Measurement of hadronic event shapes in high-pT multijet final states at \( \sqrt{s} = 13 \) TeV with the ATLAS detector

The ATLAS collaboration

DOI
10.1007/JHEP01(2021)188

Publication date
2021

Document Version
Final published version

Published in
Journal of High Energy Physics

License
CC BY

Citation for published version (APA):
Measurement of hadronic event shapes in high-$p_T$ multijet final states at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

ABSTRACT: A measurement of event-shape variables in proton-proton collisions at large momentum transfer is presented using data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector at the Large Hadron Collider. Six event-shape variables calculated using hadronic jets are studied in inclusive multijet events using data corresponding to an integrated luminosity of $139 \text{fb}^{-1}$. Measurements are performed in bins of jet multiplicity and in different ranges of the scalar sum of the transverse momenta of the two leading jets, reaching scales beyond 2 TeV. These measurements are compared with predictions from Monte Carlo event generators containing leading-order or next-to-leading order matrix elements matched to parton showers simulated to leading-logarithm accuracy. At low jet multiplicities, shape discrepancies between the measurements and the Monte Carlo predictions are observed. At high jet multiplicities, the shapes are better described but discrepancies in the normalisation are observed.

KEYWORDS: Hadron-Hadron scattering (experiments)

ArXiv ePrint: 2007.12600
1 Introduction

Event shapes [1, 2] are a class of observables that describe the dynamics of energy flow in multijet final states. Normally, event-shape observables are defined such that they vanish for $2 \rightarrow 2$ processes and increase to a maximum for final states with uniformly distributed energy. These observables are sensitive to different aspects of the theoretical description of these strong-interaction processes. They are usually defined to be infrared and collinear safe, which enables their calculation in perturbative Quantum Chromodynamics (QCD). Hard, wide-angle radiation is studied by investigating the tails of the event-shape distributions. These configurations are sensitive to higher-order corrections to the dijet cross section, which are available up to next-to-next-to-leading order (NNLO) [3]. Other regions of the event-shape distributions provide information about anisotropic, back-to-back configurations, which are sensitive to the details of the resummation of soft logarithms in the theoretical predictions.

Event-shape observables have been measured in $e^+e^-$ collisions at LEP [4–7], $ep$ collisions at HERA [8], and $p\bar{p}$ collisions at the Tevatron [9]. More recently, they have been measured in $pp$ collisions at $\sqrt{s} = 7$ TeV by the ALICE, CMS and ATLAS Collaborations [10–13]. ALICE also published measurements at $\sqrt{s} = 0.9$ and 2.76 TeV [10], and the CMS Collaboration published a measurement at $\sqrt{s} = 13$ TeV [14].
In this study, several different event-shape variables are investigated, probing the properties of the multijet energy flow at the large, \(O(\text{TeV})\), energy scales provided by the Large Hadron Collider (LHC) at \(\sqrt{s} = 13\,\text{TeV}\). Measurements are compared with fixed-order matrix elements matched to parton shower Monte Carlo (MC) predictions for a selection of observables that cover various aspects of the physics of multijet processes. These observables are defined in detail in section 3. The measurements are performed for different energy regimes, given by the scalar sum of transverse momenta of the two leading jets, \(H_{T2} = p_{T1} + p_{T2}\), and the jet multiplicity, \(n_{\text{jet}}\). The phase-space region explored in this analysis is defined by \(H_{T2} > 1\,\text{TeV}\), with jet \(p_T > 100\,\text{GeV}\) to reduce experimental uncertainties and non-perturbative effects. This paper extends the currently available measurements by studying the dependence of event shapes on \(n_{\text{jet}}\), which is not usually found in the literature. This approach provides inputs with which to compare future higher-order QCD predictions, since the power of \(\alpha_s\) in the perturbative expansion of the cross section increases with each additional jet emission in the final state. In addition, it allows the definition of phase-space regions sensitive to processes beyond-the-Standard-Model, which are often characterised by isotropically distributed multijet final states [15–17]. Measurements of the differential cross sections as a function of \(n_{\text{jet}}\) for different energy scales are also reported.

The paper is organised as follows. The ATLAS detector is described in section 2. The measurement strategy and the definitions of the observables are discussed in section 3, followed by the details of the data sample and MC simulations in section 4. Section 5 is dedicated to the object and event-selection criteria. The correction to particle level is described in section 6, followed by a discussion of systematic uncertainties in section 7. The particle-level results are compared with MC predictions in section 8 and the summary and conclusions are provided in section 9.

2 ATLAS detector

The ATLAS detector [18] at the LHC covers nearly the entire solid angle around the collision point.\(^1\) It consists of an inner charged-particle tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting toroidal magnets.

The inner-detector system is immersed in a 2T axial magnetic field and provides charged-particle tracking in the range \(|\eta| < 2.5\). Closest to the interaction point, the high-granularity silicon pixel detector covers the vertex region and typically provides four measurements per track, with the first hit normally recorded in the insertable B-layer installed before Run 2 [19, 20]. It is followed by the silicon microstrip tracker, which usually provides eight measurements per track. These silicon detectors are complemented by the

\(^1\)ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upwards. Cylindrical coordinates \((r, \phi)\) are used in the transverse plane, \(\phi\) being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle \(\theta\) as \(\eta = -\ln(\tan(\theta/2))\).
transition radiation tracker (TRT), which enables radially extended track reconstruction up to $|\eta| = 2.0$. The TRT also provides electron identification information based on the fraction of hits (typically 30 in total) above a high energy-deposit threshold that corresponds to transition radiation.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, electromagnetic calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) calorimeters, with an additional thin LAr presampler that covers $|\eta| < 1.8$, to correct for energy loss in material upstream of the calorimeters. Hadronic calorimetry is provided by the steel/scintillator-tile calorimeter, segmented into three barrel structures in the region $|\eta| < 1.7$, and two copper/LAr calorimeters in the endcap regions ($1.5 < |\eta| < 3.2$). The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimised for electromagnetic and hadronic measurements, respectively. Surrounding the calorimeters is a muon spectrometer that consists of three air-core superconducting toroidal magnets and tracking chambers, providing precision tracking for muons with $|\eta| < 2.7$ and trigger capability for $|\eta| < 2.4$.

A two-level trigger system is used to select events for offline analysis [21]. Interesting events are selected by the first-level trigger system implemented with custom electronics which uses a subset of the detector information. This is followed by selections made by algorithms implemented in a software-based high-level trigger. The first-level trigger accepts events from the 40 MHz bunch crossings at a rate below 100 kHz, which the high-level trigger further reduces in order to record events to disk at about 1 kHz.

3 Observable definitions and measurement strategy

This paper presents measurements for six event-shape variables using hadronic jets. For each event, the thrust axis $\hat{n}_T$ is defined as the direction with respect to which the projection of the jet momenta is maximised [22, 23]. The transverse thrust $T_\perp$ and its minor component $T_m$ can be expressed with respect to $\hat{n}_T$ as

$$T_\perp = \frac{\sum_i |\vec{p}_{T,i} \cdot \hat{n}_T|}{\sum_i |\vec{p}_{T,i}|}; \quad T_m = \frac{\sum_i |\vec{p}_{T,i} \times \hat{n}_T|}{\sum_i |\vec{p}_{T,i}|},$$

where the index $i$ runs over all jets in the event. It is also useful to define $\tau_\perp = 1 - T_\perp$, so lower values of $\tau_\perp$ indicate a back-to-back, dijet-like configuration. The range of allowed values for these variables is, by construction, $0 \leq \tau_\perp < 1 - 2/\pi$ and $0 \leq T_m < 2/\pi$. Higher values of $\tau_\perp$ indicate a larger energy flow orthogonal to the thrust axis, while large values of $T_m$ indicate a large energy flow outside the plane spanned by the thrust and the beam axes.

The sphericity $S$ and aplanarity $A$ encode information on the isotropy of the final-state energy distribution. These two observables are defined in terms of the eigenvalues of the linearised sphericity tensor of the event [24, 25], given by

$$\mathcal{M}_{xyz} = \frac{1}{\sum_i |\vec{p}_i|} \sum_i \frac{1}{|\vec{p}_i|} \left( \begin{array}{ccc} p_{x,i}^2 & p_{x,i} p_{y,i} & p_{x,i} p_{z,i} \\ p_{y,i} p_{x,i} & p_{y,i}^2 & p_{y,i} p_{z,i} \\ p_{z,i} p_{x,i} & p_{z,i} p_{y,i} & p_{z,i}^2 \end{array} \right).$$
Its eigenvalues \( \{ \lambda_k \} \), which satisfy \( \sum_k \lambda_k = 1 \) by definition, are ordered so that \( \lambda_1 > \lambda_2 > \lambda_3 \), and the corresponding event shapes are defined as

\[
S = \frac{3}{2} (\lambda_2 + \lambda_3); \quad A = \frac{3}{2} \lambda_3.
\] (3.3)

\( S \) takes values between 0 and 1, with larger values indicating more spherical events. \( A \) takes values between 0 and 1/2 and is a measure of the extent to which the radiation is contained in the plane defined by the two first eigenvectors of the sphericity tensor defined in eq. 3.2. The larger the value of \( A \), the less planar the event.

The transverse linearised sphericity tensor is constructed using only the transverse momentum components:

\[
\mathcal{M}_{xy} = \frac{1}{\sum_i |\vec{p}_i|} \sum_i \frac{1}{|\vec{p}_i|} \begin{pmatrix} p_{x,i}^2 & p_{x,i}p_{y,i} & p_{y,i}^2 \\ p_{y,i} & p_{x,i} & p_{y,i} \end{pmatrix}.
\]

Its eigenvalues \( \{ \mu_k \} \), which satisfy \( \sum_k \mu_k = 1 \) by definition, are ordered so that \( \mu_1 > \mu_2 \) and the corresponding transverse sphericity event shape is defined as

\[
S_\perp = \frac{2\mu_2}{\mu_1 + \mu_2}.
\] (3.4)

It takes values between 0 and 1, with large (small) values indicating isotropic (back-to-back) events in the transverse plane.

To illustrate the meaning of the event-shape variables, figure 1 shows two different multijet final states. The first represents a three-jet event with large values of \( \tau_\perp \) and \( S_\perp \). The second represents a five-jet event with low values of \( \tau_\perp \) and \( S_\perp \).

The quantities in eq. 3.3 correspond to linear combinations of the eigenvalues of the sphericity tensor. However, one may consider quadratic and cubic combinations of the
eigenvalues \{\lambda_i\} \text{[26]}, such as
\[
C = 3(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3),
\]
\[
D = 27(\lambda_1\lambda_2\lambda_3).
\]

The quantities defined in eqs. 3.5 and 3.6 are restricted to the range \([0, 1]\). These are also useful observables to characterise multijet events. Since \(C\) is defined by products of eigenvalue pairs, it vanishes for two-jet events, while \(D\), which is defined by multiplying the three eigenvalues, vanishes for events in which all jet momenta lie on the same plane.

To study the dependence of the observables on the event topology and energy scale, each of the six event-shape observables is measured as a function of \(n^{\text{jet}}\) and \(H_T^2\). Events that satisfy the selection requirements are classified in bins of \(n^{\text{jet}}\) (\(= 2, 3, 4, 5\) and \(\geq 6\)) and \(H_T^2\) (\(1\) TeV \(< H_T^2 \< 1.5\) TeV, \(1.5\) TeV \(< H_T^2 \< 2.0\) TeV, \(H_T^2 \> 2\) TeV).

A measurement of the multijet production cross section in the different \(H_T^2\) bins is performed in the same fiducial phase space in which the event-shape observables are measured, i.e. in events with 2, 3, 4, 5 or \(\geq 6\) jets. Since many of the experimental uncertainties that affect the measurement of the event-shape observables are correlated between \(n^{\text{jet}}\) bins, these measurements are presented normalised to the inclusive dijet cross section in bins of \(H_T^2\). In this way, the experimental uncertainties discussed in section 7 are significantly reduced while preserving important physics information, such as the relative shape of the distributions.

4 Data and Monte Carlo samples

The dataset used in this analysis comprises the data taken from 2015 to 2018 at a centre-of-mass energy of \(\sqrt{s} = 13\) TeV. After applying quality criteria to ensure good ATLAS detector operation, the total integrated luminosity useful for data analysis is 139 fb\(^{-1}\). The average number of inelastic pp interactions produced per bunch crossing for the dataset considered, hereafter referred to as ‘pile-up’, is \(\langle \mu \rangle = 33.6\).

Several MC samples were used for this analysis; they differ in the matrix element (ME) calculation and/or the parton shower (PS). In order to populate all regions of the spectra, these samples are divided into subsamples with differing kinematic characteristics. The samples were produced using the PYTHIA [27, 28], SHERPA [29], HERWIG [30–32] and MadGraph5 \_aMC@NLO (hereafter referred to as MG5 \_aMC) [33] generators.

The PYTHIA sample was generated using PYTHIA 8.235. The matrix element (ME) was calculated for the \(2 \rightarrow 2\) process. The parton shower algorithm includes initial- and final-state radiation based on the dipole-style \(p_T\)-ordered evolution, including \(\gamma \rightarrow q\bar{q}\) branchings and a detailed treatment of the colour connections between partons [27]. The renormalisation and factorisation scales were set to the geometric mean of the squared transverse masses of the two outgoing particles (labelled 3 and 4), i.e. \(\sqrt{m_{T3}^2 \cdot m_{T4}^2} = \sqrt{(p_T^2 + m_3^2) \cdot (p_T^2 + m_4^2)}\). The NNPDF 2.3 LO PDF set [34] was used in the ME generation, in the parton shower, and in the simulation of multi-parton interactions (MPI). The ATLAS A14 [35] set of tuned parameters (tune) is used for the parton shower and MPI,
whilst hadronisation was modelled using the Lund string model [36, 37]. The Pythia sample contains per-event weights that allow the estimation of uncertainties due to the parton shower parameters, including the variations of the renormalisation scale for QCD initial- and final-state radiation (ISR – FSR) and variations of the non-singular terms of the splitting functions.

The Sherpa sample was generated using Sherpa 2.1.1. The ME calculation is included for the $2 \rightarrow 2$ and $2 \rightarrow 3$ processes at leading order (LO), and the Sherpa parton shower [38, 39] with $p_T$ ordering was used for the showering. Matrix element renormalisation and factorisation scales for $2 \rightarrow 2$ processes were set to the harmonic mean of the Mandelstam variables $s$, $t$ and $u$ [40], whereas the Catani-Marchesini-Webber (CMW) [41] scale was chosen for the additional emission in $2 \rightarrow 3$ processes. The CT14 NNLO [42] PDF set was used for the matrix element calculation, while the parameters used for the modelling of the MPI and the parton shower were set according to the CT10 tune [43]. The Sherpa sample makes use of the dedicated Sherpa AHADIC model for hadronisation [44], which is based on the cluster fragmentation algorithm [45].

The MG5_aMC+Pythia 8 sample was generated using MadGraph5_aMC@NLO 2.3.3. The calculation includes matrix elements computed at leading order for up to four final-state partons, using the NNPDF 3.0 NLO [46] PDF set, and merged with the CKKW-L prescription [47, 48]. The renormalisation and factorisation scales were set to the transverse mass of the $2 \rightarrow 2$ system that results from the $k_t$ clustering [49] and the merging scale was set to 30 GeV. The parton shower and hadronisation were handled by Pythia 8.212. The ATLAS A14 tune with the NNPDF 2.3 LO PDF set was used for the shower and MPI, and the Lund string model was used for the modelling of hadronisation.

Finally, two Herwig samples were generated using Herwig 7.1.3 at next-to-leading order. This includes NLO accuracy for the $2 \rightarrow 2$ process and LO accuracy for the $2 \rightarrow 3$ process. The ME was calculated using Matchbox [50] with the MMHT2014 NLO PDF [51]. The renormalisation and factorisation scales were set to the $p_T$ of the leading jet. The first sample uses an angle-ordered parton shower, while the second sample uses a dipole-based parton shower. In both cases, the parton shower was interfaced to the ME calculation using the MC@NLO matching scheme. The angle-ordered shower evolves on the basis of $1\rightarrow 2$ splittings with massive DGLAP functions using a generalised angular variable and employs a global recoil scheme once showering has terminated. The dipole-based shower uses $2\rightarrow 3$ splittings with Catani-Seymour kernels with an ordering in transverse momentum and so is able to perform recoils on an emission-by-emission basis. For both Herwig samples, the parameters that control the MPI and parton shower simulation were set according to the H7-UE-MMHT tune [51], and the hadronisation was modelled by means of the cluster fragmentation algorithm.

The main features of the samples described above are summarised in table 1. The Pythia, Sherpa and Herwig samples were passed through the Geant4-based [52] ATLAS detector-simulation program [53] since they were also used to unfold the measurements to the particle level, as described in section 6. They are reconstructed and analysed with the same processing chain as the data. The MG5_aMC samples are used for comparison at particle level.
Table 1. Properties of the Monte Carlo samples used in the analysis, including the perturbative order in $\alpha_s$, the number of final-state partons, the PDF set, the parton shower algorithm, the renormalisation and factorisation scales and the value of $\alpha_s(m_Z)$ for the matrix element.

<table>
<thead>
<tr>
<th>Generator</th>
<th>ME order</th>
<th>FS partons</th>
<th>PDF set</th>
<th>Parton shower</th>
<th>Scales $\mu_R, \mu_F$</th>
<th>$\alpha_s(m_Z)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PYTHIA</td>
<td>LO</td>
<td>2</td>
<td>NNPDF 2.3 LO</td>
<td>$p_T$-ordered</td>
<td>$(m_T^2 \cdot m_T^4)^{3/2}$</td>
<td>0.140</td>
</tr>
<tr>
<td>SHERPA</td>
<td>LO</td>
<td>2,3</td>
<td>CT14 NNLO</td>
<td>CSS (dipole)</td>
<td>$H(s,t,u)$ [2 $\rightarrow$ 2]</td>
<td>0.118</td>
</tr>
<tr>
<td>MG5_aMC</td>
<td>LO</td>
<td>2,3,4</td>
<td>NNPDF 3.0 NLO</td>
<td>$p_T$-ordered</td>
<td>$m_T$</td>
<td>0.118</td>
</tr>
<tr>
<td>HERWIG</td>
<td>NLO</td>
<td>2,3</td>
<td>MMHT2014 NLO</td>
<td>Angle-ordered Dietol</td>
<td>$\max_i {p_T}_{i=1}^N$</td>
<td>0.120</td>
</tr>
</tbody>
</table>

The generation of the simulated event samples includes the effect of multiple $pp$ interactions per bunch crossing, as well as the effect on the detector response of interactions from bunch crossings before or after the one containing the hard interaction. In addition, during the data-taking, some modules (so-called “dead-tile modules”) situated in various $\eta$-$\phi$ regions of the ATLAS hadronic calorimeter were found to be malfunctioning for some periods of time, leading to poorly reconstructed jets in these regions. The resulting dead regions in the hadronic calorimeter were included in the simulation for the PYTHIA, SHERPA and HERWIG samples.

5 Event selection and object reconstruction

Events with high-$p_T$ jets are preselected using a single-jet trigger with a minimum $p_T$ threshold of 460 GeV. Events are required to have at least one reconstructed vertex that contains two or more associated tracks with transverse momentum $p_T > 500$ MeV. The reconstructed vertex that maximises $\sum p_T^2$, where the sum is performed over tracks associated with the vertex, is chosen as the primary vertex.

Jets are reconstructed using the anti-$k_t$ algorithm [54] with radius parameter $R = 0.4$ using the FastJet program [55]. The inputs to the jet algorithm are particle-flow objects [56], which make use of both the calorimeter and the inner-detector information to precisely determine the momenta of the input particles. The jet calibration procedure includes energy corrections for pile-up, as well as angular corrections. Effects due to energy losses in inactive material, shower leakage, the parameterisation of the magnetic field and inefficiencies in energy clustering and jet reconstruction are taken into account. This is done using a simulation-based correction, in bins of $\eta$ and $p_T$, derived from the relation of the reconstructed jet energy to the energy of the corresponding particle-level jet, not including muons or non-interacting particles. In a final step, an in situ calibration corrects for residual differences in the jet response between the MC simulation and the data using $p_T$-balance techniques for dijet, $\gamma$+jet, $Z$+jet and multijet final states.

The selected jets must have $p_T > 100$ GeV and $|\eta| < 2.4$. These requirements reject pile-up jets and reduce experimental uncertainties. In addition, jets are required to satisfy quality criteria that reject beam-induced backgrounds (jet cleaning) [57]. The efficiency of this requirement for selecting good jets with $p_T > 100$ GeV is larger than 99.5%. Events are required to have at least two selected jets. The two leading jets are further required
to satisfy $H_{T2} > 1$ TeV. This requirement ensures a trigger efficiency of $\approx 100\%$. About 57.5 million events in data satisfy the selection criteria. Figure 2 shows the detector level distributions for the selected jets $p_T$ and $H_{T2}$, along with MC predictions. The binning of the event-shape distributions is chosen as a compromise between maximising the number of bins while minimising migration between bins due to the resolution of the measured variables.

6 Unfolding to particle level

In order to make meaningful comparisons with particle-level MC predictions, the event-shape distributions need to be corrected for distortions induced by the response of the ATLAS detector and associated reconstruction algorithms. The fiducial phase-space region is defined at particle level for all particles with a mean decay length $c\tau > 10$ mm; these particles are referred to as ‘stable’. Particle-level jets are reconstructed using the anti-$k_t$ algorithm with $R = 0.4$ using stable particles, excluding muons and neutrinos. The fiducial phase space closely follows the event selection criteria defined in section 5. Particle-level jets are required to have $p_T > 100$ GeV and $|y| < 2.4$, where $y$ represents the rapidity. Events with at least two particle-level jets are considered. These events are also required to have $H_{T2} > 1$ TeV.

The unfolding is performed using an iterative algorithm based on Bayes’ theorem [58]. The algorithm takes into account inefficiencies and resolution effects due to the detector response that lead to bin migrations between the detector-level and particle-level phase spaces. For each observable, the method makes use of a transfer matrix, $M_{ij}$, obtained from MC simulation, that parameterises the probability of an event generated in bin $i$ to be reconstructed in bin $j$. The correction can thus be written as a linear equation

$$R_i = \sum_{j=1}^{N} \frac{E_j}{P_i} M_{ij} T_j.$$
The quantities $R_i$ and $T_j$ are the contents of the detector-level distribution in bin $i$ and the contents of the particle-level distribution in bin $j$, respectively, while the factors $E_j$ and $P_i$ are the efficiency and the purity, which are estimated using the MC simulation. Migrations between detector- and particle-level phase spaces due to different values of $n^{\text{jet}}$ and $H_{T2}$ are taken into account in the unfolding procedure for event-shape observables. Due to the fine binning of the measured distributions, detector resolution is the primary cause of bin migrations between detector- and particle-level phase spaces, followed by the jet energy resolution, which leads to different values of $n^{\text{jet}}$ between the detector and particle levels. The efficiency $E_j$ is used to correct for events in the particle-level phase space which are not reconstructed at detector level. The binned efficiency is defined by the number of MC events that satisfy all the selection requirements at both the detector and particle levels and are generated and reconstructed in the same event-shape bin, with the same $n^{\text{jet}}$ and $H_{T2}$ range, divided by the number of events generated in the same event-shape, $n^{\text{jet}}$ and $H_{T2}$ bin. For low values of the event-shape variables the efficiency is typically close to 80%, while for large values it decreases to $\approx$ 40%. In addition, the binned efficiency tends to have lower values at higher $n^{\text{jet}}$ for the same event-shape bin. The purity $P_i$ is used to correct for events in the detector-level phase space that do not have a particle-level counterpart. The binned purity is defined as the number of events that satisfy selection requirements at particle and detector levels divided by those reconstructed in the same event-shape, $n^{\text{jet}}$ and $H_{T2}$ bin. Similarly to the bin-by-bin efficiency, the purity is close to 80% for low values of the event-shape variables and decreases to 40–50% for large values. The bin-by-bin purity tends to have lower values for the same event-shape bin at higher $n^{\text{jet}}$. Since the bin-by-bin purity and efficiency distributions are similar, and the migrations in the transfer matrices are mainly between neighbouring bins of the event-shape distributions, the impact of the unfolding on the shape of the distributions is modest. The results obtained by unfolding the data with PYTHIA are used to obtain the nominal differential cross sections, whereas SHERPA and HERWIG MC predictions are used to estimate the systematic uncertainty due to the model dependence, as discussed in section 7.

7 Experimental uncertainties

The dominant sources of systematic uncertainty in the measurements arise from imperfect knowledge of the jet energy scale and resolution and the modelling of the strong interaction. The systematic uncertainties are propagated through the unfolding via their impact on the transfer matrices. The inclusive cross section for events with at least two jets is recomputed for each systematic variation and used consistently in the analysis. As an example, the breakdown of the relative systematic uncertainties in the measurement of the normalised differential cross sections as functions of $\tau_{\perp}$ and $A$ is shown in figure 3.

A detailed description of the systematic uncertainties and their values for normalised event-shape observables is given below:

- Jet Energy Scale and Resolution: the jet energy scale (JES) and jet energy resolution (JER) uncertainties are estimated as described in ref. [59]. The JES is calibrated on the basis of the simulation, including in situ corrections obtained from data. The
JES uncertainties are estimated using a decorrelation scheme comprising a set of 44 independent components, which depend on the jet $p_T$ and $\eta$. The total JES uncertainty in the $p_T$ value of individual jets is $<2\%$ at $p_T = 100$ GeV with a mild dependence on $\eta$. The JER uncertainty is estimated using a decorrelation scheme involving 26 independent components. The effect of the total JER uncertainty is evaluated by smearing the energy of the jets in the MC simulation by about 1.5% at $p_T = 100$ GeV to about 0.5% for $p_T$ of several hundred GeV. In this measurement, the JES and JER uncertainties are propagated by varying the energy and $p_T$ of each jet by one standard deviation of each of the independent components. The total uncertainty in the normalised event-shape distributions varies from 1% in the lowest $n^{\text{jet}}$ bins to 7% for the highest $n^{\text{jet}}$, while it ranges from 7% to 14% on the fiducial cross sections. This is the dominant source of uncertainty for high $n^{\text{jet}}$. 

Figure 3. Breakdown of the systematic uncertainties as a function of $\tau_\perp$ (top) and $A$ (bottom) for selected regions of $H_{T2}$ and $n^{\text{jet}}$. 
• Jet Angular Resolution: the jet angular resolution (JAR) uncertainty is estimated conservatively by smearing the angular coordinates \((\eta, \phi)\) of the jets by the resolution in MC simulation. The \(\eta\) and \(\phi\) variations are done with the \(p_T\) component of the jets held constant. The value of the JAR uncertainty is below 0.5% for the normalised event-shape distributions in all regions of the phase space, while it ranges from 1% to 2% for the fiducial cross sections, as \(n^{\text{jet}}\) increases.

• Pile-up: the uncertainty from pile-up is evaluated by varying the pile-up reweighting procedure (see section 4) to cover the difference between the predicted inelastic cross section and the measured value \([60]\). The impact of this uncertainty is below 0.5% in all regions of the phase space, for both the event-shape and fiducial cross-section measurements. As a cross-check, the event-shape distributions were compared in different slices of \(\langle \mu \rangle\) and between different data-taking periods, yielding compatible results.

• Unfolding: the mismodelling of the data in the MC simulation is accounted for as an additional source of uncertainty. This is assessed by reweighting the particle-level distributions so that the detector-level event shapes predicted by the MC samples match those in the data. The modified detector-level distributions are then unfolded using the method described in section 6. The difference between the modified particle-level distribution and the nominal one is taken as the uncertainty. This uncertainty ranges from 0.2% to a few per cent with increasing \(n^{\text{jet}}\), depending on the observable under study.

• Modelling: the modelling uncertainty is estimated by comparing the unfolded distributions using PYTHIA, SHERPA and HERWIG. In order to not double count the effect of having different priors in the unfolding procedure (the so-called unfolding uncertainty), PYTHIA, SHERPA and HERWIG MC predictions are weighted to describe the data. These weighted MC samples are then used to perform the unfolding and the envelope of the differences between the estimated cross sections defines the systematic uncertainty. The value of this uncertainty for the normalised event-shape measurement increases with \(n^{\text{jet}}\) and \(H_{12}\) and varies between 1% and 4–5%, depending on the observable under study. For the fiducial cross-section measurement this uncertainty is below 5%.

• Luminosity: the uncertainty in the combined 2015–2018 integrated luminosity is 1.7% \([61]\), obtained using the LUCID-2 detector \([62]\) for the primary luminosity measurements. The measurements of event-shape observables are unaffected by this uncertainty, given that they are normalised to the inclusive dijet cross section, thus cancelling out the contribution of the luminosity.

• Dead-tile modules: a systematic uncertainty is derived to address the possible bias on the measurements due to the residual mismodeling of disabled portions of the tile calorimeter by the MC simulations. New differential cross-sections are derived by vetoing events in data and MC simulation where at least one of these non-operating modules is found within the selected jets. The difference between this result and the nominal one is taken as the uncertainty. The value of this uncertainty is below 1%.
in most regions of the phase space, although it can reach values up to 4% in some regions for the highest jet multiplicity bin.

The total systematic uncertainty is estimated by adding in quadrature the effects previously listed. In addition, the statistical uncertainty of the data and MC simulation is propagated to the differential cross sections through the unfolding procedure using pseudo-experiments in order to properly take into account the statistical correlations between bins of the event-shape variables, \( n_{\text{jet}} \) and \( H_{T2} \) ranges. Moreover, the pseudo-experiments are also used to estimate the statistical component of each systematic uncertainty. This statistical component is reduced using the Gaussian Kernel smoothing technique [63]. The values provided above are quoted after application of this procedure.

In general the modelling uncertainty tends to dominate at low values of \( n_{\text{jet}} \) while, at high \( n_{\text{jet}} \), the JES uncertainty dominates. The total systematic uncertainty is typically constant for the measured differential cross sections, but increases as a function of \( n_{\text{jet}} \) from \( \sim 1\% \) to \( \sim 6\% \). For the fiducial cross sections, the uncertainties increase from \( \sim 5\% \) to \( \sim 9\% \) as a function of \( n_{\text{jet}} \).

8 Results

The differential cross-section measurements are presented and compared with the MC predictions described in section 4. The cross section as a function of \( n_{\text{jet}} \) is shown in figure 4, while unfolded and normalised event-shape distributions are shown in figures 5–10. The full set of observables is presented in each \( H_{T2} \) and \( n_{\text{jet}} \) bin in which the measurement is performed. The ratio of the MC prediction to the yield in data in each bin is also shown.

Figure 4 shows the fiducial cross section as a function of \( n_{\text{jet}} \) in different \( H_{T2} \) ranges. The fiducial cross section is measured in the same phase-space regions as the differential measurements. The MC predictions are normalised to the measured integrated cross section in each \( H_{T2} \) range to compare the shape of these predictions to the data. The normalisation factors for HERWIG based on angle-ordered showers and SHERPA predictions increase as a function of \( H_{T2} \), whereas a very small dependence of these factors on \( H_{T2} \) is observed for HERWIG 7 based on dipole showers, MG5_aMC and PYTHIA predictions. The normalisation factors for LO accuracy MC predictions such as PYTHIA, MG5_aMC and SHERPA are expected to strongly depend on the MC tune. In particular, the PYTHIA prediction overestimates the inclusive dijet production cross section in the studied phase-space region by 30%, which can be attributed to the large value of \( \alpha_S \) included in the PYTHIA tune (see table 1), whereas the MG5_aMC prediction underestimates it by 35%. The SHERPA prediction gives an adequate description of the measured integrated cross sections. In addition, an excellent description of the inclusive dijet production cross section is found for the HERWIG 7 prediction based on dipole showers, whereas HERWIG 7 prediction based on angle-ordered showers underestimates it, at most by 9%. The PYTHIA prediction provides a good description of the shape of the differential cross section as a function of \( n_{\text{jet}} \). The HERWIG 7 prediction based on angle-ordered showers gives a good description of the fiducial cross sections as a function of \( n_{\text{jet}} \). SHERPA tends to overestimate
the cross sections for $n^{\text{jet}} > 4$. The \textsc{Herwig} 7 based on dipole showers and \textsc{MG5} \_\textsc{aMC} predictions, while giving a good description of the cross section at low $n^{\text{jet}}$, underestimate the measurements at high $n^{\text{jet}}$.

Figure 5 shows the normalised cross section as a function of $\tau_\perp$. For low $n^{\text{jet}}$, the MC simulations tend to underestimate the measurement in the intermediate region of $\tau_\perp$, with the exception of the \textsc{Herwig} 7 predictions. At high values of $\tau_\perp$, where the population of isotropic events is expected to be larger, all MC predictions underestimate the measurements. In particular, the largest deviation is found for the \textsc{Pythia} prediction where the high-$p_T$ third jet is less likely to be produced isotropically. The shape of the distributions tends to agree with data for larger $n^{\text{jet}}$. \textsc{Pythia} tends to overestimate the measurements at low values of $\tau_\perp$, whereas the \textsc{Herwig} 7 prediction based on dipole showers highly underestimates the measurements in such region. The behaviour of $\tau_\perp$ as a function of $H_{T2}$ indicates more isotropic events at low energies, with increasing alignment of jets with the thrust axis for higher energy scales. Figure 6 shows the normalised cross section as a function of $T_m$, with very similar conclusions.

Figure 7 shows the normalised cross section as a function of $S_\perp$. In line with the observations for $\tau_\perp$ and $T_m$, the results show that, for low $n^{\text{jet}}$, \textsc{Pythia} and \textsc{Sherpa} simulations predict fewer isotropic events than in data, while the \textsc{Herwig} 7 and \textsc{MG5} \_\textsc{aMC} predictions are closer to the measurements. In addition, while \textsc{Pythia} gives an adequate description in the intermediate region of $S_\perp$, it overestimates the measurements at low $S_\perp$. For larger $n^{\text{jet}}$, the description of the shape is improved in MC simulations, while a discrepancy in the normalisation is observed for different predictions.

Figure 8 shows the normalised cross section as a function of $A$. In this case, the \textsc{Herwig} 7 prediction with angle-ordered parton shower aligns with the rest of the MC simulations, predicting more planar events than data at low $n^{\text{jet}}$, while the \textsc{Herwig} 7 prediction with the dipole parton shower predicts higher cross sections at high $A$ for low $n^{\text{jet}}$.

The normalised cross sections for the quadratic observable $C$ are shown in figure 9. Here, larger spreads are shown in the lower tails of the distributions for low $n^{\text{jet}}$. \textsc{Pythia} and \textsc{Herwig} 7 with the dipole-based parton shower predict a smaller cross section than data in these regions, while the other MC predictions overestimate the measurements. As with other event-shape observables, a larger spread is found in the normalisations at high $n^{\text{jet}}$.

The cubic observable $D$ is presented in figure 10, with conclusions similar to those for $A$. The \textsc{Herwig} 7 prediction with dipole-based parton shower predicts a higher cross section than data at high values of $D$ for low $n^{\text{jet}}$, while the other MC predictions have the opposite behaviour. For higher $n^{\text{jet}}$, the description of the shape becomes more similar among different MC predictions, with differences observed in the normalisation of the measurements.

Theoretical uncertainties on the MC predictions are not included in the discussion of the results, since the ME for the current predictions has leading order accuracy in the description of inclusive three-jet cross sections. This makes the usual variations of the renormalisation and factorisation scales not reliable for the estimation of theoretical uncertainties. To identify the phase-space regions of the event-shape variables which are
more sensitive to the MC tune, variations of the parton shower parameters were examined in 
PYTHIA as detailed in section 4. The effect of these variations on event-shape observables typically increases with jet multiplicity. Varying the FSR energy-scale parameter by a factor of two leads to differences of up to 40\% at high values of the event-shape variables i.e. regions where the contribution of isotropic events is larger. On the other hand, varying the ISR energy-scale parameter typically leads to differences from 10 to 30\% at low values of the event shapes for high jet multiplicities. Finally, varying the non-singular terms of the splitting functions contributes a 5–10\% difference that is typically constant as a function of the event-shape variables.

In summary, none of the MC predictions investigated provide a good description of the data in all regions of the phase space. In general, at low \( n^{\text{jet}} \), PYTHIA and SHERPA predictions underestimate the measurements at high values of the event-shape distributions, i.e. events in data follow a more isotropic distribution of the energy flow than those from these two predictions. PYTHIA shows discrepancies of up to 80\%. This shows the limited ability of parton shower models to simulate hard and wide angle radiation. The addition of \( 2 \rightarrow 3 \) processes in the ME allows to improve the description of the measurements in such regions. While SHERPA can differ by up to 30\%, the HERWIG and MG5_aMC predictions show discrepancies of 10–20\%. Moreover, both HERWIG 7 predictions overestimate the central regions of \( \tau_L \) by up to 20\%, while differing in the description of the aplanarity: the angle-ordered parton shower gives rise to more planar events than in data, while the dipole-based parton shower overestimates the measurements at high values of \( A \) for \( n^{\text{jet}} = 3 \). In addition, the HERWIG 7 prediction with dipole-based parton shower underestimates the measurements at low values of the event-shape distributions, whereas the HERWIG 7 angle-ordered prediction gives a better description in these regions. Overall, the description of the measurements made by the HERWIG 7 prediction based on angle-ordered parton showers is better than the description by the dipole-based parton shower. This may be due to the fact that the HERWIG 7 dipole-based shower model is recently released, and the parameters have not been tuned as thoroughly as those of the more mature angle-ordered showers. Moreover, the MG5_aMC prediction which includes up to four final-state partons in the ME gives the best overall description of the shape of the measurements in the studied \( n^{\text{jet}} \) and \( H_{T2} \) bins. This shows the importance of including in the ME beyond LO terms to describe the dynamics of high-\( p_T \) multijet final states. At high \( n^{\text{jet}} \), all MC simulations tend to give a similar prediction for the shape of the distributions. However, the normalisation of the predictions shows a large spread between different MC simulations in each bin of the jet multiplicity. The SHERPA prediction gives an adequate description of the normalisation for \( n^{\text{jet}} \leq 4 \), although it overestimates the cross sections up to 30\% for high \( n^{\text{jet}} \). The MG5_aMC and the HERWIG 7 with dipole-based parton showers simulations predict cross sections up to 30\% lower for events with at least six jets. Finally, HERWIG 7 with angle-ordered parton shower and PYTHIA predictions give a reasonably good description of the normalisation of the differential cross sections for the studied jet multiplicity and \( H_{T2} \) bins.

All measurements can be found in Hepdata [64], including these measurements binned in inclusive jet multiplicity.
Figure 4. Fiducial cross section as a function of jet multiplicity. The MC predictions are normalised to the measured integrated cross section for \( n_{\text{jet}} \geq 2 \) in each \( H_{T2} \) bin using the factors indicated in parentheses. The right panels show the ratios of the MC distributions to the data distributions. The error bars show the total uncertainty (statistical and systematic added in quadrature) and the grey bands in the right panels show the systematic uncertainty.
Figure 5. Comparison between data and MC simulation as a function of the transverse thrust $\tau_\perp$ (see eq. 3.1) for different jet multiplicities and energy scales. For illustration purposes, the corresponding differential cross section for each jet multiplicity is multiplied by $10^2$ ($n_{\text{jet}} = 3$), $10^1$ ($n_{\text{jet}} = 4$), $10^0$ ($n_{\text{jet}} = 5$), $10^{-1}$ ($n_{\text{jet}} \geq 6$). The right panels show the ratios between the MC and the data distributions. The error bars show the total uncertainty (statistical and systematic added in quadrature) and the grey bands in the right panels show the systematic uncertainty.
Figure 6. Comparison between data and MC predictions as a function of the transverse minor $T_m$ (see eq. 3.1) for different jet multiplicities and energy scales. For illustration purposes, the corresponding differential cross section for each jet multiplicity is multiplied by $10^2$ ($n_{\text{jet}} = 3$), $10^4$ ($n_{\text{jet}} = 4$), $10^6$ ($n_{\text{jet}} = 5$), $10^{-1}$ ($n_{\text{jet}} \geq 6$). The right panels show the ratios between the MC and the data distributions. The error bars show the total uncertainty (statistical and systematic added in quadrature) and the grey bands in the right panels show the systematic uncertainty.
Figure 7. Comparison between data and MC predictions as a function of the transverse sphericity $S_\perp$ (see eq. 3.4) for different jet multiplicities and energy scales. For illustration purposes, the corresponding differential cross section for each jet multiplicity is multiplied by $10^2$ ($n_{\text{jet}} = 3$), $10^4$ ($n_{\text{jet}} = 4$), $10^6$ ($n_{\text{jet}} = 5$), $10^{-3}$ ($n_{\text{jet}} \geq 6$). The right panels show the ratios between the MC and the data distributions. The error bars show the total uncertainty (statistical and systematic added in quadrature) and the grey bands in the right panels show the systematic uncertainty.
Figure 8. Comparison between data and MC predictions as a function of the aplanarity $A$ (see eq. 3.3) for different jet multiplicities and energy scales. For illustration purposes, the corresponding differential cross section for each jet multiplicity is multiplied by $10^2 (n_{\text{jet}} = 3), 10^1 (n_{\text{jet}} = 4), 10^0 (n_{\text{jet}} = 5), 10^{-1} (n_{\text{jet}} \geq 6)$. The right panels show the ratios between the MC and the data distributions. The error bars show the total uncertainty (statistical and systematic added in quadrature) and the grey bands in the right panels show the systematic uncertainty.
Figure 9. Comparison between data and MC predictions as a function of $C$ (see eq. 3.5) for different jet multiplicities and energy scales. For illustration purposes, the corresponding differential cross section for each jet multiplicity is multiplied by $10^2$ ($n^{\text{jet}} = 3$), $10^4$ ($n^{\text{jet}} = 4$), $10^6$ ($n^{\text{jet}} = 5$), $10^{-1}$ ($n^{\text{jet}} \geq 6$). The right panels show the ratios between the MC and the data distributions. The error bars show the total uncertainty (statistical and systematic added in quadrature) and the grey bands in the right panels show the systematic uncertainty.
Figure 10. Comparison between data and MC predictions as a function of $D$ (see eq. 3.6) for different jet multiplicities and energy scales. For illustration purposes, the corresponding differential cross section for each jet multiplicity is multiplied by $10^2 (n_{\text{jet}} = 3)$, $10^1 (n_{\text{jet}} = 4)$, $10^0 (n_{\text{jet}} = 5)$, $10^{-1} (n_{\text{jet}} \geq 6)$. The right panels show the ratios between the MC and the data distributions. The error bars show the total uncertainty (statistical and systematic added in quadrature) and the grey bands in the right panels show the systematic uncertainty.
9 Summary and conclusions

A measurement of event shapes in multijet final states is presented. The data correspond to 139 fb$^{-1}$ of $pp$ collisions at $\sqrt{s} = 13$ TeV, recorded by the ATLAS detector at the LHC during the period 2015–2018.

The measurements are classified in bins of the jet multiplicity and energy scale, and six event-shape observables are measured and corrected to particle level through an unfolding procedure. The distributions of event-shape observables are normalised to the inclusive two-jet cross section to reduce correlated experimental uncertainties. The main uncertainties in the measurement are due to the jet calibration procedure and the MC model used in the unfolding, and the total uncertainty is at the percent level.

The measurements are compared with MC predictions with matrix elements calculated up to different orders in perturbative QCD, matched to different parton shower algorithms. All the predictions qualitatively describe the main features of the data, but none of them gives a good description of all distributions within the experimental uncertainties. The discrepancies between data and all the MC samples investigated show that further refinement of the current MC predictions is needed to describe the data in some regions, particularly at high jet multiplicities. Moreover, these discrepancies show that these data provide a powerful testing ground for the understanding of the strong interaction at high energies.

Acknowledgments

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 and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; SMRT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRT, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRN, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; JINR; MES of Russia and NRC KI, Russian Federation; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada, CRC and IVADO, Canada; Beijing Municipal Science & Technology Commission, China; COST, ERC, ERDF, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AVH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; La Caixa Banking Foundation, CERCA Pro-
gramme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafssons Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (U.K.) and BNL (U.S.A.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [65].

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References


