Measurement of multi-jet cross sections in proton–proton collisions at a 7 TeV center-of-mass energy

The ATLAS Collaboration
CERN, 1211 Geneva 23, Switzerland

Abstract Inclusive multi-jet production is studied in proton–proton collisions at a center-of-mass energy of 7 TeV, using the ATLAS detector. The data sample corresponds to an integrated luminosity of 2.4 pb$^{-1}$. Results on multi-jet cross sections are presented and compared to both leading-order plus parton-shower Monte Carlo predictions and to next-to-leading-order QCD calculations.

1 Introduction

At hadron colliders, events containing multiple jets in the final state are plentiful and provide a fertile testing ground for the theory of the strong interaction, quantum chromodynamics (QCD). At high transverse momentum ($p_T$), the production of jets is modeled by QCD as the hard scattering of partons and the subsequent parton showering, followed by a hadronization process. Within this framework, the jet energy is related to the energy of partons produced in hadron collisions. Consequently, the study of energy distributions for multi-jet events provides a fundamental and direct test of QCD at hadron colliders.

In addition to their role in testing QCD, multi-jet events are often an important background in searches for new particles and new interactions at high energies. In particular, systematic uncertainties that contribute to multi-jet cross section measurements can carry over into search analyses. Even though the impact of multi-jets on such analyses will vary according to the specific data selection criteria, a study of multi-jet events serves as an important cross check of models used to estimate backgrounds originating from jets.

Measurements of multi-jet cross sections at the Tevatron have been performed by the CDF [1, 2] and D0 [3, 4] collaborations in proton–antiproton collisions at 1.8 TeV center-of-mass energy. The CMS collaboration has recently released measurements of the three-jet to two-jet cross sections at a 7 TeV center-of-mass energy [5]. In this paper, a first study is performed of multi-jet events from proton–proton collisions at 7 TeV center-of-mass energy using the ATLAS detector at the Large Hadron Collider (LHC) at CERN. The data sample used for the analysis was collected from April until August 2010 and represents a total integrated luminosity of 2.4 pb$^{-1}$. Approximately half a million events with at least two jets in the final state are selected using this data sample.

Two primary motivations for the multi-jet study in this paper are to evaluate how robust leading-order perturbative QCD (LO pQCD) calculations are in representing the high jet multiplicity events, and to test next-to-leading-order perturbative QCD (NLO pQCD) calculations. For the leading-order comparisons, events with up to six jets in the final state are studied, and for the next-to-leading-order perturbative QCD study, the focus is on three-jet events and their comparison to two-jet events. At present, there is no four-jet NLO pQCD calculation available.

The paper is organized as follows. Section 2 presents a description of the ATLAS detector. Section 3 discusses the cross sections and kinematics. In Sect. 4, theoretical calculations, to which the measurements are compared, are described. Sections 5 and 6 discuss the event selection and data corrections. The main uncertainty coming from the jet energy scale is discussed in Sect. 7, followed by the results and conclusions.

2 The ATLAS detector

The ATLAS detector

The ATLAS experiment consists of an approximately 45-meter long, 25-meter cylindrically shaped detector centered on the proton–proton interaction point. A detailed description of the ATLAS experiment can be found elsewhere [6]. High-energy particles produced in collisions initially pass through an inner tracking system embedded in a
2 T solenoidal magnetic field. The field is located in a region of diameter 2.3 meters and 7 meters long also centered at the interaction point. The design of this tracking system allows the measurement of charged particle kinematics within the pseudorapidity\(^1\) range of \(|\eta| < 2.5\). Precision tracking using the pixel detector with a space point resolution as small as 10 microns by 70 microns (in the beam direction) begins at a radial distance of 5 cm from the interaction point [7]. The identification of the vertex from which the jet originates, performed with the inner tracker, is of interest in the study of multi-jet events.

Just outside the inner tracker system are liquid argon and scintillating tile calorimeters used for the measurement of particle energies. A liquid-argon/lead electromagnetic calorimeter covers the pseudorapidity range of \(|\eta| < 3.2\). This calorimeter is complemented by hadronic calorimeters, built using scintillating tiles and iron for \(|\eta| < 1.7\) and liquid argon and copper in the end-cap (1.5 < |\eta| < 3.2). Forward calorimeters extend the coverage to \(|\eta| = 4.9\). The calorimeters are the primary detectors used to reconstruct the jet energy in this analysis and allow the reconstruction of the jet \(p_T\) with a fractional resolution of better than 0.10 for jets of \(p_T = 60\text{ GeV}\) and 0.05 for jets of \(p_T = 1\text{ TeV}\).

Outside the calorimeters is a toroidal magnetic field that extends to the edge of the detector. Additional tracking detectors designed for measuring muon kinematics are placed within this magnetic field. The impact of muons in the analysis presented in this paper is negligible.

The ATLAS trigger system employs three trigger levels, of which only the hardware-based first level trigger is used in this analysis. Events are selected using the calorimeter based jet trigger. The first level jet trigger [8] uses coarse detector information to identify areas in the calorimeter where energy deposits above a certain threshold occur. A simplified jet finding algorithm based on a sliding window of size \(\Delta\phi \times \Delta\eta = 0.8 \times 0.8\) is used to identify these areas. This algorithm uses coarse calorimeter towers with a granularity of \(\Delta\phi \times \Delta\eta = 0.2 \times 0.2\) as inputs.

3 Cross section definitions and kinematics

In this analysis, the anti-\(k_t\) algorithm [9, 10], with jet constituents combined according to their four-momenta, is used to identify jets. For high multiplicity studies, which includes events with up to six jets, the resolution parameter in the jet reconstruction is fixed to \(R = 0.4\) to contend with the limited phase space and to reduce the impact of the underlying event in the jet energy determination. For testing NLO pQCD calculations, the study focuses on three-jet events, a resolution parameter of \(R = 0.6\) is preferred, since a larger value of \(R\) is found to be less sensitive to theoretical scale uncertainties. The anti-\(k_t\) algorithm was chosen for a variety of reasons. It can be implemented in the NLO pQCD calculation, is infra-red and collinear safe to all orders, and reconstructs jets with a simple geometrical shape.

Jet measurements are corrected for all experimental effects such that they can be compared to particle-level predictions. At the particle level, jets are built using all final-state particles with a proper lifetime longer than 10 ps. These corrections are described in Sect. 6. The NLO pQCD calculation is not interfaced to a Monte Carlo simulation with hadronization and other non-perturbative effects. The correction for non-perturbative effects applied to the NLO pQCD calculation is described in Sect. 4.

Cross sections are calculated in bins of inclusive jet multiplicity, meaning that an event is counted in a jet multiplicity bin if it contains a number of jets that is equal to or greater than that multiplicity. For example, an event with three reconstructed jets will be counted both in the two-jet and three-jet multiplicity bins. Inclusive multiplicity bins are used because they are stable in the pQCD fixed-order calculation, unlike exclusive bins. Only jets with \(p_T > 60\text{ GeV}\) and \(|y| < 2.8\) are counted in the analysis. These cuts are chosen to ensure that the jets are reconstructed with high efficiency. The leading jet is further required to have \(p_T > 80\text{ GeV}\) to stabilize the NLO pQCD calculations in the dijet case [11].

4 Theoretical predictions

Measurements are compared to pQCD calculations at leading order and next-to-leading order.

Many different effects are included in leading-order Monte Carlo simulations of jets at the LHC. These include the modeling of the underlying event and hadronization, which can affect the cross section calculation through their impact on the jet kinematics [12]. Effects arising from differences between the matrix-element plus parton-shower (ME+PS) calculation (with up to 2 \(\rightarrow n\) matrix-element scattering diagrams) and the parton-shower calculation alone (with only 2 \(\rightarrow 2\) matrix-element scattering diagrams) also need to be understood. These topics are not easily separable, since tuning of some of the effects (such as the underlying event) to data is needed, and the tuning process fixes other inputs in the Monte Carlo simulation.
such as the proton parton distribution functions (PDF), the parton-shower model, and the hadronization model. The inability to separate out some effects makes it difficult to obtain a full estimate of the theoretical uncertainty associated with the leading-order Monte Carlo predictions. Furthermore, leading-order Monte Carlo predictions are affected by large normalization uncertainties.

In this study, the goal is to test the performance of the different leading-order Monte Carlo simulations, so that they can be used to estimate multi-jet backgrounds for new particle searches, not to discern whether deviations with respect to QCD are present in the data. The latter goal is best achieved by comparing with NLO pQCD calculations (discussed later in this section). For these reasons, the leading-order Monte Carlo predictions are all normalized to the measured inclusive two-jet cross section and then used for shape comparisons. No attempt is made to assign a theoretical uncertainty to these leading-order predictions, Instead, numerous different Monte Carlo simulations and currently available tunes have been studied in order to investigate the impact of each of these effects on the measurements. Only a representative subset is shown in the results, even though conclusions are drawn on the basis of all simulations studied.

For the leading-order analysis, ALPGEN [13] is used to generate events with up to six partons in the final state using the leading-order set of proton PDFs CTEQ6L1 [14]. A factorization and renormalization scale, $Q$, that varies from event to event is used in the event generation, where $Q^2 = \sum p_T^2$. The sum runs over all final state partons. ALPGEN is interfaced to PYTHIA 6.421 [15, 16] and, alternatively, to HERWIG/JIMMY [17–20] to sum leading logarithms to all orders in the parton-shower approximation and to include non-perturbative effects such as hadronization and the underlying event. The ATLAS generator tunes from 2009 (MC09 [21]) and from 2010 (AUET1 [22]) are used. Additional tunes have been investigated to assess the impact of the underlying-event and parton-shower tuning. With comparable underlying-event tunes and ALPGEN parameters, the comparison between ALPGEN+PYTHIA and ALPGEN+HERWIG/JIMMY uncovers differences that may arise from different parton-shower implementations and hadronization models.

SHERPA [23] with its default parameters and renormalization scale scheme from version 1.2.3 is also used to generate events with up to six partons in the final state. This provides an independent matrix-element calculation with a different matching scheme between the matrix element and the parton shower. Detailed studies of individual tunes using SHERPA, however, are not performed in this paper.

The PYTHIA event generator is also compared to the data to study the limitations of leading-order $2 \rightarrow 2$ matrix-element calculations. This generator implements a leading-order matrix-element calculation for $2 \rightarrow 2$ processes, $p_T$-ordered parton showers, an underlying-event model for multiple-parton interactions and the Lund string model for hadronization. The MRST2007 modified leading order [24, 25] PDFs interfaced with the AMBT1 [21] generator tune are used in the sample generation.

For the purpose of understanding detector effects, the particles generated in the leading-order Monte Carlo generators are passed through a full simulation of the ATLAS detector and trigger [26] based on GEANT4 [27]. Additional proton–proton collisions are added to the hard scatter in the simulation process to reproduce realistic LHC running conditions. Events and jets are selected using the same criteria in data and Monte Carlo simulations.

For the next-to-leading-order pQCD study, the calculation implemented in NLOJet++ 4.1.2 [28] is used. The renormalization and factorization scales are varied independently by a factor of two in order to estimate the impact of higher order terms not included in the calculation. An additional requirement that the ratio of the renormalization and factorization scales did not differ by more than a factor of two was imposed. Two next-to-leading-order PDF sets, CTEQ 6.6 [29] and MSTW 2008 NLO [25], are used for calculating the central values. Only results obtained with the MSTW 2008 NLO PDF set are shown in the paper since the results obtained with the CTEQ 6.6 PDF set are compatible. The 90% confidence-limit error sets are used in the evaluation of the PDF uncertainties. The uncertainty in the calculations due to the uncertainty in the value of $\alpha_S$ is determined by varying the value of $\alpha_S$ by $\pm 0.002$ for each PDF set.

The NLOJet++ program implements a matrix-element calculation, and therefore it lacks a parton-shower interface and does not account for non-perturbative effects. To compare to particle-level measurements, a correction factor is required. PYTHIA and HERWIG++ [30] are used to generate samples without underlying event. Jets in these samples are reconstructed from partons after the parton shower, and observables are compared to those obtained at the particle level in the standard HERWIG++ and PYTHIA samples. A multiplicative correction is calculated

$$C_{\text{non-pert}} = \frac{\sigma_{\text{particle}}}{\sigma_{\text{parton}}} \cdot \frac{\sigma_{\text{no UE}}}{\sigma_{\text{no UE}}} ,$$

where $\sigma$ is the observable of interest calculated at the particle or parton level in the samples with and without underlying event. The correction factor takes the next-to-leading-order pQCD calculations to the particle level. This correction is calculated in three different samples. The correction obtained using the PYTHIA AMBT1 sample is taken as the
default value for the analysis, and the systematic uncertainty is estimated from the maximum spread compared to the results from the other models (marked with an asterisk in Table 1). The size of this correction is less than 5% in all observables studied in the next-to-leading-order pQCD analysis. The total uncertainty quoted on the next-to-leading-order pQCD calculations comes from the quadrature sum of the uncertainties from the renormalization and factorization scales, the proton PDFs, $\alpha_S$ and the non-perturbative corrections.

Table 1 presents a summary of the different Monte Carlo generators and tunes that the data are compared to in this paper.

<table>
<thead>
<tr>
<th>Generator</th>
<th>PDF</th>
<th>Tune</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALPGEN+HERWIG/JIMMY</td>
<td>CTEQ6L1 [14]</td>
<td>AUET1 [22]</td>
</tr>
<tr>
<td>ALPGEN+PYTHIA</td>
<td>CTEQ6L1 [14]</td>
<td>MC09' [21]</td>
</tr>
<tr>
<td>SHERPA</td>
<td>CTEQ66 [29]</td>
<td>Default (v1.2.3)</td>
</tr>
<tr>
<td>HERWIG++$^*$</td>
<td>MRSTMC al [24, 25]</td>
<td>Default (v2.5)</td>
</tr>
</tbody>
</table>

For jets trigger object, and the event passes a two-jet trigger. studied by requiring that the second leading jet is matched calculated by requiring that the leading jet passes the single-ficiency of the trigger to fire on the second leading jet is calculated using the minimum bias triggers. Then, the ef-triggers. The efficiency for triggering on the leading jet is estimated from the maximum spread compared to the re-

The single-jet triggers with a 10 GeV level-1 threshold have been shown to be fully efficient for events with at least one anti-$k_t$ jet with $R = 0.4$ and calibrated $p_T > 60$ GeV [31] using events triggered with the minimum bias triggers. The efficiency for triggering on the leading jet is calculated using the minimum bias triggers. Then, the ef-ficiency of the trigger to fire on the second leading jet is calculated by requiring that the leading jet passes the single-jet trigger. Similarly, the efficiency of the third leading jet is studied by requiring that the second leading jet is matched to a jet trigger object, and the event passes a two-jet trigger. For $p_T > 60$ GeV, events are selected on the trigger plateau.

Figure 1 shows the efficiency for the third leading jet to fire the three-jet trigger as a function of the reconstructed jet $p_T$ for jets of $R = 0.4$ (a) and $R = 0.6$ (b). The efficiencies calculated in data are compared to those from the Monte Carlo detector simulation. The efficiency as a function of jet rapidity is also shown for $R = 0.4$ jets (c) for $p_T > 60$ GeV. A small inefficiency is present in the data at $y = \pm 1.5$. In this transition region between the barrel and end-cap calorimeters the level-1 trigger energy sums did not span between the calorimeters for the early data used here, resulting in this small efficiency drop, which is not modeled by the Monte Carlo simulation. The simulation is not corrected for this effect, since its impact in the measurements is negligible, and included as part of the systematic uncertainties in the data correction described in Sect. 6.

The event-level efficiency as a function of the closest distance between two selected $R = 0.4$ offline jets for events selected using the three-jet trigger is shown in Fig. 1(d). The study probes possible topological dependences in the trigger. A dependence at low $\Delta R$ is observed, where $\Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2}$ represents the minimum separation between selected jets in the event. The dependence on $\Delta R$ is well described by the Monte Carlo simulation. For the calculation of the efficiency in the data, the two leading jets are associated with level-1 jet objects and an assumption is made that any topological inefficiency will only affect one of the level-1 jet objects. Figure 1(d) indicates that events in which two jets are separated by $\Delta R < 0.6$ have an efficiency of less than 100%. This inefficiency appears to depend weakly on the jet $p_T$ and is well described in the detector simulation for events where the closest distance between selected jets is greater than 0.45. The inefficiency is accounted for in the Monte Carlo-based data correction described in Sect. 6. Such an inefficiency is not observed in the analysis of jets reconstructed using the anti-$k_t$ algorithm with resolution parameter $R = 0.6$.

The three-jet trigger operated without pre-scaling for the entire data collection period used in this paper. All events falling in the three-jet inclusive multiplicity bin are, therefore, selected using the three-jet trigger with a jet threshold of 10 GeV on the level-1 jet objects. On the other hand, a large pre-scaling was applied to certain two-jet triggers. In order to select events in the two-jet inclusive multiplicity bin, several two-jet triggers were used. Three two-jet triggers with symmetric transverse energy thresholds of 10, 15 and 30 GeV were combined independently, weighted by the integrated luminosity associated with each trigger. The three
Fig. 1 Jet trigger efficiency for the third leading jet as a function of $p_T$ for anti-$k_t$ jets with $R = 0.4$ (a), and $R = 0.6$ (b). Jet trigger efficiency as a function of $y$ of the third leading jet with $p_T > 60$ GeV and $R = 0.4$ (c). Jet trigger efficiency as a function of the minimum separation $\Delta R$ between the two closest jets (d). The efficiency is shown both as calculated in data, as described in the text, and in Monte Carlo simulations for the three-jet trigger with a level-1 cut on the jet transverse energy of 10 GeV.

5.2 Vertex reconstruction

The primary vertex or vertices are found using tracks that originate near the beam collision spot [32], satisfy quality criteria [33] and have transverse momentum above 150 MeV. A vertex is seeded by searching for the global maximum in the distribution of $z$ coordinates of reconstructed tracks. The vertex is fitted using the position of this seed along with neighboring tracks. Tracks incompatible with the reconstructed vertex are used to seed new vertices until no tracks are left. This analysis only uses events in which at least one primary vertex with at least five associated tracks has been reconstructed. No cut on the primary vertex position is applied. The event vertex is defined as the vertex in the event for which the sum of the $p_T$ of the tracks associated to that vertex is largest.

5.3 Jet reconstruction

Topological clusters of calorimeter energy evaluated at the electromagnetic scale [31] are used as inputs to the jet finding algorithm. These clusters use the baseline calibration derived from test beams and from $Z \rightarrow ee$ data [34], which reconstructs the energy of particles interacting electromagnetically. The anti-$k_t$ algorithm [9] with resolution parameters $R = 0.4$ and $R = 0.6$ and full four-momentum recombination is used to reconstruct jets from clusters. The jet four-momentum is calculated assuming that the jet origin is at the position of the event vertex. The jet reconstruction is fully efficient in the Monte Carlo simulation for jets with transverse momentum above 30 GeV. The reconstruction efficiency in the simulation compares well with the one measured with data [31].

5.4 Jet energy scale calibration

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. The difference between a hadron-level jet and an electromagnetic-scale jet is due to the different calorimeter response to electromagnetic objects compared to strongly interacting objects, detector induced showering and energy deposition in regions of the detector that are not instrumented. A Monte Carlo-based calibration that corrects for these effects as a function of $p_T$ and $y$ is used to obtain jets with the correct energy scale [35].

5.5 Jet selection criteria

Jets considered in the analysis are selected using the following kinematic and data quality selection criteria:

1. The event must contain at least one jet with $|y| < 2.8$ and a $p_T$ greater than 80 GeV.
2. Jets are required to have $|y| < 2.8$ and $p_T > 60$ GeV in order to be counted.
3. A series of jet cleaning cuts were applied to eliminate various detector effects and suppress beam and other non-collision backgrounds. Overall, these cuts reduce the total number of jets by less than 0.1%. These cuts have been shown to be efficient in eliminating noise, while rejecting a negligible number of true jets.
4. In order to reduce the effects from pileup events, jets are only accepted if at least 70% of their charged particle $p_T$ comes from the event vertex. The charged particle $p_T$ is calculated as the scalar sum of the $p_T$ of reconstructed tracks within a $\Delta R$ equal to the resolution parameter used in the jet reconstruction. Overall, this cut lowers the number of selected two-jet events by 0.4%, and its effect increases with jet multiplicity. The cut reduces the number of selected six-jet events by 3.4%. All observables show a negligible dependence on the number of reconstructed primary vertices once this cut is applied [36]. Jets with no charged particle content are accepted, but only constitute a few percent of events at low $p_T$.
5. Only events with at least two selected jets are used in the analysis.

For illustrative purposes, Fig. 2 presents an event display of a six-jet event passing all selection cuts. The transverse energy deposition in the calorimeter is shown as a function of $\eta$ and $\phi$. For this event, the six selected jets are well separated spatially.

Table 2 presents the total number of multi-jet events versus inclusive jet multiplicity. No correction for trigger pre-scales in the two-jet bin has been applied to the numbers in the table.

<table>
<thead>
<tr>
<th>Inclusive multiplicity</th>
<th>Number of events</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥2</td>
<td>500,148</td>
</tr>
<tr>
<td>≥3</td>
<td>112,740</td>
</tr>
<tr>
<td>≥4</td>
<td>10,999</td>
</tr>
<tr>
<td>≥5</td>
<td>1,100</td>
</tr>
<tr>
<td>≥6</td>
<td>115</td>
</tr>
</tbody>
</table>

6 Data correction for efficiencies and resolution

A correction is needed to compare the measurements to theoretical predictions. The correction, which accounts for trigger inefficiencies, detector resolutions and other detector effects that affect the jet counting, is performed in a single
Fig. 3 (Color online) Bin-by-bin correction factors for the cross sections (a) and for the $n$ to $n-1$ cross-section ratios (b) as a function of the inclusive jet multiplicity. The correction factors calculated using the ALPGEN+HERWIG/JIMMY AUET1 sample are shown with the systematic uncertainty as a yellow band around the points. See the text for an explanation of the legend labels.

To perform the correction, the ALPGEN+HERWIG/JIMMY AUET1 Monte Carlo simulation is used. The sample includes, on average, two additional soft proton–proton collision events overlapping with the hard scatter simulated by ALPGEN. The data have fewer overlapping collisions, as revealed by the distribution of the number of selected vertices, and the Monte Carlo simulation is subsequently weighted to match the distribution from the data. The truth distribution is independent of the additional collisions, since jets are built using particles simulated by the ALPGEN+HERWIG/JIMMY Monte Carlo simulation only. Distributions in the Monte Carlo simulation are not further reweighted to match the data. The impact of differences in shapes between data and Monte Carlo simulation on the calculation of the correction factors is instead considered part of the systematic uncertainties in these factors.

The uncertainty in the correction factors is estimated taking into account several effects. One arises from the spread in correction factors coming from different generators (ALPGEN+HERWIG/JIMMY AUET1 and PYTHIA AMBT1). A second detailed study is performed in which the simulated jet $p_T$, $y$ and $\phi$ resolution is varied according to their measured uncertainties [37, 38]. Third, the shape of the simulated distributions is varied within limits set by the present measurements in order to account for possible biases caused by the input distributions. Samples with a trigger inefficiency in the crack region, with different pile-up rejection cuts and different primary vertex multiplicity distributions are also used to estimate the uncertainty arising from trigger effects and from the impact of overlapping proton–proton collisions. All these effects impact the systematic uncertainties in the correction factors, and their uncertainties are ultimately added in quadrature to provide the final systematic uncertainty in the bin-by-bin correction.

The uncertainty in the correction factors is estimated taking into account several effects. One arises from the spread in correction factors coming from different generators (ALPGEN+HERWIG/JIMMY AUET1 and PYTHIA AMBT1). A second detailed study is performed in which the simulated jet $p_T$, $y$ and $\phi$ resolution is varied according to their measured uncertainties [37, 38]. Third, the shape of the simulated distributions is varied within limits set by the present measurements in order to account for possible biases caused by the input distributions. Samples with a trigger inefficiency in the crack region, with different pile-up rejection cuts and different primary vertex multiplicity distributions are also used to estimate the uncertainty arising from trigger effects and from the impact of overlapping proton–proton collisions. All these effects impact the systematic uncertainties in the correction factors, and their uncertainties are ultimately added in quadrature to provide the final systematic uncertainty in the bin-by-bin correction. Although only important for particular bins, statistical uncertainties on the correction factors are added to the total uncertainty.

The integrated luminosity of the dataset used in this paper is measured to be $2.43 \pm 0.08 \text{ pb}^{-1}$ [39] and the associated uncertainty is not shown in the figures.

7 Uncertainty on the jet energy scale

The jet energy scale uncertainty is the dominant uncertainty for most results presented in this paper. The fact that cross sections fall steeply as a function of jet $p_T$ implies that even a relatively small uncertainty in the determination of the jet
The jet energy scale and its uncertainty [35] have been determined for jets from a dijet sample without nearby activity in the calorimeter. For a multi-jet analysis, additional systematic uncertainties need to be considered. These uncertainties arise from the difference in the calorimeter response to jets of different flavors as well as the impact of nearby activity in the calorimeter on the jet energy measurement.

Figure 4 shows the calorimeter $p_T$ response for light-quark and gluon jets in the region $|y| < 0.8$ as a function of the true $p_T$ calculated using the PYTHIA AMBT1 Monte Carlo simulation sample. The response for jets in the two-jet inclusive multiplicity bin is also shown. Light-quark and gluon jets were tagged using the highest-energy parton found in the Monte Carlo simulation particle record within a cone of radius equal to the resolution parameter of the jet algorithm. Only jets that had no additional reconstructed jet of $p_T > 7$ GeV evaluated at the electromagnetic scale within $\Delta R = 1.0$ from the jet axis were used in order to decouple effects in the response caused by jet flavor from effects related to the presence of nearby calorimeter activity.

The Monte Carlo simulation shows a slightly higher fraction of jets matched to gluons for high-multiplicity final states, particularly in the ALPGEN samples. To the extent that the Monte Carlo simulation reflects the data, the difference in response as a function of multiplicity is accounted for in the bin-by-bin correction for efficiencies and resolution.

An additional jet energy scale uncertainty, however, 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 jet energy scale uncertainty could be different. In what follows, this uncertainty is referred to as the ‘flavor response’ uncertainty. This uncertainty is estimated using Monte Carlo simulations [35] by studying the difference between the gluon and light-quark jet response under various assumptions. However, the relative change of the light-quark jet response with respect to the gluon jet response is found to be negligible in all simulations studied [40], so the effect can be safely ignored.

In addition, the fraction of light-quark and gluon jets in multi-jet samples 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 precision with which the flavor composition of the sample is known thus also affects the precision of the jet energy measurement. The flavor composition depends on many theoretical aspects in the event production (parton distribution functions, limitations of leading-order calculations, initial and final state radiation tuning) and the uncertainty in the predictions is not easy to estimate using Monte Carlo simulations. The uncertainty is determined using a data-driven method that provides a measurement of the flavor composition up to the four-jet inclusive multiplicity bin and for jets of $p_T < 210$ GeV [40]. The method uses template fits to the distribution of jet widths and to the number of tracks associated with jets in bins of $\eta$, $p_T$, jet isolation and jet multiplicity. The templates are obtained using Monte Carlo simulations modified to match the distributions found in the two-jet bin. Using these template fits, the measurement of the flavor composition is determined to an accuracy of $\approx 10\%$.

Overall, ALPGEN predicts the correct flavor composition to within $30\%$ in bins where the number of collected events is enough to perform the fits. At high $p_T$ and high multiplicities the flavor composition is assumed to be unknown when calculating the jet energy scale uncertainty.

Jets with nearby activity have different properties than the jets used to estimate the jet energy scale uncertainty. In addition, the fraction of jets with nearby activity increases with jet multiplicity. Figure 5 gives the probability of a selected jet occurring within $\Delta R = 1.0$ of a reconstructed jet with $p_T > 7$ GeV at the electromagnetic scale as a function of inclusive jet multiplicity. The overlap probability increases with jet multiplicity, a trend which is reproduced by the simulations.

Jets with nearby activity have a different jet energy scale, as has been demonstrated in Monte Carlo simulations [41]. The systematic uncertainty on their energy scale has been evaluated by studying the correlation between the $p_T$ of the tracks associated to the jet and the $p_T$ measured in the calorimeter, and contributes to the final uncertainty in the jet energy scale used in this analysis.

Approximately $40\%$ of the selected events have more than one vertex in the interaction, indicating the presence
of additional proton–proton interactions. The vertex multiplicity is low enough that, with a luminous region of several mm and a vertex reconstruction resolution of a few hundred μm, the impact of merged vertices on the analysis is negligible. For the instantaneous luminosities considered in this paper, the probability that two hard events would occur at the same time is negligible. However, a soft interaction occurring in parallel with the hard interaction can produce a contamination of energy from a nearby soft jet. The average effect of these overlapping interactions on the jet energy scale is accounted for by an offset correction, and the systematic uncertainty on that correction has been evaluated [42]. The impact of this uncertainty on the overall jet energy scale uncertainty used in this analysis is negligible for the vast majority of events. The overlapping interactions can also impact the jet counting since the resolution of the jet energy reconstruction depends on the instantaneous luminosity. The effect becomes small after performing a cut on the fraction of charged particle $p_T$ that originates from the event vertex and that is associated to the jet, as described in Sect. 5. The Monte Carlo simulation has been shown to describe tracks within jets [43] and general features of events with pile-up interactions [42]. An uncertainty due to the efficiency of the cut has been estimated in Sect. 6.

In summary, the jet energy scale uncertainty is primarily made of three components: the uncertainty calculated for isolated jets, the uncertainty caused by the presence of nearby calorimeter deposits, and the flavor composition uncertainty. The uncertainty on the energy scale of isolated jets is the largest contributor to the total uncertainty in most bins, except for jets in the five and six-jet bins and of $p_T < 200$ GeV, for which the flavor composition uncertainty is comparable. The positive systematic uncertainty on the jet energy scale of isolated jets falling in the barrel and in high-multiplicity bins varies from 5% at 60 GeV to 2.5% at 1 TeV. In the three-jet and four-jet bins, where the flavor composition is better constrained, the systematic uncertainty is at most 3.5%. The negative systematic uncertainty is smaller and $≈ 3\%$ across all $p_T$ in the barrel. The impact of nearby calorimeter deposits is small, increasing the overall uncertainty by at most 1%. The uncertainty is propagated to the measured distributions using the ALPGEN+HERWIG/JIMMY Monte Carlo simulation and varying the $p_T$ of all jets in the event up or down according to the estimated uncertainties. The use of the same procedure in the data yields comparable results, but the results obtained in the Monte Carlo simulation are favored to eliminate the impact of statistical uncertainties in the data in bins with few events.

8 Results

In this section, measurements corrected to the particle level are compared to theoretical predictions. For comparisons to leading-order Monte Carlo simulations, the anti-$k_t$ algorithm with resolution parameter $R = 0.4$ is used to define a jet. In Figs. 6–10 and 12(b), the darker (orange) shaded error band bracketing the measured cross section corresponds to the total systematic uncertainty, evaluated by adding the individual systematic uncertainties in quadrature but excluding the uncertainty coming from the luminosity measurement. The ratio of the predictions from the Monte Carlo simulations to the measurements is shown at the bottom of each figure. For Figs. 6, 8 and 9, the lighter (grey) error band that appears in the ratio of the predictions from the Monte Carlo simulations to the measurements represents the total systematic uncertainty on the shape of the measured distributions.

Only a few representative Monte Carlo simulations that were studied are shown in the figures and tables. All Monte Carlo simulations are normalized to the measured inclusive two jet cross section. The normalization factors applied to the Monte Carlo simulations studied are given in Table 3, and distinctive features of some of the Monte Carlo simulations not shown are discussed when relevant. Most ALPGEN Monte Carlo simulations predict an inclusive multi-jet cross section similar to the measured cross section, while the PYTHIA Monte Carlo simulation requires scaling factors which differ the most from unity. The differences in the normalization factors between ALPGEN+PYTHIA MC09 and ALPGEN+HERWIG/JIMMY AUET1 illustrate differences between PYTHIA and HERWIG/JIMMY and their interplay

---

3All measurements in this section have been compiled in tables that can be found in HEPDATA. The NLO pQCD calculation results are also presented in the tables when applicable.
Table 3 Normalization factors applied to each of the Monte Carlo simulations in order to match the measured inclusive two-jet cross section

<table>
<thead>
<tr>
<th>Leading-order Monte Carlo</th>
<th>Normalization factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALPGEN+HERWIG AUET1</td>
<td>1.11</td>
</tr>
<tr>
<td>ALPGEN+PYTHIA MC09′</td>
<td>1.22</td>
</tr>
<tr>
<td>PYTHIA AMBT1</td>
<td>0.65</td>
</tr>
<tr>
<td>SHERPA</td>
<td>1.06</td>
</tr>
</tbody>
</table>

Fig. 6 Total inclusive jet cross section as a function of multiplicity. The data are compared to leading-order Monte Carlo simulations (ALPGEN+HERWIG AUET1, ALPGEN+PYTHIA MC09′, PYTHIA AMBT1 and SHERPA) normalized to the measured inclusive two-jet cross section. The darker (orange) shaded error bands correspond to the systematic uncertainties on the measurement, excluding the luminosity uncertainty. The lighter (grey) shaded error band corresponds to the systematic uncertainty on the shape of the measured distribution. A plot of the ratio of the different Monte Carlo simulations to the data is presented at the bottom of the figure.

with the matrix-element and parton-shower matching implemented in ALPGEN. The normalization factor for SHERPA is found to be the closest to unity.

Figure 6 shows the results for the cross section as a function of the inclusive jet multiplicity. The measurement systematics are dominated by the jet energy scale uncertainty and range from 10–20% at low multiplicities to almost 30–40% at high multiplicities. The Monte Carlo simulation predictions agree with the measured results across the full inclusive multiplicity spectrum, even when comparing just to the shape of the distributions.

A study that reduces significantly the impact of systematic uncertainties is the ratio of the n-jet to (n − 1)-jet cross section as a function of multiplicity. In this ratio, the impact of the jet energy scale uncertainty is significantly reduced and the uncertainty due to the luminosity cancels out. Figure 7 presents the results for such a study. Both the uncertainties in the data correction for efficiencies and resolutions and the jet energy scale contribute comparably to the total systematic uncertainty, whereas the statistical uncertainties are smaller than the systematic uncertainties, and negligible in most bins. All Monte Carlo simulations are consistent with the measurements at the present precision, yet there is a noticeable spread in the predictions. Differences at the level of 15% are observed between PYTHIA AMBT1 and ALPGEN+PYTHIA MC09′ in the first bin. These differences most likely arise from the difference between the pure parton-shower (with 2 → 2 matrix elements) implemented in PYTHIA and the parton-shower-matched matrix-element calculation (with up to 2 → 6 matrix elements) implemented in ALPGEN. All ALPGEN+PYTHIA tunes studied are comparable in this measurement.

The differential cross section for multi-jet events as a function of the jet $p_T$ is useful for characterizing kinematic features. The comparison reveals significant differences between the leading order calculations and the measurements. Figure 8 presents the $p_T$-dependent differential cross sections for the leading, second leading, third leading and fourth leading jet in multi-jet events. The systematic uncertainty in the measurement is 10–20% across $p_T$ and increasing up to 30% for the fourth leading jet differential cross section. The jet energy scale systematic uncertainty remains the dominant uncertainty in the measurement. However, the uncertainty is less than 10% (grey shaded error band) for the leading and second leading jet $p_T$ distributions.

All Monte Carlo simulations agree reasonably well with the data (orange darker shaded error band). However, the PYTHIA AMBT1 Monte Carlo simulation predicts a somewhat steeper slope compared to the data as a function of the leading jet $p_T$ and the second leading jet $p_T$, whereas the SHERPA and ALPGEN Monte Carlo simulations predict a less steeply falling slope compared to the data. When using
The differential cross section as a function of leading jet $p_T$ for events with $N_{\text{jets}} \geq 2$ (a), 2nd leading jet $p_T$ for events with $N_{\text{jets}} \geq 2$ (b), 3rd leading jet $p_T$ for events with $N_{\text{jets}} \geq 3$ (c) and 4th leading jet $p_T$ for events with $N_{\text{jets}} \geq 4$ (d). The results are compared to different leading-order Monte Carlo simulations normalized to the measured inclusive two-jet cross section. Other details are as in the caption to Fig. 6.

The three-to-two-jet ratio as a function of the leading jet $p_T$ can be used to tune Monte Carlo simulations for effects due to final state radiation. Figure 10 presents the results on the measurement of the three-to-two-jet cross section ratio as a function of leading jet $p_T$ for jets built with the anti-$k_t$ algorithm using the resolution parameter $R = 0.6$ and with different minimum $p_T$ cuts for all non-leading jets$^4$.

The cut on the $p_T$ of the leading jet in the event selection is also increased with the minimum $p_T$ cut ($p_T^{\text{lead}} > 110$ GeV is used in Fig. 10(b) and $p_T^{\text{lead}} > 160$ GeV in Fig. 10(c)). The systematic uncertainties on the measurement are small ($\sim 5\%$), except in the lowest $p_T$ bin, where uncertainties in the data correction for efficiencies and resolutions and the jet energy scale dominate. ALPGEN+HERWIG AUET1 and ALPGEN+PYTHIA MC09$^7$ describe the data well, and the agreements are largely independent of the tunes chosen. SHERPA also describes the data well. PYTHIA AMBT1 predicts a higher ratio than that measured over the $p_T$ range from 200 GeV to 600 GeV. The disagreement is similar

$^4$Results (not shown) were also obtained using $R = 0.4$ and are compiled in tables in HEPDATA.
Fig. 9 Differential cross section as a function of $H_T$ for events with at least two selected jets (a), three selected jets (b) and four selected jets (c). The results are compared to different leading-order Monte Carlo simulations normalized to the measured inclusive two-jet cross section. Other details are as in the caption to Fig. 6.

Fig. 10 Three-to-two-jet differential cross-section ratio as a function of the leading jet $p_T$. In the figures, a resolution parameter $R = 0.6$ is used. The three figures contain a minimum $p_T$ cut for all non-leading jets of (a) 60 GeV, (b) 80 GeV and (c) 110 GeV. The results are compared to leading-order Monte Carlo simulations. Other details are as in the caption to Fig. 6.
Fig. 11 (Color online) Three-to-two-jet differential cross-section ratio as a function of the leading jet $p_T$. In the figures a resolution parameter $R = 0.6$ is used. The three figures contain a minimum $p_T$ cut for all non-leading jets of (a) 60 GeV, (b) 80 GeV and (c) 110 GeV. The results are compared to a NLO pQCD calculation and (b) several leading-order Monte Carlo simulations. The systematic uncertainties on the theoretical prediction for the NLO pQCD calculations are shown as dotted red lines above and below the theoretical prediction.

Figure 11 presents the same measurement results as Fig. 10, except the data are now compared to the NLO pQCD calculations corrected for non-perturbative effects. The MSTW 2008 NLO PDF set has been used, but comparable results are obtained with the CTEQ 6.6 PDF set. The systematic uncertainties on the theoretical predictions are shown as dotted red lines above and below the theoretical prediction. The NLO pQCD calculations describe the data well, except in the lowest $p_T$ bin, where there is a large discrepancy. 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.

Fig. 12 (Color online) Three-to-two-jet differential cross-section ratio as a function of the sum of the $p_T$ of the two leading jets ($H_T^{(2)}$) using $R = 0.6$. The two figures present the same measurements and error bands. The data are compared to (a) a NLO pQCD calculation and (b) several leading-order Monte Carlo simulations. The systematic uncertainties on the theoretical prediction for the NLO pQCD calculations are shown as dotted red lines above and below the theoretical prediction. The data error bands are identical to the results shown in Fig. 10. The systematic uncertainties on the theoretical prediction are shown as dotted red lines above and below the theoretical prediction when other $2 \rightarrow 2$ Monte Carlo simulations with different tunes and PDFs are used. The systematic uncertainty in the lowest $p_T$ bin decreases significantly as the minimum $p_T$ cut is raised to 80 GeV for all jets.
Additional NLO pQCD calculations of the three-to-two-jet cross section ratio were performed as a function of different kinematic variables, such as $H_T$, the sum of the $p_T$ of the two leading jets ($H_T^{(2)}$) and the sum of the $p_T$ of the three leading jets. The NLO pQCD calculation for the ratio as a function of $H_T^{(2)}$ was found to give the smallest theoretical scale uncertainty and is, therefore, most sensitive to input parameters such as $\alpha_S$. Figure 12 shows a comparison of the measurement to both (a) NLO pQCD and (b) leading-order calculations for $R = 0.6$. Scale uncertainties of the NLO pQCD calculations are larger for jets with $R = 0.4$ than with $R = 0.6$. The theoretical uncertainty of the NLO pQCD calculations shown in Fig. 12 is comparable to the measurement uncertainties, but is significantly reduced compared to the theoretical uncertainties presented in Fig. 11. With the reduced theoretical uncertainty, the disagreement between data and the NLO pQCD calculations in the lowest $H_T^{(2)}$ bin is now enhanced. Due to the kinematic cuts applied in the analysis, the NLO pQCD calculations only account for the lowest-order contribution to the two-jet cross section in the region where the sum of the first and second leading jet $p_T$ is less than 160 GeV. Consequentially, this effective leading-order estimation is subject to large theoretical uncertainties, which might be responsible for the observed discrepancy.

A comparison of the same measurement to leading-order Monte Carlo simulations is given in Fig. 12(b). The general agreement between leading-order Monte Carlo simulations with the measurements follows the same general trends as the comparison of the three-to-two-jet ratio versus leading jet $p_T$ shown in Fig. 10.

9 Summary and conclusion

A first dedicated study of multi-jet events has been performed in proton–proton collisions at a center-of-mass energy of 7 TeV using the ATLAS detector with an integrated luminosity of 2.4 pb$^{-1}$. Leading-order Monte Carlo simulations have been compared to multi-jet inclusive and differential cross sections. The present study extends up to a multiplicity of six jets, up to jet $p_T$ of 800 GeV and up to event $H_T$ of 1.6 TeV.

For events containing two or more jets with $p_T > 60$ GeV, of which at least one has $p_T > 80$ GeV, a reasonable agreement is found between data and leading-order Monte Carlo simulations with parton-shower tunes that describe adequately the ATLAS $\sqrt{s} = 7$ TeV underlying-event data. The agreement is found after the predictions of the Monte Carlo simulations are normalized to the measured inclusive two-jet cross section.

All models reproduce the main features of the multijet data. The $2 \rightarrow 2$ calculations show some departure from the data for the three-to-two jet cross-section ratios, predicting a higher ratio than observed. The $2 \rightarrow n$ calculations describe the measured ratios, independent of the tune or parton shower implementation. The shape of the differential cross sections as a function of $p_T$ and $H_T$, studied in the inclusive two-jet and three-jet bins, falls off less (more) steeply in the $2 \rightarrow n (2 \rightarrow 2)$ calculations.

A measurement of the three-to-two-jet cross section ratio as a function of the leading jet $p_T$ and the sum of the two leading jet $p_T$s is described well by ALPGEN, SHHERPA and a NLO pQCD calculation, albeit with a significant discrepancy in the lowest $p_T$ bin for the latter comparison. Future comparisons with NLO pQCD calculations will be useful for constraining parameters, such as parton distribution functions or the value of the strong coupling constant, $\alpha_S$. Systematic uncertainties from the measurement are presently comparable to the theoretical uncertainties, but should be reduced with larger data samples and higher energy collisions.

Acknowledgements We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NERC and NRC, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DFKR, DNSRC and Lundbeck Foundation, Denmark; ARTEMIS, European Union; IN2P3-CNRS, CEA-DSM/IRFU, France; GNAS, Georgia; BMBF, DFG, HGF, MPG and AvH Foundation, Germany; GSRT, Greece; ISF, MINHERVA, GIF, DIP and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNiSW, Poland; GICECS and FCT, Portugal; MERSYS (MECTS), Romania; MES of Russia and ROSATOM, Russian Federation; JINR; MSTD, Serbia; MSSR, Slovakia; ARRS and MVZT, Slovenia; DST/NRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SER, SNSF and Cantons of Bern and Geneva, Switzerland; NSEC, Taiwan; TAEK, Turkey; STFC, the Royal Society and Leverhulme Trust, United Kingdom; DOE and NSF, United States of America.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN and the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KTJ/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA) and in the Tier-2 facilities worldwide.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

The ATLAS Collaboration

G. Aad\textsuperscript{48}, B. Abbott\textsuperscript{111}, J. Abdallah\textsuperscript{11}, A.A. Abdelalim\textsuperscript{49}, A. Abdesselam\textsuperscript{118}, O. Abdinov\textsuperscript{10}, B. Abi\textsuperscript{112}, M. Abolins\textsuperscript{88}, H. Abramowicz\textsuperscript{153}, H. Abreu\textsuperscript{115}, E. Acerbi\textsuperscript{164a,164b}, B.S. Acharya\textsuperscript{164a,164b}, D.L. Adams\textsuperscript{24}, T.N. Addy\textsuperscript{56}, J. Adelman\textsuperscript{175}, M. Aderholz\textsuperscript{99}, S. Adomeit\textsuperscript{75}, P. Adragna\textsuperscript{75}, T. Adye\textsuperscript{129}, S. Aeberhardt\textsuperscript{22}, J.A. Aguilar-Saavedra\textsuperscript{124a,124b}, M. Aharrouche\textsuperscript{81},