Measurement of the jet fragmentation function and transverse profile in proton-proton collisions at a center-of-mass energy of 7 TeV with the ATLAS detector

Aad, G.; et al., [Unknown]; Bentvelsen, S.; Colijn, A.P.; de Jong, P.; de Nooij, L.; Doxiadis, A.D.; Ferrari, P.; Garitaonandia, H.; Geerts, D.A.A.; Gosselink, M.; Kayl, M.S.; Koffeman, E.; Lee, H.; Linde, F.; Mechnich, J.; Mussche, I.; Ottersbach, J.P.; Tsiakiris, M.; van der Kraaij, E.; van der Leeuw, R.; van Kesteren, Z.; van Vulp, I.; Vermeulen, J.C.; Vreeswijk, M.

DOI
10.1140/epjc/s10052-011-1795-y

Publication date
2011

Document Version
Final published version

Published in
European Physical Journal C

Citation for published version (APA):
Aad, G., et al., U., Bentvelsen, S., Colijn, A. P., de Jong, P., de Nooij, L., Doxiadis, A. D., Ferrari, P., Garitaonandia, H., Geerts, D. A. A., Gosselink, M., Kayl, M. S., Koffeman, E., Lee, H., Linde, F., Mechnich, J., Mussche, I., Ottersbach, J. P., Tsiakiris, M., ... Vreeswijk, M. (2011). Measurement of the jet fragmentation function and transverse profile in proton-proton collisions at a center-of-mass energy of 7 TeV with the ATLAS detector. European Physical Journal C, 71(11), [1795]. https://doi.org/10.1140/epjc/s10052-011-1795-y
Measurement of the jet fragmentation function and transverse profile in proton–proton collisions at a center-of-mass energy of 7 TeV with the ATLAS detector

The ATLAS Collaboration
CERN, 1211 Geneva 23, Switzerland

Received: 27 September 2011 / Revised: 26 October 2011 / Published online: 30 November 2011
© CERN for the benefit of the ATLAS collaboration 2011. This article is published with open access at Springerlink.com

Abstract The jet fragmentation function and transverse profile for jets with 25 GeV < \( p_{T\,\text{jet}} \) < 500 GeV and |\( \eta_{\text{jet}} \)| < 1.2 produced in proton–proton collisions with a center-of-mass energy of 7 TeV are presented. The measurement is performed using data with an integrated luminosity of 36 pb\(^{-1}\). Jets are reconstructed and their momentum measured using calorimetric information. The momenta of the charged particle constituents are measured using the tracking system. The distributions corrected for detector effects are compared with various Monte Carlo event generators and generator tunes. Several of these choices show good agreement with the measured fragmentation function. None of these choices reproduce both the transverse profile and fragmentation function over the full kinematic range of the measurement.

1 Introduction and overview

This paper presents measurements of jet properties in proton–proton (\( pp \)) collisions at a center of mass energy of 7 TeV at the CERN LHC using the ATLAS detector. Jets are identified and their momenta measured using the calorimeters. Charged particles measured by the tracking system are then associated with these jets using a geometric definition. The structure of the jets is studied using these associated particles.

Jets produced at large transverse momentum in proton–proton collisions arise from the scattering of proton constituents leading to outgoing partons (quarks and gluons) with large transverse momenta. These manifest themselves as jets of hadrons via a “fragmentation process”. While the scattering of the proton constituents is well described by perturbative QCD and leads, at lowest order, to final states of \( gg \), \( gq \), and \( qq \), the fragmentation process is more complex. First, fragmentation must connect the outgoing partons with the rest of the event as the jet consists of colourless hadrons while the initiating parton carries colour. Second, the process involves the production of hadrons and takes place at an energy scale where the QCD coupling constant is large and perturbation theory cannot be used. Fragmentation is therefore described using a QCD-motivated model with parameters that must be determined from experiment. The fragmentation function \( D_h^i(z, Q) \) is defined as the probability that a hadron of type \( h \) carries longitudinal momentum fraction \( z \) of the momentum \( p_i \) of a parton of type \( i \)

\[
z \equiv \frac{p_i \cdot p_h}{|p_i|^2}.
\]

\( D(z, Q) \) depends on \( z \) and on the scale \( Q \) of the hard scattering process which produced the parton. While the value of \( D_h^i(z, Q) \) cannot be calculated in perturbative QCD, the variation with \( Q \) can be predicted provided \( Q \) is sufficiently large [1–6].

In this paper a quantity related to \( D_h^i(z, Q) \) is measured. After jets have been reconstructed, the data are binned for fixed ranges of jet transverse momenta (\( p_{T\,\text{jet}} \)), each bin containing \( N_{\text{jet}} \) jets; \( z \) is then determined for each charged particle associated with the jet

\[
z = \frac{p_{\text{jet}} \cdot p_{\text{ch}}}{|p_{\text{jet}}|^2},
\]

where \( p_{\text{jet}} \) is the momentum of the reconstructed jet and \( p_{\text{ch}} \) the momentum of the charged particle. The following quantity is measured

\[
F(z, p_{T\,\text{jet}}) \equiv \frac{1}{N_{\text{jet}}} \frac{dN_{\text{ch}}}{dz},
\]

where \( N_{\text{ch}} \) is the number of charged particles in the jet. \( F(z, p_{T\,\text{jet}}) \) is a sum over \( D_h^i(z, Q) \) weighted by the rate at which each parton species (\( i \)) is produced from the hard scattering process. As particle identification is not used,
$h$ is summed over all charged hadrons. The hard scattering scale $Q$ is of the same order of magnitude as $p_{T\text{jet}}$. At small $p_{T\text{jet}}$, gluon jets dominate due to the larger gluon parton densities in the proton and larger scattering rates for $gg \to gg$. In the pseudorapidity range used for jets in this analysis ($\mid \eta_{\text{jet}} \mid < 1.2$) the fraction of jets originating from a hard scattering that produces a gluon falls from 80% for $p_{T\text{jet}} \sim 25$ GeV to 50% for $p_{T\text{jet}} \sim 300$ GeV according to the PYTHIA [7] event generator.

The jets measured experimentally also contain particles produced from the hadronization of the beam remnants (the “underlying event”). It should be emphasized that because colour fields connect all the strongly interacting partons in the $pp$ event, no unambiguous assignment of particles to the hard scattering parton or underlying event is possible. The integral of $F(z, p_{T\text{jet}})$ with respect to $z$ corresponds to the multiplicity of charged particles within the jet. A clear summary of fragmentation phenomenology is provided in [8] (Sect. 17) whose notation is followed here.

The derivation of $D^k_j(z, Q)$ from $F(z, p_{T\text{jet}})$ is beyond the scope of this paper, but comparisons of $F(z, p_{T\text{jet}})$ with the predictions of several Monte Carlo (MC) generators will be made. Different features of the Monte Carlo models are probed by these studies. At low values of $p_{T\text{jet}}$, the comparisons are most sensitive to the non-perturbative models of fragmentation, the connection of the partons to the remainder of the event and to the accretion of particles from the underlying event into the jet. As $p_{T\text{jet}}$ rises, the impact of these effects is diluted and, if all the Monte Carlo models implemented perturbative QCD in the same way, $F(z, p_{T\text{jet}})$ would become similar. In particular the increase of the total particle multiplicity with the hard scattering energy, here $p_{T\text{jet}}$, is predicted by perturbative QCD [9].

Two other related quantities that describe the transverse shape of the jets are also studied here. The variable $p^{\text{rel}}_T$ is the momentum of charged particles in a jet transverse to that jet’s axis:

$$p^{\text{rel}}_T = \frac{|p_{\text{ch}} \times p_{\text{jet}}|}{|p_{\text{jet}}|}. \quad (4)$$

The following distribution is measured

$$f(p^{\text{rel}}_T, p_{T\text{jet}}) \equiv \frac{1}{N_{\text{jet}}} \frac{dN_{\text{ch}}}{dp^{\text{rel}}_T}. \quad (5)$$

Finally, the density of charged particles in $y-\phi$ space, $\rho_{\text{ch}}(r, p_{T\text{jet}})$, is measured as a function of the angular distance $r$ of charged particles from the axis of the jet that contains them, where $r$ is given by

$$r = \Delta R(\text{ch. jet}) = \sqrt{(\phi_{\text{ch}} - \phi_{\text{jet}})^2 + (y_{\text{ch}} - y_{\text{jet}})^2}. \quad (6)$$

Thus $\rho_{\text{ch}}(r, p_{T\text{jet}})$ is given by:

$$\rho_{\text{ch}}(r, p_{T\text{jet}}) \equiv \frac{1}{N_{\text{jet}} 2\pi r dr} \frac{dN_{\text{ch}}}{dr}. \quad (7)$$

As in the case of the longitudinal variables, a comparison of these transverse quantities with Monte Carlo generators is sensitive to many of their features. The non-perturbative hadronization processes produce particles that have limited transverse momentum with respect to the parton direction. The mean value of this transverse momentum is of order a few hundred MeV, the scale where the QCD coupling constant becomes non-perturbative. At low $p_{T\text{jet}}$ this effect dominates. If there were no other contributions, $p^{\text{rel}}_T$ would remain constant with increasing $p_{T\text{jet}}$. Therefore more of the energy would be concentrated in the core of the jet as $p_{T\text{jet}}$ increases and the jets would become narrower. However, as $p_{T\text{jet}}$ increases contributions from processes controlled by perturbative QCD radiation become more important, contributing to jet broadening and causing the mean value of $p^{\text{rel}}_T$ to rise slowly (approximately logarithmically).

The phenomena described above are incorporated in all the Monte Carlo generators used to describe jet production in $pp$ collisions, although there are significant differences in how these effects are implemented. For example, PYTHIA describes non-perturbative hadronization using a string model while HERWIG [10] uses a cluster model. In PYTHIA, coherent colour effects are described partly by string fragmentation. These effects are also produced in HERWIG and PYTHIA from gluon radiation. Treatments of the proton remnants are also described using different phenomenological approaches. For both generators, the implementations require that a number of input parameters be tuned to the data. The results presented in this paper will test whether these Monte Carlo models and their current input parameters adequately describe jets produced at the LHC. As the results are presented in bins of $p_{T\text{jet}}$, the explicit dependence on $p_{T\text{jet}}$ in the variables defined in (3), (5) and (7) is often suppressed in the following.

The measurement is performed using data with an integrated luminosity of 36 pb$^{-1}$ recorded in 2010 with the ATLAS detector at the LHC at a center-of-mass energy of 7 TeV. The measurement covers a kinematic range of 25 GeV $< p_{T\text{jet}} < 500$ GeV and $|\eta_{\text{jet}}| < 1.2$. Events are triggered using a minimum bias trigger and a combination of calorimeter jet triggers. A complementary ATLAS analysis [11] studying the jet fragmentation function and transverse profile of jets reconstructed from charged particle tracks using a total integrated luminosity of 800 µb$^{-1}$ has been com-

---

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 coinciding with the axis of the beam pipe. The X-axis points from the IP to the centre of the LHC ring, and the Y-axis points upward. Cylindrical coordinates $(r, \phi)$ are used in the transverse plane, $\phi$ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. The rapidity $y$ for a track or jet is defined by $y = 0.5 \ln \left( (E + p_T) / (E - p_T) \right)$ where $E$ denotes the energy and $p_T$ is the momentum along the beam direction. For tracks, the energy is calculated assuming the particle is a pion.
pleted. It explores the properties of jets at lower transverse momentum than those typically studied in this paper.

Previous measurement of jet fragmentation functions have been made in $e^+e^-$ collisions [12–15], in $p\bar{p}$ collisions [16, 17] and in $ep$ collisions [18, 19].

This paper is organized as follows. The ATLAS detector is described briefly in Sect. 2. The Monte Carlo generator samples are discussed in Sect. 3. The event and object selections are described in Sect. 4. Section 5 contains a description of the analysis. In Sect. 6 the treatment of systematic uncertainties is presented. Results and conclusions are shown in Sects. 7 and 8.

2 The ATLAS detector

The ATLAS detector is described in detail in [20]. The subsystems relevant for this analysis are the inner detector (ID), the calorimeter and the trigger. The ID is used to measure the momentum of charged particles. It consists of three subsystems: a pixel detector, a silicon strip tracker (SCT) and a transition radiation straw tube tracker (TRT). These detectors are located inside a solenoid that provides a 2 T axial field. The ID has full coverage in the azimuthal angle $\phi$ and over the pseudorapidity range $0 < \eta < 2.5$.

The electromagnetic calorimeters use liquid argon as the active detector medium. They consist of accordion-shaped electrodes and lead absorbers and cover the pseudorapidity range $|\eta| < 3.2$. The technology used for the hadronic calorimeters varies with $\eta$. In the barrel region ($|\eta| < 1.7$) the detector is made of scintillating tiles with steel radiator. In the endcap region ($1.5 < |\eta| < 3.2$) the detector uses liquid argon and copper. A forward calorimeter consisting of liquid argon and tungsten/copper absorbers serves as both electromagnetic and hadronic calorimeter at large pseudorapidity and extends the coverage to $|\eta| < 4.9$.

The calorimeters are calibrated at the electromagnetic scale which correctly reconstructs the energy deposited by electrons and photons. The calorimeters are not compensating and the response of hadrons is lower than that of electrons ($e/h > 1$). Some fraction of the hadronic energy can also be deposited in the material in front of and in-between calorimeters. The response for hadronic jets [21] is $\sim 50\%$ of the true energy for $p_{T\text{jet}} = 20$ GeV and $|\eta_{\text{jet}}| < 0.8$ and rises both with $p_{T\text{jet}}$ and $\eta_{\text{jet}}$. For $|\eta_{\text{jet}}| < 0.8$, the response at $p_{T\text{jet}} = 1$ TeV is $\sim 80\%$.

The ATLAS trigger consists of three levels of event selection: Level-1 (L1), Level-2 (L2), and Event Filter. The L2 and event filter together form the High-Level Trigger (HLT). The L1 trigger is implemented using custom-made electronics, while the HLT is based on fast data reconstruction online algorithms running on commercially available computers and networking systems. The triggers relevant for this analysis are the L1 minimum bias triggers (MBTS) and the L1 and HLT calorimeter triggers. The minimum bias trigger is based on signals from 32 scintillation counters located at pseudorapidities $2.09 < |\eta| < 3.84$. Because non-diffractive events fire the MBTS with high efficiency and negligible bias, this trigger can be used to study jets with low $p_{T\text{jet}}$. However, MBTS triggers were highly prescaled at large instantaneous luminosities, making them unsuitable for studies of high $p_{T}$ jets that are produced at low rate. A series of single jet inclusive triggers with different jet $E_{T}$ thresholds and prescales were deployed to ensure that significant data samples were taken over the full range of $p_{T\text{jet}}$ [22].

3 Monte Carlo samples

Several Monte Carlo samples are used in this analysis. Some samples were processed with the ATLAS full detector simulation [23] which is based on the GEANT4 toolkit [24]. The simulated events are then passed through the same reconstruction software as the data. These are used to model the response of the detector and to correct the data for experimental effects. The baseline Monte Carlo sample used to determine these corrections is produced using PYTHIA [7] 6.421 with the ATLAS tune AMBT1 which uses the MRST2007LO* PDFs [25] and was derived using the measured properties of minimum bias events [26]. Several other fully simulated samples are used to assess systematic uncertainties: PYTHIA using the PERUGIA2010 tune [27] (CTEQ5L PDFs [28]); Herwig 6.5 [10] using Jimmy 3.41 [29] and Herwig++ 2.4.2 [30] (MRST2007LO* PDFs).

Additional Monte Carlo generator samples are used to compare with the final corrected data: PYTHIA6.421 with the ATLAS MC09 tune [31] (MRST2007LO* PDFs), Herwig++ 2.5.1 [32] (MRST2007LO* PDFs), Sherpa [33] (CTEQ6L [34] PDFs) and PYTHIA8 (8.105) [35] (MRST2007LO* PDFs).

4 Reconstruction and event selection

Events are required to have at least one primary vertex reconstructed using ID tracks. If the event has multiple primary vertices, the vertex with the largest $\sum(p_{T\text{track}})^{2}$ is tagged as the hard-scattering vertex.

Jets are reconstructed using the infrared- and collinear-safe anti-$k_{t}$ algorithm [36] with radius parameter $R_{t} = 0.6$ using the FastJet package [37]. The detector input is based on topological clusters [38]. A topological cluster is defined to have an energy equal to the energy sum of all the included calorimeter cells, zero mass and a reconstructed direction calculated from the weighted averages of the pseudorapidities and azimuthal angles of the constituent cells. The
weight used is the absolute cell energy and the positions of
the cells are relative to the nominal ATLAS coordinate sys-
tem. The energy of these clusters is measured at the electro-
magnetic scale, which provides the appropriate calibration
for electrons and photons. A \( p_T^{\text{jet}} \) and \( n_{\text{jett}} \) dependent cali-
tration is then applied to each jet [21]. These calibrations are
based on comparing the response from simulated calorim-
eter jets to that of jets reconstructed using generator particles
and matched to the reconstructed jets in \( \eta-\phi \) space. The \( \eta-\phi \)
position of the jet (and hence its momentum) is corrected to
account for the fact that the primary vertex of the interaction
is not at the geometric centre of the detector. Quality crite-
ria are applied to ensure that jets are not produced by noisy
calorimeter cells, and to avoid problematic detector regions.

The jet energy is corrected for the presence of additional
pp interactions in the same bunch crossing using correction
constants measured in-situ that depend on the number of re-
constructed primary vertices.

Jets are required to have \(|\eta_{\text{jet}}| < 1.2\). For events selected
with the MBTS trigger, jets are required to pass a minimum cut
of \( p_T^{\text{jet}} > 20 \text{ GeV} \). For events selected using jet triggers,
a trigger-dependent minimum \( p_T^{\text{jet}} \) threshold is imposed on
jets used in the final measurements to ensure a jet trigger
efficiency larger than 99%.

Tracks are selected using the following cuts:

\[
p_{T\text{track}} > 0.5 \text{ GeV}, \quad N_{\text{pixel}} \geq 1, \quad N_{\text{SCT}} \geq 6, \quad |d_0| < 1.5 \text{ mm}, \quad |z_0 \sin \theta| < 1.5 \text{ mm},
\]

where \( N_{\text{pixel}} \) and \( N_{\text{SCT}} \) are the number of hits from the pixel
and SCT detectors, respectively, that are associated with the
track and \( d_0 \) and \( z_0 \) are the transverse and longitudinal impact
parameters measured with respect to the hard-scattering
vertex.

Tracks are associated with jets using a simple geometric algorithm. If the distance in \( \eta-\phi \) between the track and the jet is less than the radius parameter used in the jet recon-
struction (\( R_c = 0.6 \)), the tracks are considered to belong
to the jet. Track parameters are evaluated at the perigee to
the primary vertex and are \textit{not} extrapolated to the calorim-
eter. This simple association algorithm facilitates comparison
with particles from the event generator whose parameters
correspond to those measured at the primary vertex.

5 Analysis

The results presented here are obtained using four mea-
sured distributions: the jet transverse momentum spec-
trum, \( dN_{\text{jet}}(p_T^{\text{jet}})/dp_T^{\text{jet}} \), and three differential distributions
of the number of charged tracks, \( dN_{\text{tracks}}(z, p_T^{\text{jet}})/dz \),
\( dN_{\text{tracks}}(p_T^{\text{rel}}, p_T^{\text{jet}})/dp_T^{\text{rel}} \) and \( dN_{\text{tracks}}(r, p_T^{\text{jet}})/dr \). To fa-
cilitate comparison with the predictions of Monte Carlo
event generators, these distributions are corrected for de-
tector acceptance, reconstruction efficiency and migration
due to track and jet momentum resolution effects. This
correction procedure is called unfolding. The distributions
\( F(z, p_T^{\text{jet}}) \), \( f(p_T^{\text{rel}}, p_T^{\text{jet}}) \) and \( \rho_{\text{ch}}(r, p_T^{\text{jet}}) \) are obtained
from the charged particle differential distributions by nor-
malizing the distribution for each \( p_T^{\text{jet}} \) range to the value
of \( N_{\text{jet}}(p_T^{\text{jet}}) \) obtained from the unfolding of the jet trans-
verse momentum spectrum. This paper presents results for
\( p_T^{\text{jet}} > 25 \text{ GeV} \); however, to decrease the systematic uncer-
tainty associated with the modeling of the \( p_T^{\text{jet}} \) spectrum,
jets with \( 20 \text{ GeV} < p_T^{\text{jet}} < 25 \text{ GeV} \) are also used in the
unfolding.

A Bayesian iterative unfolding method [39] implemented
in the RooUnfold [40] software package is used. This pro-
cedure takes as its input the measured distributions and a re-
sponse matrix obtained from simulated data that provides a
mapping between reconstructed objects and those obtained
directly from the event generator. This response matrix is
not unitary because in mapping from generator to recon-
struction some events and objects are lost due to inefficien-
cies and some are gained due to misreconstruction or mi-
gration of truth objects from outside the fiducial acceptance
into the reconstructed observables. It is therefore not pos-
sible to obtain the unfolded distributions by inverting the
response matrix and applying it to the measured data. In-
stead, an assumed truth distribution (the "prior") is selected,
the response matrix is applied and the resulting trial recon-
struction set is compared to the observed reconstruction set.
A new prior is then constructed from the old prior and the
difference between the trial and the observed distributions.
The procedure can iterated until this difference becomes
small. Monte Carlo based studies of the performance of
the procedure demonstrate that in this analysis no iteration
is necessary. The initial truth prior is taken to be the predic-
tion of the baseline Monte Carlo generator. Systematic un-
certainties associated with this choice and with the modeling
of the response matrix are discussed in Sect. 6.

6 Systematic uncertainties

The following sources of systematic uncertainties are con-
idered:

1. The jet energy scale (JES) and resolution (JER) uncer-
tainties which affect the measurement of the number of
jets in a given \( p_T^{\text{jet}} \) bin and consequently the measured
value of \( z \).
2. The track reconstruction efficiency and momentum re-
construction uncertainties which affect the number of
tracks in each \( z, p_T^{\text{rel}} \) and \( N_{\text{ch}}(r) \) bin.
3. The uncertainty in the response matrix which is derived
using a particular Monte Carlo sample and depends on
the details of the event generator.
4. Potential bias due to the failure of the unfolding procedure to converge to the correct value.

These systematic uncertainties are addressed using Monte Carlo methods.

The first two systematic uncertainties, potential bias due to incorrect Monte Carlo modeling of the JES and/or JER and potential bias due to mismodeling by the simulation of the track reconstruction efficiency and/or resolution, are studied by modifying the detector response in simulated data. These modified Monte Carlo events are then unfolded and compared to the baseline. The systematic uncertainty on the JES is studied by varying the jet energy response by its uncertainty. The JES uncertainty varies from 4.6% at $p_{T,jet} = 20$ GeV to 2.5% at $p_{T,jet} = 500$ GeV [21]. Systematic uncertainties on the JER are studied by broadening the jet energy resolution with an additional $\eta_{jet}$ and $p_{T,jet}$ dependent Gaussian term. The uncertainty on the JER is below 14% for the full $p_{T,jet}$ and $\eta_{jet}$ range used in this analysis [41].

The uncertainty on the tracking efficiency is studied by randomly removing a fraction of the tracks in the simulated data. Uncertainties on the tracking efficiency are $\eta$-dependent and vary between 2% and 3% for the relevant range of $\eta_{track}$ [42], dominated by the accuracy of the description of the detector material in the simulation. In addition, there can be a loss of tracking efficiency in the core of jets at high $p_{T,jet}$ due to a single pixel hit receiving contributions from more than track. Studies of such hit sharing show that the simulation and data agree well and that the resulting systematic uncertainty is negligible for $p_{T,jet} < 500$ GeV. Uncertainties on the track momentum resolution are parametrized as an additional $\eta$-dependent broadening of the resolution in curvature with values that vary from 0.0004 GeV$^{-1}$ to 0.0009 GeV$^{-1}$ [43].

While the studies described above account for systematic uncertainties associated with the accuracy of the detector simulation, they do not account for the fact that the response matrix itself depends on the fragmentation properties of the jets and hence on the physics description in the event generator. Because the response of the calorimeter to hadrons depends on the hadron momentum [44], the JES depends at the few per cent level on the momentum spectrum of particles within the jet. Because the probability that a track will share hits in the ID with another track is dependent upon the local density of particles within the jet, the tracking resolution depends weakly on the transverse profile of particles within the jet. These effects have been studied by unfolding fully simulated Monte Carlo samples created from PERUGIA2010, Herwig 6.5 (with Jimmy 3.41) and Herwig++ using the baseline response matrix obtained with PYTHIA AMBT1. Differences between the unfolded results for each tune and the true distributions obtained from that same tune are studied as a function of $z$, $p_{T,rel}$ and $N_{ch}(r)$ for each bin in true $p_{T,jet}$ and used to assess the systematic uncertainty.

Potential bias in the unfolding procedure itself is studied by creating 1000 pseudo-experiments where the “data” are drawn from the baseline fully simulated Monte Carlo samples via a bootstrap method [45] and unfolding these “data” using the standard procedure. The mean results obtained from these samples show negligible bias and have a spread that is consistent with the reported statistical uncertainties. The systematic uncertainty due to the unfolding procedure is thus deemed to be negligible in comparison to the other uncertainties.

The resulting systematic uncertainties on $F(z, p_{T,jet})$, $f(p_{T,rel}, p_{T,jet})$ and $\rho_{ch}(r, p_{T,jet})$ for the $25$ GeV $< p_{T,jet} < 40$ GeV (left) and $400$ GeV $< p_{T,jet} < 500$ GeV (right) are shown in Figs. 1, 2, 3. For $F(z, p_{T,jet})$, uncertainties
Fig. 2 Systematic uncertainty in $f(p_{T}^{rel}, p_{Tjet})$ from uncertainties in the jet energy scale and resolution, the track reconstruction efficiency and momentum resolution and the response matrix for 25 GeV < $p_{Tjet}$ < 40 GeV (left) and 400 GeV < $p_{Tjet}$ < 500 GeV (right). The total uncertainty from the combination is also shown.

Fig. 3 Systematic uncertainty in $\rho_{ch}(r, p_{Tjet})$ from uncertainties in the jet energy scale and resolution, the track reconstruction efficiency and momentum resolution and the response matrix for 25 GeV < $p_{Tjet}$ < 40 GeV (left) and 400 GeV < $p_{Tjet}$ < 500 GeV (right). The total uncertainty from the combination is also shown.

7 Results

This section presents comparisons of acceptance-corrected, unfolded data to the predictions of several Monte Carlo generators. The gray band on all the figures indicates the total uncertainty which is dominated by the systematic uncertainty. Figure 4 shows distributions of $F(z)$ in two bins of $p_{Tjet}$. Figure 5 shows distributions of $F(z)$ in all bins of $p_{Tjet}$ compared to AMBT1 Monte Carlo. Comparisons of the data and the Monte Carlo samples are shown in Fig. 6. All the PYTHIA 6 tunings show good agreement with the data. Herwig+Jimmy disagrees with the data at large $z$ for $p_{Tjet}$ > 200 GeV. Herwig++ 2.5.1 is below the data at low $z$ for $p_{Tjet}$ > 100 GeV while Herwig++ 2.4.2 has too many particles at low $z$ for $p_{Tjet}$ < 100 GeV. PYTHIA8 and Sherpa provide a poor description of the data.

Figure 7 (left) shows the distribution of $\langle z \rangle$ for the data and for a selection of Monte Carlo samples as a function of $p_{Tjet}$. A comparison with the Monte Carlo generators shows that the AMBT1 and MC09 PYTHIA and PERUGIA2010 datasets show good agreement with the data over
Fig. 4 Distributions of $F(z)$ for 25 GeV < $p_T^{\text{jet}}$ < 40 GeV (left) and 400 GeV < $p_T^{\text{jet}}$ < 500 GeV (right). The gray band indicates the total uncertainty.

Fig. 5 Distributions of $F(z)$ in bins of $p_T^{\text{jet}}$. The circles show unfolded data and the lines are the predictions from AMBT1 PYTHIA.

The charged particle multiplicity as a function of $p_T^{\text{jet}}$ is shown in Fig. 7 (right). The PYTHIA 6 tunes show reasonable agreement, with AMBT1 being higher than the others. Herwig+Jimmy has slightly too few particles for $p_T^{\text{jet}}$ > 200 GeV. Herwig++ 2.4.2 (2.5.1) has too many (few) particles for $p_T^{\text{jet}}$ < 200 (> 300) GeV. Sherpa describes the data well while PYTHIA8 has ~8\% too many particles at all $p_T^{\text{jet}}$.

The transverse profile of the jets is described by the $\rho_{\text{ch}}(r)$ and $f(p_{T \text{rel}})$ distributions. Figure 8 shows the distribution of $\rho_{\text{ch}}(r)$ in two bins of $p_T^{\text{jet}}$. The sharp decrease in population in the last bin is a feature of the jet algorithm, which tends to incorporate particles close to the radius parameter into the jet. The effect is also seen in [11] (Fig. 6) where distributions for two radius parameters are shown. Figure 9 shows the distribution of $f(p_{T \text{rel}})$ in the same two $p_T^{\text{jet}}$ bins. Figures 10 and 11 show distributions of $\rho_{\text{ch}}(r)$ and $f(p_{T \text{rel}})$, respectively, in all $p_T^{\text{jet}}$ bins together with the predictions of the AMBT1 Monte Carlo. Comparisons of $\rho_{\text{ch}}(r)$ for all data and Monte Carlo are shown in Fig. 12. Sherpa, Herwig++ 2.4.2 and PYTHIA8 disagree significantly with the data over the full range of the measurement. PYTHIA8 is consistent with the data only over a very restricted range of $p_T^{\text{jet}}$ around 80 GeV. Herwig++ 2.5.1 shows good agreement except at small $r$ and for $p_T^{\text{jet}}$ > 200 GeV. Herwig+Jimmy is consistent with the data only for $p_T^{\text{jet}}$ > 160 GeV. All the PYTHIA6 tunings except AMBT1 agree; AMBT1 shows disagreement for $p_T^{\text{jet}}$ > 200 GeV. Comparison of $f(p_{T \text{rel}})$ for all data and
Fig. 6 The ratio of $F(z)$ predicted by various Monte Carlo generators to that measured. The gray band indicates the combined statistical and systematic uncertainties. Monte Carlos are shown in Fig. 13. None of the generators agree with the data within the systematic uncertainties.

The mean value of $p_{\text{T}}^{\text{rel}}$ as a function of $p_{\text{T}}^{\text{jet}}$ is shown in Fig. 14. Herwig++ 2.5.1 has much too large a value of $\langle p_{\text{T}}^{\text{rel}} \rangle$ for $p_{\text{T}}^{\text{jet}} > 100$ GeV and 2.4.2 has too small a value for $p_{\text{T}}^{\text{jet}} < 80$ GeV. AMBT1 has too small a value at all $p_{\text{T}}^{\text{jet}}$. Herwig++ Jimmy has too large a value for $p_{\text{T}}^{\text{jet}} > 200$ GeV. Agreement of the remaining Monte Carlos is quite good.

8 Conclusion

A measurement of the jet fragmentation properties for charged particles in proton–proton collisions at a center-of-mass energy of 7 TeV is presented. The dataset recorded with the ATLAS detector at the LHC in 2010 with an integrated luminosity of 36 pb$^{-1}$ is used. Systematic uncertainties for the fragmentation function which describes how the jet momentum is distributed amongst its constituents vary between approximately 4% and 40% depending on $z$ and $p_{\text{T}}^{\text{jet}}$. The uncertainties increase strongly with $z$ and are largest at small $p_{\text{T}}^{\text{jet}}$. The measurements of the distributions $\rho_{\text{ch}}(r, p_{\text{T}}^{\text{jet}})$ and $f(p_{\text{T}}^{\text{rel}}, p_{\text{T}}^{\text{jet}})$ which describe the shape of jets transverse to the jet direction have uncertainties that fall as $p_{\text{T}}^{\text{jet}}$ increases, increase at large values of $p_{\text{T}}^{\text{rel}}$ and are almost independent of $r$. They are less than 5% except in the lowest $p_{\text{T}}^{\text{jet}}$ range and for $p_{\text{T}}^{\text{rel}} > 1$ GeV.
The measurements are sensitive to several properties of QCD as implemented in and modeled by Monte Carlo event generators. The additional QCD radiation present as $p_{T,jet}$ increases is modeled by perturbative QCD and results in a growth of the particle multiplicity. This growth is very well modeled by all the Monte Carlo generators used here. Two
Fig. 9  Distributions of $f(p_T^{\text{rel}})$ for $25 \text{ GeV} < p_T^{\text{jet}} < 40 \text{ GeV}$ (left) and $400 \text{ GeV} < p_T^{\text{jet}} < 500 \text{ GeV}$ (right). The gray band indicates the total uncertainty.

Fig. 10  Distributions of $\rho_{ch}(r)$. The circles show unfolded data. The lines are the predictions from AMBT1 PYTHIA.

Fig. 11  Distributions of $f(p_T^{\text{rel}})$. The circles show unfolded data. The lines are the predictions from AMBT1 PYTHIA.
The ratio of $\rho_{ch}(r)$ predicted by various Monte Carlo generators to that measured. The gray band indicates the combined statistical and systematic uncertainties.

Other effects that cannot be described by perturbative QCD impact the measured distributions. The hadronization of partons produced in a QCD radiative shower into the observed hadrons must be modeled in the Monte Carlo generators and is described by a large number of parameters which are tuned to agree with data. Particles produced from remnants of the initial protons (underlying event) can be incorporated into jets whose constituents mainly come from the hard scattering, so the measured jet properties can be sensitive to this modeling.

The measured fragmentation functions agree well with the AMBT1 PYTHIA and PERUGIA2010 Monte Carlo predictions within statistical and systematic uncertainties. Other tunes and generators show less good agreement indicating that the non-perturbative physics is not adequately modeled in these cases. Measurements of the transverse distributions $f(p_T^{rel}, p_{Tjet})$ and $\rho_{ch}(r, p_{Tjet})$ are also presented.

For the $p_T^{rel}$ distribution, none of the generators agree with data within systematic uncertainties over the full kinematic range. For the $\rho_{ch}(r, p_{Tjet})$ distribution, Herwig+Jimmy, PYTHIA MC09 and PERUGIA2010 are in reasonable agreement with the data.

In summary, none of the Monte Carlo generators studied provide a good description of all the data. The measurements presented here provide valuable inputs to constrain future improvements in Monte Carlo modeling of fragmentation. The full results are available in the HEPDATA database [46], and a Rivet [47] module for the analysis is also available.

Acknowledgements We honour the memory of our young colleague Christoph Ruwiedel, who was closely involved in the work described here and died shortly before its completion. 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.
Fig. 13 The ratio of \( f(\rho_{\text{rel}}) \) predicted by various Monte Carlo generators to that measured. The gray band indicates the combined statistical and systematic uncertainties.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF, Austria; ANAS, Azerbaijan; SSTD, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF, 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, MINERVA, GIF, DIP and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNiSW, Poland; GRICES 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; NSC, 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), KIT/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.
Fig. 14  Comparison of the measured value of the average value of $p_{rel}$ as a function of $p_T$ with various Monte Carlo expectations

References

1. V.N. Gribov, L.N. Lipatov, Sov. J. Nucl. Phys. 15, 675–684 (1972)
2. G. Altarelli, G. Parisi, Nucl. Phys. B 126, 298 (1977)
3. Y.L. Dokshitzer, Sov. Phys. JETP 46, 641–653 (1977)
4. H. Georgi, H.D. Politzer, Nucl. Phys. B 136, 445 (1978)
5. J.F. Owens, Phys. Lett. B 76, 85 (1978)
6. T. Uematsu, Phys. Lett. B 79, 97 (1978)
7. T. Sjöstrand, S. Mrenna, P.Z. Skands, J. High Energy Phys. 06, 026 (2006)
8. K. Nakamura et al. (Particle Data Group), J. Phys. G 37, 075021 (2010)
9. A.H. Mueller, Phys. Lett. B 104, 161–164 (1981)
10. G. Corcella et al., J. High Energy Phys. 01, 010 (2001)
11. The ATLAS Collaboration, Phys. Rev. D 84, 054001 (2011)
12. The OPAL Collaboration, Z. Phys. C 68, 203–214 (1995)
13. The DELPHI Collaboration, Z. Phys. C 73, 11–60 (1996)
14. The ALEPH Collaboration, Phys. Lett. B 357, 487–499 (1995)
15. P. Mittig, Phys. Rep. 177, 141 (1989)
16. G. Amison et al., Phys. Lett. B 132, 223 (1983)
17. The CDF Collaboration, Phys. Rev. Lett. 65, 968–971 (1990)
18. ZEUS Collaboration, Phys. Lett. B 414, 428–443 (1997)
19. H1 Collaboration, Phys. Lett. B 654, 148–159 (2007)
20. The ATLAS Collaboration, J. Instrum. 3, S08003 (2008)
21. The ATLAS Collaboration, Jet energy scale and its systematic uncertainty in proton–proton collisions at $\sqrt{s} = 7$ TeV with ATLAS 2010 data. ATLAS-CONF-2011-032, https://cdsweb.cern.ch/record/1337782
22. The ATLAS Collaboration, Eur. Phys. J. C 71, 1512 (2011)
23. The ATLAS Collaboration, Eur. Phys. J. C 70, 823–874 (2010)
24. S. Agostinelli et al. (GEANT4 Collaboration), Nucl. Instrum. Methods A 506, 250–303 (2003)
25. A. Sherstnev, R.S. Thorne, Eur. Phys. J. C 55, 553–575 (2008)
26. The ATLAS Collaboration, New J. Phys. 13, 053033 (2011)
27. P.Z. Skands, Phys. Rev. D 82, 074018 (2010)
28. J. Pumplin et al., J. High Energy Phys. 07, 012 (2002)
29. J.M. Butterworth, J.R. Forshaw, M.H. Seymour, Z. Phys. C 72, 637–646 (1996)
30. M. Bahr et al., Eur. Phys. J. C 58, 639–707 (2008)
31. The ATLAS Collaboration, ATLAS Monte Carlo tunes for MC09. ATLAS-CONF-2010-002, http://cdsweb.cern.ch/record/1247375/files/ATL-PHYS-PUB-2010-002.pdf
32. S. Gieseke et al., Herwig++ 2.5 Release Note (2011)
33. T. Gleisberg et al., J. High Energy Phys. 02, 007 (2009)
34. J. Pumplin, A. Belyaev, J. Huston, D. Stump, W.K. Tung, J. High Energy Phys. 02, 032 (2006)
35. T. Sjöstrand, S. Mrenna, P.Z. Skands, Comput. Phys. Commun. 178, 852–867 (2008)
36. M. Cacciari, G.P. Salam, G. Søyze, J. High Energy Phys. 04, 063 (2008)
37. M. Cacciari, G.P. Salam, Phys. Lett. B 641, 057 (2006)
38. W. Lampl et al., Calorimeter clustering algorithms: Description and performance. ATL-LARG-PUB-2008-002, https://cdsweb.cern.ch/record/1099735
39. G. D’Agostini, Nucl. Instrum. Methods A 362, 487–498 (1995)
40. T. Adye, RooUnfold: unfolding framework and algorithms. Presentations to the Oxford and RAL ATLAS Groups, May 2008. http://hepunx.rl.ac.uk/~adye/talks/080520-unfold.pdf
41. The ATLAS Collaboration, Jet energy resolution and reconstruction efficiencies from in-situ techniques with the ATLAS detector using proton–proton collisions at a center of mass energy $\sqrt{s} = 7$ TeV. ATLAS-CONF-2010-054
42. The ATLAS Collaboration, New J. Phys. 13, 053033 (2011)
43. The ATLAS Collaboration, Muon momentum resolution in first pass reconstruction of pp collision data recorded by ATLAS in 2010. ATLAS-CONF-2011-003, http://cdsweb.cern.ch/record/1337075
44. The ATLAS Collaboration, ATLAS calorimeter response to single isolated hadrons and estimation of the calorimeter jet scale uncertainty. ATLAS-CONF-2011-028, http://cdsweb.cern.ch/record/1337075
45. B. Efron, Ann. Stat. 7, 1 (1979)
46. A. Buckley, M. Whalley, PoS ACAT2010, 067 (2010)
47. A. Buckley et al., arXiv:1003.0694 (2010)

The ATLAS Collaboration

G. Aad, B. Abbott, J. Abdallah, A.A. Abdelalim, A. Abdessalam, O. Abdinov, B. Abi, M. Abolins, H. Abramowicz, H. Abreu, E. Acerbi, B.S. Acharya, D.L. Adams, T.N. Addy, J. Adelman, M. Aderholz, S. Adomeit, P. Adragna, T. Adye, S. Aefsky, J.A. Aguilar-Saavedra, M. Aharrouche, B. Efron, Ann. Stat. 7, 1 (1979)
