IL NUOVO CIMENTO Vol. 33 C, N. 4 Luglio-Agosto 2010
DOI 10.1393/ncc/i2010-10683-0

Colloquia: TOP2010

Reconstruction of jets in top events with CMS

R. Wolf on behalf of the CMS Collaboration
Institut für Experimentalphysik, Universität Hamburg - Hamburg, Germany

(ricevuto il 21 Luglio 2010; approvato il 21 Luglio 2010; pubblicato online il 29 Ottobre 2010)

Summary. — The different reconstruction algorithms for jets and the calibration scheme for the jet energy scale (JEC) at CMS are summarized. The performance of the description of jet quantities in first data is shown. The importance of jets in the reconstruction of top events is emphasized.

PACS `14.65.Ha` – Top quarks.

1. – Introduction

The Large Hadron Collider LHC started to deliver proton proton collision data at center-of-mass energies of 7 TeV with significant luminosity at the beginning of 2010. Up to spring 2010 a luminosity of $17\,\mathrm{nb}^{-1}$ had been collected. This is just one step before the appearance of first events containing reconstructed top quark pairs. The current work of the top quark analysis groups within CMS is to contribute to the understanding of the basic and high level physics objects used for the analysis of top quark pairs. This document is concerned with the importance of jets for top analyses: the jet algorithms used within CMS are listed the calibration scheme for the jet energy scale (JEC) is summarized and a comparison is shown between current data and the expectation from the simulation which indicates the current degree of understanding of jets within CMS.

2. – Jet reconstruction and correction of the jet energy scale

Within the CMS Collaboration three jet algorithms with different opening angles (recombination parameters) are chosen as standard: The seedless infrared-safe cone (`siscone`) algorithm with opening angle 0.5 (0.7); the `kt` algorithm with recombination parameter 0.4 (0.6); the `antikt` (`ak`) algorithm with recombination parameter 0.5 (0.7), where the latter with recombination parameter 0.5 is viewed as the main jet algorithm for common analyses in CMS [1, 2]. Especially in the beginning of data taking CMS foresees to reconstruct jets based on different input objects like pure calorimeter objects (referred to as *calorimeter jets*), particle flow objects (*particle flow jets*) and bare tracks (which will not be covered in this document). In addition to these jet types

Fig. 1. – Scheme for the correction of the jet energy scale. The correction scheme is split in seven major steps leading from reconstructed jets to corrected jets. The jets are corrected for energy offsets due to event pile-up or noise in the calorimeters, different response as a function of the pseudorapidity of the jet, and the absolute energy scale. These correction steps are mandatory. Some more correction steps are optional. The correction steps are completely factorized but nested. This correction scheme has been developed for the correction of jets based on calorimeter objects only (*calorimeter jets*). Apart from some optional correction steps it also holds for the residual correction of *track supported jets* and *particle flow jets*.

calorimeter jets are reconstructed making use of the additional information of tracks enclosed in the cone of the *calorimeter jet* to profit of the superior energy resolution of the CMS track detector (this jet type is further on referred to as *track supported jets*). Jets are corrected for the expected energy scale using a factorized approach as illustrated in fig. 1 [3]. This correction scheme is split in seven major steps leading from reconstructed jets to corrected jets: The jets are corrected for energy offsets due to event pile-up or noise in the calorimeters, different response as a function of the pseudorapidity of the jet, and the absolute energy scale. These correction steps are mandatory. Some more correction steps are optional. The correction steps are completely factorized but nested, which means that for an absolute energy scale correction jets have to corrected for the different response as a function of the pseudorapidity of the jet. This correction scheme has originally been developed for the correction of jets based on calorimeter objects only (*calorimeter jets*). Apart from some optional correction steps, that might turn obsolete it also holds for the residual correction of *track corrected jets* and *particle flow jets*. In fig. 2 an illustration is shown (fig. 2 left) of the correction for the different jet energy response as a function of the pseudorapidity of the jet and (fig. 2 right) of the correction for the absolute energy scale of the jet. Up to present time all corrections are still derived from the simulation, while they are going to be replaced by data-driven methods depending on the validity of the description and the available statistics on data. For the illustrated corrections the data-driven ansatzes are summarized below:

- The correction for different response of the jet energy as a function of the pseudorapidity of the jet is foreseen to be derived from events with a clean di-jet signature. The relative calibration is taken from the difference of the transverse momentum of the two jets, where one of the two jets is required to be well contained in the central detector, while the other jet should be located in the more forward directions of the calorimeter.

- The correction for the absolute energy scale is foreseen to be taken from event with a clean photon/Z boson and jet signature, where the photon/Z boson sets the energy scale that should be balanced by the single back-to-back jet. The resolution will be dominated by the resolution of the jet energy scale. The Z boson might be reconstructed from the decay into two muons or electrons. These correction factors might be used up to a transverse momentum of the photon/Z boson of 150–200 GeV. Extrapolations to larger values might be taken from the simulation or from tri-jet topologies, where two jets are chosen to balance a third high transverse momentum jet.

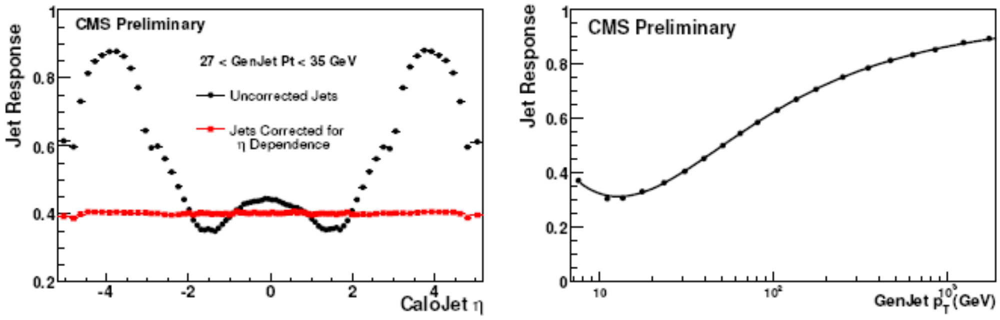

Fig. 2. – Illustration of the effect of the correction (left) of the jet energy response as a function of the pseudorapidity of the jet and (right) of the absolute jet energy scale. The histograms show the response of reconstructed *calorimeter jets* from simulated events before and after applying the corrections as derived from the simulation.

Due to non-linearities of the CMS calorimeters the response of the jet energy scale depends on the flavor of the jets. The jet energy scale is extrapolated to the expected flavor mixture in di-jet topologies using simulation. For startup a correction up the the absolute energy scale is assumed with an uncertainty of ±10%. A first comparison of data with the expectation from the simulation shows that these assumptions are well justified and maybe even conservative. It supports the use of the jet energy correction factors as determined from the simulation with residual corrections from jet energy corrections as derived from the data.

3. – Performance of the description of jets in first data

By spring 2010 the collected and approved statistics were still too small to select and reconstruct events containing top quark pairs. It was possible though to confirm a good performance of the description of the reconstructed high-level analysis objects in the current simulation [4]. In fig. 3 a selection of variables for the cleaning and identification of jets is shown. All quantities are shown after a loose di-jet event selection. The *antikt* algorithm has been used with a recombination parameter of 0.5 for these and all following histograms. The simulated events have been produced using the Pythia event generator [5] and passed through the full simulation of the CMS detector and trigger system. For the comparisons the simulated events have been normalized to the yield in data. In the upper row of fig. 3 the electromagnetic fraction of the reconstructed *calorimeter jets* is shown for jets with transverse momentum larger than 25 GeV (left) and 50 GeV (right). A reasonable agreement can be seen after a cut on the transverse momentum larger than 25 GeV which can even be improved by applying harder requirements on the transverse momentum on the jets, where the simulation is expected to be more reliable in the description of the data. In fig. 4 the basic jet kinematics for all jets after a loose di-jet selection are shown for the three different jet types as described above. In the upper row the pseudorapidity of all jets in the central detector with (corrected) transverse momentum larger than 25 GeV is shown. In the lower row the difference of the two leading jets in the azimuthal angle ϕ is shown. The distributions are shown in the left column for *calorimeter jets*, in the middle column for *track supported jets* and the right column for *particle flow jets*. The simulation has been normalized to the yield from

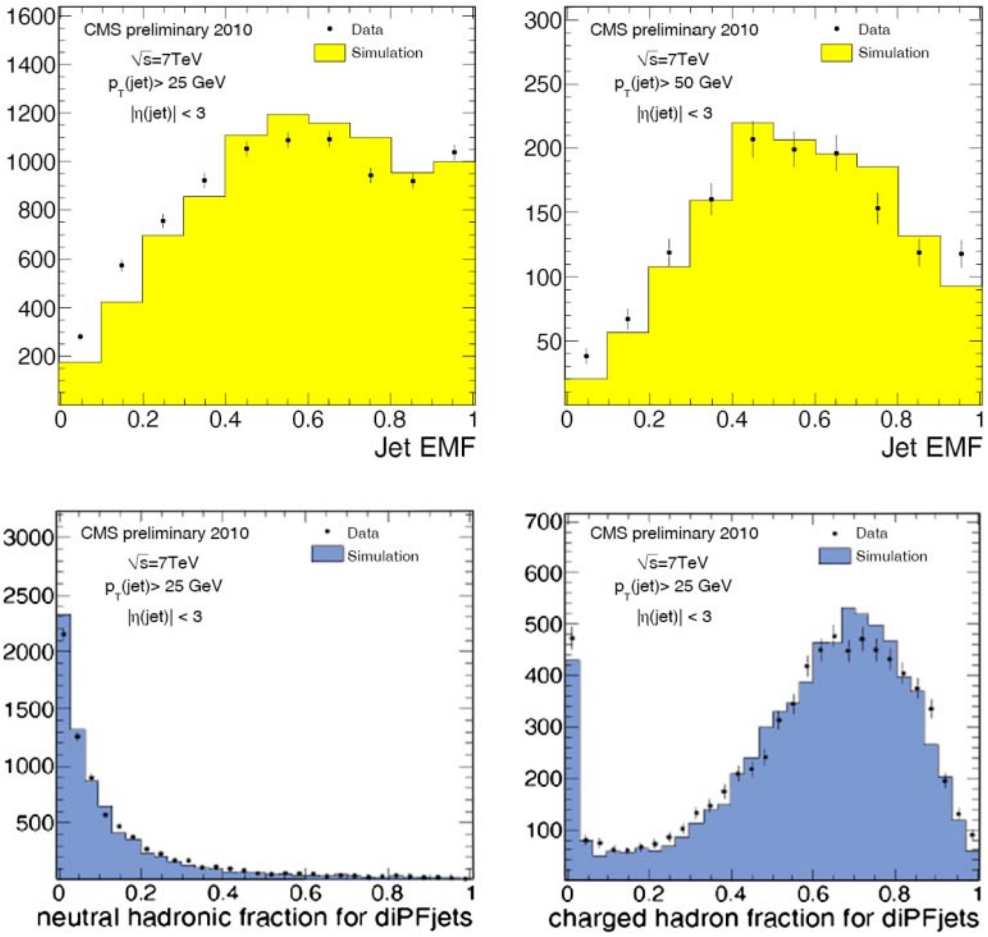

Fig. 3. – Comparison of jet cleaning and identification quantities for *calorimeter jets* and for *particle flow jets* using the *antikt* algorithm with recombination parameter of 0.5. Shown are the electromagnetic energy fraction for *calorimeter jets* with transverse momentum larger than (upper left) 25 GeV and (upper right) 50 GeV and the energy fraction carried by (lower left) neutral and (lower right) charged particle flow hadrons. For these comparisons the simulation has been normalized to the yield in data. We state reasonable to good description of the given shapes.

data. The back-to-back topology of the di-jet system is clearly visible and an overall good description can be seen for all jet types.

4. – Conclusions

The Large Hadron Collider LHC that will be a factory of top quarks resumed operation in autumn 2009. By the time of the conference the CMS experiment had collected and recorded an integrated luminosity of 17 nb^{-1}. Unfortunately this amount of luminosity is still too small to select and reconstruct candidates for top quark pairs. It allows though

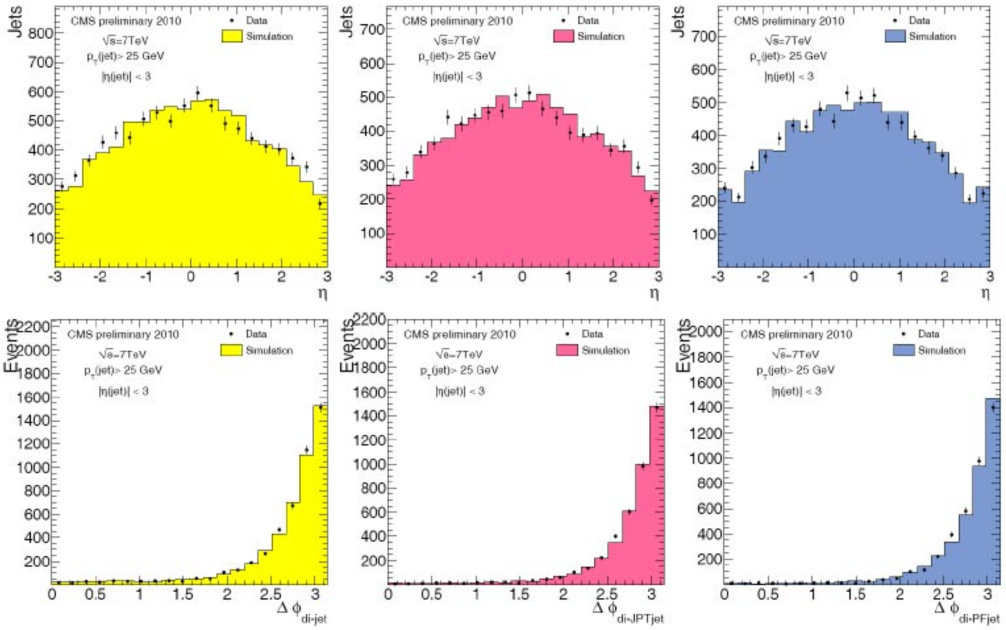

Fig. 4. – Comparison of the basic jet kinematics, pseudorapidity (upper row) and transverse momentum (lower row) for *calorimeter jets* (left column), *track supported jets* (middle column) and *particle flow jets* (right column) using the *antikt* algorithm with recombination parameter of 0.5. For these comparisons the simulation was normalized to the yield in data. We state a good agreement for all input objects.

to check the understanding of the objects involved in the selection and reconstruction of top quark pairs. The importance of jets among those objects is evident. In this paper the jet types and algorithms which are taken into consideration for measurements with the CMS detector and the correction scheme for the energy scale of these jets are summarized. The standard algorithm will be the *antikt* algorithm with recombination parameter of 0.5, while also other algorithms will be deployed. Three different types of jets, *calorimeter jets*, *track supported jets* and *particle flow jets* are foreseen, which has different impacts on the size of the jet energy corrections and the resolution of the jet energy. First comparisons show a good agreement of the simulation with data for jet kinematics, jet identification and cleaning variables and di-jet topologies.

REFERENCES

[1] Salam G., *Eur. Phys. J. C*, **67** (2010) 637.
[2] CMS Collaboration, *Physics Technical Design Report, Vol. I* (2006).
[3] CMS Collaboration, *CMS PAS JME-07/002.*
[4] CMS Collaboration, *CMS DP-2010/014.*
[5] Sjstrand T., *Monte Carlo generators for the LHC (1/4)*, CERN Lecture (2005).