Performance of the particle flow algorithm in CMS

Daniele Benedetti for the CMS Collaboration

CERN, 1211 Geneva, Switzerland

DOI: http://dx.doi.org/10.3204/DESY-PROC-2010-01/221

The aim of the CMS particle flow algorithm is to identify and reconstruct individually each particle arising from the LHC proton-proton collision, by combining the information from all subdetectors. The resulting particle-flow event reconstruction leads to an improved performance for the reconstruction of jets and MET, and for the identification of electrons, muons, and taus.

1 The particle-flow algorithm

The CMS [1] particle-flow event reconstruction [2] combines the information from all subdetectors to identify and individually reconstruct all particles produced in the collision, namely charged hadrons, photons, neutral hadrons, muons, and electrons. The resulting list of particles can then be used to build jets, to determine the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$), to reconstruct and identify taus from their decay products, and to quantify charged lepton isolation with respect to other particles.

2 Performance of the particle-flow event reconstruction in simulated data

The typical jet energy fractions carried by charged particles, photons and neutral hadrons are 65%, 25% and 10% respectively [2]. These fractions ensure that 90% of the jet energy can be reconstructed with good precision by the particle-flow algorithm with the CMS detector, thanks to the excellent tracking efficiency and electromagnetic calorimeter resolution [1], while only 10% of the energy is affected by the poor hadron calorimeter resolution and by calibration corrections of the order of 10% to 20%. As a consequence, the jets made of reconstructed particles are expected to be much closer, in energy and direction, to jets made of Monte-Carlo-generated particles than jets made from the sole calorimeter information.

2.1 Jet energy response and resolution

Jets are reconstructed from the QCD-multijet event sample with the iterative-cone algorithm [1] with a cone size of 0.5 in the (η, ϕ) plane, from several types of inputs: all generated stable particles ("gen-jets"), particles reconstructed with the particle-flow algorithm ("particle-flow jets") and calorimeter towers ("calo-jets"). The reconstructed jets are then matched to the closest gen-jet in the (η, ϕ) plane. The jet response, defined as the Gaussian mean of the

$(p_T^{rec} - p_T^{gen})/p_T^{gen}$ distribution (where "rec" and "gen" hold for reconstructed and generated jets, respectively), is shown in Fig. 1 (left) for several p_T bins. The particle-flow-jet response benefits from the reconstruction of all particles in the event from a combination of all CMS sub-detectors, which ensures that little energy is lost over the whole acceptance. The particle-flow jet-energy resolutions, obtained by dividing the Gaussian width σ by the average jet response, in each p_T bin, are compared to the fully corrected calo-jets in Fig. 1 (right). Up to three times better resolution for jets is obtained using the particle-flow event reconstruction [2].

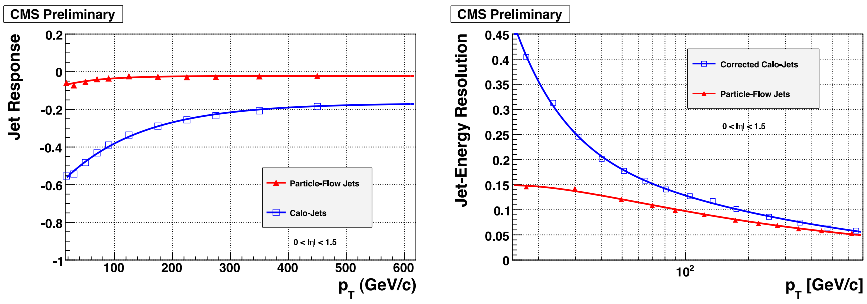

Figure 1: Jet response (left) and jet-energy resolution (right) as a function of p_T in the CMS barrel region ($|\eta| < 1.5$) for the particle-flow jets (triangles) and the calo-jets (squares).

3 Commissioning of the particle-flow event reconstruction with the first LHC collisions

3.1 Particles: photons, charged and neutral hadrons

The absolute photon-energy calibration and the uniformity of the electromagnetic calorimeter (ECAL) response can be checked with the abundant π^0s in the data recorded at $\sqrt{s} = 900$ GeV. The photon-pair invariant-mass distribution is shown in Fig. 2 (left) and is fit with a Gaussian for the π^0 signal added to an exponential function of the invariant mass for the combinatorial background. The agreement for the measured mass values in data and simulation with the world average of 135 MeV/c^2 [3] to within $\pm 2\%$ demonstrates the suitability of the simulation-based absolute ECAL cluster calibration for low-energy photons in the data [4].

The energy response of the calorimeters to hadrons and its calibration is also important for the particle-flow algorithm. An improper calorimeter calibration would lead to a systematic mis-estimation of both the energy and multiplicity of neutral hadrons [4]. Consistency for the charged-hadron calibration ensures the proper energy calibration for neutral hadrons as well. To verify the calibration procedure the average calibrated calorimeter response, integrated over the pseudorapidity range $|\eta| < 2.4$, is displayed in Fig. 2 (right) as a function of the measured track momentum, from 1 to 30 GeV/c. This figure demonstrates that *(i)* the calorimeters respond to charged hadrons as predicted from the simulation; and *(ii)* the hadron cluster calibration obtained from the simulation of the CMS detector is adequate, on average, for use of the particle-flow event reconstruction in data.

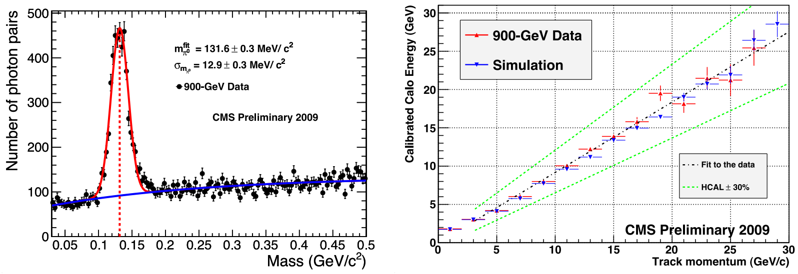

Figure 2: Left: photon-pair invariant-mass distribution in the barrel ($|\eta| < 1.0$) for the data. Right: average calibrated calorimeter response as a function of the track momentum for the 900 GeV data (light upwards triangles) and for the simulation (dark downwards triangles). The dashed lines show the same quantity when the HCAL raw response is changed by $\pm 30\%$

3.2 Jets

To demonstrate the reliability of the particle-flow event description of the jet constituents, the jet energy fraction [4] as a function of pseudorapidity is shown in Fig. 3 for the data and the simulation. In the tracker-covered region, charged hadrons are found to carry on average 65% of the jet energy, photons 15% and neutral hadrons 20%. The higher fraction of neutral hadrons, with respect to the one indicated in Sec. 2, is produced by clusters arising from the hadronic calorimeter noise and due to the very low p_{T} threshold applied in this jet selection [4]. The data and simulation are found to be in good agreement.

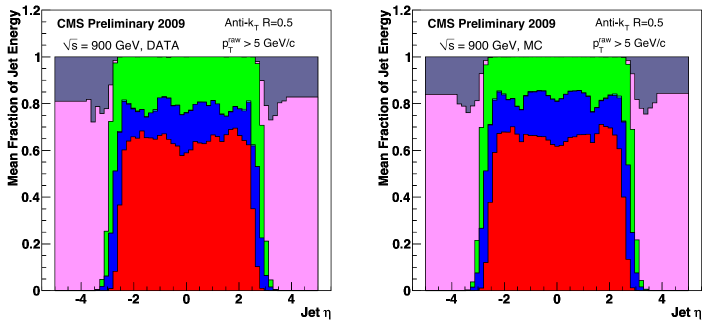

Figure 3: Reconstructed jet energy fractions as a function of pseudorapidity in the data (left), and in the simulation (right). From bottom to top in the central region: charged hadrons, photons, electrons (less than 1%), and neutral hadrons. In the forward region: hadronic deposits, electromagnetic deposits.

3.3 Missing transverse energy

The missing transverse energy allows for an indirect detection of invisible particles produced in proton–proton collisions, such as neutrinos or neutralinos. In the particle-flow event reconstruction, the missing transverse energy vector is computed as the opposite of the transverse-momentum sum of all particles reconstructed in the event, $E_{\mathrm{T}}^{\mathrm{miss}}$ is its modulus and its projections on the x and y axes are denoted $E_{\mathrm{x}}^{\mathrm{miss}}$ and $E_{\mathrm{y}}^{\mathrm{miss}}$, respectively. As a large $E_{\mathrm{T}}^{\mathrm{miss}}$ is one of the most promising signatures for new physics, it is important to ensure that experimental artifacts do not give rise to fake particles with large energies. Unlike $E_{\mathrm{T}}^{\mathrm{miss}}$, where experimental effects somewhat cancel out due to calculations involving differences in momentum, all detector effects are added up in ΣE_{T}, i.e. the scalar-sum of the transverse energies over all reconstructed particles. ΣE_{T} represents an excellent benchmark for evaluating the performance of the generator, the detector simulation, and the reconstruction algorithm.
The distribution of $E_{\mathrm{T}}^{\mathrm{miss}}/\Sigma E_{\mathrm{T}}$ is displayed for events with $\Sigma E_{\mathrm{T}} > 3$ GeV in Fig. 4 (left). Ideally, events with no expected $E_{\mathrm{T}}^{\mathrm{miss}}$ (as is the case in minimum-bias collisions) should have very low values of $E_{\mathrm{T}}^{\mathrm{miss}}/\Sigma E_{\mathrm{T}}$. This figure confirms that, for a given estimate of the ΣE_{T}, the particle-based $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is, on average, twice better than the calorimeter reconstruction [4]. Another way to visualise the improved $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is to parametrise it as a function of ΣE_{T}. To do so, the distribution of $E_{\mathrm{x}}^{\mathrm{miss}}$ and $E_{\mathrm{y}}^{\mathrm{miss}}$ was fit to a Gaussian, for several bins of ΣE_{T}. The resulting width $\sigma(E_{\mathrm{x,y}}^{\mathrm{miss}})$, shown as a function of ΣE_{T} in Fig. 4 (right), was fit by the functional form $a \oplus b\sqrt{\Sigma E_{\mathrm{T}}}$. This fit yields $a = 0.55$ GeV and $b = 45\%$ for the particle-based reconstruction.

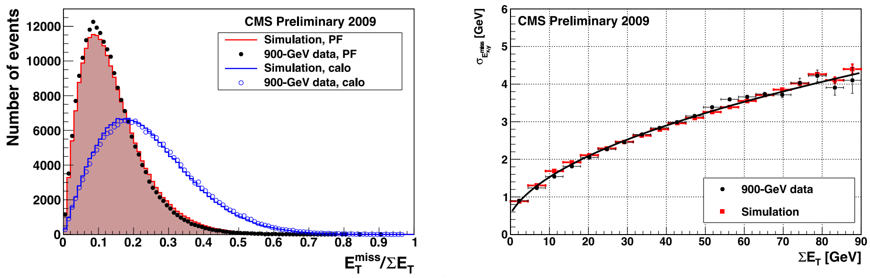

Figure 4: Left: distribution of the particle-based (solid) and calorimeter-based (hollow) $E_{\mathrm{T}}^{\mathrm{miss}}/\Sigma E_{\mathrm{T}}$ in the data (dots) and in the simulation (histogram). Right: resolution of the particle-based $E_{\mathrm{x,y}}^{\mathrm{miss}}$ as a function of the particle-based in the data (dots) and in the simulation (squares).

References

[1] CMS Collaboration, R. Adolphi *et al.*, "The CMS experiment at the CERN LHC," JINST **3** (2008) S08004.

[2] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus and $E_{\mathrm{T}}^{\mathrm{miss}}$", CMS PAS **PFT-09-001** (2009).

[3] C. Amsler *et al.* [Particle Data Group], Phys. Lett. B **667** (2008) 1.

[4] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", CMS PAS **PFT-10-001** (2010).