The Compact Muon Solenoid Experiment
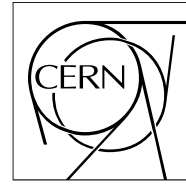
Analysis Note

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

$\not{E}_T$ Performance in CMS

S. Esen, G. Landsberg

Brown University, Providence, RI, USA

M. Titov

CEA Saclay, Gif-Sur-Yvette, France

A. DeRoeck[a], M. Spiropulu, M. Tytgat[b]

CERN, Geneva, Switzerland

D. Puigh, P. Wittich

Cornell University, Ithaca, NY, USA

K. Terashi

Rockefeller University, New York, NY, USA

A. Gurrola, T. Kamon, C.N. Nguyen, A. Safonov

Texas A&M University, College Station, TX, USA

J.W. Gary, F. Liu, H. Nguyen, J. Sturdy

University of California, Riverside, CA, USA

R. Cavanaugh, R. Remington, M. Schmitt, B. Scurlock

University of Florida, Gainesville, FL, USA

E.A. Albayrak, T. Yetkin[c]

University of Iowa, Iowa City, IA, USA

M. Zielinski

University of Rochester, Rochester, NY, USA

[a] Also: University of Antwerp, Antwerp, Belgium

[b] Also: University of Gent, Gent, Belgium

[c] Also: Mersin University, Mersin, Turkey

Abstract

We present the details of missing transverse energy ($\not{E}_T$) reconstruction and provide performance benchmarks. $\not{E}_T$ corrections for jet energy scale, muons, taus, and electrons are discussed. Examples are given for a strategy for cleaning up $\not{E}_T$ in real data and for studies of the effects on $\not{E}_T$ from miscalibration of, or problems with, the detector are presented. First results on optimization of $\not{E}_T$ resolution by varying calorimeter cell energy thresholds are presented.

1 Introduction: $\not{E}_T$ in CMS

While a typical general-purpose collider detector has nearly 4π solid angle coverage, it is not completely hermetic, as it must have an opening in the very forward direction to feed the beams through it. Therefore, low p_T energetic interaction products, which move in very forward direction can escape detection, thus making it impossible to utilize total energy balance in the collision as a useful constraint. However, although these escaping forward particles may carry significant longitudinal momentum (i.e., momentum in the direction along the beams), they nevertheless can not carry large transverse momentum. Indeed, in the case of the Compact Muon Solenoid (CMS) detector [1], the calorimeter system covers the pseudorapidity (η) range from -5 to $+5$, so in order for a particle to escape the detection it must have $|\eta| > 5$, and therefore its transverse momentum:

$$p_T = E/\cosh\eta < E/\cosh(5) < 0.013 \times E.$$

Even if the particle carries the entire energy of the beam, i.e. 7 TeV, its transverse momentum is < 100 GeV. In practice, due to rapidly falling parton distribution functions, typical momentum of forward particles in hard collisions of interest for most of physics goals of the experiment is significantly less than the full beam energy. Consequently, the transverse momentum carried away by the particles beyond the acceptance of the calorimeter is very small, and thus the detector allows for rather precise test of the 2D-momentum conservation in the plane perpendicular to the direction of the beams. As a result, any significant imbalance in transverse momentum measured in the calorimeter is indicative of a production of a weakly interacting particle in the collision, which by itself indicates a process of interest. Among the standard model (SM) particles, such an imbalance would indicate the presence of either a muon or a neutrino. The momentum of the muon can be precisely measured combining the central tracker and the muon system, and the calorimeter-based missing transverse momentum can be corrected for its presence. The only SM particle that would truly escape the detection is a neutrino, and its presence could therefore be inferred from the remaining imbalance in total transverse momentum as measured by the calorimeter and the muon system.

Many extensions of the SM predict the existence of other weakly interacting stable or quasi-stable particles. Therefore, if an excess of events with significant transverse momentum imbalance is still observed after accounting for all the SM processes, it would constitute a strong evidence for new physics beyond the SM. This makes total transverse momentum imbalance (or *missing transverse momentum*) an important variable for searches for new physics.

Unfortunately, the flip-side of this coin is that the missing transverse momentum is extremely sensitive to various detector malfunctions and particles hitting poorly instrumented regions of the detector. Any dead or malfunctioning element in the detector may result in an artificial imbalance, thus mimicking the signal for new physics. Consequently, great care is required to understand the distribution in missing transverse momentum as measured by the detector and to ensure that it's a trustworthy variable for searches.

Historically, missing transverse momentum is often referred to as Missing Transverse Energy ($\not{E}_T$). This notation is somewhat confusing, as it commonly refers to either the 2D-vector of missing transverse momentum, or to its magnitude. In this note we will use the following notations: $\not{E}_T$ is the scalar variable, which describes the magnitude of the missing transverse momentum vector, while $\vec{\not{E}_T}$ refers to the 2D-vector itself. We will also use $\not{E}_x$, $\not{E}_y$ to denote the two components of the $\not{E}_T$ 2D-vector. In some cases, when the two projections correspond to a specially chosen axis and the directions perpendicular to it, we will use $\not{E}_\parallel$, $\not{E}_\perp$ to denote the two projections. Finally, we will use $\phi(\not{E}_T)$ to denote the azimuthal direction of the $\not{E}_T$ vector.

As discussed above, the missing transverse energy vector is an important variable for electroweak measurements and for searches for new physics with CMS. The $\not{E}_T$ values in processes of interest in CMS range from "small," e.g. in Higgs decays or Standard Model processes such as top and W production, to "large" in the decays of SUSY particles or due to escaping gravitons in scenarios with large extra dimensions. In the former case a good understanding of the QCD and environmental backgrounds is needed, as well as a good low-energy resolution. For large $\not{E}_T$ values the understanding of the tails is the main experimental challenge.

In this note we describe the present state of the $\not{E}_T$ studies. A discussion of the $\not{E}_T$ reconstruction is presented, as well as its performance. The primordial $\not{E}_T$ is calculated from energy deposits in the CMS Calorimeter Towers (CaloTowers), but to improve its central value and resolution several corrections have been studied, notably those due to jet energy scale, and due to the presence of muons, electrons, and taus in the event. The validation of the $\not{E}_T$ object is discussed as well. A few methods of accounting for detector effects, some already utilized in the CMS PTDR volume II [2] are presented. The results of the first study on optimization of the tower thresholds to improve $\not{E}_T$ resolution are discussed. Finally we conclude with a number of points for future studies.

Monte Carlo (MC) samples used for the analyses presented in the note are based on the CMS software framework, CMSSW, versions from 1_3_1 to 1_5_2. All of these samples have been generated without pile-up, as none is expected in the early CMS running.

2 $\not{E}_T$ Reconstruction

2.1 Definition of $\not{E}_T$ Quantities

The Missing Transverse Energy ($\not{E}_T$) is determined in the recent versions of the CMS software framework, CMSSW, from the transverse vector sum over energy deposits in uncorrected, projective Calorimeter Towers:

$$\vec{\not{E}}_T = -\sum_n (E_n \sin\theta_n \cos\phi_n \hat{\mathbf{i}} + E_n \sin\theta_n \sin\phi_n \hat{\mathbf{j}}) = \not{E}_x \hat{\mathbf{i}} + \not{E}_y \hat{\mathbf{j}} \tag{1}$$

where the index n runs over all calorimeter input objects (e.g. energy deposits in towers, reconstructed hits, or generator-level particle energies). Here $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ are the unit vectors in the direction of the x and y axis of the CMS right-handed coordinate system, where z is pointed in the direction of the beam, and x is horizontal. Note that in the absence of $\not{E}_T$ from physics sources in the event, $\not{E}_x$ and $\not{E}_y$ are expected to be distributed as Gaussians with the mean of zero and the standard deviation of σ, while $\not{E}_T$ has a more complicated shape described by $\frac{\sqrt{2\pi}}{\sigma}\theta(\not{E}_T)\not{E}_T \times G(\not{E}_T, 0, \sigma)$, where $\theta(x)$ is the θ function (i.e., 1 for $x \geq 0$ and 0 otherwise), $G(x, \mu, \sigma) = \exp(-(x-\mu)^2/2\sigma^2)/\sqrt{2\pi}\sigma$ is a Gaussian with the mean μ and standard deviation of σ. Note that σ in the Gaussian describing $\not{E}_T$ is the same as the standard deviation in the $\not{E}_T$ projection on an arbitrary axis. It is further trivial to show that:

$$\langle \not{E}_T \rangle = \sigma\sqrt{\frac{\pi}{2}} \approx 1.253\sigma; \tag{2}$$

$$\sigma(\not{E}_T) = \sigma\sqrt{\frac{4-\pi}{2}} \approx 0.655\sigma, \tag{3}$$

where $\sigma(\not{E}_T)$ is the r.m.s. of the $\not{E}_T$ distribution.

The Scalar Transverse Energy (ΣE_T), or total visible E_T of the event is defined as the scalar sum:

$$\Sigma E_T = \sum_n E_n \sin\theta_n,$$

where the index n runs over the same input objects as for $\not{E}_T$. In the barrel, energy in a single projective Calorimeter Tower is formed from the unweighted sum over energy deposits in the 5×5 matrix of the EM Calorimeter (ECAL) crystals with that in a single Hadron Calorimeter (HCAL) tower. In the forward regions, a more complex association of ECAL crystals with HCAL towers is required.

$\not{E}_T$ is a global observable which depends on precise symmetric cancellations and hence is a good indicator of the health of the underlying input. The global nature of $\not{E}_T$ also means that disentangling and understanding the different factors affecting the performance can be very challenging. Most generally, the resolution of $\not{E}_T$ can be parameterised according to the following form:

$$\sigma(\not{E}_T) = A \oplus B\sqrt{\Sigma E_T - D} \oplus C (\Sigma E_T - D), \tag{4}$$

where the A ("noise") term represents effects due to electronic noise, pile-up (PU), and underlying-event (UE); the B ("stochastic") term represents the statistical sampling nature of the energy deposits in individual Calorimeter Towers; the C ("constant") term represents residual systematic effects due to non-linearities, cracks, and dead material; and D ("offset") represents the effects of noise and pile-up on $\sum E_T$. It is important to emphasize that the above parameterization factorizes A, B, C into *a priori* uncorrelated[1] effects. In particular, the stochastic and constant terms do not depend,to first order, on the effects of noise, pile-up, or underlying event. Hence, the offset due to noise must be explicitly taken into account when comparing the $\not{E}_T$ performance between samples having different noise thresholds and pile-up conditions – this can often be a dramatic effect, if not properly addressed.

---

[1] Of course, these terms become *a posteriori* correlated after fitting $\sigma(\not{E}_T)$ as a function of ΣE_T; while A and D are strongly correlated, separating them improves the description of $\not{E}_T$ in the case of noise uniform in ϕ, which does not affect $\not{E}_T$, but does offset $\sum E_T$.

In the case of QCD dijets events, no or little $E\!\!\!/_T$ is expected from physical sources (neutrinos), so any energy imbalance measured in the event is due to the finite calorimeter resolution and acceptance. In this case, the $E\!\!\!/_T$ uncertainty is proportional to the average $E\!\!\!/_T$, measured over many events (see Eqs. (2,3)): $\sigma(E\!\!\!/_T) \approx 0.52\langle E\!\!\!/_T\rangle$. Hence, a useful variable is the average $E\!\!\!/_T$ significance of an event, defined to be:

$$\frac{\langle E\!\!\!/_T\rangle}{\sigma(E\!\!\!/_T)} \approx \frac{1.91}{B}\frac{\langle E\!\!\!/_T\rangle}{\sqrt{\langle \Sigma E_T\rangle}} \approx 1.91,$$

where B is approximately 1.4 for QCD dijets in CMS. Assuming that only stochastic effects are important and D is small, the $E\!\!\!/_T$ significance estimates the number of standard deviations of the measured event $E\!\!\!/_T$ from the $E\!\!\!/_T = 0$ hypothesis. If stochastic effects dominate, then the $E\!\!\!/_T$ significance of QCD dijet events should be constant for all ΣE_T and close to 2. Any residual dependence of the $E\!\!\!/_T$ significance on ΣE_T is an indication of the importance of non-linearities (i.e. the constant term) in the $E\!\!\!/_T$ resolution.

2.2 Corrections to $E\!\!\!/_T$

The "raw" (or uncorrected) $E\!\!\!/_T$ may be corrected (due to jets, muons, etc) and the resulting corrected $E\!\!\!/_T$ object stored in the event. In such a case, the corrections that are applied to the "raw" $E\!\!\!/_T$ are stored as "deltas" to the x and y components of the $E\!\!\!/_T$ 2D-vector ($\Delta E\!\!\!/_x, \Delta E\!\!\!/_y$), together with the corresponding correction to the scalar sum E_T ($\Delta\Sigma E_T$). Since any given $E\!\!\!/_T$ object stored in the event may have undergone several different sequential corrections, the full list of corrections for each quantity is also stored in the $E\!\!\!/_T$ object itself. By taking the sum over all the corrections, $\sum_i \Delta(E\!\!\!/_{x,y})_i$ and subtracting that sum from the corrected $E\!\!\!/_x, E\!\!\!/_y$, quantities one is able to recover the original, uncorrected x and y components of the $E\!\!\!/_T$ 2D-vector,

$$\vec{E\!\!\!/}_T^{\,\mathrm{raw}} = \vec{E\!\!\!/}_T - \sum_i \Delta(\vec{E\!\!\!/}_T)_i,$$

as well as any intermediate corrected $E\!\!\!/_T$ quantity in the sequence.

2.3 Types of $E\!\!\!/_T$ in CMSSW

The class structure is such that all types of $E\!\!\!/_T$ inherit from the base $E\!\!\!/_T$ class, which itself inherits from the RecoCandidates class, which provides methods for generic kinematical quantities.

There are currently three "types" of $E\!\!\!/_T$ available in CMSSW: (1) a basic (generic) $E\!\!\!/_T$ (common to all types) whose class serves as a "base" class for other particular $E\!\!\!/_T$ classes. (2) $E\!\!\!/_T$ derived from energy deposits in the reconstructed calorimeter towers (CaloMET). This class inherits from the $E\!\!\!/_T$ class and includes calorimeter specific quantities such as the total contribution to ΣE_T made by HCAL. (3) $E\!\!\!/_T$ derived from energies of the generator-level particles. This class inherits from the $E\!\!\!/_T$ class and includes generator-level-specific quantities, such as the invisible energy content of an event.

3 $E\!\!\!/_T$ Performance in CMSSW_1_5_2

The data samples used to study the $E\!\!\!/_T$ performance were produced during the "pre-CSA07" production exercise. The GEN-SIM step was performed using CMSSW_1_4_x, and the DIGI-RECO step was performed in CMSSW_1_5_2. The samples correspond to QCD dijets produced in eleven bins of $\hat{p}_T$ between 20 GeV and not more than 800 GeV and contained no pile-up. Here $\hat{p}_T$ is the transverse momentum of the particles in the center-of-mass frame of the $2 \rightarrow 2$ parton scattering. At leading order, it is close to the transverse momentum of the jets produced in the scattering in either the center-of-mass or detector frame.

In order to reduce the effects of electronic noise, so-called "Scheme-B" thresholds were applied (by default) to all Calorimeter Towers (Table 1):

Figure 1 illustrates the ΣE_T and $E\!\!\!/_x$ distributions for QCD dijet events for $\hat{p}_T < 15$ GeV. The ΣE_T distribution has a mean of 41.5 GeV and a RMS of 36.4 GeV. The $E\!\!\!/_x$ resolution fit is observed to be 3.1 GeV.

Figure 2 illustrates the ΣE_T and $E\!\!\!/_T$ distributions for QCD dijet events for $20 < \hat{p}_T < 800$ GeV without pile-up. We note the peak ΣE_T value for low $\hat{p}_T$ samples is about 20 GeV and increases to 1200 GeV for the highest $\hat{p}_T$ sample.

Table 1: Calorimeter Tower thresholds, as defined in [4] for the barrel (HB), Outer (HO), endcap (HB), and forward (HF, Short and Long fibers) HCAL, barrel (EB) and endcap (EE) ECAL, in GeV. HF thresholds technically were not tuned as part of threshold optimization procedure, which resulted in the above three schemes. The thresholds listed under Scheme B are the ones used in CMSSW. No HF thresholds were used in ORCA.

Scheme	HB	HO	HE	HF (S)	HF (L)	ΣEB	ΣEE
A	0.70	0.85	0.90			0.20	0.45
B	0.90	1.10	1.40	1.8	1.2	0.20	0.45
C	1.20	1.30	1.80			0.20	0.45
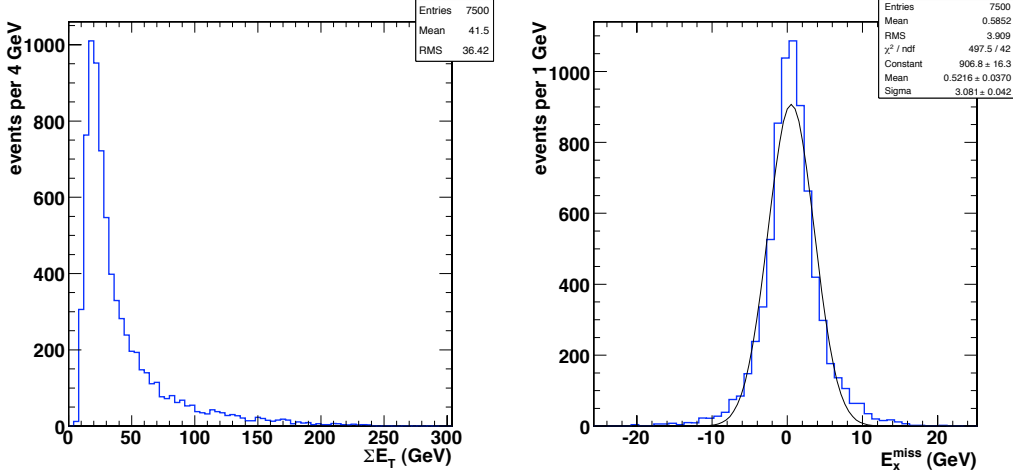
Figure 1: Left: ΣE_T distribution with the mean of 41.5 GeV and a RMS of 36.4 GeV, and Right: $\rlap{/}E_x$ distribution with a sigma of 3.9 GeV for QCD dijet events for $\hat{p}_T < 15$ GeV without pile-up, using CMSSW_1_5_2.

Figure 3 illustrates the $\sigma(\rlap{/}E_x)$ and $\langle \rlap{/}E_T \rangle$ vs. ΣE_T for QCD dijet samples with $20 < \hat{p}_T < 800$ GeV. To facilitate comparison of the stochastic and constant terms in these fits to those obtained in previous studies (which include different thresholds and events with pile-up), an offset, $\langle \Sigma E_T \rangle_{\min}$, is explicitly included as a variable parameter in the fit. The fit to the full range of data points yields the following functions:

$$\sigma^2(\rlap{/}E_x) = (1.53\ \text{GeV})^2 + (104\%\ \text{GeV}^{1/2})^2(\Sigma E_T - 77\ \text{GeV}) + (2.55\% \times (\Sigma E_T - 77\ \text{GeV}))^2;$$
$$\sigma^2(\rlap{/}E_T) = (1.48\ \text{GeV})^2 + (103\%\ \text{GeV}^{1/2})^2(\Sigma E_T - 82\ \text{GeV}) + (2.32\% \times (\Sigma E_T - 82\ \text{GeV}))^2.$$

The first terms correspond to the effects of electronic noise in an "empty" (or collision-less) event. Due to such noise, "empty" events will have a minimum average visible total transverse energy, $\langle \Sigma E_T \rangle_{\min}$, which is seen to be around 80 GeV. It is important to note that because the dataset conditions are very different in this study with respect to the PTDR [3] (the datasets studied in this work have different calorimeter cell thresholds and no pile-up, for example), the A term can not be easily compared with the earlier PTDR results. The second, B, term represents the stochastic effects of a sampling calorimeter and can be more easily compared with the earlier results of the PTDR: in the case of the PTDR, $\sigma(\rlap{/}E_x)$ had a stochastic term of 97% and, $< \rlap{/}E_T >$ had a stochastic term of 123% [3]. The third (C or constant) term can also be compared with the earlier results of the PTDR: in the case of the PTDR, $\sigma(\rlap{/}E_x)$ had a constant term of 1.2% and, $< \rlap{/}E_T >$ had a constant term of 1.9%. The fact that the stochastic and constant terms for both of these quantities are higher in the above equations is due to the fact that, compared with the software used in the PTDR, a much more realistic detector is now simulated (e.g. realistic gains are taken from sourcing data, etc.) in the version of the CMS software used for this study.

Figure 4 shows the ratio $\langle \rlap{/}E_T \rangle / sqrt\langle \Sigma E_T \rangle$ as a function of the total visible transverse energy, ΣE_T, of the event. To optain the $\rlap{/}E_T$ significance, $\rlap{/}E_T / \sigma(\rlap{/}E_T)$, the ratio $\langle \rlap{/}E_T \rangle / sqrt\langle \Sigma E_T \rangle$ in the plot should be scaled by a factor $1.91/B$. If the stochastic calorimeter-sampling term dominates the $\rlap{/}E_T$ uncertainty, then the $\rlap{/}E_T$ significance should be a constant for all values of ΣE_T. Deviations from a constant behavior indicate the presence of other non-stochastic effects, represented by the C ("constant") and A ("noise") terms in Eq. (4). Figure 4 demonstrates that a slowly rising $\sqrt{\Sigma E_T}$ behavior is immediately observed, indicating that the noise term has little effect and that the non-stochastic (or the "constant" C) term begins to dominate the $\rlap{/}E_T$ uncertainty rather quickly. The left pane shows the fit to the sum of the constant and sampling terms, and provides an adequate description of the
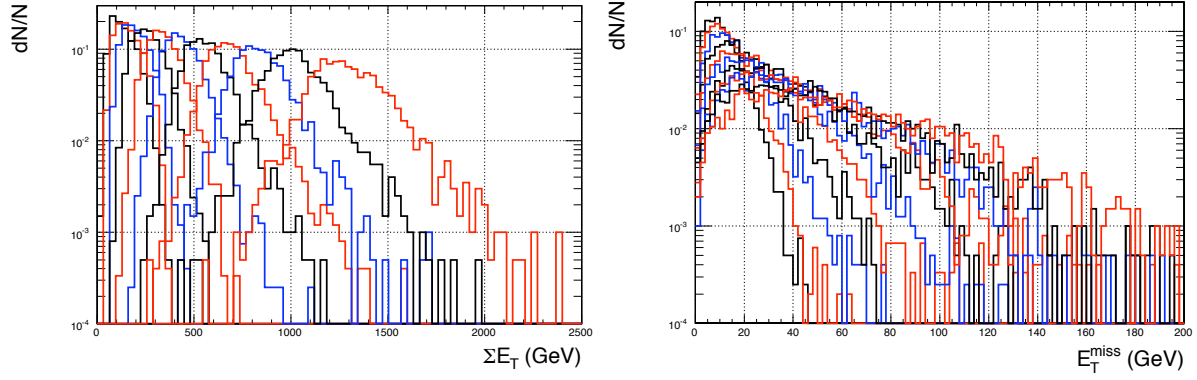
Figure 2: ΣE_T (left) and $\not{E}_T$ (right) distributions for QCD dijet events without pile-up for $\hat{p}_T$ ranges 20-30, 30-50, 50-80, 80-120, 120-170, 170-230, 230-300, 300-380, 380-470, 470-600, and 600-800 GeV, going from left to right.
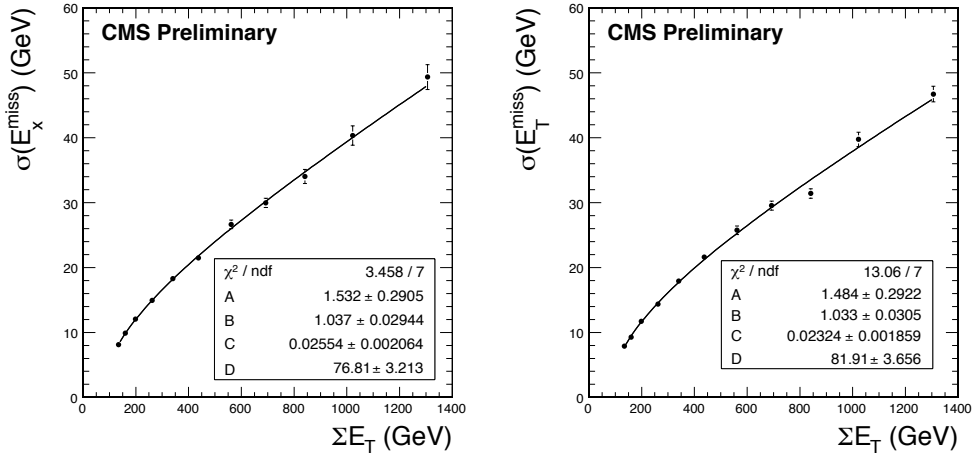

Figure 3: $\sigma(\not{E}_x)$ (left) and $\sigma(\not{E}_T)$ (right) vs. ΣE_T for QCD dijet samples without pile-up for $20 < \hat{p}_T < 800$ GeV (black dots). The resolution fit is shown by the black line. The fit parameters correspond to the corresponding terms in Eq. (4).

spectrum both at low and high ΣE_T. The right pane shows the fit to all three terms, which, while improving the overall consistency of the fit, also shows that the noise term has a little effect on the $\not{E}_T$ resolution. which, while improving the overall consistency of the fit, shows that the noise term has a little effect on the $\not{E}_T$ resolution.

4 $\not{E}_T$ Validation

An automated $\not{E}_T$ validation package has been created to systematically check each of the available quantities returned by the $\not{E}_T$ object for changes resulting from software modification. Because $\not{E}_T$ is sensitive to changes in the lower level objects used in its reconstrction, the $\not{E}_T$ validation package also checks for changes in CaloTowers, HCALRecHits, and ECALRecHits. The validation package (found in Validation/RecoMET) consists of four modules corresponding to $\not{E}_T$, CaloTower, HCALRecHit, and ECALRecHit collections. Every quantity provided by the reconstructed objects has a corresponding histogram, which is checked for changes for each new version of CMSSW. A shell script, also found in Validation/RecoMET, is used to automatically open histogram ROOT files, run statistical tests on each histogram, and build the web-interface used to display the validation results. The standard $\not{E}_T$ validation is performed using officially produced "RelVal" datasets. In particular, the standard set of good $\not{E}_T$ probes is used, and include $Z \rightarrow \mu^+\mu^-$, $Z' \rightarrow$ dijets, and QCD dijet events. Figure 5 shows the output produced by the $\not{E}_T$ validation package comparing CMSSSW_1_2_0 to CMSSW_1_3_1. On the y-axis is listed each of the available $\not{E}_T$ quantities that can be examined, with the test score shown on the x-axis. The score is set equal to the higher of the χ^2-test score and Kolmogorov-test score: a score is lower than 10^{-6} is considered failing. In the example output shown in Fig. 5, we observe a failure in the comparison of the $\phi(\not{E}_T)$ distribution
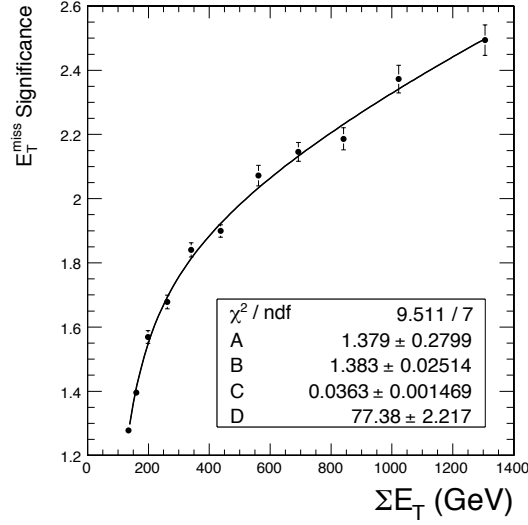
Figure 4: $\not{E}_T$ significance as a function of the total visible transverse energy, ΣE_T, of the event. For low values of ΣE_T, the $\not{E}_T$ significance asymptotically tends towards 1.9. For higher values of ΣE_T, the "constant" term dominates the $\not{E}_T$ significance and leads to a slowly rising $\sqrt{\Sigma E_T}$ behavior. The fit includes all four terms A–D in Eq. (4).

(CaloMETPhi), and a very low score in y-component of $\not{E}_T$ (CaloMEy). The $\phi(\not{E}_T)$ distribution is shown in Fig. 5.
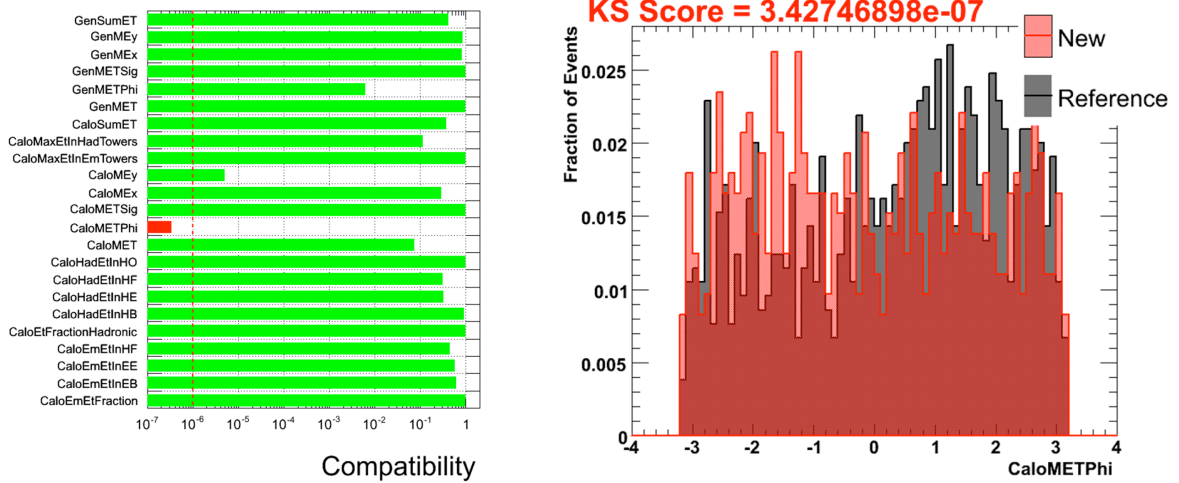
Figure 5: Left: $\not{E}_T$ validation summary for QCD dijet events for $\hat{p}_T$ ranges 15-20 GeV. Right: $\not{E}_T$-ϕ distributions for QCD dijet events for the $\hat{p}_T$ range 15-20 GeV.

The particular cause of this failure is understood by examining the input to calorimeter $\not{E}_T$ (Calorimeter Towers). The summary validation output for the CaloTower quantities is shown in Fig. 6. This plot shows the available CaloTower quantities along the y-axis, with range of compatibility scores along the x-axis. Because CaloTowers quantities are examined within η-rings, the range of scores illustrated by a colored bar (as shown in Fig. 6) extends from the lowest to highest score across all η-rings. This is further demonstrated by Fig. 7. Here the statistical test score for $\phi(\not{E}_T)$ is shown as a function of η-ring. From this plot, it is clear that for η-rings > 28 (HF) there has been a change in reconstruction performance. In this case, it was due to a problem in the Shower Library in HF, which by now has been fixed. The distribution of $\phi(\not{E}_T)$ in η-ring +40 is shown in Fig. 8.
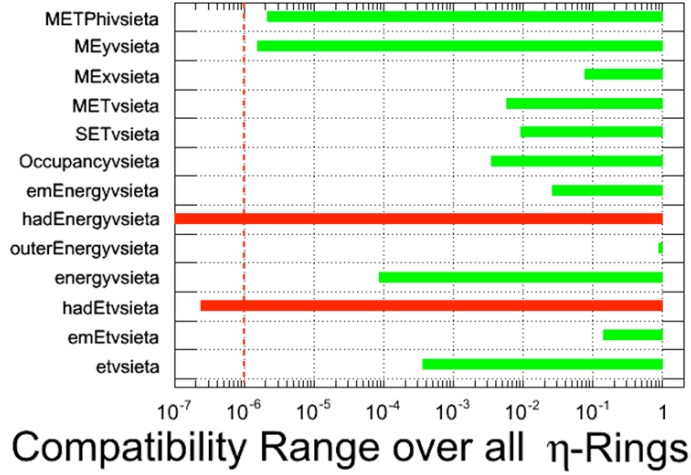
8

Figure 6: Calorimeter tower validation summary for QCD dijet events for the $\hat{p}_T$ range 15-20 GeV.

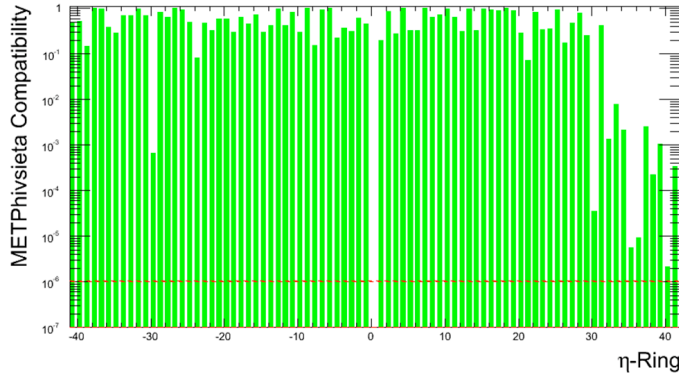Figure 7: Calorimeter tower validation scores for $\phi(\not{E}_T)$ as a function of η ring number in QCD dijet events for the $\hat{p}_T$ range 15-20 GeV.

5 $\not{E}_T$ Corrections

5.1 Introduction

The Calorimeter Tower based $\not{E}_T$ calculation can be improved by correcting $\not{E}_T$ for several effects. First there are the Jet Energy Scale corrections, also often called Type I corrections. These corrections take the measured raw energy values and adjust them for the difference between the raw jet energy and the true jet energy, as defined by the Jet Energy Scale (JES) group [5].

Events may also contain muons. In the majority of cases the muon deposits only very little energy in the calorimeters. Hence to correct for the muon response the actual muon momentum measurement from the central tracker and muon system is used to replace the energy measured along the muon trajectory in the calorimeter. Electrons are reconstructed in the detector as well, and the electron energy scale can be used to adjust $\not{E}_T$, as well as corrections for mismeasurements due to uninstrumented regions.

Isolated taus yield jets that differ substantially from average QCD jets. Specific corrections for τ-jets based on particle flow methods can be employed in this case.

Besides the corrections to the high p_T objects as listed above, there are effects due to the soft underlying event, pile-up etc. These so called Type II corrections, studied in [6], were found to improve $\not{E}_T$ bias and resolution in several benchmark processes studied. It was also suggested that perhaps a different definition of Type I corrections, taking into account the particles that are swept out of the jet cone due to the strong magnetic field of CMS, would be useful for future developements. The Type II corrections have not been studied within the CMSSW framework as of yet. Developing these corrections is one of the high-priority future tasks for the JetMET group.
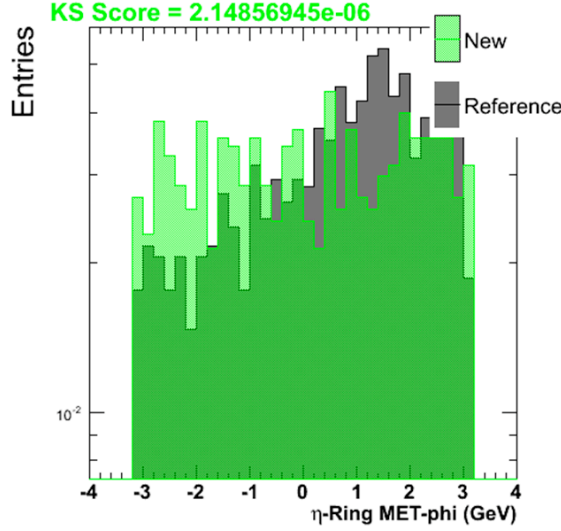
Figure 8: $\phi(\not{E}_T)$ distribution in η-ring +40 for QCD dijet events for the $\hat{p}_T$ range 15-20 GeV.

The $\not{E}_T$ corrections aim to bring the measured $\not{E}_T$ value closer to the true $\not{E}_T$ on event by event basis, and to improve the resolution of the $\not{E}_T$ variable in general. The individual corrections are described in this section.

5.2 Type I Jet Energy Scale Corrections

We discuss the effect of (Type I) JES corrections on $\not{E}_T$. Type I corrections to $\not{E}_T$ are based on the energy response of jets reconstructed in the event. They are essential to remove biases in the determination of $\not{E}_T$ due to the non-linear response of the calorimeter to jet energy deposits at different values of E_T and rapidity. We first define JES corrections and then we discuss how they affect $\not{E}_T$ bias and resolution.

Reliable MC jet corrections are the first step toward realistic jet corrections based on data. They allow us to test various methods and develop techniques to quickly derive the corrections with data. Standard MC-based jet (MCJet) corrections [7] resolve jets as a function of two variables: transverse energy (E_T^j) and pseudorapidity ($|\eta_j|$). The MCJet-corrections are derived by fitting the relative response of the calorimeter, defined as the ratio of transverse energies of the reconstructed and MC-generated jets, $E_T^j(\text{Rec})/E_T^j(\text{MC})$, with a Gaussian in each E_T^j, $|\eta_j|$ bin. This technique takes into account non-uniformity and nonlinearity of the calorimeter [7].

More advanced, EMF-dependent MC-based jet corrections additionally take into account non-compensation in the calorimeter by utilizing the fraction of jet energy deposited in the ECAL (EM fraction, or EMF) [8]. These three variables contain most of the information about a jet. They are derived based on the absolute response of the calorimeter, $E_T^j(\text{Rec}) - E_T^j(\text{MC})$, which is more Gaussian than the relative response, especially at low jet E_T. The current scheme of JES corrections in CMS, calls for factorization of various corrections, and the EMF-based correction is applied on top of the MCJet one [5] and in what follows, we refer to it as MCJet+EMF correction. It has been demonstrated [8] that accounting for the jet EMF does improve jet energy resolution by about 10% compared to the result of just the MCJet corrections.

When applying the Type I corrections, care has to be taken to avoid applying the jet response corrections to electrons reconstructed as jets with high fraction of EM energy; such jets can be identified, e.g., by matching to reconstructed electrons. In addition, average MCJet corrections are also not applicable to true jets with high EM fraction; furthermore the MCJet+EMF-corrections for such jets are already small. It is thus advantageous to eliminate jets with large EMF from the Type I MCJet correction procedure by only correcting for jets with EMF below a certain threshold. Finally, jet corrections are poorly known for jets at very low p_T^j, so it is beneficial to exclude jets below certain p_T^j threshold from Type I $\not{E}_T$ corrections as well.

The effects of these cuts are illustrated in Fig. 9. When Type I MCJet corrections are applied to a sample of $W \to e\nu$ events without an EMF cut, the $\not{E}_T$ distribution is clearly distorted relative to the generated distribution; after the EM fraction cut of 0.9 a good agreement is recovered. The effect of the jet p_T threshold is small in this sample, although one can note a smaller tail in the $\not{E}_T$ distribution when the 20 GeV jet threshold is applied.

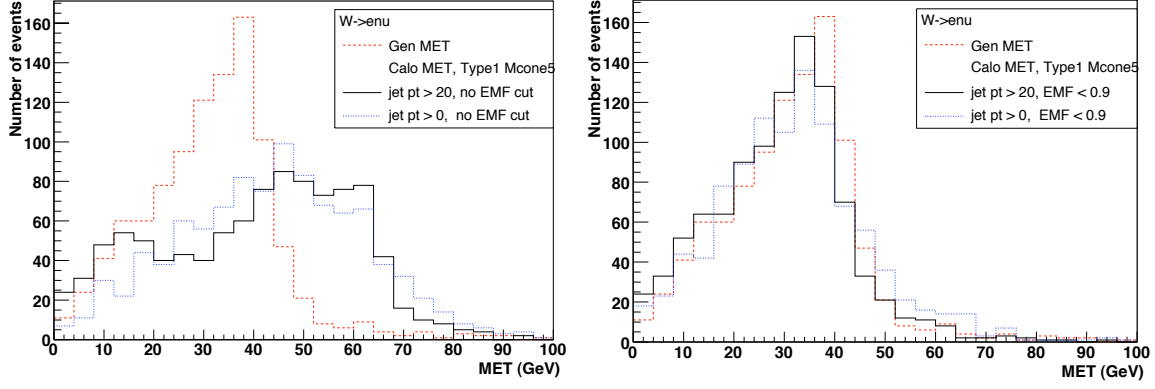In the recent CMSSW versions $\not{E}_T$ corrections for JES are implemented via the `JetMETCorrections` package.

Figure 9: Effect of the EMF or p_T cuts (left) and EMF< 0.9 and p_T cuts (right) on $\not{E}_T$ in the $W(e\nu)$ sample.

The following three jet algorithms are used to correct $\not{E}_T$:

```
iterativeCone5CaloJets,
midPointCone5CaloJets, and
midPointCone7CaloJets.
```

The package corrects jets according to one of the standard algorithms used for JES corrections (the default is MCJet). $\not{E}_T$ is corrected according to the following formula:

$$\vec{\not{E}}_T^{\,\mathrm{corr}} = \vec{\not{E}}_T - \sum_{i=1}^{\mathrm{N_{jets}}} \left[\vec{p}_{T_i}^{\,\mathrm{corr}} - \vec{p}_{T_i}^{\,\mathrm{raw}} \right].$$

The sum runs over all the jets with p_T^{raw} greater than `jetPTthreshold` and EMF less than `jetEMfracLimit`. The above two parameters are specified in the:

`JetMETCorrections/Type1MET/data/MetType1Corrections.cff`

file and are set by default to 20 GeV and 0.9, respectively. Note that these default numbers have not been thoroughly optimized and can be considered an ad hoc choice at the moment.

We applied both MCJet and MCJet+EMF JES corrections to events with true $\not{E}_T$ using $W(e\nu)+$ jets samples produced in CMSSW_1_3_X. This sample allows to see the effect of the corrections as a function of either $\sum E_T$ or $\not{E}_T$ in a large range. We used default cuts on jets minimum transverse energy and EMF for the MCJet corrections. For the MCJet+EMF corrections we removed the EMF cut and lowered the jet p_T threshold to 10 GeV. Also we removed the electrons from the jet list since the electrons should not be corrected as jets, as discussed above.

In order to demonstrate the effect of Type I $\not{E}_T$ corrections, we define the parallel component of $\not{E}_T$ ($\not{E}_\parallel$), which is the projection of the reconstructed $\not{E}_T$ on the direction of the neutrino produced in the W decay, as determined at the generator level (p_ν). This direction is the direction of "true" $\not{E}_T$ in the event, and thus allows to study both the bias and resolution in the presence of true $\not{E}_T$.

Most importantly, both MCJet and MCJet+EMF corrections largely eliminate significant bias in $\not{E}_\parallel$ vs. $\not{E}_T$(MC). This is particularly important at large $\not{E}_T$(MC) (see the left plot in Fig. 10). The absolute $\not{E}_\parallel$ resolution as a function of $\not{E}_T$(MC) is shown in the middle plot in Fig. 10. The $\not{E}_\parallel$ resolution after applying MCJet corrections does not improve compared to the no-correction case for low $\not{E}_T$, but the MCJet+EMF-corrections show significantly improved resolution compared to both the MCJet-corrections and uncorrected $\not{E}_T$ at large $\not{E}_T$, consistent with the improvement achieved for jet energy resolutions. The reason the absolute $\not{E}_\parallel$ resolution does not improve significantly when JES corrections are applied is that their application is equivalent to rescaling energy in the event by a factor greater than one. Since absolute $\not{E}_T$ resolution worsens with the increased total energy in the event, as shown in Fig. 3, JES corrections are not very effective in improving absolute $\not{E}_\parallel$resolution. More significant improvement is achieved in relative $\not{E}_\parallel$resolution, defined as $\sigma(\not{E}_\parallel)/\not{E}_\parallel$ (see the right plot in Fig. 10), primarily due to removed $\not{E}_T$ bias, which increases overall $\not{E}_\parallel$, which is found in the denominator of the expression for relative resolution.

Similar situation is seen when the $\not{E}_\parallel$ resolution is plotted as a function of H_T in the event, defined as the scalar sum over the transverse energies of all the jets. While the absolute resolution in $\not{E}_\parallel$ does not improve appreciably

after applying MCJet corrections and does improve only slightly after applying MCJet+EMF corrections (see Fig. 11, left pane), an improvement is more visible in the relative $\mathbb{E}_\parallel$ resolution, defined as $\sigma(\mathbb{E}_\parallel)$ divided by the H_T (see Fig. 11, right pane). This figure of merit takes into account that for large H_T the resolution is dominated by the constant term, and thus grows linearly with H_T (see Eq. (4)). Again, the MCJet+EMF corrections perform better than the MCJet ones.
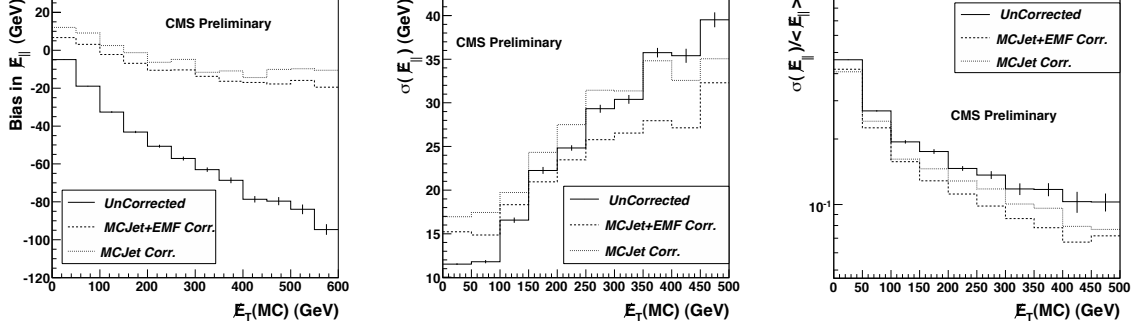

Figure 10: $\mathbb{E}_\parallel$ bias and resolution in the $W(e\nu)+$ jets sample. Left: $\mathbb{E}_\parallel$ bias as a function of $\mathbb{E}_T(\text{MC})$. Center: absolute resolution, $\sigma(\mathbb{E}_\parallel)$, as a function of $\mathbb{E}_T(\text{MC})$. Right: relative resolution, $\sigma(\mathbb{E}_\parallel)/\mathbb{E}_\parallel$, as a function of $\mathbb{E}_T(\text{MC})$.
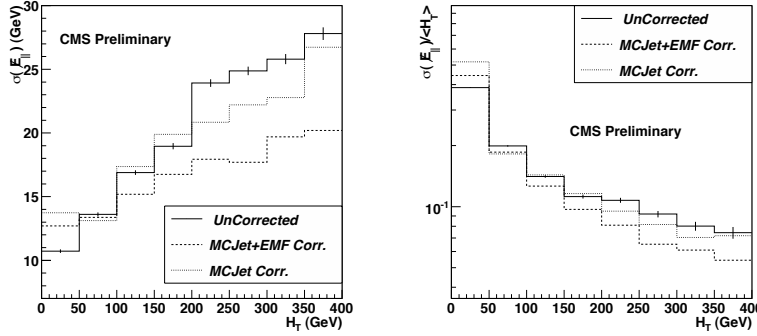

Figure 11: $\mathbb{E}_\parallel$ resolution in the $W(e\nu)+$ jets sample. Left: Absolute resolution in $\mathbb{E}_\parallel$, $\sigma(\mathbb{E}_\parallel)$, as a function of reconstructed H_T, defined as the scalar sum over transverse energies of all jets in the event. Right: relative resolution, $\sigma(\mathbb{E}_\parallel)/H_T$, as a function of reconstructed H_T.

5.3 Muon corrections

Missing transverse energy has to be also corrected for the muon objects in the event. As a muon is a minimum ionizing particle over the wide range of a particle momentum, it deposits only a small amount of energy in the calorimeter, typically a few GeV, and thus can create a fake $\mathbb{E}_T$. Therefore, the transverse momenta of all muons in the event are subtracted vectorially from the $\mathbb{E}_T$ vector, after deduction of the energy deposition of the muons in the calorimeter:

$$\vec{\mathbb{E}}_T = - \sum_{i=1}^{\text{towers}} \vec{E}_T^{\,i} - \sum^{\text{muons}} \vec{p}_T^{\,\mu} + \sum_{i=1}^{\substack{\text{deposit} \\ \text{towers}}} \vec{E}_T^{\,i}.$$

The muon $\mathbb{E}_T$ corrections are studied using $Z \to \mu^+\mu^-$ Monte Carlo samples reconstructed with CMSSW_1_3_1. To study the dependence of $\mathbb{E}_T$ from the p_T of the Z, two samples were generated with different p_T^Z requirements: 1) "inclusive" sample with a cut of $p_T^Z > 0$ GeV and 2) "high p_T" sample with $230 < p_T^Z < 300$ GeV. Table 2 gives the summary of the MC samples. The "raw" $\mathbb{E}_T$ in the event is reconstructed from energies in the Calorimeter Towers after applying scheme B thresholds, given in table1.

Muons are identified combining information from the inner tracker and the muon system. The analysis starts

Table 2: $Z \to \mu^+\mu^-$ MC Samples used to study $E\!\!\!/_T$ corrections due to minimum ionizing muons.

Z-boson p_T (GeV)	#Events	Description
$p_T^Z > 0$ "Inclusive"	10.8K	LPC Physics group production CMSSW_1_3_1 (GEN-SIM-DIGI-RECO) No pile-up
$230 < p_T^Z < 300$ "High p_T"	6.0K	CMS Spring 2007 production CMSSW_1_2_3 (GEN-SIM), 1_3_1 (DIGI-RECO) No pile-up

with muons from "globalMuons" collection in CMSSW_1_3_1 with $p_T^\mu > 10$ GeV and $|\eta^\mu| < 2.4$. To ensure a good quality of muon objects the following selection criteria are also applied: $\Delta p_T^\mu / p_T^\mu < 0.5$, $\chi^2 < 10$ and $N_{hits} \geq 13$, where Δp_T^μ is the error of p_T^μ, χ^2 and N_{hits} are the reduced χ^2 of the fit and the number of hits found in the combined muon and tracker systems, respectively. Further, the muons are required to be isolated by demanding that $[\sum^{R<0.2} p_T^{track} - p_T^\mu]/p_T^\mu < 0.2$, where $\sum^{R<0.2} p_T^{track}$ is the sum of p_T of tracks in a cone with radius $R = 0.2$ around the muon track. Such procedure selects muons mostly stemming from Z decays, which are isolated from jets. Jets are reconstructed using the midpoint clustering algorithm with a cone size of $R = 0.5$.

We calculate the direction of the Z boson by selecting events with two good muons, and divide $E\!\!\!/_T$ into two components: parallel ($E\!\!\!/_\parallel$) and perpendicular ($E\!\!\!/_\perp$) to the Z direction. The "raw" $E\!\!\!/_T$ tends to lie along the Z direction, as the dominant $E\!\!\!/_T$ source is muon energy which is not registered in the calorimeter. The estimated $E\!\!\!/_T$ resolution in the orthogonal direction $E\!\!\!/_\perp$ is dominated by the underlying event and pile-up activity and nominally Gaussian distributed around zero. Figure 12 (13) shows the $E\!\!\!/_\parallel$ ($E\!\!\!/_\perp$) distribution of $Z \to \mu\mu$ events. The "raw" $E\!\!\!/_\parallel$ distributions (dashed histograms) are shifted towards positive values, especially for large values p_T of the Z boson.
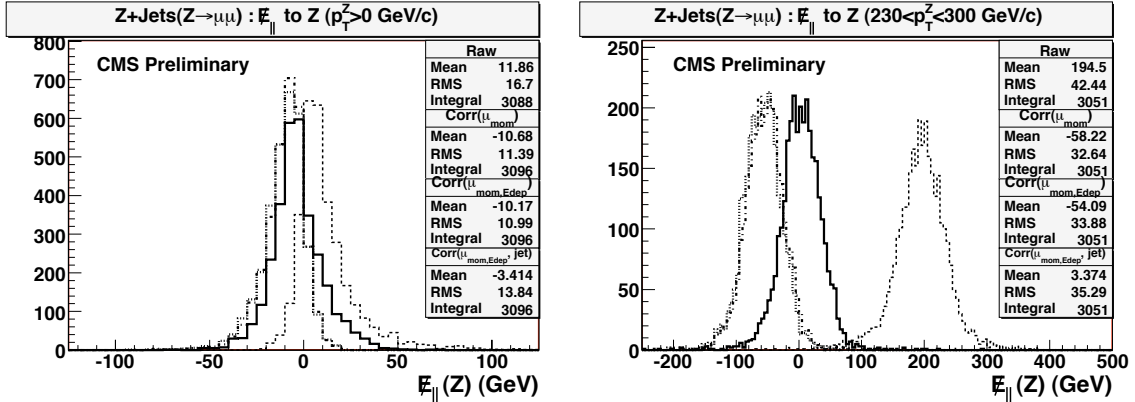

Figure 12: Distributions of $E\!\!\!/_\parallel$ ($E\!\!\!/_T$ component parallel to the reconstructed Z direction) in the inclusive (left) and high p_T (right) $Z \to \mu\mu$ samples. The histograms represent $E\!\!\!/_\parallel$ distributions at different correction levels: uncorrected "raw" $E\!\!\!/_\parallel$ (thin solid line), muon momentum correction (dotted line), muon momentum and muon energy deposit in the calorimeter (dash-dotted line), all muons and jet energy scale corrections (thick solid line). See text for more details.

The first step is to correct $E\!\!\!/_T$ for the reconstructed muon energy. The 2-dimensional (p_x^μ, p_y^μ) muon momentum vectors are taken from "globalMuons" collection, and subtracted from the "raw" $E\!\!\!/_T$ 2D vector. Distributions of the $E\!\!\!/_\parallel$ and $E\!\!\!/_\perp$ after the muon momentum correction are shown in Figs. 12 and 13. After applying the optimal muon correction, the $E\!\!\!/_\parallel$ distribution is shifted to the negative values, which indicates that the lower jet energy response causes the mismeasurement of $E\!\!\!/_T$, as the jets are produced opposite to the Z in azimuthal angle.

To avoid the double counting, it is necessary to take into account the muon energy deposited in the calorimeters. In this study, we use the `TrackAssociator` tool [9] to identify muon energy in the towers, which has been developed for muon isolation studies by the Muon POG group. The `TrackAssociator` tool uses `SteppingHelixPropagator` [10], which accounts for the magnetic field, energy loss and multiple scattering effects in the detector material. The tool allows us to obtain the energy in several different ways (e.g. energies in RecHits or CaloTowers crossed by muon track trajectory, or a $N \times N$ area of calorimeter units around the track
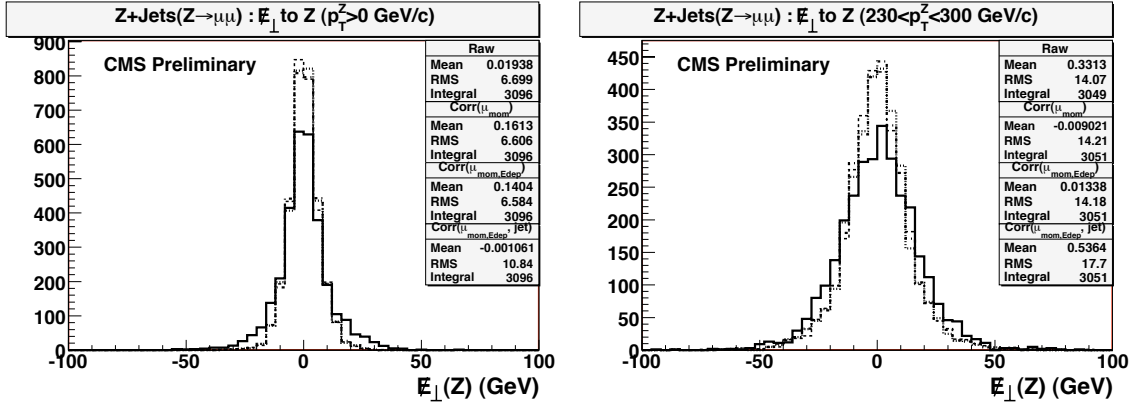
Z+Jets(Z→μμ) : $E\!\!\!/_\perp$ to Z (p_T^Z>0 GeV/c)

CMS Preliminary

Raw	
Mean	0.01938
RMS	6.699
Integral	3096
Corr(μ_{mom})	
Mean	0.1613
RMS	6.606
Integral	3096
Corr($\mu_{mom,Edep}$)	
Mean	0.1404
RMS	6.584
Integral	3096
Corr($\mu_{mom,Edep}$, jet)	
Mean	-0.001061
RMS	10.84
Integral	3096

$E\!\!\!/_\perp$(Z) (GeV)

Z+Jets(Z→μμ) : $E\!\!\!/_\perp$ to Z ($230<p_T^Z<300$ GeV/c)

CMS Preliminary

Raw	
Mean	0.3313
RMS	14.07
Integral	3049
Corr(μ_{mom})	
Mean	-0.009021
RMS	14.21
Integral	3051
Corr($\mu_{mom,Edep}$)	
Mean	0.01338
RMS	14.18
Integral	3051
Corr($\mu_{mom,Edep}$, jet)	
Mean	0.5364
RMS	17.7
Integral	3051

$E\!\!\!/_\perp$(Z) (GeV)

Figure 13: Distributions of $E\!\!\!/_\perp$ ($E\!\!\!/_T$ component perpendicular to the reconstructed Z direction) in the inclusive (left) and high p_T (right) $Z \to \mu\mu$ samples. Color coding of the histograms is same as that of Fig. 12.

direction). Based on separate studies performed using muon events (see e.g., Ref [11]), we choose to use the energy deposited in the CaloTowers. Applying the `TrackAssociator` tool to $Z \to \mu\mu$ samples, we can identify the energy in the towers for each reconstructed muon. Projecting the tower energy onto the muon track direction, one obtain a 2D vector of the tower energy, which is then further propagated to the $E\!\!\!/_T$ 2D vector. The $E\!\!\!/_\parallel$ and $E\!\!\!/_\perp$ distributions after correction for the muon energy deposit in the calorimeter are shown in Figs. 12 and 13, respectively. The muon energy correction does not significantly alter the $E\!\!\!/_T$ distributions, as one would expect.

The last step is the correction for the jet energy scale (JES). This is performed as described in Section 5.2 (MCJet corrections). The results after applying this correction are shown in Fig. 12 (13). After applying the JES correction, the $E\!\!\!/_\parallel$ distribution is centered around zero, demonstrating that the jet energy correction successfully restores the average $E\!\!\!/_T$ scale. The $E\!\!\!/_\perp$ is seen to become somewhat broader, consistent with the results discussed in Section 5.2.

Muon $E\!\!\!/_T$ correction code resides in the `JetMETCorrections/Type1MET` package, which has been used to store Type I jet energy scale correction code for $E\!\!\!/_T$. Two new classes, based on Type I jet correction code, have been added in `Type1MET`:

- `MuonMET` class, which adds muon-corrected $E\!\!\!/_T$ in an event;

- `MuonMETAlgo` class, which defines the algorithm part of the correction. The algorithm first selects good quality muons from the standard muon collection produced in the reconstruction process, and then corrects $E\!\!\!/_T$ for the muon momentum and muon energy deposit in the calorimeters. All muon selection cuts and the switch for the muon energy deposition correction are configurable in `MetMuonCorrections.cff` in the data sub-directory of the `Type1MET` package.

The current implementation of the muon energy deposition correction is based on running `TrackAssociator` in every event at the correction level. For CMSSW releases 1_5_X or later, the energies in both crossed RecHits and a 3×3 area of RecHits can be obtained from reconstructed muon objects. The correction code will be modified in the near future to use the energies attached to the muon objects. A validation of the correction package was performed using a $Z \to \mu\mu$ RelVal sample produced with CMSSW_1_6_0. The results are seen to be in good agreement with the corresponding results obtained with CMSSW_1_3_1.

It may be worth noting that the optimization of muon quality selection for $E\!\!\!/_T$ correction will be further exploited, as uncertainties in the behavior of badly reconstructed high p_T^μ muons may lead to the un-physical long tails in the $E\!\!\!/_T$ distribution. The corrections due to the very low $p_T^\mu < 10$GeV muons are less important because their contribution to the $E\!\!\!/_T$ resolution is smaller than the average measurement error based on the calorimeter response. Finally, preselection cuts are to be developed in order to reduce the "fake" $E\!\!\!/_T$ due to the cosmic muons.

5.4 Electron corrections

Corrections due to electrons are expected to be small, due to the excellent energy resolution and coverage of the electromagnetic calorimeter. Some possible reasons for electron corrections are listed below:

- Electron energy scale is incorrect;

- Mismeasurement of electron energies, due to energy deposits in uninstrumented regions.

We have reproduced some of the electron corrections derived in the old CMS software framework, ORCA [12]. The electron corrections documented in [12] are of the form:

$$\Delta \not{E}_T = -\Sigma(\vec{p}_T^{\,e} - \vec{E}_T^{\,sc}) \quad \text{(Cor. 4 in the figures)}; \tag{5}$$

$$\Delta \not{E}_T = -\Sigma(\vec{p}_T^{\,e} - \vec{E}_T^{\,seed}) \quad \text{(Cor. 5 in the figures)}. \tag{6}$$

and

$$\not{E}_T^{\,final} = \not{E}_T^{\,raw} + c \times \Delta \not{E}_T, \tag{7}$$

where c is a scale factor that minimizes the width of the $\not{E}_T$ distribution. In Eq. (5), $\vec{E}_T^{\,sc}$ refers to a vector pointing from the beam spot to the supercluster position whose magnitude is given by the energy of the supercluster. The momentum $\vec{p}_T^{\,e}$ is measured at the vertex. In this correction, we are exploring whether the track momentum is a better estimate of the total energy of the electron than the supercluster.

In Eq. (6), $\vec{E}_T^{\,seed}$ refers to a vector pointing from the beam spot to the seed cluster position whose magnitude is given by the energy of the seed. The momentum $\vec{p}_T^{\,e}$ here is measured at the face of the calorimeter. Here, we are exploring whether the momentum at the face of the calorimeter is a better measure of the energy when it enters the calorimeter than the seed cluster, which does not include any energy that the electron has radiated.
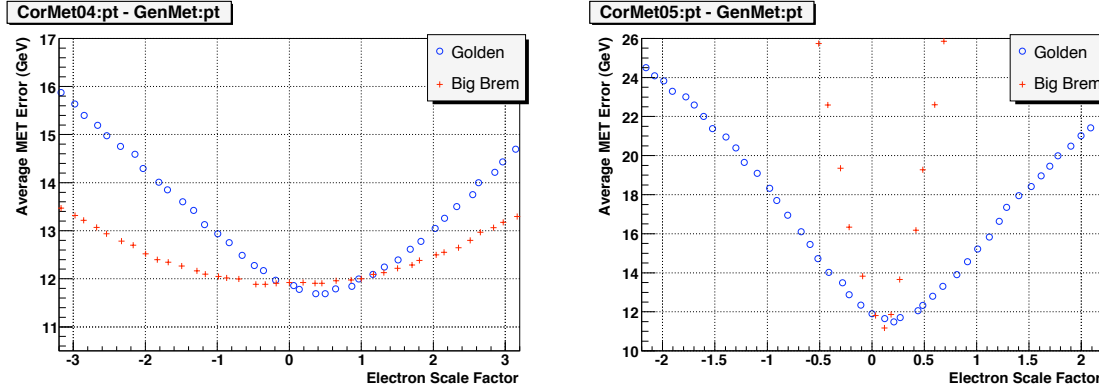

Figure 14: Type 4 (left) and Type 5 (right) corrections for $Z \rightarrow e^+e^-$ events, i.e., true $\not{E}_T = 0$. The y-axis shows the mean of the difference between the generated and reconstructed $\not{E}_T$ value. The x-axis shows the value of the scale factor c shown in Eq. (7). The best value of c for CorMet04 (CorMetO5) for these events is $c = 0.4$ ($c = 0.15$).

In Figs. 14, 15, we explore the effect of these corrections as a function of the scale factor c. As a measure of the effectiveness, we look at the mean difference between the reconstructed $\not{E}_T$ and the generated $\not{E}_T$. The optimal value of the correction corresponds to the point where the mean difference is a minimum or when it is close to an average error of zero, as this is a signed quantity. We can see how much the correction improves the resolution by looking at the difference between no correction ($c = 0$) and the minimum.

In the Figs. 14, we look at $Z \rightarrow e^+e^-$, i.e., cases where $\not{E}_T = 0$. In the Figs. 15, we look at $W \rightarrow e\nu$, i.e., cases where $\not{E}_T \neq 0$. We consider the "Golden" and "Big brem" classes of electrons separately. Electrons are classified, in part, by their bremsstrahlung fraction: $f_{brem} = (p_{in} - p_{out})/p_{in}$, where p_{in} and p_{out} are the magnitude of the track momenta at the vertex and the outermost state, respectively [13]. Golden electrons are the least affected by radiation emission, and they have a measured f_{brem} below 0.2. Big brem electrons have a large bremsstrahlung fraction with f_{brem} above 0.5, though they still have a good matching between the reconstructed supercluster energy and the inner track momentum ($0.9 < E_{SC}/p_{in} < 1.1$). Electrons with f_{brem} between 0.2 and 0.5 fall into the "Narrow" class. These electrons were not considered in this study. We see that for the golden and big brem electrons, there is very little difference between the mean differences in events with no $\not{E}_T$. Similarly, in the second row, for events with non-zero $\not{E}_T$, we see that the difference between no corrections ($c = 0$) and the minimum ($c \neq 0$) is small. More investigation is required to see if the same scale factor c can be used for every sample.
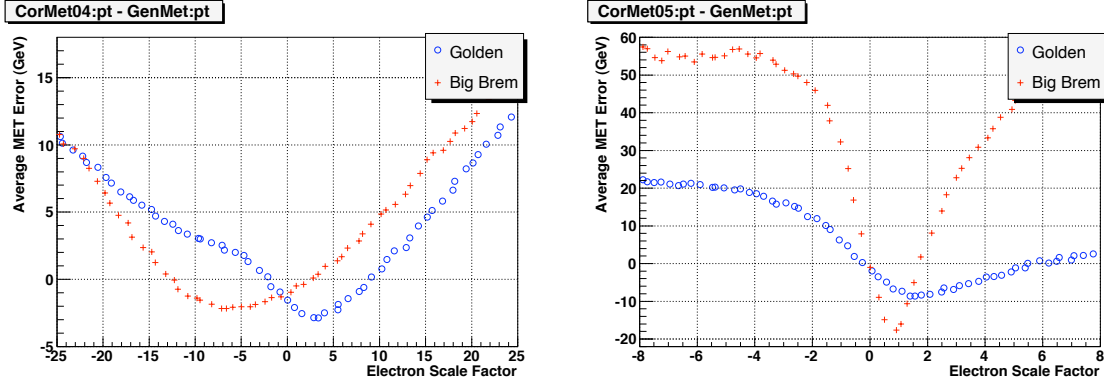
Figure 15: Type 4 (left) and Type 5 (right) corrections for $W \to e\nu$ events, i.e., true $\not{E}_T \neq 0$. The y-axis shows the mean of the difference between the generated and reconstructed $\not{E}_T$ value. The x-axis shows the value of the scale factor c shown in Eq. (7). The best value of c for CorMet04 is $c = -2$ for Golden electrons and $c = 3$ for Big brem electrons. For CorMet05, the best value of c is $c = -0.4$ for Golden electrons and $c = -0.1$ for Big brem electrons.

The electron corrections in the CMSSW have not been released officially, as the purpose of this exercise was mainly to compare the results based on the new software framework with the ORCA studies. In the future we plan to develop less ad hoc $\not{E}_T$ corrections for electrons and will release them at a later time. Nevertheless, it is clear from the present study that the corrections are small and won't affect most of the analyses.

5.5 Tau corrections

The reason to correct $\not{E}_T$ for hadronically decaying taus is the same as for ordinary QCD jets. At a minimum, the corrections remove biases associated with the non-compensating nature of the CMS calorimeter. To first order, the desired correction is given by the difference between the energy deposited by a hadronic tau jet in the calorimeter and the best estimate of its true visible energy. Despite their similarities, taus are substantially different from ordinary jets, making standard jet corrections inappropriate for taus. These differences arise from the fact that hadronic taus typically have a small number of fairly energetic particles while QCD jets of the same energy have higher multiplicity and a larger fraction of energy carried by soft particles. Applying standard jet corrections to hadronic tau jets results in a significant over-correction of $\not{E}_T$. The is illustrated in Fig. 16 (left) which shows the distribution of $(E_T^{\mathrm{reco}} - E_T^{\mathrm{true}})/E_T^{\mathrm{true}}$ for visible hadronic tau energy in $W \to \tau_{\mathrm{had}}\nu$ events for the cases without JES corrections (dashed line) and when standard MCJet corrections are applied (dash-dotted line). Neither case yields a satisfactory description of the jet p_T.

We define the $\not{E}_T$ correction procedure for hadronic tau jets, which satisfy certain identification requirements. To minimize biases, we use loose selections criteria to identify taus. The most powerful and discriminant against QCD jets is the narrow widths of the tau jet. We require that all tau decay products be contained within a signal cone of 10 degrees, and that no tracks with $p_T > 1$ GeV/c appear between the signal cone and a cone of approximately 30 degrees centered on the signal cone. The $\not{E}_T$ correction is calculated as follows:

$$\Delta \vec{\not{E}}_T = \sum_{reg} \vec{E}_T^{\mathrm{cal}} - \vec{E}_T^{\tau} - \sum_{reg} \vec{E}_T^{\mathrm{UE}} - \sum_{\mathrm{reg}} \vec{E}_T^{\mathrm{PU}},$$

where the sum is over a region (reg) of the CaloTowers that fully contains the energy deposits of the tau decay products, typically corresponding to a cone of $\Delta R = 0.5$. The first term sums up energy deposition in the selected region of the calorimeter (with the same thresholds used in the calculation of raw $\not{E}_T$), and the second term is the true visible tau transverse energy. The third and fourth terms sum up true transverse energies of particles from underlying event and pile-up in the same calorimeter region. The latter is required to avoid the double counting of energy (energy depositions of these particles are counted towards raw $\not{E}_T$, but not in the tau energy). This formulation of the correction, however, implies that a correction for the underlying event (UE) and pile-up (PU) is applied everywhere in the detector, which is not typical for most analyses. Therefore, we modify the correction to include only residual UE and PE terms:

$$\Delta \vec{\not{E}}_T = \sum \vec{E}_T^{\mathrm{cal}} - \vec{E}_T^{\tau} - \vec{E}_T^{\mathrm{UE\ res}} - \vec{E}_T^{\mathrm{PU\ res}}, \tag{8}$$

16

Residual PU and UE energy is defined as energy measured in the calorimeter and (i) accounted for in the raw $\not{E}_T$ calculation, (ii) contributing to the first term, and (iii) not contributing to the second term.

The first term in Eq. (8) can be approximated by the transverse energy of a jet obtained using standard cone-based jet clustering algorithm with the cone size of 0.5 (modulo possible difference in energy thresholds used in calculating raw $\not{E}_T$ and in the jet energy summation). The second term should be approximated by the best available measurement of the tau energy, and is discussed below. The last two terms can be parameterized as a function of the number of primary vertices and measured in e.g. $Z \to ee$ events by looking in the direction away from the two leptons. However, these corrections will be small for early running conditions and are neglected here. Also, higher tower energy thresholds used in jet reconstruction further diminish the size of the residual correction due to UE and PU.

In selecting a specific technique to determine the tau energy, one has a choice between the particle flow (PF) measurement and the calorimeter measurement with tau specific corrections. The PF algorithm provides a very accurate measurement of the tau energy except at very high energies. The solid line in Fig. 16 (left) shows the PF tau energy resolution for a sample of $W \to \tau\nu$ events. However, the PF algorithm is not currently available in the trigger, making it important to develop and understand corrections based on both algorithms. The calorimeter based corrections were studied earlier and were found to improve the $\not{E}_T$ scale, but not the resolution [14]. Therefore, we will focus on the algorithm utilizing PF based tau energy measurement.
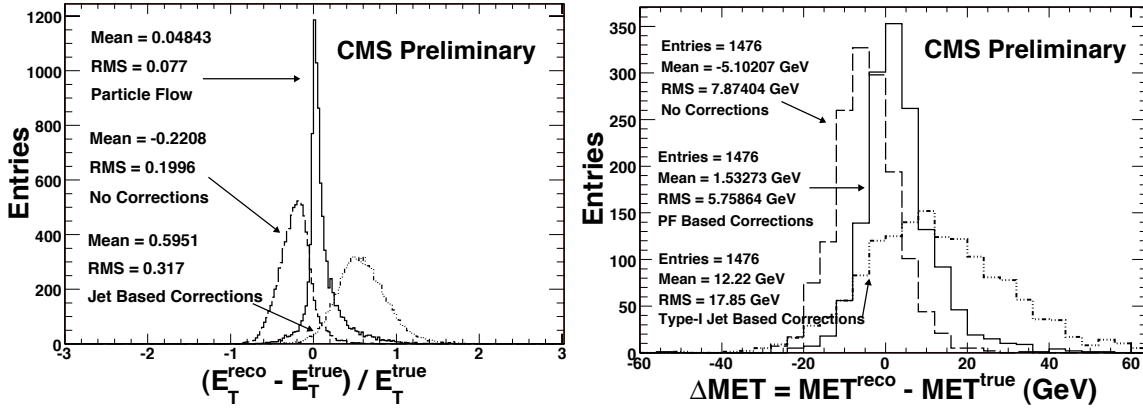

Figure 16: Left: Distribution of $(E_T^{\text{reco}} - E_T^{\text{true}})/E_T^{\text{true}}$ for hadronically decaying tau leptons in $W(+0jets) \to \tau_{\text{had}}\nu$ events for three cases: no correction (dashed line), standard jet correction (dash-dotted) and PF-based (solid) correction. Right: Distribution of $(\not{E}_T^{\text{reco}} - \not{E}_T^{\text{true}})$ in the same events. Notations for the lines is the same.

We define a minimal PF based $\not{E}_T$ correction for τ's as:

$$\Delta \vec{\not{E}}_T = \sum \vec{E}_T^{\text{cal jet 0.5}} - \vec{E}_T^{\text{PF } \tau},$$

To gauge performance of this correction and disentangle effects associated with other corrections (e.g. for jets, electrons and muons), we select a sample of $W \to \tau\nu$ events and require that there be no additional jets with $E_T > 5$ GeV. While this is a harsh requirement, it is effective in removing other effects associated with mismeasurements of recoil jet energies. Figure 16 (right) shows the $(\not{E}_T^{\text{reco}} - \not{E}_T^{\text{true}})$ distribution for three cases: no correction (dashed line), standard jet correction (dash-dotted) and PF-based (solid) correction. The PF based calculation is seen to yield the result with the smallest bias and the best resolution. The absence of visible bias confirms that effects of the UE are small. The residual efffects of the UE will nonetheless be studied and approperiate corrections will soon be available.

We also would like to discuss the limits of applicability of the dedicated tau corrections. First, dedicated tau corrections improve the $\not{E}_T$ scale and resolution only in events that have real taus, and then only if these taus are identified as such. Applying these corrections to events where a QCD jet is misidentified as a tau will not lead to better $\not{E}_T$ resolution or scale. Therefore, in the vast majority of physics analyses that are not specifically targeting events with real taus, one should not use these corrections merely because a jet was tagged as a tau candidate. Second, even in analyses that target final states with hadronic taus, one will only improve $\not{E}_T$ in signal events and not in the backgrounds where a jet is misidentified as a hadronic tau. However, applying tau corrections to $\not{E}_T$ is appropriate in the second case as it improves separation between signal and backgrounds. It is also worthwhile mentioning that applying standard jet corrections to jets misidentified as taus will not be generally correct either

because tau tagging procedures select rather peculiar jets, while standard jet corrections are derived from average jets. Finally, the correction procedure we describe is defined only for hadronically decaying taus that satisfy certain identification requirements. If one uses a different identification procedure, these corrections may need to be recalculated. Fortunately, this will only become an issue if the tau tagging method is substantially different from what was used in obtaining these corrections.

The current implementation in CMSSW is based on the code (`TauMET.cc`) that calculates the $\not{\!\!E}_T$ correction for each calorimeter jet. The default use assumes as input CaloJets and PFJets and produces output a METCollection of LorentzVectors with the same indices as PFJets. Since tau tagging methods vary for each analysis, the above functionality is powerful because it allows the user to choose only a particular jet to correct the $\not{\!\!E}_T$. Furthermore, it is important that the user be able to calculate $\not{\!\!E}_T$ corrections starting from raw $\not{\!\!E}_T$ or $\not{\!\!E}_T$- corrected for jets using standard Type-I corrections (`Type1MET.cc`). Therefore, the user can choose to calculate the $\not{\!\!E}_T$ corrections using raw calorimeter jets or corrected calorimeter jets. The default sequence produces the $\not{\!\!E}_T$ corrections using both CaloJetCollections (`JetMETCorrections/Type1MET/data/TauMetCorrections.cff`). An example of running the TauMET code can be found in:

`JetMETCorrections/Type1MET/test/testTauMetCorr.cfg.`

An example is also provided:

`JetMETCorrections/Type1MET/test/TauMETAnalyzer.cc,`
`JetMETCorrections/Type1MET/test/TauMETAnalyzer.cfg`

that shows how one can use the code to correct $\not{\!\!E}_T$. In CMSSW_1_6_0 and subsequent versions, one needs to checkout `JetMETCorrections/Type1MET` and add the producer to the path to create the collection.

6 Detector Effects on $\not{\!\!E}_T$

The $\not{\!\!E}_T$ calculation is one of the simplest that will be performed with CMS. However, the $\not{\!\!E}_T$ measurement is very complex and depends on other objects. Many sources may contribute to the $\not{\!\!E}_T$ measurements. Among them are detector failures. In this section we study the effect of certain types of detector failures on $\not{\!\!E}_T$.

Three event samples were used: SUSY mSUGRA test point LM1 (see Ref. [2] for details) events, QCD events with $0 < \hat{p_T} < 3000$ GeV, and $t\bar{t}$+3 jets. The CMSSW version used was _1_3_x. No event selection other than that used to generate the samples was applied.

We used the `CaloMiscalibTool` package provided in `CMSSW` to study readout failures. In general this package provides a mechanism to produce a new `RecHit` collection and new high-level objects such as $\not{\!\!E}_T$ and jets by using an external `XML` file and this new `RecHit` collection. The `XML` file, which contains DetId information and corresponding gain, is created by the user and placed in a configuration file for the `CMSSW` run. We considered the following cases for readout failures:

- 1%, 3% and 5% of the ECAL and HCAL readout channels in the barrel region are disabled, *i.e.* the output signal is zeroed. The channels are selected randomly.

- 1%, 3% and 5% of the ECAL and HCAL readout channels in the barrel region are radically enhanced, *i.e.* the output signal is multiplied by a large factor (5). The channels are selected randomly.

- Disabled and enhanced channels are mixed together (half and half) to give 1%, 3% and 5% of the ECAL and HCAL readout channels in the barrel region. The channels are selected randomly.

- The readout modules (RM) in $\phi = 3, 4, 5$ in HB and the geometrically corresponding crystals in EB are disabled or the signal output is enhanced. (One RM reads eighteen channels in one HCAL ϕ segment and there are twenty five ECAL crystals in front of one HCAL tower.) Therefore the number of channels in one ϕ segment corresponds to about 1.4% of the total. One RM at a time (only $\phi = 3$), two RMs at a time ($\phi = 3, 4$) and three RMs at a time ($\phi = 3, 4, 5$) cases are applied. The difference of this case from the previous two is that readout problems are grouped into a readout region.

The number of channels which are disabled or enhanced in each case are shown in table 3. The $\not{\!\!E}_T$ is reconstructed from the vector sum over all towers (CaloTowers) and no corrections are applied. The following cases are studied:

- The signal output from some of the ECAL and HCAL channels are zeroed: Fig. 17 and table 4.

Table 3: The number of channels for each case considered. The channels are selected randomly.

Case	ECAL	HCAL
1%	600	20
3%	1821	72
5%	3043	118
$\phi = 3$	400	18
$\phi = 3, 4$	800	36
$\phi = 3, 4, 5$	1200	54

- The signal output from some of the ECAL and HCAL channels are enhanced by factor of five: Fig. 18 and table 5.

- The signal output from some of the ECAL and HCAL channels are zeroed and enhanced by factor of five at the same time to have a mixture of the two cases: Fig. 19 and table 6.

- The signal output from some of the HCAL readout modules and geometrically corresponding ECAL crystals are zeroed: Fig. 20 and table 7.

- The signal output from some of the HCAL readout modules and geometrically corresponding ECAL crystals are enhanced by factor of five: Fig. 21 and table 8.

The $\not{E}_T$ is calculated both before and after the introduction of the detector effects. The results are shown in the figures and the shift in the mean and RMS values are given in the tables below.
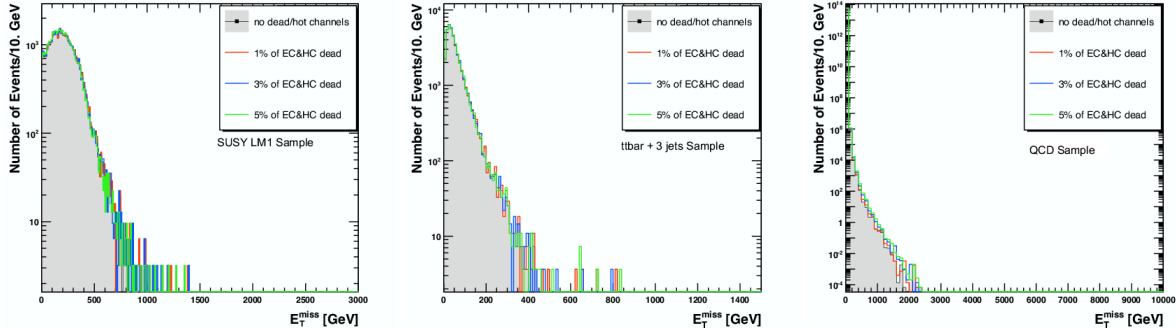
Figure 17: $\not{E}_T$ distributions for LM1 (left), $t\bar{t}$+3 jets (middle), and QCD (right) events. 1% (red), 3% (blue) and 5% (green) of the ECAL and HCAL readout channels are dead. Shaded histograms are the reference distributions without detector problems. The contribution from dead channels is seen as increase in the number of events in the tail.

Table 4: The mean and RMS values in GeV after dead channels are applied.

Sample/Case	nominal	1% dead	3% dead	5% dead
LM1	209.7 (128.6)	208.2 (127.9)	205.3 (126.1)	202.4 (124.2)
$t\bar{t} + 3$ jets	53.9 (48.5)	53.9 (48.3)	53.7 (47.4)	53.8 (47.8)
QCD	3.76 (3.17)	3.77 (3.18)	3.76 (3.18)	3.75 (3.17)

$\not{E}_T$ is calculated from CaloTowers without a requirement on E_T and η suppression. Therefore, the number of CaloTowers is determined by the tower thresholds used in `CaloTowersCreator` and by the event topology. Since the tower thresholds are larger than zero, the towers with `RecHit` energy equal to zero are not entered into the `CaloTowersCreator` algorithm. The imbalance in the total calorimeter transverse energy and therefore the shift in $\not{E}_T$ in the dead channels case is produced when the tower actually registers signal but it is killed by the user. For the hot channels the situation is different. A `RecHit` with an energy above threshold is always used in the tower creator algorithm. If it coincides with the channel in the user `XML` file its energy is enhanced before it enters the algorithm and more signal is seen in the detector.
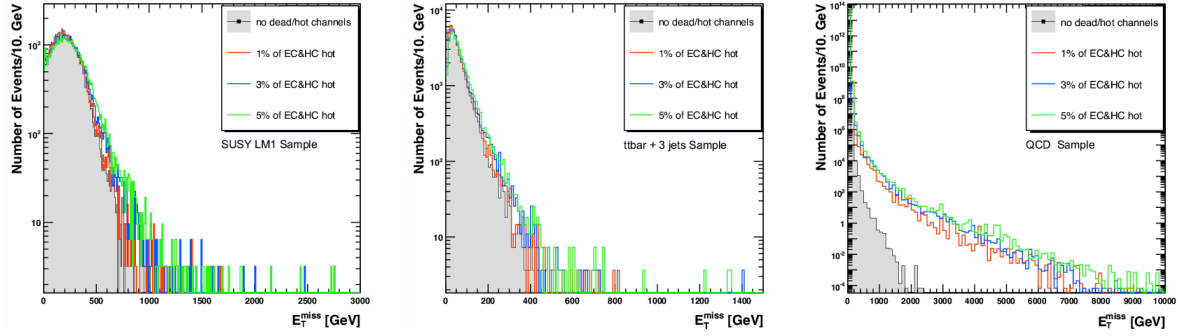
Figure 18: E_T distributions for LM1 (left), $t\bar{t}$+3 jets (middle), and QCD (right) events. 1% (red), 3% (blue) and 5% (green) of the ECAL and HCAL readout channels are hot. Shaded histograms are the reference distributions without detector problems. The contribution from dead channels is seen as increase in the number of events in the tail.

Table 5: The mean and RMS values in GeV after hot channels are applied.

Sample/Case	nominal	1% hot	3% hot	5% hot
LM1	209.7 (128.6)	219.1 (139.3)	235.7 (156.4)	249.8 (172.6)
$t\bar{t}$ + 3 jets	53.9 (48.5)	57.8 (53.1)	63.7 (59.5)	69.2 (65.5)
QCD	3.76 (3.17)	5.37 (3.58)	7.54 (4.38)	8.03 (4.73)

In the case of the dead channels, a small increase in the number of events in the tail is observed. The shift in the mean value of E_T is negligible. In case of hot channels, an increase in the number of events in the tail is clearly observed. The larger the value of E_T, the larger the number of events in the tail. Since the enhancement factor of five used to define hot channels is quite large, QCD events, with large E_T are affected more than LM1 and $t\bar{t}$ events. In the case of the mixture of hot and dead channels, as one expects for a real detector, the E_T behavior is seen to lie between the two cases of only dead and only hot channel cases.

A more realistic case is a readout module failure where an entire ϕ segment is not read out or is noisy. Such a case is simulated in this study by using dead or hot readout modules. In reality a dead module would likely be spotted quickly and necessary actions will be taken. A hot module may be difficult to identify since noise levels will not be uniform between modules and there may be correlation between the noise levels of different modules. In this study these correlations are not taken into account and the noise level of all units (Hybrid Photo-Diodes in this case) is taken to be the same. In case of dead readout modules, a small increase in the number of tail events are observed. The shift in mean value of E_T is negligible although it slightly decreases with the increasing number of dead readout modules for LM1 sample, while there is a small increase for $t\bar{t}$+3 jets and QCD samples. It is also seen that an increase in the number of dead readout modules gives rise in absolute shift. An increase in the number
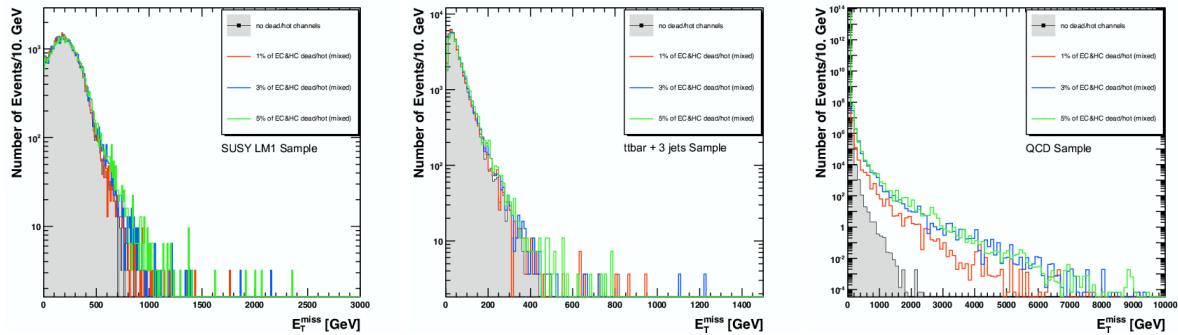

Figure 19: E_T distributions for LM1 (left), $t\bar{t}$+3 jets (middle), and QCD (right) events. 1% (red), 3% (blue) and 5% (green) of the ECAL and HCAL readout channels are mixed with hot and dead channels. Shaded histograms are the reference distributions without detector problems. The contribution from these channels is seen as increase in the number of events in the tail.

20

Table 6: The mean and RMS values in GeV after mixture of hot and dead channels are applied.

Sample/Case	nominal	1% mixed	3% mixed	5% mixed
LM1	209.7 (128.6)	212.9 (128.6)	220.3 (140.5)	227.7 (151.4)
$t\bar{t}$ + 3 jets	53.9 (48.5)	55.6 (51.4)	58.6 (53.6)	61.5 (56.4)
QCD	3.76 (3.17)	4.30 (3.27)	5.99 (3.75)	7.13 (4.18)
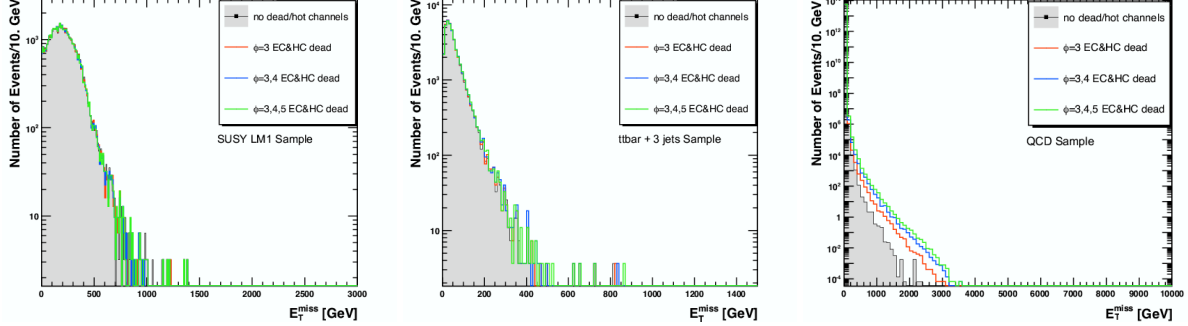
Figure 20: $E\!\!\!/_T$ distributions for LM1 (left), $t\bar{t}$+3 jets (middle), and QCD (right) events. $\phi = 3$ (red),$\phi = 3, 4$ (blue) and $\phi = 3, 4, 5$ (green) of HCAL readout modules and corresponding ECAL crystals are made dead. Shaded histograms are the reference distributions without detector problems. The contribution from these channels is seen as increase in the number of events in the tail.

of events in the tail is observed very clearly for the hot readout modules. One observes that larger $E\!\!\!/_T$ values lead to higher number of events in the tail. The shift in mean value of $E\!\!\!/_T$ becomes visible and increases with the increasing number of hot modules.

To show the effect of dead and hot channels on the tails, events are selected with more than 200 GeV $E\!\!\!/_T$ for all samples, and the ratio of number of events passing this requirement to the total number of events is shown. The results are presented in table 9.

7 $E\!\!\!/_T$ Cleanup Using Event Electromagnetic and Charged Fractions

$E\!\!\!/_T$ in an event can receive contributions from, e.g. cosmic muons or from machine or detector related background such as beam halo or electronic noise. A cleanup procedure on the event level for these types of background was developed in [15] in the old software framework, ORCA, using the event electromagnetic and charged fractions.
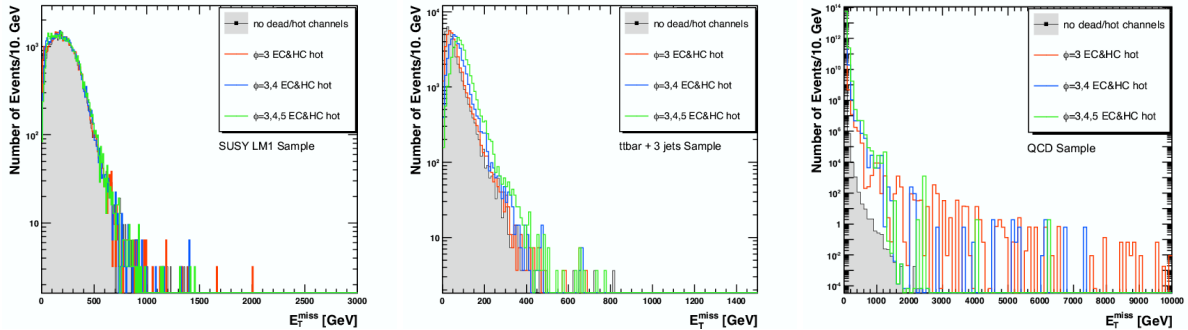

Figure 21: $E\!\!\!/_T$ distributions for LM1 (left), $t\bar{t}$+3 jets (middle), and QCD (right) events. $\phi = 3$ (red),$\phi = 3, 4$ (blue) and $\phi = 3, 4, 5$ (green) of HCAL readout modules and corresponding ECAL crystals are made hot. Shaded histograms are the reference distributions without detector problems. The contribution from these channels is seen as increase in the number of events in the tail.

Table 7: The mean and RMS values in GeV after dead ϕ segments are applied.

Sample/Case	nominal	$\phi = 3$ dead	$\phi = 3, 4$ dead	$\phi = 3, 4, 5$ dead
LM1	209.7 (128.6)	208.9 (128.2)	208.5 (128.0)	208.3 (128.1)
$t\bar{t} + 3$ jets	53.9 (48.5)	54.2 (48.5)	54.9 (49.4)	55.7 (50.3)
QCD	3.76 (3.17)	3.77 (3.18)	3.77 (3.19)	3.78 (3.19)

Table 8: The mean and RMS values in GeV after hot ϕ segments are applied.

Sample/Case	nominal	$\phi = 3$ hot	$\phi = 3, 4$ hot	$\phi = 3, 4, 5$ hot
LM1	209.7 (128.6)	209.5 (128.5)	211.4 (126.8)	216.1 (127.8)
$t\bar{t} + 3$ jets	53.9 (48.5)	61.3 (49.1)	76.8 (50.7)	96.3 (55.5)
QCD	3.76 (3.17)	21.3 (6.60)	42.0 (8.86)	62.7 (10.7)

The event electromagnetic fraction is defined as

$$F_{\text{EM}} = \frac{\sum_{j=1}^{N_{\text{jet}}} p_T^j \times f_{\text{EM}}^j}{\sum_{j=1}^{N_{\text{jet}}} p_T^j},$$

and is calculated for jets with $p_T > 30$ GeV within the central calorimeter acceptance $|\eta| \leq 3$ and with f_{EM}^j the jet electromagnetic fraction. Jets in normal collision events are expected to have an electromagnetic fraction between 0 and 1, but not close to the edges. Energy depositions from cosmics or beam halo in either the hadronic or electromagnetic calorimeter, when clustered as jets, will have f_{EM} close to 0 or 1. Electrons and photons that are clustered as jets will have f_{EM} close to 1.

The event charged fraction is

$$F_{\text{ch}} = \left\langle f_{\text{ch}}^j \right\rangle \bigg|_{N_{\text{jet}}} = \left\langle \frac{(\sum_{i}^{tracks} p_{Ti})^j}{p_T^j} \right\rangle \bigg|_{N_{\text{jet}}},$$

Table 9: The fraction of events passing $\not{E}_T > 200$ GeV requirement in each sample. Individual cross-section values are taken into account for 20 QCD bins.

Sample/Case	nominal	1% dead	2% dead	3% dead
LM1	0.49	0.46	0.48	0.47
$t\bar{t} + 3$ jets	0.018	0.018	0.017	0.017
QCD	7.0×10^{-6}	7.9×10^{-6}	1.0×10^{-5}	1.2×10^{-5}
	nominal	1% hot	2% hot	3% hot
LM1	0.49	0.51	0.54	0.57
$t\bar{t} + 3$ jets	0.018	0.024	0.034	0.040
QCD	7.0×10^{-6}	3.6×10^{-4}	1.6×10^{-3}	2.6×10^{-3}
	nominal	1% mixed	2% mixed	3% mixed
LM1	0.49	0.49	0.51	0.52
$t\bar{t} + 3$ jets	0.018	0.020	0.024	0.029
QCD	7.0×10^{-6}	9.9×10^{-5}	1.5×10^{-3}	1.6×10^{-3}
	nominal	ϕ=3 dead	ϕ=3,4 dead	ϕ=3,4,5 dead
LM1	0.49	0.48	0.48	0.48
$t\bar{t} + 3$ jets	0.018	0.018	0.019	0.020
QCD	7.0×10^{-6}	1.6×10^{-5}	5.3×10^{-5}	1.1×10^{-4}
	nominal	ϕ=3 hot	ϕ=3,4 hot	ϕ=3,4,5 hot
LM1	0.49	0.48	0.49	0.50
$t\bar{t} + 3$ jets	0.018	0.021	0.029	0.047
QCD	7.0×10^{-6}	9.8×10^{-3}	7.5×10^{-2}	0.11

where the sum for every jet runs over the particle tracks that can be matched in (η, ϕ) space to that jet and where the average over the jet charged fraction f_{jet}^{j} is taken over jets which have at least four associated tracks. For the jet-track matching, the primary vertex of the event is determined as the reconstructed vertex with the highest associated track $\sum p_T$ [2]. The tracks associated to that vertex are required to satisfy $p_T > 1.2$ GeV, $N_{\text{hits}} \geq 5$, transverse impact parameter $|d_0| \leq 600\,\mu\text{m}$, $|z_{\text{pv}} - z_{\text{track}}| < 1$ mm and $|\eta_{\text{track}}| < 2.4$. Only jets with $p_T > 30$ GeV and $|\eta'| < 1.7$ are considered, where η' is the recalculated jet pseudorapidity using the primary vertex. Here, jets reconstructed using the Iterative Cone algorithm are taken with jet cone size of 0.5, while for the jet-track matching $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} < 0.75$ is used. Background particles that do not originate from the collision vertex may pass through the tracker, but the produced tracks are not expected to point to jets in the event nor to be associated with the primary vertex. Note that the requirement on the minimum number of tracks in a jet (four), as well as the $\Delta R < 0.75$ cut have not been thoroughly optimized yet; the same values as used in previous study have been used in this note for the purpose of comparison with ORCA results. In the future these requirements will be optimized.
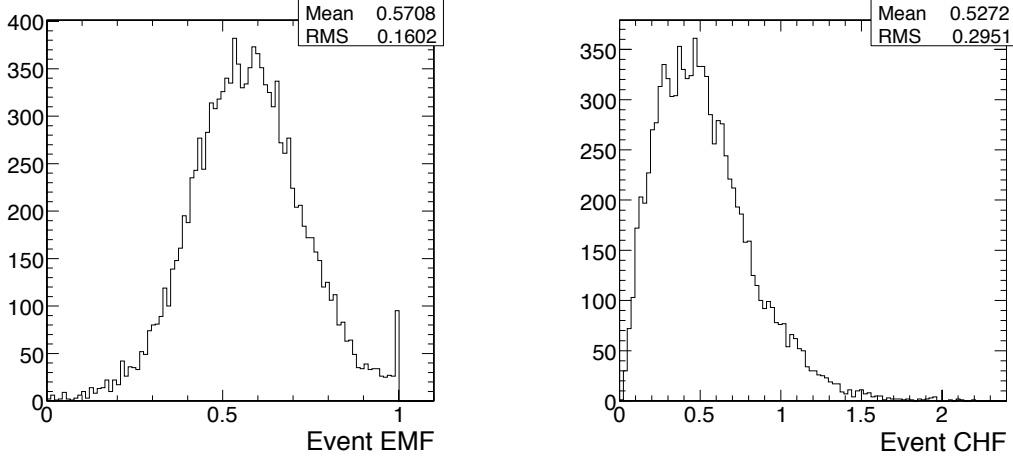
Figure 22: F_{EM} (left) and F_{ch} (right) for $t\bar{t}$ production.

The computation of the jet and event charged and electromagnetic fraction have been ported to the CMSSW framework. For illustration, Fig. 22 the F_{EM} and F_{ch} shows for a $t\bar{t}$ sample (CMSSW_1_3_1). An example of a selection criterium given in [15] selects events with at least one reconstructed vertex, at least one jet with $|\eta| < 1.7$ and $N_{\text{tracks}} \geq 4$ and with $F_{\text{ch}} \geq 0.175$ and $F_{\text{EM}} \geq 0.1$. An event filter with such selection is presently implemented and can be used as part of an event Skim Filter. Validation of this filter with cosmic and beam halo Monte Carlo events or measured cosmics data will be required to fine-tune the exact definition of F_{EM} and F_{ch} and to determine the optimal event selection criteria.

8   Optimization of $\not{\!\!E}_T$ Resolution

In this section, we present a study of the CMS resolution in missing transverse energy $\not{\!\!E}_T$ as a function of cell energy thresholds in the hadron calorimeter (HCAL). Previous optimization studies of HCAL thresholds have been performed in the context of jet reconstruction [4, 16]. The optimal HCAL thresholds for $\not{\!\!E}_T$ are not necessarily the same as those for jets, however.

At the event generator level, E_i, θ_i and ϕ_i in eq.1 for the $\not{\!\!E}_T$ calculation specify the energy and direction of particle i, and the sum is restricted to all generator level particles with lifetimes greater than 10^{-10} s except that muons and stable non-interacting particles such as neutrinos and the SUSY LSP are excluded. In the following, $\vec{\not{\!\!E}}_T$ defined at the reconstruction (generator) level is denoted $\vec{\not{\!\!E}}_T^{\text{ rec}}$ ($\vec{\not{\!\!E}}_T^{\text{ gen}}$).

The resolution in $\not{\!\!E}_T$ is evaluated by examining Monte Carlo (MC) events at the generator and reconstruction levels. For each event, we calculate the component of $\vec{\not{\!\!E}}_T^{\text{ rec}}$ in the direction of $\vec{\not{\!\!E}}_T^{\text{ gen}}$:

$$\not{\!\!E}_{\parallel}^{\text{rec}} \equiv \vec{\not{\!\!E}}_T^{\text{ rec}} \cdot \left(\frac{\vec{\not{\!\!E}}_T^{\text{ gen}}}{\not{\!\!E}_T^{\text{ gen}}} \right),$$

---

[2] This doesn't really matter for the Monte Carlo events without the pile-up used in these studies, but the method is generic enough to be used at any luminosities.
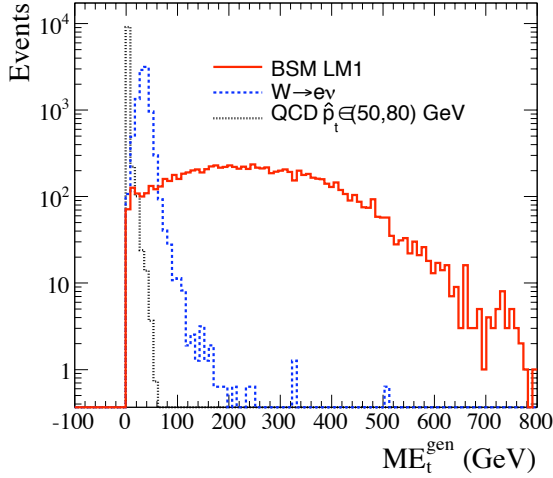
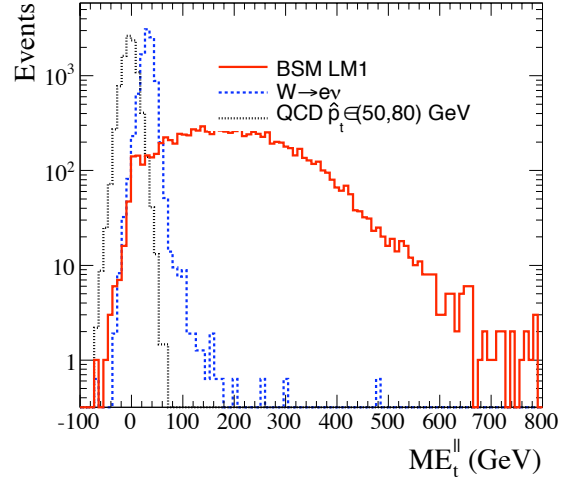Figure 23: Distribution of $\not{E}_T{}^{\text{gen}}$ in QCD, $W + jets$, and SUSY LM1 events.

Figure 24: Distributions of $\not{E}_\parallel{}^{\text{rec}}$ based on the Scheme B [4] thresholds for HB, HO, and HE, with a threshold of 1.2 GeV (1.8 GeV) for the long (short) fibers in HF.

with $\not{E}_T{}^{\text{gen}}$ the magnitude of $\vec{\not{E}}_T{}^{\text{gen}}$. We then determine the difference $\Delta\not{E}_T$:

$$\Delta\not{E}_T = \not{E}_\parallel{}^{\text{rec}} - \not{E}_T{}^{\text{gen}}.$$

Following previous practice, the resolution of $\not{E}_T$ is defined to be the RMS of the $\Delta\not{E}_T$ distribution. The resolution defined in this manner specifies how well $\not{E}_T$ is measured in events with true missing energy. We also examine the component of $\vec{\not{E}}_T{}^{\text{rec}}$ perpendicular to $\vec{\not{E}}_T{}^{\text{gen}}$

$$\not{E}_\perp{}^{\text{rec}} \equiv \vec{\not{E}}_T{}^{\text{rec}} \times \left(\frac{\vec{\not{E}}_T{}^{\text{gen}}}{\not{E}_T{}^{\text{gen}}} \right),$$

which can provide a measure of the $\not{E}_T$ resolution in situations without true missing energy.

We base our study on three MC event samples. QCD dijet events with $50 < \hat{p}_T < 80$ GeV/c provide a sample with essentially no true $\not{E}_T$. Inclusive $W+$ jets events with $W \to e\nu$ provide a sample with $\not{E}_T$ at the electroweak scale. Events with $\not{E}_T$ at BSM scales are represented by SUSY events with the LM1 (see Ref. [2] for detail) set of parameters. The QCD and $W(e\nu)+$ jets samples are produced using CMSSW_1_5_2. In contrast, the SUSY-LM1 sample is produced using CMSSW_1_3_1. We use CMSSW_1_3_1 in this latter case because, at the time of writing, we do not have SUSY events reconstructed with later versions of the software. CMSSW version 1_3_1 contains a less realistic simulation of HCAL noise than version 1_5_2. However, using QCD events, we find that the optimal threshold values in HCAL for the $\not{E}_T$ reconstruction (see below) are about the same in the two CMSSW versions, even though the overall noise level is about 20% larger in 1_5_2 than in 1_3_1. We therefore presume that the optimal thresholds will be about the same for SUSY events as well, a presumption that will be tested in the near future when SUSY events processed with later versions of the reconstruction software become available.

The distributions of $\not{E}_T{}^{\text{gen}}$ for the three samples are presented in Fig. 23. The corresponding distributions of $\not{E}_\parallel{}^{\text{rec}}$ are shown in Fig. 24. Figure 25 presents the $\Delta\not{E}_T$ distributions, i.e., the event-by-event difference between the results in Figs. 23 and 24. Finally, the $\not{E}_\perp{}^{\text{rec}}$ distributions are shown in Fig. 26. For the results in Figs. 24-26, the so-called Scheme B [4] set of HCAL energy thresholds are applied, see Table 1 for details.

The $\not{E}_T$ resolution as a function of the energy threshold in HB cells is shown by the circle symbols in Fig. 27. The squares show the corresponding number of HB cells with energies above the threshold, i.e., the number of HB cells that contribute to the determination of $\not{E}_T$. The error bars show the statistical uncertainties. The top figure shows the results for the QCD events. The middle and bottom figures show the results for the $W+$ jets and SUSY LM1 events. The thresholds of all subdetectors other than HB are fixed to the values corresponding to Figs. 24-26.

The reason for the distributions of the number of cells and $\not{E}_T$ resolution for the SUSY LM1 sample do not look smooth is the fact that exactly the same gain was used for all the calorimeter cells of a particular type in CMSSW_1_3_1, which was the software version for production of the SUSY LM1 sample. Thus, the step-like
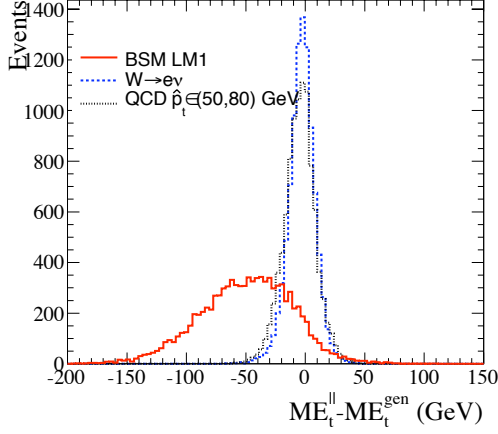
Figure 25: Distributions of $\Delta \not{E}_T$.
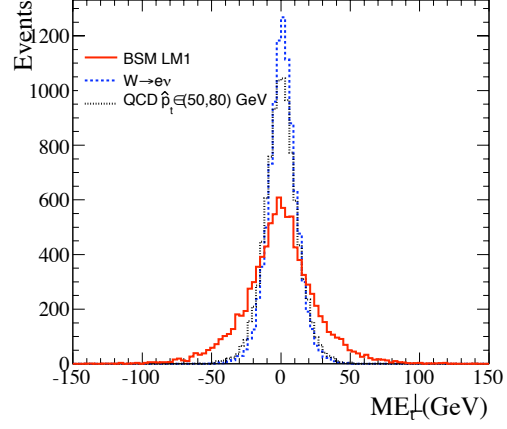

Figure 26: Distributions of $\not{E}_\perp^{\text{rec}}$.

shape of the distributions in this and following figures for this sample reflects crossing particular number of ADC counts. For the other samples produced with more modern version of the CMSSW software channel-by-channel variation of the gain has been introduced, thus resulting in smooth distributions.

For energy thresholds below about 0.2 GeV, it is seen that a large number ($\sim$ 700) of HB cells contribute to $\not{E}_T$. As the threshold is increased to about 0.5 GeV, most of these cells are eliminated and the $\not{E}_T$ resolution improves. For larger values of the threshold, the $\not{E}_T$ resolution becomes worse. These features are observed for all three event samples. Thus the optimal energy threshold value for HB cells is around 0.5 GeV for the purposes of the $\not{E}_T$ calculation. This threshold was also found optimum to use in jet reconstruction, thus making $\not{E}_T$ and jet reconstruction, as well as JES corrections to $\not{E}_T$ self-consistent.

The analogous results for the 5 degree HE cells ($1.305 < \eta < 1.740$) are shown in Fig. 28. The best resolution for the QCD and SUSY LM1 events is obtained with a threshold around 0.6 GeV. The best resolution for the W+jets events occurs for a threshold around 0.8 GeV. Thus a threshold of 0.7 GeV would be a reasonable choice for the HE 5 degree cells.

The results for the 10 degree HE cells ($1.740 < \eta < 3.000$) are shown in Fig. 29. The $W+$ jets and SUSY-LM1 events exhibit similar behavior to that seen for the corresponding samples in Fig. 28, but with optimal energy thresholds near 0.5 GeV. The QCD events show a qualitatively different behavior: the $\not{E}_T$ resolution worsens as the threshold increases from 0 to around 0.5 GeV, and then improves as the threshold increases to large values. Thus, for the QCD events, the $\not{E}_T$ resolution is improved if the HE 10 degree cells are not used at all. Similar results to these are found for the HF long fibers with all three event samples (Fig. 30).

The results for the HF short fibers are very similar to those shown for the long fibers. Thus the best $\not{E}_T$ resolution is obtained if HF is not used at all. Note that this conclusion is based only on processes with central production, which have been studied so far. For forward processes with $\not{E}_T$, e.g. invisible Higgs production via vector boson fusion, the role of HF is important.

As the next step of this optimization study, $\not{E}_T$ resolution as a function of thresholds in transverse energy in various calorimeters will be studied.

9 Conclusions

This note gives an account of the status of the missing transverse energy calculation and studies in CMS. Next are the study of the precision reach of data driven calibration and resolution estimate techniques versus luminosity. Future developements include:

- the study of pathological $\not{E}_T$ events in the Monte Carlo simulation of high-p_T QCD events;

- the study of different Type I corrections, more tailored to $\not{E}_T$ needs;

- the study of optimum muon and tau ID and kinematic cuts for the purpose of $\not{E}_T$ corrections;
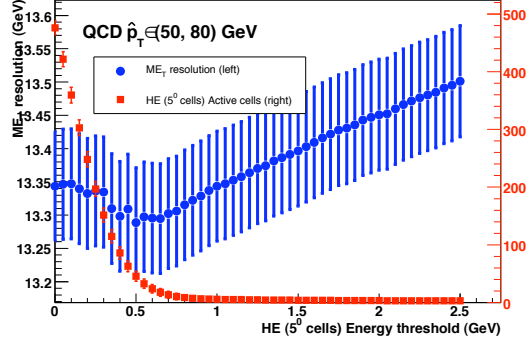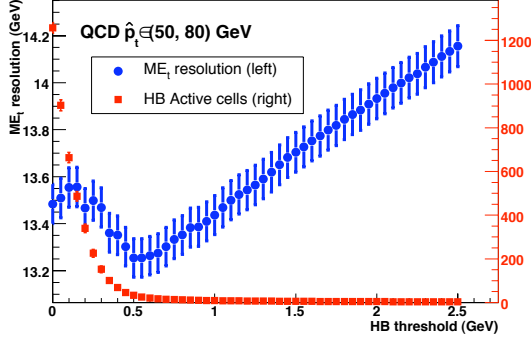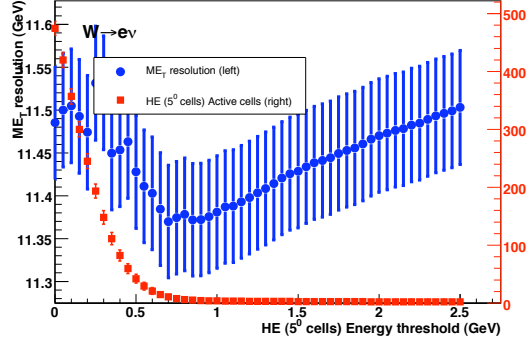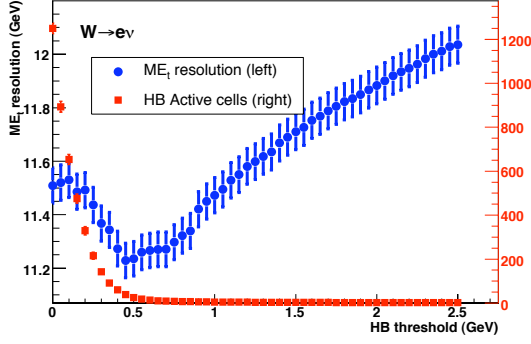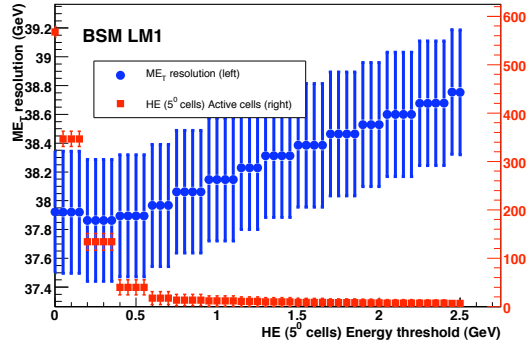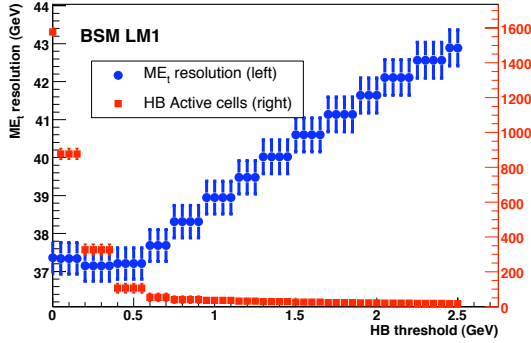
Figure 27: $\not{E}_T$ resolution versus cell energy threshold in HB.

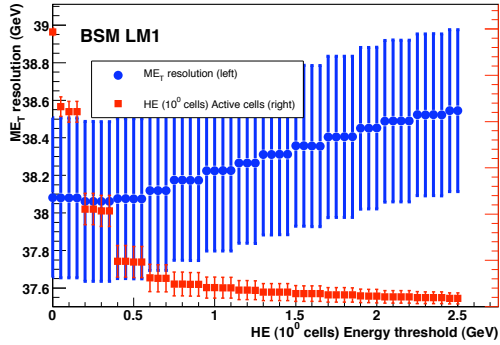Figure 28: $\not{E}_T$ resolution versus energy threshold in the HE 5 degree cells.

- the development and study of Type II corrections for the soft underlying event contribution to $\not{E}_T$;

- the impact of pile-up on $\not{E}_T$ and the effect of mis-estimating $\not{E}_T$ due to the selction fo the wrong primary vertex;

- The usage of (more) particle flow objects to improve the detector resolution and therefore the $\not{E}_T$ resolution;

- the study of the $\not{E}_T$ significance likelihood variable;

- the study of beam halo and cosmics background effects on $\not{E}_T$;

- the study of $\not{E}_T$ with real data during CMS cosmics runs in early 2008, followed by first real collider data in late 2008.

The $\not{E}_T$ variable is not an easy object to understand, so a long and dedicated study will be necessary to turn it into a useful physics analysis variable.
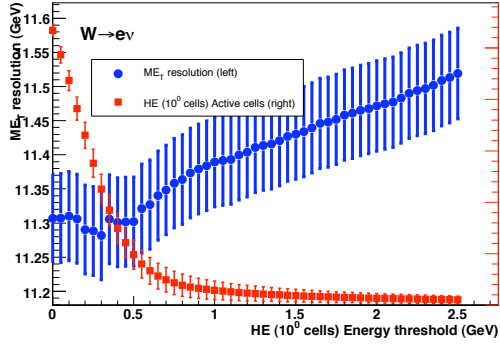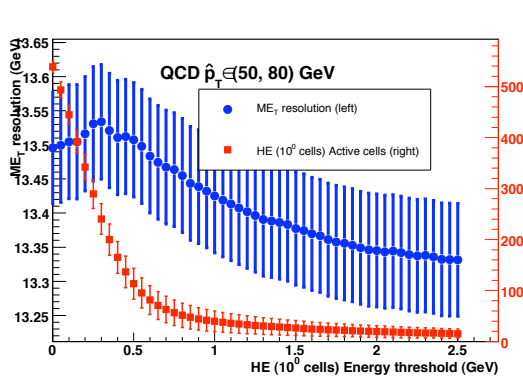
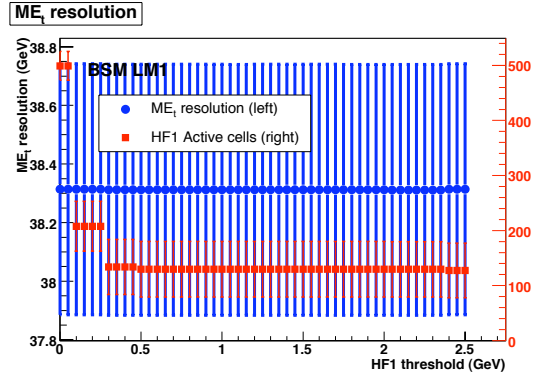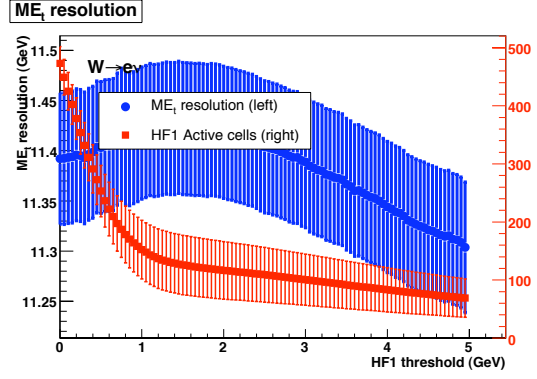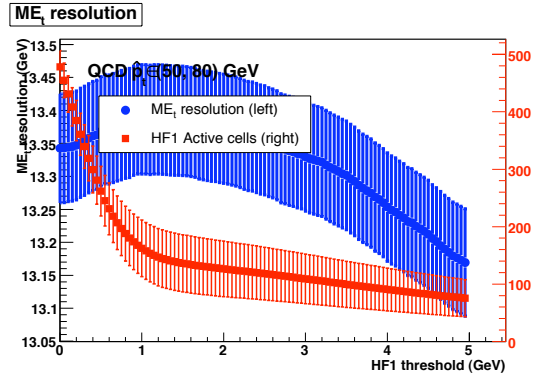Figure 29: $\not{E}_T$ resolution versus energy threshold in the HE 10 degree cells.

Figure 30: $\not{E}_T$ resolution versus energy threshold in the HF long fibers.

References

[1] For a Web page describing the CMS detector, see, e.g., `http://cms.cern.ch/iCMS`.

[2] CMS Collaboration *"CMS Physics Technical Design Report v.2: Physics Performance,"* CERN LHCC/2006–021 (2006); J. Phys. G: Nucl. Part. Phys. **34** (2007) 995.

[3] CMS Collaboration *"CMS Physics Technical Design Report v.1: Detector Performance and Software,"* CERN LHCC/2006–001 (2006).

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

[5] S. Esen *et al.*, **CMS Analysis Note in Preparation**, *"Plans for Jet Energy Corrections at CMS."*

[6] A. Nikitenko, S. Kunori, and R. Kinunnen, **CMS Note 2001/040**, *"Missing Transverse Energy Measurement with Jet Energy corrections."*

[7] M. Vazquez Acosta *et al.*, **CMS Note IN-2007/053**, *"Jet and MET Performance in CMSSW_1_2_0."*

[8] S. Esen and G. Landsberg, **CMS Internal Note in Preparation**, *"3D MCJet Corrections in CMSSW_1_2_0."*

[9] `TrackAssociator` TWiki page:
https://twiki.cern.ch/twiki/bin/view/CMS/TrackAssociator.

[10] `SteppingHelixPropagator` TWiki page:
https://twiki.cern.ch/twiki/bin/view/CMS/SteppingHelixPropagator.

[11] K. Terashi, *"$\not{E}_T$ Studies using $Z+$ Jets Events"* talk given at the JetMET group meeting 5/31/07,
http://indico.cern.ch/materialDisplay.py?contribId=6&materialId=slides&confId=16893.

[12] H. Pi *et al.*, **CMS Note 2005/52**, *"Study on Missing Transverse Energy Correction of $t\bar{t}$ and $W+$ jets events in CMS."*

[13] S. Baffioni *et al.*, **CMS Note 2006/40**, *"Electron Reconstruction in CMS."*

[14] CMS Collaboration, *"Tau jet reconstruction and tagging with CMS"*, Eur. Phys. J C, **45**, Suppl. 1. Phys. Rev. D **13**, 3214 (1976).

[15] M. Spiropulu and T. Yetkin, **CMS Note IN-2006/010**, *"Jet and Event Electromagnetic and Charged fraction in CMS."*

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