ATLAS NOTE

ATLAS-CONF-2015-025

July 23, 2015

Measurement of jet charge in dijet events from $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector

The ATLAS Collaboration

Abstract

The momentum-weighted sum of the charges of tracks associated to a jet is sensitive to the charge of the initiating quark or gluon. This paper presents a measurement of the distribution of momentum–weighted sums, called *jet charge*, in dijet events using 20.3 fb^{-1} of data recorded with the ATLAS detector at $\sqrt{s} = 8$ TeV in pp collisions at the LHC. The jet charge distribution is unfolded to remove distortions from detector effects and the resulting particle level distribution is compared with several models. The p_T-dependence of the jet charge distribution average and standard deviation are compared to predictions obtained with several LO and NLO parton density functions and the best description of the data is found with CTEQ6L1. The data are also compared to different Monte Carlo predictions of QCD using various settings of the free parameters within these models. The choice of the strong coupling constant α_s used to calculate QCD radiation is found to have a significant impact on the predicted jet charge. There is evidence for a p_{T}-dependence of the jet charge distribution for a given jet flavor. In agreement with perturbative QCD predictions, the data show that the average jet charge of quark-initiated jets decreases in magnitude as the energy of the jet increases.

ATLAS-CONF-2015-025
24 July 2015

1 Introduction

Quarks and gluons produced in high–energy particle collisions hadronize before their electric charge can be directly measured. However, information about the electric charge is embedded in the resulting collimated sprays of quantum chromodynamically (QCD) colorless particles known as jets. One jet observable sensitive to the electric charge of quarks and gluons is the momentum–weighted charge sum constructed from charged particle tracks in a jet [1]. Called the *jet charge*, this observable was first used experimentally in deep inelastic scattering studies [2–8] to establish a relationship between the quark model and hadrons. Since then, jet charge observables have been used to tag the charge of b–quark jets [9–20] and hadronically decaying W bosons [21–23] as well as to distinguish hadronically decaying W bosons from generic QCD jets [24] and quark jets from gluon jets [22, 25–27].

The study presented in this paper is a measurement of the jet charge distribution in inclusive dijet events from pp collisions at the LHC. Inclusive dijet events provide a useful environment for measuring the jet charge as they are an abundant source of both gluon–initiated and quark–initiated jets. There are fewer theoretical ambiguities from close–by jets and large–angle radiation associated with assigning the jet flavour in events with two jets than in events with higher jet multiplicities. Furthermore, the transverse momentum (p_{T}) range accessible in dijet events spans a broad range, $O(10)$ GeV up to $O(1000)$ GeV. Since the initial state at the LHC has a net positive charge, the probability for positively charged quarks to be produced in pp collisions is higher than that for negatively charged quarks. The probability for collisions to involve a positively charged valence up quark in the proton increases with the parton center–of–mass energy $\sqrt{\hat{s}}$. Thus the average jet charge is expected to increase with $\sqrt{\hat{s}}$ if it is correlated with the quark charge. The Parton Distribution Functions (PDF) encode the probabilities to find gluons and certain flavours of quarks at given momentum fractions x of the proton. The momentum fractions of the two initial partons x_1, x_2, and the proton and parton center–of–mass energies $\sqrt{s}$ and $\sqrt{\hat{s}}$ are related by $\sqrt{\hat{s}} = \sqrt{x_1 x_2 s}$. PDFs are reasonably constrained [28–32] in the x ranges relevant for this study, $0.005 - 0.5$. However, the extent to which the energy–dependent evolution of PDFs can explain the trends in the jet charge distribution provides a consistency check using largely orthogonal information. PDFs are not the only non–perturbative input needed to model the jet charge distribution and its evolution with $\sqrt{\hat{s}}$. As a momentum–weighted sum over jet constituents, the jet charge is sensitive to the modeling of fragmentation. Previous studies have shown that there are qualitative differences in the charged particle track multiplicity inside jets predicted by the leading models of hadron production and the data [25]. Thus, a measurement of the jet charge distribution with a range of quark/gluon compositions can provide a constraint on models of jet formation.

While the change in jet parton flavor due to PDFs predicts most of the variation in the jet charge distribution as a function of $\sqrt{\hat{s}}$, there is a second contribution due to the energy–dependence of the fragmentation functions. Ratios of the charge distribution moments at different values of $\sqrt{\hat{s}}$ can be calculated perturbatively. Recent calculations [33, 34] in the context of Soft Collinear Effective Theory [35–38] show a significant reduction in the magnitude of the average jet charge for a given jet flavour. Information from PDFs can be used to extract the energy–dependence of the average jet charge in the data for direct comparisons to the predictions.

This paper presents a measurement of the p_{T}–dependence of the jet charge distribution mean and standard deviation in dijet events in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. The jet p_{T} is a measurable quantity that is strongly related to $\sqrt{\hat{s}}$. The average jet charge is extracted for both the leading and subleading jet and they are differentiated based on their relative orientation in rapidity. After a description of the ATLAS detector (Sec. 2), the data and simulated samples (Sec. 3) and the detector and

particle–level objects and selections used in the analysis (Sec. 4), Sec. 5 details the construction of the jet charge and some of its properties. In order for the measured jet charge distribution to be compared with particle–level models, the data are unfolded to remove distortions from detector effects, as described in Sec. 6. Systematic uncertainties on the measured jet charge spectra are discussed in Sec. 7 and the results are presented in Sec. 8.

2 The ATLAS Detector

ATLAS [39] is a general purpose detector designed to measure the properties of particles produced in high–energy pp collisions with nearly a full 4π coverage in solid angle. The innermost subsystem of the detector is a series of tracking devices used to measure charged particle trajectories bent in a 2 T axial field provided by a solenoid whose axis is parallel with the beam direction. This inner detector (ID) consists of a silicon pixel detector surrounded by a semiconductor microstrip detector (SCT) and a straw–tube tracker. It has full coverage[1] in ϕ and can detect particles out to $|\eta| < 2.5$. Charged particle tracks are reconstructed from all three ID components, providing measurements of the transverse momentum of tracks with a resolution $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 0.05\%\ \mathrm{GeV}/p_{\mathrm{T}} \oplus 1\%$. The track reconstruction algorithm fits five track parameters: d_0, z_0, ϕ, θ, and q/p, where d_0 and z_0 are the transverse and longitudinal impact parameters, respectively, q is the track charge and p is the track momentum. Excellent spatial precision is required to maintain a well–performing track reconstruction out to and exceeding charged particle p_{T} of 1 TeV, where track sagittas are $\lesssim 0.2$ mm.

Surrounding the ID and solenoid are electromagnetic and hadronic calorimeters to measure showers from both charged and neutral particles. The high granularity liquid–argon (LAr) sampling electromagnetic calorimeter is just beyond the solenoid and spans the range $|\eta| < 3.2$. Beyond the electromagnetic calorimeter is the two–component hadronic calorimeter that uses scintillator–tile sampling technology in the range $|\eta| < 1.7$ and LAr sampling technology for $1.5 < |\eta| < 3.2$. There are additional calorimeters in the forward region and a muon spectrometer with a toroidal magnet surrounding the calorimeters, but these are not used in this paper.

Due to the large event rate, not every collision can be recorded for processing offline. Events are selected using a three–level trigger system [40] that is hardware–based at the first level and software based for the two following levels. An event must satisfy all three trigger levels to be recorded for further processing. At each stage of the trigger, energy thresholds are placed on jet–like objects, with the similarity between online and offline jets increasing with each level. The first level makes decisions based on low granularity calorimeter towers with thresholds that are typically less than half of the energy required by jets at the second level. A simple jet reconstruction is used at the second level in regions around the jets identified by the first level. The third level (known as the Event Filter) clusters jets with the same algorithm as offline over the entire detector with thresholds that are typically 20–30 GeV higher than at level two. The single jet trigger thresholds increase at each level due both to differences in the jet reconstruction and calibrations, but also to meet the different bandwidth requirements at each trigger level. For low p_{T} dijets, the rate is far too large to save every event that passes the trigger selection and so most of the jet triggers are *pre–scaled* to artificially lower their recording rate.

[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 upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$.

3 Data and Simulated Samples

This measurement uses the full dataset of pp collisions recorded by the ATLAS detector in 2012, corresponding to an integrated luminosity of 20.3 fb^{-1} at a center–of–mass energy of $\sqrt{s} = 8$ TeV. Events are only considered if they are collected during stable beam conditions and pass all data–quality requirements [41]. To reject non–collision events, there must be a primary vertex reconstructed from at least two tracks each with $p_{\mathrm{T}} > 400$ MeV [42]. Due to the high instantaneous luminosity and the large total inelastic proton–proton cross section, on average there are about 21 simultaneous (*pileup*) collisions in each bunch crossing.

A set of single jet triggers is used to collect dijet events with high efficiency. Table 1 shows the collected luminosity for each trigger as well as the offline jet p_{T} ranges used, chosen such that the trigger is fully efficient. The highest p_{T} trigger is not pre–scaled.

Trigger threshold [GeV]	Offline Selection [GeV]	Luminosity [fb^{-1}]
25	[50,100]	7.84×10^{-5}
55	[100, 136]	4.42×10^{-4}
80	[136, 190]	2.32×10^{-3}
110	[190, 200]	9.81×10^{-3}
145	[200, 225]	3.63×10^{-2}
180	[225, 250]	7.88×10^{-2}
220	[250, 300]	2.61×10^{-1}
280	[300, 400]	1.16
360	≥ 400	20.3

Table 1: The single jet trigger menu used to collect dijet events with the 2012 dataset. The first column is the level three (Event Filter) jet p_{T} threshold and the second column is the offline leading jet p_{T} range corresponding to the given trigger. The luminosity collected with each trigger is in the last column. The total 2012 dataset was 20.3 fb^{-1}; the highest p_{T} trigger is not pre–scaled.

Monte Carlo (MC) simulated events are generated in p_{T} slices in order to ensure high statistics over a broad range of reconstructed jet p_{T}, given constraints on the available computing time and power. The p_{T} slices span the interval 0 to 5 TeV in ranges that approximately double with each increasing slice, starting with a range of size 8 GeV and ending with a range of size 2240 GeV. The baseline sample used for the measurement is generated with Pythia 8.175 [43] with the tune AU2 [44] and the next to leading order (NLO) PDF set[2] CT10 [47, 48]. A sample generated with a NLO matrix element from Powheg [49–52] with PDF set CT10 interfaced with Pythia 8.175 and the AU2 tune is also used for comparisons. Another high statistics sample is generated with Herwig++ 2.63 [53, 54] with tune EE3 [55] and leading order (LO) PDF set CTEQ6L1 [56] (particle–level samples with CT10 and EE4 are also used for comparisons). Both Pythia and Herwig are LO in perturbative QCD for the ($2 \rightarrow 2$) matrix element and leading logarithm (LL) in the parton shower. However, the ordering of emissions in the MC resummation in the shower differs between these two generators: p_{T}–ordered [57] for Pythia and angular–ordered [58] for Herwig. The phenomenological modeling of the non–pertubative physics also differs between Pythia and Herwig. In addition to different underlying event models (Ref. [59] for Pythia and an eikonal model [60] for Herwig) the hadronization models differ between Pythia (Lund string model [61])

[2] For a discussion on the use of NLO PDF sets with LO matrix elements, see Ref. [45, 46].

and Herwig (cluster model [62]). These two schemes are known[3] to predict different numbers of charged particles within jets and different distributions of the charged particle energies within jets, both of which are important for the jet charge.

All tunes of the underlying event that are used with Pythia and Herwig in this analysis use LHC data as input. As discussed in Sec. 1, the corrected data are compared to models with various PDF sets; for consistency, each set has a dedicated underlying event tune constructed in the same way from a fixed set of data inputs (AU2) described in detail in Ref. [44]. The PDF sets include LO sets CTEQ6L1 [56] and MSTW08LO [28] as well as NLO sets CT10 [47, 48], NNPDF21 NLO [64], and MSTW2008NLO [28].

All MC samples are processed using the full ATLAS detector simulation [65] based on GEANT4 [66].

4 Object Reconstruction and Event Selection

The reconstructed objects used for the jet charge as well as for the event selection are described in Sec. 4.1. The fiducial definition of the measurement, unfolded to particle level, is given in Sec. 4.2.

4.1 Object Reconstruction at Detector Level

Jets are clustered using the anti–k_t $R = 0.4$ [67] jet algorithm implemented in FastJet [68] from topological calorimeter–cell clusters [69], calibrated using the local cluster weighting (LCW) algorithm [70, 71]. An overall jet calibration accounts for residual detector effects as well as contributions from pileup [72]. Jets are required to be central ($|\eta| < 2.1$) so that their charged particles are within the $|\eta| < 2.5$ coverage of the ID.

When more than one primary vertex is reconstructed, the one with the highest $\sum p_{\mathrm{T}}^2$ over tracks is selected as the hard scatter vertex. Events are further required to have at least two jets with $p_{\mathrm{T}} > 50$ GeV and only the leading two jets are considered for the jet charge measurement. To select dijet topologies, the two leading jets must have $p_{\mathrm{T}}^{\mathrm{lead}}/p_{\mathrm{T}}^{\mathrm{sublead}} < 1.5$, where $p_{\mathrm{T}}^{\mathrm{lead}}$ and $p_{\mathrm{T}}^{\mathrm{sublead}}$ are the transverse momenta of the jets with the highest and second highest p_{T}, respectively. The jet with the smaller (larger) absolute pseudorapidity $|\eta|$ is classified as the more central (more forward) jet. A measurement of the more forward and more central jet charge distributions can exploit the rapidity–dependence of the jet flavor to extract information about the jet charge of a particular flavor. This is discussed in more detail in Sec. 4.2.

Tracks used to calculate the jet charge are required to have $p_{\mathrm{T}} \geq 500$ MeV, $|\eta| < 2.5$, and a χ^2 per degree of freedom (resulting from the track fit) less than 3.0. Additional quality criteria are applied to select tracks originating from the collision vertex and reject fake tracks reconstructed from random hits in the detector. In particular, tracks must be well–associated with the hard–scatter vertex with $|z_0 \sin(\theta)| < 1.5$ mm and $|d_0| < 1$ mm. Tracks must furthermore have at least one hit in the pixel detector and at least six hits in the SCT. The matching of tracks with the calorimeter-based jets is performed via Ghost association [73]: the jet clustering process is repeated with the addition of *ghost* versions of measured tracks that have the same direction but infinitesimally small p_{T}, so that they do not change the properties of the calorimeter jets. A track is associated with a jet if its ghost version is contained after re-clustering. The distribution of the number of tracks in two representative jet p_{T} ranges is shown in Fig. 1. The

[3] See for instance the appendix of Ref. [25].

number of tracks increases with jet p_T and the data fall between the predicted distributions of PYTHIA and HERWIG.

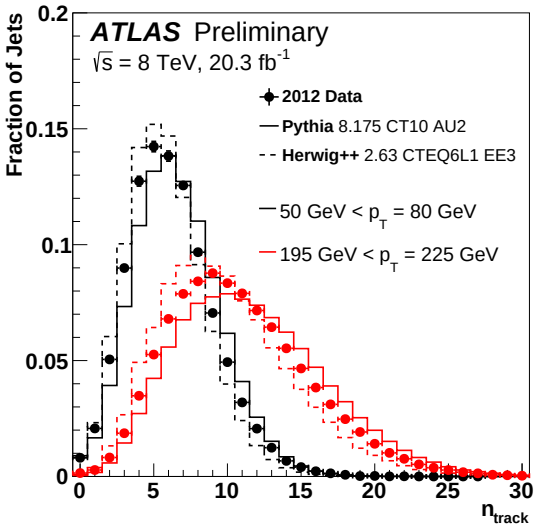
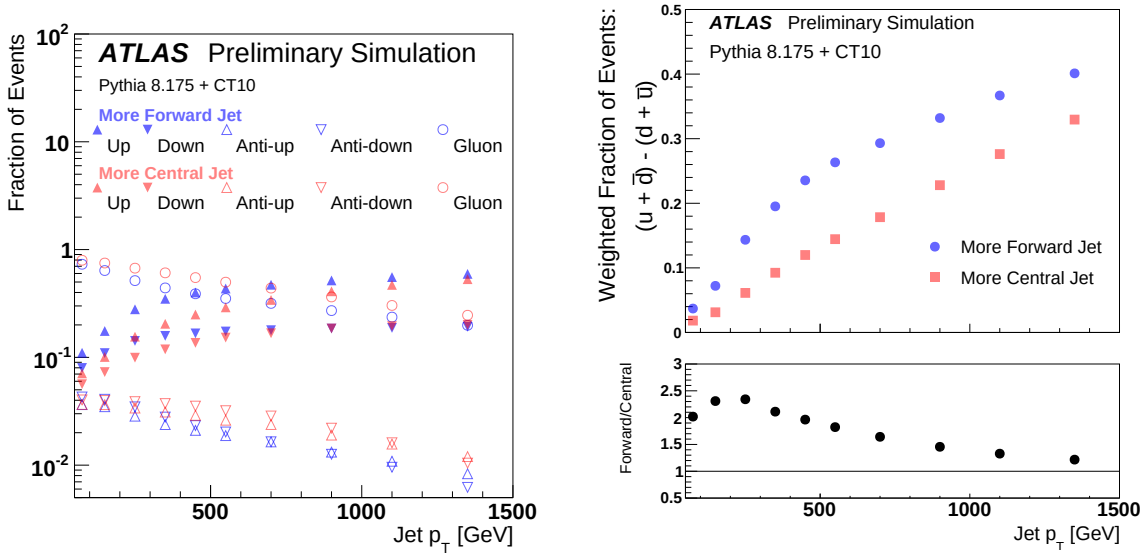


Figure 1: The distribution of the number of tracks associated with a jet in two example jet p_{T} ranges.

4.2 Object Definitions at Particle Level

The measurement is carried out within a fiducial volume matching the experimental selection to avoid extrapolation into unmeasured kinematic regions that require additional model–dependence and related uncertainties. Particle–level definitions of the reconstructed objects are chosen to be as close as possible to those described in Sec. 4.1. Particle–level jets are clustered from generated stable particles with a mean lifetime $\tau > 30$ ps, including muons and neutrinos. As with the detector–level jets, particle–level jets are clustered with the anti–k_t $R = 0.4$ algorithm. In analogy to the ghost–association of tracks to jets performed at detector–level, any charged particle clustered in a particle–level jet is considered for the jet charge calculation. There must be at least two jets with $|\eta| < 2.1$ and $p_{\mathrm{T}} > 50$ GeV. The two highest p_{T} jets must pass the same p_{T}–balance between the leading and subleading jet as at detector level ($p_{\mathrm{T}}^{\mathrm{lead}}/p_{\mathrm{T}}^{\mathrm{sublead}} < 1.5$). The flavor of a jet is defined at leading order in $m_{\mathrm{jet}}/E_{\mathrm{jet}}$, where factorization [74] results in well–defined quark/gluon cross sections order–by–order in perturbation theory. Due to the high–energy and well–separated nature of the selected jets, the hard–scatter quarks and gluons can be cleanly matched to the outgoing jets. While it is possible to classify jets as quark– or gluon–initiated beyond leading power [75], the classification is algorithm–dependent and unnecessary for the present considerations. In this analysis, the flavor of a jet is defined as that of the highest energy parton in simulation within a $\Delta R < 0.4$ cone of the particle–jet axis. The jet flavor depends on rapidity and so the two selected jets are classified as either more forward or more central; the more forward jet tends to be correlated to the higher x parton and is less likely to be a gluon jet. Figure 2 shows the flavor fraction for the more forward and more central particle–level jets passing the event selection. The p_{T} evolution of the sum of the flavor fractions weighted by the sign of the parton charge is shown in the right plot of Fig. 2. The differences between the flavor fractions are largest at low p_{T}, but the highest quark jet purity occurs

at high jet p_{T}.



(a) (b)

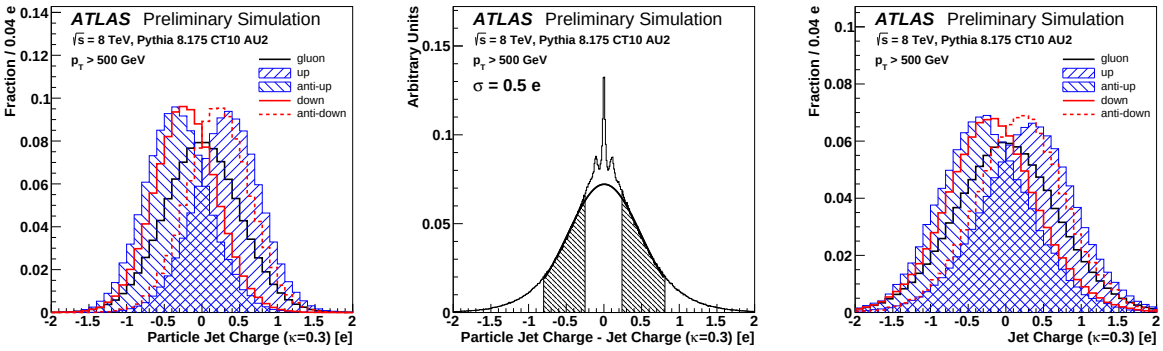
Figure 2: (a) For a given jet flavor, plotted is the fraction f of jets with that flavor passing the particle–level event selection and (b) The p_{T} evolution of the flavor fractions weighted by charge–sign: $f_{\mathrm{up}} + f_{\mathrm{anti-down}} - f_{\mathrm{anti-up}} - f_{\mathrm{down}}$. The differences between the flavor fractions are largest at low p_{T}, but the highest quark jet purity occurs at high jet p_{T}. The markers for the more forward and more central jets are distinguish by their color.

5 Constructing the Jet Charge

There is no unique way to define the jet charge. The most naïve construction is to add up the charge of all tracks associated to a jet. However, this scheme is very sensitive to lost or extraneous soft radiation. Therefore, a weighting scheme is introduced to suppress fluctuations. Using the tracks assigned to a jet by ghost association, the jet charge Q_J of a jet J is calculated using a transverse–momentum–weighting scheme [1]:

$$Q_J = \frac{1}{(p_{\mathrm{T}J})^{\kappa}} \sum_{i \in Tracks} q_i \times (p_{\mathrm{T,i}})^{\kappa}, \tag{1}$$

where *Tracks* is the set of tracks associated to jet J, q_i is the charge (in units of the positron charge) of track i with associated transverse momentum $p_{\mathrm{T,i}}$, κ is a free regularization parameter, and $p_{\mathrm{T}J}$ is the transverse momentum of the (calorimeter) jet. The distributions of Q_J for various jet flavors are shown in Fig. 3 for $\kappa = 0.3$. In the simulation, there is a clear relationship between the jet charge and the initiating parton charge, as up–quark jets tend to have a larger jet charge than gluon jets. Furthermore, gluon jets tend to have a larger jet charge than down–quark jets. However, the jet charge distribution is broad already at particle–level and the jet charge response ($Q_{\text{particle–level}} - Q_{\text{detector–level}}$) resolution is comparable to the differences in the means of the distributions for different flavors, so one can expect only small changes in the inclusive jet charge distribution for changes in the jet flavor composition. The three narrow distributions on top of the bulk response distribution in Fig. 3b are due to cases in which only one or two charged particles dominate the jet charge calculation at particle–level. The two off–center peaks are due to cases in which one of the two high p_{T}–fraction tracks is not reconstructed and the width of the two off–center and central peaks are due to the (single) track and jet p_{T} resolutions. The bulk response is fit to a Gaussian with standard deviation $\sigma \sim 0.5$ e (units of the positron charge).

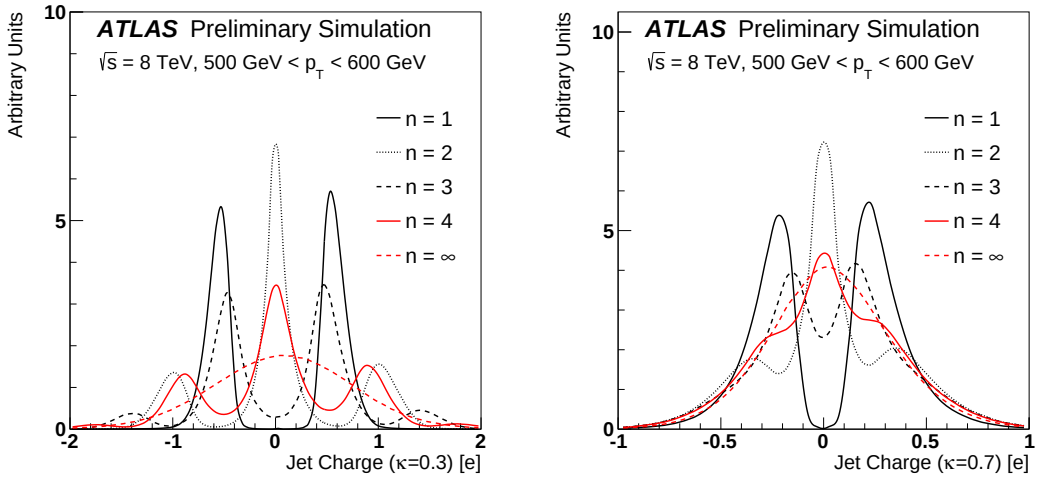


(a) (b) (c)

Figure 3: a (c): the particle–level (detector–level) jet charge distribution for various jet flavors in a sample of jets with $p_{\mathrm{T}} > 500$ GeV. b: the distribution of the jet–by–jet difference between the particle–level and detector–level jet charge distributions. The shaded region is used to fit a Gaussian to extract the bulk response resolution which is $\sigma \sim 0.5\ e$, where e is the electron charge.

The parameter κ in Eq. 1 regulates the sensitivity of the jet charge to soft radiation. For $\kappa > 0$, the jet charge is infrared safe[4]. Low values of κ enhance the contribution to the jet charge from low p_{T} particles while in the $\kappa \to \infty$ limit, only the highest p_{T} track contributes to the sum in Eq. 1. The dependence on the highest p_{T} tracks is demonstrated with the plots in Fig. 4 with the variable $Q_{j,n}$, which is the jet charge in Eq. 1, but built from the leading n tracks. The variable $Q_{J,1}$ is simply the weighted fragmentation function of the leading track p_{T} to the jet p_{T} with weight κ. The usual Q_J is recovered in the limit $n \to \infty$. Figure 4 shows the sequence $Q_{J,n}$ for $\kappa = 0.3$ and $\kappa = 0.7$. For lower values of κ, many tracks are required for the sequence of distributions to converge to the full jet charge. However, for $\kappa = 0.7$, the distribution converges quickly, indicating that only the highest p_{T} tracks are contributing. All reconstructed tracks are henceforth used when computing the jet charge, but the plots in Fig. 4 give an indication of the contribution of (relatively) high and low p_{T} tracks.

The jet–by–jet correlations of jet charge between various κ values are shown in Fig. 5. As expected from the differing sensitivity to soft and hard radiation, the further apart two κ values are, the less correlated the corresponding jet charge variables become. Dedicated studies [22] agree with theoretical predictions [33] that suggest that $\kappa \sim 0.5$ is the most sensitive to the charge of the parton initiating a jet. Therefore, the measurement presented in this paper uses $\kappa = 0.5$ in addition to $\kappa = 0.3$ and $\kappa = 0.7$ in order to maintain a broad sensitivity to both hard and soft radiation inside jets.



(a) (b)

Figure 4: The distribution of the jet charge built from the leading n tracks ($Q_{j,n}$) for $\kappa = 0.3$ (a) and $\kappa = 0.7$ (b). Note that the horizontal axis ranges are not the same. In this p_{T} range, the median number of tracks is about 15.

The reconstructed jet charge distributions for $\kappa = 0.3$ and $\kappa = 0.7$ are shown in Fig. 6 for events passing the selection described in Sec. 4. Section 6 describes how the jet charge moments are corrected for detector resolution and acceptance effects through an unfolding procedure.

[4] The jet charge is never collinear safe for $\kappa > 0$ and not even the sign of the jet charge is Lorentz invariant, though it is clearly invariant under longitudinal boosts.

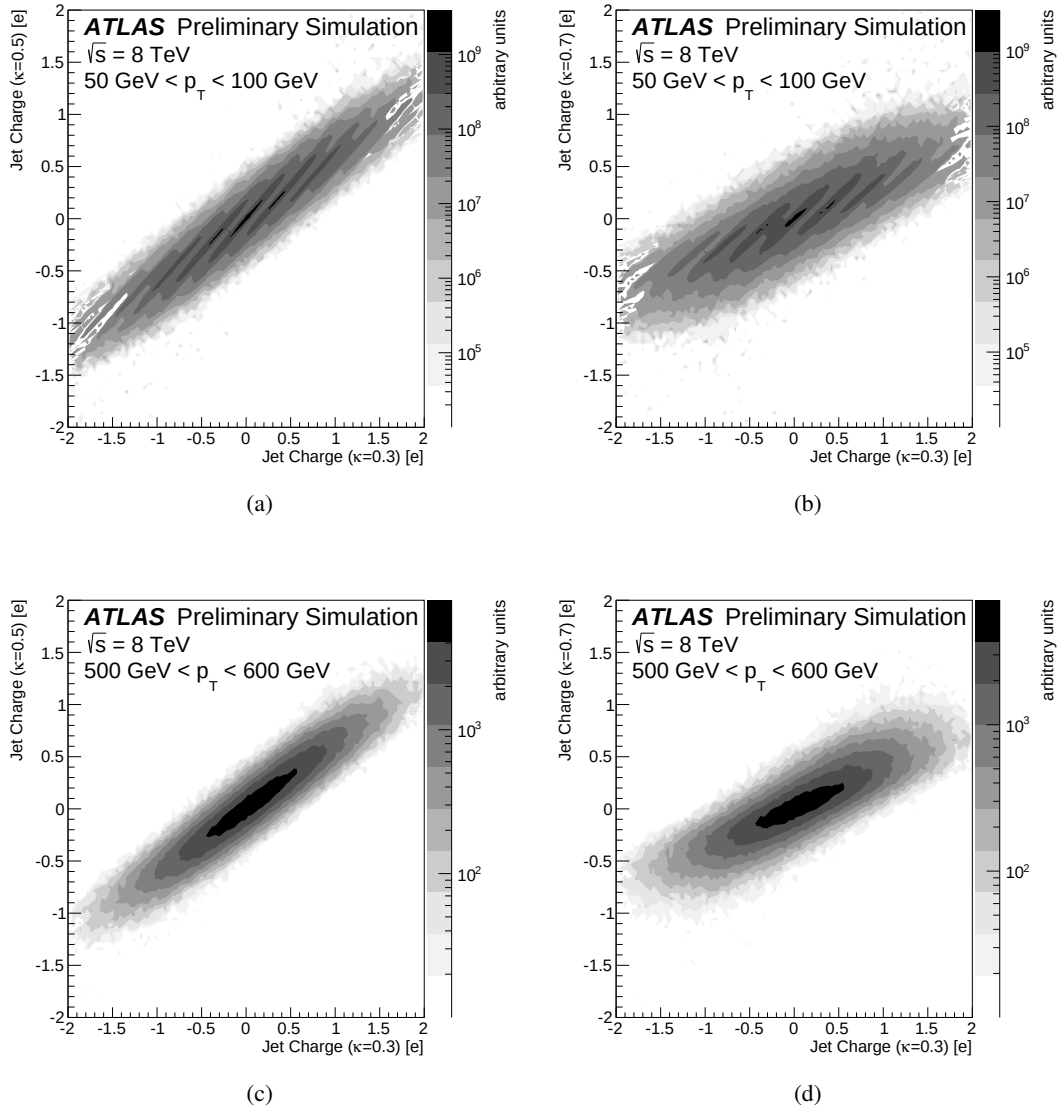


Figure 5: The jet–by–jet correlation between jet charge variables with a different momentum–weighting factor κ in simulation. (a) and (c) compare $\kappa = 0.3$ with $\kappa = 0.5$ while (b) and (d) compare $\kappa = 0.3$ with $\kappa = 0.7$. The jet charge becomes less correlated as the difference in κ values increases. The top row shows a low p_{T} bin, 50 GeV $< p_{\mathrm{T}} <$ 100 GeV while the bottom row shows a high p_{T} bin, 500 GeV $< p_{\mathrm{T}} <$ 600 GeV. The wave–like pattern in the top plots are from the low charged particle multiplicity bins (See Fig. 1).

6 Unfolding

The particle–level jet charge distribution mean and standard deviation are measured as a function of jet p_{T}. This is accomplished by unfolding a discretized two–dimensional distribution of the jet charge and the jet p_{T} and then computing the first two moments of the jet charge distribution in each bin of p_{T}. The jet charge distribution is discretized into fifteen bins in each of ten bins of the jet p_{T}. The jet charge

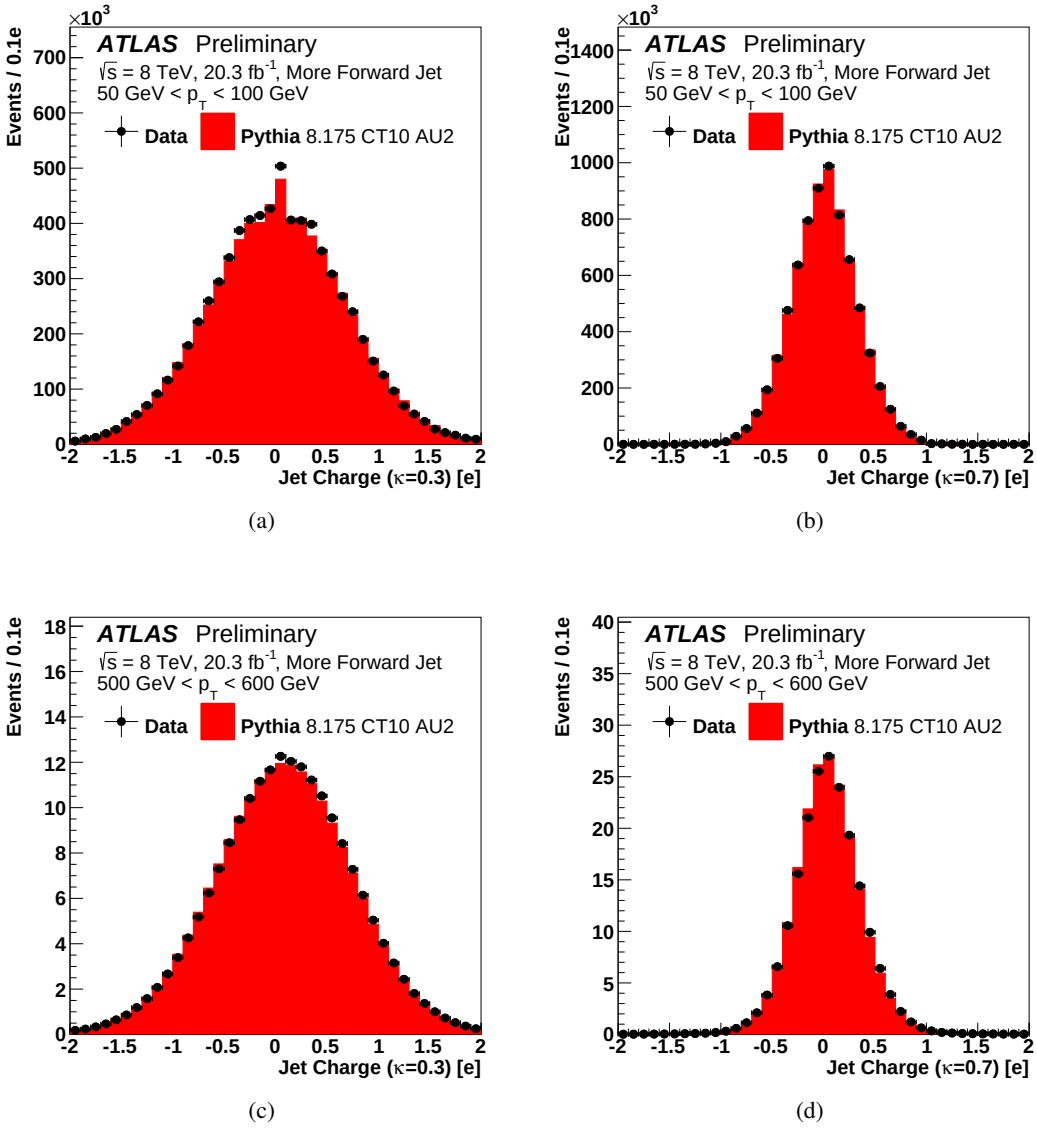


Figure 6: The detector–level jet charge distributions for 50 GeV $< p_{\mathrm{T}} <$ 100 GeV (a,b) and 500 GeV $< p_{\mathrm{T}} <$ 600 GeV (c,d) for the more forward jet and $\kappa = 0.3$ (a,c) and $\kappa = 0.7$ (b,d). The peak at zero in the top left plot is due to jets without any tracks.

mean is robust against the bias introduced from the discretization procedure with fewer than five charge bins per p_{T} bin required to have negligible bias after recovering the mean from the discretized distribution. However, the standard deviation of the jet charge distribution is sensitive to the discretization and requires about 15 charge bins[5] in order for the inherent bias due to discretization to be negligible. With only 10 charge bins, the discretization procedure introduces a percent–level bias. The jet charge spans the range $|Q_J| < 1.8$ for $\kappa = 0.3$, $|Q_J| < 1.2$ for $\kappa = 0.5$ and $|Q_J| < 0.9$ for $\kappa = 0.7$. Events in the overflow of the jet charge distribution are placed in the first or last bins. The p_{T} binning is given by: [50,100), [100,

[5] This number is what is used in the unfolding and is different than the number shown for illustration in e.g. Fig. 6.

200), [200, 300), [300, 400), [400, 500), [500, 600), [600, 800), [800, 1000), [1000, 1200), and [1200, 1500] GeV. Displayed in Fig. 7 are the detector–level data and simulation and particle–level simulation p_{T}–dependence of the jet charge distribution mean and standard deviation. The differences between the simulated detector– and particle–level distributions give a indication of the corrections required to account for detector acceptance and resolution effects in the unfolding procedure. The resolution is worse at higher p_{T} which can be seen from the larger differences in the detector– and particle–level jet charge distribution standard deviation in the highest p_{T} bins. An Iterative Bayesian (IB) technique [76] as implemented in the `RooUnfold` framework [77] is used to unfold the two–dimensional jet charge and jet p_{T} distribution. As a first step, the raw data are corrected using the simulation to account for events which pass the fiducial selection at detector–level, but not the corresponding selection at particle–level (*fake factor*). Then, the IB method iteratively applies Bayes' theorem using the *response matrix* to connect the prior to posterior at each step, with the nominal Pythia sample used for the initializing prior. The response matrix describes the bin migrations between the particle–level and detector–level two–dimensional jet charge and jet p_{T} distributions. While the response matrix is nearly diagonal, the resolution degrades at high p_{T} where more bin–to–bin migrations from particle– to detector–level occur.

The number of iterations in the IB method, which trades off unfolding bias with statistical fluctuations, is set to four in order to minimize the bias when unfolding pseudo–data Herwig with Pythia as a test of the methodology. The last step of the unfolding applies another correction from simulation to the unfolded data to account for the differential rate of events passing the particle–level selection but not the detector–level selection (*inefficiency factor*).

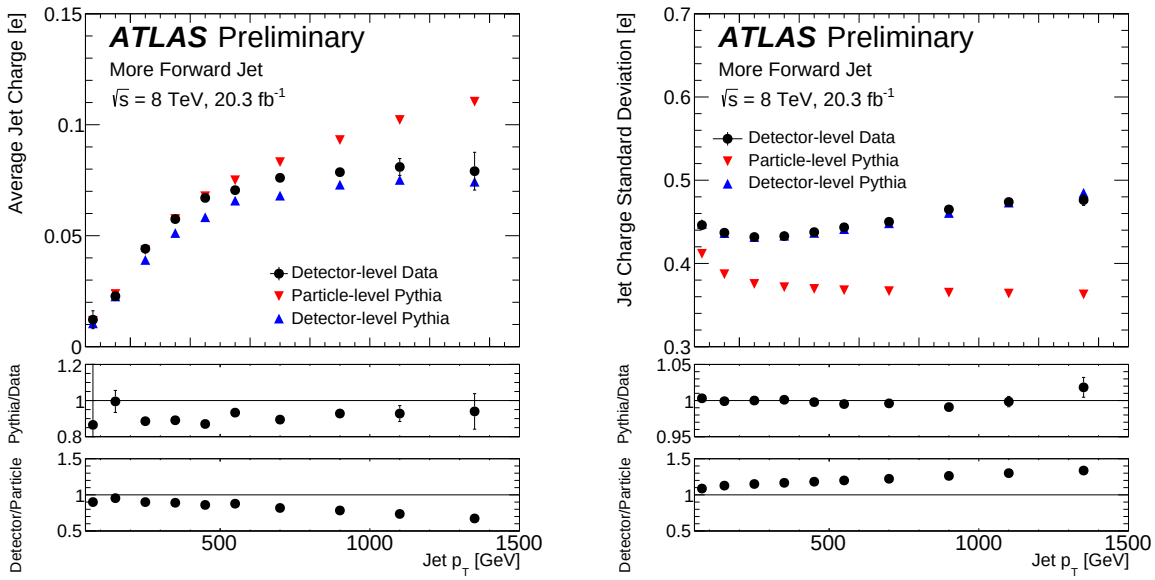


Figure 7: The detector–level (data and simulation) and particle–level jet charge distribution average (a) and standard deviation (b) as a function of the jet p_{T}. The lower ratio panel is constructed from the simulation. Bars on the data markers represent only the statistical uncertainties. For both (a) and (b), $\kappa = 0.5$.

7 Systematic Uncertainties

All stages of the jet charge measurement are sensitive to sources of potential bias. The three stages of the measurement are listed below, with an overview of the systematic uncertainties that impact the results at each stage:

Correction Factors: Fake and inefficiency factors are derived from simulation to account for the fraction of events which pass one of the reconstructed or particle–level fiducial selections, but not both. These factors are generally between 0.9 and 1.0 except in the first p_{T} bin, where threshold effects introduce corrections that can be as large as 20%. Experimental uncertainties correlated with the detector–level selection acceptance, such as the jet energy scale uncertainty, will result in uncertainties in these correction factors. An additional source of uncertainty on the correction factors is due to the explicit dependence on the particle–level spectrum of the jet charge and jet p_{T} distribution. A comparison of particle–level models (Pythia and Herwig) is used to estimate the impact on the correction factors.

Response Matrix: For events that pass both the reconstructed and particle–level fiducial selections, the response matrix describes migrations between bins when moving between the detector– and the particle–levels. The response matrix is taken from simulation and various experimental uncertainties on the jet charge and jet p_{T} spectra result in uncertainty on the matrix. Uncertainties can be divided into two classes: those impacting the calorimeter jet p_{T} and those impacting track reconstruction inside jets.

Unfolding Procedure: There is no unique way to extract the particle–level distribution of the jet charge and jet p_{T} from the reconstructed distributions. Every unfolding algorithm has regularization parameters to trade–off bias against statistical uncertainty. In the IB unfolding technique, the number of iterations and the prior distribution are the input parameters. A data–driven technique is used to estimate the potential bias from a given choice of prior and number of iterations [78]. The particle–level spectrum is re–weighted so that the simulated detector–level spectrum, when propagated through the response matrix, has improved agreement with the data. The modified detector–level distribution is unfolded with the nominal response matrix and the difference between this and the re–weighted particle–level spectrum is an indication of the bias due to the unfolding method (in particular, the choice of the prior).

The following two subsections describe the impact of the detector–related sources of systematic uncertainty in more detail. Uncertainties on the calorimeter jet p_{T} are described in Sec. 7.1 and the uncertainties related to tracking are described in Sec. 7.2. Summaries of the systematic uncertainties for the more forward jet and $\kappa = 0.5$ are found in Table 2 and Table 3 for the average jet charge and jet charge distribution standard deviation, respectively[6]. The uncertainties for the more central jet are similar.

7.1 Calorimeter Jet Uncertainties

Jets are calibrated so that the detector level p_{T} is an unbiased measurement of the particle–level jet p_{T} and various data–driven techniques are used to derive in–situ estimates of the difference in this calibration between the data and the simulation. Uncertainties in the energy scale and resolution of calibrated jets

[6] The uncertainties on the first p_T bin of the average jet charge are much larger than the other bins because the mean is small compared to the resolution.

Average Jet Charge Systematic Uncertainty [%]	Jet p_T Range [100 GeV]									
	[0.5,1]	[1,2]	[2,3]	[3,4]	[4,5]	[5,6]	[6,8]	[8,10]	[10,12]	[12,15]
Correction Factors	23	0.9	0.8	1.0	0.3	0.6	0.1	0.3	0.2	0.1
Total JES	8.8	3.8	0.9	0.8	1.1	1.1	0.7	0.7	0.4	0.9
JER	6.8	2.3	0.7	0.7	0.3	0.3	0.1	0.1	0.1	0.3
Charged Energy Loss	0.0	0.0	0.0	0.0	1.7	1.5	1.5	1.5	1.6	3.6
Track Multiplicity	1.5	0.1	0.6	1.1	0.8	0.6	1.2	1.4	2.1	2.9
Other Tracking	3.6	0.4	0.9	0.7	0.6	1.5	1.2	1.6	1.7	1.9
Unfolding Procedure	28	2.4	0.3	0.2	0.2	0.3	1.1	1.0	1.6	0.6
Total Systematic	38	5.1	1.8	2.0	2.4	2.6	2.6	2.9	3.6	5.1
Data Statistics	28	7.4	1.4	0.7	0.3	0.6	0.9	2.0	4.2	7.0

Table 2: A summary of all the systematic uncertainties and their impact on the jet charge mean for $\kappa = 0.5$ and the more forward jet. The acronym JES stands for jet energy scale, JER means jet energy resolution, and the unfolding factors are the fake and inefficiency corrections applied before/after the response matrix. The Other Tracking category includes uncertainty on the track reconstruction efficiency, track momentum resolution, charge mis-identification, and fake track rate. All numbers are given in percent.

Standard Deviation Systematic Uncertainty [%]	Jet p_T Range [100 GeV]									
	[0.5,1]	[1,2]	[2,3]	[3,4]	[4,5]	[5,6]	[6,8]	[8,10]	[10,12]	[12,15]
Correction Factors	0.9	0.1	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0
Total JES	1.9	1.5	1.1	1.1	0.9	1.0	0.8	0.7	0.5	0.5
JER	1.3	0.3	0.1	0.2	0.3	0.4	0.2	0.2	0.2	0.2
Charged Energy Loss	0.0	0.0	0.0	0.0	0.2	0.3	0.3	0.3	0.4	1.1
Track Multiplicity	0.2	0.3	0.2	0.1	0.0	0.1	0.2	0.2	0.3	0.2
Other Tracking	0.3	0.4	0.5	0.5	0.6	0.6	0.7	0.7	0.7	0.6
Unfolding Procedure	1.9	0.4	0.0	0.1	0.2	0.0	0.1	0.3	0.4	1.7
Total Systematic	3.1	1.6	1.2	1.2	1.2	1.3	1.1	1.1	1.0	2.1
Data Statistics	0.9	0.3	0.1	0.1	0.0	0.1	0.1	0.3	0.6	1.0

Table 3: A summary of all the systematic uncertainties and their impact on the jet charge distribution standard deviation for $\kappa = 0.5$ and the more forward jet. The acronym JES stands for jet energy scale, JER means jet energy resolution, and the unfolding factors are the fake and inefficiency corrections applied before/after the response matrix. The Other Tracking category includes uncertainty on the track reconstruction efficiency, track momentum resolution, charge mis-identification, and fake track rate. All numbers are given in percent.

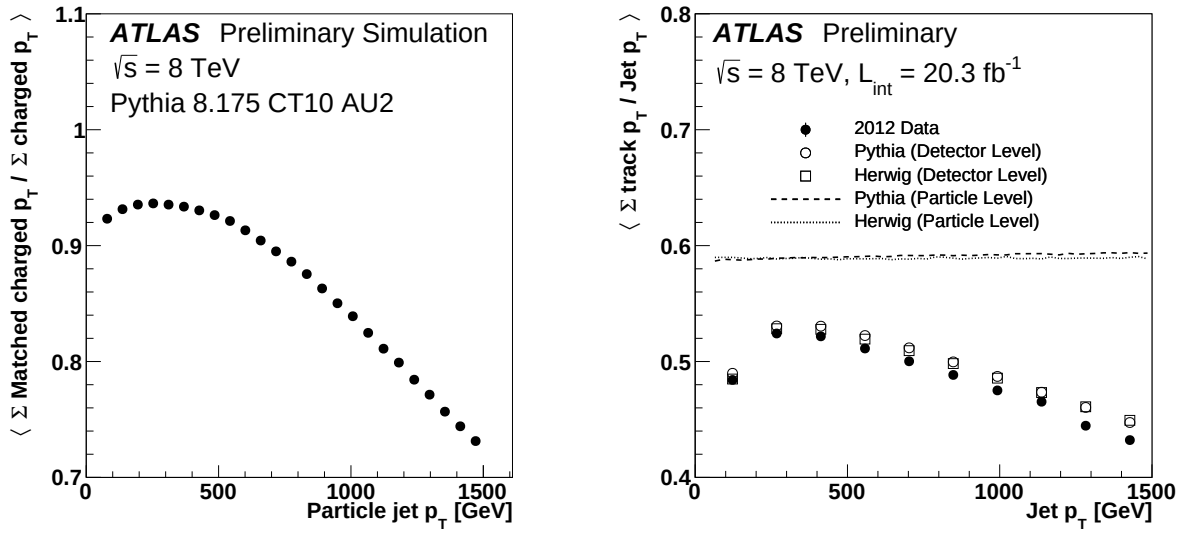
impact the jet charge in the normalization of Eq. 1 (but preserve the jet charge sign) as well as the binning for the 2D distribution. Complete details of this source of uncertainty can be found in Ref. [79]. There are many components of the jet energy scale uncertainty. The in–situ correction uncertainty is due to the modeling of Z bosons (low p_T) or photons (moderate p_T) produced in association with jets as well as the balance of multijet (high p_T) and dijet (high $|\eta|$) systems. There is also a contribution from the response of single hadrons [80]. Additional sources of uncertainty are due to the modeling of the in–time and out–of–time pileup corrections to the jet energy scale as well as differences in the response due to the flavor of the jet. To assess the impact of each component of the jet energy scale uncertainty, the jet energies in simulation are shifted according to the p_T– and η–dependent $\pm 1\sigma$ variations. For a fixed variation, the

response matrix, and fake and inefficiency factors are recomputed and the unfolding procedure is repeated. The resulting uncertainty on the jet charge distribution mean and standard deviation is about 1% or less for jet p_{T} above 200 GeV. The jet energy resolution uncertainty is derived using data–driven techniques in dijet events [81]. To assess the impact of a slightly larger jet energy resolution, jet energies are smeared according to p_{T}– and η–dependent factors and propagated through the entire unfolding machinery as for the jet energy scale uncertainty. The jet energy resolution uncertainty is subdominant to the jet energy scale uncertainty.

7.2 Tracking Uncertainties

Uncertainties on tracking are broken down into contributions related to (i) the efficiency of reconstructing tracks and (ii) measurements of those tracks that are successfully reconstructed. The uncertainty on the inclusive track reconstruction efficiency is due to the uncertainty in the material in the ID. The amount of material is known to within $\sim 5\%$ [82]. Simulated detector geometries with various levels of material in the ID within the measured uncertainties are used to estimate the track reconstruction efficiency uncertainty. These uncertainties are η– and p_{T}–dependent, ranging from $\lesssim 1\%$ for $|\eta| < 2.1$ to $\lesssim 4\%$ for $2.1 \leq |\eta| < 2.3$ and $\lesssim 7\%$ for $2.3 \leq |\eta| < 2.5$. The impact of the uncertainty is estimated by randomly dropping tracks within the p_{T}– and η–dependent probabilities leading to a $\lesssim 0.5\%$ uncertainty on the jet charge distribution mean and standard deviation. An additional uncertainty is required to assess the difference in efficiency between data and simulation due to the modeling of the track χ^2 per number of degrees of freedom (NDF) requirement. A requirement of $\chi^2/\mathrm{NDF} < 3$ is more than 99% efficient across jet and track p_{T}, but is generally higher in simulation than in data. The level of disagreement in the efficiency is $\lesssim 10\%$. The impact of this mis–modeling is evaluated by independently removing tracks with a probability that is 10% of the $\chi^2/\mathrm{NDF} < 3$ requirement inefficiency. As a result of this procedure, the jet charge distribution mean and standard deviation change by $\lesssim 0.1\%$ in most p_{T} bins.

In addition to the loss of tracks due to the material in the ID, tracks can be lost due to the busy environment inside the core of jets. This loss can be studied in simulation by comparing the reconstructed charged particle energy with the charged particle energy inside the corresponding particle–level jet. In order to remove the impact of the tracking resolution and the contribution from fake tracks already accounted for separately, reconstructed tracks in the simulation are matched with charged particles. The matching is performed by considering the energy deposited in the various layers of the ID by charged particles due to material interactions modelled with GEANT4. Figure 8a shows the ratio of charged particles that were matched to reconstructed tracks to all the charged particles as a function of the jet p_{T}. The rise at low p_{T} is due to losses as a result of hadronic interactions with the material in the ID. Beyond about 200 GeV, the fraction monotonically decreases due to the loss of tracks in the core of the jet. A related quantity is $\langle \Sigma \text{ track } p_{\mathrm{T}} \text{ / Jet } p_{\mathrm{T}} \rangle$, where the denominator is the calorimeter jet p_{T} and the numerator is a sum over tracks associated with the jet. Since the charged–to–neutral fraction of the energy is independent of p_{T}, a degradation in this ratio can provide information about the loss of tracks inside the core of a jet in data. The right plot of Fig. 8 shows the distribution of $\langle \Sigma \text{ track } p_{\mathrm{T}} \text{ / Jet } p_{\mathrm{T}} \rangle$ as a function of jet p_{T}. It exhibits very similar trends to the left plot of Fig. 8, and in fact the relative loss (fraction with respect to the peak) is similar. The MC under–estimates the loss by $\lesssim 1\%$. The impact of the charged energy loss inside the core of jets is estimated by randomly dropping tracks with a p_{T}–dependent probability such that the relative loss in the simulation matches that in the data. This uncertainty is negligible for jets with $p_{\mathrm{T}} < 400$ GeV, but is non–negligible for higher p_{T} jets, resulting in a $\lesssim 4\%$ uncertainty on the average jet charge in the highest p_{T} bin.



(a) (b)

Figure 8: (a): The p_{T}–weighted ratio of charged particles that were matched to reconstructed tracks to all the charged particles as a function of the particle–level jet p_{T}. (b): The distribution of $\langle \Sigma$ track p_{T} / Jet $p_{\mathrm{T}} \rangle$ as a function of jet p_{T} in both data and simulation. The energy ratio of charged particles to all particles is nearly 2/3 due to the number of pion species (as indicated by the straight lines for Herwig and Pythia predictions at particle level), but is not exactly so due to photons and kaons.

The momentum resolution of isolated tracks has been well measured in $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ events [83]. The scale and resolution of reconstructed muon candidates are shifted and smeared in the MC to account for differences between the data and the simulation for $m_{\mu\mu}$. Generic tracks are not corrected in the same way as muon candidates reconstructed from the ID (and ignoring the muon spectrometer). Thus, the correction factors are taken here as the systematic uncertainty on the momentum resolution. The momentum resolution is parameterized as a sum in quadrature of a p_{T}^{-1}–dependent term, a constant term, and a term linear in the track p_{T}, with coefficients r_0, r_1, and r_2, respectively. The first term accounts for fluctuations in the energy loss in the detector material, the second term captures effects due to multiple scattering, and the third term accounts for the intrinsic resolution caused by mis–alignment and the finite spatial resolution of ID hits. Unlike muon spectrometer tracks, ID tracks do not traverse a significant amount of material and so the energy–loss coefficient, r_0 and its uncertainty are neglected. The uncertainties on r_1, r_2 and the momentum scale s are estimated by randomly smearing the p_{T} of every track with $|\eta|$–dependent factors that are $\lesssim 2\%$ for track $p_{\mathrm{T}} < 100$ GeV and increase to 10–20% at 1 TeV depending on $|\eta|$. Propagating these variations through the unfolding procedure results in uncertainties that are subdominant to other uncertainties, but non–negligible ($\sim$ 2%) in the highest p_{T} bins for the average jet charge.

Aside from the track p_{T}, the other track parameter that is relevant for the jet charge is the track charge. Especially at high p_{T} when the tracks are very straight, the probability for mis–identifying the track charge can become non–negligible. Truth matched particles in the simulation, as described above, are used to study the charge flipping probability (charge mis–id rate) for non–fake tracks. The rate predicted from the simulation is $< 0.1\%$ for track $p_{\mathrm{T}} < 100$ GeV, 0.5% for 100 GeV $\leq p_{\mathrm{T}} <$ 200 GeV, 1% for

200 GeV $\leq p_{\mathrm{T}} <$ 300 GeV, 2% for 300 GeV $\leq p_{\mathrm{T}} <$ 400 GeV and 4% for $p_{\mathrm{T}} \geq$ 400 GeV. Dedicated studies of charge flipping in e.g. searches for same sign leptons [84] suggest that the mis–modeling of the charge flipping rate is (much) less than 50%. Therefore, the impact of charge flipping on the jet charge measurement is conservatively estimated by randomly flipping the charge of tracks by 50% of the charge mis–id rate. The impact on the measured jet charge mean and standard deviation is negligible.

Random combinations of hits in the detector can be combined to form a reconstructed track. Tracks resulting in particular from multi–particle trajectories that have kinks can result in very large reconstructed track p_{T}. The quality criteria are effective at mitigating the presence of fake tracks, which constitute less than $\lesssim 0.1\%$ of all reconstructed tracks. To determine the impact of fake tracks on the jet charge, fake tracks are randomly removed with a probability that is 50% of the rate in simulation. This results in a negligibly small uncertainty on the jet charge mean and a $\lesssim 0.5\%$ uncertainty on the standard deviation of the jet charge distribution.

The tracking uncertainties described so far take into account the resolution and efficiency of reconstruction of charged particle momenta. One last source of systematic uncertainty is the number of charged particles. The unfolding procedure uncertainty takes into account the uncertainty on the prior due to the charged particle multiplicity, but the jet charge resolution also changes with the charged particle multiplicity. To assess the impact on the response matrix of the mis-modeled charged particle multiplicity, the distribution of n_{track} is re-weighted in the simulation per p_{T} bin and the relative difference when unfolding the nominal Pythia with the re-weighted Pythia is taken as a systematic uncertainty[7]. This uncertainty is subdominant for the standard deviation across p_{T} and for the mean at low to moderate jet p_{T}. For the jet charge mean, the largest uncertainty is with the smallest κ and for large p_{T}, where it is 3-4% percent in the highest p_{T} bin for $\kappa = 0.3$ and $\kappa = 0.5$.

8 Results

The data passing the event selection criteria described in Sec. 4 are unfolded according to the procedure in Sec. 6 and the average and standard deviation of the jet charge distribution are computed as a function of the jet p_{T}. These results, along with the systematic uncertainties detailed in Sec. 7, are discussed in Sec. 8.1. PDF uncertainty and jet formation uncertainties in the theory predictions are compared to the unfolded data in Sections 8.2 and 8.3, respectively. Using PDF information as input, the average charge per jet flavor is extracted in Sec. 8.4 and its p_{T}–dependence is studied in Sec. 8.5.

8.1 Unfolded Jet Charge Spectrum

The unfolded jet charge mean is shown as a function of the jet p_{T} in Fig. 9 for $\kappa = 0.3$, 0.5 and 0.7. As expected, the average charge increases with jet p_{T} due to the increase in up–flavor jets from PDF effects. The average charge increases from $0.01e$ at $p_{\mathrm{T}} \sim 100$ GeV to $0.15e$ at $p_{\mathrm{T}} \sim 1.5$ TeV. Systematic uncertainties are included as bars on the unfolded data and are generally a few percent, except at low jet p_{T} where the average jet charge is close in an absolute sense to zero, and at high p_{T} where the tracking uncertainties are not negligible. The statistical uncertainty is estimated by repeating the measurement on an ensemble of bootstrapped datasets: each event is used in each pseudo–dataset n times, where n is a random number distributed according to a Poission distribution with mean one. The statistical uncertainty

[7] Since the prior is also changed, this uncertainty at least partially includes the unfolding procedure uncertainty.

is large ($\gtrsim$ 10%) in the first two p_{T} bins, then is $\sim$ 2% up until about 800 GeV where it grows again to $\sim$ 7% in the highest p_{T} bin. There are clear differences between the more forward and more central jet in the lower p_{T} bins due to the different shape of the up/down flavor fractions in those bins as shown in Fig. 2. The average jet charge in the data is generally larger in magnitude by about 10% than predicted by Pythia or Herwig with CT10.

Analogous results for the standard deviation of the jet charge distribution are shown in Fig. 10. Even though the standard deviation of the reconstructed jet charge distribution increases with jet p_{T} (Fig. 7), the particle–level value decreases and approaches an asymptote for $p_{\mathrm{T}} \gtrsim 300$ GeV. For $\kappa = 0.3$, the data falls between Pythia (larger standard deviation) and Herwig (smaller standard deviation), but this trend is less evident for larger κ values, suggesting a difference due to soft tracks.

Consistent with the expectation that the PDF and (nearly collinear) fragmentation are responsible for the jet charge distribution mean and standard deviation, there does not seem to be an effect from the Powheg NLO matrix in either Fig. 9 or Fig. 10.

8.2 Sensitivity of PDF Modeling

Variations in the PDF set impact the relative flavor fractions and thus in turn change the jet charge distribution. Such changes will not vary much with κ, since the PDF impacts the jet charge distribution mostly through the flavor fractions. Figures 11 and 12 compare the unfolded distributions of the jet charge distribution average and standard deviation with several PDF sets, with tuned predictions for Pythia for each PDF, and with the same AU2 tune. The sampling of PDF sets results in a significant spread for the average jet charge, but has almost no effect on the standard deviation. CTEQ6L1 describes the data best, however has a stronger p_{T} dependence. NLO PDFs such as CT10 agree better on the shape but are consistently below the data by about 10%.

8.3 Sensitivity of QCD Models and Tunes

The measurements presented in Sec. 8.1 show that there are qualitative differences between the data and the MC simulations, and comparisons in Sec. 8.2 suggest that variations in the PDF set cannot fully explain the differences. Differences in Sec. 8.1 between Pythia and Herwig suggest that some aspect of the modeling of fragmentation could lead to the observed differences between the simulation and the data. The plots in Fig. 13 and Fig. 14 show the measured average jet charge and jet charge distribution standard deviation, respectively, for $\kappa = 0.3$, 0.5, and 0.7, compared to various models for a fixed PDF set (CTEQ6L1). In addition to Pythia 8 and Herwig++ model predictions, Figures 13 and 14 contain the predictions for Pythia 6 using the Perugia 2012 tune [85] and the radHi and radLo Perugia 2012 tune variations. These Perugia tune variations test the sensitivity to higher/lower amounts of initial– and final–state radiation (via the scaling of α_s). For the jet charge mean, Pythia 6 with the P2012 radLo tune is very similar to Pythia 8 with the AU2 tune. The spread in the average jet charge due to the difference between the radHi and radLo tunes increases with κ, since suppression of soft radiation makes the jet charge distribution more sensitive to the modeling of the energy fraction of the leading emissions. For the jet charge distribution standard deviation, the sensitivity to the α_s scaling is large at both high and low κ. However, the sensitivity is inverted: radHi gives a larger standard deviation for $\kappa = 0.3$, but a lower standard deviation for $\kappa = 0.7$. Other Perugia 2012 tunes have been studied, testing the sensitivity to color–reconnection and multiple parton interactions, but the differences in the jet charge distribution

mean and standard deviation are small. The Perugia 2012 tunes do not fully capture the spread in non–perturbative effects, which is clear from the increasing difference between PYTHIA 8 and HERWIG++ for decreasing κ.

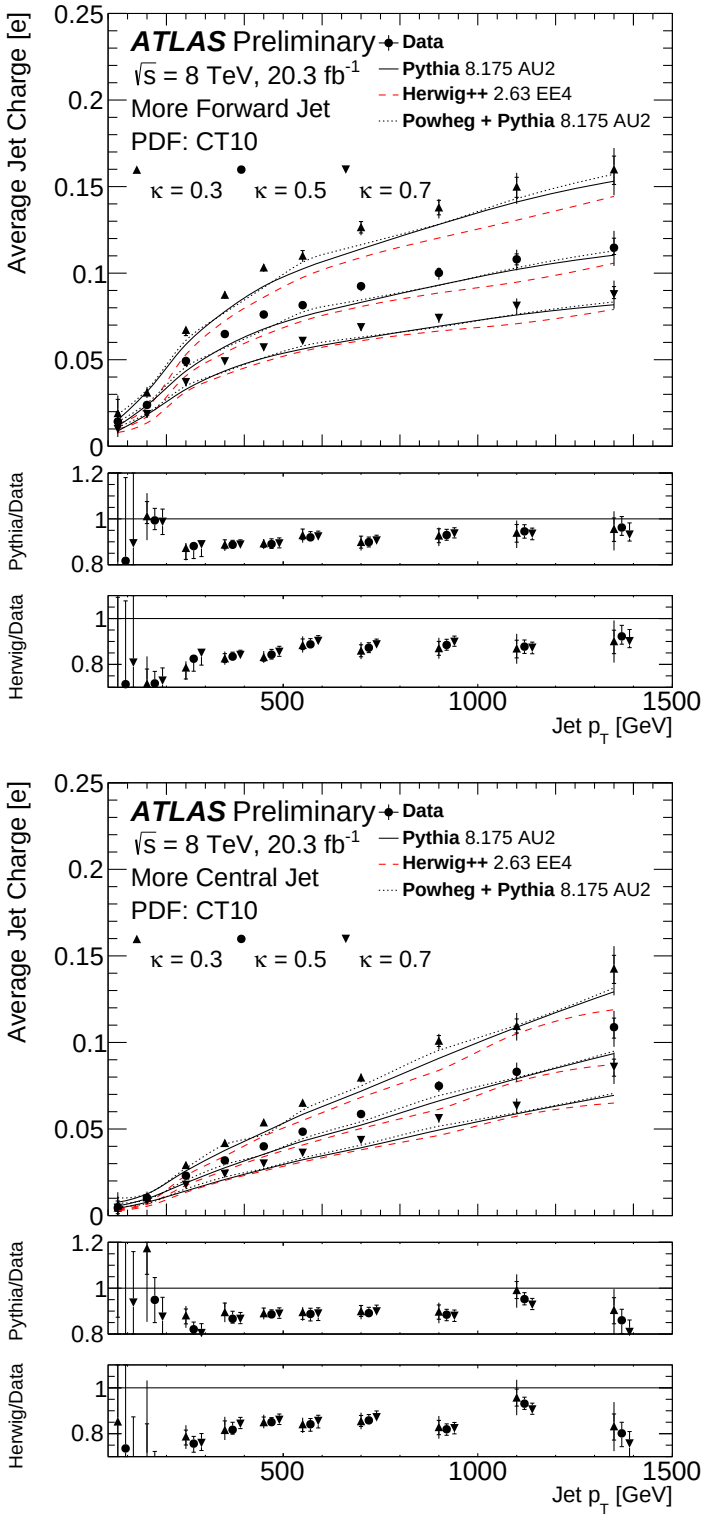


Figure 9: The measured average jet charge as a function of the jet p_{T} for $\kappa = 0.3, 0.5$, and 0.7 for the more forward jet (top) and the more central jet (bottom). The markers in the lower panel are artificially displaced horizontally to make differentiating the three κ values easier. The Powheg+Pythia curves are nearly on top of the Pythia curves. The crossed lines in the bars on the data indicate the systematic uncertainty and the full extent of the bars is the sum in quadrature of the statistical and systematic uncertainties.

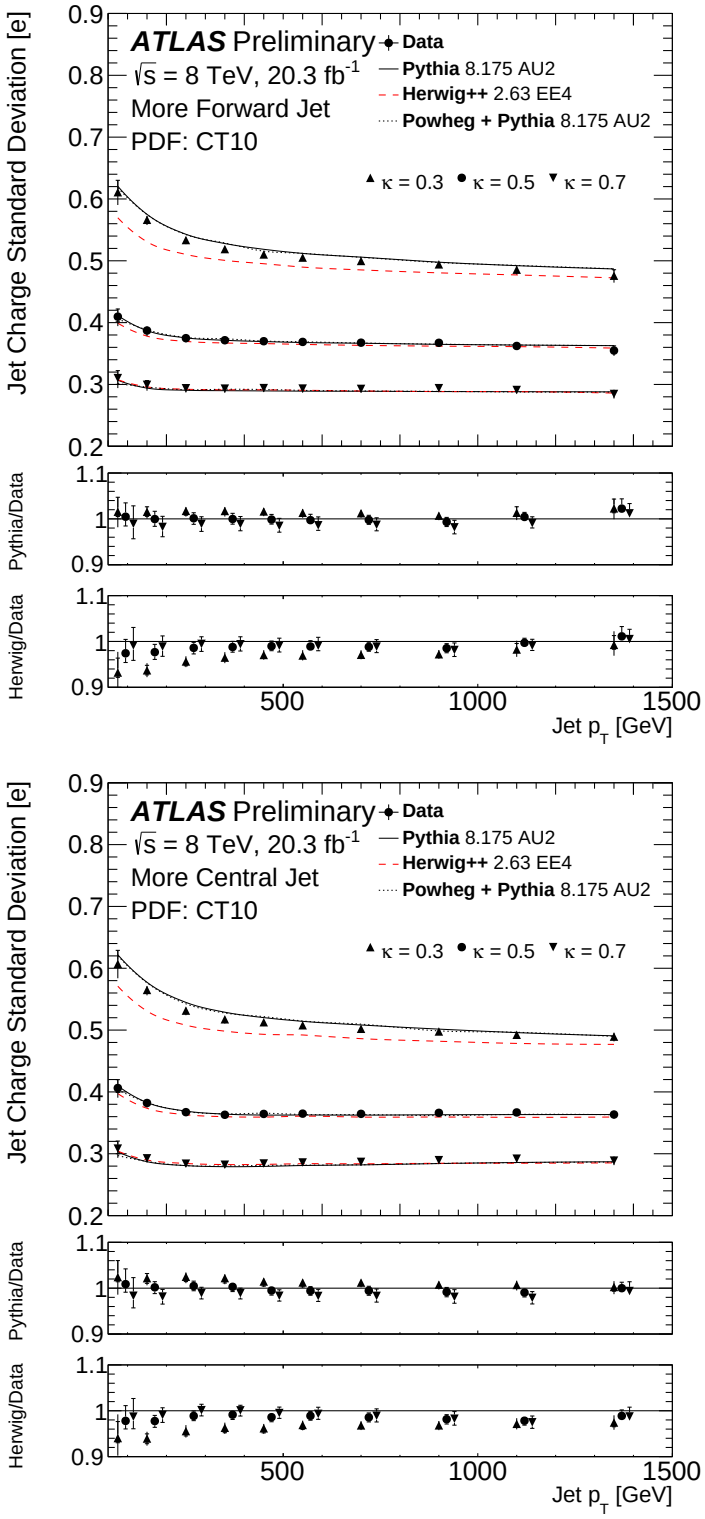
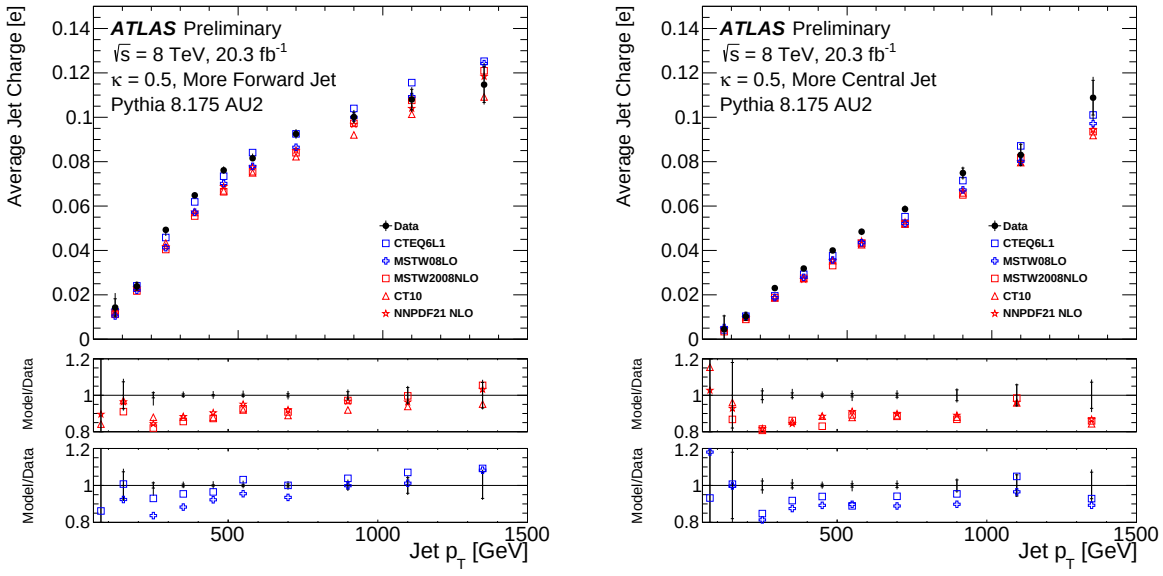
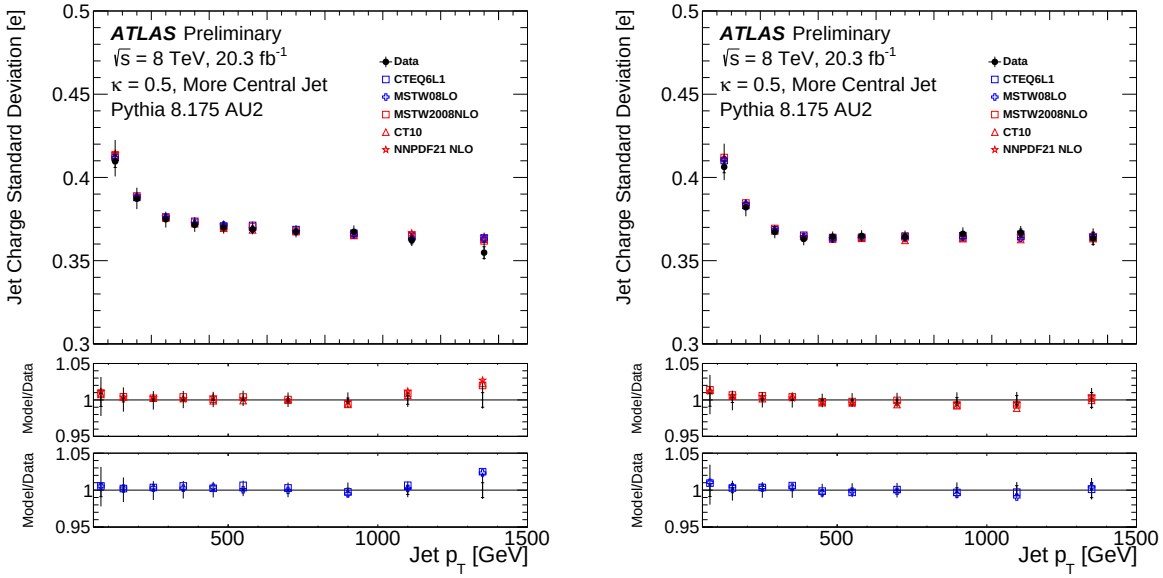


Figure 10: The measured standard deviation of the jet charge distribution as a function of the jet p_{T} for $\kappa = 0.3, 0.5,$ and 0.7 for the more forward jet (top) and the more central jet (bottom). The markers in the lower panel are artificially displaced horizontally to make differentiating the three κ values easier. The Powheg+Pythia curves are nearly on top of the Pythia curves. The crossed lines in the bars on the data indicate the systematic uncertainty and the full extent of the bars is the sum in quadrature of the statistical and systematic uncertainties.



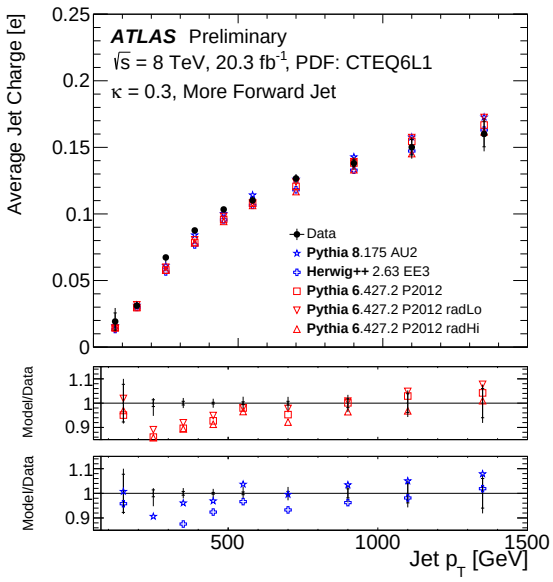
(a) (b)

Figure 11: The average jet charge for the more forward jet (a) and the more central jet (b) compared with theory predictions due to various PDF sets. The crossed lines in the bars on the data indicate the statistical uncertainty and the full extent of the bars is the sum in quadrature of the statistical and systematic uncertainties.

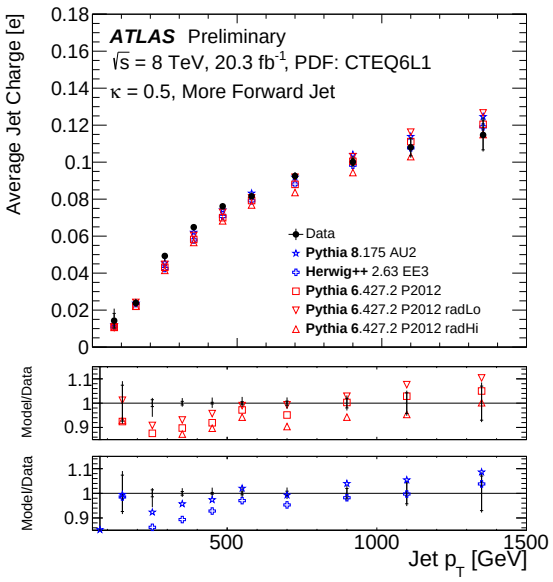


(a) (b)

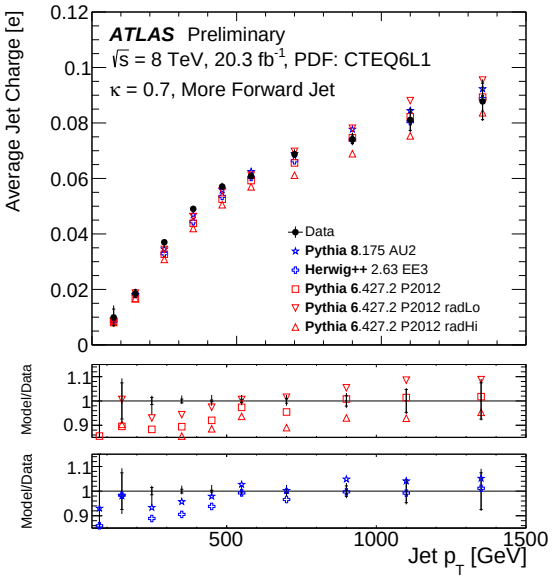
Figure 12: The standard deviation of the jet charge distribution for the more forward jet (a) and the more central jet (b) compared with theory predictions due to various PDF sets. The crossed lines in the bars on the data indicate the statistical uncertainty and the full extent of the bars is the sum in quadrature of the statistical and systematic uncertainties.



(a)

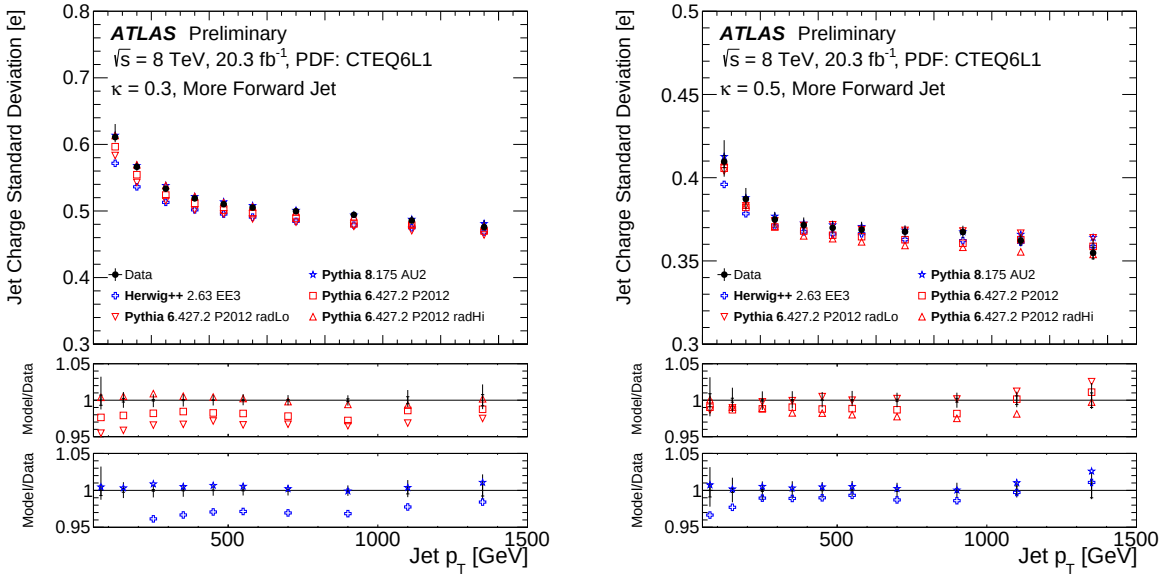


(b)

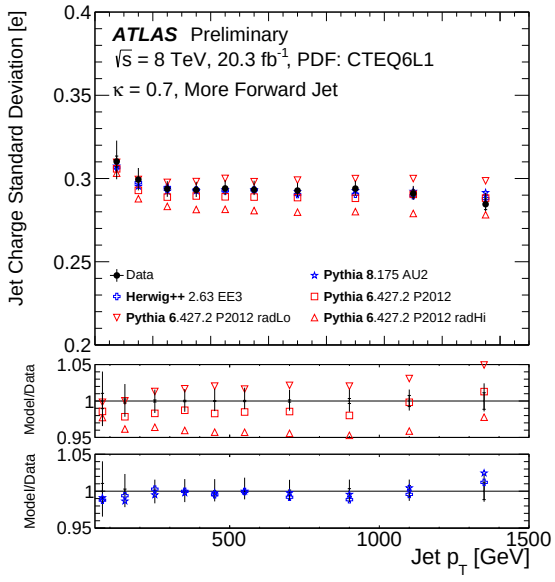


(c)

Figure 13: The average of the jet charge distribution for $\kappa = 0.3$ (a), 0.5 (b), and 0.7 (c) comparing various QCD MC models and tunes for the more forward jet. The crossed lines in the bars on the data indicate the statistical uncertainty and the full extent of the bars is the sum in quadrature of the statistical and systematic uncertainties.



(a) (b)



(c)

Figure 14: The standard deviation of the jet charge distribution for $\kappa = 0.3$ (a), 0.5 (b), and 0.7 (c) comparing various QCD MC models and tunes for the more forward jet. The crossed lines in the bars on the data indicate the statistical uncertainty and the full extent of the bars is the sum in quadrature of the statistical and systematic uncertainties.

8.4 Extraction of the Average Up–quark and Down–quark Jet Charges

In addition to understanding the trends in the jet charge distribution from PDFs, one can use PDFs to extract information about jets of a particular flavor. These *exclusive* interpretations rely on flavor fraction information in PDFs and matrix element calculations to extract the jet charge distribution for particular jet (anti–)flavors in each p_{T} bin. The required non–perturbative information is summarized in Fig. 2. Note that jets with flavors other than up/down/anti–up/anti–down/gluon are not included in Fig. 2 and give a negligible contribution ($\lesssim 2\%$) in the highest p_{T} bins.

One way of extracting the up and down average jet charges is to exploit the difference in flavor fractions shown in Fig. 2 between the more forward and the more central jets. Due to the p_{T}–balance requirement between the leading and subleading jet in the event selection, to a good approximation, the p_{T} spectrum is the same for the more forward and the more central jet. Assuming that the average jet charge of the sum of flavors that are not up/down/anti–up/anti–down is zero, in each bin i of p_{T}:

$$\begin{aligned}\langle Q_J^{\text{forward}}\rangle_i &= \left(f_{\text{up,i}}^{\text{forward}} - f_{\text{anti–up,i}}^{\text{forward}}\right) Q_i^{\text{up}} + (f_{\text{down,i}}^{\text{forward}} - f_{\text{anti–down,i}}^{\text{forward}}) Q_i^{\text{down}} \\ \langle Q_J^{\text{central}}\rangle_i &= \left(f_{\text{up,i}}^{\text{central}} - f_{\text{anti–up,i}}^{\text{central}}\right) Q_i^{\text{up}} + (f_{\text{down,i}}^{\text{central}} - f_{\text{anti–down,i}}^{\text{central}}) Q_i^{\text{down}},\end{aligned} \tag{2}$$

where $f_{y,i}^{x}$ is the fraction of flavor y in p_{T} bin i for the jet $x \in \{\text{more forward, more central}\}$ and Q_i^y is the average jet charge for such jets. The values $f_{y,i}^{x}$ are taken from simulation (Pythia CT10 AU2), which then allows an extraction of Q_i^y by solving the system of equations in Eq. 2. This extraction is performed separately in each p_{T} bin. Figure 15 shows the extracted up and down flavor jet charges in bins of jet p_{T}. At very high jet p_{T}, the absolute quark flavor fractions are large, but the difference between the more forward and more central jets is small and the statistical uncertainty is large. At low jet p_{T}, the difference between the more forward and more central jets is large, but the absolute quark flavor fraction is small and the statistical uncertainty is once again large because the mean jet charge is close to zero. The minimal uncertainty occurs in the bin 600 GeV $< p_{\mathrm{T}} <$ 800 GeV.

8.5 Dependence of the Up and Down Jet Charge on p_{T}

Using the methods of Sec. 8.4, one can examine the residual p_{T}–dependence of the average jet charge *after* accounting for PDF effects. The inclusive jet charge has been shown to increase with p_{T} due to a mixing of jet flavors and the following section investigates the p_{T}–dependence of a fixed jet flavor. Section 8.5.1 describes the theory predictions and the extracted p_{T}–dependence from the data is discussed in Sec. 8.5.2.

8.5.1 Theory Prediction

Recent theoretical studies have shown that the energy dependence of jet charge moments is calculable perturbatively [33, 34]. At leading power (treating $m_{\text{jet}}/E_{\text{jet}}$ as the expansion parameter), the charge defined with jet and track energies is equivalent to the one given in Eq. 1, since the jet opening angle is small. For jets defined by a radius R at an energy E, the average jet charge $\langle Q_J \rangle$ is given by

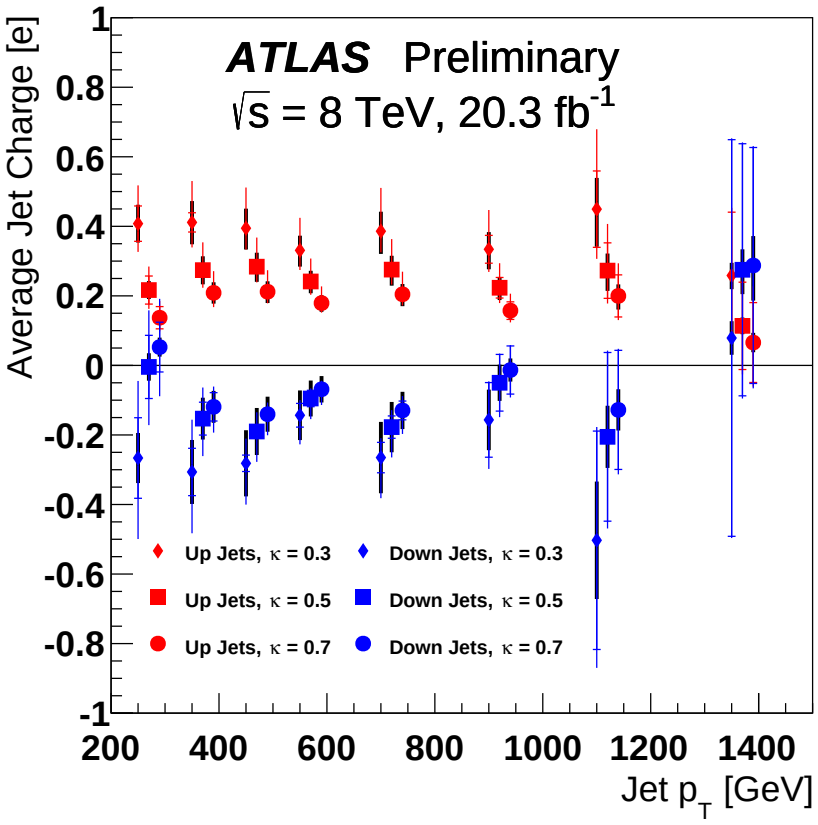


Figure 15: The extracted value of up and down quark flavor jet charges in bins of jet p_{T} for $\kappa = 0.3, 0.5$, and 0.7. The error bars include statistical, experimental systematic, and CT10 PDF uncertainties added in quadrature. The thick part of the error bar indicates the PDF contribution to the total uncertainty and the horizontal line on each error bar indicates the contribution from the statistical uncertainty. The first two p_{T} bins are excluded due to their very large uncertainties.

$$\langle Q_J \rangle = [1 + O(\alpha_s)] \sum_h Q_h \widetilde{D}_q^h(\kappa, E \times R), \tag{3}$$

where Q_h is the charge of hadron h and the functions $\widetilde{D}_q^h(\kappa, \mu)$ are the Mellin moments of the fragmentation functions

$$\widetilde{D}_q^h(\kappa, \mu) = \int_0^1 dx\, x^\kappa D_q^h(x, \mu). \tag{4}$$

The fragmentation functions $D_q^h(x, \mu)$ describe the probability for a hadron h to carry a momentum fraction x of a quark q at the energy scale μ [86, 87]. The $O(\alpha_s)$ correction is a small correction and it is

dominated by the uncertainty on the fragmentation functions [33]. Ratios of $\langle Q_J \rangle$ at different energies will result in a cancellation of the leading corrections and associated uncertainties. Soft corrections to the average jet charge are small both because of the additional suppression at $\kappa > 0$ compared to collinear radiation, but also because the leading soft emissions are made of gluons which carry no electric charge. Although gluons can split into quark and anti–quark pairs, the same number of quarks and antiquarks will go into the jet so that on average the jet charge is unchanged[8].

The leading energy dependence of the average jet charge is due to the derivative of the fragmentation functions with respect to E, which is determined by the renormalization group equations for $\widetilde{D}_q^h(\kappa, \mu)$:

$$\frac{d}{d\ln\mu}\widetilde{D}_q^h(\kappa,\mu) = \frac{\alpha_s}{\pi}\widetilde{P}_{qq}(\kappa)\widetilde{D}_q^h(\kappa,\mu) + O(\alpha_s^2), \tag{5}$$

where $\widetilde{P}_{qq}(\kappa)$ is the moment of the leading–order splitting function:

$$\widetilde{P}_{qq}(\kappa) = C_F \int_0^1 dz (z^\kappa - 1)\frac{1+z^2}{1-z}. \tag{6}$$

Recalling that $\frac{E}{p_T}$ is constant at leading power for all particles in the jet, the prediction for the *scale violation* of a quark jet is given by

$$\frac{p_T}{\langle Q_\kappa \rangle}\frac{d}{dp_T}\langle Q_\kappa \rangle = \frac{\alpha_s}{\pi}\widetilde{P}_{qq}(\kappa) \equiv c_\kappa \approx \begin{cases} -0.024 \pm 0.004 & \kappa = 0.3 \\ -0.038 \pm 0.006 & \kappa = 0.5 \\ -0.049 \pm 0.008 & \kappa = 0.7 \end{cases} \tag{7}$$

where the last form are numerical approximations setting $\mu = E \times R$ equal to 50 GeV and 500 GeV with their average giving the central value ($\alpha_s(50\text{ GeV}) = 0.130$ and $\alpha_s(500\text{ GeV}) = 0.094$). Experimentally, one measures linear combinations of quark and gluon jets with the fractions of different partonic flavors varying with energy. The next section discusses how the scaling violation parameter can be extracted from the data.

8.5.2 Extraction from the Data

Since $c_\kappa \ll 1$ from Eq. 7, one can approximate a linear dependence on c_κ:

$$\langle Q_J \rangle(p_T) = \bar{Q}(1 + c_\kappa \log(p_T/\bar{p}_T)) + O(c_\kappa^2), \tag{8}$$

where $\bar{Q} = \langle Q_J \rangle(\bar{p}_T)$ for some fixed (but arbitrary) transverse momentum, $p_{T,0}$. Therefore, for a fixed p_T bin i, the measured charge is given as a superposition of the average jet charge for various jet flavors:

[8] This is not strictly true, as the soft radiation pattern depends on the partonic colors which are correlated with electric charge, but these correlations are negligible compared to the other uncertainties.

$$\langle Q_i \rangle \approx \sum_f \alpha_{f,i} \bar{Q}_f (1 + c_\kappa \log(p_{T,i}/\bar{p}_T)), \tag{9}$$

where $\alpha_{f,i}$ is the fraction of flavor f in bin i, $\bar{Q}_f$ is the average jet charge of flavor f and $\bar{p}_T$ is a fixed transverse momentum. Fitting the model in Eq. 9 directly to the data to extract $\bar{Q}_f$ is not practical because there are three parameters and only 10 p_T bins, some of which have very little sensitivity due to low fractions α or large uncertainties on $\langle Q_J \rangle$. One way around this is to extract $\bar{Q}_f$ in one fixed bin of transverse momentum (denoted $\bar{p}_T$) as described in Sec. 8.4. Then Eq. 9 is highly constrained, with only one parameter for which each other bin of p_T gives an estimate. A χ^2 fit is performed across all p_T bins and for both the more forward and the more central jet to arrive at the results presented in Fig. 16. The systematic uncertainties are propagated through the fit treated as fully correlated between bins and the statistical uncertainty is treated coherently by bootstrapping. The data support the prediction that $c_\kappa < 0$ and $\partial c_\kappa / \partial \kappa < 0$. Linear correlations between κ values can be determined using the bootstrapped datasets: about 0.9 between $c_{0.3}$ and $c_{0.5}$ as well as between $c_{0.5}$ and $c_{0.7}$, while the correlation is about 0.7 between $c_{0.3}$ and $c_{0.7}$. Thus, the three points are quite correlated, but there is additional information from considering more than one κ value.

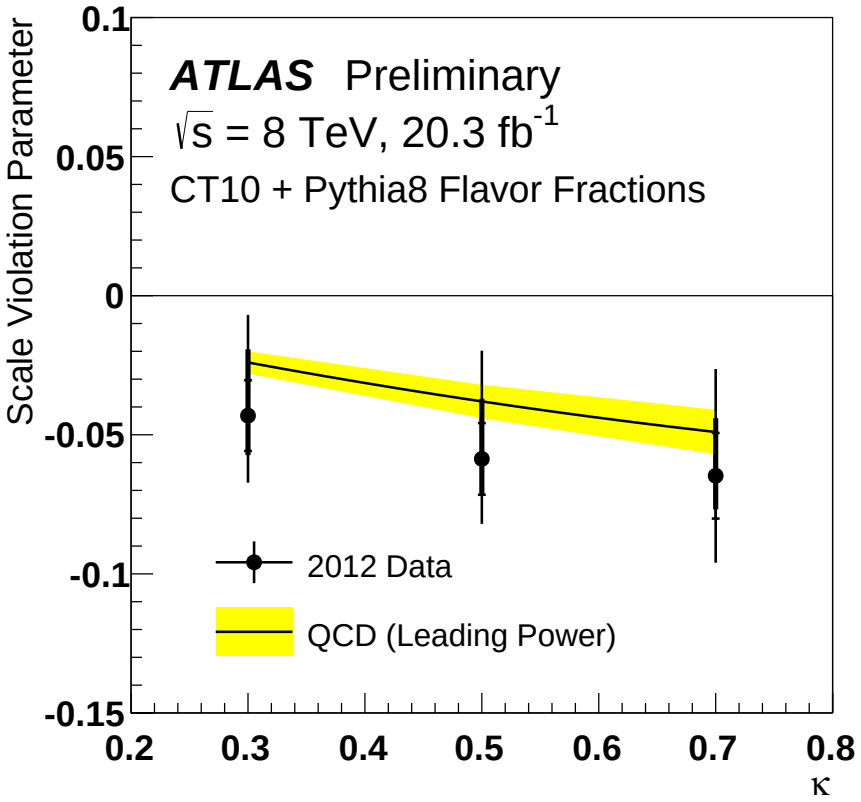


Figure 16: The extracted values of the scale violation parameter c_κ from the data compared to theoretical calculations [33, 34]. The error bars include statistical, experimental systematic, and PDF uncertainties added in quadrature. The thick part of the error bar indicates the PDF contribution to the total uncertainty and the horizontal line on each error bar indicates the contribution from the statistical uncertainty (each shown without adding quadrature with any other source of uncertainty).

9 Summary

Presented in this paper is a measurement of the particle–level p_{T}–dependence of the jet charge distribution mean and standard deviation in dijet events at $\sqrt{s} = 8$ TeV. The measured jet charge distribution is unfolded to correct for the detector acceptance and resolution for direct comparison to particle–level models. The average jet charge is systematically higher in the data than in the simulations with CT10, across jet p_{T}. This bias is reduced when CTEQ6L1 is used and for $p_{\mathrm{T}} > 500$ GeV. For large values of the momentum weighting parameter κ, the standard deviation of the jet charge distribution is well modelled within uncertainties. There are qualitative differences between Pythia, Herwig, and the data for the jet charge distribution standard deviation with the lowest measured value of $\kappa = 0.3$. Comparisons are made at particle–level between the measured jet charge distribution and various PDF sets and models of jet formation. There is no clear explanation for the differences between the data and the simulation, as PDF variations cannot explain (all of) the deviations. Taking the PDFs as inputs, the average up and down flavor jet charges are extracted as a function of p_{T} and are compared with predictions for scale violation. The data show that the average up and down quark charge decrease with p_{T} and this decrease increases with κ, as predicted. The particle–level spectra are publicly available for further interpretation and can serve as a benchmark for future measurements of the evolution of non–perturbative jet observables to validate QCD MC predictions and tune their free model parameters.

Acknowledgements

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

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, 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; EPLANET, ERC and NSRF, European Union; IN2P3-CNRS, CEA-DSM/IRFU, France; GNSF, Georgia; BMBF, DFG, HGF, MPG and AvH Foundation, Germany; GSRT and NSRF, Greece; RGC, Hong Kong SAR, China; ISF, MINERVA, GIF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; BRF and RCN, Norway; MNiSW and NCN, Poland; GRICES and FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MSTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, 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.

References

[1] R. D. Field and R. P. Feynman, *A Parameterization of the properties of Quark Jets*, Nucl. Phys. **B136** (1978) 1.

[2] J. Berge et al., *Net charge in deep inelastic antineutrino-nucleon scattering*, Phys. Lett. **B91** (1980) 311.

[3] J. Berge et al., *Quark jets from antineutrino interactions (I). Net charge and factorization in the quark jets*, Nucl. Phys. **B184** (1981) 13.

[4] P. Allen et al., *Multiplicity distributions in neutrino-hydrogen interactions*, Nucl. Phys. **B181** (1981) 385.

[5] P. Allen et al., *Charge properties of the hadronic system in νp and $\bar{\nu} p$ interactions*, Phys. Lett. **B112** (1982) 88.

[6] J. Albanese et al., *Quark charge retention in final state hadrons from deep inelastic muon scattering*, Phys. Lett. **B144** (1984) 302.

[7] S. Barlag et al., *Charged hadron multiplicities in high energy $\bar{\nu}_\mu n$ and $\bar{\nu}_\mu p$ interactions*, Z. Phys. **C11** (1982) 283.

[8] R. Erickson et al., *Charge Retention in Deep-Inelastic Electroproduction*, Phys. Rev. Lett. **42** (13 1979) 822.

[9] K. Abe et al. (**SLD** Collaboration), *Measurement of the Parity-Violation Parameter A_b from the Left-Right Forward-Backward Asymmetry of b Quark Production in Z^0 Decays Using a Momentum-Weighted Track-Charge Technique*, Phys. Rev. Lett. **74** (15 1995) 2890.

[10] W. Braunschweig et al. (**TASSO** Collaboration), *Measurement of the asymmetry of b quark production in e^+e^- annihilation of $\sqrt{s}$ = 35-GeV*, Z. Phys. **C48** (1990) 433.

[11] P. Abreu et al. (**DELPHI** Collaboration), *A Measurement of $sin^2\theta(W)$ from the charge asymmetry of hadronic events at the Z^0 peak*, Phys. Lett. **B277** (1992) 371.

[12] D. Decamp et al. (**ALEPH** Collaboration), *Measurement of charge asymmetry in hadronic Z decays*, Phys. Lett. **B259** (1991) 377.

[13] P. D. Acton et al. (**OPAL** Collaboration), *A measurement of the forward-backward charge asymmetry in hadronic decays of the Z^0*, Phys. Lett. **B294** (1992) 436.

[14] R. Akers et al. (**OPAL** Collaboration), *Measurement of the time dependence of $B^0_d \leftrightarrow \overline{B^0_d}$ mixing using a jet charge technique*, Phys. Lett. **B327** (1994) 411.

[15] P. Abreu et al. (**DELPHI** Collaboration), *Measurement of the B^0_d oscillation frequency using kaons, leptons and jet charge*, Z. Phys. **C72** (1996) 17.

[16] F. Abe et al. (**CDF** Collaboration), *Measurement of $B^0 - \overline{B}^0$ flavor oscillations using jet-charge and lepton flavor tagging in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.8 TeV*, Phys. Rev. **D60** (1999) 072003, arXiv: hep-ex/9903011 [hep-ex].

[17] V. Abazov et al. (**D0** Collaboration), *Experimental discrimination between charge 2e/3 top quark and charge 4e/3 exotic quark production scenarios*, Phys. Rev. Lett. **98** (2007) 041801, arXiv: hep-ex/0608044 [hep-ex].

[18] T. Aaltonen et al. (**CDF** Collaboration), *Exclusion of exotic top-like quarks with -4/3 electric charge using jet-charge tagging in single-lepton ttbar events at CDF*, Phys. Rev. **D88**.3 (2013) 032003, arXiv: 1304.4141 [hep-ex].

[19] **ATLAS** Collaboration, *Measurement of the top quark charge in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, JHEP **1311** (2013) 031, arXiv: 1307.4568 [hep-ex].

[20] **CMS** Collaboration, *Constraints on the Top-Quark Charge from Top-Pair Events*, CMS-PAS-TOP-11-031 (2012), URL: https://cds.cern.ch/record/1429970.

[21] R. Barate et al. (**ALEPH** Collaboration), *Measurement of triple gauge-boson couplings at 172 GeV*, Phys. Lett. **B422** (1998) 369.

[22] **ATLAS** Collaboration, *Jet Charge Studies with the ATLAS Detector Using $\sqrt{s}$ = 8 TeV Proton-Proton Collision Data*, ATLAS-CONF-2013-086 (2013), URL: https://cds.cern.ch/record/1572980.

[23] **CMS** Collaboration, *V Tagging Observables and Correlations*, CMS-PAS-JME-14-002 (2014), URL: https://cds.cern.ch/record/1754913.

[24] **CMS** Collaboration, *Identification techniques for highly boosted W bosons that decay into hadrons*, JHEP **1412** (2014) 017, arXiv: 1410.4227 [hep-ex].

[25] **ATLAS** Collaboration, *Light-quark and gluon jet discrimination in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, Eur. Phys. J. **C74**.8 (2014) 3023, arXiv: 1405.6583 [hep-ex].

[26] P. D. Acton et al. (**OPAL** Collaboration), *A Study of the electric charge distributions of quark and gluon jets in hadronic Z0 decays*, Phys. Lett. **B302** (1993) 523–532.

[27] G. Arnison at al. (**UA1** Collaboration), *Analysis of the Fragmentation Properties of Quark and Gluon Jets at the CERN SPS $p\bar{p}$ Collider*, Nucl. Phys. **B276** (1986) 253–271.

[28] G. Watt and R. Thorne, *Study of Monte Carlo approach to experimental uncertainty propagation with MSTW 2008 PDFs*, JHEP **1208** (2012) 052, arXiv: 1205.4024 [hep-ph].

[29] L. Harland-Lang et al., *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. **C75**.5 (2015) 204, arXiv: 1412.3989 [hep-ph].

[30] F. Aaron et al., *Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA*, JHEP **1001** (2010) 109, arXiv: 0911.0884 [hep-ex].

[31] H. Abramowicz et al. (**H1** and **ZEUS** Collaborations), *Combination of Measurements of Inclusive Deep Inelastic $e^{\pm}p$ Scattering Cross Sections and QCD Analysis of HERA Data* (2015), arXiv: 1506.06042 [hep-ex].

[32] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **1504** (2015) 040, arXiv: 1410.8849 [hep-ph].

[33] W. J. Waalewijn, *Calculating the Charge of a Jet*, Phys. Rev. **D86** (2012) 094030, arXiv: 1209.3019 [hep-ph].

[34] D. Krohn et al., *Jet Charge at the LHC*, Phys. Rev. Lett. **110**.21 (2013) 212001, arXiv: 1209.2421 [hep-ph].

[35] C. W. Bauer, S. Fleming, and M. E. Luke, *Summing Sudakov logarithms in B —> X(s gamma) in effective field theory*, Phys. Rev. **D63** (2000) 014006, arXiv: hep-ph/0005275 [hep-ph].

[36] C. W. Bauer et al., *An Effective field theory for collinear and soft gluons: Heavy to light decays*, Phys. Rev. **D63** (2001) 114020, arXiv: hep-ph/0011336 [hep-ph].

[37] C. W. Bauer and I. W. Stewart, *Invariant operators in collinear effective theory*, Phys. Lett. **B516** (2001) 134–142, arXiv: hep-ph/0107001 [hep-ph].

[38] C. W. Bauer, D. Pirjol, and I. W. Stewart, *Soft collinear factorization in effective field theory*, Phys. Rev. **D65** (2002) 054022, arXiv: hep-ph/0109045 [hep-ph].

[39] **ATLAS** Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[40] **ATLAS** Collaboration, *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. **C72** (2012) 1849, arXiv: 1110.1530 [hep-ex].

[41] **ATLAS** Collaboration, *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s}$ = 7 TeV*, ATLAS-CONF-2010-038 (2010), URL: https://cds.cern.ch/record/1277678.

[42] **ATLAS** Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment*, ATLAS-CONF-2010-069 (2010), URL: https://cds.cern.ch/record/1281344.

[43] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, arXiv: 0710.3820 [hep-ph].

[44] **ATLAS** Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003 (2012), URL: https://cds.cern.ch/record/1474107.

[45] J. M. Campbell, J. Huston, and W. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, Rept. Prog. Phys. **70** (2007) 89, arXiv: hep-ph/0611148 [hep-ph].

[46] A. Sherstnev and R. Thorne, *Parton Distributions for LO Generators*, Eur. Phys. J. **C55** (2008) 553–575, arXiv: 0711.2473 [hep-ph].

[47] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv: 1007.2241 [hep-ph].

[48] J. Gao et al., *The CT10 NNLO Global Analysis of QCD* (2013), arXiv: 1302.6246 [hep-ph].

[49] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **0411** (2004) 040, arXiv: hep-ph/0409146 [hep-ph].

[50] S. Frixione, P. Nason, and C. Oleari,
Matching NLO QCD computations with Parton Shower simulations: the POWHEG method,
JHEP **0711** (2007) 070, arXiv: 0709.2092 [hep-ph].

[51] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **1006** (2010) 043, arXiv: 1002.2581 [hep-ph].

[52] S. Frixione, P. Nason, and G. Ridolfi,
A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction,
JHEP **0709** (2007) 126, arXiv: 0707.3088 [hep-ph].

[53] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008) 639–707,
arXiv: 0803.0883 [hep-ph].

[54] K. Arnold et al., *Herwig++ 2.6 Release Note* (2012), arXiv: 1205.4902 [hep-ph].

[55] S. Gieseke, C. Rohr, and A. Siodmok, *Colour reconnections in Herwig++*,
Eur. Phys. J. **C72** (2012) 2225, arXiv: 1206.0041 [hep-ph].

[56] J. Pumplin et al.,
New generation of parton distributions with uncertainties from global QCD analysis,
JHEP **0207** (2002) 012, arXiv: hep-ph/0201195 [hep-ph].

[57] T. Sjostrand and P. Z. Skands,
Transverse-momentum-ordered showers and interleaved multiple interactions,
Eur. Phys. J. **C39** (2005) 129–154, arXiv: hep-ph/0408302 [hep-ph].

[58] S. Gieseke, P. Stephens, and B. Webber, *New formalism for QCD parton showers*,
JHEP **0312** (2003) 045, arXiv: hep-ph/0310083 [hep-ph].

[59] T. Sjostrand and P. Z. Skands, *Multiple interactions and the structure of beam remnants*,
JHEP **0403** (2004) 053, arXiv: hep-ph/0402078 [hep-ph].

[60] M. Bahr, S. Gieseke, and M. H. Seymour,
Simulation of multiple partonic interactions in Herwig++, JHEP **0807** (2008) 076,
arXiv: 0803.3633 [hep-ph].

[61] B. Andersson et al., *Parton Fragmentation and String Dynamics*, Phys. Rep. **97** (1983) 31.

[62] B. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*,
Nucl. Phys. **B238** (1984) 492.

[63] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*,
Phys. Rev. **D78** (2008) 013004, arXiv: 0802.0007 [hep-ph].

[64] R. D. Ball et al.,
A first unbiased global NLO determination of parton distributions and their uncertainties,
Nucl. Phys. **B838** (2010) 136–206, arXiv: 1002.4407 [hep-ph].

[65] **ATLAS** Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823–874,
arXiv: 1005.4568 [physics.ins-det].

[66] S. Agostinelli et al. (**GEANT4** Collaboration), *GEANT4: A Simulation toolkit*,
Nucl. Instrum. Meth. **A506** (2003) 250–303.

[67] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*,
JHEP **0804** (2008) 063, arXiv: 0802.1189 [hep-ph].

[68] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, Eur. Phys. J. **C72** (2012) 1896, arXiv: 1111.6097 [hep-ph].

[69] W Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002 (2008), URL: https://cds.cern.ch/record/1099735.

[70] C. Cojocaru et al., *Hadronic calibration of the ATLAS liquid argon end-cap calorimeter in the pseudorapidity region* $1.6 < |\eta| < 1.8$ *in beam tests*, Nucl. Instrum. Meth. A **531** (2004) 481, arXiv: 0407009.

[71] **ATLAS** Collaboration, *Local hadronic calibration*, ATL-LARG-PUB-2009-001 (2009), URL: https://cds.cern.ch/record/1112035.

[72] **ATLAS** Collaboration, *Pile-up subtraction and suppression for jets in ATLAS*, ATLAS-CONF-2013-083 (2013), URL: https://cds.cern.ch/record/1570994.

[73] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*, JHEP **0804** (2008) 005.

[74] J. C. Collins, D. E. Soper, and G. F. Sterman, *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1988) 1–91, arXiv: hep-ph/0409313 [hep-ph].

[75] A. Banfi, G. P. Salam, and G. Zanderighi, *Infrared safe definition of jet flavor*, Eur. Phys. J. **C47** (2006) 113–124, arXiv: hep-ph/0601139 [hep-ph].

[76] G. D'Agostini, *A Multidimensional unfolding method based on Bayes' theorem*, Nucl. Instrum. Meth. **A362** (1995) 487–498.

[77] T. Adye, *Unfolding algorithms and tests using RooUnfold* (2011), arXiv: 1105.1160 [physics.data-an].

[78] B. Malaescu, *An Iterative, dynamically stabilized method of data unfolding* (2009), arXiv: 0907.3791 [physics.data-an].

[79] **ATLAS** Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, Eur. Phys. J. **C75**.1 (2015) 17, arXiv: 1406.0076 [hep-ex].

[80] **ATLAS** Collaboration, *Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC*, Eur. Phys. J. **C73**.3 (2013) 2305, arXiv: 1203.1302 [hep-ex].

[81] **ATLAS** Collaboration, *Jet energy resolution in proton-proton collisions at* $\sqrt{s} = 7$ *TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. **C73**.3 (2013) 2306, arXiv: 1210.6210 [hep-ex].

[82] **ATLAS** Collaboration, *A study of the material in the ATLAS inner detector using secondary hadronic interactions*, JINST **7** (2012) P01013, arXiv: 1110.6191 [hep-ex].

[83] **ATLAS** Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, Eur. Phys. J. **C74**.11 (2014) 3130, arXiv: 1407.3935 [hep-ex].

[84] **ATLAS** Collaboration, *Search for anomalous production of prompt same-sign lepton pairs and pair-produced doubly charged Higgs bosons with* $\sqrt{s} = 8$ *TeV pp collisions using the ATLAS detector*, JHEP **1503** (2015) 041, arXiv: 1412.0237 [hep-ex].

[85] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. **D82** (2010) 074018, arXiv: 1005.3457 [hep-ph].

[86] J. C. Collins and D. E. Soper, *Parton Distribution and Decay Functions*, Nucl. Phys. **B194** (1982) 445.

[87] J. C. Collins and D. E. Soper, *Back-To-Back Jets in QCD*, Nucl. Phys. **B193** (1981) 381.