MEASUREMENT OF MULTI-JET PRODUCTION CROSS SECTION AT A CENTER-OF-MASS ENERGY OF 7 TeV AT THE LARGE HADRON COLLIDER WITH THE ATLAS DETECTOR

Thesis by

David López Mateos

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2011
(Defended May 13, 2011)

Acknowledgments

The road to obtaining my doctoral degree has been a long one, and leaves me indebted to many people who, through their help and advice, have made this thesis possible and my time as a graduate student a unique and enjoyable experience. First, I should thank my advisor, Emlyn W. Hughes, for his trust and encouragement through all these years, and for his availability and advice whenever doubts about how to move forward assailed me. I should also thank all the members of the ATLAS Collaboration, since without their dedication and sacrifices, the operation of the detector and the understanding of the data used in this thesis would not have been possible.

Certain members of the collaboration deserve special recognition from my part, starting with Ariel Schwartzman, who has taught me most of what I know about jets, and whose passion about experimental physics has been inspirational. Ahmimed Ouraou, Jean-François Laporte, Samira Hassani and Andreas Salzburger also deserve my gratitude, for their patience when I was a young student and their guidance through my first steps in the collaboration. I am also grateful to Paolo Morettini, Nicoletta Garelli, Iskander Ibragimov, Götz Gaycken and other members of the pixel DAQ development group for their serenity at crunch time and making me a part of their great achievements.

The analysis presented in this thesis would not have been possible without the support of several people. I should thank Kevin Einsweiler for his support and very insightful remarks all throughout the analysis process, and his personal support when I was trying to decide my future. Jon Butterworth and Koji Terashi also deserve my gratitude for their support of the analysis and helping others understand the importance of the studies. The active members of the analysis group also deserve credit. In particular, I would like to thank Matthew Tamsett and Elin Bergeaas Kuutman for their help with the editorial work for the public multi-jet results, Marc-Andre Dufour and Anne Holtsch for their inestimable expertise in performing the next-to-leading-order calculations and Zinonas Zinonos and Markus Schneider for their numerous cross-checks and valuable studies.

Many other people have helped me during my years at CERN, and naming all of them is not possible. However, Tancredi Carli, Claude Guyot and Fabiola Gianotti deserve a special recognition for their support of my applications to post-graduate research fellowships. I would also like to express my gratitude towards Peter Loch, Claire Gwenlan, James Degenhardt, Michael Duehrssen, Ahmimed Ourau and Maxwell Scherzer for reading sections of this thesis and providing helpful comments.

My doctoral experience has not only been shaped by the work I have done, but also by the people I have shared time with, who have allowed me to share my excitement in times of joy and have provided comfort at times when moving forward seemed hardest. In this context, I should thank my MIT college friends for keeping in touch and sharing their continuing passions with me at the smallest opportunity. In particular, I should thank Jon Grall, for driving often to Pasadena and for many dinners and comedy nights that always appeased my spirit. I should also thank Juan Rodriguez, for always offering a helping hand and keeping close in the distance. I also have to thank my fellow Caltech students in the physics doctoral program and the members of the Caltech rugby team, for having made my year at Caltech one with plenty of fond memories and colorful stories.

Most of my doctoral residence has been spent at CERN in Geneva, and I have been lucky to be able to consider as my friends many of the people I have worked with. David Miller and Zachary Marshall deserve my gratitude for sharing their talent, being reliable colleagues and trying to get me out of the office when the days were too long. I would also like to thank Valeria Perez Reale and Kerstin Perez, for giving me objectivity when I was needing it and keeping a cheerful atmosphere in the office at all times.

Finally, I am forever indebted to my family and, representing them, my father, Ricardo López Fernández, not only for all the immaterial and material that they have given me in the past, but also for their continued efforts to support me and my pursuits. Finally, my time in Geneva could not be understood without my best friend and partner, Géraldine Conti. For giving me so much and asking for so little, for sharing my hopes and dreams and for so many other things that would take a whole book to describe, I am forever grateful to her.

Abstract

Inclusive multi-jet production is studied using proton-proton collisions at a center-of-mass energy of 7 TeV. The data sample corresponds to an integrated luminosity of 2.43 pb^{-1}, and uses a subset of the data collected by the ATLAS detector in 2010. Measurements of multi-jet production cross sections with up to six jets are presented and compared to both leading-order predictions enhanced by a parton shower and next-to-leading-order calculations. Generally, agreement is found between data and leading-order and next-to-leading-order predictions. Regions of phase space or observables in which pure parton-shower simulations differ from the measurements are well described by leading-order calculations that combine matrix-element calculations with several outgoing partons in the final state to parton-shower simulations. Measurements compared to next-to-leading-order calculations are shown to be sensitive to the value of the strong coupling constant.

Contents

List of Figures

List of Tables

Preface

In the spring of 2006, nine months after arriving at Caltech, I joined the ATLAS Collaboration. Since June of that year, when I moved to CERN, I worked on many aspects of the ATLAS detector and the preparations for the commissioning run, then scheduled to start in late 2007.

As part of the pixel detector collaboration, that summer I helped on the setup of a cosmic-ray trigger for the “system test” of the detector. The system test consisted on the first large-scale[1] operation of the pixel detector, and was completed at the end of that year. I helped in the operation of the detector during this period, while starting work on the scalability of the data-acquisition (DAQ) system, particularly concerned about the needs for the detector calibration. That work, which involved both design and implementation of some of the controller and interfacing software for the operation of the pixel detector, continued in 2007 and was first used during that year during the “connectivity test” in the clean room where the pixel detector was assembled. The connectivity test was designed to verify the connections of the pixel read-out chain, and involved work both at the software implementation level, as well as at the operations level.

At the end of 2007, due to failures in the magnet triplets of the LHC, it became clear that the run would have to be postponed until 2008. This gave additional operational time for the pixel detector crew, which I was part of. In 2008 the connectivity test of the installed detector [1] took place, using all the components developed by the small Pixel DAQ group, including those that guaranteed the scalability of the system. An intense period of operations also took place after that connectivity test, as the detector was prepared for first data taking at the end of August 2008. The September 19 incident[2] in the LHC further

---

[1] One full end-cap of the detector, as defined in Chapter 4, was used.

[2] The incident in the LHC machine on September 19, 2008, determined the energy of the LHC beams in 2009–2010, and a detailed account can be found in the references in Chapter 3.

postponed the data collection until late 2009, and my involvement in the pixel detector operations and DAQ was then limited to the early 2009 calibration campaign and the quick-turn-around changes needed in certain calibration algorithms [2].

While involved in the pixel detector commissioning, I also worked in the event reconstruction. From 2006 until 2008 I worked in the tracking algorithms propagating muons through the calorimeters and the calculation of their energy loss [3, 4]. This work was used for many of the cosmic muon analyses and the validation framework put in place for these studies was also used in the validation of new algorithms for the propagation of muons in dense matter [5, 6]. Due to the expertise acquired in this work, I was co-editor of Chapter 4.2 [7] of Ref. [8], the last comprehensive Monte Carlo–based performance studies released by ATLAS before the arrival of collision data.

Around the end of 2007, given the lack of collision data, I also started working on jet reconstruction and calibration. My first contribution was to calculate a correction for energy carried by neutrinos in semileptonic decays of *b*-quark jets [9]. This correction has been used in recent analyses [10]. In late 2008 I proposed a simple jet calibration for early data [11, 12] that would likely have a smaller systematic uncertainty than the calibrations used so far, while giving an acceptable performance in early analyses. This calibration has been used for all physics analyses using jets, since it gives the best trade-off between jet energy resolution and jet energy scale systematic so far. In early 2009 I developed an extension to this calibration that could easily be validated with data and that had a performance that was competitive with the best jet calibrations available in ATLAS [13,14]. The techniques employed in this work were also shown to improve the performance of other calibration schemes in ATLAS [15]. A systematic uncertainty has been recently associated to this calibration scheme using the 2010 data in a public note that I co-edited [16].

With the arrival of 900 GeV and 7 TeV collision data in 2009 and 2010, the work shifted from the development of calibration strategies and understanding of different jet algorithms [17] to the basic understanding of our data. In this context I co-edited the notes studying the details of jet reconstruction and calibrations using 900 GeV [18] and 7 TeV [19] data, while contributing to the effort to establish the ATLAS jet energy scale [20]. This also resulted in work that helped in the understanding of the jet energy scale in other analyses, such as those involving the measurement of the top quark [21, 22].

The studies that I had originally planned to present in this thesis involved searches for supersymmetric particles with jets in the final state [23]. However, studies performed after the September 19 incident [24] made it obvious that the 2010 run would not deliver sufficient data to make strong statements about supersymmetry, and that the focus during this early run should be on the understanding of the Standard Model and our detector. The similarities between the Standard Model QCD multi-jet final states, and those encountered in supersymmetry, made the topic of this thesis a natural choice. In performing this analysis, I co-edited the notes containing two preliminary results [25, 26], a supporting note on the characterization of the flavor composition of the sample used [27] and a paper [28].

Only these last studies are presented in this thesis. All the other work performed during my graduate years has been very important not only for the operation of the detector and many analyses making use of it, but also for my education and my growth as an experimental particle physicist. However, it has already been documented in the references above and, therefore and for coherence, only the measurement of the Standard Model multi-jet production cross section is described in what follows. Some of the topics above are, of course, relevant to this measurement, and they are, thus, described when appropriate. Nevertheless, the discussion will necessarily be briefer than in the documents listed above. The knowledge acquired through all these years will be a great asset as I pursue the analysis of the 2011–2012 dataset, where new physics will hopefully appear.

Chapter 1

Introduction

The Standard Model of Particle Physics (SM) is a relativistic quantum field theory that provides an almost complete picture of the interactions of matter through three of the four known forces: strong, weak and electromagnetic. The symmetry of these interactions is understood through a set of three gauge groups $SU(3) \times SU(2) \times U(1)$ whose generators represent the force carriers, and whose eigenvectors represent particles with different couplings to the force carriers (quantum numbers or eigenvalues).

Leptons (electrons, muons, taus and neutrinos) interact through the $SU(2) \times U(1)$ gauge group, while quarks (up, down, strange, charm, bottom and top) couple to all force carriers. The quantum numbers of the three gauge groups building the SM are referred to as color, weak isospin and weak hypercharge. The $SU(2) \times U(1)$ symmetry is broken in such a way that the weak and electromagnetic forces arise as different mixings or combinations of the $SU(2) \times U(1)$ generators. The breaking of the symmetry gives rise to the observed neutral force carriers of the electromagnetic (photon) and the weak (Z boson) interactions. It also gives rise to the mass of the weak force carriers (W and Z bosons).

The breaking of the symmetry is believed to be produced through the Higgs mechanism, in which a scalar field changes the characteristics (the expectation value) of the vacuum. With the modified vacuum features, the observables of the theory need to be re-evaluated and, generally speaking, the fields to which the scalar field couples will acquire a mass. In the case of the $SU(2) \times U(1)$ symmetry, the quantum numbers of the scalar field can be chosen so that one of the gauge bosons of the theory is massless (the photon) while the others (the W and the Z) are massive. The scalar field that is thought to break the $SU(2) \times U(1)$ symmetry is called the Higgs boson, and it has not yet been observed.

The SM remains a landmark success of theoretical physics after numerous experimental tests and precision measurements searching for its limitations. However, the SM fails to fundamentally explain the values of many of its parameters and it is widely regarded as an “effective theory”: an approximate theory, valid at the currently understood energy scales, and arising from a more fundamental theory. In addition, the existence of the Higgs boson has only been theorized, and other realizations of the Higgs mechanism or symmetry breaking are possible.

One of the reasons why the SM is regarded as an effective theory is that, in its simplest version, two of its parameters (related to the mass of the Higgs boson) are required to be tuned to within $\sim 10^{-16}$ of each other [29,30]. This unnatural “fine tuning” is caused by the difference between the energies at which electroweak interactions happen and the Planck scale (the *terra incognita* where some new physics must help reconcile our current theories of the very small and the very big). This “hierarchy problem” is particularly interesting for collider experiments working at the few TeV range, since many of the optimal solutions to it lie lurking in this energy range.

The Large Hadron Collider (LHC), colliding protons traveling in opposite directions with very high energies, has been designed to discover (or rule out the existence of) the SM Higgs boson. Such a definite answer will come because the only free parameter of the Higgs boson in its simplest realization is its mass. This mass is bounded above, at around 1 TeV, by the unitarity of WW scattering (if the mass were higher, the probability of two Ws scattering into any other known particles would be higher than 1, breaking the self-consistency of the theory). The mass is bounded below by the Large Electron-Positron (LEP) collider, at 114 GeV. While electroweak measurements favor a low-mass SM Higgs [31], the study of the high-mass regions is important to exclude an SM Higgs in case nature has other mechanisms to accommodate for the values of the electroweak measurements. This full mass range is kinematically accessible to the LHC, and the beam intensities combined with the Higgs boson production cross section (or probability) will allow covering it in a few years, even before the design specifications of the LHC are reached.

This statement of principles hides behind it an extremely rich physics program whether the SM Higgs boson exists or not. Even if it exists, the new energy regime at the LHC requires new levels of understanding of some of the details of the SM; and, in order to have a successful discovery campaign, collaboration between theorists and experimentalists is needed to push the limits of our current theoretical calculations. This understanding is also key in trying to shed light on the hierarchy problem, which under certain models could happen much earlier than the discovery of the Higgs. If the SM Higgs boson does not exist, the search for the unknown has the potential of providing answers to questions of interest across different fields of physics such as those related to the nature of dark matter, the mechanisms for the unification of the forces or the existence of additional spatial dimensions. However, this search may also require an even more detailed knowledge of some of the features of the SM.

Among all the experimental handles available at collider experiments, jets stand as those whose impact on LHC analyses may have the widest reach. In a most simplified definition jets are collimated sets of particles and, as typically understood, with a common origin. The origin, except in certain exotic theories or extreme kinematic configurations, is either a quark or a gluon (the force carriers of the strong interaction). More precise and correct definitions of a jet will be provided in later chapters.

As shown in Figure 1.1, the production cross section of events with jets increases much faster as a function of the center-of-mass energy of the collider ($\sqrt{s}$) than the total cross section, and than that of events containing only leptons (W and Z production). In addition, the probability of creating a jet associated with other particles (W, Z, top) is also larger at the LHC than ever before. Jets are also created as part of discovery production and decay channels of some of the most promising new interactions that will be probed at the LHC. For these and other reasons, the understanding of jets and their production is as important now as ever.

The largest sources of jet production are processes mediated by the strong force and described by the theory of the strong interactions: quantum chromodynamics (QCD). A detailed knowledge of such processes is thus needed in order to successfully study some of the most interesting physics at the LHC.

While QCD is fundamentally well understood, the nature of the theory makes it hard to make precise predictions of many important observables. Models and approximations exist to increase the predictive power of the theory but, even then, some of their limitations are not always well understood and require verification. It becomes then the job of the

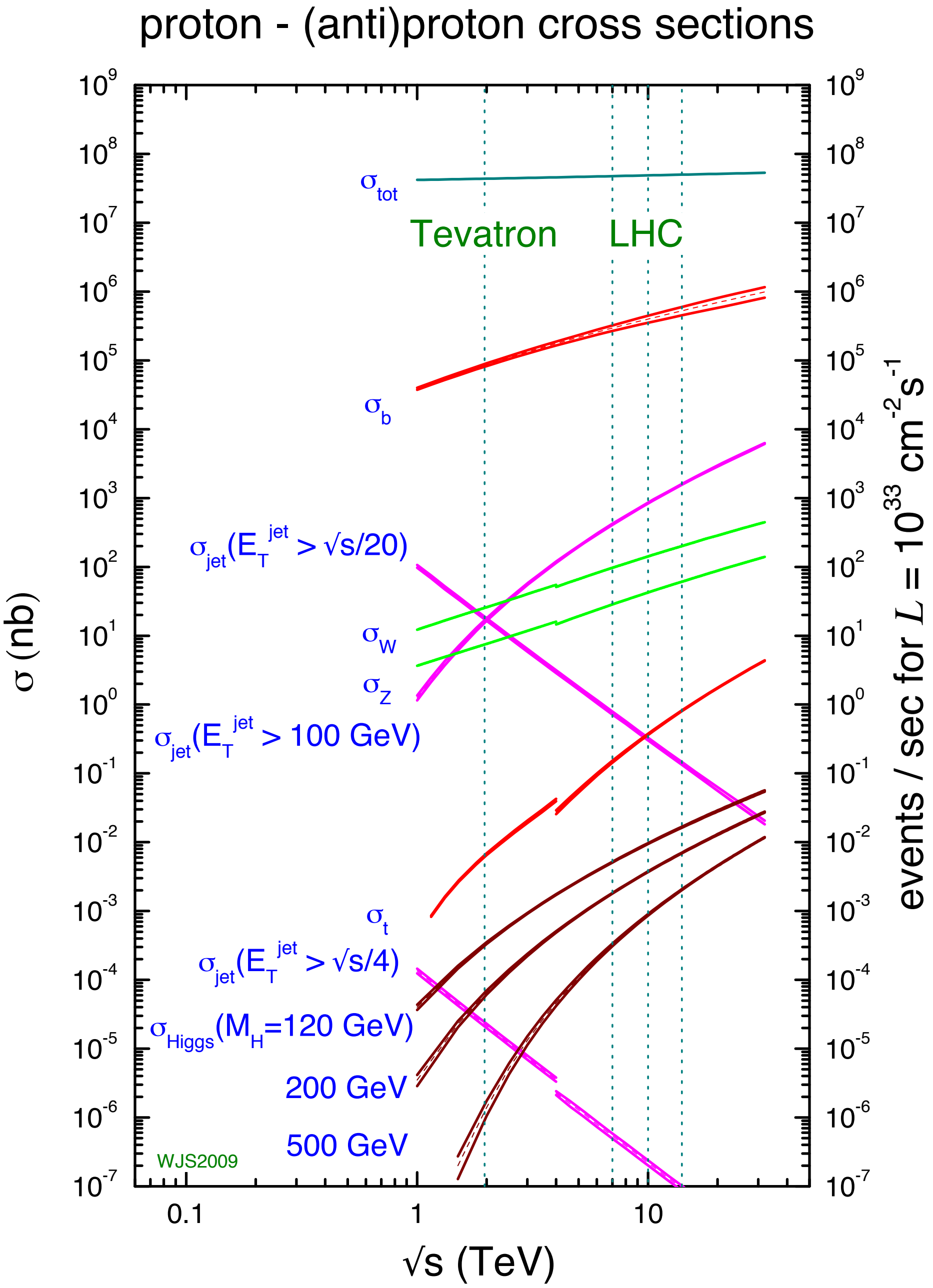

Figure 1.1: *Some of the relevant cross sections as a function of center-of-mass energy in proton-antiproton (left) and proton-proton (right) colliders [32]. The dashed lines correspond to the design center-of-mass energy at the Tevatron and the LHC. The cross section is measured in nb and events per second at the LHC design instantaneous luminosity (beam intensity).*

experimentalist to understand and characterize the production of QCD jets, to increase the knowledge of QCD and aid the many searches of new physics in which QCD is a background.

This thesis studies the production of events with multiple jets at the LHC with the ATLAS detector for the first time using early data collected between April 10 and August 30, 2010. Chapter 2 discusses the theory that is relevant to this measurement, providing an overview of both fundamental QCD, as well as the modeling and approximations used to make it useful for experimental comparisons. Chapter 3 briefly describes the LHC, focusing on those design features that impact the measurement of jets. Chapter 4 describes the ATLAS detector, focusing on the sub-detectors used for this analysis, as well as the data-acquisition chain that records the detector information. Chapter 5 describes how events are reconstructed to have access to the physics objects of interest. This includes a discussion of the jet reconstruction algorithms used both for the data and the theoretical calculations. Chapter 6 discusses how events were selected, including studies of the triggers used and a discussion of how the luminosity measurement was performed. Chapter 7 describes how measured distributions were corrected to truth-level distributions that can be calculated using Monte Carlo simulations of the multi-jet processes. Chapter 8 discusses the systematic uncertainties in the measurements, and their relative importance. Chapter 9 describes the details of the theoretical calculations used in this thesis, together with the systematic uncertainties associated to them. Finally, Chapter 10 shows the results and discusses their implications, before the measurements are summarized and future prospects are surveyed.

Chapter 2

Quantum Chromodynamics and Jet Physics

The production of jets at the LHC is dominated by processes mediated by the strong force. The theory of the strong force, Quantum Chromodynamics (QCD), arose from a series of symmetry observations of the hadron spectrum [33–38], and has been verified successfully in a wealth of fixed-target and collider experiments since the 1960s [12,39–43]. Despite its successes, QCD remains a theory were calculations are hard to make, and for which many approximations exist in order to increase its predictive power [44–47]. In this chapter, the fundamentals of the theory are first described, emphasizing the features of the theory that will be most relevant for the analysis presented in this thesis. Then, the particularities of QCD at hadron colliders, such as the approximations made and the models needed, are described, with a primary focus on the high-momentum-transfer interactions relevant to measurements of jet production.

2.1 Quantum Chromodynamics

One of the fundamental concepts in QCD is that hadrons are made of quarks. This idea originally arose from the observation of an $SU(3)$ flavor symmetry ($SU(3)_f$, not to be mistaken with the $SU(3)$ color symmetry) in the properties of low-mass mesons and baryons [33]. Quarks are particles with fractional charge and spin $\frac{1}{2}$ and some of their fundamental properties and quantum numbers are summarized in Table 2.1.

The need to resolve a conflict between the requirements of Fermi-Dirac statistics for spin-$\frac{1}{2}$ particles and the observations of hadrons of spin $\frac{3}{2}$ led to the introduction of a new

Quark	Mass (GeV)	Charge (e)	Spin
down (d)	5.05×10^{-3}	$-\frac{1}{3}$	$\frac{1}{2}$
up (u)	2.49×10^{-3}	$\frac{2}{3}$	$\frac{1}{2}$
strange (s)	1.01×10^{-1}	$-\frac{1}{3}$	$\frac{1}{2}$
charm (c)	1.27	$\frac{2}{3}$	$\frac{1}{2}$
bottom (b)	4.19	$-\frac{1}{3}$	$\frac{1}{2}$
top (t)	172	$\frac{2}{3}$	$\frac{1}{2}$

Table 2.1: *Properties of quarks [31].*

degree of freedom that could take three different values: color [48,49]. No difference between quarks of different colors was needed to explain experimental observations, and so a new symmetry was added to the theory: $SU(3)$ color. The possibility of the existence of quarks motivated direct searches using deep inelastic scattering of electrons off a proton [39,50–52]. These experiments found not only that protons were made of point-like constituents, then referred to as "partons",[1] but also that these constituents appeared to be free inside the proton (they exhibit "asymptotic freedom" [53, 54]). This came as a startling revelation, given that quarks had never been observed free (they were confined to hadrons, or exhibited "confinement").

All these elements were combined together successfully in 1973 [53,54] into a Yang-Mills quantum field theory: QCD. The classical Lagrangian density of QCD is not very different in form from that of Quantum Electrodynamics (QED):

$$\mathcal{L}_{QCD}^{\text{classical}} = -\frac{1}{4} F^{A}_{\alpha\beta} F^{\alpha\beta}_{A} + \sum_{\text{flavors}} \bar{q}_a (i\not{D} - m)^{ab} q_b. \tag{2.1.1}$$

The first term represents the strong force fields, while the second is the kinetic and mass terms of quarks, where the latin indices are summed over the four components of the Dirac spinors, and $\not{D}$ represents the covariant derivative of the four-momentum dotted with the Dirac matrices where the covariance is with the gluon field. In the quantization of this Lagrangian through Feynman path integrals, gauge-fixing and ghost terms (that avoid double-counting of physical states in the procedure) need to be added, but for the purposes of this chapter, these can be ignored.

---

[1] Today, the word parton is used to collectively refer to quarks and gluons.

Despite its apparent similarities with the QED Lagrangian, the QCD Lagrangian has a key feature that makes the physics it describes radically different. This feature has to do with the $SU(3)$ color symmetry (and the non-abelian nature of the group) and is hidden in the definition of the field strength tensor:

$$F^A_{\alpha\beta} = \partial_\alpha \mathcal{A}^A_\beta - \partial_\beta \mathcal{A}^A_\alpha - g f^{ABC} \mathcal{A}^B_\alpha \mathcal{A}^C_\beta. \tag{2.1.2}$$

Here $\mathcal{A}^A_\beta$ is the β component of the gluon field with color A (gluons have eight color degrees of freedom), g is the gauge coupling parameter (related to the "strong coupling constant" introduced later) and f^{ABC} are the totally antisymmetric structure constants of the $SU(3)$ group. The last term is not present in QED (or other abelian field theories) and represents the gluon self-interaction. This interaction gives rise to asymptotic freedom, which characterizes QCD.

Asymptotic freedom arises from the quantization of the theory. In the perturbative expansion of Feynman, loops of particles existing only virtually appear in the calculation of scattering amplitudes. Heisenberg's principle allows the energy of these particles to be arbitrarily large. We know, however, that those scales cannot be arbitrarily large, since other unknown interactions must come into play at least at the Planck scale. Such interactions should allow us reconciling our ideas of the strong and electroweak forces with those of gravity into a unified "theory of everything" in which these virtual corrections can be calculated unambiguously from first principles.

Despite our lack of knowledge about the theory of everything, quantum field theory possesses a tool that allows us to integrate out the physics at those high scales: renormalization. The price paid in this procedure is the introduction of an arbitrary scale ("renormalization" scale), and the need to introduce boundary conditions: an experimental measurement that will characterize the theory at a certain scale, allowing the perturbative calculation to provide the evolution from this scale. Physics cannot depend on the choice of this arbitrary scale, and this fact can be imposed on whichever calculation is being performed (e.g., self-coupling corrections). The imposition of this condition reveals an energy dependence of the coupling constant, which for QCD is, at leading order in perturbation theory:

$$\alpha_s(\mu) = \frac{\alpha_s(\mu_0)}{1 + (11 - \frac{2}{3}n_f)\frac{\alpha_s(\mu_0)}{2\pi}\ln\frac{\mu}{\mu_0}}. \tag{2.1.3}$$

Here, $\alpha_s = g^2/4\pi$ (for historical reasons, defined in analogy to the fine structure constant of QED), n_f is the number of flavors in the theory, μ_0 is an energy at which we know the value of the constant and μ is the variable that expresses the energy dependence. The value 11 in the denominator arises from the self-interaction of the gluons and determines that α_s decreases as μ increases (asymptotic freedom).

Renormalization allows characterizing one of the defining features of QCD. However, the introduction of an arbitrary scale has larger implications for theoretical QCD calculations that will be used later in this thesis. In particular, scattering amplitudes remain dependent on the choice of this scale. It can be shown that in the calculation of quantum corrections to the gluon propagator, a choice of scale very different from that of the energy of the gluon creates large logarithms that make the perturbative expansion unreliable. This removes part of the arbitrariness on the choice of renormalization scale, since the consistency of the procedure must be preserved. A similar effect comes into play in the calculation of any scattering amplitude in QCD, making the choice of renormalization scale an important one in theoretical calculations of QCD processes.

Confinement, or the increase of the size of α_s at low energies and large distances, also has important consequences in how theoretical calculations are performed. The increase of α_s makes perturbative calculations less and less reliable, and approximations working for specific regions of phase space need to be developed. For soft interactions at low energies, approximations have not been found and models based on experimental measurements are used to allow comparisons to QCD predictions. Despite the high energies involved at the LHC, soft interactions play a large role, whether it is in the description of the interacting partons as part of the accelerated proton, the interactions of spectator partons in the protons generating the hard interaction, or the description of the evolution of the scattered parton as it turns into hadrons that are later detected in the experiment. All these topics are discussed in detail in the next section.

Before describing the specifics of QCD calculations at hadron colliders, it is important to describe the current state of theoretical calculations of scattering amplitudes between partons. Despite the old formulation of the Feynman formalism for QCD (based on the Lagrangian from Equation 2.1.1), this remains a very active research field [55]. The increase of Feynman diagrams factorially with the number of outgoing partons in the computation of scattering cross sections allows very few QCD processes to be calculated beyond leading order (LO). The leading-order calculation exhibits large renormalization scale dependences, and has an inherent uncertainty of $\mathcal{O}(\alpha_s^{-1}) \sim 10\%$ [45]. Next-to-leading-order (NLO) calculations exhibit a much smaller uncertainty ($\mathcal{O}(\alpha_s^{-2}) \sim 1\%$). However, very few scattering cross sections have been calculated up to next-to-leading order, particularly for processes with only quarks and gluons in the final state.

The measurements performed in this thesis are compared to both leading-order and next-to-leading-order calculations (the latter, when such calculation exists). The comparisons to leading-order calculations remain interesting at the new energy frontier insofar as a breakdown of the perturbative approximation, or the existence of new physics, could yield discrepancies between data and the predictions from leading-order calculations beyond their inherent uncertainty. In addition, the measurement of certain shapes and ratios allows probing aspects of QCD that can be approximated to all orders in perturbation theory and have smaller theoretical uncertainties. Some of these aspects are discussed in the next section. The particularities of the implementations of the leading-order and next-to-leading-order calculations used in this thesis will be described in Chapter 9.

2.2 Jet Physics at Hadron Colliders

Quarks and gluons produced at hadron colliders promptly give rise to a spray of particles (mostly hadrons, but also leptons and photons). The energy required to produce this set of particles is drained from the quarks or gluons produced in the collision by the strong interaction field, which increases as the partons separate [56]. Jets are experimental objects that are built from the particles measured in the detector and that allow probing the physics of the hard interaction that created this spray of particles.

A jet is a cluster of energy depositions in the detector, particles or even partons obtained following a specific clustering (or jet) algorithm. Jet algorithms typically limit their clustering to a tunable solid angle, thus giving rise to the collimated nature of jets. The description of the specific algorithm used in this thesis, as well as the motivation for its use, will be presented in Chapter 5. For now, it suffices to clarify that a successful jet algorithm needs to satisfy certain properties that make it applicable to theoretical calculations and practically usable in experiments. The specific choice of algorithm and algorithm parameters can affect the systematic uncertainties in the measurement, and impact the ability to interpret the measurement in terms of the leading-order or next-to-leading-order calculation of parton-parton scattering.

2.2.1 Jet Formation: Parton Shower and Hadronization

Jets are formed from partons according to the Standard Model interactions and, most predominantly, through QCD processes. In the production and consequent evolution of a hard parton, certain properties of the jet can be calculated approximately in the framework of QCD through the "parton shower" approximation. The parton shower approximation allows for performing calculations of the production of collimated particles, where soft and collinear emissions of gluons and quarks are important. Perturbative calculations performed at fixed order are not successful in this regime, due to the existence of large logarithmic terms at all orders in perturbation theory [57] that impact the final result. Instead of attempting a calculation at a fixed order, which is not meaningful in this regime, the parton shower performs an approximation, summing these large logarithms at all orders while ignoring terms of lesser importance.

Figure 2.1 shows one of the Feynman diagrams that give rise to the large logarithms in the calculation of the parton evolution. The probability of this branching happening in a process with n outgoing partons can be calculated in the collinear limit summing over all possible polarizations of the outgoing gluons to arrive at the cross section for the process with $n+1$ partons in the final state:

$$|\mathcal{M}_{n+1}|^2 \sim \frac{g^2}{t}\hat{P}_{gg}(z)|\mathcal{M}_n|^2, \tag{2.2.1}$$

where $t = (1-z)E_a^2\theta_c^2/z^2$, $z = E_b/E_a$ and $\hat{P}_{gg}(z)$ is the gluon splitting function [58]. The splitting functions can be used to calculate the probability of there being an emission of

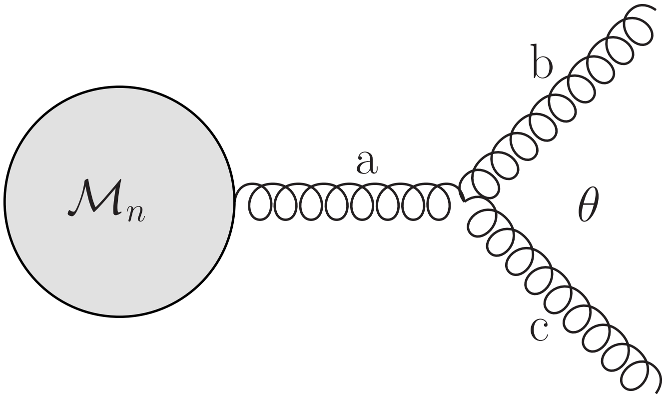

Figure 2.1: *Visualization of the $g \to gg$ term corresponding to one of the large logarithms that are important in the understanding of the evolution of a hard parton generated in the collision of two hadrons.*

some energy off a quark or a gluon, as they evolve from having energy E_0 to having energy E_1. The expression of this probability takes the form of a “Sudakov form factor” [59].

The Sudakov form factor, like the quark and gluon splittings, is singular for soft ($z = 1$ and $z = 0$) and collinear ($\theta_c = 0$) emissions. It is, thus, so far only loosely defined. An infrared cut-off can be introduced to prevent the softest emissions from happening, while removing the divergences in the form factor. This may seem as a somewhat ad-hoc procedure to solve the problem. However, there are natural scales to set this cut-off, whether they are determined by the branching kinematics or the scale of hadron formation, where the parton shower formalism breaks down. The choice of the cut-off scale and other scales involved in the evaluation of the Sudakov form factor was the subject of studies in the wake of the parton shower approximation [60], and remains an issue when attempting to properly describe the shape of jets in data [61].

The procedure of cutting off the integrals in the Sudakov form factors should not be regarded as unphysical, since the divergences are known to cancel when “virtual” corrections (Feynman diagrams with loops and no parton emission) are added to the calculation [58]. The leading virtual corrections are included in the Sudakov form factor through the unitarity requirement. This justifies the introduction of the cut-off as an approximate way of removing the collinear and soft divergences.

The interpretation of the Sudakov form factor as a probability density function, combined with the idea of branching, makes the parton shower easy to implement in Monte

Carlo simulations. A parton from a scattering calculation can be evolved according to the probabilities captured in the form factors, resulting in a "shower" of partons typically stopping close to the infrared cut-off scale. This cascade can be ordered according to the p_{T} or the virtuality ($E^2 - \mathbf{p}^2$) of the partons (each branching reducing the original parton's p_{T} or virtuality). However, it can be shown [62] that branchings at angles, θ, larger than that of the previous branching are strongly suppressed. This has given rise to parton-shower Monte Carlo simulations with angular ordering in the branching. The choice of angular versus virtuality or p_{T} ordering remains an open question in the tuning of Monte Carlo simulations, even though some observables are not sensitive to the specific choice made.

The parton-shower approximation can be used in Monte Carlo generators to evolve the partons from a scattering calculation into a set of partons that capture leading logarithms in the calculation at all orders. Some of these partons may be hard enough to form their own jets and they are loosely referred to as final state radiation (FSR). Similarly, the parton-shower formalism can be used to simulate initial state radiation (ISR). The ISR corresponds to logarithmic corrections to the incoming partons in the scattering process and is important in the context of the collision of protons, since it changes the hadronic interaction tensor [45]. This is discussed in more detail in the next section.

The formation of hadrons typically happens at energy scales at which α_s is large and perturbation theory cannot be used to obtain theoretical predictions. Several models of the hadronization process exist that attempt to connect the results of the parton shower and the final particle spectrum observed [63–66]. These models differ on the specific features of the end product of the parton shower that they try to preserve, and can be complemented by experimental observations to tune some of their predicted spectra (the "jet fragmentation" model). Despite the wealth of hadron collider data available, the different existing models of jet hadronization and fragmentation currently predict different particle spectra in jets. While the differences do not often impact significantly the predicted energy of the jet, they can impact how the energy of the jet is measured, and limit the experimental tools available to determine the quality of the jet energy measurement. Hadronization and fragmentation, thus, remain important topics to be considered in the interpretation of analyses with jets.

An additional level of complication arises in the interfacing of parton-shower Monte Carlo simulations to leading-order calculations (scattering "matrix elements") with multiple partons in the final state. In principle, such interfacing is not needed, since the parton shower is capable of creating final states with multiple partons. However, the energy scales chosen in the evaluation of the Sudakov form factors decrease as more partons are radiated, and this may not correspond accurately to the physics of the production of a final state with multiple partons. To solve this problem of choice of scale, matrix elements can be calculated for the scattering of 2 into n partons ($2 \rightarrow n$), where the choice of renormalization scale depends on the momenta of all n partons. Since these final states are already produced by the parton shower (in the $2 \rightarrow 2$ process), a mechanism to avoid this double counting needs to be devised.

There are currently two mechanisms for interfacing leading-order matrix element calculations to parton showers. Both of these mechanisms rely on a slicing of the phase space in which some part is constrained to be generated by the parton shower, while the rest is covered by the matrix element calculation. The method uses the parton shower where the approximation is most accurate, and the matrix element, where the parton shower approximation has limitations. The MLM matching procedure [67] slices the phase space using geometry and energy, such that soft and collinear emissions are left to the parton shower, while the rest is generated by the matrix element. Illustrations of “4-parton” final states generated by the matrix element, and by the parton shower under this procedure are shown in Figure 2.2. The CKKW matching procedure [68] performs a similar slicing of the phase space, but using a more formal formulation (the Durham jet algorithm [69]) in the definition of soft and collinear emissions. This procedure removes a small double-counting of final states that arises with the MLM matching procedure due to differences in the topologies generated by the matrix element and by the full event simulation. In this thesis, both matching procedures are studied in the cross-section calculations.

2.2.2 Colliding Protons: Partons in Protons and Soft Underlying Event

So far we have made the link between partons and jets, and between the calculation of parton-parton scattering amplitudes, and the measured jet cross sections. However, in a proton-proton collider, the scattered partons are part of the proton, and their energies are determined by the strong interactions that give rise to dynamics inside the proton. These dynamics need to be understood, in order to provide a cross-section calculation of

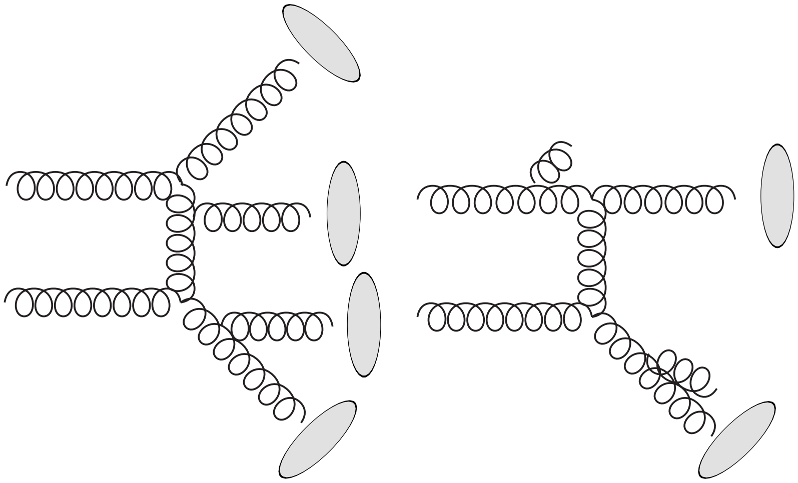

Figure 2.2: *Illustration of a "4-parton" final state generated by the* $2 \to 4$ *matrix element calculation (left) and one generated by the* $2 \to 2$ *matrix element calculation with soft (top) and collinear (bottom) parton emissions (right). The initial state gluons participating in the scattering are shown at the left of each figure.*

jet production in proton-proton collisions. In fact, as it will be explained in this section, these dynamics can generate additional parton interactions and other strong processes that modify the picture of a jet described in the previous section.

The study of the interactions between hadrons is intimately related to the "parton model". The parton model existed before QCD [70] and, in its most naïve version, assumes that the interactions between hadrons can be modeled as interactions between partons within the hadrons. In this model, the probability of a parton carrying a fraction of the total hadron momentum x is captured by the parton density functions ("pdfs"), which are related to the structure functions. The structure functions and the requirement that the model is covariant define the hadron interaction tensor, which can be used as a representation of the hadron for calculations thereafter. At leading order, the structure functions are independent of the energy at which they are probed ("Bjorken scaling") [71]. Bjorken scaling was important early experimental evidence of the existence of quarks inside the proton and that these quarks were spin-$\frac{1}{2}$ particles [39, 50–52].

In the context of QCD the probability for a quark in the proton to emit a gluon can be calculated to provide a higher-order ($\mathcal{O}(\alpha_s)$) correction to the simple parton model. This calculation adds a term to the expression of the structure function, which is proportional to $\ln \frac{Q^2}{\mu_F^2}$, where Q^2 is the scale of the parton-parton interaction, and μ_F^2 is an infrared cut-off introduced in the integrals to remove a singularity caused by collinear gluon emissions. As in any calculation, the introduction of a cut-off must come with a physical justification, and an experimentally backed means to estimate the value of the cut-off. Unlike in the parton shower evolution, this divergence is not cancelled by virtual diagrams in the next-to-leading order calculation, and has to do with non-perturbative effects inside the proton that are not calculable in perturbation theory. In parallel to how high-energy quantum effects are integrated out through renormalization, the long-range non-perturbative interactions in the proton can be captured in a multiplicative term in the structure function

$$F(x, Q^2) \sim x \int_x^1 \frac{d\xi}{\xi} q(\xi, \mu_F^2) \times \left\{ \delta(1 - x/\xi) + \frac{\alpha_S}{2\pi} P(x/\xi) \ln \frac{Q^2}{\mu_F^2} + \mathcal{O}(\alpha_S^2) \right\}, \tag{2.2.2}$$

where F is the proton structure function, P is a splitting function, $q(\xi, \mu_F^2)$ is the long-range non-perturbative term and all other notations have been already introduced. This allows factorizing perturbative and non-perturbative effects at the expense of introducing an arbitrary "factorization" scale (μ_F) and requiring data to fix the non-perturbative term (q) for each specific value of x and Q^2.

Just as the evolution of α_s with energy was calculated by imposing that the physics not depend on the renormalization scale, a similar procedure can be used to determine the evolution with energy of the structure functions. This procedure yields the DGLAP equations [72–76] that can be used to predict the pdfs at high-energy hadron colliders using data from lower-energy colliders and deep inelastic scattering experiments. These predictions cannot cover all ranges of x and Q^2 because the larger the energy, the lower the values of x (and higher the values of Q^2) that can be accessed in the perturbative regime. However, global fits to pdf data are key tools for hadron experiments and updates by new experiments are only provided in the small regions of phase space that previous experimental data did not have access to.

Despite the predictive power of the DGLAP equations for the evolution of the structure functions with energy, the structure functions remain dependent on the value chosen for the factorization scale. This dependence is smaller as higher orders in perturbation theory are taken into account in the calculation of the structure function. As for the renormalization scale, the $\ln \frac{Q^2}{\mu^2}$ dependence requires $\mu \sim Q$ so that large logarithms do not render the fixed-order calculation inaccurate. However, the specific choice of factorization scale remains a topic of study when performing fixed-order calculations of the structure function, and needs to be considered when comparing theoretical calculations to experimental measurements.

The collision of hadrons does not only present the complication of understanding how the scattered partons acquire their momenta, but also requires knowledge about the interactions between the partons that do not participate in this hard scatter. These interactions are soft and their cross section and properties have not been calculated in the perturbative regime. Several models for how these interactions happen exist, and the different models have been implemented in different Monte Carlo generators [77–79]. These models try to include several non-perturbative effects, like the probability of having an additional parton interacting, i.e., multiple parton interactions, or how soft partons created in this way may interact through the strong force with partons from the parton shower of the hard scatter to form hadrons, e.g., color connections and reconnections. These effects are collectively referred to as the underlying event (UE).

The UE impacts how the parton from the hard scatter showers and hadronizes, and may place additional energy in the direction of the formed jet. An understanding of the UE is, thus, necessary to be able to connect jet measurements to the properties of the hard scatter. The different UE models cannot currently be derived from first principles, and the parameters that enter each model need experimental constraints so that they can be determined. A UE tune is a set of model parameters determined through a fit to a specific set of collider data. Since today's UE models have limitations in the accuracy with which they describe processes they are trying to model, new UE tunes are needed for new colliders, and even the new tunes are not without uncertainties in describing the data. There is an inherent theoretical uncertainty in the interpretation of the results that arises from these limitations. Several UE tunes and models are typically used to get some handle on this uncertainty and the corresponding impact of the UE in measurements of jet production.

Figure 2.3 illustrates the most important effects that go into the calculation of the jet production cross sections at a hadron collider. This chapter has built this complex picture of a hadron-hadron collision starting with the QCD Lagrangian, and touching on the different models and approximations needed to go beyond the perturbative regime. Current limitations in theoretical calculations of these processes have also been described.

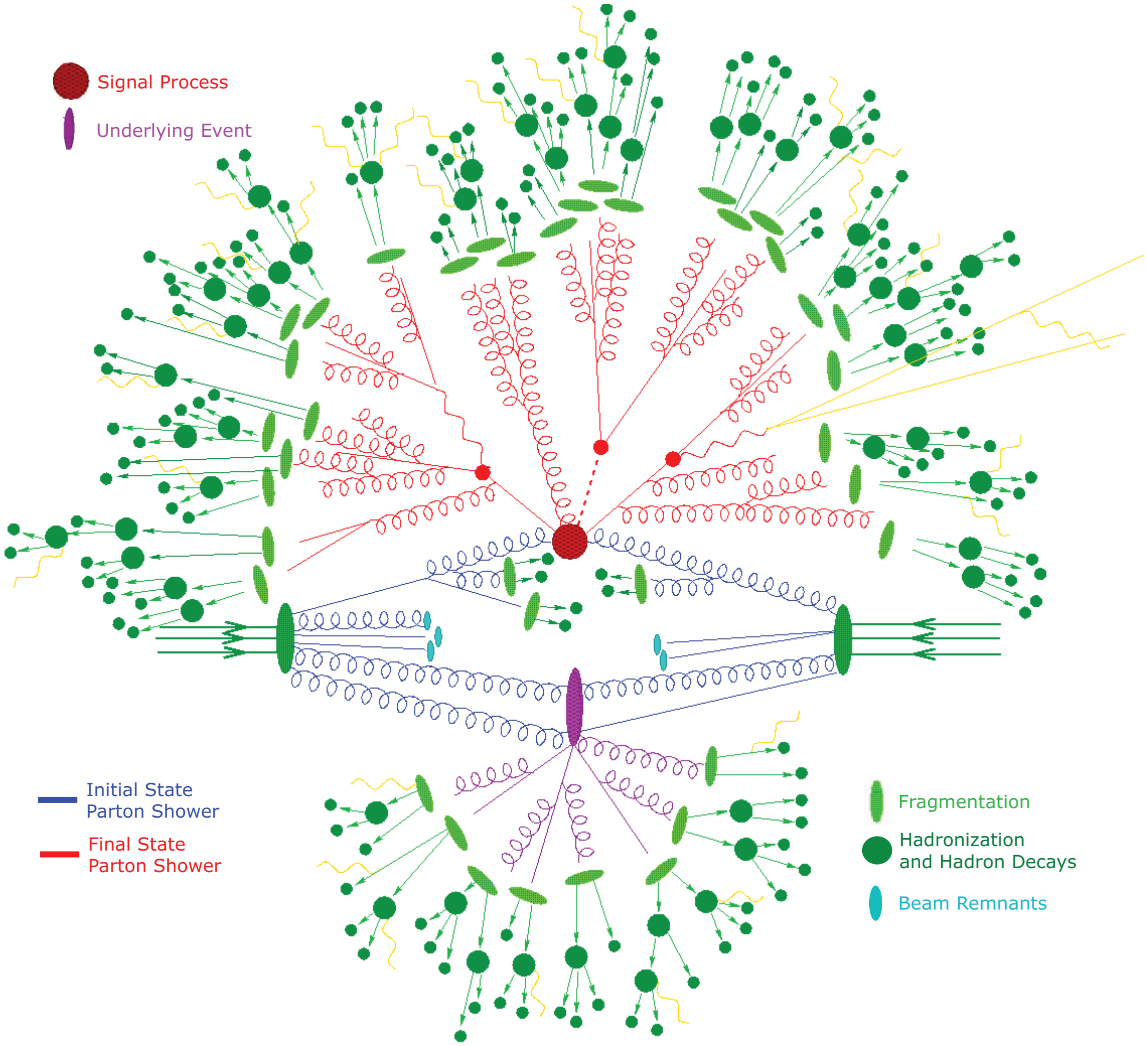

Figure 2.3: *Illustration of the complex picture of a hadron-hadron collision introduced in this chapter (adapted from Ref. [80]).*

In Chapters 9 and 10 these limitations will be addressed with a set of different Monte Carlo simulations in the effort of interpreting the measurements performed in this thesis in the context of QCD.

Chapter 3

The Large Hadron Collider (LHC)

The Large Hadron Collider (LHC) [81–84] is an unprecedented technological achievement for the fields of particle and accelerator physics. The primary design goal of the LHC is to accelerate protons up to energies of 7 TeV and produce proton-proton collisions at four locations around the accelerator at a high rate. The ATLAS detector used in this thesis is placed at one of those four locations. During the first year of operation of the LHC, protons were accelerated up to energies of 3.5 TeV. This limits the intensity of the electrical currents that create the magnetic field necessary to keep protons in the accelerator, preventing possible damage to the magnets due to high resistivity in faulty magnet connections [85].

Protons are created from hydrogen atoms that are ionized in an electric field. The acceleration chain is composed of a series of particle accelerators that progressively increase the energy of the protons. This chain is depicted in Figure 3.1. Protons are first accelerated to 50 MeV using a linear accelerator: the LINAC 2. They are then boosted by two circular accelerators (the BOOSTER) to 1 GeV, before entering the Proton Synchrotron (PS) where they are further accelerated to 26 GeV. The proton beam is then injected to the Super Proton Synchrotron (SPS) where protons reach the minimum energy at which the LHC can maintain a stable beam: 450 GeV.

Upon injection in the LHC, the two beams have a structure imposed by the number and frequency of the injections required to achieve the desired characteristics of the run. The longitudinal and transverse spread of the beams are large and the rate of collisions is low. Before the beam is accelerated, the beam is “locked” to the accelerating device, essentially consisting of an oscillating electric field in the direction of the beam with a frequency of

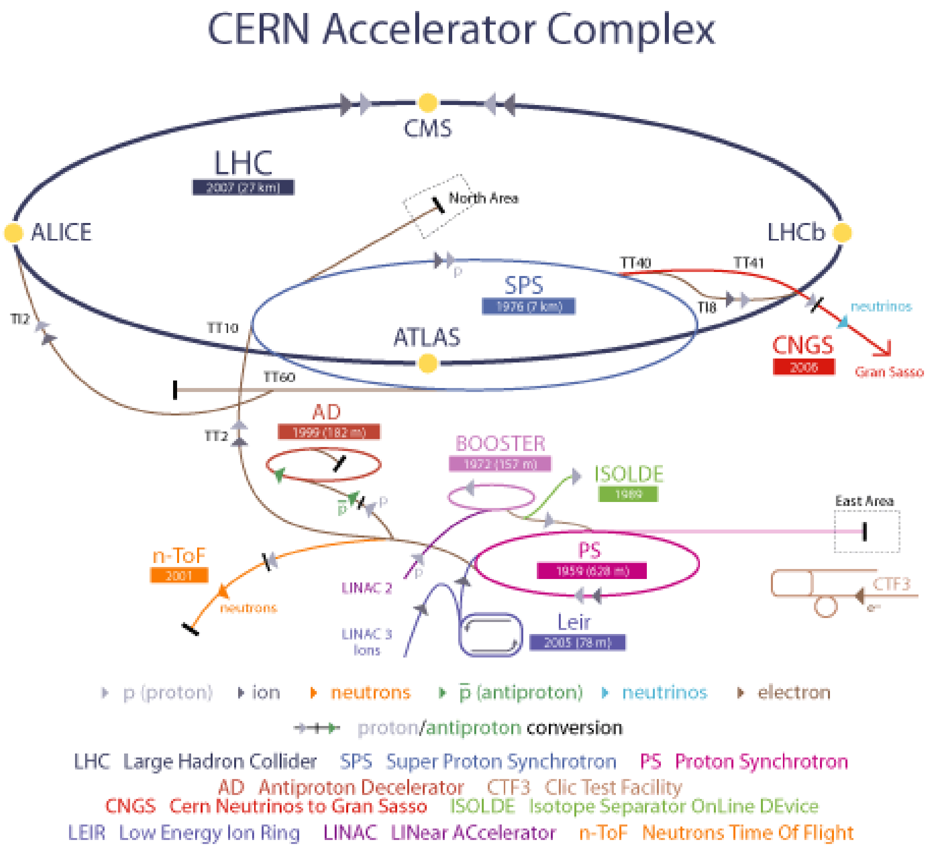

Figure 3.1: *The CERN accelerator complex, comprising the LINAC 2, the BOOSTER, the PS, the SPS and the LHC [84]. Tunnels and accelerators used for other experiments are also shown.*

oscillation that is a multiple of the revolution frequency of the accelerator: $f_{\text{osc}} = nf_{\text{rev}}$. The oscillating electric field generates stable regions of phase space (in momentum and phase of oscillation) where protons are not accelerated. Protons near those regions are accelerated or decelerated until they are bunched around the stable regions. Once the beam is locked, the magnetic field is increased adiabatically. This shifts the momentum of the stable regions of phase space, and protons are accelerated as they are pushed towards the new stable conditions, in what is referred to as synchrotron acceleration.

The stable regions correspond also to spatial regions along the longitudinal direction of the beam. Therefore, the locking of the beam changes its structure, creating bunches of protons with a reduced longitudinal extension (0.75 meters at the LHC) centered around the stable regions imposed by the synchroton acceleration. The number of these stable regions is directly related to the frequency of the oscillations of the accelerating device $n = f_{\text{osc}}/f_{\text{rev}}$, where f_{rev} is slightly over 11 kHz in this case. At the LHC n is equal to 35640, and only one-tenth of these regions can be filled with protons. At design specifications, the LHC fills 2808 of the 3564 available regions with $1.1 \cdot 10^{11}$ protons per bunch, leaving room for beam injection and abort procedures. These bunches are organized in several “bunch trains”. The spacing between the bunches within a bunch train is approximately 25 ns, while bunch trains are further apart. For the data used in this thesis at most 36 bunches were filled, with varying quantities of protons per bunch going even beyond the design $1.1 \cdot 10^{11}$. Bunches in the beams used to produce these data were separated by at least 1 μs.

Once beams are accelerated to the desired energy, stable beams are declared and the information is broadcast to the experiments around the ring, which typically start then taking data usable for analysis. The intensity of beams decreases with time due to a loss of protons in the bunches through collisions used by the experiments or through other effects influencing the trajectories of individual protons in the ring. After the intensity is reduced beyond some factor, around $1/e$, the beam is directed out of the accelerator, the beam is “dumped” into a large metal block where it is absorbed. At this point, the beam injection procedure can re-start. This beam cycle is called a “fill”, and it typically lasts around 10 hours. During the 2009 and 2010 runs, given the concurrent LHC commissioning and development, the time between fills was typically 24 hours.

The rate of collisions is not only dependent on the number of bunches and the protons per bunch, but also other features of the beam and the cross section for protons to interact with one another. All beam-dependent parameters that impact the rate of collisions are grouped under the instantaneous luminosity, $\mathcal{L}$, such that the rate of collisions, $\mathcal{R}$, can be expressed as $\mathcal{R} = \mathcal{L}\sigma_{pp}$. The luminosity can be simply calculated as

$$\mathcal{L} = \frac{f_{\text{rev}} b N^2}{A}, \tag{3.0.1}$$

where f_{rev} is the accelerator frequency, b is the number of colliding bunches, N is the number of protons per bunch and A is the transverse cross section of the overlap between the beams. This equation can be used directly to calibrate the absolute luminosity measurement in any detector. The two quantities that need to be measured are N and A. N is measured by beam current measurement systems installed on the accelerator ring [86]. The parameter A can be measured using the method developed by van der Meer [87], in which the beams are moved and the relative rate of collisions is measured as a function of beam separation. Once the absolute measurement has been performed using the “van der Meer scan”, and that measurement has been correlated to measurements in the detector, the instantaneous and integrated luminosities can be calculated for any run. The detectors and techniques used in ATLAS to measure the run-dependent luminosity are further described in Chapters 4 and 6.

The LHC produced its first proton-proton collisions on November 23, 2009 at the injection energy of 450 GeV. Less than a week later, it became the most powerful collider in the world, reaching beam energies of 1.18 TeV. Collisions at 2.36 TeV were recorded by the four experiments around the LHC ring in mid-December 2009. The LHC broke a new record on March 19, 2010 by raising the beam energy to 3.5 TeV before the first proton-proton collisions at this energy were recorded on March 30, 2010. The analysis performed in this thesis uses proton-proton collision data from the LHC collected by the ATLAS detector from April 10 to August 30 2010.

Chapter 4

The ATLAS Detector

4.1 Detector Overview

Figure 4.1 provides a general overview of the ATLAS Detector. Like many collider experiments, ATLAS is built in different layers, the sub-detectors in each layer serving a different purpose. Three different functional layers are identifiable:

1. Inner detector [89–91], used for tracking of charged particles.
2. Calorimeters [92,93], used for measuring the energies of electrons, photons and hadrons.
3. Muon spectrometer [94–96], used for tagging muons and measuring their momenta.

These three layers are discussed in separate sections in this chapter. Additional detectors exist to monitor the luminosity and the beam position. Those most relevant for the analysis presented in this thesis are described in a separate section. The data-acquisition system, which helps select interesting events and reads out all the detectors above, is described afterwards. Finally, the Monte Carlo simulation of the ATLAS detector is discussed, due to its importance as a tool to understand how the detector operates and its central role in the analysis presented in later chapters.

The ATLAS reference system is a Cartesian right-handed coordinate system, with the nominal collision point at the origin. The counter-clockwise beam direction defines the positive z axis, with the x axis pointing to the center of the LHC ring. The pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$, taking the polar angle, θ, with respect to the positive z direction. The rapidity is defined as $y = 0.5 \times \ln[(E + p_z)/(E - p_z)]$, where E denotes the

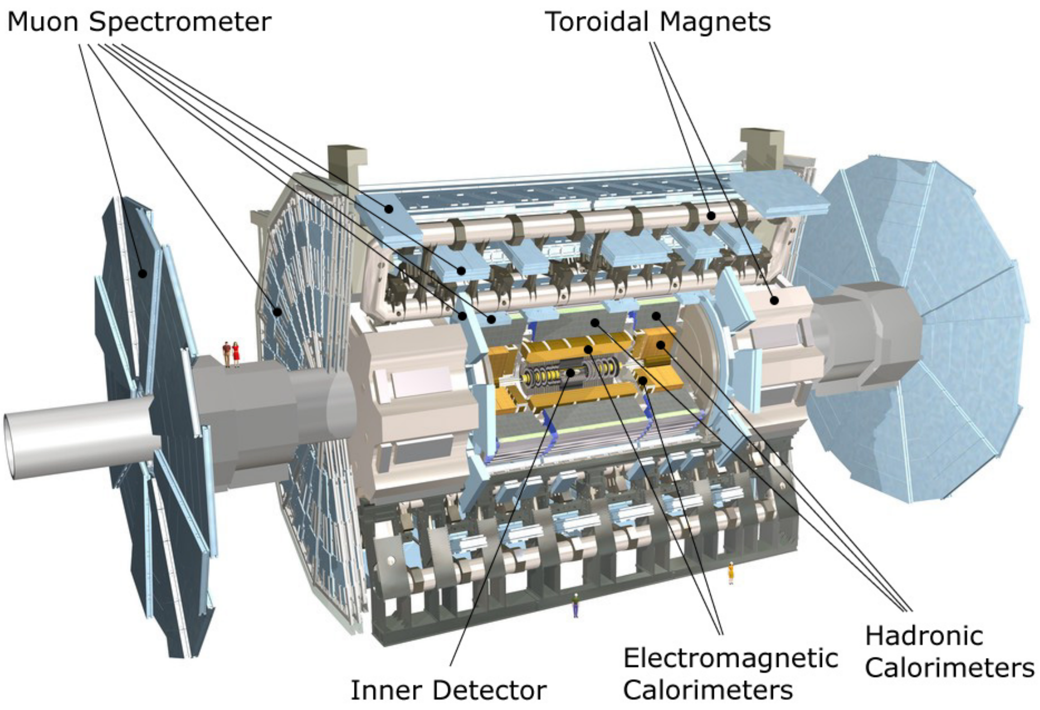

Figure 4.1: *Illustration of the ATLAS detector [88]. Some of the detectors mentioned in the text are labeled.*

energy and p_z is the component of the momentum along the beam direction. For massless objects, the rapidity and pseudorapidity are equivalent. Since the partons that give rise to the signal process of interest carry an a priori unknown fraction of the proton momentum, the overall boost of the collision is not known. For this reason, boost-invariant quantities are used in most analyses. In this thesis, the rapidity, the pseudorapidity and the projection of the momentum on the x-y plane (the transverse momentum or p_{T}), are kinematic variables that play a prominent role in the analysis and in the description of general features of the detectors and their performance.

Protons from the two LHC beams collide around the ATLAS collision point. The particles produced traverse first the three sub-detectors of the inner detector: the pixel detector, the silicon microstrip detector (SCT) and the transition radiation tracker (TRT), in that order. These detectors consist of several layers and use different technologies to determine the position of charged particles in each layer. They travel under the influence of a 2-Tesla magnetic field generated by a solenoid placed at the entrance of the calorimeter. Therefore, the trajectory of a charged particle in these detectors is curved. The curvature is inversely proportional to the particle's momentum. The position information provided in each layer, thus allows for a determination of the momenta of charged particles.

After the inner detector, particles reach the calorimeters. Particles deposit their energies first in the electromagnetic calorimeters, built using liquid argon as the active material. These calorimeters are designed to measure the energy and position of photons and electrons with high precision. Hadrons and other particles with enough energy to not be fully absorbed in the electromagnetic calorimeters deposit their remaining energy in the hadronic calorimeters. The end-cap hadronic calorimeters are built with the same active material as the electromagnetic calorimeter. In the central area of ATLAS the hadronic calorimeter uses scintillating tiles as its active material. A particularly dense liquid-argon calorimeter covers the forward region to guarantee the hermeticity of the calorimeter system, and to allow measurements of jets in the high-$|\eta|$ region.

Of all known interacting particles, only muons are penetrating enough to reach beyond the calorimeters. A tracking sub-detector is placed outside of the calorimeters to measure their momentum: the muon spectrometer. The muon spectrometer incorporates several tracking technologies to measure the position and the direction of minimum ionizing particles at different points along the particles' trajectories. The bending for this sub-detector is provided by toroidal magnets placed also beyond the calorimeters.

When particles interact with the detector components, signals are generated in the detector front-end electronics. These signals are buffered in the front-end electronics and transmitted to the off-detector electronics (outside the experimental hall or "cavern" where ATLAS lies) if an accept signal is received from the level-1 trigger. The off-detector electronics preprocess and pack data and move them to read-out servers, where data wait for the level-2 trigger decision. If the event passes the level-2 trigger selection criteria, the information is moved to the event builders. The event builders pack the data from one event in one data structure. This data structure is sent to the event filter, which performs some additional checks to accept or reject an event. If accepted, the event is written to tape where it can be accessed for analysis. For the analysis performed in this thesis, only the level-1 jet trigger is used. The other trigger levels were not necessary due to the relatively low collision rates obtained during the first year of the LHC operation.

4.2 The Inner Detector

The inner detector (ID) is a tracking detector designed to reconstruct the trajectories of charged particles traversing it. Since the ID is embedded in a 2-Tesla magnetic field, the reconstruction of the trajectory allows momentum determination of charged particles with $p_{\mathrm{T}} \gtrsim 150$ MeV [97]. The structure of the ID is summarized in Figure 4.2. Each of the three sub-detectors is divided in a barrel region, where the detector modules are laid out in cylindrical layers, and an end-cap region, where disks are used to increase the detector coverage in η without a large increase in detector surface. The ID is of primary importance in the analysis performed in this thesis for the identification of the primary vertex, and the determination of whether a jet is originating from it, as discussed in Chapters 5 and 6.

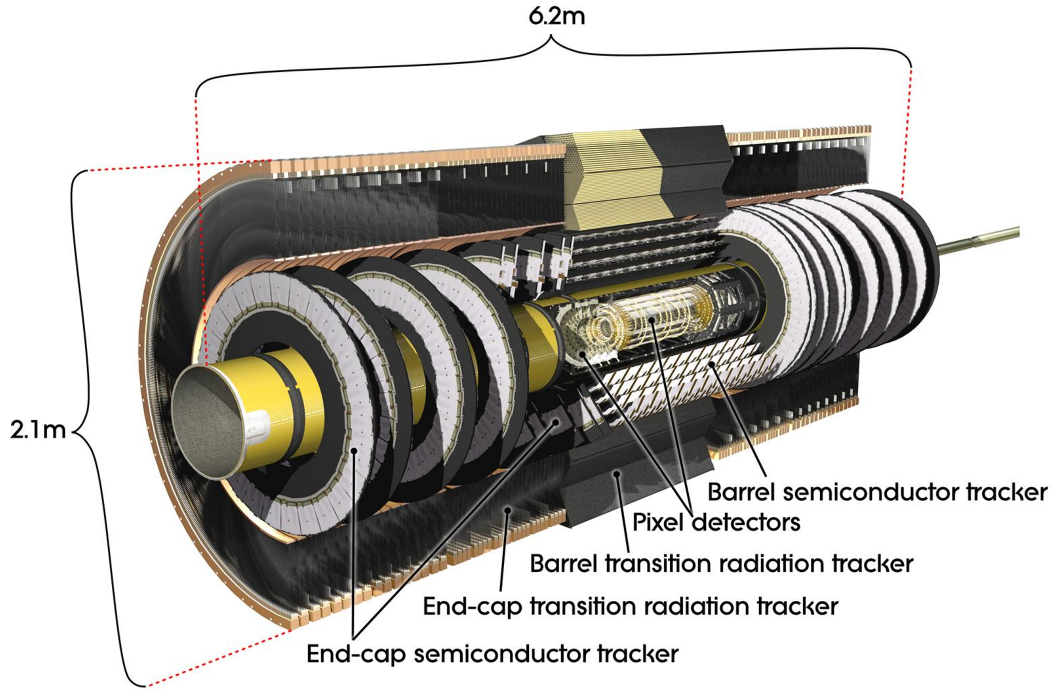

Figure 4.2: *Illustration of the ATLAS inner detector [88]. The sub-detectors mentioned in the text are labeled.*

The first two sub-detectors of the ID use semiconductor technology. When an ionizing particle traverses a semiconductor, it creates movable electron-hole pairs in numbers that are proportional to the particle's energy loss. An externally applied electric field can be used to collect that charge and generate a signal. The energy deposited, in combination with a measurement of the particle trajectory, can be interpreted successfully in order to identify the particle that traversed the detector. However, the primary use of the semiconductor detectors in the ID is to determine the particle's position. This is achieved by measuring a signal over the background (a "hit") in the detector read-out electronics. Semiconductor materials can be modified (doped) to include impurities with one more (p-doped) or one less (n-doped) valence electron. In order to improve the signal-to-background ratio, and to allow detector operation after heavy radiation damage, the ATLAS semiconductor detectors use junctions of p-doped and n-doped semiconducting materials.

The read-out electronics of semiconductor detectors can be structured in different ways according to the environment that the detector will operate in. Figure 4.3 shows some of the most common layouts. Layout (a) only allows measuring one of the hit coordinates (within the detector plane), while layouts (b), double-sided strips, and (c), pixels, provide three-dimensional position information. Double-sided strips often provide an appropriate compromise between the number of channels (and cost) and the information that can be extracted from the detector. A similar layout is used in the SCT: two single-sided strip detectors placed on top of each other with an angle between the strips. In high-density tracking environments, however, the position determination in strip detectors becomes ambiguous if two or more particles traverse the same detector module. For this reason, the pixel layout is used for the innermost ATLAS sub-detector.

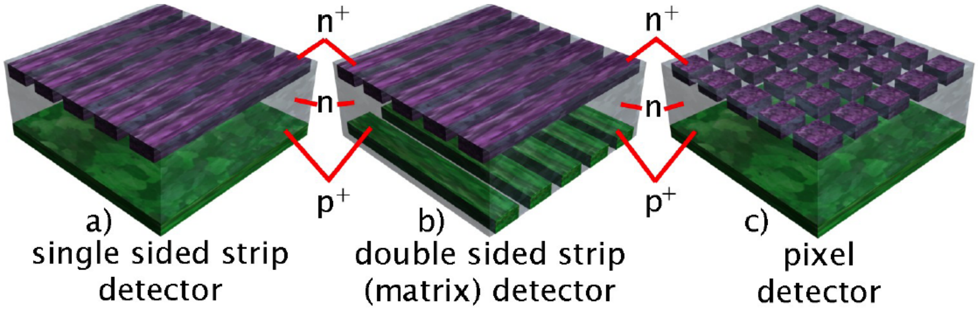

Figure 4.3: *Different semiconductor detector layouts commonly used [98]. The doping pattern is illustrative and coincides with that encountered in the ATLAS pixel detector.*

The outermost sub-detector of the ID uses small drift tubes (straw tubes) as tracking devices instead of semiconductor technology. Particles passing through the tubes ionize the gas inside the tube, and the charged atoms and electrons are pulled apart by the electric field existing between the walls of the tube and the thin wire going through the center of the tube. The detected charge is used to build a signal that is used to determine if there was a hit, and how far from the center of the straw this hit occurred. The straw tubes in the barrel are placed parallel to the beam direction, providing no information about the pseudorapidity of the particle. In the end-cap disks, a similar limitation exists, since the straw tubes are placed radially. The TRT provides many tracking points (typically around 30), and thus helps improving the measurement of the track p_{T} [99].

The TRT has an additional function in the discrimination between electrons and pions. The straws are interleaved with foils and fibers of a material with a different refraction index than that of vacuum. When any particle goes through these fibers, it radiates energy in the form of low-energy ($\sim$ keV) photons. The amount of energy radiated is proportional to $\gamma = E/m$, where E is the energy of the particle and m is its mass. Electrons thus radiate ≈ 250 times more energy than pions. These photons are primarily radiated in the direction of the particle, and can be detected by the straw tubes traversed by the particle to provide transition-radiation hits that allow the identification of a particle as an electron.

4.2.1 The Solenoid

The solenoid providing a 2-Tesla magnetic field for the ID lies inside the cryostat used for the barrel electromagnetic calorimeter (see Section 4.3.1) and in front of the calorimeters. Its operating temperature is, however, lower than that of the calorimeter, 4.5 K, and it is cooled using liquid helium. The design had to cope with stringent constraints on the amount of material, to have a small impact on the energy measurement in the calorimeters. The solenoid is made of a single-layer aluminum coil wound by a NbTi conductor. The inner radius of the solenoid is of 1.23 m, and it is placed such that the beam goes through its axis. Its total length is 5.8 m and its nominal operating current is 7730 A. This provides the 2-Tesla magnetic field that embeds the ID. This field is returned through the steel support structure of the hadronic barrel calorimeter, beyond the calorimeter volume (see Section 4.3.2).

The understanding of the solenoidal magnetic field is important for accurate momentum determination in the ID. Given the interactions between the solenoidal field and magnetic materials in and around the ATLAS detector, a complex simulation accurately describing the detector geometry is used to provide an accurate prediction of the magnitude and direction of the magnetic field at each point in space (see Section 4.7). This prediction is verified through the use of probes operating under the principles of nuclear magnetic resonance (NMR) and the Hall effect [100]. The Hall probes were only used during a detailed mapping campaign before the detector was installed [101], while four NMR probes remain glued to the cryostat wall for monitoring purposes.

4.2.2 The Pixel Detector

The pixel detector [102] provides three tracking space points up to $|\eta| = 2.5$ with a resolution of 10 μm in the $r - \phi$ direction and 115 μm in the z direction. In order to allow for high resolution in the reconstruction of the vertex coordinates the detector is placed close to the beam, with its innermost layer a mere 5 cm away from the nominal beam interaction point. The detector is thus designed for effective charge collection even after long exposure to radiation.

The detector consists of 1744 modules of approximately 2 cm×6.2 cm assembled on 3 cylindrical layers and 6 disks, 3 on each end-cap. The three cylindrical layers are placed 50.5, 88.5 and 122.5 mm away from the interaction point in the radial direction. The disks are placed at a distance of 495, 580 and 650 mm from the interaction point along the beam direction. In order to provide full ϕ coverage, layer modules are tilted 20° with respect to the tangent to the cylinder and partially overlap with the next module in the $r-\phi$ direction. Disk modules are placed on two layers on each disk, one layer shifted in ϕ by 3.75° with respect to the other for a similar reason. Layer modules are laid out in sets of 13 along the direction of the beam, spanning from $z = -400.5$ mm to $z = 400.5$ mm. Disk modules are placed perpendicular to the beam at a radial distance of 88 mm from the beam. The total extension of the pixel detector is thus roughly 1.3 m in z and 30 cm in diameter.

The pixel detector is operated at between -5°C and -10°C to reduce the damage to the sensors caused by radiation. This is achieved through the use of N_2 as an environmental gas, flushed in the pixel detector enclosure. The pixel sensors operate at a voltage ranging between 150 V and 600 V, the voltage increasing with their exposure to radiation. The voltage is typically turned off at the beginning of an LHC fill to prevent damage to the sensors from beam losses, which are more likely to occur before stable beams are declared. To dissipate the heat created by the detector electronics, cooling pipes carrying C_3F_8 in gaseous form are placed next to the modules. The cooling pipes, along with service cables (sensor read-out, voltage supplies, environmental read-out...), exit the detector through the end-caps. Additional read-out services are also placed in this region, making the overall size of the pixel detector and read-out package around 7 m in length.

A pixel module consists of 16 front-end chips, each with 2880 read-out channels. The read-out components for each pixel are connected to the pixel sensor through "bump" bonds (one per pixel), which make up the only mechanical connection between the sensors and the front-end electronics. The sensors are 250 μm thick, which provides a good compromise between the size of the region for charge collection and the amount of multiple scattering in the detector, which degrades the performance of the charged-particle momentum reconstruction. Pixels that are directly connected to the front-end chips have a size of 50×400 μm. The space between front-end chips along the beam direction is covered by longer pixels (50×600 μm), while the space on the $r - \phi$ direction is covered by pixels that share the read-out with pixels that are bump-bonded to the front-end chips. This allows a full sensor coverage of the module area at the expense of some ambiguity in the $r - \phi$ direction and higher noise in the inter-chip pixels along the beam direction. The total number of pixels per module is thus 47232, while the read-out is performed using 46080 different channels. Figure 4.4 illustrates many of the features of a pixel module mentioned above.

The read-out in the front-end chip consists of a charge sensitive pre-amplifier for each pixel, which transforms the charge collected into a pulse whose length is proportional to the amount of collected charge. Additional logic exists to add time and location (within the front-end) information to the measured signal length. This information remains in the front-end buffers and is overwritten by new events if no signal from the trigger arrives at the module after a latency of $\sim$3.2 μs. If a trigger signal arrives, the event is sent to the module control chip, which further packages the information and sends it out of the detector.

The electrical signal is converted to a more robust optical signal in the service areas of the pixel detector and read-out package beyond the pixel disks. The optical signal is then sent out over tens of meters to the read-out drivers off the detector in the underground

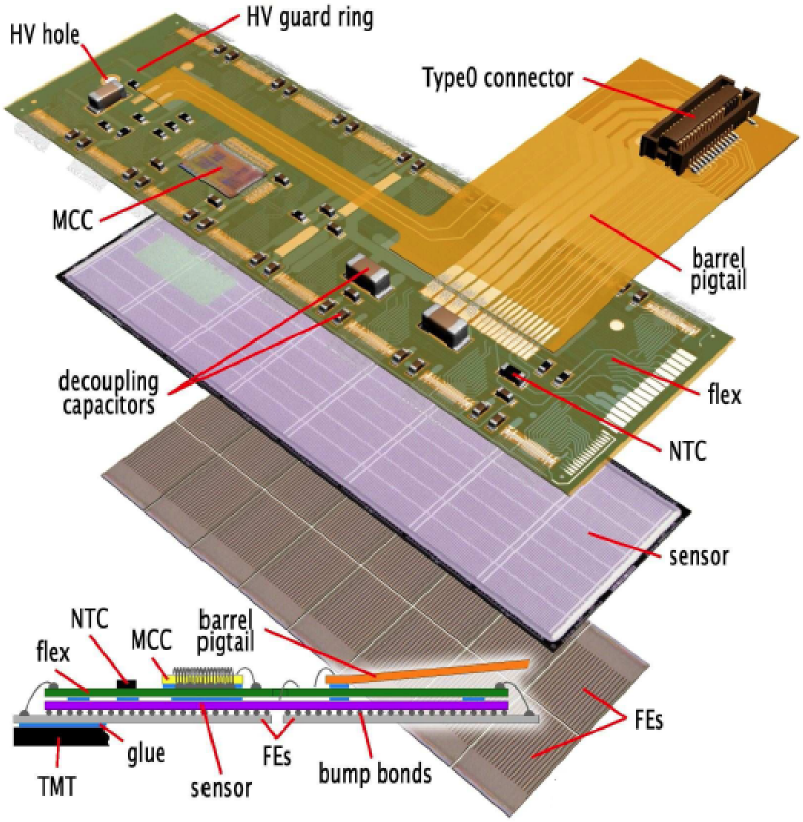

Figure 4.4: *Illustration of a module of the pixel detector showing the sensors, their connection to the front-end chips, the module control chip and other electrical components [88]. The environmental (NTC) sensors and the connector (Type 0) that routes all electrical signals into and outside of the module are also shown.*

"counting rooms". The data enter the read-out drivers as electrical signals, after the optical signals are decoded through a similar mechanism to that used in the on-detector electronics. Some of the integrated circuits performing the electronic signal-to-light signal conversion to send information from the read-out drivers to the modules failed during the 2008, 2009 and 2010 running periods. These failures have been attributed to humidity conditions in the counting rooms [103]. The reason for the failures was only found towards the end of the 2010 run, so no preventive measures could be taken. This was not critical, since the counting rooms are accessible during the run, and replacements could be performed whenever a failure occurred.

The read-out drivers are the last custom element in the read-out chain. Their role in the pixel detector read-out is small during stable data taking (mostly limited to monitoring and local event-building) and their full capabilities need only be exploited for detector calibration. The data is sent from the read-out drivers to the read-out servers, along with data from other detectors, where the data acquisition that is common to all detectors (described later in this chapter) processes the information.

4.2.3 The Silicon Microstrip Detector (SCT)

The silicon microstrip detector provides four tracking space points up to $|\eta| = 2.5$ with a resolution of 17 μm in the $r - \phi$ direction and 580 μm in the z direction. This is achieved through 8 layers of microstrip sensors laid out in pairs in the barrel, and 18 in each end-cap. The two layers in each pair are placed back-to-back and the strips in one of the layers are tilted along the plane of the strips by 0.040 radians with respect to the strips in the other layer. This provides three-dimensional hit information and determines the resolution in the z direction. Many of the design considerations for the SCT are similar to those of the pixel detector, once the smaller radiation, track density and required space-point resolution (determined by its location) are considered.

The detector consists of 4088 modules of rectangular (barrel) or trapezoidal (end-cap) shape assembled on 4 cylindrical layers (2112 modules) and 18 disks (1976 modules), 9 on each end-cap. The four cylindrical layers are placed 299, 371, 443 and 514 mm away from the interaction point in the radial direction. The disks are placed at a distance ranging from 934 to 2720 mm from the interaction point along the beam direction. As in the pixel detector, full ϕ coverage in the barrel layers is obtained through a tilt in the modules with respect to the tangent to the cylinder (of approximately 11° this time) and a partial overlap with the next module in the $r - \phi$ direction. Disk modules provide full ϕ coverage through a small tilt with respect to the plane perpendicular to the beam, and a small overlap. Disk modules are placed on up to three concentric circumferences (depending on the disk), and each circumference is shifted up or down along the beam direction to allow for overlaps between modules in the radial direction.

The SCT operates in the same environmental conditions as the pixel detector, and it shares the system providing cooling for the modules. The voltage applied on the sensors varies from 150 V up to 450 V as radiation damage increases. The SCT service cables and cooling pipes enter the detector primarily through the gap between the barrel and the end-cap, unlike those of the pixel detector, which enter at a point beyond the SCT end-cap disks. This is clearly illustrated in Figure 4.5 where the amount of material in the ID and

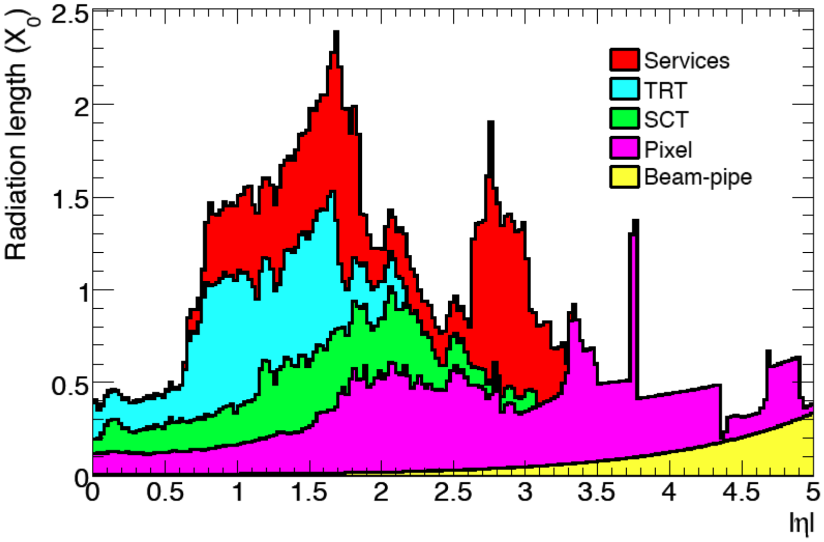

Figure 4.5: *Material inside the ID volume in units of the radiation length split by sub-detector [88]. The material taken up by the services for the SCT and TRT (*$|\eta| \sim 1$*) and those for the pixel detector (*$|\eta| \sim 2.7$*) is shown in red.*

its services as a function of $|\eta|$ is shown. The amount of material is shown in units of the radiation length, X_0, (the average distance an electron has to travel in the material to undergo bremsstrahlung). A clear increase in the service material happens at around $|\eta| = 1$ and $|\eta| = 2.7$. The former corresponds to the region where the service cables and supports for the SCT (and the TRT) are concentrated, while the latter corresponds to the service material for the pixel detector.

The SCT barrel modules use four 80 μm wide strip sensors. The read-out electronics are placed near the center of the module, with two sensors on each side. Each of the sensors is rotated by ±0.020 radians with respect to the center of the module, and contains 768 strips. The four sensors thus form a 128 mm long and 64 mm wide unit. Figure 4.6 illustrates the geometry of the barrel modules.

The SCT end-cap modules are built differently, with a trapezoidal shape, where the long side of the trapezoid is placed far away from the beam. Three types of modules are used to build the end-cap disks, the differences being in their size (varying from 52.48 mm to 126.09 mm in length) and thus the sensor size too. Each of the three types is used to cover one of each of the concentric circumferences building an end-cap disk. The strips on

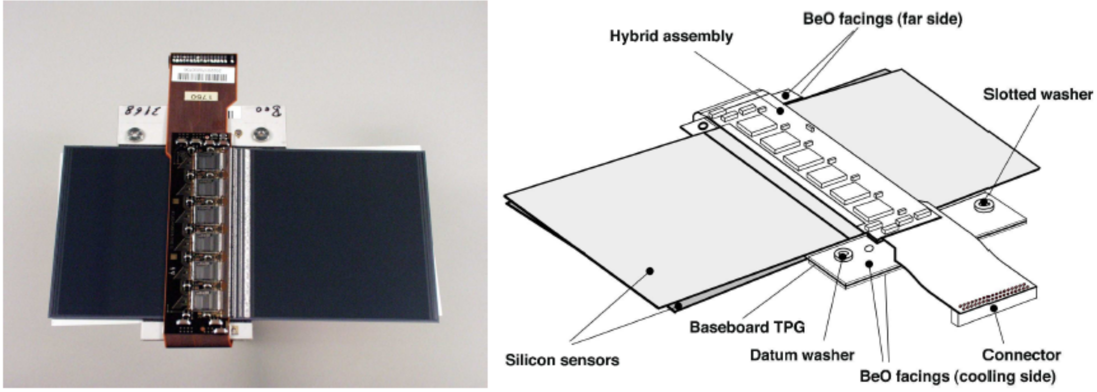

Figure 4.6: *Picture (left) and illustration (right) of a SCT barrel module showing the sensors, the read-out electronics (on the hybrid assembly), the support structure and the connector that routes all electrical signals into and outside of the module [104].*

the sensors are on average 80 μm wide, but there are 768 strips per sensor, in such a way that the strip size changes as it goes from the small base to the large base of the trapezoid. This gives strip sensors that have a width that varies from 56.9 μm to 94.2 μm. As in the barrel, there are sensors on both sides of the module rotated with respect to the center of the module by ± 0.020 radians to provide three-dimensional hit reconstruction. The sensors are 285 μm thick, providing a similar compromise in the number of radiation lengths as the pixel detector.

The read-out of the SCT is performed using an integrated circuit placed near the center of the module (in the barrel modules) or on the long base of the trapezoid (in the end-cap modules). The integrated circuit contains all input/output circuitry as well as the control circuitry needed for module configuration and calibration. A pre-amplifier and other signal processing devices exist for each channel. The integrated circuit contains buffers to store hit information for ~ 3.2 μs, and information is only packaged and sent out after a trigger request. As for the pixel detector, the electric signal is transformed into an optical signal before it is sent off the detector to the read-out drivers. The read-out drivers used for the SCT are the same as those used for the pixel detector, and the conversion from optical to electric signal (and vice-versa for sending commands to the detector) is performed by the same devices. The controlling software running on the read-out drivers and on the single-board computers installed next to them is different for each detector, matching the different requirements.

4.2.4 The Transition Radiation Tracker (TRT)

The TRT [105,106] provides on average 36 tracking space points up to $|\eta| = 2.0$ with a two-dimensional design hit resolution of 130 μm. The large number of tracking points allows almost continuous tracking that improves the pattern recognition of tracking algorithms, thus improving the precision of the momentum measurement. The TRT provides also electron/pion discrimination through the use of transition radiation foils and fibers. Unlike the pixel detector and the SCT, the TRT provides limited η information on the track. This adds ambiguity to the track reconstruction, but this ambiguity can be resolved using information from the pixel detector and the SCT.

Drift tubes of 4 mm in diameter (straw tubes) are the basic detector elements of the TRT. A total of 73 layers of 144 cm long straws make up the barrel module, while 160 layers of 37 cm long straws are used to build each of the two end-cap modules. Straws in the barrel are placed in the direction of the beam, while straws in the end-cap are placed at fixed ϕ, extending radially on the plane transverse to the beam direction. The straws in the barrel are organized in 96 modules that provide the support structure for the straws and contain 52544 straws in total. The straws are read out from each end (giving two channels each) and they are inactive in the middle, where mechanical support for the anode wire is provided. Each module is composed of a carbon-fiber shell and an array of straws embedded in a matrix of polypropylene fibers, 19 μm in diameter, that enable the transition radiation. The average spacing between the straws is 7 mm. The barrel modules extend 78 cm in the negative and positive direction along the direction of the beam, and cover the region with radii between 55.4 cm and 108.2 cm on the plane transverse to the beam. The straws in the end-cap are distributed in 20 “wheels” containing a total of 122880 straws per side. Straws are organized in disk-shaped layers perpendicular to the beam direction. Layers of straws are 8 mm apart for modules close to the interaction point and 15 mm apart for modules further away. Each layer contains 768 straws with uniform azimuthal spacing, and is rotated by 3/8 of the straw angular distance to guarantee full ϕ coverage. The spacing between end-cap layers is filled with 15 μm thick foils of polypropylene radiator.

The TRT is designed to operate at room temperature. However, CO_2 is circulated inside each module to prevent high-voltage discharges, flush away any gas leakage from inside the straws and prevent the diffusion of gas from the outside into the straws. Heat dissipated by the straws is conducted by the CO_2 to the carbon-fiber shells, which are cooled using the same system used for the cooling of the TRT front-end electronics. Given the very different operating environments between the TRT and the silicon detectors, the different sub-detectors are placed in different thermal enclosures that allow for control over their environment and isolate them from the environment in the other sub-detectors.

Both barrel and end-cap straws have 31 μm diameter gold-coated tungsten wires going through their center that serve as anodes. The wires are kept at ground potential and connected directly to the front-end electronics. The walls of the straws act as the cathode and are kept at ≈ -1500 V during operation, yielding a maximum collection time of ~ 50 ns for drift electrons. The straws are filled with a gas mixture containing 70% xenon, 27% CO_2 and 3% O_2. This mixture provides a high gain for ionization signals from charged particles, while absorbing low-energy transition radiation photons, which give much larger signals. The difference in the signal amplitude allows for particle identification using transition radiation signals and tracking using ionization signals.

The read-out of the TRT is performed using printed circuit boards that hold the wire ends and are placed on the outer circumference of the end-cap cylinders and on the sides of the barrel modules. Two integrated circuits perform the data processing and packaging before the data is sent off the detector. The first of these circuits reads out eight channels, amplifies and shapes the signal, and includes a low-threshold (for ionization signals, at ≈ 250 eV) and a high-threshold (for transition radiation signals, at ≈ 6 keV) discriminator. The second circuit reads out 16 channels, performs the measurement of the drift time to increase the precision of the position measurement, and contains the necessary digital circuitry to interface to the trigger and control the tunable parameters of the detector. Once a trigger is received electrical signals are sent to a patch panel inside the muon spectrometer, where the conversion from electrical to optical signals happens. Optical signals are then sent to the counting rooms, where custom-built read-out drivers re-package the data and send it in a common ATLAS format to the read-out servers.

4.3 The Calorimeters

The ATLAS detector uses sampling calorimeters to measure the energy of all interacting particles (except muons, which are highly penetrating and are not absorbed in the calorimeters). Figure 4.7 shows the layout of the different calorimeters around the ID. As for the ID, the different components of the calorimeters and their read-out structures roughly point to the interaction point, giving rise to distinct detector geometries in the barrel and the end-caps.

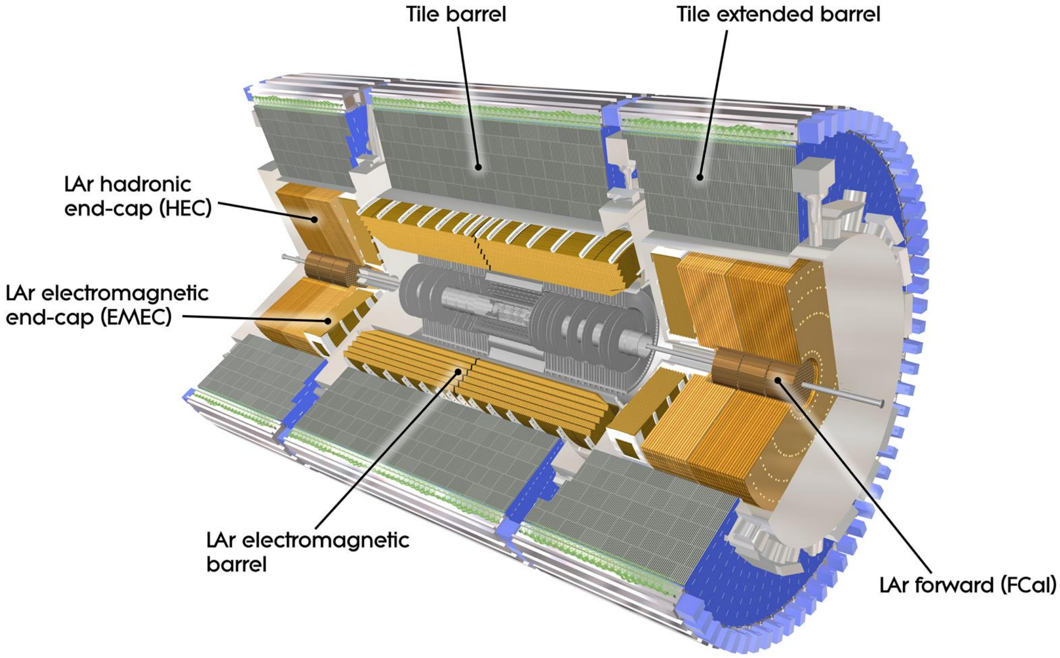

Figure 4.7: *Illustration of the ATLAS calorimeters [88]. The different calorimeters mentioned in the text are labeled.*

Sampling calorimeters are made of two types of materials: the absorber and the active material. Particles deposit their energy in both the absorber and the active material, even though only the energy deposited in the active material is measured. The absorber is typically chosen to be made of a dense material with high stopping power so that size constraints on the detector can be met while guaranteeing that particles in the energy range of interest are contained in the calorimeter. Energy deposited in the active material is deposited by particles produced through interactions of the incident particle in both the absorber and the active material. The incident particle interacts with the calorimeter material through several processes (bremsstrahlung, e^+e^- pair production, hadron-nucleon collisions,...), whose importance depends on the nature of the particle, its kinematics and the material that makes up the absorber.

For charged particles interacting electromagnetically, most of the energy lost in the absorber is lost through bremsstrahlung photons. The interplay between bremsstrahlung photons and the subsequent e^+e^- pair production leads to an exponential cascade of particles, until the produced particles have low enough energy that Compton scattering and absorption through the photoelectric effect become important [107]. Fluctuations in the shape of this cascade along the direction of the shower development cause fluctuations in the measurement, given that only interactions in the active material give rise to detector signals. The reduction of these fluctuations motivates a commonly used design of sampling calorimeters in which layers of active material are intercalated to layers of absorber material.

Even with optimal design, calorimeters do not manage to measure all the energy of the original particle, due to energy that remains in the absorber and other effects. The "response" of the calorimeter usually refers to the fraction of the energy of the incident particle that is measured. The response is related to the "sampling fraction" $E_{\text{active}}/(E_{\text{active}} + E_{\text{absorber}})$, where E refers to the energy deposited by the particle, but it also includes additional effects such as the energy lost in non-instrumented regions of the detector. A high sampling fraction is desirable to minimize measurement fluctuations. The response and the sampling fraction can be different for hadron and electrons, and they may also depend on the particle's energy. For this reason, calibrations need to be applied to properly measure the energy of the incident particle.

If the calorimeter energy measurement is linear with the energy of the incident particle, the number of particles n produced in the shower is proportional to the energy of the incident particle. Event-to-event fluctuations in the energy measurement correspond to Gaussian fluctuations on n and the frequency with which the energy loss of each particle is sampled by the active layers, N. Therefore, the relative precision of the measurement will go as $\sim 1/\sqrt{nN}$. This leads to the commonly used formula to parameterize calorimeter resolution: $\sigma_E/E = a/\sqrt{E}$. This equation is often expanded when discussing calorimeter performance on analysis objects to account for noise in the electronics, energy that might be lost in non-instrumented areas of the detector and various other calibration effects:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{4.3.1}$$

where the E^{-1} term accounts for noise in the electronics, the $E^{-1/2}$ term is commonly referred to as the stochastic term and the $\oplus$ symbolizes addition in quadrature.

The ATLAS calorimeters can be classified according to their design goals as electromagnetic and hadronic calorimeters. The forward calorimeter can be considered to fall into a different category, since its design is much more heavily driven by the radiation resistance required for detectors placed very close to the beam at the LHC and its relatively small size. The electromagnetic calorimeters are placed in front of the hadronic calorimeters, in part because particles that interact only electromagnetically require less material to be absorbed. This can be quantified in terms of the radiation length (introduced in Figure 4.5) and the nuclear interaction length (the average distance a hadron has to travel in the material before undergoing an inelastic hadron-nucleon collision). For lead, one radiation length corresponds to 0.033 nuclear interaction lengths [31].

The electromagnetic calorimeters are designed for the identification and measurement of photon and electron energies. The identification is performed to a large extent through the study of the energy deposition patterns ("shower shapes") [8]. These studies are made possible by a high granularity in η and ϕ as well as a segmentation in the read-out along the direction of the development of the shower. These calorimeters use lead as the absorber material and liquid argon as the active material. Liquid argon was chosen as the active material due to its resistance to radiation and its uniformity, which translates into spatial uniformity in the energy measurement. The stochastic term in the electromagnetic calorimeter energy measurement for electrons and photons is ≈ 0.1 GeV$^{-1/2}$, while the constant term is ≈ 0.002.

Hadrons deposit a large fraction of their energy in the electromagnetic calorimeters. However, the probability of being fully absorbed in the electromagnetic calorimeters is small, and becomes smaller with increasing hadron energy. The hadronic calorimeters are designed to contain the showers of high-energy hadrons (spanning up to 10 nuclear interaction lengths), have a linear response at high energies and provide precise measurements of jet energies down to 20 GeV. These goals are achieved through an iron-absorber calorimeter that uses scintillating plastic tiles as the active material in the barrel, and a copper-absorber liquid-argon calorimeter in the end-cap. These calorimeters, in combination with the electromagnetic calorimeters, provide a stochastic term of ≈ 0.5 GeV$^{-1/2}$ and a constant term of ≈ 0.05 for the measurement of jets. The constant term can be improved to ≈ 0.03 through the use of sophisticated jet calibrations.

The high-rapidity region is of interest because instrumenting it allows for a better measurement of the p_{T} of neutrinos or other particles that may escape the detector. Calorimetric energy measurements in this region, as provided by the forward calorimeters, also give access to certain Higgs production channels that are associated with forward jet production. At high rapidity, the proximity to the beam and the limited amount of space determine the design of the forward calorimeter. The forward calorimeter again uses liquid argon as its active material and copper and tungsten as absorbers. Its small size is driven by technical requirements (it resides in the same cryostat as the end-cap calorimeters) and by the decreasing solid angle covered by one unit of pseudorapidity with increasing η. In order to provide the absorption capabilities needed in such a confined volume, the amount of absorber is maximized, thus reducing the sampling fraction significantly. This results in larger stochastic term (≈ 1 GeV$^{-1/2}$ [108]) and noise. However, since the interest in this region is to measure very high-energy jets (with a p_{T} that is similar to that of lower-energy jets in the barrel), the performance for jets is comparable to that in the barrel and the end-cap.

4.3.1 The Electromagnetic Calorimeter

The electromagnetic (EM) calorimeter uses liquid argon as its active material and lead as its absorber. It provides electron and photon identification up to $|\eta| = 3.2$, even though the performance is degraded beyond the ID coverage ($|\eta| > 2.5$). Below $|\eta| = 1.8$ the EM calorimeter is complemented by a presampler that consists of a thin (11 mm in the barrel) layer of instrumented liquid argon that conceptually acts as the sampling layer for the material before the calorimeter, which would be the absorber.

The EM calorimeter consists of a barrel section, covering the region with $|\eta| < 1.475$, and two end-cap sections, covering the regions with $1.375 < |\eta| < 3.2$. The end-cap modules are split in two coaxial “wheels”, with slightly different design specifications. The outer wheel covers the region up to $|\eta| = 2.5$ while the inner wheel extends to $|\eta| = 3.2$. The barrel module extends radially from ≈ 1250 mm up to ≈ 2050 mm, even though the volume of the cryostat that houses it and the solenoid providing the magnetic field for the ID starts at 1150 mm. The barrel detector extends 3100 mm from the center of the detector in each direction along z.

The end-cap wheels are 630 mm thick and have external and internal radii of 2098 mm and 330 mm, interfacing at ≈ 520 mm. The end-cap wheels are placed ≈ 3700 mm away from the interaction point in the beam direction. The inside wall of the cryostats that house each wheel, together with the hadronic end-cap and forward calorimeters, are 3500 mm away from the interaction point.

The barrel presampler is placed in front of the EM calorimeter, inside the cryostat, over the full η range of the barrel EM calorimeter. The end-cap presampler is placed also in front of the EM calorimeter, but over a much smaller η range ($1.5 < |\eta| < 1.8$) to provide additional measurement information in a region with larger amount of non-instrumented material due to the routing of services through the gap between the barrel and end-cap calorimeters. The EM calorimeter, with its cryostats and the other calorimeters, are depicted in Figure 4.7, where the structures around the liquid-argon calorimeters represent the cryostats.

The calorimeter read-out is segmented longitudinally along the direction of the particle shower development in three layers for $|\eta| < 2.5$ and two layers elsewhere. An accurate position measurement for photons (for which there is no ID information) is obtained through a fine segmentation in the η direction of the first layer. The first layer is approximately four radiation lengths (X_0) thick only, and thus only a fraction of the energy of a particle is deposited there. Most of the energy is deposited in the second layer ($\approx 17X_0$ thick). The last layer (about $1X_0 - 10X_0$ thick) only contains the tails of the electromagnetic shower, and has a coarser segmentation. The read-out cell size in the end-cap varies also within a layer and between layers. A summary of the calorimeter read-out segmentation in $[\eta, \phi]$ and depth is shown in Table 4.1.

The liquid-argon calorimeters are housed in cryostats that maintain a temperature of ≈ 89 K, at which argon is a liquid at ambient pressure. They are designed to maintain a

Detector and Layer	Segmentation in $[\eta, \phi]$	Radiation Lengths
Pre-sampler	$[0.025, 0.1]$	-
EM Barrel		
1^{st} Layer	$[0.003, 0.0245]$-$[0.025, 0.0245]$	≈ 4
2^{nd} Layer	$[0.025, 0.0245]$-$[0.075, 0.0245]$	≈ 17
3^{rd} Layer	$[0.05, 0.0245]$	$\approx 1 - 10$
EM End-cap		
1^{st} Layer ($\lvert\eta\rvert < 2.5$)	$[0.003, 0.1]$-$[0.05, 0.1]$	≈ 4
1^{st} Layer ($\lvert\eta\rvert > 2.5$)	$[0.1, 0.1]$	≈ 25
2^{nd} Layer ($\lvert\eta\rvert < 2.5$)	$[0.025, 0.0245]$-$[0.05, 0.0245]$	≈ 26
2^{nd} Layer ($\lvert\eta\rvert > 2.5$)	$[0.1, 0.1]$	$\approx 3 - 11$
3^{rd} Layer ($\lvert\eta\rvert < 2.5$)	$[0.05, 0.0245]$	$\approx 5 - 15$

Table 4.1: *Summary of the segmentation in $[\eta, \phi]$ of the different read-out layers of the EM calorimeters in the barrel and the end-cap.*

constant temperature and high purity of the liquid argon. These two aspects are critical to achieve stability in the calorimeter calibration. Liquid nitrogen (N_2) is used as refrigeration fluid, and circulated inside the cryostat through a number of valves and cooling loops. Signal and high-voltage feed-throughs exist to bring the necessary services from outside the cryostat into the cold volume. The signal feed-throughs are distributed radially along the lateral ends of the barrel cryostats and on the back of the end-cap cryostat to minimize the length of the connections. Two high-voltage feed-throughs exist in the barrel cryostat and one in each end-cap. Figure 4.8 illustrates the position of the feed-throughs (and the front-end crates) in the end-cap cryostats.

The EM calorimeter sampling layers have a peculiar structure that allows a finely segmented read-out and provides naturally full ϕ coverage without cracks. This “accordion” structure consists on absorber plates with a zig-zag shape running radially with respect to the direction of the beam. A schematic view of this shape could already be observed in the illustration of the barrel EM calorimeter in Figure 4.7. In the barrel, in order to maintain a constant sampling fraction with depth, the folding angle of the accordion decreases with increasing depth. In the end-cap, the amplitude of the accordion is also varied for the same reason. The absorber layers are 1.53 mm thick for $|\eta| < 0.8$ and 1.13 mm for $|\eta| > 0.8$, again in an attempt to keep a uniform sampling fraction. In the end-cap something similar

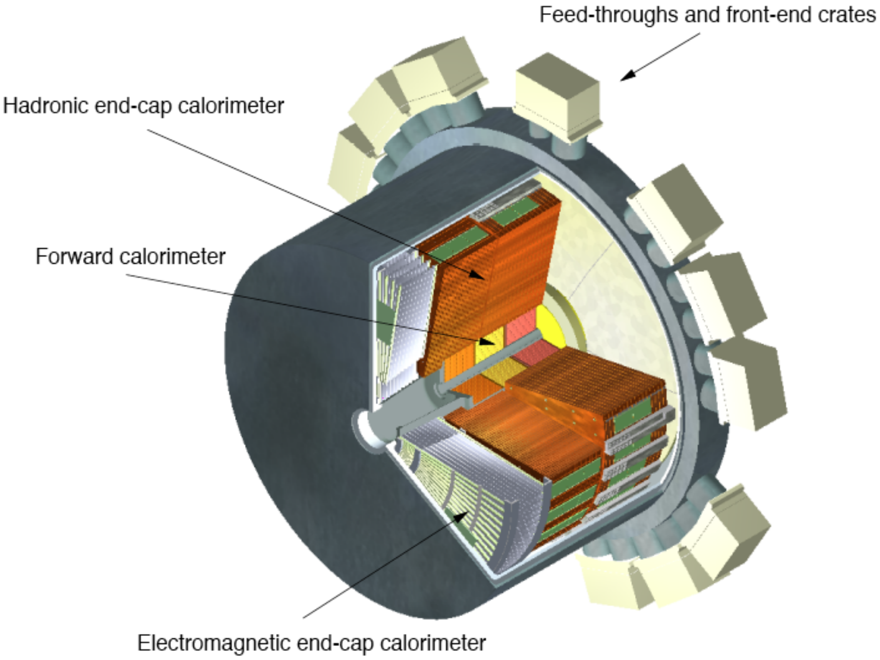

Figure 4.8: *Illustration of the end-cap cryostat, with the signal and high-voltage feed-throughs and the front-end crates [88].*

happens, with plates of thickness 1.7 mm for $|\eta| < 2.5$ and 2.2 mm above.

The accordion absorbers are interleaved with the electrodes that perform the read-out of the particle shower. These electrodes are placed in the middle of the gap between absorber layers supported by honey-comb-shaped spacers. The barrel accordion absorbers with electrodes and spacers are shown in Figure 4.9, in a picture taken before their installation. The electrodes in the barrel are placed a distance of 2.1 mm from the absorbers, which at the nominal operating voltage of 2000 V gives an average drift time of ≈ 450 ns. This is much longer than the nominal LHC bunch spacing, and impacts the energy reconstruction in the calorimeters during operation in nominal conditions. For the data used in this thesis the spacing between bunches was at least 1 μs, and this was not an issue. The electrodes can be read out at the front of the calorimeter (for the first longitudinal barrel layer) or at the back (for the second and third barrel layers.) In the end-cap, the read-out proceeds similarly for the outer wheel, while for the inner wheel both layers are read out through the back due to the high radiation levels at the front.

Once charge is collected in the read-out electrodes, summing boards located at the end of the electrodes add the signals from adjacent gaps, forming a cell. The signals from several

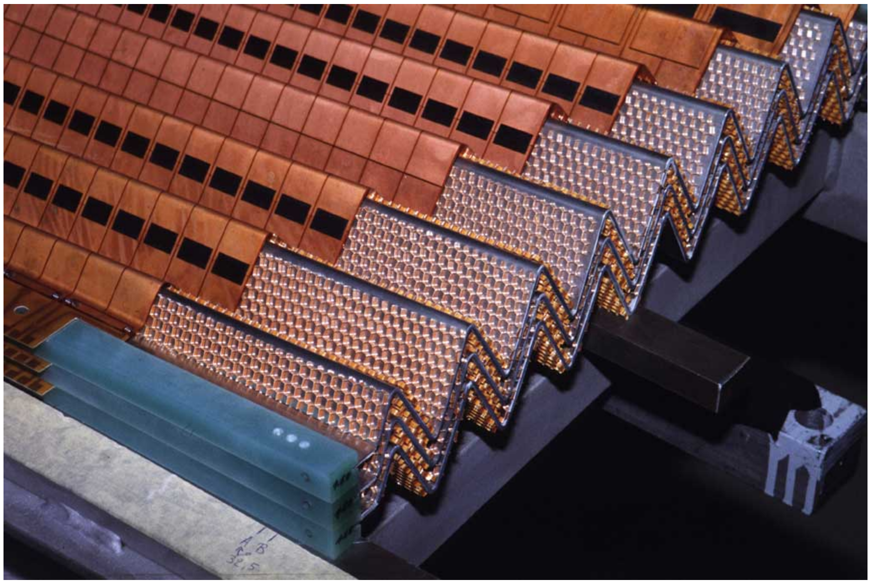

Figure 4.9: *Accordion absorbers with the electrodes, supported by honey-comb-shaped spacers [88].*

summing boards are collected on a mother-board that routes the outputs to the read-out cables, and also provides calibration functionality. The presampler modules use their own mother-boards. Signals exit the cryostats through the feed-throughs and reach the front-end crate where the pre-amplifier and shaping circuitry resides. The signal is transformed into a bipolar pulse shape in the front-end boards. This shape is illustrated in Figure 4.10.

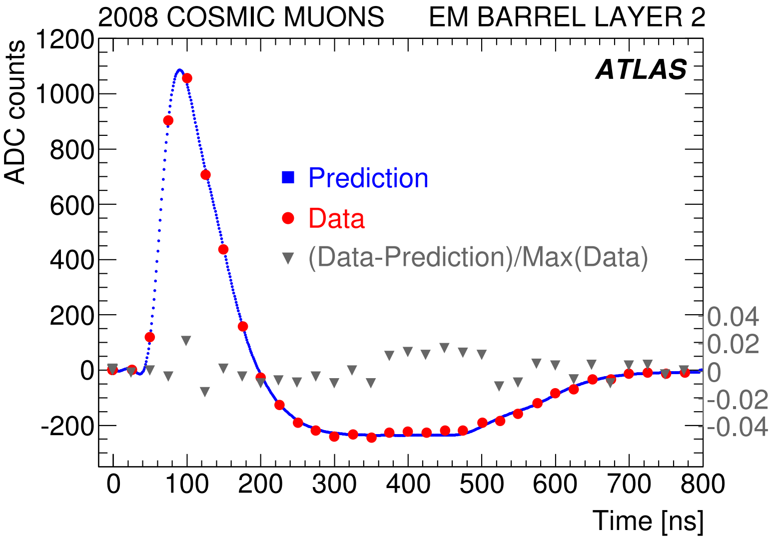

Figure 4.10: *Predicted and measured pulse shape for a signal deposited by a cosmic muon in the EM barrel calorimeter [109].*

While the pulse is continuous, its shape is only sampled every 25 ns up to 32 times (red points.) During regular data taking, the signal is extracted from the first five time samples only, and the rest of the signal suppresses energy deposited by very soft interactions in the bunch crossings that follow. This pulse shaping, with its zero-area time integral, thus cancels on average signal contributions from soft interactions that happen simultaneously with the hard interaction of interest.

The front-end boards also contain the electrical-to-optical conversion electronics, which use the same electronics as those used for the pixel detector and SCT. A small number of failures of these optical links in 2010 led to small regions in the detector with lower-quality energy reconstruction, even though an approximate energy measurement for the affected regions is provided by the independent trigger electronics. Once the electrical-to-optical conversion happens, the digital optical signals are sent to the read-out drivers. The read-out drivers not only repackage the data before sending it to the read-out servers, but also calibrate the signals to deposited energies in units of MeV [110]. The calibration also provides timing information and a simplified quality measure, χ^2, for the measured pulse shape with respect to the expected pulse shape. This χ^2 is later used in the selection of jets for data quality assurance.

The front-end crates also contain circuitry to perform the hardware-level trigger decisions. These include the jet triggers used in this thesis, and are further described in Section 4.6 and Chapter 6.

4.3.2 The Hadronic Barrel (Tile) Calorimeter

The tile calorimeter uses scintillating plastic tiles as its active material and steel as its absorber. It complements the energy measurement of the EM calorimeter for hadrons in the region with $|\eta| < 1.7$. It consists of one barrel module (5.8 meters in length) and two extended barrel modules (each 2.6 meters in length). The barrel module surrounds the EM barrel calorimeter, while the extended barrel modules surround the end-cap cryostats housing the EM end-cap, hadronic end-cap and forward calorimeters. In the gaps between the two, scintillating tiles are placed before the modules providing a similar functionality to that obtained with the presampler: they act as the active material for the non-instrumented services in these gaps.

The tile calorimeter extends radially from 2.28 m to 4.25 m, including the steel support structure on its outer shell, which provides for the flux return of the solenoidal field and houses the read-out electronics. The tile calorimeter operates at room temperature, so no cooling is required for the modules. However, a cooling system operates in its support structure to dissipate the power consumed by the read-out electronics. As for the other calorimeters, the read-out is segmented radially, providing information about the shape of the shower that is important, for example, to estimate the energy lost by particles in the cryostat walls right before it. The read-out cells have an almost "projective" geometry (that is, with their boundaries at approximately constant η and ϕ) and provide a granularity of 0.1×0.1 in $\eta \times \phi$ in all but the last layer, where the η granularity is 0.2. Figure 4.11 shows the thickness of each of these sampling layers in units of the interaction length.

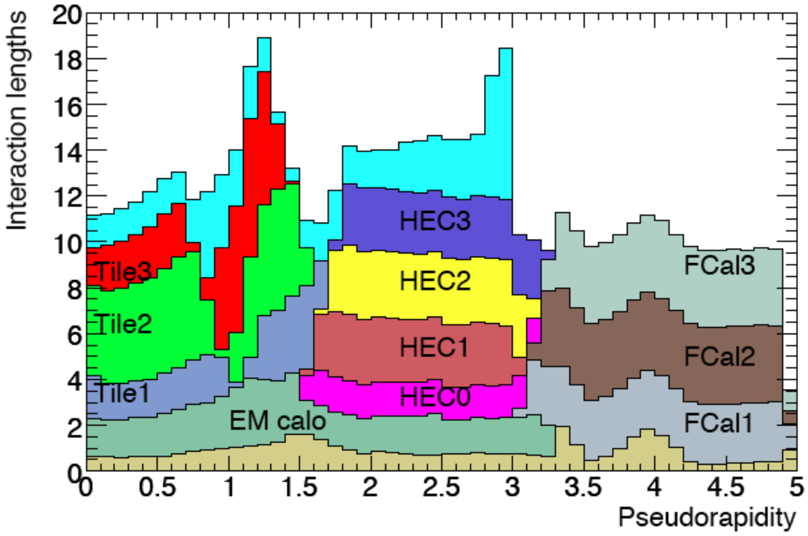

Figure 4.11: *Nuclear interaction lengths in each layer of the hadronic calorimeters [88]. The amount of material in support structures before the muon spectrometer is also shown in light blue as the last of the stacked histograms.*

The basic absorber unit of the tile calorimeter is a 5 mm thick master plate going from the front to the back of the calorimeter onto which small 4 mm thick spacer plates are glued in a staggered fashion along the radial direction. The area not covered by the spacer plates is where the 3 mm thick scintillating tiles are placed. This geometry is illustrated in Figure 4.12. The polystyrene scintillating tiles produce ultraviolet light when particles shower in the calorimeter. This ultraviolet light is converted to visible light by 1 mm wide

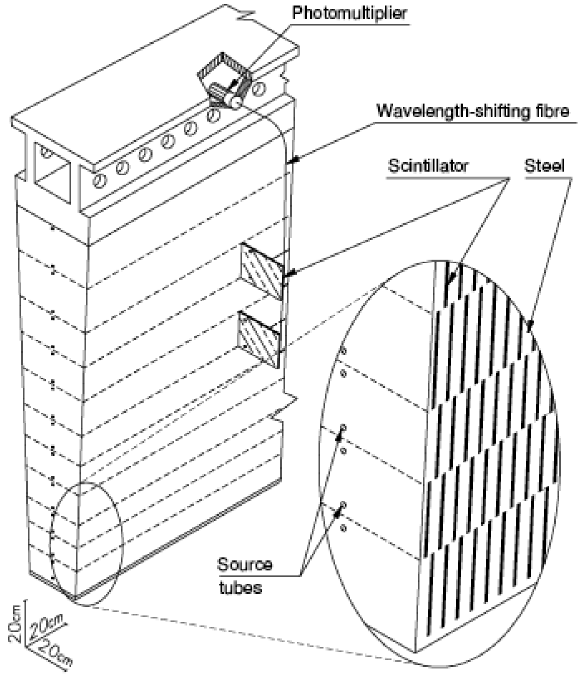

Figure 4.12: *Structure of the absorber and active material in the tile calorimeter [88]. The beam is 2.28 m below the lower extreme of the module as depicted in the figure. The support structure is shown at the top of the figure. The small holes on the side are used to introduce a radioactive source for calibration purposes.*

wavelength-shifting fibers that carry the signal from the scintillators to photomultipliers housed in the calorimeter support structure. This is also illustrated in Figure 4.12. Fibers are grouped as they enter the photomultipliers, defining the cell structure and granularity of the calorimeter.

The read-out components of the tile calorimeter, housed in its support structure, consist primarily of a photomultiplier block and a mother-board. The photomultiplier block includes the photomultipliers and different components to improve the photon collection and shape the signal appropriately (in a unipolar shape, unlike for the liquid-argon calorimeters.) Trigger signals are also produced in the photomultiplier block, even though they are added in the mother-board before being sent out for further processing with other muon and calorimeter signals. The mother-boards distribute power to the photomultiplier blocks, provide analog-to-digital signal conversions and transform digitized signals into optical signals that are sent to the read-out drivers upon a trigger accept. The read-out drivers have the same architecture and functionality as those of the other calorimeters, with some modifications to add further monitoring capabilities required to resolve a fail-safe redundancy provided in the signal transmission off the detector.

4.3.3 The Hadronic End-cap Calorimeter (HEC)

The HEC uses liquid argon as its active material and copper as its absorber. It enables the measurement of energy and absorption of hadrons up to $|\eta| = 3.2$. It shares many of its features with the EM calorimeter, already described in Section 4.3.1, such as its operating environment, read-out service routing and read-out infrastructure. For this reason, none of those topics are discussed here.

The HEC consists of four end-cap wheels, covering the region with $1.5 < |\eta| < 3.2$. Each end-cap has two wheels placed one after the other along the beam direction. Both wheels are cylindrical and have an outer radius of 2030 mm. The wheels have inner radii of 372 (for the first nine absorber plates of the front wheel) and 475 mm. The HEC extends from $|z| \approx 4300$ mm to $|z| \approx 6100$ mm. The copper-absorber plates are flat and placed perpendicular to the beam direction. Their thickness varies from 12.5 mm for the very front plate to 50 mm for most plates in the back wheel. The gaps in between all plates have a thickness of 8.5 mm. The gaps are split into four regions of 1.8 mm each by three electrodes. The middle electrode serves as the read-out electrode and defines the $\eta \times \phi$ segmentation (of 0.1×0.1 for $|\eta| < 2.5$ and 0.2×0.2 elsewhere), while the other two have a voltage of 1800 V applied to them. This yields a drift time of ~ 400 ns, comparable to that in the EM calorimeter. The HEC read-out is segmented along the direction of the shower in four read-out layers. The thickness of these layers in units of the nuclear interaction length have already been shown in Figure 4.11. Unlike the signals produced in the EM calorimeters, the signals produced in the HEC are already amplified in the cold volume, before they are sent out to the front-end crates through the feed-throughs.

4.3.4 The Forward Calorimeter (FCAL)

The FCAL uses liquid argon as its active material and copper and tungsten as its absorbers. It allows for the energy measurement and absorption of all particles (except muons) up to $|\eta| = 4.9$. Like the HEC, it shares many of its read-out and environmental features with the EM calorimeter, already described in Section 4.3.1.

The FCAL consists of two cylindrical structures placed in the two high-pseudorapidity regions ($3.1 < |\eta| < 4.9$) of the ATLAS detector, with the beam passing through the center of the cylinder. Each structure has three 45 cm deep modules placed along the beam direction. The modules have inner and outer radii of $\approx$ 7 cm and 45 cm, and the surface of the first module is located at $z = 470$ cm. Beyond the third module (still in the cryostat) a non-instrumented block of a copper alloy provides further shielding for the muon spectrometer. To optimize the resolution, the first module uses a copper absorber, while the other two modules use tungsten to limit the extension of the particle shower.

The modules are made of plates stacked one after the other with small holes all around the plates lining up to make module-long cylindrical holes. These holes are filled with the detector electrodes. The electrodes consist of tubes of diameter close to that of the cylindrical holes in the modules. Cylindrical rods are placed inside the tubes as to leave a small gap between the tube wall an the rod. These gaps are filled with liquid argon, and vary in size from 0.269 mm for the first module to 0.508 mm for the last module. A voltage difference of up to 600 V is established between the electrodes during operation. This gives a drift time ranging from $\approx$ 60 ns to $\approx$ 120 ns. This structure is illustrated in Figure 4.13. Signals are read out from each module independently, through the front of the first module and through the back of the two other modules. Signals are routed through the side of the modules and added up in summing boards mounted on the back of the HEC calorimeter, before being sent to the front-end crates.

4.4 The Muon Spectrometer

The muon spectrometer is an air-toroid system built with four types of muon detectors providing stand-alone muon reconstruction down to $\approx$ 5 GeV. The muon chambers are tracking devices for precision measurements of the muon position over long distances, in order to provide a momentum resolution of 10% for muons with momenta of 1 TeV. Their momentum reconstruction resolution is best at around 100 GeV (3%, dominated by the effect of multiple scattering in the material in the muon spectrometer) and increases to about 4% at low momenta due to fluctuations in the energy lost by muons in the calorimeters. The resolution at low momenta can be improved to $\approx$ 2% if measurements from the muon

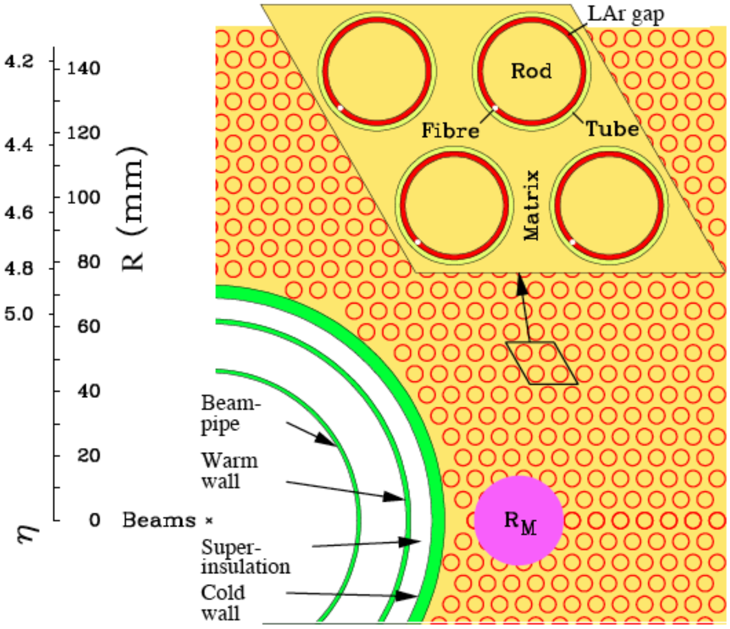

Figure 4.13: *Structure of the FCAL calorimeter absorber, active material gap and electrodes [88]. The dimensions correspond to those of the first module. The Molière radius (R_M: the radius within which an electron or photon deposits 90% of its energy) for the first FCAL module is shown for illustrative purposes.*

spectrometer are combined with measurements from the ID.

The chambers provide coverage up to $|\eta| = 2.7$, even though they have no coverage at $|\eta| \sim 0$, where all the services from the other detectors are routed out of ATLAS. The chambers extend from the end of the calorimeters, about 5 meters away from the beam, to about 10 meters radially. Along the direction of the beam, they extend from around 7 m to over 20 m in each direction. Mechanical stability and good alignment systems are, thus, required to provide the position measurement accuracy required.

A toroidal magnetic field of 0.5–1 Tesla on average provides the bending power for the momentum measurement. This toroidal field is created by eight coils in the barrel and eight coils in each end-cap. Since this magnetic field bends particles only on a plane of constant azimuth, ϕ, that passes through the beam axis, no precise ϕ information is needed to reconstruct the muon momentum. Precision muon chambers are built that focus on the measurement of the coordinate of interest. These are the monitored drift-tube chambers, which are laid out in three layers around the detector. The drift-tube chambers do not operate properly in very high particle density environments, like those existing close to the beam. For this reason, a different technology is chosen for the first layer of precision chambers in the region with $2.0 < |\eta| < 2.7$: a cathode-strip detector.

The precision chambers have a long charge collection time, which results in a long readout latency. This latency does not guarantee that the signals can be used for triggering purposes. That is, by the time the presence of a muon could be verified with the precision chambers, the rest of the detectors would no longer hold the data from the bunch crossing of interest. For this reason, the precision chambers are complemented by trigger chambers. The trigger chambers provide a coarser η measurement (and a ϕ measurement) but have a fast response time that allow their use in the hardware-level trigger decisions. Their position information can also be used by the offline reconstruction algorithms. Two types of technologies are used for the trigger chambers, following radiation hardness and spatial resolution requirements. In the barrel, resistive plate chambers are used, while thin-gap chambers are used in the end-cap. The positions of all the chambers introduced and the toroidal magnets are illustrated in Figure 4.14.

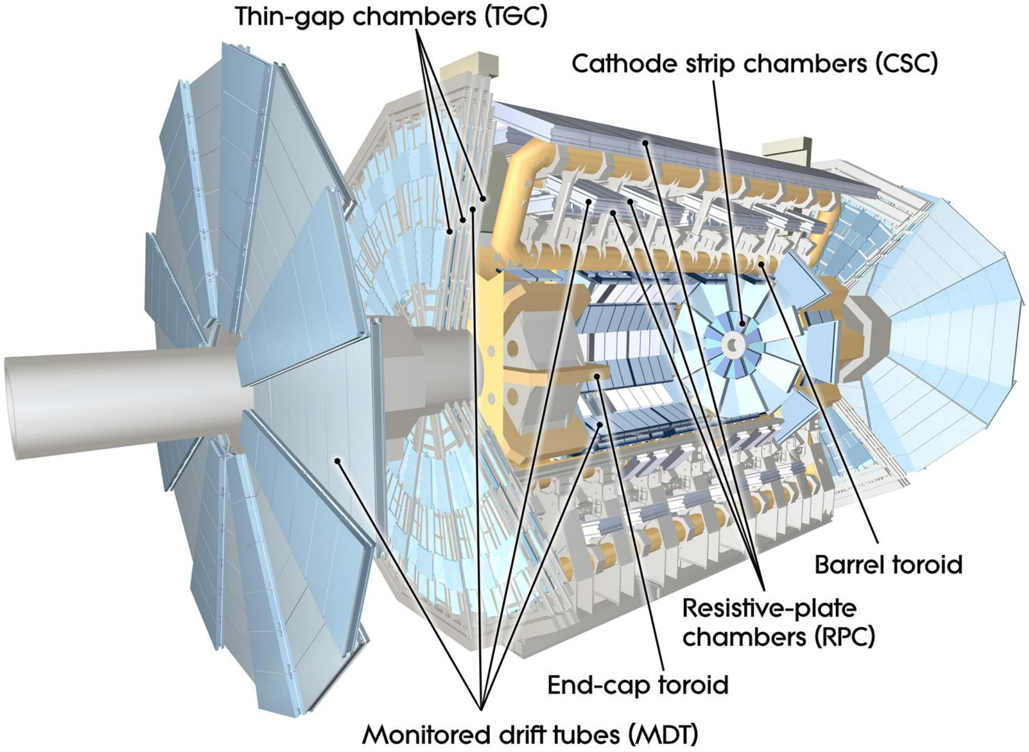

Figure 4.14: *Illustration of the ATLAS muon spectrometer [88]. The different types of chambers mentioned in the text are labeled. The magnet toroids are colored in orange.*

4.4.1 The Toroidal Magnet System

The toroidal magnet system providing bending power for the muon spectrometer consists of eight rectangular coils in the barrel and eight in each end-cap. The coils in the barrel are housed by one cryostat each, while the coils in one end-cap are all housed in the same cryostat. The coils in the barrel are 25.3 m long and over 10 m wide, once inside their cryostats. The coils in the end-cap have a similar width, but are placed closer to the beam (at ≈ 1.65 m) and have a length of 5 m. The coils consist of a conductor made of a mixture of aluminum, niobium, titanium and copper wound into the required coil shape. Like the solenoid, they operate at a temperature of 4.5 K and are cooled by liquid helium. They operate at a current of 20.5 kA and attain a maximum field of ≈ 4 Tesla in the coil windings.

Unlike what occurs in the ID and the solenoidal magnetic field, uncertainties in the magnetic field created by the toroid do not translate into an overall shift in the momentum scale of the measurement performed in the muon spectrometer. Instead, any uncertainty will result in a reduction of the precision of the muon momentum reconstruction. This is due to the non-uniformity of the toroidal magnetic field, which causes particles to behave differently depending on their trajectory. In order to understand this field a combination of accurate simulations and in situ measurements using NMR and Hall probes is used, as for the solenoidal magnetic field. The system of Hall probes is mounted on the detector structure permanently and consists of 1730 probes distributed through the muon spectrometer volume. The input from these probes and the knowledge of the material around the muon spectrometer is used to create a complex magnetic field map that is used for the reconstruction and simulation of muon trajectories through the spectrometer.

4.4.2 The Precision Chambers

The precision chambers are built using two different technologies: monitored drift tubes (MDTs) and cathode strips (CSCs). These provide three tracking points all the way up to $|\eta| = 2.7$, with a position resolution of the bending coordinate of $\sim 50\ \mu$m. This allows stand-alone momentum reconstruction with a resolution of 10% for 1 TeV muons and charge identification capabilities up to ~ 3 TeV.

The precision chambers are placed on three concentrical cylinders in the barrel, a radial distance of 5 m, 7.5 m and 10 m from the interaction point. In the end-cap, they are placed on four wheels, the last one extending up to 12 m radially. The three main wheels are placed 7 m, 13 m and 21 m away from the interaction point along the beam direction. An additional, smaller wheel is placed on top of the end-cap toroid at a distance of 11 m from the interaction point, to guarantee that three tracking points are available in the full η coverage without having to increase the radial size of the last wheel significantly. A gap exists in the region of $\eta \sim 0$, where cables and other services for the inner detector and the calorimeter exit the ATLAS detector. The gap has a size of 0.04 to 0.08 in each η direction. This does not translate into a complete inefficiency except for very high momentum tracks, which are essentially straight.

The first and last barrel layers are placed right next to the toroid cryostats. This imposes a similar geometric constraint on the muon spectrometer, which is divided into 16 sectors. Long and short chambers are used in each sector with an up and down radial displacement that provides chamber overlaps for full ϕ coverage. This is illustrated in Figure 4.15. No gap at $\eta \sim 0$ exists in the ϕ regions covered by the small chambers. The first and second main wheels in the end-cap are also placed right next to the end-cap toroid cryostats, one in front and one in the back. The MDTs are used to cover most of this space, except in

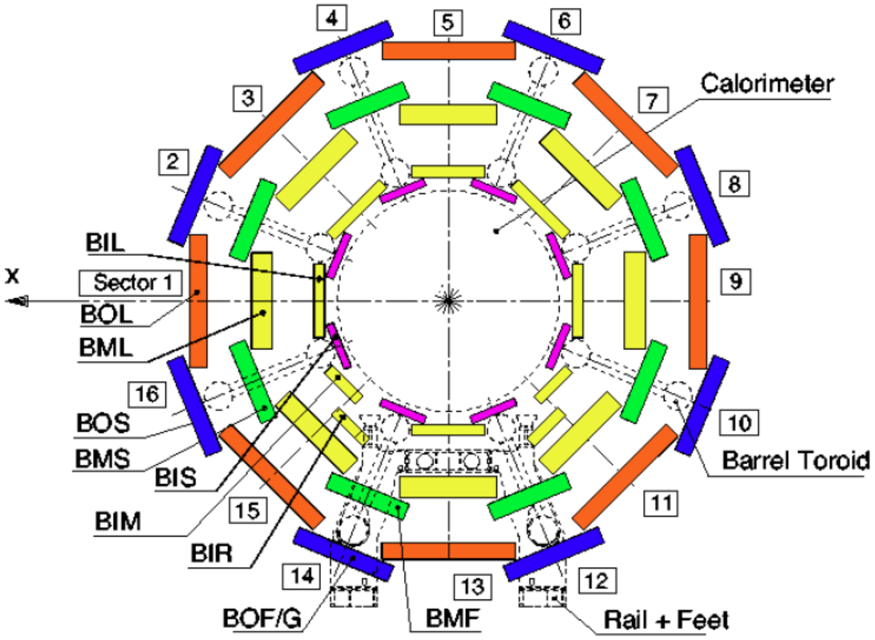

Figure 4.15: *Illustration of the layout of the muon chambers in the barrel [88].*

the first end-cap wheel, in the region $2.0 < |\eta| < 2.7$, where particle density is higher, and a higher-rate detector is needed.

The MDTs consist of a number of pressurized drift tubes of ≈ 3 cm in diameter, using an argon/CO_2 mixture as the drift gas. A tungsten-rhenium wire collects the charge produced through the ionization caused by the muons traversing the tube. The wire is at a nominal voltage of ≈ 3 kV. This provides a maximum drift time of 700 ns. The measured drift time is used to determine the coordinate of the muon with the required precision. The drift tubes are placed in rectangular chambers in the barrel, and trapezoidal chambers in the end-cap, always tangential to a cylinder whose axis corresponds to the beam direction. Tubes in an end-cap chamber vary in length since the trapezoid's small base is closest to the beam. In order to improve the tracking performance, a chamber consists of two multi-layers of three (four, for the innermost barrel chambers) drift tube layers each, spaced by a mechanical structure. This is illustrated in Figure 4.16.

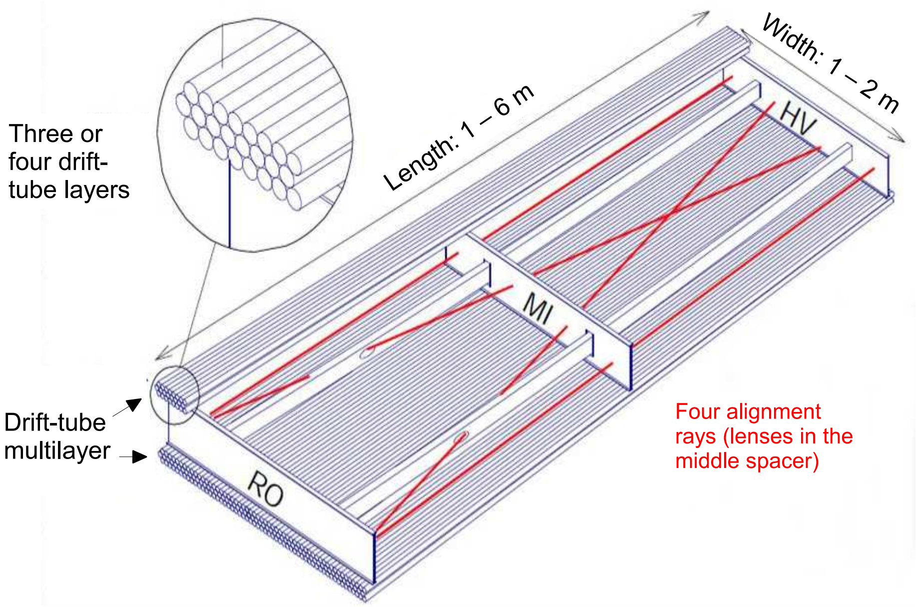

Figure 4.16: *Illustration of the layout of the drift tubes in a barrel muon chamber [88]. The interface to the read-out (RO) services and the high-voltage supplies (HV) lie on the mechanical support structure. The optical alignment system is also schematically shown.*

This figure also illustrates an optical alignment system within a chamber. This system guarantees a good understanding of changes in the chamber mechanical structure due to thermal gradients and other factors. Changes in the mechanical structure due to the gravitational pull have also been taken into account in the chamber construction, and a mechanical system to correct for it exists in the barrel chambers. In addition, an optical alignment system exists to monitor the position of each chamber with respect to its neighbors [111, 112]. This also applies to the CSCs. This system allows for a relative alignment and needs to be complemented by track-based alignment to attain full knowledge of the chamber position. All these systems guarantee that the intrinsic drift-tube accuracy for position reconstruction can be reached in spite of the large mechanical structures and distances needed in this sub-detector.

The CSCs are also segmented into large and small chambers and the full system consists of two disks (shifted along the z direction) with eight chambers each in each end-cap. The CSCs are multiwire proportional chambers. That is, wires at a voltage (1900 V for the CSCs) lie sandwiched between strip cathodes in a gas mixture (argon/CO_2 in this case), creating an electric field that can be used to measure charge created by an ionizing particle. Wires run radially and the read-out of both cathodes is segmented, providing a measurement in both the bending direction and in ϕ. The strip size in the bending direction is ≈ 5 mm, which is sufficient to achieve a ≈ 60 μm position resolution. In the non-bending direction, the achievable position resolution is ≈ 5 mm. The spacing between the cathodes is 5 mm, which results in a maximum electron drift time of 40 ns.

The MDT read-out happens through one of the ends of the drift tubes as illustrated in Figure 4.16. Analog electronics on a chip perform the amplification, shaping and discrimination of the tube signal, before the information is sent to the time-to-digital converter (TDC), where further signal processing, time-stamping and other event-level information is added to the signal and the signal is placed in a buffer awaiting the trigger decision. All these electronics are placed on the chamber, one TDC corresponding to 24 drift tubes. Each chamber is controlled by a service module located in the cavern that routes trigger signals, packages the signals from a chamber and performs the electrical-to-optical conversion before sending out the signal to the read-out drivers outside the cavern. The custom read-out drivers have the primary tasks of data monitoring and packaging during stable data taking [113].

The CSC read-out has a different design to accommodate the large radiation and rates in the region near the beam. The functionality is, however, similar, with a pre-amplifier and signal shaper at a first stage (at the end of the cathode strip), and the analog-to-digital conversion, as well as electrical-to-optical conversion, at a later stage (mounted on the lateral frame of the chamber). The read-out drivers are also located outside the cavern and also custom-built in order to perform the appropriate signal calibration and packaging [114].

4.4.3 The Trigger Chambers

The trigger chambers need to provide a fast approximate measurement of the muon track position and p_{T}. Their coverage goes up to $|\eta| = 2.4$. The increasing momentum with rapidity at fixed p_{T}, and the differences between the magnetic field in the barrel and end-cap regions call for a different spatial resolution to maintain a p_{T} resolution that is roughly independent of η. For this reason, two different technologies have been chosen to cover each region. Resistive-plate chambers (RPCs) provide muon triggering up to $|\eta| = 1.05$, while thin-gap chambers (TGCs) cover the higher $|\eta|$ regions.

Three layers of RPCs exist, all placed on the same support structure as some of the barrel MDTs. The first two layers sandwich the middle layer of the MDTs, while the last layer is on the outer (for large chambers) or inner (for small chambers) sides of the last MDT layer. Four TGC layers exist, and unlike the RPCs, they have their own support structure. Each layer is placed in two concentric disks within a wheel. One layer is placed on the inner side of the first end-cap wheel, while the other three are placed in the proximity of the third wheel, behind the end-cap toroid magnet cryostat. One of those three is placed in front of the MDTs of the third wheel, while the other two are behind the MDTs, a few centimeters apart. Each RPC layer and most TGC layers have two (up to three for the TGCs) detector layers to provide redundancy that reduces the noise and other backgrounds, such as photons and low-energy neutrons in the cavern, and thus the trigger rates. The layout of the trigger chambers is illustrated in Figure 4.17.

The operating principle of the RPCs is that of a gaseous detector in which the field is established between two parallel plates, where the read-out electrodes are placed. The plates are kept at a distance of 2 mm and a gas mixture of primarily $C_2H_2F_4$ is used. A voltage difference of 9.8 kV is established between the two plates to reduce the drift time, which is ≈ 5 ns. The read-out strips have a width of 25–35 mm and are placed in the η and ϕ directions on the outside of each of the two resistive plates in each detector.

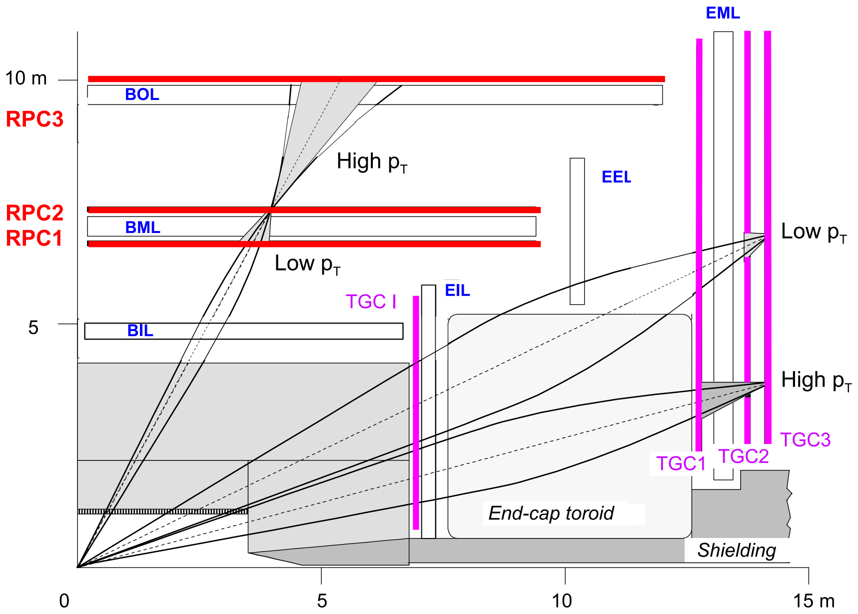

Figure 4.17: *Illustration of the layout of the trigger chambers within the muon chambers [88]. The last RPC layer is shown on the outside of the last MDT layer, corresponding to the layout for ϕ sectors with large MDTs.*

The operating principle of the TGCs is similar to that of the CSCs; that is, they are multi-wire proportional chambers. The bending coordinate is determined using the anode wire signal, while the ϕ position is determined using the cathode-strip signal. The distance between the wires is 1.8 mm, while the distance from the wire to the cathode is 1.4 mm. The read-out of the wires happens in wire groups, and the number of wires per group varies to maintain a uniform momentum resolution with η. The gas mixture contains CO_2 and n-C_5H_{12} in similar proportions, and the anode wires are kept at a nominal potential of 2.9 kV. The drift time achieved combined with the signal propagation time in the electrodes guarantee that signals arrive to the read-out chain within 25 ns.

The RPCs are read out using a pre-amplifier, a signal shaper and a comparator, which defines the threshold needed for a hit, located on the side of the chamber and soldered directly to the strips. The signal from a RPC needs to be combined with signals from two more RPCs within a sector. These comparisons are made by a system of coincidence units near the chambers. These units also perform the electrical-to-optical conversion, before sending a signal to the counting rooms for use by the trigger system. A similar system exists for the TGCs, even though some additional logic is implemented in the counting rooms, outside the cavern. The trigger read-out is performed through different wires and servers from those used to read out the detector if the event is accepted by the trigger. This is further discussed in Section 4.6.

4.5 Luminosity Detectors

The ATLAS detector is complemented by other detectors, some of them outside the ATLAS detector volume presented earlier in this chapter. With these detectors, additional measurements for physics or monitoring purposes are performed. The most important for the studies in this thesis are the luminosity detectors. The detectors used for the luminosity determination in this thesis were the minimum bias trigger scintillators [115] (MBTS), the beam condition monitor [116] (BCM), and the luminosity Cherenkov integrating detector [117] (LUCID.) The measurements from these detectors are complemented by measurements in the ID to provide the final luminosity estimate. All these detectors perform a relative luminosity measurement, and thus had to be combined with a van der Meer scan for their absolute calibration, as will be discussed in Chapter 6.

The MBTS has been designed to detect and trigger on collision activity during a bunch crossing in the early operation of ATLAS. In this thesis, it is not only used in the luminosity determination, but also as a reference trigger to calculate the efficiency of the jet triggers. The detector consists of 32 2 cm thick scintillator paddles assembled on two disks, one on each end-cap. The disks are placed on the face of the end-cap calorimeter cryostats nearest to the interaction point. The paddles, of a trapezoidal shape, are organized in two concentric rings within a disk. This is illustrated in Figure 4.18. This figure also shows the wavelength-shifting fibers that collect the light produced in each scintillator and transport it to the photomultiplier tubes for signal amplification. The signals out of the photomultiplier tubes are processed using the tile calorimeter read-out electronics. They are shaped and amplified, and a threshold of 50 mV is used to define a hit. The signals are then sent to the central trigger processor, and if the event is recorded they become available for analysis.

The BCM has the main function of monitoring the beam for beam losses that could cause detector damage, particularly in the pixel detector. If large beam losses are detected, it sends a signal to the LHC, which causes an abort and a controlled emptying of the

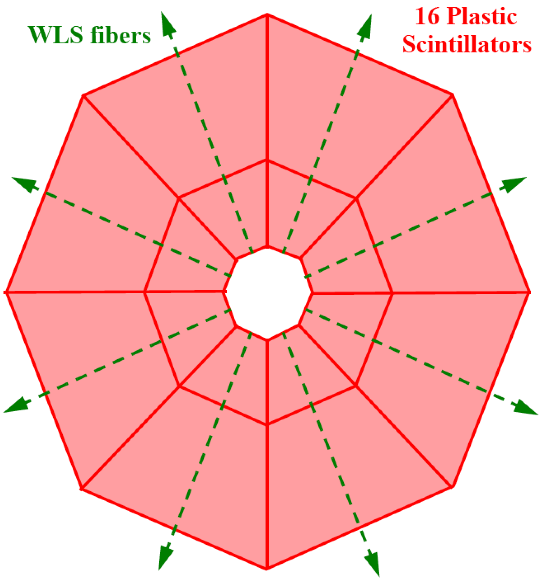

Figure 4.18: *Illustration of the layout of the scintillator paddles on an MBTS disk [118]. The read-out fibers transporting light to the photomultiplier tubes are also shown.*

accelerator ring. This signal is also sent to the ATLAS detector, which can take action (e.g.: turn off the high voltage on the pixel detector) to minimize the damage produced by the beam loss. Its readings are also available for trigger and reconstruction algorithms, and not all affected by dead-time in the data acquisition read-out.

The BCM consists of two sets of diamond sensors located 184 cm away from the interaction point in the direction of the beam and 5.5 cm away in the radial distance on the transverse plane. Each side has four modules with two sensors each. The sensors are 1 cm×1 cm in size and 500 μm thick, and operate at a voltage of 1000 V. Diamond is chosen as the sensor material due to its radiation hardness, necessary for a detector placed at such high η, and the high mobility of electrons in the diamond, which provides a fast signal, as required for prompt action in case of beam losses. The use of diamond also allows for operating the detector at room temperature, without the need for cooling services.

LUCID is dedicated to the luminosity measurement. It consists of 32 aluminum tubes located around the beam at a distance of 17 m from the interaction point on each side of the detector. The tubes are 1.5 m long and 15 mm in diameter, and point towards the interaction point to minimize the signals created by particles coming from other interactions in the cavern. The tubes are filled with C_4F_{10} gas where Cherenkov light is created. The Cherenkov photons are reflected in the tube walls, until they reach the photomultiplier tubes located at the end of the tubes furthest from the interaction point. The signal from the photomultipliers is amplified and a discriminator is used to determine if a tube was hit.

The hit information is sent to a programmable front-end card, which combines the hits from different tubes using different algorithms, and associates the hits to a specific bunch crossing. This data is made available to the trigger processor in ATLAS, to the data-acquisition system for event building and to any processes, performing monitoring or online analysis, that are running in the environment of the ATLAS control room.

4.6 The Trigger and Data Acquisition

The actions taken throughout the data acquisition chain are driven by the trigger decisions. The ATLAS trigger consists of three trigger levels. The level-1 trigger [119] is a hardware-level trigger and is designed to accept as much as 75000 events per second. The decisions made at the level-1 trigger must reach the detector front-end electronics within 2.5 μs of the event production (including 1 μs needed for cable propagation), to guarantee that the event is still stored in the front-end buffers and, thus, can be sent further along the read-out chain. When the front-end electronics receive a level-1 trigger signal, the data are sent to the read-out drivers and, after packaging and processing, to the read-out servers, where they await the decision from the high-level trigger [120, 121] (level-2 and event filter.)

The level-1 trigger uses information of reduced granularity from the calorimeters and the trigger chambers. This allows for selections based on the presence of muons, electrons, photons, τ leptons, jets, event transverse energy and missing transverse energy. The MBTS is also used at the level-1 to select events in which a collision happened. The level-1 trigger uses no information from the ID. The level-1 trigger information is processed in the electronics on or near the detectors, and it is sent to the central trigger processor [122] (CTP.) The level-1 algorithms used for reconstructing jets and selecting the events in this thesis are discussed in the Chapter 6.

The CTP contains the level-1 trigger menu. The level-1 trigger menu includes the thresholds of the trigger objects that should be accepted, certain simple level-1 object combinations that should be used for the trigger decisions and pre-scales for each trigger. The pre-scales determine how often an event that has passed a specific trigger should be accepted. A pre-scale of 1000 implies that events passing the trigger will only be accepted one out of every thousand times. Pre-scales are needed to control the rate at which commonly occurring events are recorded. If an event passes the level-1 trigger, the CTP distributes the signal to the full detector through an optical network and the on-detector trigger processors forward the position information from the level-1 trigger algorithms, which is not used by the CTP, to the level-2 trigger processing units.

The level-2 trigger is designed to reduce the trigger rate to 3.5 kHz. The level-2 trigger processing units use the position information on the level-1 "regions of interest" (RoIs) to determine from which read-out servers they should retrieve the data. The level-2 algorithms have, thus, a partial picture of the event, and cannot enhance the level-1 reconstruction of global variables, such as the missing transverse energy. Once the level-2 has analyzed the RoIs with its more complex algorithms, a decision is made and sent to data-flow managers that distribute the data from the read-out servers to the event builders. In the event builders, all detector data from an accepted event is put together in one data structure. An event, is then sent to the event filter for further analysis.

The event filter analyzes events using more complex algorithms to reduce the output rate to 200-300 Hz, which is the rate at which ATLAS events can be written to disk. The event filter also assigns events according to the triggers passed to one or several of a pre-determined set of event streams.

The event streams are designed to aid in the data analysis by separating events of interest for different analyses in different datasets. Since one event is written to as many event streams as it belongs to according to the stream definition, the streams are defined to minimize their overlap. Events that take too long to process in the event filter are rare, and thus kept for analysis in the "debug" stream. While the event filter was not operating to select events with jets in the data used for this thesis, other objects (muon, photons,...) were selected using the event filter algorithms. This caused some events passing the selection criteria to fall in the debug stream, and the events had to be included in this thesis' analysis as described in Chapter 6. An express stream also exists, where a fraction of the data is written for fast processing. These data are used for calibration purposes and establish the values of certain variables that are needed for the bulk data processing. Specific examples of the usage of the express stream are given in Chapter 5.

The trigger system is flexible to real-time configuration changes that may be necessary due to changing beam conditions or hardware problems. This flexibility is provided by a data-taking unit of time referred to as a "luminosity block". Luminosity blocks are sub-units of a "run", which refers to a period between the start and the end of the data acquisition. Optimally, runs start at the beginning of an LHC fill and end when the beam is dumped, even though problems in the data acquisition sometimes require having several runs in a fill. Luminosity blocks are typically two minutes long and certain settings of the trigger, such as the pre-scales, can be changed at their start. The changes are recorded in the event data structure and can be later used for analysis tasks needing time-dependent run information. Some of those tasks, such as the calculation of the luminosity, the determination of the jet origin, or the exclusion of periods with operational problems in the detector, are relevant to the analysis presented in this thesis and will be discussed in Chapters 5 and 6, when pertinent.

4.7 The ATLAS Detector Simulation and Conditions

The ATLAS collaboration has developed a detailed simulation of the detector based on GEANT4 [123] with an accurate description of the detector material and geometry [124]. The tunable parameters in GEANT4 have been fixed according to the results from "test-beam" analyses [125–127]. Test-beam analyses used data collected from individual detector modules or fractions of the detector radiated with beams of particles. The nature and energy of these particle beams was chosen to be representative of the nature and energy of particles created in proton-proton collisions at the LHC. The accuracy of the detector geometry and the accuracy in the description of the basic properties of detector signals in the simulation have also been evaluated using 2009 and 2010 data [18, 19, 128–132].

The simulation not only reproduces the detector geometry and the interactions of particles with the detector, but also the functions of the hardware that reads out the signal. For example, the decisions of the level-1 trigger algorithms for a given event, or the impact of varying the distance between bunches in the EM calorimeter energy readings can be reproduced. Certain detector failures and beam conditions can also be introduced in the simulation, such as the failures in the electrical-to-optical signal conversion discussed in

Section 4.3.1 or the intensity of the beam.

Since some of these conditions vary from run to run, the simulation conditions are fixed to be representative of the runs to which a given simulated dataset can be compared. Furthermore, certain experimental procedures can be used to reduce the differences between the simulation and the data conditions, and checks can be performed to guarantee that any remaining differences between data and simulation have a negligible impact on a specific analysis (see Chapter 7).

The preparation of the simulated events started at the same time as the collection of the data used in this thesis was completed. For this reason, the detector conditions in the simulation match the detector conditions for the last few runs. This and the exponential increase in the beam intensity during the first few months of the LHC operation guarantee that the majority of the data used in this thesis have the same conditions as the simulated events. The good match in the detector conditions between collected and simulated data is illustrated in Table 4.2, where the number of faulty optical connections in the EM calorimeter and the number of tile calorimeter cells with read-out problems are listed for the last run and the simulated events. Only problematic channels in the calorimeters are shown since the calorimeters are the most important detectors for the measurements presented in later chapters. The impact of these problematic channels on the energy measurement and in the analysis is discussed in Chapters 5 and 7. For the purposes of this section, it suffices to point out that the near perfect agreement between the simulation and the data makes it

Subdetector	Data (Run 162882)	Simulation	Surface Coverage
Tile Layer 1	73	73	3.9%
Tile Layer 2	62	62	3.4%
Tile Layer 3	25	25	2.9%
Presampler	1	1	1.3%
EM Calorimeter Layer 1	12	13	1.7%
EM Calorimeter Layer 2	10	10	1.7%
EM Calorimeter Layer 3	4	4	1.5%

Table 4.2: *Number of cells with read-out problems (tile calorimeter) and number of problematic electrical-to-optical connections (EM calorimeter) in different read-out layers of the tile and EM calorimeters in the last run and the simulated events used in this thesis. The approximate detector coverage affected by these problems is also listed.*

easier for analyses to account for the effects caused by these failures.

The position of these regions in the calorimeters as represented in the detector simulation is shown in Figure 4.19 for illustrative purposes. A few additional regions in the tile calorimeter are shown, corresponding to faulty read-out connections for the scintillators lying between the barrel and extended barrel modules. The region in the first layer of the EM calorimeter shown at around $(\eta, \phi) = (0, 0)$ was not problematic in the data used in this thesis, and corresponds to the only discrepancy in the list of faulty channels between the detector simulation and the real detector.

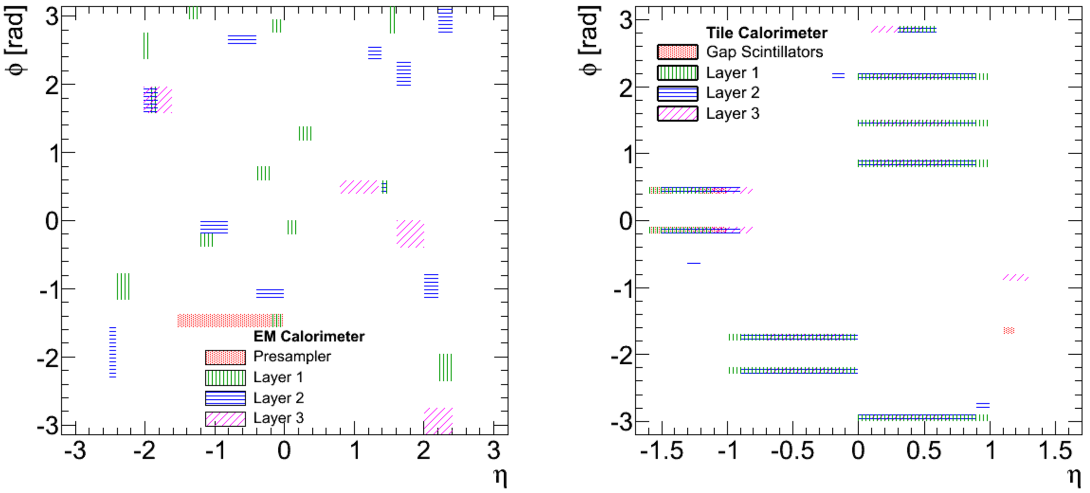

Figure 4.19: *Regions of the EM (left) and tile (right) calorimeters with read-out problems as represented in the detector simulation.*

Chapter 5

Event Reconstruction

The information recorded by the detector needs to be processed to extract the kinematics of particles and, more generally, build up experimental objects that can be used in an analysis. This process is typically referred to as "offline event reconstruction". The adjective "offline" refers to the fact that this process does not need to be performed in real time, as the detector collects data. Three experimental objects are of importance for this thesis: tracks, vertices and jets. In addition, the reconstruction of the luminous region ("beam-spot") where collisions can happen is also important to understand track and vertex reconstruction. Given the complexity and variety of reconstruction algorithms existing for other objects, namely electrons, photons, muons, τ leptons, b-quark jets and missing transverse energy, the reconstruction of these objects is not described here, and detailed information can be found elsewhere [8, 88].

5.1 Track Reconstruction

Tracks represent the trajectories of charged particles, and are reconstructed using hit information from the ID. Tracks can be defined by their charge, momentum and position at any given point along the track's trajectory. Their position and momentum at any other point can then be calculated given the magnetic field and material that they traverse. Track coordinates in this thesis are expressed at the distance of closest approach to the center of the beam-spot, unless it is explicitly mentioned that they are expressed with respect to the event vertex, as defined in the next section. The two coordinates of primary interest for the studies in later chapters are d_0 and z_0, that is, the radial and longitudinal distances at the point of closest approach. The track reconstruction algorithms also provide with an estimate of the uncertainty in the reconstructed parameters, $\sigma(d_0)$ and $\sigma(z_0)$.

Tracks are reconstructed in several steps [133]. First, nearby raw hits from the pixel detector and the SCT are used to build clusters that are in turn transformed into three-dimensional space-points. Second, several tracking algorithms are run. Finally, the vertex reconstruction algorithm is run and tracks are associated to "primary" (compatible with the beam-spot) or "secondary" vertices.

The majority of tracks are reconstructed by the "inside-out" algorithm. This algorithm uses three space points from the semiconductor detectors to define a "road". Hits along that road are added to the space-point candidates that are then fit using a Kalman fitting algorithm [134]. At this stage, hits can belong to several track candidates. A score is then assigned to each track candidate according to the quality of the fit, and poor candidates are rejected until no shared space-points exist. The new set of space points is refit to improve the quality of the fit, and extended to the TRT to find straw hits compatible with the track trajectory. Once all such hits are found, the final estimate of the track parameters is obtained performing an additional fit to all hits associated to the track.

An additional "outside-in" algorithm is used to recover tracks that may not have been reconstructed using semiconductor hits as seeds due to detector inefficiencies, or simply because the tracks originated from particles that decayed well into the semiconductor detectors. The track in this case is seeded by TRT segments and hits in the pixel detector and SCT that were not used by tracks reconstructed using the inside-out algorithm are checked for compatibility with the seed. Compatible hits are then used to perform a fit similarly to how it was done for the inside-out algorithm.

Both these algorithms have a threshold of 500 MeV for track reconstruction. In order to reconstruct tracks all the way down to 100 MeV the inside-out reconstruction is re-run on the remaining hits. The presence of less hits in this second pass reduces the ambiguity in the track reconstruction and thus the fakes. In this thesis, we are most interested with tracks in jets. The efficiency for tracks in jets is affected by the higher hit density and this is most obvious in the first layer of the pixel detector. This inefficiency has been studied in data looking at the efficiency for a track in a jet to have a hit in this first layer, and has been found to be in good agreement with the prediction from the Monte Carlo detector simulation [135].

5.2 Vertex Reconstruction

Tracks compatible with originating in the beam-spot, satisfying quality criteria [136] and having a transverse momentum above 150 MeV, are used for finding the event primary vertex or vertices. The track selection criteria in the primary vertex reconstruction algorithm are:

1. $\sigma(d_0) < 5$ mm,
2. $\sigma(z_0) < 10$ mm,
3. at least four hits in the SCT,
4. at least six hits in the pixel and SCT detectors.

A vertex is seeded by looking for the maximum in the distribution of the number of tracks along the z axis. The vertex is fit using this seed along with neighboring tracks. The fit quality of those tracks is properly accounted for in this vertex fit. Tracks incompatible with the reconstructed vertex are used to seed new vertices until no tracks are left. The incompatibility is determined using the errors, σ, on the position estimated by the track reconstruction algorithm, by imposing a 7σ cut between the track and the vertex positions. Vertices are required to have at least two tracks associated to them. The vertex in the event for which $\sum p_T^{\text{tracks}}$ is highest is considered to be the hard-scatter or “event” vertex from which jets are considered to originate. Figure 5.1 shows the vertex reconstruction resolution in the z direction as a function of the number of tracks associated to the vertex. This resolution was determined splitting reconstructed vertices in two and studying the distance between the split vertices using data. This figure shows that vertices with at least five tracks, of concern to the analysis presented in this thesis, can be reconstructed with a resolution better than 1 mm in the z direction.

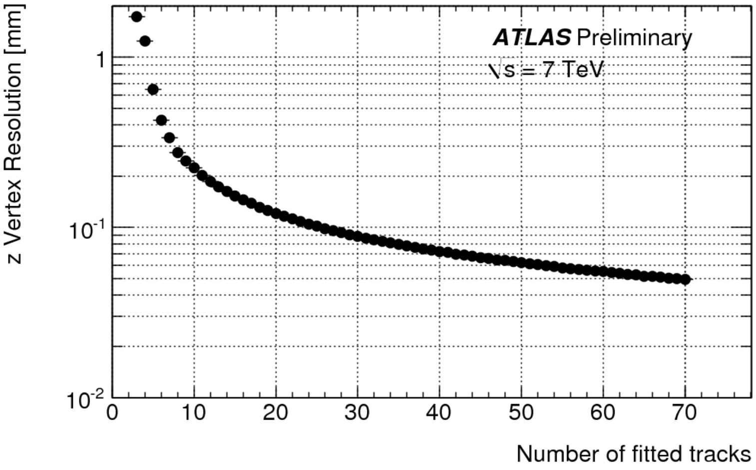

Figure 5.1: *Resolution for the reconstruction of the z position of the vertex as a function of the number of tracks in the vertex [136].*

5.3 Reconstruction of the Beam-spot

The luminous region or beam-spot refers to the place where the beams can interact. The beam-spot can vary from luminosity block to luminosity block, and need not be exactly at the nominal collision point, at (0,0,0) in the ATLAS coordinate system. This is illustrated in Figure 5.2, where its position along the y axis is shown for several LHC fills in 2009.

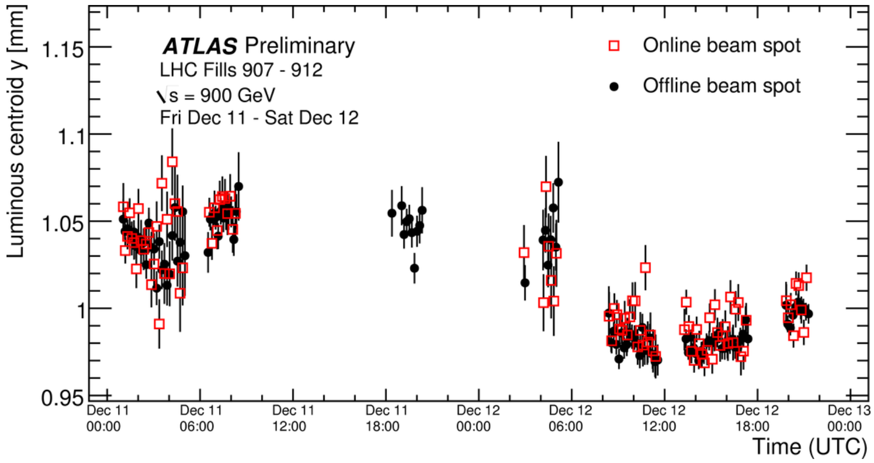

Figure 5.2: *Position of the center of the beam-spot along the y axis as a function of time as measured with the offline and online beam-spot algorithms during five LHC fills in 2009 [137].*

The beam-spot is reconstructed both by level-2 trigger algorithms and by offline algorithms, even though only the offline algorithms are used to constrain the vertex reconstruction. The beam-spot is determined for every luminosity block, using the express stream, so that its position can be used in the bulk data reconstruction. A mixture of events, coming from the MBTS triggers as well as high-p_{T} lepton and jet triggers, is used. Tracks reconstructed as detailed above, are used to build vertices without the beam-spot constraint. All reconstructed vertices in the data in this manner are input to a maximum likelihood fit that extracts the mean position of the beam-spot, its average size and its tilt. The vertex position resolution, calculated from data as described in the previous section, is also input to this maximum likelihood fit, since it determines the relationship between the measured and the truth beam-spot. Further details of this study can be found in Ref. [137].

The size of the luminous region as determined for a $\sqrt{s} = 7$ TeV run in 2010 is shown in Figure 5.3, and it is large enough compared to the vertex reconstruction resolution, see Figure 5.1, that vertex merging has a negligible impact in the analysis presented in this thesis. Only vertices with at least 10 tracks were used in this figure, unlike for the vertex reconstruction. This figure also shows a slight tilt in the luminous region along the z direction. This tilt is created by the LHC beam-steering to minimize beam-to-beam interactions near the interaction point that could affect the stability of the beam.

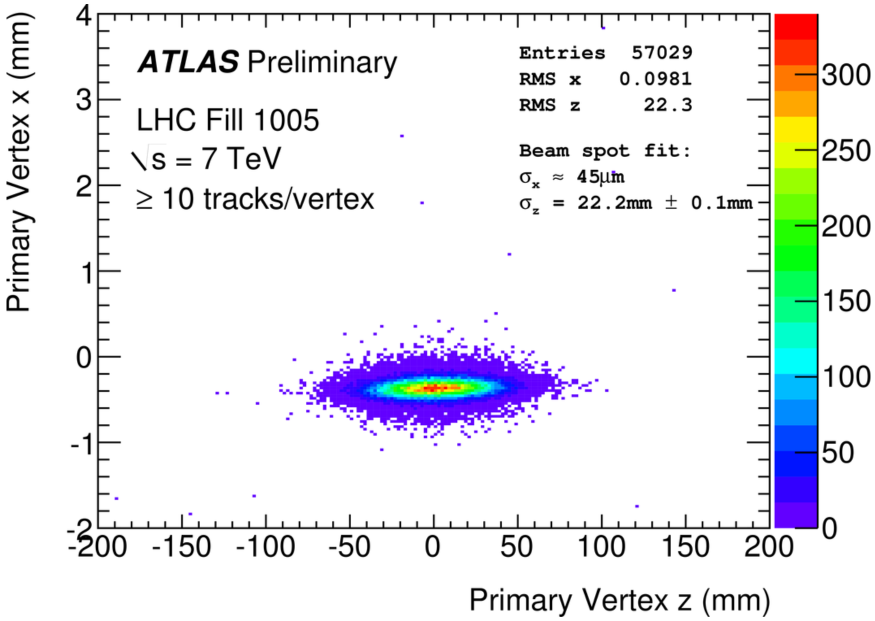

Figure 5.3: *Distribution of the number of vertices with at least 10 tracks as a function of x and z in a dataset used to determine the position and size of the beam-spot [137].*

5.4 Jet Reconstruction

Many jet reconstruction algorithms are available in the ATLAS offline software and can be used for different purposes. In this thesis, only the anti-k_t algorithm [138] is used. The anti-k_t algorithm belongs to a class of “sequential recombination” algorithms [69, 139–142] that build jets combining the inputs to the jet reconstruction into larger objects according to some pre-defined condition, until that condition is no longer satisfied by any of the remaining inputs.

Sequential recombination algorithms build jets that have, in principle, no pre-determined shape, their boundaries adapting to include soft and collinear radiation. The algorithms are, thus, “infrared-” and “collinear-safe”, that is, the addition of collinear or infrared radiation does not change the resulting jets reconstructed by the algorithm. This property is necessary to run a jet algorithm on matrix element calculations so that those calculations do not diverge. For the next-to-leading-order comparisons performed in this thesis, the soft and collinear regions are not filled by the parton shower, and thus are subject to divergent terms. These terms only disappear when using infrared- and collinear-safe algorithms that properly combine the terms coming from the virtual and the real corrections to the leading-order calculation.

A flexible boundary is, however, not a desirable experimental feature for a jet algorithm for several reasons. One of the most obvious reasons has to do with the changing geometry across the calorimeter, which results in different response to particles in the jet depending on where the particle hits the calorimeter. A jet with a flexible boundary would thus require a complex calibration that depends on its position along the calorimeter, its size and its shape, presumably parameterized as a function of several variables.

The anti-k_t algorithm maintains many of the features of other sequential recombination algorithms, like infrared and collinear safety, while providing a resilient boundary that is approximately circular in $[\eta, \phi]$ and presenting other desirable features for experiments [17]. The anti-k_t algorithm starts by assigning a distance to all input objects and all pairwise combinations of those objects:

$$d_i = ({p_{\mathrm{T}}}^i)^{-2}, \tag{5.4.1}$$

$$d_{ij} = \min\left[({p_{\mathrm{T}}}^i)^{-2}, ({p_{\mathrm{T}}}^j)^{-2}\right] \frac{\Delta_{ij}^2}{R^2}. \tag{5.4.2}$$

In this equation $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, with y, ϕ and p_{T} being the rapidity, azimuth and transverse momentum of the input. R is the resolution parameter of the algorithm and controls the size of the jet. It corresponds roughly to the radius of the jet in $[\eta, \phi]$. The sequential combination proceeds by selecting the smallest of all these distances. If the smallest is one of the d_i values, that input is called a jet and removed from the list of inputs. If the smallest is one of the d_{ij} values, the two inputs, i and j, are combined into one, adding their two four-vectors. The distances are then recalculated for the new list of input objects and the recombination procedure continues. At the end of the process a list of jets is produced, each jet with its list of inputs, none of which is shared between two jets.

This algorithm allows for two jets to be within a distance of each other of less than $2R$. For such jets typically one of the jets is roughly circular in $[\eta, \phi]$, while the other one will have the shape of a crescent moon, even though the exact shapes depend on the energy density patterns within the jets [138].

In the studies in later chapters, two resolution parameters are used: $R = 0.4$ and $R = 0.6$. Studies with $R = 0.4$ minimize the impact of the underlying event, since the area of the jet is smaller, and have a wider applicability. Most analyses searching for physics beyond the Standard Model in ATLAS use the $R = 0.4$ parameter. On the other hand, next-to-leading-order calculations are more precise if $R = 0.6$ is used.

5.4.1 Inputs to Jet Reconstruction

Jet algorithms can be run on any set of four-vectors. In this thesis, for the definition of a "truth jet" in the leading-order calculation, they are run on all stable particles with a lifetime longer than 10 ps. For the study of the performance of the jet energy reconstruction, a truth jet is instead defined by running the jet algorithm on all stable particles except muons and neutrinos, since they do not leave any signal in the calorimeter (except ≈ 3 GeV for muons [3]). For the next-to-leading-order calculation, the algorithm is run on partons, whether they come from the matrix element calculation or from the parton shower, as detailed in Chapter 9.

At the detector level, jet reconstruction is performed using "topological calorimeter clusters" as inputs to the jet algorithm. Topological clusters are objects that are built in an attempt to capture the shower of a particle in the calorimeter, while suppressing calorimeter noise. The construction of a topological cluster can be summarized in the following steps:

1. A cell with $|E_{\text{cell}}| > 4\sigma_{\text{noise}}$ seeds the cluster.

2. Any cells with $|E_{\text{cell}}| > 2\sigma_{\text{noise}}$ contiguous to the cluster are added to the cluster until no such cells remain.

3. An additional layer of all contiguous cells (independent of their energy) is added to finalize the construction of the cluster.

The expected level of noise is calculated for each cell using data triggered randomly in the absence of a beam. The clustering accepts cells with negative energies, whose energy readings arise from electronic noise, to prevent a bias in the reconstructed energy that would arise from an asymmetric cut. Clusters are assigned a mass of 0. Their energy is calculated as the sum of the energies of its constituent cells, and their rapidity and azimuth is calculated as the energy-weighted average of the rapidity and azimuth of their constituent cells in the ATLAS detector coordinate system.

Clusters used by the jet reconstruction have an energy that corresponds to the "electromagnetic (EM) scale". This means that the calorimeter signals are calibrated to properly reproduce the energy lost in the calorimeter by an electron, if the energy deposit came from an electron. This calibration was originally verified in the ATLAS test beams [125–127,143], has been recently refined for the EM calorimeter using $Z \to ee$ data and is periodically verified using cosmic ray muons and radioactive sources for the other calorimeter systems [109,144].

Objects with a rigid structure, referred to as noise-suppressed towers, can also be used as inputs to the jet reconstruction at the detector level. Noise-suppressed towers are built using cells that belong to topological clusters, effectively suppressing calorimeter noise. However, towers have specific positions within the calorimeter, and their position does not change depending on the energy deposition patterns in the calorimeter. Towers define a 0.1×0.1 grid in $[\eta, \phi]$. When cells larger than 0.1 exist, like in the third layer of the tile calorimeter, they are split between towers, and so is their energy, in a proportional manner. The difference between towers and clusters, and how towers are built from clusters is illustrated in Figure 5.4. Jets built of towers are not used in this thesis' analysis, even though they are used to derive one of the components of the jet energy calibration for jets built of clusters as described later in this chapter.

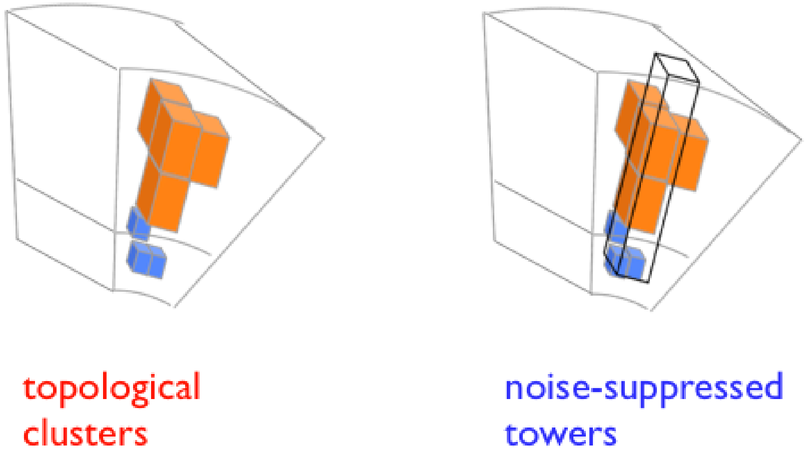

Figure 5.4: *Illustration of how cells may build a topological cluster (left) and how noise suppressed towers are built from those cells (right).*

Table 5.1 summarizes all the inputs to the jet reconstruction that are used directly or indirectly in this thesis. All these usages are described in more detail in the next chapters.

Type of input	Usage	Described in section
Topological Clusters	Detector-level reconstruction	5.4.1
Towers	Jet energy calibration	5.4.3
All stable particles	LO calculations	5.4.1
All stable particles (no μ, ν)	Jet energy calibration	5.4.3
ME partons	NLO calculations	9.2
PS partons	NLO calculations	9.2.2

Table 5.1: *Types of inputs to jet reconstruction used throughout this thesis. The usage given to each type of jet and the chapter where that usage is described in detail are also listed.*

5.4.2 Corrections for Cells with Problematic Read-out

As discussed in Section 4.7, certain areas of the calorimeter were not operational during the period when the data analyzed in this thesis was collected. The small surface covered by those regions and the overall good agreement between the detector simulation and the data make their impact in the analysis small, even if no attempt is made at recovering their energy measurements. Nevertheless, such an attempt is made.

The readings lost due to faulty optical connections in the EM calorimeter are recovered with coarser position granularity and larger noise using the level-1 trigger readings. As described in Section 4.3.1, the read-out of the signal used for the level-1 trigger is independent from that used in the nominal data acquisition chain. The read-out chain for the two purposes splits before the failing optical connections, such that the level-1 trigger readings are available. The level-1 energy readings serve as a substitute for the nominal detector readings when the latter are not present, providing an approximate correction for this effect.

The read-out problems in the tile calorimeter are not so easily recoverable. An attempt is made using the cell energy density information of cells around the problematic regions. The average cell energy density is used as an estimate of the energy density in the problematic region, and the volume of the region allows the energy density estimate to be transformed into an energy estimate. This correction works well on average, even though varying calorimeter response to jets has been observed depending on how close the jet is to the problematic region. Since most of the energy of jets is deposited in the EM calorimeter up to very high jet energies (~ 1 TeV), this effect is not studied in any more detail.

5.4.3 Jet Energy Scale Calibration

The kinematics of jets reconstructed from EM calorimeter clusters need to be corrected for a number of effects in order to reproduce the kinematics of the truth jets they originate from. A number of complex calibration schemes have been developed in ATLAS to take into account correlations between the different energy depositions along the longitudinal shower development and on the transverse extension of the jet [8, 14, 19]. However, these algorithms have only very recently become fully validated with data [16] and a simpler calibration scheme is used in this analysis.

Figure 5.5 shows the algorithmic flow in this calibration scheme. Jets built from EM clusters are first corrected by the offset correction, to remove the dependence of the reconstructed jet energy on the instantaneous luminosity of the bunch crossing (remove the "in-time pile-up"). The resulting jet energy is then corrected using a p_T- and η-dependent correction to the energy that the corresponding truth jet would have. In the simulation, the truth jet is built with the same jet algorithm as in the reconstruction, using all stable particles with a lifetime longer than 10 ps in the center-of-mass frame of the collision except muons and neutrinos. A final correction is applied to correct the position measurement in

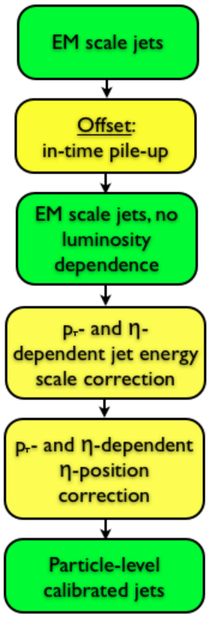

Figure 5.5: *Jet energy calibration sequence, going from reconstructed jets at the electromagnetic scale to calibrated jets whose energy is independent of the instantaneous luminosity in a bunch crossing.*

the regions of transition between calorimeters. The position measurement at the EM scale in those regions is inaccurate due to inhomogeneities in the calorimeter geometry.

5.4.3.1 Offset Correction

In the presence of several proton-proton collisions within one bunch crossing, the energy measured in the topological clusters is not all deposited by particles belonging to the jet. Furthermore, even if only one proton-proton interaction happened in a given bunch crossing, proton-proton interactions in a previous bunch crossing can also affect the energy measurement if the previous bunch crossing happened within the read-out window of the calorimeters. These two effects are typically referred to as “in-time” and “out-of-time pile-up”, and create a dependence of the measurement on the instantaneous luminosity of a bunch crossing and the LHC bunch structure, respectively. Only in-time pile-up concerns this thesis due to the long spacing between bunches in the early data.

The energy contribution from in-time pile-up is estimated using events from a trigger stream containing triggers for all calorimeter objects [145]. While jets, electrons and other objects are thus, by definition, present in these events, most of the calorimeter is empty, and the events can be used as if those objects were not present. Towers are used to calculate the average energy deposited in the calorimeter by in-time pile-up for simplicity. The subtraction of pile-up from clusters is possible, but more complex, since the cluster shape and extension needs to be accounted for. The number of reconstructed primary vertices with at least five tracks is used as a measure of the instantaneous luminosity. This is possible due to the high position resolution of the vertex reconstruction along the z axis combined with a large luminous region, as shown in Sections 5.2 and 5.3. Figure 5.6 (left) shows the in-time pile-up transverse energy contribution per tower as a function of η and the number of reconstructed primary vertices, N_{PV}. The energy is reconstructed at the EM scale. For

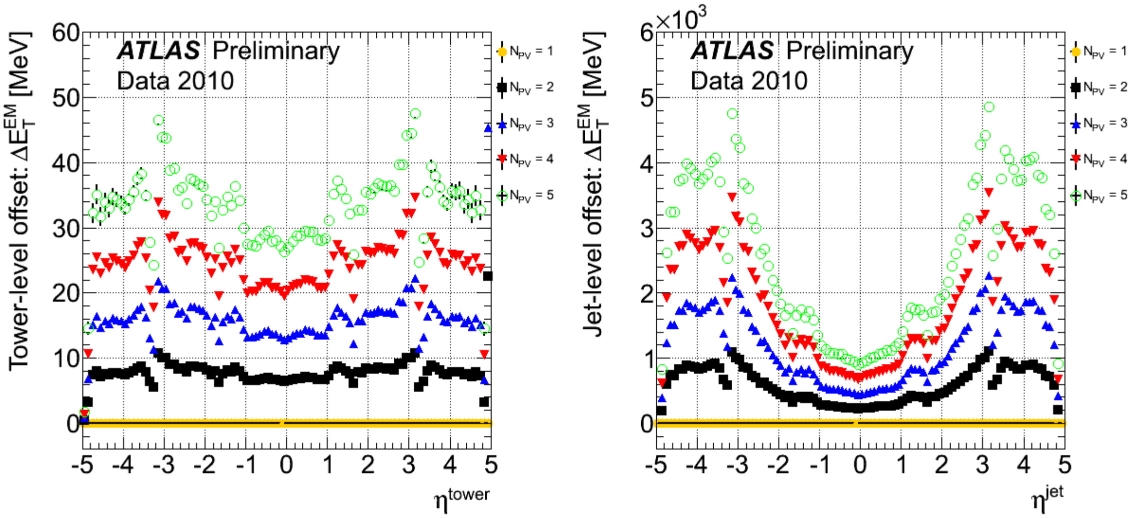

Figure 5.6: *Average transverse energy deposited in towers (left) and anti-k_t jets built of towers (right) with $R = 0.6$ by pile-up interactions for events with different N_{PV} [145].*

$N_{\mathrm{PV}} = 1$ the pile-up energy is by definition zero. The discontinuities in the shape of this curve correspond to transitions between calorimeters which cause a change in the noise thresholds and thus the sensitivity to reconstruct pile-up energy after noise suppression.

The energy deposited through in-time pile-up is subtracted from the EM scale energy of the jet and the dependence of the jet energy on N_{PV} is thus removed

$$E_{\mathrm{EM,corr}}^{\mathrm{jet}} = E_{\mathrm{EM}}^{\mathrm{jet}} - E_{\mathrm{offset}}^{\mathrm{jet}}(N_{\mathrm{PV}}). \tag{5.4.3}$$

However, only the energy of jets built of towers can be corrected precisely and in a straight forward way, since $E_{\text{offset}}^{\text{jet}} = N_{\text{towers}} \times E_{\text{offset}}^{\text{tower}}$, where N_{towers} is the number of towers in the jet. The average energy subtracted from a jet using this procedure is shown as a function of η for jets of ${p_{\text{T}}}^{\text{EM}} > 7$ GeV in Figure 5.6 (right).

The energy of jets built of clusters is subtracted using the offset energy of jets built of towers. While this procedure loses precision due to the varying number of towers in a jet at fixed energy, it does not lose accuracy due to the large similarities between jets built of clusters and jets built of towers [18, 19]. The mean number of towers in a jet of ${p_{\text{T}}}^{\text{EM}} > 7$ GeV is used for this calculation. Since the mean number of towers in a jet decreases with increasing p_{T}, this is an overestimate of the energy that should be subtracted at high p_{T}. However, the impact of the offset correction at high p_{T} is negligible, and the bias incurred by this simplification is accounted for in the systematic uncertainty in the jet energy scale (see Chapter 8).

5.4.3.2 Jet Energy Scale Correction

Jets reconstructed at the electromagnetic scale are measured to have an energy which is lower than the true energy of interacting particles within the jet. This is caused by the different response of the calorimeter to hadrons and electrons and by energy lost in dead material, deposited outside the jet cone or falling outside the cell clustering algorithm. The size of these effects is illustrated in Figure 5.7, where the "jet response" before calibration is shown for jets from the Monte Carlo simulation as a function of $|\eta|$. The Monte Carlo simulation used did not include additional proton-proton interactions, so the influence of pile-up does not need to be considered. The jet response is defined as the mean of a Gaussian fit to the peak of the response distribution $E_{\text{reco}}/E_{\text{truth}}$, where E_{truth} is the energy of the closest jet built of all truth stable interacting particles (except muons and neutrinos), within $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$. This value was chosen because it results in a reconstructed-to-truth jet match more than 99% of the time for jets with ${p_{\text{T}}}^{\text{truth}} > 20$ GeV [11]. Only jets of ${p_{\text{T}}}^{\text{truth}} > 20$ GeV were used in this figure, and that is why the curves at fixed energy are truncated at a specific value of $|\eta|$.

The jet energy response as defined above (or the jet p_{T} response, which will be used later and uses the ${p_{\text{T}}}^{\text{reco}}/{p_{\text{T}}}^{\text{truth}}$ distribution) is different, but related to, the calorimeter response defined in Section 4.3. In that section the definition was the same, but it applied to

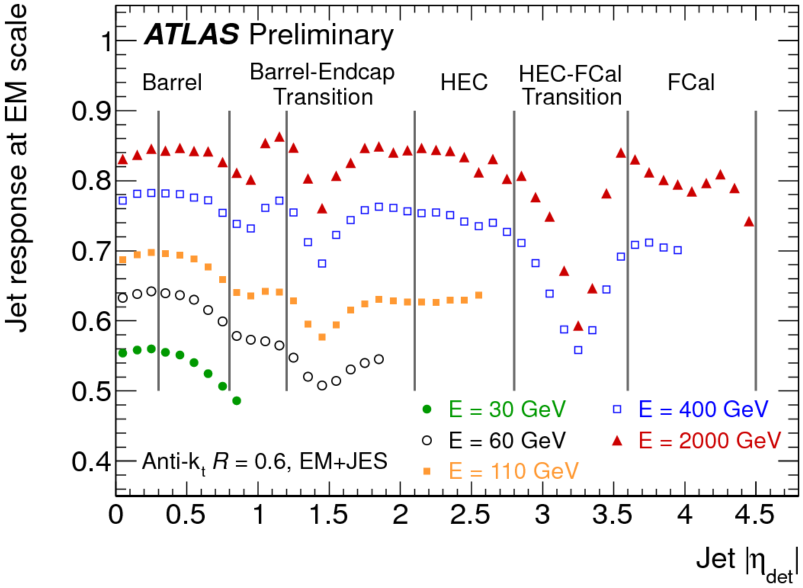

Figure 5.7: *Jet response at the EM scale as a function of $|\eta|$ for different jet energies [146]. The different regions of the calorimeter with increasing $|\eta|$ are illustrated, as well as the boundaries that will be later used for other jet performance plots, which represent regions of the calorimeter with relative homogeneity.*

particles and was only affected by the calorimeter specifications. In extending the definition to jets, many other effects come into play beyond the calorimeter specifications. Some of those effects are related to the ATLAS detector design, such as the bending of particles out of the jet area, or the ID material, where particles deposit some of their energy. Other effects are related to the particle composition of the jet, and affected by the physics of a proton-proton collision (e.g., how many soft particles the jet has). All these new effects are taken into account in understanding the limitations of the jet energy measurement and calibration.

A Monte Carlo–based calibration that corrects the jet energy response as a function of p_{T} and η is used to obtain jets with the correct energy scale. The calibration constants consist in the inverse of the response: $C({p_{\mathrm{T}}}^{\mathrm{EM}}) = R^{-1}_{\mathrm{reco}}({p_{\mathrm{T}}}^{\mathrm{EM}})$, where C is the calibration constant, R_{reco} is the response calculated as a function of reconstructed jet p_{T} and ${p_{\mathrm{T}}}^{\mathrm{EM}}$ has been defined earlier. They are derived as a function of ${p_{\mathrm{T}}}^{\mathrm{truth}}$, to remove the impact of the underlying p_{T} spectrum on the response. The response determined as a function of ${p_{\mathrm{T}}}^{\mathrm{truth}}$, R_{truth}, is used to apply the constants as a function of ${p_{\mathrm{T}}}^{\mathrm{EM}}$, that is $R_{\mathrm{reco}}({p_{\mathrm{T}}}^{\mathrm{EM}}) = R_{\mathrm{truth}}(R_{\mathrm{truth}} \cdot {p_{\mathrm{T}}}^{\mathrm{truth}})$ [11]. This relationship is valid in ATLAS due to the linearity of the jet response as a function of p_T. The calibration constants are derived in 0.1-wide $|\eta|$ bins to capture the changing calorimeter geometry observed in Figure 5.7. This calibration is typically referred to as the "EM+JES" calibration. Figure 5.8 shows the mean value of the constants as a function of calibrated jet p_T for three $|\eta|$ ranges for jets built of clusters with the anti-k_t algorithm with resolution parameter $R = 0.6$.

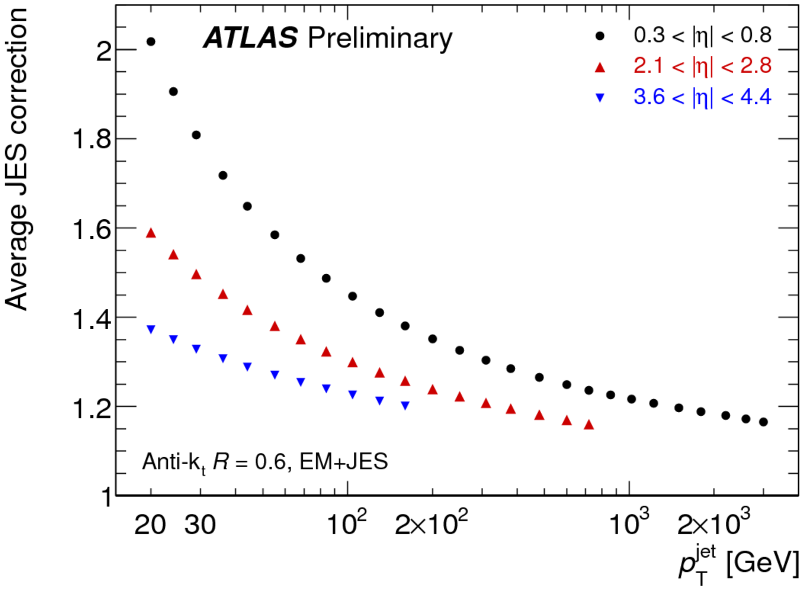

Figure 5.8: *Average jet energy scale corrections in the EM+JES calibration as a function of calibrated jet* p_T *in three different* $|\eta|$ *ranges [146].*

5.4.3.3 Jet Position Corrections

In order to obtain the full jet kinematics additional corrections are needed. Most generally, the energy calibration does not suffice to calibrate the jet mass and ϕ, η positions. The jet mass is correlated with the transverse extension of the jet and different physics models predict different values for it [147]. For this reason, no correction has been applied to the jet mass, and this results in a systematic uncertainty in the jet p_T measurement. The symmetry in ϕ in the ATLAS detector allows for an unbiased measurement in the ϕ direction at the EM scale, so no corrections of that kinematic variable are performed either.

The η measurement is, however, affected by the calorimeter response dependence on η. This results in an energy- and η-dependent bias in the measurement that is depicted in

Figure 5.9. This bias is corrected as a function of jet p_{T} using a similar technique to that used to correct the jet energy, and it is small for the jets used in this thesis.

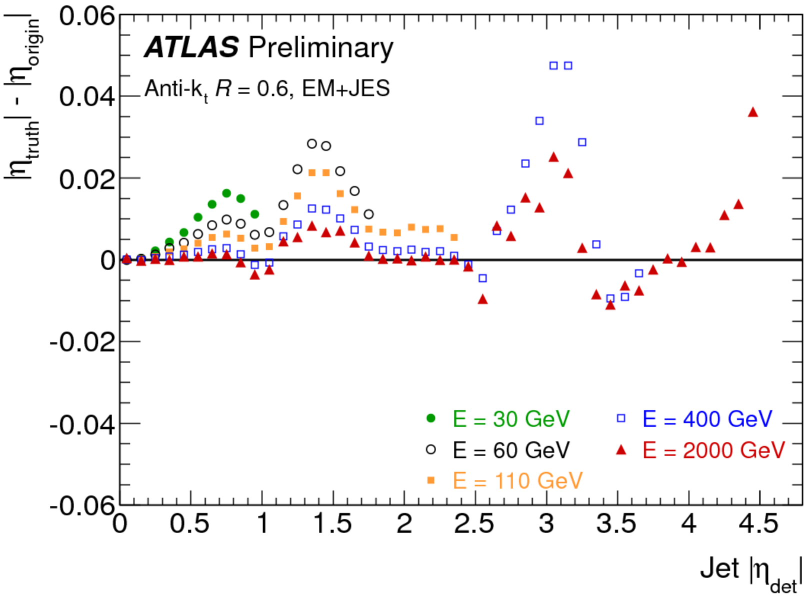

Figure 5.9: *Bias in the reconstructed η as a function of $|\eta|$ for jets of different energies [146].*

An additional position correction is needed to consider the jet origin information provided by the tracking algorithms. The jet four-momentum is thus corrected assuming that the jet origin is at the position of the event vertex. This is a valid assumption as long as the jet originates from that vertex, and not from pile-up. The impact of this correction is small for jets in the p_{T} and rapidity ranges used in this thesis.

Chapter 6

Event Selection

THE DATA USED IN THIS THESIS come from the first 2010 data collection period that occurred from April 10 to August 30. The data from earlier parts of the 2010 run (starting on March 30) were not used because the commissioning of the jet triggers was not complete. Some trigger towers were not properly timed to send data from the same bunch crossing as other towers, and the jet triggers exhibited inefficiencies for regions of the detector where these towers lie. These data contribute negligibly to the data set used in this thesis.

The commissioning of the trigger tower timing continued through all the runs studied here. However, after April 10, the commissioning concentrated on the transition region between the barrel and the end-cap calorimeters, and the result was just a small inefficiency in the trigger selection, as discussed in detail in Section 6.2.2, for most of the summer. From July 29 to August 9, only the presampler was read-out by the trigger decision algorithms to appropriately time its readings in the region between the barrel and the end-cap calorimeters. Since little energy is deposited in the presampler, the result was a significant drop in efficiency for jets of all p_{T}s falling in the transition region. For this reason, data from that period of time are not used in this analysis. This corresponds to a loss of ≈ 0.5 pb^{-1} in the total integrated luminosity of the dataset.

Not all collected data were used, given transient problems in the detectors or data acquisition systems needed to perform the analysis. In particular, stable operation of the CTP, level-1 calorimeter trigger, the ID, the solenoidal magnet, the full calorimeter system and the luminosity detectors was required [148]. The definition of stable operation was determined by detector experts and includes periods with problems that can be understood or eliminated during the data analysis. Similarly, offline software experts determined that the detector conditions did not impact in any detrimental way the luminosity, jet, tracking or missing transverse energy measurements. No requirement on the operation of the muon detectors or the muon data acquisition or software was imposed.

After these "data quality" requirements, the data used in this thesis correspond to an integrated luminosity of 2.43 pb^{-1}. The evolution of the integrated luminosity as a function of time for the period when these data were collected is shown in Figure 6.1. Both the luminosity collected with the ATLAS detector on, and that after the data quality selection above was applied are shown. The next section describes how the measurement of the luminosity for each run, shown in this figure and used for the cross-section measurements, was performed.

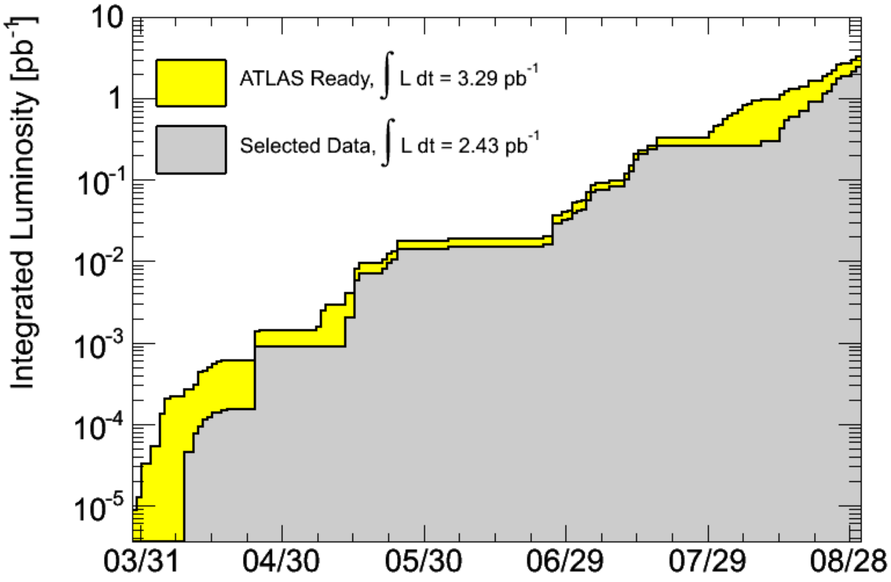

Figure 6.1: *Integrated luminosity of data collected by ATLAS as a function of time in the summer of 2010. The data collected with the ATLAS detector on (ATLAS Ready) is shown in yellow, and the data after the data quality selection used for the analysis is shown in gray.*

6.1 Luminosity Measurement

The luminosity measurement is common to all ATLAS analyses and has been described in detail in Refs. [115, 149]. The relationship between the luminosity and a set of beam parameters was already expressed in Equation 3.0.1. In the presence of multiple proton-proton interactions in each bunch crossing, this can be expressed most conveniently as

$$\mathcal{L} = \frac{\mu f b}{\sigma_{\text{inel}}}, \tag{6.1.1}$$

where μ is the mean number of interactions per bunch crossing and σ_{inel} is the cross section for inelastic scattering. The detectors used to measure the luminosity are not fully efficient for determining whether there was an inelastic scattering. That inefficiency affects both the measured μ (μ_{vis}) and σ_{inel} (σ_{vis}) in such a way that the luminosity can be extracted from the measured quantities.

The measurement of μ_{vis} is straight forward in the limit of $\mu_{\text{vis}} \ll 1$, and amounts to the number of selected events divided by the total number of bunch crossings. As the number of events per bunch crossing increases, proper consideration needs to be given to the Poisson statistics governing the event production. The measurement of σ_{vis} is done using several runs in which van der Meer scans (described in Chapter 3) were performed. These runs happened on April 26, May 9 and October 1 and are used to calibrate the luminosity measurements in prior and later dates. The machine parameters together with the event rates are used to calculate the luminosity of those runs, and the measured luminosities are used to calculate σ_{vis} from μ_{vis} and Equation 6.1.1. The results from all the runs in which a van der Meer scan happened are combined and also used in the evaluation of the systematic uncertainty in the measurement.

Several algorithms and detectors are used to measure μ_{vis}. Two simple algorithms use LUCID and report whether one or both sides of the detector had at least one hit. A similar algorithm uses the BCM and reports whether at least one module reported a hit. The MBTS detector is used by requiring at least a hit in each side within 10 ns of each other. Finally, the ID is used to measure the number of events with a primary vertex with at least four tracks with $p_{\text{T}} > 150$ MeV. Only the algorithm requiring at least one hit on one side of the LUCID detector is used for the luminosity calculation in this thesis. The other algorithms are used to establish the long-term stability of the LUCID measurement. The systematic uncertainties on the luminosity measurement are further described in Section 8.1.

The luminosity is calculated for every luminosity block. A given trigger is, however, not collecting data all the time due to dead-time in the data acquisition caused by different effects (e.g., a sub-detector being unable to send data due to a longer-than-average event processing time.) The collected luminosity is, thus, calculated for each trigger independently as

$$L_{\text{trigger},j} = \int \mathcal{L}_{\text{trigger},j} dt = \sum_i \frac{\Delta t_i^j \cdot \mathcal{L}_i}{P_i^j}, \tag{6.1.2}$$

where j represents the trigger in question, the sum runs over all luminosity blocks i, Δt_i^j is the amount of time that trigger j was able to take data in a given luminosity block, $\mathcal{L}_i$ is the average instantaneous luminosity in the luminosity block and P_i^j is the pre-scale that was applied to trigger j in luminosity block i. The calculation of a trigger-level luminosity simplifies the combination of measurements from different triggers. The luminosity shown in Figure 6.1 corresponds to that of any jet trigger running without pre-scales. All such triggers collect the same luminosity because the effects causing their dead-time are the same, i.e., Δt_i^j is the same for all of them for all luminosity blocks.

6.2 Trigger Selection and Efficiencies

Events in the analysis are selected using a set of level-1 multi-jet triggers. These triggers require several jets reconstructed by the level-1 algorithms described in Section 6.2.1. However, only level-1 jets with $|\eta| \lesssim 3$ are used by these triggers, the forward region of the detector having dedicated triggers. This already constrains the jet selection for this thesis, which focuses on jets with $|y| \leq 2.8$, as described in Section 6.3. Only multi-jet triggers in which all jets share the same threshold are used. Of all triggers used, those with the lowest threshold apply a p_{T} cut of 10 GeV on the level-1 jet objects. Only the two-jet and three-jet triggers are needed for the analysis, since the three-jet trigger ran without pre-scales for the entirety of the data collection period used in this thesis.

Events in the analysis falling in the three-jet “inclusive” multiplicity bin or higher are selected using the three-jet trigger with a jet threshold of 10 GeV (3J10 trigger, hereafter).

In this context, inclusive means that an event must have *at least* n jets to fall in the n-jet bin. Events with three or more jets that pass the 3J10 trigger are not used to fill the two-jet inclusive multiplicity bin, and events falling in this bin are only selected using two-jet triggers. Due to large pre-scales, a combination of several two-jet triggers is used to select events falling in the two-jet inclusive multiplicity bin. Symmetric two-jet triggers with thresholds of 10, 15 and 30 GeV (2J10, 2J15 and 2J30, hereafter) are used to maximize the available statistics. In the two-jet bin, triggers are combined exclusively. This means that events falling in a specific region of phase space are only accepted if one specific trigger fired. The phase space regions assigned to the 2J10, 2J15 and 2J30 triggers are defined by the p_{T} of the second leading jet as reconstructed by the offline software. These regions are defined by the following cuts: [60,80) (2J10), [80,110) (2J15) and ${p_{\mathrm{T}}}^{\mathrm{2nd}} \geq 110$ GeV (2J30). This splitting of the phase space is illustrated in Figure 6.2. This figure also illustrates an additional cut on the leading jet p_{T} of 80 GeV. Events collected by the different triggers are combined using the inverse of the trigger-specific integrated luminosity,

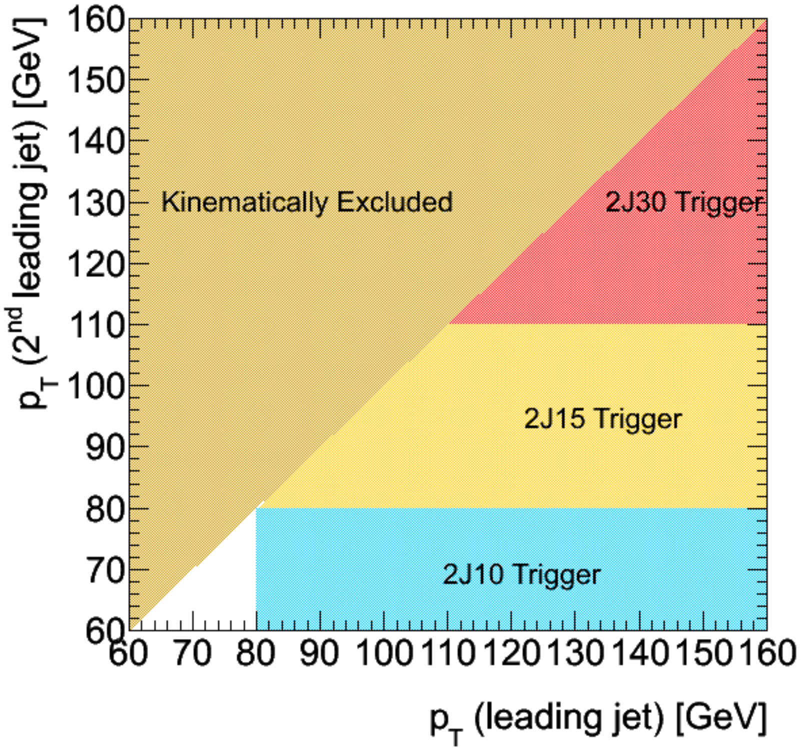

Figure 6.2: *Illustration of the regions of phase space selected using each of the two-jet triggers for events falling in the two-jet inclusive multiplicity bin.*

defined in Equation 6.1.2, as a weight. This effectively corrects the number of events to a detector-level cross section.

A small algorithmic modification of the trigger selection is needed to account for the fact that the 2J30 trigger was not on the level-1 trigger menu before July 28. Fortunately, 2J15 ran without pre-scales before that date in such a way that it can be used as a *virtual* 2J30 trigger. Therefore, events with ${p_{\mathrm{T}}}^{\mathrm{2nd}} \geq 110$ GeV falling in the inclusive two-jet bin are selected using the 2J15 trigger before July 28 and using the 2J30 trigger afterwards. The integrated luminosity used for the 2J30 trigger is then $L_{\mathrm{2J30}} = L_{\mathrm{2J15}}^{<\mathrm{July28}} + L_{\mathrm{2J30}}^{\geq\mathrm{July28}}$. This integrated luminosity is equal to that of triggers that ran without pre-scales through the entire data taking period, such as the 3J10 trigger.

Most events passing a level-1 jet trigger were written to a common stream by the event filter, and that stream is used to select most of the data. However, some events could not be assigned to the proper stream due to a time-out in the event-filter algorithms. The time-out was typically caused by particles in a jet that are not contained in the calorimeter and result in many hits in the muon spectrometer. This increased activity in the muon system causes the event-filter muon algorithms to run for longer than their typical allocated time, thus timing out. These events are written to the debug stream, which is also used to select events for this analysis. The contribution from this stream to the measurements performed is negligible as outlined in Section 6.3.4.

6.2.1 Level-1 Jet Trigger Reconstruction

Signals from the liquid argon and tile calorimeters are used to build level-1 jets. The signal that is output by the pre-amplifiers in the front-end electronics is added to signals of nearby cells in the front-end board to build a “trigger tower” that is used by the level-1 trigger algorithms. The level-1 trigger signal, thus, bypasses the shaping circuitry, and has a read-out that is different from that used by the offline reconstruction algorithms. Trigger towers are 0.1×0.1 in size in $[\eta, \phi]$. They are built for the EM and hadronic calorimeters independently and are sent off the detector to the counting rooms near the ATLAS cavern through electrical connections for further trigger processing.

The jet/energy module in the counting rooms combines EM and hadronic trigger towers into one calorimeter tower, and into groups of four, yielding an effective tower size of

0.2×0.2 in $[\eta, \phi]$. Only towers with $|\eta| < 3.2$ are used by the barrel/end-cap jet algorithms that provide the inputs to the multi-jet triggers used in the studies in this thesis. A sliding-window jet algorithm is run on these towers, searching for transverse energy signals above pre-defined thresholds within a window of 0.8×0.8 in $[\eta, \phi]$. To avoid multiple counting of jet candidates, the transverse energy deposited at the center of the window (in a region of size 0.4×0.4) is required to be a local maximum. This is illustrated in Figure 6.3, where the effective calorimeter towers are shown together with the window that corresponds to a level-1 trigger jet and the towers used to enforce the local maximum. The information from the jet/energy module, including transverse energy and position of the jet candidates is sent to the CTP, where trigger decisions based on the trigger menu are performed and distributed to the on-detector read-out electronics.

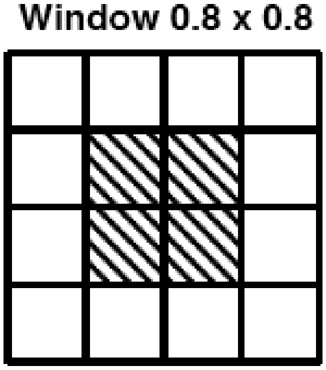

Figure 6.3: *Illustration of the geometry of a level-1 trigger jet [88], with the jet trigger towers of size 0.2×0.2 in $[\eta, \phi]$. The towers used to establish a local maximum that prevents multiple counting of candidates are shaded.*

The size of the window in the level-1 jet algorithm together with the requirement of a local maximum in the center of the jet can result in a mismatch between the level-1 jet trigger object and the offline reconstructed jet. This is self-evident for anti-k_t jets of resolution parameter $R = 0.4$ that may be reconstructed very close to each other. In this case, two jets may be reconstructed as one, leading to a mismatch between the jet multiplicities in the offline reconstruction and those in the the level-1 reconstruction. This leads to an inefficiency in the trigger selection for multi-jet events that is discussed in detail in the next section.

The level-1 jet trigger uses not only a simplified jet algorithm and coarser granularity in

$[\eta, \phi]$, but also a calorimeter signal with increased noise and no noise suppression algorithm. For this reason, and to limit the rate of events collected by the jet triggers, only trigger towers with energies higher than 1 GeV are used in the jet building, and the lowest trigger threshold is set at 10 GeV for jet trigger algorithms using a window of size 0.8 in η and ϕ. The large noise and the simplified noise suppression scheme cause a very rough jet energy measurement with large fluctuations. In order to guarantee that the trigger was fully efficient, jets must have a p_{T} that is much larger than the threshold used in the level-1 algorithm. This effect and disagreements between data and Monte Carlo simulation in the size of the jet trigger inefficiencies for jets of lower p_{T} determine the kinematic cuts used in the analysis performed in this thesis.

6.2.2 Trigger Efficiencies

The efficiency of the two-jet and three-jet triggers has been determined. Since up to three jet trigger objects are required, the efficiency to trigger on each of these objects needs to be studied. This is done using a "bootstrap" method, in which events are selected with a trigger that is fully efficient for the process of interest, and thus unbiased.

The efficiency of the trigger to accept the leading jet is calculated using the minimum bias trigger scintillators. In particular, at least one hit is required in the MBTS detector (i.e., the MBTS_1 trigger is used). This trigger is fully efficient for events in which a large number of reconstructed tracks exist. Events with jets in the kinematic range of interest for this analysis have typically a very large number of tracks, so the MBTS_1 trigger is fully efficient to select them. In particular, a vertex with at least five tracks with $p_{\mathrm{T}} > 150$ MeV will be required as part of the offline selection for the analysis (see Section 6.3). Figure 6.4 shows the MBTS_1 trigger efficiency as a function of the number of reconstructed tracks in the event, demonstrating its high efficiency for selected events.

The efficiency for the leading jet in the event to pass the J10 trigger depends on the jet p_{T}, since the threshold is applied to the reconstructed transverse energy, and potentially η, due to the changing calorimeter geometry. Since the jet rapidity is of interest to the analysis, the efficiency is studied as a function of rapidity instead of η. Figure 6.5 shows the efficiency of the J10 trigger to select events as a function of the reconstructed offline leading jet p_{T} within $|y| < 2.8$. The efficiency is shown for jets reconstructed with the

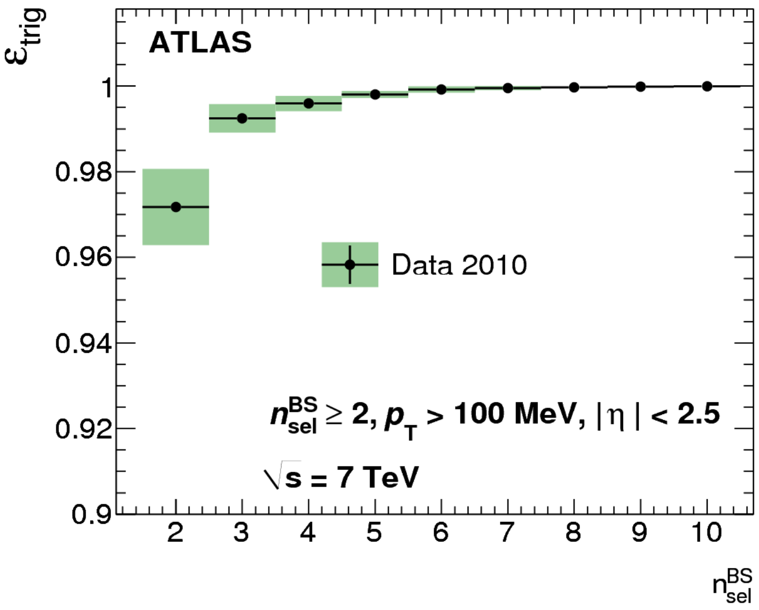

Figure 6.4: *Efficiency of the MBTS_1 trigger to select events as a function of the number of reconstructed tracks in the event with* $p_T > 100$ MeV *[97].*

anti-k_t algorithm with $R = 0.4$ and $R = 0.6$. The efficiency as a function of p_T shows a

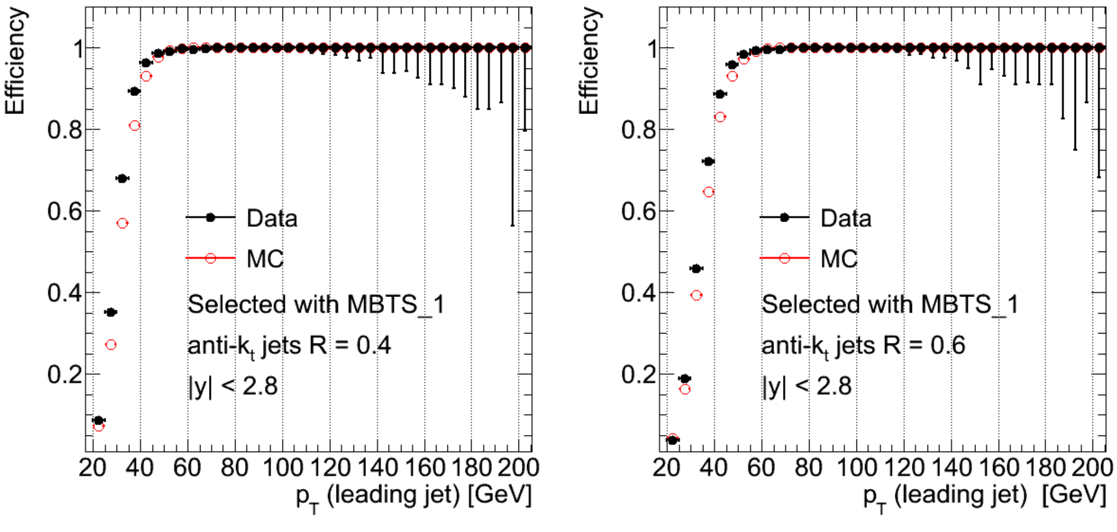

Figure 6.5: *Efficiency of the J10 trigger to select events as a function of the leading jet* p_T *within* $|y| < 2.8$ *for jets reconstructed with the anti-*k_t *algorithm with* $R = 0.4$ *(left) and* $R = 0.6$ *(right). The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the unbiased bootstrap method in data, using events selected with the MBTS_1 trigger.*

characteristic shape, with a "turn-on" region where the efficiency increases quickly, and a "plateau" region where the efficiency becomes 1, and no longer varies as a function of p_T.

The width of the turn-on region is large (30-40 GeV) and relates to the E_{T} resolution of the level-1 jet objects, which is coarse. The error bars on the data are calculated using a uniform bayesian prior going from 0 to 1 that accounts for the fact that the efficiency cannot be higher than 1.

The trigger reaches an efficiency higher than 99% at $p_{\mathrm{T}} \approx 50$ GeV for jets reconstructed with $R = 0.4$ and at $p_{\mathrm{T}} \approx 60$ GeV for jets reconstructed with $R = 0.6$. This is expected, given that larger jets collect more energy, and that the level-1 jet objects are the same in both cases. There is agreement between data and Monte Carlo simulation in the plateau region. Differences between the two exist on the efficiencies in the turn-on region. A di-jet Pythia sample was used for the Monte Carlo simulation in these studies (see Section 9.1 for details on this sample). However, the use of other Monte Carlo simulations (e.g., with different jet fragmentations) is not expected to change the value of the jet p_{T} where the plateau starts, since this is dominated by the large E_{T} resolution of the level-1 jet objects.

Figure 6.6 shows the efficiency of the J10 trigger to select events as a function of the reconstructed offline leading jet rapidity, when the leading jet has $p_{\mathrm{T}} > 60$ GeV. The efficiency is 100% across all rapidities except in the region of transition between the barrel

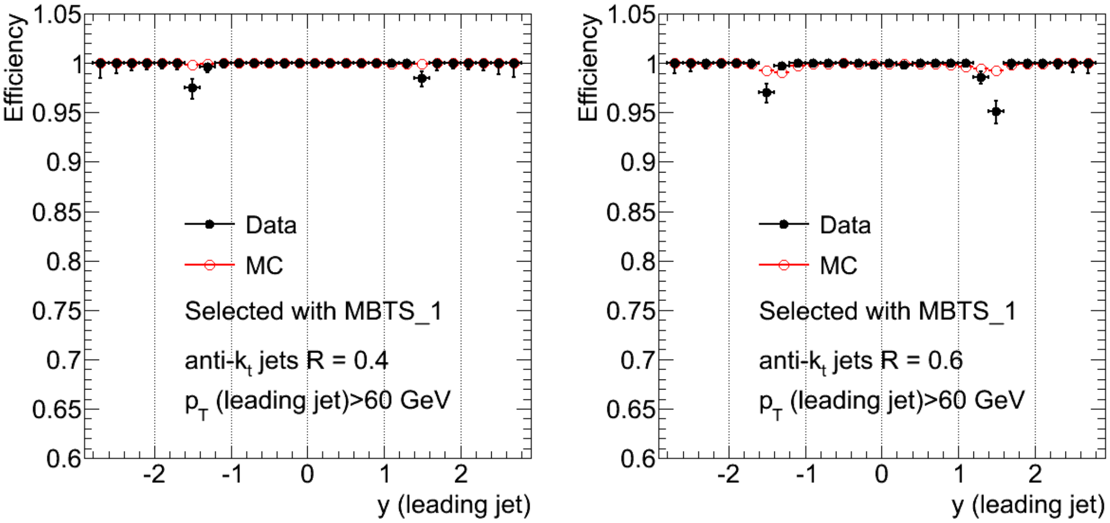

Figure 6.6: *Efficiency of the J10 trigger to select events as a function of the leading jet rapidity for events where the leading jet* p_{T} *is above* 60 GeV *for jets reconstructed with the anti-*k_t *algorithm with* $R = 0.4$ *(left) and* $R = 0.6$ *(right). The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the unbiased bootstrap method in data, using events selected with the MBTS_1 trigger.*

and the end-cap calorimeters, where inefficiencies of up to 5% are observed in the data. Such inefficiencies are not present in the Monte Carlo simulation, and they are related to a mismatch in the level-1 jet trigger configuration between data and Monte Carlo simulation. As described at the beginning of this chapter, the level-1 jet trigger timing was commissioned during the period when these data were taken. This implied feeding only parts of the calorimeter measurement to the level-1 tower builders. This results in a lower energy measurement and the plateau being reached at higher p_{T}s. This translates into an inefficiency observed in Figure 6.6. However, this inefficiency only appears at around the value of the p_{T} where the plateau is reached, as demonstrated in Appendix A, and has a small impact in the measurements presented in Chapter 10. The size of this impact is quantified in Section 8.2.

A similar study can be performed for the J15 and J30 triggers. Figure 6.7 shows the efficiencies of the J15 and J30 triggers as a function of leading jet p_{T} for jets reconstructed with the anti-k_t algorithm with $R = 0.6$. Only the results for $R = 0.6$ are shown because the plateau is reached at higher p_{T}s than for $R = 0.4$, and the results are comparable. The results for $R = 0.4$ together with the y dependence of the efficiency for these two triggers are shown for completeness in Appendix A. The small inefficiencies observed in the data

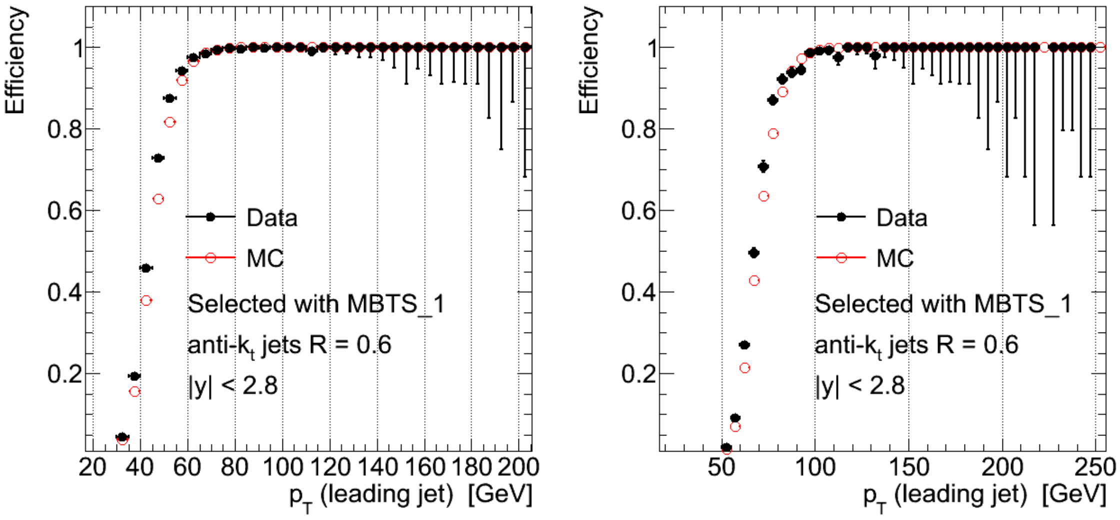

Figure 6.7: *Efficiency of the J15 (left) and J30 (right) triggers to select events as a function of the leading jet* p_{T} *within* $|y| < 2.8$ *for jets reconstructed with the anti-*k_t *algorithm with* $R = 0.6$*. The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the unbiased bootstrap method in data, using events selected with the MBTS_1 trigger.*

in the plateau region are caused by inefficiencies in the region of transition between the barrel and end-cap calorimeters, which are similar to those shown in Figure 6.6 for the J10 trigger. The J15 and J30 trigger are fully efficient for events with the leading jet p_{T} above 80 and 110 GeV in the Monte Carlo simulation. This is also true in data, once the jets with $|y| \sim 1.5$ are not considered.

The basic behavior of the J10, J15 and J30 triggers can be used to understand the 2J10, 2J15, 2J30 and 3J10 triggers used to select multi-jet events. Indeed, a two-jet event in which the leading jet falls on the plateau region of one of the single-jet triggers will be fully efficient to select two-jet events, independent of the second jet p_{T}. Similarly once the plateau region for the second leading jet is understood, the two-jet triggers can be used to select three-jet events with full efficiency and calculate the three-jet trigger efficiency. Figure 6.8 shows the efficiency of the 2J10 and 3J10 triggers as a function of the p_{T} of the second leading and third leading jet. The result is only shown for jets with $R = 0.6$, even though similar conclusions can be reached for jets with $R = 0.4$, as shown in Appendix A. The leading or the two leading jets, respectively, are not only required to be on the plateau, but also to be matched to a level-1 jet. The matching is done using a maximum distance of 0.3 in η and ϕ, which is highly efficient for identifying whether a level-1 jet was corresponds

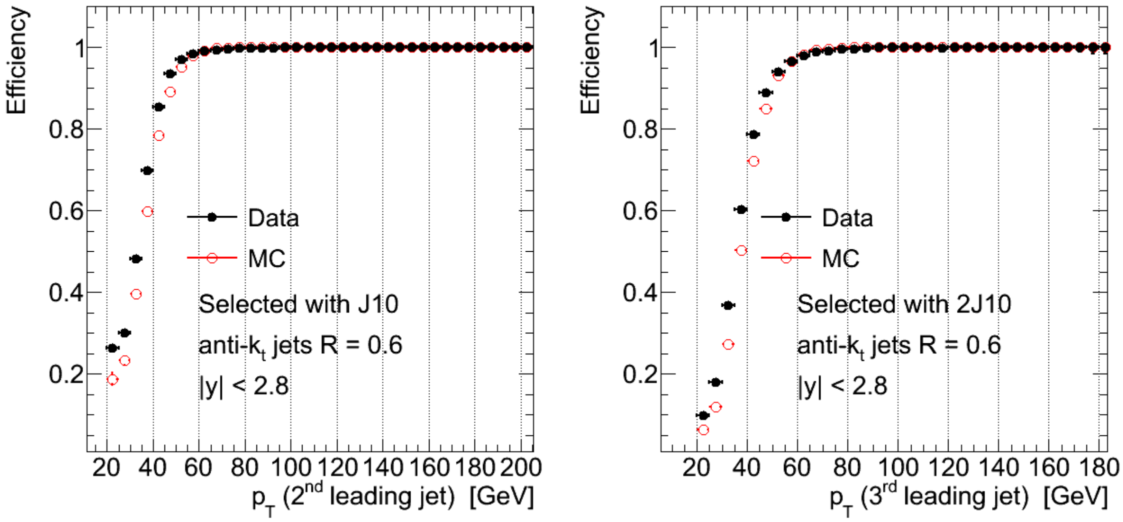

Figure 6.8: *Efficiency of the 2J10 (left) and 3J10 (right) triggers to select events as a function of the second and third leading jet* p_{T} *within* $|y| < 2.8$ *for jets reconstructed with the anti-*k_t *algorithm with* $R = 0.6$. *The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the unbiased bootstrap method in data, using events selected with the J10 and 2J10 triggers.*

to an offline jet [150]. This additional requirement guarantees that the small inefficiency in the region of transition between the barrel and end-cap calorimeters does not affect these efficiency measurements.

Figure 6.8 shows that the 2J10 and 3J10 triggers have an efficiency bigger than 99% to select two-jet events and three-jet events if the second and third leading jet p_{T} are above 60 GeV. A similar conclusion can be reached for the 2J15 and 2J30 triggers, when thresholds of 80 GeV and 110 GeV are used to select jets, as demonstrated in Appendix A.

The previous figures only explore the dependence of the trigger efficiency on jet p_{T} and rapidity. This is sufficient for the single-jet triggers, given how the level-1 jet algorithm works. However, once several jets are present in the event, the mismatch between the level-1 and offline jet algorithms can cause topological dependences in the multi-jet trigger efficiencies. These dependences necessarily do not affect the measurement of the total multi-jet cross section significantly, or otherwise a hint of them would have been observed in Figure 6.8 in the form of a lower efficiency somewhere in the plateau region. The fact that nothing is observed suggests that any inefficiency has to be correlated with low-probability topological configurations. However, understanding these inefficiencies is still important in order to perform measurements that might directly or indirectly probe those topological configurations.

The energy density in a jet originating from QCD processes is typically peaked around the center of the jet, falling quickly as the distance from the jet axis increases [147]. The local maximum constraint on the sliding-window algorithm tries to capture this so that there is a direct map between an offline jet and level-1 jet. However, the energy density pattern can be different, leading to two jets reconstructed next to each other. In such cases, the sliding window algorithm might not be able to capture the complexity of the energy deposition pattern, leading to one level-1 jet instead of two, and a corresponding inefficiency in the multi-jet trigger. This is illustrated in Figure 6.9 for the 3J10 trigger. The efficiency of the 3J10 trigger to select events in which two of the selected jets are found nearby ($0.4 < \Delta R < 0.6$) is shown as a function of the third leading jet p_{T}. Only jets with $|y| < 2.8$ reconstructed with the anti-k_t algorithm with $R = 0.4$ are used. The efficiency is shown as a function of the third leading jet p_{T} because in a three-jet event the third leading jet is almost always one of the two jets found close to one another.

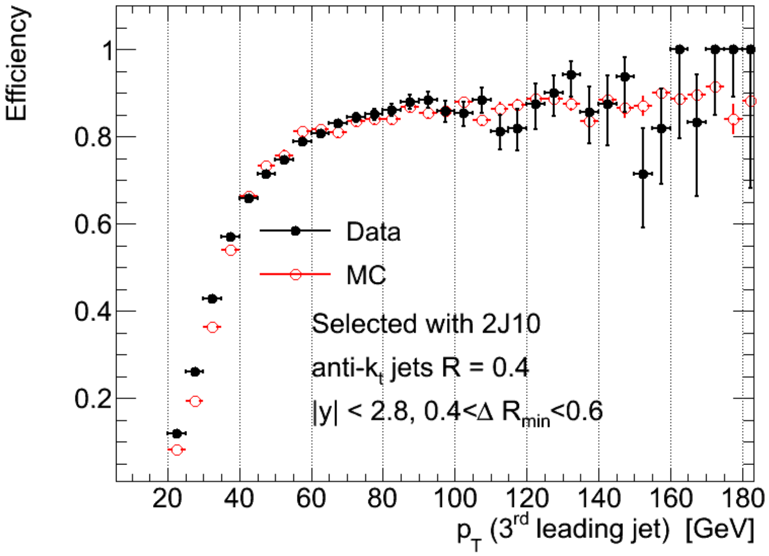

Figure 6.9: *Efficiency of the 3J10 trigger to select events in which two of the selected jets are found a distance* ΔR *between 0.4 and 0.6 apart from each other as a function of the third leading jet* p_{T}. *Only jets with* $|y| < 2.8$ *reconstructed with the anti-*k_t *algorithm with* $R = 0.4$ *are used. The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the unbiased bootstrap method in data, using events selected with the 2J10 triggers.*

An inefficiency in the trigger is observed as expected, and this inefficiency seems to depend weakly on the p_{T} of the third leading jet once this jet has $p_{\mathrm{T}} \gtrsim 60$ GeV. The inefficiency is well described by the Monte Carlo simulation. Such an inefficiency is not relevant in the use of two-jet triggers, since two-jet events in which the two jets are produced nearby cannot be produced by QCD processes.

The inefficiency is expected to be significant for anti-k_t jets with $R = 0.4$ because their size is smaller than that of the sliding window. Figure 6.10 shows the efficiency of the 3J10 trigger as a function of the minimum ΔR between the first, second and third leading jets for both jets built with $R = 0.4$ and $R = 0.6$, when all three jets have a $p_{\mathrm{T}} > 60$ GeV. In fact, no inefficiency is observed when jets are built using $R = 0.6$, while the inefficiency observed in Figure 6.9 increases as the distance between the closest two jets in the event decreases. A good description of this effect is found in the Monte Carlo simulation.

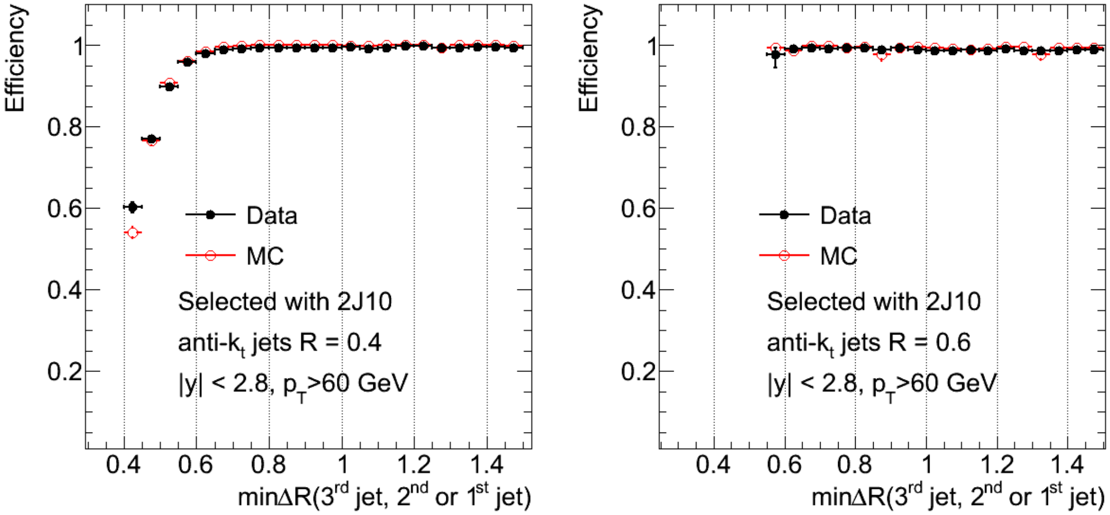

Figure 6.10: *Efficiency of the 3J10 trigger to select events as a function of the smallest distance in* $[\eta, \phi]$, ΔR, *between selected jets. Only jets with* $|y| < 2.8$ *and* $p_{\mathrm{T}} > 60$ GeV *reconstructed with the anti-*k_t *algorithm with* $R = 0.4$ *(left) and* $R = 0.6$ *(right) are used. The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the unbiased bootstrap method in data, using events selected with the 2J10 triggers.*

6.3 Offline and Jet Selection

Events in the analysis are selected using an additional set of cuts on the reconstructed offline objects, including jet kinematic and jet-specific data quality cuts. A vertex cut is also included, requiring at least one primary vertex with five or more associated tracks in the event. This cut serves as a first rejection for events originating from cosmic rays and particles produced in interactions of the beam with particles in the beam tunnel ("beam halo" and "beam gas"). However, its impact in the analysis is negligible once all other cuts are considered.

6.3.1 Kinematic Cuts and Observables

A high cut on the minimum jet p_{T} is implemented to select jets in the region where the triggers used in the analysis are most efficient. While this is not strictly necessary, it simplifies significantly the analysis, particularly given the differences between the Monte Carlo simulation and the data for the turn-on trigger efficiency curves. A high p_{T} cut also minimizes the impact of the underlying event on the measurement.

Jets are required to have a minimum p_{T} of 60 GeV. Jets are also required to have

$|y| \leq 2.8$. This simplifies the trigger selection, and leaves out a part of the detector where the jet energy scale is understood less precisely [151]. At the same time, the part of phase space that is left out would otherwise not contribute significantly to the measurements performed, since the cross section for forward jet production is much lower than for central jet production at the same p_{T} [152].

An additional cut is imposed on the p_{T} of the leading jet, which should at least be 80 GeV. This is added to the event selection to ease the comparison with next-to-leading-order calculations. Kinematic cuts on the outgoing partons of the matrix-element calculation can, in fact, cause instabilities in the results of the calculations due to suppression of terms that are needed to cancel other very large terms [153]. This effect can be mitigated by avoiding fully symmetric p_{T} cuts on the outgoing partons. A cut of 80 GeV on the p_{T} of the leading jet/parton is thus chosen to exclude the region of phase space where the next-to-leading-order calculation gives unphysical results. This is further discussed in Section 9.2.

Finally, two additional sets of p_{T} cuts are used for certain studies, namely, $[{p_{\mathrm{T}}}^{\mathrm{jet}} \geq 80\ \mathrm{GeV}, {p_{\mathrm{T}}}^{\mathrm{lead}} \geq 110\ \mathrm{GeV}]$ and $[{p_{\mathrm{T}}}^{\mathrm{jet}} \geq 110\ \mathrm{GeV}, {p_{\mathrm{T}}}^{\mathrm{lead}} \geq 160\ \mathrm{GeV}]$. The asymmetry in the cuts is maintained to allow for comparisons to next-to-leading-order calculations. The main objective of these studies is to help in the tuning of the parameters that describe final state radiation in leading-order and next-to-leading-order Monte Carlo generators. These different sets of cuts are summarized in Table 6.1.

<table>
<tr><th colspan="2">Standard Kinematic Cuts</th></tr>
<tr><td colspan="2">${p_{\mathrm{T}}}^{\mathrm{jet}} \geq 60$ GeV
$|y^{\mathrm{jet}}| \leq 2.8$
${p_{\mathrm{T}}}^{\mathrm{lead}} \geq 80$ GeV</td></tr>
<tr><th colspan="2">Kinematic Cuts for FSR Tuning</th></tr>
<tr><td>${p_{\mathrm{T}}}^{\mathrm{jet}} \geq 80$ GeV
$|y^{\mathrm{jet}}| \leq 2.8$
${p_{\mathrm{T}}}^{\mathrm{lead}} \geq 110$ GeV</td><td>${p_{\mathrm{T}}}^{\mathrm{jet}} \geq 110$ GeV
$|y^{\mathrm{jet}}| \leq 2.8$
${p_{\mathrm{T}}}^{\mathrm{lead}} \geq 160$ GeV</td></tr>
</table>

Table 6.1: *Kinematic cuts applied to jets used in the analysis. One set of "standard" kinematic cuts is used for most studies, while two sets are used in studies aimed at helping in the tuning of final state radiation in Monte Carlo generators.*

In addition to some basic kinematic distributions, the following variables are used in some of the different measurements performed in this thesis:

- $H_{\mathrm{T}} = \sum_{\text{jets}} p_{\mathrm{T}}$, where the sum runs over all selected jets.
- ${H_{\mathrm{T}}}^{(2)} = {p_{\mathrm{T}}}^{\text{lead}} + {p_{\mathrm{T}}}^{\text{2nd}}$, where the superscripts refer to the ordering in p_{T} of the selected jets in the event.
- $|\Delta\phi_{ij}|$, where the subscripts also refer to the ordering in p_{T} of the selected jets in the event.
- $|\Delta\eta_{ij}|$, defined similarly to the previous quantity.

6.3.2 Jet Data Quality

The data quality cuts applied are established by studying data collected with a random trigger, and with jet triggers selecting events in which either one or no proton bunches were passing through the ATLAS interaction point [154]. The former set of events was used for the study of noise and other detector problems, while the latter was used for the study of cosmic and beam backgrounds. The following additional cuts are applied to reject fake jets.

Cuts to remove energy spikes in the HEC due to defective signals.

- A jet is rejected if the fraction of the total energy of the jet deposited in the HEC is above 0.5 and the absolute value of the HEC jet quality variable is higher than 0.5. The HEC quality variable is calculated as the fraction of the jet energy coming from cells with a high χ^2 value for the fit to the shape of the calorimeter pulse performed in the read-out drivers.
- A jet is also rejected if the absolute value of energy contributions to the jet from cells with negative energy readings is higher than 60 GeV.

Cuts to remove coherent noise in the EM calorimeter.

- A jet is rejected if the fraction of the total energy of the jet deposited in the EM calorimeter is greater than 0.95, and the absolute value of the EM jet quality variable, calculated like the HEC jet quality variable, is higher than 0.8.

Cuts to remove cosmic and beam ("non-collision") backgrounds.

- A jet is rejected if the energy-weighted average raising time of the signals from the cells that build the jet is not within 25 ns of the time when the bunch crossing happened.
- A jet is rejected if it has $|\eta| < 2$, the fraction of the jet energy deposited in the electromagnetic calorimeter is less than 0.05, and the "charged fraction" of the jet is below 0.05. The charged fraction is defined as the ratio of the scalar sum of the p_{T}s of tracks associated to the jet to the calibrated jet p_{T}. Only tracks of p_{T} above 500 MeV and within ΔR from the jet axis equal to the resolution parameter of the jet are used.
- A jet is rejected if it has $|\eta| \geq 2$, and the fraction of the energy of the jet deposited in the EM calorimeter is less than 0.05.
- Finally, a jet is also rejected if it has $|\eta| < 2$ and the maximum fraction of the jet energy deposited in any one calorimeter layer is greater than or equal to 0.99.

The first three of these non-collision-background cuts are designed to reject cosmic muons, which could overlap a bunch crossing in which a legitimate QCD event was produced. The last cut is designed to reject beam halo or beam gas muons that travel parallel to the beam.

All these cuts are found to be effective at rejecting fake jets and other spurious backgrounds. A tighter set of cuts also exists that targets high p_{T} fake jets, which occur with a smaller probability [154]. However, these cuts do not need to be applied in this analysis.

The jet data quality cuts also need to be as efficient as possible in accepting jets created in proton-proton collisions. Some of the effects described above are not modeled well by the Monte Carlo simulations (e.g. the fit to the signal shape in the HEC and the EM calorimeter is by definition perfect, since the expected shape is used to produce the signal from simulated energy depositions). For this reason, the effects on the signal sample need to be studied using data.

A calculation of the efficiency is performed using a di-jet sample in which the two jets approximately balance. At least one of the jets passed the cuts described above, to increase the probability that a jet is produced in the collision. The jet is also required to have $|\eta| < 2$, where the jet energy measurement is better understood. This is the "tag" jet, and allows the "probe" jet, then presumed to be a jet generated in the collision, to be used to study the efficiency. If both jets pass the data quality cuts and have $|\eta| < 2$, then the event is used twice. Events are selected with single-jet triggers, following the studies in Ref. [16]. The event selection cuts are:

- $|\Delta\phi(\text{tag}, \text{probe})| > 2.6$, and
- $|{p_{\mathrm{T}}}^{\text{tag}} - {p_{\mathrm{T}}}^{\text{probe}}|/{p_{\mathrm{T}}}^{\text{avg}} < 0.4$, where ${p_{\mathrm{T}}}^{\text{avg}}$ is the mean p_{T} of the two jets.

The dependence of the results on the actual values of these cuts has been investigated in Appendix B and is found to be negligible.

Figure 6.11 shows the efficiency as a function of jet p_{T} in several η bins that capture the changing calorimeter geometry (and the change in the data quality cuts). The efficiency for each set of data quality cuts is not shown individually. Jets built with a resolution parameter of $R = 0.4$ are used in this figure. Results for jets built with $R = 0.6$ are similar and can be found in Appendix B. The results can be extended to higher p_{T}s and show 100% efficiency, even though the statistical uncertainties become large due to the few events collected in those kinematic regions in this dataset. The data quality cuts are essentially fully efficient for jets produced in proton-proton collisions with $p_{\mathrm{T}} > 30$ GeV, and thus fully efficient for jets used in this analysis.

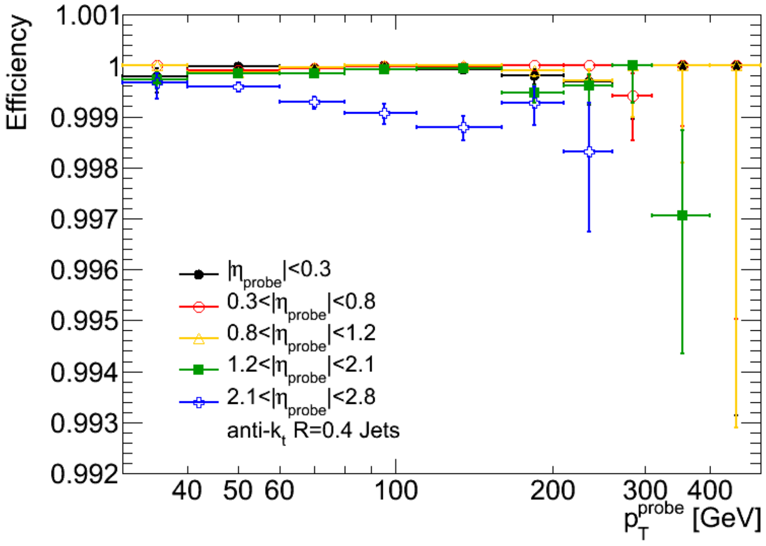

Figure 6.11: *Efficiency of the jet data quality cuts to accept jets originating from proton-proton collisions as a function of jet* p_{T}*. Results are shown for jets built with the anti-*k_t *algorithm with* $R = 0.4$ *and for jets falling in different* $|\eta|$ *bins.*

6.3.3 Pile-up Rejection

The data collected in this analysis contains multiple proton-proton interactions in a significant number of selected events. In particular, if inclusive two-jet events (selected with the two-jet triggers) are considered, 58% of them had one reconstructed vertex, 31% of them had two, 9.1% had three, and 1.9% had four or more. The impact of the pile-up events on the energy of selected jets is controlled through the offset correction, described in Section 5.4.3.

However, pile-up interactions not only have an average effect, but also a smearing effect in the jet energy reconstruction. This is caused by the fact that not all jets have the same contribution from pile-up. This smearing effect reduces the jet energy resolution proportionally to the instantaneous luminosity of the bunch crossing. Measurements that are sensitive to the jet energy resolution thus remain dependent on the number of reconstructed primary vertices, even after the offset correction is applied. This effect is particularly important for low p_{T} jets, but is also expected to increase with jet multiplicity, since high-multiplicity final states are more sensitive to fluctuations on the jet energy measurement.

In order to reduce the pile-up fluctuations, a variable that is correlated to the jet-by-jet pile-up contribution can be used. This variable is the jet vertex fraction (JVF) [155]. The jet vertex fraction is defined as

$$\mathrm{JVF} = \frac{\sum\limits_{\mathrm{event-vertex}} {p_{\mathrm{T}}}^{\mathrm{matched-tracks}}}{\sum {p_{\mathrm{T}}}^{\mathrm{matched-tracks}}}, \tag{6.3.1}$$

where the sum in the numerator runs over all tracks that are matched to the jet and that originate at the event vertex, and the sum in the denominator runs over all tracks that are matched to the jet, independent of their origin. The jet-track matching is performed using a ΔR in $[\eta, \phi]$ equal to the resolution parameter of the jet, and using the track coordinates at their distance of closest approach to the beam. Jets with no tracks associated to them are assigned a JVF of -1. Jets with JVF$= -1$ are always accepted in the studies that follow, because most of such jets arise from the high-rapidity region, which is beyond the acceptance of the ID.

Optimally, the jet vertex fraction variable should be used to change the prediction of the jet energy on a jet-by-jet basis [156]. This would allow a reduction of the pile-up fluctuations, and an appropriate jet selection. However, such techniques have not been fully commissioned with data in ATLAS yet. Instead, a JVF cut is used to reject pile-up backgrounds. Such a cut will necessarily have an efficiency for signal jets and some rejection power for pile-up jets. In order to appropriately select the cut for the analysis, the right compromise between these two needs to be found so that the measurement is independent of the instantaneous luminosity of a bunch crossing.

The Kolmogorov distance [157], Δ, between the measurements performed using low-luminosity bunch crossings (with just one reconstructed primary vertex, $N_{\mathrm{PV}} = 1$) and those performed using higher-luminosity bunch crossings is used to find the optimal value of the JVF cut for each measurement. Figure 6.12 shows the Kolmogorov distance as a function of the cut on the absolute value of JVF (always accepting jets with no tracks associated to them) between measurements performed with events reconstructed with $N_{\mathrm{PV}} = 1$ and those reconstructed with $N_{\mathrm{PV}} \geq 1$. The distance is shown for the inclusive multiplicity distribution and the H_{T} distribution for events with $N_{\mathrm{jets}} \geq 2$. Only jets with $|\mathrm{JVF}| \geq 0.5$, i.e. with approximately most of their energy originating at the event vertex, are used for this study. The distance is typically small, even though it increases with N_{PV}. The $|\mathrm{JVF}|$

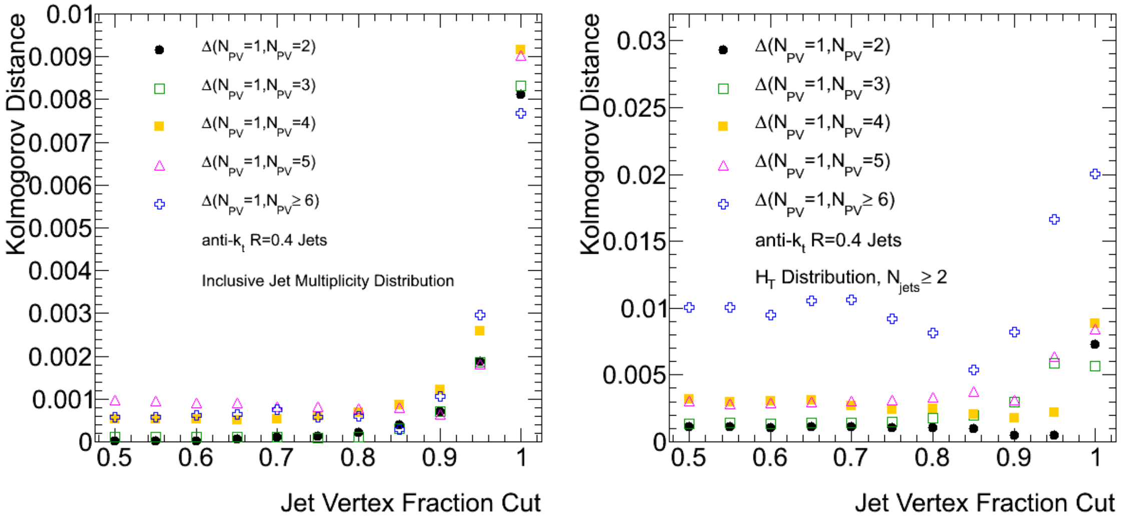

Figure 6.12: *Kolmogorov distance as a function of the cut on the absolute value of JVF between measurements performed with events reconstructed with* $N_{\mathrm{PV}} = 1$ *and those reconstructed with* $N_{\mathrm{PV}} \geq 1$. *The distance is shown for the inclusive multiplicity distribution (left) and the* H_{T} *distribution for events with* $N_{\mathrm{jets}} \geq 2$ *(right).*

distribution and the H_{T} distribution for events with $N_{\mathrm{jets}} \geq 2$. Only jets with $|\mathrm{JVF}| \geq 0.5$, i.e. with approximately most of their energy originating at the event vertex, are used for this study. The distance is typically small, even though it increases with N_{PV}. The $|\mathrm{JVF}|$

cut value at which the Kolmogorov distance reaches a minimum varies for different values of N_{PV} and even for different observables. This is not unexpected, since efficiency and rejection curves most generally depend on jet multiplicities and kinematics.

Similar studies have been performed for all observables of interest in this thesis, and a summary of those studies can be found in Appendix B. Based on those studies, jets are required to have $|\mathrm{JVF}| \geq 0.7$. Given the varying optimal values of $|\mathrm{JVF}|$ for events with different N_{PV} and for the different observables, the $|\mathrm{JVF}|$ cut is varied from 0.5 to 0.9 to understand its impact quantitatively. Differences in the obtained distributions are then used to estimate the systematic uncertainties associated to this cut, as detailed in Section 8.2.

The impact of the jet vertex fraction cut in the analysis remains small. This is due to the high p_{T} of selected jets, which makes pile-up contributions generally small in comparison, and the fact that 90% of the events used in the analysis have minimal or no pile-up contributions. A more quantitative measure of its impact is given in the next section. However, the jet vertex fraction cut succeeds at improving the measurement for events with high N_{PV}, thus removing the dependence of the measurement on the instantaneous luminosity of the bunch crossing. This is illustrated in Figure 6.13, where the inclusive jet multiplicity distribution is shown after the jet vertex fraction cut for $N_{\mathrm{PV}} = 1$ and $N_{\mathrm{PV}} = 4$. Both measurements are the same within statistical uncertainties and no dependence on N_{PV} is observed.

6.3.4 Selected Events

In summary, the following cuts are applied on reconstructed objects to select events for the analysis:

- The event must contain at least one reconstructed primary vertex with five or more tracks associated to it.
- An event must contain at least one jet with a p_{T} greater than or equal to 80 GeV and $|y| \leq 2.8$.
- Jets are required to be within $|y| \leq 2.8$ and have a p_{T} greater than or equal to 60 GeV in order to be counted.

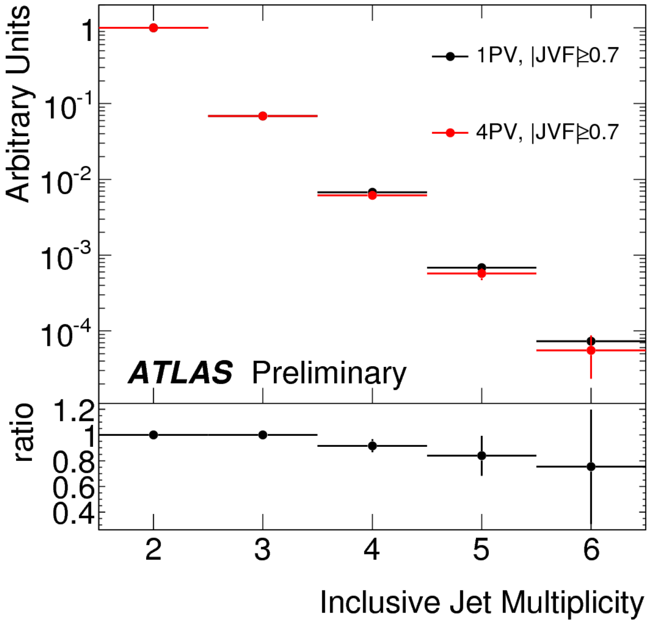

Figure 6.13: *The inclusive jet multiplicity distribution for events containing 1 and 4 primary vertices after the JVF cut. The distributions are normalized to the first bin. The ratio of the two distributions is shown in the bottom plot.*

- Jets are only accepted if they pass selection cuts optimized to reject cosmic and beam backgrounds, coherent noise in the electromagnetic calorimeter and defective signals in the hadronic end-cap calorimeter.
- Jets are only accepted if 70% ($|\mathrm{JVF}| > 0.70$) of their jet-matched charged particle p_{T} comes from the event vertex. Jets with no associated tracks are not rejected.
- Only events with a minimum of two selected jets are used.

In addition to this standard set of cuts, two more sets of kinematic cuts on the jets are applied for studies of final state radiation, as summarized in Table 6.1.

For illustrative purposes, Figure 6.14 presents an event display of a six-jet event passing the standard selection cuts. For this event, the six selected jets are well separated spatially. The main display shows a projection in the $r - \phi$ plane of the ATLAS detector, with a fish-eye perspective that allows seeing details in the ID and calorimeters, despite their size being much smaller than that of the muon spectrometer. The EM calorimeters are shown in green, while the hadronic calorimeters are shown in red. The calorimeter read-out segmentation in cells is delimited by gray boundaries, and energy depositions within those cells are shown as

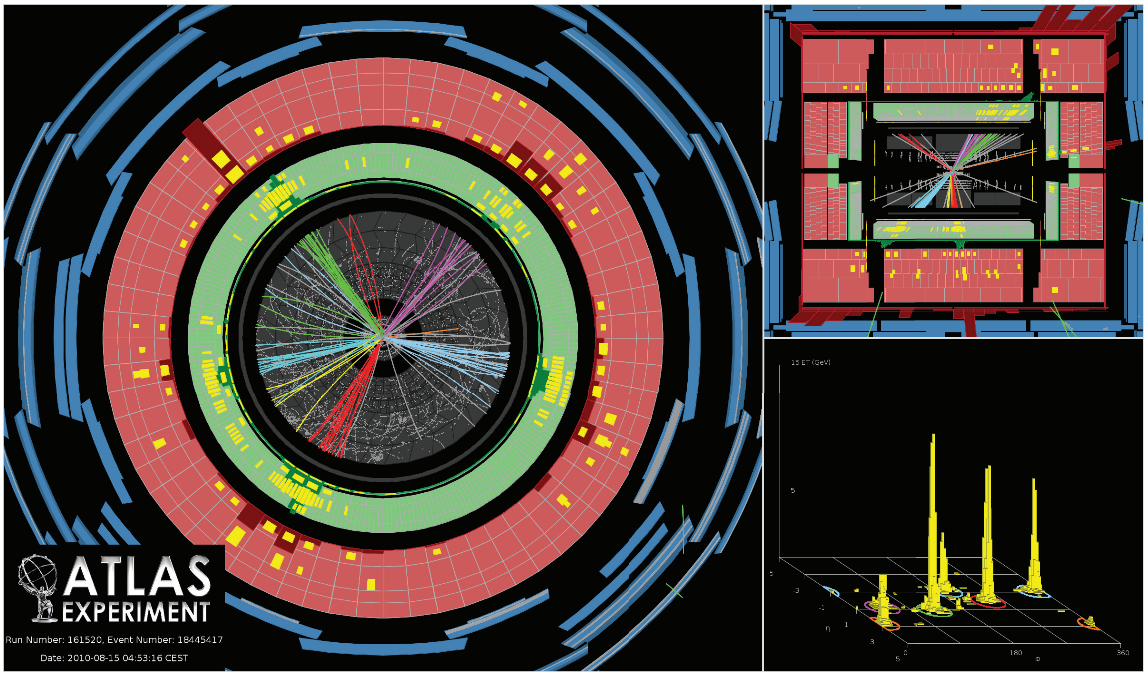

Figure 6.14: *Event display of a six-jet event passing the selection cuts.*

yellow squares of size proportional to the energy deposited. A similar image, but showing the projection on the $r - \eta$ plane, is shown on the top right corner. The transverse energy depositions within a jet are plotted as a function of η and ϕ on a lego plot on the bottom left corner.

A total of approximately half a million multi-jet events were collected in the analysis after the standard selection cuts. Table 6.2 presents the total number of multi-jet events as a function of the inclusive jet multiplicity. No correction for pre-scales in the two-jet bin has been applied to the numbers on this table. The numbers are shown after only the vertex and kinematic cuts are applied (left), once the jet quality cuts are applied (left center) and after all cuts, including the pile-up rejection cut, are applied (right center). The number of events coming from the debug stream after all cuts is also shown (right), and those events contribute $\sim 0.1\%$ to the total. While most of these events have jets with high p_{T}, they still contribute negligibly to all measurements performed (see Appendix B). The impact of the pile-up rejection cut increases with multiplicity going from 0.4% in the two-jet bin to 3.5% in the six-jet bin. Most of the impact is caused by rejecting jets whose $|\mathrm{JVF}|$ is less than 0.5, and that originate from a vertex different from the event vertex. The evolution of the impact of the cut with multiplicity is expected, since pile-up resolution effects controlled by

Number of jets	Number of selected events			
	Only Kinematic Cuts	Jet Quality Cuts	All Cuts	Debug Stream
≥ 2	502,277	502,106	500,148	233
≥ 3	113,776	113,733	112,740	79
≥ 4	11,141	11,130	10,999	16
≥ 5	1,120	1,119	1,100	0
≥ 6	119	119	115	0

Table 6.2: *Number of selected events using the criteria described in this chapter as a function of inclusive jet multiplicity (for jets built with anti-*k_t *algorithm and* $R = 0.4$*) before correcting for pre-scales. Events listed in the first column are selected if they have a leading jet* $p_{\rm T} \geq 80$ GeV *and* $|y| \leq 2.8$*. Additional jets are counted for that event, if they have* $p_{\rm T} \geq 60$ GeV *and* $|y| \leq 2.8$*. The second column applies an additional set of jet quality cuts as described in Section 6.3.2. The third column applies all selection cuts, including the pile-up rejection cut described in Section 6.3.3. The final column shows the number of events selected from the debug stream after all cuts are applied.*

the pile-up rejection cut increase with multiplicity. Overall, the effect is small, particularly when compared to the systematic uncertainties in the measurement. The effect of the jet data quality cuts is even smaller. Therefore, while all these cuts are desirable to control known effects, they do not change the physical interpretation of the results that will be shown in Chapter 10.

Chapter 7

Unfolding of Detector Effects

The measurements performed with the detector generally do not correspond to the observables obtained with jets built from truth particles, "truth jets", in the Monte Carlo simulation. If the jet reconstruction and event and jet selection are fully efficient, the mismatch is exclusively caused by resolution effects.

An example of such resolution effects arises from the fact that jet production through QCD processes decays exponentially as a function of jet p_{T}. Symmetric fluctuations in the measurement of the jet p_{T} make the number of jets with measured p_{T} above a given cut be higher than the true number of jets above that cut. More jets are measured above the cut because the underlying p_{T} distribution is not symmetric around the cut, favoring the migration of low p_{T} truth jets to higher bins of reconstructed p_{T}.

In addition, sometimes one truth jet is reconstructed as two jets due to inaccuracies in the reconstruction of the energy deposition patterns in the calorimeter. This affects the measured jet multiplicity, creating again a mismatch between measured and truth distributions. These detector effects, which fold into the truth distributions, can be corrected through an "unfolding correction" that enables the retrieval of the truth distribution from the measurement. The correction is applied to measured distributions so that theoretical comparisons are possible, and independent of the detector used to perform the measurement.

The calculation of the unfolding correction needs a way of mapping particles and other objects from a full event Monte Carlo simulation to the objects measured in the detector after event reconstruction. A sophisticated ATLAS detector simulation exists and has been discussed in Section 4.7. This simulation enables performing this mapping taking into account many details that would otherwise be hard to include and understand.

The detector simulation needs to provide a description of interactions within the real detector, and the event simulation should provide a description of the aspects of the real event that can impact the unfolding. If both descriptions are accurate, then the unfolding correction becomes very versatile, allowing for a simultaneous correction of resolution effects and other effects, such as inefficiencies or miscalibrations.

If the events are mostly reconstructed in the same bin (of a given observable) that they belong to, the unfolding correction has a small effect. In addition, correlations between the measurement in different bins also become small. This simplifies the unfolding procedure, since the correlations and the significance of the different correlations in the final result are diminished [158]. Under this assumption, the correction can be limited to a multiplicative correction factor for each bin, calculated as $c = \mathrm{obs}_{\mathrm{truth}}/\mathrm{obs}_{\mathrm{reco}}$ for any given observable, in what is typically known as "bin-by-bin unfolding."

7.1 Event and Detector Simulations

Two event generators are used to estimate the size of the unfolding factors and their systematic uncertainties: Pythia [79] with the AMBT1 underlying event tune parameters [159] and Alpgen [160] interfaced to Herwig [161] to produce the parton shower and hadronization and Jimmy [77] to produce the underlying event using the AUET1 tune parameters [162]. These generators are also used in leading-order theoretical calculations for comparisons to the measurements.

Approximately 2 million events are generated with Alpgen, while roughly 4.5 million events make the Pythia sample. The samples are generated using different slices in jet multiplicity (for the Alpgen sample only) and p_{T} to guarantee that enough events are generated in the regions of phase space of interest to the analysis. The details of the slicings can be found in Ref. [163]. The Alpgen sample provides the central value of the unfolding correction because more events remain in the sample after selection cuts than in the Pythia sample and because it succeeds better than Pythia at describing important features of the event. The Pythia sample is thus only used to understand the impact of certain systematic effects on the corrections.

An average of two non-diffractive interactions are overlaid to the generated events following Poisson statistics to simulate the effect of pile-up. The non-diffractive interactions are generated using Pythia with the AMBT1 tune for both the Pythia and the Alpgen samples, and they are generated without any kinematic constrains, thus generally producing soft particles.

Truth distributions are built using the same jet reconstruction algorithms as for the detector-level reconstruction using as input all stable particles with a lifetime above 10 ps in the center-of-mass frame of the main event interaction. This includes muons and neutrinos created in hadronic decays, even though their interactions in the calorimeter are small or essentially non-existent. Only particles from the main interaction are used in this process, since the use of particles from the additional pile-up interactions would modify the observables of interest and the unfolding correction would then not be able to account for the effects of pile-up.

The generated particles are passed through the ATLAS detector simulation [124], assuming a large separation between bunches. This allows for the read-out of a given event to be insensitive to the activity in other events, and corresponds to the parameters of the beam during the period when the data for the analysis was collected.

7.2 Simulation Description of the Data

The simulation needs to succeed at reproducing the data if it is to be used to calculate the unfolding corrections. The detector simulation needs to be able to properly represent the interactions of particles within the detector, how those are interpreted by the read-out electronics, and the characteristics of the beam. The event generator also needs to be able to reproduce aspects of the event that may impact the unfolding correction, such as certain topological configurations or particle spectra within the jet that may affect the jet energy scale.

In the context of jet triggers, the former has been studied in detail in the previous chapter, showing a good agreement between Monte Carlo simulation and data, except in the small region of transition between the barrel and end-cap calorimeters. In the context of the offline reconstruction, detailed comparisons between data and Monte Carlo simulations have been performed using 2009 and 2010 data. Observables that are relevant in the reconstruction of tracks [133,164–166], vertices [136,137] and jets [18,19,167] have been shown to be in reasonable agreement between data and Monte Carlo simulations. The descriptions of the jet energy scale and resolution have been studied separately [151,168–172] and also found to be in good agreement within the systematic uncertainties of the studies.

Figure 7.1 shows two examples of the studies on jet reconstruction. On the left, the mean energy measured at the electromagnetic scale in the second layer of the EM calorimeter is shown as a function of calibrated jet $p_{\rm T}$ for anti-k_t jets built of clusters with $R = 0.6$. This layer has the most energy deposited of all layers for jets in the $p_{\rm T}$ range shown, and this study demonstrates that some of the important features of the longitudinal shower development are well described in the Monte Carlo simulation.

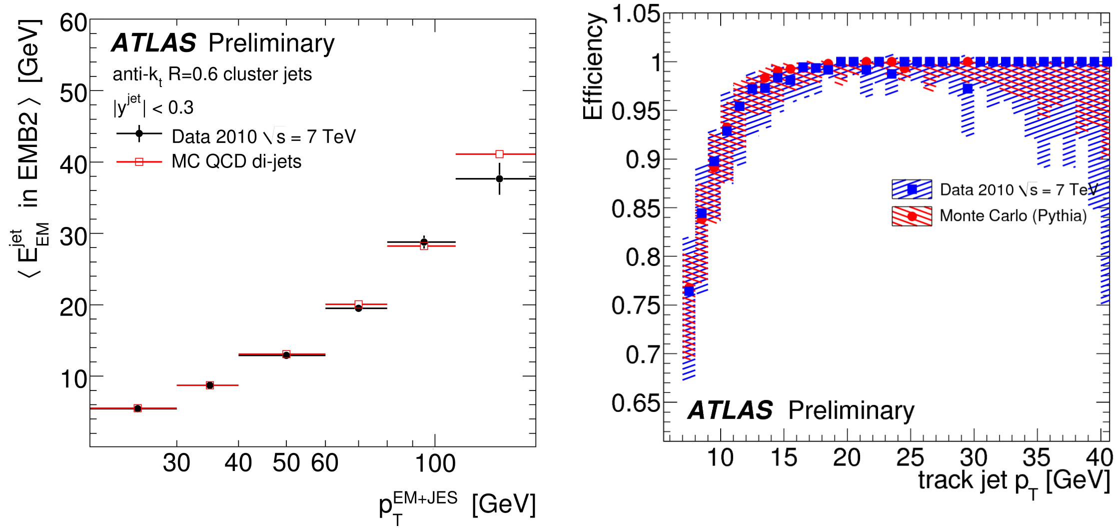

Figure 7.1: *Comparisons between the results obtained in data and Monte Carlo simulations for some properties of the jet reconstruction. The mean energy deposited in the second layer of the electromagnetic calorimeter as a function of calibrated jet* $p_{\rm T}$ *is shown on the left [19]. The jet reconstruction efficiency, calculated comparing jets reconstructed from calorimeter clusters to jets built of tracks, is shown as a function of the* $p_{\rm T}$ *of the jet built of tracks [169].*

On the right, the jet reconstruction efficiency measured using jets built of tracks ("track jets") as a baseline is shown as a function of the $p_{\rm T}$ of the track jets. A di-jet sample is used to guarantee the physical presence of a jet, using a similar technique to that used in Section 6.3.2. The error bands represent the addition in quadrature of statistical and systematic uncertainties in the measurement. The latter are estimated varying the details of the event selection in the study [169]. This study shows that the jet reconstruction is fully efficient for track jet p_{T}s above ≈ 20 GeV, and that the jet reconstruction efficiency is well described by the Monte Carlo simulation. This study also demonstrates that, in essence, the jet reconstruction efficiency is 100% for jets used in the analysis, given that the fraction of the jet momentum carried by charged particles is ≈ 50% for jets of $p_{\mathrm{T}} \sim 60$ GeV [19].

The parameters of the beam, such as the average number of proton-proton collisions in a bunch crossing, or the size of the luminous region should also be well described, since they can impact the size of pile-up effects and the jet reconstruction. Figure 7.2 shows measures of both these parameters in data and in the two Monte Carlo simulations used to calculate the unfolding correction factors. The events in the data are selected using the inclusive two-jet triggers and weighted according to the luminosity of each trigger. The Monte Carlo simulation distributions are normalized to the area under the distributions from data.

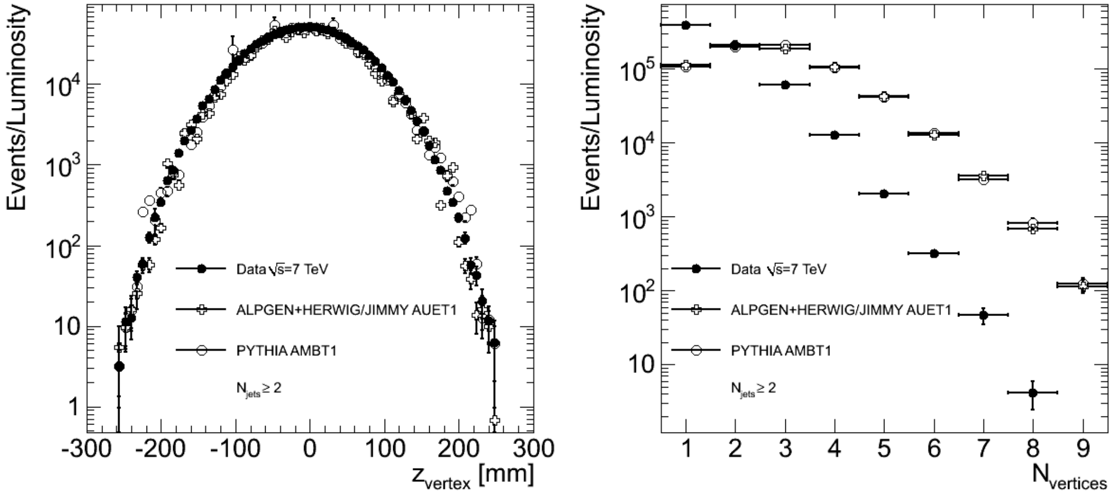

Figure 7.2: *Distributions of the position of the reconstructed primary vertices along the z axis (left) and the number of vertices reconstructed in an event (right) in data and the two Monte Carlo simulations used for the unfolding of detector effects.*

On the left, the distribution of the z position of the reconstructed vertices, with at least five associated tracks, is shown. This illustrates the size of the luminous region along the direction of the beam, where the size is largest. A good agreement is found between the simulated and the reconstructed distributions. Large statistical fluctuations for some values of the PYTHIA simulation arise from events with large weights that can occur due to the slicing in p_{T} used in the generation and discussed in the previous section.

On the right, the distribution of the number of reconstructed vertices in an event, with at least five tracks associated to each, is shown. The number of reconstructed vertices is correlated to the number of proton-proton collisions in an event, and is higher on average in the Monte Carlo simulation than in the data. This is expected, since the Monte Carlo simulation was optimized for the luminosity conditions towards the end of the 2010 run, which had a higher instantaneous luminosity per bunch than in the data used in the analysis. To improve the agreement between the Monte Carlo simulations and the data, events in the simulations are weighted to match the vertex distribution observed in the data. Two slightly different weighting functions are used for the ALPGEN and PYTHIA simulations. The normalization of the re-weighting function is arbitrary, since only the relationship between the measured and truth distributions is of interest for the calculation of the unfolding factors.

The vertex multiplicity distribution shown is specific to the selected inclusive two-jet sample. A correlation between trigger pre-scales and luminosity exists, which makes this distribution be dependent on the trigger or combination of triggers used for the event selection. For example, the 2J10 trigger was strongly pre-scaled in the later runs due to its high rate, and most of the data collected with it comes from early runs, where the instantaneous luminosity was low. The vertex multiplicity distribution is thus more strongly peaked at 1 for the 2J10 trigger. The impact of this effect is not taken into account in the unfolding correction, and instead included as a (small) systematic uncertainty in the correction, as detailed in Section 8.2.

The studies above show that the simulation describes (or can be modified to describe) well the interaction of particles within the detector, and some conditions of the data collected that can impact the analysis. However, the event generator should also be successful at describing the physics that can impact the unfolding correction. For example, a very soft energy distribution of particles within a jet can affect significantly the average measurement of the jet energy, and thus the size of the unfolding corrections. Effects that can impact the jet energy scale are discussed separately in Section 8.3 and properly accounted for in the systematic uncertainties in the jet energy scale.

An example that is more relevant to this section is that of event topologies and, in particular, the rates of selected jets reconstructed next to each other. A good prediction of these rates is desirable so that the topological inefficiencies in the three-jet triggers are properly accounted for in the unfolding correction. Figure 7.3 shows the fraction of selected jets with another selected jet within a ΔR of 0.6 in $[\eta, \phi]$ as a function of the inclusive jet multiplicity. Jets reconstructed with a resolution parameter $R = 0.4$ are used, since only

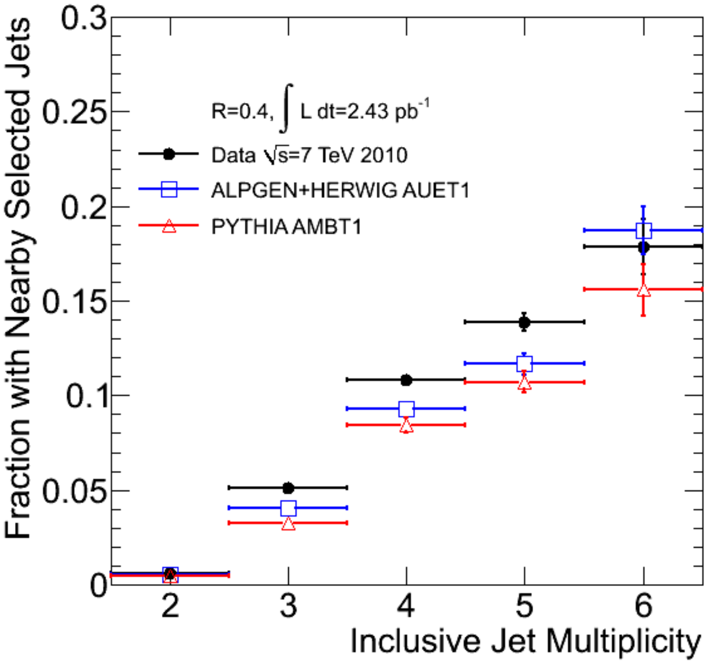

Figure 7.3: *Fraction of selected jets with another selected jet within a* ΔR *of 0.6 in* $[\eta, \phi]$ *as a function of the inclusive jet multiplicity of the event, for jets built with the anti-*k_t *algorithm with resolution parameter* $R = 0.4$.

for this size is a topological inefficiency in the trigger algorithm observed. The fraction of jets increases with inclusive jet multiplicity, and this is expected, since the phase space fills up as the required number of jets in the event is increased. Neither of the two simulations agree with the data, even though the prediction from ALPGEN is closer than that from PYTHIA. The difference between ALPGEN and the data is comparable to the difference between PYTHIA and ALPGEN in the three-jet bin, where this effect is most relevant in determining the inefficiency in the trigger selection. Given the small difference between data and ALPGEN, and the small fraction of events affected by this inefficiency, no attempt is made to make ALPGEN look more similar to the data in this aspect for the purposes of unfolding. The difference is considered as part of the systematic uncertainty in the unfolding factors and accounted for using the difference in the unfolding factors calculated in Alpgen and Pythia, as discussed in detail in Section 8.2.

Finally, a discussion of event streaming in the simulation is relevant for this section. The event filter streaming present in the data is not simulated. In particular, the event-filter algorithms cannot time out in the simulation, and events that would fall in the debug stream in the data streaming are not distinguished in the Monte Carlo simulations. These events have, however, been observed in Monte Carlo simulations too. The energy of a jet in these events is not fully measured in the calorimeter, since particles within the jet reach the muon spectrometer. However, the lost energy can be considered a source of fluctuations in the jet energy measurement that is included in the unfolding correction. Detailed studies of these events in Monte Carlo simulation and in data do not exist, and little more can be said about the accuracy with which the effect is modeled in the simulation. However, few events in the analysis are affected by this. The impact of any such mis-modeling would, thus, be negligible.

7.3 Bin-by-bin Unfolding

A reasonable agreement is found between Monte Carlo simulations and data in aspects that impact the unfolding. Small differences are not corrected in the calculation of the unfolding factors, and instead considered in the calculation of systematic uncertainties in Section 8.2. The unfolding method used is a simple bin-by-bin unfolding method in which the multiplicative factors for an observable, obs, are calculated for each bin i as

$$c^{i}_{\mathrm{unfold}} = \frac{\mathrm{obs}^{i}_{\mathrm{truth}}}{\mathrm{obs}^{i}_{\mathrm{reco}}} = \frac{\mathrm{obs}^{i,\mathrm{truth}}_{i,\mathrm{reco}} + \sum\limits_{i\neq j} \mathrm{obs}^{i,\mathrm{truth}}_{j,\mathrm{reco}}}{\mathrm{obs}^{i,\mathrm{truth}}_{i,\mathrm{reco}} + \sum\limits_{i\neq j} \mathrm{obs}^{j,\mathrm{truth}}_{i,\mathrm{reco}}}, \tag{7.3.1}$$

where in the second expression the superscript refers to truth bins and the subscript to reconstructed bins in which events may fall. The second expression is useful to appropriately take into account correlations between the numerator and the denominator in the calculation of the statistical uncertainties in the unfolding factors, since it splits the two into fully correlated and uncorrelated components.

Bin-by-bin unfolding does not take into account correlations between bins in the unfolding process, and is thus sensitive to changes in the shape of the input distributions, biasing the corrected distribution towards the distribution in the Monte Carlo simulation [158]. This leads to a systematic uncertainty driven by the limitations in the knowledge of the shape of the distribution of interest. The effect of this bias is, however, small as long as migrations between truth and reconstructed bins are small, i.e. as long as the migration matrix is almost diagonal. Migrations between bins can be characterized by two quantities:

$$\text{Efficiency} = \frac{\text{obs}_{i,\text{reco}}^{i,\text{truth}}}{\text{obs}_{\text{truth}}^{i}}; \quad \text{Purity} = \frac{\text{obs}_{i,\text{reco}}^{i,\text{truth}}}{\text{obs}_{\text{reco}}^{i}}. \tag{7.3.2}$$

The efficiency thus represents how many events in the truth bin i fall in the reconstructed bin i, while the purity represents the fraction of events in the reconstructed bin i that come from the corresponding truth bin. For certain observables, the bin widths can be chosen to minimize bin migrations due to resolution effects, thus minimizing also the systematic uncertainties in the unfolding factors and adding robustness to the method. Bins that are very wide are useful for this purpose, but they can hide physics of interest for the measurement. For this reason, a compromise needs to be found in the choice of bin widths for each observable. In the present analysis purities and efficiencies of typically 0.7 were obtained.

Figure 7.4 shows the purity and the efficiency for the measurement of the production cross section as a function of inclusive jet multiplicity and that of the differential cross section as a function of event H_{T} for events with $N_{\text{jets}} \geq 2$. The results are shown for both the PYTHIA and ALPGEN samples, and are comparable. The results for the other observables studied in the analysis can be found in Appendix C. The efficiencies and purities are generally around 0.7 or higher. They decrease with increasing jet multiplicity due to several effects, including the trigger topological inefficiency, resolution effects that become more important for steeply decaying p_{T} distributions in the sub-leading jets and the increased probability of multiplicity mis-reconstruction as the available phase space for jet production fills up. The purity and efficiency are roughly constant as a function of H_{T} due to a choice of bin widths that roughly follows the variation of the resolution of the H_{T} reconstruction with increasing H_{T}. Small differences without further transcendence are found between ALPGEN and PYTHIA.

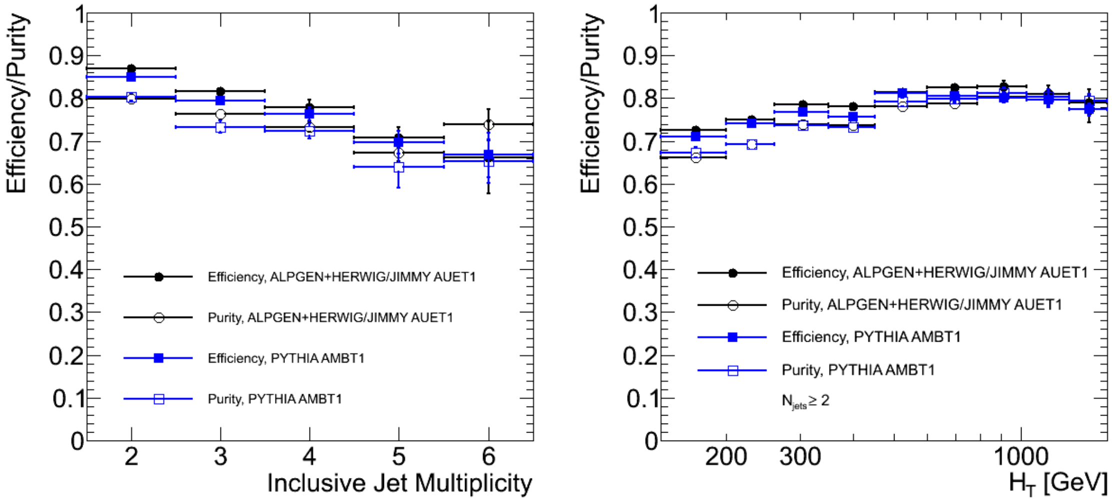

Figure 7.4: *Efficiency (solid markers) and purity (open markers) as a function of jet multiplicity (left) and event H_{T} for events with $N_{\mathrm{jets}} \geq 2$ (right) as calculated in the* ALPGEN *(black markers) and* PYTHIA *(blue markers) samples.*

The unfolding of the normalized differential cross sections as a function of $|\Delta\eta|$ and $|\Delta\phi|$ between the different jets in the event requires slight modifications to the bin-by-bin unfolding technique. The modifications are necessary due to the large statistical fluctuations for large angular separations found in the Monte Carlo samples used. These statistical fluctuations are much larger than those of the data sample. If the bin-by-bin unfolding was applied without any modifications, the statistical uncertainties would propagate to the measurement, increasing artificially its systematic uncertainties.

Using very wide bins in the measurement is possible, but it is not necessary if the statistical uncertainty in the data or the bin migrations caused by resolution effects are considered. A similar problem has been encountered in Ref. [173]. A slow variation of the unfolding constants enables the use of larger bins to calculate the unfolding constants than used to bin the data. This method adds smoothness to the unfolding constants without any loss of accuracy. A possible underestimation of the systematic uncertainties can happen in this method if features in the unfolding are believed to appear with much smaller granularity than the binning used. In order to verify that such features are not present, several cross-checks have been performed that study the impact of problematic areas of the calorimeter with small granularity (e.g. dead optical connections) on these measurements. These studies can be found in Appendix C.

Figure 7.5 shows the unfolding factors for $\sigma_{3\mathrm{jet}}^{-1} d\sigma/d|\Delta\eta_{12}|$ and $\sigma_{3\mathrm{jet}}^{-1} d\sigma/d|\Delta\phi_{12}|$ for events with three or more jets before and after the rebinning is performed. The gray bands

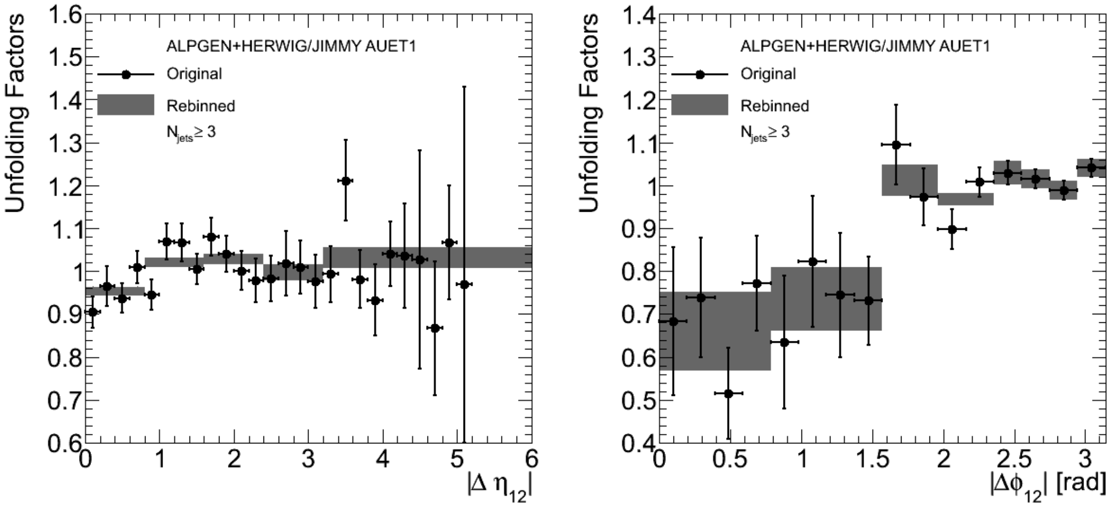

Figure 7.5: *Unfolding factors for $\sigma_{3\mathrm{jet}}^{-1} d\sigma/d|\Delta\eta_{12}|$ (left) and $\sigma_{3\mathrm{jet}}^{-1} d\sigma/d|\Delta\phi_{12}|$ (right). The gray band is centered around the final unfolding factors and represents the statistical uncertainty in them. The points show the unfolding factors when the same binning as for the data is used.*

represent the statistical uncertainty in the final unfolding factors, and are centered around the final values of the unfolding factors. These values are consistent with those obtained with the bin widths used for the measurement. The slowly varying size of the unfolding factors together with the much reduced statistical uncertainties obtained justify the use of the method. Similar figures for the other angular variables studied in this thesis can be found in Appendix C.

The unfolding factors for the measurements as a function of $|\Delta\eta|$ are close to 1, as expected when bins that provide high purity and efficiency in the unfolding are used. This is also true for most other measurements performed in the analysis. The unfolding factors for the measurement of the cross section as a function of the jet multiplicity vary from ≈ 0.9 to ≈ 1.1, never differing from 1 by more than $\approx 10\%$. This also applies to the differential cross sections and cross-section ratios investigated. Some of these unfolding factors are shown in Section 8.2 with their corresponding systematic uncertainties, while the rest are shown for completeness in Appendix C.

An exception to this are the angular distributions as a function of $|\Delta\phi|$, one of which was shown in Figure 7.5. At small $|\Delta\phi_{12}|$, the value of the unfolding factor is ≈ 0.7. This corresponds to a region of phase space with nearby jets where the trigger shows inefficiencies. However, it also corresponds to a region of phase space where mis-measurements of the jet energy and angular position are enhanced. For example, many of the events with small $|\Delta\phi_{12}|$ correspond to events in which the second leading jet p_{T} is reconstructed to be lower than that of the third leading jet. This is self-evident since transverse momentum in the event would otherwise not be conserved. For this reason, the purity and efficiency of those bins are low, unlike for all other measurements. This results in a large systematic uncertainty in those regions that could be remedied through the use of more complex unfolding methods. This is not pursued here, given the large statistical uncertainties in the unfolding factors, which already limit the precision of this particular measurement.

Chapter 8

Systematic Uncertainties in the Measurements

Several effects may impact the measurements performed in the analysis and are not or cannot be corrected for. Effects that cannot be corrected for arise from limitations in our knowledge about physics models or the detector, which can impact significantly the interpretation of the detector signals, and thus the final measurements. The size of these effects can be estimated and their impact on the measurements evaluated to provide the main component of the systematic uncertainties in the measurements.

The systematic uncertainties in the multi-jet measurements performed in the analysis can be most conveniently split into three essentially uncorrelated components that capture the impact of systematic effects on three important aspects of the measurement: the luminosity, the unfolding factors and the jet energy scale. The uncertainty in the luminosity measurement has been evaluated using redundancies in the luminosity detectors as well as studies of the stability of the measurement with time [149]. The result is a 3.4% uncertainty, or 0.08 pb^{-1} in the 2.43 pb^{-1} used for the analysis. This is the smallest of the three components that contribute to the total systematic uncertainty. Certain measurements that study relative rates, or ratios of cross sections are not affected by the uncertainties in the luminosity measurement.

The uncertainty in the unfolding factors includes effects that arise from limitations in the unfolding procedure, whether those limitations are in the descriptions provided by the Monte Carlo simulations, caused by simplifications in the procedure, or inherent to the bin-by-bin unfolding method that was used. Some of these effects have already been briefly described in previous chapters. They will be developed further in this chapter, and their impact on the analysis evaluated. While some of the effects (e.g. the uncertainties in the physics models) can also impact the estimate of the jet energy scale, the correlations are small and do not dominate the systematic uncertainties. For this reason, and in the interest of simplicity, those correlations are ignored. The uncertainty arising from systematic effects on the unfolding factors is typically much smaller than that arising from effects on the jet energy scale. An exception to this occurs in regions of phase space with low statistics in the Monte Carlo simulations, where statistical fluctuations in the unfolding become an important systematic effect.

The uncertainty in the jet energy scale has the largest impact on most of the measurements performed in this thesis. This is due to the complexity of jets and the interactions of particles within them in the calorimeter. Many studies have been performed that attempt to understand different aspects of these interactions [16,20,27,131,145,146,168,172,174–176]. In spite of the depth in such studies the uncertainty in the measurement of the jet energy remains at a few percent and will probably not become less than $\approx 1\%$ even within the lifetime of the ATLAS experiment. This combined with the steeply falling distributions typical of QCD jet production leads to large uncertainties in simple production cross section measurements. The impact of systematic effects on the jet energy scale is significantly reduced in the measurement of shapes and ratios, which can thus be performed with higher precision. The three components of the systematic uncertainty are discussed in detail in separate sections in this chapter, as well as the impact of each component on the different measurements performed in the analysis.

8.1 Systematic Uncertainties in the Measurement of the Luminosity

The measurement of the luminosity is directly affected by limitations in the calibration of the luminosity detectors and long-term variations of that calibration. The largest contribution to the systematic uncertainty in the luminosity measurement arises from the former. Figure 8.1 compares the mean number of interactions per bunch crossing, μ, measured by two luminosity algorithms as a function of time to the measurement provided by the algorithm used for the luminosity calculation. In the figure, LUCID_EventOR refers to the

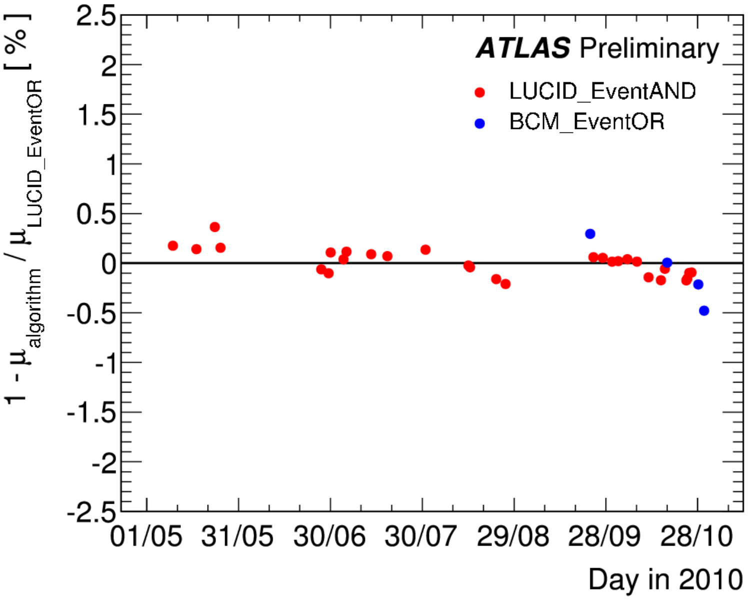

Figure 8.1: *Difference, in percentage of the average measured number of interactions per bunch crossing, between two luminosity algorithms and the luminosity algorithm used for the luminosity measurement as a function of time during the 2010 run [149].*

algorithm requiring at least one hit on either side of the LUCID detector, LUCID_EventAND refers to the algorithm requiring at least one hit on each side of the LUCID detector, and BCM_EventOR refers to the algorithm requiring at least one hit on either side of the BCM detector. Based on this study a 0.5% uncertainty is added to the luminosity measurement to account for long-term calibration changes in the LUCID detector. A similar study has been performed to determine the stability of the calibrations of the luminosity detectors with μ. While the two are correlated, the effect can be estimated conservatively, and it currently contributes an additional 0.5% to the total uncertainty in the luminosity measurements. The contributions of these uncertainties to the total systematic uncertainty in the luminosity are small.

The main uncertainty in the luminosity measurements comes from uncertainties in the calibrations performed during runs in which van der Meer scans happened. During the van der Meer scan, the visible cross section for a given detector can be calculated as

$$\sigma_{\text{vis}} = \mu_{\text{vis}}^{\text{MAX}} \frac{2\pi\Sigma_x\Sigma_y}{n_1 n_2}, \tag{8.1.1}$$

where $\mu_{\text{vis}}^{\text{MAX}}$ is the maximum average number of collisions observed with a given detector and algorithm when the beams collide head on, $\Sigma_{x,y}$ are the beam transverse sizes assuming that the beam density corresponds to an ellipsoid in the $[x, y]$ plane and $n_{1,2}$ are the number of protons per bunch in each beam. The understanding of all these quantities allows performing the measurement of the luminosity with any detector using Equation 6.1.1.

Several effects may impact the measurement of σ_{vis} through Equation 8.1.1. An obvious one is the measurement of $n_1 n_2$, performed through several devices, "current transformers", that measure the charge carried by each beam and are installed around the beam at eight locations along the LHC ring. The current transformers have a calibration that can drift from fill to fill due to several reasons, including environmental and mechanical changes over time. In addition, beam conditions can also impact the calibration. These effects have been studied using specific calibration fills [177,178]. Variations over different calibration fills are used to establish a 2.7% uncertainty in the measurement arising from the uncertainty in the calibration of the current transformers. In addition, variations from bunch to bunch and the presence of protons outside of the colliding bunches have been studied with the BCM and dedicated current transformers. Differences between the two also add to the uncertainty in the measurement of $n_1 n_2$, which is 3.1% in total and is the dominating uncertainty in the luminosity measurement.

Several effects impact the measurement of Σ_x and Σ_y. These include the knowledge of the orientation of the beam ellipsoid, the knowledge of the exact position of the beams as the scan progresses, the stability of the beam size within a scan, and the accuracy of the fits to the collision rates as a function of beam separation. The most important of all these is the first one, which yields a 0.9% uncertainty. All the others contribute less than 0.5% to the total uncertainty. After all these effects are considered, the total systematic uncertainty on the luminosity measurement is determined to be 3.4% [149].

8.2 Systematic Uncertainties in the Unfolding Corrections

The procedure of unfolding detector effects is justified as long as the Monte Carlo simulation is able to describe properly the relevant features of the detector that account for differences between the measured and the truth spectra. Many relevant features have been found to be in reasonable agreement between data and Monte Carlo simulation in the previous chapter. However, the level of agreement and its impact on the analysis needs to be quantified.

Bin migrations included in the unfolding corrections arise primarily from the following effects:

1. Trigger inefficiencies.
2. Pile-up interactions and impact of the pile-up rejection cuts.
3. Jet energy and position measurement fluctuations.
4. Splitting or merging of jets found nearby in the Monte Carlo truth.
5. Systematic effects on the jet energy measurement.

The impact of our understanding of the first four effects on the unfolding correction is quantified in this section. In addition, the uncertainties arising from limitations in the bin-by-bin unfolding technique are evaluated.

8.2.1 Trigger Inefficiencies

The unfolding correction accounts for inefficiencies in the trigger selection for events with selected jets near one another. The Monte Carlo simulation describes well this inefficiency as shown in Section 6.2.2. However, neither the ALPGEN nor the PYTHIA samples that can be used for the unfolding have a perfect prediction of the rate of such events. This, in turn, leads to an underestimate of the size of the effect. Given the small rate of such events, no correction is applied to the samples to perform a more accurate calculation of the unfolding factors. Instead, this effect is taken as a systematic effect on the calculation of the unfolding correction. A systematic uncertainty is estimated using the differences in the unfolding factors between the ALPGEN and PYTHIA samples, since the difference in the rate predicted with each sample is comparable to the difference between the ALPGEN sample and the data. The total difference in the unfolding factors between the two samples is taken as a systematic uncertainty. Therefore, the fact that both samples have a smaller rate of nearby selected jets than that observed in the data does not preclude the validity of the method used to calculate the impact of this systematic effect.

An additional inefficiency in the trigger affects jets falling in the region of $|\eta| \sim 1.5$ near the region of p_{T} where the η-inclusive efficiency curve reaches a plateau, as shown in Section 6.2.2. This inefficiency is not modeled by the Monte Carlo simulation. The inefficiency is, however, small and limited to a reduced part of the phase space for jet production. For this reason, it is again not corrected for, and an uncertainty is estimated for this systematic effect. In particular, an inefficiency that overestimates the size of the effect, based on the results shown in Section 6.2.2 and Appendix A, is added to the calculation of the unfolding factors. This inefficiency is 5% and affects the 2J10 and 3J10 triggers if a reconstructed jet in the event has $1.4 < |y| < 1.6$ and $60\text{ GeV} < p_{\mathrm{T}} < 80\text{ GeV}$. The inefficiency also affects the 2J15 and 2J30 triggers, if the jet falls in the same rapidity region and has $80\text{ GeV} < p_{\mathrm{T}} < 100\text{ GeV}$ and $110\text{ GeV} < p_{\mathrm{T}} < 130\text{ GeV}$. However, a larger difference exists between the simulation and the data in the description of the efficiency of these triggers. For this reason, the added inefficiency is 6% and 16%, respectively. The difference in the unfolding factors between the nominal ALPGEN sample and this modified ALPGEN sample with added inefficiencies is used to estimate the size of the effect. The absolute difference between the unfolding factors calculated in the two samples is considered as a systematic uncertainty, and it accounts for the discrepancy between data and Monte Carlo simulation targeted by this study.

8.2.2 Pile-up Interactions and Impact of Pile-up Rejection Cuts

Pile-up interactions impact the measurement in many ways. They shift the measured energy of jets upwards on average, and increase the fluctuations in that measurement. The average shift is corrected through an offset correction, described in Section 5.4.3, and the uncertainty in this correction is evaluated later in this chapter. The impact of the fluctuations in the measurements is controlled through a pile-up rejection cut on the quantity JVF. This cut reduces bin migrations due to jet energy measurement fluctuations induced by the presence of additional proton-proton interactions, thus making the value of the unfolding factors closer to 1. This is achieved through a balance between the efficiency of the cut to select jets that are present in the absence of pile-up interactions versus its rejection power for jets whose energy is significantly boosted by the interactions. This balance depends on many input parameters, such as the underlying p_{T} distributions or the vertex multiplicity in the event. This motivates the re-weighting of the vertex distribution in the Monte Carlo simulation and causes the dependence of the optimal cut on the observable and the number of reconstructed primary vertices in the event.

The re-weighting of the vertex distribution is needed for the balance achieved by the pile-up rejection cut to be the same in the simulation and in the data, so that bin migrations are well described in the simulation. The distribution of the number of reconstructed vertices is correlated to the instantaneous luminosity during the period when data was collected, and thus the rates at which QCD production happened. High event rates in ATLAS lead to the application of pre-scales on certain triggers, in such a way that the distribution of the number of reconstructed vertices depends on the pre-scales applied to different triggers, and is, thus, different for each trigger.

Figure 8.2 shows the distributions of the number of reconstructed primary vertices for events collected with the different triggers used in the analysis. The distribution for events collected in the two-jet inclusive multiplicity bin, with appropriate luminosity weights, is shown as a solid line for reference, since it is used to weight the distributions in the Monte Carlo simulations for the calculation of the unfolding factor central values. The distribution obtained using events collected with the 2J10 trigger is more strongly peaked at low-vertex multiplicities than the other distributions. The distribution obtained using events collected with the 3J10 trigger shows the highest average multiplicity. The 3J10 ran without pre-scales for the full data collection period, while the 2J10 trigger was heavily pre-scaled in the last runs. This implies that most of the data from the 2J10 trigger comes from early runs, where the instantaneous luminosity was lower, while the opposite is true for the 3J10 trigger.

The Monte Carlo simulation does not simulate varying pre-scales, and for this reason, events have a similar vertex multiplicity distribution independent of the trigger used. Since different triggers populate different regions of phase space, different re-weighting factors would be needed for each region of phase space to properly take into account pile-up bin migrations. The effect is, however, small and instead just treated as a systematic effect.

The vertex multiplicity distributions in events selected with the 2J10 and 3J10 triggers are used to calculate two alternative sets of vertex re-weighting factors. These re-weighting factors are, in turn, used to calculate two different sets of unfolding factors. These two

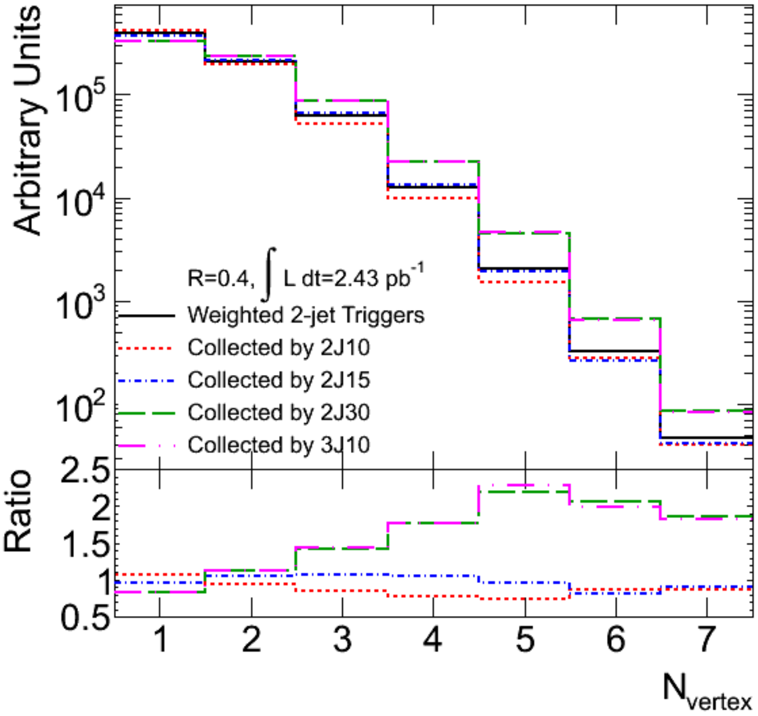

Figure 8.2: *Distribution of primary vertices reconstructed with all two-jet triggers weighted by luminosity (solid line) and the individual triggers used in the analysis. Events were selected according to selection cuts on jets reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.4$. The curves are normalized to the measured detector-level two-jet inclusive cross section. Ratios of the distributions for each trigger to the combined two-jet trigger distribution are shown at the bottom of the figure.*

vertex distributions are chosen because they represent a bound on the impact that pile-up interactions have on the values of the unfolding factors. For each bin, the unfolding factors obtained using these two alternative vertex re-weighting factors are compared to the nominal factor, and the largest of the two differences is taken as a systematic uncertainty.

The impact of pile-up effects on the measurements after the pile-up rejection cut are small. The Monte Carlo simulation has been shown to describe reasonably the properties of tracks within jets [166, 179] and general features of events with pile-up interactions [145]. Bin migrations after the cut are, thus, reasonably well described in the simulation. However, differences in track reconstruction efficiencies or jet reconstruction could cause changes in the size of these migrations.

Given how small the effect of the cut is (see Section 6.3.4) and the lack of more detailed studies of the impact of pile-up on the jet reconstruction, these potential effects are evaluated conservatively by varying the value of the JVF cut. In particular, the unfolding factors are calculated using two different values of the JVF cut, 0.5 and 0.9, that bracket the value used to calculate the nominal unfolding factors, 0.7. These two values are chosen to approximately span the set of optimal variables found for different observables in the data in Appendix C. This variation in the cut, of $\approx 20\%$, is much larger than any of the differences between data and Monte Carlo simulations observed in the available studies of tracking inside jets and is, thus, considered conservative. For each bin, the unfolding factors obtained using these two alternative cuts are compared to the nominal factor, and the largest of the two differences is taken as a systematic uncertainty to account for the potential effects arising from the limited knowledge of the impact of the pile-up rejection cut.

A cross-check has been performed studying the difference in the unfolding factors in Pythia Monte Carlo samples simulated with and without additional proton-proton interactions [163]. This cross-check shows that the difference is well within the systematic uncertainties in the unfolding factors, validating the treatment of pile-up effects described.

8.2.3 Jet Energy and Position Measurement Fluctuations

Fluctuations in the measurement of the jet kinematics cause bin migrations that are included in the unfolding corrections. Fluctuations on the jet position measurement are generally small, and their impact in the analysis is mostly limited to the $\Delta\eta$ and $\Delta\phi$ measurements. The description of such fluctuations has been tested using single-particle data and propagating the results to a jet environment using Monte Carlo simulations, as detailed in Ref. [180]. These studies show good agreement between the data and the Monte Carlo simulations in the size of the fluctuations within less than 10%. The angular resolution of the ATLAS calorimeters increases with jet p_{T}, and for jets above 60 GeV is at most 0.05 in η and 0.05 radians in ϕ [11]. In order to estimate the impact of limitations in the description of the angular resolution a sample is produced in which the jet position is smeared conservatively by 0.005 in each direction. The absolute difference between the unfolding factors calculated in this sample and those calculated in the nominal sample is considered as a systematic uncertainty.

Fluctuations in the measurement of the jet p_{T} are more important for the analysis, since the steep p_{T} distributions involved cause more bin migrations due to resolution effects. The agreement between Monte Carlo simulation and data has been studied using di-jet samples. Two independent methods have been used, as described in Ref. [169]: the di-jet balance and bi-sector methods. The two methods have different sensitivities to the radiation that can impact the balance in the di-jet system, and thus the determination of the resolution.

Figure 8.3 shows the fractional jet p_{T} resolution measured by the two methods using the Pythia sample and the data collected in the full 2010 run. The measurement is performed for jets reconstructed using the anti-k_t algorithm with resolution parameter $R = 0.6$ and with $|y| < 0.8$. The measurement is performed as a function of the average jet p_{T} of the di-jet system, but a one-to-one mapping exists between the measured resolution and the resolution as a function of jet p_{T}. A good agreement is found between the two methods used, and between data and Monte Carlo simulation in the kinematic range explored in this study, once the systematic uncertainties in the measurement ($\approx 10\%$) are considered [181]. This study has been performed up to $|y| = 2.8$ and similar results are obtained.

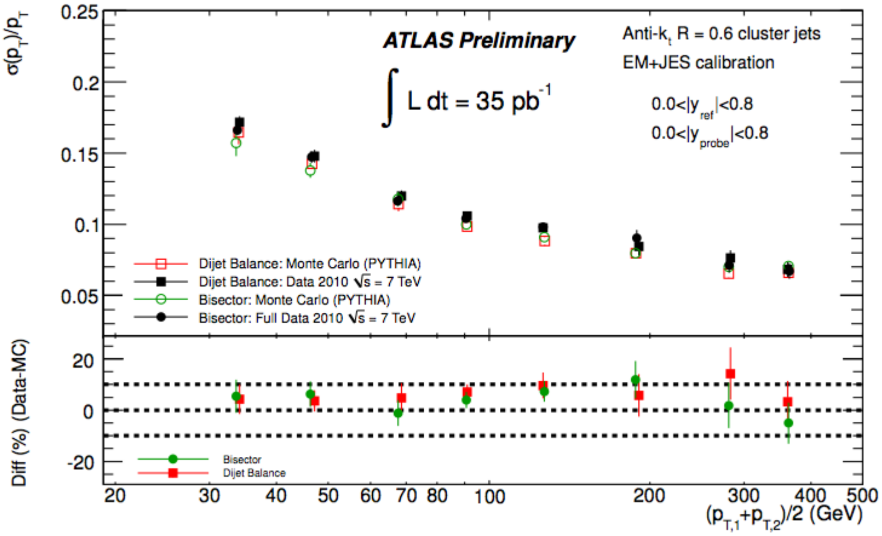

Figure 8.3: *Calorimeter jet p_{T} resolution measured using two different methods in data and the* Pythia *Monte Carlo sample [181]. Jets reconstructed using the anti-k_t algorithm with resolution parameter $R = 0.6$ and with $|y| < 0.8$ are used.*

The data and Monte Carlo simulations agree well within the systematic uncertainties in the measurement of the jet p_{T} resolution. These uncertainties, thus, characterize the potential limitations of the Monte Carlo simulations. An η-dependent number is used to smear the reconstructed p_{T} of jets in the ALPGEN sample. This number is taken as the uncertainty in the jet p_{T} resolution derived in the studies presented in Ref. [181], and varies from 5% in the barrel of the detector to 14% in the end-cap. The sample with the smeared jet p_{T} is used to calculate a new set of unfolding factors. The absolute difference between these factors and those calculated in the nominal sample is considered as a systematic uncertainty.

8.2.4 Splitting or Merging of Jets Found Nearby at the Particle Level in the Monte Carlo Simulation

Detector effects can cause a truth jet to be reconstructed as two jets in the calorimeter or the opposite. Due to the potentially complex structure of a jet, this can happen simply through angular or energy smearing of the energy depositions in the calorimeter. Monte Carlo simulations show that this does not happen often for jets in the range of p_{T} under study. This is expected, in light of the results shown in Figure 7.3. This figure showed that even in six-jet events the rate of selected jets that are reconstructed nearby is less than 20%. Given that only a fraction of these can arise from a reconstruction effect, it is clear that the rate with which a truth jet is reconstructed as two jets is small. Since the effects causing a truth jet to be reconstructed as two jets are the same as those causing two truth jets to be reconstructed as one jet, a similar rate is expected for the latter as for the former.

Of all possible mismatches in the jet reconstruction, the case in which a truth jet turns into two selected jets is most relevant for the analysis. Events in which a truth jet is reconstructed as two jets, but one is below the p_{T} threshold for jet selection do not change the jet multiplicity, and only impact the analysis through resolution effects on the energy measurement. Events with n truth jets in which two selected truth jets are reconstructed as one change the measurement in the n-jet bin only, but given the small rate of such occurrences the changes do not impact the analysis significantly. On the other hand, the effects of migrations caused by a truth jet being reconstructed as two jets can be large, due to the steeply falling spectrum of the jet multiplicity distribution.

Figure 8.4 shows the probability for jets produced in a multi-jet event to be reconstructed as two jets and the opposite effect as a function of jet multiplicity. These are calculated as a function of the particle-level inclusive jet multiplicity after kinematic cuts in both the

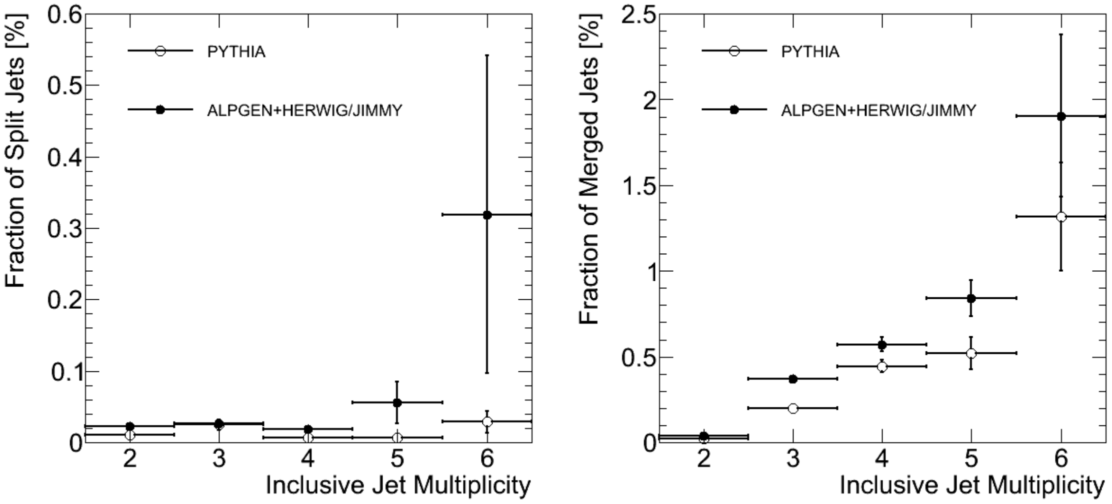

Figure 8.4: *Probability for jets produced in a multi-jet event to be reconstructed as two jets (left) and for two jets to be reconstructed as one jet (right) as a function of jet multiplicity. Results are shown as obtained in both the* PYTHIA *and the* ALPGEN *samples. Jets are built using the anti-k_t jet reconstruction algorithm with resolution parameter $R = 0.4$.*

ALPGEN and PYTHIA samples. A jet is split in the calorimeter if there are two selected jets reconstructed next to each other (within a ΔR in $[\eta, \phi]$ of twice the resolution parameter) and the nearest truth jet within $\Delta R = 0.25$ is the same for both jets. This matching cut in ΔR is at least 5 times the angular resolution of the calorimeter for selected jets, guaranteeing that reconstructed jets are correctly matched to their corresponding truth jet. A jet is merged in the calorimeter if there are two selected truth jets next to each other and they are matched, as before, to the same reconstructed jet. Merged jets are counted twice for this figure, once for each truth jet involved in the merging. The results are shown for jets built using the anti-k_t jet reconstruction algorithm with resolution parameter $R = 0.4$ but similar results are found when using $R = 0.6$.

The effect is small, as expected, and differences of the same order as the size of the effect are seen between the two samples studied. Given the small size of the effect, it is not given any further consideration. The difference in the unfolding factors between the two samples, already used in Section 8.2.1, is considered to account for the uncertainties related to the dependence of these effects on the physics model. Potential limitations in the reproduction of these effects by the detector simulation are also considered covered by this uncertainty. However, no evidence of mis-modeling of these effects has been observed in detailed studies of the properties of jets reconstructed near other jets.

8.2.5 Shape of the Truth Distributions

The shape of the truth distributions that are measured strongly affect the size of bin-to-bin migrations. An optimal unfolding algorithm would take the measurement and all the information about the different effects that cause bin migrations, and it would extract the truth distribution without any bias. However, the simplicity of the bin-by-bin unfolding method comes at the price of ignoring all correlations between bins. This results in an inherent bias in the unfolded distribution towards the distribution used to calculate the unfolding factors [158], i.e. the distribution in the event simulation.

In this analysis, the bias is ignored, and instead a systematic uncertainty is calculated to cover for all possible shapes of the input distributions. Shapes of distributions are well constrained by the measurements performed in the analysis. In particular, the shapes of the distributions are limited to those allowed within the systematic uncertainties in the measurements. Since these uncertainties are not known, a priori, and they include uncertainties arising from the effect of the shape bias, the uncertainty is estimated for this study. The assumption that this estimate is accurate can then be verified once the measurement is performed and all uncertainties have been properly accounted for.

Since measurements are binned in inclusive jet multiplicity, the shape of the inclusive jet multiplicity distribution impacts the unfolding factors for all measurements. In order to study this impact, the truth multiplicity distribution is weighted using the following weighting function:

$$w = a^{N_{\text{jets}}}, \tag{8.2.1}$$

where N_{jets} is the number of jets reconstructed at the particle level. The value of a is taken to be 1.10 to simulate a less steep multiplicity distribution and 0.95 to simulate a steeper distribution. Events that do not pass selection cuts at the particle level receive a weight of 1. Unfolding factors are calculated using these two alternative samples. For each bin, the unfolding factors obtained with these samples are compared to the nominal factor, and the largest of the two differences is taken as a systematic uncertainty to account for the bias in the unfolding method to the input jet multiplicity distribution.

The shapes of kinematic distributions also impact the calculation of the unfolding factors for each distribution. A similar procedure is used to estimate the uncertainty in those factors arising from the bias in the unfolding method. The following weighting function is used:

$$w = a^{x/[100\ \text{GeV}]}, \tag{8.2.2}$$

where x refers to the kinematic variable under study, whether it is the leading, or sub-leading jet p_{T}, the event H_{T} or the event ${H_{\mathrm{T}}}^{(2)}$, as defined in Section 6.3.1. As before, two values of a, $a = 0.95,\ 1.05$, are used to cover the two extreme shapes that the truth distributions can have, based on the systematic uncertainties in the measurement. Unfolding factors are calculated using these two alternative samples. For each bin, the unfolding factors obtained with these samples are compared to the nominal factor, and the largest of the two differences is taken as a systematic uncertainty, in addition to that arising from the bias coming from the shape of the jet multiplicity distribution.

Figure 8.5 shows the re-weighting functions used for the jet multiplicity distribution and the kinematic distributions. The corresponding measurements are presented in Chapter 10 with their associated systematic uncertainties. Those results show that these re-weighting factors safely cover the range of shapes consistent with the measurement.

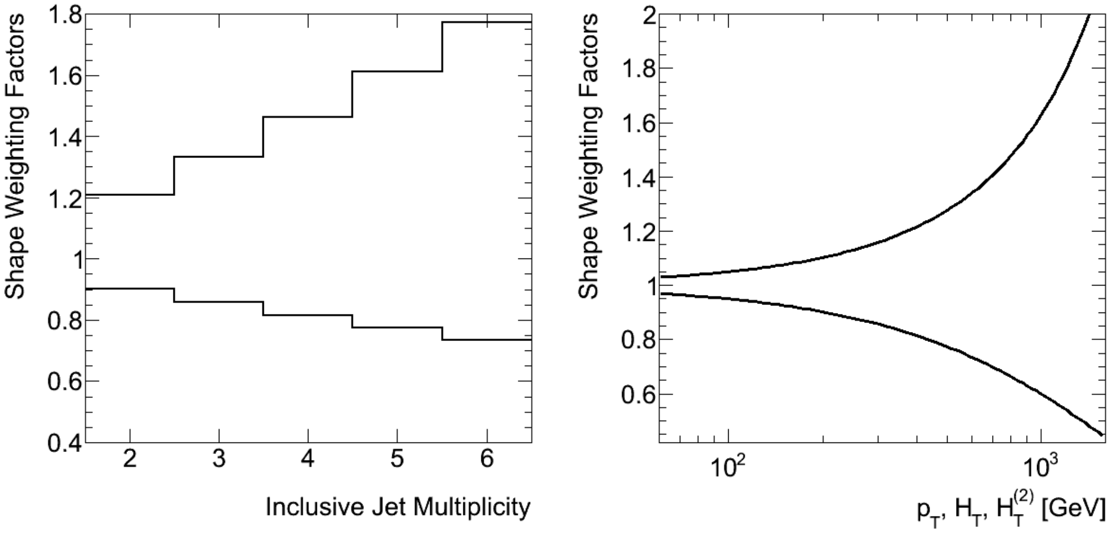

Figure 8.5: *Re-weighting functions used to modify the jet multiplicity distribution (left) and the kinematic distributions (right) to associate a systematic uncertainty due to the bias of the bin-by-bin unfolding method towards the truth distribution in the simulation.*

The $\Delta\eta$ and $\Delta\phi$ distributions studied are affected only weakly by the two re-weighting factors above. In fact, only the jet multiplicity re-weighting is used to associate a systematic uncertainty from this effect, since the re-weighting of the kinematic distributions does not achieve the purpose of studying the bias of the bin-by-bin unfolding method to the input $\Delta\eta$ and $\Delta\phi$ shapes.

As shown in the next section, the uncertainty in the unfolding factors applied to these distributions is dominated by the statistical uncertainties in the Monte Carlo simulations, which are large. This implies that a change in the input shape would not have a large impact in the systematic uncertainties of the unfolding factors for these distributions. In addition, Chapter 10 will show that there are very significant differences in some of the shapes of the distributions predicted by ALPGEN and PYTHIA. This implies that differences in the unfolding factors calculated with ALPGEN and PYTHIA are at least partially caused by the different biases in the bin-by-bin unfolding factors in the two samples. For these two reasons, no additional uncertainty is associated to the limitations of the bin-by-bin unfolding for the measurements of the $\Delta\eta$ and $\Delta\phi$ distributions.

8.2.6 Combination of Uncertainties

Table 8.1 presents a summary of the samples used in the studies presented before in this chapter, and the effect or effects they address. Most of those studies address specific issues that are independent from each other and not correlated. The exception is the use of differences in the unfolding factors calculated with the PYTHIA and ALPGEN samples, which addresses several issues at the same time. Certain cross-checks performed in Ref. [163] verify that some of these individual effects do not cancel each other. However, some cancelations cannot be completely discarded.

In addition, the differences between the PYTHIA and ALPGEN samples are also partially caused by differences in the shapes of the truth distributions. This causes the effect described in the previous section to be double-counted. Since the re-weighting functions that are used span larger shape differences than those found between PYTHIA and ALPGEN, the effect is not fully double-counted. Furthermore, the contribution of shape uncertainties to the total systematic uncertainty is small, implying that fully accounting for the correlations between the two studies would change negligibly the estimate of the systematic

Sample	Systematic Effects
Pythia	Trigger Topological Inefficiencies Split/merged Jet Rates Shape Differences (angular measurements) JES of non-isolated jets (described in the next section)
Alpgen($\lvert\eta\rvert \sim 1.5$ inefficiency)	Trigger $\lvert\eta\rvert \sim 1.5$ inefficiency
Alpgen(JVF$= 0.5, 0.9$)	Modeling of pile-up rejection cut variables
Alpgen(2J10,3J10 vertex weights)	Impact of pile-up for different triggers
Alpgen(jet angular position smearing)	Jet angular position resolution
Alpgen(jet p_{T} smearing)	Jet p_{T} resolution
Alpgen(inclusive multiplicity weights)	Unfolding bias to input distribution
Alpgen(kinematic distribution weights)	Unfolding bias to input distribution

Table 8.1: *Samples used in the studies of systematic effects on the unfolding corrections and effects they address.*

uncertainties in the unfolding factors. For these reasons, such correlations are ignored, and the uncertainties estimated using all the studies presented in previous sections are added in quadrature in order to obtain an estimate of the systematic uncertainty in the unfolding.

The statistical uncertainties in the unfolding factors obtained with the Alpgen sample are not negligible in certain bins. Since this is a limitation of the unfolding procedure, they are added to the systematic uncertainties in the unfolding factors. A million pseudo-experiments are used for this purpose, in order to consider that statistical uncertainties are uncorrelated between bins. The use of pseudo-experiments also allows to properly account for the statistical uncertainties in the Pythia sample. For this purpose, the difference between Pythia and Alpgen is fluctuated for each pseudo-experiment, according to the statistical fluctuations in the unfolding factors calculated in the Pythia sample. Such procedure is not needed when comparing the different Alpgen samples used to study all the other systematic effects, since the statistical uncertainties between the samples are nearly fully correlated.

In the rest of this section the unfolding factors and associated systematic uncertainties for representative observables studied in the analysis are discussed. The results for all other observables can be found in Appendix C for completeness. Figure 8.6 shows the unfolding factors and the associated systematic uncertainties for the inclusive jet multiplicity mea-

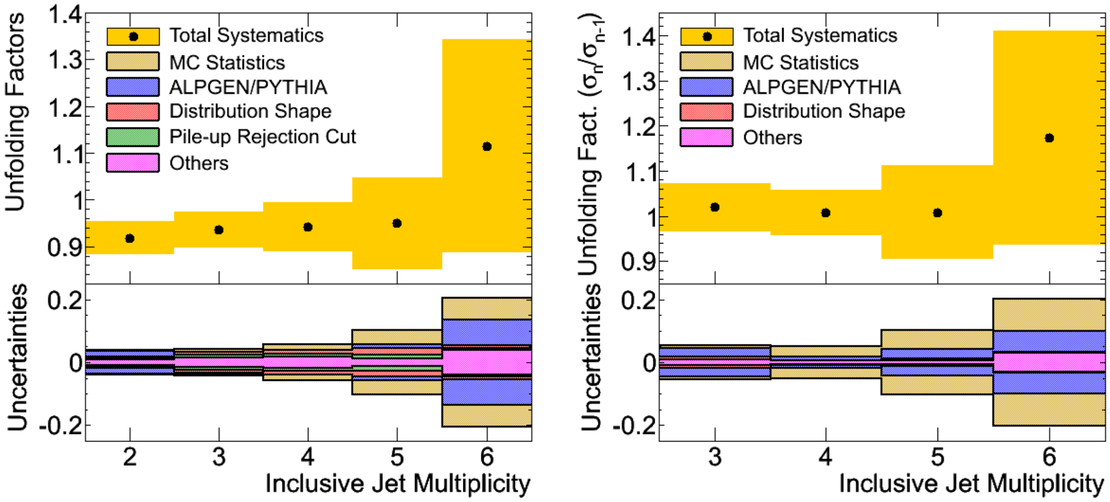

Figure 8.6: *Unfolding factors for the inclusive jet multiplicity cross section (left) and n-to-$(n-1)$-jet cross-section ratios (right) measurements and their associated uncertainties. The main effects contributing to the systematic uncertainties are shown at the bottom of each figure. Jets reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.4$ were used for these results.*

surement and the n-to-$(n-1)$-jet cross-section ratios. The main effects contributing to the systematic uncertainties are shown at the bottom of each figure as they are combined to build the total systematic uncertainty. The band corresponding to the statistical uncertainties in the Monte Carlo simulation includes effects caused by the limited statistics in both the Alpgen and Pythia samples. The unfolding factors are close to 1, as discussed in the previous chapter. The differences between the Alpgen and Pythia samples and the statistical uncertainties in the simulation dominate the systematic uncertainties in the unfolding factors.

The systematic uncertainties in the measurement of the cross section as a function of the inclusive jet multiplicity are dominated by the uncertainties in the jet energy scale. However, when ratios of cross sections are taken, the impact of unfolding uncertainties become more important. Figure 8.7 shows the unfolding factors and their associated uncertainties for the measurement of the three-to-two-jet cross-section ratio as a function of the ${H_{\mathrm{T}}}^{(2)}$ and the leading jet p_{T} in the event. The unfolding factors are very close to 1. The uncertainties in these factors are the dominating systematic uncertainties in these measurements. The statistical uncertainties in the simulation contribute less significantly to the systematic uncertainties in the unfolding than in the previous figure. The difference between the

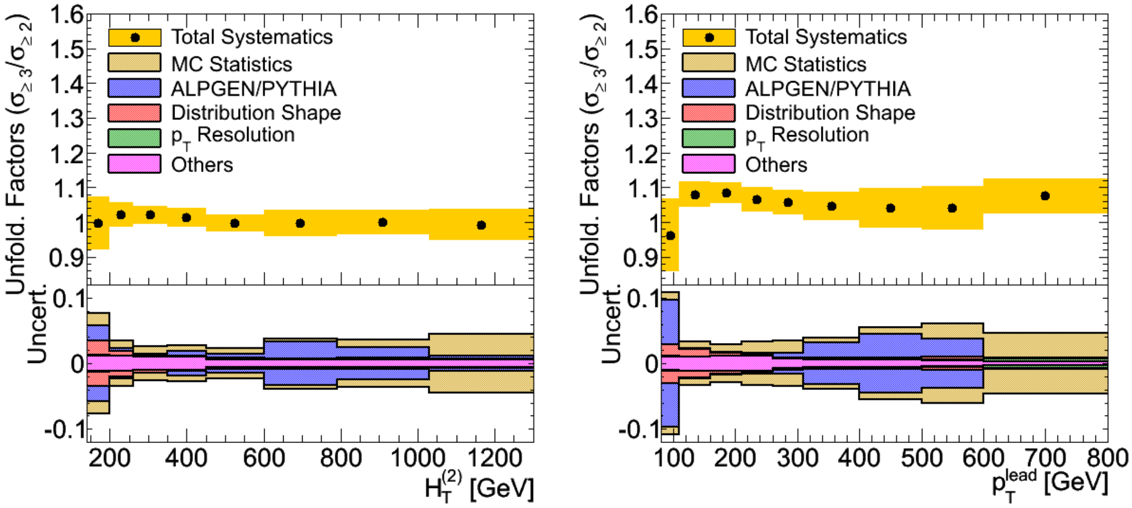

Figure 8.7: *Unfolding factors of three-to-two-jet cross-section ratio measurements as a function of* ${H_T}^{(2)}$ *(left) and the leading jet* p_T *(right) in the event and their associated uncertainties. The main effects contributing to the systematic uncertainties are shown at the bottom of each figure. Jets reconstructed with the anti-*k_t *algorithm with resolution parameter* $R = 0.6$ *were used for these results.*

ALPGEN and PYTHIA samples remains a large contributor, even though at low values of p_T and ${H_T}^{(2)}$ other effects, such as the shape of the input distributions, are also important.

The unfolding factors for the measurements of angular distributions have their own defining characteristics. Figure 8.8 shows the unfolding factors for the measurement of the shape of the distributions of $|\Delta\phi|$ and $|\Delta\eta|$ between the two leading jets in events with at least three selected jets. The systematic uncertainties in the factors, as well as the main effects contributing to those uncertainties are shown as in previous figures. The size of the unfolding factors has already been discussed in the previous chapter. The systematic uncertainties are large, and dominated by differences between the ALPGEN and PYTHIA samples and the statistical uncertainties in the unfolding factors. This is not surprising, since these measurements probe regions of phase space in which the inefficiency in the trigger and the rate of split and merged jets are important. In addition, shape biases are part of the differences observed between the two samples, further contributing to the differences observed between the two sets of unfolding factors.

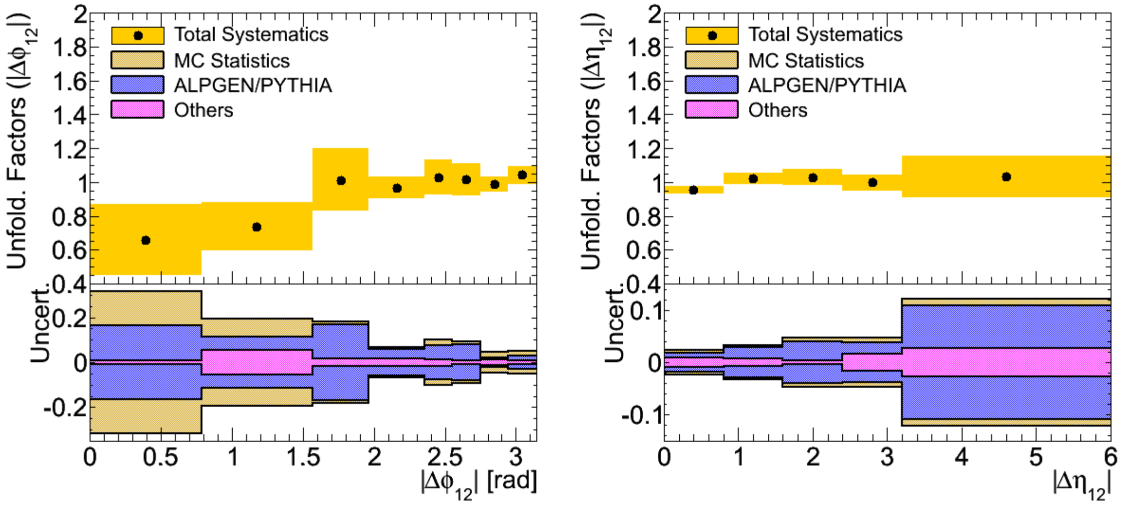

Figure 8.8: *Unfolding factors for the measurement of the shape of the distributions of $|\Delta\phi|$ (left) and $|\Delta\eta|$ (right) between the two leading jets in events with at least three selected jets and their associated uncertainties. The main effects contributing to the systematic uncertainties are shown at the bottom of each figure. Jets reconstructed with the anti-k_t algorithm with resolution parameter* $R = 0.4$ *were used for these results.*

8.3 Systematic Uncertainties in the Jet Energy Scale

A precise measurement of the jet energy and p_{T} is important for many of the measurements performed in the analysis. Limitations in that understanding give rise to a systematic uncertainty in the jet energy scale. This uncertainty can lead to large uncertainties in the measurements due to the steeply decaying p_{T} spectra characterizing QCD jet production. For this reason, many studies are needed to understand the jet energy scale in detail, both in data and Monte Carlo simulations.

The calculation of the uncertainty in the jet energy scale is performed here first through the study of a well-defined jet sample: the inclusive sample of isolated jets. In this context, inclusive refers to the fact that jets are studied independently of the other activity in the event. That sample is overwhelmingly built by jets produced through QCD processes, primarily as di-jet events. Isolated in this context refers to jets that have little or no activity nearby. This is more precisely defined as jets with no other reconstructed jet of $p_{\mathrm{T}} > 7$ GeV before calibration within a ΔR in $[\eta, \phi]$ of 2.5 times the resolution parameter of the jet algorithm used in the event reconstruction.

The study of isolated jets from the inclusive sample suffices for many analyses in which such jets characterize well the jets found in the final state. However, the isolation criterion is sensitive to the topology and production process of a given event. In particular, as the jet multiplicity in the event increases, the probability of a jet being isolated decreases. The properties of non-isolated jets are different from those of isolated jets. For these reasons, additional studies need to complement those performed in the inclusive sample of isolated jets to understand the precision of the energy measurement in multi-jet final states.

The complexity of the calorimeter response and its geometry implies that jets with different properties are seen to carry different energies on average [13]. This results in a degraded resolution for most purposes. However, this also implies that the response of jets depends on their physical origin. In a multi-jet analysis, the most important consequence of this effect is that the calorimeter has a different response for jets originating from light quarks and gluons. This leads to additional systematic uncertainties in the measurement of the jet energy. These uncertainties arise from the differences of light-quark/gluon jets from the admixture found in the inclusive sample or from the imperfect knowledge about the flavor composition of the final state of interest.

8.3.1 Systematic Uncertainties in Jets in the Inclusive Sample

The baseline uncertainty for isolated jets in the inclusive sample is obtained using charged hadrons in data. This study has been described in detail in Refs. [151, 172]. The data used in the study were collected using the MBTS triggers and belong to the very early data collection period in 2010. Tracks are used as an independent measurement of the charged hadron kinematics and matched to energy deposited in calorimeter clusters and within a $\Delta R = 0.2$ in $[\eta, \phi]$ of the impact point of the track in the calorimeter. To guarantee that the energy deposited in the calorimeter is not contaminated by energy from other charged hadrons, the track is required to have no other track reconstructed within a $\Delta R = 0.4$ in $[\eta, \phi]$, or be isolated. Only isolated tracks with p_{T} between 2 and 20 GeV are used in this study, even though a similar study using data collected in 2009 has been used to access lower-track p_{T}s [131].

The isolation cut does not guarantee that energy from neutral particles does not contaminate the measurement of the energy of the charged particle in the calorimeter. The amount of energy deposited by neutral particles in the EM calorimeter is estimated in data using charged hadrons that behave like minimum ionizing particles and thus deposit very little energy in the EM calorimeter. Such particles are identified through the large energy depositions in the hadronic calorimeter since this is where the hadron starts showering. The energy deposited around the impact point of hadrons selected in this way is used to estimate the background contribution and subtracted from the energy measured around the impact point of other charged hadrons. The amount of energy deposited by neutral hadrons in the hadronic calorimeter has also been estimated, and found to be small [168]. For this reason, it is simply ignored in the subtraction.

Figure 8.9 shows the average value of the ratio of the energy measured in the calorimeter

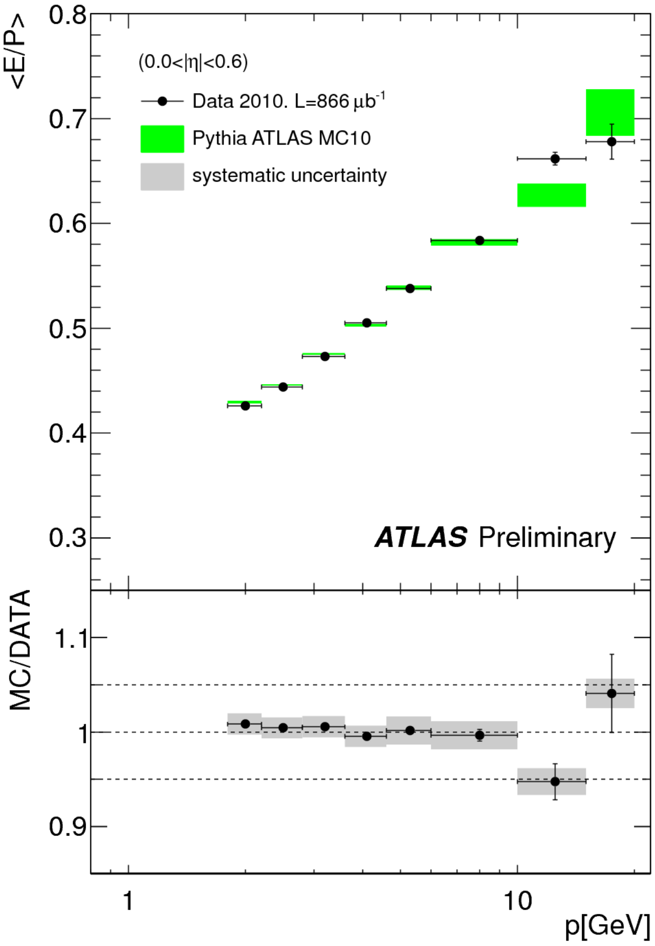

Figure 8.9: *Average of the ratio of calorimeter clustered energy to track momentum for charged hadrons after neutral particle background subtraction in data and Monte Carlo simulation [172]. The bottom of the figure shows the ratio of the results in Monte Carlo simulation to the results in data, together with the systematic uncertainty in the measurement, as a gray error band. Tracks with* $|\eta| < 0.6$ *were used in this figure.*

to the momentum of the track as a function of track momentum. The results are shown as obtained in data and in a Pythia sample generated with the AMBT1 tune for tracks reconstructed in the barrel region of the calorimeter. The bottom of the figure shows the ratio of the simulation to the data, and the systematic uncertainties in the measurement as an error band. The systematic uncertainties arise from the track selection, the uncertainty in the track momentum [164] and limitations inherent to the background subtraction method. A good agreement is observed between the results in the simulation and the data, within the systematic uncertainties in the measurement.

This systematic uncertainty, however, cannot be directly translated into a systematic uncertainty in the measured jet energy, since jets are not only composed of charged hadrons of momenta below 20 GeV. Also, variations in the amount of material in front of the calorimeter impact the frequency with which particles in a jet reach the calorimeter. In addition, the response of the calorimeter can be different in the context of a jet, in particular due to noise suppression, which may remove more signal in an isolated environment compared to the busy environment inside a jet.

The last of these effects has been investigated removing the noise suppression mechanism in the study above [168] and studying the dependence of the E/p ratio on the isolation cut [171]. Both studies show that this effect is negligible. The impact of material in front of the calorimeter has been studied using dedicated Monte Carlo simulations with increased amounts of material, and an uncertainty is assigned for this effect.

The calorimeter response for high-energy particles has been studied in test beams that reproduce the calorimeter geometry [125–127] and a similar agreement to results shown in Figure 8.9 has been found. A similar response to protons and anti-protons is also found using test beams and collision data [171].

Differences in the calorimeter conditions in the test beam and the final detector configuration are the only additional effects contributing to the total uncertainty in the calorimeter response to these particles. Limitations in the detector description for high-momentum hadrons are studied using dedicated Monte Carlo simulations with additional material in the ID and in the regions in between calorimeters. These simulations are used to assign an uncertainty due to the differences between the detector configuration in test beams and the final detector configuration. This uncertainty is shown independently of other uncertainties in this single-particle analysis in the figures shown later in this section.

No data exists from test beams on the calorimeter response to neutral hadrons. On average $\approx 10\%$ of the jet energy is carried by these particles, in particular K^0_S, K^0_L and (anti-)neutrons. Most K^0_S decay to pions before reaching the calorimeter, so the uncertainty in the charged hadron study above suffices to characterize the uncertainty in the calorimeter response to K^0_S. For (anti-)neutrons, a 5–10% uncertainty is assumed, depending on their momenta, based on studies with GEANT4 [123]. Similar studies allow assigning an additional 20% uncertainty to the simulated calorimeter response to K^0_L.

All the studies above are combined using the particle identities and kinematics inside jets predicted by the Pythia simulation with the AMBT1 tune to obtain an estimate of the jet energy scale uncertainty based on these single-particle studies. This estimate is $\approx 2\%$ for low-p_{T} jets and $\approx 4\%$ for high-p_{T} jets.

The single-particle uncertainty is only calculated for jets falling in the barrel region of the detector, $|\eta| < 0.8$, due to the limitations of the detector geometry available for test beams. Therefore, these studies need to be complemented by in situ measurements of the jet response as a function of η. Such detailed study has been performed using di-jet events in Refs. [151,174]. In essence, this study uses the p_{T} balance in a di-jet system to study the relative response of a jet in any η region with respect to the response of a jet falling in the barrel region, where the measurement is well understood.

Figure 8.10 shows the results of such a study for a particular bin of the average p_{T} of the di-jet system. The average p_{T} of the di-jet system is used for the same reasons as those discussed in Section 8.2.3, and the trivial mapping can be used to arrive at the jet p_{T} bin probed by the study. The results are shown as obtained using data and several Monte Carlo simulations. The data results fall somewhere between the predictions from the different Monte Carlo simulations. The uncertainty is thus estimated as the standard deviation of the predictions from the different Monte Carlo simulations. This additional systematic uncertainty is added in quadrature to that obtained in the single-particle analysis, and is small for jets used in the analysis, being most important for jets with low p_{T} and high $|\eta|$.

The single-particle analysis assumes a specific particle composition and spectrum: that predicted by the Pythia sample with the AMBT1 tune. Different event simulations predict different particle spectra within a jet, and that can affect the response. Since no model is

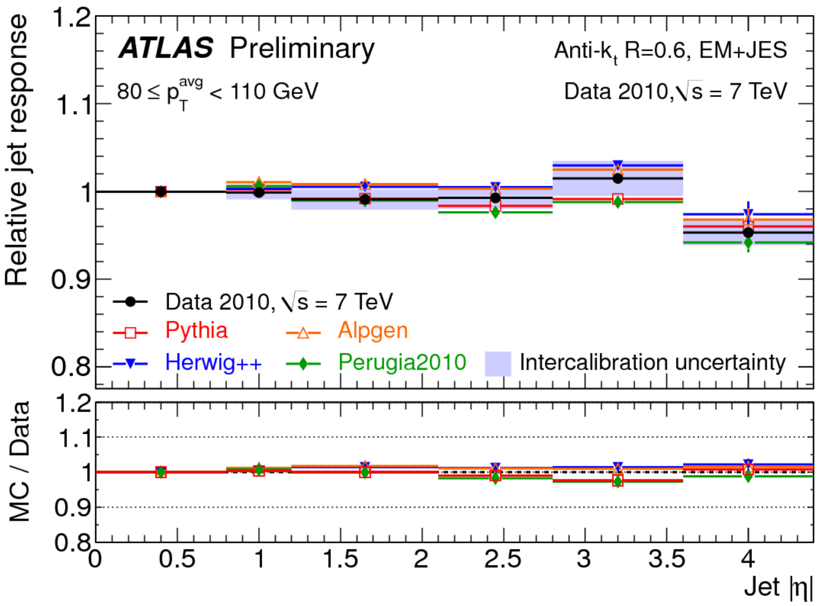

Figure 8.10: *Response of the calorimeter to jets as a function of jet $|\eta|$ relative to a jet falling in the barrel as measured in data and Monte Carlo simulations [151]. The uncertainty in the measurement is shown as an error band. Results for several Monte Carlo simulations are also shown. The label "*Pythia*" refers to the* Pythia *event generator using the AMBT1 tune. Jets reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.6$ are used in this figure.*

preferred over another, such shifts in the response are included in the systematic uncertainty in the jet energy scale as an additional uncorrelated component.

Figure 8.11 shows the calorimeter p_T response to jets after the calibration is applied as a function of jet p_T in the Pythia and Alpgen samples and a Pythia sample simulated with the Perugia2010 [61] tune. Differences in the response are not negligible between samples, particularly for the lowest p_T bins shown. These differences are taken as additional uncorrelated components to the systematic uncertainty in the jet energy scale. In the multi-jet analysis, these differences are also accounted for when considering the difference in the unfolding factors calculated in the Alpgen and Pythia samples. However, the difference in the response in the two samples contributes negligibly to the total uncertainty in the jet energy scale for jets with $p_T > 60$ GeV. For this reason, this small double counting can be tolerated.

A few other effects are considered in the calculation of the final jet energy scale uncertainty as detailed in Ref. [151]. The noise description in the Monte Carlo simulation has been compared to the data. The differences found are used to estimate the impact of

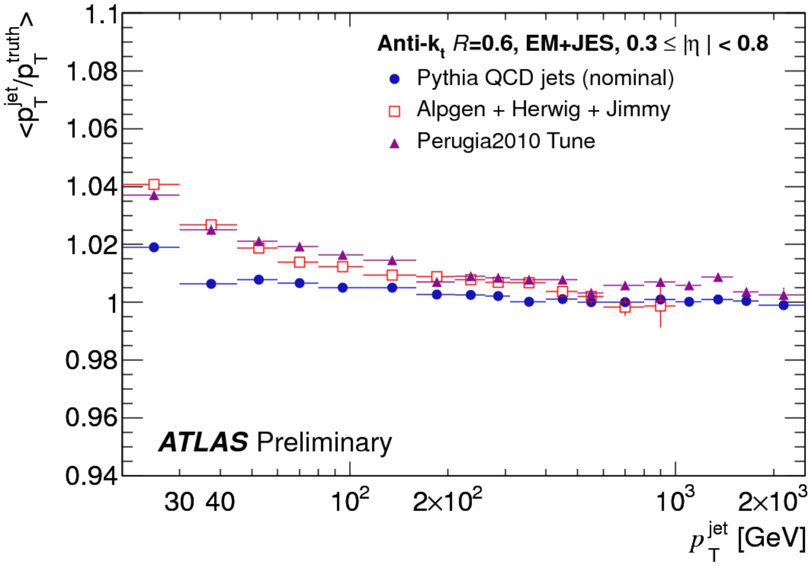

Figure 8.11: *Calorimeter p_T response to jets after the calibration is applied as a function of jet p_T in the* Pythia *and* Alpgen *samples and a* Pythia *sample simulated with the Perugia2010 tune [151]. Jets reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.6$ and with $0.3 < |\eta| < 0.8$ are used in this figure.*

this mis-modeling in the calibration factors in the Monte Carlo simulation, and added in quadrature to the other systematic uncertainties.

An additional "non-closure" term is added to the uncertainty in the jet energy scale. This term accounts for limitations in the calibration procedure, which does not fully correct the jet mass as described in Section 5.4.3. The result of this is that the jet energy and p_T responses cannot be simultaneously 1. Since the physics causing this effect, i.e. the description of the jet mass, is not currently well constrained, an additional uncorrelated systematic uncertainty is added. The value of this systematic uncertainty is calculated as the difference from 1 in the p_T or energy response. For each bin, the largest of the two differences is taken as the systematic uncertainty. The contribution of this component to the total uncertainty in the jet energy scale is small for jets with $p_T > 60$ GeV.

Figure 8.12 shows the different components contributing to the systematic uncertainty in the jet energy scale as a function of jet p_T. Results are shown for jets reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.6$, but similar results are obtained for jets reconstructed using $R = 0.4$. The uncertainties are shown for jets with $0.3 \leq |\eta| < 0.8$, but have been calculated for jets up to $|\eta| = 4.5$ [151]. The addition in quadrature of all

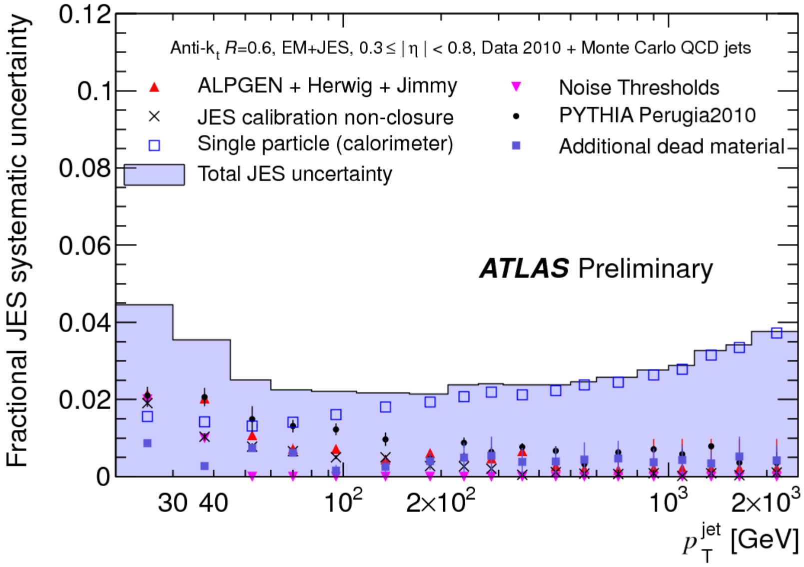

Figure 8.12: *Different components contributing to the jet energy scale systematic uncertainty as a function of jet* p_T *for jets reconstructed with the anti-*k_t *algorithm with resolution parameter* $R = 0.6$ *with* $0.3 \leq |\eta| < 0.8$ *[151]. The addition in quadrature of all the contributions is shown as a filled histogram.*

these contributions is also shown as a filled histogram. For jets with $p_T > 60$ GeV the largest contributions to the uncertainty come from the single-particle analysis and differences in the calorimeter response to jets simulated using PYTHIA with the AMBT1 and Perugia2010 tunes.

The total systematic uncertainty shown in this figure is appropriate for jets in events with no additional proton-proton interactions. In the presence of pile-up, the application of the offset correction calls for a systematic uncertainty associated to that correction. The correction is calculated using data, as described in Section 5.4.3. Uncertainties in the correction are thus associated to the methodology of the calculation and the approximations made in the application of the correction. Three effects are considered that fall in these categories: trigger selection, variations with jet tower multiplicity and variations observed in the validation of the method using tracks. These effects and their estimation are described in detail in Ref. [145].

The first effect is studied looking at the change in the offset correction estimated using events selected with different triggers. It accounts for the dependence of the offset correction on the different environments where it can be applied. Different environments change the interplay between the noise suppression and the energy depositions from particles created in pile-up interactions, and thus the size of the offset correction. The second effect accounts for a simplification in the offset correction algorithm, which estimates the number of towers in a jet using mostly low-p_{T} jets. The third effect accounts for differences found when the offset correction is calculated using an alternative method that considers correlations between the momenta of tracks associated to the jet and the energy of the jet. The differences between the corrections obtained using this or the default method are not understood and are taken as a systematic uncertainty in the correction.

Figure 8.13 shows the uncertainty in the jet energy scale arising from the uncertainty in the offset correction as a function of jet p_{T} for events with two reconstructed primary vertices. The results are shown for jets reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.6$ falling in three different $|\eta|$ bins. The uncertainty is most significant at low p_{T} and increases with the pseudorapidity of the jet. The uncertainty decreases with jet p_{T} and for jets used in the analysis it is at most 0.6%. The offset correction and its uncertainty increase linearly with the number of reconstructed primary vertices in the event. However, given the vertex distributions for events selected in the

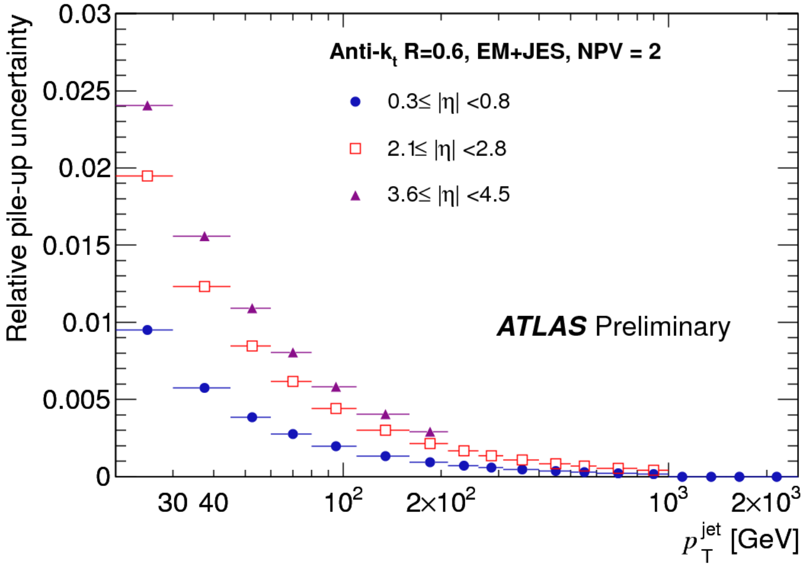

Figure 8.13: *Uncertainty in the jet energy scale arising from the uncertainty in the pile-up offset correction as a function of jet p_{T} for events with two reconstructed primary vertices [145]. The results are shown for jets reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.6$ falling in three different $|\eta|$ bins.*

analysis, the contribution of this term to the overall jet energy scale systematic uncertainty remains small for the vast majority of events.

The jet energy scale has been validated with many techniques using data. These techniques are complementary and they use correlations between the track p_{T} and the jet p_{T} [182], the photon and the jet p_{T} in γ+jet events [176], or the high-p_{T} jet and other jets produced in the opposite detector hemisphere in multi-jet events [175]. Results from these studies are shown in Figure 8.14. All the studies coincide in the agreement between the data and detector simulations within the uncertainties in the jet energy scale calculated as described in this section, and shown as an error band in the figure. Each of these studies has its own systematic uncertainties, which are larger than those presented above. For this reason, they are not used to calculate the systematic uncertainty in the jet energy scale, and instead are simply interesting as independent cross-checks of the systematic uncertainty in the inclusive sample of isolated jets.

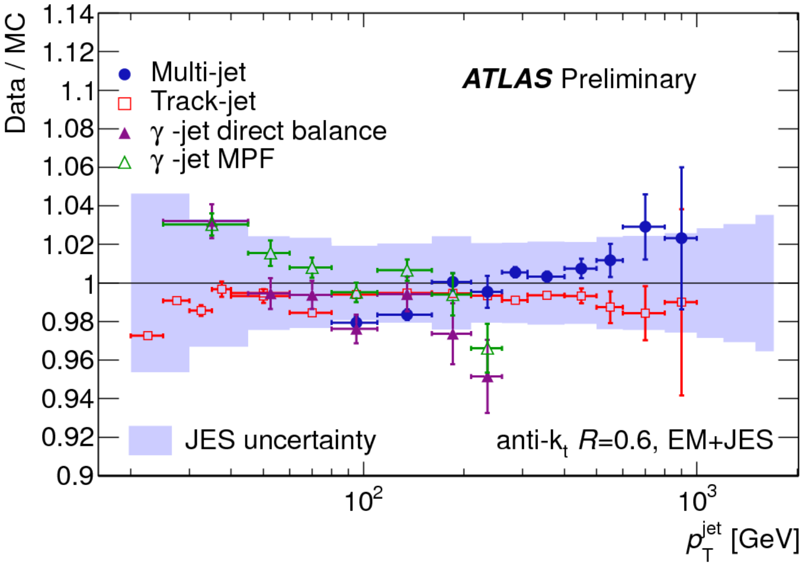

Figure 8.14: *Ratio of results obtained in data and the* Pythia *Monte Carlo simulation with the AMBT1 tune in studies probing the jet energy scale through different methods [151]. The systematic uncertainty calculated as described in this section is shown as an error band. These results correspond to jets reconstructed using the anti-k_t algorithm with resolution parameter $R = 0.6$ and $|\eta| < 1.2$.*

8.3.2 Systematic Uncertainties in Jets with Nearby Calorimeter Activity

In this section, non-isolated jets, or jets with nearby calorimeter activity, are defined as jets for which ΔR_{min} is less than 2.5 times the resolution parameter of the jet algorithm used in the jet reconstruction. ΔR_{min} is defined as the distance in $[\eta, \phi]$ to the closest reconstructed jet of ${p_{\mathrm{T}}}^{\mathrm{EM}} > 7$ GeV. Non-isolated jets are excluded from the studies in the previous section, because their properties are somewhat different due to the nearby energy depositions. While few jets fall on this category in the inclusive jet sample, they become important when studying multi-jet events.

Figure 8.15 shows the fraction of non-isolated jets as a function of inclusive jet multiplicity for jets selected in the analysis. Results are shown for detector-level jets in the data and the Alpgen and Pythia Monte Carlo simulations. As expected the fraction of

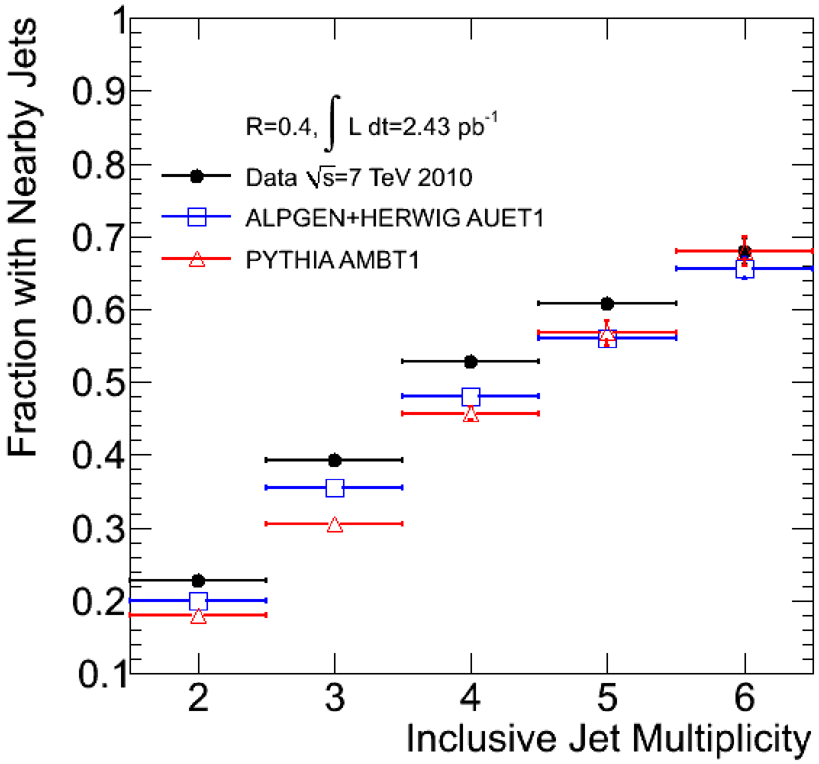

Figure 8.15: *Fraction of selected jets in the analysis with nearby activity in the calorimeter as a function of inclusive jet multiplicity. Results are shown for jets reconstructed with the anti-k_t jet algorithm with resolution parameter $R = 0.4$ in data and the* Alpgen *and* Pythia *Monte Carlo simulations.*

non-isolated jets is relatively small for two-jet events, but increases quickly to be dominant already in four-jet events. The trend is reproduced by both simulations, even though both predict a lower fraction than that observed in the data. The results in the Alpgen sample are closer to the data than those in the Pythia sample.

The understanding of the jet energy scale of non-isolated jets is, thus, critical to understand multi-jet events. Monte Carlo simulations show that the properties of non-isolated jets are different than those of isolated jets. Differences become larger as the nearby reconstructed jet becomes closer to the selected jet. This is demonstrated in Figure 8.16 where the calorimeter jet p_{T} response as a function of jet p_{T} is shown for different values of ΔR_{min}. This response is calculated using a Pythia sample generated with the AMBT1

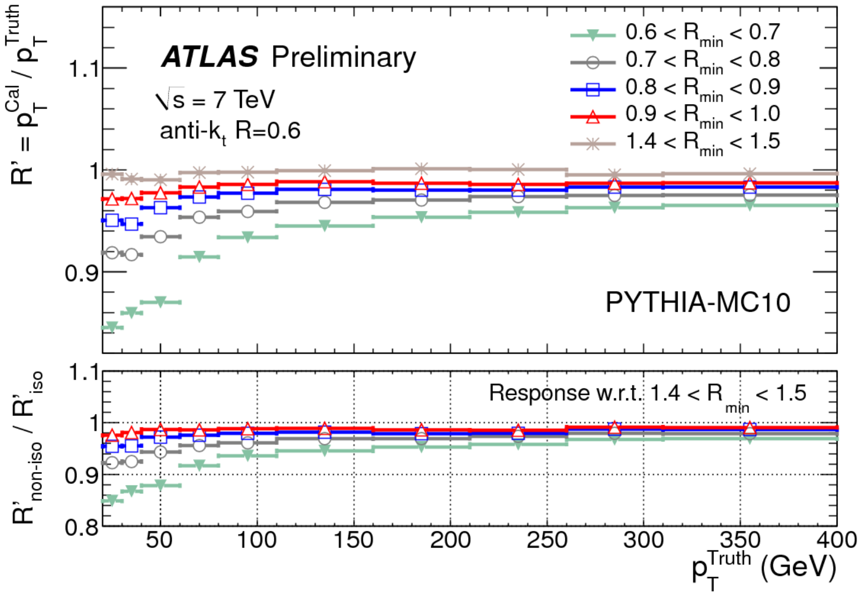

Figure 8.16: *Calorimeter jet p_{T} response as a function of jet p_{T} for jets with nearby calorimeter activity and different values of ΔR_{min} [183]. Results are shown for jets built using the anti-k_t algorithm with resolution parameter $R = 0.6$ and with $|\eta| < 2.0$. The bottom of the figure shows the response normalized to the response of isolated jets.*

tune without pile-up interactions. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.6$ and with $|\eta| < 2.0$ are used in this figure, but similar results have been found for jets up to $|\eta| = 2.8$ and jets built with $R = 0.4$ [183]. The response is clearly different for non-isolated jets, and lower as the nearby jet gets closer to the jet under study. The effect is larger at low p_{T}, but it remains important for jets of $p_{\mathrm{T}} \approx 60$ GeV.

No correction to the jet energy scale is applied to remove the effect observed in this figure. Instead, the unfolding correction takes care of bin migrations caused by this effect. Since the Alpgen sample has limitations in the description of the rates of non-isolated jets, the unfolding correction is unable to fully account for this effect. However, differences between the Pythia and Alpgen samples are comparable to differences between the Alpgen sample and the data, and this limitation is considered as part of the systematic uncertainty in the unfolding factors, described previously in this chapter.

Despite the effect being corrected through the unfolding correction, the distinctive features of non-isolated jets motivate an independent study of their jet energy scale. This study attempts to characterize the level of agreement between the data and the Monte Carlo simulation in the description of the calorimeter response to these jets and has been described in detail in Ref. [183]. The study uses jets reconstructed using tracks of $p_\mathrm{T} > 500$ MeV and passing a set of standard quality cuts. At least two tracks in the jet are required for the jet to be used. Selected track jets are matched to jets built of calorimeter jets and the ratio of their p_Ts is studied, following a similar study performed for isolated jets [182], whose results were shown in Figure 8.14.

Figure 8.17 shows this ratio as a function of calorimeter jet p_T for jets with different values of ΔR_min in data and the Pythia AMBT1 Monte Carlo simulation. The full 2010 dataset was used in this study. The ratio is always higher than 1, as expected given the fragmentation of the jet. The ratio increases at low p_T, except in the very first bin, due to the kinematic cuts applied in the study. The fraction is smaller for jets with other nearby jets, and the Monte Carlo simulation is in agreement with the data.

This study only compares relative scales between track jets and calorimeter jets. However, there is agreement in this same ratio for isolated jets, as demonstrated in Figure 8.14. An additional systematic uncertainty in the jet energy scale of non-isolated jets can, thus, be estimated, and it is uncorrelated with the uncertainties presented in the previous section.

The uncertainty is estimated using the ratio of the results in data and Monte Carlo simulation. Differences and statistical fluctuations in the study are both covered using an uncertainty that depends on ΔR_min only and is different for jets built with different resolution parameters. This uncertainty takes values that vary from 1.6% to 2.8% [183], and its contribution to the total energy scale uncertainty is shown at the end of this chapter.

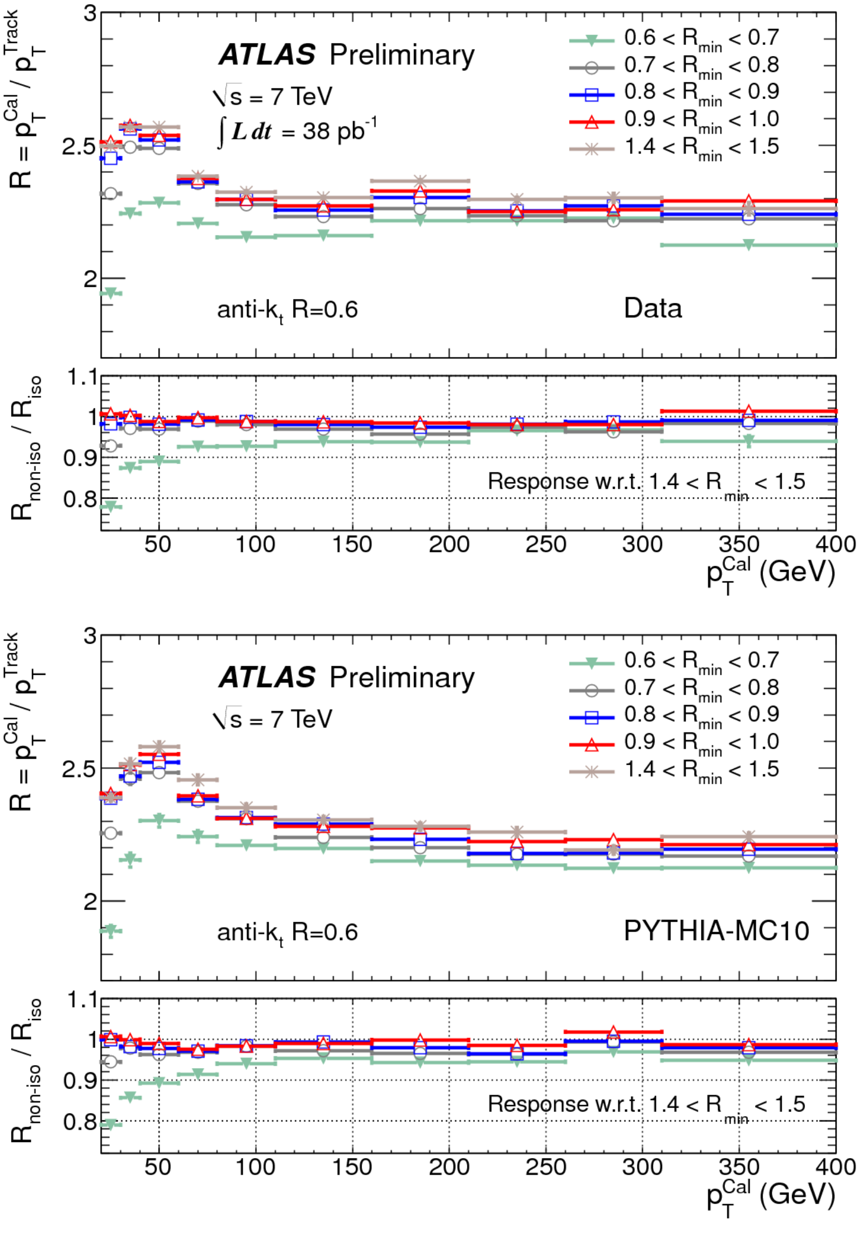

Figure 8.17: *Average ratio of calorimeter jet* p_{T} *to track jet* p_{T} *as a function of calorimeter jet* p_{T} *in data (top) and Monte Carlo simulation (bottom) [183] for non-isolated jets. The ratio between the values obtained for isolated and non-isolated jets is shown at the bottom of each figure. Jets reconstructed with the anti-*k_t *algorithm with resolution parameter* $R = 0.6$ *are used in these figures.*

8.3.3 Systematic Uncertainties Related to the Flavor Composition of the Sample

Differences in the properties of jets cause differences in the calorimeter response. This implies that for a multi-jet sample, the response is not the same as for the inclusive sample, even if the impact of non-isolated jets is ignored. This effect can be conveniently parameterized for the purposes of this analysis as a difference in the flavor composition of the sample.

The tagging of light-quark and gluon jets in a Monte Carlo event simulation carries a number of theoretical ambiguities [184, 185]. However, if the tagging is simply used to parameterize differences in the properties of the jets that can impact the jet response, then the discussion of such ambiguities becomes largely irrelevant. Therefore, any tagging algorithm that is able to capture the varying properties of jets from a di-jet sample to a γ+jet sample or a multi-jet sample suffices.

In this section, jets in the event simulation are identified as light-quark (LQ) jets or gluon jets using the partons in the generator event record. The highest energy parton within a ΔR equal to the resolution parameter used in the jet reconstruction determines the flavor of the jet. Only a small fraction of the jets ($< 1\%$ at low p_{T} and less at high p_{T}) is not assigned a flavor. Jets identified as originating from c- and b-quarks (heavy flavors) are considered separately from light-quark jets. The tagging algorithm has limitations in identifying heavy-quark jets when they originate, for example, from a gluon splitting. Such limitations are not important for the studies presented in this section, since the contribution of heavy flavors to multi-jet QCD final states is small.

Jets identified in the Monte Carlo simulation as originating from a light quark have a significantly different response from that of jets identified as originating from a gluon. This is illustrated in Figure 8.18. The p_{T} response is shown for isolated jets built with the anti-k_t algorithm with resolution parameter $R = 0.4$ falling in the barrel and in the end-cap. The response is shown for jets tagged as light-quark jets and those tagged as gluon jets, as well as the average response of all jets in the sample. Similar features have been observed for jets reconstructed with $R = 0.6$. Results for tagged non-isolated jets have been verified to show similar features, once normalized to the overall response of the sub-sample of non-

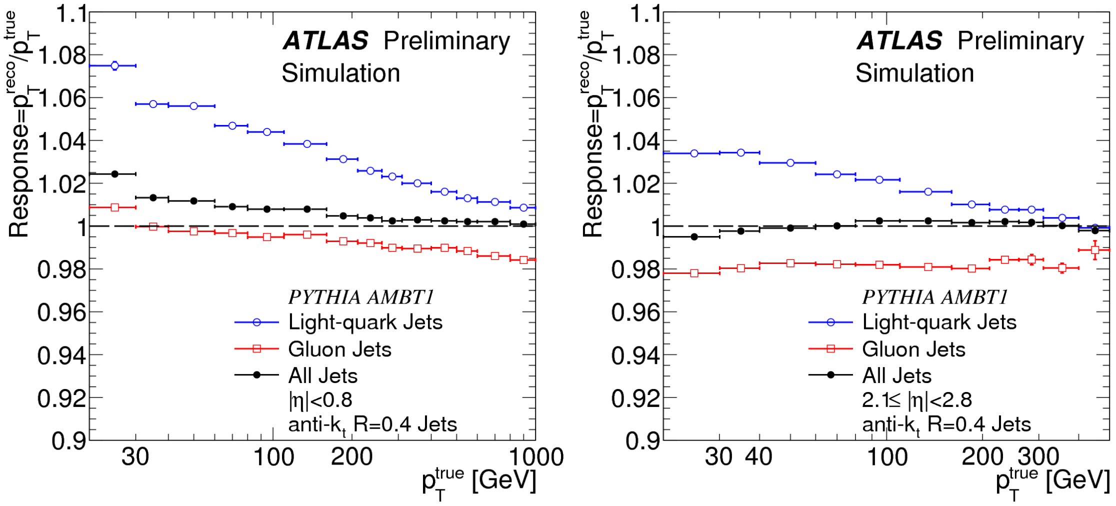

Figure 8.18: *Response as a function of particle jet p_T for all jets in the* Pythia *AMBT1 sample, gluon jets, and light-quark jets falling in the barrel (left) and in the end-cap (right) [27]. Jets are reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.4$.*

isolated jets. This conveniently allows separating the study of the systematic effects on the response due to non-isolation and those due to flavor. This implies that changes in the jet response due to flavor can be studied independent of the isolation of the sample, and the results apply to both isolated and non-isolated jets if the average effect on non-isolated jets is properly treated.

The difference in response between jets of the two different flavors is, at least partially, a result of the different properties of these two types of jets. Jets identified as gluon jets tend to have more and softer particles than light-quark jets. Additionally, gluon jets tend to have a lower energy density in the center of the jet, being typically wider. This is demonstrated in Figure 8.19, where the n_{trk} and width distributions of jets tagged as originating from a light quark and a gluon are shown. These results are shown for jets with $|\eta| < 0.8$ and 80 GeV $\leq p_T <$ 110 GeV built using a resolution parameter $R = 0.6$. In this context and for the rest of this section, n_{trk} refers to the number of tracks with $p_T >$ 1 GeV found within a $\Delta R = 0.6$ from the jet. Tracks are further required to pass standard quality cuts (see Chapter 5), and have an impact parameter with respect to the event vertex of less than 200 μm in the radial direction and 450 μm along the beam direction. These last cuts guarantee that tracks are originating from the hard interaction and make the variable robust

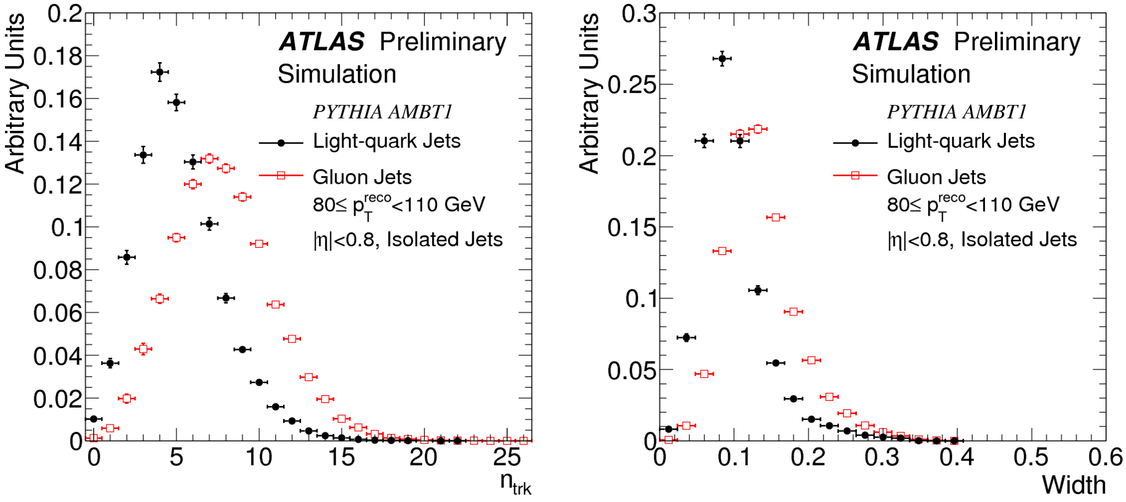

Figure 8.19: *Distribution of number of tracks associated to the jet (left) and jet width (right) for isolated jets tagged as light-quark and gluon jets in the* Pythia *Monte Carlo simulation with the AMBT1 tune [27]. Jets with* $|\eta| < 0.8$ *and* $80\text{ GeV} \leq p_{\text{T}} < 110\text{ GeV}$ *built with the anti-*k_t *algorithm with resolution parameter* $R = 0.6$ *are shown. The distributions are normalized to a unit area.*

in the presence of pile-up interactions. In this figure and also in the rest of this section, the jet width is defined as

$$\text{Width} = \frac{\sum\limits_{\text{clusters}} {p_{\text{T}}}^i \Delta R(i, \text{jet})}{{p_{\text{T}}}^{\text{jet}}}, \tag{8.3.1}$$

where the sum runs over all clusters in the jet, and ΔR refers to the distance in $[\eta, \phi]$ from the cluster to the jet axis.

To the extent that the Monte Carlo simulation reflects the data, the difference in the response as a function of jet multiplicity due to the effect of flavor is accounted for in the unfolding correction. However, an additional jet energy scale uncertainty could arise, since the standard jet energy scale was derived for a particular admixture of light-quark and gluon jets. For a different admixture, the uncertainty in the jet energy scale could be different. In addition, the fraction of light-quark and gluon jets in the multi-jet sample in the data could differ from the fraction predicted by the Monte Carlo simulations, thus leading to a systematic shift in the jet energy scale.

The first effect has been considered using the different Monte Carlo simulations used in the calculation of the jet energy scale uncertainty in the inclusive sample [27]. The response of light-quark and gluon jets with respect to the response of the admixture has been studied in each of those samples, and has been found to be the same as in the standard Pythia sample generated with the AMBT1 tune, for the available statistics. This implies that the overall shifts in energy scale found in those samples impacts similarly light-quark and gluon jets. For this reason, no additional uncertainty is considered to account for this effect.

The second effect depends on many theoretical aspects in the event production (parton distribution functions, limitations of leading-order calculations, initial and final state radiation tuning) and is not easy to study consistently using different Monte Carlo simulations. Instead, the effect can be addressed in a data-driven way using template fits to determine the flavor composition of the multi-jet sample.

This study has been described in detail in Ref. [27]. It exploits the sensitivity of n_{trk} and the jet width to the flavor of a jet, illustrated in Figure 8.19. In essence, Monte Carlo simulations are used to build template distributions of n_{trk} and the jet width for jets of different flavors. Different templates are built for different values of jet p_{T}, η, and the variable characterizing the jet isolation in the previous section, ΔR_{min}. These templates are then used to fit for the fraction of jets of each flavor in the multi-jet sample.

The n_{trk} and jet width distributions in the Monte Carlo simulations are, however, not in perfect agreement with those observed in data [19]. In addition, several Monte Carlo simulations provide a different level of agreement. In order to build the templates, the di-jet sample is used as a baseline, and the distributions in the di-jet Monte Carlo simulations are re-weighted to match those found in the data. Only events with one selected primary vertex are used to build the width template distributions, since these distributions are sensitive to pile-up interactions.

This procedure assumes that the description of the flavor composition in the di-jet data sample is the same as in the Monte Carlo simulations, and that the discrepancy between data and Monte Carlo simulations is caused by effects that impact the light-quark and gluon templates in the same way. An uncertainty is associated to these assumptions through the use of several Monte Carlo simulations using different pdfs, fragmentation and hadronization models. In order to understand the impact of pdfs on the study, the flavor composition from the Alpgen sample is alternatively used in the re-weighting procedure. This leads to different templates that help estimate the differences in the study caused by differences between the pdfs used in the ALPGEN and PYTHIA generation. The PYTHIA samples with the AMBT1 and Perugia2010 tunes and a HERWIG++ sample are used to estimate the uncertainty associated to hadronization and fragmentation effects. More details on these simulations are provided in the next chapter.

Figure 8.20 shows the n_{trk} and width templates built using this re-weighting procedure. Only the light-quark and gluon templates are shown, together with the inclusive distribution from the data. The templates are shown as a band and the width of the band represents the maximum variation among the results obtained with the PYTHIA AMBT1 and Perugia2010 tunes and the HERWIG++ simulations, used to calculate a systematic uncertainty in the template fit procedure.

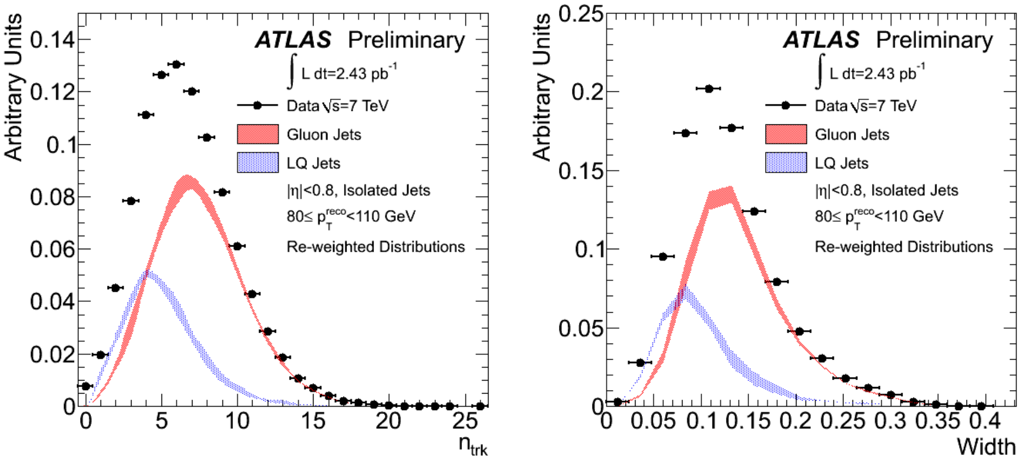

Figure 8.20: *Distribution of number of tracks associated to the jet, n_{trk} (left), and jet width (right) for isolated jets in data and Monte Carlo simulations after the re-weighting procedure [27]. The distributions for light-quark jets are shown in blue, and those for gluon jets are shown in red. The width of the band represents the maximum variation among the* PYTHIA *AMBT1 and Perugia2010 tunes and the* HERWIG++ *simulations. Jets are reconstructed with the anti-k_t algorithm and have* $|\eta| < 0.8$ *and* $80\ \mathrm{GeV} \leq p_{\mathrm{T}} < 110\ \mathrm{GeV}$.

In order to verify that these variations properly account for the systematic effects on the re-weighting procedure, a template fit is performed to a γ+jet sample. The cuts used to select of sample are described in detail in Ref. [27] and are very similar to those used in the results from the γ+jet studies shown in Figure 8.14. This sample is used because it has a significantly different flavor composition than that of the di-jet sample. Therefore, fits in this sample provide a window into the limitations of the templates obtained through the re-weighting procedure. The template fit uses the standard set of gluon, light-quark and heavy-quark templates built using Pythia with the AMBT1 tune. The heavy-quark flavor composition is fixed using the prediction from a Pythia sample that simulates the γ+jet process to allow the fit to converge.

Figure 8.21 shows the results of this fit using the width distribution. Isolated jets built with the anti-k_t algorithm with $R = 0.6$ with $|\eta| < 0.8$ and 60 GeV $\leq p_{\mathrm{T}} <$ 80 GeV are used in this figure. The template distributions are shown stacked on top of each other multiplied by the corresponding fraction obtained through the fit. The distribution in the data is also shown. The fit predicts 16% gluon jets with a 10% statistical uncertainty. The uncertainty is larger once systematic effects are considered. The Monte Carlo simulation predicts 6% gluon jets, in agreement with the results of the fit. Such an agreement is not obtained if the n_{trk} templates are used. This implies that the estimated systematic uncertainty in the light-quark n_{trk} templates is not sufficient, and the predictions of the fits are likely to be inaccurate in samples dominated by light-quark jets, like the γ+jet sample. Given that the multi-jet sample is expected to be dominated by gluons, the n_{trk} fits have been carried

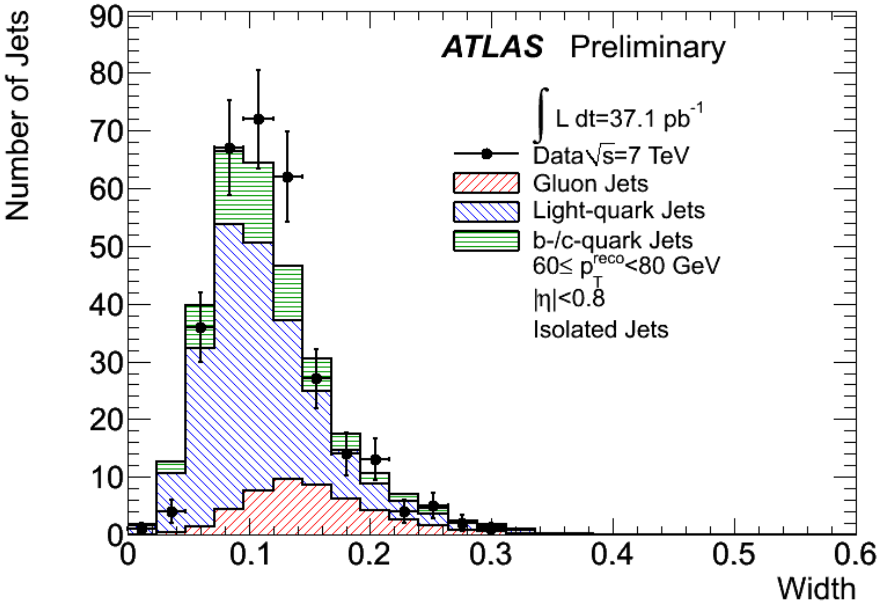

Figure 8.21: *The template fit performed using a γ+jet data sample and templates derived in the di-jet Monte Carlo sample created using* Pythia *with the AMBT1 tune [27]. Isolated jets built with the anti-k_t algorithm with $R = 0.6$ with $|\eta| < 0.8$ and 60 GeV $\leq p_{\mathrm{T}} <$ 80 GeV are shown.*

out [27]. Their results are consistent with those obtained with the width fits and are not shown in this section.

Once the templates are built and validated, fits can be performed to the multi-jet sample to verify that its flavor composition is well modeled in the simulation. As in the example before, the fraction of heavy-flavor jets is fixed to help the fits converge. The fraction is fixed to that predicted by the Alpgen sample, and a systematic uncertainty is associated to this choice by varying the fraction by a factor of 2. This ad-hoc choice relies on the contribution of heavy-flavor jets to the multi-jet sample being small. This is true in the Monte Carlo simulation, and also has been shown to be true in studies of QCD di-jet events [10]. It is thus considered a reasonable assumption for the purpose of this measurement.

Two additional uncertainties are considered in the fits to the multi-jet sample. Both target reconstruction effects that can change the template distributions. Templates are built in bins of reconstructed jet p_{T}. They are thus subject to bin migrations caused by the underlying p_{T} spectrum. The underlying p_{T} spectrum depends on the jet multiplicity and optimally templates should be calculated for each multiplicity or even each leading and sub-leading jet. This is not done, and instead an uncertainty is associated to it by building templates with two extreme shapes: a flat p_{T} distribution and the steeply decaying distribution of the sixth leading jet in a six-jet event, as predicted in the Alpgen sample.

Differences in the template distributions can also be caused by mis-measurements of the jet energy. Even though the width distributions in di-jet simulations and data agree by construction, this does not mean that the templates are not affected by a jet energy mis-measurement. To conservatively account for this effect, templates are built shifting up and down the jet energy scale according to the results presented in previous sections. This new set of templates is used to extract new predictions on the flavor composition of the sample and these predictions quantify one of the uncertainties in the fits.

Figure 8.22 shows the results of these fits in three-jet events. The selection criteria are the same as those presented in Chapter 6, except that only events with one reconstructed vertex are used, so that the impact of pile-up interactions on the width distributions can be ignored. The fits are shown for non-isolated jets (with $0.8 < \Delta R_{\mathrm{min}} < 1.0$) with $|\eta| < 0.8$ reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.6$. The different systematic uncertainties associated to the measurement are shown at the bottom of the

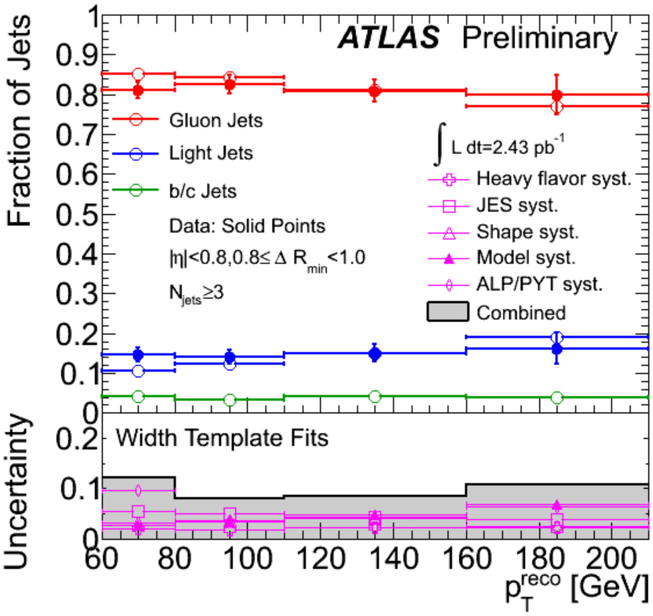

Figure 8.22: *Fitted values of the light-quark and gluon jet fraction in events with three or more jets, as a function of jet p_{T}, calculated using jet width templates [27]. Non-isolated jets ($0.8 < \Delta R_{\mathrm{min}} < 1.0$) with $|\eta| < 0.8$ are shown. The fraction of heavy-quark jets has been fixed to that of the* ALPGEN *simulation. The fitted fractions are shown with solid markers, while the values obtained from the Monte Carlo simulation are shown with empty markers. Jets reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.6$ were used for this study. On the bottom of the figure, the impact of the different systematic effects is shown with magenta markers. The combined systematic uncertainty is shown at the bottom as a gray band.*

figure, separating those caused by limitations in the understanding of the hadronization and fragmentation models (“model systematics” in the labels) and those caused by limitations in the knowledge of the flavor composition of the di-jet sample (“ALP/PYT systematics” in the labels). The addition in quadrature of all these components is shown as a gray band. Agreement is found within the systematic uncertainties of the measurement between the data and the predictions from ALPGEN.

This exercise has been performed for all bins up to the four-jet bin. Similar results in the four-jet bin are shown in Figure 8.23. Again, the extracted fractions are consistent with the predictions from the ALPGEN Monte Carlo simulation, and the total systematic uncertainty is similar to that found in the three-jet bin. The statistical uncertainties in the fit are much larger, as expected due to the decreasing cross section with jet multiplicity. The available statistics do not allow performing the fits in bins of higher multiplicity.

The fits have been performed in several bins of p_{T}, η and isolation, only a few of which

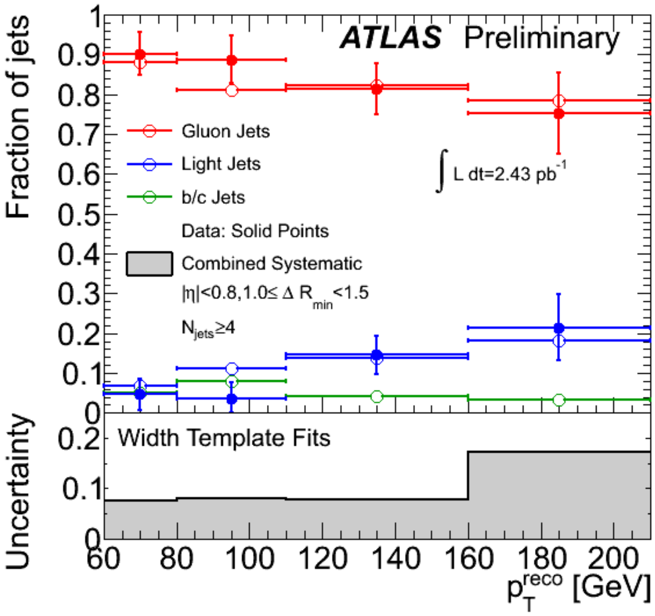

Figure 8.23: *Fitted values of the light-quark and gluon jet fraction in events with four or more jets, as a function of jet* p_{T}*, calculated using jet width templates [27]. Non-isolated jets (*$1.0 < \Delta R_{\mathrm{min}} < 1.5$*) with* $|\eta| < 0.8$ *are shown. The fraction of heavy-quark jets has been fixed to that of the* Alpgen *simulation. The fitted fractions are shown with solid markers, while the values obtained from the Monte Carlo simulation are shown with empty markers. Jets reconstructed with the anti-*k_t *algorithm with resolution parameter* $R = 0.6$ *were used for this study. The combined systematic uncertainty is shown at the bottom as a gray band.*

are shown in the figures above. The fits in all these bins can be combined to provide the extracted flavor composition as a function of jet multiplicity. This result is shown in Figure 8.24. The fractions are consistent between the data and the Alpgen Monte Carlo simulation, and the total systematic uncertainty is around 10% for each multiplicity bin. The fit and associated uncertainties are also shown in the two-jet bin, even though the use of a di-jet sample to construct the templates does not allow providing an independent prediction in that bin.

This study shows that the Alpgen Monte Carlo simulation correctly describes the flavor composition of the multi-jet sample up to the four-jet bin and up to jet p_{T}s of 210 GeV. The description is accurate within a ≈10% systematic uncertainty on average. However, additional statistical uncertainties in the fit need to be considered, particularly at high p_{T} in the four-jet bin. In addition, the fits for certain non-isolated bins do not agree so well, as shown in Appendix D. In order to account for these effects, and for simplicity, a 30%

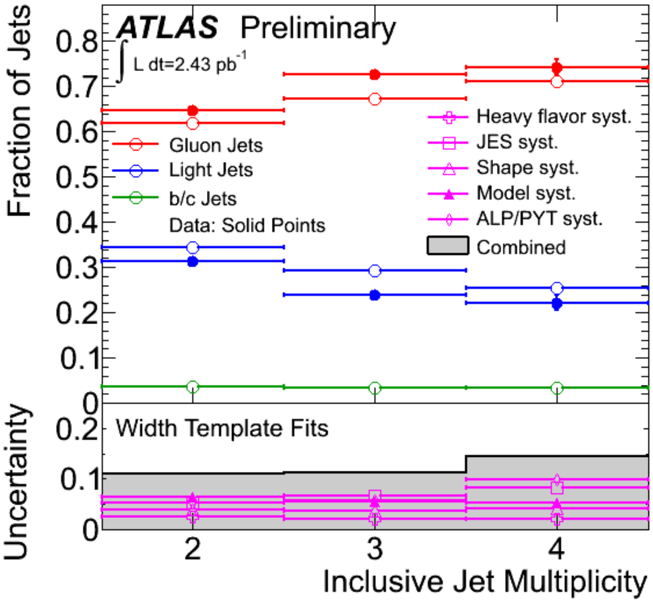

Figure 8.24: *Fitted values of the light-quark and gluon jet fraction as a function of inclusive jet multiplicity, calculated using jet width templates [27]. The fraction of heavy-quark jets has been fixed from the* ALPGEN *Monte Carlo simulation. The fitted fractions are shown with solid markers, while the values predicted by the* ALPGEN *Monte Carlo simulation are shown with empty markers. Jets reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.6$ were used for this study. On the bottom of the figure, the impact of the different systematic effects is shown with magenta markers. The combined systematic uncertainty is shown at the bottom as a gray band.*

uncertainty in the flavor composition is used to estimate the jet energy scale uncertainty for jets of $p_{\mathrm{T}} < 210$ GeV in events with three and four selected jets. A 100% uncertainty is assumed conservatively in five- and six-jet events, where statistics do not allow such a detailed study. In the two-jet sample, no additional uncertainty is added, since this sample has a large overlap with the inclusive sample, and any uncertainties in the flavor composition have already been accounted for through the studies in previous sections. The uncertainties in the flavor composition are used to calculate the final uncertainty in the jet energy scale.

Differences exist in the flavor composition between the ALPGEN and the PYTHIA samples. Those differences are not large for low-multiplicity final states, but increase up to ≈ 10% for high-multiplicity final states. These differences are already accounted for in the systematic uncertainties in the unfolding factors, leading to a partial double-counting of this effect. However, 10% differences in the flavor composition do not impact the jet energy scale significantly, so the impact of this double-counting on the final systematic uncertainty in the measurements is small and can be neglected.

8.3.4 Combination of Uncertainties

Three sets of studies have been presented addressing three different aspects of the uncertainties in the jet energy scale. The three aspects are independent and the different effects they address are uncorrelated. For this reason, the uncertainties derived in each of these studies are added in quadrature to provide the final systematic uncertainty in the jet energy scale.

The uncertainty in the flavor composition derived in the previous section needs to be translated into an uncertainty in the jet energy scale. This can be done easily as follows. Let us refer to the response of a particular sub-sample of jets in the multi-jet analysis as R_s and to the fraction of jets in that sample with flavor x as f_x. Then, the response in that sub-sample can be expressed as:

$$\begin{aligned} R_s &= f_g R_g + f_q R_q + f_b R_b + f_c R_c && (8.3.2) \\ &= 1 + f_g(R_g - 1) + f_q(R_q - 1) + f_b(R_b - 1) + f_c(R_c - 1), && (8.3.3) \end{aligned}$$

where the q subscript refers to light-quark jets and $f_g + f_q + f_b + f_c = 1$ has been used to arrive at the second expression.

Variations in the response, δR_s, can be calculated as

$$\begin{aligned} \delta R_s = {} & \delta f_g(R_g - 1) + \delta f_q(R_q - 1) + \delta f_b(R_b - 1) + \delta f_c(R_c - 1) && (8.3.4) \\ & + f_g \delta R_g + f_q \delta R_q + f_b \delta R_b + f_c \delta R_c. && (8.3.5) \end{aligned}$$

The uncertainty in the fraction of heavy-quark jets has been already included in the template fits, and is small, so this expression can be simplified to

$$\delta R_s \approx \delta f_g(R_g - 1) + \delta f_q(R_q - 1) + f_g \delta R_g + f_q \delta R_q + f_b \delta R_b + f_c \delta R_c. \tag{8.3.6}$$

The uncertainty in the response is the same for light-quark and gluon jets and equal to the uncertainty found in the inclusive sample. Since few heavy-quark jets are present in the sample, a similar result can be assumed for these jets. Therefore, variations in the response across jets of all flavors are the same and fully correlated, $\delta R_x = \delta R_j$. Also, since f_b and f_c are small, $\delta f_q = -\delta f_g$ so

$$\begin{aligned} \delta R_s &\approx -\delta f_q(R_g - 1) + \delta f_q(R_q - 1) + f_g \delta R_j + f_q \delta R_j + f_b \delta R_j + f_c \delta R_j, && (8.3.7) \\ &= \delta f_q(R_q - R_g) + \delta R_j, && (8.3.8) \end{aligned}$$

where the expression $f_g + f_q + f_b + f_c = 1$ has again been used. This variational expression can be related to the systematic uncertainty in R_s, which is formed by the addition in quadrature of two uncorrelated terms. The second term corresponds to the systematic uncertainty calculated in the inclusive sample, and the first term corresponds to the uncertainty due to limitations in the knowledge of the flavor composition of the sample. In relative terms, the latter can be expressed as

$$\frac{\delta R_s}{R_s} = \delta f_q \frac{R_q - R_g}{R_s}. \tag{8.3.9}$$

This expression provides a way of translating uncertainties in the flavor composition, estimated in the previous section, to uncertainties in the jet energy scale. R_s has been studied in the multi-jet sample as a function of multiplicity, p_{T} and η using Monte Carlo simulations. The differences in flavor composition with respect to the inclusive sample are not enough to change R_s significantly, so $R_s = 1$ is used for simplicity. It should be noted that δf_q is bounded below by f_q and above by $1 - f_q$. For this reason, the uncertainties in the jet energy scale caused by this effect are asymmetric for typical values of f_q of 20% in the multi-jet sample. The consequences of this are most noticeable in the five-jet and six-jet bins. A 100% uncertainty in the flavor composition is assumed, which in fact means that $\delta f_q \approx^{+0.80}_{-0.20}$.

Figure 8.25 shows the positive and negative uncertainties in the jet energy scale after all effects are combined in the inclusive sample of isolated jets. It also shows the uncertainty for non-isolated jets with two values of ΔR_{min} for jets in events with two selected jets. Jets reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.4$ and with $0.3 < |\eta| < 0.8$ were used in this figure. Results for jets built with $R = 0.6$ and for other $|\eta|$ bins are comparable and can be found in Appendix D.

The asymmetry in the flavor composition uncertainty is obvious. The results in the three- and four-jet bins are the same as in the five- and six-jet bins for the downwards

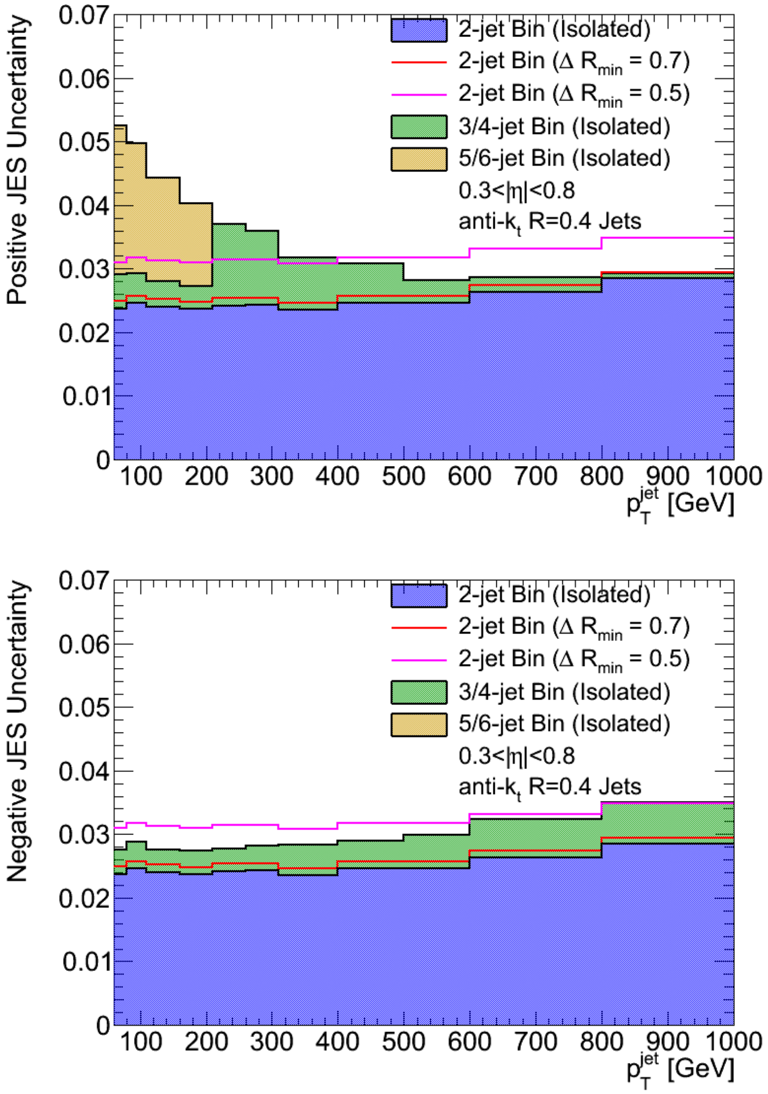

Figure 8.25: *Total upwards (top) and downwards (bottom) jet energy scale uncertainties as a function of jet* p_T *for isolated jets in different multiplicity bins, and non-isolated jets in events with two selected jets. Jets reconstructed with the anti-*k_t *algorithm with resolution parameter* $R = 0.4$ *and with* $0.3 < |\eta| < 0.8$ *were used in this figure.*

uncertainty, since $f_q \approx 0.20$, thus constraining the value of δf_q in all bins. The impact of the flavor composition uncertainty on the upwards uncertainty in the jet energy scale reaches up to 3% where template fits were not available, while it is just 0.5% where the fits could be performed. The effect of nearby jets is also shown. It increases with decreasing ΔR_min, impacting the overall uncertainty in the jet energy scale by up to $\approx$ 1% in the two-jet bin.

Chapter 9

Theoretical Calculations and Associated Uncertainties

There are two primary motivations for the measurements performed in this thesis. The first is to evaluate how robust leading-order (LO) QCD calculations are in representing the high jet multiplicity events, relevant as backgrounds in many searches of new physics. The second is to test next-to-leading-order perturbative QCD (NLO pQCD) calculations, as a first step towards constraining some of the parameters entering them.

Leading-order predictions are available in the form of calculations of the scattering matrix elements coupled to Monte Carlo simulations of parton showers, hadronization and underlying event, among other effects. These predictions, thus, provide a full description of a particle-level event. Unfortunately, large uncertainties are inherent to the choice of renormalization and factorization scales in the calculation of the scattering matrix elements. In addition, certain parameters of the predictions, such as those defining the underlying event, require tuning to data and automatically fix some of the other parameters. Since the number of available tunes is limited, it is hard to treat systematic uncertainties in the predictions consistently. For these reasons, no attempt is made in this thesis to assign an uncertainty to the leading-order prediction. Instead, a large number of leading-order calculations and tunes is studied. This allows understanding the effects of the various parameters and approximations entering each calculation, and evaluating which agrees best with the data.

Next-to-leading-order predictions have an inherently lower uncertainty due to the choice of renormalization and factorization scale. This makes these predictions more precise than those provided by leading-order calculations, and the estimation of uncertainties in the predictions becomes more relevant. However, the choice of the exact value for this scale is somewhat ad-hoc, and the impact of such choice needs to be understood. Currently, calculations of multi-jet production cross sections at next-to-leading order only exist up to three-jet final states, and do not sum leading logarithms to all orders or include corrections for non-perturbative effects. The effects of these limitations also need to be understood to allow for comparisons with data.

The details of the leading-order and next-to-leading-order calculations used in this thesis are described in separate sections of this chapter. The description of the different leading-order calculations includes brief summaries of each model studied, and a discussion of what can be learned from the study of the model. The description of the next-to-leading-order calculation includes studies that motivate the choice of renormalization and factorization scales. It also presents a calculation of a correction for non-perturbative effects, and the estimate of the uncertainties in the calculation.

9.1 Leading-order Calculations

Many different effects contribute to the calculations implemented in leading-order Monte Carlo simulations of jets. The impact of these effects in multi-jet production at the LHC needs to be understood, given the high center-of-mass energy of the accelerator, which challenges some of the approximations and models going into these simulations.

Some of the relevant aspects of the LO calculations for the studies performed in this analysis include the parton-shower parameters, and the advantages of a matrix-element/parton-shower (ME+PS) calculation ($2 \to n$) over a pure parton-shower calculation evolving from a $2 \to 2$ matrix element. The impact of the underlying event, whose importance at the LHC increases with respect to earlier accelerators [186], or the specifics of the hadronization model cannot be neglected.

In order to study differences between a ME+PS calculation and a pure PS calculation, Alpgen [160] and Pythia [79] event simulations are used. The Alpgen samples are generated with the `CTEQ6L1` [187] parton density functions. Scattering diagrams are calculated with up to six outgoing partons, and renormalization and factorization scales, Q, that are different for each event:

$$Q^2 = \sum {p_{\mathrm{T}}}^2, \tag{9.1.1}$$

where the sum runs over all outgoing partons in the event. The calculation is constrained to produce only partons with $p_{\mathrm{T}} > 10$ GeV, $|\eta| < 6$ and separated by at least $\Delta R = 0.7$ in $[\eta, \phi]$. Samples with doubled and halved values of the renormalization and factorization scales have also been produced to study variations in the choice of scale.

Alpgen is a pure LO ME calculation, and thus needs to be interfaced to a PS implementation with hadronization and other effects to provide a particle-level prediction with re-summation of large logarithms. Both the Pythia and the Herwig [161] event generators can be interfaced to Alpgen using the MLM [67] method for matching matrix elements to parton showers. Pythia implements p_{T}-ordered parton showers, an underlying-event model for multiple-parton interactions and the Lund string model for hadronization. One of the newest Pythia tunes available for the CTEQ6L1 pdf set is used: the ATLAS MC09′ tune [25, 188]. This tune is optimized to reproduce measurements of minimum bias, underlying event and angular correlations in di-jet events at the Tevatron [189–194], and has been shown to perform well at reproducing similar measurements in ATLAS [97, 195]. An equivalent Pythia sample with the same pdf and tune has been produced to isolate differences between the ME+PS and the pure PS calculations.

Alpgen is also interfaced to Pythia with different parton-shower and underlying-event tunes. In particular, samples are generated with Alpgen+Pythia and the Perugia6 [61] and D6 tunes [79]. The Perugia6 tune is created using a different methodology than that used to produce the MC09′ tune. Data from the Tevatron, the Large Electron Positron collider (LEP) [196, 197] and the Super Proton Synchrotron (SPS) [198] are used in this tune. The tune has been shown to be less successful at describing ATLAS data than the MC09′ tune [97], and can be used to establish the sensitivity of measurements performed in this analysis to the underlying-event tune. The D6 tune is calculated using Tevatron data, but it uses virtuality ordering in the PS evolution, unlike the other Pythia tunes considered here, which use p_{T} ordering. It, thus, allows for an investigation of the impact of the ordering of the parton shower on the observables under study.

The different Pythia tunes have been derived using results from the Pythia event generator alone. The interactions between the tune parameters and the MLM matching procedure have not been considered in the tuning process. Such interactions have been shown to be non-negligible and can change the intended effects of a tune [199]. The study of several tunes, thus, also allows for an investigation of such interactions.

Alpgen is also interfaced to Herwig, which implements angular-ordered parton showers and a cluster model for hadronization. Herwig is combined with the Jimmy simulation, that provides a realistic description of multiple-parton interactions. An Alpgen+Herwig/Jimmy sample is generated using the Herwig/Jimmy MC09 tune [188], which is derived using a similar methodology as the MC09′ tune, but instead using tunable parameters in the Herwig/Jimmy models. This allows for a direct comparison with the Alpgen+Pythia MC09′ sample, establishing the impact of angular ordering and hadronization in the measured quantities. An additional Alpgen+Herwig sample with the AUET1-CTEQ6L1 tune [162] (also referred as MC10) is generated. The AUET1 tune is similar to the MC09 tune, except that it uses additional data from ATLAS measurements of underlying-event properties in the tuning process. Since this tune is most successful at describing the underlying event data in ATLAS, it is not only used for theoretical comparisons, but also to produce the baseline sample used for the unfolding of detector effects.

In Chapter 2, two methods for performing the matching of parton showers and matrix elements were introduced. The CKKW method is implemented in the Sherpa event generator [80], and an additional sample is produced to investigate its behavior. Few tunes currently exist for this generator, so the default tune in version 1.2.3 of the generator is used. This tune has been successfully validated in comparisons with ATLAS measurements [200].

The tune is derived using the CTEQ6.6 [187] NLO pdf, different from the pdf used to generate the Alpgen samples. The renormalization and factorization scales are calculated using the maximum transverse mass of the two objects obtained when clustering outgoing partons with the k_t algorithm into a $2 \rightarrow 2$ configuration. These are some of the important differences between the Sherpa and Alpgen generators. Such differences and the lack of more Sherpa tunes make it impossible to systematically study the impact of the CKKW and MLM matching procedures in the observables under study. The comparisons with Sherpa, thus, constitute a global evaluation of the many parameters and approximations entering this generator.

The generation of samples with some of the newer pdfs was not possible in the ALPGEN version used in this thesis, due to limitations in the software implementation. For this reason, additional PYTHIA samples with the MC09 [188], AMBT1 [97] and Perugia2010 [61] tunes are generated to study the impact of the use of different pdfs, and associated tunes in the predictions of the PS Monte Carlo simulations.

The PYTHIA MC09 tune is similar to the MC09′ tune, except that it uses the MRST2007 lomod [201] pdf instead. This pdf uses calculations of α_S at next-to-leading order in an attempt to approximate the results obtained with the leading-order calculation to results that one would obtain with a NLO matrix-element calculation [202]. The AMBT1 tune (also sometimes referred to as ATLAS MC10 tune), uses the same pdf and methodology as the MC09 tune, but includes minimum-bias measurements performed with the ATLAS detector in the tuning algorithm. Finally the Perugia2010 tune is obtained using a similar methodology as the Perugia6 tune, but it uses the CTEQ5L pdf [203], includes measurements of event and jet shapes in the tuning procedure and varies certain parameters of the parton-shower component of the tune. These additional measurements and the additional parameters surveyed cause a change in the initial-state radiation in the tune, possibly affecting the predictions of multi-jet cross sections.

An additional implementation of the 2 → 2 scattering with the parton-shower enhancement is briefly investigated: HERWIG++ [204]. This generator improves many of the features of HERWIG, incorporating a realistic model of multiple-parton interactions. HERWIG++ describes the internal structure of jets in ATLAS comparably well as the generators and tunes introduced above, but showing different features [147]. Since this generator is relatively new, few tunes are available. Only the default tune in version 2.5, which uses the MRSTMCal pdf is used in comparisons to the measurements. As for SHERPA, few conclusions are drawn from these comparisons, which only provide a global evaluation of the performance of the generator and this particular tune at describing the multi-jet data.

Table 9.1 summarizes the LO Monte Carlo simulations used to compare to data as well as the motivation for each choice.

Generator	pdf	Tune	Purpose
ALPGEN +PYTHIA	CTEQ6L1 [187]	MC09′ [26, 188]	-ME+PS accuracy -Direct comparisons to PS
ALPGEN +PYTHIA	CTEQ6L1 [187]	Perugia6 [61]	-Impact of UE -MLM/UE interactions
ALPGEN +PYTHIA	CTEQ6L1 [187]	D6 [79]	-Impact of ordering in PS -MLM/UE interactions
ALPGEN +HERWIG/JIMMY	CTEQ6L1 [187]	MC09 [188]	-Impact of ordering in PS
ALPGEN +HERWIG/JIMMY	CTEQ6L1 [187]	AUET1 [162]	-As previous entry, tuned to ATLAS data
SHERPA	CTEQ6.6 [187]	Default	-Independent ME+PS implementation
PYTHIA	CTEQ6L1 [187]	MC09′ [26, 188]	-Direct comparisons to ME+PS
PYTHIA	MRST2007lomod [201]	MC09 [188]	-Impact of different pdfs -UE systematic effects in NLO calculation
PYTHIA	MRST2007lomod [201]	AMBT1 [97]	-As previous entry, tuned to ATLAS data
PYTHIA	CTEQ5L [203]	Perugia2010 [61]	-Impact of different pdfs -Impact of ISR tuning
HERWIG++	MRSTMCal [201]	Default	-Independent PS implementation -Hadronization effects in NLO calculation

Table 9.1: *Summary of leading-order calculations used for comparisons to data, and the physics probed by each of them.*

9.2 Next-to-leading-order Calculations

For the next-to-leading-order study, the calculation implemented in NLOJet++ [205] is used. The NLOJet++ program implements a matrix-element calculation to next-to-leading order for QCD processes with two and three outgoing partons. This calculation lacks an interface to a parton-shower calculation and does not account for non-perturbative effects. The inclusion of leading logarithms to all orders is not possible in the current NLOJet++

framework, and this can cause limitations in the predictions in regions of phase space where such contributions are important. On the other hand, non-perturbative effects can be estimated from LO Monte Carlo simulations, and the NLO calculation can be corrected to provide a particle-level prediction that can be compared to the measurements.

Additional limitations in the NLO prediction may arise from the treatment of virtual diagrams, which typically appear in the calculation with a negative sign. Virtual diagrams are considered as events that contribute with a negative weight to the total cross section. An obvious example of a limitation caused by this treatment appears when a two-jet event is defined by the presence of two jets above a certain p_{T} cut. Predictions of the cross sections for events with the two jets near that cut may be negative. The cut effectively throws away any diagram in which a soft real emission is radiated off one of the original partons, because two of the three partons no longer pass the cut. However, these diagrams cancel certain terms in virtual diagrams, with internal loops, that are not thrown away by the cut. In order to avoid this instability in the calculation the leading jet p_{T} is required to be higher than the p_{T} of other jets in the analysis. This avoids unphysical predictions, but may not be enough to prevent the calculation from being limited in certain regions of phase space.

There are very few parameters entering the NLO calculation, namely, the renormalization and factorization scales, the set of parton distribution functions and the strong coupling constant, α_S. The renormalization and factorization scales are typically taken to be the same, while α_S is fixed once the pdf set is chosen. Then, only the choice of scale and pdf is left to the user. Two sets of pdfs are used in this thesis to understand the dependence of the results on that choice: MSTW2008nlo [32] and CTEQ6.6 [187].

The renormalization and factorization scales should be chosen to be around some typical scale of the event, to guarantee the validity of the fixed-order calculation. The value of that "typical scale" depends on the process, and needs to be carefully determined. The choice of the wrong scale causes large logarithms to enter the calculation at fixed order. This implies that if an appropriate scale is chosen, the fixed-order calculation becomes most stable under small variations of the scale, providing an algorithm for empirically making this choice.

With all input parameters chosen, the NLO calculation can be used to make a prediction of observables in the two- and three-jet bins or combinations of those observables. Limitations in the fits to the non-perturbative component of the pdfs, or uncertainties in the size of the non-perturbative correction applied on top of the NLOJet++ calculation contribute to the systematic uncertainty in the theoretical prediction. In addition, an inherent limitation in any fixed-order prediction exists, as evidenced by the dependence of the prediction on the renormalization and factorization scale choice. The impact of such limitation in the prediction can be approximately quantified and contributes to the final systematic uncertainty in the prediction.

9.2.1 Choice of Scale

In the analysis, NLO comparisons focus on the three-to-two jet cross-section ratios. These measurements have the smallest systematic uncertainties, and are promising for probing some of the input parameters of the NLO calculations. For this reason, the scale is chosen to optimize the measurement of such ratios.

Four renormalization and factorization scales are studied:

- $Q^2 = \sum_{\text{jets}} {p_{\text{T}}}^2$,
- $Q = p_{\text{T,lead}}$,
- $Q = \frac{1}{3} \sum_{\text{jets}} p_{\text{T}}$,
- $Q = \frac{1}{N_{\text{jets}}} \sum_{\text{jets}} p_{\text{T}}$.

In these expressions, the sums run over jets built using outgoing partons from the matrix-element calculation. While not being exhaustive, this set of scales probes several characteristic momenta in the event, and provides a reasonable span of the space of all possible scale choices.

In order to determine the optimal choice of scale, the predictions are studied as the scales are varied around their central values. Six variations with respect to the nominal scales, $\mu_F = \mu_R = Q$, are considered, such that $(\mu_F^{\text{var}}, \mu_R^{\text{var}}) = (\kappa_1 Q, \kappa_2 Q)$. The values of κ in each of these variations are: $(0.5, 0.5)$, $(0.5, 1)$, $(1, 0.5)$, $(2, 1)$, $(1, 2)$, $(2, 2)$. Such ad hoc scale modifications have been used in the past [12] to estimate the impact of the limitations in the fixed-order calculation on the predictions. With that aim, a systematic uncertainty arising from such limitations is estimated as the largest difference between all the six calculations and the central value of the calculation in each bin.

A small systematic uncertainty corresponds to stability in the calculation against variations of the scale, and indicates an appropriate choice of scale. Figure 9.1 shows the relative scale uncertainty in the prediction of the three-to-two-jet cross-section ratio as a function of the p_{T} of the leading jet in the event. The uncertainty for each of the scale choices is shown as an error band around 1. The results are shown for jets reconstructed using the anti-k_t algorithm with resolution parameters $R = 0.4$ and $R = 0.6$. Clearly, the scale choice of $Q^2 = \sum_{\text{jets}} {p_{\mathrm{T}}}^2$ provides the most stable predictions under variations of that scale for both choices of R.

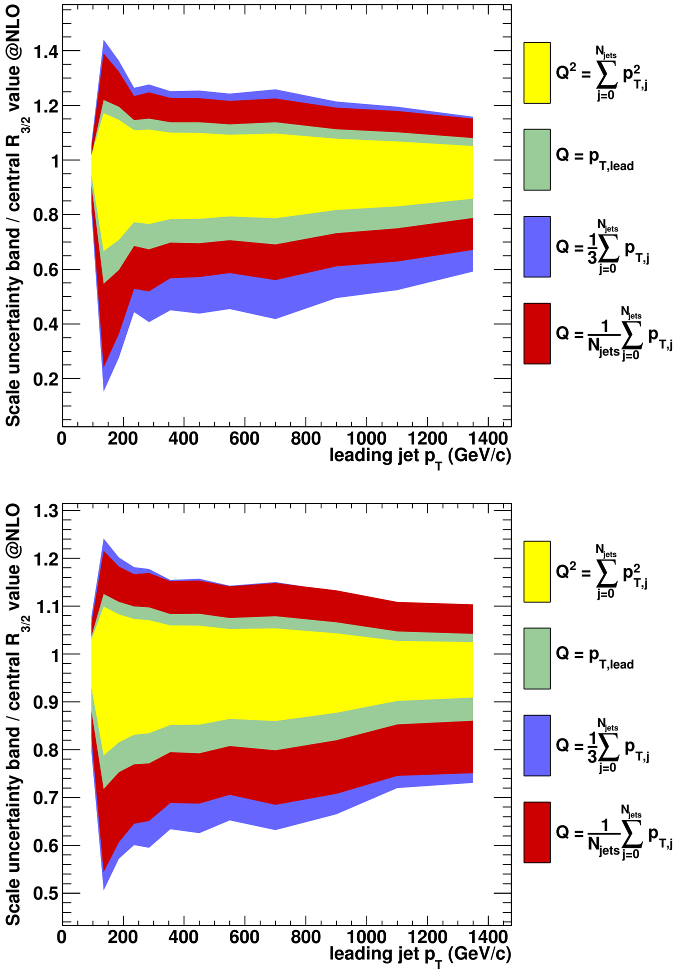

Figure 9.1: *Scale systematic uncertainty in the NLO prediction of the three-to-two-jet cross-section ratio as a function of leading jet* p_{T} *for several scale choices [163]. The results are shown for jets reconstructed using the anti-*k_t *algorithm with resolution parameters* $R = 0.4$ *(top) and* $R = 0.6$ *(bottom).*

The uncertainty estimated in this way is significantly smaller when $R = 0.6$ is used, and it varies from $\approx^{+10}_{-20}$ % to $\approx^{+3}_{-10}$ % (compared to a range of $\approx^{+15}_{-35}$ % to $\approx^{+6}_{-15}$ % if $R = 0.4$ is used). The dependence with R is not unexpected, because the resolution parameter limits the parton radiation to certain angles. This introduces terms in the matrix element calculation that go as $\ln R$ and can make smaller jets less stable under scale variations for fixed-order calculations [206].

These studies help choosing the scale providing the most stable predictions. However, the calculated uncertainty in the NLO prediction remains large, even if a resolution parameter $R = 0.6$ is used. If the measurement is to be sensitive to other parameters entering the NLO calculation, this uncertainty, which is inherent to the fixed-order calculation, should be smaller. Without higher-order calculations, the size of this systematic uncertainty can be reduced through the choice of scale or jet resolution parameter, and also through the choice of the observable.

The three-to-two-jet cross-section ratio is studied because the systematic uncertainties in the measurement are small. However, its study as a function of the leading jet p_{T}, as a representative scale of the event, can be regarded as historical [207] and its study as a function of other scales in the event is equally valid, and, in fact, provides predictions with smaller uncertainties.

In order to reduce the uncertainties in the prediction, the NLO calculation of the three-to-two-jet cross-section ratio has been performed as a function of several additional scales in the event:

- ${H_{\mathrm{T}}}^{(2)} = {p_{\mathrm{T}}}^{\mathrm{lead}} + {p_{\mathrm{T}}}^{\mathrm{2nd}}$,
- ${H_{\mathrm{T}}}^{(3)} = {p_{\mathrm{T}}}^{\mathrm{lead}} + {p_{\mathrm{T}}}^{\mathrm{2nd}} + {p_{\mathrm{T}}}^{\mathrm{3rd}}$,
- $H_{\mathrm{T}} = \sum_{\mathrm{jets}} p_{\mathrm{T}}$.

In these expressions the superscripts refer to the ordering in p_{T} of selected jets in the event, and the last sum runs over all selected jets. Figure 9.2 shows the relative scale uncertainty in the prediction of the three-to-two-jet cross-section ratio as a function of the ${H_{\mathrm{T}}}^{(2)}$ of the event. The results are shown for jets reconstructed using the anti-k_t algorithm with resolution parameters $R = 0.4$ and $R = 0.6$. The results for the other variables studied

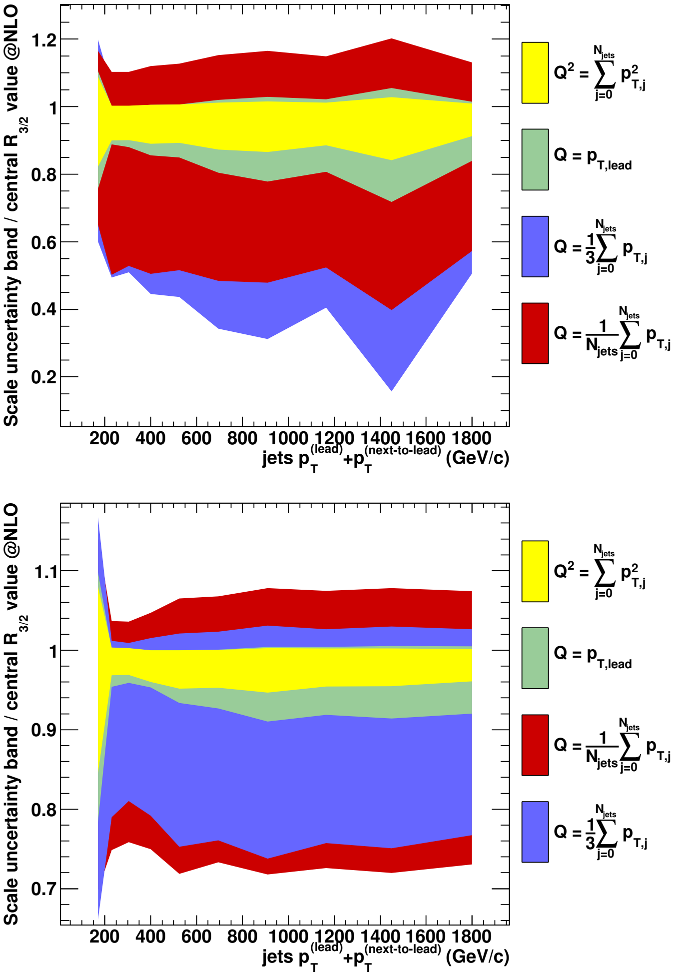

Figure 9.2: *Scale systematic uncertainty in the NLO prediction of the three-to-two-jet cross-section ratio as a function of* $H_{\mathrm{T}}{}^{(2)} = p_{\mathrm{T}}{}^{\mathrm{lead}} + p_{\mathrm{T}}{}^{\mathrm{2nd}}$ *for several scale choices [163]. The results are shown for jets reconstructed using the anti-*k_t *algorithm with resolution parameters* $R = 0.4$ *(top) and* $R = 0.6$ *(bottom).*

are shown in Appendix E, but the calculated systematic uncertainties are larger than in Figure 9.2. The choice of scale should be the same for this observable as for the previous observable shown, based on these results. The scale uncertainty is also lower when $R = 0.6$ is used, and varies from $\approx^{+8}_{-15}$ % to $\approx^{+0}_{-4}$ %. The estimated uncertainty in this prediction is significantly reduced with respect to the previous observable, as desired.

In the analysis, the two measurements, of the ratios as a function of leading jet p_{T} and event $H_{\mathrm{T}}{}^{(2)}$, are performed. The first measurement allows comparisons to similar measurements in previous hadron colliders, and is interesting from the point of view of tuning parameters of the leading-order calculation. The second measurement represents a targeted study for comparisons to NLO calculations, and can be used in the future to constrain parameters entering these calculations.

9.2.2 Non-perturbative Corrections

Non-perturbative effects, such as the soft interactions between partons in the proton not participating in the hard scatter or the hadronization, are not included in the NLO calculation provided by NLOJet++. In order to produce a particle-level NLO prediction, additional corrections for these effects need to be included in the calculation.

Leading-order simulations that include these effects are used for this purpose. Pythia samples with the AMBT1 tune are generated with and without simulation of the underlying event for this purpose, using the Rivet package [208]. This package is used to easily generate high-statistic samples with leading-order event simulations that are optimal for this calculation. Alpgen could not be interfaced to the package, so it is not used, despite potentially providing a more realistic description of the event topology. However, the calculations have been performed with Alpgen samples with fewer events and are consistent with those obtained with the Pythia samples (see Appendix E).

Jets are built using stable particles and partons before the hadronization step in the two samples. This allows comparing the predictions at the parton level in a sample without underlying event and at the particle level in a sample with underlying event. This comparison captures the non-perturbative effects that are not included in NLOJet++ and provides a correction for these effects. In particular, for a prediction, p, a multiplicative correction that is applied to the results from NLOJet++ is calculated as

$$c_{\mathrm{non-pert}} = \frac{p_{\mathrm{UE}}^{\mathrm{particle}}}{p_{\mathrm{no\ UE}}^{\mathrm{parton}}}. \tag{9.2.1}$$

The size of this multiplicative constant is typically within 5% of 1 for the observables studied. The specific values of these constants for each bin and observable are shown in the next section and Appendix E.

9.2.3 Systematic Uncertainties in the Calculation

Several uncertainties are associated to the NLO calculation and some of them have already been briefly discussed in previous sections. The uncertainty in the NLOJet++ calculation stems from uncertainties in its input parameters, pdfs and α_S, and an uncertainty that tries to quantify limitations in the fixed-order calculation through the study of its dependence on the values of the renormalization and factorization scales. An additional uncertainty in the non-perturbative correction that completes the prediction also needs to be considered.

Two effects are considered in the non-perturbative corrections: the underlying event and hadronization. Additional leading-order simulations are used to estimate the uncertainties due to limitations in our understanding of these effects. The PYTHIA sample with the MC09 tune is used to measure the impact of uncertainties in the underlying-event model. Given the differences between this tune and the data [97], which is well described by the AMBT1 tune, differences between the two PYTHIA samples are considered as representative of the uncertainties due to underlying-event effects.

The HERWIG++ sample is used to estimate the uncertainties caused by limitations in the understanding of hadronization models. The hadronization model in this sample is different from the model implemented in PYTHIA, and it represents an alternative model that also provides a reasonable agreement with the data in measurements that are sensitive to the hadronization [147]. Differences in the underlying-event tunes and models between the HERWIG++ and PYTHIA samples are not disentangled, and partial cancellation of effects or double-counting cannot be excluded.

Non-perturbative correction factors are calculated using the two alternative samples. The systematic uncertainty is calculated in each bin as the largest difference between the factors calculated using the PYTHIA AMBT1 sample and those calculated using each of the other two samples. Figure 9.3 shows the size of the corrections to the calculation of the three-to-two-jet cross-section ratio as a function of ${H_\mathrm{T}}^{(2)}$ in each of the three samples. Results are shown for jets reconstructed using resolution parameters $R = 0.4$ and $R = 0.6$. The estimated systematic uncertainty relative to the value of the correction is shown as an error band centered around 1 on the bottom of the figure. Similar results for all other NLO predictions used in this thesis can be found in Appendix E. The values of the correction

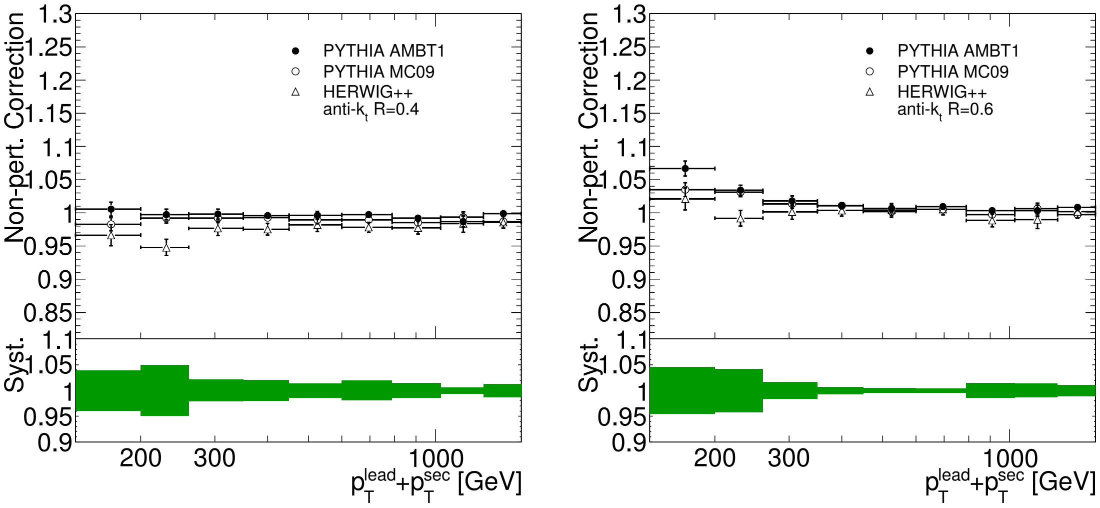

Figure 9.3: *Non-perturbative corrections to the three-to-two-jet cross-section ratio NLO prediction as a function of* $H_{\mathrm{T}}{}^{(2)}$ *and associated systematic uncertainties. The factors are shown as obtained using different leading-order Monte Carlo simulations for jets reconstructed with resolution parameters* $R = 0.4$ *(left) and* $R = 0.6$ *(right).*

factors are near 1, as expected, and the systematic uncertainties vary from ≈ 5% to ≈ 1%.

The three components of the uncertainty in the NLOJet++ calculation are calculated independently. The uncertainty due to limitations in the fixed-order calculation is calculated through variations of the renormalization and factorization scales, as described in a previous section.

The uncertainty caused by the uncertainty in the value of α_S is calculated varying the central value required by each pdf set ($\alpha_S = 0.118$ for CTEQ6.6 and $\alpha_S = 0.12018$ for MSTW2008nlo) by ± 0.002, or approximately by 2%. Since the value of α_S impacts the pdf fits, alternative pdf sets provided by the pdf fit authors are used in each case. Differences in the calculation obtained with these alternative values are taken as a systematic uncertainty.

The uncertainties caused by limitations in the fits of the non-perturbative component of the pdfs to data are calculated using the 90% confidence-limit pdf error sets provided by the authors of the pdf fits used. These error sets consist of 40 (44 for CTEQ6.6) alternative pdf sets that are obtained by varying the central value of the pdf fit by 2σ along each one of the principal components of the fit [32, 187]. The calculation is performed for each of these sets, and the results are compared to the results obtained for each central value, providing a systematic uncertainty associated with each set. Since each pdf set in the error set corresponds to a different uncorrelated degree of freedom in the fit, the systematic uncertainties obtained for each pdf set can be added in quadrature to obtain the total uncertainty in the calculation coming from the uncertainties in the fit.

Figure 9.4 shows the NLO calculation for the three-to-two-jet cross-section ratio as a function of ${H_{\mathrm{T}}}^{(2)}$ and the estimated systematic uncertainty. The size of the three individual uncertainties that contribute to the total uncertainty is shown at the bottom of the figure. Results are shown for jets built using the anti-k_t algorithm with resolution parameter $R = 0.6$ and are obtained using the MSTW2008nlo pdf set. Results for the other pdf set and other measurements performed in this thesis can be found in Appendix E.

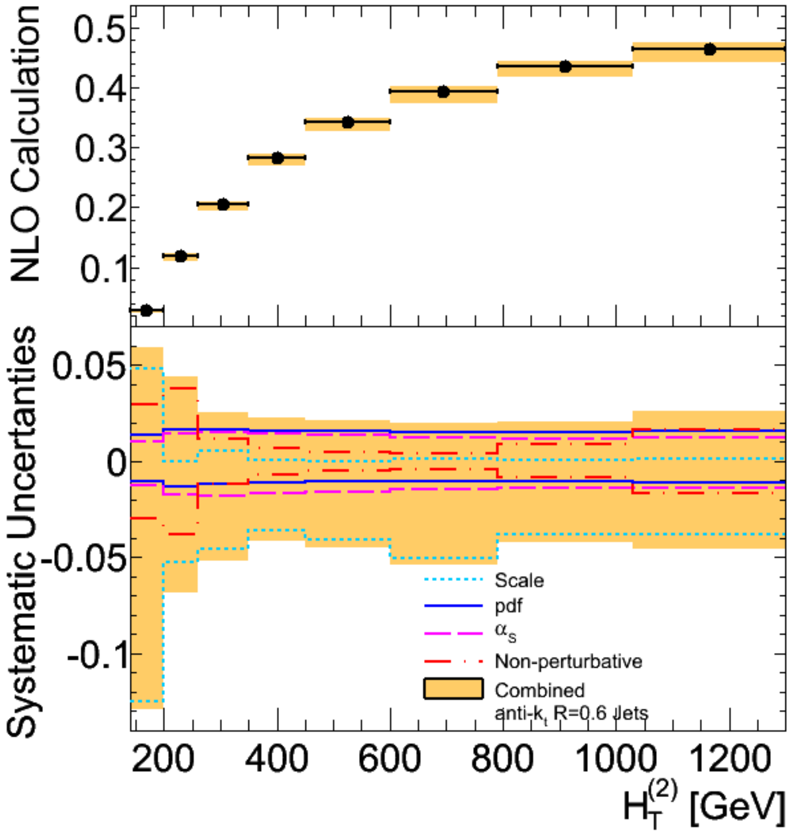

Figure 9.4: *NLO calculation of the three-to-two-jet cross-section ratio as a function of* ${H_{\mathrm{T}}}^{(2)}$ *and the estimated systematic uncertainty (band). Error bars on the predicted points corresponding to the statistical uncertainty in the calculation are small and not visible. The size of the three individual uncertainties that contribute to the total uncertainty is shown at the bottom. Results are shown for jets built using the anti-*k_t *algorithm with resolution parameter* $R = 0.6$ *and are obtained using the MSTW2008nlo pdf set.*

The scale uncertainty is ≈ 5% for most of the kinematic range of interest, and contributes the most to the total downwards uncertainty. The uncertainties caused by the uncertainties in the value of α_S and in the pdf fits are ≈ 1% and make up the majority of the upwards uncertainty in the prediction. Given that the scale uncertainties are larger than the 90% confidence-limit pdf uncertainties, this measurement does not seem promising to constrain the pdf. However, a variation in α_S of 2% changes the prediction by an amount comparable to the size of the other estimated uncertainties. This shows the potential of this observable to constrain the value of α_S at the 5–10% level. Table 9.2 summarizes the results of this calculation, including the non-perturbative corrections for illustrative purposes.

${H_{\mathrm{T}}}^{(2)}$ [GeV]	140-200	200-260	260-350	350-450
Final Prediction	0.028	0.12	0.21	0.28
Non-perturbative Correction	1.067	1.034	1.018	1.011
Statistical Uncertainty	3.8×10^{-4}	1.2×10^{-3}	1.7×10^{-3}	1.9×10^{-3}
Scale Uncertainty	$^{+1.3}_{-3.5} \times 10^{-3}$	$^{+0}_{-6.2} \times 10^{-3}$	$^{+1.0}_{-9.3} \times 10^{-3}$	$^{+0.000065}_{-0.010}$
pdf Uncertainty	$^{+3.8}_{-2.9} \times 10^{-4}$	$^{+2.0}_{-1.6} \times 10^{-3}$	$^{+3.3}_{-2.4} \times 10^{-3}$	$^{+4.5}_{-3.2} \times 10^{-3}$
α_S Uncertainty	$^{+2.9}_{-3.6} \times 10^{-4}$	$^{+1.7}_{-2.1} \times 10^{-3}$	$^{+3.1}_{-3.7} \times 10^{-3}$	$^{+4.0}_{-4.8} \times 10^{-3}$
Non-pert. Corr. Uncertainty	8.3×10^{-4}	4.5×10^{-3}	2.4×10^{-3}	2.0×10^{-3}
Total Theoretical Uncertainty	$^{+1.6}_{-3.6} \times 10^{-3}$	$^{+5.2}_{-8.1} \times 10^{-3}$	$^{+5.2}_{-11} \times 10^{-3}$	$^{+6.3}_{-12} \times 10^{-3}$
Total Theoretical Uncertainty [%]	$^{+5.7}_{-13}\%$	$^{+4.3}_{-6.8}\%$	$^{+2.5}_{-5.2}\%$	$^{+2.2}_{-4.3}\%$

${H_{\mathrm{T}}}^{(2)}$ [GeV]	450-600	600-790	790-1030	1030-1300
Final Prediction	0.34	0.39	0.44	0.46
Non-perturbative Correction	1.007	1.009	1.003	1.003
Statistical Uncertainty	2.5×10^{-3}	3.0×10^{-3}	2.8×10^{-3}	4.8×10^{-3}
Scale Uncertainty	$^{+0}_{-0.014}$	$^{+0.00056}_{-0.020}$	$^{+0.00024}_{-0.016}$	$^{+0.00057}_{-0.018}$
pdf Uncertainty	$^{+5.4}_{-3.7} \times 10^{-3}$	$^{+6.0}_{-4.1} \times 10^{-3}$	$^{+6.4}_{-4.5} \times 10^{-3}$	$^{+7.2}_{-5.3} \times 10^{-3}$
α_S Uncertainty	$^{+4.6}_{-5.4} \times 10^{-3}$	$^{+4.9}_{-5.8} \times 10^{-3}$	$^{+5.1}_{-6.0} \times 10^{-3}$	$^{+5.6}_{-6.4} \times 10^{-3}$
Non-pert. Corr. Uncertainty	1.7×10^{-3}	1.7×10^{-3}	3.8×10^{-3}	7.8×10^{-3}
Total Theoretical Uncertainty	$^{+7.3}_{-16} \times 10^{-3}$	$^{+7.9}_{-21} \times 10^{-3}$	$^{+9.0}_{-19} \times 10^{-3}$	$^{+0.012}_{-0.021}$
Total Theoretical Uncertainty [%]	$^{+2.1}_{-4.7}\%$	$^{+2.0}_{-5.4}\%$	$^{+2.0}_{-4.3}\%$	$^{+2.6}_{-4.6}\%$

Table 9.2: *NLO predictions and systematic uncertainties for each bin of the three-to-two-jet cross-section ratio calculation as a function of* ${H_{\mathrm{T}}}^{(2)}$*. Results shown use jets built with the anti-*k_t *algorithm with resolution parameter* $R = 0.6$*.*

Chapter 10

Results and Theoretical Interpretation

SEVERAL OBSERVABLES ARE USED in this chapter to characterize QCD multi-jet production at the LHC. The inclusive jet cross section and kinematic distributions of jets in multi-jet events represent natural starting points for such characterization. Given that the available $[\eta, \phi]$ space for jet production decreases with increasing multiplicity, small jet resolution parameters are typically used in studies and searches reaching high jet multiplicities, and the parameter $R = 0.4$ is used here.

These measurements can be extended to understand angular correlations between produced jets. These correlations probe specific aspects of QCD and are interesting, for example, to study in the limit of high rapidity separation between jets, where the limitations of the fixed-order calculation may be most significant [209]. Studies of such correlations have already been performed in an inclusive jet sample [210, 211], so the focus in this chapter is on a sample in which $N_{\text{jets}} \geq 3$. The study of samples with higher multiplicities requires more data and is not pursued here.

The ratios of production cross sections as a function of leading jet p_{T} and event ${H_{\text{T}}}^{(2)}$ are also studied. The use of a ratio reduces the impact of the uncertainty in the jet energy scale significantly, allowing for a more precise test of QCD. The focus is on the study of the ratio of the inclusive three-jet cross section and the inclusive two-jet cross section. The study of the relationship between three-jet and two-jet production is sensitive to final state radiation and can aid in the tuning of parton-shower Monte Carlo simulations. In addition, calculations are available to next-to-leading order, and the measurement as a function of

${H_T}^{(2)}$ is sensitive to the value of α_S. These studies are performed using jets built with the anti-k_t algorithm with resolution parameters $R = 0.4$ and $R = 0.6$. The small parameter is used for consistency with all other measurements performed, while the large parameter allows for a more precise theoretical prediction.

For the leading-order comparisons, the predictions provided by the different Monte Carlo simulations are normalized to the measured inclusive two-jet cross section and then used for shape comparisons, since no theoretical uncertainty is assigned to the predictions. While many leading-order predictions have been studied, only a few representative predictions are shown in the figures for clarity. Comparisons to the other simulations can be found in Appendix F for completeness. The normalization for all Monte Carlo simulations studied is given in Table 10.1, and distinctive features of some of the Monte Carlo simulations not shown are discussed when relevant.

Leading-order Monte Carlo	Normalization factor
ALPGEN+HERWIG/JIMMY AUET1	1.11
ALPGEN+HERWIG/JIMMY MC09	1.12
ALPGEN+PYTHIA MC09′	1.22
ALPGEN+PYTHIA Perugia6	1.15
ALPGEN+PYTHIA D6	1.62
PYTHIA AMBT1	0.65
PYTHIA Perugia2010	0.86
PYTHIA MC09	0.66
PYTHIA MC09′	0.91
SHERPA	1.06
HERWIG++	0.86

Table 10.1: *Normalizations applied to match the predicted and measured inclusive two-jet cross sections in each of the leading-order Monte Carlo simulations studied.*

Most ALPGEN Monte Carlo simulations predict an inclusive two-jet cross section similar to the measured cross section, except for the ALPGEN+PYTHIA D6 sample, while the PYTHIA simulations require scaling factors that are further from 1 to describe the inclusive multiplicity spectrum. The normalization factor for SHERPA is the closest to unity.

The differences in the normalization factors required for ALPGEN+PYTHIA MC09′ and ALPGEN+HERWIG/JIMMY AUET1 largely illustrate differences between PYTHIA and HERwig/Jimmy and how they interplay with the matrix-element and parton-shower matching implemented in Alpgen. Differences between the different Alpgen+Pythia tunes illustrate the large impact of the underlying-event and parton-shower tune on the leading-order predictions. In particular, the normalization is largest for the Alpgen+Pythia D6 tune, which is the only Pythia tune that uses virtuality ordering in the parton shower. Based on the results in Ref. [199], it has been hypothesized that the different parton-shower models available in Pythia and Herwig may, through the MLM matching used in Alpgen, have a significant effect on the overall prediction of the cross sections. A thorough understanding of this issue requires further studies, and perhaps dedicated tuning of Alpgen+Pythia or Alpgen+Herwig/Jimmy.

Table 10.2 details the measurements shown in this chapter, and the section in which they are presented. Certain kinematic distributions are only studied for events with up to $N_{\mathrm{jets}} \geq 4$ due to the limited number of events collected in higher-multiplicity bins. The studies of the cross-section ratio as a function of leading jet p_{T} are performed with three different sets of kinematic cuts as described in Section 6.3.1, to provide additional data points in future studies attempting to tune final state radiation in parton-shower leading-order Monte Carlo simulations. These studies and those of the cross-section ratio as a function of ${H_{\mathrm{T}}}^{(2)}$ are performed using jets built with the anti-k_t algorithm with resolution parameters $R = 0.4$ and $R = 0.6$, but only the studies for $R = 0.6$ are shown in Section 10.3. Results for $R = 0.4$ lead to similar conclusions and can be found in Appendix F.

10.1 Inclusive Multiplicity and Kinematic Distributions

Measurements presented in this section are performed using jets built with the anti-k_t algorithm with $R = 0.4$ and compared to leading-order Monte Carlo simulations. In the figures, the orange error band bracketing the measured cross section corresponds to the total systematic uncertainty in the measurement excluding the uncertainty in the luminosity. The ratio of the predictions from the Monte Carlo simulations to the measurements is shown at the bottom of each figure.

Figure 10.1 presents the measurement of the jet cross section as a function of inclusive jet multiplicity. The measurement, with the systematic uncertainties split in their three main

Observable	Resolution Parameter	Calculation	Section
$\sigma_{\geq n\text{ jets}}$ for $n = 2-6$ $\sigma_{\geq n\text{ jets}}/\sigma_{\geq n-1\text{ jets}}$ for $n = 3-6$ $d\sigma/dp_{\mathrm{T}}$ for leading jet ($N_{\text{jets}} \geq 2$) $d\sigma/dp_{\mathrm{T}}$ for 2nd leading jet ($N_{\text{jets}} \geq 2$) $d\sigma/dp_{\mathrm{T}}$ for 3rd leading jet ($N_{\text{jets}} \geq 3$) $d\sigma/dp_{\mathrm{T}}$ for 4th leading jet ($N_{\text{jets}} \geq 4$) $d\sigma/dH_{\mathrm{T}}$ for events with $N_{\text{jets}} \geq 2$ $d\sigma/dH_{\mathrm{T}}$ for events with $N_{\text{jets}} \geq 3$ $d\sigma/dH_{\mathrm{T}}$ for events with $N_{\text{jets}} \geq 4$	$R = 0.4$	LO	10.1
$\sigma^{-1}_{\geq 3\text{ jets}} d\sigma/d\|\Delta\phi_{12}\|$ ($N_{\text{jets}} \geq 3$) $\sigma^{-1}_{\geq 3\text{ jets}} d\sigma/d\|\Delta\phi_{13}\|$ ($N_{\text{jets}} \geq 3$) $\sigma^{-1}_{\geq 3\text{ jets}} d\sigma/d\|\Delta\phi_{23}\|$ ($N_{\text{jets}} \geq 3$) $\sigma^{-1}_{\geq 3\text{ jets}} d\sigma/d\|\Delta\eta_{12}\|$ ($N_{\text{jets}} \geq 3$) $\sigma^{-1}_{\geq 3\text{ jets}} d\sigma/d\|\Delta\eta_{13}\|$ ($N_{\text{jets}} \geq 3$) $\sigma^{-1}_{\geq 3\text{ jets}} d\sigma/d\|\Delta\eta_{23}\|$ ($N_{\text{jets}} \geq 3$)	$R = 0.4$	LO	10.2
$[d\sigma/d{p_{\mathrm{T}}}^{\text{lead}}]^{\geq 3\text{ jets}}/[d\sigma/d{p_{\mathrm{T}}}^{\text{lead}}]^{\geq 2\text{ jets}}$, with ${p_{\mathrm{T}}}^{\text{lead}} \geq 80$ GeV, $p_{\mathrm{T}} \geq 60$ GeV	$R = 0.6$ ($R = 0.4$)	LO/NLO	10.3
$[d\sigma/d{p_{\mathrm{T}}}^{\text{lead}}]^{\geq 3\text{ jets}}/[d\sigma/d{p_{\mathrm{T}}}^{\text{lead}}]^{\geq 2\text{ jets}}$, with ${p_{\mathrm{T}}}^{\text{lead}} \geq 110$ GeV, $p_{\mathrm{T}} \geq 80$ GeV	$R = 0.6$ ($R = 0.4$)	LO/NLO	10.3
$[d\sigma/d{p_{\mathrm{T}}}^{\text{lead}}]^{\geq 3\text{ jets}}/[d\sigma/d{p_{\mathrm{T}}}^{\text{lead}}]^{\geq 2\text{ jets}}$, with ${p_{\mathrm{T}}}^{\text{lead}} \geq 160$ GeV, $p_{\mathrm{T}} \geq 110$ GeV	$R = 0.6$ ($R = 0.4$)	LO/NLO	10.3
$[d\sigma/d{H_{\mathrm{T}}}^{(2)}]^{\geq 3\text{ jets}}/[d\sigma/d{H_{\mathrm{T}}}^{(2)}]^{\geq 2\text{ jets}}$,	$R = 0.6$ ($R = 0.4$)	LO/NLO	10.3

Table 10.2: *Summary of the observables measured in this thesis, including the resolution parameter in the jet algorithm used for the measurement, the theoretical calculations the measurements are compared to and the section where they are presented.*

components described in Chapter 8, is given in Table 10.3. The measurement is limited by the systematic uncertainties in all bins. The systematic uncertainty is dominated by the uncertainty in the measurement of the jet energy and ranges from 10-20% at low multiplicity to almost 30–40% at high multiplicity. The predictions from the Monte Carlo simulations fall on the measured results across the full inclusive multiplicity spectrum after the appropriate normalizations are applied. Differences between the different simulations are $\approx 10\%$ except for events with six jets, for which differences reach $\approx 30\%$. Such differences essentially fall within the systematic uncertainties in the measurement.

In order to allow for a more precise understanding of the shapes of these distributions,

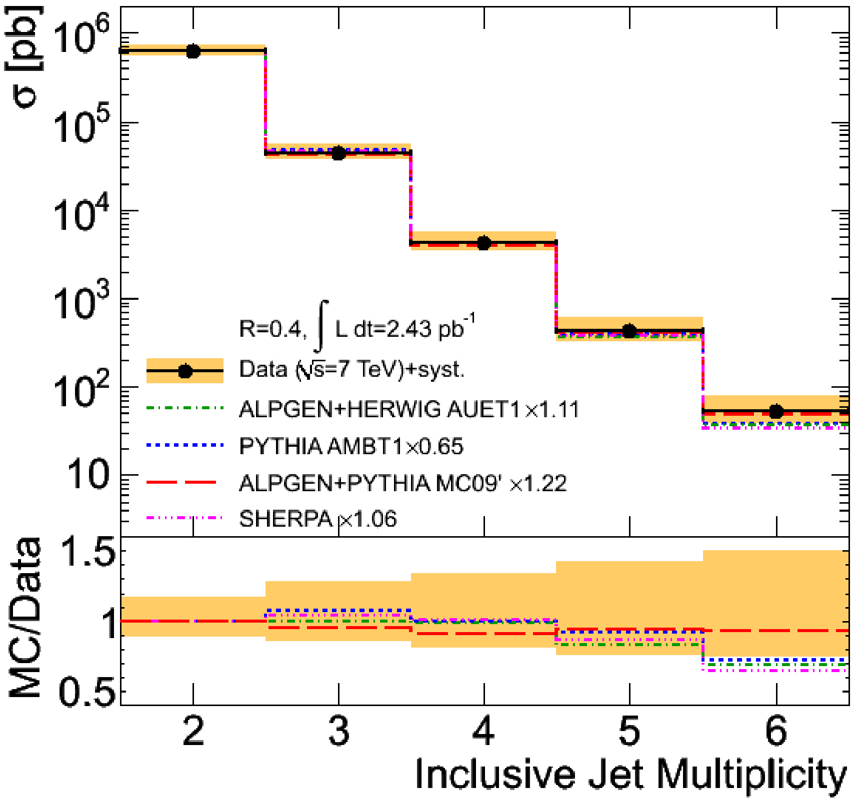

Figure 10.1: *Cross section as a function of inclusive jet multiplicity for jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$. The results are compared to leading-order Monte Carlo simulations normalized to the measured two-jet inclusive cross section. The error bands correspond to the systematic uncertainty in the measurement. A plot of the ratio of the different Monte Carlo simulations to the data is presented at the bottom of the figure.*

Multiplicity	≥ 2	≥ 3	≥ 4	≥ 5	≥ 6
Measured Value	6.2×10^5	4.3×10^4	4.3×10^3	4.3×10^2	53
Statistical Uncertainty	1.3×10^3	1.3×10^2	41	13	4.9
JES Uncertainty	${}^{+11}_{-6.6}\times10^4$	${}^{+12}_{-6.2}\times10^3$	${}^{+14}_{-7.9}\times10^2$	${}^{+180}_{-94}$	${}^{+24}_{-8.8}$
Unfolding Uncertainty	2.4×10^4	1.7×10^3	2.4×10^2	44	10
Luminosity Uncertainty	2.1×10^4	1.5×10^3	1.5×10^2	15	1.8
Total Systematic Uncertainty	${}^{+11}_{-7.3}\times10^4$	${}^{+12}_{-6.6}\times10^3$	${}^{+14}_{-8.4}\times10^2$	${}^{+1.9}_{-1.0}\times10^2$	${}^{+27}_{-14}$

Table 10.3: *Measured cross section as a function of inclusive jet multiplicity for jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ and corresponding uncertainties.*

the n-jet to $(n-1)$-jet cross-section ratio can be used. The impact of the systematic uncertainty in the jet energy scale on this measurement is small and the measurement of the luminosity is not necessary. Figure 10.2 shows the results of such a study. The results

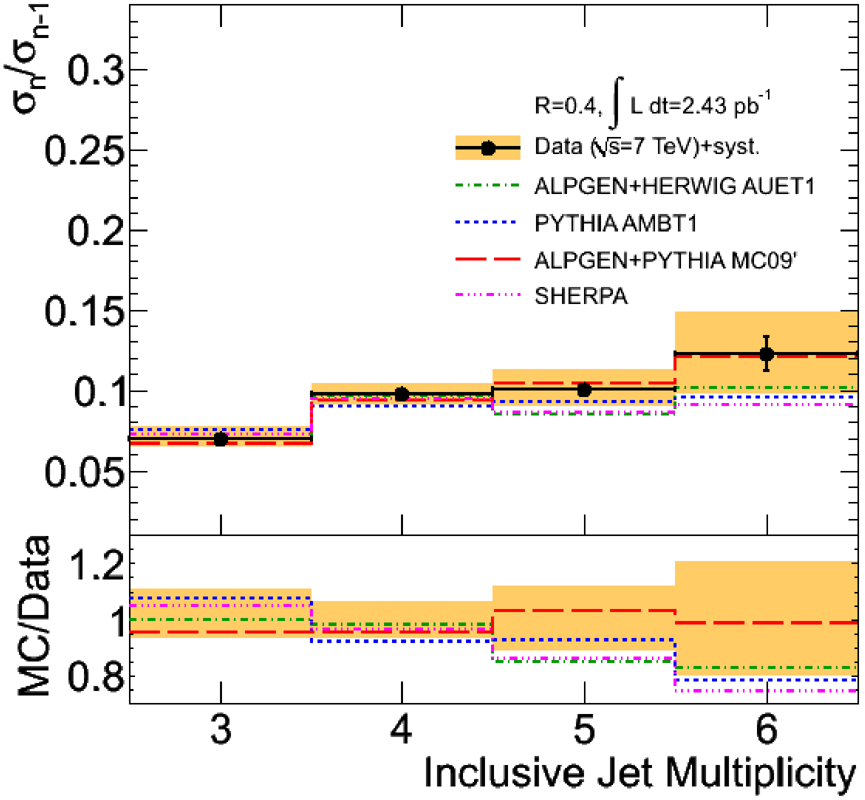

Figure 10.2: *Ratio of the n-jet cross section to the $(n-1)$-jet cross section for values of n varying from three to six for jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$. The results are compared to leading-order Monte Carlo simulations. The error bands correspond to the systematic uncertainty in the measurement. A plot of the ratio of the different Monte Carlo simulations to the data is presented at the bottom of the figure.*

with the corresponding uncertainties are given in Table 10.4. Both the unfolding and the jet energy scale uncertainties contribute comparably to the total systematic uncertainty, whereas the statistical uncertainties are smaller than the systematic uncertainties, and negligible in most bins. All leading-order predictions agree reasonably well with the data within systematic uncertainties.

The Pythia AMBT1 Monte Carlo simulation predicts a somewhat higher (lower) three-to-two-jet (four-to-three-jet) cross-section ratio, while the Alpgen+Herwig/Jimmy and

Ratio ($\sigma_{\geq n}/\sigma_{\geq n-1}$)	3	4	5	6
Measured Value	0.070	0.098	0.101	0.123
Statistical Uncertainty	2.4×10^{-4}	9.0×10^{-4}	2.9×10^{-3}	0.011
JES Uncertainty	${}^{+6.5}_{-2.9}\times10^{-3}$	${}^{+3.8}_{-4.8}\times10^{-3}$	${}^{+6.4}_{-4.2}\times10^{-3}$	${}^{+8.2}_{-0}\times10^{-3}$
Unfolding Uncertainty	3.7×10^{-3}	4.9×10^{-3}	0.010	0.025
Total Systematic Uncertainty	${}^{+7.4}_{-4.7}\times10^{-3}$	${}^{+6.3}_{-6.8}\times10^{-3}$	${}^{+0.012}_{-0.011}$	0.025

Table 10.4: *Measured n-jet to $(n-1)$-jet cross-section ratio as a function of inclusive jet multiplicity for jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ and corresponding uncertainties.*

Sherpa predictions are closer to the data at low multiplicities, but predict a lower ratio at the highest multiplicities. Differences at the level of 15% are observed between Pythia AMBT1 and Alpgen+Pythia MC09′ in the first bin. These differences arise from the difference between the ME+PS and the pure parton-shower Monte Carlo simulation, since the prediction in the Pythia MC09′ sample has been verified to be similar to that from the Pythia AMBT1 sample. All Alpgen+Pythia tunes studied produce similar results, despite the different normalizations given in Table 10.1. These additional studies can be found in Appendix F.

The differential cross sections as a function of jet p_{T} for multi-jet events are essential observables in the characterization of the kinematic features of such events. Figure 10.3 presents the differential cross sections as a function of the leading, second leading, third leading and fourth leading jet p_{T} in multi-jet events. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ are used in this study. The systematic uncertainty in the measurement is 10–20% across p_{T} and increasing up to 30% for the differential cross section as a function of the fourth leading jet. The jet energy scale systematic uncertainty remains the dominant uncertainty in the measurement. Despite the relatively large uncertainties, differences between the Monte Carlo simulations and the data are significative since the uncertainty in the measurement is strongly correlated across bins.

The Alpgen+Herwig/Jimmy AUET1 and Alpgen+Pythia MC09′ Monte Carlo simulations are in agreement with the data, even though they tend to predict a shape that decays less steeply in two-jet and three-jet events. This effect is slightly larger in the Sherpa Monte Carlo simulation, whose behavior is otherwise similar to that of the Alpgen tunes. The shapes of the other Alpgen+Pythia Monte Carlo simulations studied are similar to

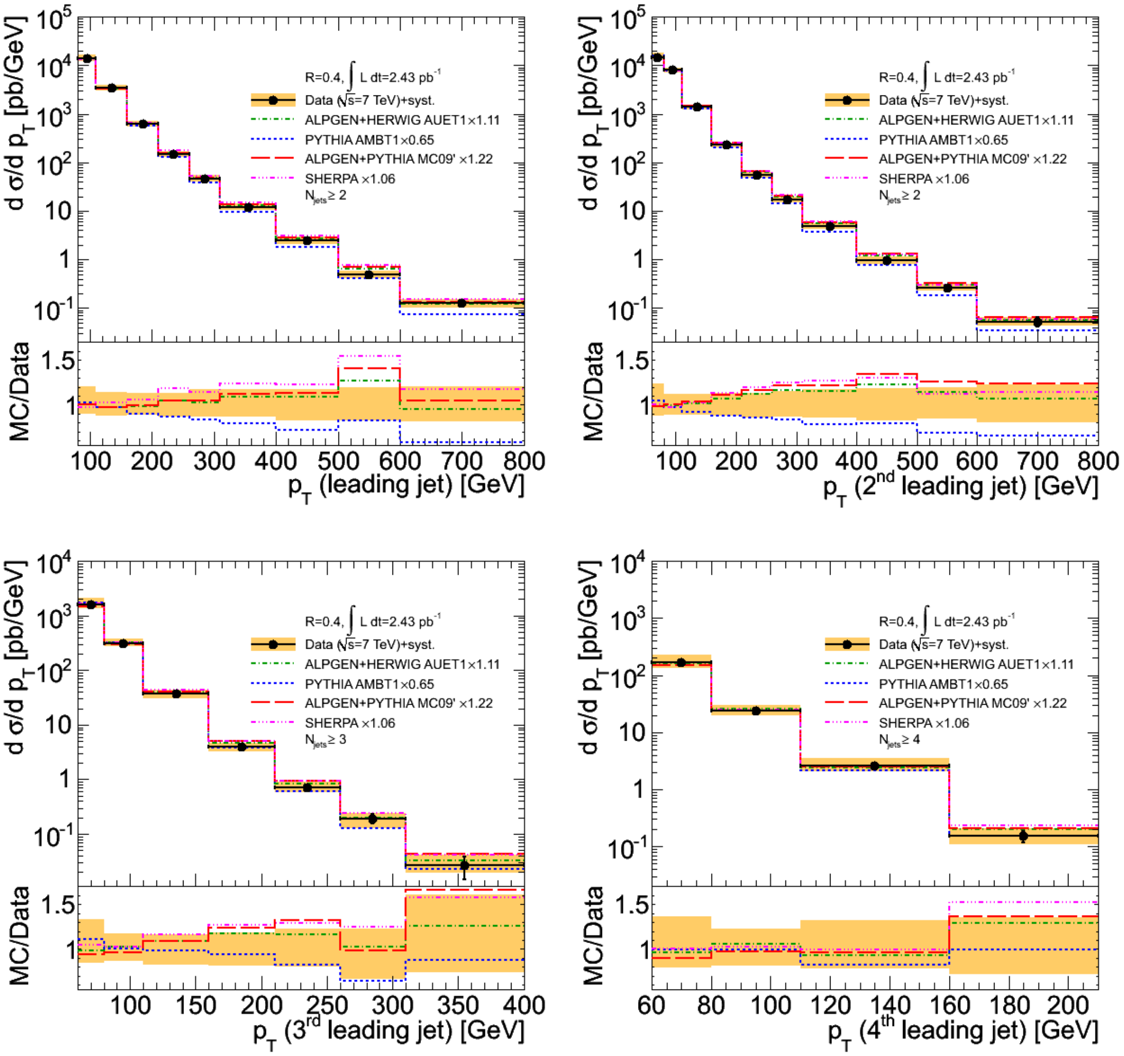

Figure 10.3: *Differential cross section as a function of leading jet $p_{\rm T}$ for events with $N_{\rm jets} \geq 2$ (top left), second leading jet $p_{\rm T}$ for events with $N_{\rm jets} \geq 2$ (top right), third leading jet $p_{\rm T}$ for events with $N_{\rm jets} \geq 3$ (bottom left) and fourth leading jet $p_{\rm T}$ for events with $N_{\rm jets} \geq 4$ (bottom right). Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ are used in this study. The results are compared to different leading-order Monte Carlo simulations normalized to the measured two-jet inclusive cross section. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

those of the Alpgen tunes shown, except for the D6 tune, which follows the data even more closely. The shape of the Pythia AMBT1 is steeper than that of the data in the two-jet bin. A similar feature is observed in the Pythia Perugia2010 and Herwig++ predictions, the effect being even more pronounced in the two-jet bin when using Herwig++.

A similar study, targeting a global event variable is the study of the differential cross section for multi-jet production as a function of event H_{T}, that is, the scalar sum of the p_{T}s of selected jets in the event. These distributions are commonly used in studies of the top quark and searches for new physics. Figure 10.4 presents the results of the measurement of the differential production cross section as a function of event H_{T} for events falling in the first three multiplicity bins. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ are used in this study. The systematic uncertainties are similar to those presented in the previous figure, increasing with jet multiplicity. Similar trends are observed in the description of the data provided by the different Monte Carlo simulations as those discussed for the previous figure.

10.2 Angular Distributions

The results for the studies of distributions of the angular separation between selected jets for events with at least three jets are shown in this section. The focus is on the understanding of the shapes of these distributions. For this reason, the measurement is normalized to the total cross section for production of events with three or more jets. This cancels luminosity and to a large extent jet energy scale uncertainties, and the uncertainties in the unfolding corrections dominate the uncertainties in the measured quantities. Unlike in the studies shown in the rest of this chapter, statistical uncertainties in the leading-order predictions are not small for the results presented in this section, even though they are not shown for clarity in the presentation.

Figure 10.5 shows the measured distribution of the angular separations in the ϕ direction between the first and second, first and third and second and third leading jets in the event. The distributions are truncated to regions where nearby selected jets do not contribute significantly to the cross section due to a lack of studies of the angular resolution in the reconstruction of such jets. These studies are needed to appropriately establish a system-

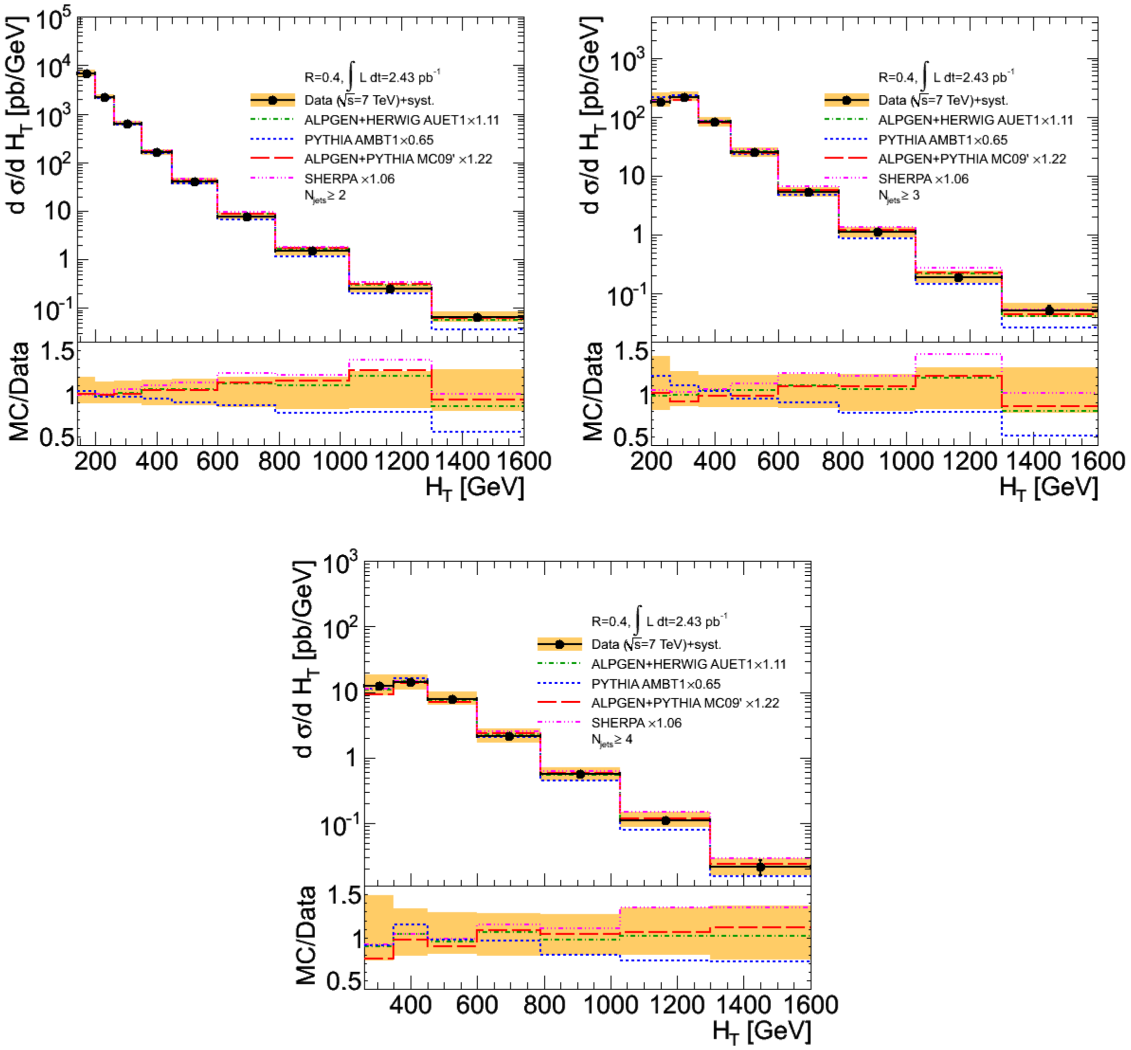

Figure 10.4: *Differential cross section as a function of H_{T} for events with at least two (top left), three (top right) and four (bottom) selected jets. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ are used in this study. The results are compared to different leading-order Monte Carlo simulations normalized to the measured two-jet inclusive cross section. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

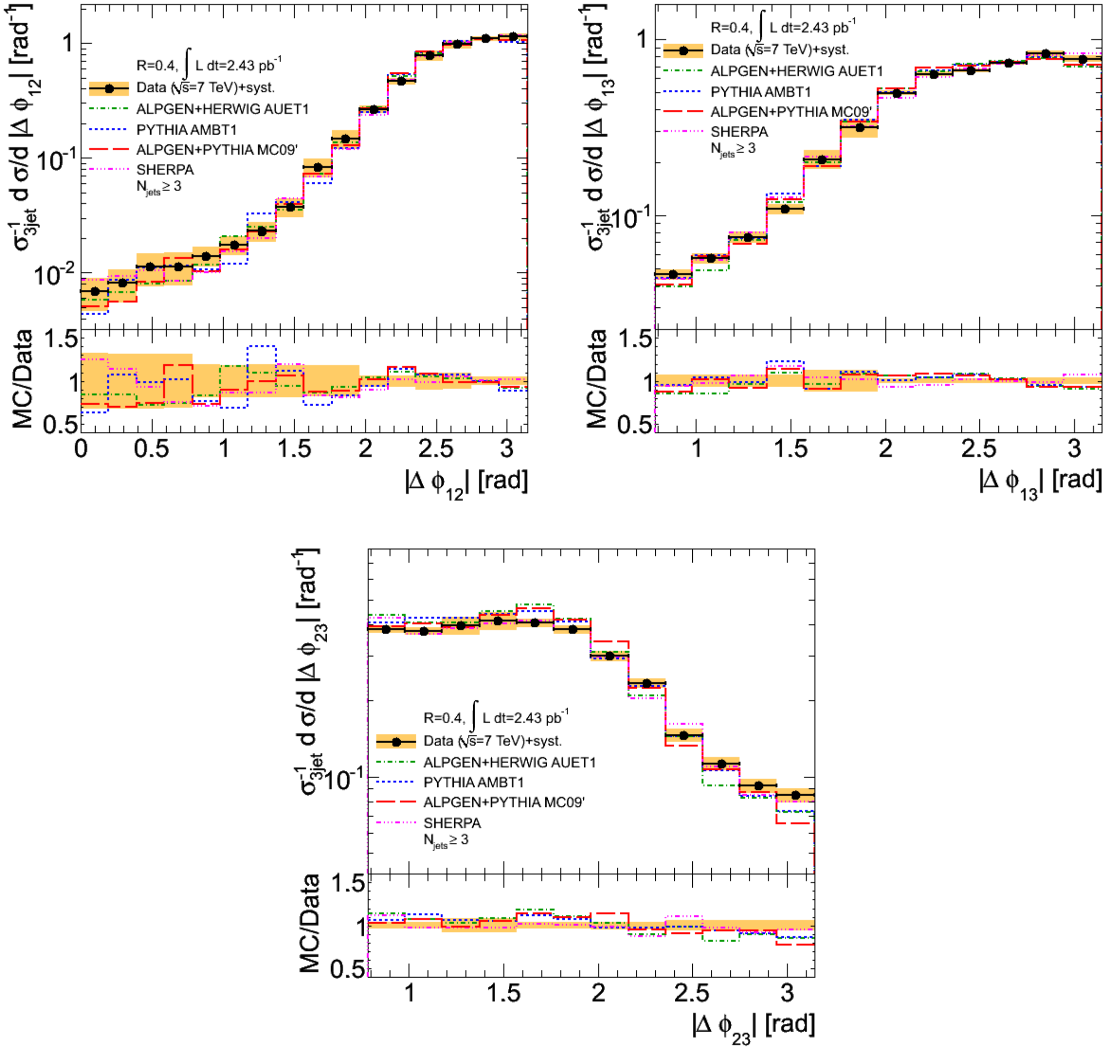

Figure 10.5: *Relative differential cross section as a function of $|\Delta\phi|$ between the leading and second leading (top left), leading and third leading (top right) and second leading and third leading (bottom) jets for events with three or more jets in data and leading-order Monte Carlo simulations. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ are used in this study. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

atic uncertainty in the measurement, and those regions, where the systematic uncertainty estimate is not reliable, are thus not shown. Given the statistical uncertainties in the Monte Carlo simulations, noticeable through the bin-to-bin fluctuations in the predictions, agreement between the different predictions and the data is observed. No significant differences between the different leading-order predictions are observed either.

Figure 10.6 shows the measured distribution of the angular separations in the η direction between the first and second, first and third and second and third leading jets in the event. The distributions are again truncated to regions where nearby selected jets do not contribute

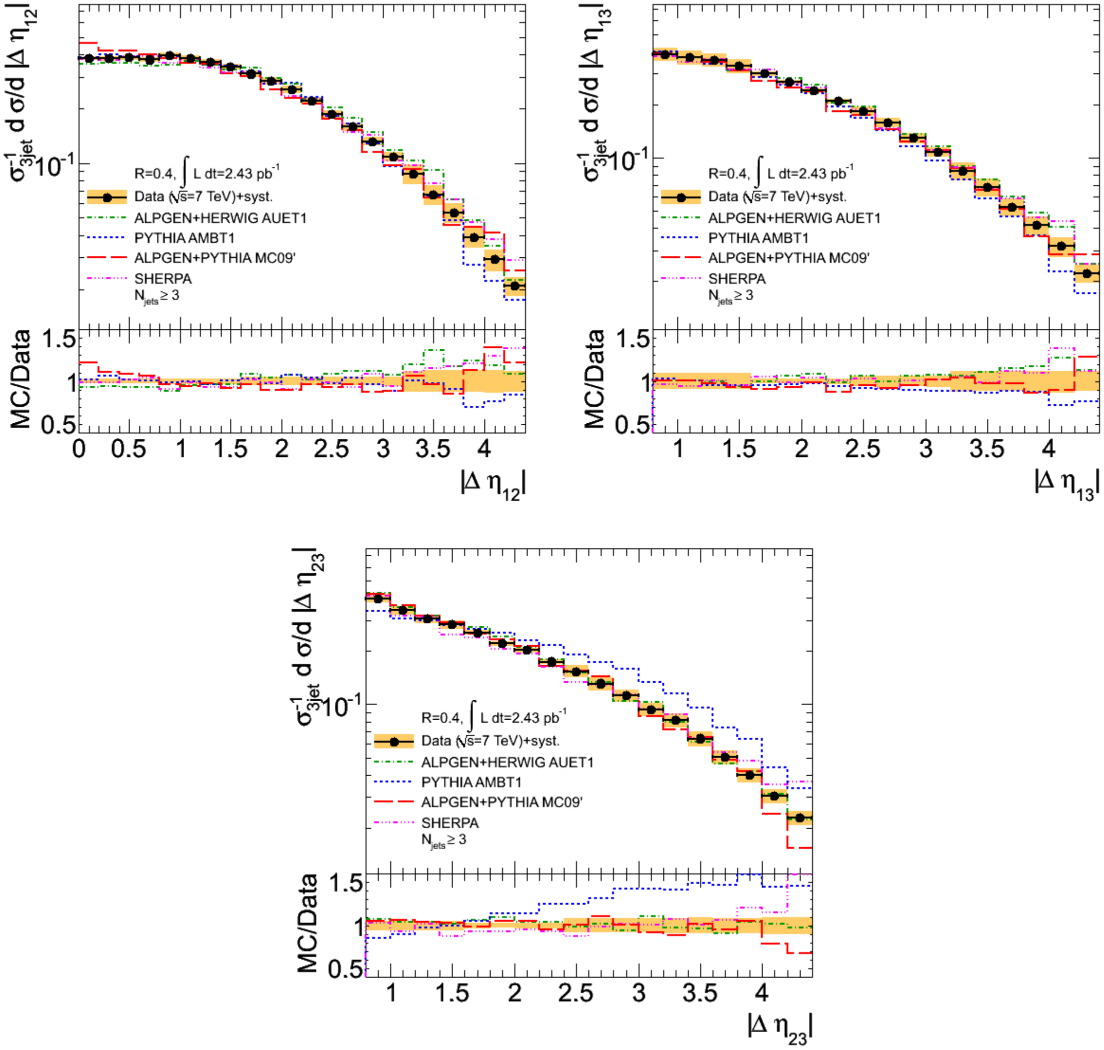

Figure 10.6: *Relative differential cross section as a function of $|\Delta\eta|$ between the leading and second leading (top left), leading and third leading (top right) and second leading and third leading (bottom) jets for events with three or more jets in data and leading-order Monte Carlo simulations. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ are used in this study. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

significantly to the cross section. Note that the selection cut on the jet rapidity constrains the values of these variables to be $\lesssim$5.6. However, the $|\Delta\eta|$ ranges shown in Figure 10.6 are truncated earlier due to the extremely low Monte Carlo simulation statistics available at very large $\Delta\eta$, which do not allow for meaningful comparisons.

The Alpgen and Sherpa Monte Carlo simulations provide a reasonably good description of all the three $\Delta\eta$ distributions. At low $|\Delta\eta|$ between the first and second leading jet, Alpgen+Pythia predicts a higher rate. This effect is not observed in the Pythia simulation, but can be also observed in the samples produced with the Alpgen+Pythia D6 and Perugia6 tunes. The effect is slightly smaller for the sample generated with the Perugia6 tune, hinting that it arises from the interplay between Alpgen and Pythia, but it is also sensitive to the details of the tune. The Pythia simulation describes the distributions involving the leading jet well, but predicts a much higher rate at large angular separations between the second and third leading jets than the measured rate. This effect is not observed in any of the Alpgen tunes studied. A similar prediction has been observed in the other Pythia samples studied [212], pointing to a limitation in the $2 \to 2$ calculation.

10.3 Three-to-Two-Jet Cross-Section Ratios

The measurement of the three-to-two-jet cross-section ratio as a function of kinematic variables in the event has particular sensitivity to limitations in the leading-order Monte Carlo simulations and a NLO pQCD calculation can be performed. As in the measurements presented in the previous section, the uncertainty in the luminosity determination cancels out and uncertainties in the jet energy scale are reduced in this measurement.

The three-to-two-jet cross-section ratio as a function of the leading jet p_{T} has been shown to be sensitive to effects from final state radiation in previous experiments [207]. Figure 10.7 presents the results on this measurement using three different sets of kinematic cuts on the jet selection, as detailed in Section 6.3.1. Jets built with the anti-k_t algorithm with resolution parameter $R = 0.6$ are used in this study. Similar results are obtained with jets built with resolution parameter $R = 0.4$ and this is shown in Appendix F.

The systematic uncertainties in the measurement are small ($\approx$5%), except in the lowest p_{T} bin in the measurement with the default kinematic cuts, where the uncertainties in the unfolding correction and the jet energy scale are both important. Whereas the Alpgen+Herwig/Jimmy AUET1 and Alpgen+Pythia simulations describe the data well, the Pythia AMBT1 simulation predicts a higher ratio than what is measured over the p_{T}

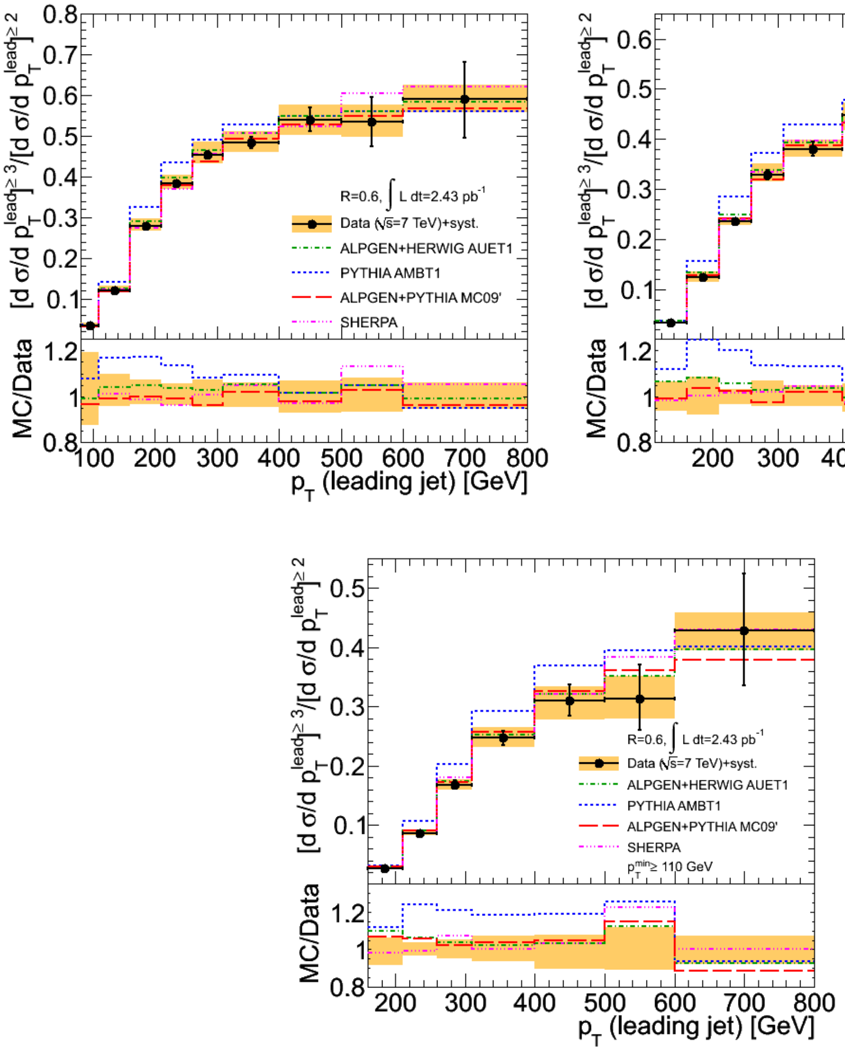

Figure 10.7: *Ratio of the three-jet to two-jet differential cross section as a function of the leading jet p_{T} using three different kinematic cuts on the jet selection. Jets built with the anti-k_t algorithm using a resolution parameter $R = 0.6$ are used in this study. The results are compared to leading-order Monte Carlo simulations. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

range from 200 GeV to 800 GeV. The disagreement is similar when other Pythia tunes or Herwig++ are used, pointing for the need of final state radiation tuning for these models to remain viable to predict such observables. A good agreement between data and simulation is also found if Sherpa or any other of the Alpgen simulations are used. Similar conclusions are reached for all the simulations as the minimum cut on the p_{T} of selected jets is increased.

These measurements can also be used to test the NLO predictions. Figure 10.8 presents the comparison of the measurements to the NLO calculations. The systematic uncertainties

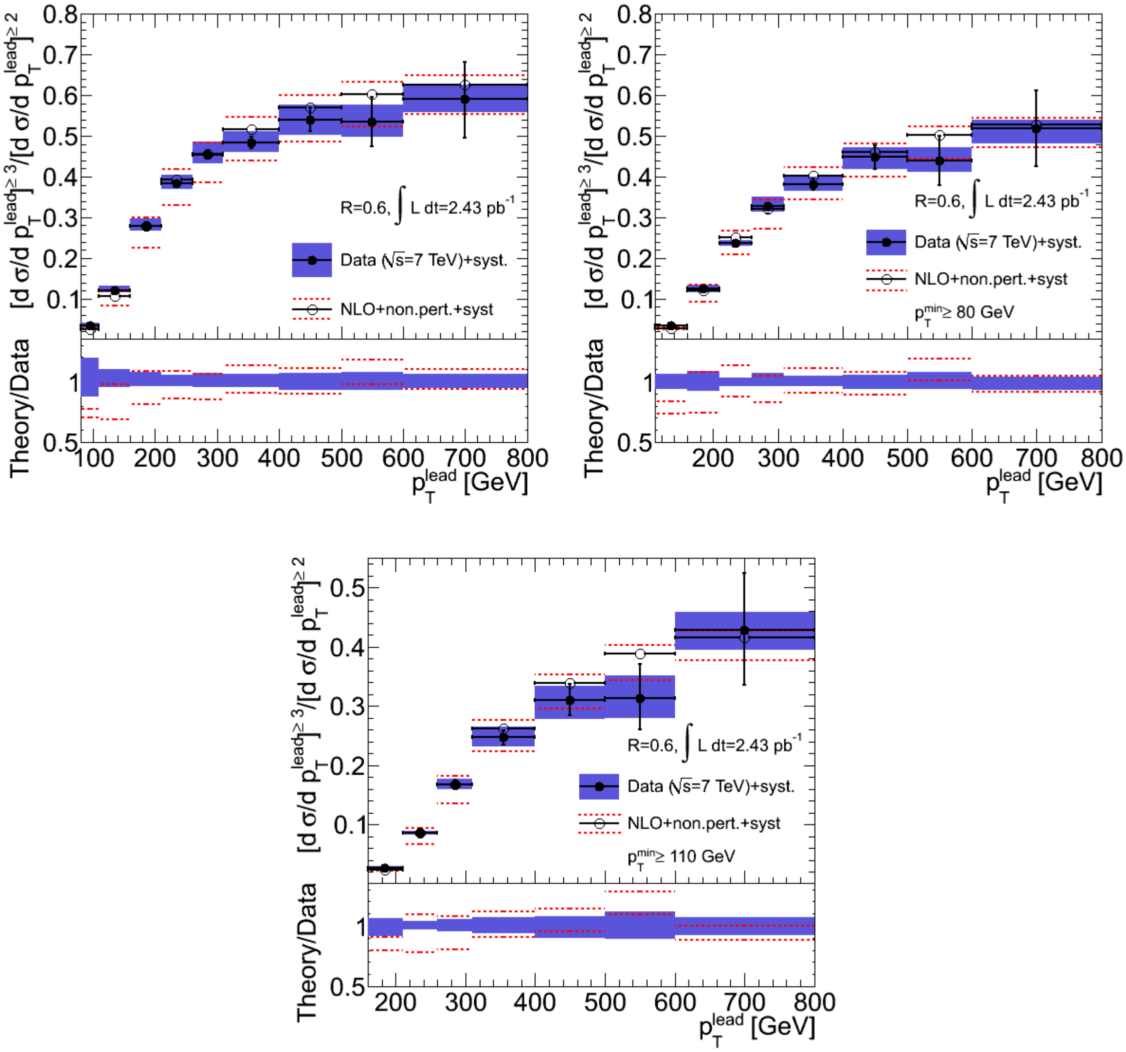

Figure 10.8: *Ratio of the three-jet to two-jet differential cross-section as a function of the leading jet p_{T} using three different kinematic cuts on the jet selection. Jets built with the anti-k_t algorithm using a resolution parameter $R = 0.6$ are used in this study. The results are compared to a next-to-leading-order calculation. The error bands correspond to the systematic uncertainty in the measurement, while the dotted lines represent the uncertainties in the theoretical prediction. Plots of the ratios of the NLO calculation to the data are presented at the bottom of each figure.*

in the theoretical predictions are shown as dotted red lines above and below the theoretical prediction. The `MSTW2008nlo` pdf set has been used in this figure, but comparable results are obtained with the `CTEQ 6.6` pdf set (see Appendix F). Jets built with the anti-k_t

algorithm using a resolution parameter $R = 0.6$ are used in this study, but similar results are obtained when using $R = 0.4$, except that the systematic uncertainties in the NLO predictions are larger.

The NLO calculation describes the data well, except in the lowest p_{T} bin of each ratio, where there is a discrepancy. This discrepancy is likely due to the treatment of virtual diagrams in the calculation, or other limitations in the NLO calculation. The discrepancy diminishes significantly once the minimum p_{T} for all jets is raised to 110 GeV and the p_{T} of the leading jet is required to be greater than 160 GeV. This is probably caused by the increased asymmetry between leading and sub-leading jet p_{T} cuts, which decreases the region of phase space in which diagrams with real emissions are excluded from the calculation through the kinematic cuts.

The systematic uncertainty in the theoretical prediction is larger in most bins than the uncertainty in the measurement of the ratios as a function of ${p_{\mathrm{T}}}^{\mathrm{lead}}$. Since most of the uncertainty in the prediction arises from the intrinsic limitation of a fixed-order calculation due to missing higher-order terms, this measurement cannot be used to constrain the parameters of the calculation. As shown in the previous chapter, the measurement of the three-to-two-jet cross-section ratio as a function of the event ${H_{\mathrm{T}}}^{(2)}$ can be used for that purpose. Figure 10.9 shows this measurement compared to both leading-order Monte Carlo simulations and the NLO calculation. The MSTW2008nlo pdf set has been used in this figure, but comparable results are obtained with the CTEQ 6.6 pdf set (see Appendix F). Jets built with the anti-k_t algorithm using a resolution parameter $R = 0.6$ are used in this study, but similar results are obtained when using $R = 0.4$, except that the systematic uncertainty in the NLO prediction is larger.

The leading-order predictions show similar features to those observed in the previous study. The theoretical uncertainty in the NLO calculation is comparable to the uncertainty in the measurement, and is significantly reduced with respect to the theoretical uncertainty in the previous study. A large disagreement between the measurement and the NLO calculation in the lowest ${H_{\mathrm{T}}}^{(2)}$ bin is still visible, and it is caused by the same effects discussed earlier in this section. Elsewhere, the NLO calculation agrees with the measurement.

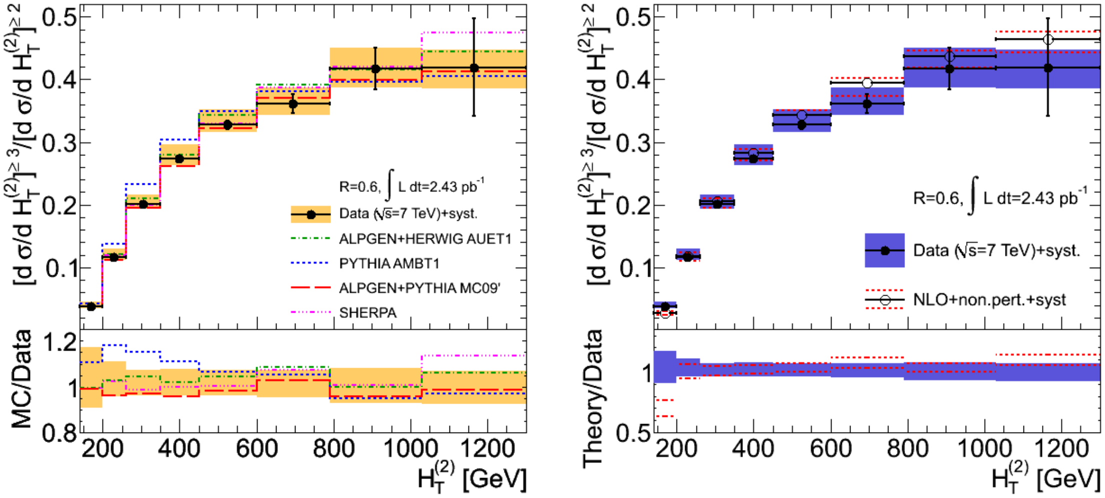

Figure 10.9: *Ratio of the three-jet to two-jet differential cross section as a function of* ${H_T}^{(2)}$. *Jets built with the anti-*k_t *algorithm using a resolution parameter* $R = 0.6$ *are used in this study. The results are compared to leading-order Monte Carlo simulations (left) and a next-to-leading-order calculation (right). The error bands correspond to the systematic uncertainty in the measurement, while the dotted lines on the right figure represent the uncertainty in the theoretical prediction. Plots of the ratios of the different Monte Carlo simulations (left) and the NLO calculation (right) to the data are presented at the bottom of each figure.*

Conclusions and Future Prospects

The physics program of the LHC is currently at its early stages of development, and promises to deepen our understanding of the Standard Model and potentially change many of the current prevalent ideas in particle physics. In this thesis, the characterization of a fundamental final state in this physics program has been performed in the context of the study of QCD processes. Such processes occur at the LHC with high probability, and produce events that are backgrounds to many searches of new particles. In particular, the production of events with multiple jets has been studied, and several production cross sections and related observables have been measured.

Comparisons of the measurements to leading-order and next-to-leading-order calculations have been performed. Comparisons to leading-order calculations reveal certain limitations in the description provided by parton-shower Monte Carlo simulations, even though other observables are well described. Such limitations disappear when matrix-element calculations of the leading-order production diagrams with many partons in the final state are combined with parton-shower simulations. The dependence of the predictions on different parameters entering the leading-order Monte Carlo simulations has also been investigated, revealing stability against variations of such parameters. These studies are already helping the tuning effort for the Monte Carlo simulations that will be used in analyses of ATLAS data collected during 2011 [213].

Comparisons to next-to-leading-order calculations reveal limitations in the calculations in certain regions of phase space. Outside of those regions agreement between the predicted values and the data is found. Observables have been found that can help in constraining some of the input parameters of the NLO calculation through measurements and they have been studied with data. The analysis of the uncertainties in the NLO calculation reveals that these measurements are mostly sensitive to the value of the strong coupling constant

α_S, even though sensitivity to other parameters, such as the parton distribution functions, cannot be excluded if the kinematic range of the study is increased. This sensitivity has not been exploited in this thesis to constrain the value of α_S, but this should be possible in the near future with an uncertainty of $\approx$ 5–10%.

The conclusions reached about the different calculations apply only to the observables investigated in this thesis. Additional observables are likely to be studied as more data become available. Such studies may include more in-depth studies of the angular correlations in multi-jet events, or jet invariant–mass distributions in multi-jet final states. These studies should help deepen our understanding of multi-jet event topologies in ways that will be required to observe lurking new physics in the new energy regime.

The studies shown in this thesis can be used to apply k-factors to Monte Carlo–based QCD background estimates for other analyses in which QCD is a background. However, the studies concentrate on specific regions of phase space, and k-factors can vary according to analysis cuts. For this reason, the studies in this thesis are not necessarily most useful to that effect, and data-driven background estimates will be preferred most often. In that case, the analysis performed in this thesis has shown how to treat many effects that are important in the measurement of multi-jet final states and will serve as a guide for other analyses using events with multiple jets.

Some of the important effects studied in this analysis include the impact of complex event topologies in the jet triggers and the jet energy scale. Equally important are the studies that provide an understanding of the impact of flavor on the measurement of the jet energy. A technique to characterize the flavor composition of a sample of light quarks and gluons has been used to reduce the systematic uncertainty in the jet energy scale. This technique can also be used to characterize jets in a sample with new particles decaying into jets, and such characterization could help in understanding the nature of such particles.

The 2010 run at the LHC established many experimental techniques and probed the performance of the ATLAS detector and SM physics at the new energy frontier. These studies will be useful for analyses in years to come, and form the building blocks for future analyses. This thesis was devoted to the development of some of these techniques as well as some initial physics studies in the hope that these exercises will help uncover new physics at the LHC in the near future.

Appendix A

Additional Selection Efficiencies for Two-jet and Three-jet Triggers

The studies in this appendix complement those shown in Section 6.2.2 and justify some of the procedures used in Section 8.2.1. Figure A.1 shows the efficiency of the J15 and J30 triggers to fire on the leading jet of the event as a function of the p_{T} of that jet, calculated using events selected with the MBTS_1 trigger. Jets reconstructed with the anti-k_t algorithm with a resolution parameter $R = 0.4$ were used in this figure, and these results complement those shown in Figure 6.7.

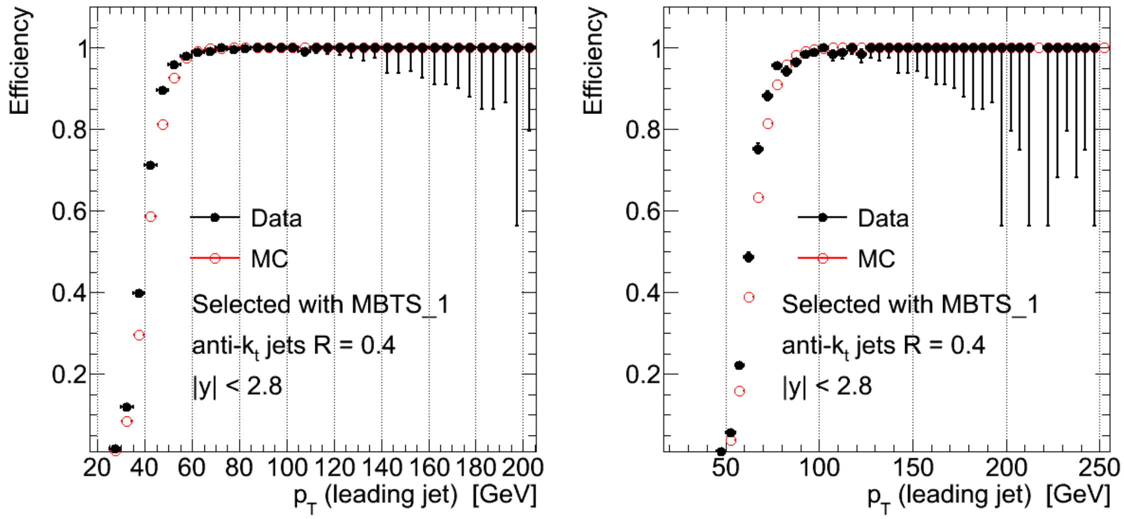

Figure A.1: *Efficiency of the J15 (left) and J30 (right) triggers to select events as a function of the leading jet* p_{T} *within* $|y| < 2.8$ *for jets reconstructed with the anti-*k_t *algorithm with* $R = 0.4$. *The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the unbiased bootstrap method in data, using events selected with the MBTS_1 trigger.*

Figure A.2 shows the efficiency of the J15 and J30 triggers as a function of the rapidity of the leading jet when the jet has $p_{\mathrm{T}} > 80$ GeV and $p_{\mathrm{T}} > 110$ GeV. The results are similar to those shown in Figure 6.6, except that the inefficiency in the data around $|\eta| = 1.5$ seems to get larger with increasing jet trigger threshold. This does not affect the bootstrap method, which requires an offline jet to be matched to a trigger jet in order to consider the event. However, it has implications for the efficiencies of the two-jet triggers that are

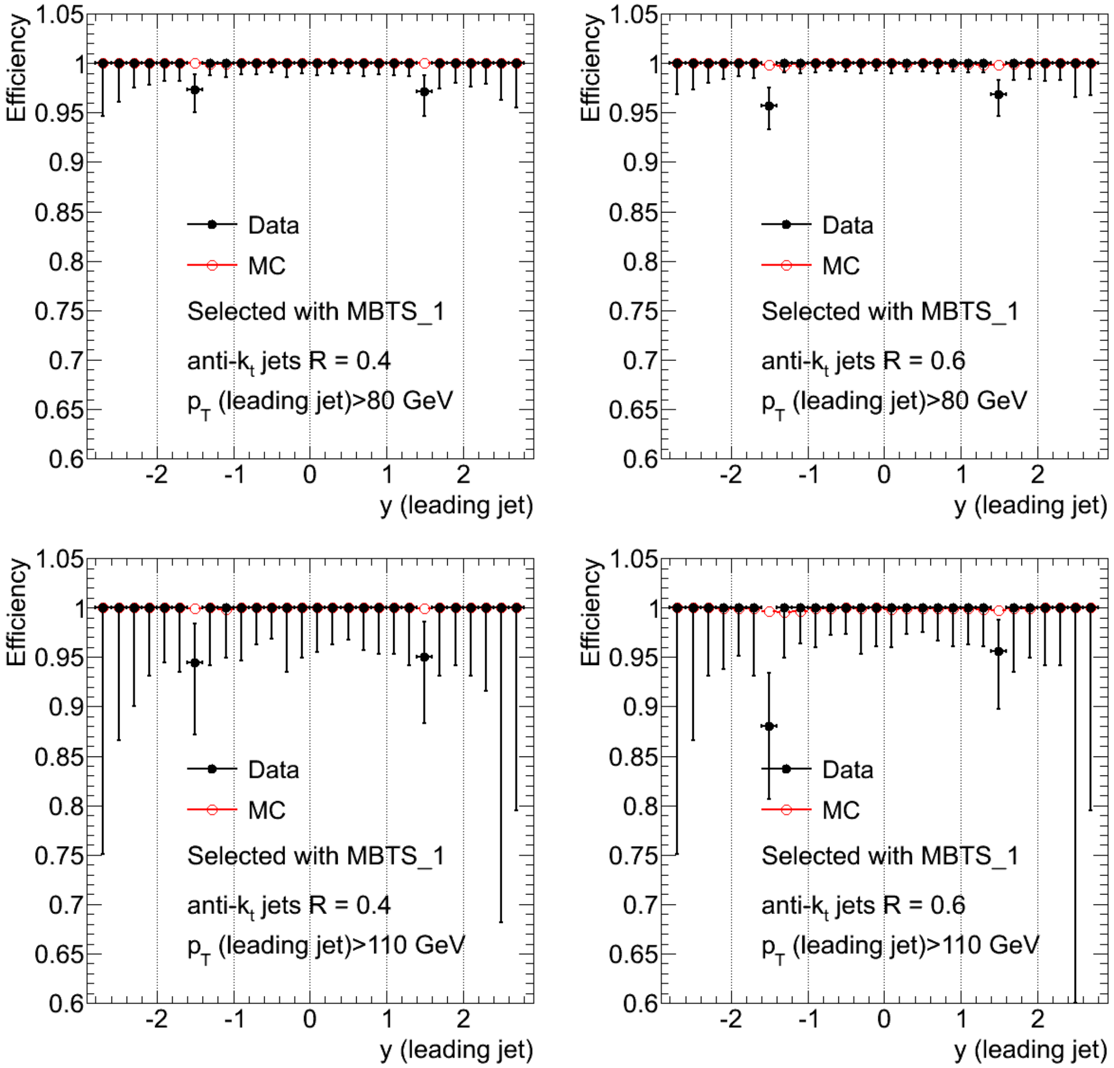

Figure A.2: *Efficiency of the J15 (top) and J30 (bottom) triggers to select events as a function of the leading jet rapidity for events where the leading jet p_{T} is above* 80 GeV *and* 110 GeV *for jets reconstructed with the anti-k_t algorithm with $R = 0.4$ (left) and $R = 0.6$ (right). The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the unbiased bootstrap method in data, using events selected with the MBTS_1 trigger.*

presented later in this appendix.

Figure A.3 shows the same results as those presented in Figure 6.8, but using jets reconstructed with a resolution parameter $R = 0.4$ for completeness.

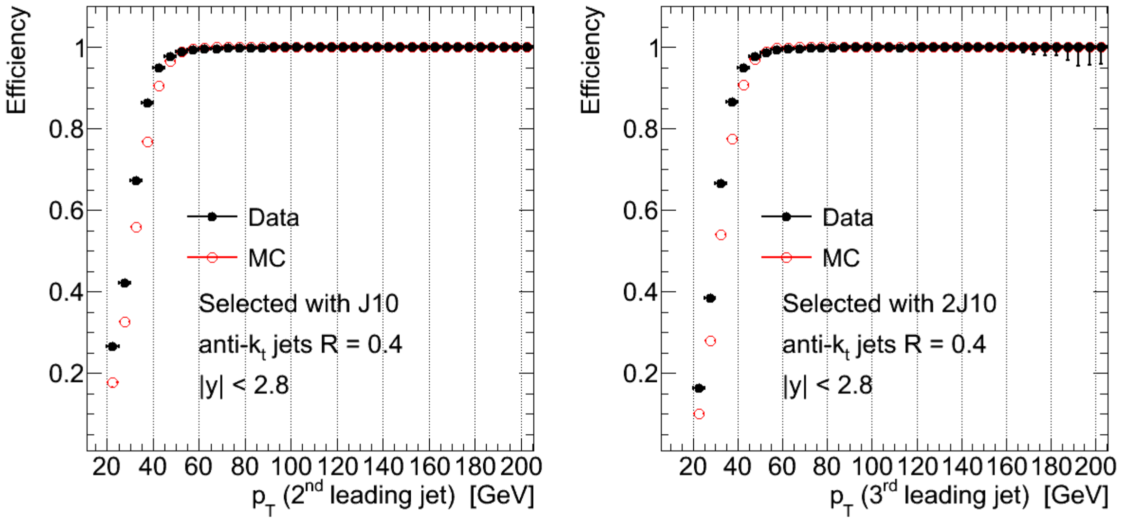

Figure A.3: *Efficiency of the 2J10 (left) and 3J10 (right) triggers to select events as a function of the second and third leading jet* p_{T} *within* $|y| < 2.8$ *for jets reconstructed with the anti-*k_t *algorithm with* $R = 0.4$. *The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the unbiased bootstrap method in data, using events selected with the J10 and 2J10 triggers.*

Figure A.4 shows the efficiency of the 2J15 and 2J30 triggers to fire as a function of the p_{T} of the second leading jet in the event. Jets reconstructed with the anti-k_t algorithm with $R = 0.6$ are shown only, since the results are representative of those obtained using $R = 0.4$ (except that the plateau is reached at lower p_{T} when a small resolution parameter is used). The results show that the 2J15 and 2J30 triggers reach a 99% efficiency for events in which the second leading jet p_{T} is above 80 GeV and $\approx$ 110 GeV. A small difference is observed between data and the Monte Carlo simulation in the plateau region of the 2J30 trigger. This difference is caused by differences in the region of $|\eta| \approx 1.5$.

The efficiency as a function of rapidity is studied in Figure A.5 for all the triggers used to select events in the analysis. The results are shown for jets reconstructed using a resolution parameter $R = 0.6$, but results are similar for jets reconstructed with $R = 0.4$, except that the inefficiencies observed are smaller. The efficiency is high (above 98% for the 2J30 trigger and above 99% for the rest) in the barrel and the end-cap with the kinematic cuts used. Agreement between Monte Carlo simulation and data is found outside of the

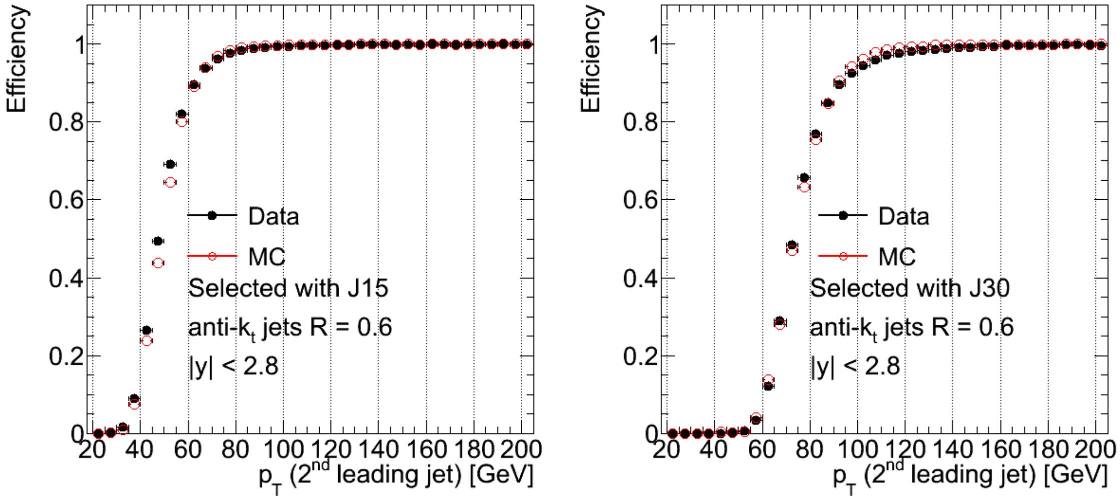

Figure A.4: *Efficiency of the 2J15 (left) and 2J30 (right) triggers to select events as a function of the* p_{T} *of the second leading jet for jets reconstructed with the anti-*k_t *algorithm with* $R = 0.6$*. The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the unbiased bootstrap method in data, using events selected with the J15 and J30 triggers.*

region of transition between the barrel and end-cap calorimeters. In that region, differences of 1–2% are found for $1.0 < |y| < 1.4$, while larger differences are found in the region of $1.4 < |y| < 1.6$. The size of the differences in the region of $1.4 < |y| < 1.6$ change from trigger to trigger, and motivate the estimate of an additional unfolding uncertainty, as described in Section 8.2.1.

The inefficiency in the region of transition between the barrel and the end-cap calorimeters and the differences between data and Monte Carlo simulation mostly occur near the region of p_{T} where the plateau is reached. This is illustrated in Figure A.6 for the 3J10 trigger. The efficiency is shown as a function of the third leading jet p_{T} when that jet has $1.2 < |y| < 1.7$ for jets reconstructed with the anti-k_t algorithm with $R = 0.4$ (left) and $R = 0.6$ (right). Differences between data and Monte Carlo simulation are particularly visible around $p_{\mathrm{T}} = 60$ GeV. The convergence between the data and Monte Carlo simulation curves at higher p_{T}, on the plateau, justifies the use of limited p_{T} ranges in the estimate of the impact of this inefficiency in the analysis, as described in Section 8.2.1. Similar results have been found for the two-jet triggers used in the analysis.

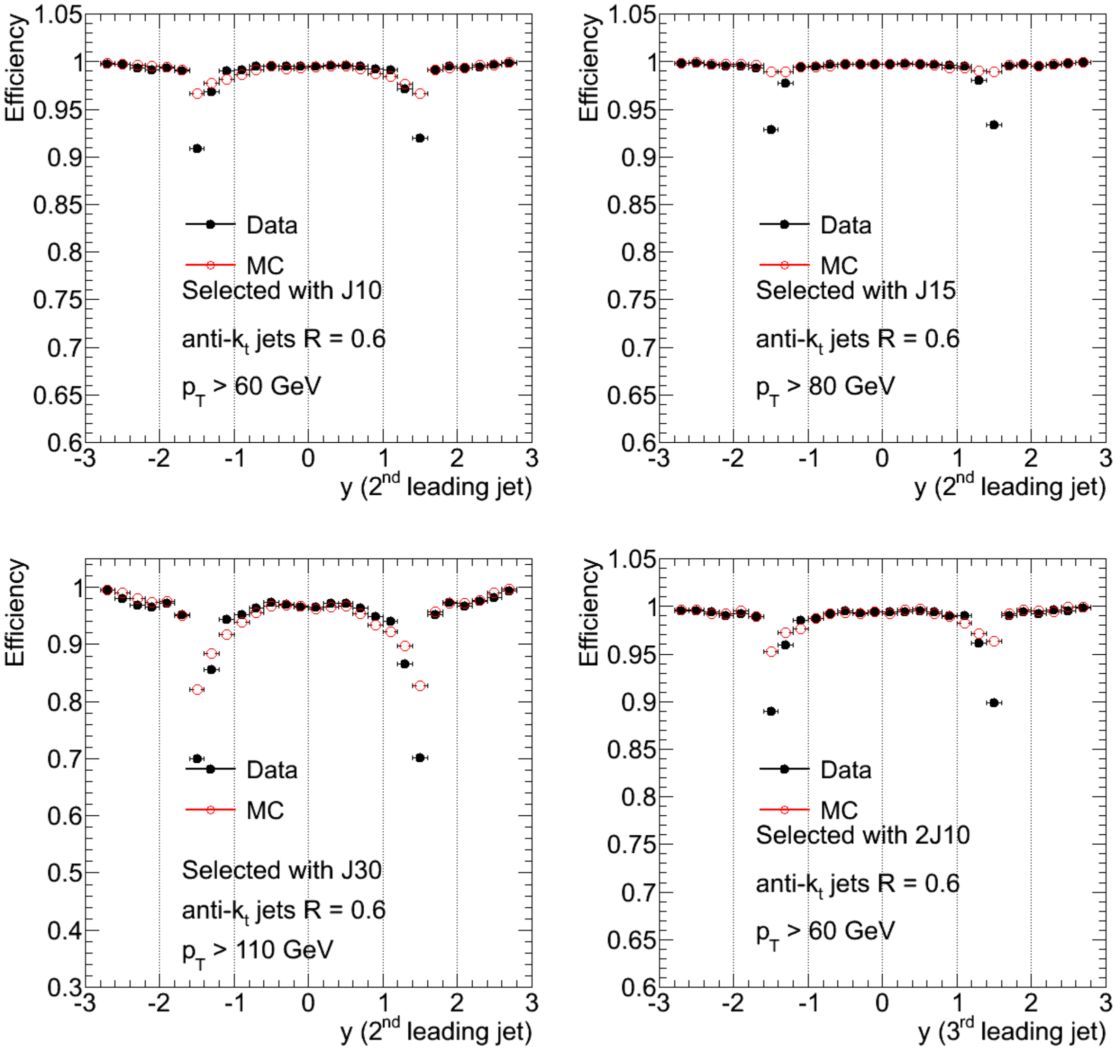

Figure A.5: *Efficiency of the 2J10 (top left), 2J15 (top right), 2J30 (bottom left) and 3J10 (bottom right) triggers to select events as a function of the second or third leading jet rapidity for events in which the other jets in the event passed the trigger for jets reconstructed with the anti-k_t algorithm with $R = 0.6$. The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the unbiased bootstrap method in data, using events selected with the single-jet and di-jet calorimeter triggers.*

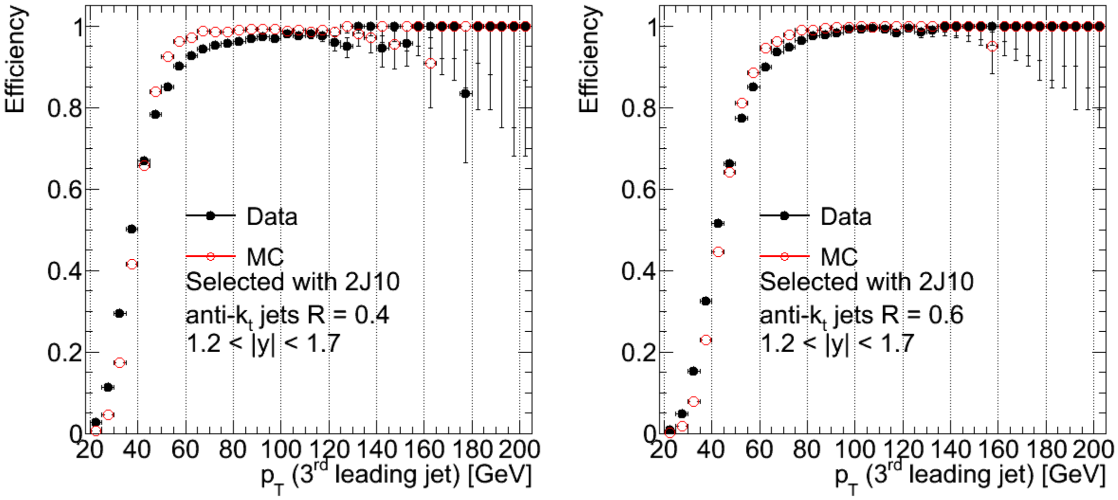

Figure A.6: *Efficiency of the 3J10 trigger to select events as a function of the third leading jet* p_{T} *when this jet has* $1.2 < |y| < 1.7$ *for jets reconstructed with the anti-*k_t *algorithm with* $R = 0.4$ *(left) and* $R = 0.6$ *(right). The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the unbiased bootstrap method in data, using events selected with the 2J10 trigger.*

Appendix B

Additional Studies on the Offline Event Selection

The studies performed in this section complement those shown in Sections 6.3.2 and 6.3.3. Figure B.1 shows similar results as those shown in Figure 6.11 for jets built with the anti-k_t algorithm with resolution parameter $R = 0.6$. The conclusions reached through this study have already been discussed in the corresponding section.

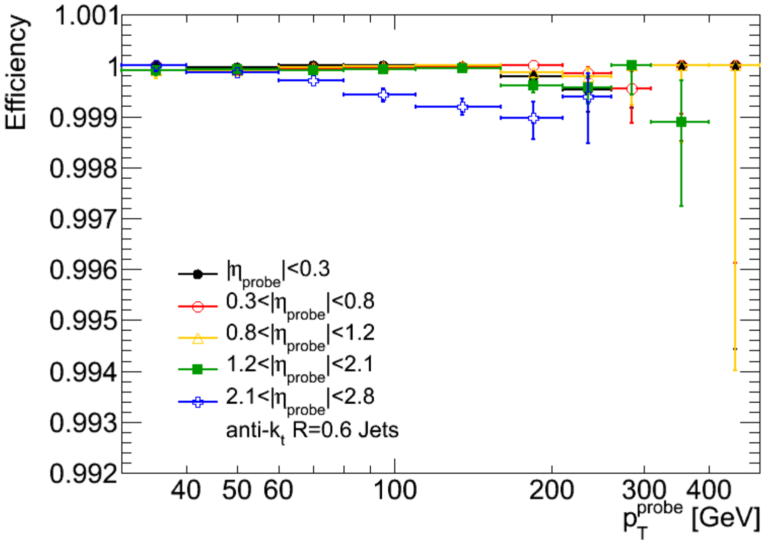

Figure B.1: *Efficiency of the jet data quality cuts to accept jets originating from proton-proton collisions as a function of jet* p_{T}. *Results are shown for jets built with the anti-*k_t *algorithm with* $R = 0.6$ *and for jets falling in different* $|\eta|$ *bins.*

The studies presented in Figures 6.11 and B.1 have been repeated with three additional sets of selection cuts, to study the stability of the results to the specifics of the event selection. Loose and tight cuts on the di-jet balance are used, corresponding to $\Delta\phi$ and imbalance cuts on the di-jet system of 2.4 and 0.5 for the loose cuts and 2.8 and 0.3 for the tight cuts. An additional set of cuts that combines the standard cuts in the study and a cut on the third leading jet p_{T} (required to be less than 15% of the average p_{T} of the di-jet system) is used to investigate the impact on the studies of an explicit veto in the radiation in the event. The efficiencies of the jet quality cuts calculated with these three sets of cuts are shown in Figure B.2 for jets built with the anti-k_t algorithm with both $R = 0.4$ and $R = 0.6$. The results are almost identical to those obtained with the standard set of cuts,

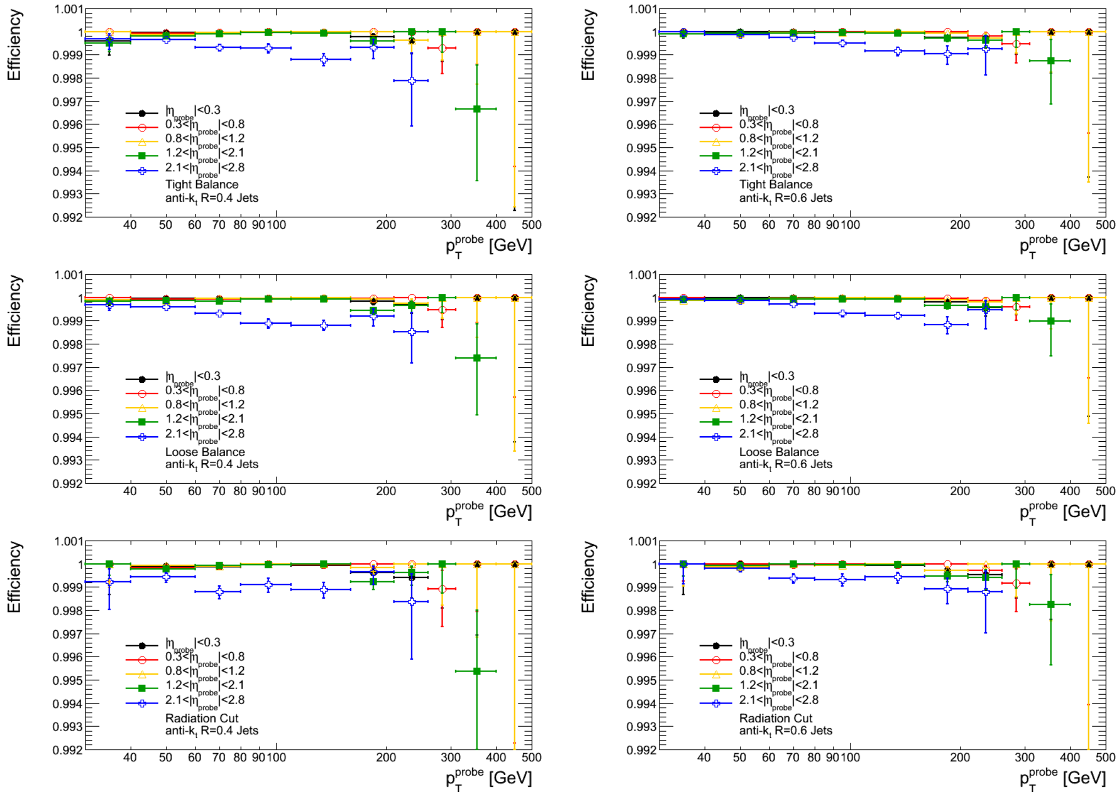

Figure B.2: *Efficiency of the jet data quality cuts to accept jets originating from proton-proton collisions as a function of jet p_{T}. Results are shown for jets built with the anti-k_t algorithm with $R = 0.4$ (left) and $R = 0.6$ (right), and for jets falling in different $|\eta|$ bins. The efficiency is calculated in a data-driven way with three different sets of event selection cuts, as described in the text.*

validating those results.

In Section 6.3.4 it is argued that events with jets that are not contained in the calorimeter, thus recorded in the debug stream, do not contribute significantly to the measurements performed in this thesis. This is illustrated in Figures B.3 and B.4 for a representative set of the kinematic distributions studied. Jets reconstructed with the anti-k_t algorithm with $R = 0.4$ are used in this study. The fractions of events collected in the debug stream are shown at the bottom of each figure. The fraction increases as a function of jet p_{T} and event H_{T}. This is expected, since the probability of a jet not being contained in the calorimeter increases with the energy of the jet. The contribution of these events to the measured cross

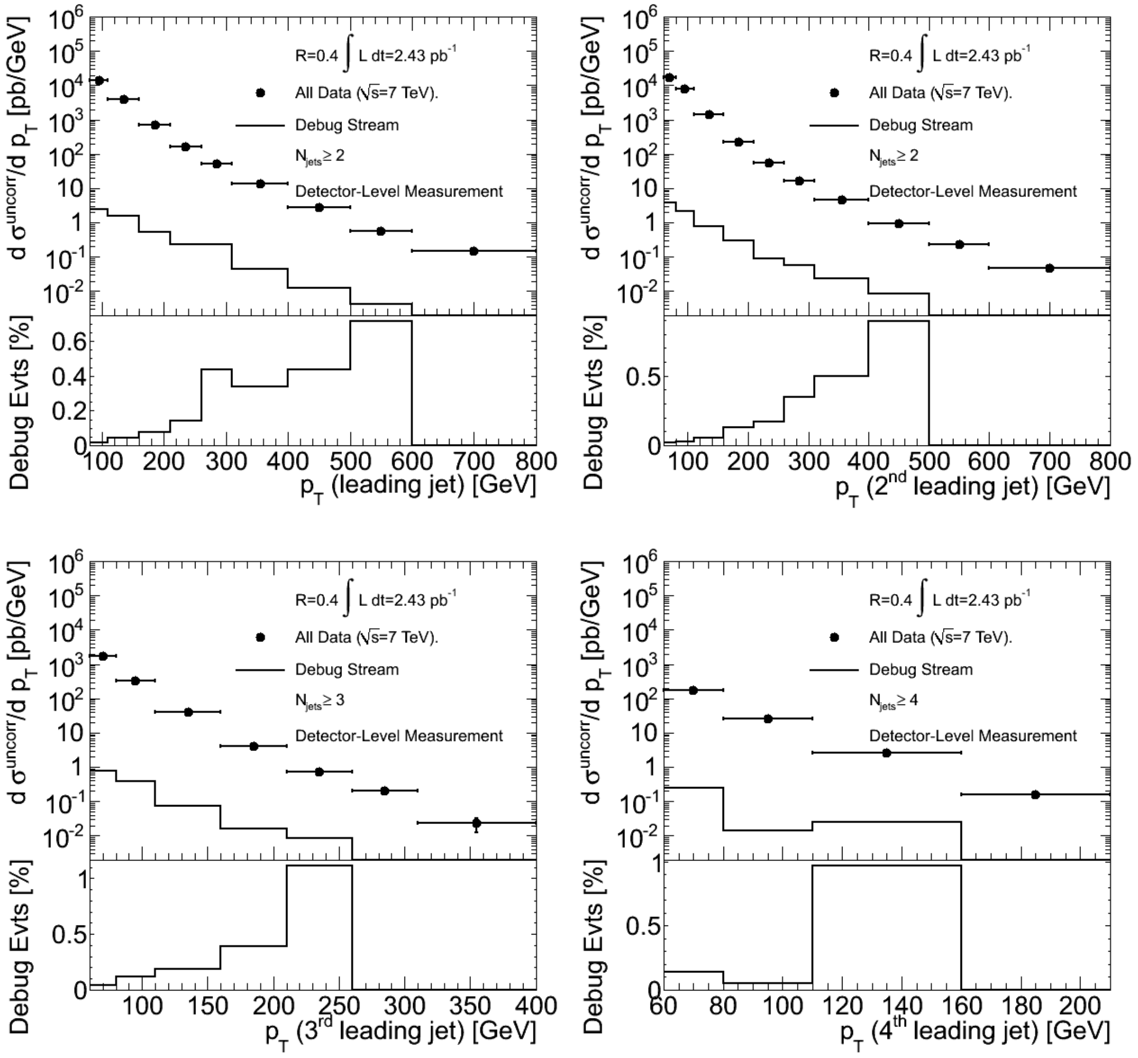

Figure B.3: *Detector-level measurements of the differential cross sections and events collected in the debug stream as a function of leading jet p_{T} in events with $N_{\mathrm{jets}} \geq 2$ (top left), second leading jet p_{T} in events with $N_{\mathrm{jets}} \geq 2$ (top right), third leading jet p_{T} in events with $N_{\mathrm{jets}} \geq 3$ (bottom left) and fourth leading jet p_{T} in events with $N_{\mathrm{jets}} \geq 4$ (bottom right). The fractions of events collected in the debug stream are shown at the bottom of each figure as a percentage.*

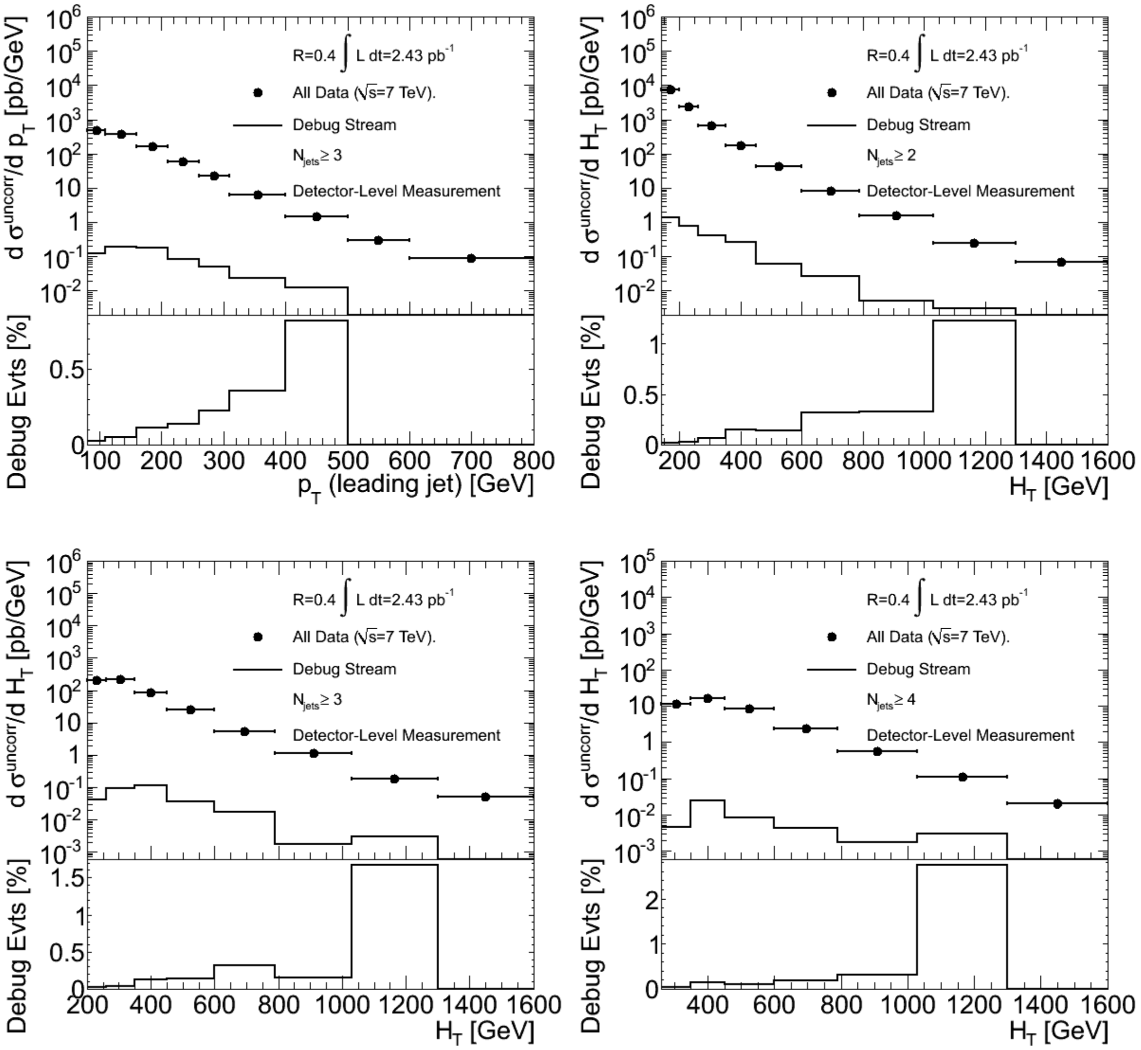

Figure B.4: *Detector-level measurements of the differential cross sections and events collected in the debug stream as a function of leading jet* p_T *in the three-jet bin (top left) and as a function of event* H_T *for events with* $N_{\text{jets}} \geq 2$ *(top right),* $N_{\text{jets}} \geq 3$ *(bottom left) and* $N_{\text{jets}} \geq 4$ *(bottom right). The fractions of events collected in the debug stream are shown at the bottom of each figure as a percentage.*

sections is at the sub-percent level for all bins of all distributions studied.

Finally, Table B.1 presents the value of the cut on the absolute value of the jet vertex fraction that achieves the global minimum in the Kolmogorov distance between the measured observable in events with one reconstructed primary vertex and that measured in events with more than one primary vertex. In the table, Δ_i refers to the Kolmogorov distance, where the subscript indicates the number of primary vertices reconstructed in the events that are used to measure the observable. These results summarize studies performed

	min Δ_2	min Δ_3	min Δ_4	min Δ_5	min $\Delta_{\geq 6}$
Inclusive multiplicity	0.50	0.75	0.65	0.90	0.85
${p_{\mathrm{T}}}^{\mathrm{lead}}$, $N_{\mathrm{jets}} \geq 2$	0.85	0.75	1.00	0.95	0.95
${p_{\mathrm{T}}}^{2^{\mathrm{nd}}}$, $N_{\mathrm{jets}} \geq 2$	0.50	0.50	0.85	0.65	0.85
${p_{\mathrm{T}}}^{3^{\mathrm{rd}}}$, $N_{\mathrm{jets}} \geq 3$	0.65	1.00	0.85	0.80	1.00
${p_{\mathrm{T}}}^{4^{\mathrm{th}}}$, $N_{\mathrm{jets}} \geq 4$	0.90	0.95	1.00	0.95	0.95
3-to-2-jet Ratio vs ${H_{\mathrm{T}}}^{(2)}$, $R = 0.4$	1.00	0.50	0.85	0.70	0.50
3-to-2-jet Ratio vs ${H_{\mathrm{T}}}^{(2)}$, $R = 0.6$	1.00	0.50	0.95	0.65	0.85
H_{T}, $N_{\mathrm{jets}} \geq 2$	0.95	0.50	0.90	0.55	0.85
H_{T}, $N_{\mathrm{jets}} \geq 3$	0.80	1.00	0.50	0.50	0.85
H_{T}, $N_{\mathrm{jets}} \geq 4$	0.95	0.65	0.80	0.95	0.95
$\lvert\Delta\eta_{12}\rvert$	1.00	0.65	0.70	0.90	0.90
$\lvert\Delta\eta_{13}\rvert$	0.90	0.70	0.85	0.80	0.85
$\lvert\Delta\eta_{23}\rvert$	0.95	0.95	0.85	0.95	0.70
$\lvert\Delta\phi_{12}\rvert$	0.85	0.50	0.70	0.85	0.50
$\lvert\Delta\phi_{13}\rvert$	0.90	0.90	0.50	0.95	0.60
$\lvert\Delta\phi_{23}\rvert$	0.50	0.80	0.85	0.95	0.90

Table B.1: *Cut value on the absolute value of the jet vertex fraction that minimizes the Kolmogorov distance between the observable measured in events with one reconstructed primary vertex and those measured in events with multiple primary vertices for a representative set of the observables measured in this thesis.*

as presented in Figure 6.12 for all observables studied. The minimum value is different for the different variables and vertex multiplicities. Such differences are not surprising, given the interplay between the pile-up rejection cut and the different kinematic variables and distributions under study. However, it should be noted that the values obtained are also sensitive to statistical fluctuations in the data. These results are used in Section 8.2.2 to select a wide range of cuts on the jet vertex fraction for the estimation of an associated systematic uncertainty. The range used (0.5 to 0.9) does not fully cover the set of optimal cuts presented in this table. Values above 0.9 significantly increase the Kolmogorov distance for many of the variables studied, while the variables and vertex multiplicities for which an optimal cut value is found above 0.9 display an almost optimal Kolmogorov distance when a cut of 0.9 is used. This justifies the choice of range for the systematic uncertainty calculation described in detail in Section 8.2.2.

Appendix C

Additional Studies on Unfolding and Systematic Uncertainties in the Unfolding Factors

This appendix complements the studies shown in Chapter 7 and Section 8.2. Figures C.1 through C.5 show the purity and efficiency of the binning used for a representative set of the observables studied in this thesis. The display of the results has the same format as that of Figure 7.4, described in the corresponding chapter. The purities and efficiencies are around 70% for all observables, guaranteeing that the impact of bin migrations is small and the limitations of the bin-by-bin unfolding do not impact the uncertainties in the unfolding factors significantly. One exception to this statement can be observed in Figure C.4. Despite the low statistics in the Monte Carlo sample, a trend is observed towards lower purities and efficiencies, particularly for low $|\Delta\phi_{12}|$, but also for low $|\Delta\phi_{13}|$ and high $|\Delta\phi_{23}|$. The reasons and impact of such trends have already been discussed in Chapter 7.

The re-binning procedure used in the bin-by-bin unfolding of angular variables relies on the unfolding factors varying slowly. Naturally slow-varying factors guaranteed by the geometry of the detector can be altered by calorimeter failures, which can add small-granularity features at preferred values of $|\Delta\phi|$ or $|\Delta\eta|$. To investigate this, the measured distributions are recalculated excluding jets falling within a distance in η and ϕ of 0.2 from problematic regions of the EM calorimeter (labelled as dead OTX in the figures that follow) and jets falling on (within a distance in η and ϕ of 0.05) or near (within a distance in η and ϕ of 0.15) the problematic regions of the tile calorimeter. These regions have been shown in Figure 4.19. The excluded regions have a size smaller than 0.4 because most of the energy

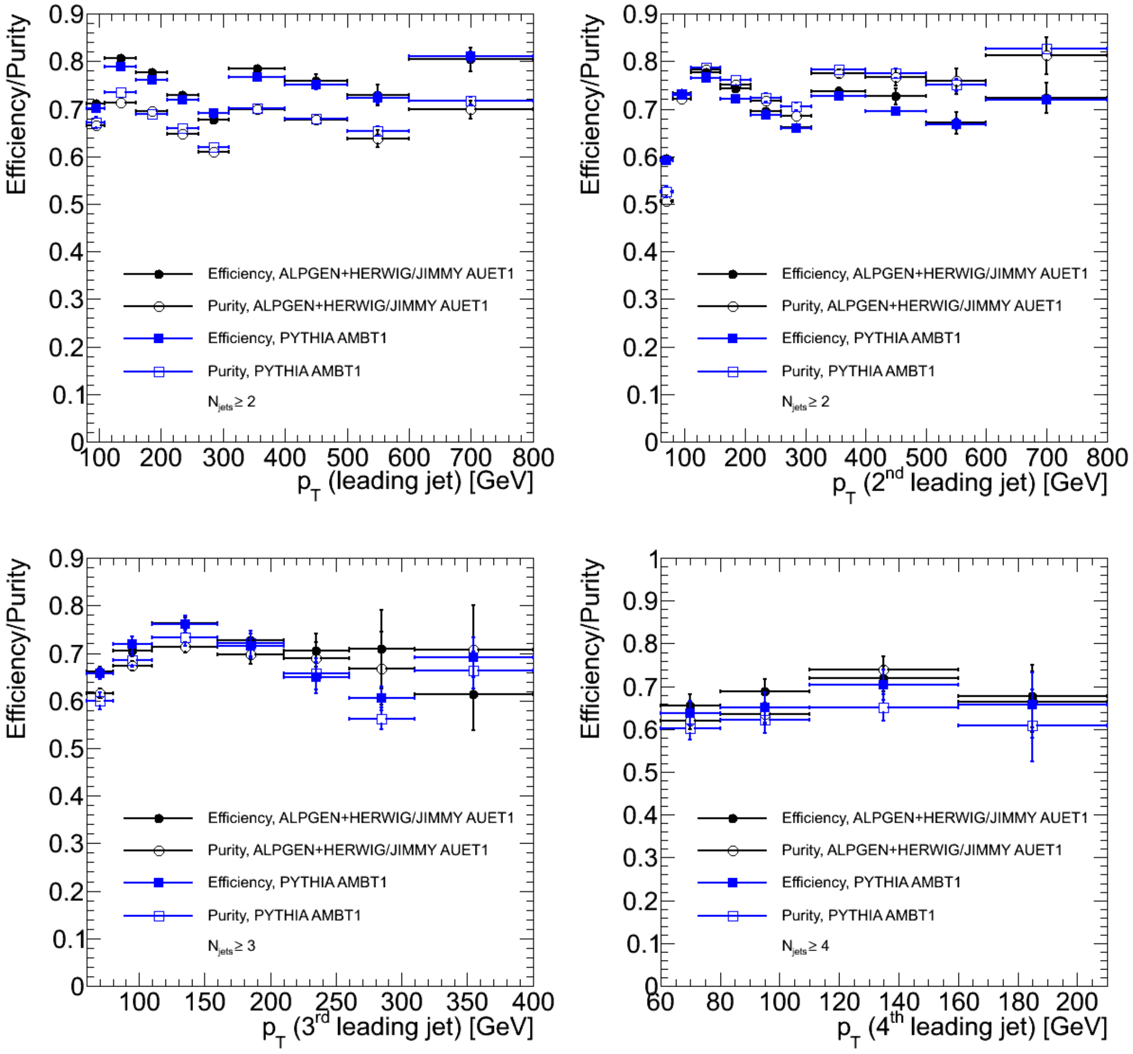

Figure C.1: *Efficiency (solid markers) and purity (open markers) as a function of leading jet* p_{T} *for events with* $N_{\mathrm{jets}} \geq 2$ *(top left), second leading jet* p_{T} *for events with* $N_{\mathrm{jets}} \geq 2$ *(top right), third leading jet* p_{T} *for events with* $N_{\mathrm{jets}} \geq 3$ *(bottom left) and fourth leading jet* p_{T} *for events with* $N_{\mathrm{jets}} \geq 4$ *(bottom right) as calculated in the* ALPGEN *(black markers) and* PYTHIA *(blue markers) samples. Jets reconstructed with the anti-*k_t *algorithm with* $R = 0.4$ *are used in this study.*

of the jet is concentrated in the central area of the jet, after the kinematic cuts in the analysis are applied [147].

Different studies are performed for each type of failure because the failures are treated differently in the software as discussed in Section 5.4.2. The studies overestimate the impact of these failures on the unfolding factors, since they assume that jets falling in these regions are lost, while small effects on the jet response and resolution are expected instead. Figures C.6 and C.7 show the results of these studies. The $|\Delta\phi|$ distributions are unaf-

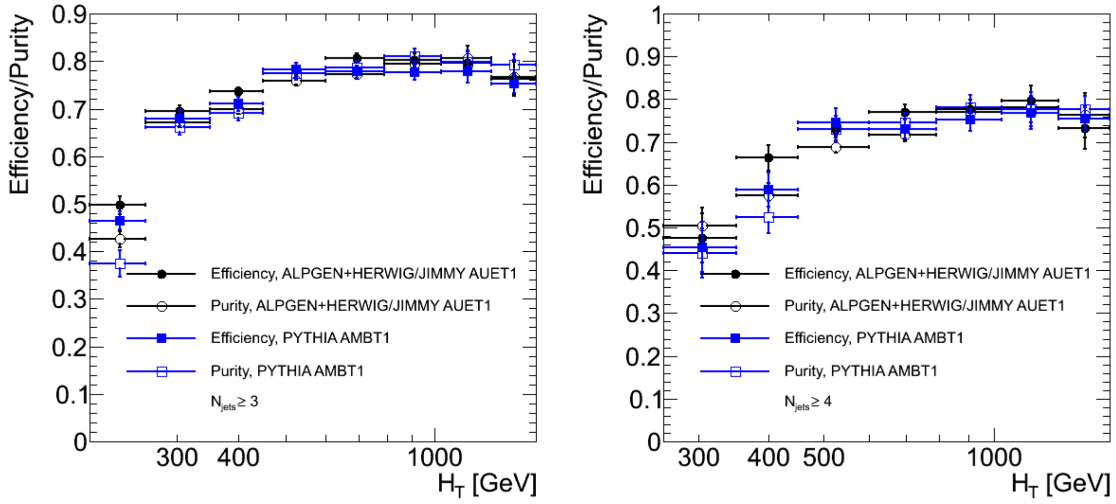

Figure C.2: *Efficiency (solid markers) and purity (open markers) as a function of event H_{T} for events with $N_{\mathrm{jets}} \geq 3$ (left), and events with $N_{\mathrm{jets}} \geq 4$ (right) as calculated in the* Alpgen *(black markers) and* Pythia *(blue markers) samples. Jets reconstructed with the anti-k_t algorithm with $R = 0.4$ are used in this study.*

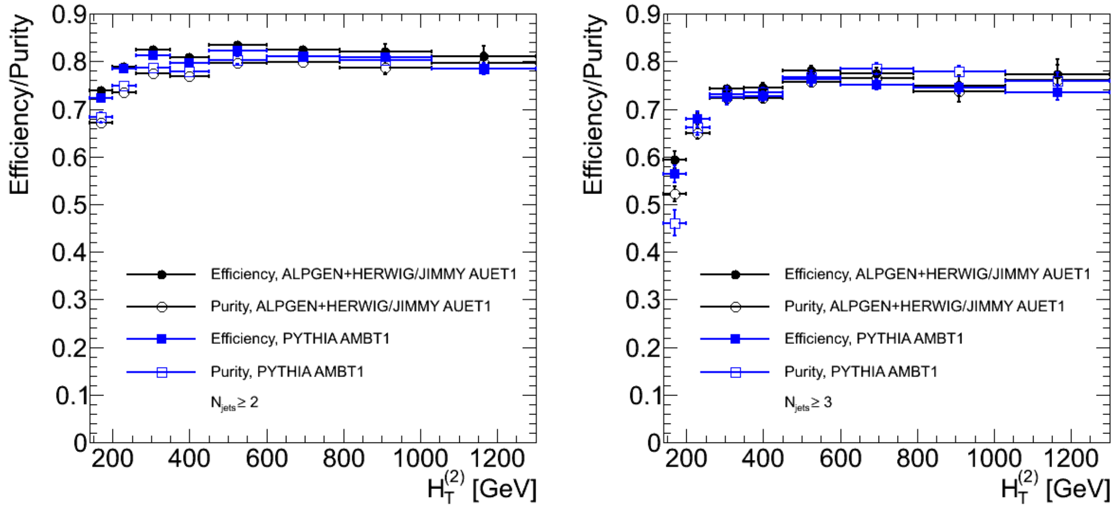

Figure C.3: *Efficiency (solid markers) and purity (open markers) as a function of event ${H_{\mathrm{T}}}^{(2)}$ for events with $N_{\mathrm{jets}} \geq 2$ (left), and events with $N_{\mathrm{jets}} \geq 3$ (right) as calculated in the* Alpgen *(black markers) and* Pythia *(blue markers) samples. Jets reconstructed with the anti-k_t algorithm with $R = 0.4$ are used in this study.*

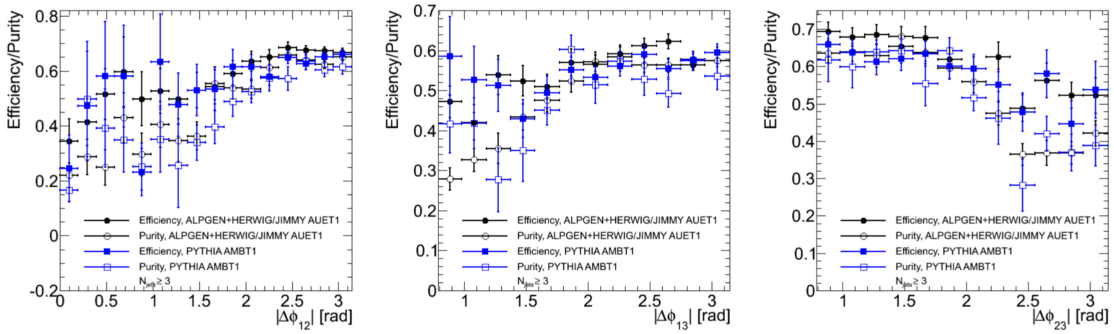

Figure C.4: *Efficiency (solid markers) and purity (open markers) as a function of $|\Delta\phi|$ between different jets for events with $N_{\text{jets}} \geq 3$ as calculated in the* ALPGEN *(black markers) and* PYTHIA *(blue markers) samples. Jets reconstructed with the anti-k_t algorithm with $R = 0.4$ are used in this study.*

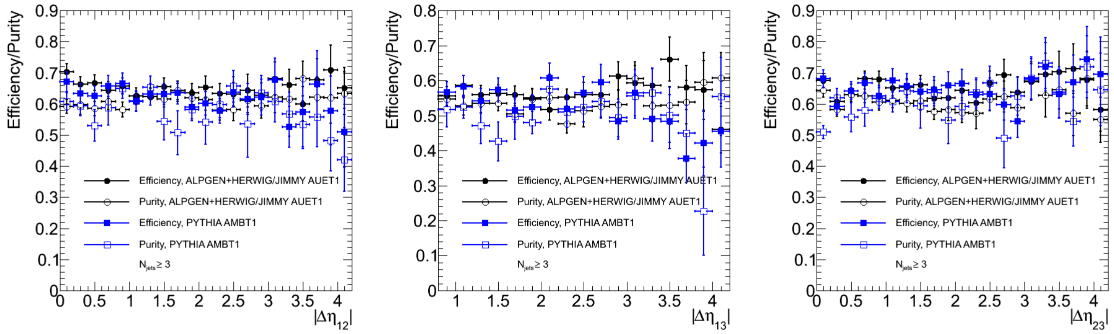

Figure C.5: *Efficiency (solid markers) and purity (open markers) as a function of $|\Delta\eta|$ between different jets for events with $N_{\text{jets}} \geq 3$ as calculated in the* ALPGEN *(black markers) and* PYTHIA *(blue markers) samples. Jets reconstructed with the anti-k_t algorithm with $R = 0.4$ are used in this study.*

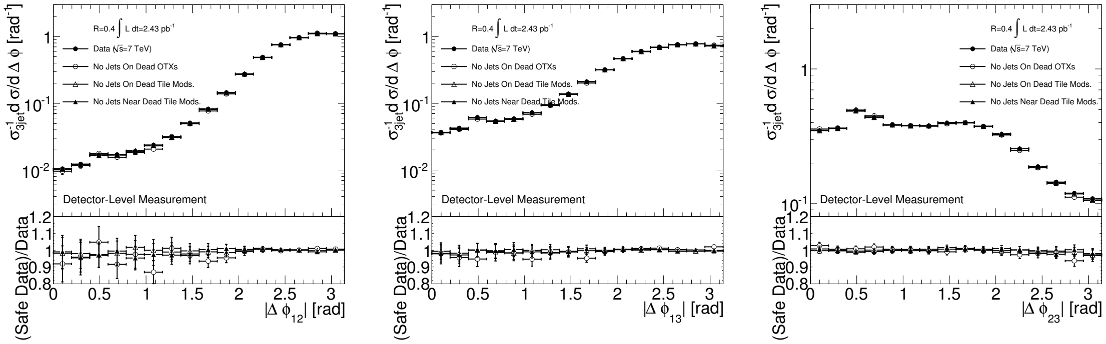

Figure C.6: *Detector-level measured $|\Delta\phi|$ distributions for events with $N_{\text{jets}} \geq 3$ after cuts that remove jets in problematic regions of the calorimeters. The ratios of the results after the cuts over those before the cuts are shown at the bottom of each figure.*

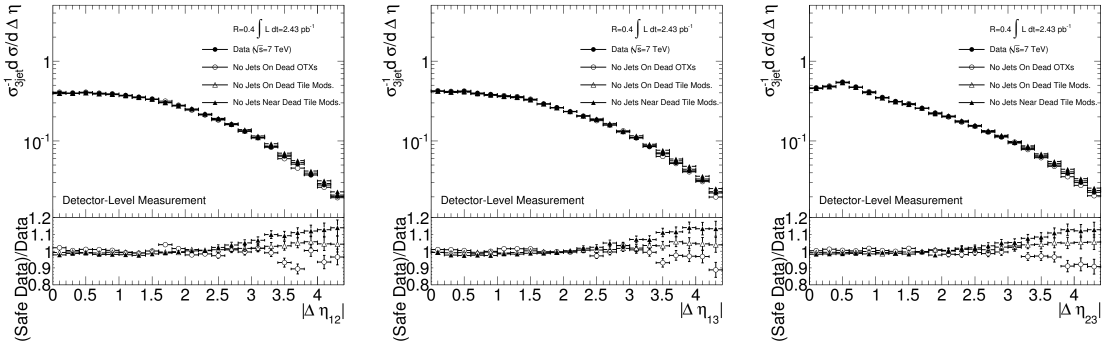

Figure C.7: *Detector-level measured $|\Delta\eta|$ distributions for events with $N_{\text{jets}} \geq 3$ after cuts that remove jets in problematic regions of the calorimeters. The ratios of the results after the cuts over those before the cuts are shown at the bottom of each figure.*

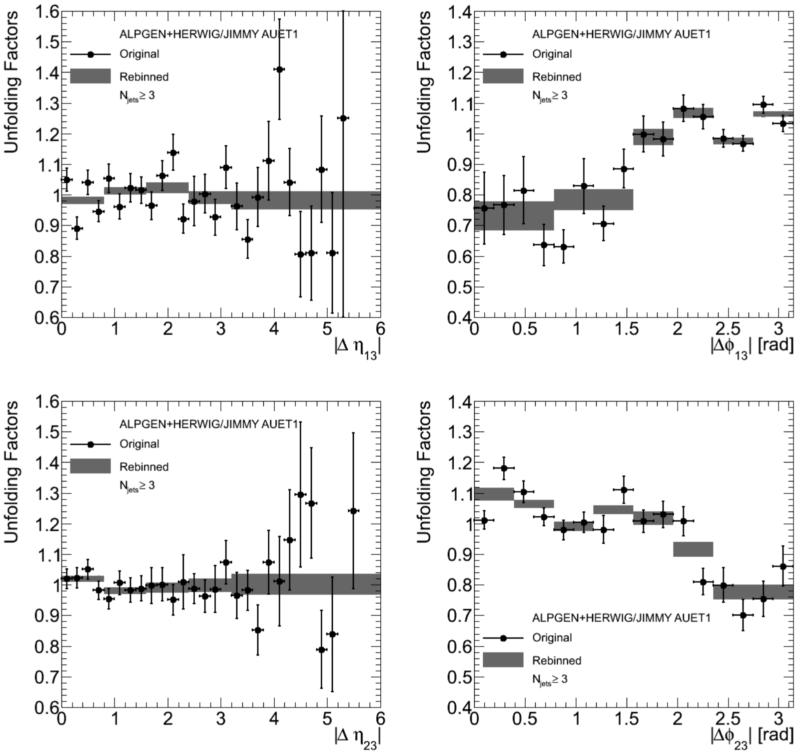

Figure C.8: *Unfolding factors for $\sigma_{\text{3jet}}^{-1} d\sigma/d|\Delta\eta_{13}|$ (top left), $\sigma_{\text{3jet}}^{-1} d\sigma/d|\Delta\phi_{13}|$ (top right), $\sigma_{\text{3jet}}^{-1} d\sigma/d|\Delta\eta_{23}|$ (bottom left), and $\sigma_{\text{3jet}}^{-1} d\sigma/d|\Delta\phi_{23}|$ (bottom right) for events with $N_{\text{jets}} \geq 3$. The gray band is centered around the final unfolding factors and represents the statistical uncertainty in them. The points show the unfolding factors if the same binning as for the data was used.*

fected by the cuts. More significant effects are observed in the $|\Delta\eta|$ distributions at large $|\Delta\eta|$. The changes in the distributions after each set of cuts, shown at the bottom of each figure, are within the systematic uncertainties in the measurement. In addition, the effects of the EM and the tile calorimeter failures go in opposite directions, pointing to a possible cancellation of effects. This and the fact that these studies provide an upper bound on the effect validate the re-binning procedure.

Figure C.8 shows the unfolding factors before and after the change of binning discussed in Chapter 7 for the measurements of angular variables not shown in figure 7.5. The unfolding factors after the re-binning procedure are compatible with those prior to the procedure.

Finally, Figures C.9 through C.12 show the unfolding factors and the associated systematic uncertainties for variables not shown in Section 8.2 but studied in Chapter 10 for completeness. The general trend of these figures and the different studies they summarize have already been discussed in that section.

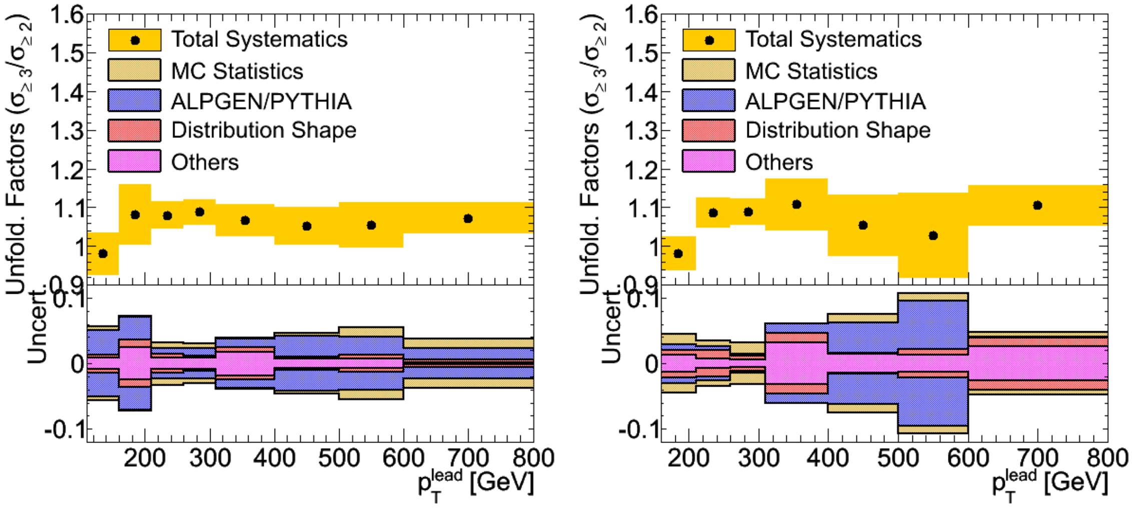

Figure C.9: *Unfolding factors of three-to-two-jet cross-section ratio measurements as a function of leading jet* p_{T} *for events selected using* $p_{\mathrm{T}} \geq 80$ *(left) and* $p_{\mathrm{T}} \geq 110$ *(right) in the jet selection and their associated uncertainties. The main effects contributing to the systematic uncertainties are shown at the bottom of each figure. Jets reconstructed with the anti-*k_t *algorithm with resolution parameter* $R = 0.6$ *were used for these results.*

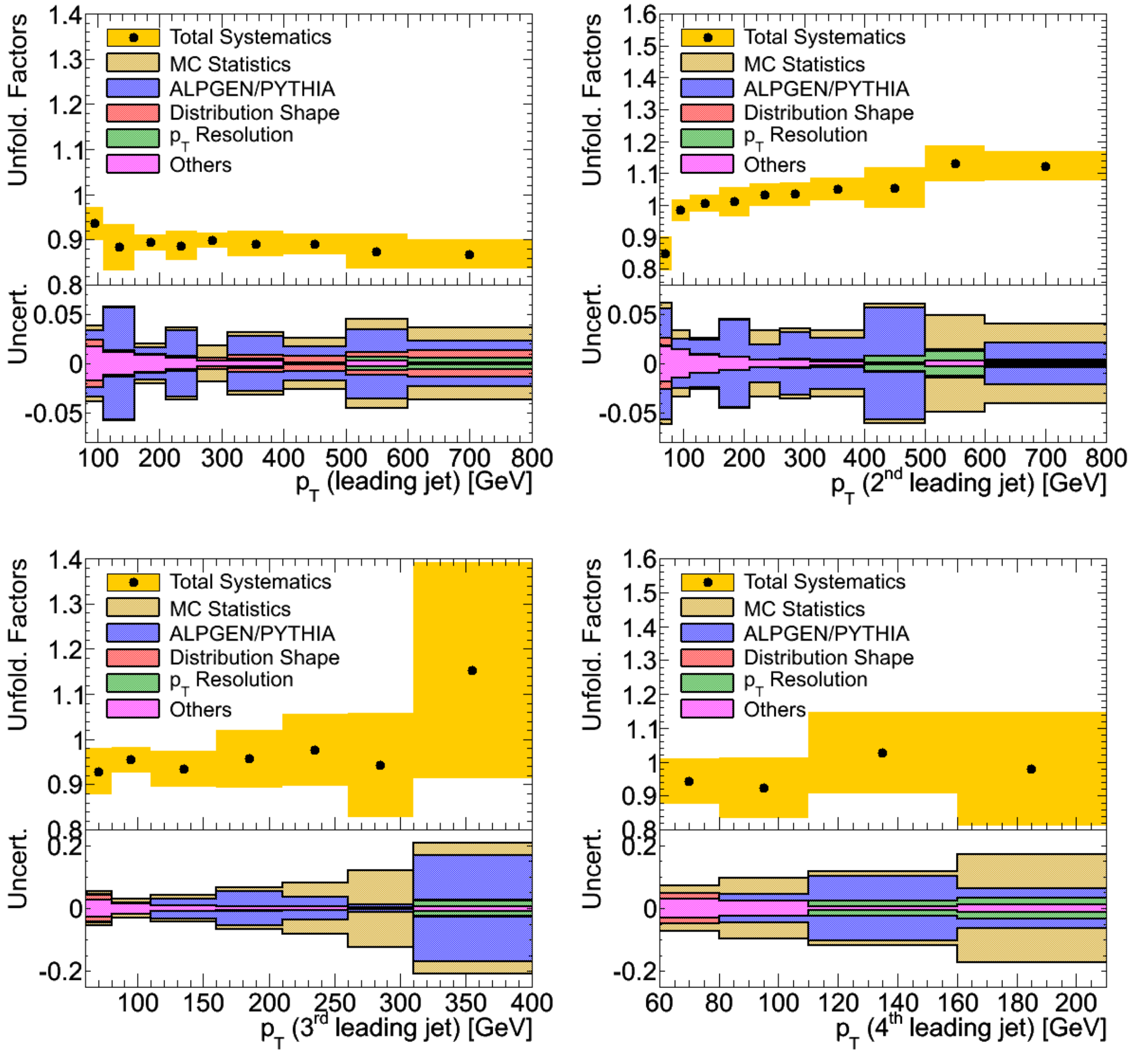

Figure C.10: *Unfolding factors for the measurements of the differential cross sections as a function of the leading (top left), second leading (top right), third leading (bottom left) and fourth leading (bottom right) jet* p_T *in the event and their associated uncertainties. The main effects contributing to the systematic uncertainties are shown at the bottom of each figure. Jets reconstructed with the anti-*k_t *algorithm with resolution parameter* $R = 0.4$ *were used for these results.*

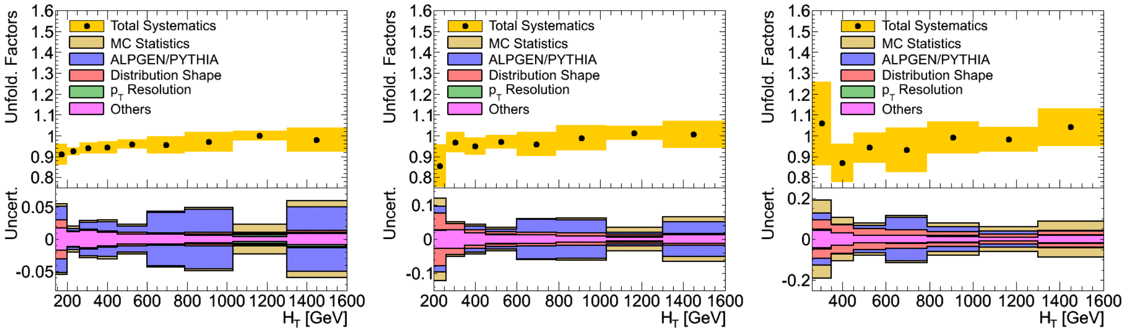

Figure C.11: *Unfolding factors for the measurements of the differential cross sections as a function of the H_T for events with $N_{\text{jets}} \geq 2$ (left), $N_{\text{jets}} \geq 3$ (center) and $N_{\text{jets}} \geq 4$ (right) and their associated uncertainties. The main effects contributing to the systematic uncertainties are shown at the bottom of each figure. Jets reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.4$ were used for these results.*

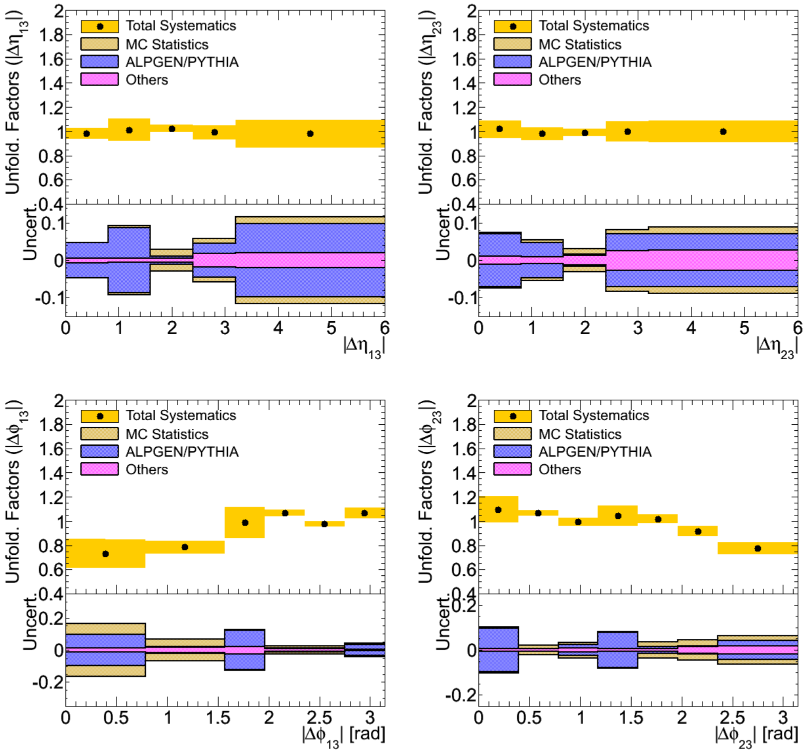

Figure C.12: *Unfolding factors for the measurements of the angular distributions $|\Delta\eta_{13}|$ (top left), $|\Delta\eta_{23}|$ (top right), $|\Delta\phi_{13}|$ (bottom left) and $|\Delta\phi_{23}|$ (bottom right) in events with $N_{\text{jets}} \geq 3$ and their associated uncertainties. The main effects contributing to the systematic uncertainties are shown at the bottom of each figure. Jets reconstructed with the anti-k_t algorithm with resolution parameter $R = 0.4$ were used for these results.*

Appendix D

Additional Studies on the Systematic Uncertainties in the Jet Energy Scale

This appendix complements the studies shown in Section 8.3. Figure D.1 shows the fit to the light-quark and gluon fraction in events with three or more jets for non-isolated jets

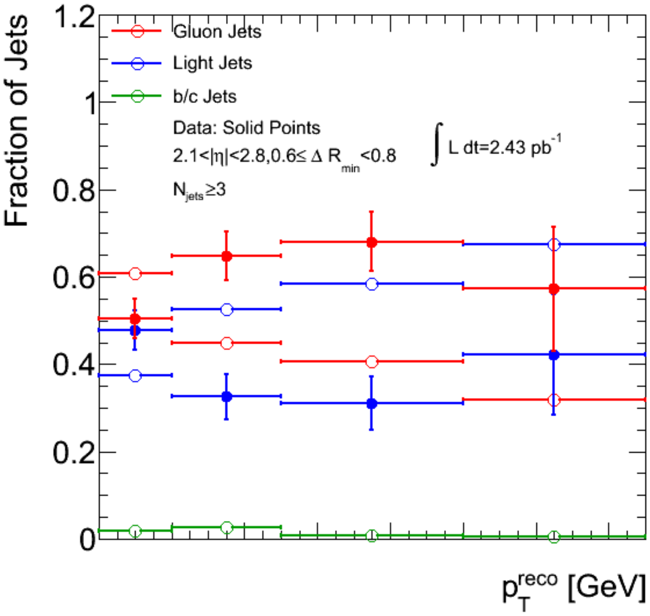

Figure D.1: *Fitted values of the light-quark and gluon jet fraction in events with three or more jets, as a function of jet* p_{T}*, calculated using jet width templates [27]. Non-isolated jets (*$0.6 < \Delta R_{\mathrm{min}} < 0.8$*) with* $2.1 < |\eta| < 2.8$ *are shown. The fraction of heavy-quark jets has been fixed to that of the* Alpgen *simulation. The fitted fractions are shown with solid markers, while the values obtained from the Monte Carlo simulation are shown with empty markers.*

falling in the end-cap. This $|\eta|$ and isolation bin correspond to one of the bins showing the largest disagreement between the flavor composition predicted in the ALPGEN sample and that fitted in the data. The differences remain below 30%, justifying the choice of 30% as the uncertainty in the flavor composition of the multi-jet sample, as discussed in Section 8.3.

Figures D.2 through D.6 show the final upwards and downwards uncertainties in the jet energy scale as in Figure 8.25 for different values of $|\eta|$. Results are shown for jets reconstructed with both $R = 0.4$ and $R = 0.6$ for completeness, and the features and conclusions are similar to those discussed in Section 8.3.4.

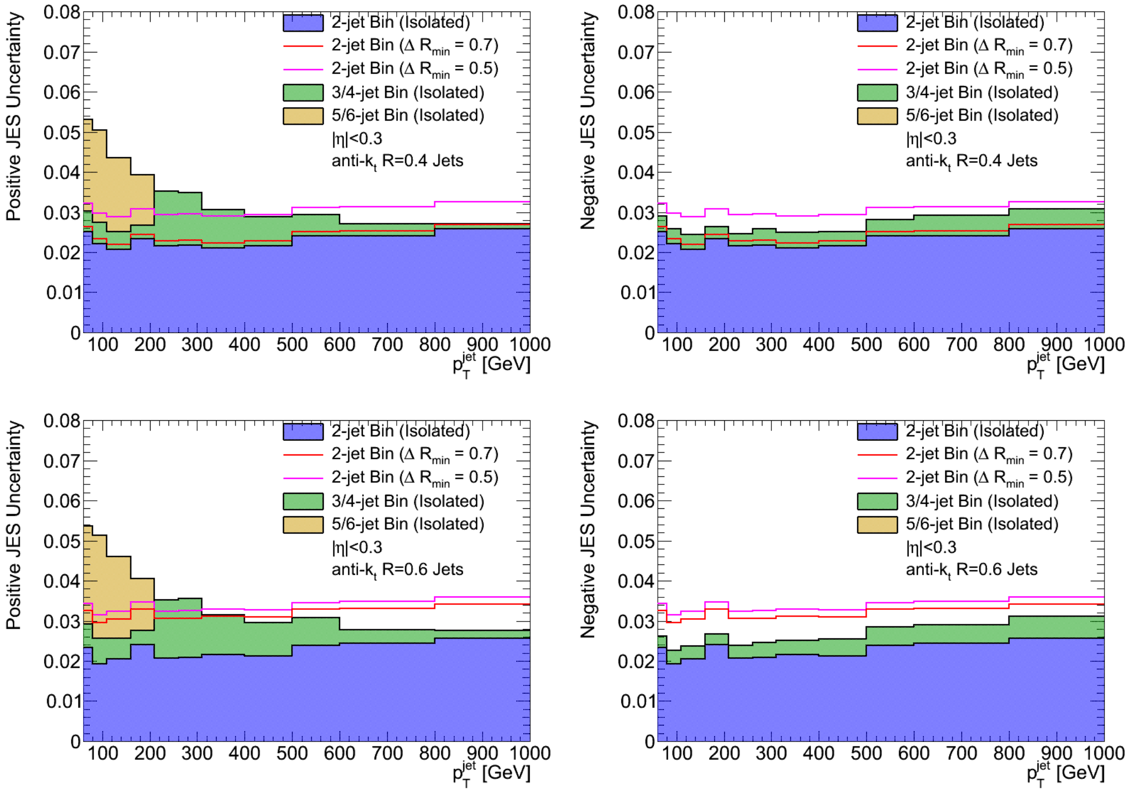

Figure D.2: *Total upwards (right) and downwards (left) jet energy scale uncertainties as a function of jet* p_{T} *for isolated jets in different multiplicity bins, and non-isolated jets in events with two selected jets. Jets reconstructed with the anti-*k_t *algorithm with resolution parameters* $R = 0.4$ *(top) and* $R = 0.6$ *(bottom) and with* $|\eta| < 0.3$ *were used in this figure.*

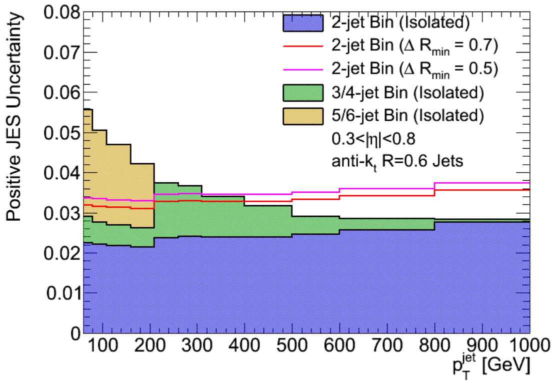

Figure D.3: *Total upwards (left) and downwards (right) jet energy scale uncertainties as a function of jet* p_T *for isolated jets in different multiplicity bins, and non-isolated jets in events with two selected jets. Jets reconstructed with the anti-*k_t *algorithm with resolution parameter* $R = 0.6$ *and with* $0.3 < |\eta| < 0.8$ *were used in this figure.*

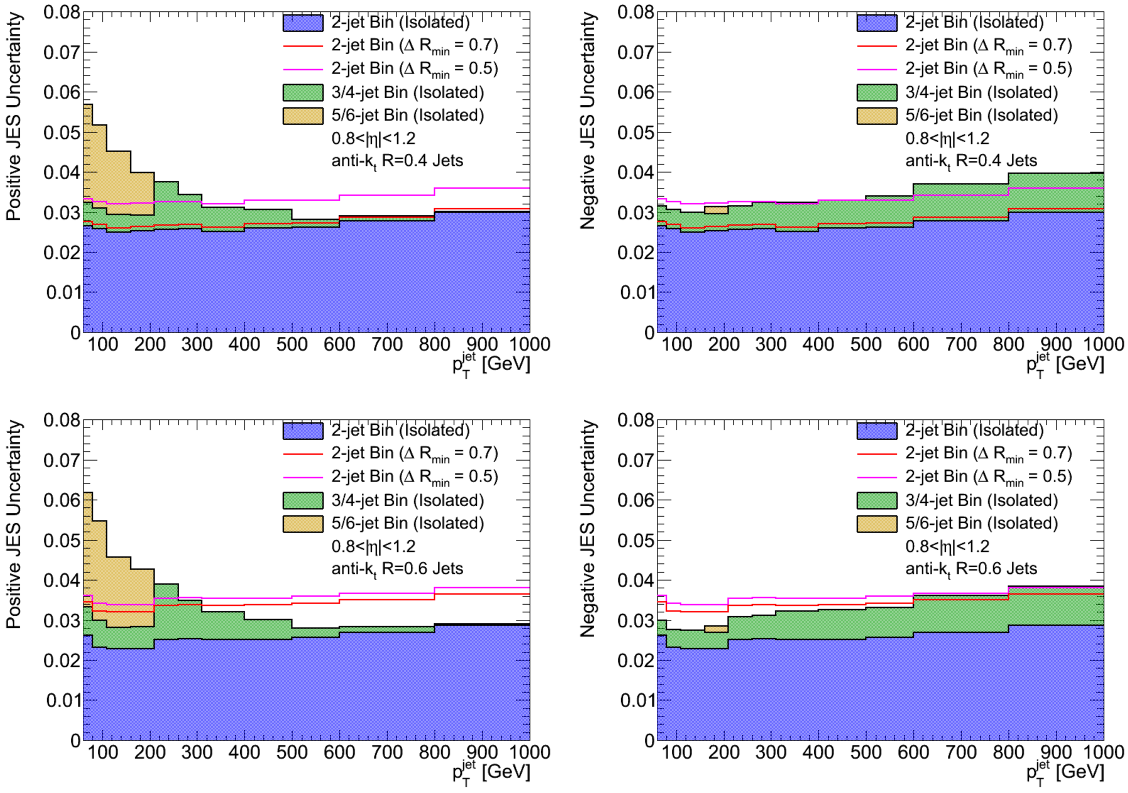

Figure D.4: *Total upwards (top) and downwards (bottom) jet energy scale uncertainties as a function of jet* p_T *for isolated jets in different multiplicity bins, and non-isolated jets in events with two selected jets. Jets reconstructed with the anti-*k_t *algorithm with resolution parameters* $R = 0.4$ *(top) and* $R = 0.6$ *(bottom) and with* $0.8 < |\eta| < 1.2$ *were used in this figure.*

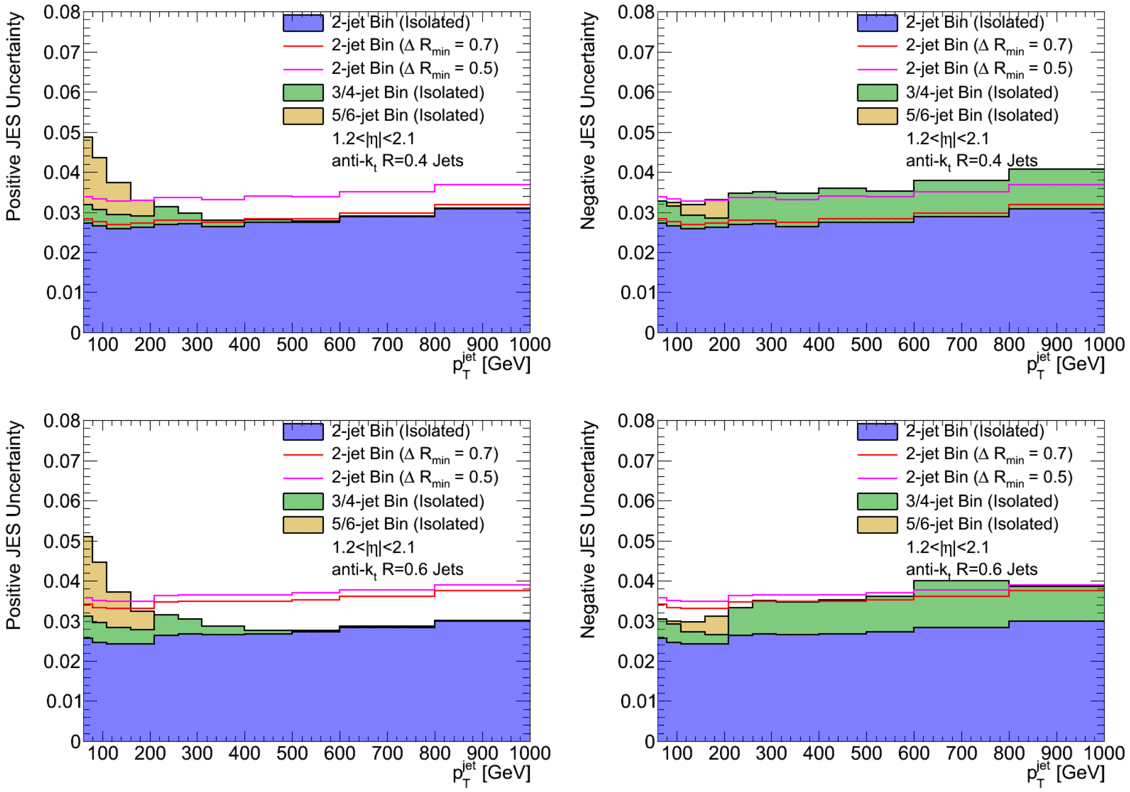

Figure D.5: *Total upwards (top) and downwards (bottom) jet energy scale uncertainties as a function of jet* p_{T} *for isolated jets in different multiplicity bins, and non-isolated jets in events with two selected jets. Jets reconstructed with the anti-*k_t *algorithm with resolution parameters* $R = 0.4$ *(top) and* $R = 0.6$ *(bottom) and with* $1.2 < |\eta| < 2.1$ *were used in this figure.*

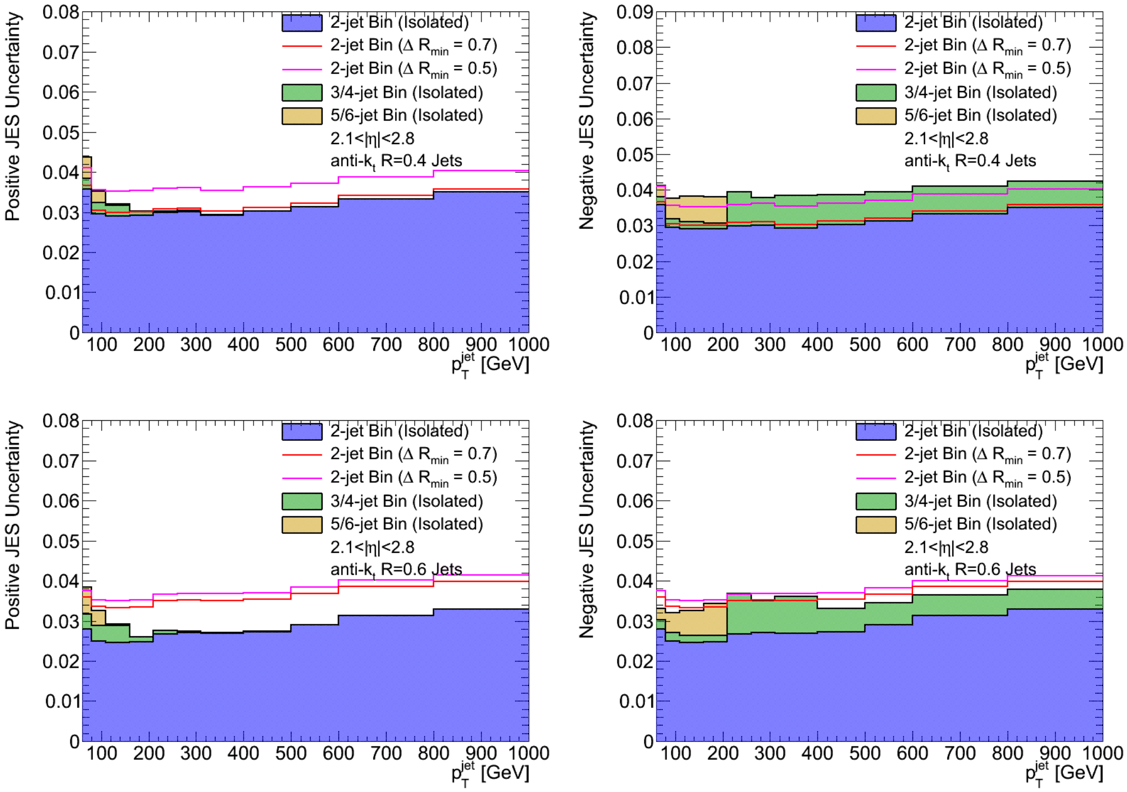

Figure D.6: *Total upwards (top) and downwards (bottom) jet energy scale uncertainties as a function of jet p_{T} for isolated jets in different multiplicity bins, and non-isolated jets in events with two selected jets. Jets reconstructed with the anti-k_t algorithm with resolution parameters $R = 0.4$ (top) and $R = 0.6$ (bottom) and with $2.1 < |\eta| < 2.8$ were used in this figure.*

Appendix E

Additional Studies on the Next-to-Leading-Order Calculations

This appendix presents results that complement those shown in Section 9.2. Figure E.1 shows the scale uncertainty in the NLO calculation of the three-to-two-jet cross-section ratios as a function of different characteristic event scales. This figure complements the results shown in Figures 9.1 and 9.2 and justifies the choice of ${H_\mathrm{T}}^{(2)}$ for the measurements performed in this thesis.

Figure E.2 compares the non-perturbative correction factors for the three-to-two-jet cross-section ratios as a function of ${H_\mathrm{T}}^{(2)}$ as calculated using Alpgen+Herwig/Jimmy and Pythia. The results are consistent with each other, justifying the use of Pythia for the final calculation, due to the higher statistics available in the samples generated with Rivet.

Figures E.3 through E.5 show for completeness the non-perturbative corrections and systematic uncertainties calculated for the three-to-two-jet cross-section ratios as a function of leading jet p_T, as done for the same variable as a function of event ${H_\mathrm{T}}^{(2)}$ in Figure 9.3. A summary of these results can be found in Section 9.2.2.

Finally, Figures E.6 through E.9 present results similar to those presented in Figure 9.4 for all other observables compared to next-to-leading-order calculations in Chapter 10. The next-to-leading-order calculations and uncertainties are shown as obtained using both the CTEQ6.6 and MSTW2008nlo pdf sets, and for jets built using a resolution parameter $R = 0.4$ and $R = 0.6$. A summary of these results can be found in Section 9.2.3.

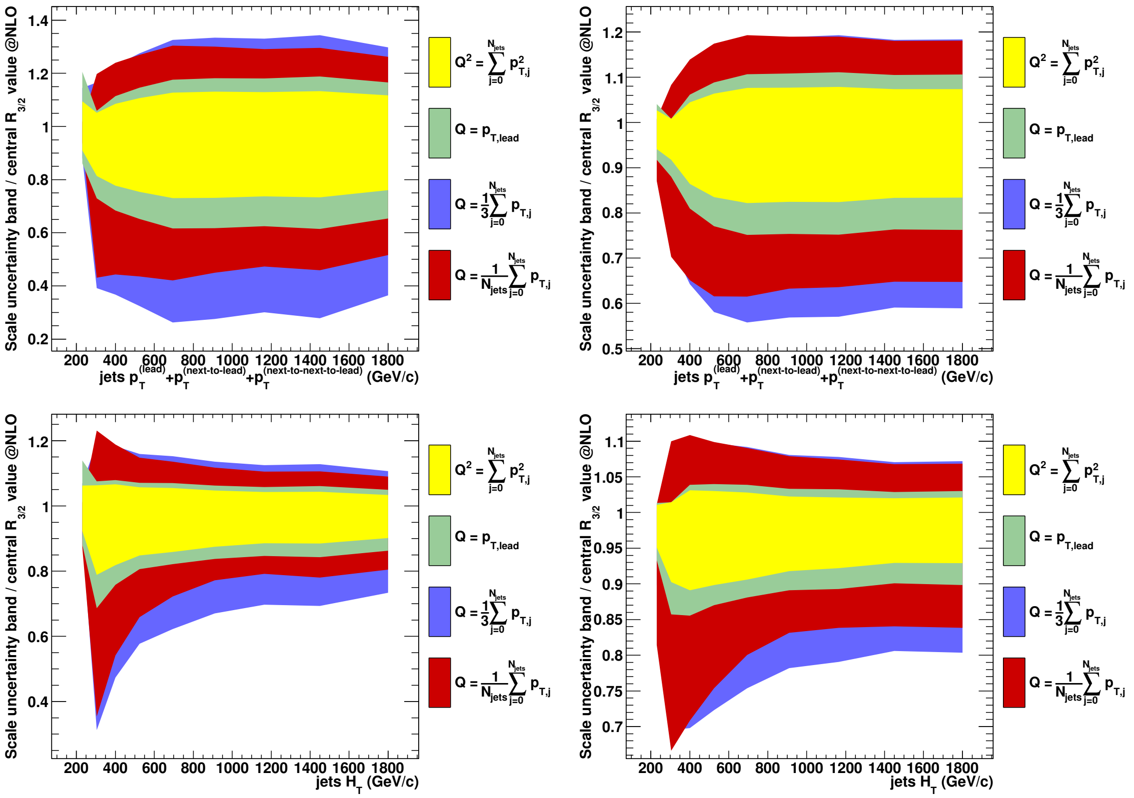

Figure E.1: *Scale systematic uncertainty in the NLO prediction of the three-to-two-jet cross-section ratio as a function of* $H_{\mathrm{T}}{}^{(3)} = p_{\mathrm{T}}{}^{\mathrm{lead}} + p_{\mathrm{T}}{}^{\mathrm{2nd}} + p_{\mathrm{T}}{}^{\mathrm{3rd}}$ *(left) and* H_{T} *(right) for several scale choices [163]. The results are shown for jets reconstructed using the anti-*k_t *algorithm with resolution parameters* $R = 0.4$ *(top) and* $R = 0.6$ *(bottom).*

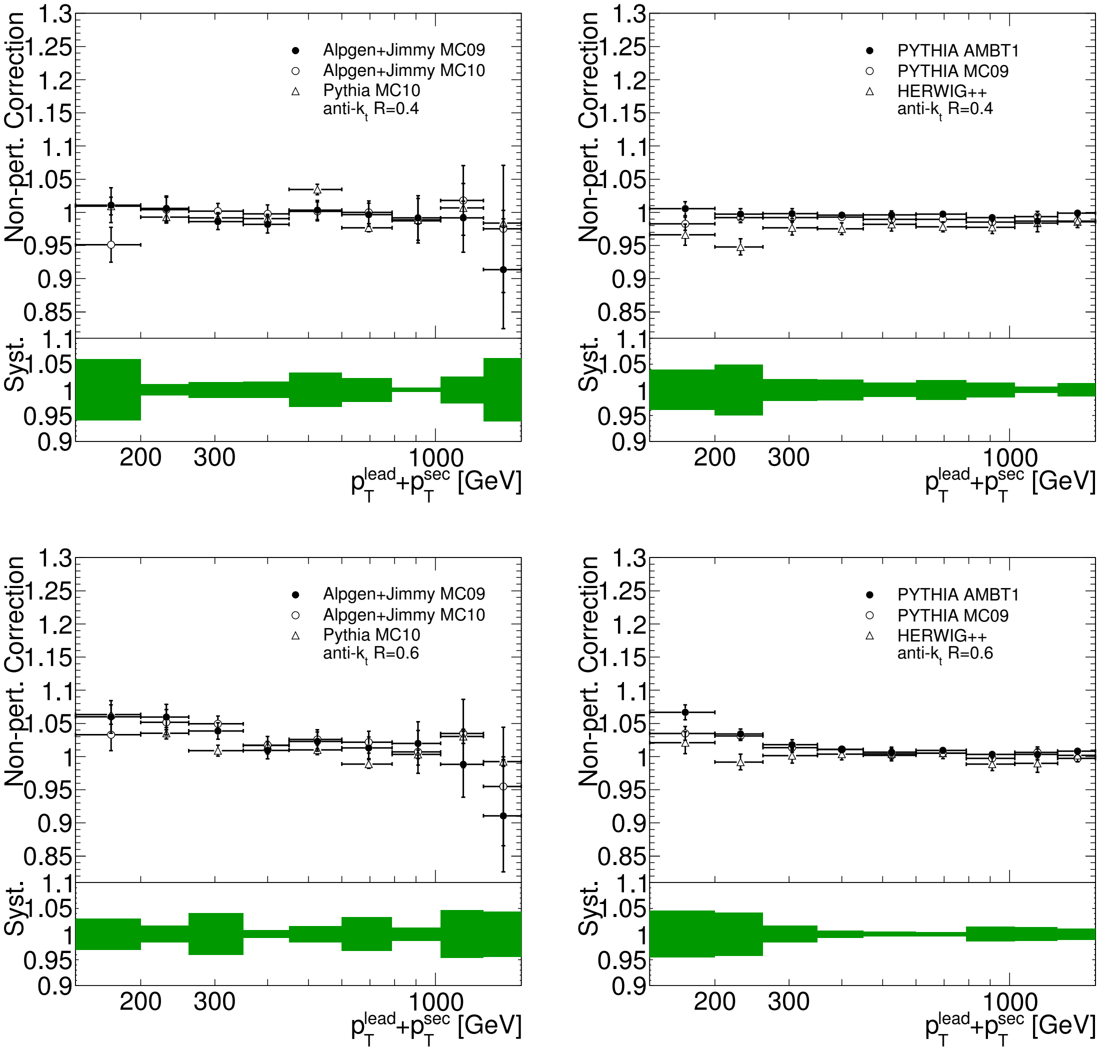

Figure E.2: *Non-perturbative corrections to the three-to-two-jet cross-section ratio NLO prediction as a function of* ${H_{\mathrm{T}}}^{(2)}$ *and associated systematic uncertainties. The factors are shown as obtained using different leading-order Monte Carlo simulations. The results are obtained with* ALPGEN *samples (left) and the default samples used in the calculation (right). The* PYTHIA *sample used on the left corresponds to the official ATLAS production, while that on the right was generated using Rivet. Results are shown for jets built with a resolution parameter* $R = 0.4$ *(top) and* $R = 0.6$ *(bottom).*

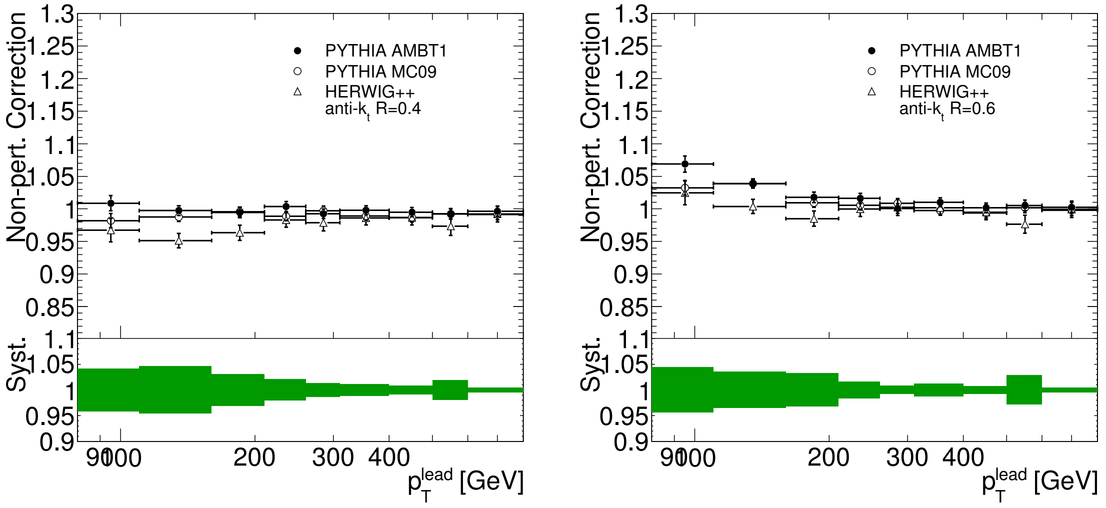

Figure E.3: *Non-perturbative corrections to the three-to-two-jet cross-section ratio NLO prediction as a function of leading jet p_{T} and associated systematic uncertainties. The factors are shown as obtained using different leading-order Monte Carlo simulations for jets reconstructed with resolution parameters $R = 0.4$ (left) and $R = 0.6$ (right).*

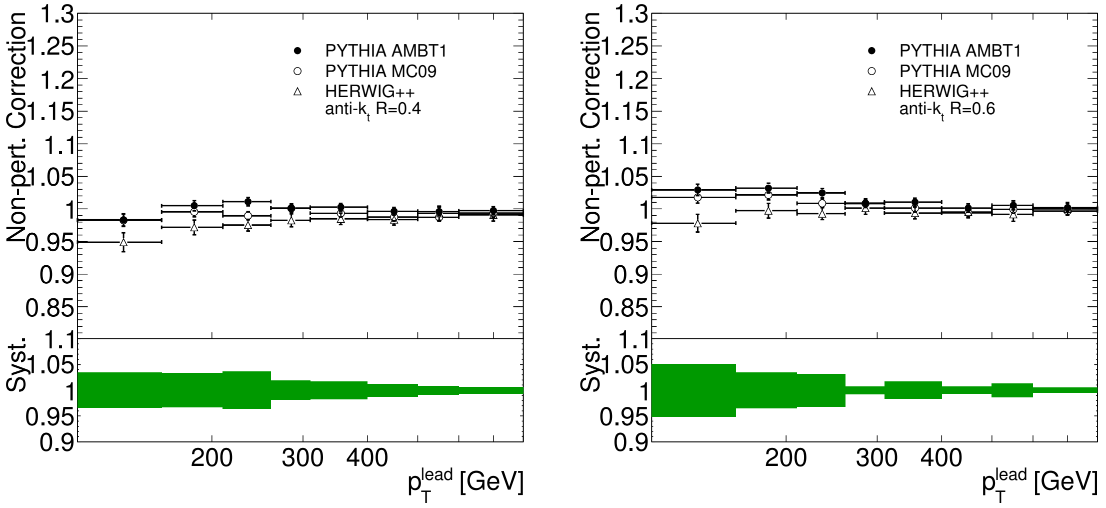

Figure E.4: *Non-perturbative corrections to the three-to-two-jet cross-section ratio NLO prediction as a function of leading jet p_{T} for the second set of kinematic cuts used for these studies and associated systematic uncertainties. The factors are shown as obtained using different leading-order Monte Carlo simulations for jets reconstructed with resolution parameters $R = 0.4$ (left) and $R = 0.6$ (right).*

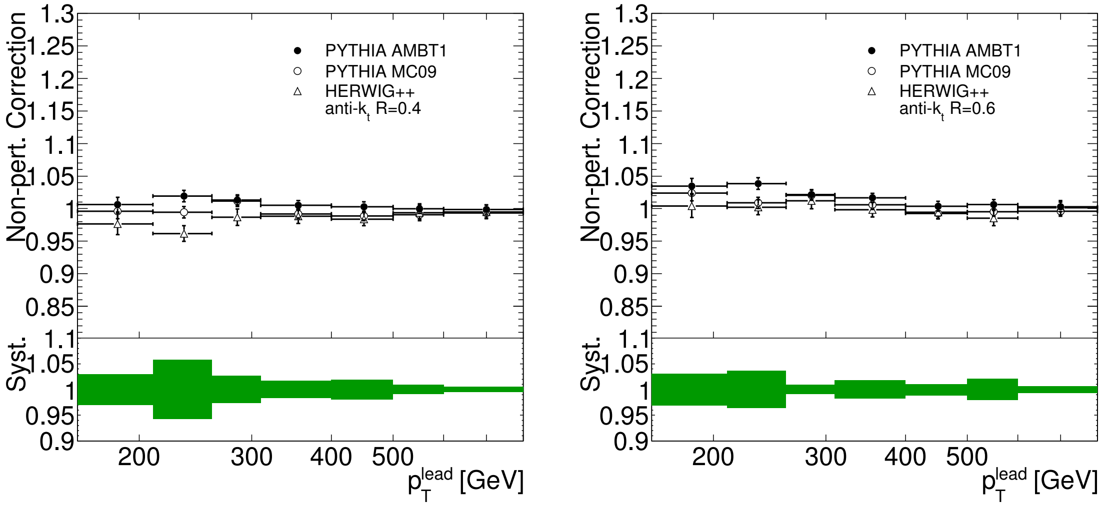

Figure E.5: *Non-perturbative corrections to the three-to-two-jet cross-section ratio NLO prediction as a function of leading jet* p_{T} *for the third set of kinematic cuts used for these studies and associated systematic uncertainties. The factors are shown as obtained using different leading-order Monte Carlo simulations for jets reconstructed with resolution parameters* $R = 0.4$ *(left) and* $R = 0.6$ *(right).*

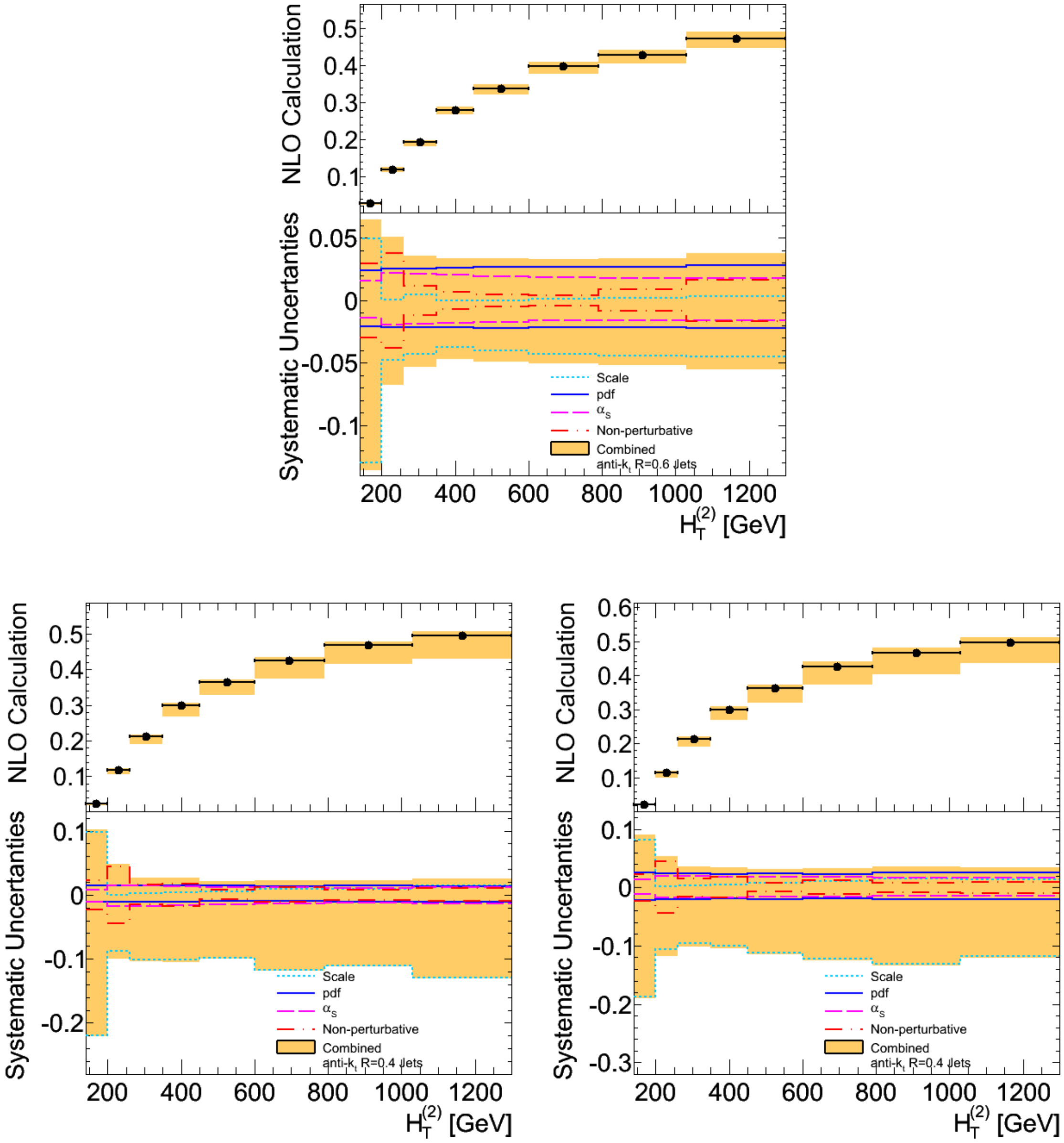

Figure E.6: *NLO calculation of the three-to-two-jet cross-section ratio as a function of* $H_{\mathrm{T}}{}^{(2)}$ *and the estimated systematic uncertainty (band). Error bars on the predicted points corresponding to the statistical uncertainty in the calculation are small and not visible. The size of the three individual uncertainties that contribute to the total uncertainty is shown at the bottom. Results are shown for jets built using the anti-*k_t *algorithm with resolution parameters* $R = 0.6$ *(top) and* $R = 0.4$ *(bottom) and are obtained using the* CTEQ6.6 *(top and bottom right) and* MSTW2008nlo *(bottom left) pdf sets.*

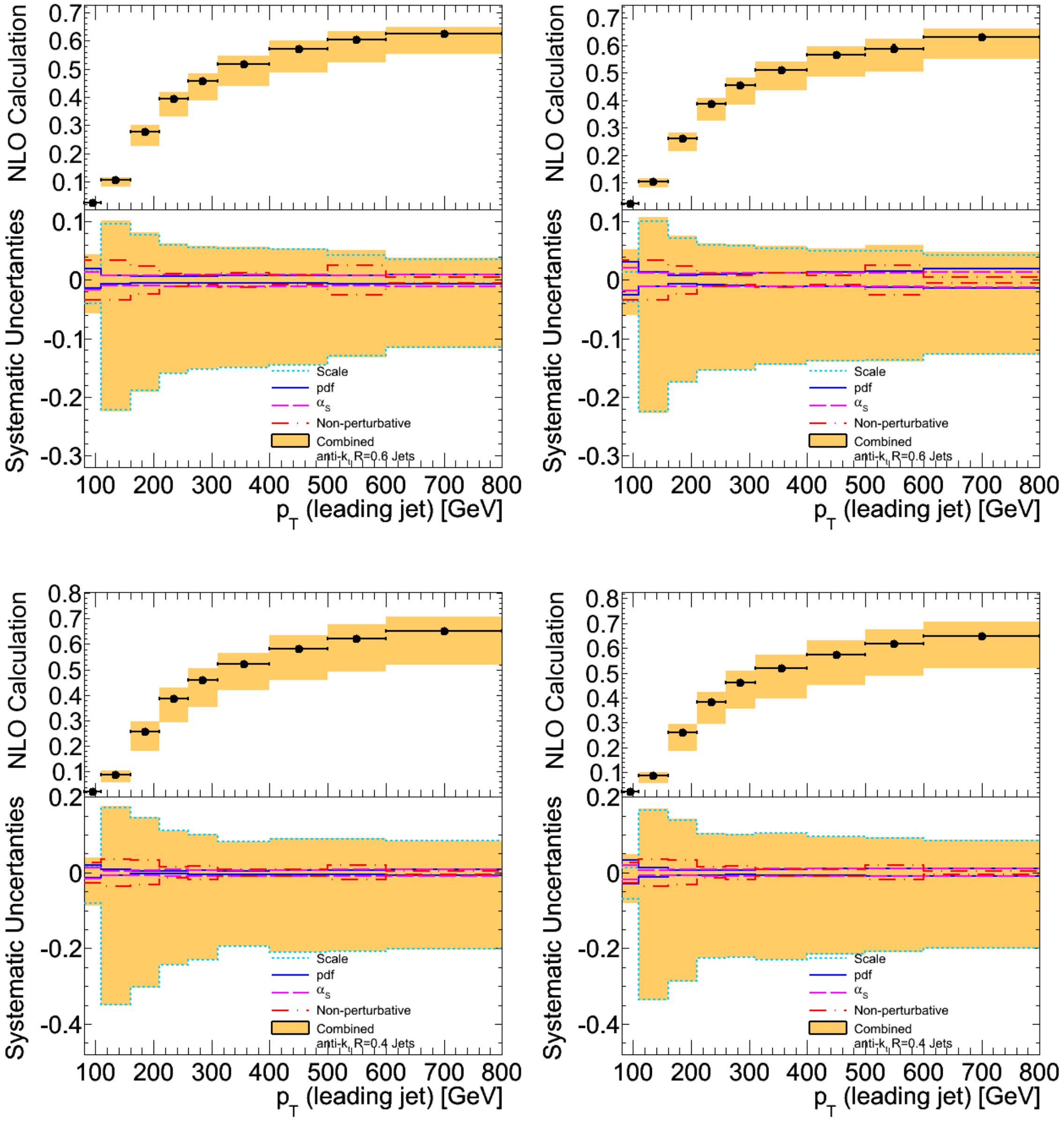

Figure E.7: *NLO calculation of the three-to-two-jet cross-section ratio as a function of leading jet* p_{T} *and the estimated systematic uncertainty (band). Error bars on the predicted points corresponding to the statistical uncertainty in the calculation are small and not visible. The size of the three individual uncertainties that contribute to the total uncertainty is shown at the bottom. Results are shown for jets built using the anti-*k_t *algorithm with resolution parameter* $R = 0.6$ *(top) and* $R = 0.4$ *(bottom) and are obtained using the* MSTW2008nlo *(left) and* CTEQ6.6 *(right) pdf sets.*

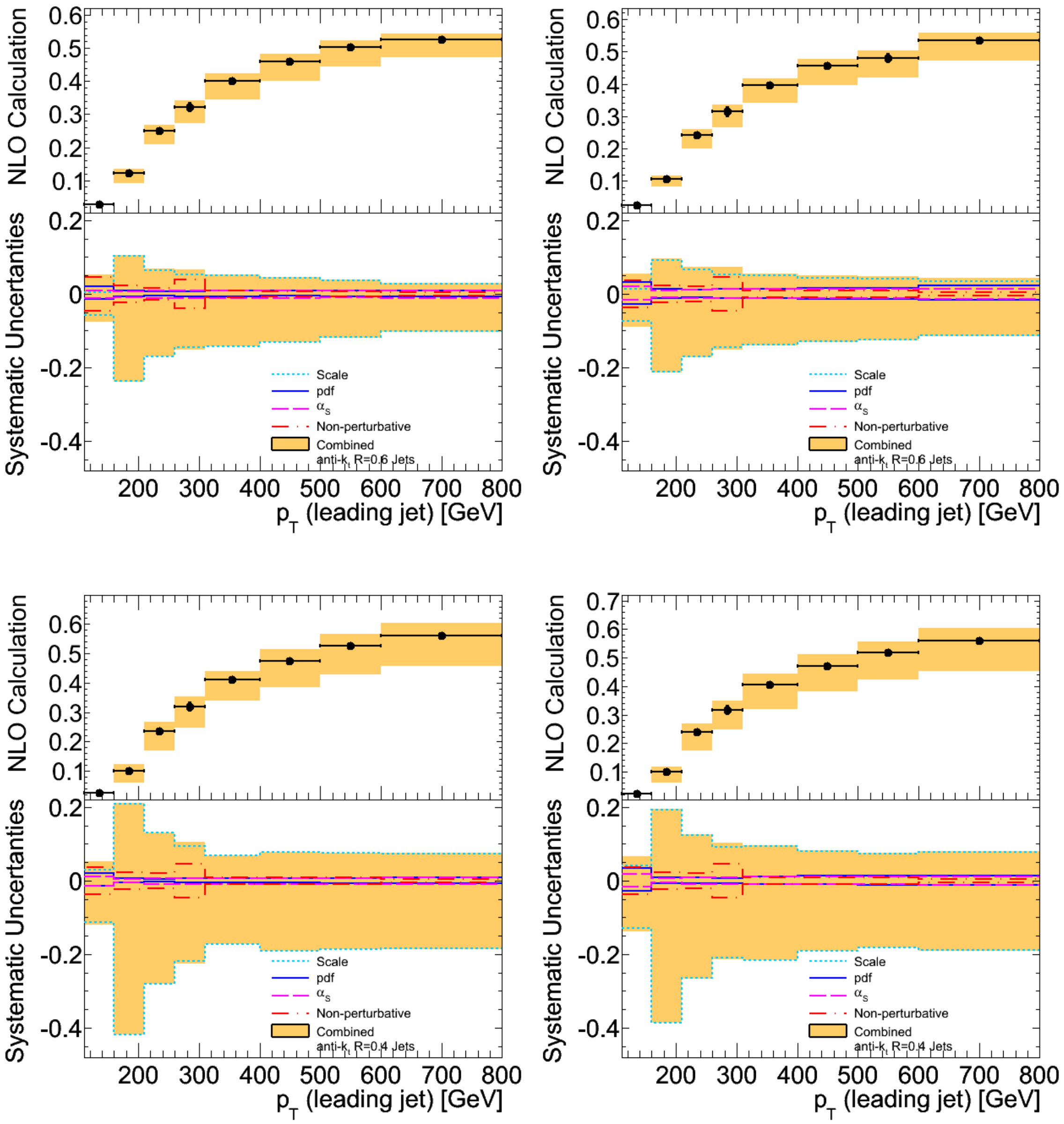

Figure E.8: *NLO calculation of the three-to-two-jet cross-section ratio as a function of leading jet* p_T *for the second set of kinematic cuts studied and the estimated systematic uncertainty (band). Error bars on the predicted points corresponding to the statistical uncertainty in the calculation are small and not visible. The size of the three individual uncertainties that contribute to the total uncertainty is shown at the bottom. Results are shown for jets built using the anti-*k_t *algorithm with resolution parameter* $R = 0.6$ *(top) and* $R = 0.4$ *(bottom) and are obtained using the* MSTW2008nlo *(left) and* CTEQ6.6 *(right) pdf sets.*

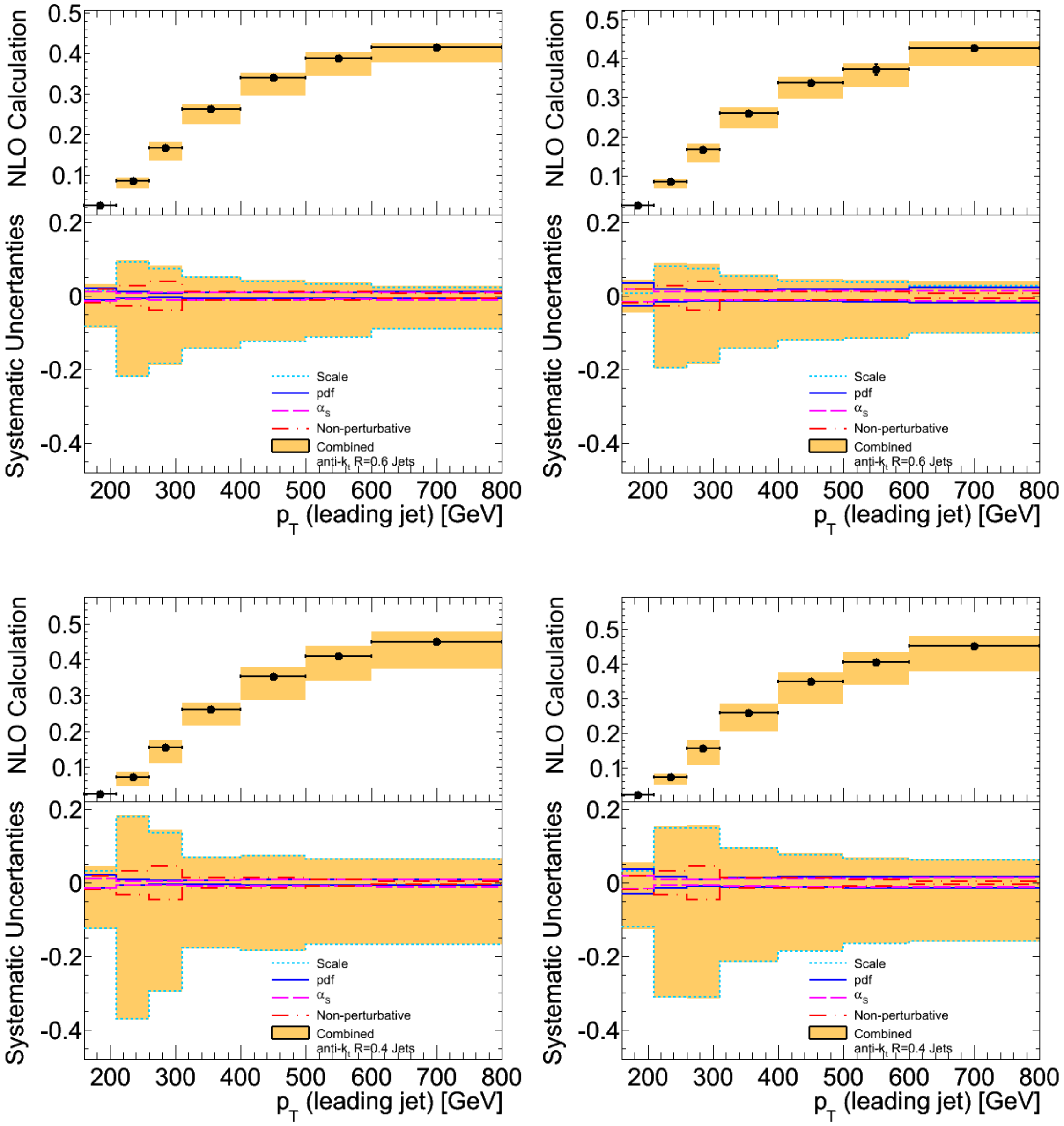

Figure E.9: *NLO calculation of the three-to-two-jet cross-section ratio as a function of leading jet* p_{T} *for the third set of kinematic cuts studied and the estimated systematic uncertainty (band). Error bars on the predicted points corresponding to the statistical uncertainty in the calculation are small and not visible. The size of the three individual uncertainties that contribute to the total uncertainty is shown at the bottom. Results are shown for jets built using the anti-*k_t *algorithm with resolution parameter* $R = 0.6$ *(top) and* $R = 0.4$ *(bottom) and are obtained using the MSTW2008nlo (left) and CTEQ6.6 (right) pdf sets.*

Appendix F

Additional Leading-Order and Next-to-Leading-Order Comparisons to the Data

This appendix presents additional comparisons to leading-order and next-to-leading-order calculations. Section F.1 presents additional comparisons of the data to the leading-order calculations presented in Table 9.1. The calculations shown in Chapter 10 are also shown for reference. The PYTHIA MC09 and MC09$'$ calculations are not shown for clarity, but they follow closely the AMBT1 prediction. In addition, the results obtained with ALPGEN+HERWIG/JIMMY samples generated with different renormalization and factorization scales are shown as dotted lines above and below the ALPGEN+HERWIG/JIMMY AUET1 prediction in the left side figures. The renormalization and factorization scales are varied simultaneously by a factor of two up and down from the nominal value used in the ALPGEN samples to obtain these curves. While this does not constitute an uncertainty in the calculation, for reasons detailed in Chapter 10, it does give an idea of the intrinsic limitations in the leading-order calculation.

Section F.2 presents comparisons of kinematic observables in the two and three-jet bin in data and next-to-leading-order calculations. The calculations are not renormalized to the data, thus providing an idea of the accuracy of the next-to-leading-order predictions. The predictions for jets reconstructed with $R = 0.6$ are typically closer to the data than those for jets reconstructed with $R = 0.4$. Also, large uncertainties in the theoretical calculations arise near the kinematic thresholds for observables in the two-jet bin. These are caused by the samples generated with a lower renormalization scale, and point to a limitation in the next-to-leading-order calculation and also the renormalization scale chosen, which would not be optimal for the study of these observables.

Finally, Section F.3 presents additional comparisons of measurements of the three-to-two-jet cross-section ratios to next-to-leading-order calculations, including comparisons to the calculations performed with the `CTEQ6.6` pdf set and using jets reconstructed with $R = 0.4$. The conclusions drawn from these figures have already been discussed in Chapter 10.

F.1 Additional Comparisons to Leading-Order Monte Carlo Simulations

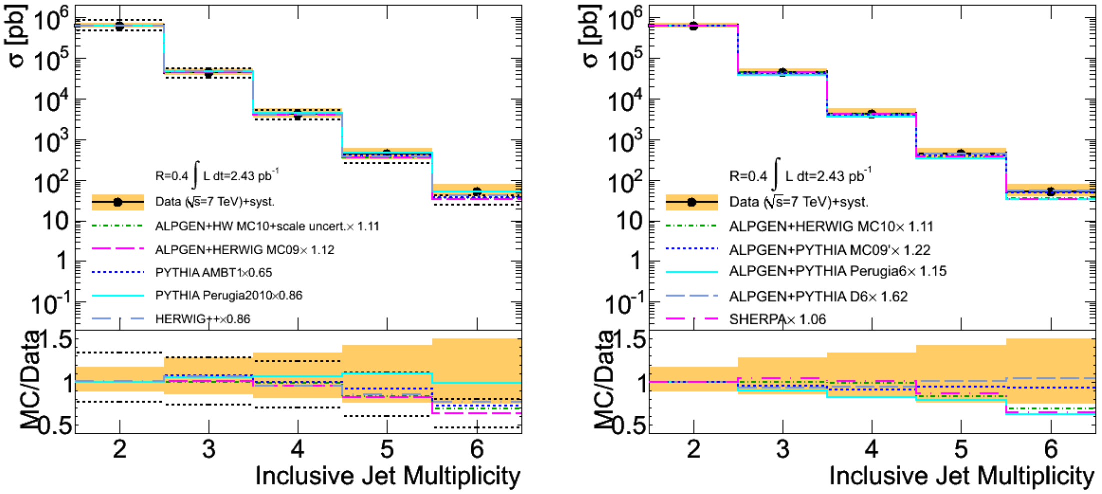

Figure F.1: *Cross section as a function of inclusive jet multiplicity for jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$. The results are compared to additional leading-order Monte Carlo simulations normalized to the measured two-jet inclusive cross section. The error bands correspond to the systematic uncertainty in the measurement. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

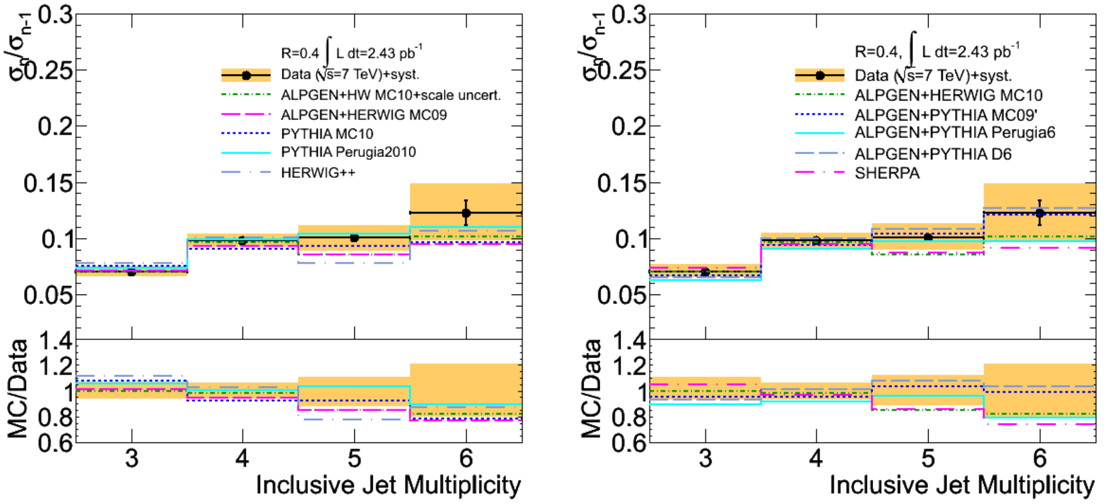

Figure F.2: *Ratio of the n-jet cross section to the $(n-1)$-jet cross section for values of n varying from three to six. The results are compared to additional leading-order Monte Carlo simulations for jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$. The error bands correspond to the systematic uncertainty in the measurement. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

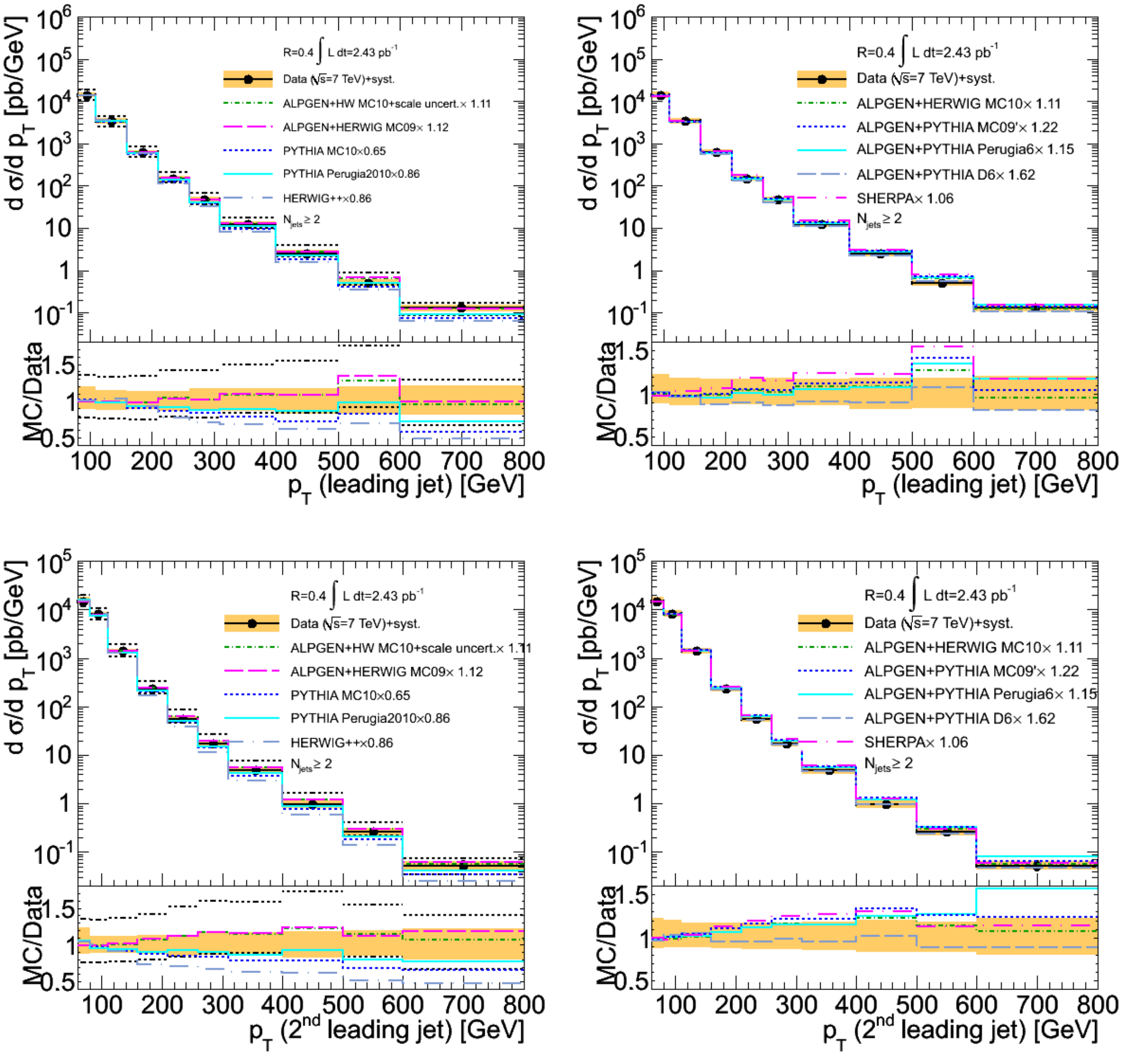

Figure F.3: *Differential cross section as a function of leading jet* p_T *(top), and second leading jet* p_T *(bottom) for events with* $N_{\text{jets}} \geq 2$*. Jets built using the anti-*k_t *algorithm with resolution parameter* $R = 0.4$ *are used in this study. The results are compared to additional leading-order Monte Carlo simulations normalized to the measured two-jet inclusive cross section. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

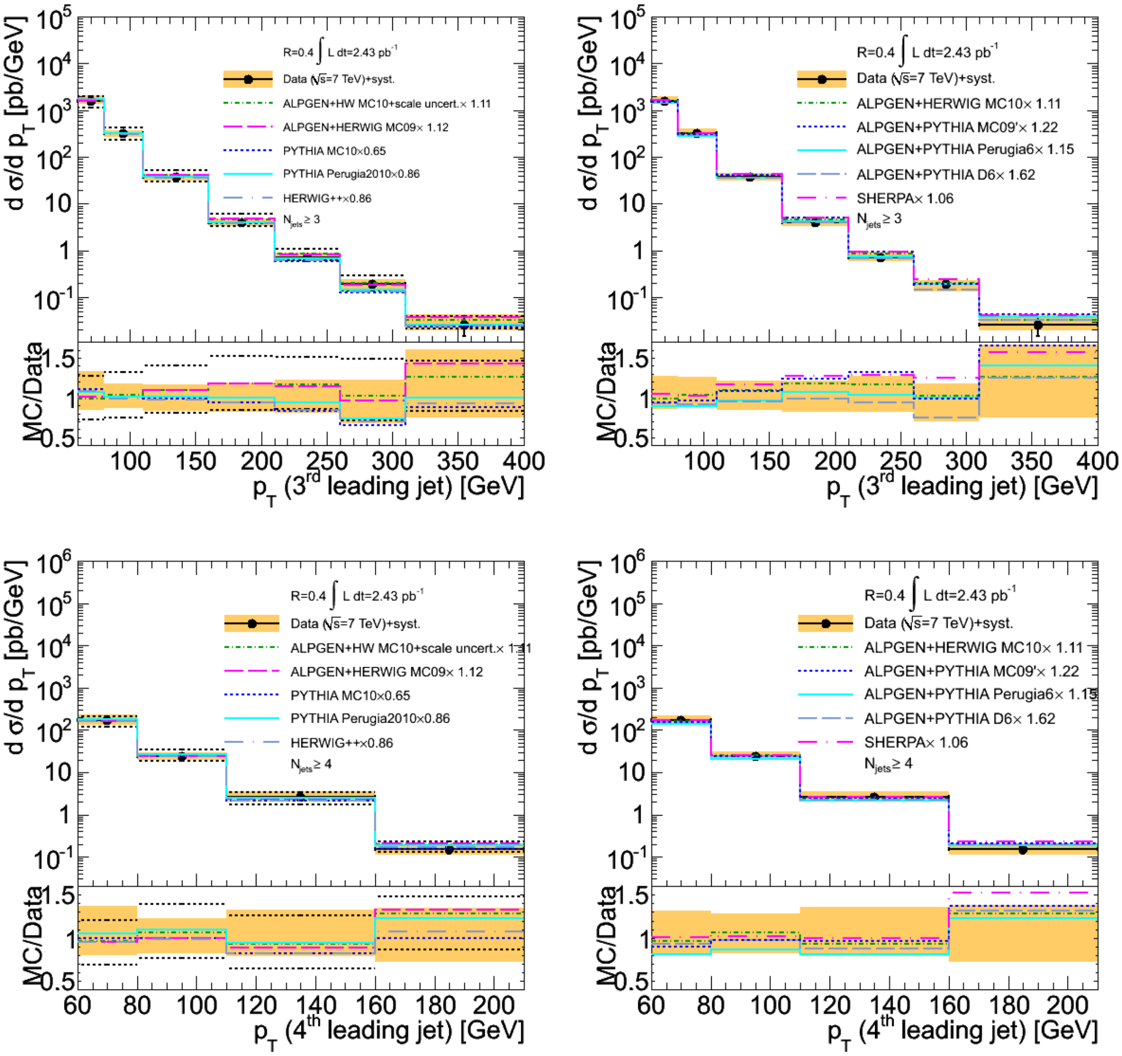

Figure F.4: *Differential cross section as a function of third leading jet p_{T} for events with $N_{\mathrm{jets}} \geq 3$ (top) and fourth leading jet p_{T} for events with $N_{\mathrm{jets}} \geq 4$ (bottom). Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ are used in this study. The results are compared to additional leading-order Monte Carlo simulations normalized to the measured two-jet inclusive cross section. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

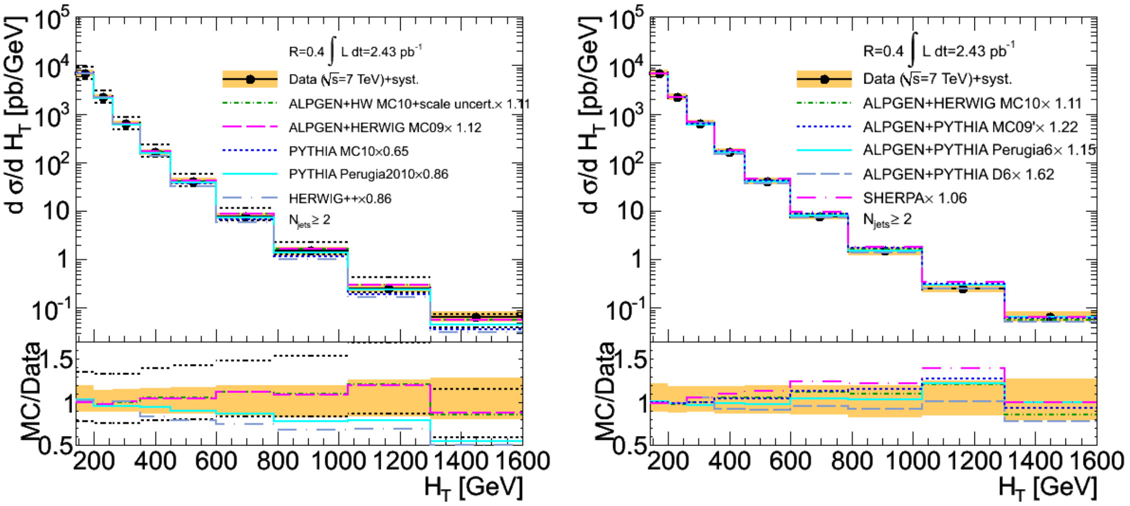

Figure F.5: *Differential cross section as a function of H_{T} for events with at least two selected jets. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ are used in this study. The results are compared to additional leading-order Monte Carlo simulations normalized to the measured two-jet inclusive cross section. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

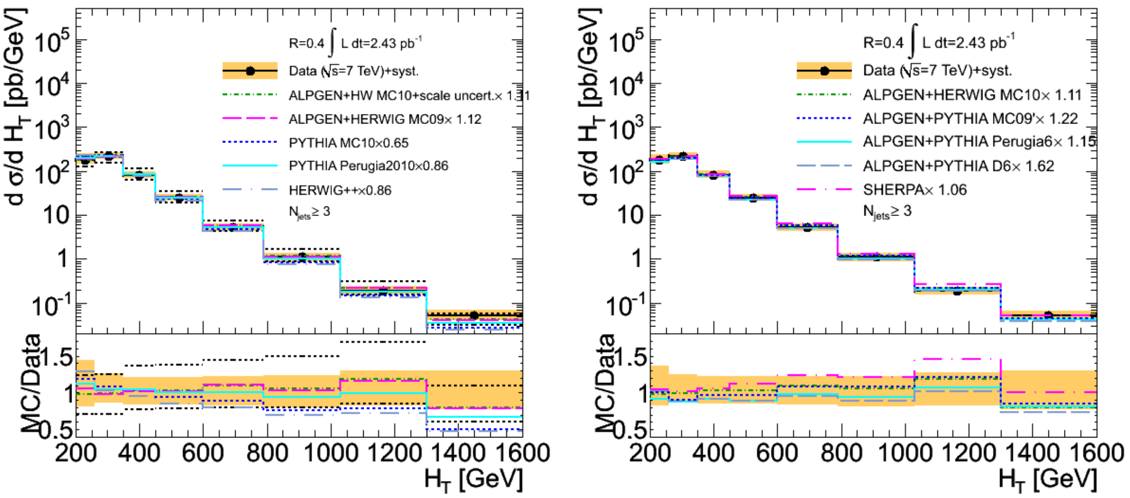

Figure F.6: *Differential cross section as a function of H_{T} for events with at least three selected jets. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ are used in this study. The results are compared to additional leading-order Monte Carlo simulations normalized to the measured two-jet inclusive cross section. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

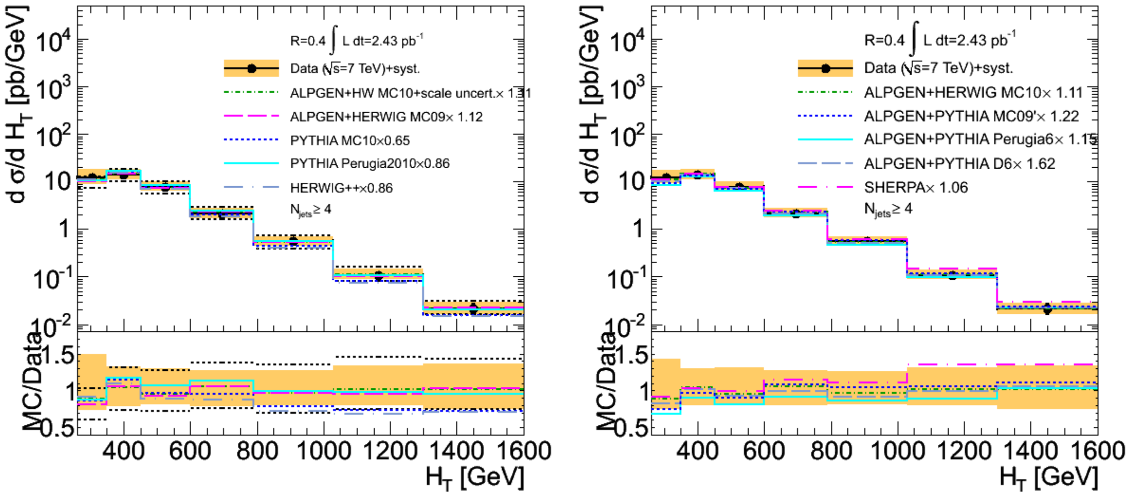

Figure F.7: *Differential cross section as a function of H_{T} for events with at least four selected jets. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ are used in this study. The results are compared to additional leading-order Monte Carlo simulations normalized to the measured two-jet inclusive cross section. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

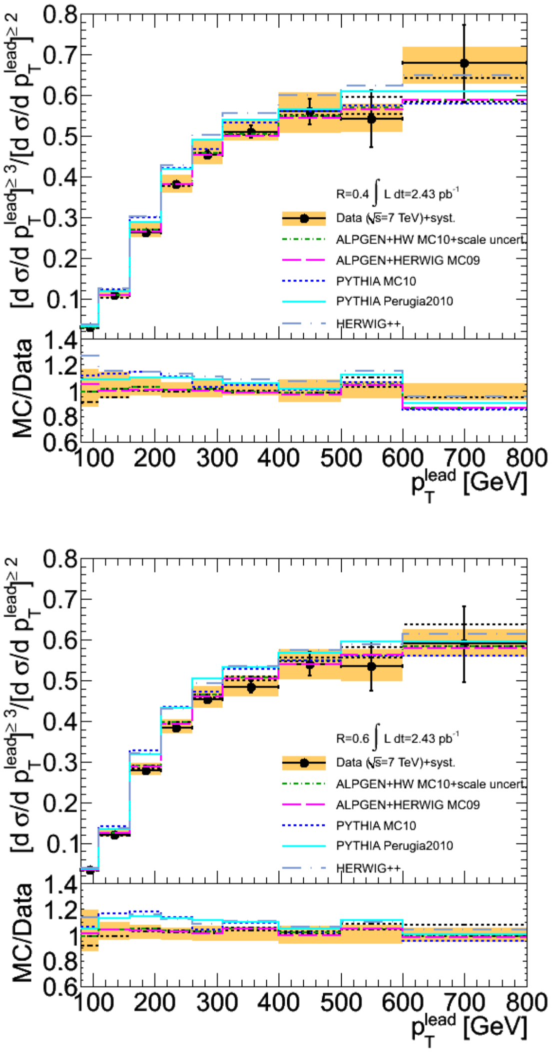

Figure F.8: *Ratio of the three-jet to two-jet differential cross section as a function of the leading jet* p_{T} *using the default kinematic cuts on the jet selection. Jets built with the anti-*k_t *algorithm using a resolution parameter* $R = 0.4$ *(top) and* $R = 0.6$ *(bottom) are used in this study. The results are compared to additional leading-order Monte Carlo simulations. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

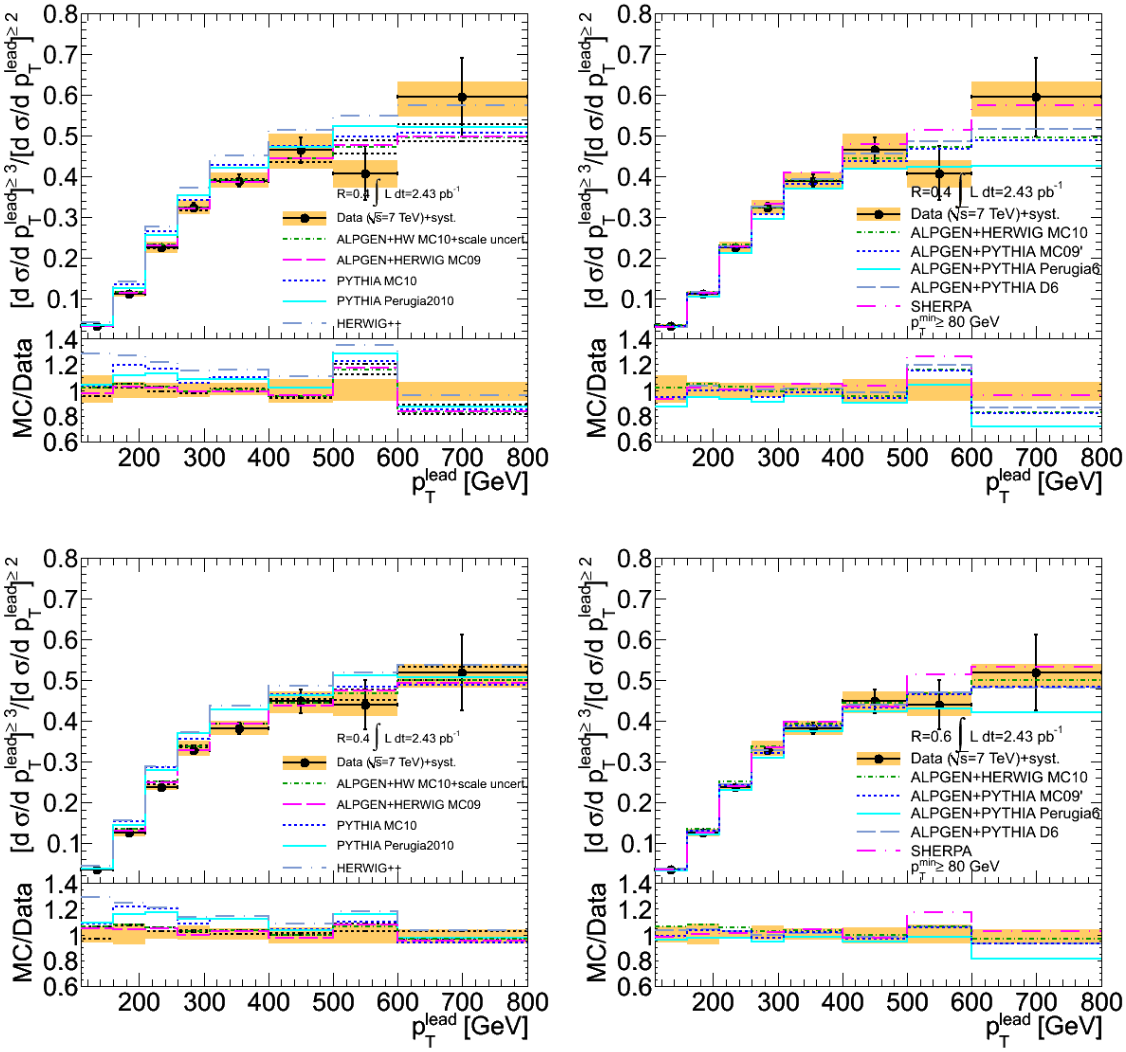

Figure F.9: *Ratio of the three-jet to two-jet differential cross section as a function of the leading jet* p_{T} *using the second set of kinematic cuts on the jet selection. Jets built with the anti-*k_t *algorithm using a resolution parameter* $R = 0.4$ *(top) and* $R = 0.6$ *(bottom) are used in this study. The results are compared to additional leading-order Monte Carlo simulations. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

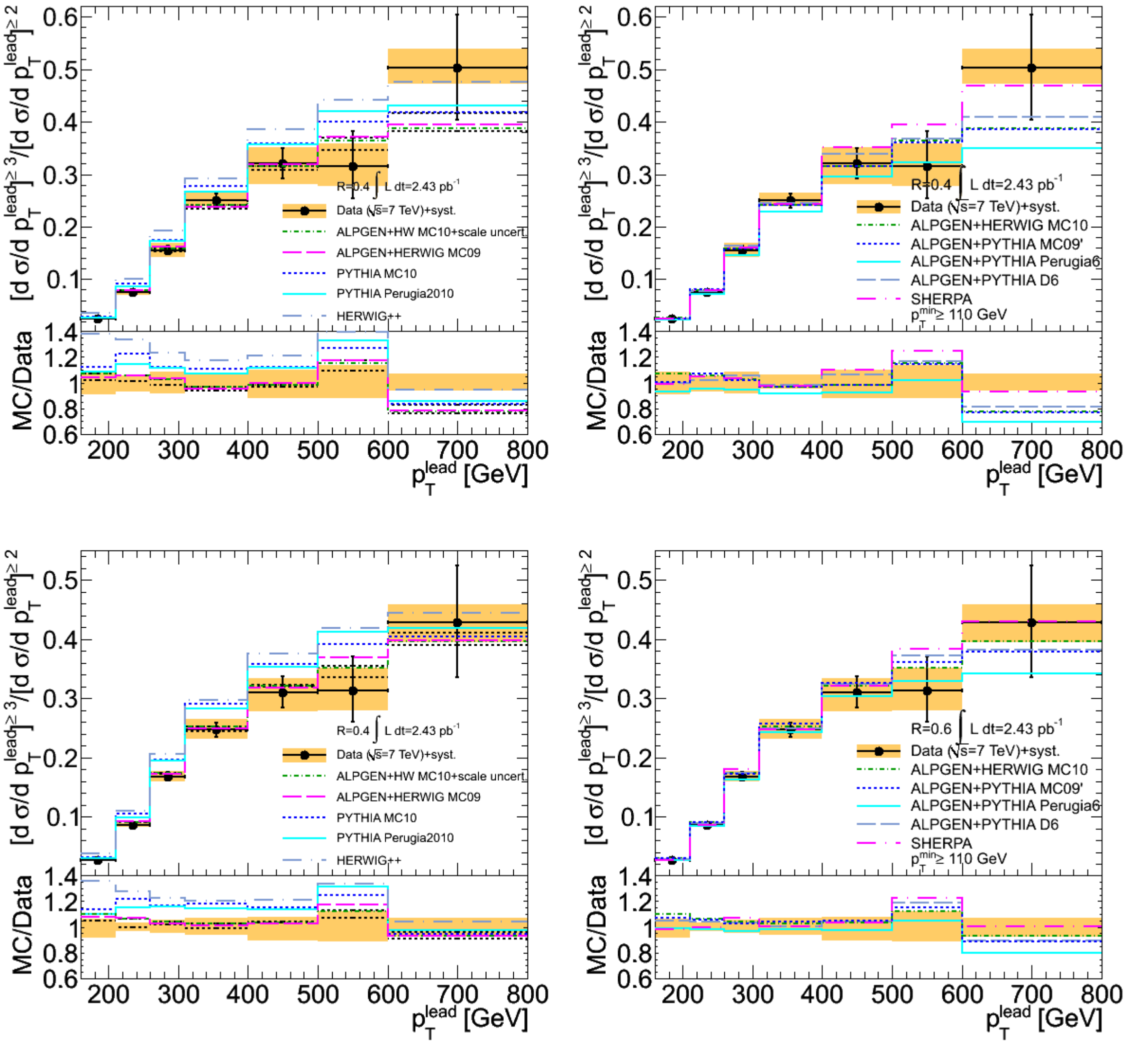

Figure F.10: *Ratio of the three-jet to two-jet differential cross section as a function of the leading jet* p_{T} *using the third set of kinematic cuts on the jet selection. Jets built with the anti-*k_t *algorithm using a resolution parameter* $R = 0.4$ *(top) and* $R = 0.6$ *(bottom) are used in this study. The results are compared to additional leading-order Monte Carlo simulations. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

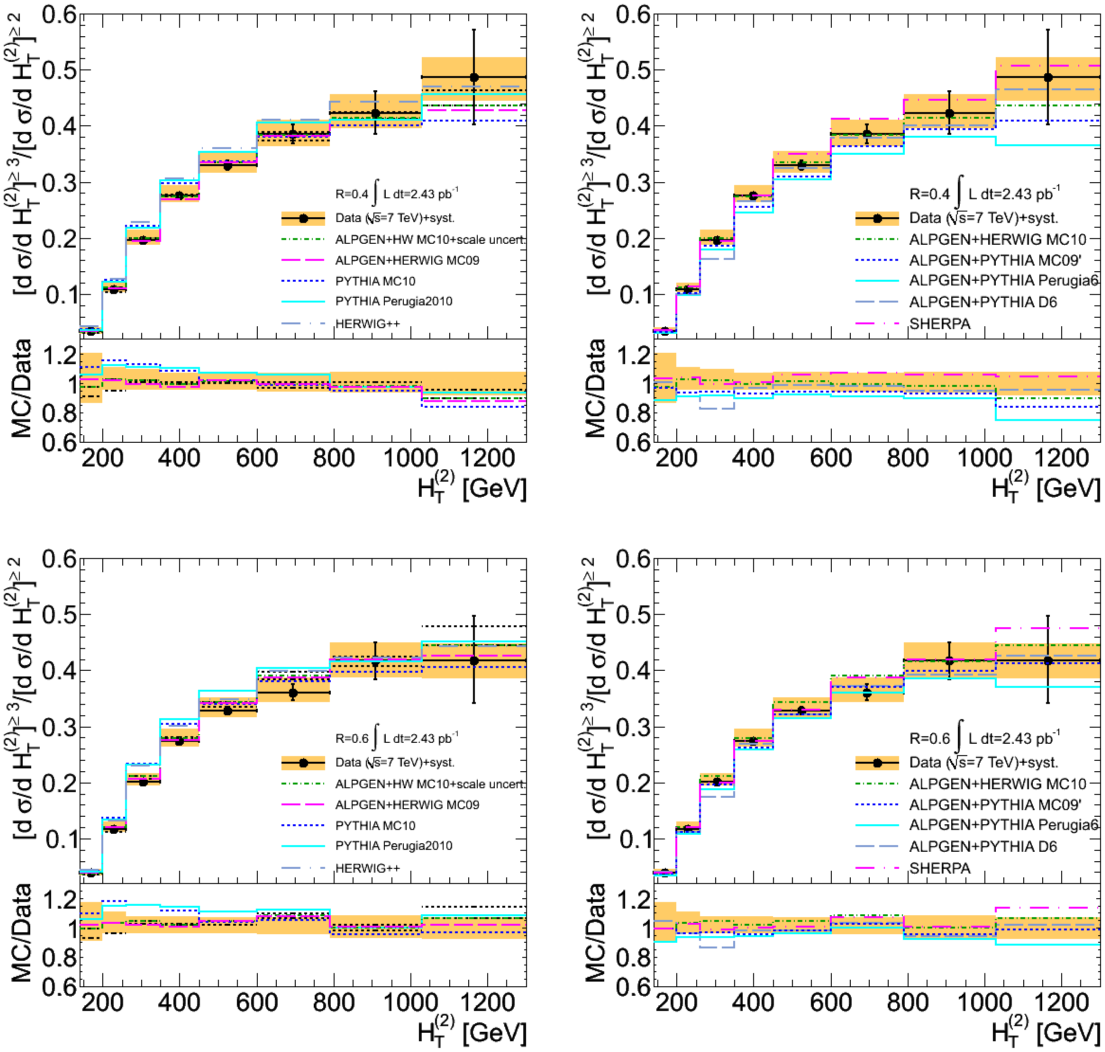

Figure F.11: *Ratio of the three-jet to two-jet differential cross section as a function of the event* ${H_T}^{(2)}$*. Jets built with the anti-*k_t *algorithm using a resolution parameter* $R = 0.4$ *(top) and* $R = 0.6$ *(bottom) are used in this study. The results are compared to additional leading-order Monte Carlo simulations. The error bands correspond to the systematic uncertainties in the measurements. Plots of the ratios of the different Monte Carlo simulations to the data are presented at the bottom of each figure.*

F.2 Comparisons of Kinematic Variables to Next-to-Leading-Order Calculations

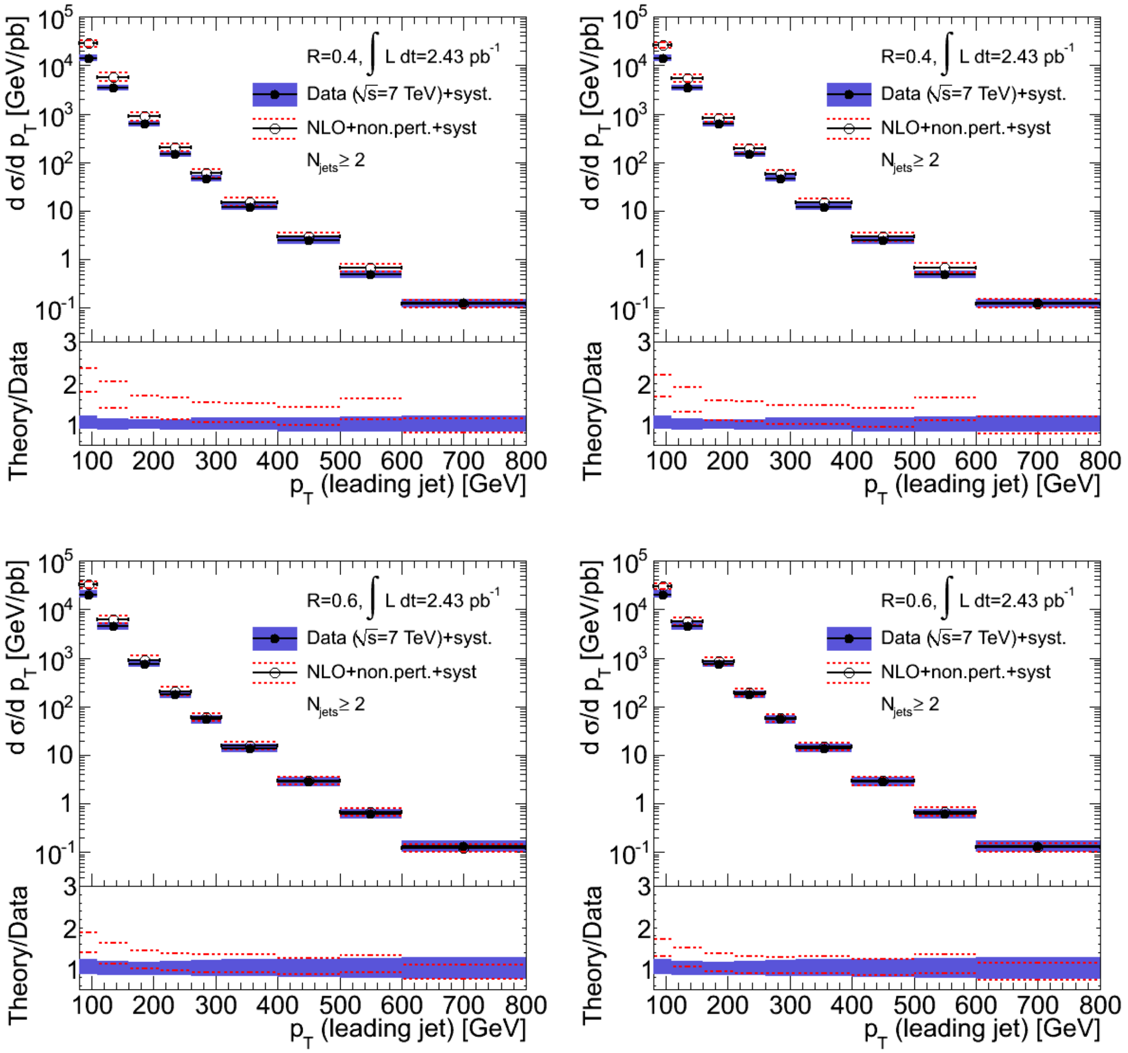

Figure F.12: *Differential cross section as a function of leading jet p_{T} in two-jet events compared to a next-to-leading-order calculation. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ (top) and $R = 0.6$ (bottom) are used in this study. The results are compared to a next-to-leading-order calculation performed with the* MSTW2008nlo *(left) and the* CTEQ6.6 *(right) pdf sets. The error bands correspond to the systematic uncertainties in the measurements. The systematic uncertainty in the next-to-leading-order calculation is shown as dotted lines above and below the calculation. Plots of the ratios of the theoretical calculations to the data are presented at the bottom of each figure.*

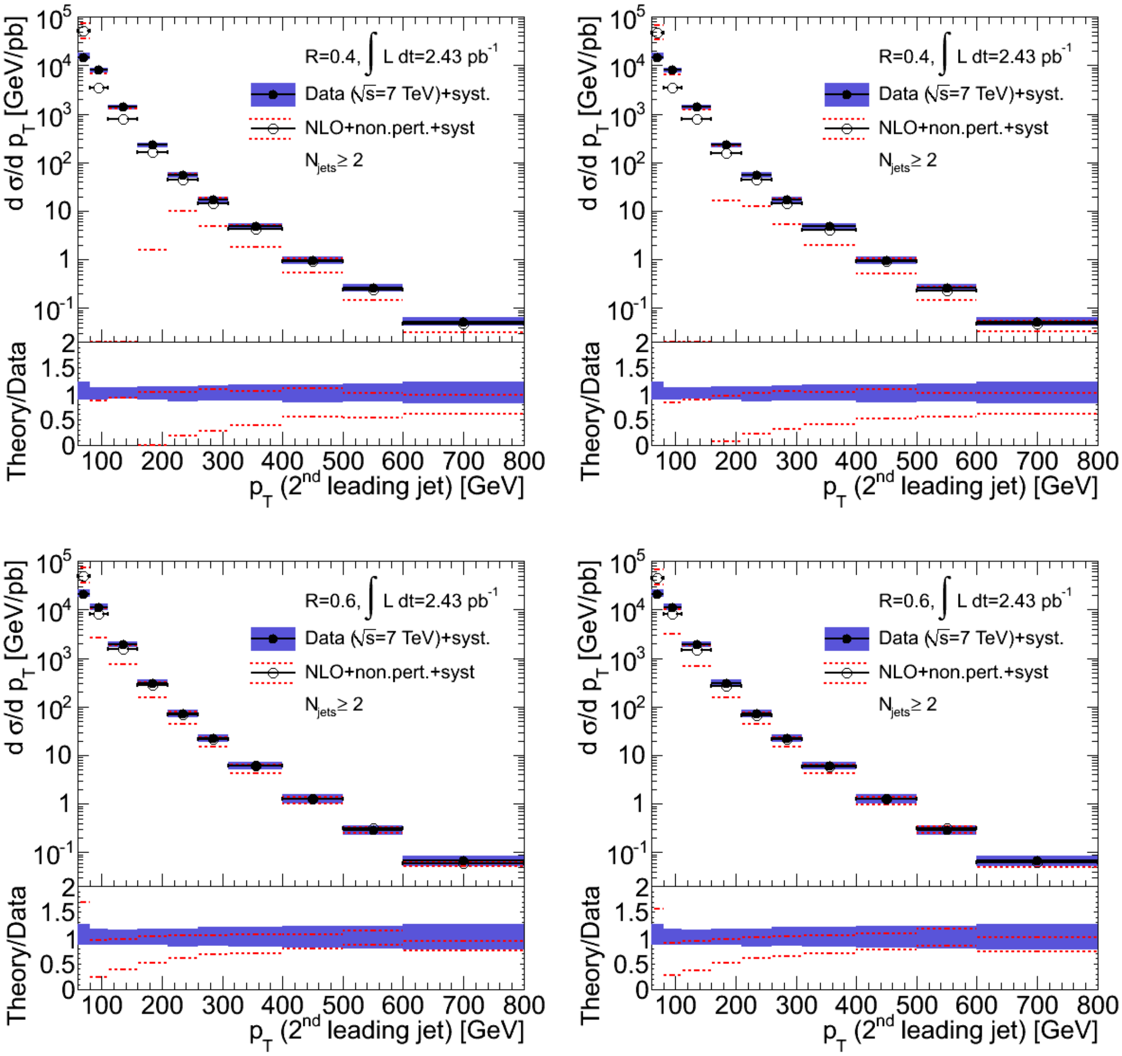

Figure F.13: *Differential cross section as a function of second leading jet $p_{\rm T}$ in two-jet events compared to a next-to-leading-order calculation. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ (top) and $R = 0.6$ (bottom) are used in this study. The results are compared to a next-to-leading-order calculation performed with the MSTW2008nlo (left) and the CTEQ6.6 (right) pdf sets. The error bands correspond to the systematic uncertainties in the measurements. The systematic uncertainty in the next-to-leading-order calculation is shown as dotted lines above and below the calculation. Plots of the ratios of the theoretical calculations to the data are presented at the bottom of each figure.*

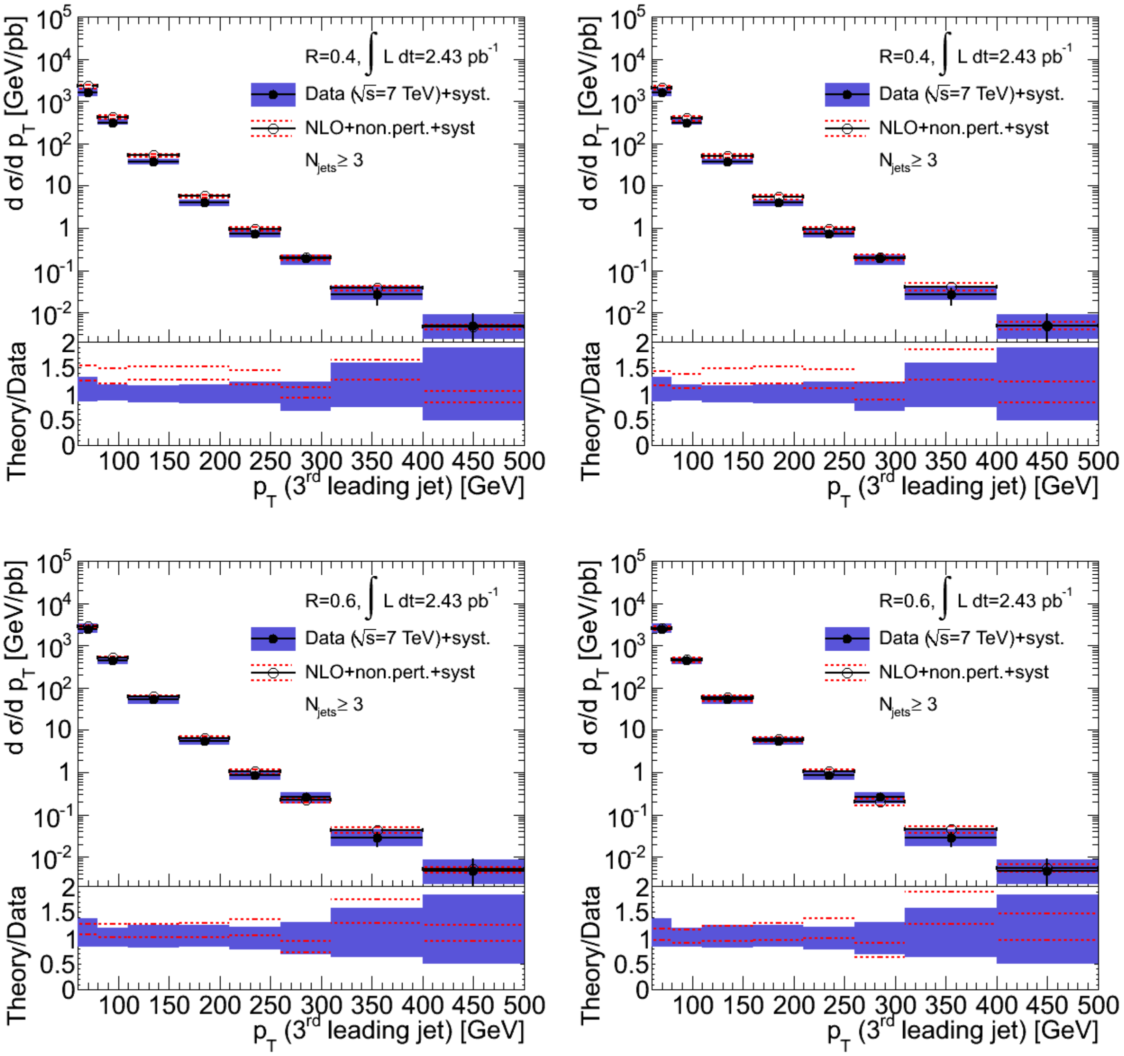

Figure F.14: *Differential cross section as a function of third leading jet p_{T} in three-jet events compared to a next-to-leading-order calculation. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ (top) and $R = 0.6$ (bottom) are used in this study. The results are compared to a next-to-leading-order calculation performed with the* MSTW2008nlo *(left) and the* CTEQ6.6 *(right) pdf sets. The error bands correspond to the systematic uncertainties in the measurements. The systematic uncertainty in the NLO calculation is shown as dotted lines above and below the calculation. Plots of the ratios of the theoretical calculations to the data are presented at the bottom of each figure.*

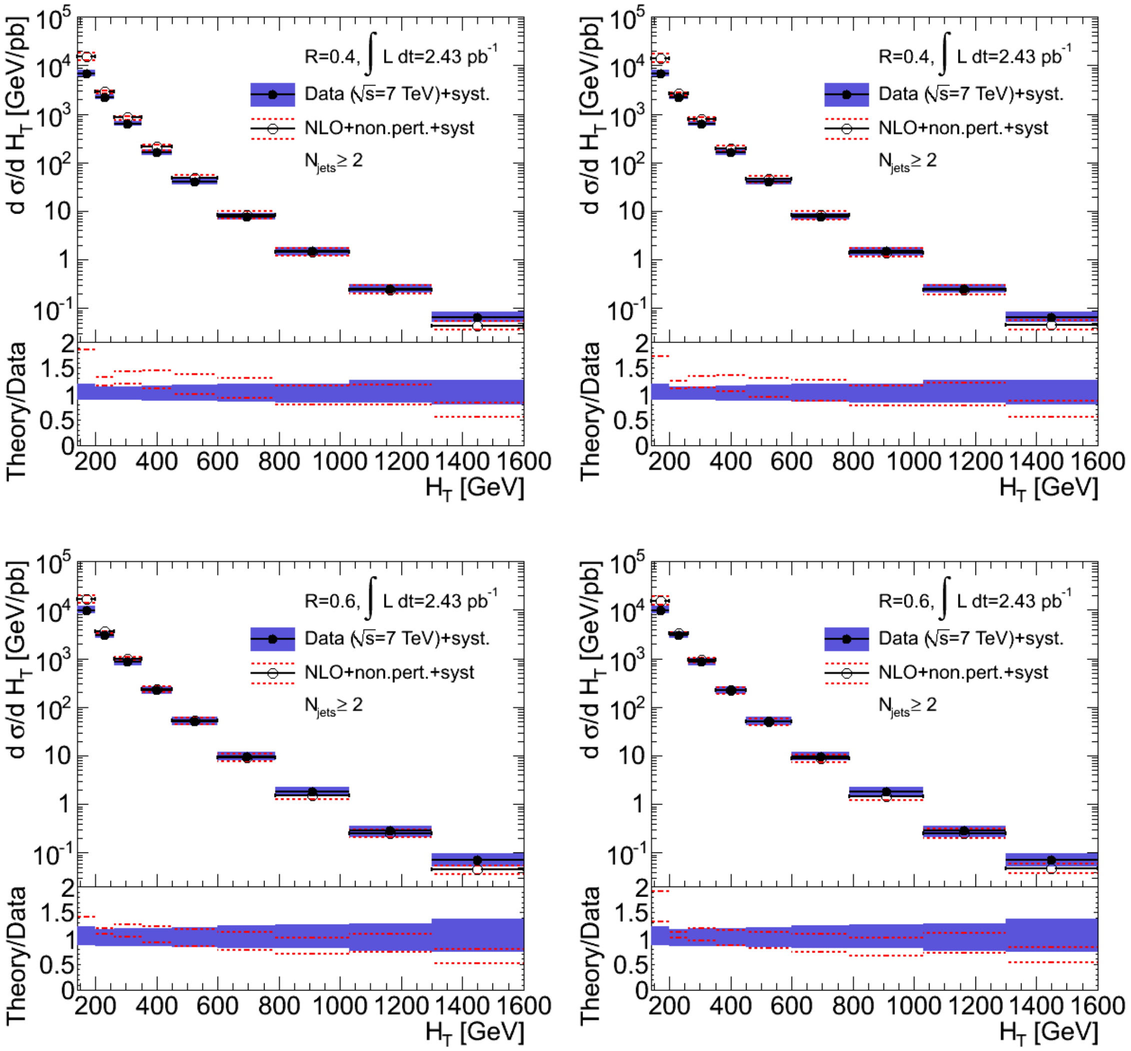

Figure F.15: *Differential cross section as a function of event H_{T} in two-jet events compared to a next-to-leading-order calculation. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ (top) and $R = 0.6$ (bottom) are used in this study. The results are compared to a next-to-leading-order calculation performed with the* MSTW2008nlo *(left) and the* CTEQ6.6 *(right) pdf sets. The error bands correspond to the systematic uncertainties in the measurements. The systematic uncertainty in the NLO calculation is shown as dotted lines above and below the calculation. Plots of the ratios of the theoretical calculations to the data are presented at the bottom of each figure.*

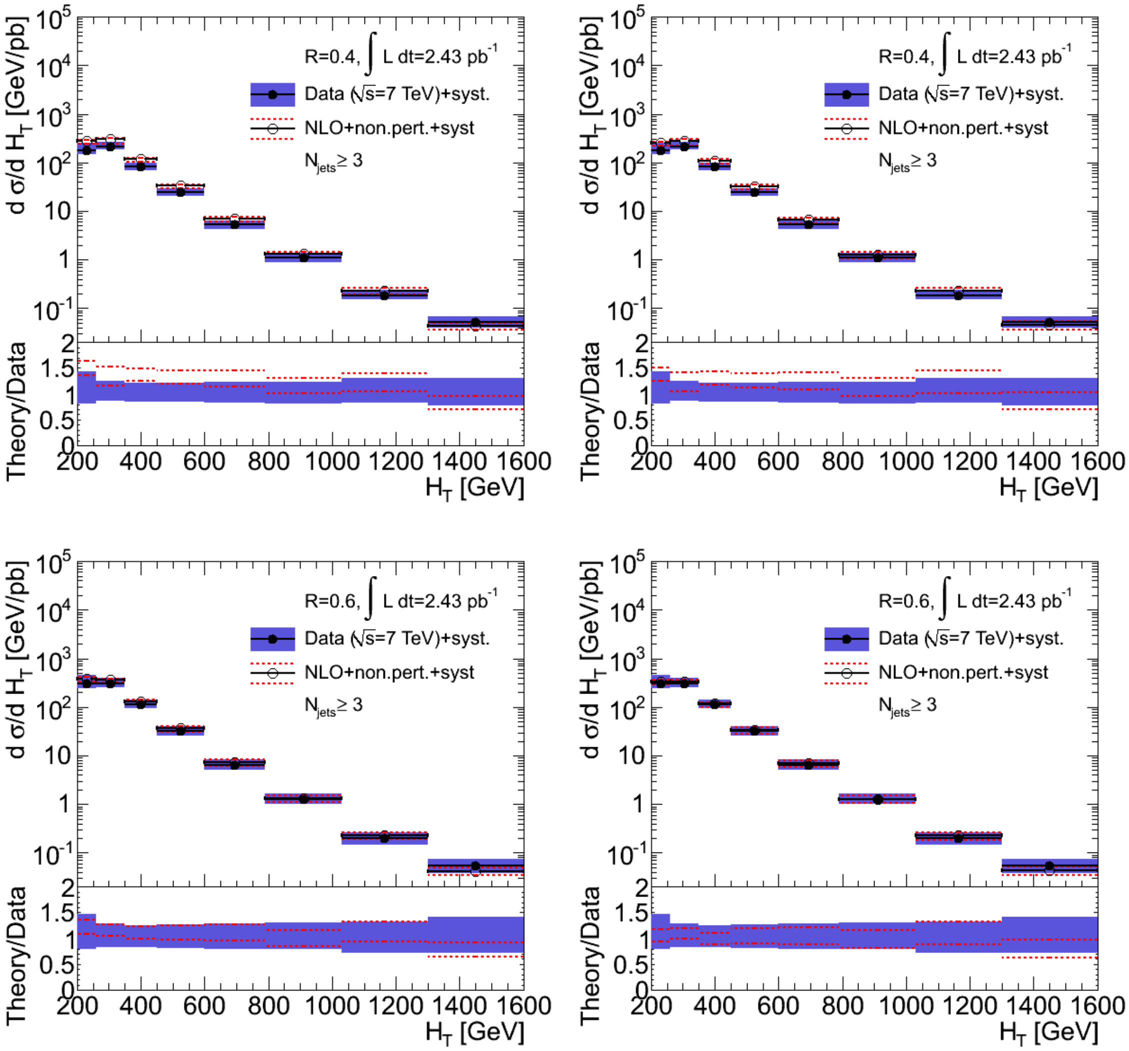

Figure F.16: *Differential cross section as a function of event H_{T} in three-jet events compared to a next-to-leading-order calculation. Jets built using the anti-k_t algorithm with resolution parameter $R = 0.4$ (top) and $R = 0.6$ (bottom) are used in this study. The results are compared to a next-to-leading-order calculation performed with the* MSTW2008nlo *(left) and the* CTEQ6.6 *(right) pdf sets. The error bands correspond to the systematic uncertainties in the measurements. The systematic uncertainty in the next-to-leading-order calculation is shown as dotted lines above and below the calculation. Plots of the ratios of the theoretical calculations to the data are presented at the bottom of each figure.*

F.3 Additional Comparisons of Ratios to Next-to-Leading-Order Calculations

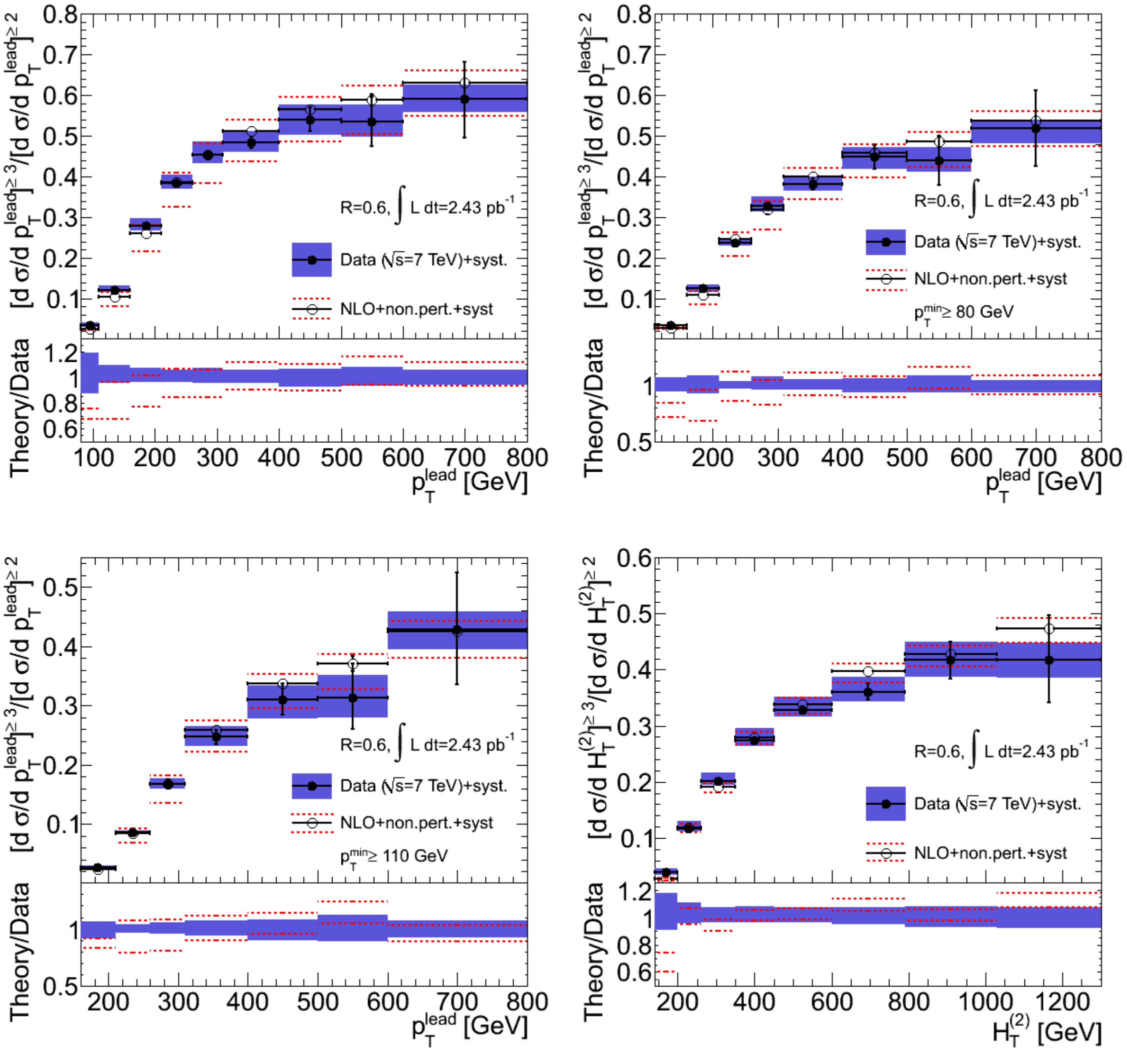

Figure F.17: *Ratio of the three-jet to two-jet differential cross section as a function of the leading jet* $p_{\rm T}$ *(top, bottom left) using three different kinematic cuts on the jet selection and* ${H_{\rm T}}^{(2)}$ *(bottom right). Jets built with the anti-*k_t *algorithm using a resolution parameter* $R = 0.6$ *are used in this study. The results are compared to a next-to-leading-order calculation performed with the* CTEQ6.6 *pdf set. The error bands correspond to the systematic uncertainties in the measurements. The systematic uncertainty in the next-to-leading-order calculation is shown as dotted lines above and below the calculation. Plots of the ratios of the theoretical calculation to the data are presented at the bottom of each figure.*

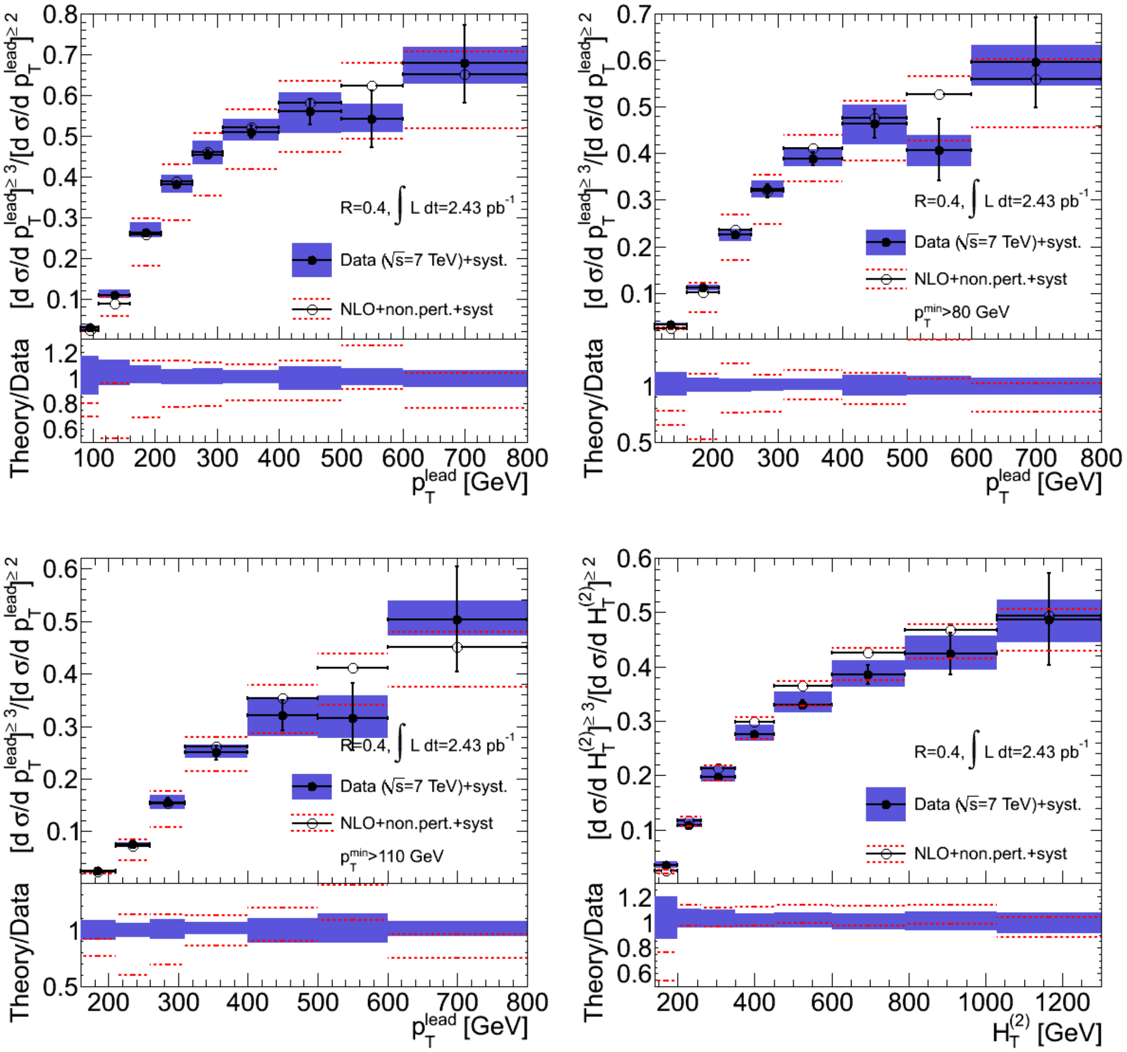

Figure F.18: *Ratio of the three-jet to two-jet differential cross section as a function of the leading jet* p_{T} *(top, bottom left) using three different kinematic cuts on the jet selection and* ${H_{\mathrm{T}}}^{(2)}$ *(bottom right). Jets built with the anti-*k_t *algorithm using a resolution parameter* $R = 0.4$ *are used in this study. The results are compared to a next-to-leading-order calculation performed with the* MSTW2008nlo *pdf set. The error bands correspond to the systematic uncertainties in the measurements. The systematic uncertainty in the next-to-leading-order calculation is shown as dotted lines above and below the calculation. Plots of the ratios of the theoretical calculation to the data are presented at the bottom of each figure.*

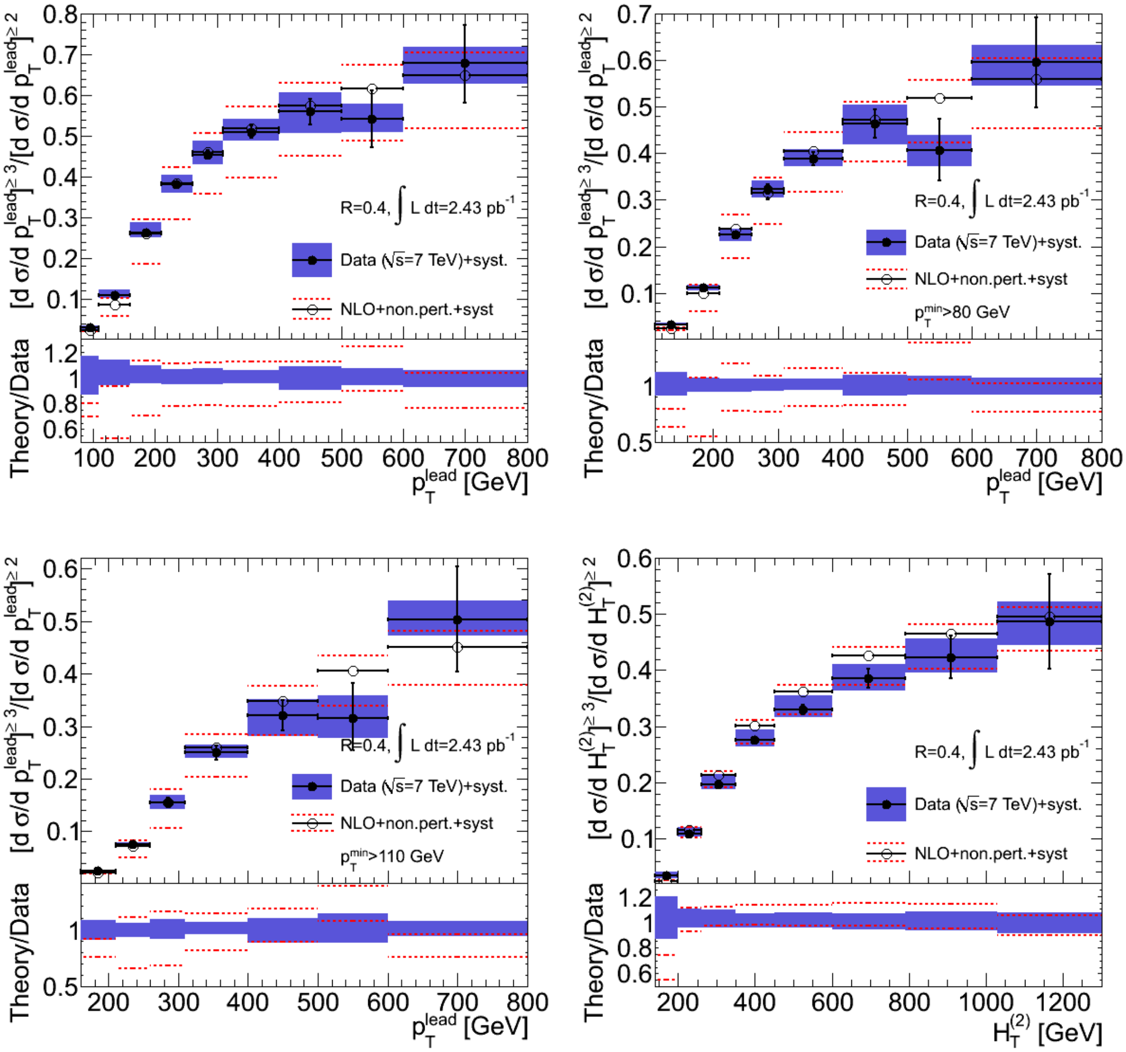

Figure F.19: *Ratio of the three-jet to two-jet differential cross section as a function of the leading jet* p_{T} *(top, bottom left) using three different kinematic cuts on the jet selection and* ${H_{\mathrm{T}}}^{(2)}$ *(bottom right). Jets built with the anti-*k_t *algorithm using a resolution parameter* $R = 0.4$ *are used in this study. The results are compared to a next-to-leading-order calculation performed with the* CTEQ6.6 *pdf set. The error bands correspond to the systematic uncertainties in the measurements. The systematic uncertainty in the next-to-leading-order calculation is shown as dotted lines above and below the calculation. Plots of the ratios of the theoretical calculation to the data are presented at the bottom of each figure.*

Bibliography

[1] E. Anderssen, J.-F. Arguin, B. Atehortua-Escobar, M. Citterio et al., *Results from the Connectivity Test During the Pixel Package Integration.* ATL-COM-INDET 2008-016 (2008).

[2] P. Behara, G. Gaycken, C. Horn, A. Khanov et al., *Threshold Tuning of the ATLAS Pixel Detector.* ATL-INDET-PUB 2010-001 (2010).

[3] D. Lopez Mateos, E. W. Hughes, and A. Salzburger, *A Parameterization of the Energy Loss of Muons in the ATLAS Tracking Geometry.* ATL-MUON-PUB 2008-002 (2008).

[4] D. Lopez Mateos, S. Hassani, E. W. Hughes, J. F. Laporte et al., *A Bayesian Method for Estimating the Energy Loss of Muons in Calorimeters.* ATL-MUON-INT 2009-001 (2008).

[5] E. Lund, L. Bugge, E. W. Hughes, D. Lopez Mateos et al., *Treatment of Energy Loss and Multiple Scattering in the Context of Track Parameter and Covariance Matrix Propagation in Continuous Material in the ATLAS Experiment.* ATL-SOFT-PUB 2008-003 (2008).

[6] E. Lund, L. Bugge, T. G. Cornelissen, M. Elsing et al., *Refitting of Combined Inner Detector and Muon Spectrometer Tracks from Monte Carlo Samples by Using the Kalman Fitter and the STEP Algorithm in the ATLAS Experiment.* ATL-SOFT-PUB 2008-005 (2008).

[7] K. A. Assamagan, K. Bachas, T. Carli, T. Davidek et al., *Muons in the ATLAS Calorimeters: Energy Loss Corrections and Muon Tagging.* ATL-COM-MUON 2008-009 (2008).

[8] G. Aad et al., the **ATLAS** Collaboration, *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*. CERN-OPEN 2008-020, Geneva (2008), [`ISBN: 9789290833215`].

[9] D. Lopez Mateos, E. W. Hughes, and A. Schwartzman, *Jet Energy Scale Correction to Semileptonic b-jets from Missing Neutrino Energy*. ATL-PHYS-INT 2008-047 (2008).

[10] G. Add et al., the **ATLAS** Collaboration, *Measurement of the b-jet Production Cross Section Using Muons in Jets with ATLAS in pp Collisions at* $\sqrt{s} = 7TeV$. ATLAS-CONF 2011-057 (2011).

[11] D. Lopez Mateos, E. W. Hughes, and A. Schwartzman, *A Simple* p_T*- and* η*-Dependent Monte Carlo–Based Jet Calibration*. ATL-PHYS-INT 2009-077 (2009).

[12] G. Aad et al., the **ATLAS** Collaboration, *Measurement of Inclusive Jet and Dijet Cross Sections in Proton-Proton Collisions at 7 TeV Centre-of-mass Energy with the ATLAS Detector*. Eur. Phys. Jour. C **71**, (2010) 1–59, [`arXiv:1009.5908`], URL `http://dx.doi.org/10.1140/epjc/s10052-010-1512-2`.

[13] D. Lopez Mateos, E. W. Hughes, and A. Schwartzman, *Jet Energy Resolution Improvement After Calibration Using Longitudinal Calorimeter Segmentation in ATLAS*. ATL-PHYS-INT 2009-051 (2009).

[14] D. Lopez Mateos, E. W. Hughes, and A. Schwartzman, *A Sequential Multi-Variate Jet Calibration Based On Global Properties of the Jet Structure*. ATL-PHYS-INT 2011-023 (2010).

[15] D. Lopez Mateos, E. W. Hughes, and A. Schwartzman, *Reduction of the Flavor Dependence of the Jet Energy Scale Using a Correction that Depends on the Transverse Structure of the Jet*. ATL-PHYS-INT 2010-046 (2010).

[16] G. Aad et al., the **ATLAS** Collaboration, *Validation of the Jet Global Sequential Calibration Using Proton-Proton Collisions at* $\sqrt{s} = 7$ *TeV and Evaluation of its Associated Systematic Uncertainty*. ATLAS-COM-CONF 2011-080 (2011).

[17] L. Asquith, B. Brelier, J. M. Butterworth, M. Campanelli et al., *Performance of Jet Algorithms in the ATLAS Detector*. ATL-PHYS-INT 2010-129 (2010).

[18] G. Aad et al., the **ATLAS** Collaboration, *Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at* $\sqrt{s} = 900$ *GeV.* ATLAS-CONF 2010-016 (2010).

[19] G. Aad et al., the **ATLAS** Collaboration, *Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at* $\sqrt{s} = 7$ *TeV.* ATLAS-CONF 2010-053 (2010).

[20] G. Aad et al., the **ATLAS** Collaboration, *Jet Energy Scale and its Systematic Uncertainty for Jets Produced in Proton-Proton Collisions at* $\sqrt{s} = 7$ *TeV and Measured with the ATLAS Detector.* ATLAS-CONF 2010-056 (2010).

[21] J. F. Arguin, D. Boumediene, M. Bondioli, G. Cortiana et al., *Jet Selection for Top Physics.* ATL-PHYS-INT 2010-134 (2010).

[22] S. Cabrera Urban, M. J. Costa, N. Ghodbane, I. M. Nugent et al., *Jets, Missing Transverse Energy and Taus for Top Physics Analyses in Release 16 with the 2010 Dataset.* ATL-COM-PHYS 2011-132 (2011).

[23] D. Lopez Mateos, *Inclusive Measurements as an mSUGRA Signal with ATLAS.* In *Proceedings of the International School of Subnuclear Physics, Erice, Sicily, Italy, 29 August–7 September 2006*, 575 (World Scientific Publishing Co.) (2008), [`ISBN: 978-981-283-245-0`].

[24] H. M. Gray, E. W. Hughes, D. Lopez Mateos, Z. Marshall et al., *Aspects of an Inclusive Search for Supersymmetry in the 0-lepton Channel at* $\sqrt{s} = 10$ *TeV.* ATL-COM-PHYS 2009-622 (2009).

[25] G. Aad et al., the **ATLAS** Collaboration, *Measurements of Multijet Production Cross Sections in Proton-Proton Collisions at 7 TeV Center-of-mass Energy with the ATLAS Detector.* ATLAS-CONF 2010-084 (2010).

[26] G. Aad et al., the **ATLAS** Collaboration, *Measurement of Multi-jet Cross Sections in Proton-Proton Collisions at 7 TeV Center-of-mass Energy.* ATLAS-CONF 2011-043 (2011).

[27] G. Aad et al., the **ATLAS** Collaboration, *Light-quark and Gluon Jets in ATLAS: Calorimeter Response, Jet Energy Scale Systematics, and Sample Characterization.* ATLAS-CONF 2011-053 (2011).

[28] G. Aad et al., the **ATLAS** Collaboration, *Measurement of Multi-jet Cross Sections in Proton-Proton Collisions at 7 TeV Center-of-mass Energy with the ATLAS Detector* (in preparation).

[29] E. Gildener, *Gauge-symmetry Hierarchies.* Phys. Rev. D **14**(6), (1976) 1667–1672, [`doi: 10.1103/PhysRevD.14.1667`].

[30] S. Weinberg, *Gauge Hierarchies.* Physics Letters B **82**(3-4), (1979) 387–391, ISSN 0370-2693, [`doi: 10.1016/0370-2693(79)90248-X`].

[31] C. Amsler et al., the **Particle Data Group** Collaboration. Physics Letters **B667**, (2008) 1.

[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, [`doi: 10.1140/epjc/s10052-009-1072-5`], [`arXiv:0901.0002`], figures from the MSTW Website.

[33] M. Gell-Mann, *Symmetries of Baryons and Mesons.* Phys. Rev. **125**(3), (1962) 1067–1084, [`doi: 10.1103/PhysRev.125.1067`].

[34] H. Yukawa, *On the Interaction of Elementary Particles.* Proc. Phys. Math. Soc. Jap. **17**, (1935) 48–57.

[35] M. Gell-Mann, *A Schematic Model of Baryons and Mesons.* Phys. Lett. **8**, (1964) 214–215, [`doi: 10.1016/S0031-9163(64)92001-3`].

[36] G. Zweig, *An SU(3) Model for Strong Interaction Symmetry and its Breaking.* CERN-TH 401 (1964).

[37] G. Zweig, *An SU(3) Model for Strong Interaction Symmetry and its Breaking 2.* CERN-TH 412 (1964).

[38] M. Gell-Mann and Y. Ne'eman, *The Eightfold Way* (W.A. Benjamin, New York) (1964), [`ISBN: 9780738202990`].

[39] E. D. Bloom et al., *High-Energy Inelastic $e-p$ Scattering at 6° and 10°*. Phys. Rev. Lett. **23**(16), (1969) 930–934, [`doi: 10.1103/PhysRevLett.23.930`].

[40] I. H. Park et al., *Experimental Evidence for the Non-Abelian Nature of QCD from a Study of Multijet Events Produced in e^+e^- Annihilation.* Phys. Rev. Lett. **62**(15), (1989) 1713–1716, [`doi: 10.1103/PhysRevLett.62.1713`].

[41] S. Chekanov et al., *Dijet Photoproduction at HERA and the Structure of the Photon.* Eur. Phys. Jour. C **23**, (2002) 615–631, [`doi: 10.1007/s100520200936`].

[42] G. Arnison et al., *Angular Distributions and Structure Functions from Two-jet Events at the CERN SPS Collider.* Phys. Lett. B **136**(4), [`doi: 10.1016/0370-2693(84)91164-X`].

[43] F. Abe et al., *Measurement of the Inclusive Jet Cross Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV.* Phys. Rev. Lett. **62**(6), (1989) 613–616, [`doi: 10.1103/PhysRevLett.62.613`].

[44] R. P. Feynman, *Photon Hadron Interactions* (W.A. Benjamin, New York) (1972), [`ISBN: 9780201360745`].

[45] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics* (Cambridge University Press, Cambridge) (1996), [`ISBN: 9780521545891`], [`doi: 10.2277/0521545897`].

[46] W. Bietenholz et al., *Exploring the Nucleon Structure from First Principles of QCD.* In *Proceedings of the XXXIII Symposium on Nuclear Physics, Cocoyoc (Morelos, Mexico), January 5-8, 2010*, edited by Libertad Barrón-Palos, Roelof Bijker, Ruben Fossion, and David Lizcano, 012011 (IOP Journal of Physics: Conference Series) (2010), [`doi: 10.1088/1742-6596/239/1/012011`], [`arXiv:1004.2100`].

[47] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, *An Effective Field Theory for Collinear and Soft Gluons: Heavy to Light Decays.* Phys. Rev. D **63**(11), (2001) 114020, [`doi: 10.1103/PhysRevD.63.114020`].

[48] O. W. Greenberg, *Spin and Unitarity-Spin Independence in a Paraquark Model of Baryons and Mesons*. Phys. Rev. Lett. **13**, (1964) 598–602, [`doi: 10.1103/PhysRevLett.13.598`].

[49] M. Y. Han and Y. Nambu, *Three-Triplet Model with Double $SU(3)$ Symmetry*. Phys. Rev. **139**, (1965) B1006–B1010, [`doi: 10.1103/PhysRev.139.B1006`].

[50] J. I. Friedman, *Deep Inelastic Scattering: Comparisons with the Quark Model*. Rev. Mod. Phys. **63**(3), (1991) 615–627, [`doi: 10.1103/RevModPhys.63.615`].

[51] H. W. Kendall., *Deep Inelastic Scattering: Experiments on the Proton and the Observation of Scaling*. Rev. Mod. Phys. **63**(3), (1991) 597–614, [`doi: 10.1103/RevModPhys.63.597`].

[52] R. E. Taylor, *Deep Inelastic Scattering: The Early Years*. Rev. Mod. Phys. **63**(3), (1991) 573–595, [`doi: 10.1103/RevModPhys.63.573`].

[53] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions*. Phys. Rev. Lett. **30**, (1973) 1346–1349, [`doi: 10.1103/PhysRevLett.30.1346`].

[54] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*. Phys. Rev. Lett. **30**, (1973) 1343–1346, [`doi: 10.1103/PhysRevLett.30.1343`].

[55] C. F. Berger et al., *Precise Predictions for $W+3$ Jet Production at Hadron Colliders*. Phys. Rev. Lett. **102**(22), (2009) 222001, [`doi: 10.1103/PhysRevLett.102.222001`].

[56] F. Bissey, A. I. Signal, and D. B. Leinweber, *Comparison of Gluon Flux-tube Distributions for Quark- Diquark and Quark-Antiquark Hadrons*. Phys. Rev. **D80**, (2009) 114506, [`doi: 10.1103/PhysRevD.80.114506`], [`arXiv:0910.0958`].

[57] T. Sjöstrand, *A Model for Initial State Parton Showers*. Phys. Lett. B **157**(4), (1985) 321 – 325, [`doi: 10.1016/0370-2693(85)90674-4`].

[58] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics* (Cambridge University Press, Cambridge) (1996), [`ISBN: 9780521545891`], [`doi: 10.2277/0521545897`].

[59] V. V. Sudakov. Sov. Phys. J.E.T.P **30**, (1956) 65.

[60] D. Amati et al., *A Treatment of Hard Processes Sensitive to the Infrared Structure of QCD*. Nucl. Phys. **B173**, (1980) 429 – 455, [`doi: 10.1016/0550-3213(80)90012-7`].

[61] P. Z. Skands, *The Perugia Tunes*. MCnet 10/08 (2010), [`arXiv:1005.3457`].

[62] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics* (Cambridge University Press, Cambridge) (1996), [`ISBN: 9780521545891`], [`doi: 10.2277/0521545897`].

[63] B. Andersson et al., *Parton Fragmentation and String Dynamics*. Phys. Rep. **97**, (1983) 31, [`doi: 10.1016/0550-3213(86)90471-2`].

[64] G. Gustafson B. Andersson and B. Söderberg, *A Probability Measure on Parton and String States*. Nucl. Phys. B **264**, (1986) 29.

[65] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*. Nucl. Phys. **B238**, (1984) 492–528, [`doi: 10.1016/0550-3213(84)90333-X`].

[66] F. Krauss J.-C. Winter and G. Soff, *A Modified Cluster-Hadronization Model*. Eur. Phys. J. **C36**, (2004) 381–395, [`doi: 10.1140/epjc/s2004-01960-8`], [`hep-ph/0311085`].

[67] M. L. Mangano, M. Moretti, F. Piccini, and M. Treccani, *Matching Matrix Elements and Shower Evolution for Top-Pair Production in Hadronic Collisions*. JHEP **0701**, (2007) 013, [`doi: 10.1088/1126-6708/2007/01/013`], [`hep-ph/0611129`].

[68] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD Matrix Elements + Parton Showers*. JHEP **0111**, (2001) 063, [`doi: 10.1088/1126-6708/2001/11/063`], [`hep-ph/0109231`].

[69] S. D. Ellis and D. Soper, *Successive Combination Jet Algorithm for Hadron Collisions*. Phys. Rev. **D48**, (1993) 3160, [`doi: 10.1103/PhysRevD.48.3160`], [`hep-ph/9305266`].

[70] R. P. Feynman, *Very High-Energy Collisions of Hadrons*. Phys. Rev. Lett. **23**(24), (1969) 1415–1417, [`doi: 10.1103/PhysRevLett.23.1415`].

[71] J. D. Bjorken, *Asymptotic Sum Rules at Infinite Momentum.* Phys. Rev. **179**, (1969) 1547–1553, [`doi: 10.1103/PhysRev.179.1547`].

[72] V. N. Gribov and L. N. Lipatov, *Deep Inelastic e-p Scattering in Perturbation Theory.* Sov. J. Nucl. Phys. **15**, (1972) 438–450.

[73] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language.* Nucl. Phys. **B126**, (1977) 298, [`doi: 10.1016/0550-3213(77)90384-4`].

[74] 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.

[75] H. Georgi and H. D. Politzer, *Electroproduction Scaling in an Asymptotically Free Theory of Strong Interactions.* Phys. Rev. D **9**(2), (1974) 416–420, [`doi: 10.1103/PhysRevD.9.416`].

[76] D. J. Gross and F. Wilczek, *Asymptotically Free Gauge Theories II.* Phys. Rev. D **9**(4), (1974) 980–993, [`doi: 10.1103/PhysRevD.9.980`].

[77] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton Interactions in Photoproduction at HERA.* Zeit. f. Phys. **C72**, (1996) 637–646, [`doi: 10.1007/s002880050286`], [`hep-ph/9601371`].

[78] C. M. Buttar et al., *The Underlying Event.* In *HERA and the LHC: A workshop on the implications of HERA for LHC physics*, edited by A. de Roeck and H. Jung (DESY-PROC-2005-001, DESY, Germany) (2005), [`ISBN: 9789290832652`], [`hep-ph/0510198v1`].

[79] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 Physics and Manual.* JHEP **05**, (2006) 026, [`doi: 10.1088/1126-6708/2006/05/026`], [`hep-ph/0603175`].

[80] T. Gleisberg et al., *SHERPA 1.α, A Proof-of-Concept Version.* JHEP **02**, (2004) 056, [`doi: 10.1088/1126-6708/2004/02/056`], [`hep-ph/0311263`], figures from the Sherpa Website.

[81] O. S. Bruening et al. (eds.) *LHC Design Report*, volume v.1: The LHC Main Ring (CERN, Geneva) (2004), [`ISBN: 9789290832249`], also CERN-2004-003-V-1.

[82] O. S. Bruening et al. (eds.) *LHC Design Report*, volume v.2: The LHC Infrastructure and General Services (CERN, Geneva) (2004), [`ISBN: 9789290832249`], also CERN-2004-003-V-2.

[83] M. Benedikt et al. (eds.) *LHC Design Report*, volume v.3: The LHC Injector Chain (CERN, Geneva) (2004), [`ISBN: 9789290832393`], also CERN-2004-003-V-3.

[84] L. Evans and P. Bryant, *LHC Machine*. JINST **3**, (2008) S08001, URL `http://stacks.iop.org/1748-0221/3/i=08/a=S08001`, figures from the LHC Website.

[85] M. Bajko, F. Bertinelli, N. Catalan-Lasheras, S. Claudet et al., *Report of the Task Force on the Incident of 19th September 2008 at the LHC*. CERN-LHC-PROJECT-Report 1168, Geneva.

[86] D. Belohrad et al., *Commissioning and First Performance of the LHC Beam Current Measurement Systems*. In *Proceedings of the 1st International Particle Accelerator Conference, Kyoto, Japan, 23–28 May 2010*, 1110–1112 (2010), [`ISBN: 9789290833529`].

[87] S. van der Meer, *Calibration of the Effective Beam Height in the ISR*. CERN-ISR-PO 68-31 (1968).

[88] G. Aad et al., the **ATLAS** Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*. JINST **3**, (2008) S08003, [`doi: 10.1088/1748-0221/3/08/S08003`], figures also from the ATLAS website.

[89] The **ATLAS** Collaboration, *ATLAS Inner Detector Technical Design Report*. CERN/LHCC 97-016 (1997).

[90] The **ATLAS** Collaboration, *ATLAS Central Solenoid: Technical Design Report*. CERN/LHCC 97-021 (1997).

[91] A. Yamamoto, Y. Makida, R. Ruber, Y. Doi et al., *The ATLAS Central Solenoid*. Nucl. Instr. Meth. A **584**(1), (2008) 53–74, ISSN 0168-9002, [`doi:`

`10.1016/j.nima.2007.09.047`], URL `http://www.sciencedirect.com/science/article/B6TJM-4PVPVHG-2/2/7bef321f21f1d2784dbd7f62b7a025cc`.

[92] The **ATLAS** Collaboration, *ATLAS Liquid-Argon Calorimeter: Technical Design Report.* CERN/LHCC 96-041 (1996).

[93] The **ATLAS** Collaboration, *ATLAS Tile Calorimeter: Technical Design Report.* CERN/LHCC 96-042 (1996).

[94] The **ATLAS** Collaboration, *ATLAS Muon Spectrometer: Technical Design Report.* CERN/LHCC 97-022 (1997).

[95] The **ATLAS** Collaboration, *ATLAS Barrel Toroid: Technical Design Report.* CERN/LHCC 97-019 (1997).

[96] The **ATLAS** Collaboration, *ATLAS End-cap Toroids: Technical Design Report.* CERN/LHCC 97-020 (1997).

[97] G. Aad et al., the **ATLAS** Collaboration, *Charged-Particle Multiplicities in pp Interactions Measured with the ATLAS Detector at the LHC* [`arXiv:1012.5104`], (accepted by New J. Phys.).

[98] D. A. Dobos, C. Goessling, and R. Klingenberg, *Commissioning Perspectives for the ATLAS Pixel Detector.* CERN-THESIS 2008-022 (2007).

[99] G. Aad et al., the **ATLAS** Collaboration, *Performance of the ATLAS Inner Detector Track Reconstruction Algorithms.* ATLAS-COM-PHYS 2010-848 (2010).

[100] F. Bergsma, *Calibration of Hall Sensors in Three Dimensions.* In *Proceedings of the 13th International Magnetic Measurement Workshop, Stanford, USA* (2003), URL `http://cdsweb.cern.ch/record/1072471`.

[101] M. Aleksa et al., *Measurement of the ATLAS Solenoid Magnetic Field.* J. Instrum. **3**(04), (2008) P04003, URL `http://stacks.iop.org/1748-0221/3/i=04/a=P04003`.

[102] The **ATLAS** Collaboration, *ATLAS Pixel Detector Technical Design Report.* CERN/LHCC 98-013 (1998).

[103] G. Aad et al., the **ATLAS** Collaboration, *The ATLAS Inner Detector Commissioning and Calibration*. Eur. Phys. J. **C70**, (2010) 787–821, [`doi: 10.1140/epjc/s10052-010-1366-7`], [`arXiv:1004.5293`].

[104] A. Ahmad et al., *The Silicon Microstrip Sensors of the ATLAS Semiconductor Tracker*. Nucl. Instr. Meth. A **578**(1), (2007) 98–118, ISSN 0168-9002, [`doi: 10.1016/j.nima.2007.04.157`], URL `http://www.sciencedirect.com/science/article/B6TJM-4NMWPRH-1/2/876aef58f2c94cc682482449f54c560f`.

[105] E. Abat et al., *The ATLAS TRT Barrel Detector*. J. Instrum. **3**(02), (2008) P02014, URL `http://stacks.iop.org/1748-0221/3/i=02/a=P02014`.

[106] E. Abat et al., *The ATLAS TRT End-cap Detector*. J. Instrum. **3**(10), (2008) P10003, URL `http://stacks.iop.org/1748-0221/3/i=10/a=P10003`.

[107] R. Wigmans, *Calorimetry: Energy Measurement in Particle Physics*. International series of monographs on physics (Clarendon Press, Oxford) (2000).

[108] A. Artamonov et al., *The ATLAS Forward Calorimeter*. J. Instrum. **3**(02), (2008) P02010, URL `http://stacks.iop.org/1748-0221/3/i=02/a=P02010`.

[109] G. Aad et al., the **ATLAS** Collaboration, *Readiness of the ATLAS Liquid Argon Calorimeter for LHC Collisions*. Technical report (2010), [`doi: 10.1140/epjc/s10052-010-1354-y`], [`arXiv:0912.2642`].

[110] W. E. Cleland and E. G. Stern, *Signal Processing Considerations for Liquid Ionization Calorimeters in a High Rate Environment*. Nucl. Instr. Meth. A **338**(2-3), (1994) 467–497, ISSN 0168-9002, [`doi: 10.1016/0168-9002(94)91332-3`], URL `http://www.sciencedirect.com/science/article/B6TJM-473FMY7-1K3/2/4f5609319aef05d05bde624f68a3bcae`.

[111] J. C. Barriere, F. Bauer, M. Fontaine, A. Formica et al., *The Alignment System of the Barrel Part of the ATLAS Muon Spectrometer*. ATL-MUON-PUB 2008-007 (2008).

[112] S. Aefsky, C. Amelung, J. Bensinger, C. Blocker et al., *The Optical Alignment System of the ATLAS Muon Spectrometer Endcaps*. ATL-MUON-PUB 2008-003 (2008).

[113] H. Boterenbrood et al., *The Read-out Driver for the ATLAS MDT Muon Precision Chambers*. IEEE Trans. Nucl. Sci. **53**, (2006) 741.

[114] J. Dailing, N. Drego, A. Gordeev, V. Gratchev et al., *Off-detector Electronics for a High-Rate CSC Detector*. In *6th Workshop on Electronic for LHC Experiments, Cracow, Poland* (2000), URL `http://cdsweb.cern.ch/record/479707`.

[115] G. Aad et al., the **ATLAS** Collaboration, *Luminosity Determination in pp Collisions at sqrt(s)=7 TeV Using the ATLAS Detector at the LHC* [`arXiv:1101.2185`], accepted by Eur. Phys. Jour. C.

[116] V. Cindro, D. Dobos, I. Dolenc, H. Frais-Kölbl et al., *The ATLAS Beam Conditions Monitor*. J. Instrum. **3**, (2008) P02004.

[117] P. Jenni, M. Nordberg, M. Nessi, and K. Jon-And, *ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity*. CERN-LHCC 2008-004 (2008).

[118] G. Aad et al., the **ATLAS** Collaboration, *Performance of the Minimum Bias Trigger in p-p Collisions at* $\sqrt{s} = 900\ GeV$. ATLAS-CONF 2010-025 (2010).

[119] The **ATLAS** Collaboration, *ATLAS Level-1 Trigger: Technical Design Report*. CERN/LHCC 98-014 (1998).

[120] M. Grothe et al., *Architecture of the ATLAS High Level Trigger Event Selection Software*. Nucl. Instrum. Methods Phys. Res. **A 518**, (2003) 537–541, proceedings of Frontier Detectors for Frontier Physics, La Biodola, Isola d'Elba, Italy, 25–31 May 2003.

[121] P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS High-Level Trigger, Data-Acquisition and Controls: Technical Design Report*. CERN-LHCC 2003-022 (2003).

[122] S. Ask et al., *The ATLAS Central Level-1 Trigger Logic and TTC System*. J. Instrum. **3**, (2008) P08002, [`doi: 10.1088/1748-0221/3/08/P08002`].

[123] S. Agostinelli et al., *Geant Simulation*. Nucl. Instr. and Meth. A **506**, (2003) 250.

[124] G. Aad et al., the **ATLAS** Collaboration, *The ATLAS Monte Carlo Project.* ATLAS-SOFT 2010-01-004 (2010), accepted for publication in Eur. Phys. J. C, [`arXiv:1005.4568`].

[125] P. Adragna et al., *Measurement of Pion and Proton Longitudinal Shower Profiles up to 20 Nuclear Interaction Lengths with the ATLAS Tile Calorimeter.* Nucl. Instr. and Meth. A **615**, (2010) 158–181, [`doi: 10.1016/j.nima.2010.01.037`].

[126] P. Adragna et al., *Testbeam Studies of Production Modules of the ATLAS Tile Calorimeter.* Nucl. Instr. and Meth. A **606**, (2009) 362–394, [`doi: 10.1016/j.nima.2009.04.009`].

[127] E. Abat et al., *Response and Shower Topology of 2 to 180 GeV Pions Measured with the ATLAS Barrel Calorimeter at the CERN Test-beam and Comparison to Monte Carlo Simulations.* ATL-CAL-PUB 2010-001 (2010).

[128] G. Aad et al., the **ATLAS** Collaboration, *Mapping the Material in the ATLAS Inner Detector using Secondary Hadronic Interactions in 7 TeV Collisions.* ATLAS-CONF 2010-058 (2010).

[129] G. Aad et al., the **ATLAS** Collaboration, *Photon Conversions at* $\sqrt{s} = 900$ *GeV Measured with the ATLAS Detector.* ATLAS-CONF 2010-007 (2010).

[130] G. Aad et al., the **ATLAS** Collaboration, *Performance of the Missing Transverse Energy Reconstruction in Minimum Bias Events at* $\sqrt{s}$ *of 900 GeV and 2.36 TeV with the ATLAS Detector.* ATLAS-CONF 2010-008 (2010).

[131] G. Aad et al., the **ATLAS** Collaboration, *Response of the ATLAS Calorimeters to Single Isolated Hadrons Produced in Proton Proton Collisions at a Center of Mass Energy of* $\sqrt{s} = 900$ *GeV.* ATLAS-CONF 2010-017 (2010).

[132] G. Aad et al., the **ATLAS** Collaboration, *Study of the Material Budget in the ATLAS Inner Detector with* K^0_S *decays in collision data at* $\sqrt{s} = 900$ *GeV.* ATLAS-CONF 2010-019 (2010).

[133] G. Aad et al., the **ATLAS** Collaboration, *Performance of the New Tracking Silicon Pattern Recognition Algorithm in Data and Simulation at* $\sqrt{s} = 7$ *TeV*. ATLAS-CONF 2010-072 (2010).

[134] R. Fruhwirth, *Application of Kalman Filtering to Track and Vertex Fitting*. Nucl. Instr. Meth. A **262**(2-3), (1987) 444–450, ISSN 0168-9002, [`doi: 10.1016/0168-9002(87)90887-4`].

[135] J. Biesiada, A. Gaponenko, I. Hinchliffe, M. Shapiro et al., *Properties of Jets from Tracks in Proton-Proton Collisions at Center-of-mass Energy* $\sqrt{s}$= *7 TeV with the ATLAS Detector*. ATL-COM-PHYS 2011-063 (2011).

[136] G. Aad et al., the **ATLAS** Collaboration, *Performance of Primary Vertex Reconstruction in Proton-Proton Collisions at* $\sqrt{s} = 7$ *TeV in the ATLAS Experiment*. ATLAS-CONF 2010-069 (2010).

[137] G. Aad et al., the **ATLAS** Collaboration, *Characterization of Interaction-Point Beam Parameters Using the pp Event-Vertex Distribution Reconstructed in the ATLAS Detector at the LHC*. ATLAS-CONF 2010-027 (2010).

[138] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-*k_t *Jet Clustering Algorithm*. JHEP **0804**, (2008) 063, [`doi: 10.1088/1126-6708/2008/04/063`], [`arXiv:0802.1189`].

[139] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, *Better Jet Clustering Algorithms*. JHEP **9708**, (1997) 001, [`doi: 10.1088/1126-6708/1997/08/001`], [`hep-ph/9707323`].

[140] M. Wobisch and T. Wengler, *Hadronization Corrections to Jet Cross Sections in Deep-Inelastic Scattering*. In *Proceedings of the Workshop on Monte Carlo Generators for HERA Physics*, edited by A.T. Doyle, G. Grindhammer, G. Ingelman, and H. Jung (DESY-PROC-1999-02, DESY, Germany) (1999), [`hep-ph/9907280`].

[141] S. Catani, Y. L. Dokshitzer, and B. R. Webber, *The* $k_\perp$*-Clustering Algorithm for Jets in Deep Inelastic Scattering and Hadron Collisions*. Phys. Lett. **B285**, (1992) 291–299, [`doi: 10.1016/0370-2693(92)91467-N`].

[142] S. Catani, Y. L. Dokshitzer, B. R. Webber, and M. H. Seymour, *Longitudinally-Invariant $k_\perp$-Clustering Algorithms for Hadron-Hadron Collisions*. Nucl. Phys. **B406**, (1993) 187–224, [`doi: 10.1016/0550-3213(93)90166-M`].

[143] J. Pinfold et al., *Performance of the ATLAS Liquid Argon Endcap Calorimeter in the Pseudorapidity Region* $2.5 < |\eta| < 4.0$ *in Beam Tests*. Nucl. Inst. and Meth. A **593**(3), (2008) 324 – 342, ISSN 0168-9002, [`doi: 10.1016/j.nima.2008.05.033`].

[144] G. Aad et al., the **ATLAS Collaboration** Collaboration, *Readiness of the ATLAS Tile Calorimeter for LHC Collisions*. Eur. Phys. J. **C70**, (2010) 1193–1236, [`doi: 10.1140/epjc/s10052-010-1508-y`], [`arXiv:1007.5423`].

[145] G. Aad et al., the **ATLAS** Collaboration, *In-situ Jet Energy Scale and Jet Shape Corrections for Multiple Interactions in the First ATLAS Data at the LHC*. ATLAS-CONF 2011-030 (2011).

[146] G. Aad et al., the **ATLAS** Collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 7$ *TeV in ATLAS 2010 data*. ATLAS-CONF 2011-032 (2011).

[147] G. Aad et al., the **ATLAS** Collaboration, *Study of Jet Shapes in Inclusive Jet Production in pp Collisions at* $\sqrt{s} = 7$ *TeV using the ATLAS Detector* [`arXiv:1101.0070`], (accepted by Phys. Rev. D).

[148] J. Adelman et al., *ATLAS Offline Data Quality Monitoring*. In *Proceedings of the 17th International Conference on Computing in High Energy and Nuclear Physics, Prague, Czech Republic, 21–27 March 2009*, 042018 (IOP Journal of Physics: Conference Series) (2010), [`doi: 10.1088/1742-6596/219/4/042018`].

[149] G. Aad et al., the **ATLAS** Collaboration, *Updated Luminosity Determination in pp Collisions at* $\sqrt{s} = 7$ *TeV Using the ATLAS detector*. ATLAS-CONF 2011-011 (2011).

[150] G. Aad et al., the **ATLAS** Collaboration, *Performance of the ATLAS jet trigger with p-p collisions at* $\sqrt{s}$ *=900 GeV*. ATLAS-CONF 2010-028 (2010).

[151] G. Aad et al., the **ATLAS** Collaboration, *Jet Energy Scale and its Systematic Uncertainty in Proton-Proton Collisions at* $\sqrt{s}$*=7 TeV in ATLAS 2010 Data*. ATLAS-CONF 2011-032 (2011).

[152] G. Aad et al., the **ATLAS** Collaboration, *Measurement of Inclusive jet and Dijet Cross Sections in Proton-Proton Collision Data at 7 TeV Centre-of-mass Energy Using the ATLAS Detector*. ATLAS-CONF 2011-047 (2011).

[153] S. Frixione and G. Ridolfi, *Jet Photoproduction at HERA*. Nucl. Phys. **B507**, (1997) 315.

[154] G. Aad et al., the **ATLAS** Collaboration, *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 of* $\sqrt{s}$ *=7 TeV*. ATL-CONF 2010-038 (2010).

[155] D. W. Miller, A. Schwartzman, and D. Su, *Jet-Vertex Association Algorithm*. ATL-COM-PHYS 2008-008 (2008).

[156] D. W. Miller, A. Schwartzman, and D. Su, *Pile-up Jet Energy Scale Corrections Using the Jet-Vertex Fraction Method*. ATL-PHYS-INT 2009-090 (2009).

[157] A. Kolmogorov, *Sulla Determinazione Empirica di una Legge di Distribuzione*. G. Inst. Ital. Attuari **4**, (1933) 83–91.

[158] G. D. Cowan, *Statistical Data Analysis* (Oxford Univ. Press, Oxford) (1998).

[159] G. Aad et al., the **ATLAS** Collaboration, *Charged Particle Multiplicities in pp Interactions at* $\sqrt{s}$*=0.9 and 7 TeV in Diffractive Limited Phase Space Measured with the ATLAS Detector at the LHC and a New* PYTHIA*6 Tune*. ATLAS-CONF 2010-031 (2010).

[160] M. L. Mangano et al., *ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions*. JHEP **07**, (2003) 001, [`doi: 10.1088/1126-6708/2003/07/001`], [`hep-ph/0206293`].

[161] G. Corcella et al., *HERWIG 6: An Event Generator for Hadron Emission Reactions with Interfering Gluons (Including Supersymmetric Processes)*. JHEP **01**, (2001) 010, [`doi: 10.1088/1126-6708/2001/01/010`], [`hep-ph/0011363`].

[162] G. Aad et al., the **ATLAS** Collaboration, *First Tuning of HERWIG/JIMMY to ATLAS Data*. ATLAS-PHYS-PUB 2010-014 (2010).

[163] M. Tamsett et al., *Measurement of Multi-jet Cross-Sections in Proton-Proton Collisions at 7 TeV Center-of-mass Energy*. ATLAS-PHYS-COM 2011-138 (2011).

[164] G. Aad et al., the **ATLAS** Collaboration, *Estimating Track Momentum Resolution in Minimum Bias Events using Simulation and* K_s^0 *in* $\sqrt{s} = 900$ *GeV Collision Data*. ATLAS-CONF 2010-009 (2010).

[165] G. Aad et al., the **ATLAS** Collaboration, *Tracking Results and Comparison to Monte Carlo Simulation at 900 GeV*. ATLAS-CONF 2010-011 (2010).

[166] G. Aad et al., the **ATLAS** Collaboration, *Properties of Tracks in Jets in Proton-Proton Collisions at* $\sqrt{s} = 7$ *TeV in the ATLAS Detector*. ATLAS-CONF 2010-061 (2010).

[167] G. Aad et al., the **ATLAS** Collaboration, *Properties and Internal Structure of Jets Produced in Proton-Proton Collisions at* $\sqrt{s} = 900$ *GeV*. ATLAS-CONF 2010-018 (2010).

[168] G. Aad et al., the **ATLAS** Collaboration, *ATLAS Calorimeter Response to Single Isolated Hadrons and Estimation of the Calorimeter Jet Scale Uncertainty*. ATLAS-CONF 2010-052 (2010).

[169] G. Aad et al., the **ATLAS** Collaboration, *Jet Energy Resolution and Reconstruction Efficiencies from In-situ Techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy* $\sqrt{s} = 7$ *TeV*. ATLAS-CONF 2010-054 (2010).

[170] G. Aad et al., the **ATLAS** Collaboration, *In-situ Pseudorapidity Intercalibration for Evaluation of Jet Energy Scale Uncertainty Using Dijet Events in Proton-Proton Collisions at 7 TeV*. ATLAS-CONF 2011-014 (2011).

[171] G. Aad et al., the **ATLAS** Collaboration, *Response of isolated particles identified using resonances in proton-proton collisions at* $\sqrt{s}=7$ *TeV with the ATLAS detector.* ATLAS-CONF 2011-019 (2011).

[172] G. Aad et al., the **ATLAS** Collaboration, *ATLAS Calorimeter Response to Single Isolated Hadrons and Estimation of the Calorimeter Jet Scale Uncertainty.* ATLAS-CONF 2011-028 (2011).

[173] M. Begel, S. Majewski, M. Campanelli, J. Gray et al., *Azimuthal Decorrelations in Dijet Events at* $\sqrt{s}=7$ *TeV.* ATL-COM-PHYS 2010-1044 (2010).

[174] G. Aad et al., the **ATLAS** Collaboration, *In-situ Pseudo-rapidity Inter-calibration to Evaluate Jet Energy Scale Uncertainty and Calorimeter Performance in the Forward Region.* ATLAS-CONF 2010-055 (2010).

[175] G. Aad et al., the **ATLAS** Collaboration, *Probing the Jet Energy Measurement at the TeV-Scale Using the Multi-jet Balance Technique in Proton-Proton Collisions at* $\sqrt{s}=7$ *TeV.* ATLAS-CONF 2011-029 (2011).

[176] G. Aad et al., the **ATLAS** Collaboration, *Determination of the ATLAS Jet Energy Measurement Uncertainty Using Photon-jet Events in Proton-Proton Collisions at* $\sqrt{s}=7$ *TeV.* ATLAS-CONF 2011-031 (2011).

[177] G. Anders, N. Bacchetta, V. Balabara, C. Barschel et al., *LHC Bunch Current Normalisation for the April–May 2010 Luminosity Calibration Measurements.* CERN-ATS-Note 2011-004 PERF (2011).

[178] G. Anders, N. Bacchetta, V. Balabara, C. Barschel et al., *LHC Bunch Current Normalisation for the October 2010 Luminosity Calibration Measurements.* CERN-ATS-Note 2011-016 PERF (2011).

[179] G. Aad et al., the **ATLAS** Collaboration, *Measurement of Differential Cross Section and Fragmentation of Jets from Tracks in Proton-Proton Collisions at Center-of-mass Energy* $\sqrt{s}=7$ *TeV with the ATLAS Detector.* ATLAS-CONF 2010-049 (2010).

[180] G. Aad et al., the **ATLAS** Collaboration, *Azimuthal Decorrelations in Dijet Events at* $\sqrt{s}=7$ *TeV.* ATLAS-CONF 2010-083 (2010).

[181] G. Aad et al., the **ATLAS** Collaboration, *Jet Energy Resolution from In-situ Techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy* $\sqrt{s} = 7$ *TeV*. ATLAS-COM-PHYS 2011-240 (2011).

[182] G. Aad et al., the **ATLAS** Collaboration, *Validation of the ATLAS Jet Energy Scale Uncertainties Using Tracks in Proton-Proton Collisions at* $\sqrt{s} = 7$ *TeV*. ATLAS-CONF 2011-067 (2011).

[183] G. Aad et al., the **ATLAS** Collaboration, *Close-by Jet Effects on Jet Energy Scale Calibration in pp Collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS Detector*. ATLAS-CONF 2011-062 (2011).

[184] G. P. Salam, *Towards Jetography*. Eur. Phys. J. **C67**, (2010) 637–686, [`doi: 10.1140/epjc/s10052-010-1314-6`], [`arXiv:0906.1833`].

[185] M. Cacciari, J. Rojo, G. P. Salam, and G. Soyez, *Quantifying the Performance of Jet Definitions for Kinematic Reconstruction at the LHC*. JHEP **12**, (2008) 032, [`doi: 10.1088/1126-6708/2008/12/032`], [`arXiv:0810.1304`].

[186] A. Moraes, *Modeling the Underlying Event: Generating Predictions for the LHC*. In *Proceedings of the First International Workshop on Multiple Partonic Interactions at the LHC (MPI@LHC08), Perugia, Italy, October 27–31, 2008*, edited by P. Bartalini and L. Fanò, 124–130 (DESY-PROC-2009-06, DESY, Germany) (2010), [`arXiv:1003.4220`].

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

[188] G. Aad et al., the **ATLAS** Collaboration, *ATLAS Monte Carlo Tunes for MC09*. ATL-PHYS-PUB 2010-002 (2010).

[189] T. Affolder et al., the **CDF** Collaboration, *Charged Jet Evolution and the Underlying Event in Proton-Antiproton Collisions at 1.8 TeV*. Phys. Rev. D **65**(9), (2002) 092002, [`doi: 10.1103/PhysRevD.65.092002`].

[190] D. Acosta et al., the **CDF** Collaboration, *The Underlying Event in Hard Interactions at the Tevatron anti-p p Collider.* Phys. Rev. **D70**, (2004) 072002, [`doi: 10.1103/PhysRevD.70.072002`], [`hep-ex/0404004`].

[191] D. Acosta et al., the **CDF** Collaboration, *Soft and Hard Interactions in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1800 and 630 GeV.* Phys. Rev. D **65**(7), (2002) 072005, [`doi: 10.1103/PhysRevD.65.072005`].

[192] V. M. Abazov et al., the **D0** Collaboration, *Measurement of Dijet Azimuthal Decorrelations at Central Rapidities in p anti-p Collisions at $\sqrt{s} = 1.96$ TeV.* Phys. Rev. Lett. **94**, (2005) 221801, [`doi: 10.1103/PhysRevLett.94.221801`], [`hep-ex/0409040`].

[193] T. Aaltonen et al., the **CDF** Collaboration, *Measurement of Particle Production and Inclusive Differential Cross Sections in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV.* Phys. Rev. **D79**, (2009) 112005, [`doi: 10.1103/PhysRevD.79.112005`], [`arXiv:0904.1098`].

[194] K. Abe et al., the **SLD** Collaboration, *Measurement of the b-Quark Fragmentation Function in Z0 Decays.* Phys. Rev. **D65**, (2002) 092006, [`doi: 10.1103/PhysRevD.65.092006`], [`hep-ex/0202031`].

[195] A. Buckley, J. Katzy, B. Kersevan, L. Mijovic et al., *ATLAS Monte Carlo Tunes for MC09.* ATL-PHYS-INT 2010-005 (2010).

[196] The LEP Injector Study Group, *LEP design report, Vol. 1, The LEP Injector Chain.* CERN-LEP-TH 83-29, Geneva (1983).

[197] The LEP Injector Study Group, *LEP design report, Vol. 2, The LEP Main Ring.* CERN-LEP 84-01, Geneva (1984).

[198] J. B. Adams and E. J. N. Wilson, *Design Studies for a Large Proton Synchrotron and its Laboratory.* Nucl. Instrum. Methods (87), (1970) 157–179.

[199] H. Schulz, H. Hoeth, A. Buckley, and L. Mijovic, *Global Tunes versus Process-Dependent Results.* Talk, from the Vector Boson plus Jets as a Signal and Background Topical Workshop.

[200] F. Siegert, *Status of Sherpa Tuning.* Talk, from the 2011 ATLAS Monte Carlo Tuning Workshop.

[201] A. Sherstnev and R.S. Thorne, *Parton Distributions for LO Generators.* Eur. Phys. J. **C55**, (2008) 553–575, [`doi: 10.1140/epjc/s10052-008-0610-x`], [`arXiv:0711.2473`].

[202] A. Sherstnev and R. S. Thorne, *Different PDF Approximations Useful for LO Monte Carlo Generators* [`doi: 10.3360/dis.2008.149`], [`arXiv:0807.2132`].

[203] H. L. Lai et al., the **CTEQ** Collaboration, *Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions.* Eur. Phys. J. **C12**, (2000) 375–392, [`doi: 10.1007/s100529900196`], [`hep-ph/9903282`].

[204] M. Bahr et al., *Herwig++ Physics and Manual.* Eur. Phys. J. **C58**, (2008) 639–707, [`doi: 10.1140/epjc/s10052-008-0798-9`], [`arXiv:0803.0883`].

[205] Z. Nagy, *Next-to-Leading Order Calculation of Three Jet Observables in Hadron Hadron Collisions.* Phys. Rev. D **68**(9), (2003) 094002, [`doi: 10.1103/PhysRevD.68.094002`].

[206] J. Huston, *(Private communication).*

[207] L. Sawyer, *D0 Results on Three-Jet Production, Multijet Cross-Section Ratios, and Minimum Bias Angular Correlations.* PoS **DIS2010**, (2010) 135.

[208] A. Buckley et al., *Rivet User Manual* [`arXiv:1003.0694`].

[209] J. R. Andersen and J. M. Smillie, *Constructing All-Order Corrections to Multi-Jet Rates.* JHEP **01**, (2010) 039, [`doi: 10.1007/JHEP01(2010)039`], [`arXiv:0908.2786`].

[210] G. Aad et al., the **ATLAS** Collaboration, *Measurement of Dijet Azimuthal Decorrelations in pp Collisions at sqrt(s)=7 TeV* [`arXiv:1102.2696`], (accepted by Phys. Rev. Lett.).

[211] G. Aad et al., the **ATLAS** Collaboration, *Dijet Production with a Jet Veto in pp Collisions at* $\sqrt{s} = 7$ *TeV*. ATLAS-CONF 2010-085 (2010).

[212] M. Tamsett et al., *Measurement of Multi-jet Cross-Sections in Proton-Proton Collisions at 7 TeV Center-of-mass Energy, Updated.* ATLAS-PHYS-COM 2011-399 (2011).

[213] A. Buckley, C. Gwenlan, J. Katzy, and P. Loch, *(Private communication).*