet Energy Scale

The final part of the jet calibration involves applying a jet energy scale (JES) correction to account for the fact that the jets are reconstructed at the electromagnetic (EM) scale. This is known as the EM+JES calibration, and it corrects for calorimeter non-compensation, energy losses in inactive regions, out-of-cone showering effects as well as inefficiencies in the calorimeter clustering and jet reconstruction. This calibration is primarily dependent on energy, since the calorimeter response is energy-dependent, and the jet direction, due to the changing calorimeter technology and to the varying amounts of dead material in front of the calorimeters.

The EM+JES calibration is derived from simulated events, specifically the AMBT1 PYTHIA dijet sample. To derive the correction factors in Monte Carlo, isolated particle jets, reconstructed using final-state particles, are matched with isolated detector level jets, reconstructed using the full calorimeter level information. The particle jet energy is then divided by the EM-scale energy of the matching calorimeter jet in order to obtain the appropriate correction factor.

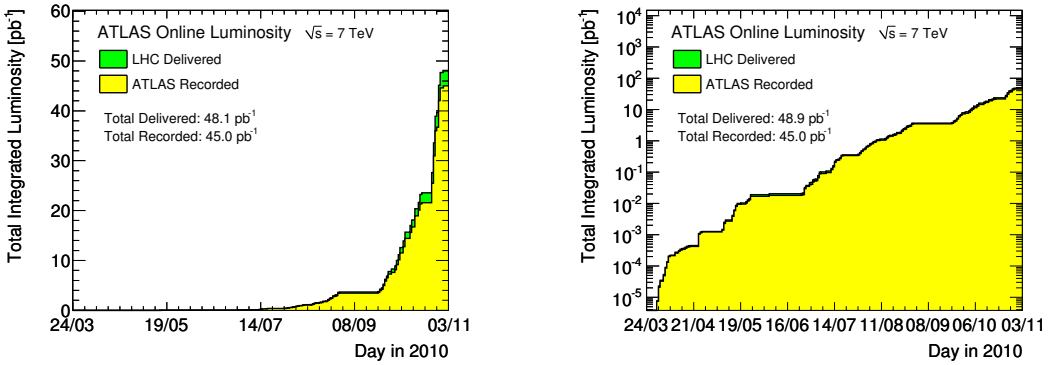
Following this, a small η-dependent correction is applied to remove a bias in the reconstructed η of jets that occurs when jets fall in poorly instrumented regions of the calorimeter that have a lower response than the regions around it. The reconstructed direction of the jet will be biased since the clusters that fall in these regions have a lower response when their four-vectors are added up to build the jet four-vector, and hence a smaller overall weight. As a consequence of this, the jet is pulled toward the region with the higher response. This η-correction is parameterized as a function of jet energy and pseudorapidity, and is small, $\Delta\eta < 0.01$, in most regions of the calorimeter, although larger in the crack regions: up to $\Delta\eta = 0.07$ for low p_{T} jets in the HEC-FCAL transition region.

3.3. Generic Event Selection Considerations

3.3.1. Luminosity

All of the studies presented in this thesis were carried out at a centre-of-mass energy of $\sqrt{s} = 7$ TeV using the 37.3 ± 1.2 pb^{-1} of integrated luminosity that was collected during 2010 (see Figure 3.1).

The luminosity was measured independently by multiple detectors and algorithms, each of which had differing acceptances, systematic uncertainties and background sensitivities. In 2010, the primary methods used were LUCID, a dedicated Čerenkov detector and counting hits in the MBTS. For both of these, the case in which hits were registered on the A-side AND the C-side of the detector was treated as a separate measurement from the case in which the requirement was only for a hit on at least one of these.



(a) 2010 luminosity evolution **(b)** 2010 luminosity evolution, logarithmic scale

Figure 3.1.: Total integrated luminosity available in ATLAS and recorded by the LHC in 2010. Identical information is shown in (a) and (b), with the only difference being that a logarithmic scale is used in the latter case [8].

The van der Meer method discussed in Section 1.9 was used to obtain σ_{vis} for each of these processes. The differences between the luminosities obtained from each method were monitored as a function of time and of μ. The measurements were finally combined to produce the overall ATLAS luminosity determination, together with its uncertainty [50].

A single run of proton-proton collisions is divided into luminosity blocks, each of which typically represents about two minutes of data taking, within which conditions

Period	A	B	C	D	E
Start of data taking	Mar 30	Apr 23	May 18	Jun 24	Jul 29
End of data taking	Apr 19	May 17	Jun 05	Jul 18	Aug 17
Luminosity [nb^{-1}]	0.380	8.07	8.46	201	1000
Period	E5	F	G	H	I
Start of data taking	Aug 10	Aug 19	Sep 22	Oct 08	Oct 24
End of data taking	Aug 17	Aug 30	Oct 06	Oct 18	Oct 29
Luminosity [nb^{-1}]	445	1810	6870	7250	19100

Table 3.1.: Dates and total integrated luminosity for each of the nine data taking periods for data collected by ATLAS in 2010. Separate numbers are presented for period E and period E5 (which excludes the first four sets of runs in this period) due to a software problem which affected forward jet triggers at this time.

such as bunch spacing and beam intensity are constant. The instantaneous luminosity, bunch size and bunch shaping each evolved over the course of this period of data taking; data collected in ATLAS is therefore divided into different run periods, with the boundaries of these periods being defined by changes in running conditions. The dates and total integrated luminosity corresponding to each of these run periods are summarised in Table 3.1.

3.3.2. Data Quality

Certain generic cuts, held in common between all analyses are necessary in order to extract useful events from the recorded data. First, the event is required to belong to one of a set of "good runs". These are specific luminosity blocks in which the relevant detector subsystems, trigger and reconstructed physics objects have passed a data-quality assessment and are deemed suitable for physics analysis. For the analyses discussed in this thesis, this means the central trigger processor, solenoid magnet, inner detectors (Pixel, SCT, and TRT), calorimeters (barrel, end-cap, and forward) and the luminosity recording system.

3.3.3. Primary Vertex

To reject events arising from cosmic-ray muons and other non-collision backgrounds, events are selected as collision candidates by requiring that they have at least one primary vertex that is consistent with the beamspot position and that has at least five tracks associated to it, each with $p_{\mathrm{T}} > 0.5$ GeV. This vertex definition is consistent with that used to evaluate pile-up vertices in the offset correction, as discussed in Section 3.2.1. The efficiency for collision events to pass this vertex requirement, although obviously analysis dependent, is generally well over 99%.

3.3.4. Jet Cleaning

Standard jet cleaning criteria have been developed in order to identify fake jets which arise due to noise or to out-of-time energy depositions. Jets failing these criteria are flagged as either "bad", likely to be fake, or "ugly", likely to be mismeasured due to falling into less well instrumented regions [51, 52]. Three main issues are addressed, with a dedicated set of selection criteria for each:

- **Single-cell jets in the HEC.** Most misreconstructed jets arise from noise bursts in the HEC. This results in jets with most of their energy coming from single calorimeter cells.
- **Bad quality jets in the EM calorimeter.** Noise bursts in the EM calorimeter, although rarer than in the HEC, result in jets with most of their energy coming from the EM calorimeter and whose cells have bad reconstruction "quality[1]".
- **Out-of-time jets.** When large out-of-time energy deposits appear in the calorimeter, possibly from photons produced by cosmic rays, jets will be reconstructed with timing that is incompatible with the event time.

A series of per-jet variables, seen in Table 3.2 are used to reject jets falling into one of these categories.

[1] Q-factor, one of the inputs to jet cleaning, is a measure of the quality of the signal in a given LAr cell, analogous to χ^2, parameterising how well the expected pulse shape fits the digitised samples and thus how well the amplitude is measured in this cell. Cutting on this variable aims to remove large but badly measured cell amplitudes which could otherwise fake a jet.

EMf	fraction of energy coming from the EM calorimeter
FMax	maximum energy fraction in one calorimeter layer
HECf	energy fraction in the HEC
LArQ	the fraction of energy coming from LAr cells having Q-factor > 4000
HECQ	same as the LArQuality except calculated only with the HEC
NegE	negative energy in the jet
t	the mean timing difference between cells in the jet and the event time
η	η at the EM-scale
Chf	the ratio of the $\sum {p_{\mathrm{T}}}^{track}$ associated to the jet divided by jet p_{T}

Table 3.2.: Per-jet variables used as an input to jet cleaning.

Three levels of bad jet rejection have been determined by the ATLAS Jet/Etmiss Working Group. The most lenient of these is termed "loose" cleaning, with "medium" and "tight" successively applying stricter criteria in identifying additional jets as bad. The jet cleaning cuts used in each of these cases are shown in Table 3.3.

As well as this removal of bad jets, ugly jets are identified as those jets with more than 50% of their energy coming either from TileGap3, the transition region between the barrel and end-cap, or from known dead cells, which are assigned an energy value based on the values of their neighbouring cells.

3.3.5. Jet Energy Scale Uncertainty

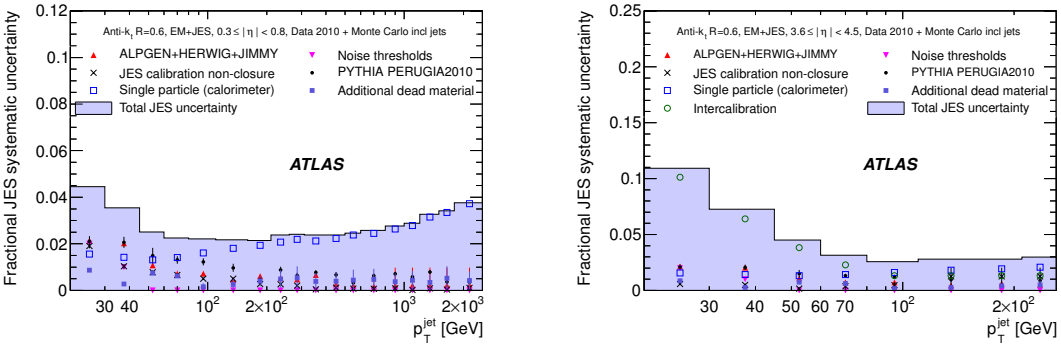
As discussed in Section 3.2.3, the jet energy scale (JES) is derived in Monte Carlo before being validated using a series of in-situ measurements. Evaluating the JES uncertainty therefore necessitates combining uncertainties arising from each of these sources: in-situ and single pion test-beam measurements, uncertainties on precise details of material distribution in the ATLAS detector, the Monte Carlo modelling used in event simulation and electronic noise must all be considered [9, 10].

Important individual sources of uncertainty include: non-closure when the JES correction is applied to reconstructed Monte Carlo; the single particle calorimeter response determined from in-situ measurements, as described in Section 2.3.1; the accuracy of detector simulations obtained by varying calorimeter noise in Monte Carlo

	Loose	Medium = Loose OR	Tight = Medium OR
HEC spikes	HECf > 0.5 & $\lvert$HECQ$\rvert > 0.5$ OR $\lvert$NegE$\rvert > 60$ GeV	HECf $> 1 - \lvert$HECQ$\rvert$	
EM coherent noise	EMf > 0.95 & $\lvert$LArQ$\rvert > 0.8$ & $\lvert\eta\rvert < 2.8$	EMf > 0.9 & $\lvert$LArQ$\rvert > 0.8$ & $\lvert\eta\rvert < 2.8$	EMf > 0.98 & $\lvert$LArQ$\rvert > 0.05$ OR $\lvert$LArQ$\rvert > 0.95$
Non-collision background and cosmics	$\lvert t\rvert < 25$ ns OR EMf < 0.05 & Chf < 0.05 & $\lvert\eta\rvert < 2$ OR EMf < 0.05 & $\lvert\eta\rvert \geq 2$ OR FMax > 0.99 & $\lvert\eta\rvert < 2$	$\lvert t\rvert < 10$ ns OR EMf < 0.05 & Chf < 0.1 & $\lvert\eta\rvert < 2$ OR EMf > 0.95 & Chf < 0.05 & $\lvert\eta\rvert < 2$	EMf < 0.1 & Chf < 0.2 & $\lvert\eta\rvert < 2$ OR EMf > 0.9 & Chf < 0.02 & $\lvert\eta\rvert < 2$ OR EMf < 0.1 & $\lvert\eta\rvert \geq 2$

Table 3.3.: The cuts used to remove bad jets as part of jet cleaning in ATLAS. Each of the three sets of cuts are shown: loose, medium and tight. A jet is considered "bad" if it passes any of these cuts. The medium cuts comprise the loose cuts plus additional requirements, while the tight cuts have the same relationship to the medium cuts. This means that any jet considered bad by the loose (medium) cuts will automatically be considered bad when using the medium (tight) cuts.

samples; the uncertainty associated with physics modelling, which is obtained from comparing the detector response in different Monte Carlo generators and finally the relative jet calibration obtained through η-intercalibration, as discussed in Chapter 4. The level of JES uncertainty is shown in Figure 3.2 as a function of jet p_{T} for central and for forward jets.



(a) $0.3 \leq |\eta| < 0.8$ **(b)** $3.6 \leq |\eta| < 4.5$

Figure 3.2.: Fractional jet energy scale systematic uncertainty is shown as a function of p_{T} for jets in two pseudorapidity regions: (a) $0.3 \leq |\eta| < 0.8$ and (b) $3.6 \leq |\eta| < 4.5$. In the forward region, the JES uncertainty is extrapolated from the barrel uncertainty, with the uncertainty contribution from the η-intercalibration between central and forward jets in data and Monte Carlo added in quadrature. The total uncertainty is shown as the solid light blue area. The individual sources are also shown, with uncertainties from the fitting procedure where applicable [9, 10].

3.3.6. Jet Trigger Threshold Evolution

As data conditions have changed over time, the trigger has changed to reflect this: both through enabling the HLT, which was initially used in passthrough mode, and by changing trigger prescales to keep the overall recorded event rate within acceptable bounds. This was primarily done by prescaling triggers with lower E_{T} thresholds, while the triggers with the highest E_{T} thresholds remained unprescaled.

In general, the aim of jet-based analyses is to retain as high a proportion of events as possible while minimising the biases and systematic uncertainties arising from the trigger. The easiest way to achieve this is to use the trigger in the plateau region where its efficiency is close to 100%. In order to accept as many events as possible, hence reducing statistical uncertainty, the trigger thresholds used in analyses are

usually chosen to be the lowest ones for which trigger efficiency is still larger than 99%, thus ensuring that the effects of prescaling are as small as possible.

For high p_{T} jets in the central region, this means that the trigger of choice changes over time, as successive triggers become increasingly prescaled. For low p_{T} jets, the majority of the recorded data comes from periods A–C, recorded between March and June 2010, when much of the trigger bandwidth was allocated to minimum bias triggers. For forward jets, the first four periods (A–D) could not be used, as the forward jet trigger had not yet been commissioned, so the majority of analyses can use only periods E–I.

In 2010 only L1 information was used to select events in the early periods, up until the summer, while L2 was used from the summer to the end of the year. The jet trigger did not reject events at the EF stage in 2010. In the early part of Period A, before run 152777, a mistiming in the L1 central jet trigger hardware caused large inefficiencies although MBTS triggers were not affected by this. Additionally, further calibration problems mean that forward jet triggers cannot be used during the early runs of period E.

3.4. Unfolding Detector Effects

Unfolding is a procedure which attempts to correct distributions made using information from the detector output back to the equivalent distributions that would have been seen given an ideal detector which was able to perfectly measure all final-state particles in the event. Unfolding aims to compensate for smearing effects in the detector as well as for event selection inefficiencies. This allows easy comparison to any theoretical calculation even if, at some future point, the precise details of the detector simulation are lost.

Various different unfolding methods are used in different analyses, but they all share one essential feature: they rely on comparison, using one or more Monte Carlos generators, between final-state particle level (hadron level) information and the equivalent information after detector simulation and reconstruction have been applied (detector level).

Bin-by-bin unfolding relies entirely on the shape of distributions in Monte Carlo in order to compute the correction factors. Firstly, for each relevant distribution, the ratio between the hadron level and detector level predictions is calculated. This ratio is then applied as a correction factor to the measured data. This technique can only be applied when migrations between bins are small in comparison to the bin contents, often forcing the use of large bins. Efficiency, the proportion of events which remain in the same bin between hadron level and detector level, and purity, the proportion of events in a given detector level bin which were in that bin at hadron level are important in assessing the impact of inter-bin migrations.

For the Singular Value Decomposition (SVD) method, the first step is to construct a 2D transfer matrix, providing full information about movements from one bin to another between hadron level and detector level. A singular value decomposition is then used in order to prevent fluctuations that a simple inversion of this transfer matrix would introduce. However, the regularisation procedure in SVD also uses a constraint on the curvature of the unfolded spectrum, which introduces long-range correlations in the result that produce an artificial smoothing of the unfolded distribution and which can also cause biases.

The Iterative, Dynamically Stabilised (IDS) method uses the same transfer matrix to compute the matrix of unfolding probabilities, which encodes the probability for an event reconstructed in a given bin i to be generated in bin j. The unfolding matrix is improved in a series of iterations, where the hadron level Monte Carlo is reweighted to the shape of the corrected data spectrum. The regularisation, preventing statistical fluctuations from being amplified by the successive iterations, is provided by the use of the significance of the differences between data and Monte Carlo in each bin. The final unfolding matrix, after the optimal number of iterations, is used to correct the reconstructed spectrum for detector effects.

Finally, unfolding based on Bayes' theorem uses the transfer matrix to obtain a series of transition probabilities: given a certain hadron level result, what are the probabilities for each detector level measured outcome. Bayes' theorem is then used to calculate the reverse probability: given a detector level outcome, what is the probability that the hadron level result could have come from each bin of the measurement: this process is then repeated iteratively [53].

Chapter 4.

η-intercalibration

"The best and safest thing is to keep a balance in your life. If you can do that, and live that way, you are really a wise man."
— Euripides

4.1. Introduction

As discussed in Section 2.2.1, the ATLAS calorimeters use different technology in different detector regions, with varying amounts of dead material in front of the calorimeters. In order to ensure that the calorimeter response to jets is uniform throughout $\eta - \phi$ space, it is therefore necessary to apply a jet-level calibration. This calibration is determined, at least in part, using Monte Carlo samples, however, given the non-compensating nature of the calorimeters together with the complex calorimeter geometry and material distribution, it is clear that such corrections need to be validated in-situ. The relative response of the calorimeter system to jets can be determined by studying the balance of transverse momenta in dijets.

4.2. Intercalibration using Events with Dijet Topologies

4.2.1. Intercalibration using a Central Reference Region

In the central region of the calorimeter, it is possible to use tracking information to reconstruct jets. This entails using "good" tracks in the inner detector; those with $p_{\mathrm{T}} > 500$ MeV, a transverse impact parameter less than 1.5 mm from the primary vertex, a longitudinal impact parameter satisfying $|z_0 \sin\theta| < 1.5$ mm and with at least six hits in the SCT. These tracks are then used as input to the anti-k_t algorithm, with the jets thus obtained being termed track jets. The calibration of calorimeter jets can be cross-checked using these track jets, in particular by looking for any systematic differences in their p_{T} or η distributions.

As a result, the standard approach for η-intercalibration with dijet events is to use the central region of the barrel, $|\eta| < 0.8$, as a reference region. The relative calorimeter response of jets in other calorimeter regions is quantified by the p_{T} balance between the reference jet and the probe jet, exploiting the fact that, in the absence of any additional radiation in the event, these jets are expected to have equal p_{T} due to conservation of transverse momentum. The p_{T} balance is characterised by the asymmetry $\mathcal{A}$, defined as

$$\mathcal{A} = \frac{p_{\mathrm{T}}^{\mathrm{probe}} - p_{\mathrm{T}}^{\mathrm{ref}}}{p_{\mathrm{T}}^{\mathrm{avg}}} \tag{4.1}$$

with $p_{\mathrm{T}}^{\mathrm{avg}} = (p_{\mathrm{T}}^{\mathrm{probe}} + p_{\mathrm{T}}^{\mathrm{ref}})/2$. If both jets fall into the reference region, each jet is used, in turn, to probe the other. As a consequence, the average asymmetry in the reference region will be zero by construction; although this may not be true in the regions $-0.8 \leq \eta < 0$ and $0 \leq \eta < 0.8$ when these are considered individually.

The asymmetry is then used to measure an η-intercalibration factor, c, for the probe jet or, conversely, the response of the probe jet relative to the reference jet, $1/c$, using the relation

$$\frac{p_{\mathrm{T}}^{\mathrm{probe}}}{p_{\mathrm{T}}^{\mathrm{ref}}} = \frac{2+\mathcal{A}}{2-\mathcal{A}} = 1/c \tag{4.2}$$

This analysis is performed in bins of jet η and $p_{\mathrm{T}}^{\mathrm{avg}}$. Using the standard method outlined above, there is an asymmetry distribution $\mathcal{A}_{ik}$ for each probe jet η-bin i and each $p_{\mathrm{T}}^{\mathrm{avg}}$-bin k. Intercalibration factors are calculated for each bin according to Equation (4.3)

$$c_{ik} = \frac{2-\langle\mathcal{A}_{ik}\rangle}{2+\langle\mathcal{A}_{ik}\rangle} \tag{4.3}$$

where $\langle\mathcal{A}_{ik}\rangle$ is the mean value of the asymmetry distribution in bin ik The uncertainty on $\langle\mathcal{A}_{ik}\rangle$ is taken to be the RMS$/\sqrt{N}$ of each distribution. For the data, N is the number of events in the bin, while for the Monte Carlo, N is the effective number of events calculated using the Monte Carlo event weights[1]. This can be seen in Figure 4.1.

To enhance events which have a $2 \to 2$ topology, the following selection criteria are applied:

$$p_{\mathrm{T}}^{\mathrm{avg}} > 20 \text{ GeV}, \quad \Delta\phi(j_1, j_2) > 2.6 \text{ rad}, \quad p_{\mathrm{T}}(j_3) < \max(0.15 p_{\mathrm{T}}^{\mathrm{avg}}, 7 \text{ GeV}) \tag{4.4}$$

where j_i denotes the ith highest p_{T} jet in the event and $\Delta\phi(j_1, j_2)$ is the azimuthal angle between the two leading jets. It should be noted that the lowest p_{T}-bins suffer from biases; firstly, the jet reconstruction efficiency is worse for low p_{T} jets; furthermore, there are larger soft corrections resulting from jets at around the jet reconstruction threshold of 7 GeV. The selection criterion on the third jet, which is used to suppress the unbalancing effects of soft radiation, is not as efficient here, since, for events with $p_{\mathrm{T}}^{\mathrm{avg}} = 20$ GeV, the strongest possible third jet rejection threshold, of

[1]The effective number of events is calculated as $N_{\mathrm{eff}} = (\sum w_i)^2 / \sum w_i^2$, where the sum is over all events, and w_i is the event weight for event i.

7 GeV, corresponds to 35% of $p_{\mathrm{T}}^{\mathrm{avg}}$. For events with $p_{\mathrm{T}}^{\mathrm{avg}} > 45$ GeV, rejections at the 15% level becomes possible

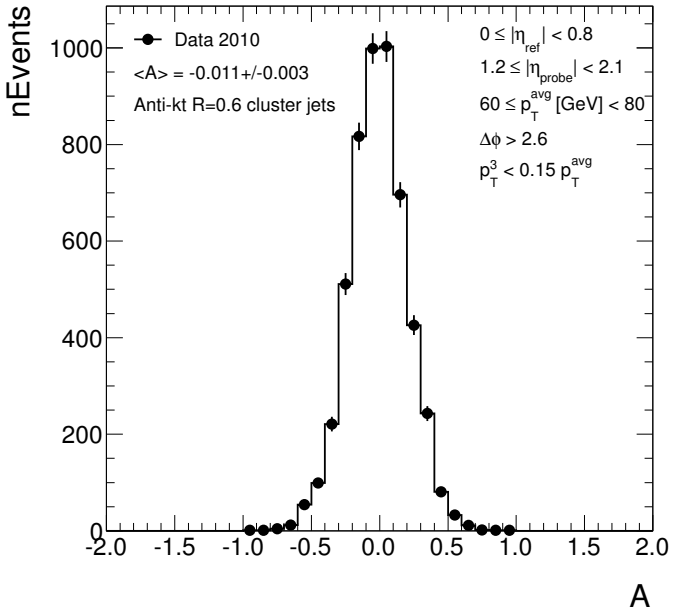


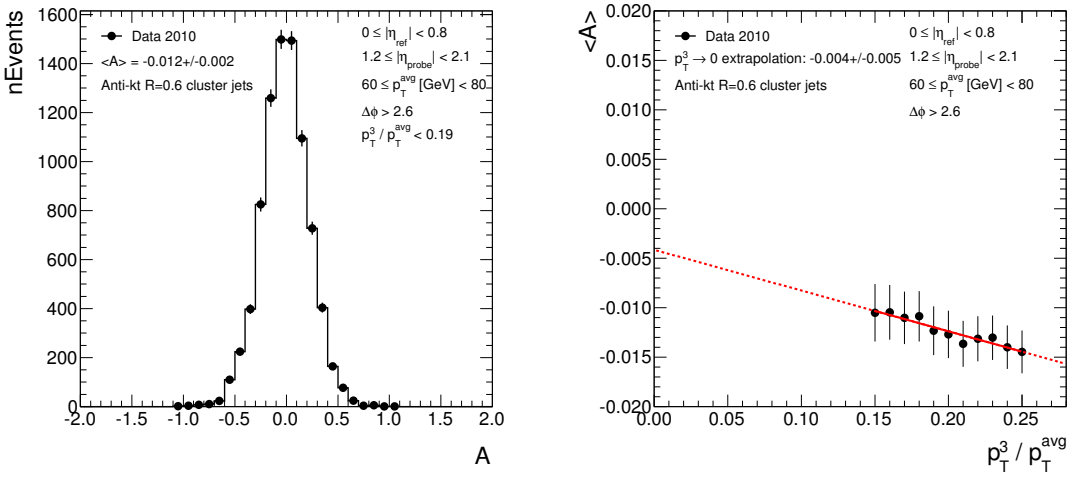
Figure 4.1.: A sample $\mathcal{A}$ distribution using anti-k_t jets with $R = 0.6$ and requiring that $\Delta\phi(j_1, j_2) > 2.6$ and $60 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 80$ GeV with the reference jet falling into the region $|\eta| < 0.8$ and the probe jet into the region $1.2 \leq |\eta| < 2.1$.

4.2.2. Intercalibration using a Central Reference Region with a Soft-Radiation Correction

As already discussed, a disadvantage with the method outlined above is that the effects of soft radiation are hard to quantify and hence hard to correct for. One solution to this is to apply a series of increasing cuts on the p_{T} of the third hardest jet in the event and to extrapolate the information to the case in which the third jet has zero p_{T}; in other words, the case in which there is no soft radiation in the event. As can be seen from Figure 4.2b, this can be done using a simple linear fit. The contents of each bin are, by definition, highly correlated, since each bin is a superset of the previous one. Accordingly, the best fit line was obtained by taking the covariance between the points into account and minimising:

$$\chi^2(\theta) = \sum_{i,j=1}^{n} (y_i - f(x_i,\theta))\, V_{ij}^{-1}\, (y_j - f(x_j,\theta)) \tag{4.5}$$

with respect to θ. Here $f(x_i,\theta)$ is a simple linear function in x_i, with θ representing the two free variables. The covariance matrix V is estimated using $V_{ij} = \frac{c_{ij}}{\sqrt{n_i n_j}}$, where c_{ij} is the number of events held in common between points i and j while n_i and n_j are the total number of events contributing to each of these points.



(a) $\mathcal{A}$ distribution for $p_{\mathrm{T}}(j_3)/p_{\mathrm{T}}^{\mathrm{avg}} < 0.19$ **(b)** Extrapolation to $p_{\mathrm{T}}(j_3) = 0$

Figure 4.2.: A sample $\mathcal{A}$ distribution using anti-k_t jets with $R = 0.6$ and requiring that $\Delta\phi(j_1, j_2) > 2.6$ and $60 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 80$ GeV, with the reference jet falling into the region $|\eta| < 0.8$ and the probe jet into the region $1.2 \leq |\eta| < 2.1$. (a) is the $\mathcal{A}$ distribution for events which additionally satisfy $p_{\mathrm{T}}(j_3)/p_{\mathrm{T}}^{\mathrm{avg}} < 0.19$, while (b) shows the extrapolation from several different cuts on $p_{\mathrm{T}}(j_3)/p_{\mathrm{T}}^{\mathrm{avg}}$ to the case in which $p_{\mathrm{T}}(j_3) = 0$.

4.3. Event Selection

As discussed in Section 3.3, events are required to belong to a good run and to have at least one good primary vertex. For this analysis, the vertex requirement is tightened: only events with exactly one good primary vertex are considered.

Events are required to possess at least two jets above the jet reconstruction threshold of 7 GeV. The event is rejected if either of the two leading jets are flagged as "bad" or "ugly" by the standard loose jet cleaning cuts, discussed in Section 3.3.4.

A trigger is assigned for each event, based on the run period and on $p_{\mathrm{T}}^{\mathrm{avg}}$ of the dijet pair. If this trigger is passed then the event is accepted. In later periods, instead of a single trigger per event, two triggers are assigned, one from the forward trigger system and one from the central trigger system. The thresholds are chosen such that the trigger efficiency for each specific region of $p_{\mathrm{T}}^{\mathrm{avg}}$ is greater than 99% and is approximately flat as a function of the pseudorapidity of the probe jet. To cover the low p_{T} region, $p_{\mathrm{T}} < 40$ GeV, as well as in early data taking when the trigger rates were low, triggers from the minimum bias stream are used. These require at least one hit in the Minimum Bias Trigger Scintillators (MBTS), which cover the region $2.08 \leq |\eta| < 3.8$. Forward triggers are not used before period E5, due to calibration problems.

Table 4.1 summarises the triggers used in the η-intercalibration measurement as a function of $p_{\mathrm{T}}^{\mathrm{avg}}$ and run period. As outlined in Section 3.3.6, the lowest prescaled fully efficient jet trigger in each $p_{\mathrm{T}}^{\mathrm{avg}}$ bin changes over time. All of these triggers are over 99% efficient in the relevant regions.

The Monte Carlo datasets which are compared to the data were generated using PYTHIA, HERWIG++, ALPGEN and the Perugia PYTHIA tune, using the parameters described in Section 1.11.

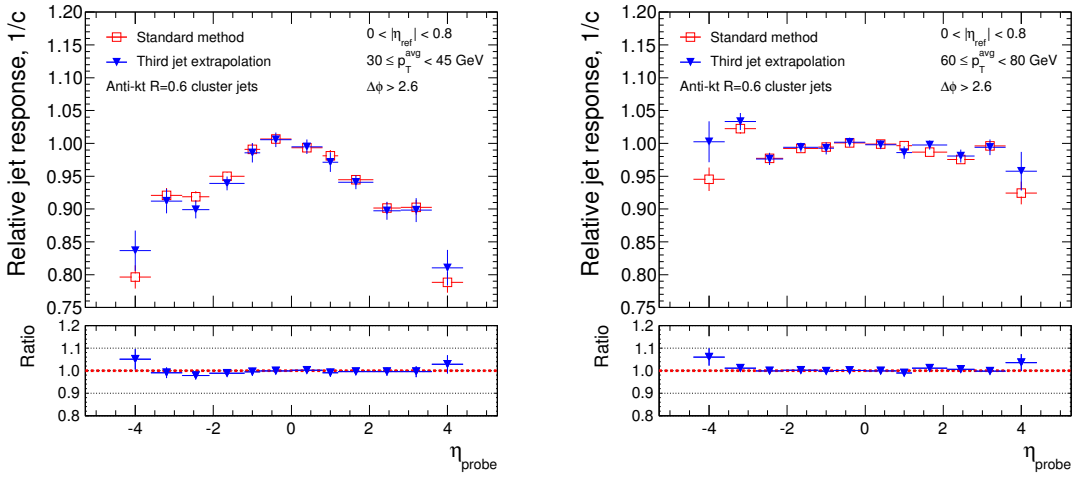
4.4. Dijet Balance Results

In this section, the relative jet response obtained with the extrapolation method is compared to the relative jet response obtained using the standard method with a fixed cut on maximum $p_{\mathrm{T}}(j_3)$. Figure 4.3 shows the jet response relative to central jets, c, for two $p_{\mathrm{T}}^{\mathrm{avg}}$-bins: $30 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 45$ GeV and $60 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 80$ GeV. These results indicate that the response observed using the extrapolation method is compatible with that obtained using the standard method. The results shown here are representative of all the phase space regions considered in this analysis; the extrapolation method is therefore used to give the final results for data due to its more advanced treatment

$p_{\mathrm{T}}^{\mathrm{avg}}$ range [GeV]	Period A*	Period B–D	Period E1–4
20–40	L1_MBTS_1	L1_MBTS_1	L1_MBTS_1
40–50	L1_MBTS_1	L1_J5	L1_J5
50–110	L1_J10	L1_J10	L1_J10
110–160	L1_J30	L1_J30	L1_J30
160+	L1_J55	L1_J55	L1_J55
$p_{\mathrm{T}}^{\mathrm{avg}}$ range [GeV]	Period E5–F	Period G–H	
20–40	L1_MBTS_1	EF_mbMbts_1_eff	
40–50	L1_J5	EF_mbMbts_1_eff	
50–110	L1_J10 or L1_FJ10	EF_j30_jetNoEF or EF_fj30_jetNoEF	
110–160	L1_J30 or L1_FJ30	EF_j50_jetNoEF or EF_fj50_jetNoEF	
160+	L1_J55 or L1_FJ30	EF_j75_jetNoEF or EF_fj75_jetNoEF	

Table 4.1.: The trigger chains used for the η-intercalibration analysis. The forward jet trigger could not be used in the first four periods (A–D) as it had not yet been commissioned, while additional problems made it unreliable for subperiods E1–4. L1_MBTS_1 was also used to trigger all jets before run 152777. The period after this timing change is denoted here as "A*".

of residual soft effects; in the forward region, where statistics in some Monte Carlo datasets are poor, this unfortunately results in large errors.

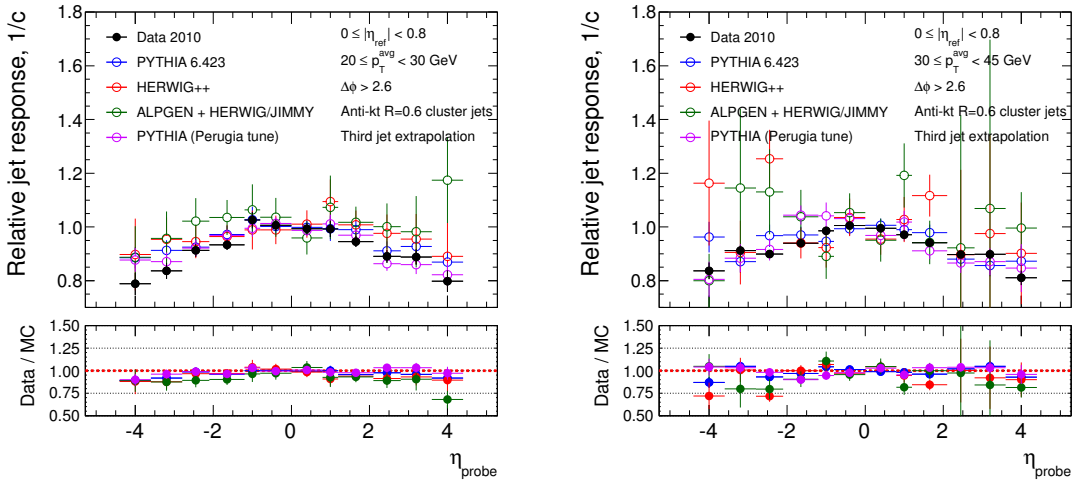


(a) $30 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 45$ GeV **(b)** $60 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 80$ GeV

Figure 4.3.: Relative jet response, $1/c$, as a function of the pseudorapidity of the probe jet. Results are presented for two bins of $p_{\mathrm{T}}^{\mathrm{avg}}$: (a) $30 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 45$ GeV and (b) $60 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 80$ GeV.

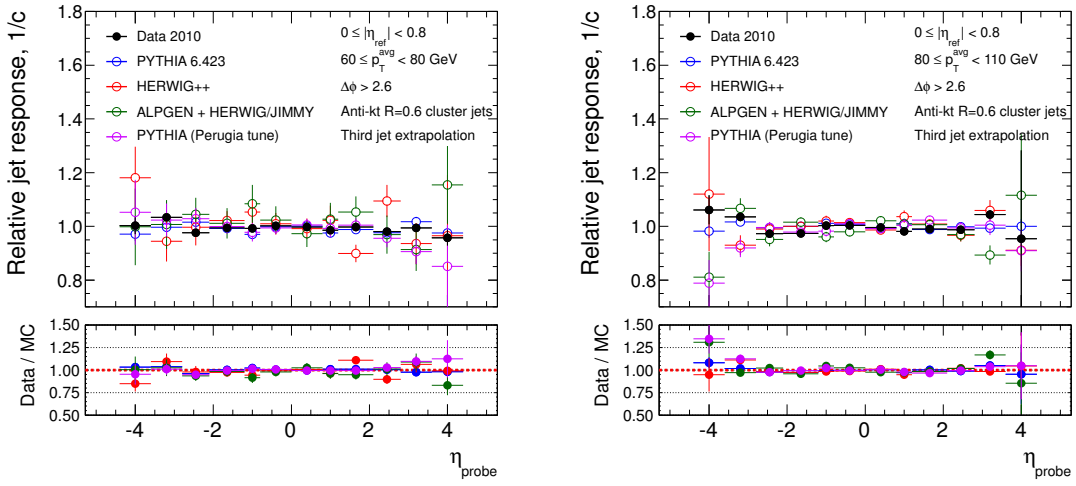
Figure 4.4 shows the relative response obtained with the extrapolation method as a function of the jet pseudorapidity for data and the Monte Carlo event generator simulations. Four different $p_{\mathrm{T}}^{\mathrm{avg}}$ regions are shown: $20 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 30$ GeV, $30 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 45$ GeV, $60 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 80$ GeV and $80 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 110$ GeV. The response in data is reasonably well reproduced by the Monte Carlo simulations for $p_{\mathrm{T}} > 60$ GeV, with the Monte Carlo and data typically agreeing at better than the 2% level in the central region ($|\eta| < 2.8$) and better than 5% in the forward region ($|\eta| > 2.8$). At lower values of $p_{\mathrm{T}}^{\mathrm{avg}}$, the data do not agree so well with the Monte Carlo simulations and the Monte Carlo simulations themselves show a large spread around the data. For $20 \leq p_{\mathrm{T}} < 30$ GeV, the Monte Carlo deviates from the data by about 10% for $|\eta| > 2.8$, with the different Monte Carlo simulations predicting both higher and lower responses than that observed in the data. The main differences, due to residual low p_{T} jet effects (see Section 4.2.1), occur between PYTHIA/Perugia and ALPGEN/HERWIG++. The reason is that the PYTHIA and Perugia predictions are based upon a p_{T}-ordered parton shower, Lund String hadronisation and the PYTHIA underlying event model, whereas the HERWIG++ and ALPGEN predictions are based on an angular-ordered parton shower, cluster hadronisation and the JIMMY

underlying event model. These differences therefore reflect a difference in physics modelling among the event generators.



(a) $20 \le p_{\mathrm{T}}^{\mathrm{avg}} < 30$ GeV

(b) $30 \le p_{\mathrm{T}}^{\mathrm{avg}} < 45$ GeV

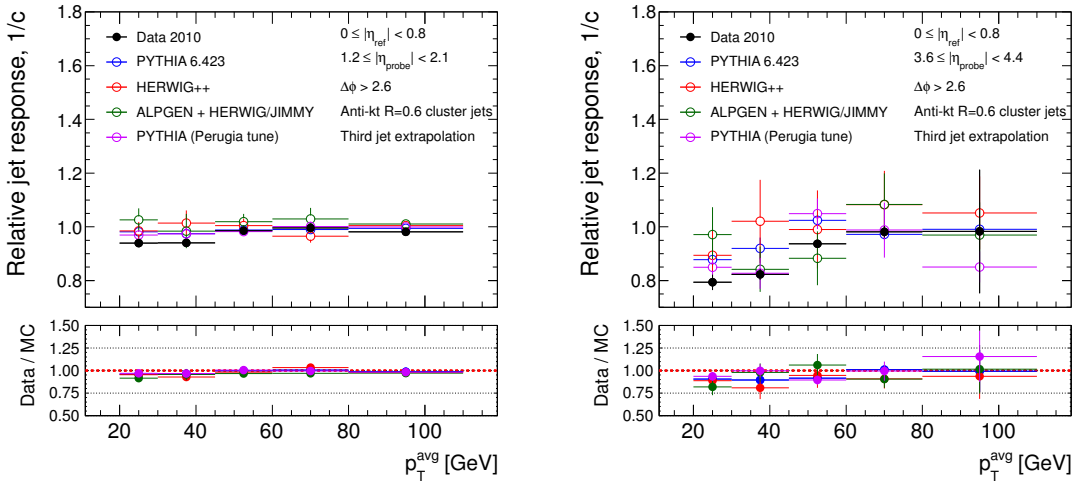


(c) $60 \le p_{\mathrm{T}}^{\mathrm{avg}} < 80$ GeV

(d) $80 \le p_{\mathrm{T}}^{\mathrm{avg}} < 110$ GeV

Figure 4.4.: Relative jet response, $1/c$, as a function of the η of the probe jet. Results are presented for four bins of $p_{\mathrm{T}}^{\mathrm{avg}}$: (a) $20 \le p_{\mathrm{T}}^{\mathrm{avg}} < 30$ GeV, (b) $30 \le p_{\mathrm{T}}^{\mathrm{avg}} < 45$ GeV, (c) $60 \le p_{\mathrm{T}}^{\mathrm{avg}} < 80$ GeV and (d) $80 \le p_{\mathrm{T}}^{\mathrm{avg}} < 110$ GeV.

Figure 4.5 shows the relative response as a function of $p_{\mathrm{T}}^{\mathrm{avg}}$. The distributions are shown for jets in the region $1.2 \le |\eta| < 2.1$ and also for those in the region $3.6 \le |\eta| < 4.4$. Again, the response is reasonably well described by the Monte Carlo for all calorimeter regions at high $p_{\mathrm{T}}^{\mathrm{avg}}$ and for the more central region at low $p_{\mathrm{T}}^{\mathrm{avg}}$.



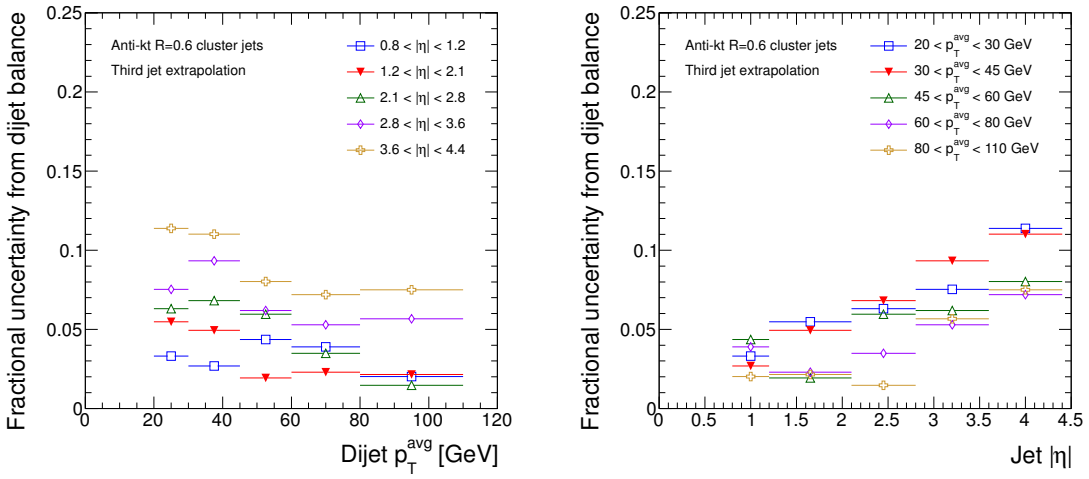
(a) $1.2 \leq |\eta| < 2.1$ **(b)** $3.6 \leq |\eta| < 4.4$

Figure 4.5.: Relative jet response, $1/c$, as a function of the η of the probe jet. For low $p_{\mathrm{T}}^{\mathrm{avg}}$ and early data periods, the data is collected using the minimum bias trigger stream. For higher $p_{\mathrm{T}}^{\mathrm{avg}}$, the data is collected using the calorimeter trigger stream. Results are presented for two bins of $|\eta|$: (a) $1.2 \leq |\eta| < 2.1$ and (b) $3.6 \leq |\eta| < 4.4$.

4.5. Uncertainty due to Intercalibration

In the previous section it was shown that the Monte Carlo predictions for the relative jet response diverge at low values of $p_{\mathrm{T}}^{\mathrm{avg}}$, with the data itself lying between the different predictions for central values of η. The uncertainty on the relative jet response must reflect this disagreement because there is no *a priori* reason to believe one theoretical prediction over another. The uncertainty on the relative response is taken to be the RMS deviation of the Monte Carlo predictions from the data. At high $p_{\mathrm{T}}^{\mathrm{avg}}$, where the spread of the Monte Carlo predictions is small, the uncertainty mainly reflects the true difference between the response in data and simulation. At low $p_{\mathrm{T}}^{\mathrm{avg}}$ and large $|\eta|$, the uncertainty mainly reflects the physics modelling uncertainty, although the detector-based differences between data and simulation are also accounted for. The RMS spread of Monte Carlo predictions around the data measurement is less sensitive to statistical fluctuations in comparison to other possible measures of the deviation, such as the maximal difference between the predictions or between data and Monte Carlo. The latter quantities tend to give unstable results since all of the Monte Carlo samples used in the analysis, aside from PYTHIA, are generated with low statistics (for details see Section 1.11); Monte Carlo

generators with fewer than ten entries in a particular bin of $|\eta_{\text{probe}}|$ and $p_{\text{T}}^{\text{avg}}$ are excluded from that bin of the RMS calculation for this reason.



(a) Uncertainty in the jet response as a function of dijet $p_{\text{T}}^{\text{avg}}$ in five regions of the calorimeter.

(b) Uncertainty in the jet response as a function of jet $|\eta|$ for five values of dijet $p_{\text{T}}^{\text{avg}}$.

Figure 4.6.: Fractional uncertainty in the relative jet response, $1/c$, arising from dijet balance as a function of (a) the $p_{\text{T}}^{\text{avg}}$ and (b) the η of the dijet system.

Figure 4.6 shows the uncertainty in the jet response, relative to jets in the region $|\eta| < 0.8$, as a function of the dijet $p_{\text{T}}^{\text{avg}}$ and $|\eta|$.

4.6. Summary

These results indicate that the relative response to jets with $|\eta| \leq 0.8$ is well understood both for high p_{T} jets across the full calorimeter range and for central jets across the p_{T} range considered. Deviations between data and Monte Carlo are at their smallest for high p_{T} central jets, at about 3%, and at their worst for low p_{T} forward jets, at about 12%. Additionally, different Monte Carlo generators show a large spread of predictions in the low p_{T} region, reflecting a real physics modelling uncertainty. This work was published as an ATLAS conference note [54] and also formed one of the inputs for the evaluation of the jet energy scale uncertainty [9] (see Section 3.3.5 for details) which was submitted to European Physical Journal C [10]. The results shown here are in agreement with these published results, which

use a slightly different method of determining the response factors, $1/c$, providing increased statistical precision.

Chapter 5.

Inclusive Jet Cross-sections

"It's not exclusive, but inclusive, which is the whole spirit of jazz."
— Herbie Hancock

5.1. Introduction

As discussed in Section 1.7, jet production is the dominant high p_{T} process at the LHC. Jet cross-sections are important observables in high-energy particle physics and are one of the first measurements that can be performed in a hadron collider experiment. These cross-sections are an important tool for understanding the strong interaction, in particular through providing precise measurements of α_{S}. In addition to testing QCD in a new kinematic regime, the data also provide sensitivity to different parton distribution functions (PDFs) in a region where they are currently poorly constrained.

This analysis presents inclusive double-differential jet cross-sections; studied as a function of jet p_{T} and y. The measurement covers 20 GeV $\leq p_{\mathrm{T}} <$ 1.5 TeV in the rapidity range $|y| < 4.4$. The results are compared to theoretical predictions from next-to-leading order QCD corrected for non-perturbative effects, as well as to next-to-leading order Monte Carlo simulations. The rapidity slices used in the measurement have boundaries which follow the calorimeter geometry.

5.2. Cross-section Definition

Clearly, jet cross-sections can only be defined for a specific jet algorithm; here cross-sections have been calculated for the ATLAS standard: anti-k_t jets with $R = 0.4$ and $R = 0.6$ (see Section 1.10.2 and Section 3.2). The jet cross-section measurements are corrected for all experimental effects (see Section 3.4), and so refer to the ideal particle level final state of a proton-proton collision [55]. In the Monte Carlo, particle level jets are identified using the same jet algorithms applied to data, using stable particles, all those with a proper lifetime longer than 10 ps, as their input. This definition includes muons and neutrinos from decaying hadrons which cannot be included in calorimeter jets reconstructed in data.

5.3. Event Selection

Following the outline in Section 3.3, events are required to belong to a good run and to have at least one good primary vertex. All jets in the event which satisfy $p_{\mathrm{T}} \geq 20$ GeV and $|y| < 4.4$ are retained unless they are flagged as "bad" or "ugly" by the standard medium jet cleaning cuts (see Section 3.3.4). Each jet is then assigned a specific, fully-efficient trigger which depends on its transverse momentum and rapidity, as well as the run period that the event belongs to. For the jet to contribute towards the inclusive cross-section, this trigger must be passed. All jets which pass their trigger are retained, with a weight reflecting the amount of luminosity seen by the trigger, which, due to prescaling, is usually less than then the total luminosity recorded by ATLAS.

5.4. Trigger Strategy

This measurement uses three different trigger systems (see Section 2.4): the Minimum Bias Trigger Scintillators (MBTS) are used to select events containing jets with $20 \leq p_{\mathrm{T}} < 60$ GeV, the central jet triggers are used for the remainder of the phase space region $|y| < 3.6$ while the forward jet triggers are used for jets with $3.2 \leq |y| < 4.9$. In the region $3.2 \leq |y| < 3.6$ a combination of central and forward jet

	Period A*–F		Period G–I	
p_{T} range [GeV]	Central $\lvert y\rvert < 2.8$ except crack	Central crack $1.2 \leq \lvert y\rvert < 2.1$	Central $\lvert y\rvert < 2.8$ except crack	Central crack $1.2 \leq \lvert y\rvert < 2.1$
60–80	L1_J5	L1_J5	EF_j20_jetNoEF	EF_j20_jetNoEF
80–110	L1_J15	L1_J5	EF_j35_jetNoEF	EF_j20_jetNoEF
110–160	L1_J30	L1_J15	EF_j50_jetNoEF	EF_j35_jetNoEF
160–210	L1_J55	L1_J30	EF_j75_jetNoEF	EF_j50_jetNoEF
210–260	L1_J75	L1_J55	EF_j95_jetNoEF	EF_j75_jetNoEF
260–310	L1_J95	L1_J75	EF_L1J95_NoAlg	EF_j95_jetNoEF
310–400	L1_J95	L1_J95	EF_L1J115_NoAlg	EF_L1J95_NoAlg
400+	L1_J95	L1_J95	EF_L1J115_NoAlg	EF_L1J115_NoAlg

Table 5.1.: The trigger chains used for the inclusive jet analysis in the region $|y| < 2.8$. The L1_MBTS_1 trigger is used for the $20 \leq p_{\mathrm{T}} < 60$ GeV over the range $|y| < 2.8$. Due to mistimings in the Level-1 central jet trigger hardware, L1_MBTS_1 was also used to trigger all jets before run 152777. The period after this timing change is here denoted as "A*".

triggers is used. The L1_MBTS_1 trigger has been demonstrated to have a negligible inefficiency in selecting the low p_{T} events for which it is used here.

For this measurement, the rapidity interval $|y| < 4.4$ was divided into seven bins: the central bins with $|y| < 2.8$, the HEC-FCAL transition bin with $2.8 \leq |y| < 3.6$ and the forward bin with $|y| \geq 3.6$. As discussed in Section 3.3.6, increasing instantaneous luminosity throughout 2010 means that the lowest prescaled fully efficient trigger in each p_{T} bin changes over time.

Tables 5.1–5.2 summarise the triggers that are used in each p_{T} bin of the cross-section measurement as a function of the data taking period. For the central bins, these are central jet triggers and for the forward bin, forward jet triggers. In both cases, certain bins are augmented by data taken using the L1_MBTS_1 trigger. It is confirmed that all of these triggers are over 99% efficient in these regions.

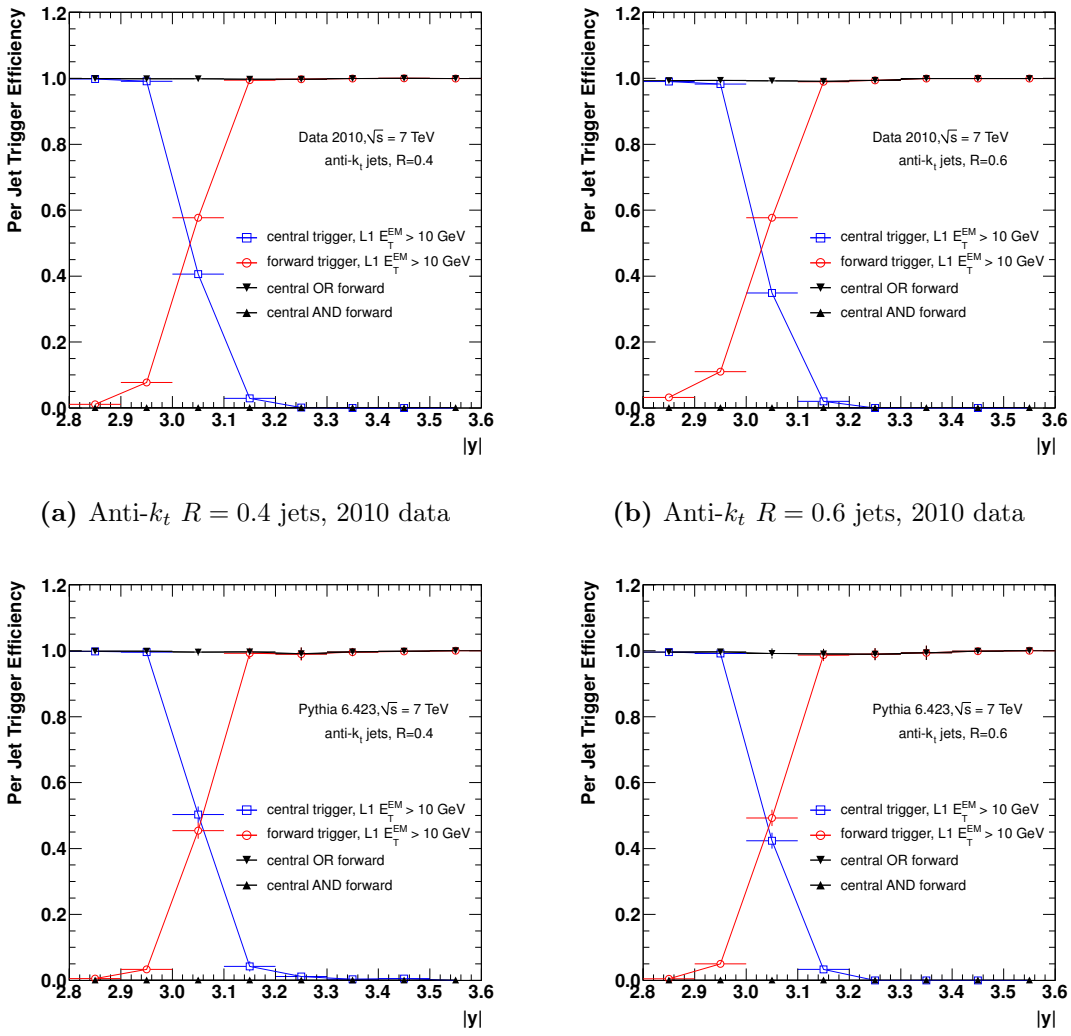
p_T range [GeV]	Period A–C	Period E5–F	Period G–I
20–30	L1_MBTS_1	n/a	n/a
30–45	n/a	L1_FJ10	n/a
45–60	n/a	L1_FJ10	n/a
60–80	n/a	L1_FJ10	EF_fj30_jetNoEF
80–110	n/a	L1_FJ30	EF_fj30_jetNoEF
110–160	n/a	L1_FJ55	EF_fj50_jetNoEF
160–210	n/a	L1_FJ55	EF_fj75_jetNoEF
210–260	n/a	L1_FJ55	EF_fj75_jetNoEF
260+	n/a	L1_FJ55	EF_fj75_jetNoEF

Table 5.2.: The trigger chains used for the inclusive jet analysis in the forward region, $3.6 \leq |y| < 4.4$. The first four periods (A–D) could not be used, as the forward jet trigger had not yet been commissioned, while additional problems made it unreliable for subperiods E1–4. L1_MBTS_1 was found to be fully efficient for forward jets and hence was used in early periods to trigger low p_T forward jets.

5.4.1. Trigger Strategy in the Transition Region

A jet is accepted as having passed its particular inclusive trigger requirement if at least one jet in the event has passed that trigger. Due to the reduced η-granularity available in the trigger system, it is therefore possible for jets in the forward region to belong to an event in which only central trigger chains have fired, or *vice versa*. Events of this type will predominantly fall into the HEC-FCAL transition region, $2.8 \leq |y| < 3.6$. Additionally, it is possible that jets in this region might trigger both the central and forward trigger systems. It can be seen in Figure 5.1 that, when taking central and forward triggers with the same E_T^{EM} threshold, the logical OR of the two triggers is fully efficient across this transition region although neither trigger is on its own.

Sample turn-on curves for the central OR forward combination can be seen in Figure 5.2, demonstrating that this combination is fully efficient at sufficiently high jet p_T. Table 5.3 summarises the triggers that are used in each p_T bin of the cross-section measurement as a function of the data taking period.



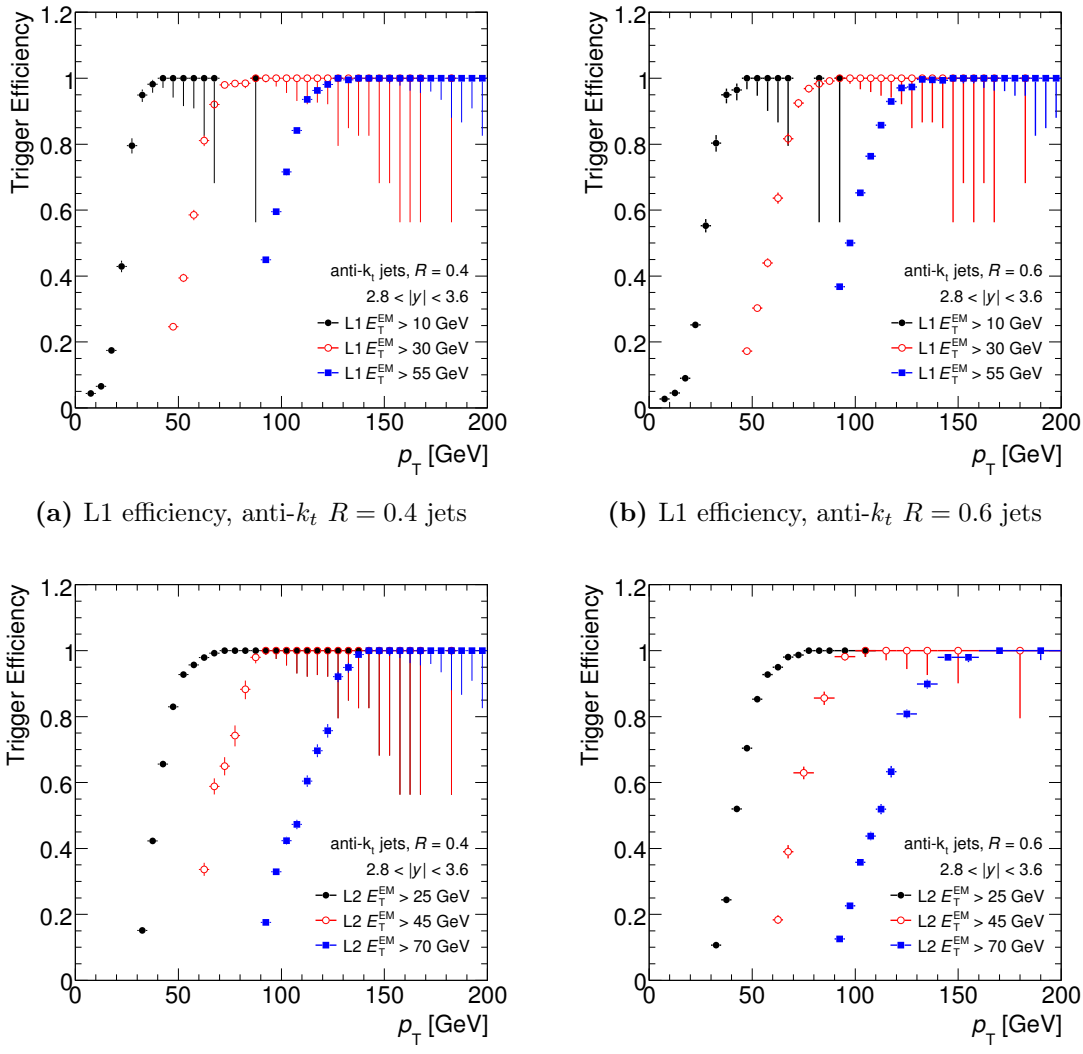
(a) Anti-k_t $R = 0.4$ jets, 2010 data

(b) Anti-k_t $R = 0.6$ jets, 2010 data

(c) Anti-k_t $R = 0.4$ jets, PYTHIA

(d) Anti-k_t $R = 0.6$ jets, PYTHIA

Figure 5.1.: Efficiency of the central and forward trigger signatures J10 and FJ10, their AND and their OR, as a function of rapidity, y, of the offline jet. Trigger efficiency is shown in data (top) and in Monte Carlo (bottom) while anti-k_t $R = 0.4$ jets (left) and anti-k_t $R = 0.6$ jets (right) are also considered separately.



(a) L1 efficiency, anti-k_t $R = 0.4$ jets

(b) L1 efficiency, anti-k_t $R = 0.6$ jets

(c) L1+L2 efficiency, anti-k_t $R = 0.4$ jets

(d) L1+L2 efficiency, anti-k_t $R = 0.6$ jets

Figure 5.2.: Jet trigger efficiency in the HEC-FCAL transition region $2.8 \leq |y| < 3.6$ at L1 (top) and combined L1+L2 (bottom). Trigger efficiencies are presented as a function of reconstructed jet p_{T} for anti-k_t jets with $R = 0.4$ (left) and $R = 0.6$ (right), shown for three trigger thresholds. The trigger thresholds are at the electromagnetic scale, while the jet p_{T} is at the calibrated scale (see Section 3.2).

p_T range [GeV]	Period A–C	Period E5–F
20–30	L1_MBTS_1	n/a
30–45	L1_MBTS_1	n/a
45–60	L1_MBTS_1	n/a
60–80	n/a	L1_J10 or L1_FJ10
80–110	n/a	L1_J10 or L1_FJ10
110–160	n/a	L1_J30 or L1_FJ30
160–210	n/a	L1_J55 or L1_FJ55
210–260	n/a	L1_J55 or L1_FJ55
260+	n/a	L1_J55 or L1_FJ55

p_T range [GeV]	Period G–I
20–30	n/a
30–45	n/a
45–60	n/a
60–80	n/a
80–110	EF_j30_jetNoEF or EF_fj30_jetNoEF
110–160	EF_j50_jetNoEF or EF_fj50_jetNoEF
160–210	EF_j50_jetNoEF or EF_fj50_jetNoEF
210–260	EF_j75_jetNoEF or EF_fj75_jetNoEF
260+	EF_j75_jetNoEF or EF_fj75_jetNoEF

Table 5.3.: The trigger chains used for the inclusive jet analysis in the transition region, $2.8 \leq |y| < 3.6$. The first four periods (A–D) could not be used, as the forward jet trigger had not yet been commissioned, while additional problems made it unreliable for subperiods E1–4. L1_MBTS_1 was found to be fully efficient for transition jets and hence was used in early periods to trigger low p_T jets here.

Period	Jet $\|y\|$	Jet p_{T} [GeV]	L1_J10	L1_FJ10	Event accepted?
F	3.0	30	passed	passed	no
F	3.0	70	passed	passed	yes
F	3.0	70	failed	passed	yes
F	3.0	70	passed	failed	yes
F	2.6	70	failed	passed	no
F	3.8	70	failed	passed	yes

Table 5.4.: Sample table demonstrating trigger decisions and event acceptance for a series of possible jets in period F.

Table 5.4 provides a series of example trigger decisions for jets in or around the transition region, according to their p_{T} and which triggers are passed. For the sake of this example, all of these imaginary events are considered to belong to period F.

To avoid double counting in taking the OR of central and forward chains, these events have to be considered separately with respect to those only passing a central or a forward threshold.

Consider the set of all events which pass the OR of the central and forward chain, in other words, those which are taken by either the central trigger, forward trigger or both at once. In these events, which are already triggered and accepted, it is possible to examine the complete set trigger information; in particular it can be determined which triggers each of these events would have passed *before prescale was applied.*

Triggered jets in the transition bin can therefore be divided into three classes, according to whether they would have passed, before prescale, the central threshold, the forward threshold or both. For each of these classes, an equivalent luminosity is calculated, summing over all luminosity blocks. The luminosity for jets selected by the central trigger is given by

$$\mathcal{L}_{\mathrm{J}} = \sum_{\mathrm{LB}} \frac{\mathcal{L}_{\mathrm{LB}}}{P^{\mathrm{J}}_{\mathrm{LB}}} \tag{5.1}$$

where $P^{\mathrm{J}}_{\mathrm{LB}}$ is the prescale of the central trigger for each luminosity block, and $\mathcal{L}_{\mathrm{LB}}$ its luminosity; for jets selected by the forward trigger the equivalent luminosity will be

$$\mathcal{L}_{\mathrm{FJ}} = \sum_{\mathrm{LB}} \frac{\mathcal{L}_{\mathrm{LB}}}{P^{\mathrm{FJ}}_{\mathrm{LB}}} \tag{5.2}$$

Finally, for events taken, before prescale, by both central and forward trigger, the equivalent luminosity is

$$\mathcal{L}_{\mathrm{JFJ}} = \sum_{\mathrm{LB}} \frac{\mathcal{L}_{\mathrm{LB}}}{P^{\mathrm{J}}_{\mathrm{LB}} P^{\mathrm{FJ}}_{\mathrm{LB}} / (P^{\mathrm{J}}_{\mathrm{LB}} + P^{\mathrm{FJ}}_{\mathrm{LB}} - 1)} \tag{5.3}$$

Let N_{J}, N_{FJ} and N_{JFJ} denote respectively the number of events taken, by the central trigger, by the forward trigger, and by both triggers. The cross-section, before any other correction, is then given by

$$\sigma = \frac{N_{\mathrm{J}}}{\mathcal{L}_{\mathrm{J}}} + \frac{N_{\mathrm{FJ}}}{\mathcal{L}_{\mathrm{FJ}}} + \frac{N_{\mathrm{JFJ}}}{\mathcal{L}_{\mathrm{JFJ}}} \tag{5.4}$$

ensuring that events passing two triggers are properly treated in a separate category and not double-counted.

5.4.2. Trigger Efficiencies

In general, the trigger strategy has been designed to ensure that all jets are on the 99% plateau; however, due to a known problem with a dead trigger tower in a region of the FCAL, the efficiency of some forward jet triggers reaches a plateau at less than 100%.

The effects of these inefficiencies are small, less than 5%, which is helped by the fact that a per-event efficiency definition is used, so that an offline jet falling into a lower efficiency region can still be accepted if there is another jet in the

event. Accordingly, a systematic uncertainty is applied rather than restricting the phase-space of the measurement. Figure 2.8 shows the trigger efficiency in the most forward rapidity region, $3.6 \leq |y| < 4.4$, where the jets are fully contained by the FCAL. Irregularities in the trigger plateau arising from the problematic trigger tower can be seen here.

5.5. Unfolding Detector Effects

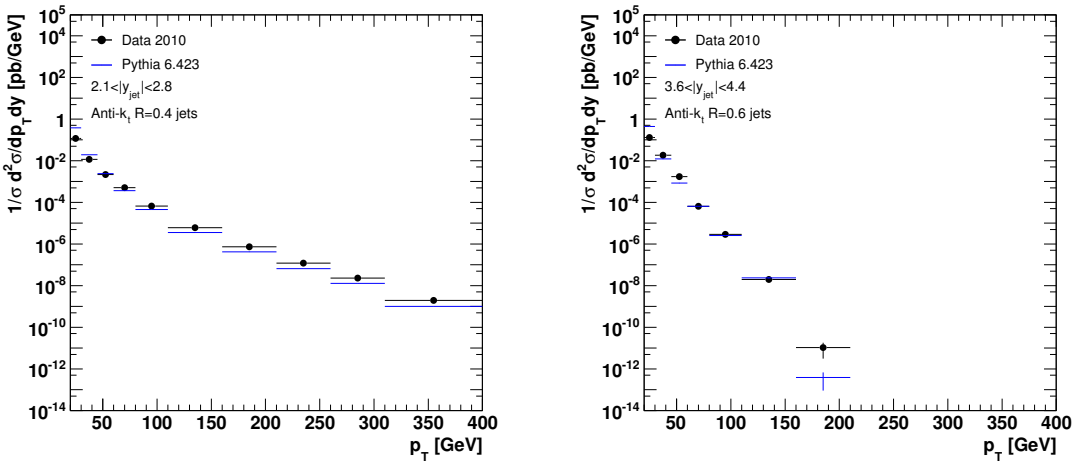
Detector inefficiencies and resolutions, apart from those corrected for in by the jet calibration scheme, are corrected for using an iterative unfolding method: the IDS scheme detailed in Section 3.4. The same binning is used as for the final distributions and the unfolding is performed separately for each rapidity bin.

A detector level cross-section is constructed in each case, by combining the set of events passing each trigger and correcting, in each case, for the appropriate integrated luminosity by that trigger. This detector level cross-section is then used as the input for the unfolding procedure.

The effect of any potential mismodelling of the cross-section shape in Monte Carlo is examined by comparing the shapes of detector level spectra in Monte Carlo and in data. These are used to derive an event-by-event reweighting function that is applied to the Monte Carlo particle level spectra, in each bin of y and p_{T}, before the unfolding factors are calculated. Figure 5.3 shows the level of agreement between PYTHIA and the data for two sample distributions.

5.6. Systematic Uncertainties

Track jets, reconstructed using only information from the tracking detector, are used to provide an in-situ estimate of jet reconstruction efficiency. The frequency with which a calorimeter jet was reconstructed, given the existence of a nearby track jet, was studied in data and Monte Carlo and used to infer an uncertainty on the calorimeter jet reconstruction efficiency measured in Monte Carlo. This disagreement of 1%, or 2% for the lowest p_{T} jets, was taken as a systematic uncertainty.



(a) Data to Monte Carlo comparison for anti-k_t $R = 0.4$ jets, in the region $2.1 \leq y < 2.8$

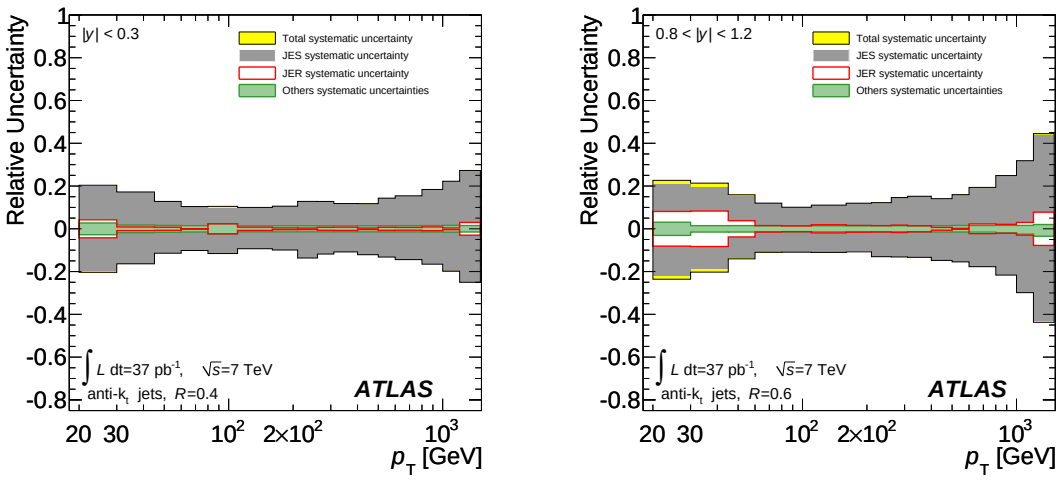
(b) Data to Monte Carlo comparison for anti-k_t $R = 0.6$ jets, in the region $3.6 \leq y < 4.4$

Figure 5.3.: Control distributions, used to demonstrate the level of agreement between data and Monte Carlo; data is shown in black, with PYTHIA in blue. (a) shows the comparison for anti-k_t $R = 0.4$ jets in the region $2.1 \leq y < 2.8$, while (b) shows the comparison for anti-k_t $R = 0.6$ jets in the region $3.6 \leq y < 4.4$.

The jet energy scale uncertainty, evaluated as described in Section 3.3.5, is the largest single contributor to systematic uncertainties due to the steeply falling cross-section as a function of jet p_{T}. The JES uncertainty was treated in the same way as energy resolution, measured from dijet balancing in data, and angular resolution, estimated in Monte Carlo by matching particle level and detector level jets; each of these quantities was varied up and down by one standard deviation and the relative per-bin shift in jet yield was taken as a systematic uncertainty for that bin. In the central region, $|y| < 0.8$, the uncertainty is lower than 4.6% for all jets with $p_{\mathrm{T}} \geq 20$ GeV, while for jets with $60 \leq p_{\mathrm{T}} < 800$ GeV the uncertainty is below 2.5%, as can be seen from Figure 3.2.

By propagating these uncertainties through the unfolding procedure, an estimate of the systematic uncertainty due to unfolding can also be obtained. This is approximately 5% at low and high p_{T} and smaller at intermediate p_{T} values. An additional uncertainty of 3.4% comes from the luminosity measurement.

Systematic uncertainties on the final cross-section are obtained by summing all uncertainties in quadrature to give a total uncertainty on the unfolded data; an example of this process for two $|y|$ bins can be seen in Figure 5.4. However, the



(a) Anti-k_t $R = 0.4$ jets, $|y| < 0.3$ **(b)** Anti-k_t $R = 0.6$ jets, $0.8 \leq |y| < 1.2$

Figure 5.4.: Summary of relative systematic effects affecting the inclusive jets cross-section. (a) shows the uncertainty for anti-k_t $R = 0.4$ jets in the region $|y| < 0.3$ while (b) shows the corresponding uncertainty for anti-k_t $R = 0.6$ jets in the region $0.8 \leq |y| < 1.2$. The jet energy scale uncertainty provides the dominant systematic in both cases.

steeply falling jet p_{T} spectrum, especially at large rapidity, unavoidably converts even small uncertainties in p_{T} into large errors on the measured cross-section.

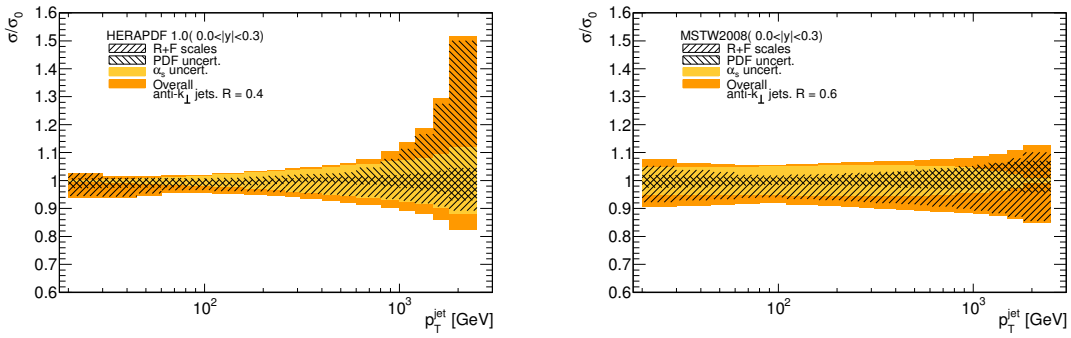
5.7. Theoretical Predictions

The measured inclusive jet cross-sections are compared to both NLO pQCD predictions, with corrections for non-perturbative effects, and to NLO Monte Carlo.

5.7.1. Next-to-Leading Order Perturbative QCD Calculations

Next-to-leading order (NLO) perturbative QCD (pQCD) predictions are produced using the NLOJET++ 4.1.2 [56] program together with the CT10 [57] NLO PDFs. The main uncertainties on the NLO prediction come from the uncertainties on the PDFs, the choice of factorisation and renormalisation scales (as discussed in Section 1.5) and the uncertainty in the value of the strong coupling constant, α_{S}.

To estimate the uncertainty on the NLO prediction due to neglected higher-order terms, each observable was recalculated while varying the renormalisation scale by a factor of two with respect to the default choice, defined to be the p_{T} of the hardest jet in the event. Similarly, to estimate the sensitivity to the choice of scale where the PDF evolution is separated from the matrix element, the factorisation scale was separately varied by a factor of two. The experimental uncertainties which propagate through the PDF fits, together with the associated uncertainties on the value of $\alpha_{\mathrm{S}}(M_Z)$ were used to determine uncertainty bands on the theoretical predictions; a summary of these corrections can be seen in Figure 5.5. Uncertainties due to the choice of PDF were evaluated by creating predictions using four different PDF sets and presenting each of these on the final plots.



(a) Anti-k_t $R = 0.4$ jets, $|y| < 0.3$ **(b)** Anti-k_t $R = 0.6$ jets, $|y| < 1.2$

Figure 5.5.: Summary of theory uncertainties arising from the scale choice, uncertainties inherent to the chosen PDF and uncertainties arising from the value of α_{S}. (a) shows the ratio of the cross-section to the nominal cross-section for HERAPDF, using anti-k_t $R = 0.4$ jets in the region $|y| < 0.3$ while (b) shows the same quantity for MSTW2008, using anti-k_t $R = 0.6$ jets in the region $|y| < 0.3$.

Parton level cross-sections, obtained from fixed-order NLO calculations must be corrected for non-perturbative effects before they can be compared with data. This is done by using leading-logarithmic parton shower generators (in this case, PYTHIA with the AUET2B CTEQ6L1 tune [39]) to evaluate the ratio of cross-sections with and without hadronisation and underlying event. The parton level cross-sections are then multiplied, bin-by-bin, by this ratio; tacitly assuming that the effects of soft and hard physics can be factorised. The uncertainty is estimated as the maximum spread of the correction factors obtained from performing this procedure using different PYTHIA tunes.

5.7.2. Next-to-leading Order Monte Carlo with Parton Shower

The POWHEG generator (see Section 1.11.4) is used to provide an NLO matrix element prediction. The use of an event generator with NLO matrix elements, including the simulation of the parton shower, the hadronisation, and the underlying event, creates a more coherent theoretical prediction and overcomes the need for separate non-perturbative corrections.

5.8. Inclusive Jet Cross-sections

The double-differential inclusive jet cross-section is shown in Figure 5.6 and Figure 5.7 for jets reconstructed using the anti-k_t algorithm with $R = 0.4$ and $R = 0.6$ respectively. The measurement covers the jet p_{T} range from 20 GeV to 1.5 TeV: spanning two orders of magnitude in p_{T} and seven orders of magnitude in cross-section. Statistical and systematic errors on the data are shown, as discussed in Section 5.6, and the unfolded data is compared to NLO pQCD predictions which are corrected for non-perturbative effects as discussed in Section 5.7.1. It can be seen from Figure 5.8 that the data and the theory predictions are generally in good agreement within the experimental and theoretical uncertainties, with some minor differences visible at high jet p_{T} and $|y|$.

The comparison of the data with the POWHEG prediction, using the CT10 PDF set, is shown for anti-k_t jets with $R = 0.4$ and $R = 0.6$ in different rapidity regions in Figure 5.9. Here the data, POWHEG predictions interfaced either PYTHIA (AUET2B and Perugia 2010 tunes) or HERWIG (AUET2 tune) as well as fixed order POWHEG with non-perturbative corrections are all compared. The ratio of each of these is shown with respect to the baseline NLO pQCD prediction, again using the CT10 PDF set.

In general, the non-perturbative corrections appear asymmetric here, particularly at higher values of p_{T} since the steeply falling cross-section means that, when such corrections are applied, migrations from lower bins to higher predominate over the reverse case. This results asymmetric shifts with respect to the nominal case, and hence an asymmetric error band on the ratio.

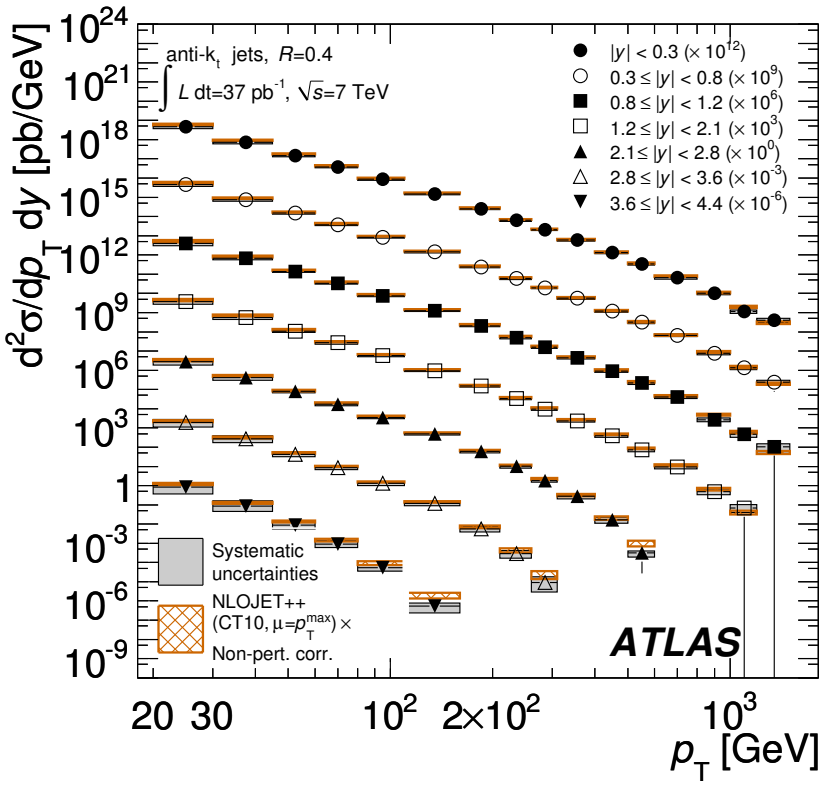


Figure 5.6.: Inclusive jet double-differential cross-section as a function of jet p_{T} in different regions of $|y|$ for jets identified using the anti-k_t algorithm with $R = 0.4$. For convenience, the cross-sections are multiplied by the factors indicated in the legend. The data are compared to NLO pQCD calculations to which non-perturbative corrections have been applied. The error bars indicate the statistical uncertainty on the measurement, and the dark-shaded band indicates the quadratic sum of the experimental systematic uncertainties, dominated by the jet energy scale uncertainty. There is an additional overall uncertainty of 3.4% due to the luminosity measurement that is not shown. The theory uncertainty (light cross-hatched band) shown is the quadratic sum of uncertainties from the choice of renormalisation and factorisation scales, parton distribution functions, α_{S} (M_Z), and the modelling of non-perturbative effects, as described in the text.

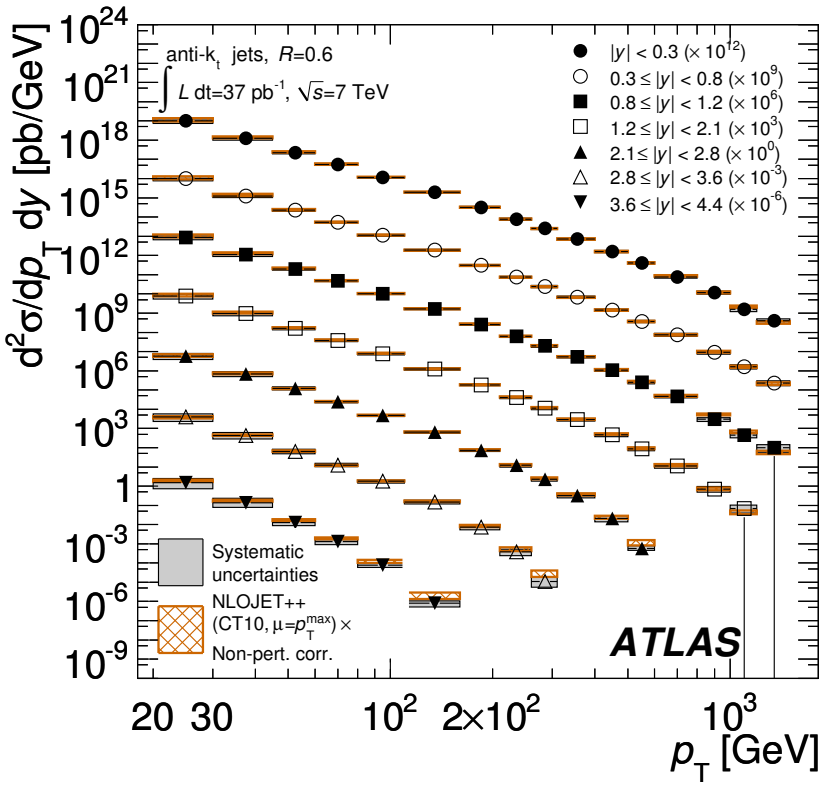
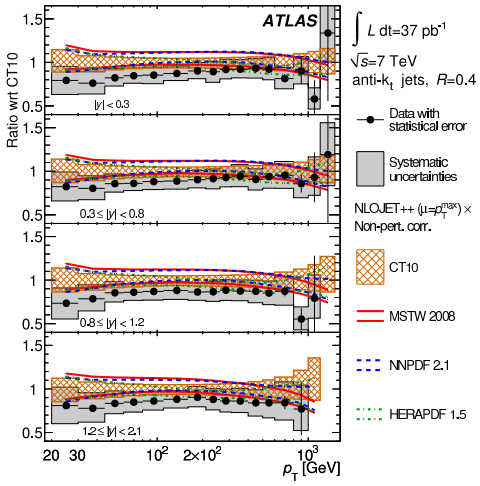
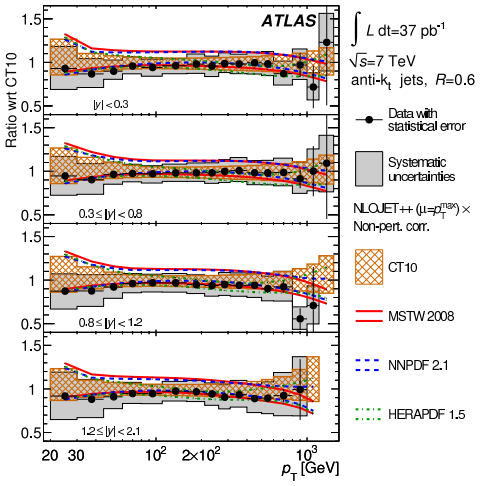


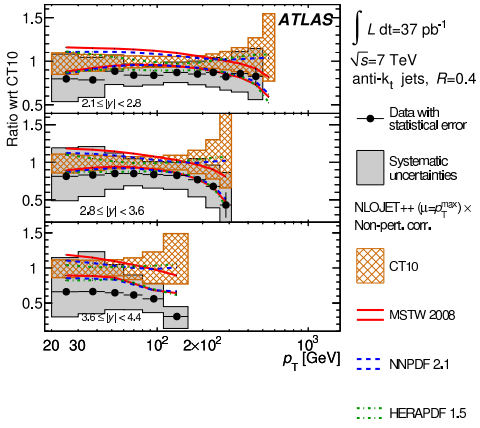
Figure 5.7.: Inclusive jet double-differential cross-section as a function of jet p_{T} in different regions of $|y|$ for jets identified using the anti-k_t algorithm with $R = 0.6$. For convenience, the cross-sections are multiplied by the factors indicated in the legend. The data are compared to NLO pQCD calculations to which non-perturbative corrections have been applied. The theoretical and experimental uncertainties indicated are calculated as described in Figure 5.6.



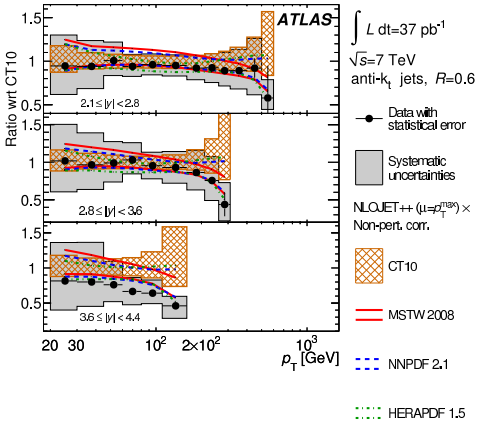
(a) Anti-k_t $R = 0.4$ jets, central rapidities



(b) Anti-k_t $R = 0.6$ jets, central rapidities



(c) Anti-k_t $R = 0.4$ jets, forward rapidities



(d) Anti-k_t $R = 0.6$ jets, forward rapidities

Figure 5.8.: Inclusive jet double-differential cross-section as a function of jet p_{T} in different regions of $|y|$ for jets identified using the anti-k_t algorithm with $R = 0.4$ (left) and $R = 0.6$ (right) for central rapidities (top) and forward rapidities (bottom). The ratio of the data to the theoretical prediction is shown, and the total systematic uncertainties on the theory and measurement are indicated. The theoretical and experimental uncertainties are calculated as described in Figure 5.6. Statistically insignificant data points at large p_{T} are omitted in this ratio.

POWHEG interfaced with PYTHIA describes the data better than when it is interfaced with HERWIG. Since the same matrix element is then passed through the PYTHIA and HERWIG parton showers, it can be deduced that the observable differences between these predictions must be connected to the specifics of their parton shower implementations and can be taken as an indication of the uncertainty arising from the leading-logarithmic approximation used in parton showering.

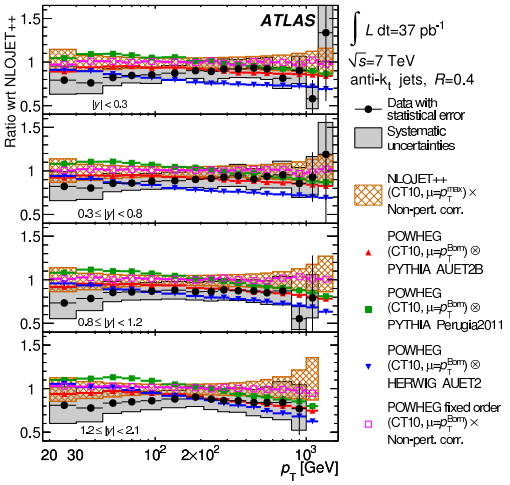
It can also be seen that the POWHEG NLO predictions, after parton shower, are in good agreement with the pure parton level matrix element calculation from NLOJET++ after this has been corrected for non-perturbative effects. No direct comparison has been carried out between the NLO parton level predictions of POWHEG and NLOJET++. This is unlikely to be feasible since the POWHEG formalism guarantees to generate the hardest third jet in the event while NLOJET++ only guarantees that a third jet will be generated. This is also an added complication when interfacing POWHEG with a parton shower, as a matching process needs to be performed to ensure this condition is adhered to.

Within the present uncertainties, the POWHEG predictions are consistent with both the data and NLOJET++ calculations. There is a trend for POWHEG to predict larger cross-sections than both the data and NLOJET++ at low p_{T}, and smaller cross-sections than NLOJET++ (but closer to the data) in the high-p_{T} region. These are also the regions where the scale uncertainty in NLOJET++ increases. At low p_{T} the non-perturbative corrections have a significant influence, and their uncertainty can be large.

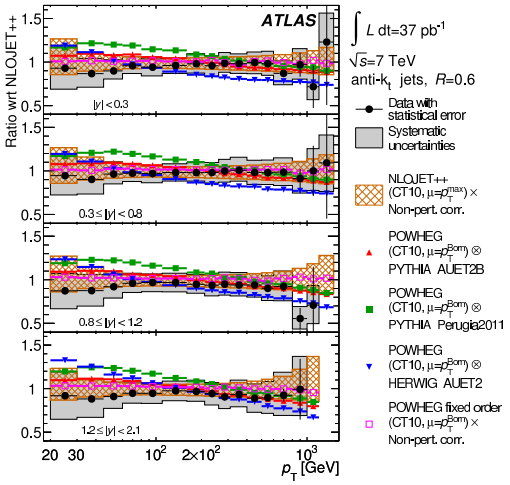
5.9. Summary

These results represent one of the most comprehensive tests of QCD ever performed. Data taken using minimum bias and jet triggers has been combined in order to measure cross-sections across a wide range of p_{T} and rapidity. In particular, the forward region has never previously been explored with such precision at a hadron collider.

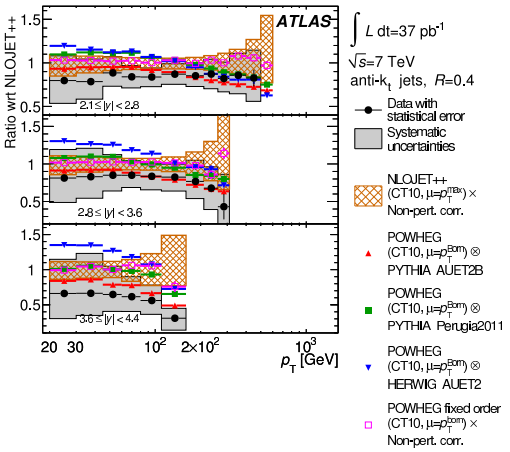
Jets are reconstructed using the anti-k_t algorithm with $R = 0.4$ and $R = 0.6$ in order to probe the relative effects of the parton shower, hadronisation and underlying



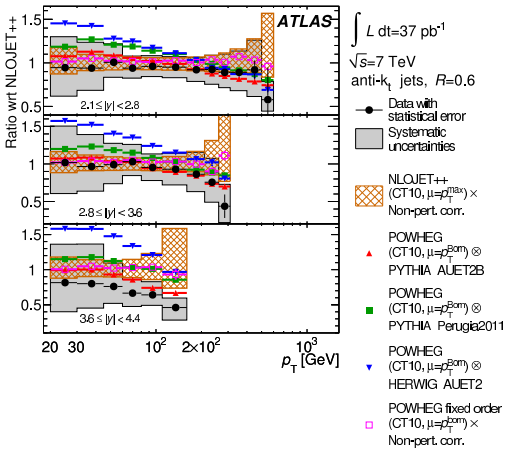
(a) Anti-k_t $R = 0.4$ jets, central rapidities



(b) Anti-k_t $R = 0.6$ jets, central rapidities



(c) Anti-k_t $R = 0.4$ jets, forward rapidities



(d) Anti-k_t $R = 0.6$ jets, forward rapidities

Figure 5.9.: Ratios of inclusive jet double-differential cross-section to the theoretical prediction obtained using NLOJET++ with the CT10 PDF set. The ratios are shown as a function of jet p_{T} in different regions of $|y|$ for jets identified using the anti-k_t algorithm with $R = 0.4$ (left) and $R = 0.6$ (right) for central rapidities (top) and forward rapidities (bottom). The ratios of POWHEG predictions interfaced with either PYTHIA or HERWIG to the NLOJET++ predictions corrected for non-perturbative effects are shown and can be compared to the corresponding ratios for data. Only the statistical uncertainty on the POWHEG predictions is shown. The total systematic uncertainties on the theory and the measurement are indicated. The NLOJET++ prediction and the POWHEG ME calculations use the CT10 PDF set. Statistically insignificant data points at large p_{T} are omitted in the ratio.

event. Overall, the agreement of the NLO perturbative QCD predictions with the measurements extends over seven orders of magnitude in cross-section. These measurements probe and may constrain the previously unexplored area of parton distribution functions at large x and high momentum transfer.

In some regions of phase space, the experimental and theoretical uncertainties are similar in size, thereby providing some sensitivity to different theoretical predictions. Such differences as have been observed are of the same order as the NLO scale variation, meaning that no definitive conclusions can yet be drawn. This work has been published in the European Physical Journal C [58] and an updated version has been submitted to Physical Review D [59].

Chapter 6.

Leading Dijet Cross-sections

"A short reign does not spare the masses."
— Statius

6.1. Introduction

As discussed in Chapter 5, jet cross-section measurements are an important tool for probing theoretical predictions from QCD. This analysis examines the leading dijet cross-section as a function of the dijet mass, m_{12}, in the region $y^* < 4.4$, where the variable y^* is defined as half the rapidity difference between the two leading (highest p_{T}) jets, $y^* = |y_1 - y_2|/2$, and is the absolute rapidity of the dijets in their centre-of-mass frame. The results are compared to expectations based on next-to-leading order (NLO) QCD, corrected for non-perturbative effects, as well as to NLO Monte Carlo predictions.

6.2. Cross-section Definition

As for the inclusive jet cross-section (Section 5.2), the dijet cross-section is defined for the ATLAS standard jet algorithms: anti-k_t jets with $R = 0.4$ and $R = 0.6$. The jet cross-section measurements are, again, unfolded back to the ideal particle level final state of the proton-proton collision. Identically to the inclusive jet cross-section,

particle level jets in the Monte Carlo are identified with the same jet algorithms as for data, using all stable particles and hence including muons and neutrinos from decaying hadrons.

6.3. Event Selection

Similarly to the inclusive jets cross-section, events are required to belong to a good run and to have at least one good primary vertex (see Section 3.3). The dijet pair are selected as the two highest p_{T} jets in the event, provided that these satisfy $p_{\mathrm{T}}(j_1) \geq 30$ GeV, $p_{\mathrm{T}}(j_2) \geq 20$ GeV and that both jets are within the acceptance region, $|y| < 4.4$. If either of these jets are flagged as "bad" or "ugly" by the standard medium jet cleaning cuts, as detailed in Section 3.3.4, then the event is discarded. A pair of triggers appropriate to the transverse momentum and rapidity of the dijets is then assigned - the event must pass one of these triggers in order to contribute towards the cross-section. Events passing this trigger selection are retained, with a weight reflecting the amount of luminosity seen by the trigger combination in question. Approximately 45% of events passing these cuts contain at least one additional jet which would also pass the cuts, however, the definition used here, in which only the leading pair of jets is considered, ensures that there is no ambiguity in such situations.

6.4. Trigger Strategy

The trigger considerations for dijet measurements are more complicated than for the inclusive jet case because the dijet cross-section is measured as a function of the invariant mass of the two-jet system and the rapidity separation y^*, while jets are triggered according to the jet transverse momentum. A per-event strategy based on invariant mass, only considering events taken by triggers with $> 99\%$ efficiency, would force the use of highly prescaled low-threshold triggers to cover the mass ranges of interest.

The trigger strategy used for the inclusive jet cross-section would be inefficient if directly transferred to the dijet case since only the two hardest jets in the event are

important, rather than all jets over a given p_{T} threshold. Similarly, a trigger strategy which determined a single trigger that the event should pass, based on the p_{T} of the leading jet in the event, would unnecessarily lose events due to the effects of prescale.

The strategy adopted is to define two possible trigger chains, based on the p_{T} of the two leading jets, and to accept the event if either or both of these are passed. A similar trigger strategy to that discussed in Section 5.4 is used: again dependent on the run period and the p_{T} and y of the two leading jets. One important difference is that whereas for the single jet cross-section, triggers were selected based on whether the *per-event* efficiency was on plateau, here it must be required that the *per-jet* efficiency should be on plateau.

Using similar definitions to those discussed in Section 5.4, the rapidity interval in which jets can be accepted, $|y| < 4.4$, is divided into different regions: the central region, $|y| < 2.9$; the transition region, $2.9 \leq |y| < 3.3$; the forward region $3.3 \leq |y| < 3.6$ and the far-forward region $3.6 \leq |y| < 4.4$.

The trigger requirements are summarised in Table 6.1 for central jets, Table 6.2 for forward jets and Table 6.3 for far-forward jets. The transition region presents additional complications which are discussed in Section 6.4.1.

Having determined the two triggers that each event should pass, based on the p_{T} of the two leading jets in the event, it is now required that one or both should have been passed before the event can be accepted. All events which satisfy this condition are placed into a trigger category defined by those specific two triggers, in other words, one trigger category is created for each possible combination of the two triggers.

The equivalent luminosity seen by each trigger category can be calculated (see Section 6.4.3) and the event can then be weighted by the inverse of this effective luminosity before the final cross-section measurements are produced.

p_{T} [GeV]	Run<152777	Periods A–C	Periods D–F (minus E1–4)	Periods G–I
20–42.5	L1_MBTS_1	L1_MBTS_1	L1_MBTS_1	EF_mbMbts_1_eff
42.5–70	L1_MBTS_1	L1_J5	L1_J5	EF_j20_jetNoEF
70–97.5	L1_MBTS_1	L1_J15	L1_J15	EF_j35_jetNoEF
97.5–152.5	L1_MBTS_1	L1_J30	L1_J30	EF_j50_jetNoEF
152.5–197.5	L1_MBTS_1	L1_J55	L1_J55	EF_j75_jetNoEF
197.5–217.5	L1_MBTS_1	L1_J55	L1_J55	EF_j95_jetNoEF
217.5+	L1_MBTS_1	L1_J55	L1_J55	EF_L1J95_NoAlg

Table 6.1.: The trigger chains used in the dijet analysis for the central region, $|y| < 2.9$.

p_{T} [GeV]	Periods A–C	Periods E–F (minus E1–4)	Periods G–I
20–42.5	L1_MBTS_1	L1_MBTS_1	EF_mbMbts_1_eff
42.5–62.5	L1_MBTS_1	L1_FJ10	EF_mbMbts_1_eff
62.5–72.5	L1_MBTS_1	L1_FJ10	EF_fj30_jetNoEF
72.5–95	L1_MBTS_1	L1_FJ30	EF_fj30_jetNoEF
95–160	L1_MBTS_1	L1_FJ30	EF_fj50_jetNoEF
160+	L1_MBTS_1	L1_FJ30	EF_fj75_jetNoEF

Table 6.2.: The trigger chains used in the dijet analysis for the forward region, $3.3 \leq |y| < 3.6$.

p_{T} [GeV]	Periods A–E4	Periods E5–F	Periods G–I
20–42.5	L1_MBTS_1	L1_FJ10	EF_mbMbts_1_eff
42.5–50	L1_MBTS_1	L1_FJ10	EF_fj30_jetNoEF
50–67.5	L1_MBTS_1	L1_FJ30	EF_fj30_jetNoEF
67.5–100	L1_MBTS_1	L1_FJ30	EF_fj50_jetNoEF
100+	L1_MBTS_1	L1_FJ30	EF_fj75_jetNoEF

Table 6.3.: The trigger chains used in the dijet analysis for the far-forward region, $3.6 \leq |y| < 4.4$.

6.4.1. Trigger Complications in the Transition Region

As discussed in Section 5.4.1, there is an ambiguity over whether to associate jets which are reconstructed in the transition region, $2.9 \leq |y| < 3.3$, with a central or a forward jet trigger. In this region, which bridges the boundary between the central and forward calorimeters, rapidity differences between L1 or L2 trigger jets and offline jets can lead to central trigger jets that are reconstructed offline as forward jets, and *vice versa*. It can be seen in Figure 5.1 that the inefficiencies for central and forward triggers overlap by approximately 0.4 in rapidity.

Since each trigger is required to be on a per-jet efficiency plateau, different p_{T} bin boundaries are used in the central and forward regions to optimise acceptance; it is, therefore, no longer enough to use a the simple central OR forward logic that could be applied to the inclusive jet cross-section. Accordingly, an angular "matching" is performed between offline jets which fall into this region and trigger objects, to identify the closest trigger jet and hence whether the offline jet should be classified as central or forward.

Matching Offline Jets to L1 Triggers

During periods A–F, when no HLT information was used for event selection, offline jets must be matched to L1 trigger objects. However, η information is not retained by the FCAL at L1 and so simple ΔR matching cannot be used without risking biases in trigger identification. First the closest *forward* L1 trigger object is identified as that with the smallest $\Delta\phi$ with respect to the offline jet. Next the closest *central* L1 trigger object is identified as that with the smallest ΔR with respect to the offline jet; this takes advantage of the fact that η information is available for L1 objects in the central calorimeter. Having identified the closest *forward* and the closest *central* L1 objects, whichever one has the smaller $\Delta\phi$ separation with respect to the offline jet is assigned as the best match.

Matching Offline Jets to L2 Triggers

In contrast to L1 objects, L2 trigger jets have η information in all calorimeter systems. For periods G–I therefore, when the HLT was in rejection mode, each offline jet can

be matched to the L2 trigger jet that is closest to it in ΔR; in other words, the trigger jet that has the smallest ΔR with respect to the offline jet. Since each L2 trigger jet is seeded by a L1 ROI, it can be uniquely determined whether the L2 jet was seeded from a central or forward L1 ROI and hence whether the offline jet should be associated with a central or forward trigger.

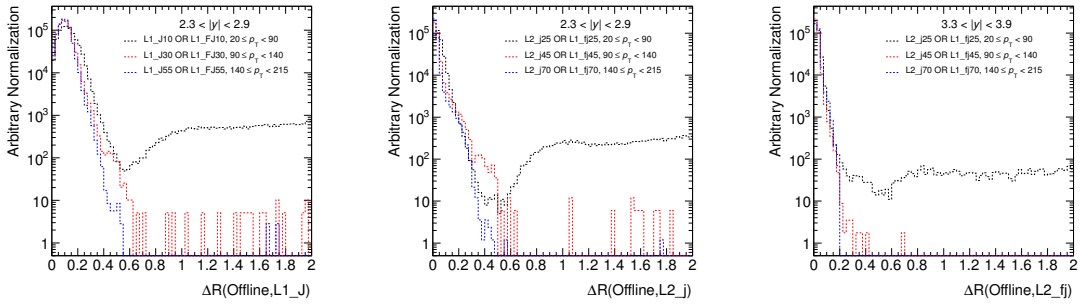
This procedure makes optimal use of such angular information as is available, avoiding bias by using $\Delta\phi$ to discriminate among forward L1 ROIs, where ΔR would tend to incorrectly assign forward jets as central. However it still takes advantage of the η information where possible in order to discriminate between central L1 trigger jets.

Determining ΔR and $\Delta\phi$ for Matching

The matching procedures described above are subject to the constraint that an offline jet is considered to be matched to a trigger jet only if they are separated by less than a maximum value of $\Delta R < 0.5$ or $\Delta\phi < 0.4$, depending which metric is used. The value of the maximum ΔR allowed for a match was chosen by inspecting the ΔR distribution of the closest trigger jet, which is shown in Figure 6.1. The distribution has a Gaussian core that extends to approximately $\Delta R \simeq 0.5$. This Gaussian population indicates a genuine correspondence between the offline jet and the closest trigger jet, where the smearing arises from the fundamentally different way in which offline and trigger jets are constructed. However, beyond this, in the region $\Delta R > 0.5$, a large non-Gaussian tail dominates. The non-Gaussian behaviour of this population indicates that these are mismatches, and that the offline jet and its closest trigger jet in fact have no real association. The boundary between these two populations occurring at $\Delta R = 0.5$ thus distinguishes between real matches and mismatches, so it is the optimal value for the ΔR matching cut.

A similar analysis was performed for the $\Delta\phi$ distribution shown in Figure 6.2 in order to derive the optimal matching cut of $\Delta\phi < 0.4$. In both the ΔR and $\Delta\phi$ cases, the studies were performed outside the region of interest, looking separately at more central jets, with $2.3 \leq |y| < 2.9$ and more forward jets, with $3.3 \leq |y| < 3.9$. This was done in order to ensure that the choice of cut was not biased by inefficiencies in triggering that would be present in the transition region. The distributions show a noticeable dependence on the p_{T} region considered, with the lowest p_{T} region

showing a large number of events in which the best matched trigger jet is a large distance away from the offline jet; this effect is an artefact of the low p_{T} offline jets considered here. The low edge of the p_{T} bins used for this study is not yet on the efficiency plateau of the relevant triggers, and so there will be cases in which a real offline jet has no counterpart among the collection of trigger jets. In such cases, the closest matched trigger jet will necessarily be an uncorrelated object which may therefore be a large distance away. Since the jets used in the analysis are always on their trigger plateau this is not an issue for the measurement. In fact the percentage of jets which are rejected by this matching procedure as a result of the closest match being above the cut is of the order of 0.5%, depending on p_{T}, and a systematic error is assigned to account for this.



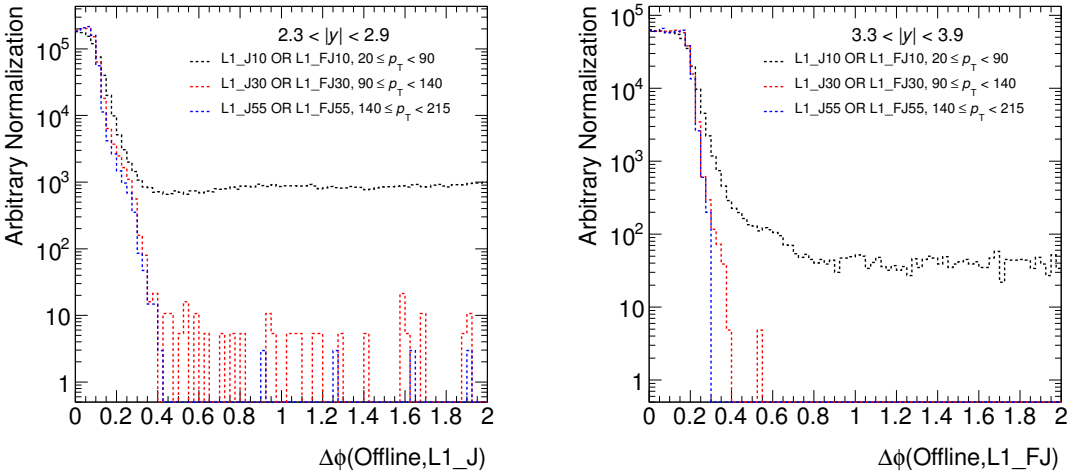
(a) L1 central jets **(b)** L2 central jets **(c)** L2 forward jets

Figure 6.1.: Angular separation ΔR between an offline jet and the closest trigger jet at L1 and L2. (a) L1 central jets, (b) L2 central jets and (c) are shown separately. The distribution is not shown for the forward region at L1 since the FCAL has no η measurement at L1. The distributions indicate that the optimal ΔR cut to distinguish real matches from mismatches is $\Delta R < 0.5$.

6.4.2. Per-Jet Trigger Inefficiencies

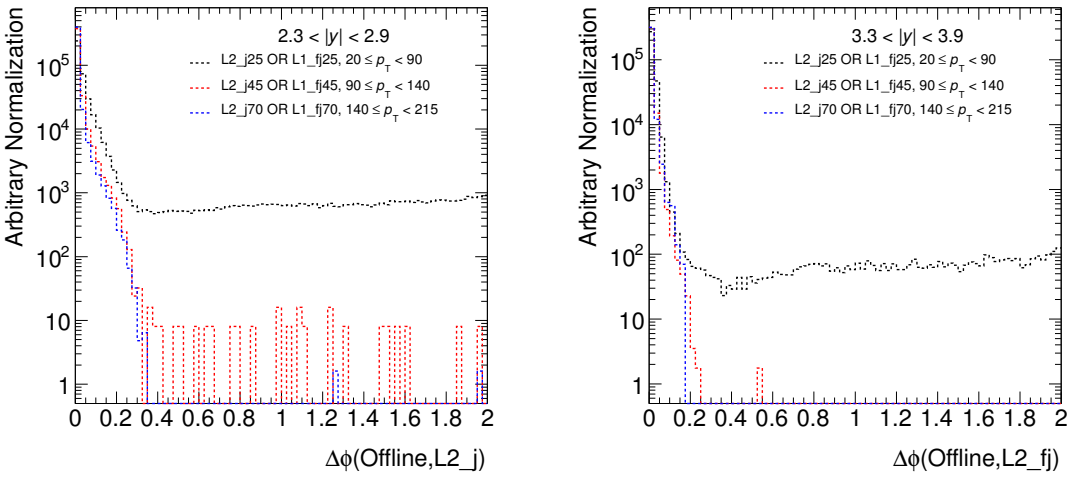
Unlike the case of the inclusive jet efficiency, there is a "crack" region between the calorimeter barrel and end-cap regions ($1.3 \leq |y| \leq 1.6$) where the per-jet trigger efficiency never becomes fully efficient due to calorimeter inhomogeneities. The per-jet trigger efficiency in this crack region is shown in Table 6.4.

Additionally, due to a dead FCAL trigger tower that spans a width of $\Delta\phi = \pi/4$ in the rapidity region $\eta > 3.1$, the forward jet triggers are not fully efficient in this



(a) L1 central jets

(b) L1 forward jets



(c) L2 central jets

(d) L2 forward jets

Figure 6.2.: Angular separation $\Delta\phi$ between an offline jet and the closest trigger jet at L1 (top) and L2 (bottom) for central jets (left) and forward jets (right). The distributions indicate that the optimal $\Delta\phi$ cut to distinguish real matches from mismatches is $\Delta\phi < 0.4$.

p_{T} [GeV]	Run< 152777	Run 152777–Period F	Period G–I
20–42.5	1.00	1.00	1.00
42.5–70	1.00	0.89	0.96
70–97.5	1.00	0.88	0.87
97.5–152.5	1.00	0.81	0.83
152.5–197.5	1.00	0.83	0.82
197.5–217.5	1.00	0.83	0.80
217.5+	1.00	0.83	0.81

Table 6.4.: The plateau per-jet trigger efficiency in the crack region $1.3 \leq |y| < 1.6$. Trigger inefficiency arises due to inhomogeneities in the crack region and is corrected for in the dijet measurement.

p_{T} [GeV]	Periods A–D	Periods E–F (minus E1–4)	Period G–I
20–42.5	1.00	1.00	1.00
42.5–62.5	1.00	1.00	1.00
62.5–72.5	1.00	1.00	0.99
72.5–95	1.00	0.97	0.99
95–160	1.00	0.97	0.99
160+	1.00	0.97	1.00

Table 6.5.: The per-jet trigger efficiency for the jet rapidity region $3.1 \leq |y| < 3.6$. Trigger inefficiency arises due to the dead FCAL tower and is corrected for in the dijet measurement.

region. The per-jet efficiency is shown for the rapidity region $3.1 \leq |y| < 3.6$ in Table 6.5 and for the forward region $3.6 \leq |y| < 4.4$ in Table 6.6.

A trigger efficiency correction is applied to any of the dijets that fall into either of these regions; in each case, the jet is weighted by the inverse of its trigger efficiency. A systematic error is assigned for this procedure, equal to the maximal correction.

p_{T} [GeV]	Periods A–D	Periods E–F (minus E1–4)	Period G–I
20–42.5	1.00	0.95	1.00
42.5–50	1.00	0.95	0.99
50–67.5	1.00	0.95	0.99
67.5–100	1.00	0.95	0.97
100+	1.00	0.95	0.97

Table 6.6.: The per-jet trigger efficiency for the jet rapidity region $3.6 \leq |y| < 4.4$. Trigger inefficiency arises due to the dead FCAL tower and is corrected for in the dijet measurement.

6.4.3. Calculating Effective Luminosity

In Section 5.4.1 it was possible to determine a cross-section for the case in which multiple triggers were used by dividing events up according to which triggers they would have passed before prescale. Here, however, the number of trigger categories is large and hence it would be preferable not to divide up the number of events any further.

Given that the triggers are selected in such a way that they are on plateau in the relevant region, it is expected that, in the absence of efficiency or prescale effects, all events should pass both of the appropriate triggers. Defining the case where the leading jet passes its trigger as T_{10} and the case where the second jet passes its trigger as T_{01}, this would mean that all events belong to case T_{11}.

In the hypothetical case where the leading and second jet triggers had respective prescales P^{L} and P^{S} but no inefficiencies, some events would move to T_{10}, some to T_{01} and some to T_{00}. The effective luminosity for the events remaining in one of T_{10}, T_{01} and T_{11} would be, after summing over all luminosity blocks, LB:

$$\mathcal{L}_{\mathrm{effective}} = \sum_{\mathrm{LB}} \frac{\mathcal{L}_{\mathrm{LB}}}{P^{\mathrm{L}}_{\mathrm{LB}} P^{\mathrm{S}}_{\mathrm{LB}} / (P^{\mathrm{L}}_{\mathrm{LB}} + P^{\mathrm{S}}_{\mathrm{LB}} - 1)} \tag{6.1}$$

Triggers passed (after prescale)	Efficiency weighting
Leading jet (T_{10})	$1/e_L$
Second jet (T_{01})	$1/e_S$
Both triggers (T_{11})	$1/e_L + 1/e_S - 1/(e_L * e_S)$

Table 6.7.: The efficiency weightings for different combinations of passed triggers. e_L is the efficiency of the leading jet trigger and e_S is the efficiency of the second jet trigger.

where $\mathcal{L}_{\mathrm{LB}}$ is the luminosity of each luminosity block. This is simply the degenerate case of Equation (5.3) in which all events belong to the last category. Unlike that situation, however, it is no longer important which events belong to which category before prescale: an identical effective luminosity can be applied to all events in the trigger category.

With the addition of trigger inefficiencies, the situation becomes slightly more complicated. However, after compensating for inefficiency, only the total number of events needs to be determined. It can be demonstrated, see Appendix A, that, if events are weighted as shown in Table 6.7, then this total number of events can be correctly recovered: although the individual numbers no longer directly indicate which triggers were passed, the total number passing one of the two triggers is correct.

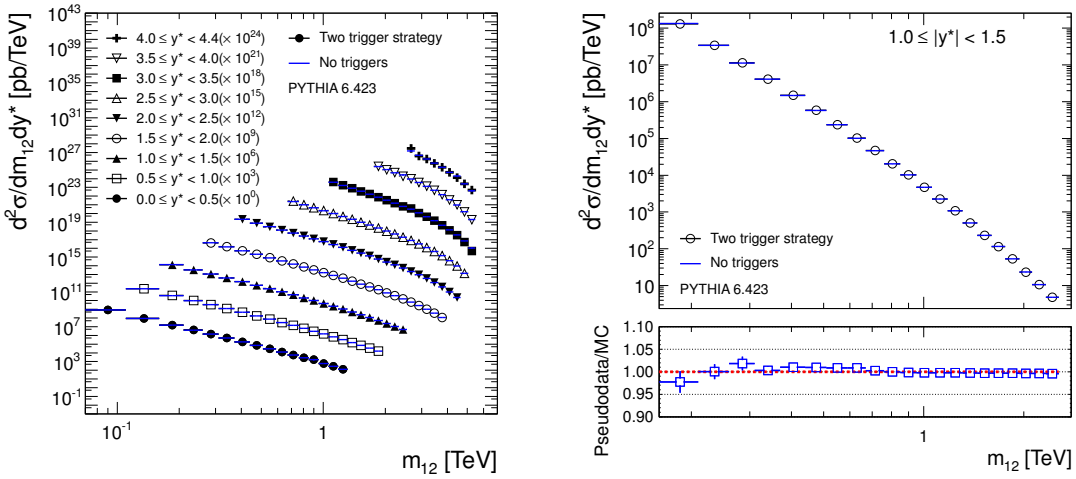
6.5. Validating the Two-Trigger Strategy

This trigger procedure was validated to ensure that it gave compatible results with a simpler single-trigger strategy. Additionally, investigations were made to determine optimum values for the matching cuts, which are discussed in Section 6.4.1. Finally a closure test was performed with Monte Carlo, where the effects of the trigger have been emulated using the prescale values of a typical run.

The implementation of this method is validated using a closure test with Monte Carlo, where the effects of the trigger are emulated using the prescale values of a typical run. Monte Carlo events are discarded at random according to the prescales of the triggers appropriate to the dijet pair, with the surviving events constituting a

"triggered pseudodata" sample, which is then analysed with the same procedure as is used for data. In particular, the pseudo-triggered events are assigned to trigger categories, the dijet mass histogram from each trigger category is divided by the appropriate equivalent luminosity, and finally the resulting cross-sections from all the trigger categories are combined to obtain the final detector level results.

The resulting dijet mass spectrum in slices of the rapidity separation y^* from 0.0 to 4.4 is shown in Figure 6.3 for both the pseudo-triggered and untriggered Monte Carlo samples. The distributions obtained after trigger emulation and correction using the trigger analysis code are compatible, within the available statistics, with the same distributions from the original Monte Carlo sample without the trigger requirement, hence validating the two-jet trigger method.



(a) Dijet closure $0.0 \le y^* < 4.4$ **(b)** Dijet closure, $1.0 \le y^* < 1.5$

Figure 6.3.: Summary of the closure test for a dijet Monte Carlo sample for jets identified using the anti-k_t algorithm with $R = 0.6$. The black dots represent the result of emulating the trigger in the Monte Carlo and then correcting for it using the same technique used on data (the pseudodata sample), while the blue solid lines (overlaid) represent the result of analysing all events, without trigger corrections (the Monte Carlo sample). The results are presented in (a) as a function of m_{12} in different y^* bins while in (b) the case $1.0 \le y^* < 1.5$ is shown separately, together with the ratio of pseudodata to Monte Carlo in each mass bin.

	L1_MBTS_1	L1_J5	L1_J15	L1_J30	L1_J55
L1_MBTS_1	0.0005827	0.02627	0.2723	0.2723	0.2723
L1_J5		0.2723	0.2723	0.2723	0.2723
L1_J15			0.2723	0.2723	0.2723
L1_J30				0.2723	0.2723
L1_J55					0.2723

Table 6.8.: Effective luminosity in pb^{-1} of different trigger combinations used in Periods A–D.

	L1_MBTS_1	L1_J5	L1_J15	L1_J30
L1_MBTS_1	0.00008220	0.002549	0.02488	1.136
L1_J5		0.002473	0.02728	1.137
L1_J15			0.02489	1.140
L1_J30				1.136
	L1_J55	L1_FJ10	L1_FJ30	
L1_MBTS_1	2.251	0.01606	2.251	
L1_J5	2.251	0.01844	2.251	
L1_J15	2.251	0.04019	2.251	
L1_J30	2.251	1.137	2.251	
L1_J55	2.251	2.251	2.251	
L1_FJ10		0.01600	2.251	
L1_FJ30			2.251	

Table 6.9.: Effective luminosity in pb^{-1} of different trigger combinations used in Periods E5–F.

6.6. Effective Luminosities

For the dijet measurements, the luminosities are calculated for the two-jet trigger scheme used (see Tables 6.1–6.3). The effective luminosities for different trigger combinations are shown for Periods A–D in Table 6.8, for Periods E5–F in Table 6.9, and for Periods G–I in Table 6.10.

	mbMbts_1_eff	j20_jetNoEF	j35_jetNoEF	j50_jetNoEF
mbMbts_1_eff	0.0001021	0.002504	0.05213	0.2541
j20_jetNoEF		0.002418	0.05437	0.2578
j35_jetNoEF			0.05203	0.2782
j50_jetNoEF				0.2556
	j75_jetNoEF	fj30_jetNoEF	fj50_jetNoEF	fj75_jetNoEF
mbMbts_1_eff	6.507	0.1643	3.871	35.69
j20_jetNoEF	6.508	0.1665	3.873	35.69
j35_jetNoEF	6.521	0.1891	3.889	35.69
j50_jetNoEF	6.569	0.3424	3.951	35.69
j75_jetNoEF	6.507	6.558	7.906	35.69
fj30_jetNoEF		0.1642	3.937	35.69
fj50_jetNoEF			3.871	35.69
fj75_jetNoEF				35.69

Table 6.10.: Effective luminosity in pb^{-1} of different trigger combinations used in Periods G–I. The initial "EF_" has been left off all trigger names for brevity.

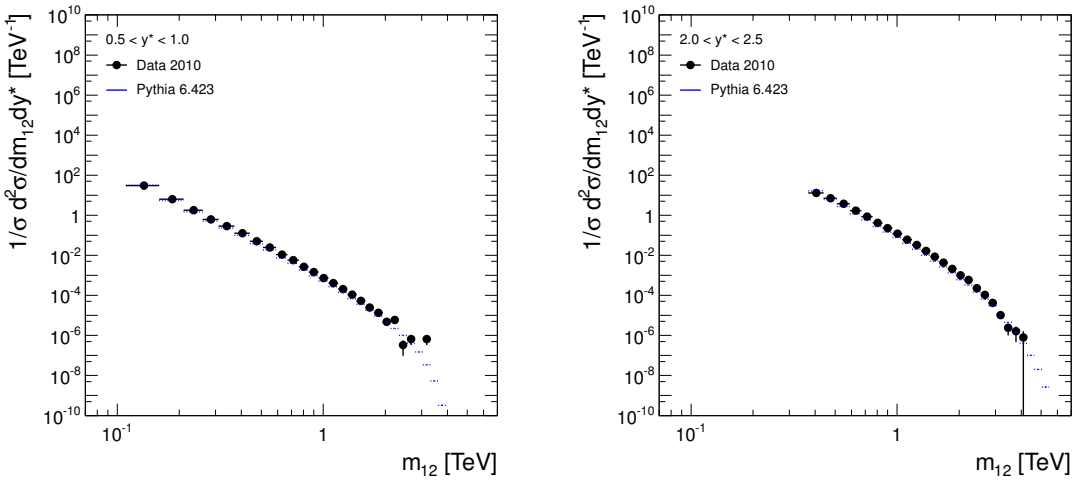
6.7. Unfolding Detector Effects

Unfolding of the dijet cross-sections is performed using the IDS method as described in Section 5.5. The same binning is used as for the final distributions and the particle level spectra in Monte Carlo are reweighted to correct for any shape mismodelling as discussed previously.

Figure 6.4 shows the level of agreement between PYTHIA and the data for two sample distributions; an event-by-event reweighting function is calculated and applied to the particle level spectra in Monte Carlo to eliminate any effects of shape differences between data and Monte Carlo.

6.8. Systematic Uncertainties

The systematic uncertainties for the dijet cross-section measurements are calculated, propagated and combined as described in Section 5.6. The jet energy scale uncertainty



(a) Data to Monte Carlo comparison for anti-k_t $R = 0.4$, in the region $0.5 \leq y^* < 1.0$

(b) Data to Monte Carlo comparison for anti-k_t $R = 0.6$, in the region $2.0 \leq y^* < 2.5$

Figure 6.4.: Control distributions, used to demonstrate the level of agreement between data and Monte Carlo. Data is shown in black, with PYTHIA in blue. (a) shows the comparison for anti-k_t $R = 0.4$ jets in the region $0.5 \leq y^* < 1.0$, while (b) shows the comparison for anti-k_t $R = 0.6$ jets in the region $2.0 \leq y^* < 2.5$.

is again the largest single contributor to the overall uncertainty, dominating the other contributions, particularly in the high mass region when the cross-section is steeply falling.

6.9. Theoretical Predictions

Measured dijet cross-sections are compared to both NLO pQCD predictions, with corrections for non-perturbative effects, and to NLO Monte Carlo. The NLO pQCD predictions are produced using NLOJET++ 4.1.2 [56] with the CT10 [57] NLO PDFs while NLO Monte Carlo predictions come from POWHEG, interfaced with both PYTHIA and HERWIG. Uncertainties are determined as outlined in Section 5.7.

6.10. Leading Dijet Cross-sections

The dijet double-differential cross-section is measured as a function of the dijet invariant mass for nine bins of the variable y^*, defined as half the absolute value of

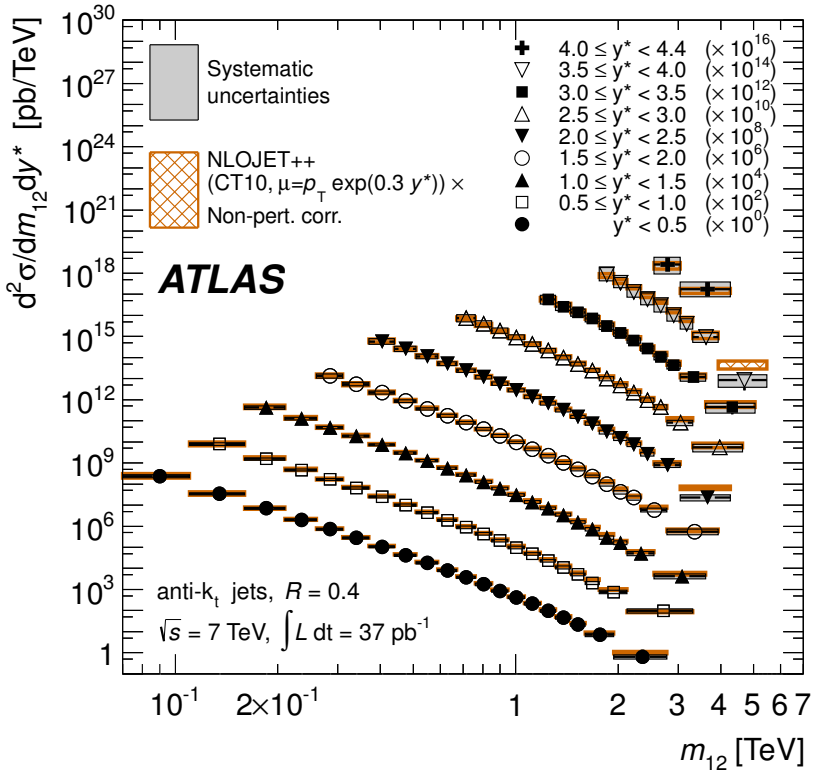


Figure 6.5.: Dijet double-differential cross-section as a function of dijet mass, binned in half the rapidity separation between the two leading jets, $y^* = |y_1 - y_2|/2$, for jets identified using the anti-k_t algorithm with $R = 0.4$. For convenience, the cross-sections are multiplied by the factors indicated in the legend. The data are compared to NLO pQCD calculations from NLOJET++, using the CT10 PDF set and to which non-perturbative corrections have been applied. The error bars, which are usually smaller than the symbols, indicate the statistical uncertainty on the measurement. The dark-shaded band indicates the quadratic sum of the experimental systematic uncertainties, which are dominated by the jet energy scale uncertainty. There is an additional overall uncertainty of 3.4% due to the luminosity measurement that is not shown. The theory uncertainty, shown as the light cross-hatched band, is the quadratic sum of uncertainties from the choice of the renormalisation and factorisation scales, PDFs, $\alpha_S(M_Z)$, and the modelling of non-perturbative effects, as described in the text.

the rapidity difference of the two leading jets, ranging from 0 to 4.4. The results are shown in Figures 6.5 and 6.6 for anti-k_t jets with $R = 0.4$ and $R = 0.6$, respectively. The cross-section falls rapidly with mass, and extends up to dijet masses of nearly 5 TeV.

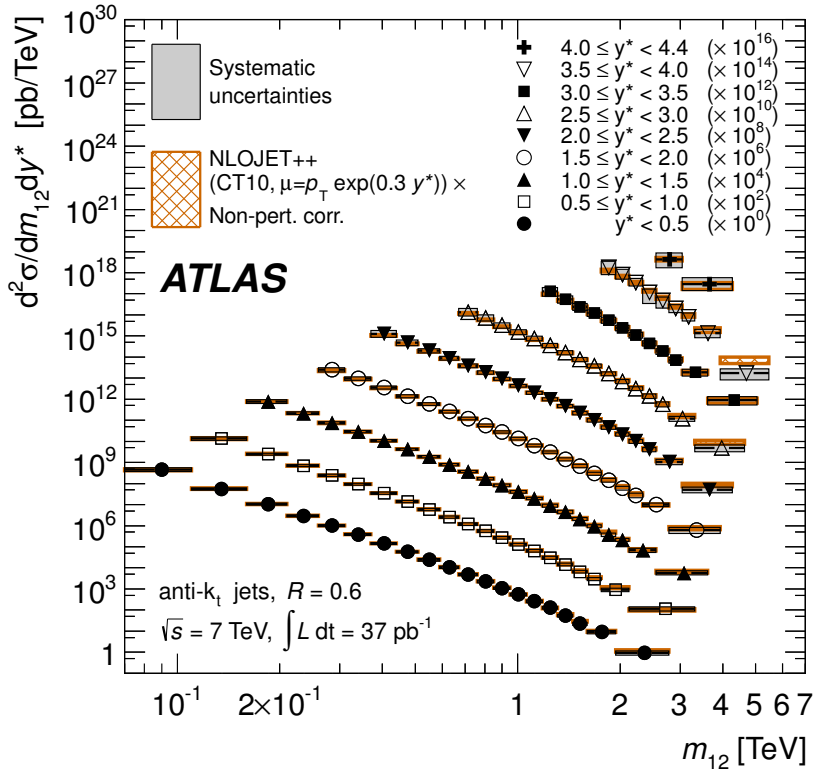


Figure 6.6.: Dijet double-differential cross-section as a function of dijet mass, binned in half the rapidity separation between the two leading jets, $y^* = |y_1 - y_2|/2$, for jets identified using the anti-k_t algorithm with $R = 0.6$. For convenience, the cross-sections are multiplied by the factors indicated in the legend. The data are compared to NLO pQCD calculations from NLOJET++, using the CT10 PDF set and to which non-perturbative corrections have been applied. The systematic uncertainties are calculated as described in Figure 6.5.

The comparison of the data with the POWHEG prediction, using the CT10 PDF set, is shown in different y^* regions in Figures 6.7 and 6.9 for anti-k_t jets with $R = 0.4$ and in Figures 6.8 and 6.10 for anti-k_t jets with $R = 0.6$. Here the data, POWHEG predictions interfaced with PYTHIA (AUET2B and Perugia 2010 tunes) and HERWIG (AUET2 tune) and fixed order POWHEG with non-perturbative corrections are all compared. The ratio of each of these with respect to the NLO pQCD prediction (CT10 PDF set) baseline is shown.

As in the inclusive jets case, the POWHEG prediction agrees better with data after being interfaced with PYTHIA than with HERWIG. Since the same matrix element, which agrees with the NLOJET++ prediction, is used in both cases, this provides further evidence that the PYTHIA parton shower approach more accurately reproduces the data than the angle-ordered parton shower used by HERWIG. The Perugia tune of PYTHIA also performs badly, but this is a deliberately extreme tune, which is not expected to agree well with the data. In general, the HERWIG tunes used for these comparisons are perhaps not correctly optimised for LHC data - additional tunes may provide improved agreement in future.

6.11. Summary

Through the use of a complicated trigger strategy involving the combination of multiple triggers from different trigger systems, it has been possible to measure dijet cross-sections across a wide range of mass and rapidity separations, reaching higher in y^* than has previously been possible.

Jets are reconstructed using the anti-k_t algorithm with $R = 0.4$ and $R = 0.6$ in order to probe the relative effects of the parton shower, hadronisation and underlying event. Overall, the agreement of the NLO perturbative QCD predictions with the measurements extends over seven orders of magnitude in cross-section: presenting an impressive validation of QCD over a range of phase space.

As discussed in Section 5.9, the experimental uncertainties are comparable to the theoretical uncertainties in some regions of phase space, providing sensitivity to different predictions. Differences between the NLO pQCD calculation and NLO Monte Carlo predictions are, however, of the same order as the NLO scale variation

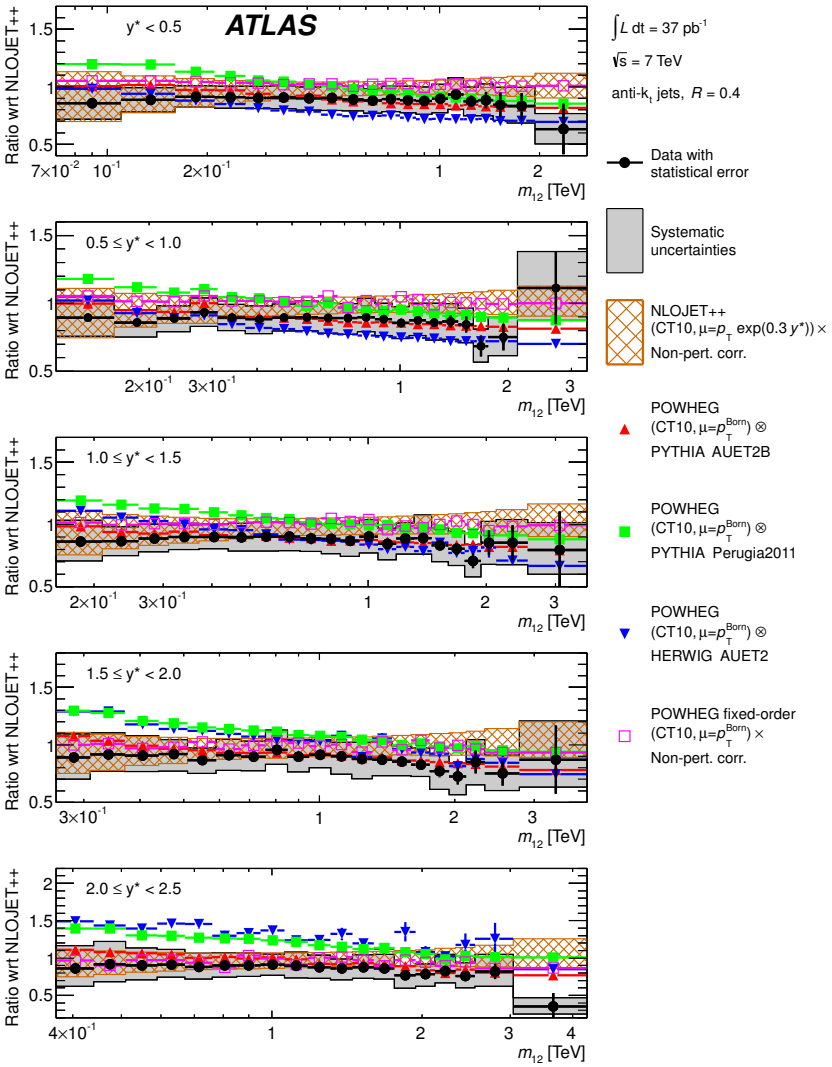


Figure 6.7.: Ratios of inclusive dijet double-differential cross-section to the theoretical prediction obtained using NLOJET++ with the CT10 PDF set. The ratios are shown as a function of dijet mass, binned in half the rapidity separation between the two leading jets, $y^* = |y_1 - y_2|/2$, for $0.0 \leq y^* < 2.5$. Jets are identified using the anti-k_t algorithm with $R = 0.4$. The plot shows the ratios of POWHEG predictions, interfaced with either PYTHIA (AUET2B tune), PYTHIA (Perugia 2011 tune) or HERWIG, to the NLOJET++ predictions, after these have been corrected for non-perturbative effects. The corresponding ratios for data are also shown for comparison. Additionally, POWHEG matrix-element calculations, also using the CT10 PDF set are shown. The total systematic uncertainties on the theory and the measurement are indicated. Only the statistical uncertainty on the POWHEG predictions is shown. The experimental uncertainties are calculated as described in Figure 6.5.

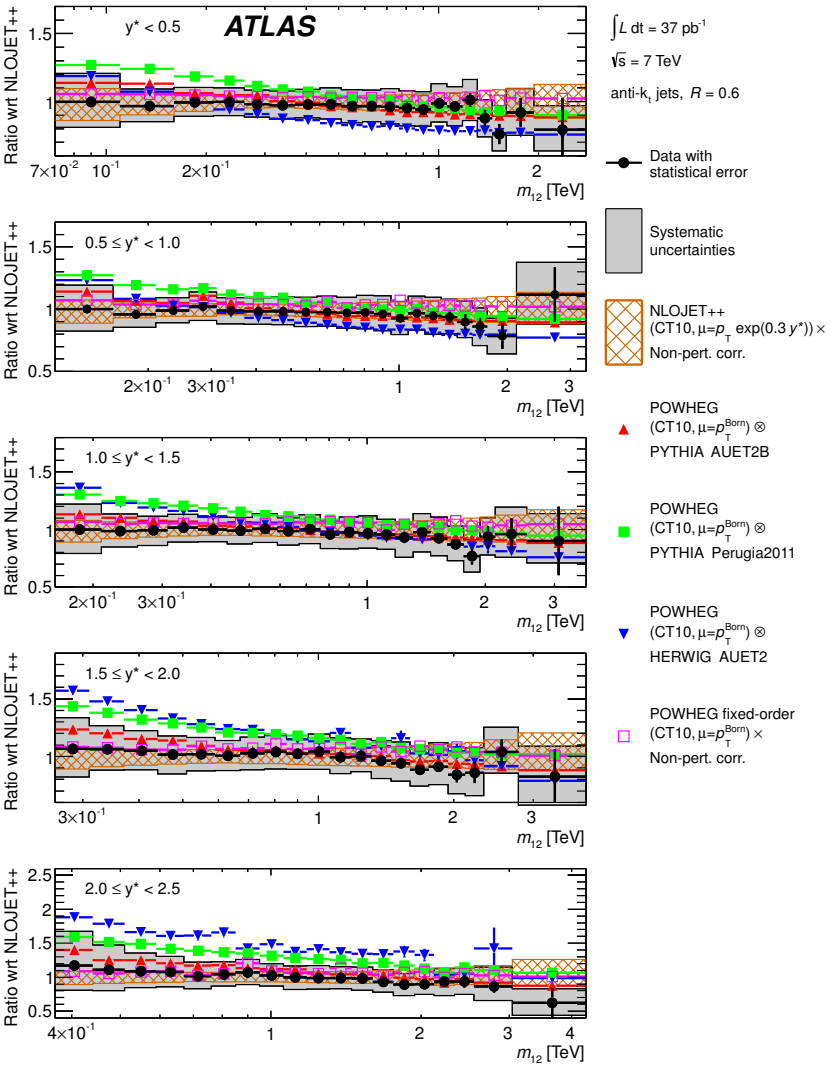


Figure 6.8.: Ratios of inclusive dijet double-differential cross-section to the theoretical prediction obtained using NLOJET++ with the CT10 PDF set. The ratios are shown as a function of dijet mass, binned in half the rapidity separation between the two leading jets, $y^* = |y_1 - y_2|/2$, for $0.0 \leq y^* < 2.5$. Jets are identified using the anti-k_t algorithm with $R = 0.6$. The plot shows the ratios of POWHEG predictions, interfaced with either PYTHIA (AUET2B tune), PYTHIA (Perugia 2011 tune) or HERWIG, to the NLOJET++ predictions, after these have been corrected for non-perturbative effects. The corresponding ratios for data are also shown for comparison. Additionally, POWHEG matrix-element calculations, also using the CT10 PDF set are shown. Uncertainties are as described in Figure 6.7.

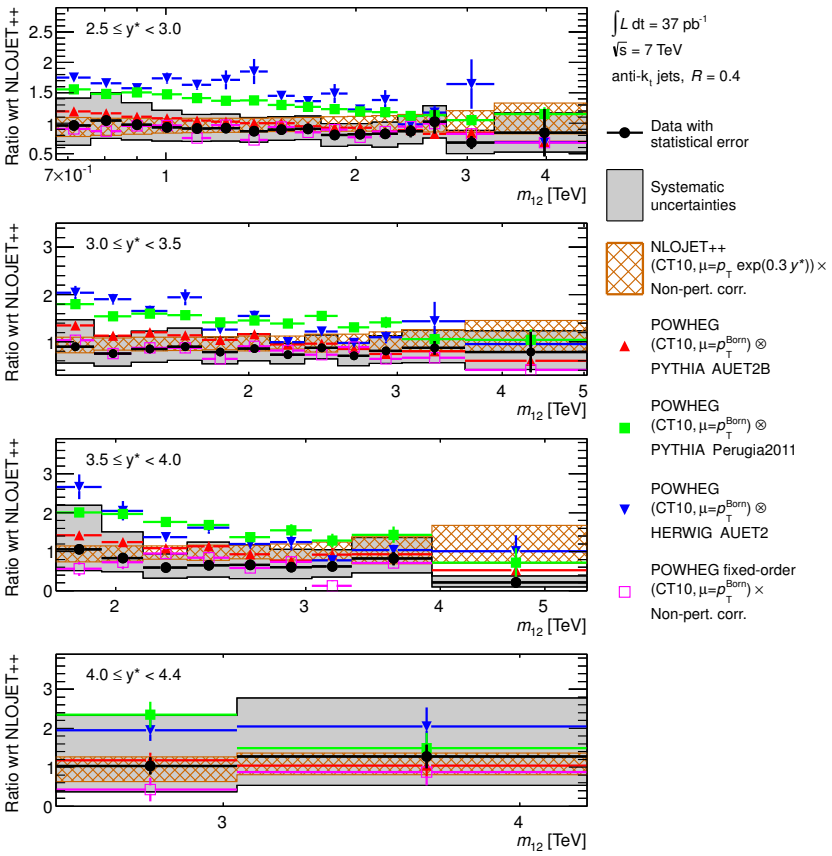


Figure 6.9.: Ratios of inclusive dijet double-differential cross-section to the theoretical prediction obtained using NLOJET++ with the CT10 PDF set. The ratios are shown as a function of dijet mass, binned in half the rapidity separation between the two leading jets, $y^* = |y_1 - y_2|/2$, for $2.5 \leq y^* < 4.4$. Jets are identified using the anti-k_t algorithm with $R = 0.4$. The plot shows the ratios of POWHEG predictions, interfaced with either PYTHIA (AUET2B tune), PYTHIA (Perugia 2011 tune) or HERWIG, to the NLOJET++ predictions, after these have been corrected for non-perturbative effects. The corresponding ratios for data are also shown for comparison. Additionally, POWHEG matrix-element calculations, also using the CT10 PDF set are shown. Uncertainties are as described in Figure 6.7.

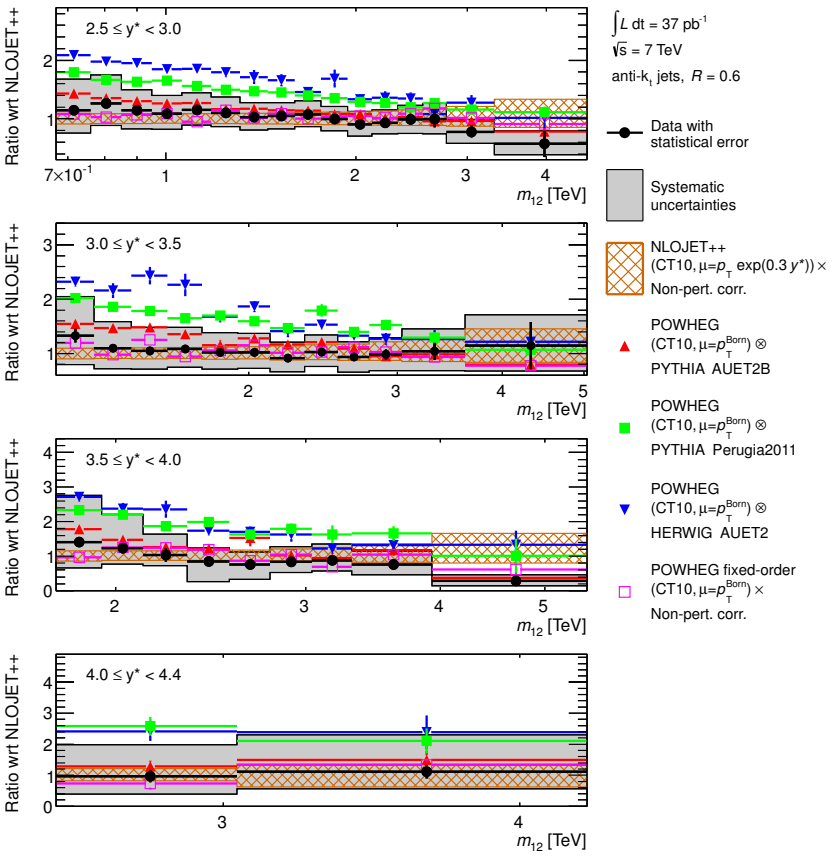


Figure 6.10.: Ratios of inclusive dijet double-differential cross-section to the theoretical prediction obtained using NLOJET++ with the CT10 PDF set. The ratios are shown as a function of dijet mass, binned in half the rapidity separation between the two leading jets, $y^* = |y_1 - y_2|/2$, for $2.5 \leq y^* < 4.4$. Jets are identified using the anti-k_t algorithm with $R = 0.6$. The plot shows the ratios of POWHEG predictions, interfaced with either PYTHIA (AUET2B tune), PYTHIA (Perugia 2011 tune) or HERWIG, to the NLOJET++ predictions, after these have been corrected for non-perturbative effects. The corresponding ratios for data are also shown for comparison. Additionally, POWHEG matrix-element calculations, also using the CT10 PDF set are shown. Uncertainties are as described in Figure 6.7.

and are hence inconclusive. This work has been published, together with the inclusive jet cross-sections in the European Physical Journal C [58] with an updated version submitted to Physical Review D [59].

Chapter 7.

Dijet Events with a Jet Veto

"If you ask me to compromise on principle, I will get out the veto pen."

— Bill Owens

7.1. Introduction

The production of dijets in which a veto is placed on additional radiation in the rapidity interval between the jets has previously been studied at HERA [60–62] and the Tevatron [63–67]. At the LHC, however, it is possible to study this process with an increased centre-of-mass energy and a greater rapidity coverage, allowing for wider gaps to be studied. In previous measurements, the main purpose of such measurements has been to search for evidence of colour singlet exchange. In order to do this, colour octet exchange contributions have to be suppressed, typically by imposing a low cut, of the order of 1 GeV, on the total radiation between the jets.

This level of veto is not feasible for jets in the ATLAS environment due to the effects of underlying event; instead, this analysis examines dijet systems in which radiation in the rapidity interval between the jets is suppressed using a third-jet veto. This allows a variety of different QCD phenomena to be examined. In the case in

which the rapidity separation between the jets is large, BFKL-like dynamics[1] are expected to become more important [70–72]; conversely, in the limit that the average dijet transverse momentum is much larger than the veto scale, the effects of wide-angle soft-gluon radiation can be examined [73, 74]. In summary, this measurement aims to study the effects of QCD radiation in regions of phase space which standard event generators sometimes struggle to describe adequately.

Jet veto studies are also relevant to Higgs boson production: searches for Higgs production via vector boson fusion, for example the Higgs-plus-two-jet analysis, often use jet vetoes as a method to reject background events, and this measurement provides an early test of the technique. In the long term, when extracting the couplings of the Higgs boson, the contribution from gluon fusion has a large theoretical uncertainty [75, 76] and jet veto studies provide one way to constrain the theoretical modelling.

7.2. Measurement Definition

For all data, Monte Carlo and theoretical distributions, the jets used are created using the anti-k_t algorithm with distance parameter $R = 0.6$, one of the standard ATLAS jet collections. Jets are required to have transverse momentum $p_{\mathrm{T}} \geq 20$ GeV and rapidity, $|y| < 4.4$, ensuring that they are in a region in which the jet energy scale has been validated (see Section 7.5.1).

In order to study the radiation in the rapidity region bounded by a dijet system, a scheme must be defined by which the dijets are identified. The analysis is performed using two different definitions of these "boundary" dijets, with the aim of probing different physics in the two cases. The first approach, referred to here as "selection A", identifies the boundary jets as the two highest transverse momentum jets in the event. The second approach, "selection B", identifies the boundary jets as the most forward and most backward in rapidity among all jets in the event with $p_{\mathrm{T}} \geq 30$ GeV.

[1] Balitsky-Fadin-Kuraev-Lipatov (BFKL) dynamics propose an evolution in $\ln(1/x)$, where x is the Bjorken variable, as opposed to the DGLAP evolution in $\ln(Q^2)$, where Q^2 is the parton virtuality [68, 69].

7.3. Event Selection

As discussed in Section 3.3, events are required to belong to a good run and to have at least one good primary vertex. As for the η-intercalibration measurement in Chapter 4, only events with exactly one good primary vertex are considered in order to cut down on the number of events affected by in-time pile-up. The fraction of events retained by this single vertex requirement is 92% in the first periods of data taking, falling to 20% in the last period. Events are rejected if they contain any jets with $p_{\mathrm{T}} \geq 20$ GeV that are flagged as "bad" or "ugly" by the standard loose jet cleaning cuts.

After applying these cuts, the inclusive sample of events is defined as those for which both boundary jets satisfy $p_{\mathrm{T}} \geq 30$ GeV and additionally the average transverse momentum, $\overline{p_{\mathrm{T}}}$, of the boundary jets is greater than 60 GeV.

Gap events are defined as that subset of inclusive events that do not contain an additional jet with p_{T} greater than the veto scale, Q_0. The default value of $Q_0 = 20$ GeV was chosen because jets with p_{T} lower than this are not fully accounted for by the ATLAS jet energy scale uncertainty tools. For the selection B criteria, the case $Q_0 = \overline{p_{\mathrm{T}}}$ is also considered.

The major focus of study in this analysis is the gap fraction, the fraction of inclusive events which do not possess a third jet with $p_{\mathrm{T}} \geq Q_0$. The gap fraction is investigated as a function of $\overline{p_{\mathrm{T}}}$ and of the rapidity separation between the boundary jets, Δy.

Using the gap fraction has the advantage that some of the most important experimental uncertainties cancel out; in particular the jet energy scale uncertainty does not have such as large effect as, for instance, in a jet cross-section measurement. This fact allows the use of the forward part of the calorimeter, where the jet energy scale uncertainty is larger than in the central part.

7.4. Trigger Strategy

For this analysis, data are taken using the jet trigger system. As discussed in Section 3.3.6, L1 calorimeter triggers were used for early data and L2 triggers for

$\overline{p_{\mathrm{T}}}$ range [GeV]	Period B–D	Period E5–F	Period G–I
50–70	L1_J5	EF_j20_NoCut	EF_j20_NoEF
70–90	L1_J10	EF_j30_NoCut	EF_j30_NoEF
90–120	L1_J15	EF_j35_NoCut	EF_j35_NoEF
120–150	L1_J35	EF_j50_NoCut	EF_j50_NoEF
150–180	L1_J55	EF_j75_NoCut	EF_j75_NoEF
180–210	L1_J75	EF_j95_NoCut	EF_j95_NoEF
210+	L1_J95	EF_L1J95_NoAlg	EF_L1J95_NoAlg

Table 7.1.: The trigger chains used for the gaps between jets analysis. Trigger items are used exclusively in specific regions of $\overline{p_{\mathrm{T}}}$. Periods E1–4 were not used due to problems with the FCAL.

later periods. Due to the regions of phase space considered, it is likely that most events considered here will include at least one central jet. Given this fact, only central jet triggers are used for this analysis in order to simplify the trigger strategy. Tests using Monte Carlo demonstrate that the inefficiency caused by this strategy is negligible.

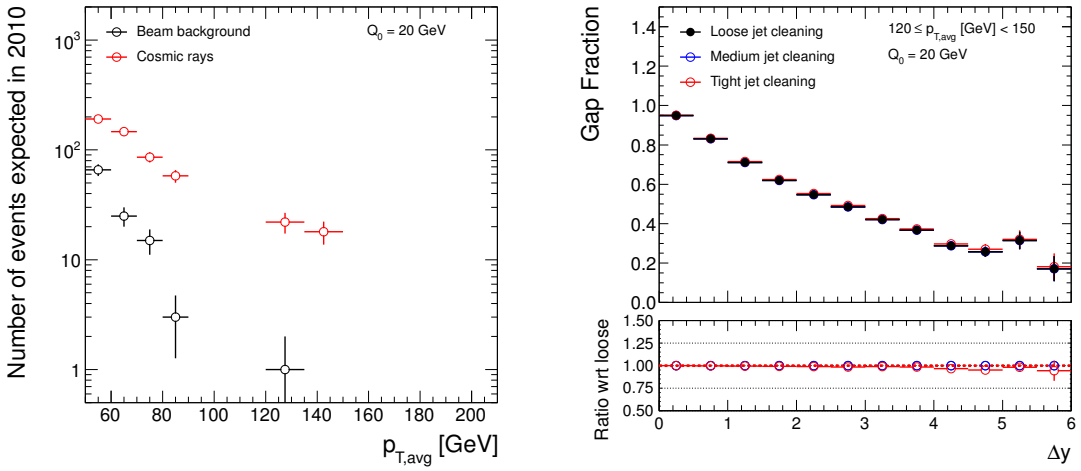
Since the analysis is performed in distinct slices of $\overline{p_{\mathrm{T}}}$, the mean transverse momentum of the two boundary jets, the trigger strategy is specified in terms of this variable. For each $\overline{p_{\mathrm{T}}}$ bin, only events passing the appropriate trigger are used in the analysis. The $\overline{p_{\mathrm{T}}}$ bin boundaries are determined by requiring the trigger to be on the 99% efficiency plateau over the whole Δy region under consideration. In other words, $\overline{p_{\mathrm{T}}}$ is calculated for each event, using the two leading jets, and this quantity is used to define a single trigger requirement for the event: if this trigger is passed then the event is accepted. The $\overline{p_{\mathrm{T}}}$ regions and corresponding trigger items are given in Table 7.1. Since this measurement studies ratios, there is no need to determine luminosities or to use weighted events.

7.5. Systematic Uncertainties

Several sources of systematic uncertainties are investigated for this measurement. The most important of these uncertainties are those associated with the absolute

and relative jet energy scale (see Section 7.5.1) and with the detector unfolding (see Section 7.5.2).

The effectiveness of the single vertex requirement at reducing the influence of pile-up is also studied. Despite a marked increase in pile-up activity in the latter data taking periods, the gap fraction is observed to be independent of data taking period. In addition, cosmic and beam related backgrounds are estimated using events from appropriate data streams and their effects on the final measurement found to be small in comparison to signal events (see Figure 7.1a). The three different sets of jet cleaning cuts discussed in Section 3.3.4 are compared and the choice of jet cleaning is found to have no impact on the final measurement (see Figure 7.1b). Finally, the bias due to the central-only trigger strategy is evaluated using data taken by the minimum bias trigger stream and found to be negligible in all regions of phase space.

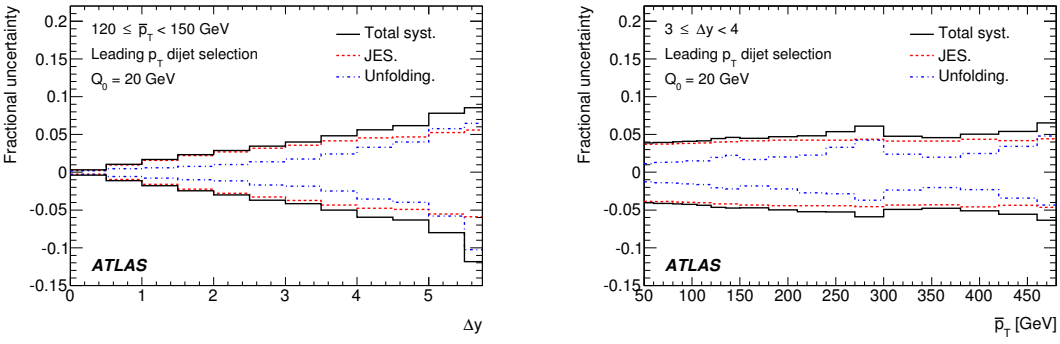


(a) Background contribution, inclusive events **(b)** Effect of jet cleaning cuts on gap fraction

Figure 7.1.: Control distributions, used to demonstrate the effects of some potential sources of systematic error. (a) shows the Δy distribution of events arising from beam background and cosmic rays. The overall number of accepted signal events was approximately 5×10^5 inclusive and 4×10^5 gap events, so these backgrounds represent only a small perturbation and do not need to be explicitly corrected for. (b) shows the effect on the gap fraction as a function of Q_0 of applying different jet cleaning cuts; the differences arising in the final distributions are found to be negligible.

The final systematic uncertainty is obtained by summing each of these components in quadrature. Figure 7.2 shows an example of how the relative systematic uncertainty

changes with the Δy of the boundary jets, both for the gap fraction and the mean number of jets in the rapidity interval between the boundary jets.



(a) Gap fraction systematics **(b)** $\overline{p_{\mathrm{T}}}$ systematics

Figure 7.2.: Example of the contributions to the systematic uncertainty for selection A data with $Q_0 = 20$ GeV: (a) on the gap fraction, for $120 \leq \overline{p_{\mathrm{T}}} < 150$GeV, (b) on the mean number of jets in the rapidity interval between the boundary jets, for $3 \leq \Delta y < 4$.

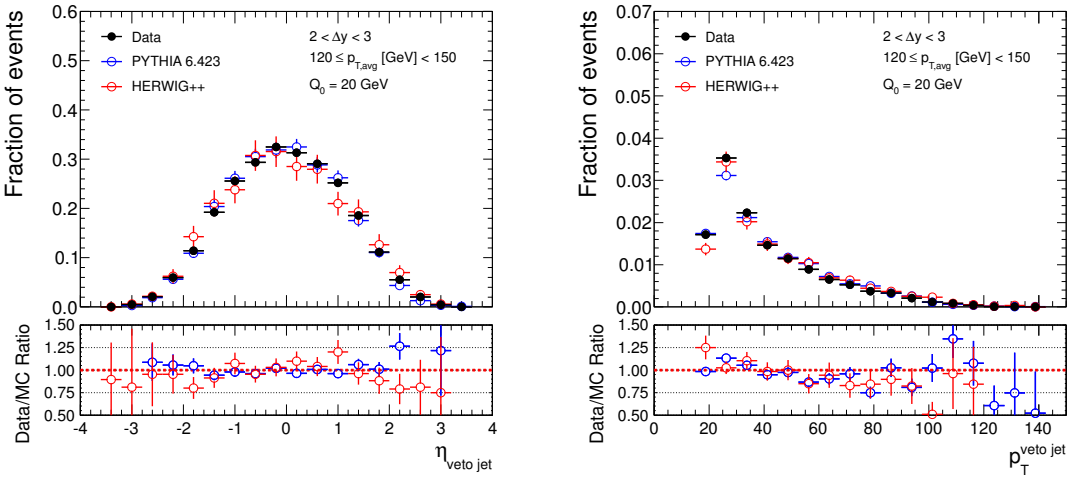
The irregularity at approximately 290 GeV that can be seen in the uncertainty arising from the unfolding procedure is caused by the way in which the relevant Monte Carlo samples are produced. Samples are created in bins of p_{T}, with statistics at the high end of one sample lower than the statistics available at the low end of the next sample. Such differences at the boundaries between samples, often cause statistical fluctuations when the results are combined.

7.5.1. Effect of Jet Energy Scale Uncertainty

Jet energy scale (JES) uncertainties are available as a function of jet transverse momentum and rapidity from a dedicated ATLAS tool [9]. Both absolute and relative jet energy scale (JES) uncertainty affect this measurement. The effect of the absolute JES uncertainty can be studied by shifting the energy of all jets by $\pm 1\sigma$. The effect of this shift on the gap fraction is small, ranging from 2% to 5% across the $\overline{p_{\mathrm{T}}}$ spectrum and from 0.5% to 5% across the Δy spectrum.

Relative JES uncertainty is the uncertainty in calibration between different rapidity regions of the detector. The largest effect occurs when there is a decorrelation of the JES uncertainty between the boundary jets and those jets which lie between them in rapidity. This occurs because of the way that gap and non-gap events are

distinguished using a third jet veto. To estimate the maximum uncertainty arising from this, it is assumed that the veto jets are central (see Figure 7.3a) and that the maximum decorrelation between the boundary jets and the veto jets is therefore 3% when the most forward boundary jet is central ($|\eta| < 2.8$) and 10% when the most forward boundary jet is forward ($|\eta| > 2.8$). These numbers come from the ATLAS JES uncertainty provider [9] and are based, in the forward region, on the results of the dijet intercalibration (see Figure 4.6). It is important to note that the data lies within the spread of Monte Carlo predictions in the forward region and hence that this is a conservative overestimate of the true JES uncertainty. The uncertainty on the gap fraction deriving from the effects of relative JES uncertainty is approximately 2% across the $\overline{p_{\mathrm{T}}}$ distribution and between 0.5% and 10% across the Δy distribution.



(a) Veto jet η distribution

(b) Veto jet p_{T} distribution

Figure 7.3.: The distributions of jets which are above the veto threshold, $Q_0 = 20$ GeV, and in the rapidity gap between the boundary jets in (a) η and (a) p_{T}.

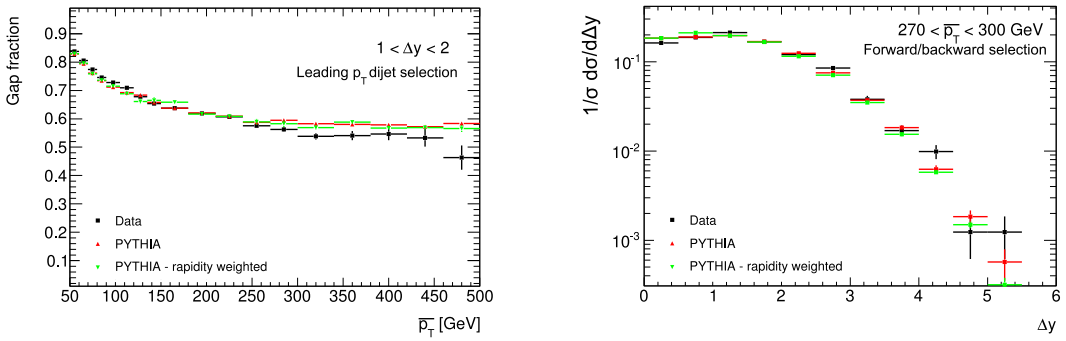
7.5.2. Unfolding Detector Effects and Associated Systematic Uncertainty

For this analysis, bin-by-bin unfolding (see Section 3.4) is used as the level of migration between bins for the gap fraction is small as a function of both $\overline{p_{\mathrm{T}}}$ and Δy.

The calculation of the unfolding correction factors is performed with PYTHIA, HERWIG++ and ALPGEN (see Section 1.11). Each of the resultant correction factors

is close to unity, with a typical deviation being smaller than 0.01; as such, the effect of unfolding is much smaller than the other systematic uncertainties in the analysis. Because of this, an unfolding correction is applied; instead a systematic error associated with unfolding is assigned as the quadrature sum of the following three quantities: the deviation from unity of the unfolding factor obtained with PYTHIA; the statistical error on the unfolding factor obtained with PYTHIA; the difference in unfolding factors obtained when the shapes of the $\overline{p_{\mathrm{T}}}$, Δy and $p_{\mathrm{T},3}$ truth distributions are changed by the maximal amount which produces detector level distributions that are still compatible with the observed data.

Figure 7.4 shows the level of agreement between PYTHIA and the data for two sample distributions; this provides a cross-check that using PYTHIA for the unfolding does not introduce any strong biases.



(a) Selection A. Data to Monte Carlo comparison as a function of $\overline{p_{\mathrm{T}}}$

(b) Selection B. Data to Monte Carlo comparison as a function of Δy for inclusive events

Figure 7.4.: Control distributions, used to demonstrate the level of agreement between data and Monte Carlo. Data is shown in black, with PYTHIA in red and specially generated PYTHIA samples which are weighted in y in green. (a) shows the comparison in the gap fraction as a function of $\overline{p_{\mathrm{T}}}$ for selection A, with $1 \leq \Delta y < 2$. (b) shows the comparison as a function of Δy for inclusive events in selection B, with $270 \leq \overline{p_{\mathrm{T}}} < 300$ GeV.

The final uncertainty in unfolding is typically much smaller than the JES uncertainty, except in the largest $\overline{p_{\mathrm{T}}}$ and Δy bins, where the Monte Carlo statistics are poor. In these bins, the unfolding uncertainty can be larger than 5% (see Figure 7.2). The correction factors predicted independently by HERWIG++ and ALPGEN agree within the statistical uncertainty of each sample.

7.6. Theoretical Predictions

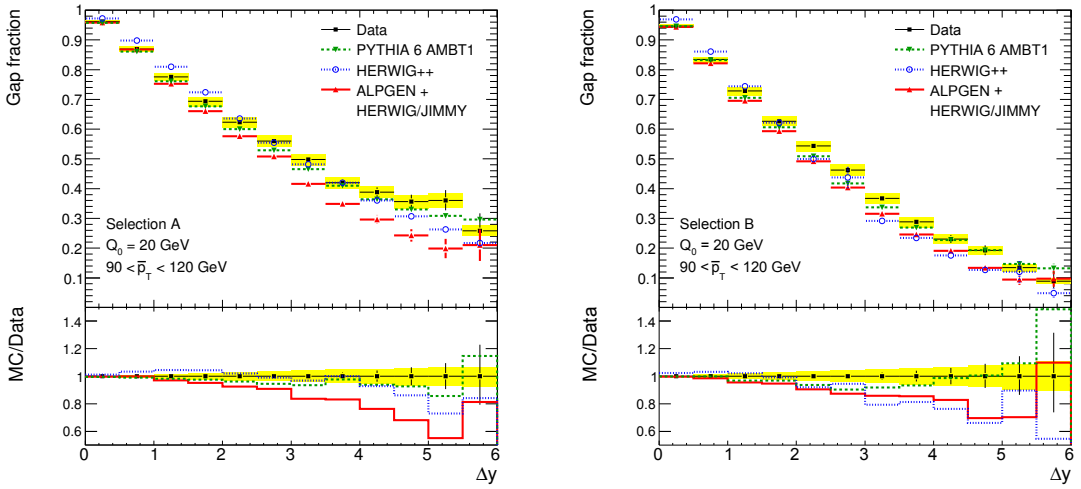
The measurements presented here probe perturbative QCD in the region where the energy scale of the dijet system is larger than the scale of any additional radiation. At large values of $\overline{p_{\mathrm{T}}}/Q_0$ or of Δy, it is expected that fixed order calculations are unlikely to describe the data and that a resummation to all orders in perturbation theory is necessary. These measurements are not particularly sensitive to non-perturbative physics because Q_0 is chosen to be much greater than Λ_{QCD}. The net effect of corrections arising from non-perturbative physics is estimated by turning the hadronisation and underlying event on and off in PYTHIA: the resulting shift in the gap fraction is less than 2% and the change in the mean number of jets in the rapidity interval bounded by the dijet system is less than 4%.

Theoretical predictions are produced using the next-to-leading order (NLO) POWHEG generator (see Section 1.11.4) and HEJ, a parton level event generator that provides an all-order description of wide-angle emissions of similar transverse momentum (see Section 1.11.6). In this BFKL-inspired limit, HEJ reproduces the full QCD results and should be especially suited for events with at least two jets separated by a large rapidity interval. HEJ events are generated with the MSTW 2008 NLO PDF set [31, 32] and with the renormalisation and factorisation scales chosen to be the p_{T} of the leading parton. The uncertainty due to this scale choice was estimated by increasing and decreasing each scale by a factor of two. Uncertainties coming from the choice of PDF are estimated using the full set of eigenvector errors provided by MSTW and also by changing the PDF to CTEQ6L1 [39]. The overall uncertainty in the HEJ calculation is dominated by the scale choice and is typically 5% for the gap fraction and 8% for the mean number of jets in the rapidity interval bounded by the dijet system. These uncertainties are larger than the non-perturbative physics corrections which were therefore incorporated into the systematic uncertainty in quadrature, rather than being applied explicitly.

7.7. Jet Veto Results

The unfolded data, for specific phase space regions, is compared against three of the leading order Monte Carlo event generators that are commonly used for

predictive purposes in ATLAS, namely PYTHIA, HERWIG++ and ALPGEN. Figure 7.5 shows the gap fraction as a function of Δy given that the boundary jets satisfy $90 \leq \overline{p_{\mathrm{T}}} < 120$ GeV.

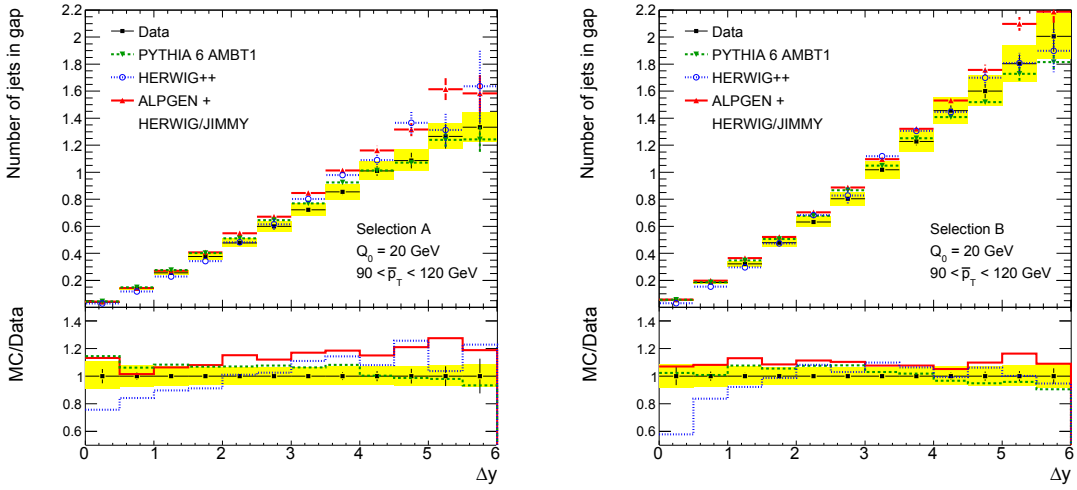


(a) Selection A. Leading p_{T} dijets **(b)** Selection B. Forward/backward dijets

Figure 7.5.: Gap fraction as a function of Δy for boundary jets that satisfy $90 \leq \overline{p_{\mathrm{T}}} < 120$ GeV for (a) selection A and (b) selection B. The (unfolded) data are the black points, with error bars representing the statistical uncertainty. The systematic uncertainty on the measurement is represented by the yellow band. The green curve represents the PYTHIA prediction (tune AMBT1), the blue curve represents the HERWIG++ prediction (tune for LO* PDFs) and the red curve represents the ALPGEN +HERWIG/JIMMY prediction (tune AUET1).

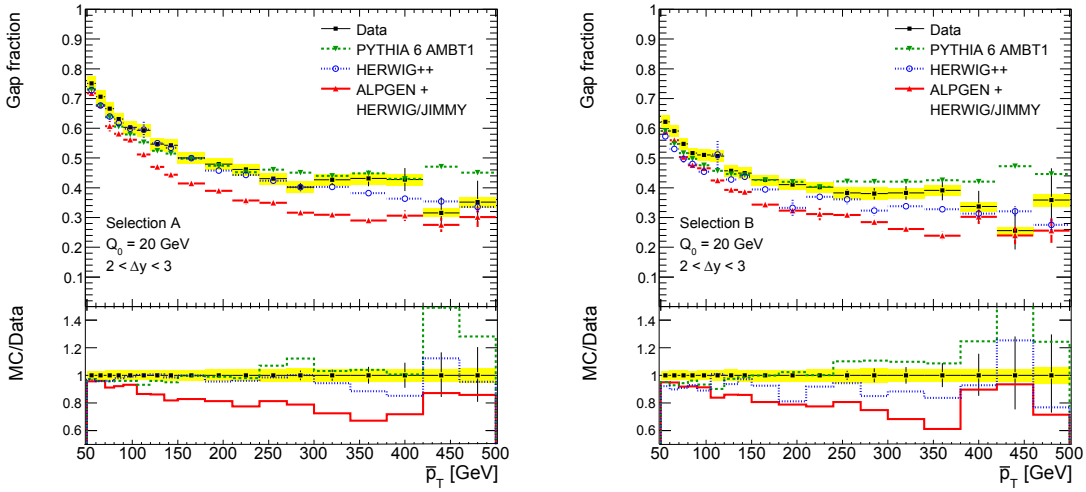
Figure 7.6 shows the mean number of jets in the rapidity interval between the boundary jets for the same cuts. Figure 7.7 shows the gap fraction as a function of $\overline{p_{\mathrm{T}}}$ given that $2 < \Delta y < 3$ while Figure 7.8 shows the mean number of jets in the rapidity interval between the boundary jets as a function of $\overline{p_{\mathrm{T}}}$ for the same phase space region.

In all cases, the Monte Carlo event generators provide different predictions. PYTHIA tends to slightly overestimate jet activity, and hence underestimate the gap fraction at low Δy and low $\overline{p_{\mathrm{T}}}$, but in general gives the best description of the data. HERWIG++ predicts the gap fraction reasonably well at low values of Δy, but not the number of jets in the same region, and predicts too much jet activity, and so too small a gap fraction, at large values of Δy. ALPGEN shows the largest deviation from the data, predicting too much jet activity, and thus too low a gap fraction, at large values of Δy and $\overline{p_{\mathrm{T}}}$. Some of this deviation can be attributed to the use of HERWIG



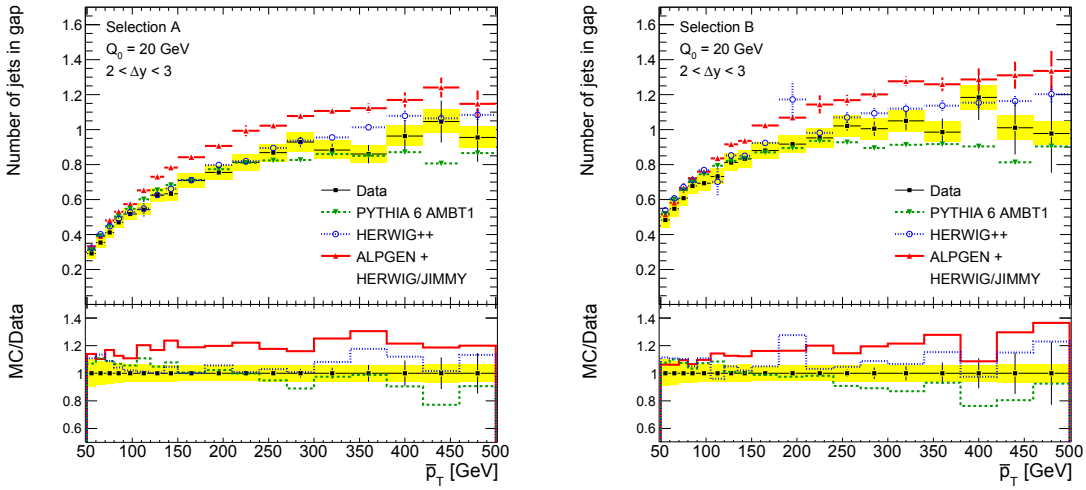
(a) Selection A. Leading p_{T} dijets **(b)** Selection B. Forward/backward dijets

Figure 7.6.: Mean number of jets in the rapidity interval between the boundary jets as a function of Δy, for boundary jets that satisfy $90 \leq \overline{p_{\mathrm{T}}} < 120$ GeV, for (a) selection A and (b) selection B. The unfolded data are compared to predictions from three leading-order Monte Carlo generators. The data and theory are presented in the same way as in Figure 7.5.



(a) Selection A. Leading p_{T} dijets **(b)** Selection B. Forward/backward dijets

Figure 7.7.: Gap fraction as a function of $\overline{p_{\mathrm{T}}}$ for boundary jets that satisfy $2 \leq \Delta y < 3$ for (a) selection A and (b) selection B. The unfolded data are compared to predictions from three leading-order Monte Carlo generators. The data and theory are presented in the same way as in Figure 7.5.



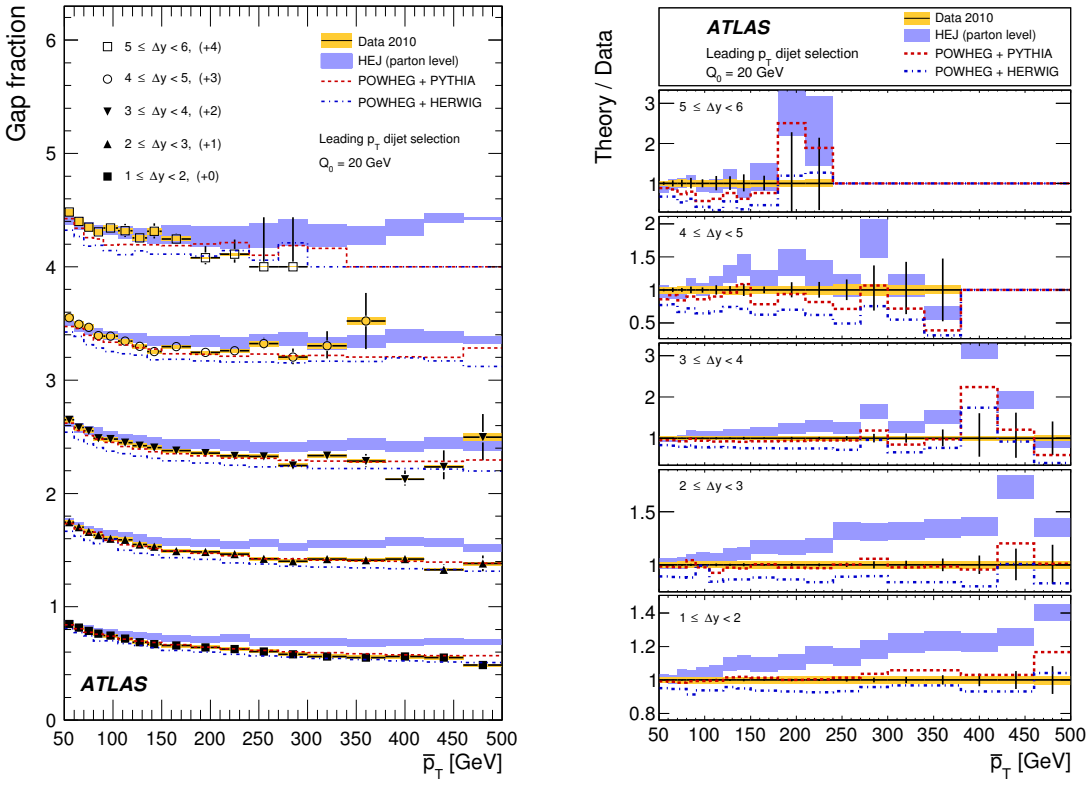
(a) Selection A. Leading p_{T} dijets **(b)** Selection B. Forward/backward dijets

Figure 7.8.: Mean number of jets in the rapidity interval between the boundary jets as a function of $\overline{p_{\mathrm{T}}}$, for boundary jets that satisfy $2 \leq \Delta y < 3$, for (a) selection A and (b) selection B. The unfolded data are compared to predictions from three leading-order Monte Carlo generators. The data and theory are presented in the same way as in Figure 7.5.

and JIMMY for parton shower, hadronisation and underlying event - HERWIG and JIMMY implement similar algorithms to those in HERWIG++, which also predicts more jet activity than is observed in the data.

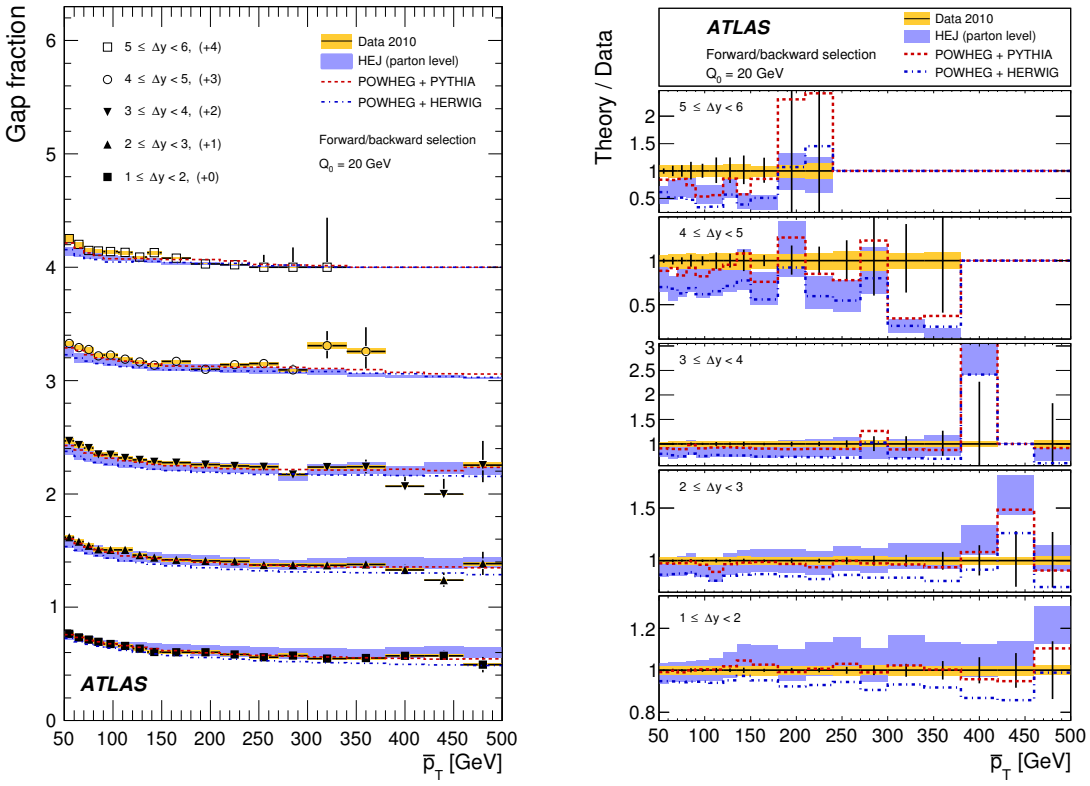
The unfolded data, for all phase space regions, is compared to the theoretical predictions supplied by POWHEG and HEJ. The HEJ predictions are presented as a band in order to represent the theoretical uncertainty due to scale and PDF choices. The POWHEG predictions are presented after parton showering, hadronisation and underlying event simulation with either PYTHIA or HERWIG. The difference between these two predictions is much larger than the uncertainty obtained by varying the PDFs or the renormalisation and factorisation scales and both curves are therefore shown as a conservative estimate of the uncertainty due to higher order effects. The experimental uncertainties are much smaller than the theoretical uncertainty on the HEJ prediction except when both Δy and $\overline{p_{\mathrm{T}}}$ are large.

Figures 7.9 and 7.10 show the gap fraction as a function of $\overline{p_{\mathrm{T}}}$ for selections A and B, respectively. A particularly striking feature is that the parton level HEJ prediction, based on BFKL resummation, predicts too little jet activity, and hence too large a gap fraction, at large values of $\overline{p_{\mathrm{T}}}/Q_0$ for selection A. For selection B, however,



(a) Δy slices **(b)** Ratio to theory

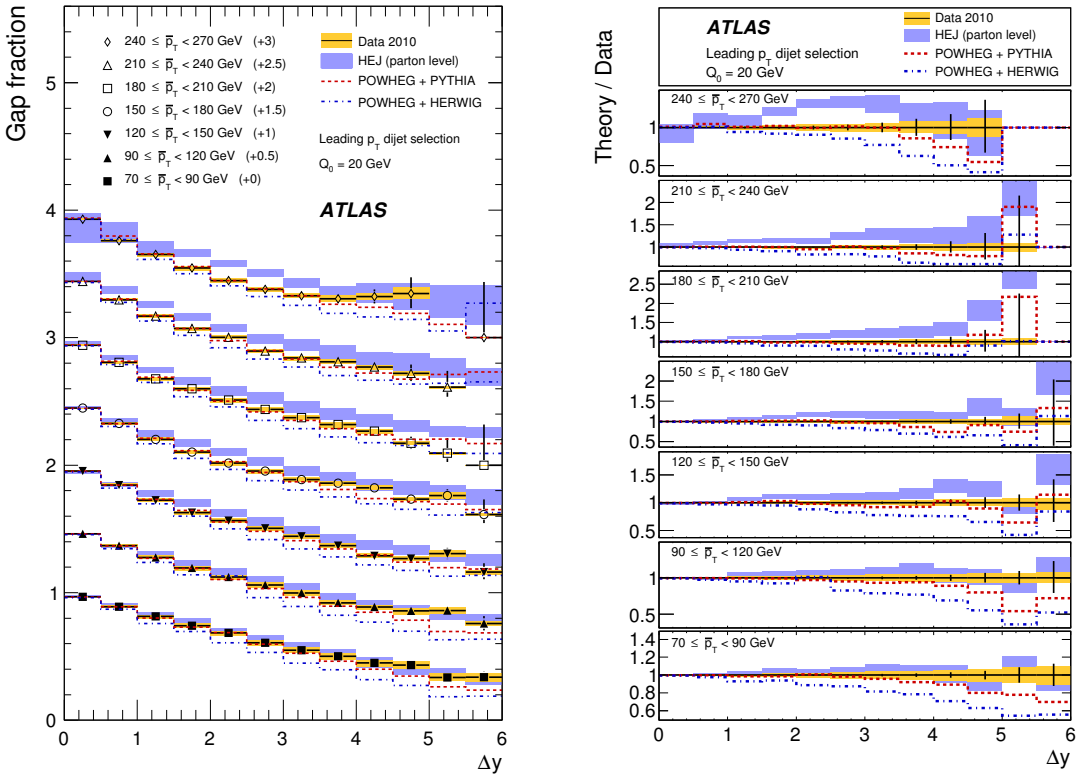
Figure 7.9.: Gap fraction as a function of $\overline{p_{\mathrm{T}}}$ for five Δy slices. (a) shows selection A data against the HEJ and POWHEG generators, while (b) shows the ratio of the theory predictions to the data. The unfolded data are the black points, with error bars representing the statistical uncertainty. The systematic uncertainty on the measurement is represented by the yellow band. The light, solid band represents the theoretical uncertainty in the HEJ calculation from variation of the PDF and renormalisation/factorisation scales. The red and blue dotted lines represent the POWHEG predictions after showering, hadronisation and underlying event simulation with PYTHIA (tune AMBT1) and HERWIG +JIMMY (tune AUET1), respectively.



(a) Δy slices **(b)** Ratio to theory

Figure 7.10.: Gap fraction as a function of $\overline{p_{\mathrm{T}}}$ for five Δy slices. (a) shows selection B data against the HEJ and POWHEG generators, while (b) shows the ratio of the theory predictions to the data. The data and theory are presented in the same way as Figure 7.9.

the HEJ prediction shows much better agreement with the data; this behaviour is expected since the HEJ formalism assumes a large rapidity gap between the boundary jets.

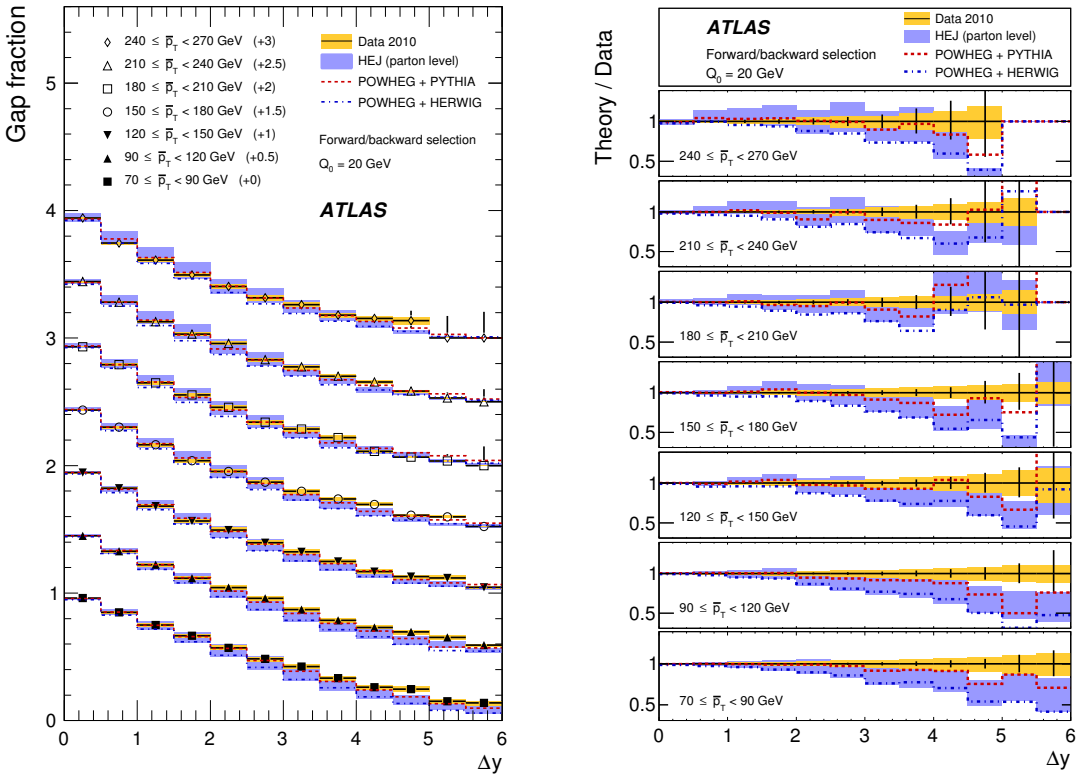


(a) $\overline{p_{\mathrm{T}}}$ slices **(b)** Ratio to theory

Figure 7.11.: Gap fraction as a function of Δy for seven $\overline{p_{\mathrm{T}}}$ slices. (a) shows selection A data against the HEJ and POWHEG generators, while (b) shows the ratio of the theory predictions to the data. The data and theory are presented in the same way as Figure 7.9.

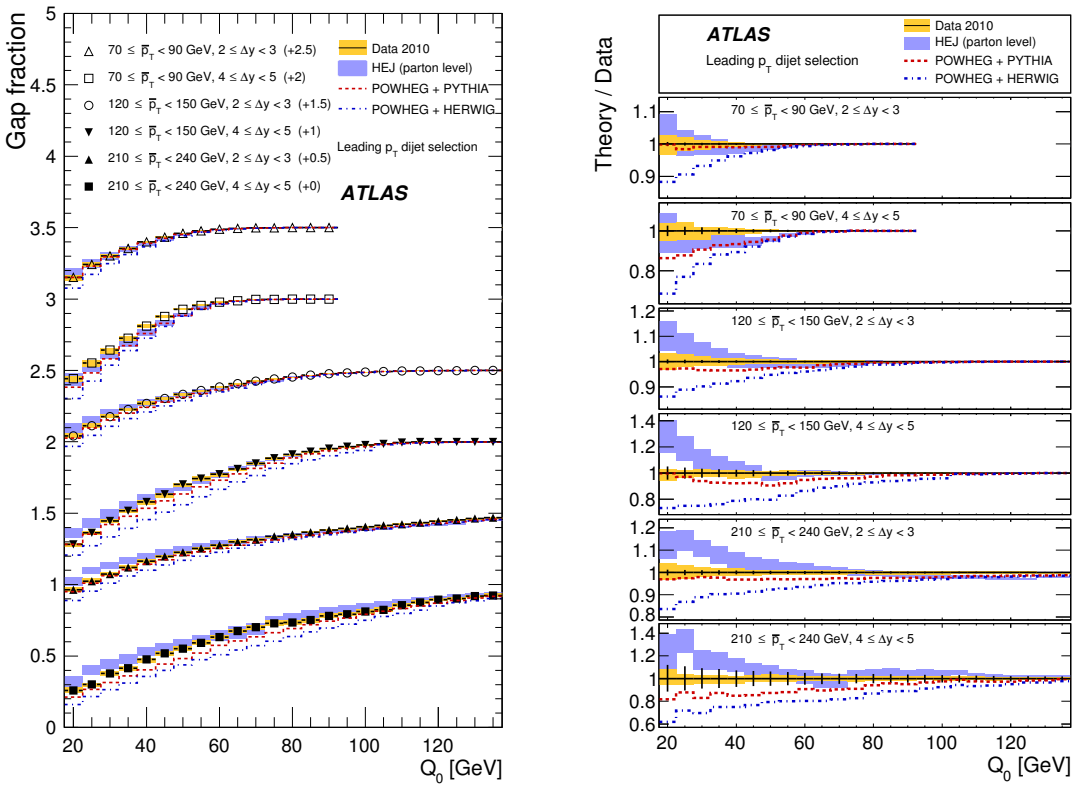
Figures 7.11 and 7.12 show the gap fraction as a function of Δy. Again, the HEJ prediction deviates from the data, here specifying too much jet activity, and so too low a gap fraction at large Δy when the boundary jets are selected as the most forward and most backward jets in the event.

Figure 7.13 shows the gap fraction as a function of the veto scale, Q_0 for selection A data. Figures 7.14 and 7.15 show the mean number of jets in the gap region as a function of $\overline{p_{\mathrm{T}}}$ for selections A and B, respectively. Again HEJ can be seen to perform poorly at high $\overline{p_{\mathrm{T}}}$ and high Δy.



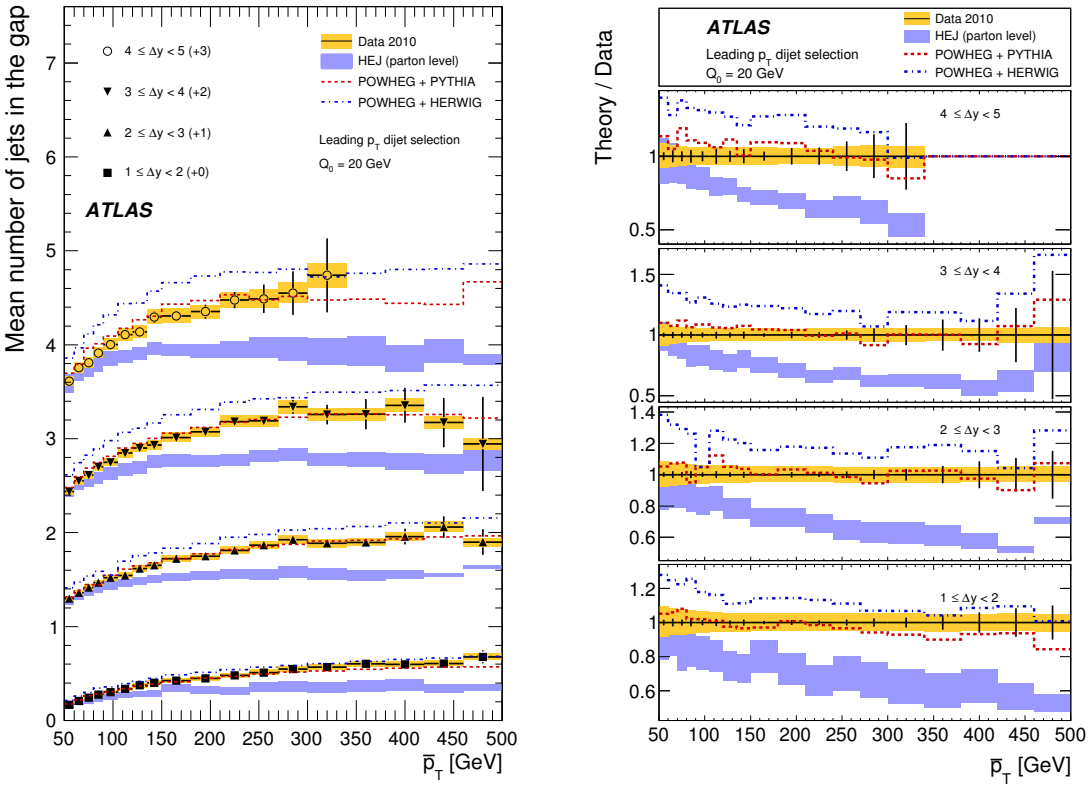
(a) $\overline{p_{\mathrm{T}}}$ slices **(b)** Ratio to theory

Figure 7.12.: Gap fraction as a function of Δy for seven $\overline{p_{\mathrm{T}}}$ slices. (a) shows selection B data against the HEJ and POWHEG generators, while (b) shows the ratio of the theory predictions to the data. The data and theory are presented in the same way as Figure 7.9.



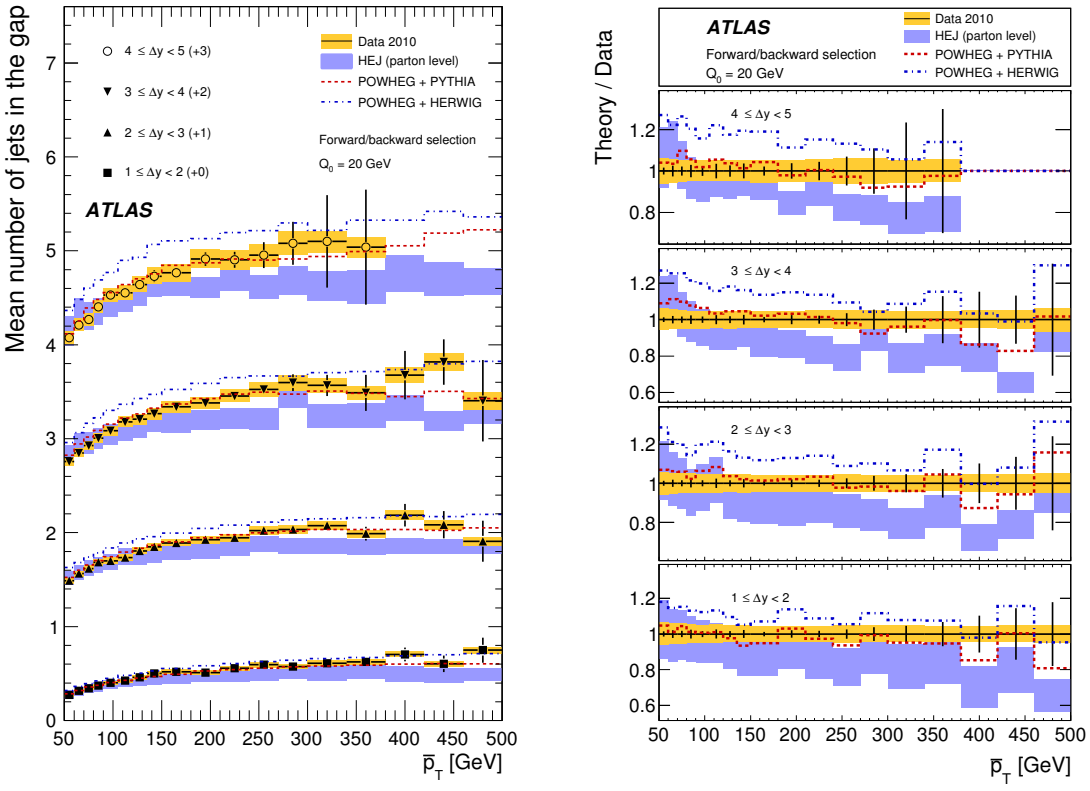
(a) $\overline{p_{\mathrm{T}}}$ and Δy slices **(b)** Ratio to theory

Figure 7.13.: Gap fraction as a function of Q_0 for six different $\overline{p_{\mathrm{T}}}$ and Δy slices. (a) shows selection A data against the HEJ and POWHEG generators, while (b) shows the ratio of the theory predictions to the data. The data points for $Q_0 > \overline{p_{\mathrm{T}}}$ have been removed because the gap fraction is always equal to one for this dijet selection, by definition. The data and theory are presented in the same way as Figure 7.9.



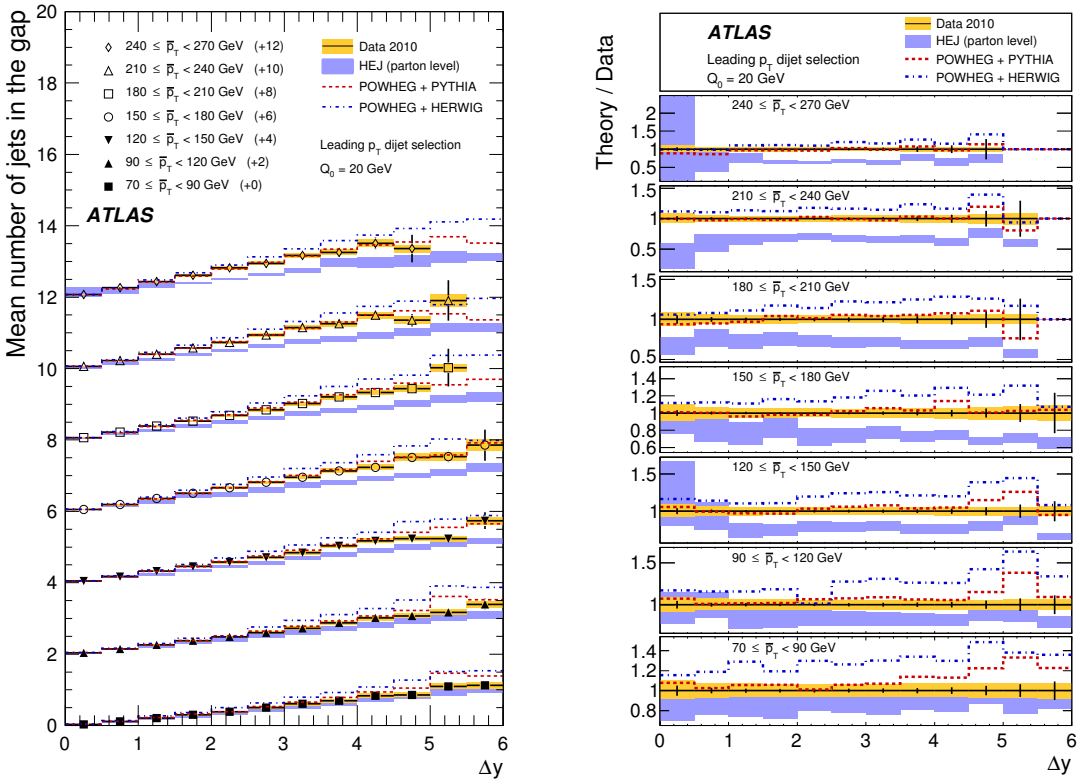
(a) Δy slices **(b)** Ratio to theory

Figure 7.14.: Mean number of jets in the gap as a function of $\overline{p_{\mathrm{T}}}$ for four Δy slices. (a) shows selection A data against the HEJ and POWHEG generators, while (b) shows the ratio of the theory predictions to the data. The data and theory are presented in the same way as Figure 7.9.



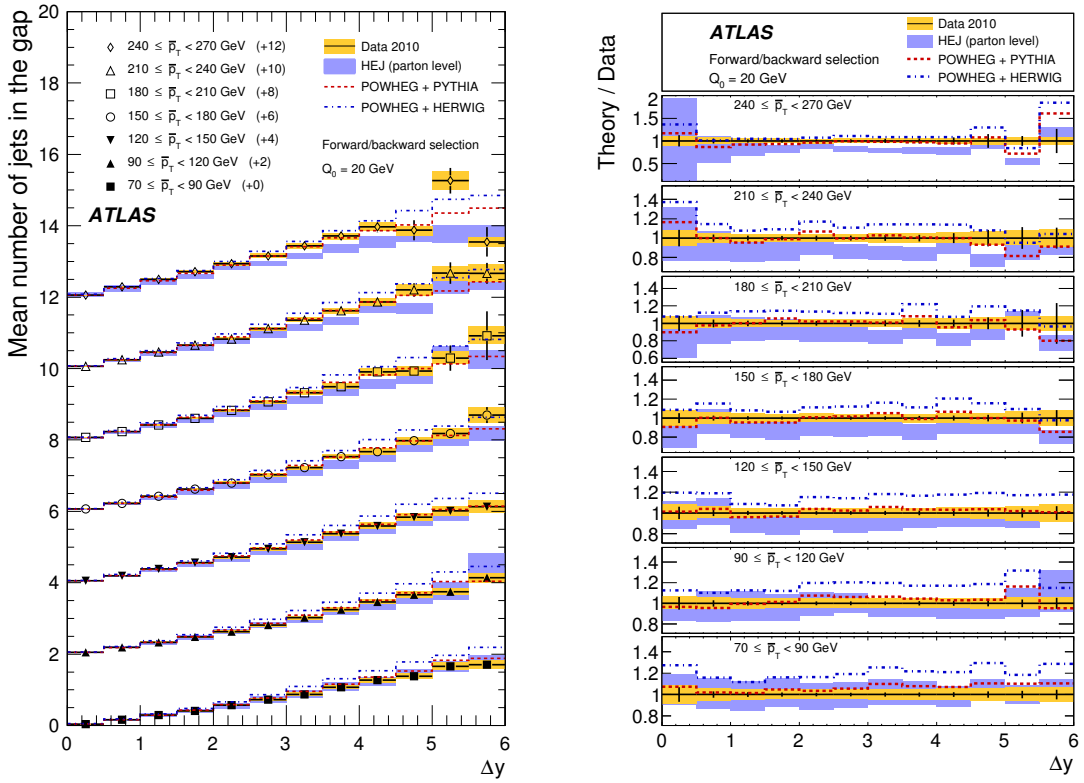
(a) Δy slices **(b)** Ratio to theory

Figure 7.15.: Mean number of jets in the gap as a function of $\overline{p_{\mathrm{T}}}$ for four Δy slices. (a) shows selection B data against the HEJ and POWHEG generators, while (b) shows the ratio of the theory predictions to the data. The data and theory are presented in the same way as Figure 7.9.



(a) $\overline{p_{\mathrm{T}}}$ slices **(b)** Ratio to theory

Figure 7.16.: Mean number of jets in the gap as a function of Δy for seven $\overline{p_{\mathrm{T}}}$ slices. (a) shows selection A data against the HEJ and POWHEG generators, while (b) shows the ratio of the theory predictions to the data. The data and theory are presented in the same way as Figure 7.9.



(a) $\overline{p_{\mathrm{T}}}$ slices **(b)** Ratio to theory

Figure 7.17.: Mean number of jets in the gap as a function of Δy for seven $\overline{p_{\mathrm{T}}}$ slices. (a) shows selection B data against the HEJ and POWHEG generators, while (b) shows the ratio of the theory predictions to the data. The data and theory are presented in the same way as Figure 7.9.

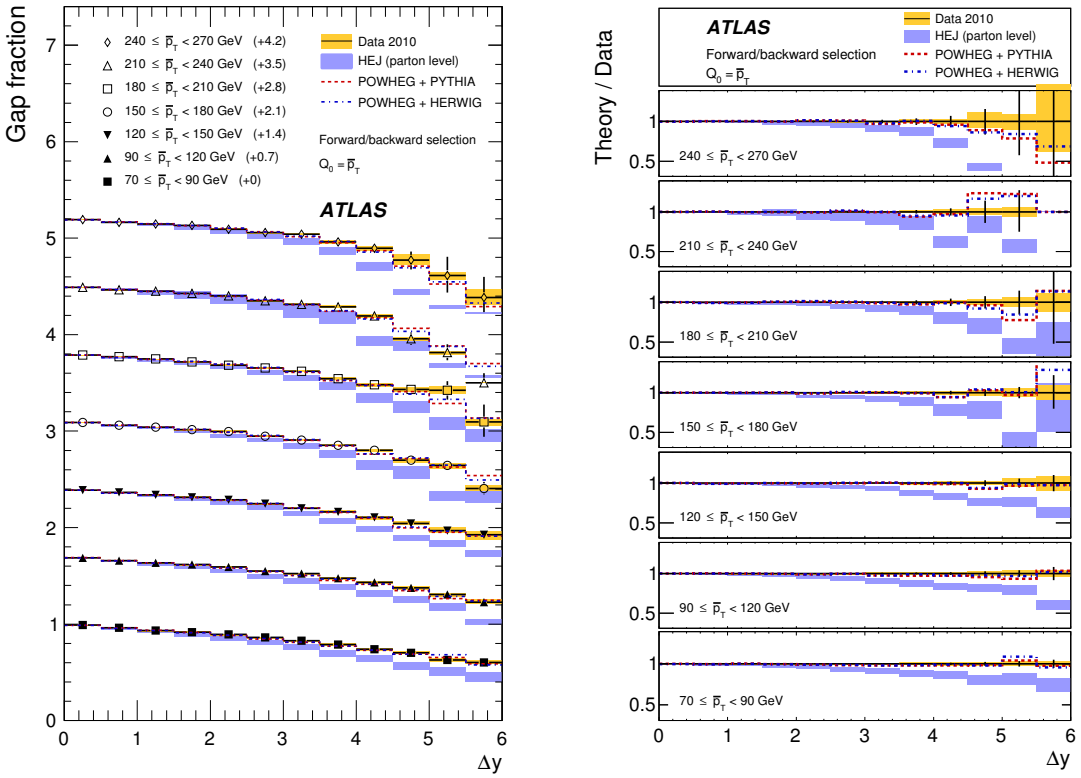
Finally Figures 7.16 and 7.17 show the mean number of jets in the gap region as a function of Δy. HEJ does not describe the data at large values of $\overline{p_{\mathrm{T}}}$ for selection A (Figures 7.9 and 7.14). This is expected, since the underlying calculation is an all order resummation, which accounts for terms proportional to Δy but does not contain all the terms that become important as $\overline{p_{\mathrm{T}}}/Q_0$ increases. The large $\overline{p_{\mathrm{T}}}$ region is slightly better described for selection B (Figures 7.10 and 7.15). The HEJ generator agrees with the data in the low to medium Δy range for both selections (Figures 7.11 and 7.12), while underestimating the gap fraction at large Δy for selection B. On the other hand, the mean number of jets is better described by selection B (Figure 7.17), with larger discrepancies seen in selection A (Figure 7.16). Work is currently in progress to interface the HEJ generator with a parton shower and hadronisation programs. In principle such matching could help describe the QCD radiation for large values of $\overline{p_{\mathrm{T}}}/Q_0$ [77].

In general, POWHEG describes the data well. There is often a substantial difference between the POWHEG +PYTHIA and POWHEG +HERWIG predictions, with the former showing a better agreement for almost all distributions while POWHEG +HERWIG tends to overproduce jets in the rapidity interval between the boundary jets. The only serious disagreement is observed at large Δy, where POWHEG slightly underestimates the gap fraction in both cases. This indicates that the parton showers may not be recovering terms in the resummation that are important as Δy increases.

Finally, for selection B data, Figure 7.18 shows the gap fraction and Figure 7.19 the mean number of jets in the gap region as a function of Δy, but with the veto scale now set to $Q_0 = \overline{p_{\mathrm{T}}}$. In this case, POWHEG +PYTHIA and POWHEG +HERWIG both describe the data well, implying a smaller dependence on the generator modelling of parton shower, hadronisation and underlying event. The HEJ description, however, becomes a little worse with the increase in veto scale.

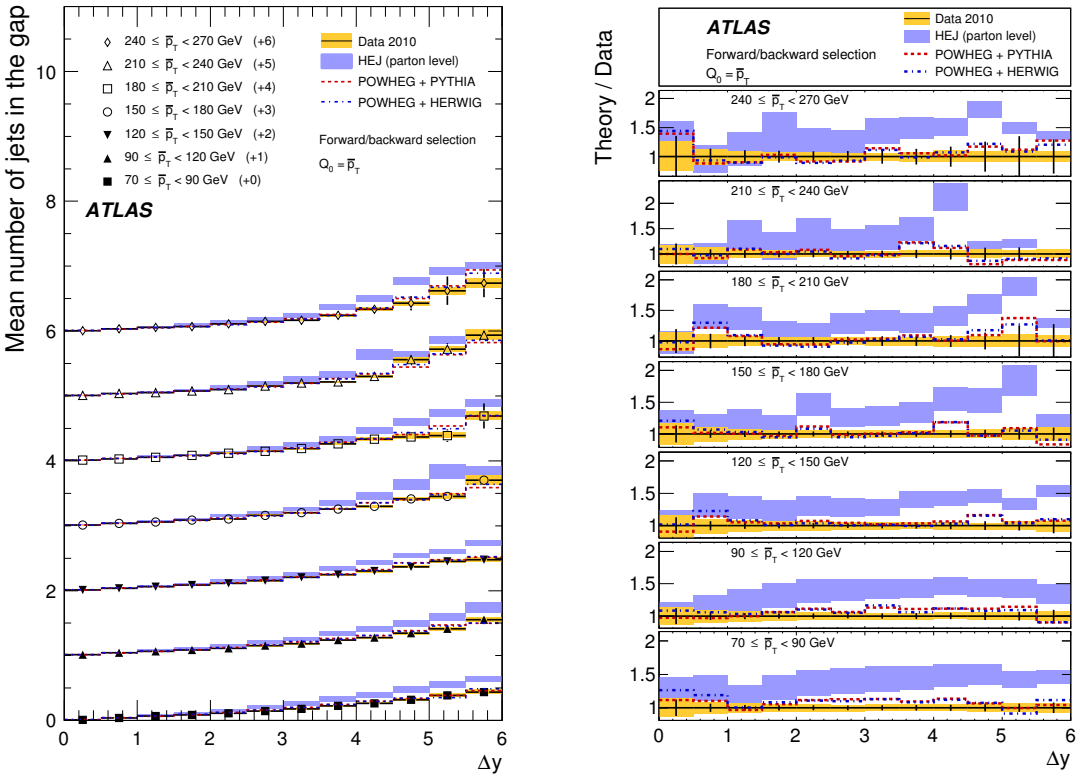
7.8. Summary

These results show the expected behaviour of a reduction of gap events for harder jets and for larger rapidity gaps [74] for both selections. There are some interesting deviations between the data and the various theory calculations, in particular, the parton level HEJ prediction deviates from the data at large Δy and large $\overline{p_{\mathrm{T}}}/Q_0$;



(a) $\overline{p_{\mathrm{T}}}$ slices **(b)** Ratio to theory

Figure 7.18.: Gap fraction as a function of Δy for seven $\overline{p_{\mathrm{T}}}$ slices, but with the veto scale set to $Q_0 = \overline{p_{\mathrm{T}}}$. (a) shows selection B data against the HEJ and POWHEG generators, while (b) shows the ratio of the theory predictions to the data. The data and theory are presented in the same way as Figure 7.9.



(a) $\overline{p_{\mathrm{T}}}$ slices **(b)** Ratio to theory

Figure 7.19.: Mean number of jets in the gap as a function of Δy for seven $\overline{p_{\mathrm{T}}}$ slices, but with the veto scale set to $Q_0 = \overline{p_{\mathrm{T}}}$. (a) shows selection B data against the HEJ and POWHEG generators, while (b) shows the ratio of the theory predictions to the data. The data and theory are presented in the same way as Figure 7.9.

areas of phase space in which it was expected that it should perform well. POWHEG generally describes the data well when interfaced to PYTHIA, except at large Δy, and performs much more poorly when combined with HERWIG.

Experimental uncertainties are smaller than the theoretical ones for most of the phase-space considered here. In addition, the experimental uncertainty is smaller than the spread of Monte Carlo event generator predictions. These results could, therefore, be used to constrain the modelling of QCD radiation in widely separated dijet systems, leading to an improved understanding of the behaviour of jet vetoes for future measurements. This work was published in the Journal of High Energy Physics [78].

Chapter 8.

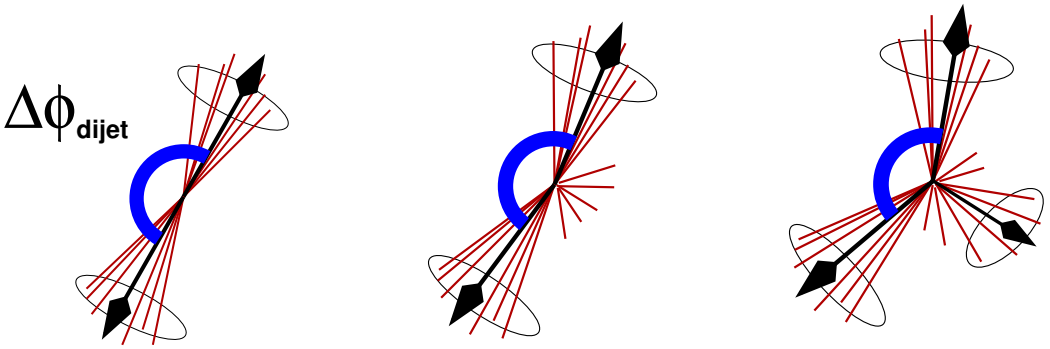
Dijet Azimuthal Decorrelations

"I believe there is a direct correlation between love and laughter."
— Yakov Smirnoff

8.1. Introduction

At leading order, dijet production results in two jets being produced which are completely anti-correlated in azimuthal angle: satisfying $\Delta\phi = \pi$. With the addition of soft radiation or higher order jet production, the opening angle between the jets will deviate from this idealised case (see Figure 8.1). Thus azimuthal decorrelation tests both higher order perturbative QCD and the modelling of non-perturbative soft processes.

Studying azimuthal decorrelations probes similar physics to the dijet with a jet veto study discussed in Chapter 7. It therefore makes sense to attempt to integrate these two measurements by constructing observables which combine angular information with jet vetoes.



(a) $\Delta\phi \simeq \pi$ for dijets **(b)** $\Delta\phi < \pi$ with soft radiation **(c)** $\Delta\phi < \pi$ with third jet

Figure 8.1.: Dijet configurations and resulting $\Delta\phi$ in jet events. As the amount of QCD radiation in the event increases (from left to right), the azimuthal angle (shown in blue) between the two leading jets in the event decreases.

8.2. Event Selection

As was the case for the previously discussed jet veto measurement, events are required to belong to a good run and to have exactly one good primary vertex, as discussed in Section 7.3. Jets are then reconstructed using one of the ATLAS standard jet parameter sets: anti-k_t jets with $R = 0.6$ (see Section 1.10.2 and Section 3.2). Events are rejected if they contain any jets with $p_{\mathrm{T}} > 20$ GeV that are flagged as "bad" or "ugly" by the standard loose jet cleaning cuts.

After these cuts have been applied, the boundary jets are taken as the two highest p_{T} jets with $|y| < 4.4$. Provided that the leading jet satisfies $p_{\mathrm{T}} > 60$ GeV and the second jet $p_{\mathrm{T}} > 50$ GeV, the event is accepted as an inclusive event. Gap events are defined as the subset of inclusive events that do not contain an additional jet with p_{T} greater than the veto scale, $Q_0 = 20$ GeV.

8.3. Observables

This analysis measures quantities in terms of the gap fraction, that fraction of accepted events which do not possess a third jet with $p_{\mathrm{T}} > Q_0$. The gap fraction is measured as a function of both Δy and the jet veto scale, Q_0.

The decorrelation between $\Delta\phi$ and Δy is primarily studied through the variables $\langle\cos(\pi - \Delta\phi)\rangle$ and $\langle\cos(2\Delta\phi)\rangle$. Theoretical predictions [79] indicate that these quantities could provide discriminating power to distinguish between DGLAP and BFKL-like evolutions: distributions of $\langle\cos(\pi - \Delta\phi)\rangle$ and $\langle\cos(2\Delta\phi)\rangle$ are therefore constructed as functions of Δy.

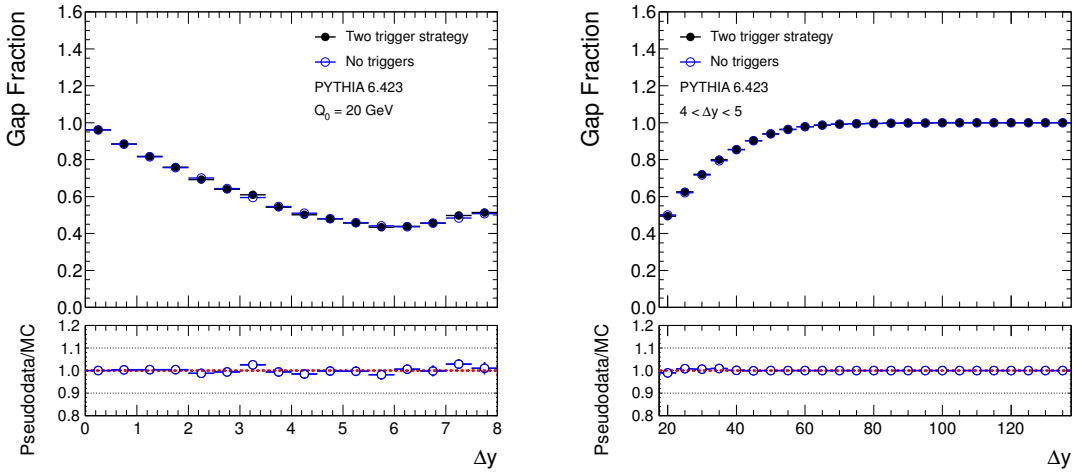
Finally, double differential cross-sections as a function of Δy and $\Delta\phi$, $\cos(\Delta\phi)$ or $\cos(2\Delta\phi)$ are calculated separately for gap and inclusive events, using a jet veto at $Q_0 = 20$ GeV.

8.4. Trigger Strategy

As this measurement extends out to $\Delta y = 8$, the strategy used in Section 7.4 cannot be used, since the assumption that all events will have at least one central jet is no longer valid. The trigger strategy used to measure the dijet cross-section is, however, applicable since this uses both the central and forward trigger systems depending on the p_{T} and rapidity of the two leading jets (see Section 6.4).

Although parameters like the optimum values for the matching cuts do not have to be rederived, a closure test is performed with Monte Carlo to demonstrate that the gap fraction is not biased when adopting this trigger strategy. The effects of the trigger are emulated using the prescale values of a typical run. Monte Carlo events are discarded at random according to these prescales and the surviving events constitute a "triggered pseudodata" sample, which is then analysed with the same procedure as is used for data.

The resulting gap fraction spectrum is shown in Figure 8.2a as a function of the rapidity difference, Δy, and in Figure 8.2b as a function of the veto scale, Q_0. The pseudodata distributions obtained after trigger emulation and correction for prescale are compatible within $\sim 1\%$ with the same distributions from the original Monte Carlo sample without the trigger requirement, validating the two trigger method for this measurement.



(a) Closure as a function of Δy

(b) Closure as a function of Q_0

Figure 8.2.: Summary plot of the closure test for a dijet Monte Carlo sample for jets identified using the anti-k_t algorithm with $R = 0.6$. The black dots represent the result of emulating the trigger in the Monte Carlo and then correcting for it using the same technique used on data, while the blue dotted lines represent the result of analysing all events, without trigger corrections. This test is conducted (a) as a function of Δy and (b) as a function ofQ_0.

Period	A	B	C	D	E	F	G	H	I
Efficiency	1.000	0.929	0.932	0.573	0.439	0.368	0.302	0.296	0.213

Table 8.1.: Efficiency of the single vertex cut as a function of data taking period. The effects of this inefficiency are corrected for by reducing the effective luminosities in each period by the same fraction.

8.5. Single Vertex Cut Efficiency

In order to measure cross-sections, the impact of the single vertex cut needs to be accounted for. In early periods, with lower beam intensities, the average number of collisions per bunch crossing was low and hence the cut efficiency was high. In later periods, increasing pile-up means that the efficiency drops to about 20%. To obtain accurate cross-section measurements, a period-by-period correction to the observed luminosity needs to be applied, to remove the effects of this inefficiency. Table 8.1 shows the efficiency of this cut as a function of period; the observed luminosities are reduced by the same fraction before the final cross-section results are produced.

8.6. Unfolding Detector Effects

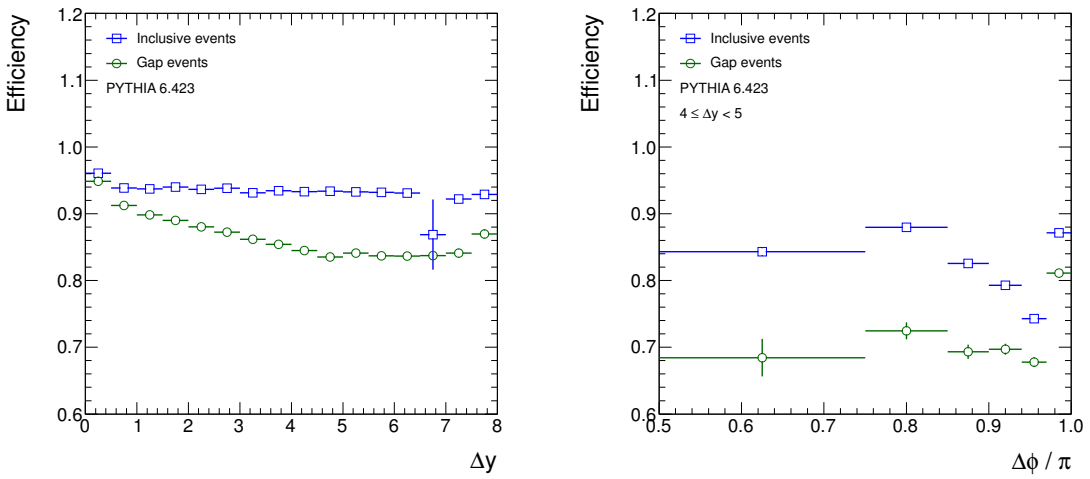
The measured distributions are corrected for detector effects using the Bayesian unfolding scheme as detailed in Section 3.4. Measuring the efficiency, the extent to which particle level events remain in the same bin at detector level, and the purity, the level of contamination in detector level bins, for each bin is essential in order to determine the optimum binning for each distribution: the bin sizes must be chosen such that the efficiency and the purity of each bin remains high.

Here, when binned in one variable, either Δy or $\Delta\phi$, bin sizes were altered until both efficiency and purity were above 0.8 in all bins. When binned in both Δy and $\Delta\phi$ simultaneously, this requirement was relaxed; both efficiency and purity were required to be above 0.6 in all bins.

Sample efficiency distributions as a function of Δy and $\Delta\phi$ can be seen in Figure 8.3, with purity distributions as a function of Δy and $\Delta\phi$ in Figure 8.4. The trend with respect to Δy that can be seen in the efficiency and purity for gap events, an initial decrease followed by a flattening out, is explained by the kinematics of the observable. Due to the steeply falling $p_{\mathrm{T},3}$ distribution, it is likely that events which satisfy the gap criteria at particle level, will contain a third jet with $p_{\mathrm{T},3}$ just below the veto scale; small differences in p_{T} between particle level and detector level can mean that such events will fail the gap event selection cuts at detector level. As Δy increases, the chances of this happening increase too, before flattening out due to the reduced phase-space which is available after requiring two hard high y jets - this effect can also be seen in Figure 8.12, which shows the distribution of number of jets in the gap.

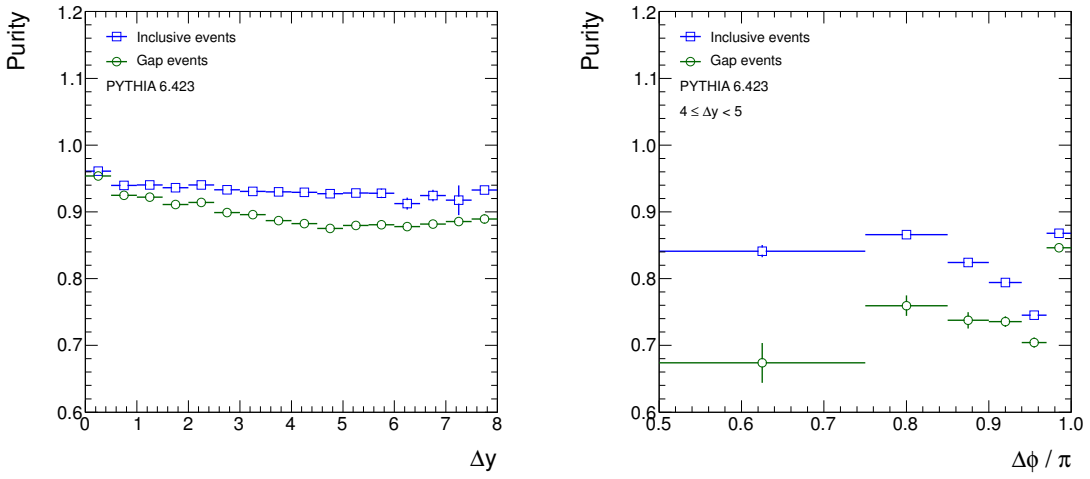
Figure 8.5 shows the level of agreement between PYTHIA, HERWIG++, ALPGEN and the data for two sample distributions: the gap fraction as a function of Δy and as a function of Q_0. The only Monte Carlo with good statistics and good agreement with data in all regions of phase space considered here is PYTHIA. Accordingly, these PYTHIA samples are used to unfold the data.

Finally, unfolded distributions are calculated using the RooUnfold framework [80]. This allows simultaneous comparison of multiple different unfolding methods, with the same Monte Carlo events entering each calculation. Here simple bin-by-bin unfolding is compared to an iterative Bayesian unfolding method [53] and to SVD.



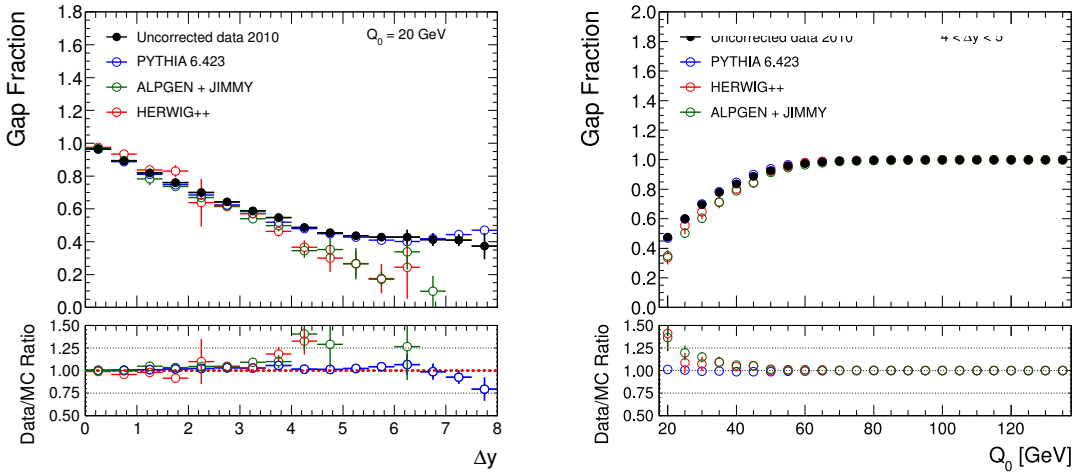
(a) Δy efficiency **(b)** $\Delta\phi$ efficiency for $4 \leq \Delta y < 5$

Figure 8.3.: Efficiency distributions showing the proportion of events at particle level which remain in each bin at detector level. (a) shows the efficiency as a function of Δy, while (b) shows the efficiency as a function of $\Delta\phi$ in the region $4 \leq \Delta y < 5$. Efficiencies are shown separately for inclusive events (blue) and gap events (green). Efficiencies are calculated using PYTHIA Monte Carlo for jets identified using the anti-k_t algorithm with $R = 0.6$.



(a) Δy purity **(b)** $\Delta\phi$ purity for $4 \leq \Delta y < 5$

Figure 8.4.: Purity distributions showing the proportion of events in each bin at detector level which came from the same bin at particle level. (a) shows is the purity as a function of Δy, while (b) shows the purity as a function of $\Delta\phi$ in the region $4 \leq \Delta y < 5$. Purities are shown for inclusive events (blue) and gap events (green). Purities are calculated using PYTHIA Monte Carlo for jets identified using the anti-k_t algorithm with $R = 0.6$.



(a) Gap fraction as a function of Δy **(b)** Gap fraction as a function of Q_0

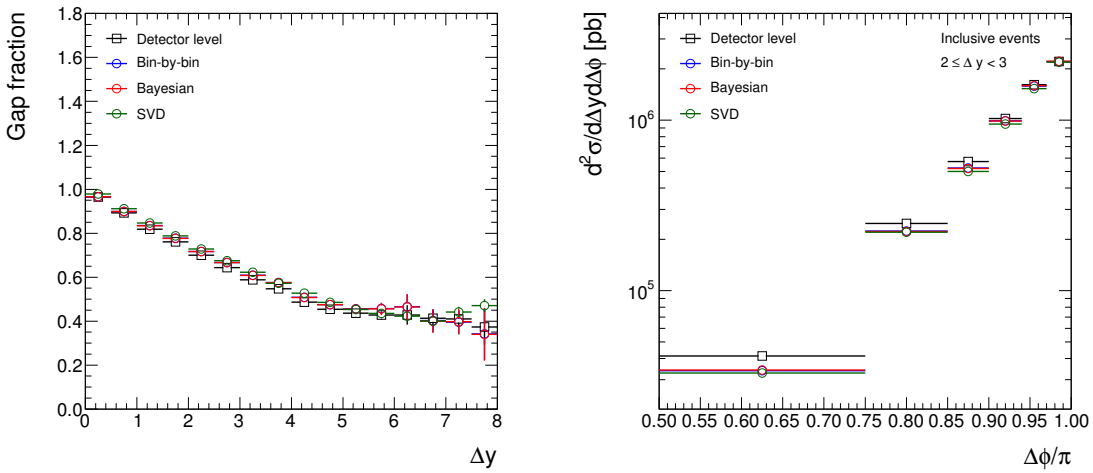
Figure 8.5.: Gap fraction distributions as a function of Q_0 and Δy for jets identified using the anti-k_t algorithm with $R = 0.6$. (a) shows the gap fraction as a function of Δy for $Q_0 = 20$ GeV, while (b) shows the gap fraction as a function of Q_0 in the region $4 \leq \Delta y < 5$. The uncorrected data are compared to the leading order PYTHIA 6.423, HERWIG++ and ALPGEN predictions after these have been passed through the ATLAS detector simulation software. The error bars indicate the statistical uncertainty on the measurement.

Figure 8.6 shows sample distributions, comparing data unfolded using each of these three methods to uncorrected, detector level data. Good agreement is seen between each of the unfolding methods, with the shape of the unfolded distributions also agreeing well with the detector level data.

8.7. Azimuthal Decorrelation Results

8.7.1. Comparison to Leading Order Monte Carlo Generators

For these measurements, data are initially compared to leading-order Monte Carlo generators. predictions from next-to-leading order generators, specifically POWHEG and HEJ as in Chapter 7, are also shown, but the focus of this thesis is, anyway, on the experimental data. Unfolded data are compared against PYTHIA, HERWIG++ and ALPGEN, all generators that are commonly used for predictive purposes in ATLAS.



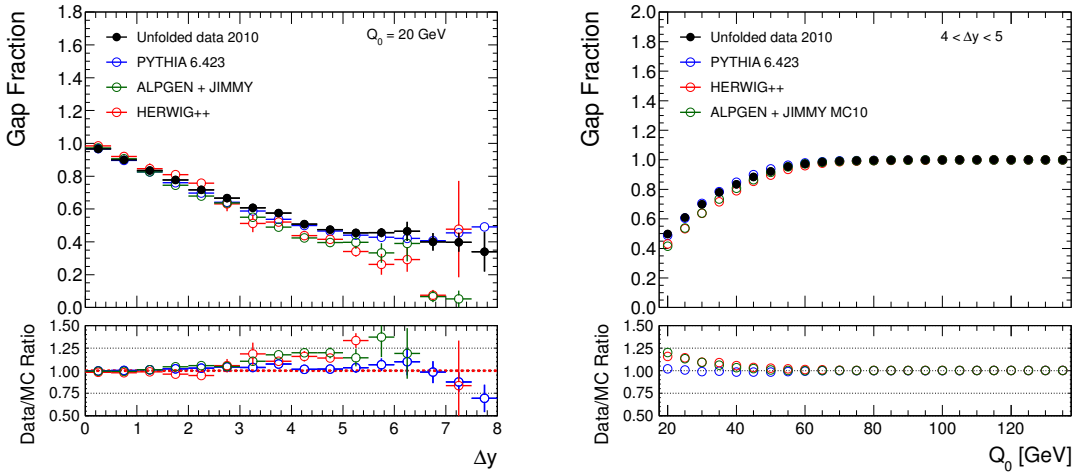
(a) Unfolded gap fraction as a function of Δy

(b) Unfolded cross-section as a function of $\Delta\phi$ for $2 \leq \Delta y < 3$

Figure 8.6.: Comparison of three different unfolding methods: bin-by-bin, Bayesian and SVD, showing the result of correcting measured data for detector effects together with the uncorrected spectrum for comparison. (a) shows the unfolded distributions as a function of Δy, while (b) shows the unfolded cross-section distribution as a function of $\Delta\phi$ for the region $2 \leq \Delta y < 3$. Unfolding is performed using PYTHIA Monte Carlo for jets identified using the anti-k_t algorithm with $R = 0.6$.

For both data and Monte Carlo, the errors shown reflect statistical uncertainties only.

Figure 8.7 shows the gap fraction as a function of Δy and as a function of Q_0 in the case that the boundary jets satisfy $4 \leq \Delta y < 5$. Divergences between the Monte Carlo predictions can be seen at high Δy and at low Q_0, although PYTHIA agrees well throughout.



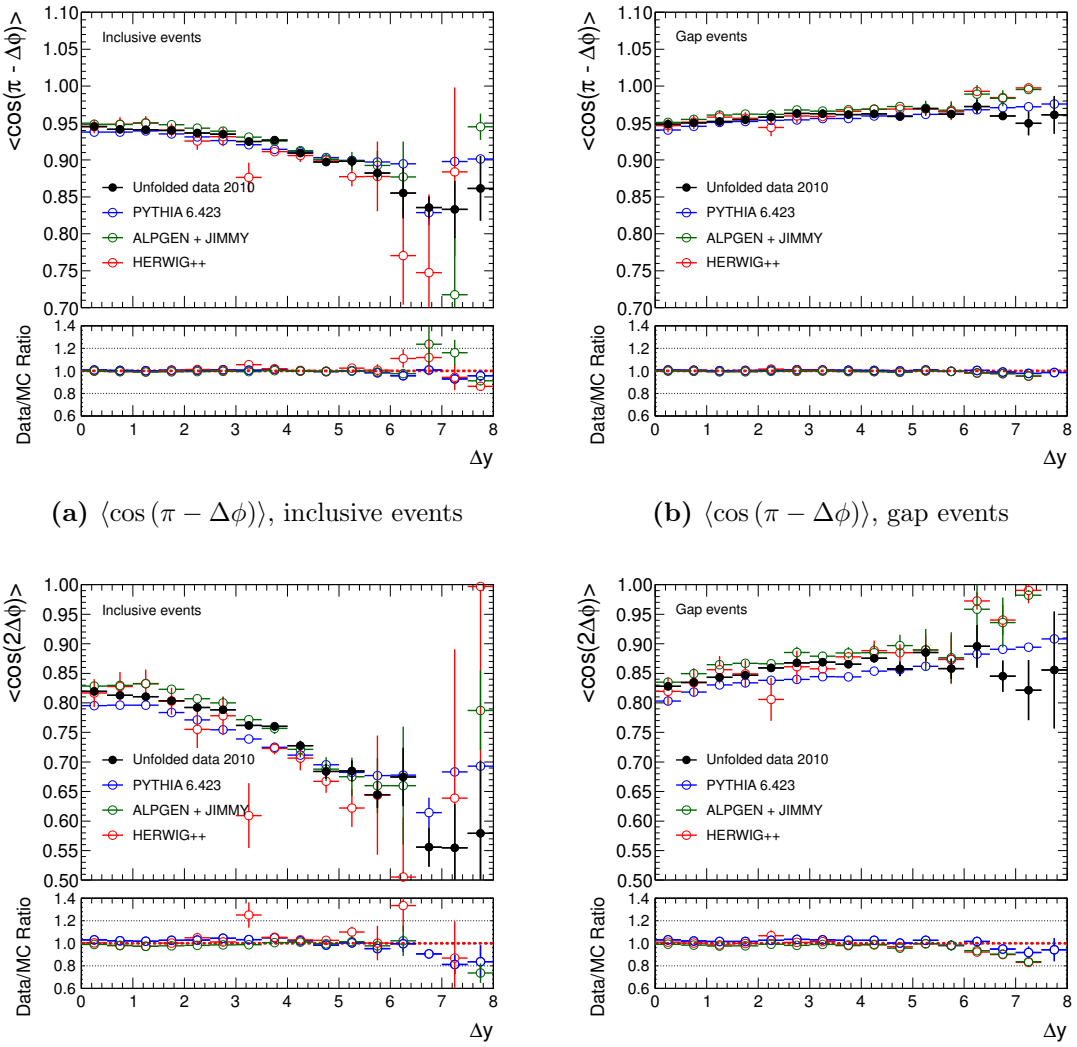
(a) Gap fraction as a function of Δy

(b) Gap fraction as a function of Q_0

Figure 8.7.: Gap fraction distributions as a function of Q_0 and Δy for jets identified using the anti-k_t algorithm with $R = 0.6$. (a) shows the gap fraction as a function of Δy for $Q_0 = 20$ GeV, while (b) shows the gap fraction as a function of Q_0 in the region $4 \leq \Delta y < 5$. The unfolded data are compared to the leading order particle level PYTHIA 6.423, HERWIG++ and ALPGEN predictions. The error bars indicate the statistical uncertainty on the measurement.

Figure 8.8 shows the $\langle \cos(\pi - \Delta\phi) \rangle$ and $\langle \cos(2\Delta\phi) \rangle$ distributions as a function of Δy for both inclusive and gap events. In general, there is good agreement between the data and the different Monte Carlo generators, with the major areas of disagreement coming at high Δy, where statistics are poorer. These disagreements are only present in the inclusive sample; once the jet veto is applied the different Monte Carlo predictions agree well.

Figures 8.9, 8.10 and 8.11 show the $\Delta\phi$, $\cos(\Delta\phi)$ and $\cos(2\Delta\phi)$ cross-sections respectively, for six different Δy bins. Inclusive events and gap events are shown separately. Only comparison between PYTHIA and data is shown here, since the



(a) $\langle \cos(\pi - \Delta\phi) \rangle$, inclusive events

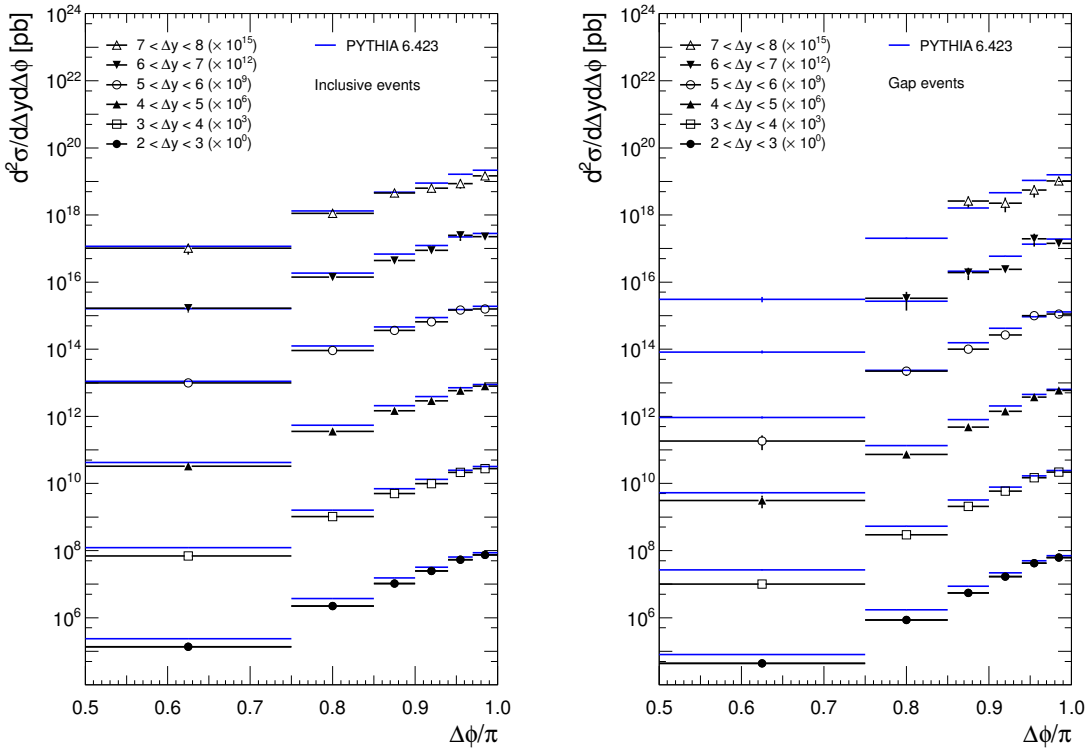
(b) $\langle \cos(\pi - \Delta\phi) \rangle$, gap events

(c) $\langle \cos(2\Delta\phi) \rangle$, inclusive events

(d) $\langle \cos(2\Delta\phi) \rangle$, gap events

Figure 8.8.: Distributions of $\langle \cos(\pi - \Delta\phi) \rangle$ (top) and $\langle \cos(2\Delta\phi) \rangle$ (bottom) shown as a function of Δy for jets identified using the anti-k_t algorithm with $R = 0.6$. Inclusive events (left) and gap events (right) are shown separately. The unfolded data are compared to the leading order particle level PYTHIA 6.423, HERWIG++ and ALPGEN predictions. The error bars indicate the statistical uncertainty on the measurement.

other leading order Monte Carlo generators suffer from poor statistics at large Δy. Bins in which only a single event is present have been removed from these plots.

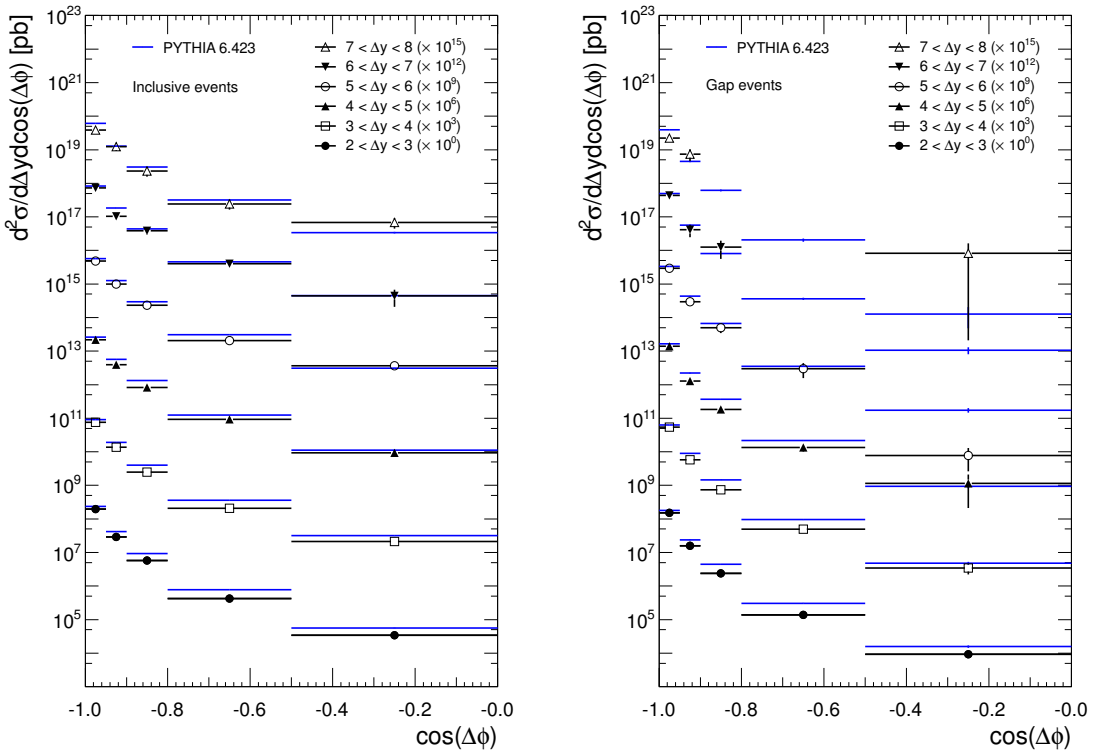


(a) Cross-section as a function of $\Delta\phi$ for inclusive events

(b) Cross-section as a function of $\Delta\phi$ for gap events

Figure 8.9.: Double-differential cross-section as a function of $\Delta\phi$ in different regions of Δy. The cross-section is shown for jets identified using the anti-k_t algorithm with $R = 0.6$. (a) shows inclusive events while (b) shows gap events. The unfolded data are compared to the leading order particle level PYTHIA 6.423 prediction. In each case, the error bars indicate the statistical uncertainty on the measurement only. Statistically insignificant data points at large Δy are omitted.

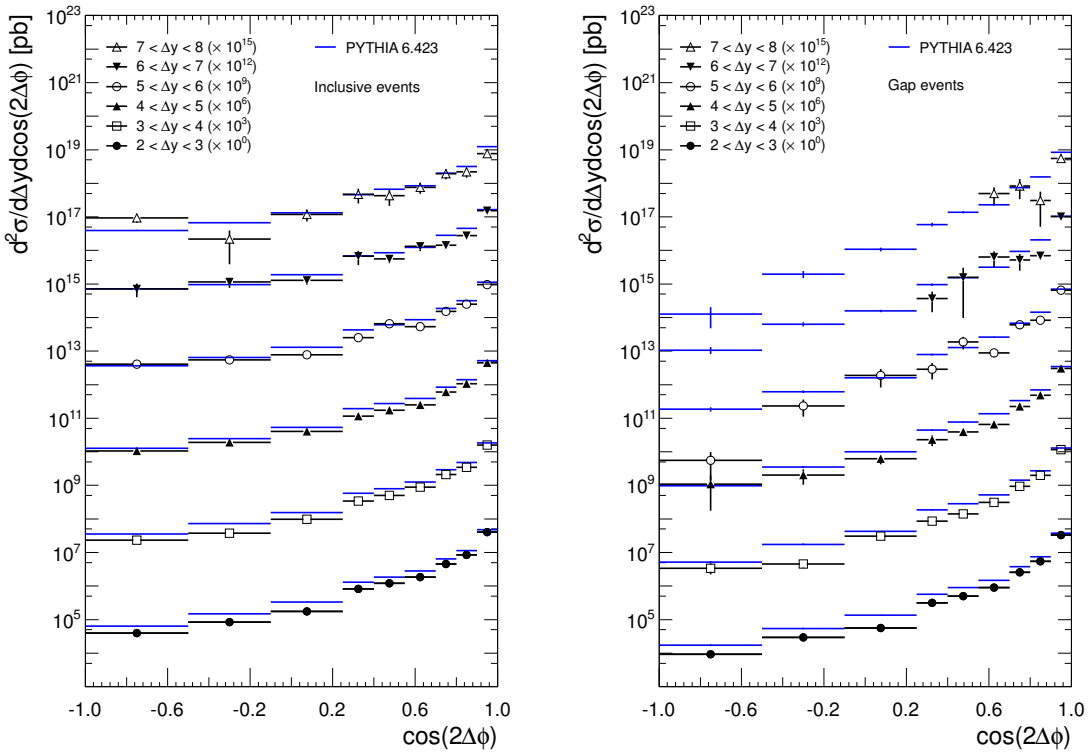
Finally, the distribution of number of jets in the gap, as a function of Δy, is shown in Figure 8.12. Here, divergences from the leading order Monte Carlo predictions can be seen at high Δy, although PYTHIA still provides the best agreement with the data.



(a) Cross-section as a function of $\cos(\Delta\phi)$ for inclusive events

(b) Cross-section as a function of $\cos(\Delta\phi)$ for gap events

Figure 8.10.: Double-differential cross-section as a function of $\cos(\Delta\phi)$ in different regions of Δy. The cross-section is shown for jets identified using the anti-k_t algorithm with $R = 0.6$. (a) shows inclusive events, while (b) shows gap events. The unfolded data are compared to the leading order particle level PYTHIA 6.423 prediction. In each case, the error bars indicate the statistical uncertainty on the measurement only. Statistically insignificant data points at large Δy are omitted.



(a) Cross-section as a function of $\cos(2\Delta\phi)$ for inclusive events

(b) Cross-section as a function of $\cos(2\Delta\phi)$ for gap events

Figure 8.11.: Double-differential cross-section as a function of $\cos(2\Delta\phi)$ in different regions of Δy. The cross-section is shown for jets identified using the anti-k_t algorithm with $R = 0.6$. (a) shows inclusive events, while (b) shows gap events. The unfolded data are compared to the leading order particle level PYTHIA 6.423 prediction. In each case, the error bars indicate the statistical uncertainty on the measurement only. Statistically insignificant data points at large Δy are omitted.

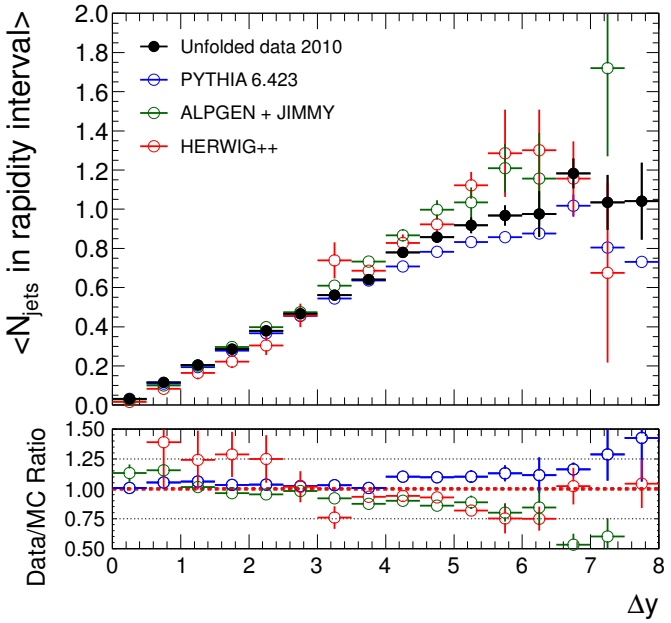


Figure 8.12.: Mean number of jets in the gap as a function of Δy. The veto scale is set to $Q_0 = 20$ GeV. The unfolded data are compared to the leading order particle level PYTHIA 6.423, HERWIG++ and ALPGEN predictions. In each case, the error bars indicate the statistical uncertainty on the measurement only.

8.7.2. Comparison to Higher Order Monte Carlo Generators

Preliminary predictions from next-to-leading order generators are shown here, using the MSTW2008 PDF set in all cases. Unfolded data are compared against HEJ interfaced with the ARIADNE parton shower [81], POWHEG showered with PYTHIA and POWHEG showered with HERWIG; the same particle level events are used for both of these POWHEG predictions. For the data, the errors shown reflect statistical uncertainties, with systematic uncertainties arising from the effects of jet energy scale, jet resolution and $\Delta\phi$ pointing resolution added in quadrature; uncertainties arising from the unfolding procedure have not yet been evaluated. For POWHEG, statistical uncertainties are combined with scale uncertainties, evaluated by varying the renormalisation and factorisation scales by factors of two in each direction; HEJ shows only the statistical uncertainties.

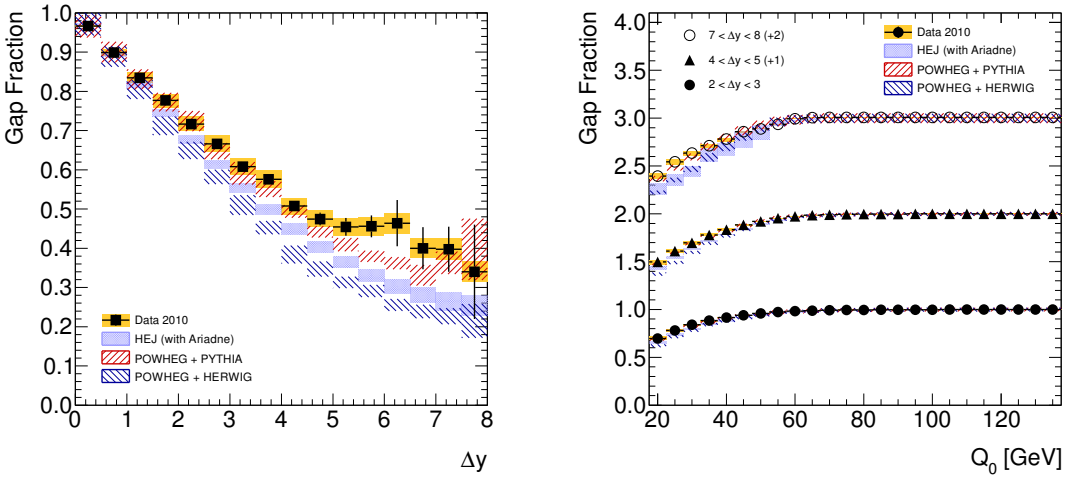
Figure 8.13 shows the gap fraction as a function of Δy and as a function of Q_0 for three different Δy slices. Large divergences between the Monte Carlo predictions can be seen throughout, these are accentuated at high Δy. The best agreement with the data comes from the POWHEG +PYTHIA prediction, although even this shows large levels of disagreement as Δy increases.

Figure 8.14 shows the $\langle\cos(\pi - \Delta\phi)\rangle$ and $\langle\cos(2\Delta\phi)\rangle$ distributions as a function of Δy for both inclusive and gap events. In general, there is good agreement between the data and the different Monte Carlo generators, with the major areas of disagreement coming at high Δy, where statistics are poorer. These disagreements are only present in the inclusive sample; once the jet veto is applied the different Monte Carlo predictions agree well.

Finally, the distribution of number of jets in the gap, as a function of Δy, is shown in Figure 8.15. As is the case for the comparison to leading order Monte Carlo predictions, divergences can be seen at high Δy, although POWHEG +PYTHIA agrees well with the data here.

8.8. Summary

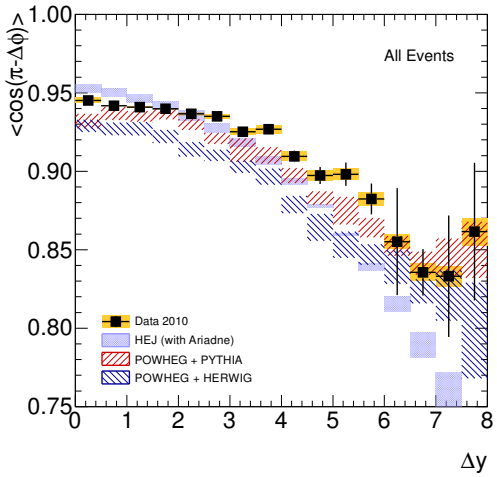
The different Monte Carlo event generators show certain consistencies in their predictions across each of these distributions. PYTHIA tends to give the best description



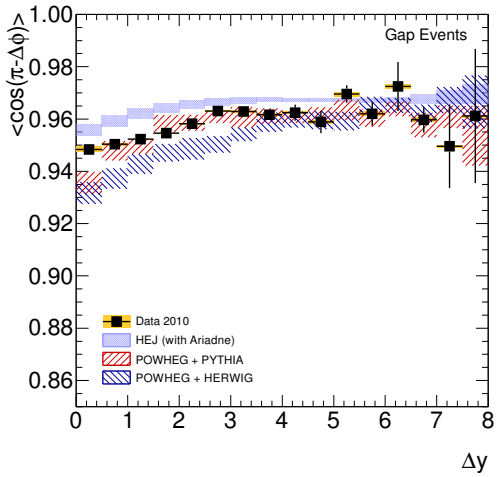
(a) Gap fraction as a function of Δy

(b) Gap fraction as a function of Q_0

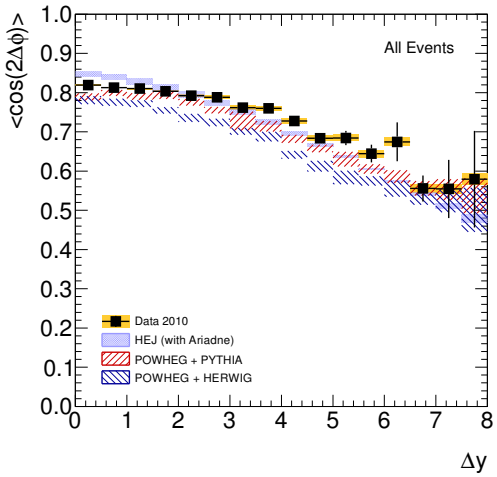
Figure 8.13.: Gap fraction distributions as a function of Q_0 and Δy for jets identified using the anti-k_t algorithm with $R = 0.6$. (a) shows the gap fraction as a function of Δy for $Q_0 = 20$ GeV, while (b) shows the gap fraction as a function of Q_0 for three different slices in Δy. The unfolded data are compared to the HEJ, POWHEG +PYTHIA and POWHEG +HERWIG. The error bars on data indicate the statistical uncertainty on the measurement with a series of systematic uncertainties summarised by the orange band. The error bands on the Monte Carlo predictions represent the statistical errors only, in the case of HEJ, with the scale uncertainties also combined in the case of POWHEG.



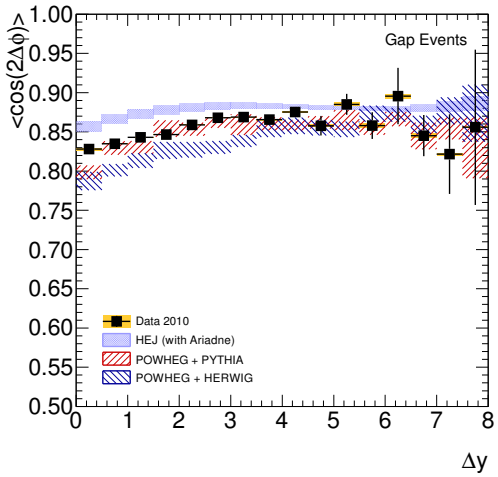
(a) $\langle \cos(\pi - \Delta\phi) \rangle$, inclusive events



(b) $\langle \cos(\pi - \Delta\phi) \rangle$, gap events



(c) $\langle \cos(2\Delta\phi) \rangle$, inclusive events



(d) $\langle \cos(2\Delta\phi) \rangle$, gap events

Figure 8.14.: Distributions of $\langle \cos(\pi - \Delta\phi) \rangle$ (top) and $\langle \cos(2\Delta\phi) \rangle$ (bottom) shown as a function of Δy for jets identified using the anti-k_t algorithm with $R = 0.6$. Inclusive events (left) and gap events (right) are shown separately. The unfolded data are compared to the HEJ, POWHEG +PYTHIA and POWHEG +HERWIG. The errors are as described in Figure 8.13.

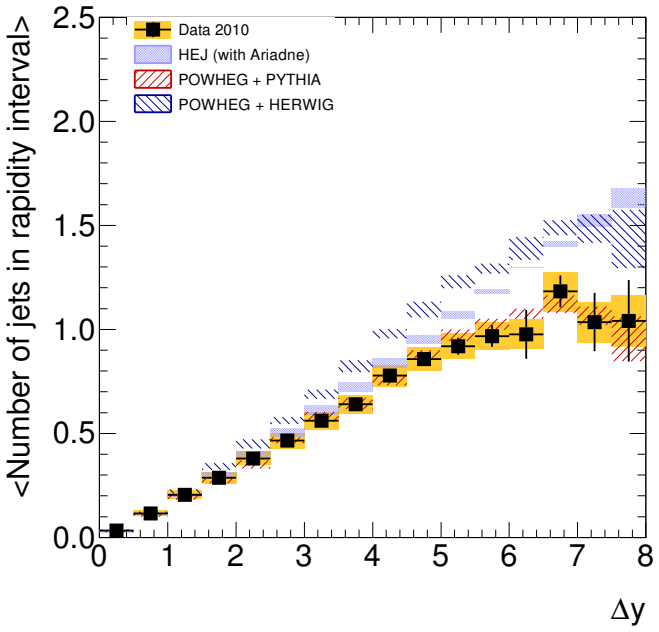


Figure 8.15.: Mean number of jets in the gap as a function of Δy. The veto scale is set to $Q_0 = 20$ GeV. The unfolded data are compared to the HEJ, POWHEG +PYTHIA and POWHEG +HERWIG. The errors are as described in Figure 8.13.

of the data while HERWIG++ and ALPGEN both overestimate activity in the gap and hence underestimate the gap fraction, particularly at high Δy and low Q_0. All generators give a reasonable description of the $\langle \cos(\pi - \Delta\phi) \rangle$ and $\langle \cos(2\Delta\phi) \rangle$ distributions, with the largest differences tending to occur at the largest Δy; the $\langle \cos(2\Delta\phi) \rangle$ distributions also shows significant shape differences between PYTHIA and unfolded data. The cross-section distributions show reasonable agreement between data and leading order Monte Carlo generator predictions.

For the cross-section distributions, reasonable agreement is seen between each of these leading order Monte Carlo generators and the unfolded data, although statistics are limited at large Δy and low $\Delta\phi$. This agreement comes despite the fact that no correction has been made for soft effects.

These distributions will soon be augmented by the replacement of leading-order Monte Carlo predictions by next-to-leading order predictions from POWHEG and HEJ and will then be considered for approval as an ATLAS publication.

Chapter 9.

Conclusions

"I conclude that there is no thing constantly observable in nature, which will not always bring some light with it, and lead us farther into the knowledge of her ways of working."

— John Locke

QCD studies are among the first measurements that can be made at hadron colliders; at the LHC jet production is the dominant high p_{T} process. Jet physics observables such as cross-sections are important for improving our understanding of the strong interaction, particularly through providing measurements of α_{S} at a range of scales which can then be used to constrain PDF fits.

Accurately calibrated jets are essential in order to make these measurements; dijet intercalibration is one of the in-situ methods used to validate the jet energy scale, which is primarily derived from Monte Carlo and test beam data. Results from intercalibration indicate that the relative response to both high p_{T} and central jets is well understood. Maximal disagreement occurs for low p_{T}, high rapidity jets. As well as providing a cross-check for the jet energy scale in the central region, intercalibration also provides one of the means by which it can be extended into the forward region.

Jet and dijet cross-section measurements conducted across the full accessible kinematic range provide information that has never been available before, particularly in the forward region, which previous hadron colliders have not been able to explore with such precision. In those regions of phase space in which experimental

and theoretical uncertainties are similar in size, these measurements provide some discriminating power between different theoretical models. Overall, after corrections for non-perturbative effects, NLO perturbative QCD predictions, agree with the measurements across seven orders of magnitude in cross-section. The greatest disagreements are seen at large values of jet transverse momentum and dijet invariant mass, where the theoretical predictions for the cross-sections tend to be larger than the measured values. These measurements probe, and may help to constrain, the previously unexplored area of parton distribution functions at large x and high Q^2, representing one of the most comprehensive tests of QCD ever performed.

In dijet events, the extent of hadronic activity in the rapidity interval between the jets can be studied with the use of a central jet veto. When considering the fraction of events which have no activity above the veto scale in this region, most experimental uncertainties cancel. Measurements of this ratio show the expected behaviour of a reduction of gap events as the dijet system becomes harder or more widely separated in rapidity. Good agreement is found with both leading-order Monte Carlo simulations and NLO predictions interfaced with a parton shower. No evidence was found that a BFKL-like description of parton evolution would provide an improved agreement between data and theoretical predictions. This data can be used to constrain the event generator modelling of QCD radiation between widely separated jets. Such a constraint is useful for the current Higgs-plus-two-jet searches and also for any future measurements sensitive to higher order QCD emissions.

Appendix A.

Combining Multiple Triggers

"It should be done with the same degree of alacrity and nonchalance that you would display in authorising a highly intelligent trained bear to remove your appendix."

— Daniel S. Greenberg

Consider the case in which there are two triggers of interest, one for each of the two leading jets. The case in which the leading jet passes the appropriate trigger is denoted T_{10} and the case where the second jet passes its trigger as T_{01}, while the case in which both do is labelled T_{11}. In the absence of any prescales or inefficiencies, we would expect all N of our events to be in category T_{11}, since our trigger boundaries are chosen to ensure that all jets fall on the trigger efficiency plateau. If the amount of data taken corresponds to a luminosity $\mathcal{L}_{true}$ then we have a cross-section of $\sigma_{true} = N/\mathcal{L}_{true}$.

A.1. Correcting for Inefficiency in the Absence of Prescale

Taking specific detector effects into account, which may lower the efficiencies of the leading and subleading triggers to e_L and e_S respectively, we can see that the distribution of events will be as summarised in Table A.1.

Trigger Category	Number of events
T_{00}	$N(1-e_L)(1-e_S)$
T_{10}	$Ne_L(1-e_S)$
T_{01}	$Ne_S(1-e_L)$
T_{11}	Ne_Le_S

Table A.1.: Number of events in each trigger category after allowing for trigger inefficiencies

Trigger Category	Number of events
T_{00}	$N(1-1/P_L)(1-1/P_S)$
T_{10}	$N/P_L(1-1/P_S)$
T_{01}	$N/P_S(1-1/P_L)$
T_{11}	$N/(P_LP_S)$

Table A.2.: Number of events in each trigger category after allowing for trigger prescales

Obviously events in the category T_{00} are not recorded, but we still see $N_{visible} = N(e_L+e_S-e_Le_S)$ events. Multiplying the luminosity by an overall efficiency correction of $e_{eff} = e_L + e_S - e_Le_S$ allows us to recover the correct cross-section:

$$\sigma_{corrected} = \frac{N_{visible}}{e_{eff}\mathcal{L}_{true}} = \frac{N(e_L + e_S - e_Le_S)}{(e_L + e_S - e_Le_S)\mathcal{L}_{true}} = \sigma_{true} \tag{A.1}$$

A.2. Correcting for Prescale in the Absence of Inefficiency

Allowing now for prescales of P_L and P_S for the leading and subleading triggers respectively, we can see that the distribution of events will be as summarised in Table A.2.

Analogously to the trigger efficiency case, we can apply a prescale correction of $1/P_{eff} = 1/P_L + 1/P_S - 1/(P_L P_S)$ to the luminosity to recover the correct cross-section.

$$\sigma_{corrected} = \frac{N_{visible}}{\mathcal{L}_{true}/P_{eff}} = \frac{N(1/P_L + 1/P_S - 1/(P_L P_S))}{\mathcal{L}_{true}(1/P_L + 1/P_S - 1/(P_L P_S))} = \sigma_{true} \tag{A.2}$$

In general, statistical precision will be improved by instead dividing events into separately weighted trigger categories, based on the trigger decision that would have been taken in the absence of prescale. This approach is discussed in more detail in Section 5.4.1.

A.3. Correcting for Inefficiency and Prescale Simultaneously

In the case in which efficiencies of e_L and e_S are present together with prescales of P_L and P_S, recovering the correct cross-section becomes more complicated.

If we could recover the prescale-only (PO) numbers, then recovering the correct cross-section would be a matter of following the prescription from Section A.2. In fact, using the simplified approach presented there, knowing only $N^{PO}_{visible} = N^{PO}_{11} + N^{PO}_{10} + N^{PO}_{01}$ would be sufficient.

Consider the relationships, shown in Equation (A.3), between the final observed (F) numbers of events observed and the numbers after prescale-only, where L denotes events passing the leading jet trigger and S denotes those passing the second trigger, while LS denotes those passing both triggers.

$$\begin{aligned} N^F_L &= N^{PO}_L e_L \\ N^F_S &= N^{PO}_S e_S \\ N^F_{LS} &= N^{PO}_{LS} e_L e_S \end{aligned} \tag{A.3}$$

Trigger Category	Required weight
T_{00}	0
T_{10}	$1/e_L$
T_{01}	$1/e_S$
T_{11}	$1/e_L/e_S$

Table A.3.: Required weights for each trigger category necessary to allow reconstruction of the total number of events that would have been observed without trigger inefficiency.

Expanding the first two of these equations out, we can see that

$$\begin{aligned} N_{10}^F + N_{11}^F &= (N_{10}^{PO} + N_{11}^{PO})e_L \\ N_{10}^F/e_L + N_{11}^F/e_L &= N_{10}^{PO} + N_{11}^{PO} \end{aligned} \tag{A.4}$$

and

$$\begin{aligned} N_{01}^F + N_{11}^F &= (N_{01}^{PO} + N_{11}^{PO})e_S \\ N_{01}^F/e_S + N_{11}^F/e_S &= N_{01}^{PO} + N_{11}^{PO} \end{aligned} \tag{A.5}$$

Combining Equation (A.4) and Equation (A.5) we have

$$N_{10}^{PO} + N_{11}^{PO} + N_{01}^{PO} + N_{11}^{PO} = N_{10}^F/e_L + N_{11}^F/e_L + N_{01}^F/e_S + N_{11}^F/e_S \tag{A.6}$$

$$N_{10}^{PO} + N_{11}^{PO} + N_{01}^{PO} + (N_{11}^F/(e_L e_S)) = N_{10}^F/e_L + N_{01}^F/e_S + N_{11}^F(1/e_L + 1/e_S) \tag{A.7}$$

$$N_{10}^{PO} + N_{01}^{PO} + N_{11}^{PO} = N_{10}^F/e_L + N_{01}^F/e_S + N_{11}^F(1/e_L + 1/e_S - 1/(e_L e_S)) \tag{A.8}$$

so the sum $N_{10}^{PO} + N_{01}^{PO} + N_{11}^{PO}$ can be reconstructed if events are weighted based on their final trigger category as shown in Table A.3.

Note that this will not allow reconstruction of the numbers of events passing each individual trigger category - it is only possible to obtain the total number of events which would have passed the logical OR of these triggers. For this reason, it is not possible to use the more statistically precise effective luminosity correction discussed in Section 5.4.1.

Colophon

This thesis was made in LaTeX 2_ε using the "hepthesis" class [82].

Bibliography

[1] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys. Lett. **B641** (2006) 57–61, arXiv:0512210 [hep-ph].

[2] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, Eur.Phys.J. **C72** (2012) 1896, arXiv:1111.6097 [hep-ph].

[3] CERN, *The accelerator complex*, http://public.web.cern.ch/public/en/research/AccelComplex-en.html, 2012.

[4] CERN, *ATLAS Detector images*, http://cdsweb.cern.ch/record/1095924, 2012.

[5] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, J. Instrum. **3** (2008) S08003.

[6] ATLAS Collaboration, G. Aad et al., *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. arXiv:1203.1302 [hep-ex].

[7] ATLAS Collaboration, G. Aad et al., *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.

[8] ATLAS Collaboration, G. Aad et al., *Updated Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2011-011, CERN, Geneva, 2011.

[9] ATLAS Collaboration, G. Aad et al., *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.

[10] ATLAS Collaboration, G. Aad et al., *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Tech. Rep. CERN-PH-EP-2011-191, CERN, Geneva, 2011. `arXiv:1112.6426 [hep-ex]`.

[11] B. Martin, *Nuclear and Particle Physics: An Introduction.* Wiley, 2009.

[12] K. Gottfried and V. Weisskopf, *Concepts of Particle Physics.* Clarendon Press, 1986.

[13] J. D. Bjorken, *Asymptotic Sum Rules at Infinite Momentum*, Phys. Rev. **179** (1969) 1547.

[14] V. N. Gribov and L. N. Lipatov, *Deep inelastic ep scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438–450.

[15] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. **B126** (1977) 298.

[16] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics*, Sov. Phys. JETP **46** (1977) 641–653.

[17] J. C. Collins and J. Qiu, *New derivation of the Altarelli-Parisi equations*, Phys. Rev. **D39** (1989) 1398–1408.

[18] S. Moch, J. Vermaseren, and A. Vogt, *The Three loop splitting functions in QCD: The Nonsinglet case*, Nucl. Phys. **B688** (2004) 101–134, `arXiv:0403192 [hep-ph]`.

[19] A. Vogt, S. Moch, and J. Vermaseren, *The Three-loop splitting functions in QCD: The Singlet case*, Nucl. Phys. **B691** (2004) 129–181, `arXiv:0404111 [hep-ph]`.

[20] G. Hanson et al., *Evidence for Jet Structure in Hadron Production by e^+e^- Annihilation*, Phys. Rev. Lett. **35** (1975) 1609–1612.

[21] G. P. Salam and G. Soyez, *A practical Seedless Infrared-Safe Cone jet*

algorithm, JHEP **05** (2007) 086, arXiv:0704.0292 [hep-ph].

[22] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, *New clustering algorithm for multijet cross sections in e^+e^- annihilation*, Phys. Lett. **B269** (1991) 432–438.

[23] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, *Longitudinally invariant K(t) clustering algorithms for hadron-hadron collisions*, Nucl. Phys. **B406** (1993) 187–224.

[24] Y. Dokshitzer, G. Leder, S. Moretti, and B. Webber, *Better jet clustering algorithms*, JHEP **08** (1997) 001, arXiv:9707323 [hep-ph].

[25] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, arXiv:9907280 [hep-ph].

[26] M. Cacciari, G. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **0804** (2008) 063, arXiv:0802.1189 [hep-ph].

[27] ATLAS Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[28] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303.

[29] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, arXiv:0603175 [hep-ph].

[30] ATLAS Collaboration, G. Aad et al., *Charged particle multiplicities in pp interactions at $\sqrt{s}$ = 900 GeV and 7 TeV in a diffractive limited phase-space measured with the ATLAS detector at the LHC and new PYTHIA6 tune*, Tech. Rep. ATLAS-CONF-2010-031, CERN, Geneva, 2010.

[31] A. Sherstnev and R. Thorne, *Different PDF approximations useful for LO Monte Carlo generators*, arXiv:0807.2132 [hep-ph].

[32] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[33] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. **D82** (2010) 074018, arXiv:1005.3457 [hep-ph].

[34] G. Corcella et al., *HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, arXiv:0011363 [hep-ph].

[35] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. **C72** (1996) 637–646, arXiv:9601371 [hep-ph].

[36] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008) 639–707, arXiv:0803.0883 [hep-ph].

[37] S. Gieseke, D. Grellscheid, K. Hamilton, A. Papaefstathiou, S. Platzer, et al., *Herwig++ 2.5 Release Note*, arXiv:1102.1672 [hep-ph].

[38] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, arXiv:0206293 [hep-ph].

[39] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv:0201195 [hep-ph].

[40] ATLAS Collaboration, G. Aad et al., *First tuning of HERWIG/JIMMY to ATLAS data*, Tech. Rep. ATL-PHYS-PUB-2010-014, CERN, Geneva, 2010.

[41] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv:0409146 [hep-ph].

[42] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv:0709.2092 [hep-ph].

[43] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **1006** (2010) 043, arXiv:1002.2581 [hep-ph].

[44] J. R. Andersen and J. M. Smillie, *Factorization of the t-channel pole in quark-gluon scattering*, Phys. Rev. **D81** (2010) 114021.

[45] ATLAS Collaboration, G. Aad et al., *Estimating Track Momentum Resolution in Minimum Bias Events using Simulation and K^0_S in $\sqrt{s}$ = 900 GeV Collision Data*, Tech. Rep. ATLAS-CONF-2010-009, CERN, Geneva, 2010.

[46] ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS jet trigger with pp collisions at* $\sqrt{s} = 900 GeV$, Tech. Rep. ATLAS-CONF-2010-028, CERN, Geneva, 2010.

[47] ATLAS Collaboration, W. Lampl et al., *Calorimeter Clustering algorithms: Description and Performance*, Tech. Rep. ATLAS-LARG-PUB-2008-002, CERN, Geneva, 2008.

[48] ATLAS Collaboration, C. Cojocaru et al., *Hadronic calibration of the ATLAS liquid argon end-cap calorimeter in the pseudorapidity region* $1.6 < |\eta| < 1.8$ *in beam tests*, Nucl. Instrum. Meth. **A531** (2004) 481–514, `arXiv:0407009v1 [physics.ins-det]`.

[49] H1 Collaboration, B. Andrieu et al., *Results from pion calibration runs for the H1 liquid Argon calorimeter and comparisons with simulations*, Nucl. Instrum. Meth. **A336** (1993) 499.

[50] ATLAS Collaboration, G. Aad et al., *Luminosity Determination in pp Collisions at* $\sqrt{s} = 7$ *TeV Using the ATLAS Detector at the LHC*, Eur. Phys. J. **C71** (2011) 1630, `arXiv:1101.2185 [hep-ex]`.

[51] ATLAS Collaboration, G. Aad et al., *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center of Mass Energy* $\sqrt{s} = 7$ *TeV*, Tech. Rep. ATLAS-CONF-2010-038, CERN, Geneva, 2010.

[52] ATLAS Collaboration, G. Aad et al., *Measurement of jet production in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2010-050, CERN, Geneva, 2010.

[53] G. D'Agostini, *Improved iterative Bayesian unfolding*, `arXiv:1010.0632 [physics.data-an]`.

[54] ATLAS Collaboration, G. Aad et al., *In-situ pseudorapidity intercalibration for evaluation of jet energy scale uncertainty using dijet events in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, Tech. Rep. ATLAS-CONF-2011-014, CERN, Geneva, 2011.

[55] C. Buttar, J. D'Hondt, M. Kramer, G. Salam, M. Wobisch, et al., *Standard Model Handles and Candles Working Group: Tools and Jets Summary Report*,

arXiv:0803.0678 [hep-ph].

[56] Z. Nagy, *Next-to-leading order calculation of three jet observables in hadron hadron collision*, Phys. Rev. **D68** (2003) 094002, arXiv:0307268 [hep-ph].

[57] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C.-P. Yuan, *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[58] ATLAS Collaboration, G. Aad et al., *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) 1–59.

[59] ATLAS Collaboration, G. Aad et al., *Measurement of inclusive jet and dijet production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector.*, Phys. Rev. **D86** (2012) no. CERN-PH-EP-2011-192, 014022, arXiv:1112.6297 [hep-ex].

[60] ZEUS Collaboration, M. Derrick et al., *Rapidity gaps between jets in photoproduction at HERA*, Phys.Lett. **B369** (1996) 55–68, arXiv:hep-ex/9510012 [hep-ex].

[61] H1 Collaboration, C. Adloff et al., *Energy flow and rapidity gaps between jets in photoproduction at HERA*, Eur. Phys. J. **C24** (2002) 517–527, arXiv:0203011 [hep-ex].

[62] ZEUS Collaboration, S. Chekanov et al., *Photoproduction of events with rapidity gaps between jets at HERA*, Eur. Phys. J. **C50(2)** (2006) 283–297, arXiv:0612008 [hep-ex].

[63] D0 Collaboration, S. Abachi et al., *Rapidity gaps between jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **72** (1994) 2332–2336.

[64] CDF Collaboration, F. Abe et al., *Observation of rapidity gaps in $\bar{p}p$ collisions at 1.8 TeV*, Phys. Rev. Lett. **74** (1995) 855–859.

[65] CDF Collaboration, F. Abe et al., *Events With A Rapidity Gap Between Jets In $\bar{p}p$ Collisions at $\sqrt{s} = 630$ GeV*, Phys. Rev. Lett. **81** (1998) 5278–5283.

[66] D0 Collaboration, B. Abbott et al., *Probing hard color-singlet exchange in $p\bar{p}$ collisions at $\sqrt{s} = 630$ GeV and 1800 GeV*, Phys. Lett. **B440** (1998) 189–202,

arXiv:9809016 [hep-ex].

[67] CDF Collaboration, F. Abe et al., *Dijet Production By Color-Singlet Exchange At The Fermilab Tevatron*, Phys. Rev. Lett. **80** (1998) 1156–1161.

[68] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, *The Pomeranchuk Singularity in Nonabelian Gauge Theories*, Sov. Phys. JETP **45** (1977) 199–204.

[69] I. I. Balitsky and L. N. Lipatov, *The Pomeranchuk Singularity in Quantum Chromodynamics*, Sov. J. Nucl. Phys. **28** (1978) 822–829.

[70] J. R. Andersen and J. M. Smillie, *Constructing All-Order Corrections to Multi-Jet Rates*, JHEP **01** (2010) 039, arXiv:0908.2786 [hep-ph].

[71] J. R. Andersen and J. M. Smillie, *High Energy Description of Processes with Multiple Hard Jets*, Nucl.Phys.Proc.Suppl. **205–206** (2010) 205–210, arXiv:1007.4449 [hep-ph].

[72] J. R. Forshaw, A. Kyrieleis, and M. H. Seymour, *Gaps between jets in the high energy limit*, JHEP **06** (2005) 034, arXiv:0502086 [hep-ph].

[73] J. R. Forshaw, A. Kyrieleis, and M. H. Seymour, *Super-leading logarithms in non-global observables in QCD*, JHEP **08** (2006) 059, arXiv:0604094 [hep-ph].

[74] J. R. Forshaw, J. Keates, and S. Marzani, *Jet vetoing at the LHC*, JHEP **07** (2009) 023, arXiv:0905.1350 [hep-ph].

[75] J. M. Campbell, R. K. Ellis, and G. Zanderighi, *Next-to-leading order Higgs + 2 jet production via gluon fusion*, JHEP **10** (2006) 028, arXiv:0608194 [hep-ph].

[76] SM and NLO Multileg Working Group Collaboration, J. R. Andersen et al., *The SM and NLO multileg working group: Summary report*, arXiv:1003.1241 [hep-ph].

[77] J. R. Andersen, L. Lonnblad, and J. M. Smillie, *A Parton Shower for High Energy Jets*, JHEP **1107** (2011) 110, arXiv:1104.1316 [hep-ph].

[78] ATLAS Collaboration, G. Aad et al., *Measurement of dijet production with a veto on additional central jet activity in pp collisions at $\sqrt{s} = 7$ TeV using the*

ATLAS detector, JHEP **09** (2011) 053.

[79] D. Colferai, F. Schwennsen, L. Szymanowski, and S. Wallon, *The first complete NLL BFKL study of Mueller Navelet jets at LHC*, `arXiv:1110.1250 [hep-ph]`.

[80] T. Adye, *Unfolding algorithms and tests using RooUnfold*, `arXiv:1105.1160 [physics.data-an]`.

[81] L. Lonnblad, *Ariadne version 4.12 program and manual*, Comput. Phys. Commun. **71** (1992) 153.

[82] A. Buckley, *A class for typesetting academic theses*, 2010.