Properties of jets measured from tracks in proton-proton collisions at center-of-mass energy $\sqrt{s} = 7$ TeV with the ATLAS detector


DOI
10.1103/PhysRevD.84.054001

Publication date
2011

Document Version
Final published version

Published in
Physical Review D. Particles, Fields, Gravitation, and Cosmology

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
Properties of jets measured from tracks in proton-proton collisions at center-of-mass energy $\sqrt{s} = 7$ TeV with the ATLAS detector

G. Aad et al.*
(ATLAS Collaboration)
(Received 17 July 2011; published 20 September 2011)

Jets are identified and their properties studied in center-of-mass energy $\sqrt{s} = 7$ TeV proton-proton collisions at the Large Hadron Collider using charged particles measured by the ATLAS inner detector. Events are selected using a minimum bias trigger, allowing jets at very low transverse momentum to be observed and their characteristics in the transition to high-momentum fully perturbative jets to be studied. Jets are reconstructed using the anti-$k_t$ algorithm applied to charged particles with two radius parameter choices, 0.4 and 0.6. An inclusive charged jet transverse momentum cross section measurement from 4 GeV to 100 GeV is shown for four ranges in rapidity extending to 1.9 and corrected to charged particle-level truth jets. The transverse momenta and longitudinal momentum fractions of charged particles within jets are measured, along with the charged particle multiplicity and the particle density as a function of radial distance from the jet axis. Comparison of the data with the theoretical models implemented in existing tunings of Monte Carlo event generators indicates reasonable overall agreement between data and Monte Carlo. These comparisons are sensitive to Monte Carlo parton showering, hadronization, and soft physics models.

DOI: 10.1103/PhysRevD.84.054001 PACS numbers: 13.85.Hd, 13.87.Ce, 13.87.Fh

I. INTRODUCTION

Quantum chromodynamics (QCD) [1,2] provides an excellent description of the kinematic distribution of high transverse momentum jets in proton-proton collisions, but does not give straightforward predictions for the properties of particles within these jets or for the properties of low-momentum jets. These quantities may be predicted by Monte Carlo event generators, whose results are dependent on phenomenological models of parton showering, hadronization, and soft (i.e. low-momentum transfer) physics. These models have free parameters that must in turn be tuned to data.

In this work the properties of low-momentum jets are measured from charged particle tracks in the ATLAS 2010 data at center-of-mass energy $\sqrt{s} = 7$ TeV. These properties will be compared to a range of Monte Carlo tunes derived from previously measured data, allowing the study of the transition between the MC generators’ separately-implemented models of soft strong interactions and perturbative QCD. This measurement is complementary to others that are used for comparison with models of soft QCD, such as inclusive charged particle kinematic distributions [3] and underlying event properties [4], and may shed light on the limited ability of existing models to describe both types of measurements simultaneously with a single set of tunable parameters [5]. At higher momenta, track-based jet measurements provide a complement to calorimeter-based measurements [6,7], with results independent of calorimeter calibrations, selections, and uncertainties.

Jets have previously been measured using charged particles in proton-proton collisions at the CERN Intersecting Storage Rings [8], and in proton-antiproton collisions at the CERN Super Proton Synchrotron [9] and Fermilab Tevatron [10]. The momentum fraction of charged particles with respect to calorimeter jets has also been measured [11,12]. Related fragmentation measurements in electron-positron and electron-proton collisions are reviewed in Ref. [13].

In this work, jets are reconstructed using the anti-$k_t$ algorithm [14], using a radius parameter $R$ of 0.4 or 0.6, applied to measured charged particles with transverse momentum $p_T > 300$ MeV; jet four-momenta are determined by adding constituent four-vectors. The distribution of jet momenta, jet track multiplicity, and kinematic properties of tracks within these track-based jets are corrected back to truth-level charged particle jets, which are defined to be the jets obtained when the same algorithm is applied to all primary charged particles emerging from the proton-proton collision with the same $p_T$ cut.

Five quantities are measured for charged particle jets:

$$\frac{d^2\sigma_{\text{jet}}}{dp_T^{\text{jet}}dy_{\text{jet}}}; \frac{1}{N_{\text{jet}}} \frac{dN_{\text{jet}}}{dN_{\text{ch}}}; \frac{1}{N_{\text{jet}}} \frac{dN_{\text{ch}}}{dz}; \frac{1}{N_{\text{jet}}} \frac{dN_{\text{ch}}}{dp_T^{\text{rel}}}; \rho_{\text{ch}}(r)$$

(1)

*Full author list given at the end of the article.

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.
The inclusive charged particle jet cross section $\sigma_{\text{jet}}$ is measured in bins of $p_{T,\text{jet}}$ for four rapidity ($y$) [15] ranges: 0–0.5, 0.5–1.0, 1.0–1.5, 1.5–1.9, where the last range is narrower than the others to ensure that measured jets are fully contained in the tracking acceptance of the ATLAS detector. The distribution of charged particle multiplicities per jet $N_{\text{ch}}$ is measured for each of these ranges, separated further into five jet transverse momentum ranges: 4–6 GeV, 6–10 GeV, 10–15 GeV, 15–24 GeV, 24–40 GeV. The longitudinal momentum fraction $z$ of charged particles in these jets is measured in the same rapidity and transverse momentum ranges. The variable $z$, also known as the fragmentation variable, is defined for each particle in a jet by

$$z = \frac{\vec{p}_{\text{ch}} \cdot \vec{p}_{\text{jet}}}{|\vec{p}_{\text{jet}}|^2},$$

(2)

where $\vec{p}_{\text{ch}}$ is the momentum of the charged particle and $\vec{p}_{\text{jet}}$ is the momentum of the jet that contains it. The $z$ distribution presented here differs from the usual definition (in [11], for example), which would include neutral particles and low-momentum charged particles in the total jet momentum. The variable $p_{T,\text{jet}}$ is the momentum of charged particles in a jet transverse to that jet’s axis:

$$p_{T,\text{jet}} = \frac{|\vec{p}_{\text{ch}} \times \vec{p}_{\text{jet}}|}{|\vec{p}_{\text{jet}}|}.$$  

(3)

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

$$r = \sqrt{(\phi_{\text{ch}} - \phi_{\text{jet}})^2 + (y_{\text{ch}} - y_{\text{jet}})^2},$$

(4)

so that

$$\rho_{\text{ch}}(r) = \frac{1}{N_{\text{jet}}} \frac{dN_{\text{ch}}}{2\pi r dr}.$$  

(5)

Note that this is a particle number density, rather than the related energy density variable used for calorimeter-based jet shape measurements. The variables $\rho_{\text{ch}}(r)$ and $p_{T,\text{jet}}$ are measured in the same rapidity and jet momentum ranges as $z$.

### II. THE ATLAS DETECTOR

The ATLAS detector [16] at the Large Hadron Collider (LHC) [17] covers almost the whole solid angle around the collision point with layers of tracking detectors, calorimeters and muon chambers. It has been designed to study a wide range of physics topics at LHC energies. For the measurements presented in this paper, the tracking devices and the trigger system are used.

The ATLAS inner detector (ID) has full coverage in the pseudorapidity range $|\eta| < 2.5$. It consists of a silicon pixel detector (Pixel), a silicon microstrip detector (SCT) and a transition radiation tracker (TRT), the last of which only covers $|\eta| < 2.0$. These detectors cover a sensitive radial distance from the interaction point of 50–150 mm, 299–560 mm and 563–1066 mm, respectively, and are immersed in a 2 T axial magnetic field. The inner-detector barrel (end-cap) parts consist of 3 (2 x 3) Pixel layers, 4 (2 x 9) double layers of single-sided silicon microstrips with a 40 mrad stereo angle, and 73 (2 x 160) layers of TRT straws. Typical position resolutions are 10, 17 and 130 $\mu$m for the $R$-$\phi$ coordinate of the pixel detector, SCT, and TRT, respectively; the resolution of the second measured coordinate is 115 (580) $\mu$m for the pixel detector (SCT). A track from a charged particle traversing the barrel detector would typically have 11 silicon hits (3 pixel clusters and 8 strip clusters) and more than 30 straw hits.

The ATLAS detector has a three-level trigger system: Level 1 (L1), Level 2 (L2) and event filter (EF). The L1 triggers use custom fast electronics; the L2 and EF further refine selections in software, with L2 using a subset of the event information for rapid processing and EF reconstructing complete events. For this measurement, the trigger relies on the L1 signals from the beam pickup timing devices (BPTX) and the minimum bias trigger scintillators (MBTS). The BPTX stations are composed of electrostatic button pick-up detectors attached to the beam pipe at $\pm 175$ m from the center of the ATLAS detector. The coincidence of the BPTX signal between the two sides of the detector is used to determine when bunches are colliding in the center of the detector. The MBTS are mounted on each side of the detector at $Z = \pm 3.56$ m. They are segmented into eight sectors in azimuth and two rings in pseudorapidity ($2.09 < |\eta| < 2.82$ and $2.82 < |\eta| < 3.84$). Data were collected for this analysis using a trigger requiring a BPTX coincidence and MBTS trigger signals. The MBTS trigger used for this paper is configured to require one hit above threshold from either side of the detector, referred to as a single-arm trigger.

### III. MONTE CARLO SAMPLES

The corrections used in this analysis, from detector-level track jets to truth-level charged particle jets, are derived from a Monte Carlo (MC) simulated sample using the PYTHIA 6.421 event generator program [18] with the ATLAS AMBTL tune, whose parameters are chosen based on single-track distributions in ATLAS minimum bias data [3]. In order to derive corrections, these events are then passed through the ATLAS detector simulation [19], based on GEANT 4 [20]. Large simulated samples of the ATLAS MC09 [21] and Perugia 2010 [22] tunes of PYTHIA 6 are also used.

Additional PYTHIA 6.421 tunings are used at the generator-level for comparison to data, and in order to explore the sensitivity of the measurement to underlying event, fragmentation, and hadronization parameters, as
described in Sec. VII. The Perugia family of tunes [22] uses the CTEQ5L [23] parton distributions; there is a central value ("Perugia 0") and several variants that attempt to bracket the possible changes in tune parameters that are permitted in fitting existing data. Two variants, "Perugia HARD" and "Perugia SOFT", are created by changing the initial-state radiation cutoff scale so that the perturbative contribution is different, while other parameters are adjusted to bring the tuned distributions back into agreement. Perugia HARD (SOFTWARE) has more (fewer) jets per event with higher (lower) average momentum, and a higher (lower) average charged multiplicity. The Perugia 2010 variant is an adjustment of Perugia 0 that improves the description of jet shapes and adjusts hadronization parameters for better consistency with LEP data, given that the original values for these parameters were based on $Q^2$ shower ordering; in particular, it has an adjusted yield of strange particles. Perugia NOCR ("no color reconnection") represents an effort to reproduce the same data with no color reconnection model used; it is used in computing systematic uncertainties, but because it could not be made to fit all the original tuning data it is not included on comparison plots. All PYTHIA tunes described thus far use $p_T$-ordered showering; the DW tune is also used in order to include the impact of $Q^2$ ordering.

Predictions from separately-implemented generators are used for additional studies and comparisons. PHOJET 1.12.1.3 [24], which relies on PYTHIA 6.115 for the fragmentation of partons, is used with its default settings. A HERWIG++ 2.4.2 [25] sample is used with its default settings, along with an additional sample tuned on 7 TeV underlying event data with HERWIG++ 2.5.1 (UE7) [26]. A PYTHIA 8.145 [27] sample is used with the 4C tune [28].

Single- and double-diffractive events are included in the PHOJET and PYTHIA 8 samples, with cross sections relative to the nondiffractive events as indicated by those generators. HERWIG++ does not include diffractive modeling, while PYTHIA 6 diffractive modeling produces a negligible yield of jets as defined in this analysis.

IV. DATA SELECTION

This measurement uses a sample of early ATLAS data at center-of-mass energy $\sqrt{s} = 7$ TeV for which the minimum bias trigger was minimally prescaled, corresponding to a total integrated luminosity of 799 $\mu$b$^{-1}$ and a peak luminosity of $6.6 \times 10^{28}$ cm$^{-2}$ s$^{-1}$. Events from colliding proton bunches are selected if the MBTS recorded one or more counters above threshold on either side. They are further required to have a primary vertex [29] reconstructed using beam-spot information [30]. Events with additional reconstructed primary vertices are rejected. The average number of collisions per bunch crossing $\mu$ depends on luminosity; the highest-luminosity data collected have $\mu \sim 0.14$, but over half the data have $\mu \approx 0.01$. The inefficiency introduced by the requirement that no additional primary vertices be present arises from two causes:

(i) Events may be lost due to multiple collisions in a single bunch crossing. The rate at which this occurs is calculated as a function of $\mu$ [32], and corresponds to a 3.3% correction to the cross section over the entire data set.

(ii) Events may be lost because multiple vertices are incorrectly reconstructed in bunch crossings with only a single collision. The rate at which this occurs is determined from simulated events, and

A. Track reconstruction

Tracks are reconstructed using the ATLAS primary silicon-seeded tracking algorithm, with a configuration similar to that described in the 900 GeV and 7 TeV minimum bias measurements [3,31]. In order to select good-quality tracks emerging from the primary vertex while maintaining a high efficiency, each track is then required to have:

(i) transverse momentum $p_T > 300$ MeV;

(ii) pseudorapidity $|\eta| < 2.5$;

(iii) transverse impact parameter with respect to the primary vertex $|d_0| < 1.5$ mm (0.2 mm) for tracks with $p_T$ less than (greater than) 10 GeV;

(iv) longitudinal impact parameter with respect to the primary vertex $z_0$ satisfying $|z_0 \sin \theta| < 1.5$ mm;

(v) if a signal (or hit) is expected in the innermost pixel detector layer (i.e. if the extrapolated track passes through a section of that layer with functioning instrumentation), then such a hit is required, with one pixel hit in any layer required otherwise;

(vi) at least 6 SCT hits.

B. Selection efficiencies

The efficiency for triggering given the presence of one or more jets is determined from data using a random subset of all events with a BPTX coincidence. In such events for which an offline track jet (as defined in Sec. V) is present, the fraction of events for which the MBTS single-arm trigger is also passed is determined. It is found to be negligibly different from 100% for all events containing jets as defined in this analysis.

The efficiency for primary vertex-finding given the presence of one or more jets is determined from data by removing the track selection cuts that use impact parameter with respect to the primary vertex described in Sec. IV.A. Other cuts are kept as described, and a requirement is added that the transverse impact parameter with respect to the beam spot satisfies $|d_0| < 4.0$ mm. Jets are then reconstructed from tracks satisfying this new selection, and the fraction of events containing jets that also have a primary vertex reconstructed is determined. It is found to be negligibly different from 100% for all events containing jets as defined in this analysis.

The inefficiency introduced by the requirement that no additional primary vertices be present arises from two causes:

(i) Events may be lost due to multiple collisions in a single bunch crossing. The rate at which this occurs is calculated as a function of $\mu$ [32], and corresponds to a 3.3% correction to the cross section over the entire data set.

(ii) Events may be lost because multiple vertices are incorrectly reconstructed in bunch crossings with only a single collision. The rate at which this occurs is determined from simulated events, and
parameterized as a function of the number of selected tracks in the event. It corresponds to a cross section correction of roughly 2%, depending weakly on jet momentum.

The efficiencies for track reconstruction and selection are derived from simulated events, whose tracking properties have been shown to be in excellent agreement with the data [3]. The uncertainties on these efficiencies, and their impact on the present results, are discussed in Sec. VII.

**V. JET DEFINITION**

Reconstructed track jets are produced by applying the anti-$k_t$ algorithm with radius parameter $R = 0.4$ or 0.6 to the selected tracks; the pion mass is assumed for all tracks in the application of the algorithm. After the algorithm is applied, track jets with $p_T > 4$ GeV and $|y| < 1.9$ are accepted, including those with only one constituent track. The rapidity cut is chosen to ensure that jets are fully contained within the tracking acceptance $|y| < 2.5$ in the $R = 0.6$ case; for the narrower $R = 0.4$ jets, the same rapidity cut is used for consistency.

Truth-level charged particle jets similarly have the anti-$k_t$ algorithm applied to primary charged simulated particles with $p_T > 300$ MeV, using each particle’s true mass in the application of the algorithm. Primary charged particles are defined as charged particles with a mean lifetime $\tau > 0.3 \times 10^{-10}$ s, which are produced in the primary collision or from subsequent decays of particles with a shorter lifetime. Thus the charged decay products of $K^0_S$ particles are not included. Charged particle jets are required to meet the $p_T$ and $y$ requirements given for track jets above.

**VI. CORRECTION PROCEDURE**

After all distributions are measured for reconstruction-level track jets, each distribution is corrected to truth-level charged particle jets. The corrections are derived from the AMBT1 sample discussed in Sec. III, and account for tracking efficiency and momentum resolution.

Reconstructed jets are binned simultaneously in $p_T$, constituent multiplicity ($N^\text{ch}_{\text{jet}}$), and rapidity ($y$), so that a three-dimensional distribution is produced for the purpose of applying corrections. Similarly, for each per-track variable $\tau$, $p_T^{\text{rel}}$, and $r$, each track in each jet is binned in the variable, the parent jet momentum, and parent jet rapidity. For both the jet-level quantities and each track-level variable, corrections are applied simultaneously in the three binned variables using the Bayesian iterative unfolding algorithm [33].

The corrections for this algorithm are based on a response matrix derived from MC events, which encapsulates the probability for a charged particle jet with a particular $(p_T, N^\text{ch}_{\text{jet}}, y)$ to be reconstructed in each possible $p_T$, $N^\text{ch}_{\text{jet}}$, and $y$ bin. A reconstructed track jet is defined to be matched to a truth-level charged particle jet if it is within $\Delta R < 0.2$ (0.3) for jet radius 0.4 (0.6). These values are chosen to avoid matching ambiguity while still giving good efficiency, which rises from roughly 55% at the very lowest momenta to greater than 95% for $p_T^{\text{jet}} > 10$ GeV. In the case of the $\tau$ response matrix, binned in $(\tau, p_T^{\text{jet}}_\tau, y^{\text{jet}}_\tau)$, if a truth jet is unmatched then all its tracks are counted as lost to inefficiency. If it is matched, then each of its constituent charged particles are matched to the closest track within $\Delta R < 0.04$; if there is no such track, then that particular particle is counted as lost to inefficiency. The $p_T^{\text{rel}}$ and $r$ matrices are filled in the same manner as the $\tau$.

The Bayesian Iterative Unfolding algorithm takes both inefficiencies and resolution effects into account, as “missed” entries and bin-to-bin transfers respectively; the probability of these occurrences is determined from the central AMBT1 MC sample. Before the Bayesian iterative unfolding algorithm can be applied, a correction must be made for unmatched reconstructed jets, and for unmatched reconstructed tracks in the $\tau$, $p_T^{\text{rel}}$, and $r$ distributions. This correction is determined from the central MC sample by the ratio between the reconstructed distributions and the matched reconstructed distributions.

Three iterations of the algorithm are used. This is validated by applying the unfolding to ensembles of data-sized MC samples, and observing the number of iterations required to obtain convergence. Statistical uncertainties on the unfolded variables are determined by a toy histogram-variation method, which produces 100 pseudoexperiment distributions by varying the contents of each input bin randomly, then calculating the covariance matrix from the variation in the output of the unfolding. This error calculation is also validated using ensembles of data-sized MC samples (with full detector simulation) by considering the distribution of pulls:

$$\text{Pull} = \frac{N^\text{corr}}{\sigma^\text{corr}} - \frac{N^\text{truth}}{\sigma^\text{truth}},$$

where $N^\text{corr}$ is the corrected number of events in a given bin, $N^\text{truth}$ is the true number of events in that bin, and $\sigma^\text{corr}$ is the statistical error for that bin as reported by the unfolding procedure. The RMS of the pull distribution should be consistent with one. In some bins it is not, so the statistical errors are scaled appropriately. This scaling is at most roughly 1.5 except for a few outlying bins in the multiplicity and $p_T^{\text{rel}}$ distributions that have very low statistics.

The ensemble studies also reveal a bias in the unfolding algorithm calculation in closure tests, which is roughly 2% in the jet $p_T$ distributions. The possibility of an additional bias from changing the prior distribution used as input to the unfolding algorithm was investigated. After three iterations, this bias vanishes.
VII. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties are computed by determining the impact of a given effect on reconstruction-level distributions, then propagating it through the unfolding procedure to determine the final impact on the measurement. A summary of all systematic uncertainties for selected bins is shown in Table I for the cross section and $p_T^{rel}$ distributions, and in Table II for the multiplicity, $z$, and $\rho(r)$ distributions. The uncertainties given in each column of the tables are:

Tracking efficiency and resolution: The systematic uncertainties associated with tracking efficiency and resolution are computed by applying efficiency scaling and resolution smearing to single tracks in generated events. The efficiencies and resolutions are derived from fully-simulated events and parameterized by a track’s $1/p_T$, $\eta$, and the $\Delta R$ to the nearest other track. These efficiency uncertainties are determined by adjusting the derived single-track efficiencies in accordance with uncertainties determined from data [3], for which uncertainties in the material description of the ATLAS inner detector dominate. The resulting distributions are compared with those found for the baseline resolution, with the increase (decrease) taken as the upper (lower) systematic uncertainty. Similarly, the uncertainties in tracking resolution are investigated by applying an additional Gaussian resolution smearing derived from studies of high-momentum muons; the resulting change is taken as a symmetric uncertainty.

Monte Carlo: The efficiency scaling and resolution smearing are similarly used for uncertainties associated with the particular event generator used to compute the unfolding corrections. For each tune described in Sec. III, an ensemble of 50 samples with the same number of events as the data has the smearing applied, and the average unfolded distribution is compared to one produced similarly (with 196 data-sized samples) for the central AMB1 MC sample. The largest increase (decrease) is taken as the upper (lower) systematic uncertainty. Although the response matrix primarily models the impact of detector efficiency and resolution, truth-level details can impact the corrections in several ways. Different jet or track momentum distributions change the population within bins; differences in the amount or distribution of activity in the underlying event can significantly change the track momentum and radius distribution, especially in the low-momentum bins, because these tracks are not correlated with the jet direction and so may appear at large radii.

| Bin [GeV] | $d^3\sigma_{\text{jet}}/dp_T^3$ | $|y_{\text{jet}}| < 0.5$ | $d^3\sigma_{\text{jet}}/dp_T^3$ | $1.5 < |y_{\text{jet}}| < 1.9$ |
|-----------|-------------------------------|----------------|-------------------------------|----------------------------------|
| 4–5       | +3.2% ± 0.07%                  | +0.07% ± 0.1%  | +0.05% ± 0.1%                | <0.005% ± 0.2%                  |
| 20–22     | +6.6% ± 0.3%                  | +0.2% ± 0.1%   | +0.1% ± 0.1%                 | <0.005% ± 0.2%                  |
| 40–45     | +6.8% ± 0.3%                  | +1.1% ± 0.1%   | +0.01% ± 0.1%                | <0.005% ± 0.2%                  |
| 80–90     | +7.1% ± 0.3%                  | +1.1% ± 0.1%   | +0.01% ± 0.2%                | <0.005% ± 0.2%                  |
| 1.5 < | $d^3\sigma_{\text{jet}}/dp_T^3$ | $|y_{\text{jet}}| < 1.9$ |
| 4 GeV < $p_T^{\text{jet}}$ < 6 GeV | +0.44% ± 0.1% | +0.05% ± 0.1% | <0.005% ± 0.2% | <0.005% ± 0.2% |
| 0.5–0.55  | +0.3% ± 0.1%                  | <0.005% ± 0.2% | <0.005% ± 0.2%                | <0.005% ± 0.2%                  |
| 0.85–0.9  | +1% ± 0.03%                   | <0.005% ± 0.2% | <0.005% ± 0.2%                | <0.005% ± 0.2%                  |
| 24 GeV < $p_T^{\text{jet}}$ < 40 GeV | +0.95% ± 0.1% | +1.2% ± 0.1% | +0.05% ± 0.1% | <0.005% ± 0.2% |
| 0.0–0.05  | +1.2% ± 0.1%                  | +0.05% ± 0.1%  | +0.05% ± 0.1%                | <0.005% ± 0.2%                  |
| 0.5–0.55  | +1.2% ± 0.1%                  | +0.05% ± 0.1%  | +0.05% ± 0.1%                | <0.005% ± 0.2%                  |
| 0.85–0.9  | +0.95% ± 0.1%                 | +0.05% ± 0.1%  | +0.05% ± 0.1%                | <0.005% ± 0.2%                  |
| 3.0–3.5   | +2% ± 0.1%                    | +0.05% ± 0.1%  | +0.05% ± 0.1%                | <0.005% ± 0.2%                  |
Differing strangeness fractions can change the distribution of long-lifetime tracks that decay and produce kinks in tracks, leading to momentum mismeasurements and/or loss of tracks due to failed hit requirements. The variation over tunes also accounts for PDF uncertainties, because MC properties are tuned with a particular PDF and several different PDF's are used.

**High-$p_T$ Tracks:** Despite the tightening of the $d_0$ cut for high-momentum tracks, 0.6% (8.3%) of selected tracks in the simulation with momentum above 10 (40) GeV do not have a matching truth particle with momentum within 50%. These high-momentum mismeasured tracks are primarily associated with wide-angle scatterings in the material of the inner detector that create the appearance of a single, straight track [3]. Were the fraction of such tracks similar in data, their impact on the measurement would be accounted for in the unfolding procedure. However, the data have a larger fraction of high-$p_T$ tracks failing the $d_0$ cut (2.3%) than does the MC (1.6%). A systematic uncertainty on the number of mismeasured tracks in data, as a function of
FIG. 1 (color online). The cross section for anti-$k_t$ charged particle jets as a function of $p_T$, with $|y| < 0.5$ and radius parameter $R$ as indicated. The shaded area is the total uncertainty for the corrected data distribution, excluding the overall 3.4% luminosity uncertainty. The data are compared to a range of theoretical results from Monte Carlo event generators, which are normalized to the data over the full momentum and rapidity range measured, using the scale factor $S$ as defined in the text. The bottom inserts show the fractional difference between these distributions and the data. The distributions for the Perugia HARD (Perugia SOFT) tune, not shown, agree qualitatively with the Perugia 2010 (Perugia 0) tune.

FIG. 2 (color online). The cross section for anti-$k_t$ charged particle jets as a function of rapidity, for selected momentum bins and radius parameter $R$ as indicated. The shaded area is the total uncertainty for the corrected data distribution, excluding the overall 3.4% luminosity uncertainty. The data are compared to a range of theoretical results from Monte Carlo event generators, which are normalized to the data separately for each momentum range. The distributions for all Perugia samples agree qualitatively, so only Perugia 0 is shown. The two HERWIG++ tunes agree, so only HERWIG++ 2.4.2 is shown.
track $p_T$, is computed and propagated to assess its impact on jets. The upper bound of this uncertainty assumes that all such extra tracks are well-measured, but did not pass the $d_0$ cut due to improperly modeled resolution. The lower bound is based on the assumption that the increase in rejected tracks corresponds to a proportional increase in accepted mismeasured tracks. These uncertainties are then propagated to each measured jet bin, with the scale factor determined by the fraction of jets in a given jet bin with their leading track in each momentum range. Thus the correction is largest for high jet momentum and low number of particles per jet, because these jets have the highest leading track momenta. This uncertainty on the reconstructed data is then used to scale the measured distributions, with the unfolding applied.
in each case and compared to the unfolded central value. The resulting differences in each measured bin give the final uncertainty due to high-momentum mismeasured tracks. This correction is also applied to the $z$, $p_T$, and $r$ distributions. Each bin is scaled in proportion to the fraction of tracks in that bin that come from a jet whose leading track is in a given range, in proportion to the uncertainty on tracks in that range. The corrections are only significant for high-$z$ bins in high-$p_T$ jets. The corrections are then propagated through the unfolding as in the jet-based distributions, and the result compared to the central value.

**Unmatched jets and tracks:** In order to assign an uncertainty to the correction for unmatched reconstructed jets and tracks, the unfolding is repeated with a correction determined from a fully-simulated Perugia 2010 in place of the baseline ATLAS AMBT1. The difference in the unfolding output between this and the baseline configuration is taken to be the uncertainty on the correction. The uncertainty is symmetrized by taking the maximum deviation.

**Split vertex:** The data are corrected for event rejection due to misreconstruction of extra primary vertices as a function of the number of selected tracks in the event, as discussed in Sec. IV B. A weight is applied on an event-by-event basis based on the probability of that event being rejected. As the correction is derived entirely from simulation, the full value of the correction is taken as a (symmetrized) uncertainty.

**Closure:** As discussed in Sec. VI, closure tests reveal a bias in the correction procedure. This is taken as a bin-by-bin systematic uncertainty, which is symmetrized.

FIG. 4 (color online). The distribution of the fragmentation variable $z$ for anti-$k_t$ jets with radius parameter $R$ as indicated, in the rapidity range $|y| < 1.9$. Figure 4(a) (c) shows the distributions for five momentum ranges with $R = 0.4$ (0.6), and Fig. 4(b) (d) shows the fractional difference between a range of Monte Carlo event generator predictions and the data. The distributions for the Perugia 2010 (Perugia SOFT) tune, not shown, agree qualitatively with the Perugia 0 (AMBT1) tune.
The total uncertainty on the cross section is dominated by tracking efficiency, with mismeasured high-momentum tracks also playing a role for the highest-momentum jets. Tracking efficiency uncertainties play a similar role for the multiplicity distributions, especially for low-multiplicity jets, with Monte Carlo uncertainties also making large contributions. For other distributions, Monte Carlo uncertainties are dominant, with closure uncertainties making large contributions in extremal bins; tracking efficiency uncertainties mostly cancel because these distributions are normalized by the number of jets.

**VIII. RESULTS AND DISCUSSION**

A selection of the distributions measured in this analysis appears in Figs. 1–6. Other rapidity ranges may be found in Ref. [34]. They are compared to the MC distributions described in Sec. III.

### A. Charged particle jet cross section

Cross sections as a function of jet $p_T$ are shown in Fig. 1. The simulated cross sections shown for comparison are scaled to the data, using the scale factor $S$ defined by

$$S = \frac{\sigma_{data}^{total}}{\sigma_{MC}^{total}}$$

where

$$\sigma_{total}^{total} = \int_{4 \text{ GeV}}^{100 \text{ GeV}} dp_T \int_{-1.9}^{1.9} dy_{jet} \frac{d^2 \sigma_{jet}}{dp_T dy_{jet}}$$

**FIG. 5 (color online).** The distribution of the charged particle transverse momentum $p_T^{\text{ch}}$ with respect to anti-$k_t$ jets with radius parameter $R$ as indicated, in the rapidity range $|y| < 1.9$. Figure 5(a) (c) shows the distributions for five momentum ranges with $R = 0.4$ (0.6), and Fig. 5(b) (d) shows the fractional difference between a range of Monte Carlo event generator predictions and the data.
The scale factors for the various MC’s are given in Table III.

In all cases except HERWIG++ 2.4.2, the scale factors $S$ for $R = 0.6$ are larger than those for $R = 0.4$. Since the larger radius parameter results in the inclusion of more particles not directly associated with perturbative scattering, this implies that the models underestimate the contribution of the underlying event required to reproduce the data. PHOJET has the best agreement between the scale factors at $R = 0.4$ and $R = 0.6$. The PYTHIA Perugia tunes also agree well for the two radii, and are most consistent with one.

The jet cross section distributions (Fig. 1) fall by 6 orders of magnitude between jet momenta of 4 and 100 GeV. The MC models considered agree broadly with this trend, but do not agree well in detail. By construction of the normalization factor $S$, all distributions agree with the data in the lowest momentum bins; most also give qualitative agreement for the shape at the lowest momenta. The MC distributions diverge from the data in the 10–20 GeV range, with some having a harder and others a softer momentum dependence. At higher $p_T$, many of the models’ momentum dependence agrees well with the data. If one identifies the higher-momentum region as dominated by perturbative modeling and the low-momentum region as dominated by soft physics, this indicates that perturbative modeling of charged particle jets is in fair agreement for most of the tunes. It is the transition from soft physics to the perturbative region that is not successfully modeled.

The PYTHIA models give a harder shape for the momentum spectrum than the data below a jet $p_T$ of about 20 GeV, after which they exhibit roughly the same momentum dependence or become slightly softer. By contrast, the
PHYSICAL REVIEW D 84, 054001 (2011)

TABLE III. Scale factors $S$ for Monte Carlo cross section normalization as defined in Eq. (7). The total Monte Carlo cross sections are normalized to the total for the data, over the full momentum and rapidity ranges investigated in this analysis, 4 GeV < $p_T\text{jet}$ < 100 GeV and |$y_{\text{jet}}$| < 1.9. The statistical uncertainties on these ratios are less than 0.1%. The systematic uncertainties are entirely correlated within columns, and largely correlated between columns.

<table>
<thead>
<tr>
<th>Tune</th>
<th>$S(R = 0.4)$</th>
<th>$S(R = 0.6)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMBT1</td>
<td>0.838</td>
<td>0.896</td>
</tr>
<tr>
<td>Perugia 0</td>
<td>0.981</td>
<td>1.087</td>
</tr>
<tr>
<td>Perugia HARD</td>
<td>0.936</td>
<td>1.058</td>
</tr>
<tr>
<td>Perugia SOFT</td>
<td>0.968</td>
<td>1.036</td>
</tr>
<tr>
<td>Perugia 2010</td>
<td>0.976</td>
<td>1.044</td>
</tr>
<tr>
<td>DW</td>
<td>0.894</td>
<td>1.045</td>
</tr>
<tr>
<td>HERWIG++ 2.4.2</td>
<td>0.753</td>
<td>0.612</td>
</tr>
<tr>
<td>HERWIG++ 2.5.1 UE7</td>
<td>0.425</td>
<td>0.458</td>
</tr>
<tr>
<td>PYTHIA 8</td>
<td>0.777</td>
<td>0.815</td>
</tr>
<tr>
<td>PHOJET</td>
<td>0.643</td>
<td>0.668</td>
</tr>
</tbody>
</table>

PHOJET and HERWIG++ models produce spectra that are softer than the data in the 10–20 GeV range but have relatively good shape agreement outside, although the HERWIG++ 2.5.1 UE7 tune has the additional feature of producing too hard a spectrum at momenta below 10 GeV.

The DW tune has a cross section that falls much too slowly in the transition region, and falls much too rapidly at high momentum; this effect is especially pronounced for $R = 0.4$ (0.6); these uncertainties are entirely correlated within columns, and largely correlated between columns.

Cross sections as a function of rapidity are shown in Fig. 2. The MC distributions are normalized to the data in each momentum bin separately rather than to the scale factor $S$. The rapidity dependence of the cross section shows generally good agreement between data and MC. The cross section decreases only slightly with increasing rapidity at low momenta, but by a somewhat higher amount at higher momenta.

B. Charged particle kinematics and multiplicity in jets

The multiplicity of charged tracks per jet, for several momentum ranges over the full range |$y_{\text{jet}}$| < 1.9, is shown in Fig. 3. The charged particle $z$, $p_T^{\text{rel}}$, and $\rho_{\text{ch}}(r)$ distributions are shown for the central rapidity range in Figs. 4–6 respectively.

No tune describes well all of the kinematic distributions and multiplicities of charged particles within jets. For $z$, ATLAS AMBT1 and Perugia SOFT give good descriptions of the data. For $\rho_{\text{ch}}(r)$, AMBT1 gives a good description. For $p_T^{\text{rel}}$ and multiplicity, no tune correctly describes all data. For all distributions, HERWIG++ 2.4.2 shows strong disagreement with the data, characterized by an excess of low-momentum particles, which is especially pronounced for the larger jet-finding parameter $R$ and at large particle $r$ [as defined in Eq. (4)]. HERWIG++ 2.5.1 UE7 represents a significant improvement, so HERWIG++ 2.4.2 will not be discussed further.

All remaining models give good agreement for the average charged particle multiplicity per jet. The AMBT1 and Perugia SOFT tunes agree well with the multiplicity distributions (Fig. 3) for the vast majority of jets, and Perugia 0, Perugia 2010, and PHOJET give fair agreement for $R = 0.4$, although the high-multiplicity tail in data is greatly underestimated by all models. HERWIG++ 2.5.1 UE7 and Perugia HARD have a significant excess of low-multiplicity jets, while PYTHIA 8.145 4C and PYTHIA DW exhibit a deficit.

The AMBT1 and Perugia SOFT tunes give good agreement with the measured longitudinal momentum fraction $z$ (Fig. 4); Perugia 0 also agrees well for $R = 0.4$. The other MC’s (except HERWIG++ 2.4.2) agree within 30% at low $z$, but diverge more significantly at high $z$. Perugia HARD has the most significant excess of high-$z$ particles, with excesses also present for PHOJET and Perugia 0. The excess is particularly large at lower jet momenta and $R = 0.6$, suggesting that the soft physics model is characterized by fewer particles with higher momentum. By contrast, PYTHIA 8.145 4C, PYTHIA DW, and HERWIG++ 2.5.1 UE7 have too few high-$z$ particles, with variations again larger for $R = 0.6$. PYTHIA DW exhibits an excess at mid-$z$ at low jet momenta, which is seen at progressively lower $z$ values as the jet momentum increases, implying an excess of particles with a momentum of roughly 2 GeV that is not associated with jet structure.

The transverse momentum $p_T^{\text{rel}}$ is in fair agreement (~20%) at low-to-moderate values for all MC generators except HERWIG++ 2.4.2. At the lowest jet momenta and highest measurable $p_T^{\text{rel}}$, PHOJET and HERWIG++ 2.5.1 UE7 have an excess of particles, while PYTHIA DW has a deficit. At higher jet momenta, the data have more high-$p_T^{\text{rel}}$ particles than any tune, with Perugia 2010 and Perugia HARD giving the closest description and Perugia SOFT the furthest. Perugia 2010 and Perugia HARD agree better than do the other Perugia tunes.

The AMBT1, PYTHIA 8.145 4C, and HERWIG++ 2.5.1 UE7 tunes provide a good description of the charged particle number density $\rho_{\text{ch}}(r)$ (Fig. 6) at all radii. PHOJET and the Perugia tunes (especially SOFT) have an excess of particles very close to the jet axis, which is most pronounced at high jet momentum and for $R = 0.6$; Perugia 2010 agrees better in this region than do the other Perugia tunes. At high $r$, PHOJET and all PYTHIA tunes except AMBT1 and Perugia SOFT have too few particles. However, the disagreement is less pronounced than is seen at high $p_T^{\text{rel}}$, implying that high-radius soft particles from the underlying event are better-described than high-radius hard radiation.
IX. CONCLUSIONS

A measurement is presented of the charged particle jet cross section as a function of transverse momentum and rapidity, along with the transverse momentum, longitudinal momentum fraction, and number density as a function of radius for charged particles within these jets, using early 7 TeV LHC collision data collected with the ATLAS detector. The study of jets with tracks allows for precise measurements of low-momentum jets and their properties, thus complementing calorimeter-based jet measurements and allowing the study of the transition from soft collisions to jet production in the perturbative regime of QCD. It also provides additional observables for consideration in the tuning of MC event generators, which complement existing “minimum bias” and underlying event measurements.

No tune or model presented here agrees with all quantities measured within their uncertainties, suggesting that future MC tunes may be improved. Difficulty in modeling the transition between soft and perturbative physics is indicated by disagreements between data and all MC distributions in the 10–20 GeV range in the dependence of the charged jet cross section on jet momentum. Dependence of contributions in the 10–20 GeV range in the dependence of the charged particle number density with respect to the jet that contains them are produced more copiously than any model predicts, as are jets with large charged particle multiplicity. The longitudinal momentum fraction $z$ is best described by the PYTHIA 6.421 AMBT1 tune. The charged particle number density $p_{ch}(r)$ is well-described by the PYTHIA 6.421 AMBT1, PYTHIA 8.145 4C, and HERWIG++ 2.5.1 UE7 tunes. With the exception of the HERWIG++ 2.4.2 default tune, which greatly disagrees with these measurements, all models appear to underestimate the contribution of the underlying event required to model the data.

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, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MITEI, CRP and VSC CR, Czech Republic; DNRF, DNRSC and Lundbeck Foundation, Denmark; ARTEMIS, European Union; IN2P3-CNRS, CEA-DSM/IRFU, France; GNAS, Georgia; BMBF, DFG, 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.

[15] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the Z-axis along the beam pipe. The X-axis points from the IP to the center of the LHC ring, and the Y axis points upward. Cylindrical coordinates $(r, \phi)$ are used in the transverse plane, with $\phi$ being the azimuthal angle around the beam pipe. The rapidity $y$ for a track or jet is defined by $y = 0.5 \ln ((E + pz)/(E - pz))$, where $E$ denotes the
energy and $p_z$ is the component of the momentum along the beam direction; the pion mass is assumed for reconstructed tracks.


[34] A complete set of tables for all distributions in all measured rapidity bins are available at the Durham HepData repository (http://hepdata.cedar.ac.uk).

(ATLAS Collaboration)

1 University at Albany, Albany New York, USA
2 Department of Physics, University of Alberta, Edmonton AB, Canada
3 Department of Physics, Ankara University, Ankara, Turkey
4 Department of Physics, Dumlupinar University, Kutahya, Turkey
5 Department of Physics, Gazi University, Ankara, Turkey
6 Division of Physics, TOBB University of Economics and Technology, Ankara, Turkey
7 Turkish Atomic Energy Authority, Ankara, Turkey
8 LAPP, CNRS/IN2P3 and Université de Savoie, Annecy-le-Vieux, France
9 High Energy Physics Division, Argonne National Laboratory, Argonne, Illinois, USA
10 Department of Physics, University of Arizona, Tucson, Arizona, USA
11 Department of Physics, The University of Texas at Arlington, Arlington, Texas, USA
12 Physics Department, The University of Athens, Athens, Greece

054001-22
Also at DSM/IRFU (Institut de Recherches sur les Lois Fondamentales de l’Univers), CEA Saclay (Commissariat a l’Energie Atomique), Gif-sur-Yvette, France

Also at Laboratoire de Physique Nucléaire et de Hautes Energies, UPMC and Université Paris-Diderot and CNRS/IN2P3, Paris, France

Also at Department of Physics, Nanjing University, Jiangsu, China

Deceased