Available on CMS information server **CMS IN 2007/xxx**

CMS Internal Note

The content of this note is intended for CMS internal use and distribution only

June 15, 2007

Jet and MET Performance in CMSSW_1_2_0

M. Vazquez Acosta

CERN, Geneva, Switzerland

L. Apanasevich, C. Dragoiu, A. Smoron, N. Varelas

University of Illinois at Chicago, Chicago, IL , USA

A. Bhatti

Rockefeller University, New York, NY, USA

R. Cavanaugh, B. Scurlock

University of Florida, Gainesville, FL , USA

R. Demina, J. Widawsky, M. Zielinski

University of Rochester, Rochester, NY, USA

G. Dissertori

ETH Zurich, Switzerland

R. Harris

Fermilab, Batavia, IL, USA

M. Jha

University of Delhi, Delhi, India

S. Petrushanko, O. Kodolova

Moscow State University, Moscow, Russia

F. Ratnikov

University of Maryland, College Park, MD, USA

Abstract

We summarize results of jet and Missing ET (MET) performance studies with CMSSW_1_2_0. This note serves to establish a baseline for comparing future releases of CMSSW and to document the status of jet and MET simulation and reconstruction established at the conclusion of the Software and Detector Performance Validation (SDPV) exercise.

1 Introduction

The majority of results presented in this note has been obtained during the course of the Software and Detector Performance Validation (SDPV) effort. The main goal of the SDPV was to establish the then-current level of performance of CMSSW simulation and reconstruction and compare to selected benchmark results based on the ORCA software and published in the CMS Physics TDR [1]. Detailed comparisons between CMSSW and ORCA results can be found in the SDPV meeting presentations [2] and the final SDPV report to the collaboration [3].

In this note we report the jet and MET performance studies based on CMSSW_1_2_0 in order to establish a baseline for comparing results from future releases of CMSSW. However, for many plots we retain the original format intended for comparisons with ORCA and include selected direct comparisons or comments, when of particular interest. When making any detailed comparisons between CMSSW and ORCA one needs to keep in mind various differences in the respective choices of settings, including:

- Tune DWT for the Underlying Event parameters in PYTHIA has been used in CMSSW (older settings in ORCA)
- Different versions of GEANT, detector material description and geometry
- Differences in calorimeter simulation/digitization/calibration, different Shower Library for the Forward Hadron calorimeter (HF)
- In CMSSW, the HCAL energy is calculated from the sum of four time-slices (while two time-slices were used in ORCA)
- Pileup settings, as discussed below (ORCA plots typically correspond to low-lumi pileup, $2\times10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$)
- Recombination E-scheme for jets is the default in CMSSW (ET-scheme used for many ORCA studies)
- Scheme-B calorimeter cell thresholds for jet and MET reconstruction (and 0.5 GeV tower E_T cut for jets) are used in CMSSW. In ORCA, a zero suppression threshold of 0.5 GeV was applied on HCAL cell energies and cuts on tower $E > 0.8$ GeV and $E_T > 0.5$ GeV were commonly used for jets.

After a brief overview of jet reconstruction in CMSSW_1_2_0 (Section 2) and Monte Carlo samples employed (Section 3), we discuss the derivation of MCJet jet energy corrections (Section 4), impact of pileup on jet reconstruction (Section 5), jet energy response, energy resolutions and position resolutions for various pileup samples (Section 6), jet efficiencies (Section 7), dijet balance approach to jet response equalization across the detector (Section 8). Missing ET performance is discussed in detail in Section 9. The paper concludes with Section 10.

2 Jet Reconstruction in CMSSW_1_2_0

CMSSW_1_2_0 contains three jet reconstruction algorithms, each of which is able to create three "flavors" of jet objects depending on the type ("flavor") of input objects. Jet reconstruction is a two-step procedure. In the first step, an arbitrary input collection is treated as a set of Lorentz vectors. Every Lorentz vector satisfying energy and/or E_T requirements is used by the jet clustering algorithm. In the second step, after jet clustering is completed, flavor specific information is extracted from objects contributing to the jet, and corresponding flavor-specific values are associated with the jet.

Three jets flavors are available:

- *BasicJets* are produced from arbitrary set of constituents. No specific constituent information is used, the produced jet contains only kinematics parameters and back references to the jet constituents.
- *CaloJets* are produced from constituents which are CaloTowers. Every CaloTower is built from one or more HCAL cells and corresponding ECAL crystals. In the barrel, a single projective Calorimeter Tower is formed from the unweighted sum of energy deposits in 5×5 ECAL crystals and in a single HCAL cell. In the forward regions, a more complex association of ECAL crystals with HCAL cells is required. In order to reduce the effects of electronic noise in CMSSW, so-called "Scheme-B" thresholds [4] were applied (by default) to all Calorimeter Towers:

Scheme	HB	HO	HE	ΣEB	ΣEE
A	0.70	0.85	0.90	0.20	0.45
B	0.90	1.10	1.40	0.20	0.45
C	1.20	1.30	1.80	0.20	0.45

In addition to basic kinematics parameters, CaloJets contain specific information collected from contributing CaloTowers, like energies deposited in differentECAL and HCAL detector regions, fractions of electromagnetic and hadronic energies of the jet, maximum electromagnetic and hadronic energies deposited in one tower etc.

- *GenJets* are produced from constituents which are Monte-Carlo generator level particles. Only particles marked as stable on the generator level contribute to a GenJet. A GenJet contains information about basic kinematic variables and energy contributions from electromagnetic particles, hadrons, and invisible particles contributing to the jet eg., neutrinos.

Three jet clustering algorithms are used in CMSSW_1_2_0 (see [1, 5] for additional details and a discussion of ORCA results):

- *Iterative Cone* algorithm ported from ORCA. A seed threshold is applied, input objects are assigned to a jet that is found first and then constituents are excluded from the input list; no merging/splitting is applied.
- *Midpoint Cone* algorithm has been developed for and is widely used in Tevatron Run II jet reconstruction. In CMSSW it is ported directly from CDF code. Seed threshold is applied and proto-jet clusters are found without removing constituents that have been already used. A splitting-merging step is invoked for the produced proto-jet clusters, controlled by the Overlap Threshold parameter. The Search Cone step was not applied.
- Kt is an implementation of the standard $K_\perp$ algorithm ported using the original code [6]. The CMSSW_1_2_0 implementation of the Kt algorithm had known problems and hence it has been substituted by the external original Kt code in following releases and finally superceded by the FastJet implementation [7].

Table 1 aummarizes the parameters used for CaloJet reconstruction in CMSSW_1_2_0.

	Iterative Cone	Midpoint Cone		Kt
Alias	IC5	MC5	MC7	KT
Size	$R = 0.5$	$R = 0.5$	$R = 0.7$	$D = 1$
Tower thresholds	$E_T > 0.5\ GeV$	$E_T > 0.5\ GeV$		$E > 1\ GeV$
Cell thresholds	Scheme B			
Seed threshold	$E_T > 1\ GeV$	$E_T > 1\ GeV$		NA
Other parameters	NA	Overlap Threshold = 0.75 maxPairSize = 2 maxIterations = 100		NA

Table 1: Parameters for CaloJet reconstruction algorithms in CMSSW_1_2_0.

The Iterative Cone algorithm is the fastest one. The Kt algorithm is well defined theoretically, but is also the slowest due to N^3 dependency on the total number of input objects. However, the new FastJet implementation of Kt, available since CMSSW_1_3_x, is much faster, comparable in speed to the Iterative Cone algorithm. Figure 1 illustrates the CPU time necessary for different algorithms to reconstruct jets for different input muliplicities.

MET reconstruction is discussed in Section 9.

3 CMSSW_1_2_0 QCD Samples

This analysis employs Monte Carlo QCD event samples produced for the SDPV exercise using CMSSW_1_2_0, and a few special test samples as described below. The particle-level events were generated with PYTHIA 6.227

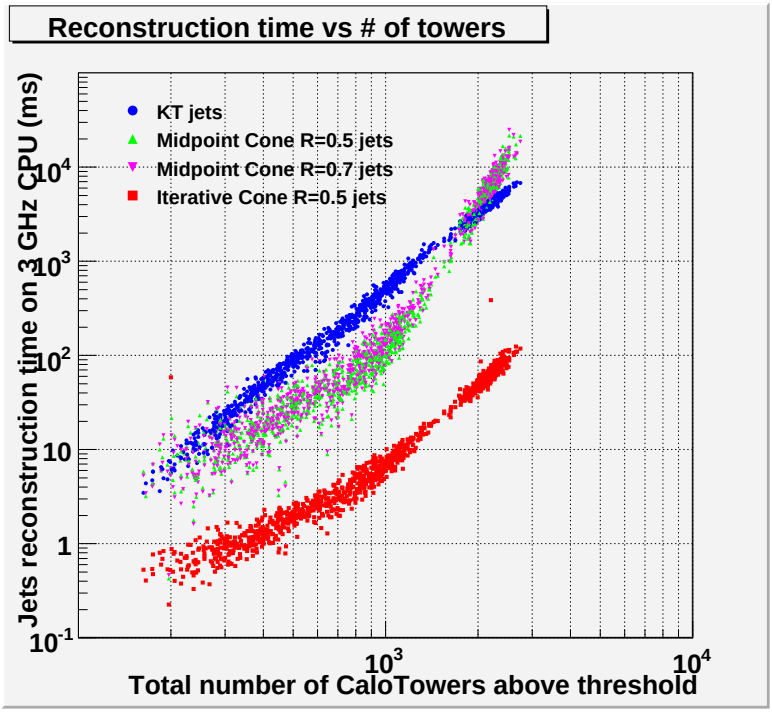


Figure 1: CPU time necessary for different algorithms to reconstruct CaloJets as a function of the total number of CaloTowers above threshold in the event. The range $N_{\mathrm{CaloTower}} < 700$ corresponds to low luminosity, the range $500 < N_{\mathrm{CaloTower}} < 1500$ corresponds to a medium multiplicity of five extra mimimum bias interactions per event, and the range $1500 < N_{\mathrm{CaloTower}} < 3000$ corresponds to a high multiplicity of 25 extra mimimum bias interactions per event.

using the Tune DWT for Underlying Event parameters [8]. The CMS detector simulation as implemented in CMSSW_1_2_0 based on the GEANT4 package was used to simulate passage of particles through the detector and the energy deposits in the sensitive volumes. Most of the results below have been derived from samples without pileup (PU), unless stated otherwise. QCD dijet samples with no pileup were generated in 21 bins of the momentum transfer in the parton hard-scatter, $\hat{p}_T$, which span the full kinematic range [9].

We include a brief discussion of various pileup settings for CMSSW_1_2_0 and illustrate the impact of PU on some selected results. To simulate additional proton-proton interactions in a beam crossing from PU, the signal events were mixed with a random number of minimum bias events in one crossing. The minimum bias events were generated with PYTHIA as inclusive QCD events. The Poission distribution with an average of five was used to simulate pileup, corresponding to the so-called "low-lumi" conditions. The QCD samples with in-time pileup were part of the official OnSel production, and were available for $\hat{p}_T < 800$ GeV (samples with no pileup were used for higher $\hat{p}_T$ values to extend the range of coverage in some plots, as needed).

Finally we also used small samples with full (in-time plus out-of-time) low-lumi pileup from a private production for $50 < \hat{p}_T < 80$ GeV [10], and with a HF Shower Library fix corresponding to CMSSW_1_2_2 (for $80 < \hat{p}_T < 120$ GeV and no pileup).

4 MC Jet Corrections

Here we describe the MC Jet corrections which correct the calorimeter level jet (CaloJet) to have the same E_T as the particle level jet (GenJet). These corrections are based on a QCD dijet Monte Carlo simulation using Pythia and the CMS detector simulation with no pileup. The methodology was developed previously [5] for OSCAR/ORCA and ported to CMSSW. We adopt this methodology in order to provide corrections to CMSSW data samples.

4.1 Jet Response for Jet Corrections

Measurements of the jet response were made with the module SimJetResponseAnalysis in the MCJet package of the JetMETCorrections subsystem of CMSSW. Jets were reconstructed using the iterative cone and midpoint cone algorithms in CMSSW with the E recombination scheme. We consider all GenJets in the event and match each GenJet with the closest CaloJet which minimizes

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}, \tag{1}$$

where η and ϕ denote jet pseudorapidity and azimuthal angle, respectively. If the closest CaloJet is not within $\Delta R = 0.25$, the GenJet is discarded. For all passing GenJets we measure

$$\text{Jet Response} = \frac{\text{CaloJet } E_T}{\text{GenJet } E_T} \tag{2}$$

in the bins of GenJet E_T and CaloJet $|\eta|$ listed in table 2 and table 3:

10	12	15	20	27	35	45	57	72	90	120
150	200	300	400	550	750	1000	1400	2000	2900	4500

Table 2: Bin edges of GenJet E_T in GeV used for measurements of the jet response.

0.0	0.226	0.441	0.751	0.991	1.260	1.496	1.757	2.046
2.295	2.487	2.690	2.916	3.284	4.0	4.4	4.8	

Table 3: Bin edges of CaloJet $|\eta|$ used for measurements of the jet response.

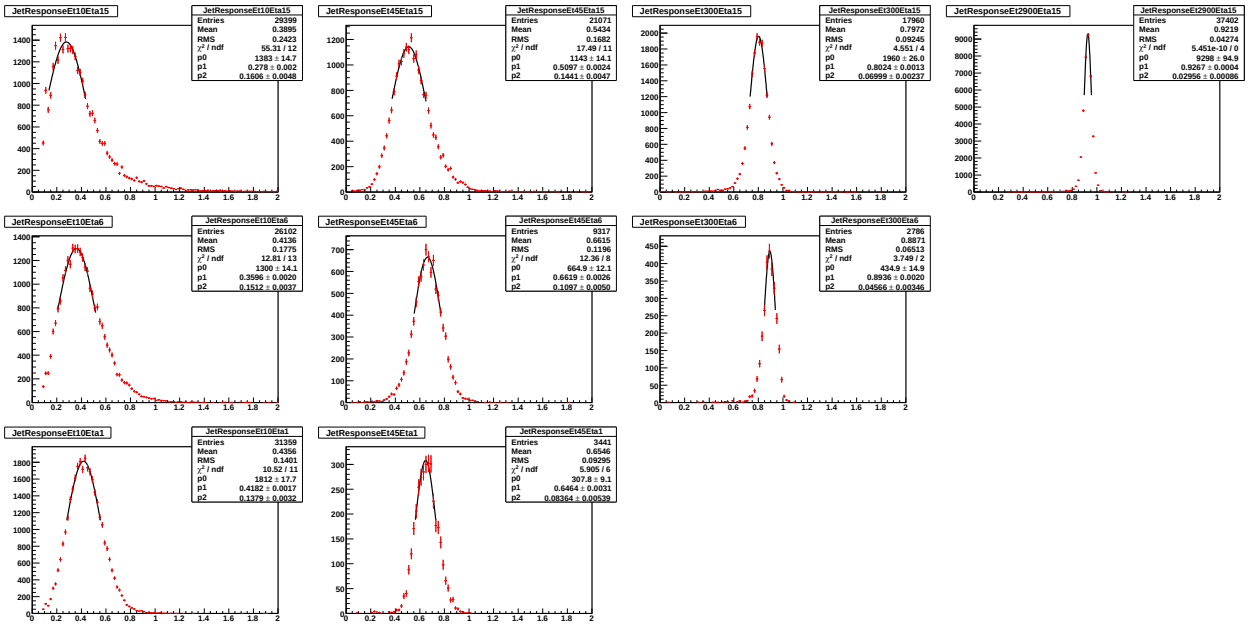


Figure 2: Jet response and fit for the iterative cone algorithm with cone size $R = 0.5$. The number of jets as a function of jet response (points) is compared to a Gaussian fit (curve) in the interval $\pm 1\sigma$ from the peak response. The three rows of plots, from top to bottom, are for the following regions of CaloJet η: $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$. The four columns of plots, from left to right, are for the following regions of GenJet E_T: $10 < E_T < 12$, $45 < E_T < 57$, $300 < E_T < 400$, and $2900 < E_T < 4500$ GeV.

Example histograms of the jet response are shown in Fig 2. Notice that for each η region (row) the response increases and the resolution improves with increasing E_T (column). To determine the peak of the jet response, the most probable value, we have fit each of the histograms in Fig. 2 with Gaussians in the interval $\pm 1\sigma$ from the peak.

If a full Gaussian fit is used instead of $\pm 1\sigma$, the mean value of the Gaussian increases by less than 3% depending on E_T. We use the mean value of the $\pm 1\sigma$ Gaussian fits to define the peak jet response at the average GenJet E_T in a specific bin. We use the peak jet response as input to determine the jet correction for the following reasons:

- The peak is easy to find and well defined.
- The peak is optimal for correcting dijet resonances.
- The peak is less sensitive to the details of the thresholds and cuts applied to the calibration samples.

Example plots of peak jet response as a function of average GenJet E_T are shown in Fig. 3. The response increases smoothly with E_T. For each of the 16 bins of CaloJet $|\eta|$, we fit the response with the same parameterization used for ORCA [5]. The parameterization is compared to the response points in Fig. 3.

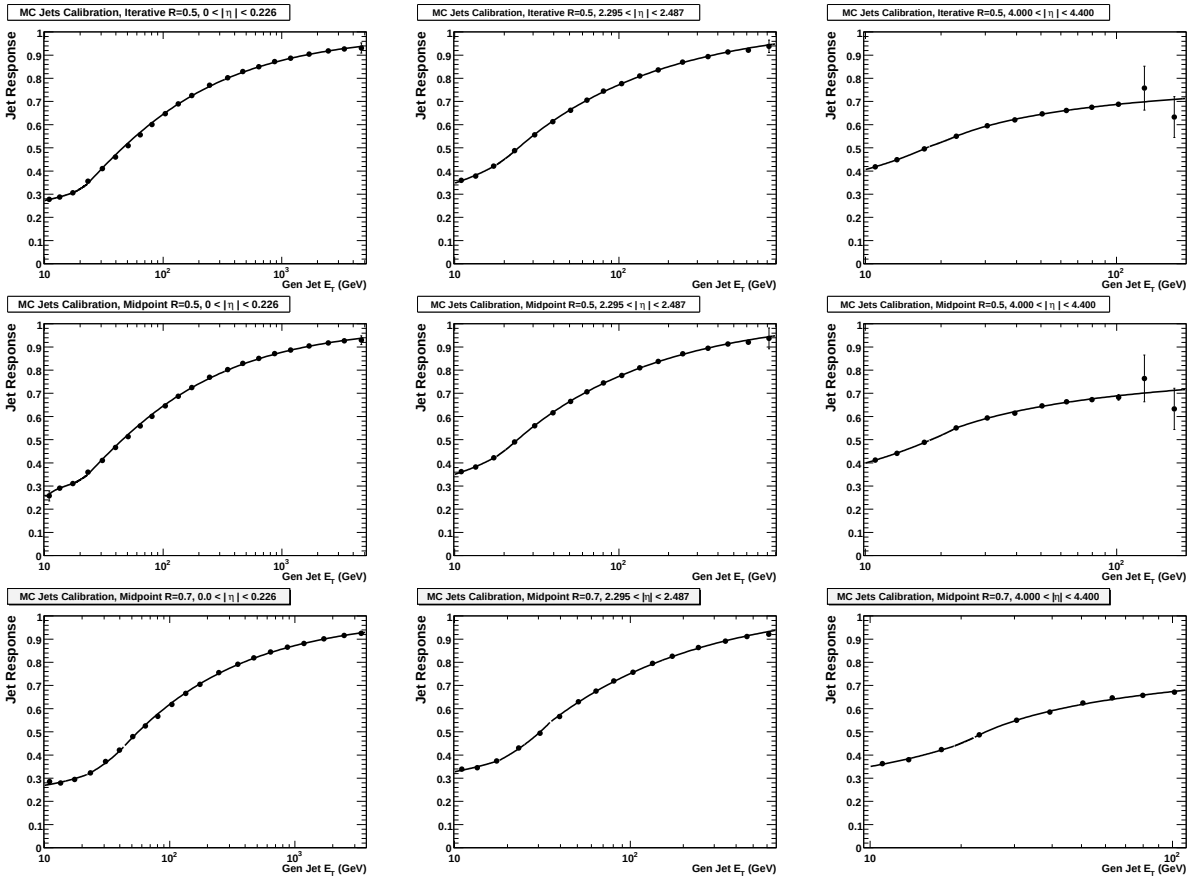


Figure 3: The jet response as a function of GenJet E_T (points) is compared to a parameterization (curve). The three rows of plots, from top to bottom, are for the iterative cone algorithm with a cone size $R = 0.5$, and for the midpoint cone algorithm with a cone size of $R = 0.5$ and $R = 0.7$. The three columns of the plot, from left to right, are for the following regions of CaloJet η: $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$.

Changes in jet response with different versions of the CMS detector simulations are illustrated in Fig. 4. The CMSSW_1_2_0 response is compared with both CSA06 [11] and ORCA [5] samples. CSA06 used CMSSW_0_8_3 simulation and CMSSW_1_0_3 reconstruction. We see that the CMSSW_1_2_0 response is higher than CSA06 in the barrel and endcap. The following changes are known to have occurred and contribute to that difference

- The HCAL Endcap (HE) response in CSA06 was low by 35% because the reconstruction used the same calorimeter sampling fraction as the HCAL Barrel (HB), and this was fixed in time for the CMSSW_1_2_0 release.

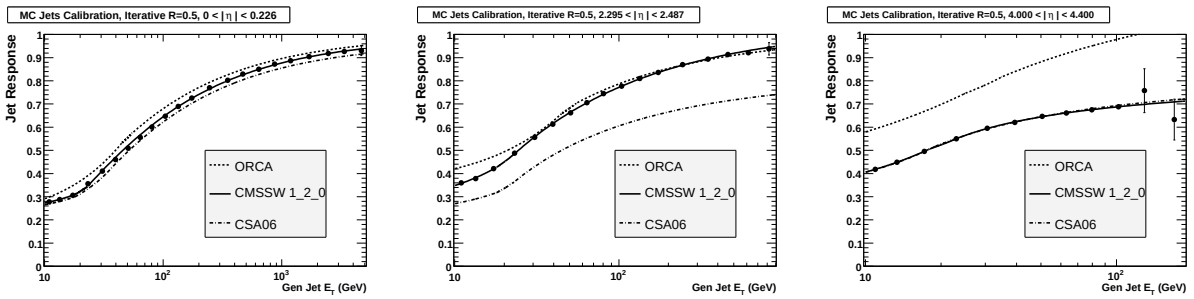


Figure 4: The jet response as a function of GenJet E_T from ORCA (dashed curve), CSA06 (dot-dashed curve), and CMSSW_1_2_0 (points and solid curve). The three plots, from left to right, are for the following regions of CaloJet η: $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$.

- The RecHit energy in HB and HE was increased by 3% in CMSSW_1_2_0 to account for signal lost outside of the 4 time-slice integration window, and this correction was not present in CSA06.

The response in CMSSW_1_2_0 remains different than seen with ORCA by a few percent in the Barrel and Endcap, and by around 25% in the Forward due to the use of an uncalibrated shower library.

4.2 Jet Energy Corrections

The jet energy correction, k, is defined as

$$k = \frac{1}{\text{Jet Response}}. \tag{3}$$

It is a multiplicative correction: the CaloJet Lorentz vector, p,is multiplied by the jet correction to obtain a corrected CaloJet Lorentz vector, p':

$$p' = kp \tag{4}$$

The parameterized jet response as a function of GenJet E_T in 16 slices of CaloJet $|\eta|$ is used as input to the MC Jet package in the JetMETCorrections subsystem of CMSSW.

The software applies a simple iteration procedure to derive the jet correction as a function of the observed CaloJet E_T from the input Jet Response which is a function of true GenJet E_T. Let i be the iteration number, then k_i is the correction obtained in the ith iteration, and is equal to

$$k_i = \frac{1}{\text{Jet Response(CaloJet } E_T \times k_{i-1})} \tag{5}$$

where $k_0 = 1$. In equation 5 we are substituting an approximation for the GenJet E_T into the Jet Response function of GenJet E_T, and with each iteration the approximation becomes more precise. The software iterates ten times to obtain a value of the jet correction as a function of CaloJet E_T which has safely converged. In other words, for each reconstructed CaloJet the software solves the non-linear equation

$$\frac{\text{CaloJet } E_T}{\text{GenJet } E_T} = \text{Jet Response(GenJet } E_T) \tag{6}$$

using a simple iteration procedure.

The jet correction as a function of CaloJet E_T is obtained for the bins of CaloJet $|\eta|$ given above. This means that for a fixed CaloJet E_T the correction is held constant within the bin of CaloJet $|\eta|$. The correction changes as a function of $|\eta|$ in discrete jumps, as one moves from one bin of CaloJet $|\eta|$ to the next. We plan to replace this in the future with a correction that varies smoothly as a function of CaloJet η.

In Fig. 5 we show the jet correction as a function of CaloJet E_T in three bins of CaloJet $|\eta|$. Inversely to the behavior of the jet response, the jet correction decreases with increasing jet E_T, asymptotically approaching 1 in the barrel and the endcap at high CaloJet E_T.

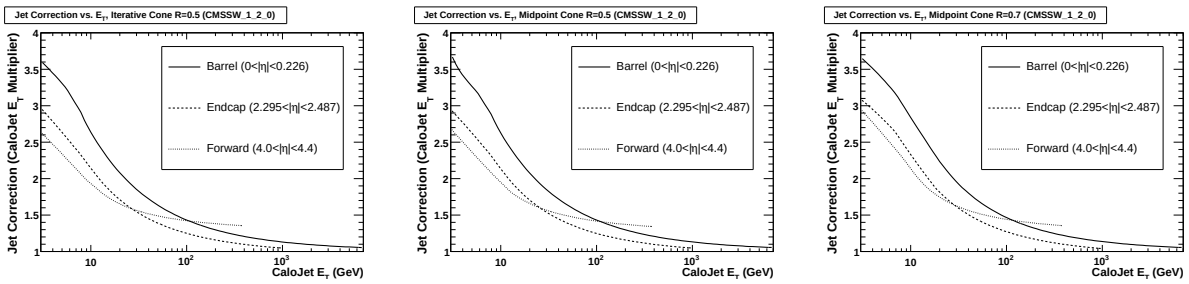


Figure 5: The jet correction as a function of CaloJet E_T for the iterative cone algorithm with a cone size of $R = 0.5$ (left plot), and the midpoint cone algorithm with a cone size of $R = 0.5$ (middle plot) and $R = 0.7$ (right plot).

4.3 Corrected Jet Response

As a closure test, the jet correction procedure described above was used in the MC Jet package to produce collections of corrected CaloJets. The corrected jet collections were then input to the same SimJetResponseAnalysis module used to measure the raw jet response. The corrected jet response

$$\text{Corrected Jet Response} = \frac{\text{Corrected CaloJet } E_T}{\text{GenJet } E_T} \tag{7}$$

was histogrammed in the same bins of GenJet E_T and CaloJet $|\eta|$.

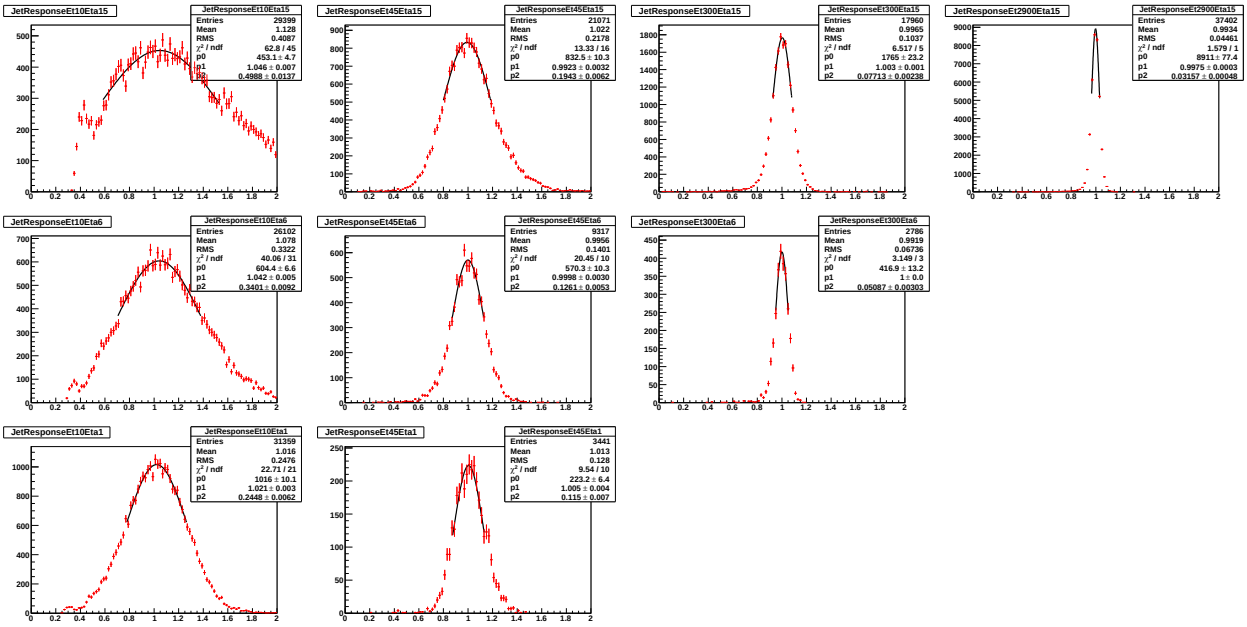


Figure 6: Corrected jet response and fit for the iterative cone algorithm with cone size $R = 0.5$. The number of jets as a function of corrected jet response (points) is compared to a Gaussian fit (curve) in the interval $\pm 1\sigma$ from the peak response. The three rows of plots, from top to bottom, are for the following regions of CaloJet η: $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$. The four columns of plots, from left to right, are for the following regions of GenJet E_T: $10 < E_T < 12$, $45 < E_T < 57$, $300 < E_T < 400$, and $2900 < E_T < 4500$ GeV.

Example histograms of corrected jet response are shown in Fig. 6. The corrected jet response peaks near 1, particularly at high GenJet E_T, indicating that the correction is working. At very low values of GenJet E_T, typically less than 30 GeV, the jet resolution is very poor, and the correction made to the jet varies significantly over the width of the jet response within the histogram. As a result the shape of the corrected jet response at low GenJet E_T in Fig. 6 is slightly different from the shape of the uncorrected jet response in Fig 2. Consequently, for these very low energy jets, the correction does not work as well. To quantify how well the correction is working,

we have determined the peak of the corrected jet response by fitting each histogram with Gaussians in the interval $\pm 1\sigma$ from the peak, as shown in the examples of Fig. 6.

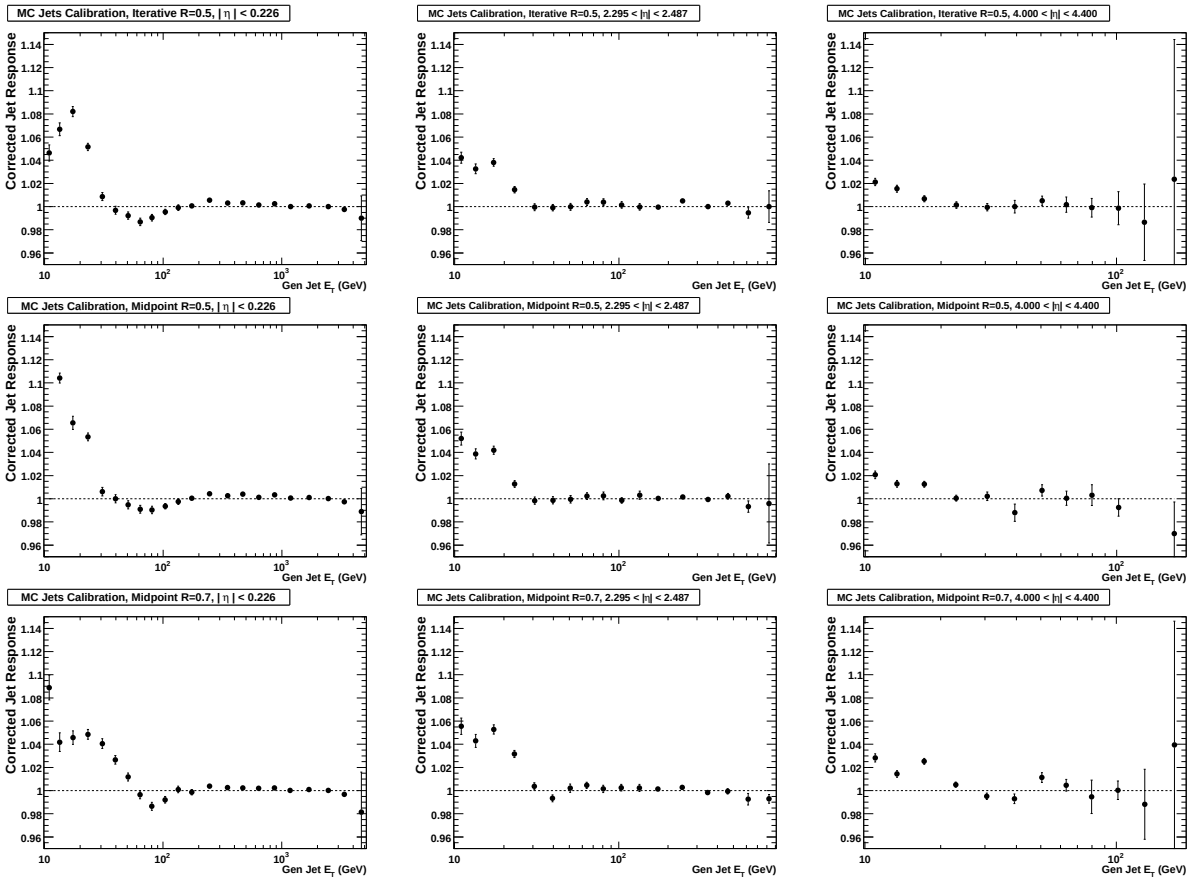


Figure 7: The corrected jet response as a function of GenJet E_T (points). The three rows of plots, from top to bottom, are for the iterative cone algorithm with a cone size $R = 0.5$, and for the midpoint cone algorithm with a cone size of $R = 0.5$ and $R = 0.7$. The three columns of plot, from left to right, are for the following regions of CaloJet η: $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$.

Example plots of the corrected jet response as a function of average GenJet E_T are shown in Fig. 7. The correction is working to within roughly 1% accuracy for GenJet $E_T > 30$ GeV. The accuracy is significantly worse at lower E_T, particularly in the barrel where the jet resolution is worse than in the endcap or forward regions for fixed E_T. At high E_T where the jet resolution is pretty good, and the correction does not vary much within a single histogram of jet response, the corrected jet response peaks at 1 to within roughly 1%. However, at low jet E_T, where the jet resolution is poor, the corrected jet response peaks around 2% to 10% above the expected value of 1. We note that GenJets with E_T less than 30 GeV, corresponding to CaloJets with uncorrected E_T less than 10 GeV in the barrel, are very difficult to understand.

An example of the jet response as a function of η before and after jet corrections is shown in Fig. 8. Before jet corrections the plot shows the response variations of the CMS detector simulation as a function of η. After the jet corrections are applied, the response is reasonably flat around 1. The vertical dotted lines show the edges of the $|\eta|$ bins in Table 3, while the points are the jet response measured in bins equal to the CaloTower η segmentation. Comparing the response variations after corrections as a function of η with the coarse binning of the jet correction into 16 bins of $|\eta|$, we get an idea as to the maximum level of improvement we can expect from a smoothly varying jet correction as a function of η over our current binned correction as a function of $|\eta|$.

For completeness we note that Fig. 8 was made in bins of GenJet P_T (not E_T) and in this figure we are plotting the mean of the jet response distribution, not the peak of a $\pm 1\sigma$ Gaussian fit. We don't expect this to change any

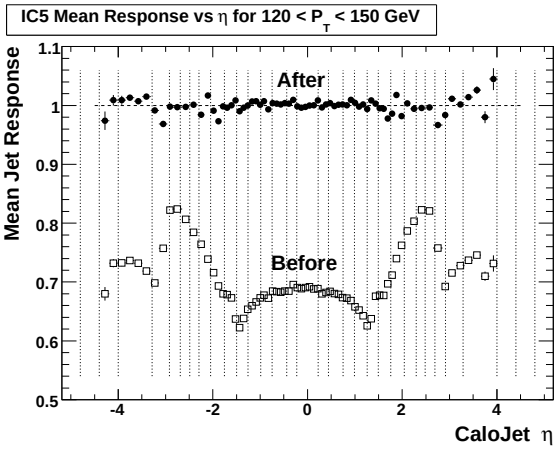


Figure 8: The jet response as a function of CaloJet η before jet corrections (boxes) and after jet corrections (closed circles), for GenJet P_T in the range $120 < P_T < 150$ GeV for the Iterative cone algorithm with cone size $R = 0.5$.

of our conclusions. Also note that the distribution is asymmetric in η for $|\eta| > 3$ due to a known problem with the η values returned by the HF simulation; this was fixed after release CMSSW_1_2_0.

5 Pileup Considerations for Jet Reconstruction

At the design luminosity (10^{34} cm^{-2}s^{-1}), the LHC is expected to deliver on average about 20 proton-proton interactions per beam crossing. Each event will receive contributions both from particles produced in the trigger crossing (in-time pileup) and from particles produced in the adjacent crossings (out-of-time pileup). Pileup (PU) of particles from different interactions will produce energy clusters in the calorimeter which will affect jet reconstruction. Below, we demonstrate the effect of in-time and full (in-time + out-of-time) pileup on jet pseudorapidity distributions.

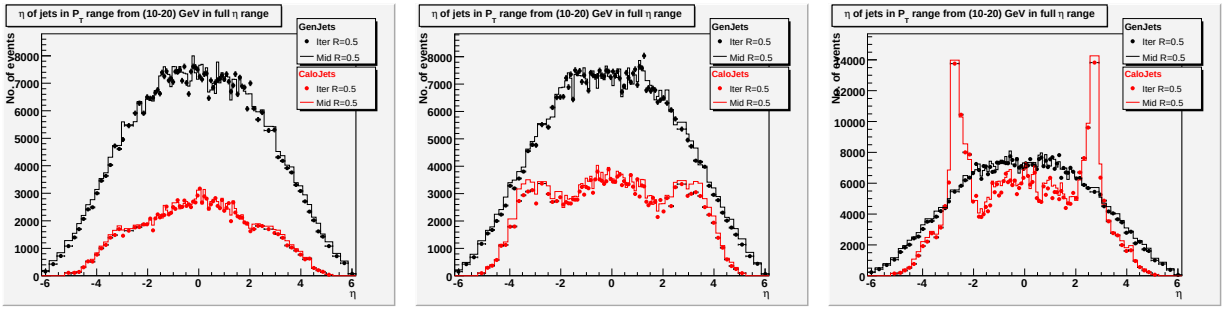
Since PU events contribute mostly low energy particles, one of the most striking effects of pileup is on the total E_T observed in the calorimeter; for the $50 < \hat{p}_T < 80$ GeV samples, the average total E_T is ≈ 180 GeV for the no-PU case, 290 GeV for in-time PU and 410 GeV for full PU. Consequently, the impact of pileup on jet reconstruction is much larger at low jet E_T than at high E_T (the effect on MET reconstruction is discussed in detail in Section 9).

The variation of the pseudorapidity distribution of the calo jets with uncorrected $10 < p_T < 20$ GeV in the QCD sample with the signal event $\hat{p}_T$ in the range $50 < \hat{p}_T < 80$ GeV for no pileup, in-time pileup and full pileup is illustrated in Fig. 9. The low number of calo jets being reconstructed in the calorimeter, as compared to the generated jet distributions, is due to the low average calorimeter response in this range of jet p_T. The forward calorimeter doesn't suffer from out-of-time pileup due to its very short response time. This can be seen in the Fig. 9(b) and Fig. 9(c). The endcap towers near the HE/HF boundary have the largest transverse size in the calorimeter and tend to accumulate significant energy flow from pileup events. It is due to this effect that the "horns" are visible in the full pileup sample (Fig. 9(c)) around $\sim |\eta| = 3$.

Figure 10 shows the η distribution for jets with $40 < p_T < 60$ GeV in the $50 < \hat{p}_T < 80$ GeV samples. It is evident from the Figs. 9(c) and 10(c) that the impact of pileup on η distributions of calo jets is reduced at higher p_T.

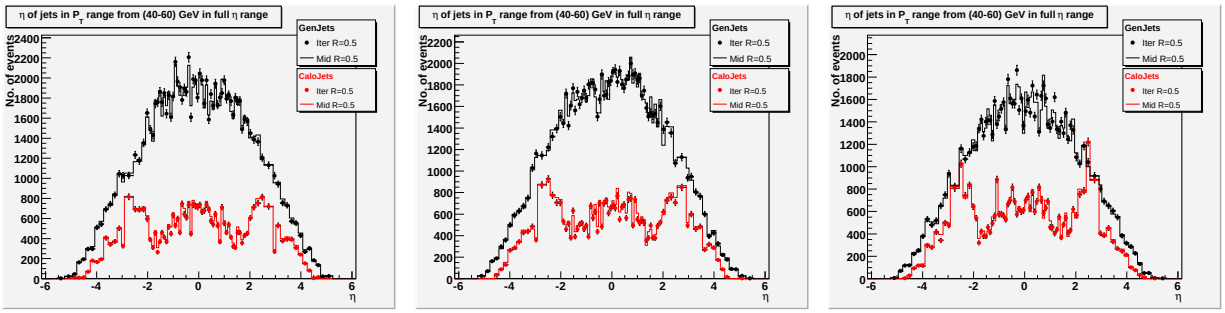
6 Jet Response and Resolutions

The goal of the following jet response study has been to provide plots in the format corresponding to previous results from ORCA; these plots were a part of the SDPV effort. Clearly, there is a significant overlap with the results and methods presented in Section 4. Below, we also include results for various pileup conditions using the same analysis code and plot formats.



(a) No PU (b) "In-time" PU (c) "Full" PU

Figure 9: The pseudorapidity(η) distribution of IC5 and MC5 jets with $10 < p_T < 20$ GeV in $50 < \hat{p}_T < 80$ GeV samples for different cases of pileup, for calo and gen jets.



(a) No PU (b) "In-time" PU (c) "Full" PU

Figure 10: The pseudorapidity distribution of IC5 and MC5 jets with $40 < p_T < 60$ GeV in $50 < \hat{p}_T < 80$ GeV samples for different cases of pileup, for calo and gen.

6.1 Analysis

For the study described in this section, the matching between calo and gen jets had to satisfy the restriction that $\Delta R_{min} < 0.25$, or else the candidate match was discarded. In addition, a minimum generated transverse energy cut of 10 GeV was imposed. In the case that a reconstructed jet was successfully matched with a generated jet, the jet response, $R_{\text{jet}} = E_T^{REC}/E_T^{GEN}$, was calculated.

Three generated pseudorapidity bins were defined, each meant to highlight a distinct region of the detector. The first, designated as a sufficient representation of the barrel region, includes jets with pseudorapidity between $0 < |\eta_{GEN}| < 1.4$. For the endcap region, the pseudorapidity range of $1.4 < |\eta_{GEN}| < 3.0$ is used, and for the forward region, the pseudorapidity range $3.0 < |\eta_{GEN}| < 5.0$ is used. In the plots of the jet response as a function of E_T^{GEN}, the horizontal axis was divided into E_T^{GEN} bins with bounds given by the equation $10^{1+n/8}$ from $n = 1$ to $n = 22$.

6.2 Jet Response

For the results discussed in this section, the histograms of the response were fit with a Gaussian function in the interval $\pm 1.5\sigma$ centered about the mean. The mean parameters of the Gaussian fits were extracted and used as data points to determine the jet response curve as a function of E_T^{GEN}. The results are shown in Figure 11 for the IC5 algorithm for samples with no pileup and in-time pileup. The effect of in-time pileup is observed only at the low

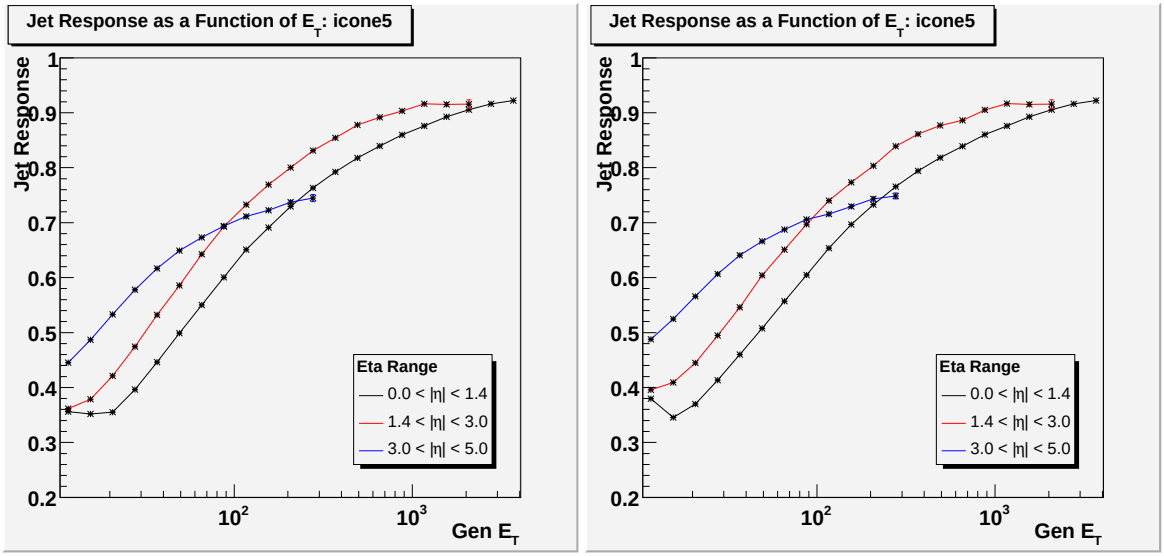


Figure 11: The jet response as a function of generated transverse energy, as reconstructed by the iterative cone algorithm with $R = 0.5$, for three detector regions and samples with no pileup (left) and for the case of in-time pileup (right).

end of the E_T range.

Again, these curves illustrate how the jet response increases with increasing transverse energy. The jet response in the barrel and the endcap levels off between 90% and 95% for jets with generated transverse energy of over 1 TeV. The response in the forward region saturates much sooner, for jets with transverse energy of about 200 GeV at about 75%, as discussed before.

A similar approach was taken when plotting the jet response as a function of generated η. The range of E_T^{GEN} was partitioned into 12 bins, covering a span of 90 GeV $< E_T^{GEN} <$ 4500+ GeV, and plotted individually. Figure 12 shows the jet response curve as a function of $|\eta_{GEN}|$ for cases with no pileup and in-time pileup. No significant impact of pileup is observed for jet $E_T > 90$ GeV. The dip of the jet response curves at $|\eta| = 1.4$ and $|\eta| = 3.0$ corresponds to the transition region of the detector.

The jet response study was repeated using the MC5 and MC7 reconstruction algorithms, as illustrated in Fig. 13.

The corrected jet response was plotted in order to assess the effectiveness of the correction factors (a uniform average value of 1 is expected for all matched jets, if the jet corrections are working properly). The results are shown for the IC5 algorithm in Fig. 14 as a function of E_T^{GEN}. The procedure was also repeated for the MC5 and MC7 algorithms (not shown) with similar results.

As discussed in Section 4, the corrected jet response curves are generally within 1-2% of unity. The exception are very low E_T jets with transverse energy less than 30 GeV which show response deviations of about 4-10% from unity.

6.3 Jet Resolutions

The jet resolutions have been calculated for jet energies as well as for jet angular positions (η and ϕ). The results were separated according to detector region and plotted as a function of E_T^{GEN}. The plots were then fitted with a function of the following form:

$$\text{Resolution} = \sqrt{\left(\frac{C_1}{E_T^{GEN}}\right)^2 + \left(\frac{C_2}{\sqrt{E_T^{GEN}}}\right)^2 + (C_3)^2} \tag{8}$$

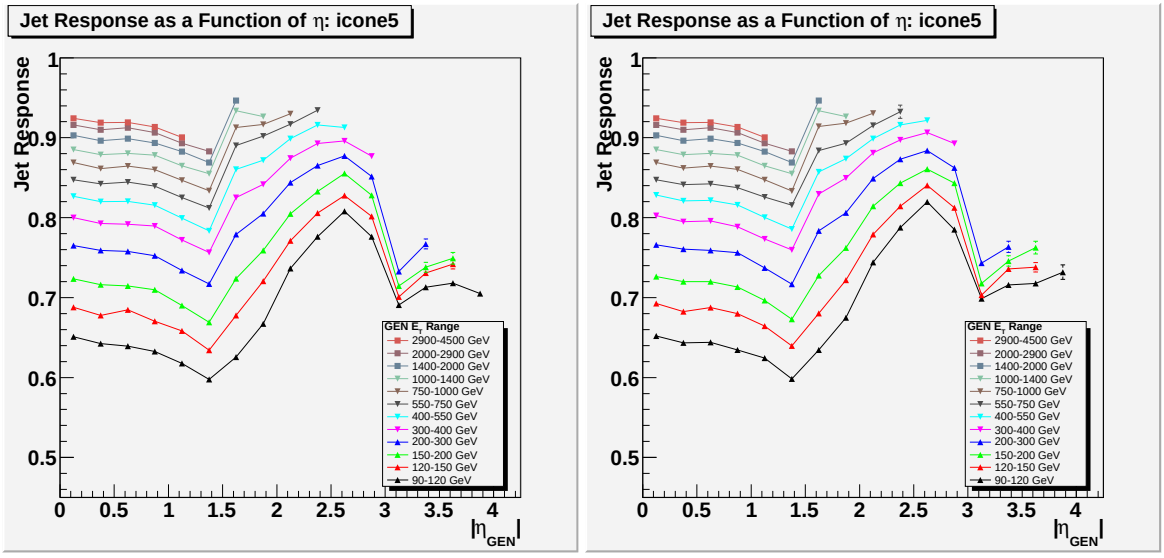


Figure 12: The jet response as a function of generated pseudorapidity, as reconstructed by the iterative cone algorithm with $R = 0.5$, in generated transverse energy bins for samples with no pileup (left) and for the case of in-time pileup (right).

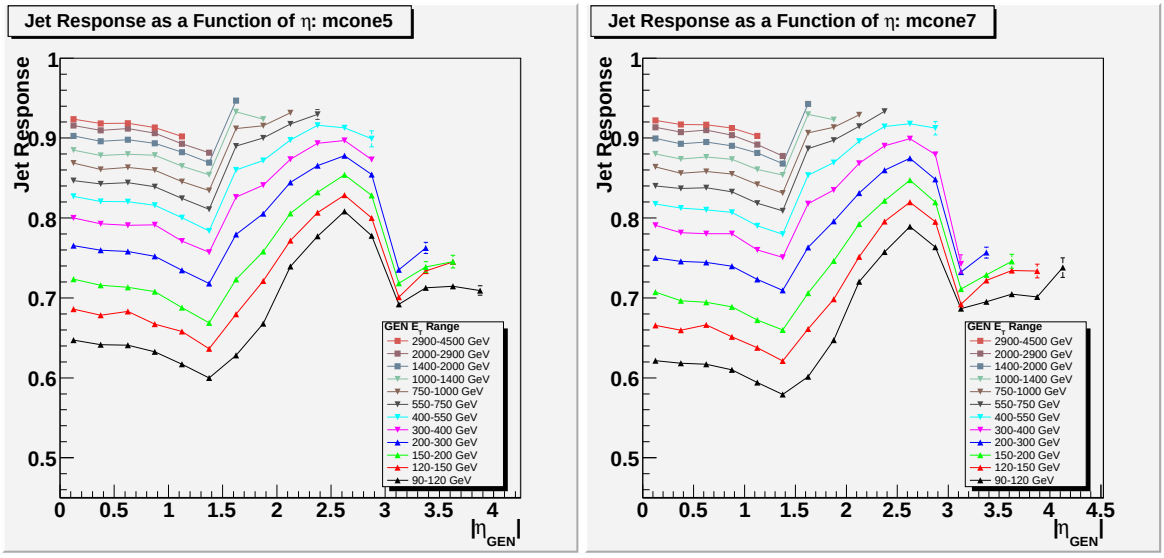


Figure 13: The jet response as a function of generated pseudorapidity, as reconstructed by the MC5 (left) and MC7 (right) algorithms with $R = 0.5$, in generated transverse energy bins for samples with no pileup.

where C_1, C_2, and C_3 are the fit parameters. For validation purposes, the plots were compared to those obtained with ORCA.

6.3.1 Resolution of the Jet Response

Using the histograms of the jet response, the fractional jet E_T resolution

$$\frac{\sigma\left(R_{jet}\right)}{\langle R_{jet}\rangle} = \frac{\sigma\left(\frac{E_T^{REC}}{E_T^{GEN}}\right)}{\left\langle \frac{E_T^{REC}}{E_T^{GEN}} \right\rangle} \tag{9}$$

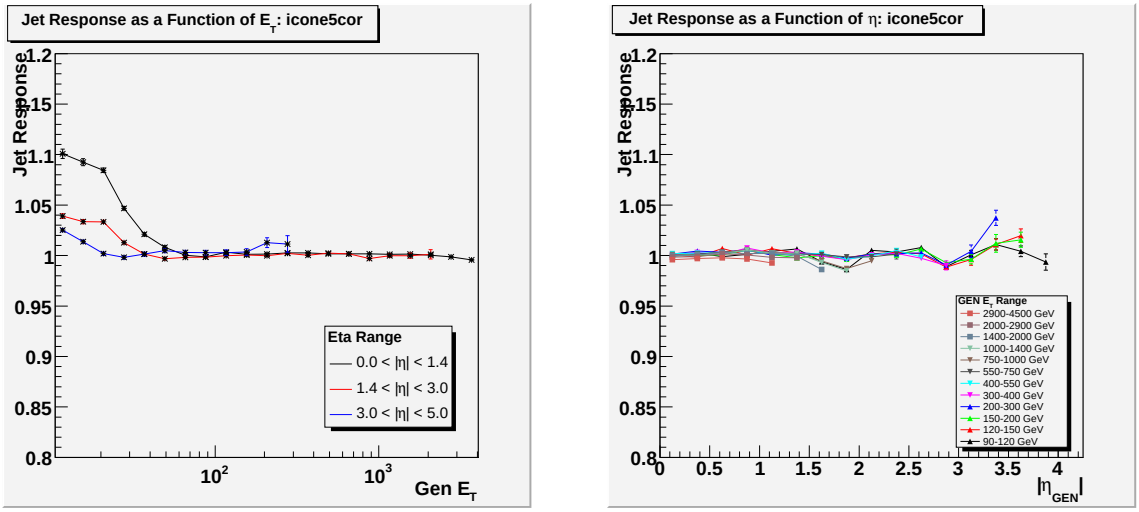


Figure 14: The jet response curves after applying the jet corrections when reconstructed with the IC5 algorithm using no-pileup samples. (left) Jet response shown as a function of E_T^{GEN} for the different regions of the detector. (right) Jet response shown as a function of $|\eta_{GEN}|$ for different generated transverse energy ranges.

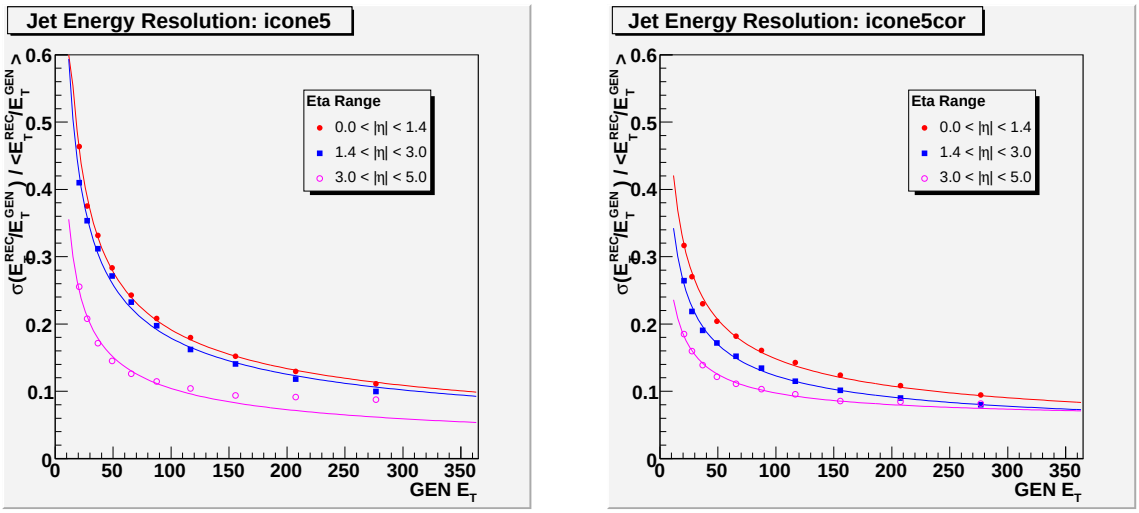


Figure 15: The jet resolution curves (no pileup) as functions of E_T^{GEN}, divided by detector region, when reconstructed with the IC5 algorithm. (left) Resolution prior to the application of the jet corrections. (right) Resolution after the jet corrections have been applied.

was calculated as a function of E_T^{GEN}, using Gaussian fits (for these fits, the full range was fitted with a Gaussian). For the IC5 algorithm, the resolution curves before and after the application of the jet energy corrections for the three detector regions are displayed in Fig. 15 for the no-pileup case. Direct comparisons between uncorrected and corrected resolutions in the three regions of the detector for the IC5 algorithm are shown in Fig. 16. As expected, the jets exhibit an improved resolution when the jet corrections are applied.

The resolutions for the MC5 and MC7 jets after respective energy corrections are shown in Fig. 17.

In Figure 18 we compare IC5 jet resolutions in the three detector regions derived from samples with in-time and full pileup. We conclude that pileup effects on the presented resolutions are small over the E_T^{GEN} range shown.

The resolution plots were compared to ORCA as part of the validation study. The comparisons of both corrected and uncorrected jets are shown in Fig. 19. Overall, the resolution curves of CMSSW match relatively well to those

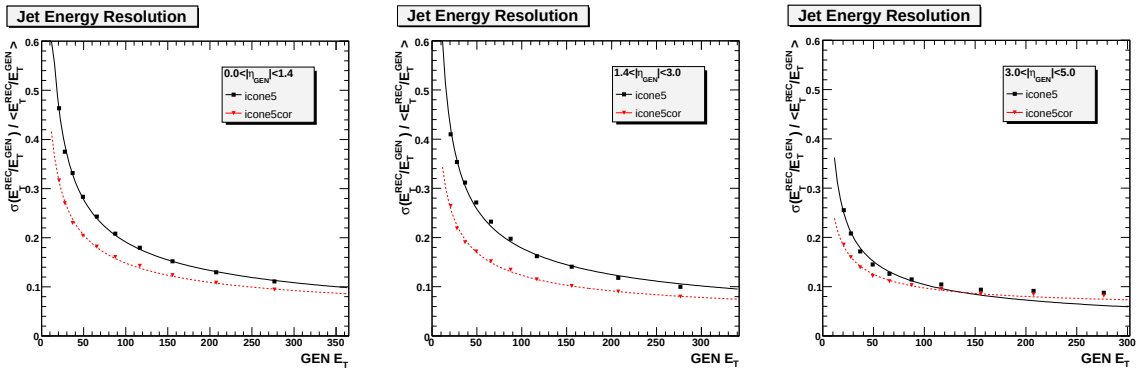


Figure 16: The comparisons of the IC5 jet resolutions before and after jet energy correction (for no pileup) as functions of E_T^{GEN} for the three detector regions.

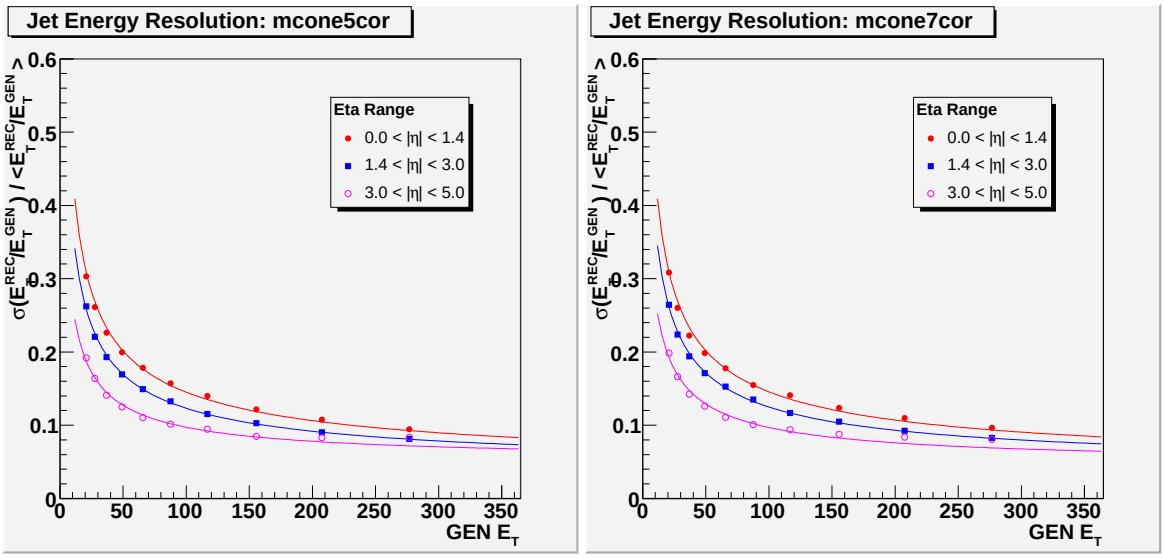


Figure 17: The jet energy resolution as a function of generated transverse energy, when reconstructed with the MC5 (left) and MC7 (right)algorithms in CMSSW_1_2_0, for each of three detector regions, after respective energy corrections and for no-pileup conditions.

of ORCA.

6.3.2 Resolutions of Jet positions in ϕ and η

The resolutions of jet positions in ϕ and η were calculated from distributions of the following variables for the matched jets:

$$\Delta\phi = \phi_{REC} - \phi_{GEN}, \tag{10}$$

$$\Delta\eta = |\eta_{REC}| - |\eta_{GEN}|, \tag{11}$$

and taking care of the $\pm\pi$ wrap-around in ϕ. We use the absolute values of generated and reconstructed jet η to eliminate a reconstruction bias that pulls the calorimeter jet towards $\eta = 0$ relative to the generated jet[12]. This effect visibly affects the results in HF. Figure 20 shows the resolutions in ϕ and η. Both are displayed as a function of E_T^{GEN} for the IC5 algorithm.

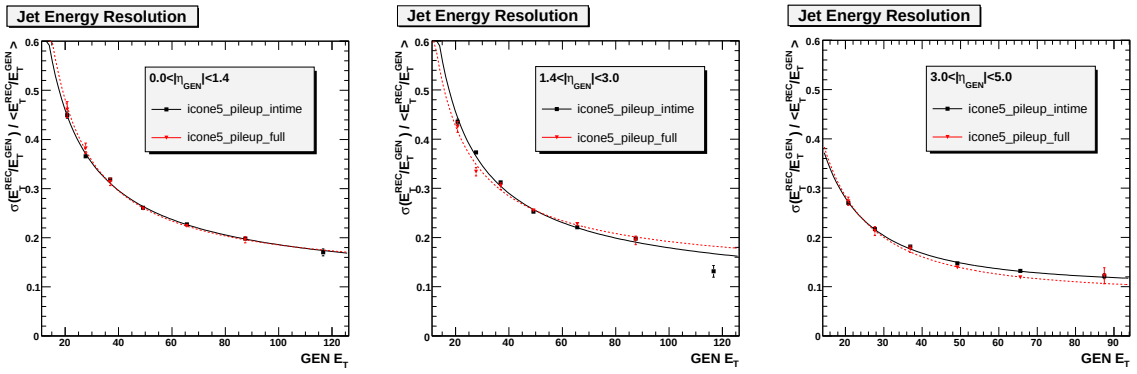


Figure 18: Comparison of IC5 jet resolutions for in-time (black squares) and full (red triangles) pileup conditions in three regions of the detector. No energy corrections have been applied.

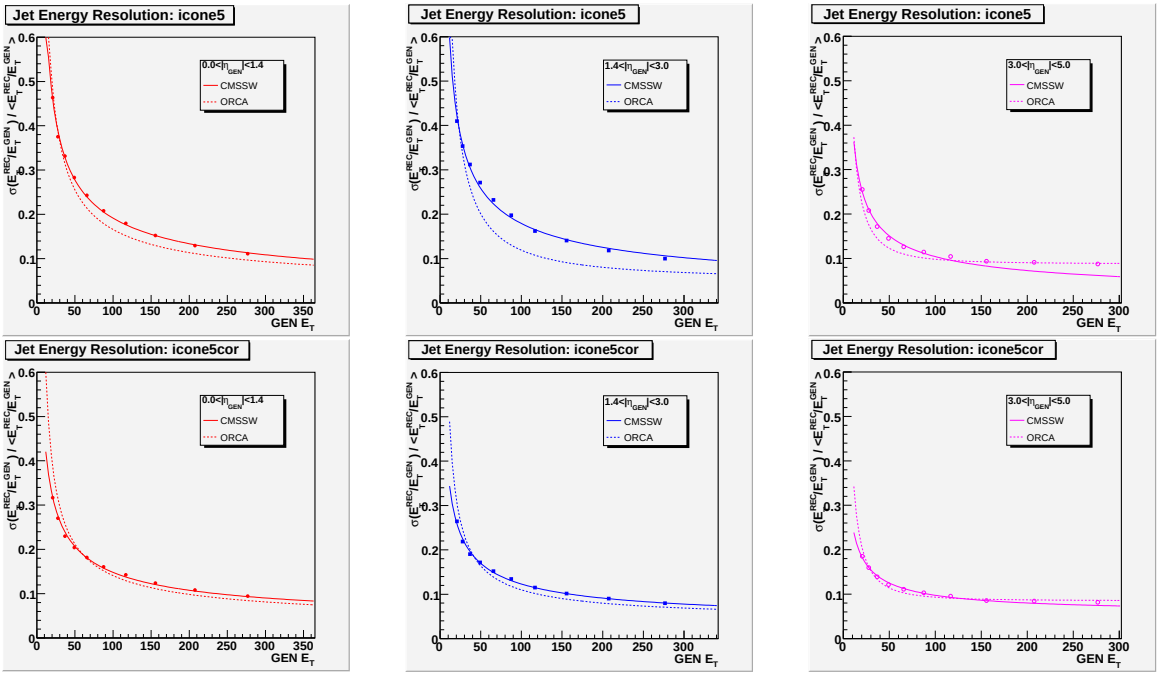


Figure 19: The jet resolution curves of CMSSW (for no pileup) compared to ORCA, separated by detector region. Both the uncorrected (top row) and corrected (bottom row) resolutions are shown.

The jet position resolutions in HF (for both $\Delta\phi$ and $\Delta\eta$) in CMSSW_1_2_0 are by a factor of 2 worse than previously determined in ORCA. This has been traced back to the Shower Library in HF and improvements have been implemented in CMSSW_1_2_2 and higher versions. For the validation exercise, we used a special small-scale private simulation based on CMSSW_1_2_2 to test the effect of the fix. These results are presented in Fig. 21. A large statistics result from the Spring07 samples using CMSSW_1_2_3 simulation and 1_3_1 reconstruction has become available recently and is presented in Fig. 22.

7 Jet Efficiencies

Jet finding efficiencies were determined for the iterative and midpoint cone jet algorithms for cone sizes of $R = 0.5$ and $R = 0.7$. The IC7 reconstruction was added with the same parameters as the standard IC5 reconstruction (except for using $R = 0.7$). Jets reconstructed using the k_T algorithm were omitted from this study because of implementation problems in CMSSW_1_2_0. The jet finding efficiency was defined as the ratio of the number of

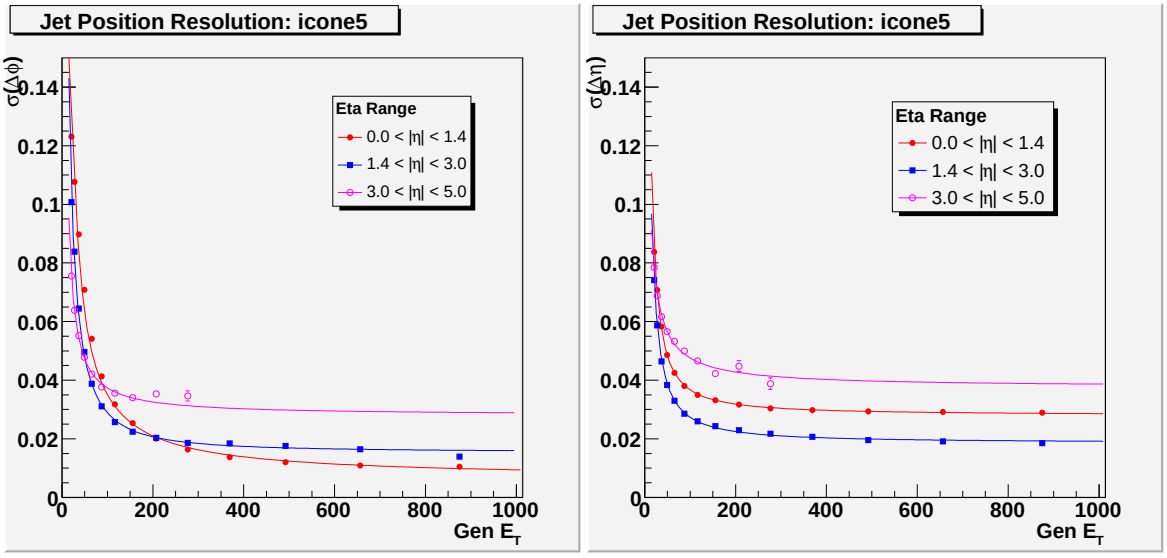


Figure 20: The ϕ (left) and η (right) resolutions as a function of generated transverse energy, when reconstructed with the IC5 algorithm in CMSSW_1_2_0, for each of three detector regions.

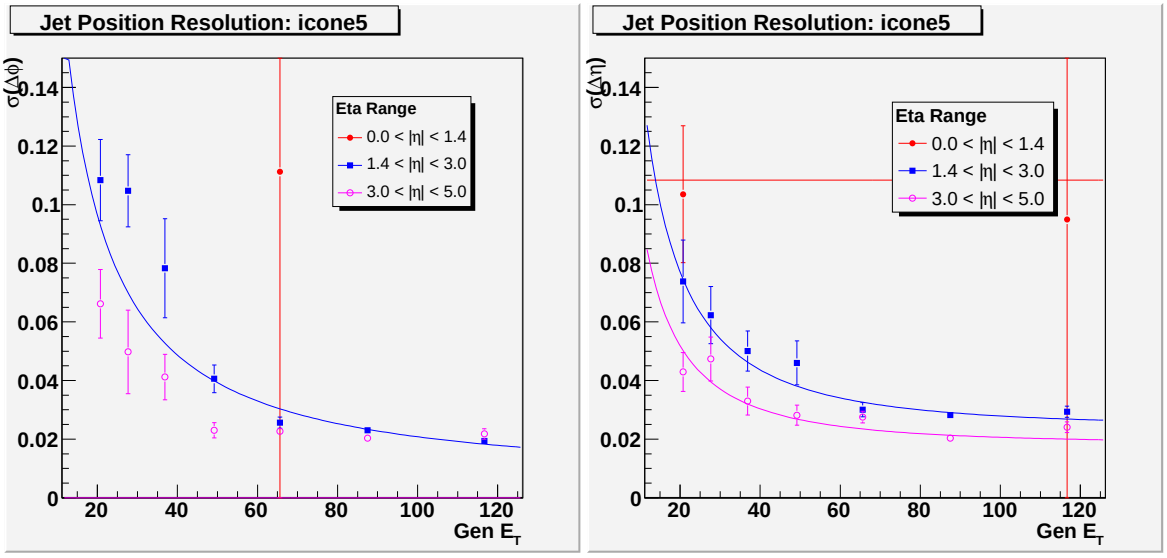


Figure 21: Results of a small scale test of the HF Shower Library fix in CMSSW_1_2_2. The ϕ (left) and η (right)resolutions as a function of generated transverse energy, when simulated with CMSSW_1_2_2 and reconstructed with the IC5 algorithm.

generated jets that matched a reconstructed jet to the total number of generated jets. The same jet definition is used for both reconstructed jets and generated jets. A generated jet was considered to be matched if the nearest reconstructed jet was within a distance $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.3$. Distributions in ΔR for the four jet definitions are shown in Fig. 23 for two $\hat{p}_T$ bins.

Figure 24 shows the jet finding efficiencies as a function of generated jet p_T for $|\eta| \leq 1.5$. The jet finding efficiency is largest for jets reconstructed using the iterative cone algorithm with cone size $R = 0.5$. There is a relatively large difference in the jet finding efficiencies for the different jet definitions. However, one should note that when comparing the jet finding efficiencies as defined above, the energies assigned to the generated jets can be different for each jet definition. The effects of the different generated jet energies can be taken out by instead using a single

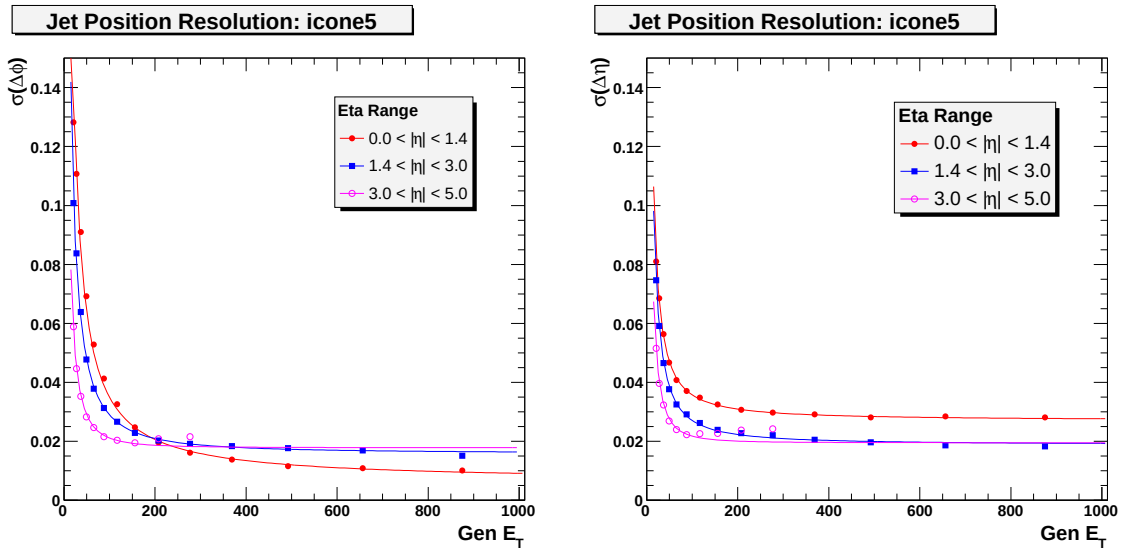


Figure 22: The ϕ (left) and η (right) resolutions as a function of generated transverse energy, when reconstructed with the IC5 algorithm in CMSSW_1_3_1, for each of three detector regions.

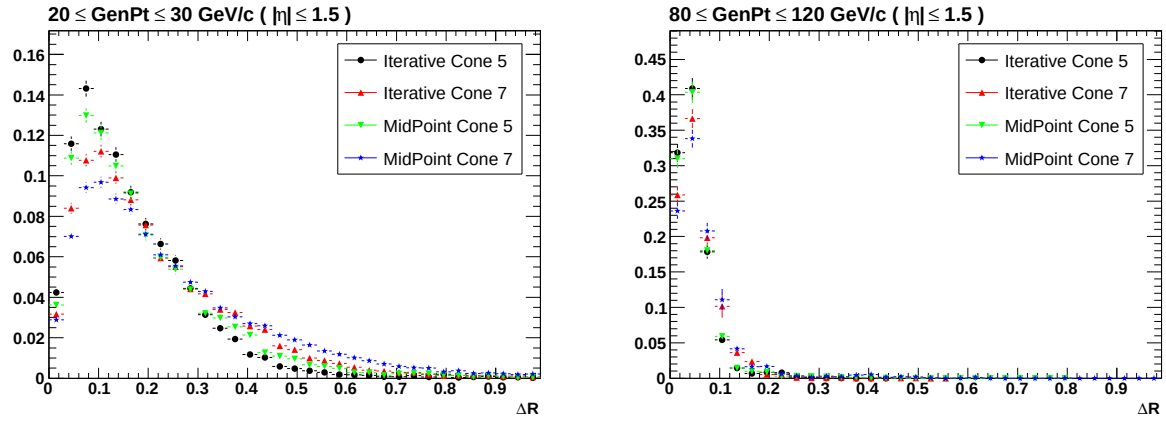


Figure 23: ΔR distributions for the jet p_T bin of $20 - 30$ GeV (left) and $80 - 120$ GeV (right).

jet definition for the generated jets, which provides a common p_T scale for the comparisons, and then finding the number of matched reconstructed jets using each of the different jet definitions. Defined in this manner, the differences in the jet matching efficiencies are much less pronounced, as shown in Fig. 25. We emphasize that the quantity shown in Fig. 25 does not represent the jet finding efficiency but is only meant to help understand the behavior of the efficiency for different algorithms.

The impact of a minimal p_T requirement on reconstructed jets on the jet finding efficiency is shown in Fig. 26. For the chosen minimal p_T requirements, the effects are small for generated p_T values above 40 GeV. For the minimal p_T requirement of 6 GeV, the jet finding efficiency is ≈50% at generated p_T = 20 GeV, while for no p_T requirement, the jet finding efficiency is ≈50% at generated p_T = 10 GeV.

The effects of pileup on the jet finding efficiency were also investigated. In Fig. 27, the jet finding efficiency for iterative cone jets with $R = 0.5$ is shown for no pileup, in-time pileup only, and full pileup scenarios. The inclusion of extra pileup interactions increases the jet finding efficiency at low p_T since these interactions occasionally add enough energy to the calorimeter towers to allow them to pass the jet threshold requirements. (Note that pileup particles are not included in generated jets.)

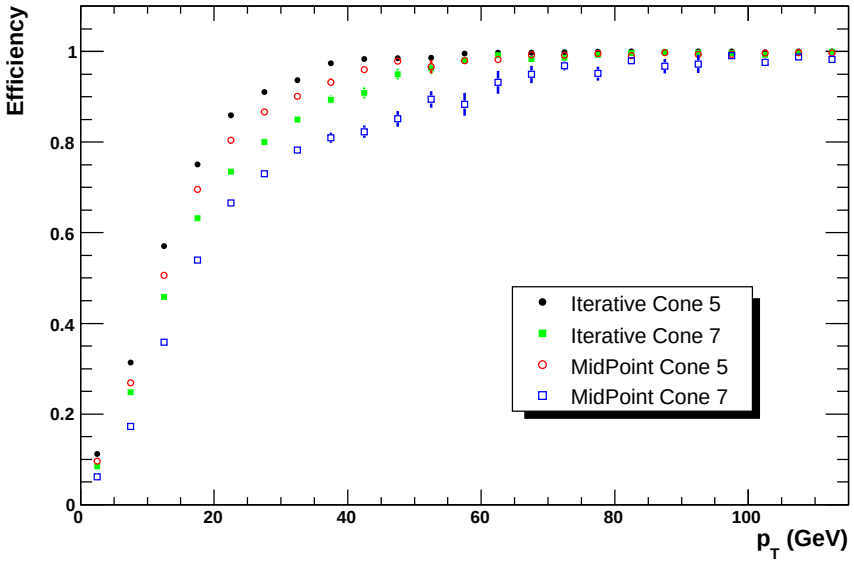


Figure 24: Jet finding efficiency as a function of generated jet p_T for the central calorimeter ($|\eta| \leq 1.5$) and a matching cone of $\Delta R = 0.3$.

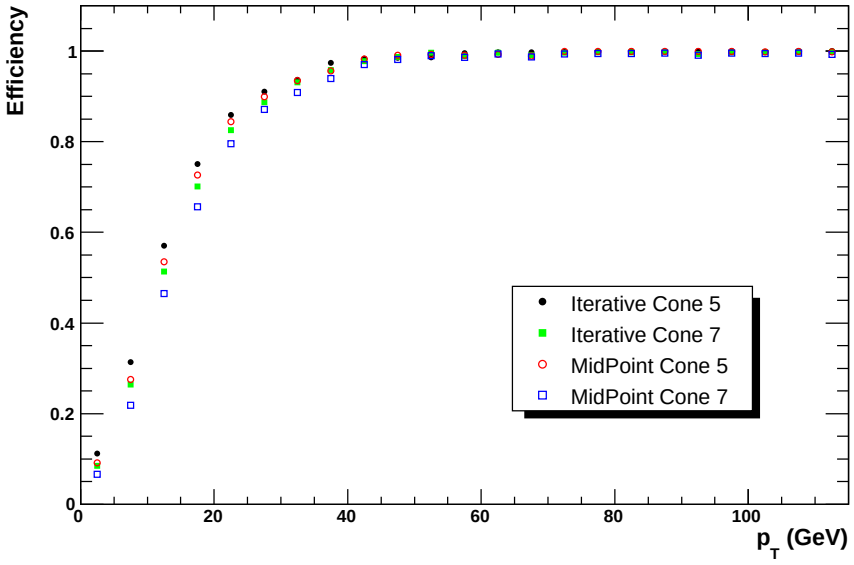


Figure 25: Efficiency of matching calorimeter jets for different jet definitions to generated iterative cone jets, for $R = 0.5$.

8 Dijet Balance

In $2 \rightarrow 2$ events, transverse momenta of two jets are equal and this property can be used to scale a jet p_T at pseudorapidity, η, to a jet p_T in a reference region. This procedure can be used to correct for the variation of the response as a function of η based on data. For the study described below, we define the reference region to be $|\eta| < 1.0$. A similar study was performed using ORCA DC04 QCD samples[14].

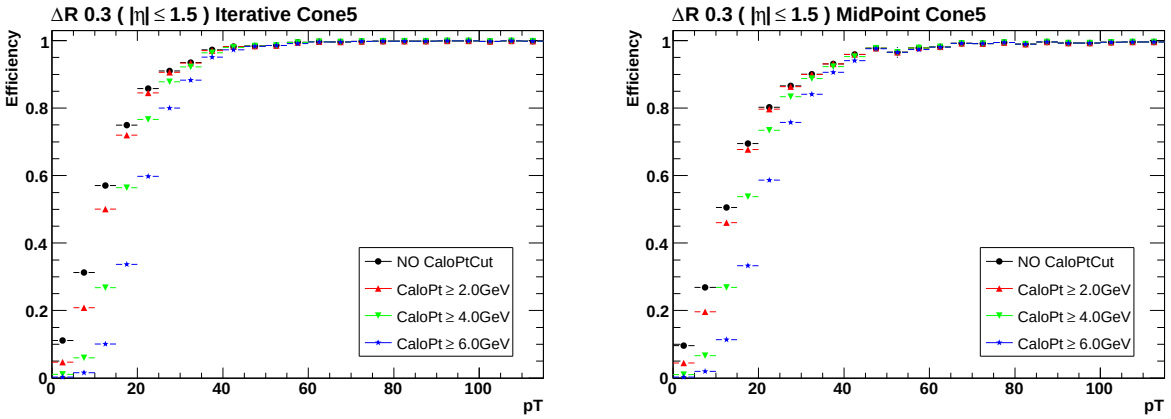


Figure 26: The effects of a minimal reconstructed jet p_T requirement on the jet finding efficiency for IC5 (left) and MC5 (right) jets.

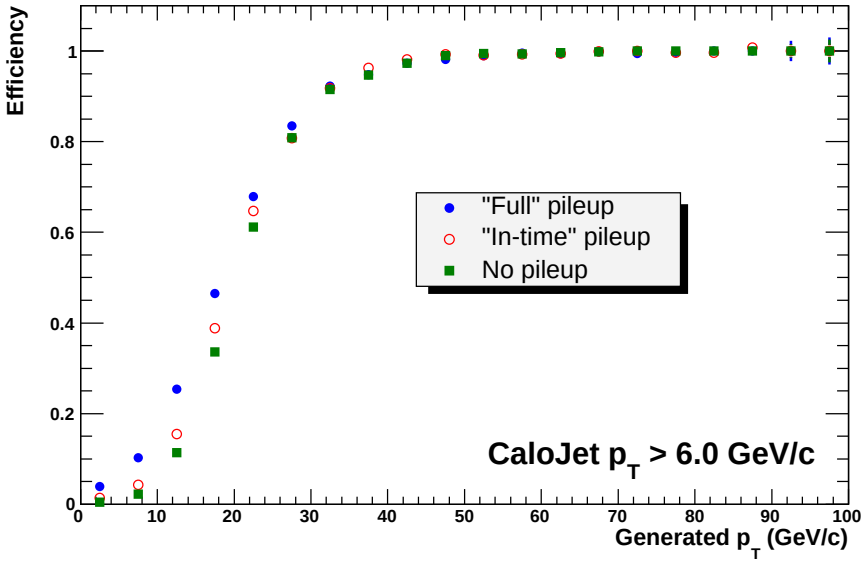


Figure 27: Jet finding efficiency for IC5 jets and various pileup scenarios.

We define the dijet balance as

$$B(\eta) = \frac{p_T^{Tag} - p_T^{Probe}(\eta)}{(p_T^{Tag} + p_T^{Probe}(\eta))/2} \tag{12}$$

where p_T^{Tag} is the transeverse momentum of the jet in the reference region and $p_T^{Probe}(\eta)$ is the transverse momentum of the probe jet at pseudorapidity η. To minimize the effect of jet energy smearing due to finite resolution, the analysis is performed in bins of the average p_T of the two leading jets, $p_T^{ave} = (p_T^{(1)} + p_T^{(2)})/2$.

The relative calorimeter response $R(\eta)$ can then be determined from Eq. 12

$$R(\eta) = \frac{p_T^{Probe}(\eta)}{p_T^{Tag}} = \frac{2- < B(\eta) >}{2+ < B(\eta) >} \tag{13}$$

where $< B(\eta) >$ is the average dijet balance for the probe jets at η. The dijet balance is histogrammed in η bins which have the same size as calorimeter towers. The response, R, is calculated from the mean of the dijet balance

histrogram. The observable $R(\eta)$ differs from r of Ref. [14] by +1. The measured $R(\eta)$ can be interpolated using a cubic spline and then used to correct the probe jet.

8.1 Analysis

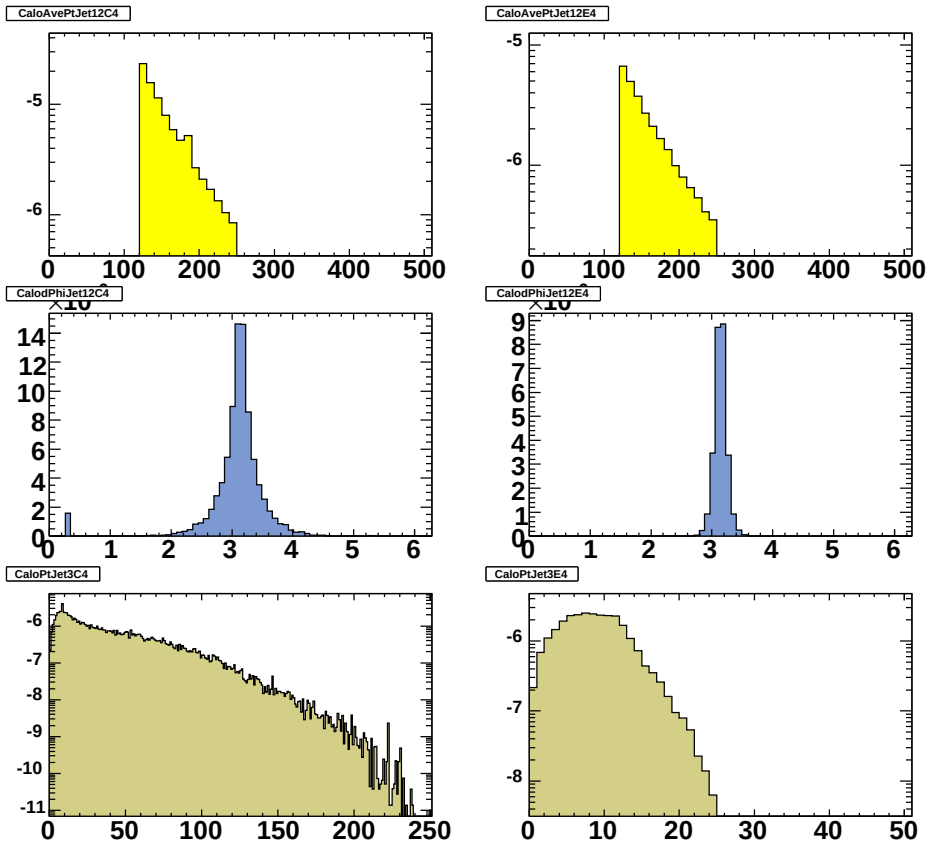


Figure 28: Average p_T, difference in azimuthal angles of the two leading jets and p_T of the third jet before (left) and after (right) applying cuts.

In order to minimize the effect of hard initial and final state radiation (ISR/FSR) which break $p_T^{(1)} = p_T^{(2)}$ condition, the event is required to satisfy following requirements:

$$\phi^{(1)} - \phi^{(2)} < 2.7 \qquad p_T^{(3)} < min(10., 0.1 \times p_T^{ave}) \tag{14}$$

where $\phi^{(1)}$ and $\phi^{(2)}$ are the azimuthal angles of the two leading jets and $p_T^{(3)}$ is the transverse momentum of the third leading jet in the event.

The impact of the radiation cuts is shown in Fig. 29 for the dijet balance vs. η calculated at the particle level without and with the conditions given by Eq. 14. The Figure demonstrates that the concept of dijet balance can be expected to result in a uniform jet energy response when the radiation cuts are applied.

We use the CMSSW 1_2_0 QCD dataset, generated in 21 $\hat{p}_T$ bins. The data are analyzed in various p_T^{ave} bins, weighting events according to production cross section. Figure 28 shows the average p_T^{ave} of two leading jets, their difference in azimuthal angle ($\Delta\phi$), and the transverse momentum of the third jet before and after the cuts. These cuts to suppress radiation were not used in [14] for the relative response study.

The relative response for $120 \leq p_T^{ave} < 230$ GeV bin using raw calorimeter jets is shown in Fig. 30(left), with and without cuts in Eq. 14. The filled circles show the response without $\Delta\phi$ and third jet cuts and can be compared to Fig. 5 in reference [14]. Let us now discuss results for the different regions of η. The relative response averaged

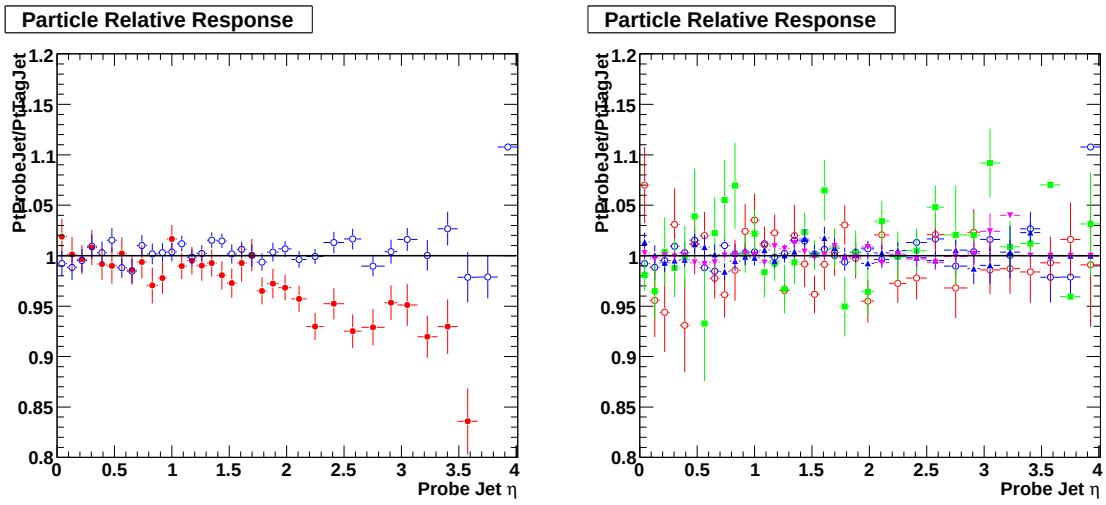


Figure 29: (left) Dijet balancing at the particle level without (filled circles) and with radiation cuts (open circles) for $120 < p_T^{ave} < 250$ GeV; (right) the dijet balancing at the particle level for different p_T^{ave} bins after radiation cuts, Eq. 14.

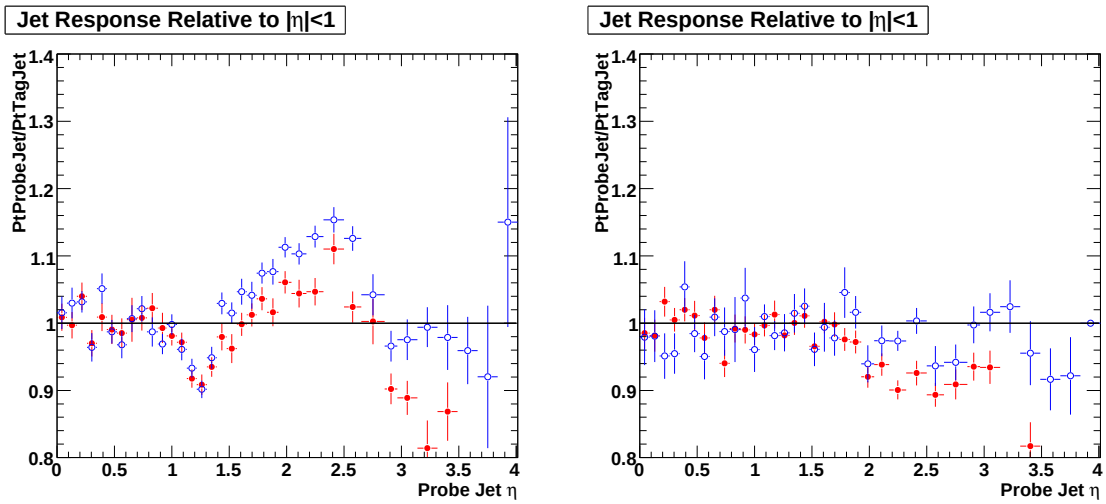


Figure 30: (left) Calorimeter jet response relative to $|\eta| < 1$ region using raw jets before (filled circles) and after (open circles) radiation cuts, Eq. 14. (right) Calorimeter jet response relative to $|\eta| < 1$ region using MCJet-corrected jets before (filled circles) and after (open circles) cuts.

over the reference region, $|\eta| < 1$, should equal to 1.0 by defintion, as is observed. The loss of response at the barrel-endcap boundary is $\sim 10\%$ in both cases. The CMSSW 1.2.0 shows a endcap relative response is $\sim 4\%$ higher in $1.6 < |\eta| < 2.8$ region compared to 5-6% observed in [14]; both results are consistent within the large statistical uncertainties. For $|\eta| > 2.8$, two responses differ due to the difference in forward calorimeter (HF) calibration between the two data sets. The relative response after $\Delta\phi$ and third jet cuts (open circles) in the Endcap calorimeter is higher.

The response for $120 \leq p_T^{ave} < 230$ GeV range, using corrected calorimeter jets is shown in Fig.30(right); the jets have been corrected using 1_2_0 MCJet Corrections. The filled (open) circles show the response before (after) the extra radiation cuts. As expected, the response is close to unity, except possibly for $|\eta| > 3.5$ region.

The calorimeter relative response may vary with the p_T of the jet. The data was divided in different p_T^{ave} bin (30-60, 60-90, 90-120, 120-250, 250-300, 300-500, 500-800, 800-1000, 1000-2000 GeV). The response for first 6

bins is shown in Fig. 31(left) and for last 4 bin in Fig. 31(right). A variation of $\sim 10\%$ is observed over 30-2000 GeV p_T range, though the statistical uncertainties are large. At large p_T, the jets become very narrow and the region where relative response is low, at barrel-endcap boundary, shrinks.

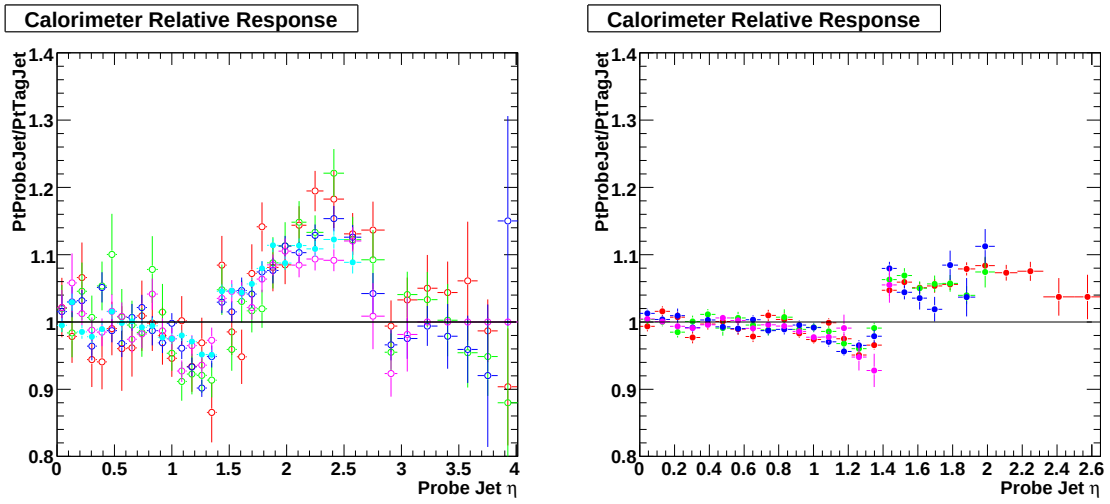


Figure 31: Calorimeter jet response relative to $|\eta| < 1$ region for six p_T^{ave} bins in the 30–500 GeV range (left), and four bins in the 500–2000 GeV range (right); see text for bin values and discussion.

9 MET Performance

The Missing Transverse Energy (MET) is determined in CMSSW_1_2_0 from the tranverse vector sum over uncorrected, projective Calorimeter Towers:

$$\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_n (E_n \sin\theta_n \cos\phi_n \hat{\mathbf{i}} + E_n \sin\theta_n \sin\phi_n \hat{\mathbf{j}}) = -E_x^{\mathrm{miss}} \hat{\mathbf{i}} - E_y^{\mathrm{miss}} \hat{\mathbf{j}} \tag{15}$$

where the index n runs over all calorimeter input towers.

In order to conform to the same conditions as the Physics TDR Vol. 1, or as similar as possible, the default CMSSW "Scheme-B" thresholds were removed, and the Calorimeter Towers were re-reconstructed applying so-called "ORCA PTDR" thresholds, which were simply 0.5 GeV for all HCAL cells.

The data samples used to study the MET performance correspond to QCD dijets produced in eleven bins of $\hat{p}_T$ between 0 GeV and not more than 800 GeV. In-time pile-up corresponding to low luminosity running (2×10^{33} cm^{-2}s^{-1}) was also included by default. The effects of out-of-time (integrated over -3 to $+5$ bunch crossings) pile-up were separately studied.

MET is a simple, global observable which depends on precise symmetric cancelations and hence a good indicator of the health of the underlying input. The global nature of MET also means that disentangling and understanding the different factors affecting the performance can be very challenging and sometimes impossible. As such, this study attempts to quantify the status of CMSSW_1_2_0 with respect to the ORCA PTDR results and to list, as far as known, the possible differences between the conditions used in the ORCA studies and this study based on CMSSW_1_2_0. The performance of MET resolution is parameterised in the total visible transverse energy of the event (ΣE_T) according to the following form:

$$\frac{\sigma(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}})}{\Sigma E_T} = \frac{A}{\sqrt{\Sigma E_T}} \oplus \frac{B}{\Sigma E_T} \oplus C \tag{16}$$

where the A ("stochastic") term represents the statistical sampling nature of the Calorimeter Towers; the B ("noise") term represents effects due to electronic noise, pile-up, underlying-event, etc; and the C ("constant")

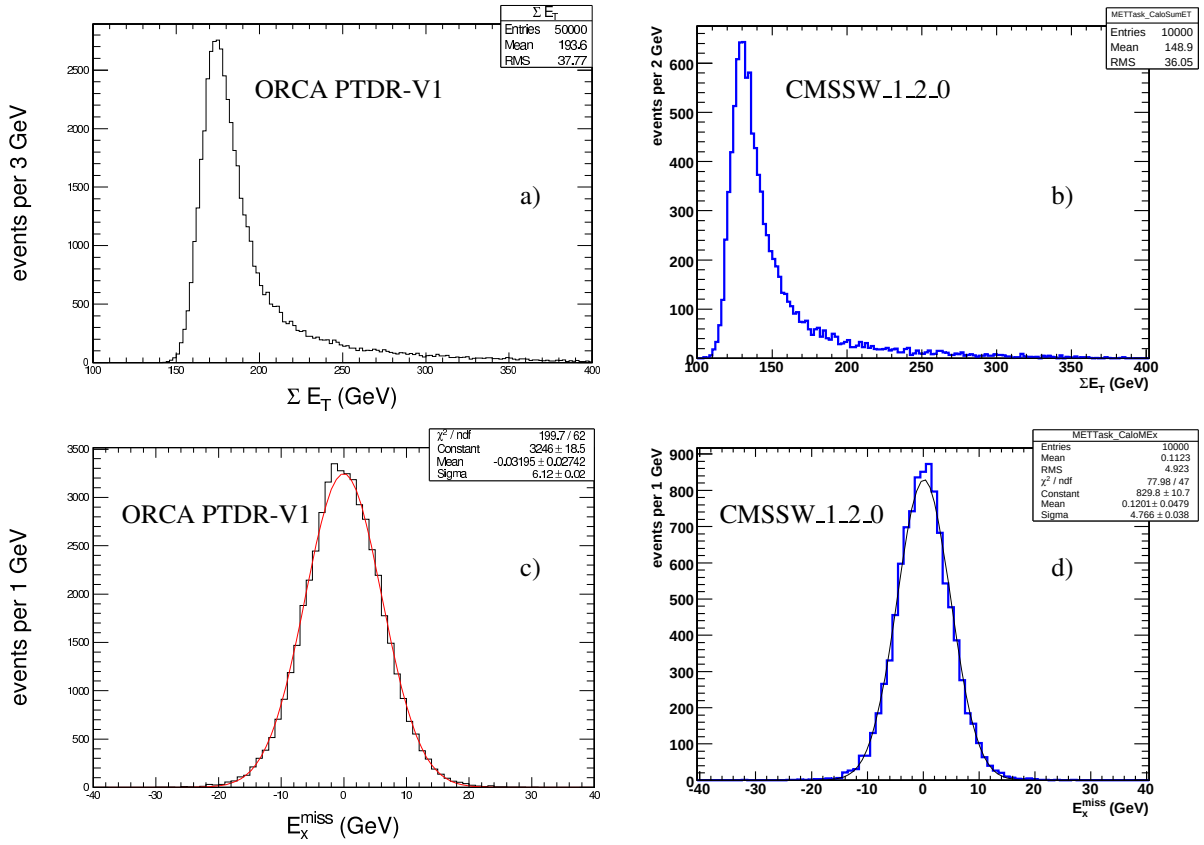


Figure 32: ΣE_T (top) and ME_x with a Gaussian fit (bottom) distributions for Minimum bias events without pile-up. ORCA CaloTower thresholds were applied.

term represents residual systematic effects due to non-linearities, cracks, dead material, etc. It is important to emphasize that the above parameterisation factorizes the MET uncertainty into independent effects A, B, C. In particular, the stochastic and constant terms do not depend on the effects due to noise, pile-up, and underlying event (to first order). Hence, when comparing the MET performance between samples having different noise thresholds and pile-up conditions (as is done in this study), the offset in ΣE_T must be explicitly taken into account.

QCD dijets represent a benchmark process for studying the performance of MET. For example, in the case of QCD dijets, no MET is expected and so any energy imbalance measured in the event is due to the finite calorimetery resolution. In this case, the average MET, measured over many events, is an estimate of the MET uncertainty: $< \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} > \approx \sigma(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}})$ (assuming Gaussian behaviour).

9.1 Missing Transverse Energy in Minimum bias events

Figure 32 compares the results presented in Figure 11.10 of the Physics TDR Vol. I and the corresponding results of this study for the total visible transverse energy ΣE_T and the x-component of the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ (labeled $E_{\mathrm{x}}^{\mathrm{miss}}$) for minimum bias events without pile-up. As pointed out in the preceeding section, the default thresholds applied in ORCA (0.5 GeV for all HCAL cells) are different with respect to the default used in CMSSW ("Scheme-B"). Replacing the default CMSSW thresholds with the ORCA thresholds recovers most of the behaviour observed in ORCA, but several other known differences remain unaccounted.

Despite these known differences between ORCA and CMSSW, similar shape behaviour for ΣE_T, displayed as plots a) and b), is seen between ORCA and CMSSW_1_2_0. An obvious difference is observed in the mean value of ΣE_T, which is about 50 GeV lower in CMSSW_1_2_0 compared with ORCA. The origin of this shift to lower visible energies may lie in many places, but one possibility may be the known lower energy response of

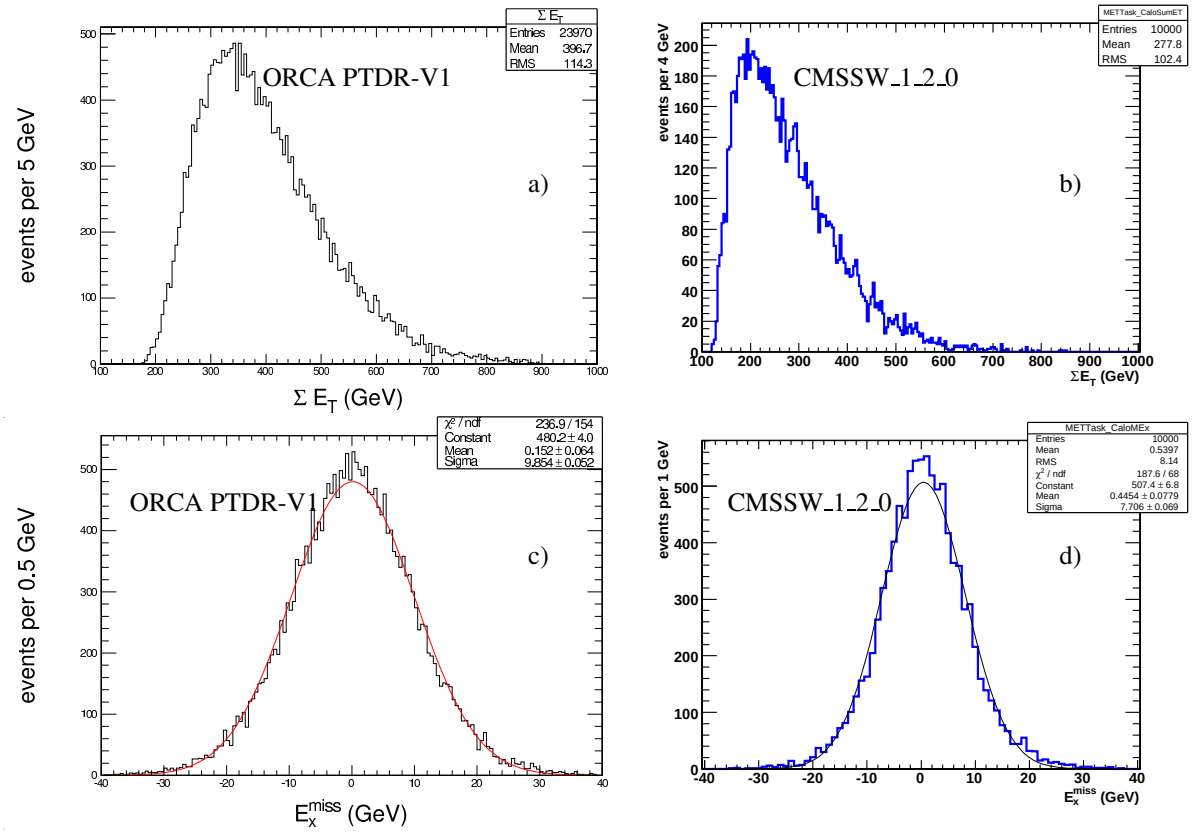


Figure 33: ΣE_T (top) and ME_x with a Gaussian fit (bottom) distributions for QCD dijet events for $\hat{p}_T < 15$ GeV using CMSSW_1_2_0.

HF in CMSSW_1_2_0. In the case of $E_{\mathrm{x}}^{\mathrm{miss}}$, displayed as plots c) and d), the width of the $E_{\mathrm{x}}^{\mathrm{miss}}$ distribution in CMSSW_1_2_0 is observed to be about 20% narrower than in ORCA and may also be an indication of the lower energy response in the forward hadronic calorimeter simulation for CMSSW_1_2_0. It should be emphasized, however, that the role of HF in these plots remains open and that the origin of the differences presented in this note are not yet fully understood.

9.2 Missing Transverse Energy in Low $\hat{p}_T$ QCD dijet events

Figure 33 compares the results presented in Figure 11.11 of the Physics TDR Vol. I and the corresponding results of this study for the total visible transverse energy ΣE_T and the x-component of the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ (labeled $E_{\mathrm{x}}^{\mathrm{miss}}$) for QCD dijet events with $\hat{p}_T < 15$ GeV. While pile-up effects corresponding to low-luminosity running are included, the ORCA studies included both in-time and out-of-time pile-up, whereas the CMSSW_1_2_0 studies only include in-time pile-up. As in the previous section, similar shape behaviour for ΣE_T, displayed as plots a) and b), is again seen between ORCA and CMSSW_1_2_0. However, shift of more than 100 GeV (with respect to ORCA) to lower average ΣE_T is observed in CMSSW_1_2_0. The larger magnitude of the shift (which should be compared with the 50 GeV shift in the previous section) is understood to be due to the fact that the CMSSW_1_2_0 simulation does not include out-of-time pile-up, whereas the ORCA studies did. As in the previous section and comparing plots c) and d) of Figure 33, the $E_{\mathrm{x}}^{\mathrm{miss}}$ resolution in CMSSW_1_2_0 is observed to be 20% narrower than in ORCA.

Figure 34 illustrates the resolution of the $E_{\mathrm{x}}^{\mathrm{miss}}$ as a function of ΣE_T for QCD dijet events with $\hat{p}_T < 15$ GeV. The blue points were obtained using CMSSW_1_2_0 and are superimposed on the ORCA results (displayed with black dots and a fitted red line). Notice that while the mean ΣE_T and width of the $E_{\mathrm{x}}^{\mathrm{miss}}$ distributions are different when comparing CMSSW_1_2_0 with ORCA, the relationship between the two variables in CMSSW_1_2_0 is not wildly dis-similar with respect to ORCA, but does seem to indicate that out-of-time pile-up plays an important role

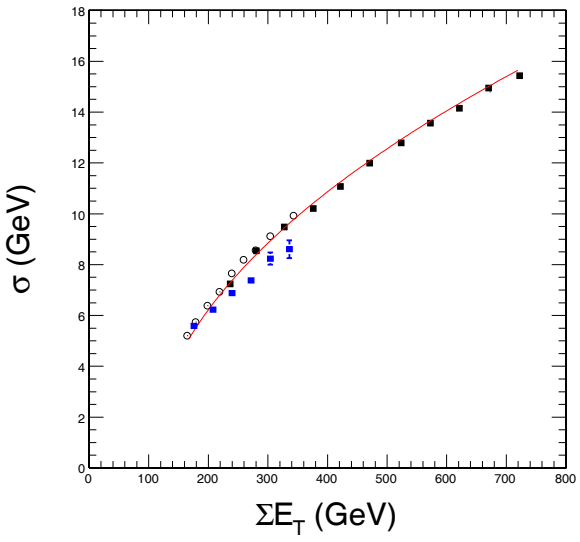


Figure 34: (ME_x vs ΣE_T distribution for QCD dijet events for $\hat{p}_T < 15$ GeV in CMSSW_1_2_0 (blue). The ORCA results are shown in black.

in the low MET resolution: A fit to the CMSSW_1_2_0 data yields a constant term of 1.8% and a stochastic term of 0.47 GeV$^{1/2}$, which should be compared with the corresponding result from ORCA reported in the PTDR of 0.65 GeV$^{1/2}$ for the stochastic term (no constant term is reported).

9.3 Missing Transverse Energy in all QCD dijet events

Figure 35 illustrates the sensitivity of the total visible transverse energy, ΣE_T, to different running and detector conditions. The black histogram represents conditions when there is no pile-up and when "Scheme-B" thresholds are applied. The gray histogram represents the case when in-time pile-up is included, but "Scheme-B" thresholds are still applied. When one applies the "ORCA PTDR" thresholds together with in-time pile-up, one obtains the red histogram. Finally, when one includes out-of-time pile-up, together with the "ORCA PTDR" thresholds, one arrives at the green histogram. For reference, the results from ORCA, as presented in the Physics TDR Vol. 1, are presented as the blue histogram. Notice that the final CMSSW (green) histogram nearly agrees with the ORCA (blue) results from the Physics TDR, but slightly overshoots. This small difference is likely due to the fact that in CMSSW the HCAL is read out over more time slices than in ORCA. Indeed, Figure 35 nicely shows that one must account for a variety of effects in CMSSW (thresholding, pile-up conditions, read-out configuration, etc) before one can adequately reproduce the ORCA results. This should also serve as a warning of the challenges in u nderstanding global quantities, such as the total visible transverse energy, in real data and for the need to exhaustively account for all detector as well as running conditions.

Figure 36 illustrates the behaviour of ΣE_T distributions for different QCD dijet samples within $20 < \hat{p}_T < 800$ GeV and corresponding to in-time pile-up only. At first glance, the ORCA distributions and CMSSW distributions look very similar, however upon closer inspection, the peak ΣE_T value in CMSSW for low $\hat{p}_T$ samples is about 300 GeV and increases to 1500 GeV for the highest $\hat{p}_T$ sample, thus indicating that, on average, ΣE_T is about 100 GeV to 200 GeV lower compared to results shown using ORCA. However, as demonstrated in Figure 35, nearly all of this difference can be accounted for by the lack of out-of-time pile-up in CMSSW_1_2_0.

Figure 37 illustrates the behaviour of $\sigma(E_{\mathrm{x}}^{\mathrm{miss}})$ with ΣE_T for QCD dijet samples within $20 < \hat{p}_T < 800$ GeV. A fit to the full range of available data points yields the function for the MET resolution: $\sigma^2 = (9.487\ \mathrm{GeV})^2 + (0.89\ \mathrm{GeV}^{1/2})^2(\Sigma E_T - 369.4\ \mathrm{GeV}) + (0.023(\Sigma E_T - 369.4\ \mathrm{GeV}))^2$ (but with a relatively high χ^2 value). Since we know that there are differences between ORCA and CMSSW in the offset of ΣE_T and we are most interested in comparing the performance of the stochastic and constant terms of the MET resolution, we explicitly fit for any

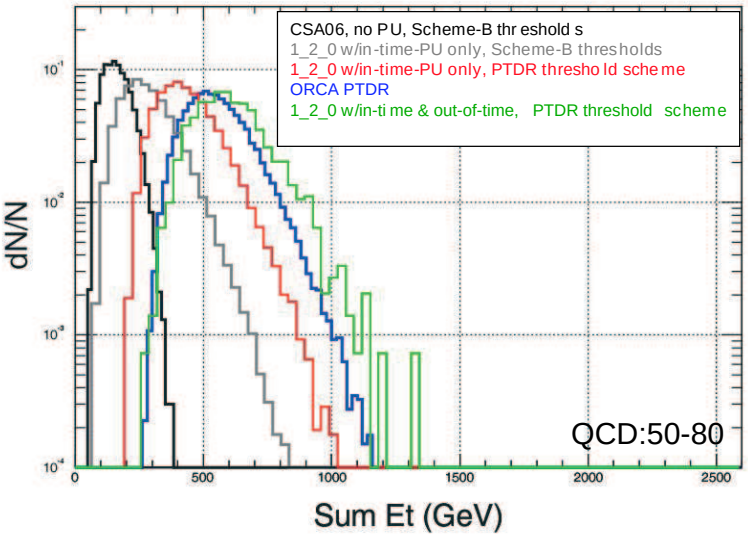


Figure 35: ΣE_T distrubtion of QCD dijet events with $50 < \hat{p}_T < 80$ GeV. Effects of losening threholds and adding pile-up can be seen. The result obtained using ORCA is seen in blue. The green curve is obtained by using ORCA thresholds and includes both in-time and out-of-time pile-up. The difference between the blue and green curves can be understood by the increased number of time slices used by HCAL in CMSSW.

offset in ΣE_T (which is seen to be 369.4 GeV). Compared with ORCA, the fit gives an 8% smaller stochastic term, but significantly larger (by about 91%) constant term. The highest data points for ΣE_T in CMSSW possibly indicate that the low ΣE_T regime may behave differently from the high ΣE_T regime. A fit to just the high ΣE_T gives better results (plot on the right), but this can only be confirmed or excluded with additional high ΣE_T data samples (which were not available for this study).

Figure 38 illustrates the $\sigma(E_{\mathrm{T}}^{\mathrm{miss}})$ distributions for QCD dijet samples within $20 < \hat{p}_T < 800$ GeV and having in-time pile-up only. The peak values for these distrubtions are in good agreement with those shown by ORCA; however, each distribution is somewhat more narrow, resulting in a smaller $\sigma(E_{\mathrm{T}}^{\mathrm{miss}})$ average value. This is behaviour expected due to the decrease in transverse energy activity present in each sample resulting from the lack of out-of-time pile-up in the CMSSW samples.

Figure 39 shows the dependence of $< E_{\mathrm{T}}^{\mathrm{miss}} >$ as a function of ΣE_T for all QCD dijet samples with $\hat{p}_T < 800$ GeV. Because there is no out-of-time pile-up included in the QCD samples, an overall offset in ΣE_T is removed by the fit. Nevertheless, compared with ORCA, lower averge MET values are still observed. A fit to the full range of available data points yields the function $< \sigma(E_{\mathrm{T}}^{\mathrm{miss}}) >^2 = (9.745GeV)^2 + (0.83GeV^{1/2})^2(\Sigma E_T - 277.8GeV) + (0.04(\Sigma E_T - 277.8GeV))^2$ or, when compared with ORCA, a 32% lower stochastic term and a constant term which is larger by more than a factor of 2. It is clear from the high χ^2, that the data points do not fit the function well. Hence, two additional fits are performed: one to the low ΣE_T regime and another to only the high ΣE_T regime. In the high ΣE_T regime, the constant term appears to reduce to about 2%, consistent with the ORCA results, however, the stochastic term increases to 1.46, or 15% higher than obtained with ORCA. Additional data points at high ΣE_T are required to confirm this behavior. The fit to the low ΣE_T regime (not shown) results in a constant term of 4% (confirming the full fit results), whereas the stochastic term reduces to 0.71, or about 40% lower than that of ORCA.

Fig. 40 shows the σ(MET-perp) vs ΣE_T distribution, where MET-perp is the component of MET perpendicular to the jet axis. The jet axis is found by defining the unit vector $\hat{n}$ in the direction of $\vec{p}_{T,jet1} - \vec{p}_{T,jet2}$. The resolution of MET-perp is determined by underlying event, pile-up, and detector noise rather than by the calorimeter response to the jets. Here we see that CMSSW is in rough agreement with ORCA.

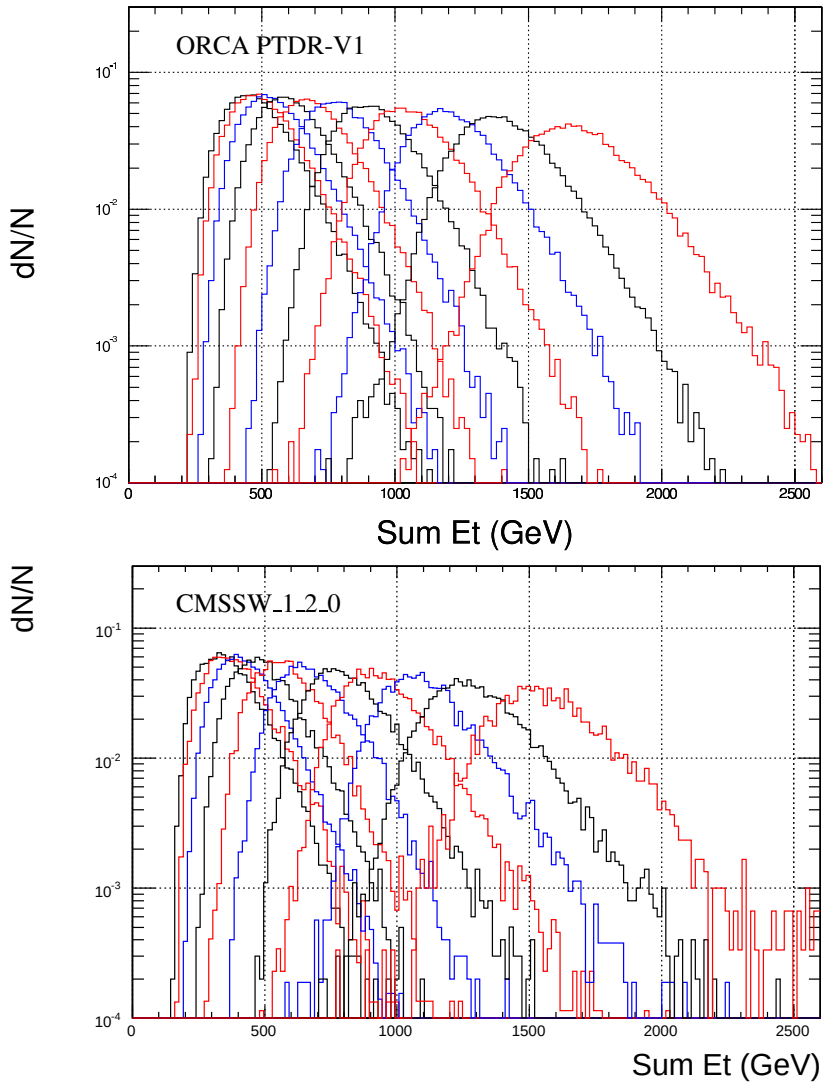


Figure 36: ΣE_T distributions for QCD dijet events for $\hat{p}_T$ ranges 20-30, 30-50, 50-80, 80-120, 120-170, 170-230, 230-300, 300-380, 380-470, 470-600, and 600-800 GeV going from left to right.

10 Conclusions

We summarized results of jet and Missing ET performance studies in CMSSW_1_2_0 to establish a baseline for comparing future releases of CMSSW.

Significant changes in the simulation and calibration of calorimeter properties are expected to be in place for the CSA07 production. The results presented here will help to evaluate the consequent impact on various aspects of jet and MET performance.

References

[1] CMS PTDR Vol.I: Detector Performance and Software, CERN/LHCC 2006-001

[2] SDPV agenda links can be found at: http://indico.cern.ch/categoryDisplay.py?categId=4l55

[3] SDPV final report is linked from: http://indico.cern.ch/conferenceDisplay.py?confId=12699

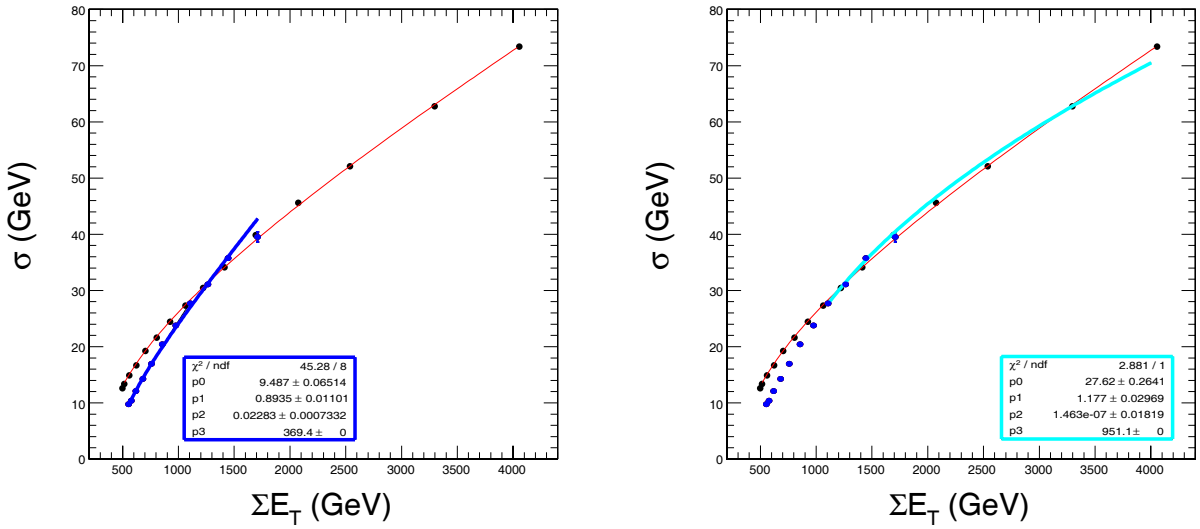


Figure 37: $\sigma(ME_x)$ vs ΣE_T for QCD dijet samples with $20 < \hat{p}_T < 800$ GeV (blue dots). Three fits are shown: the green line is the fit over the range $20 < \hat{p}_T < 380$ GeV, in blue $20 < \hat{p}_T < 800$ GeV, and in light blue $300 < \hat{p}_T < 800$ GeV. The ORCA results are indicated by black dots, with the resolution fit shown by the red line.

[4] R. Demina et al., CMS NOTE-2006/020, *"Calorimeter Cell Energy Thresholds for Jet Reconstruction in CMS"*

[5] CMS Note 2006/036, A. Heister et al., *"Measurement of Jets with the CMS Detector at LHC"*.

[6] Comp. Phys. Comm. vol 153/1 85-96 (2003), J. M. Butterworth, J. P. Couchman, B. E. Cox, B. M. Waugh, *"A C++ implementation of the $K_\perp$ clustering algorithm"*.

[7] M. Cacciari and G. P. Salam, Phys. Lett. B **641**, 57 (2006)

[8] D. Acosta et al., CMS NOTE-2006/067, *"The Underlying Event at the LHC"*.

[9] http://www.uscms.org/LPC/lpc_jetmet/edmData/edmData.html

[10] https://twiki.cern.ch/twiki/bin/view/CMS/CmsSamples

[11] CMS Note 2007/006, The CMS Collaboration, *"CMS Computing, Software and Analysis Challenge in 2006 (CSA06) Summary"*.

[12] O. Kodolova, ORCA study, private communication

[13] We thank T. Bose for help with production of special MC samples.

[14] CMS Note 2005/034, R. Harris, *"Jet Calibration from Dijet Balancing"*.

[15] H. Pi et al., CMS NOTE-2006/035, *"Measurement of Missing Transverse Energy With the CMS Detector at the LHC"*.

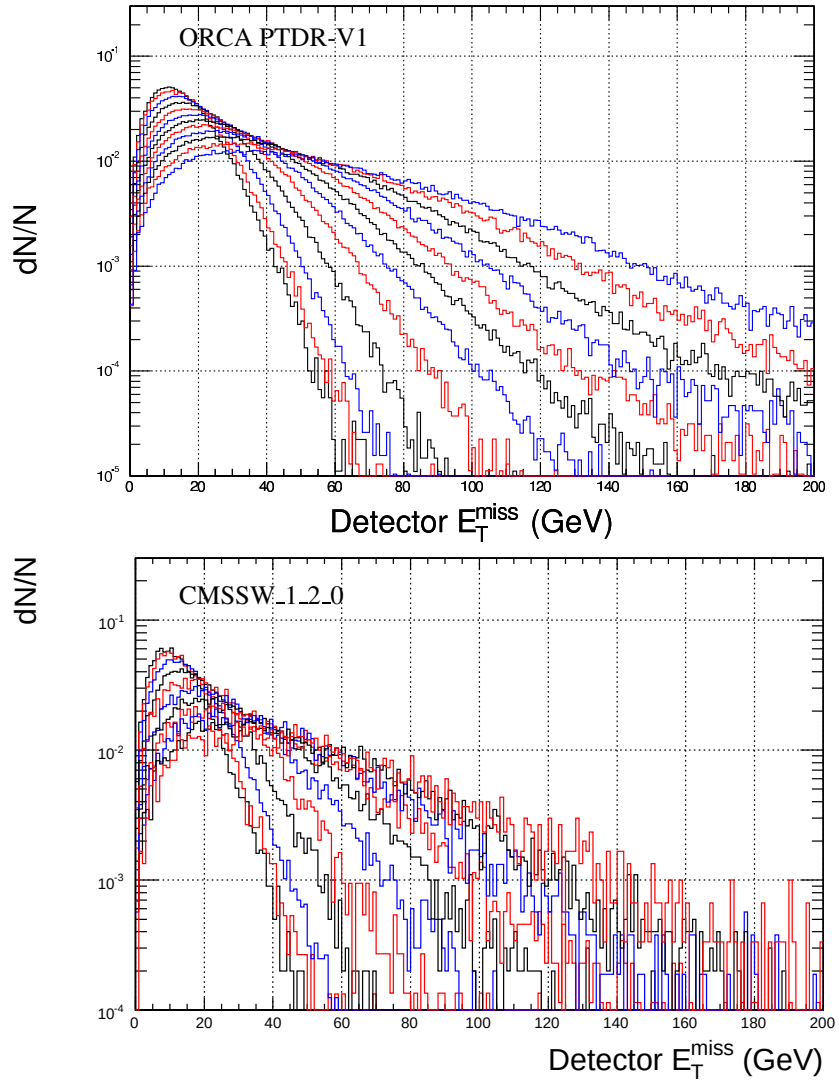


Figure 38: MET distributions for QCD dijet events for $\hat{p}_T$ ranges 20-30, 30-50, 50-80, 80-120, 120-170, 170-230, 230-300, 300-380, 380-470, 470-600, and 600-800 GeV going from left to right.

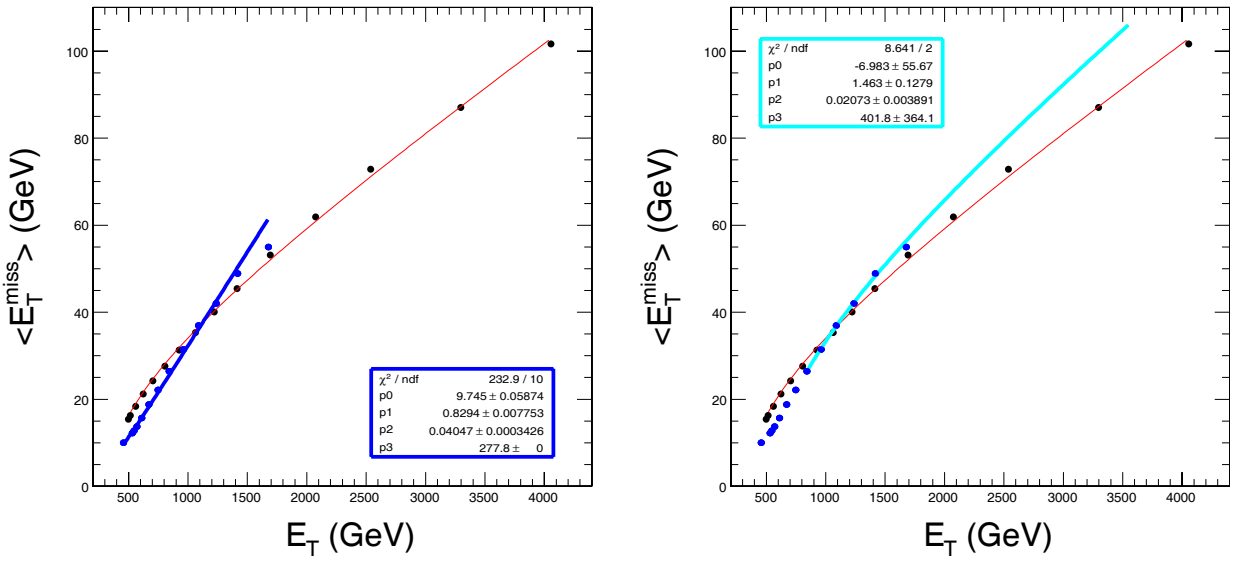


Figure 39: $< MET >$ vs ΣE_T distribution for QCD dijet samples with $\hat{p}_T < 800$ GeV (blue dots). Three fits are shown: the green line is the fit over the range $0 < \hat{p}_T < 230$ GeV, in blue $0 < \hat{p}_T < 800$ GeV, and in light blue $170 < \hat{p}_T < 800 GeV$. The ORCA results are indicated by black dots, with the resolution fit shown by the red line.

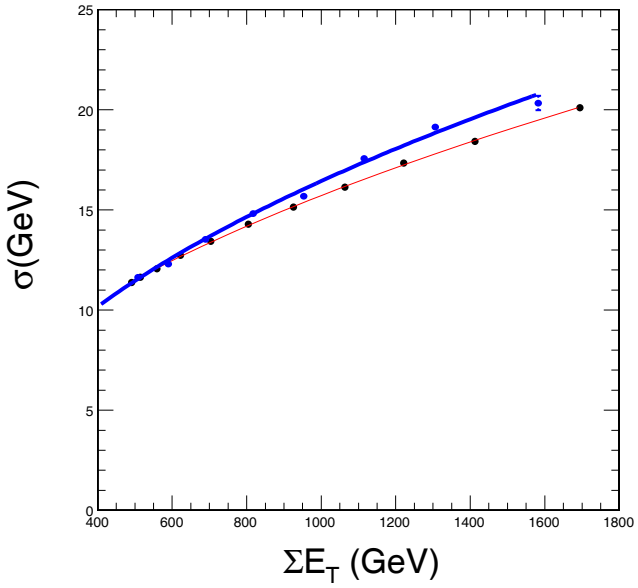


Figure 40: Distribution of σ(MET-perp) vs ΣE_T.