CMS Physics Analysis Summary

Contact: cms-pag-conveners-smp@cern.ch 2016/06/01

Measurement of jet charge observables in dijet events at $\sqrt{s} = 8$ TeV

The CMS Collaboration

Abstract

Jet charge is an estimator for the electric charge of a quark, antiquark or gluon initiating a jet. It is based on the momentum-weighted sum of the measured electric charges of the jet constituents. A measurement of three different charge observables of the leading jet is performed in dijet events. The analysis is carried out with a data sample of 19.7 fb^{-1} collected in proton-proton collisions at $\sqrt{s} = 8$ TeV. The results are presented differentially in transverse momentum of the leading jet and compared to predictions from QCD-inspired leading order event generators. This is the first CMS measurement of jet charge observables unfolded for detector effects.

1 Introduction

Jets can be initiated by high momentum colored partons, by top quarks, W bosons, Z bosons and Higgs bosons decaying to quarks, or by potential new physics resonances decaying to partons. At leading order in QCD we can distinguish the flavor of a parton that initiated a jet most of the times and refer to such jets as quark jets (q-jet), anti-quark jets ($\overline{\mathrm{q}}$-jet) or gluon-jets (g-jet). To distinguish different signal scenarios from background processes or to characterize a new physics resonance, it is crucial to identify the object initiating a jet by means of the properties of the reconstructed particles inside the jet. In particular, the charge quantum number of the original parton from which a jet is initiated can be estimated from a momentum-weighted sum of the charges of the particles inside the jet.

The idea to estimate the charge of a parton from a jet-based observable has a long history. Jet charge was suggested for the first time by Field and Feynman [1]. It was first measured in deep inelastic scattering at Fermilab [2, 3], CERN [4–7] and Cornell [8] in order to understand the correlation between quark and hadron models. Among its various applications there were the W boson charge discrimination in ALEPH [9], as well as the determination of the charge of the top quark at the Tevatron [10, 11] and the LHC [12].

Recent theory calculations [13, 14] motivate a more detailed validation of jet charge and promote its use in new applications. Studies of the performance and the discrimination power of jet charge as well as comparisons of dijet, W boson + jets and $\mathrm{t\bar{t}}$ data with simulated proton-proton (pp) collisions have been recently reported by CMS [15] and ATLAS [16]. A measurement of the mean value of the jet charge in bins of the jet transverse momentum (p_{T}) has been recently published by ATLAS [17].

In this paper, the first measurement of the entire jet charge distribution of the jet leading in p_{T} in a dijet sample is provided. The measurement is performed in different jet p_{T} ranges and corrected for detector effects. It is carried out for different definitions of jet charge in order to gain a better understanding of the underlying physics model that may be used further to constrain the predictions of Monte Carlo (MC) event generators.

2 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the superconducting solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. A preshower detector consisting of two planes of silicon sensors interleaved with lead is located in front of the ECAL at $1.653 < |\eta| < 2.6$. An iron and quartz-fiber Cerenkov hadron calorimeter (HF) covers pseudo-rapidities $3.0 < |\eta| < 5.0$. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid.

Charged particle trajectories are measured by the silicon tracker within the pseudorapidity range $|\eta| < 2.5$, where the pseudorapidity η is defined as η = - ln[tan(θ/2)], θ is the polar angle with respect to the z axis. It consists of 1440 silicon pixel and 15148 silicon strip detector modules and is located in the 3.8 T field of the superconducting magnetic field. For nonisolated particles of $1 < p_{\mathrm{T}} < 10$ GeV and $|\eta| < 1.4$, the track resolutions are typically 1.5% in p_{T} and 25–90 (45–150) μm in the transverse (longitudinal) impact parameter [18]. The calorimeters provide a coverage in pseudo-rapidity up to $|\eta| < 3.0$. In the region $|\eta| < 1.74$, the HCAL cells have widths of 0.087 in pseudorapidity and 0.087 in azimuth (ϕ). In the η-ϕ plane, and for

$|\eta| < 1.48$, the HCAL cells map on to 5×5 ECAL crystals arrays to form calorimeter towers projecting radially outwards from close to the nominal interaction point. At larger values of $|\eta|$, the size of the towers increases and the matching ECAL arrays contain fewer crystals.

The first level (L1) of the CMS trigger system, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select the most interesting events in a fixed time interval of less than 4 μs. The high-level trigger (HLT) processor farm further decreases the event rate from around 100 kHz to less than 1 kHz, before data storage. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [19].

3 Data and MC samples

The data sample for this analysis was recorded by the CMS detector in 2012 at the CERN LHC at a center of mass energy of 8 TeV. It corresponds to an integrated luminosity of 19.7 fb^{-1}. This analysis uses events collected with a high-level trigger, requiring a single jet with $p_T > 320$ GeV. This trigger identifies with 99% efficiency events with at least one jet reconstructed offline with $p_T > 400$ GeV.

Two MC event generators are used in this analysis, PYTHIA 6.4.26 [20] and HERWIG++ 2.5.0 [21]. They are based on the leading order (LO) QCD matrix element combined with parton showers (PS). The PS, used to simulate higher order processes, follows an ordering principle motivated by QCD. Successive radiation of gluons from a high energetic parton is ordered by a specific variable, e.g., p_T or angle of radiation. The two generators differ in the choice of ordering technique, as well as in the treatment of beam remnants, multiple interactions, and the hadronization model. PYTHIA 6.4.26 uses a p_T-ordered PS model. It provides a good description of parton emission when the emitted partons are close. The Z2* tune is used for the underlying event (UE). Apart from the parton distribution function (PDF) dependent energy extrapolation parameter it resembles the Z2 tune [22]. Partons are hadronized using the Lund string model [23, 24]. The HERWIG++ 2.5.0 program with EE3C tune is based on a parton shower model using the coherent branching algorithm with angular ordering of the showers [25]. The partons are then hadronized using a cluster model [26] and the multiple parton interaction (MPI) is simulated using the eikonal multiple partonic scattering model [25]. The generated events are passed through the CMS detector simulation based on GEANT4 [27].

4 Event Reconstruction and Event Selection

Jets are reconstructed from particle flow (PF) objects [28] using the anti-k_T clustering algorithm [29, 30] with a distance parameter R = 0.5. The particle flow algorithm identifies each particle with an optimized combination of information from all sub-detectors. In order to reduce the contamination from additional pp interactions (pileup), charged particles from non-primary vertices are removed before clustering. Jets are clustered from the particle flow objects and the total momenta of the jets are calculated by four-momentum summation. Because of the non-uniform and non-linear response of the CMS calorimeter, the reconstructed jets require additional energy corrections. The jet energy corrections are based on high p_T jet events generated with PYTHIA 6 [20]. Residual corrections derived from in situ measurements with dijet, γ + jet and Z boson + jet events [31] are applied.

Events are selected by requiring at least two jets passing the following selection criteria: the jets with leading and sub-leading p_T should lie within $|\eta| < 1.5$ and should have $p_T > 400$

GeV and $p_{\mathrm{T}} > 100$ GeV, respectively. Events with spurious jets from noise and non-collision backgrounds are rejected by applying a set of jet identification criteria to the jet properties [32]. Additional selection requirements are also applied to reduce beam background and noise. At least one good reconstructed primary vertex is required with the number of degrees of freedom > 4 and with $|z| < 24$ cm. The missing transverse momentum (E_T^{miss}) over the transverse momentum scalar sum ($\sum \mathrm{E}_T$) is required to pass $\mathrm{E}_T^{miss}/\sum \mathrm{E}_T < 0.3$.

The agreement between the data and the MC simulations based on PYTHIA 6 and HERWIG, is verified at reconstructed level looking at the kinematic properties of the leading jets: jet p_{T}, η, ϕ, dijet invariant mass, as well as the jet shape properties such as track multiplicity and jet charge. In each case a reasonable agreement is found. Figure 1 provides a comparison of PYTHIA 6 with the data as function of the p_{T} of the leading jet. For each event the flavor of the leading jet is identified in PYTHIA 6 using a geometrical matching procedure in angular space between the jet and a hard parton at matrix element level. The distance between the generator level hard partons and the reconstructed level jet in the η-ϕ plane is defined as $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$. The parton before showering and radiation with smallest ΔR to the jet, passing the matching criterion ($\Delta R < \Delta R_{max}$ where $\Delta R_{max} = 0.3$), is chosen as the parton initiating the jet. Jets that cannot be matched to any generator level hard parton with $\Delta R < \Delta R_{max}$ are categorized as unmatched. The matching efficiency is found to be more than 96% throughout the jet p_{T} range studied. The "others" category represents those jets which are initiated by parton flavors including the anti-up quark ($\bar{\mathrm{u}}$), the anti-down quark ($\bar{\mathrm{d}}$), the charm, strange and bottom quarks (c, $\bar{\mathrm{c}}$, s, $\bar{\mathrm{s}}$, b, $\bar{\mathrm{b}}$) and the unmatched jets.

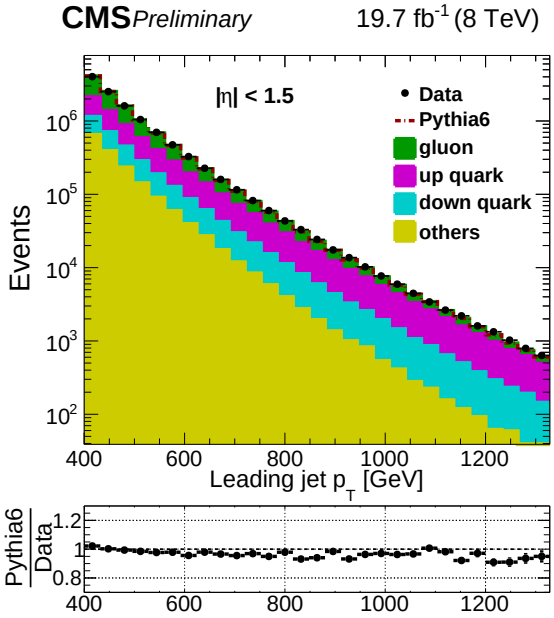

Figure 1: Leading jet p_{T} distribution in data compared to PYTHIA 6 simulation. The fraction of events where the leading jet is matched to a generator quark or gluon are indicated as filled histograms. The "others" category represents those jets which are initiated by parton flavors including the anti-up quark ($\bar{\mathrm{u}}$), the anti-down quark ($\bar{\mathrm{d}}$), the charm, strange and bottom quarks (c, $\bar{\mathrm{c}}$, s, $\bar{\mathrm{s}}$, b, $\bar{\mathrm{b}}$) and the unmatched jets.

5 Jet charge observables

Jet charge is the p_{T} weighted sum over the electric charge of the particles in a jet. Following Refs. [13, 14], we define jet charge as

$$Q^{\kappa} = \frac{1}{(p_{\mathrm{T}})^{\kappa}} \sum_{i} Q_i (p_{\mathrm{T}}^{i})^{\kappa} \tag{1}$$

where the sum is over all colour-neutral particles i in the jet with $p_{\mathrm{T}} > 1$ GeV, Q_i is the integer charge of the particle, p_{T}^{i} is the magnitude of the particle transverse momentum with respect to the beam axis, and κ is a free parameter.

Values of κ between 0.2 and 1 have been used in earlier experimental studies [3, 33]. Here three values of κ are taken: 0.3, 0.6 and 1.0, in coherence with a related study carried out by the ATLAS experiment [16]. In addition to the above given definition of jet charge, we study two other jet charge variables defined as the following,

$$Q_L^{\kappa} = \sum_{i} Q_i \left(p_{\parallel}^{i}\right)^{\kappa} / \sum_{i} \left(p_{\parallel}^{i}\right)^{\kappa} \tag{2}$$

$$Q_T^{\kappa} = \sum_{i} Q_i \left(p_{\perp}^{i}\right)^{\kappa} / \sum_{i} \left(p_{\perp}^{i}\right)^{\kappa} \tag{3}$$

where the sum over i refers to all the particles (charged and neutral) in the jet with $p_{\mathrm{T}} > 1$ GeV. The notations $p_{\parallel}^{i} = \vec{p}^{\,i} \cdot \frac{\vec{p}_{jet}}{\|\vec{p}_{jet}\|}$ and $p_{\perp}^{i} = \vec{p}^{\,i} \times \frac{\vec{p}_{jet}}{\|\vec{p}_{jet}\|}$ refer to the momentum component of the constituent i parallel and transverse to the jet axis, respectively. A cut of 1 GeV is applied on the p_{T} of the particle candidates in order to minimize the residual dependence on the number of pileup interactions in the same and different bunch crossing. It is found that jet charge variables have a slight dependence on the number of primary vertices in each event with this specific particle p_{T} requirement, such that it becomes negligible as compared to the other sources of experimental uncertainties. Compared to Q^{κ}, the definition Q_L^{κ} is more directly related with the fragmentation function $F(z)dz$ of a quark or a gluon that gives the probability to find a particle with the momentum fraction $z = p_{\parallel}^{i}/|p_{jet}|$ within dz in a q-jet or a g-jet [1]. We study all three variables Q^{κ} (default definition), Q_L^{κ} (longitudinal definition) and Q_T^{κ} (transverse definitions) to disentangle effects parametrized by the fragmentation function from other effects.

At the generator level the jet charge observables are computed in a similar way as summarized above. The generator level jet charge is computed from the generator level stable particles with $p_{\mathrm{T}} > 1$ GeV considering also the charged decay products from K_S^0 and Λ particles.

The jet charge variable can be used to distinguish a jet initiated by a gluon or a quark. It can also differentiate the jets coming from quarks with different electric charges. In Fig. 2 the jet charge distribution of the leading jet, initiated by either an up quark (u) or a down quark (d) or a gluon in PYTHIA 6 is compared to the data. The splitting with respect to the various flavors is carried out in PYTHIA 6 only. From this figure onwards Q_1^{κ}, $Q_{L,1}^{\kappa}$ and , $Q_{T,1}^{\kappa}$ indicate the jet charge values of the leading jet in the event. The jet charge distribution initiated by quarks with positive electric charge is peaked towards the positive side as opposed to the one initiated by negatively charged quarks. Good agreement between the data and PYTHIA 6, HERWIG++ predictions is observed.

As it is shown in Fig. 1, the jet flavor composition of the selected dijet sample depends on the leading jet p_{T}. Gluon jets are dominating the lower part of the spectrum, while up quarks are becoming progressively more important above 500 GeV. As a consequence, the average jet

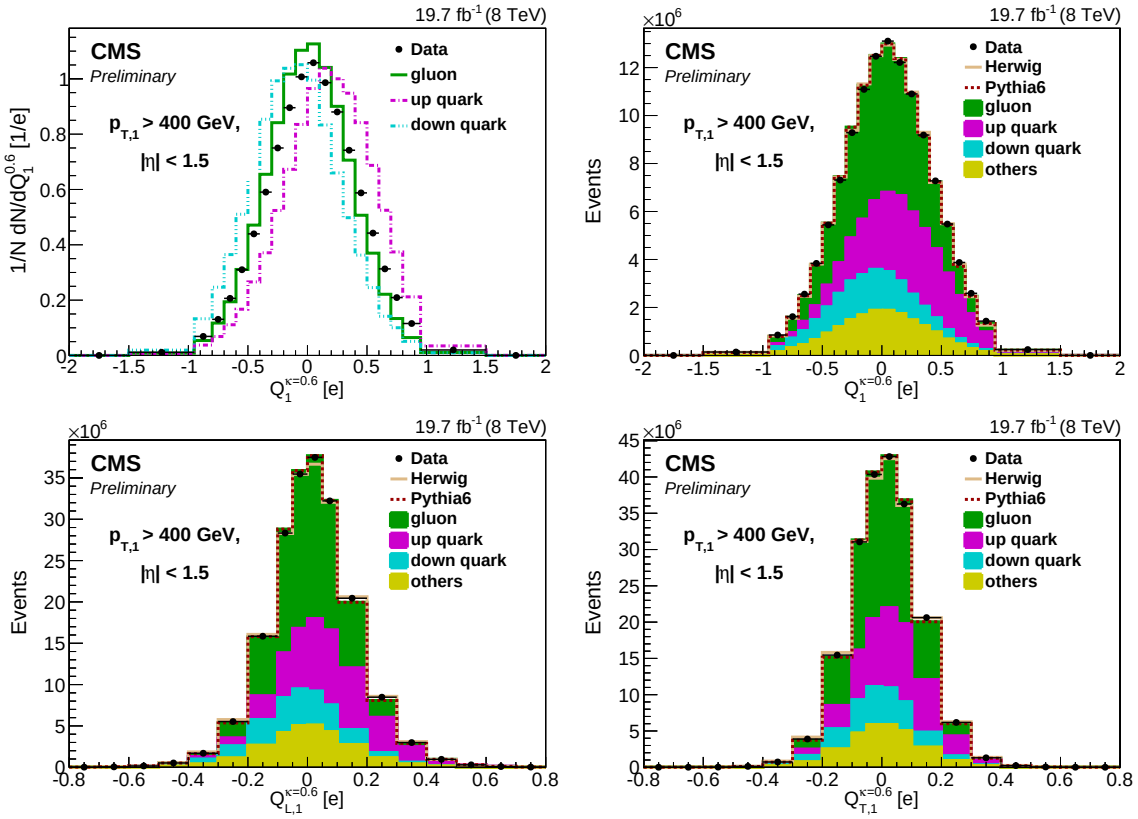

Figure 2: Distribution of the jet charge (top left and top right) Q^κ, (bottom left) Q_L^κ and (bottom right) Q_T^κ in data compared to MC simulations. The top left plot shows for the u, d and g distributions in PYTHIA 6 in comparison with data where each distribution is normalized to its total number of events. The top right, bottom left and bottom right plots compare the sum of the contributions in PYTHIA 6 and HERWIG++ to data where the flavor breakdown is carried out in PYTHIA 6.

charge increases as a function of the leading jet p_T, as it can be observed in Fig. 3. PYTHIA 6 and HERWIG++ reproduce this trend well. It is therefore interesting to divide the dijet sample according to different ranges of leading jet p_T, and measure the jet charge distribution separately in each subsample. In this way the sensitivity of the different jet charge definitions to different mixtures of parton flavors, and the quality of description by various generators, can be studied.

6 Unfolding of detector effects

In order to compare with other measurements or theoretical predictions, the measured jet charge distribution is unfolded from the detector-level to the final-state particle level. The difference between the variables obtained from the MC simulation at the detector level are not identical with those constructed using the generator level information, coming from a given theoretical input, because of the effects from the detector resolution, acceptance and bin migration. In Fig. 4, we can see that the difference between the jet charge distributions at the generator level and the reconstructed level in PYTHIA 6 increases with lowering the κ value, since the jet charge definition for low values of κ gives more weight to low p_T particles which have a track reconstruction efficiency of about 90%.

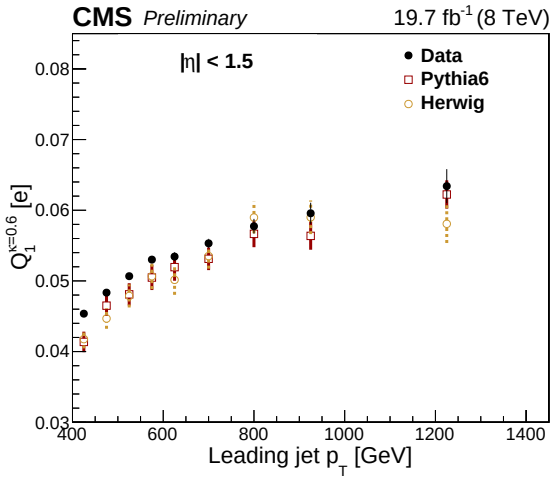

Figure 3: Variation of the average leading jet charge in PYTHIA 6, HERWIG++ and data as a function of leading jet p_{T}. The error bars for the simulation indicate the uncertainty due to the MC simulation statistics.

The unfolding is performed using particles with $p_{\mathrm{T}} > 1$ GeV. The unfolding method that is used in this analysis is the iterative Bayesian method [34]. The unfolding approach utilizes a response matrix that maps the true distribution onto the measured one. The response matrix is taken from the PYTHIA 6 simulation, and it is used to unfold the data distributions. The Bayesian algorithm follows an iterative response matrix inversion approach. The regularization in this approach is achieved by stopping the iteration just before the wild fluctuation of the inverse [34]. Another frequently used regularized unfolding algorithm, known as the Singular Value Decomposition method (SVD) [35] is also studied to cross check the results as an alternative unfolding procedure, utilizing the regularization concept. These two unfolding approaches show a good agreement (disagreement level $\approx 1\%$). For both the Bayesian and SVD approaches their implementations in the RooUnfold software package [36] are used.

7 Systematic Uncertainties

The experimental uncertainties that affect the measured result are summarized in the following section.

The uncertainties on the jet energy scale and the jet energy resolution are estimated after considering the corresponding effects in computing the jet charge definitions and then propagating the changes through the complete analysis.

The uncertainty on the jet energy scale has been estimated to be 1-2.5% [31] depending on p_{T} and η. In order to map this uncertainty on the jet charge variable, the two jets (leading and subleading jets) in the selected events are systematically shifted by their respective uncertainty and the new set of values for the jet charge variables is calculated and compared. Effects due to the scale uncertainty on jet charge variables are found to be small (less than 1%).

The jet energy resolution has been measured by comparing the asymmetry between the momenta of the two jets in the dijet events [37]. The jet energy resolution in simulation is smeared to match the measured resolutions and varied within its uncertainty. The effect on the jet charge is found to be small (less than 1%).

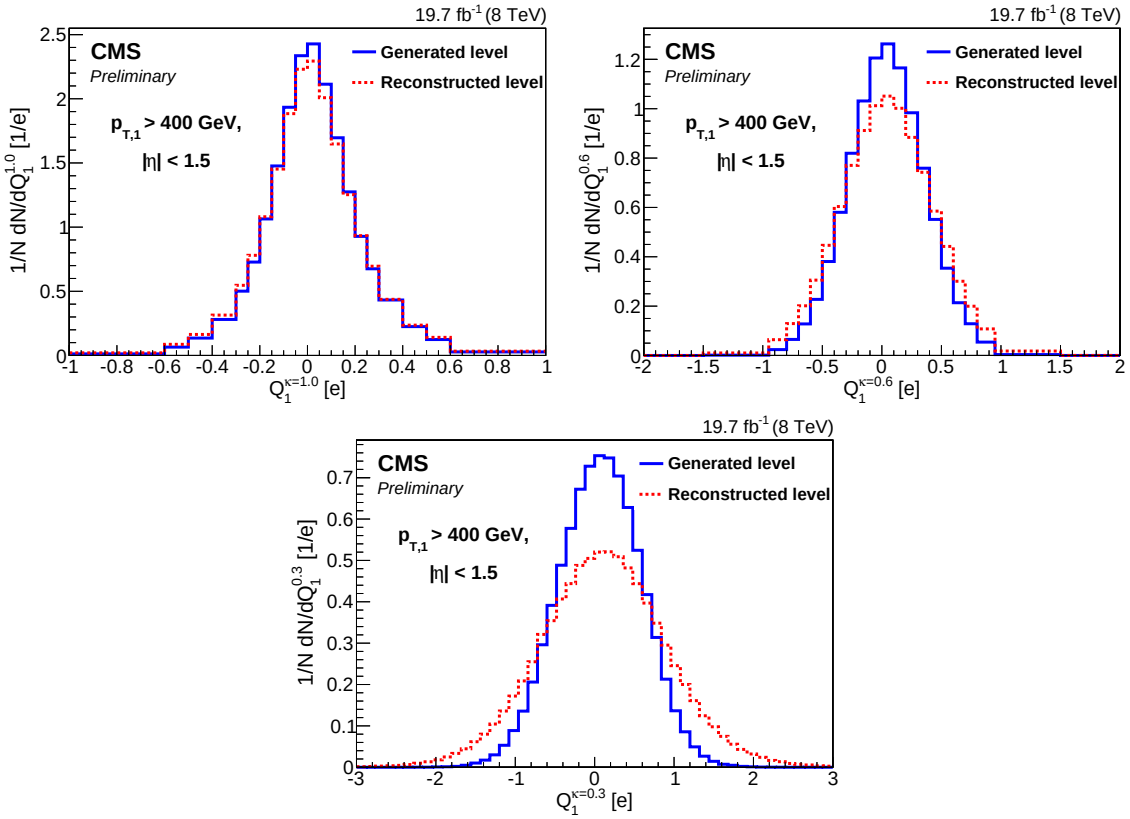

Figure 4: Distributions of the jet charge of the leading jet at reconstructed level and generated level in PYTHIA 6 with $\kappa = 1.0$ (top left), 0.6 (top right), 0.3 (bottom) respectively.

The jet charge computation is done based on the reconstructed charged hadrons built from the reconstructed tracks and calorimeter energy by the particle flow algorithm. For each track considered, the corresponding reconstruction efficiency varies with track p_T and η. The variation of the track reconstruction efficiency was estimated in Ref. [18] and is used as the weight factors to the particle flow objects. For each track the corresponding track reconstruction efficiency value as a function of η and p_T is estimated from a QCD multijet simulated sample. The resultant efficiency is varied within one standard deviation of its original value and the jet charge variable is computed for each variation of the track weight factors accordingly.

The track p_T resolution is the dominant experimental source of systematic uncertainty as compared to the above studied ones. The track p_T resolution value changes with track η. As an example, the relative p_T resolution changes for a 1 GeV track from 0.011 to 0.015 as $|\eta|$ changes from 0.5 to 1.0 [18]. For each track the corresponding p_T resolution value as a function of η and p_T is estimated from a QCD multijet simulated sample. The resultant resolution is varied within one standard deviation of its original value and the jet charge is computed for each variation of the track p_T smearing accordingly. It is confirmed that the jet energy scale and jet energy resolution have negligible correlation with track resolution and track reconstruction.

The unfolding procedure is carried out based on a response matrix derived with the PYTHIA 6 event generator. In order to study the systematic effect due to the choice of a particular generator to derive this matrix, the response matrix is made also from HERWIG++ and using both of these response matrices data are unfolded accordingly. The corresponding difference is taken as the uncertainty due to the response matrix modeling.

Another systematic effect that is taken into account for the unfolding procedure is the (MC simulation) statistical uncertainty on the matrix elements in the response matrix. They are propagated using the RooUnfold software package.

Table 1 summarizes the size of the various systematic effects. The average over all the bins of the jet charge distribution ratio of the upward and downward variations, weighted by the inverse squared statistical uncertainty is given to summarize the size of each effect. Among all the systematic effects, the most dominating ones are the uncertainties associated to track p_{T} resolution effect and the effect due to the response matrix modeling. The remaining systematic uncertainties result in small effects (less than a percent) including the jet energy scale and jet energy resolution effects. It can also be concluded that the jet charge computations with all three κ values show comparable systematic uncertainties.

Table 1: Various systematic effects and their corresponding error weighted mean of the fractional deviation in percent.

Syst. effect in percent (%)	$\kappa = 1.0$			$\kappa = 0.6$			$\kappa = 0.3$		
	Q^{κ}	Q_L^{κ}	Q_T^{κ}	Q^{κ}	Q_L^{κ}	Q_T^{κ}	Q^{κ}	Q_L^{κ}	Q_T^{κ}
Jet energy scale	0.7	<0.1	<0.1	0.4	<0.1	<0.1	0.3	<0.1	<0.1
Jet energy resolution	0.1	<0.1	<0.1	0.1	<0.1	<0.1	<0.1	<0.1	<0.1
Track reconstruction	0.4	0.4	0.5	0.5	0.4	0.5	0.5	0.4	0.4
Track p_{T} resolution	1.4	1.0	0.8	1.0	0.6	0.7	1.5	0.4	0.4
Response matrix modeling	1.6	1.6	1.8	1.0	0.8	1.3	1.5	1.3	1.3
Response matrix statistics	0.9	0.9	0.6	0.6	0.6	0.5	0.6	0.5	0.4

8 Results

In Figs. 5, 6, 7, 8 the distributions of the unfolded data are presented and compared to the generator level PYTHIA 6 and HERWIG++ model predictions with two different LO PDF sets, CTEQ6L1 [38] and NNPDF3.0 [39] used for PYTHIA 6. Each plot shows the unfolded data and the generator level predictions, as well as their ratio and a band representing the uncertainty, determined by adding in quadrature the statistical error on data and errors arising due to all systematic effects on data. In Figs. 5, 6, 7, 8 the normalization of each distribution is performed with the corresponding total number of events.

The effect of the parton shower and fragmentation model on the jet charge distribution can be seen by comparing the predictions from PYTHIA 6 and HERWIG++ simulation that make predictions based on different parton shower and fragmentation models.

The effect of the PDFs on the jet charge distribution can be seen, by comparing PYTHIA 6 simulation generated with CTEQ6L1 and re-weighted with NNPDF3.0. When comparing predictions of the jet charge distributions in data, both PDF as well as parton shower and fragmentation modeling have to be considered to explain discrepancies.

In Fig. 5, the dependence on the default and the longitudinal jet charge definitions in two different columns and on different κ values, in three different rows is demonstrated. The differences between the two generators are below 10% for the Q^{κ} and Q_L^{κ} definitions. Figure 6 represents the distributions of the transverse definition with all three different κ values. A large difference of up to 50% is observed for the Q_T^{κ}, showing different sensitivity of the variables to the fragmentation function. The impact of the PDFs on the jet charge distribution is found to be of the

order 1-5%.

Thus the jet charge modeling in the dijet sample is not very sensitive to PDFs, on which the quark vs. gluon composition of the sample depends. It is more significantly dependent on the parton shower and fragmentation function.

In general the predictions from PYTHIA 6 and HERWIG++ generators show only mild discrepancies with the data distributions, however some systematic differences can be pointed out. Experimental uncertainties are generally larger for small values of κ and also for the Q_T^κ definition, due to the larger weight given to soft particles for these cases. For the Q^κ jet charge definition (left column of Fig. 5), PYTHIA 6 and HERWIG++ for all values of κ shows similar level of agreement. For the Q_L^κ jet charge definition (right column of Fig. 5), PYTHIA 6 and HERWIG++ have a similar level of agreement with data for all values of κ. For the Q_T^κ jet charge definition (Fig. 6), both two generators diverge significantly for most of the range. The two generators are systematically different for the three jet charge definitions and thus it can be concluded that this measurement can constrain jet charge modeling in generator predictions. It should also be noted that a smaller fraction of the differences between the data and the simulation may arise due to the PDF modeling, while a larger fraction of the differences may arise from hadronization and parton shower models.

In Fig. 7, the dependence on the jet p_{T} in three different rows for the default and longitudinal jet charge definitions is demonstrated. For the transverse definition the jet p_{T} dependence is shown in Fig. 8. In the jet p_{T} range considered, the gluon fraction is expected to decrease with p_{T} from about 35% (left) to 15% (right). Generally for all jet charge definitions, the level of disagreement between the two generator predictions decreases as a function of jet p_{T}. This suggest that the description of g-jets differs more between PYTHIA 6 and HERWIG++ than the description of q-jets. The level of agreement between simulation and data stays similar as a function of jet p_{T}, while the PYTHIA 6 and HERWIG++ predictions approach each other at high p_{T}.

9 Summary

In this paper we presented the first measurement of the jet charge distribution with the CMS experiment. Data distributions from dijet events unfolded for detector effects are provided for different definitions of jet charge and in different jet p_{T} bins.

Three different definitions of jet charge provide different sensitivity to the fragmentation model. Three different κ parameter choices provide different sensitivity to the softer and harder particles in the jet.

We provided measurements in different leading jet p_{T} bins with sensitivity to the different composition of the quark and gluon jets in the dijet sample. In general the predictions from PYTHIA 6 and HERWIG++ generators show only mild discrepancies with the data distributions. However, the predictions from PYTHIA 6 and HERWIG++ generators are systematically different and could be constrained by this measurement.

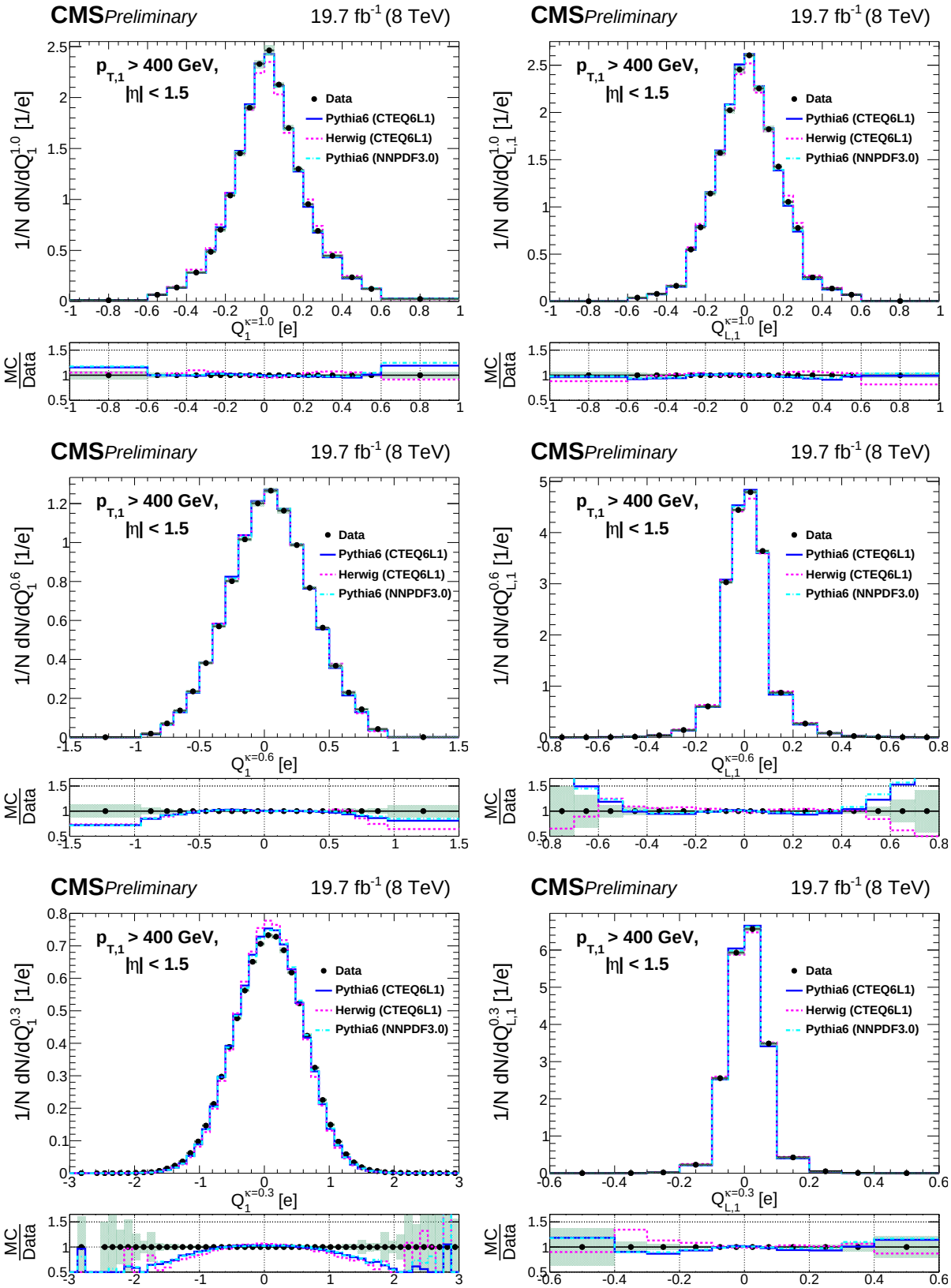

Figure 5: Comparison of unfolded leading jet charge distributions with PYTHIA 6, HERWIG++ generators. The left column shows the distributions for the default jet charge definition (Q^{κ}) with all three different κ values, while the right column shows for the longitudinal jet charge definition (Q_L^{κ}) with all three different values of κ. Shaded uncertainty bands include both statistical and systematic effects, added in quadrature.

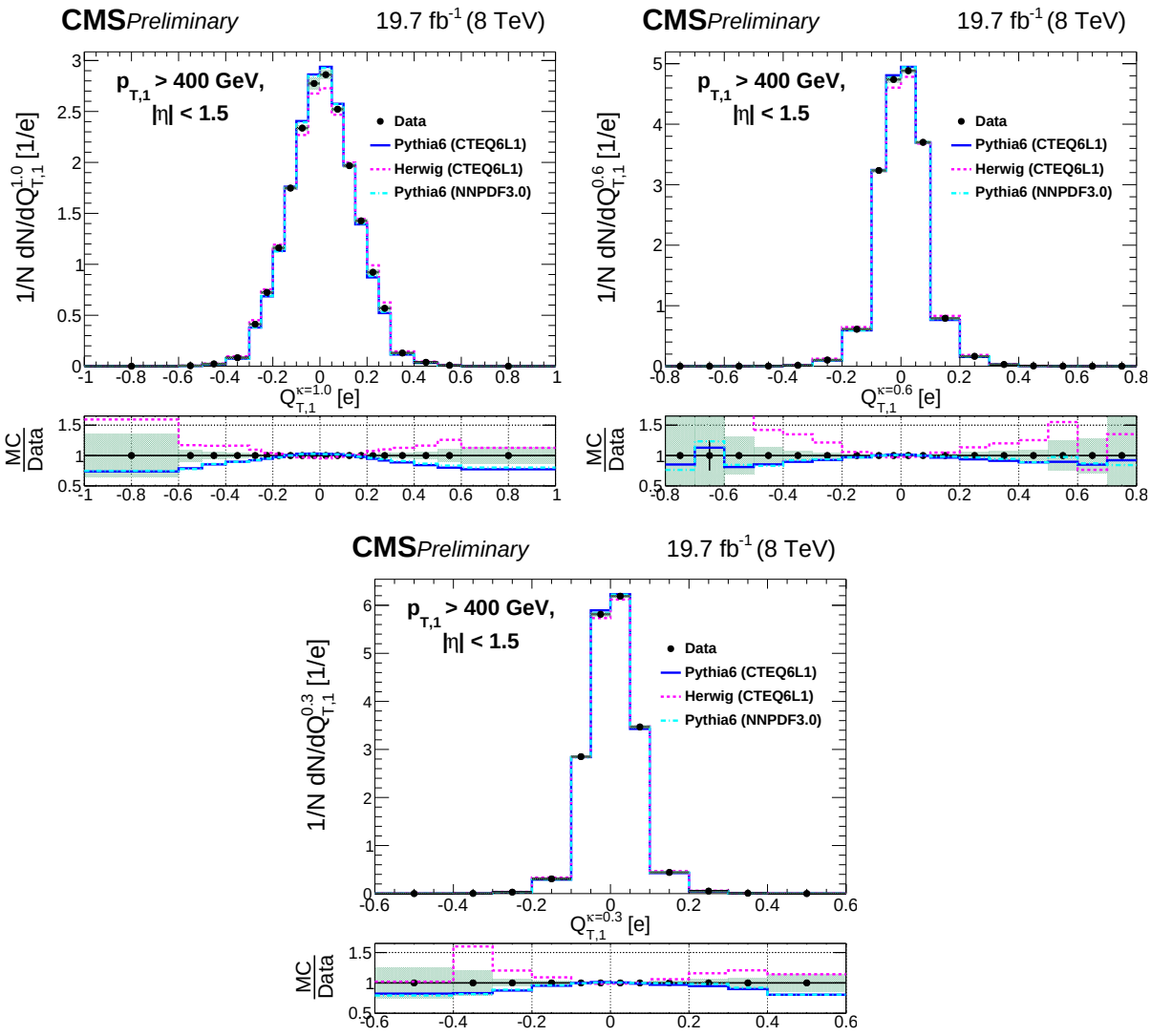

Figure 6: Comparison of unfolded leading jet charge distributions with PYTHIA 6, HERWIG++ generators for transverse jet charge definition (Q_T^{κ}) with all different κ values. Shaded uncertainty bands include both statistical and systematic effects, added in quadrature.

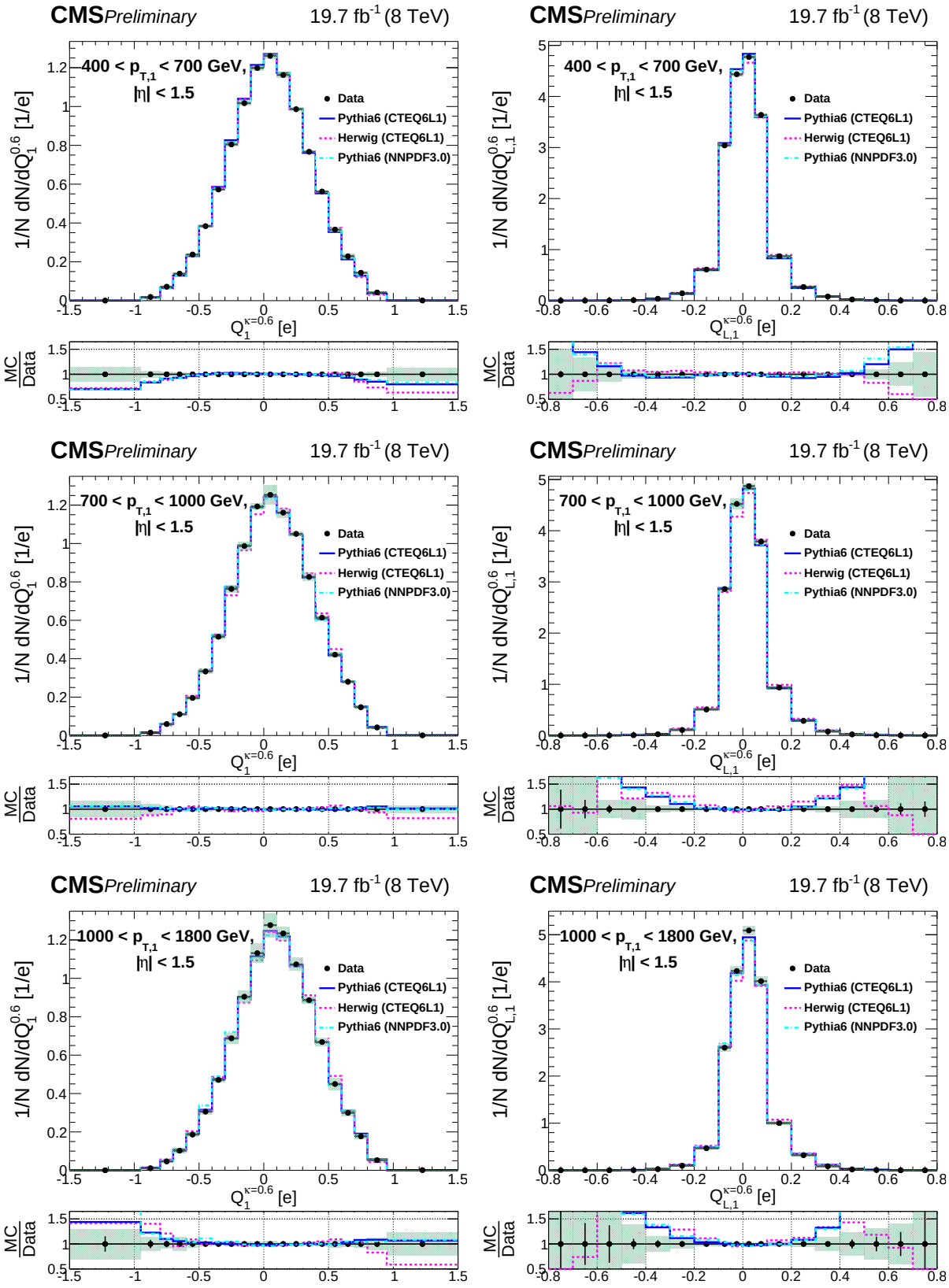

Figure 7: Comparison of unfolded leading jet charge distributions with PYTHIA 6, HERWIG++ generators in bins of leading jet p_{T}. Left column shows the jet p_{T} dependence for the default jet charge definition (Q^{κ}) with $\kappa = 0.6$. Right column shows the jet p_{T} dependence for the longitudinal jet charge definition (Q_L^{κ}) with $\kappa = 0.6$. Shaded uncertainty bands include both statistical and systematic effects, added in quadrature.

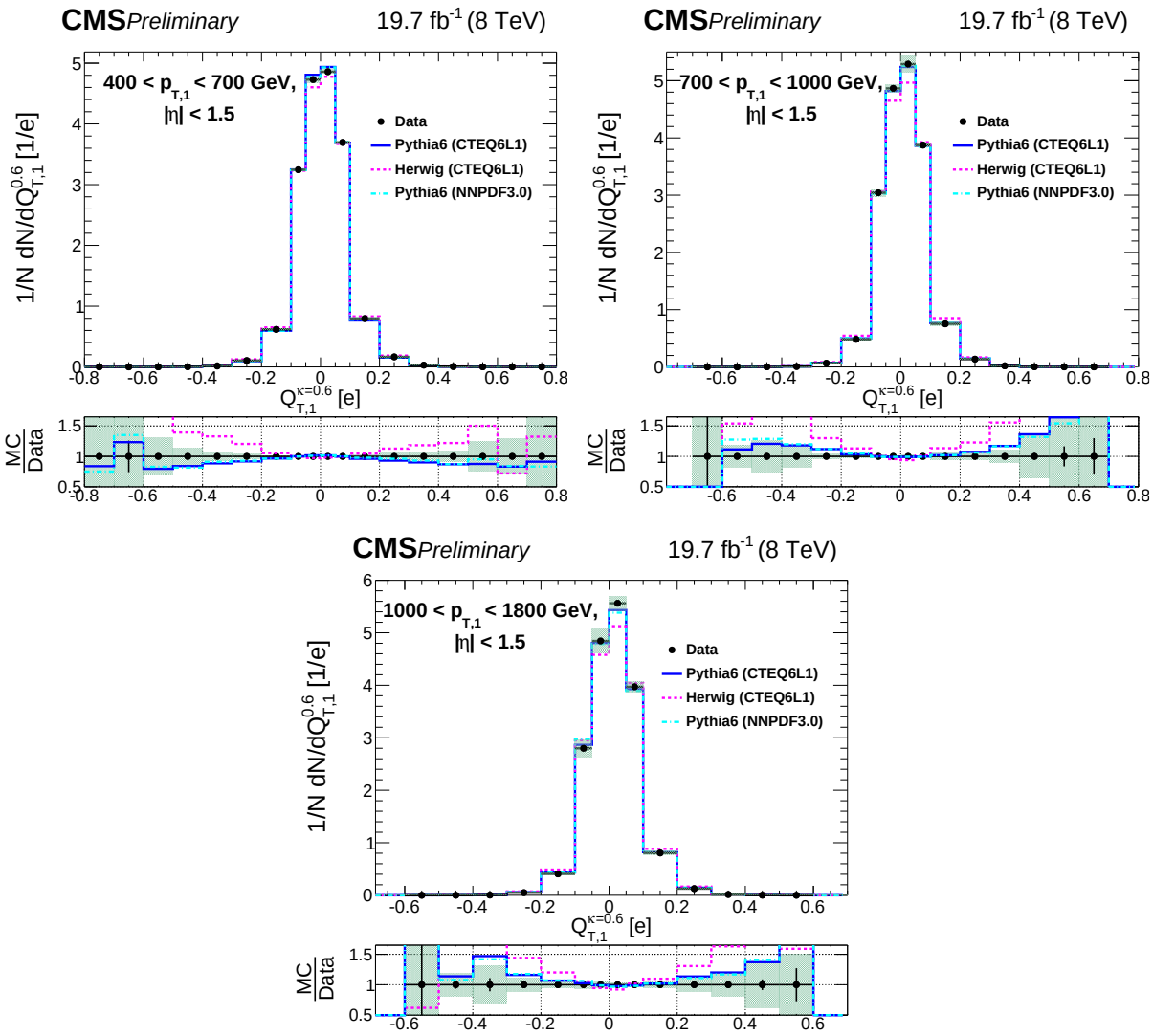

Figure 8: Comparison of unfolded leading jet charge distributions with PYTHIA 6, HERWIG++ generators in bins of leading jet p_{T} for the transverse jet charge definition (Q_T^{κ}) with $\kappa = 0.6$. Shaded uncertainty bands include both statistical and systematic effects, added in quadrature.

References

[1] R. D. Field and R. P. Feynman, "A Parametrization of the Properties of Quark Jets", *Nucl. Phys. B* **136** (1978) 1, doi:10.1016/0550-3213(78)90015-9.

[2] Fermilab-Serpukhov-Moscow-Michigan Collaboration, "Net Charge in Deep Inelastic anti-neutrino - Nucleon Scattering", *Phys. Lett. B* **91** (1980) 311–313, doi:10.1016/0370-2693(80)90456-6.

[3] J. P. Berge et al., "Quark Jets from anti-neutrino Interactions. 1. Net Charge and Factorization in the Quark Jets", *Nucl. Phys. B* **184** (1981) 13–30, doi:10.1016/0550-3213(81)90207-8.

[4] Aachen-Bonn-CERN-Munich-Oxford Collaboration, "Multiplicity Distributions in Neutrino - Hydrogen Interactions", *Nucl. Phys. B* **181** (1981) 385–402, doi:10.1016/0550-3213(81)90532-0.

[5] Aachen-Bonn-CERN-Munich-Oxford Collaboration, "Charge Properties of the Hadronic System in Neutrino p and Anti-neutrino p Interactions", *Phys. Lett. B* **112** (1982) 88, doi:10.1016/0370-2693(82)90912-1. [,307(1982)].

[6] European Muon Collaboration, "Quark Charge Retention in Final State Hadrons From Deep Inelastic Muon Scattering", *Phys. Lett. B* **144** (1984) 302–308, doi:10.1016/0370-2693(84)91825-2.

[7] Amsterdam-Bologna-Padua-Pisa-Saclay-Turin Collaboration, "Charged Hadron Multiplicities in High-energy Anti-muon Neutrino n and Anti-muon Neutrino p Interactions", *Z. Phys. C* **11** (1982) 283, doi:10.1007/BF01578279. [Erratum: Z. Phys.C14,281(1982)].

[8] R. Erickson et al., "Charge Retention in Deep Inelastic Electroproduction", *Phys. Rev. Lett.* **42** (1979) 822, doi:10.1103/PhysRevLett.42.822. [Erratum: Phys. Rev. Lett.42,1246(1979)].

[9] ALEPH Collaboration, "Measurement of triple gauge boson couplings at 172-GeV", *Phys. Lett. B* **422** (1998) 369–383, doi:10.1016/S0370-2693(98)00061-6.

[10] D0 Collaboration, "Experimental discrimination between charge 2e/3 top quark and charge 4e/3 exotic quark production scenarios", *Phys. Rev. Lett.* **98** (2007) 041801, doi:10.1103/PhysRevLett.98.041801, arXiv:hep-ex/0608044.

[11] CDF Collaboration, "The CDF Measurement of the Top Quark Charge using the Top Decay Products in Lepton+Jet channel", CDF note 10460, 2011.

[12] ATLAS Collaboration, "Measurement of the top quark charge in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *JHEP* **11** (2013) 031, doi:10.1007/JHEP11(2013)031, arXiv:1307.4568.

[13] W. J. Waalewijn, "Calculating the Charge of a Jet", *Phys. Rev. D* **86** (2012) 094030, doi:10.1103/PhysRevD.86.094030, arXiv:1209.3019.

[14] D. Krohn, M. D. Schwartz, T. Lin, and W. J. Waalewijn, "Jet Charge at the LHC", *Phys. Rev. Lett.* **110** (2013), no. 21, 212001, doi:10.1103/PhysRevLett.110.212001, arXiv:1209.2421.

[15] CMS Collaboration, "Identification techniques for highly boosted W bosons that decay into hadrons", *JHEP* **12** (2014) 017, doi:10.1007/JHEP12(2014)017, arXiv:1410.4227.

[16] ATLAS collaboration, "Jet Charge Studies with the ATLAS Detector Using $\sqrt{s} = 8$ TeV Proton-Proton Collision Data", Technical Report ATLAS-CONF-2013-086, 2013.

[17] ATLAS Collaboration, "Measurement of jet charge in dijet events from $\sqrt{s}$=8TeV pp collisions with the ATLAS detector", *Phys. Rev. D* **93** (2016), no. 5, 052003, doi:10.1103/PhysRevD.93.052003, arXiv:1509.05190.

[18] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014), no. 10, P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[19] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[20] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[21] M. Bahr et al., "Herwig++ Physics and Manual", *Eur. Phys. J. C* **58** (2008) 639–707, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.

[22] R. Field, "Min-Bias and the Underlying Event at the LHC", in *Proceedings, 31st International Conference on Physics in collisions (PIC 2011)*. 2012. arXiv:1202.0901.

[23] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, "Parton Fragmentation and String Dynamics", *Phys. Rept.* **97** (1983) 31–145, doi:10.1016/0370-1573(83)90080-7.

[24] T. Sjostrand, "The Merging of Jets", *Phys. Lett. B* **142** (1984) 420–424, doi:10.1016/0370-2693(84)91354-6.

[25] S. Gieseke, P. Stephens, and B. Webber, "New formalism for QCD parton showers", *JHEP* **12** (2003) 045, doi:10.1088/1126-6708/2003/12/045, arXiv:hep-ph/0310083.

[26] B. R. Webber, "A QCD Model for Jet Fragmentation Including Soft Gluon Interference", *Nucl. Phys. B* **238** (1984) 492–528, doi:10.1016/0550-3213(84)90333-X.

[27] GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.

[28] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", CMS Physics Analysis Summary CMS-PAS-PFT-10-001, 2010.

[29] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[30] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[31] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[32] CMS Collaboration, "Jet Performance in pp Collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-003, 2010.

[33] ALEPH Collaboration, "Measurement of charge asymmetry in hadronic Z decays", *Phys. Lett. B* **259** (1991) 377–388, doi:10.1016/0370-2693(91)90844-G.

[34] G. D'Agostini, "A Multidimensional unfolding method based on Bayes' theorem", *Nucl. Instrum. Meth. A* **362** (1995) 487–498, doi:10.1016/0168-9002(95)00274-X.

[35] A. Hocker and V. Kartvelishvili, "SVD approach to data unfolding", *Nucl. Instrum. Meth. A* **372** (1996) 469–481, doi:10.1016/0168-9002(95)01478-0, arXiv:hep-ph/9509307.

[36] T. Adye, "Unfolding algorithms and tests using RooUnfold", in *Proceedings of the PHYSTAT 2011 Workshop, CERN, Geneva, Switzerland, January 2011, CERN-2011-006*, pp. 313–318. 2011. arXiv:1105.1160.

[37] CMS Collaboration, "Jet Energy Resolution in CMS at sqrt(s)=7 TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-014, 2011.

[38] J. Pumplin et al., "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP* **07** (2002) 012, doi:10.1088/1126-6708/2002/07/012, arXiv:hep-ph/0201195.

[39] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.