Available on the CERN CDS information server **CMS PAS QCD-10-013**

CMS Physics Analysis Summary

Contact: cms-pag-conveners-qcd@cern.ch 2015/10/13

Hadronic Event Shapes in pp Collisions at 7 TeV

The CMS Collaboration

Abstract

Hadronic event shapes are studied with a data sample of 7 TeV proton-proton collisions collected with the CMS detector at the Large Hadron Collider (LHC). The size of the sample corresponds to an integrated luminosity of $78\,\mathrm{nb}^{-1}$. Jets constructed with different techniques are used as input for calculating event-shape variables, which probe the structure of the hadronic final state. Normalized event-shape distributions are shown to be robust against various sources of systematic uncertainty. It is demonstrated that these early measurements of event-shape variables are sensitive to differences in the modeling of QCD multi-jet production.

1 Introduction

Event shape variables provide geometrical information about the energy flow in hadronic events. Suitably defined event shapes can be described by the theory of Quantum Chromodynamics (QCD) [1–3]. They were among the first observables proposed to test QCD and have subsequently been closely connected with progress in the theory. At e^+e^- and ep colliders, event shapes played a crucial role in the extraction of the strong coupling. They have been essential in tuning the parton shower and non-perturbative components of Monte Carlo event generators and have provided a laboratory for developing and testing analytical insight into the hadronisation process. More recently, a large set of event-shape variables suitable for pp colliders has been proposed in [4]. An important aspect of these variables is their normalization to the total transverse momentum or energy in the event. It is thereby anticipated that energy scale uncertainties should cancel to a large extent. Event shapes have thus been proposed as a valuable tool for early measurements of the properties of QCD events at the LHC and the tuning of Monte Carlo models.

In the following, we present a first study of two event shape variables with the initial data sample of 7 TeV proton-proton collisions collected with the CMS detector at the LHC, corresponding to an integrated luminosity of $78\,\mathrm{nb}^{-1}$. The four-momenta of jets in the hadronic final state is used as input to the event-shape calculation. Jet finding is performed using the anti-k_T clustering algorithm [5], with four different sets of reconstructed objects as input for comparison.

2 Definition of the Hadronic Event-Shape Variables

This study focuses on two event shape variables: the *central transverse thrust* $T_{\perp,\mathcal{C}}$ and the *central thrust minor* $T_{m,\mathcal{C}}$. The term *central* ($\mathcal{C}$) indicates that the input to the calculation of these quantities are jets in the central region of the detector ($|\eta| < 1.3$, see Section 4).

The central transverse thrust is defined as [4]:

$$T_{\perp,\mathcal{C}} \equiv \max_{\vec{n}_T} \frac{\sum_{i\in\mathcal{C}} |\vec{p}_{\perp,i} \cdot \vec{n}_T|}{\sum_{i\in\mathcal{C}} p_{\perp,i}}. \tag{1}$$

where $p_{\perp,i}$ are the transverse momenta of all selected jets with respect to the beam axis. The transverse axis, for which the maximum is obtained, is the thrust axis $\vec{n}_{\mathrm{T},\mathcal{C}}$. The variable that is typically used in perturbative calculations is $\tau_{\perp,\mathcal{C}} \equiv 1 - T_{\perp,\mathcal{C}}$, referred to as central transverse thrust in the following.

The central thrust minor is a measure of the momentum out of the plane defined by $\vec{n}_{\mathrm{T},\mathcal{C}}$ and the beam axis. It is defined as:

$$T_{m,\mathcal{C}} \equiv \frac{\sum_{i\in\mathcal{C}} |\vec{p}_{\perp,i} \times \vec{n}_{\mathrm{T},\mathcal{C}}|}{\sum_{i\in\mathcal{C}} p_{\perp,i}}. \tag{2}$$

The sensitivity of these event shapes to the experimental absolute energy scale is reduced, because of cancellations in the ratio. It is therefore expected that these event shapes will be robust against uncertainties in the detector performance, as will be shown below.

3 Jet Reconstruction

Calorimetric energy deposits are collected in the lead-tungstate crystal electromagnetic calorimeter (ECAL) and the brass-scintillator hadronic calorimeter (HCAL). These calorimeters

provide a uniform and hermetic coverage over a large range of pseudorapidity ($|\eta| \leq 3$, where $\eta = -\ln\tan(\theta/2)$ and θ is the polar angle relative to the beam axis). The calorimeter cells are grouped in projective towers of granularity $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ (0.175×0.175) at central (forward) rapidity regions. Towers with $|\eta| < 1.3$ contain only cells from the barrel calorimeters. CMS uses a right-handed coordinate system, with the origin located at the nominal collision point, the x-axis pointing towards the center of the LHC, the y-axis pointing up (perpendicular to the LHC plane), and the z-axis along the (anticlockwise) beam direction.

Theoretical calculations typically define event shapes in terms of the particles in the event. Identification of individual particles, however, is experimentally challenging; instead, reconstructed jets (clustered energy deposits or tracks) are used to calculate the event shapes. Four different types of reconstructed objects are clustered into jets, as described in the following. In all four cases, the clustering of the objects is done with the anti-k_T algorithm [5], a sequential recombination algorithm, with jet size parameter $R = 0.5$.

Calorimeter jets [6, 7] are reconstructed only from energy deposits in the electromagnetic and hadronic calorimeter cells, combined into calorimeter towers as inputs. Jet energy corrections are applied to account for the non-linear and non-uniform response of the CMS calorimeters.

In the Jet-Plus-Tracks jets (*JPT jets*) algorithm [8], jets reconstructed from calorimetric energy deposits are corrected using the momentum of associated charged particles measured in the tracker. The jet energy and position resolutions are significantly improved with respect to that of calorimeter jets, and only small residual corrections are applied.

Track jets are based on reconstructed tracks only [9] and are thus independent of calorimeter jets. An energy correction is applied to account for neutral particles in jets.

PF jets are constructed using Particle Flow techniques [10], which aim to reconstruct each individual particle in the event by combining the information from all CMS sub-detector systems. PF particles are reconstructed as a combination of charged tracks in the tracker and clusters in the electromagnetic and hadronic calorimeters, as well as signals in either of the two CMS pre-shower detectors and the muon system. The energy of each particle is calibrated according to its type. As a result of the PF reconstruction, the input to the jet clustering is almost fully calibrated and the resulting jets only require small energy correction.

4 Data Set and Event Selection

The data analyzed in this paper were collected in April to July 2010 at a center-of-mass energy of 7 TeV. Non-collision background is removed and quality cuts are applied to ensure the presence of a well reconstructed primary vertex [11]. The integrated luminosity of the selected sample is $78\,\mathrm{nb}^{-1}$.

CMS uses a two-tiered trigger system to select events online: a hardware Level-1 (L1) trigger and an online software High Level Trigger (HLT). We select two data samples using single jet triggers that require an L1 jet with transverse momentum $p_\mathrm{T} > 6$ GeV/c (resp. 20 GeV/c) and an HLT jet with $p_\mathrm{T} > 15$ GeV/c (resp. 30 GeV/c). The jet energies used in the trigger are uncorrected for the calorimeter energy response. The low-p_T trigger became prescaled as the instantaneous luminosity of the LHC increased. As a result, the two selected data sample only share 20% of the events. The trigger efficiency, measured from a sample acquired with triggers with lower threshold, is greater than 99% for corrected jet p_T above 50 GeV/c (resp. 80 GeV/c), for all types of jet reconstruction.

Monte Carlo simulation samples are produced with four event generators for comparison with data:

- PYTHIA 6.4.22 [12] is used with the multiple parton interaction tune D6T. In this version of PYTHIA, parton showers are ordered by virtuality. The underlying event modelling is based on the Multiple Parton Interaction model.
- PYTHIA 8.135 [13] is used with the default "Tune 1" for pp collisions. In this version of PYTHIA, parton showers are ordered by p_{T}. The underlying event model is based on the Multiple Parton Interaction model of PYTHIA 6, now interleaved with initial-state radiation.
- HERWIG++ 2.4.2 [14] is used with its default tune. The parton shower in HERWIG++ is based on the coherent branching algorithm. The underlying event is simulated using an eikonal multiple parton-parton scattering model.
- MadGraph 4.4.24 [15] is used in conjunction with PYTHIA 6, with tune D6T for PYTHIA. Matrix element (ME) level events including 2 up to 4 jets with parton p_{T} above 20 GeV/c are produced with MadGraph and subsequently passed to PYTHIA to generate a parton shower (PS). The so-called MLM matching procedure [16] is used to avoid double-counting between the ME and PS calculations. For the matching, the minimum jet k_T threshold is set to 30 GeV/c.
- ALPGEN 2.13 [17] is used in a similar way to MadGraph. ALPGEN samples are produced separately for each jet multiplicity, from 2 to 5 jets with parton p_{T} above 20 GeV/c, and are weighted according to their cross section. ME level events produced with ALPGEN are passed to PYTHIA, and the MLM matching procedure is used to avoid double-counting. For the matching, the lower jet p_{T} threshold is set to 20 GeV/c and ΔR is kept to its default value of 0.7.

Events are generated with a center-of-mass energy of 7 TeV and processed with a simulation of the CMS detector response based on GEANT 4 [18].

All jets used in the event-shape calculation and that rely on energy deposits in the calorimeter are required to pass additional quality cuts, in order to remove possible instrumental or non-collision background. Calorimeter and JPT jets are required to have a minimum of 1% of their total energy detected in the ECAL; the number of calorimeter cells containing 90% of the jet energy must be greater than one; the fraction of total jet energy appearing in any given HCAL 18-channel read-out unit must be smaller than 98% [7]. PF jets are required to contain at least two particles, to have a non-zero fraction of the HCAL energy carried by charged particles, not to have all of the ECAL energy carried only by neutral or only by charged particles, and not to have all of the HCAL energy carried by neutral particles.

The tracks used as input for the track jet reconstruction are selected with a minimal set of quality requirements, ensuring a low fake rate and at the same time sufficient efficiency to reconstruct track jets down to transverse momenta of a few GeV/c [9].

Only jets with $p_{\mathrm{T}} > 30$ GeV/c and pseudorapidity $|\eta| < 2.6$ are subsequently considered. If one of the two leading jets (jets with highest p_{T}) does not pass the quality criteria, the event is rejected. This requirement rejects 2.5% of the events for calorimeter and JPT jets, and less than 0.5% for PF jets.

After this initial jet selection, the two leading jets of the event are required to be in the central region of the detector ($|\eta| < 1.3$). To maximise the trigger efficiency and minimise trigger bias, the leading jet is required to have $p_{\mathrm{T}} > 60$ GeV/c for the low-p_{T} data sample and $p_{\mathrm{T}} >$

90 GeV/c for the high-p_{T} data sample. If any of the two conditions is not fulfilled, the event is rejected. This requirement will, in particular, reject events where one of the two leading jets is in the range $1.3 < |\eta| < 2.6$.

If the event is selected, all jets within $|\eta| < 1.3$ and with $p_{\mathrm{T}} > 30$ GeV/c are used in the event shape calculation. In the low-p_{T} data sample, this selection retains 23 000 events in total, of which 85% are events with two selected jets and 15% with three or more selected jets. In the high-p_{T} data sample, 22 000 events are retained in total, of which 74% are events with two selected jets.

5 Results

The normalized event shape distributions from data and simulations are shown in the top portion of Figs. 1 and 2, for the low-p_{T} and high-p_{T} data samples, respectively. The results are shown for calorimeter jets only because we find that the event shape results are similar for the different types of reconstructed objects, as will be discussed below. The error bars on the data points represent statistical uncertainties; the shaded (yellow) bands represent the quadratic sum of systematic and statistical uncertainties. Ratio plots between data and simulations are shown in the bottom portion of Figs. 1 and 2. In the ratio plots, the jet energy and position resolution uncertainties that affect the Monte Carlo predictions are added in quadrature to the total uncertainty of the data points.

In the following, we present studies of the event shape sensitivity and sources of systematic errors. The comparison between data and Monte Carlo models is then discussed.

5.1 Sensitivity Studies

We verify the robustness of the event shape measurements with respect to the different types of reconstructed objects. A comparison of the central transverse thrust and central thrust minor distributions for the four types of jet objects in the low-p_{T} sample is illustrated in Fig. 3. The error bars on the data points represent statistical uncertainties; the shaded (yellow) bands represent the quadratic sum of systematic and statistical uncertainties. The distributions are seen to be generally consistent with each other. Therefore our final results are based on the calorimeter jets only, which is a conservative choice with respect to the evaluation of systematic uncertainties (discussed below).

5.2 Systematic Uncertainties

Event shapes are energy-normalized distributions and are therefore robust against jet energy scale uncertainties, as already mentioned. While the event shape definitions are invariant under a shift in jet energy scale, the uncertainty on the energy scale does modify the number of jets passing the p_{T} threshold, thus affecting the event shapes. At present, this constitutes the dominant systematic error.

In order to estimate the resulting systematic uncertainty on the event shape distributions, we apply an absolute $\pm 10\%$ shift and an η-dependent $\pm 2\% \times |\eta|$ shift to all jets entering the calculation. This is a conservative sensitivity test, based on preliminary studies with the available CMS data [7]. The maximum bin-by-bin difference between the original distributions and the two shifted distributions is then assigned as a systematic error due to jet energy scale uncertainty.

Figure 4 illustrates the effect of the absolute 10% shift in jet energy scale on the central trans-

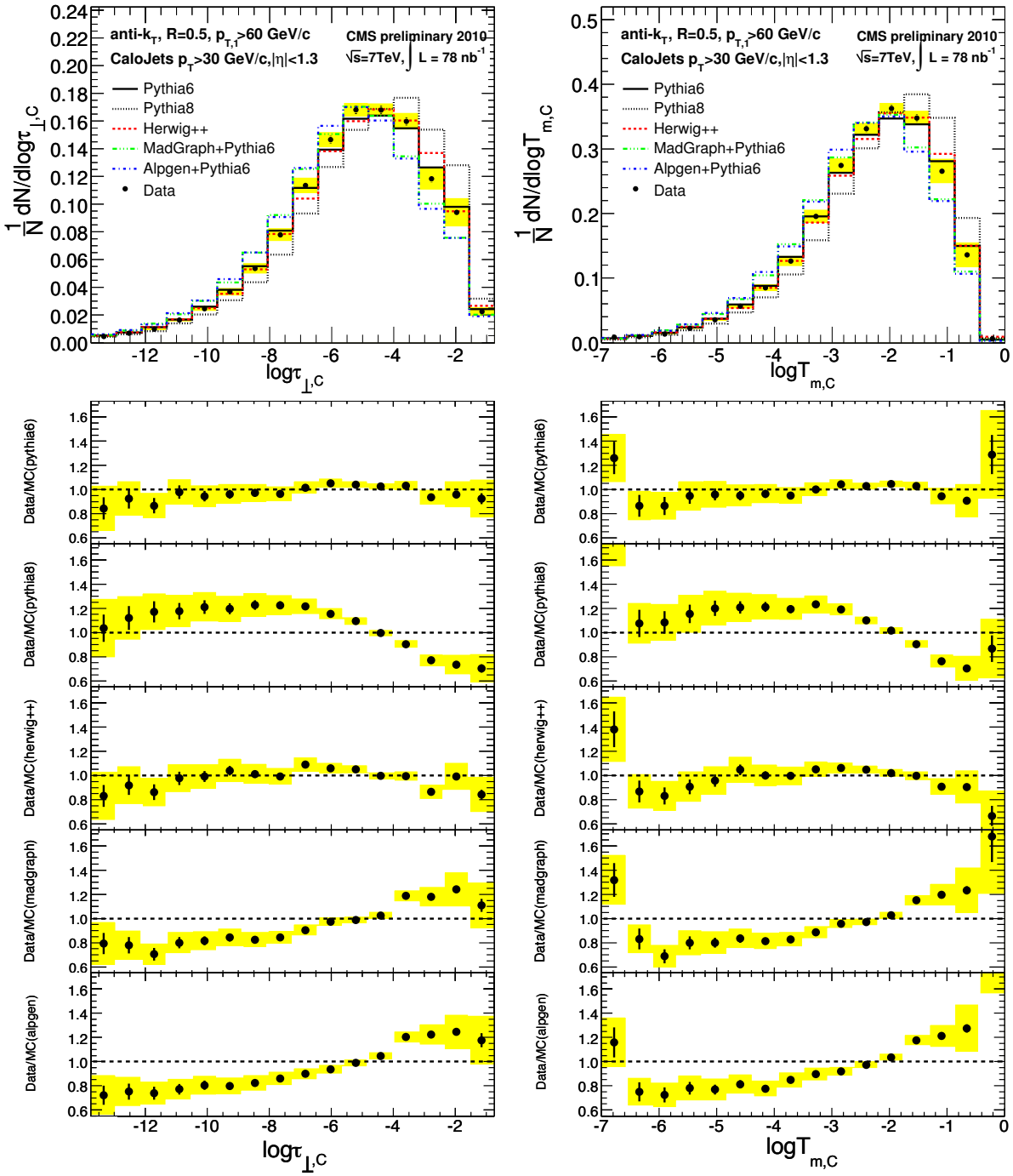

Figure 1: The central transverse thrust (left) and central thrust minor (right) distributions for calorimeter jets in events with leading jet $p_{\mathrm{T}} > 60$ GeV/c. The bars represent the statistical uncertainty on the data, and the yellow bands represent the sum of statistical and systematic errors. Bottom plots show the ratio between data and the different simulation samples, including the jet resolution uncertainty on Monte Carlo simulations.

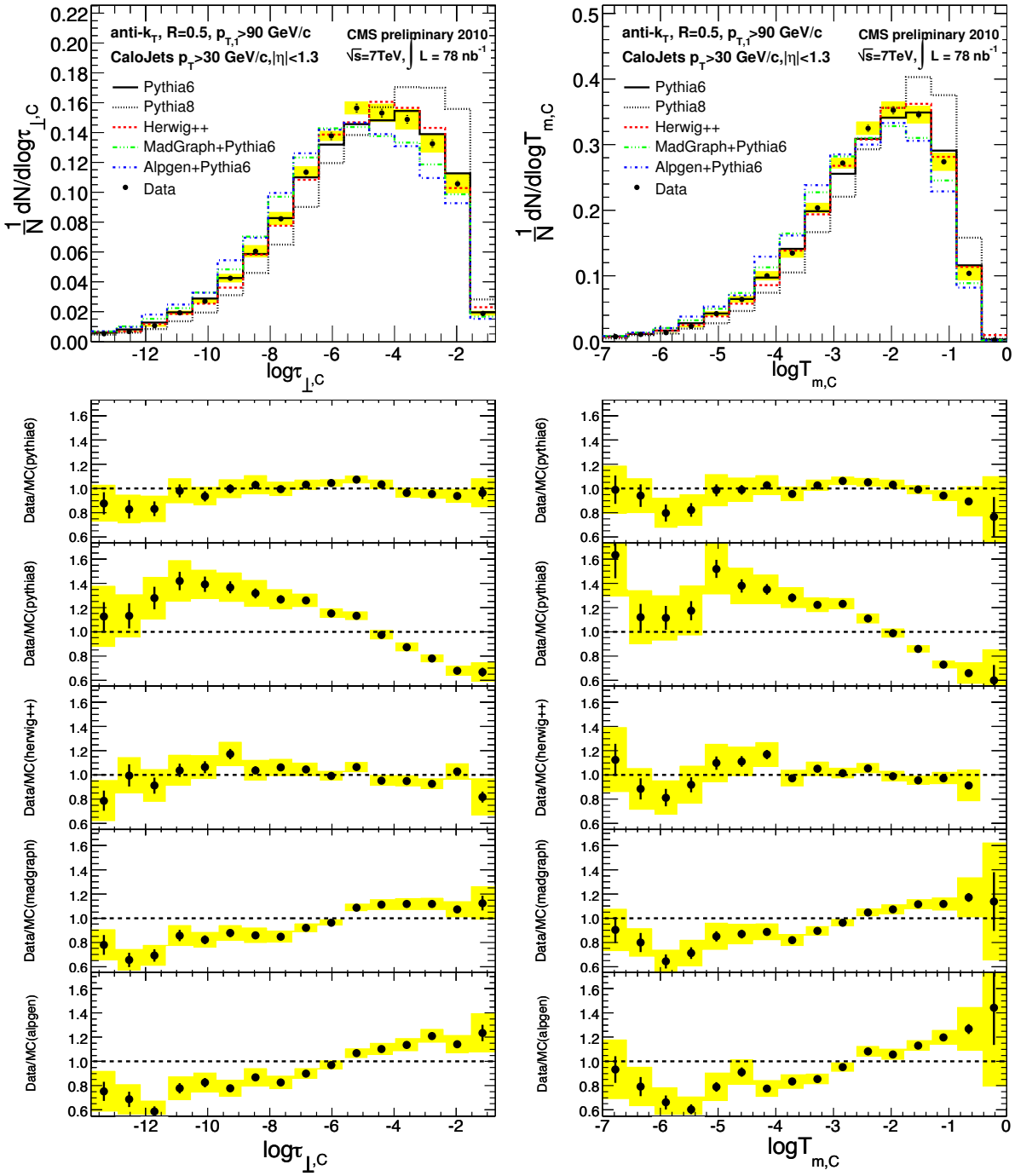

Figure 2: The central transverse thrust (left) and central thrust minor (right) distributions for calorimeter jets in events with leading jet $p_{\mathrm{T}} > 90$ GeV/c. The bars represent the statistical error on the data, and the yellow bands represent the sum of statistical and systematic errors. Bottom plots show the ratio between data and the different simulation samples, including the jet resolution uncertainty on Monte Carlo simulations.

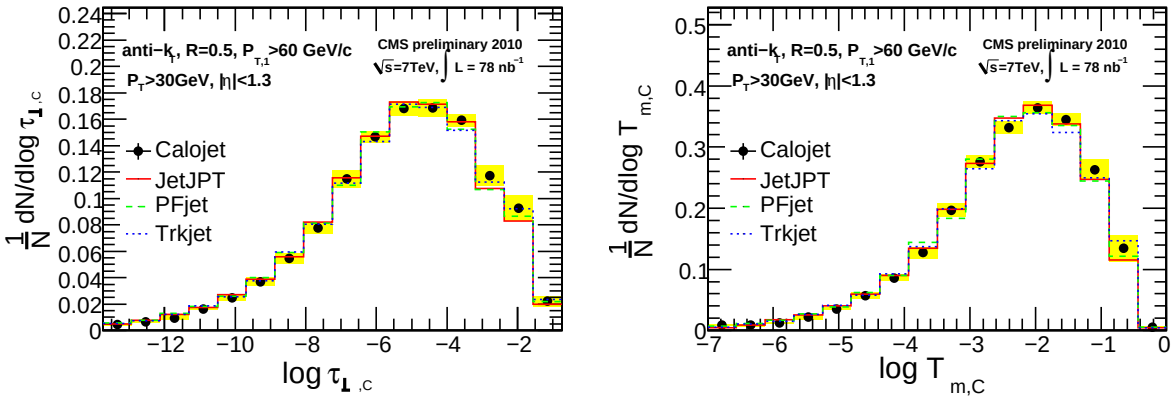

Figure 3: Comparison between calorimeter jets and the three other types of jet reconstruction for the central transverse thrust (left) and central thrust minor (right) distributions in data events with leading jet $p_{\mathrm{T}} > 60$ GeV/c. The bars represent the statistical uncertainty on the data, and the yellow bands represent the sum of statistical and systematic errors.

verse thrust and central thrust minor, in the low-p_{T} data sample. In the core of the distributions, the effect of the shift is less than 10%. The effect of the η-dependent shift is negligible with respect to the absolute shift. Fluctuations in the outermost bins are due to statistical fluctuations. These bins are also expected to be more sensitive to the shift in jet energy scale because they correspond to extreme topologies that are already significantly modified by a small shift, contrary to the topologies that enter the central bins.

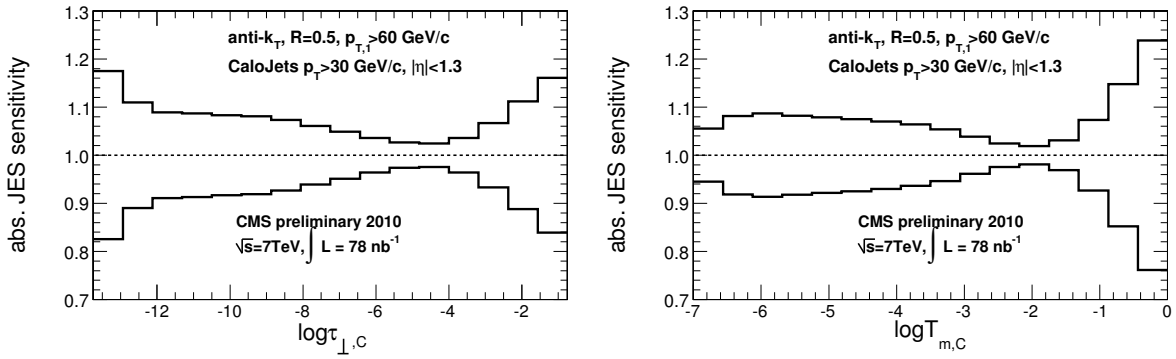

Figure 4: The effect of absolute jet energy scale uncertainty on the central transverse thrust (left) and central thrust minor (right) distributions for calorimeter jets in events with leading jet $p_{\mathrm{T}} > 60$ GeV/c

A systematic error is assigned to Monte Carlo distributions due to the imperfect modelling of the jet energy and position resolutions. In order to estimate this uncertainty, it was first verified that a Gaussian smearing of the PYTHIA 6 generator-level distribution of jet p_{T}, together with the smearing of the ϕ and η distributions results in event shape distributions that agree with the full detector simulation. The uncertainties related to jet energy and position resolution were then estimated in generator-level studies by varying the width of the respective Gaussian smearing by $\pm$10%. The bin-by-bin difference between the original and modified event shape distributions was assigned as a systematic error. This results in up to 5% systematic error on the central transverse thrust and central thrust minor distributions.

The energy scale uncertainty on data and detector resolution uncertainty on Monte Carlo sim-

ulations are added in quadrature in the ratio plots shown in Figs. 1 and 2.

5.3 Comparison with Models of QCD Multi-jet Production

The comparison of event shape distributions obtained in the data with the predictions of the PYTHIA 6, PYTHIA 8, HERWIG++, MadGraph and ALPGEN Monte Carlo generators, after full detector simulation, is shown in Figs. 1 and 2, for the low-p_{T} and high-p_{T} samples, respectively.

We observe that the ALPGEN, MadGraph and PYTHIA 8 curves deviate from the data points while the PYTHIA 6 and HERWIG++ predictions agree with the measurements within experimental uncertainties. ALPGEN and MadGraph overestimate the fraction of back-to-back di-jet events, while PYTHIA 8 underestimates it. All observations apply equally to the low-p_{T} and to the high-p_{T} data samples.

Further investigations indicate that the momentum of the second leading jet is harder in ALPGEN and MadGraph than in the data. This results in differences in the distribution of $\Delta\phi$ between the two leading jets, ALPGEN and MadGraph exhibiting a sharper peak at $\Delta\phi = \pi$. The three-jet to two-jet event ratio is also underestimated. These differences then lead to a shift of the event shape distributions towards lower values. This is especially true for two-jet events, as can be seen on Fig. 5, where we separately compared the event shape distributions in events with two selected jets, and in events with three or more selected jets, in the low-p_{T} sample. We conclude that the kinematic regime where the explicit matrix-element calculations of ALPGEN and MadGraph become superior to the parton shower treatment is still beyond reach. A tuning of the generator parameters might, however, alleviate the disagreement; it is currently under study.

In order to check if the PYTHIA 8 discrepancy was related to the shower ordering, we compared the PYTHIA 6 distributions with our default D6T tune, which uses virtuality-ordered showers, and that with the P0 tune [19], which uses p_{T}-ordered showers. The slight differences observed between the two tunes do not fully account for the differences observed in PYTHIA 8. Further tuning of the PYTHIA 8 generator is currently under study.

6 Conclusion

We presented a first study of two event-shapes variables, the central transverse thrust and the central thrust minor. We used a data sample of proton-proton collision data at 7 TeV centre-of-mass energy, accumulated by the CMS detector during the first months of 2010. The sample corresponds to an integrated luminosity of 78 nb^{-1}. Event shapes were shown to be robust against the choice of jet reconstruction object, as well as to uncertainties in the jet energy scale and experimental resolution. We compared the data with predictions from the PYTHIA 6, PYTHIA 8, HERWIG++, MadGraph and ALPGEN Monte Carlo generators, after full detector simulation. The event shape distributions from PYTHIA 6 and HERWIG++ show satisfactory agreement with the data, while discrepancies are found between the data and predictions from ALPGEN, MadGraph and PYTHIA 8. Tuning of the event generators to better model the data is currently under investigation, with the input provided by this measurement.

References

[1] O. Biebel, "Experimental tests of the strong interaction and its energy dependence in electron positron annihilation", *Phys. Rept.* **340** (2001) 165–289, doi:10.1016/S0370-1573(00)00072-7.

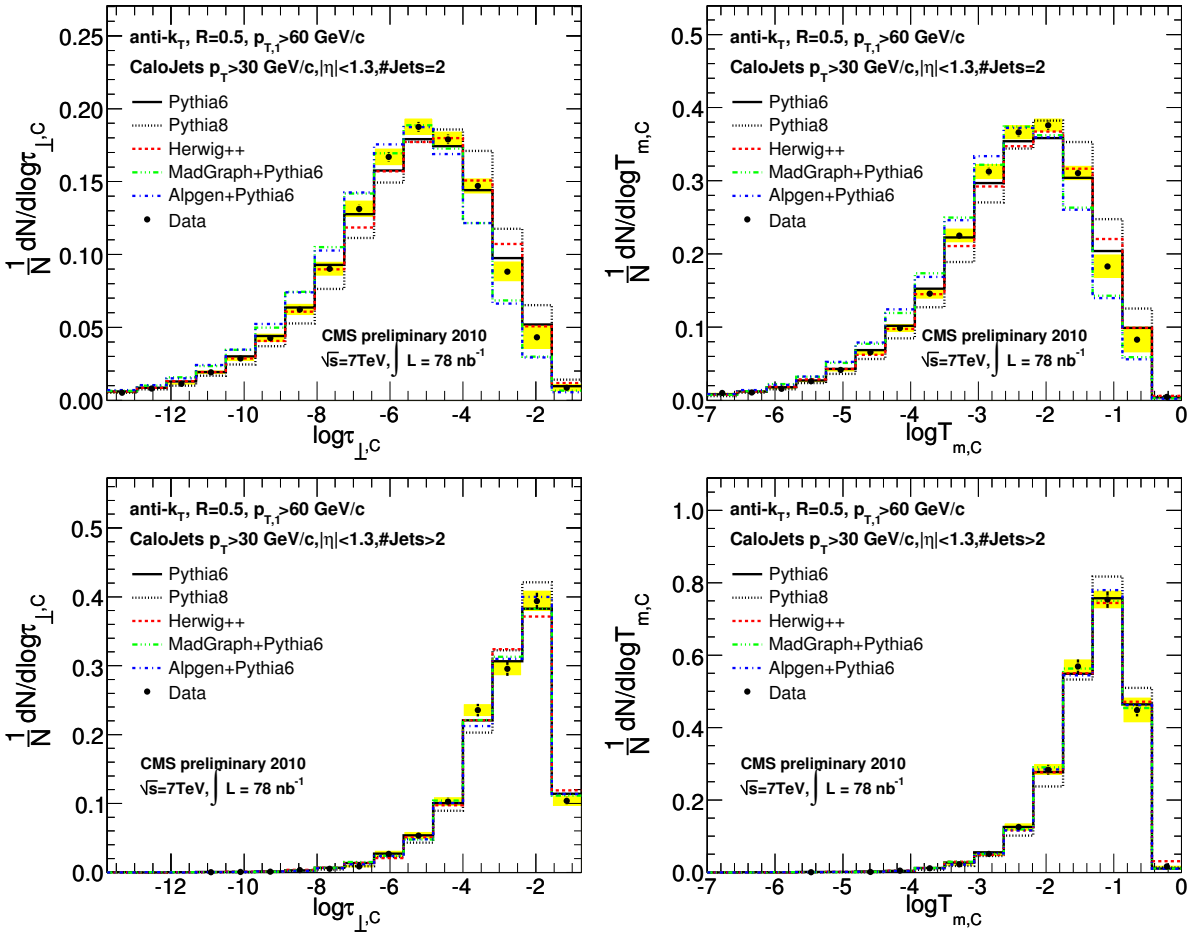

Figure 5: The central transverse thrust (left) and central thrust minor (right) distributions for calorimeter jets, in events with two selected jets (top) and three or more selected jets (bottom) with leading jet $p_{\mathrm{T}} > 60$ GeV/c. The bars represent the statistical error on the data, and the yellow bands represent the sum of statistical and systematic errors.

[2] S. Kluth et al., "A Measurement of the QCD color factors using event shape distributions at $\sqrt{s} = 14$ to 189 GeV", *Eur.Phys.J.* **C21** (2001) 199–210, doi:10.1007/s100520100742, arXiv:hep-ex/0012044.

[3] S. Kluth, "Tests of quantum chromo dynamics at e+ e- colliders", *Rept. Prog. Phys.* **69** (2006) 1771–1846, doi:10.1088/0034-4885/69/6/R04, arXiv:hep-ex/0603011.

[4] A. Banfi, G. P. Salam, and G. Zanderighi, "Phenomenology of event shapes at hadron colliders", *JHEP* **06** (2010) 038, doi:10.1007/JHEP06(2010)038, arXiv:1001.4082.

[5] 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.

[6] **CMS** Collaboration, "Performance of Jet Algorithms in CMS", *CMS-PAS* **JME 07-003** (2007).

[7] **CMS** Collaboration, "Jet Performance in pp Collisions at $\sqrt{s} = 7$ TeV", *CMS-PAS* **JME 10-003** (2010).

[8] **CMS** Collaboration, "The Jet Plus Tracks Algorithm for Calorimeter Jet Energy Corrections in CMS", *CMS-PAS* **JME 09-002** (2009).

[9] **CMS** Collaboration, "Commissioning of TrackJets in pp Collisions at $\sqrt{s} = 7$ TeV", *CMS-PAS* **JME 10-006** (2010).

[10] **CMS** Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", *CMS-PAS* **PFT 10-002** (2010).

[11] **CMS** Collaboration, "Tracking and Primary Vertex Results in First 7 TeV Collisions", *CMS-PAS* **TRK 10-005** (2010).

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

[13] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[14] M. Bahr et al., "Herwig++ Physics and Manual", *Eur. Phys. J.* **C58** (2008) 639–707, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.

[15] J. Alwall et al., "MadGraph/MadEvent v4: The New Web Generation", *JHEP* **09** (2007) 028, arXiv:0706.2334.

[16] S. Hoche et al., "Matching parton showers and matrix elements", arXiv:hep-ph/0602031.

[17] M. L. Mangano et al., "ALPGEN, a generator for hard multiparton processes in hadronic collisions", *JHEP* **07** (2003) 001, arXiv:hep-ph/0206293.

[18] GEANT4 Collaboration, "GEANT4: A simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.

[19] P. Z. Skands, "The Perugia Tunes", arXiv:0905.3418.