
CMS Physics Analysis Summary

Contact: cms-pog-conveners-jetmet@cern.ch

2010/07/21

Missing Transverse Energy Performance in Minimum-Bias and Jet Events from Proton-Proton Collisions at $\sqrt{s} = 7$ TeV

The CMS Collaboration

Abstract

The Large Hadron Collider (LHC) has successfully delivered proton-proton collision data at a centre-of-mass energy of 7 TeV since the end of March 2010. The studies presented in this note are based on a data sample of up to 11.7 nb^{-1} of integrated luminosity. We discuss the performance of several different missing transverse energy reconstruction algorithms employed by the CMS experiment using minimum-bias data and data selected to contain high transverse momentum jets. We find that, after the removal of anomalous signals in the calorimeters, the measured missing transverse energy distributions using different algorithms are in general agreement with expectations from the Monte Carlo simulation. An excellent performance of different complementary algorithms will allow a robust physics analysis using missing transverse energy.

1 Introduction

The CMS detector [1] has nearly 4π solid angle coverage and is able to detect most species of particles produced in proton-proton (pp) collisions up to $|\eta| \approx 5$. Exceptions are neutrinos and hypothetical weakly interacting particles, which escape from the detector without leaving a trace. Their presence can still be inferred from the *missing transverse momentum* (\vec{E}_T), defined as the apparent imbalance of the component of the momentum in the plane perpendicular to the beam direction, and its magnitude is referred to as *missing transverse energy* (E_T).

E_T is one of the most important variables for discriminating leptonic decays of W bosons from background events which do not contain neutrinos, such as QCD jet and Drell-Yan events. E_T is also an important variable in any search for new particles that are weakly interacting or quasi-stable. Many beyond-the-standard-model scenarios, including Supersymmetry, predict events containing large E_T . While E_T is an important variable, it is also one of the most complex variables, since E_T reconstruction is very sensitive to various detector malfunctions, particles impinging poorly-instrumented regions of the detector, cosmic-ray particles, and beam-halo particles. These effects may result in an artificial E_T . Thus, great care is required to understand the E_T distribution as measured by the detector.

E_T is generally calculated as the magnitude of the negative vector sum of the momentum transverse to the beam axis of all final-state particles reconstructed in the detector. The most traditional and common algorithm uses energies deposited in calorimeter towers and assumes mass-less objects based on energies measured in the tower and angles defined by a vector from the reconstructed primary vertex of the event to the tower. CMS has implemented four major types of algorithms to reconstruct E_T : (i) E_T based on calorimeter energies ($caloE_T$) [2], using the tower geometry of the hadron calorimeter, (ii) E_T calculated by replacing the calorimeter tower energies matched to charged hadrons with their corresponding charged-track momenta (track-corrected E_T or tcE_T) [3], (iii) E_T calculated using a complete particle-flow technique (pfE_T) [4], and (iv) E_T calculated using reconstructed jets (H_T). For $caloE_T$, a correction scheme is developed to account for the effects due to the non-compensating nature of the calorimeter by applying jet energy scale corrections for calorimeter towers within jets and another set of energy scale corrections for calorimeter towers outside of jets.

In this note, we present the performance of these E_T reconstruction techniques in minimum-bias and jet data sets collected by the CMS experiment from 7 TeV pp collisions. Minimum-bias events and QCD jet events are expected to be balanced in the plane transverse to the beam-axis, and have little true E_T . Therefore, these event samples provide a good testbed to study the finite resolution in the E_T measurement, and also examine the effects of calorimeter anomalous signals, beam-induced backgrounds, and other instrumental artifacts.

2 Data Sets and Event Selection

The data sets used for studies presented in this note were collected since the end of March 2010 and consist of about the first two months of low luminosity pp collisions at the centre-of-mass energy $\sqrt{s} = 7$ TeV. Only runs with all CMS subdetectors fully functional have been used in the analyses, and they correspond to an integrated luminosity of 11.7 nb^{-1} .

The data samples are collected by the minimum-bias, jet 6 GeV, and jet 15 GeV triggers. The minimum-bias trigger is based on a signal in the Beam Scintillation Counters (BSC), from at least one side of the CMS detector, coincident with a signal from either of the two Beam Pick-up Timing eXperiment (BPTX) devices. The minimum-bias trigger rate was sometimes randomly

reduced (prescaled) by a factor up to 500 in order to avoid saturating the bandwidth of the data acquisition system. The jet 6 GeV trigger requires at least one jet with jet energy-scale uncorrected $p_T > 6$ GeV at level-1 of the CMS trigger system, and was prescaled up to 100. The jet 15 GeV trigger requires at least one jet with uncorrected $p_T > 15$ GeV in the high level trigger system, and was not prescaled.

The offline event selection requires at least one well-identified primary vertex whose z position is less than 15 cm away from the nominal centre of the detector, ensuring that particles coming from collisions are well contained in the CMS detector. In addition, we require that the fraction of high-purity tracks [5] is greater than 25% in events with more than 10 tracks; this removes rare background events which are the result of particle bunches or showers traversing the pixel detector longitudinally. These offline event selections are referred to as the standard good collision criteria. After these selections are applied to the minimum-bias triggered events, we obtain an event sample which is referred to hereafter as the minimum-bias data sample.

The collision data are compared to samples of simulated minimum-bias or jet events that were generated using PYTHIA 8.130¹ [6] with a tuned parameter setting referred to as tune 1². The generated events are passed through the CMS detector simulation based on GEANT4 [7]. The detector geometry description includes realistic subsystem conditions such as dead channels. The offline event selection for collision data discussed above is also applied to the Monte Carlo simulated events.

In addition to the minimum-bias data sample, event samples containing high p_T hadronic jets are of particular interest in some of the studies presented here. Jet reconstruction and its performance at CMS is discussed in detail elsewhere [8]. Briefly, jets reconstructed by the anti-kt algorithm [9] with jet radius parameter $R = 0.5$ are most commonly used in CMS and specifically in the studies presented in this note. In CMS, three different methods are developed for jet clustering; calorimeter jets, jet-plus-track (JPT) jets, and particle-flow jets. The energy response of these jets is corrected on average to the particle-level using a calibration derived from Monte Carlo simulated events. The jet energy and transverse momentum mentioned hereafter are the values after this correction unless otherwise mentioned. False energies from electronics noise and other detector artifacts may cause fake jets. They are reduced by a set of jet identification criteria [8].

We use two different selections in this note. The first is defined by requiring at least two jets passing the jet ID cuts and requiring $p_T > 20$ or 25 GeV and $|\eta| < 3$. The other requires at least one jet with $p_T > 40$ GeV, the next-to-leading jet $p_T > 20$ GeV, and $|\eta| < 5$, or the next-to-leading jet $p_T > 40$ GeV, and $|\eta| < 3$. These different selections are chosen to optimize the differently focussed studies presented in this note. No additional requirement is made on a possible third jet in the event unless otherwise mentioned. Calorimeter jets are used to define the dijet event sample when studying $\text{calo}\cancel{E}_T$ and $\text{tc}\cancel{E}_T$, and particle-flow jets are used to define the dijet sample when studying $\text{pf}\cancel{E}_T$, unless otherwise mentioned. Additional requirements may be applied depending on the analysis; such requirements are clearly stated in the subsequent sections.

¹We refer to this version of PYTHIA as PYTHIA 8 hereafter.

²<http://home.thep.lu.se/~torbjorn/php8135/Welcome.php>

3 Correcting for Effects of Instrumental Anomalies and Beam-Induced Backgrounds

This section briefly describes various instrumental causes that can lead to anomalous \cancel{E}_T measurements and the methods used to correct \cancel{E}_T .

3.1 Removal of Anomalous Signals in the Calorimeters

The CMS electromagnetic calorimeter (ECAL) barrel (EB) and endcap (EE), and hadron calorimeter (HCAL) barrel (HB), endcap (HE), and forward (HF) sections are known to record anomalous signals that correspond to particles hitting the transducers, or rare random discharges. Some of these effects were already observed during cosmic data taking and the 2009 collisions at 900 GeV and 2.36 TeV. However, detailed studies of these effects have continued with the 7 TeV data taking. Those studies are documented in detail in Ref. [10] for the HCAL, and in Ref. [11] for the ECAL. In the present note we briefly summarize the known effects, and refer the interested reader to Refs. [10, 11] for more details.

The basic strategy is to identify anomalous signals based on unphysical charge sharing between neighboring channels in space and/or depth, as well as timing / pulse shape information. Each of the calorimeters in CMS records signals as a time series. The corresponding pulse shape is different for signals that develop in the calorimeters versus anomalous signals due to particles hitting the transducers directly and rare random discharges.

Once a “hit” in an HCAL tower or ECAL crystal is determined to be unphysical, we exclude it from the reconstruction of higher level objects like jets or \cancel{E}_T . We thus arrive at a reconstruction of jets and \cancel{E}_T that is consistently “cleaned” of anomalous detector effects.

In addition, there are some features of anomalous signals that can be exploited efficiently only after higher level objects such as jets are reconstructed. For example, a reconstructed high energy jet without any ECAL energy is exceedingly unlikely to be the result of the interaction of the particles that result from the hadronization of a parton in the calorimeters. Finally, there are very rare events for which one readout box (RBX) in the HB or HE fires most of its 72 electronic channels, without creating any signal in any neighboring RBX. While we could just exclude the corresponding channels from the reconstruction, as we do with the other corrections, it is unclear how useful an event is once 72 of about 7000 HCAL channels have been invalidated. For the present note, we exclude such events. We find that this requirement excludes 0.003% of otherwise good minimum-bias events. The \cancel{E}_T distributions after different stages of the anomalous signal removal procedure in the minimum-bias data sample are shown in Fig. 1, demonstrating the effectiveness of the anomalous signal removal.

3.2 Removal of the Beam-Induced Backgrounds

The effects of machine-induced backgrounds, often referred to as beam halo, can be quite detrimental to \cancel{E}_T or \cancel{E}_T -related quantities. Lessons learned from prior collider experiments motivated an intense effort to identify beam halo and preclude its potential harmful effects to the \cancel{E}_T observable.

The beam halo event filter presented in the present note relies on the information obtained from the Cathode Strip Chambers (CSCs), a subdetector with good reconstruction performance for both collision and non-collision muons. The CSCs provide extensive geometrical coverage in the transverse plane, making it difficult for beam halo particles, with mostly parallel-to-beam trajectories, to traverse the barrel calorimetry without also traversing one or both CSC endcaps. This feature of the CMS geometry is exploited in order to tag events which may have

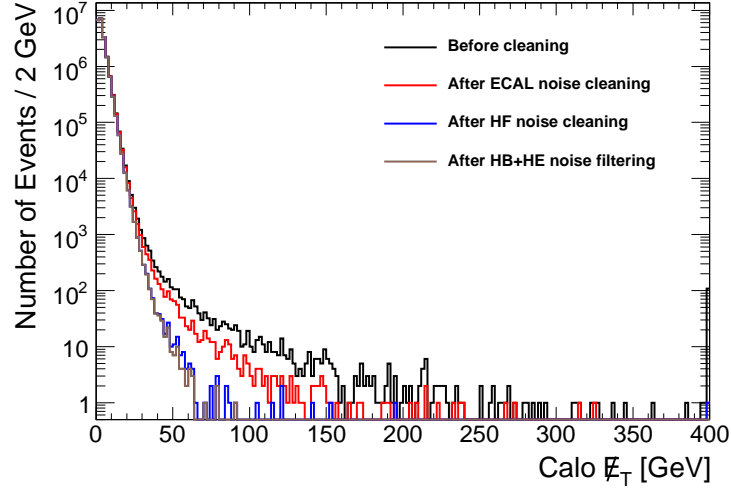
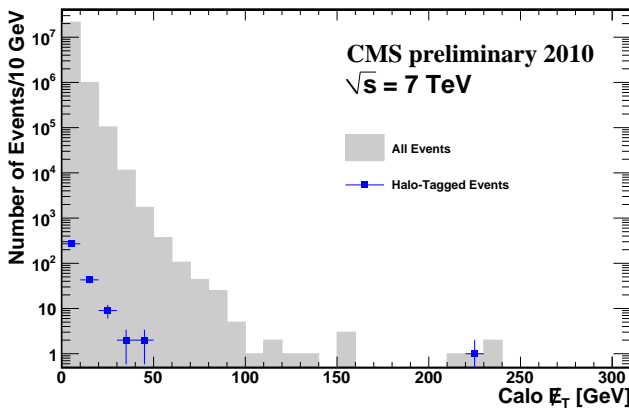


Figure 1: Calo E_T distributions after different stages of the anomalous signal removal procedure.

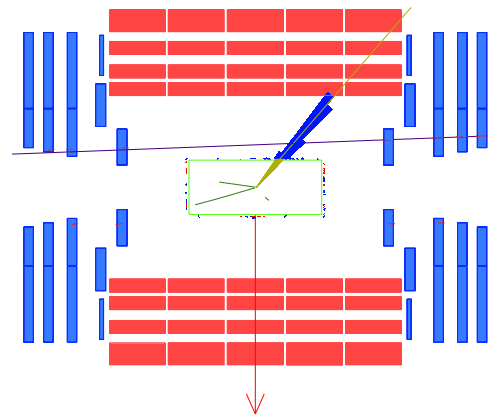
contamination from beam halo. The filter efficiency is found to be 65% in a high E_T beam halo Monte Carlo sample ($E_T > 15$ GeV), while the mistag efficiency is $1.5 \cdot 10^{-7}$ in a minimum bias Monte Carlo sample.

This CSC-based beam halo filter was applied to the good 7 TeV collision runs of 2010 which only saw a modest increase in beam intensity. Before applying the halo filter, the event was required to pass the standard good collision criteria. The beam halo tagged event fraction in minimum-bias events was found to be stable at $1 - 2 \times 10^{-5}$ for the run period considered.

Figure 2(a) shows the calo E_T distribution for all events and beam halo tagged events, passing the standard good collision criteria as defined in Section 2, collected by the minimum-bias



(a) calo E_T distribution



(b) high E_T beam halo event

Figure 2: (a) Calo E_T distribution for events before the beam halo filter is applied and for events tagged by the beam halo filter, collected with the minimum-bias trigger or jet 15 GeV trigger and passing the standard good collision criteria. (b) Event display of the beam halo tagged event (Run 135535, Event 15564358) with largest calo E_T (224 GeV).

trigger or jet 15 GeV trigger. In most events, beam halo does not significantly affect \cancel{E}_T , while a few events are observed in the high \cancel{E}_T tails. Figure 2(b) shows an event display for the halo-tagged event with the highest reconstructed $\text{calo}\cancel{E}_T$. We can readily observe one halo muon traversing the barrel calorimetry causing high \cancel{E}_T . Although the beam halo particles do not presently make significant contributions to high \cancel{E}_T events, they may become so when the beam current increases. The beam-halo filter already developed will allow good control of high \cancel{E}_T events due to beam halos.

4 \cancel{E}_T Performance in Minimum-Bias and Dijet Events

In this section, we demonstrate the performance of different \cancel{E}_T reconstruction algorithms developed by CMS in minimum-bias and dijet events.

In each case, we present data - Monte Carlo comparisons for three types of \cancel{E}_T reconstruction. The minimum-bias sample discussed in Section 4.1 allows us to study \cancel{E}_T in a least-biased manner. However, for such low- p_T events the event generator is not believed to be as reliable as for the higher- p_T dijet events discussed in Section 4.2. The jet samples also allow us to test \cancel{E}_T performance in an event topology closer to the one to be used in future searches for physics beyond the standard model. In all data - Monte Carlo comparisons of histograms presented in this note, Monte Carlo distributions are normalized to the data distributions.

In Section 4.3 we present the \cancel{E}_T scale corrections for $\text{calo}\cancel{E}_T$. Finally, in Section 4.4 comparisons of the performance for corrected $\text{calo}\cancel{E}_T$, $\text{tc}\cancel{E}_T$, and $\text{pf}\cancel{E}_T$ are presented.

4.1 \cancel{E}_T Performance in Minimum-Bias Events

For each of the three \cancel{E}_T algorithms, we study the \cancel{E}_T and $\sum E_T$ distributions in minimum-bias events. The results are presented in Fig. 3 for (a) $\text{calo}\cancel{E}_T$, (c) $\text{tc}\cancel{E}_T$, and (d) $\text{pf}\cancel{E}_T$, respectively. The $\text{pf}\cancel{E}_T$ distribution is discussed in further detail in Ref. [12]. The interested reader is also referred to the components of \cancel{E}_T along the x and y axes (\cancel{E}_x and \cancel{E}_y), where the x and y axes are in the horizontal and vertical directions perpendicular to the beam direction, shown in Appendix A (Fig. 12).

For the \cancel{E}_T distributions, general agreement is observed between data and PYTHIA 8 Monte Carlo simulated events, although the Monte Carlo distribution is somewhat narrower. The observed differences are attributed to various sources including the generator-level modeling of minimum-bias events and the imperfect modeling of the calorimeter response in the simulation. The observed difference was further investigated by separately studying the \cancel{E}_T measured by the ECAL and HCAL, and in different parts of the calorimeters (barrel, endcap, and forward). This study showed that the data and Monte Carlo differences are most evident in the HB and HE regions. As seen in Ref. [13], the energy response in the HE region is known to be underestimated in the simulation, which contributes to the difference observed between data and Monte Carlo prediction.

Events with $\cancel{E}_T > 60$ GeV are studied in detail by a visual scan. There are ten events with $\text{calo}\cancel{E}_T > 60$ GeV in about 20 million minimum-bias events. Two of these events show an indication of residual anomalous signals in HF. In one of them anomalous signals in a few adjacent channels in HF are observed, which are not effectively removed by the present cleaning procedure presented in Section 3. This event also has $\text{tc}\cancel{E}_T$ and $\text{pf}\cancel{E}_T > 60$ GeV. Among the five events with $\text{tc}\cancel{E}_T > 60$ GeV, we find one additional event that appears to have anomalous signals in HB and HE. Work is in progress to develop more effective algorithms to further reduce contributions of these anomalous signals. We observe three events with $\text{pf}\cancel{E}_T > 60$ GeV.

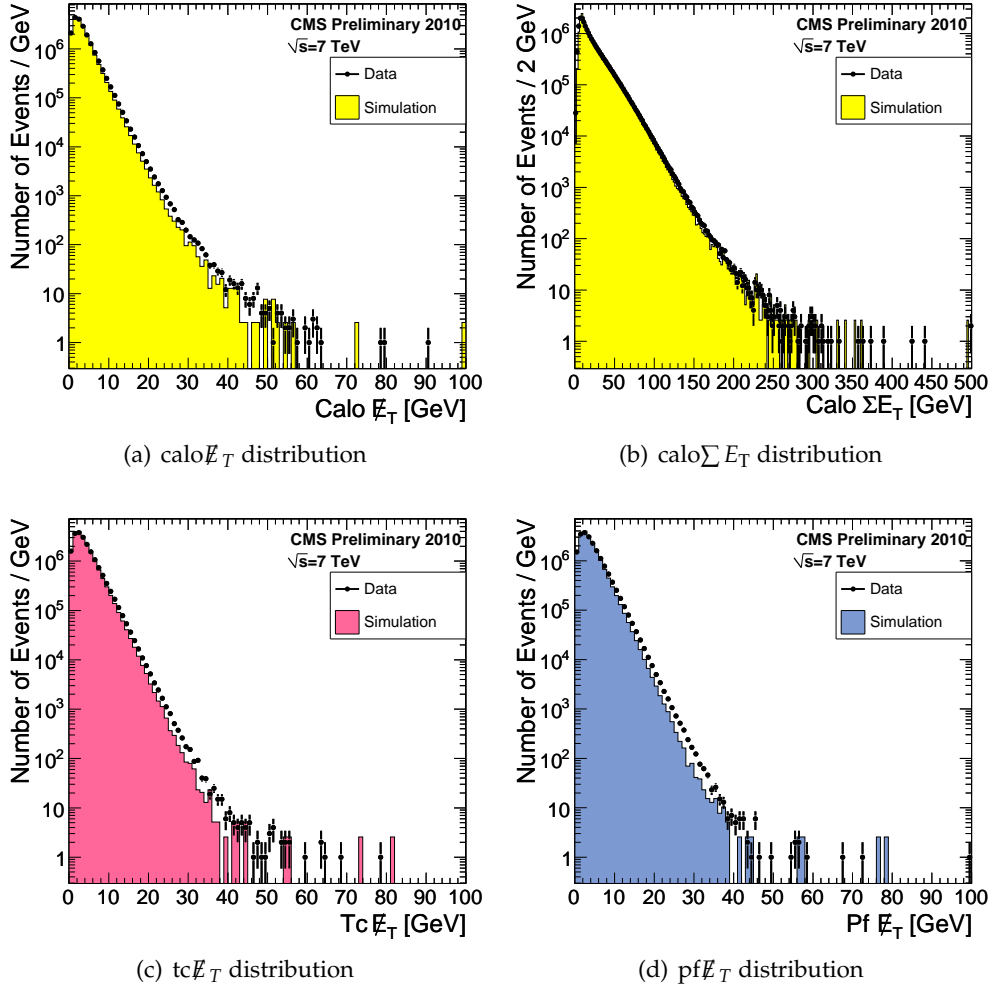


Figure 3: calorimeter E_T ($\text{calo} E_T$), $\text{calo} \Sigma E_T$, track-corrected E_T ($\text{tc} E_T$), and particle-flow E_T ($\text{pf} E_T$) distributions in the minimum-bias data compared with Monte Carlo simulation.

In addition to the event with anomalous HF signals mentioned above, the highest $\text{pf} E_T$ event has a high p_T forward muon due to poor track reconstruction of the associated track in the central tracker, which will be fixed in future analysis. The other events do now show obvious anomalous signals, and they appear to be QCD jet production events with some jet energies mismeasured. The mismeasured jets tend to be at the boundaries between different sections of the calorimeters. Therefore, the high E_T tail is no longer dominated by events with anomalous calorimeter signals after the anomalous signal cleaning procedure, and the high E_T tail is reasonably well described by the Monte Carlo simulation.

Figure 3(b) shows the $\text{calo} \Sigma E_T$ distribution (see Appendix A for $\text{tc} \Sigma E_T$ and $\text{pf} \Sigma E_T$ distributions). We find an improved agreement between data and Monte Carlo simulation using PYTHIA 8 compared to the studies in Ref. [14] using PYTHIA 6 D6T tune[15]. Traditionally, ΣE_T has been considered to be a difficult quantity to correctly simulate due to its sensitivity to details of the calorimeter simulation, in particular, to the noise treatment, and also to the generator-level modeling of soft QCD physics. The PYTHIA 8 Monte Carlo simulated events tend to have a somewhat softer ΣE_T , although the observed agreement is acceptable given the various difficulties mentioned above. It is important that the event ΣE_T , *i.e.*, the total visible

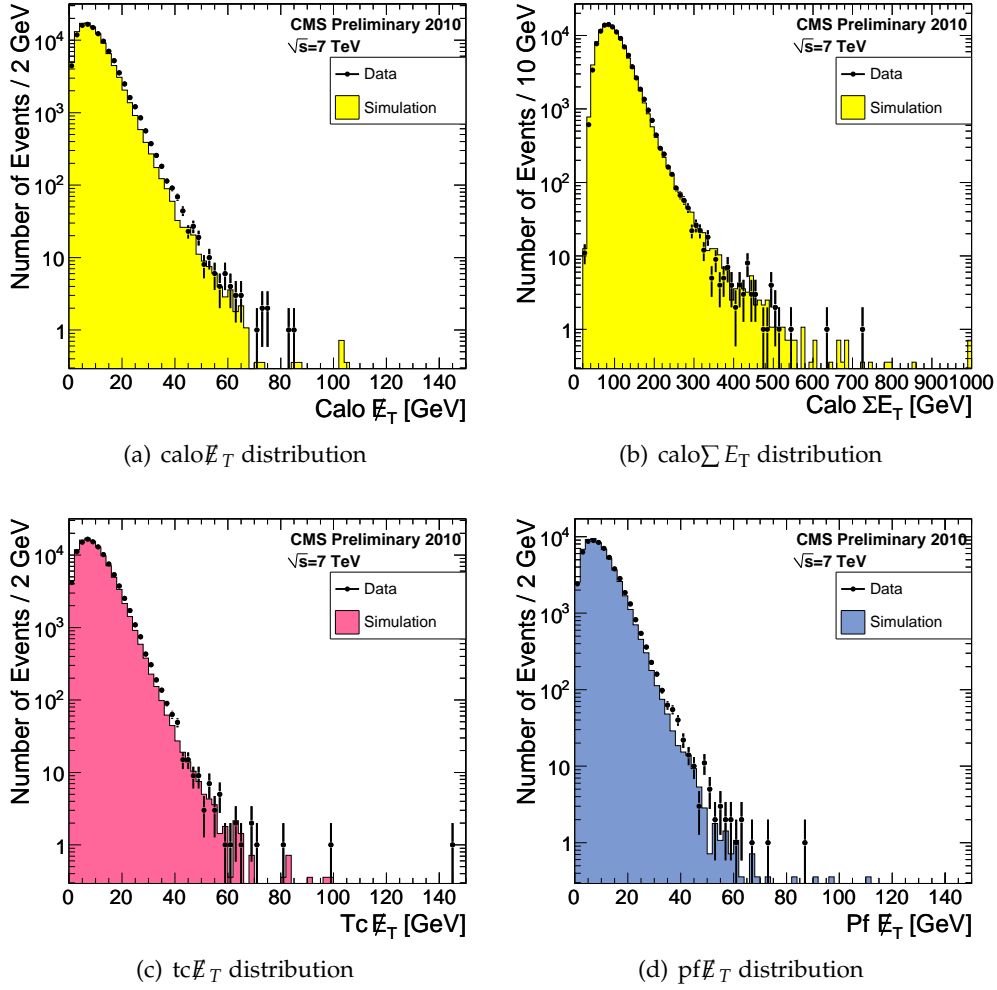


Figure 4: Calorimeter \cancel{E}_T ($\text{calo}\cancel{E}_T$), $\text{calo}\sum E_T$, track-corrected \cancel{E}_T ($\text{tc}\cancel{E}_T$), and particle-flow \cancel{E}_T ($\text{pf}\cancel{E}_T$) distributions in the inclusive dijet data ($p_T^{\text{jet}1,2} > 25$ GeV and $|\eta^{\text{jet}1,2}| < 3$) compared with Monte Carlo simulation.

transverse energy in the event, is reasonably-well modeled since the \cancel{E}_T reconstruction relies on all visible transverse energies in an event.

4.2 \cancel{E}_T Performance in Dijet Events

In this subsection, a study of \cancel{E}_T performance in events containing two high energy hadronic jets is presented. Figure 4 includes the same set of distributions as Fig. 3, but in events containing at least two jets with $p_T > 25$ GeV and $|\eta| < 3$. In general, events containing a hard scatter such as dijet events are better understood theoretically, and Monte Carlo simulated events are more reliable. Indeed, good agreement is observed. The $\text{pf}\cancel{E}_T$ distribution in dijet events is discussed in further detail in Ref. [12].

Comparisons between data and Monte Carlo for the \cancel{E}_T and $\sum E_T$ distributions in this event topology show reasonable agreement, especially for $\sum E_T$. In the case of \cancel{E}_T , the Monte Carlo distribution is somewhat narrower, consistent with the under-estimation of the \cancel{E}_T resolution in the simulation. This is similar to our observation in the minimum-bias event distribution and is expected to improve when the response in the HE region is improved in the simulation.

The $E_{x,y}$ distributions and the angular correlation between \vec{E}_T and jets in the same event sample are discussed in Appendix A.

4.3 Type I and Type II Corrections for Calorimeter E_T

Because of the non-compensating nature of the calorimeter, energies of hadrons tend to be under-measured. Among the various types of E_T discussed in this note, $\text{calo}E_T$ is most affected by this effect, and a two-step correction has been devised in order to remove the bias in the E_T scale.

The calorimeter energies within jets can be corrected to the particle level using the jet energy correction [8]. This correction with respect to \vec{E}_T is called the type-I correction and is applied to all jets with electromagnetic energy fraction < 0.9 and corrected $p_T > 20$ GeV. In order to account for the remaining soft jets below this threshold and energy deposits not clustered in any jet (unclustered energy), a second correction is applied on top of the type-I correction, which is referred to as the type-II correction.

The type-II correction is currently obtained using a $Z \rightarrow ee$ Monte Carlo sample. Z events constitute an ideal sample for determining the type-II correction, as Z 's are generally produced with low p_T and the recoil is often dominated by unclustered energy. We take the vector sum of momenta defined for all calorimeter towers not corrected in the type-I correction as a single object denoted \vec{U} . The measured value of \vec{U} is defined as:

$$\vec{U}_{\text{meas}} = -\text{calo}\vec{E}_{T\text{uncorr}} - \sum \vec{p}_{\text{jet,uncorr}} - \sum \vec{p}_{\text{ele,meas}}, \quad (1)$$

where \vec{U}_{meas} is the measured \vec{U} , $\text{calo}\vec{E}_{T\text{uncorr}}$ is the uncorrected $\text{calo}E_T$, $\vec{p}_{\text{jets,uncorr}}$ is the momentum of uncorrected jets, the sum is over all jets with the corrected $p_T > 20$ GeV, and $\vec{p}_{\text{ele,meas}}$ is the momentum of the measured electrons. All these vectors are considered only in the plane perpendicular to the beam direction.

The type-II correction for \vec{U}_{meas} is obtained using the correlation between \vec{U}_{meas} and \vec{U}_{gen} , the true unclustered energy from the Monte Carlo generator information. As the direction of the \vec{U}_{meas} may differ from the direction of \vec{U}_{gen} due to noise, underlying event, etc., the parallel component of \vec{U}_{meas} projected on the direction of \vec{U}_{gen} , $\vec{U}_{\text{meas},||}$, is used for the derivation of the correction. The response of the unclustered energy is then defined as $R(U_{\text{gen}}p_T) = U_{\text{gen}}p_T / U_{\text{meas},||}p_T$. The obtained correction U_{scale} is parametrized as $1.5 + 1.8 \cdot \exp(-0.06 \cdot U_{\text{meas}}p_T)$.

The $\text{calo}E_T$ distributions without the type-I and II corrections, after the type-I correction, and after the type-II corrections, in events containing at least two jets with $p_T > 25$ GeV and $|\eta| < 3$ are shown in Fig. 5(a). The type-II corrected $\text{calo}E_T$ distributions are wider than the uncorrected $\text{calo}E_T$ distribution, although dijet events have no intrinsic E_T . This trend simply reflects the fact that the corrected E_T yields a correct E_T scale on average in events containing intrinsic E_T .

One of the ways to validate the type-II corrected $\text{calo}E_T$ in jet events is to look at \vec{E}_T decomposed into two components based on the dijet bisector axis. The dijet bisector axis divides the azimuthal opening angle between the leading two jets in two. Positive $E_{T||}$ points towards the opening angle of the leading two jets, and positive $E_{T\perp}$ points towards the more central jet (smaller $|\eta|$).

Distributions of both $E_{T||}$ and $E_{T\perp}$ are nearly symmetric before the type-I and type-II corrections, although there is a slight asymmetry in $E_{T||}$ which is due to the fact that the bisector

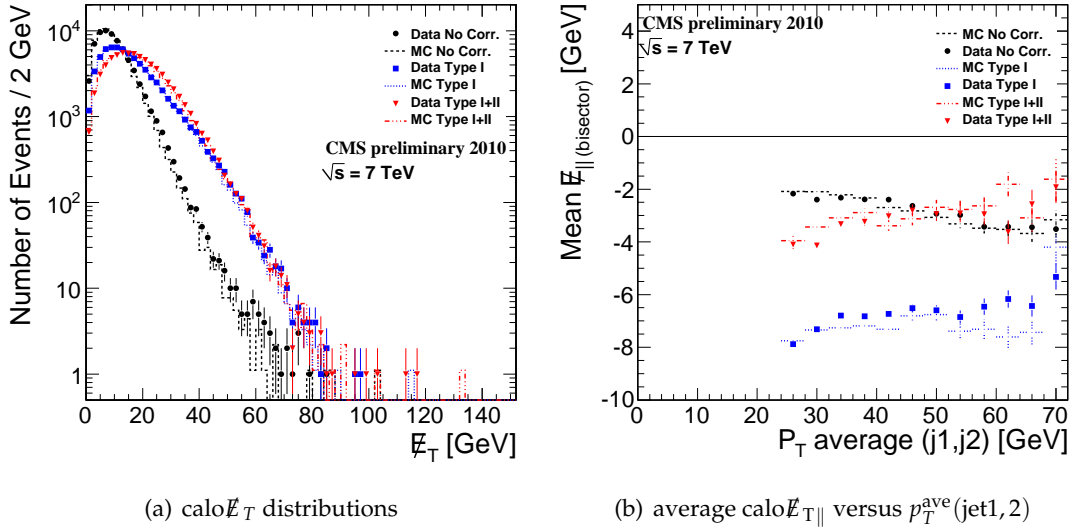


Figure 5: (a) Calo \cancel{E}_T distributions and (b) average calo $\cancel{E}_{T||}$ with respect to the dijet bisector axis versus average p_T of the leading two jets ($p_T^{\text{ave}}(\text{jet1},2)$) in dijet events compared with Monte Carlo simulation, without correction, with type I correction and type II correction.

axis always points towards the opening angle of the leading two jets. The type-I correction introduces a more significant asymmetry in the $\cancel{E}_{T||}$ distribution, because it produces artificial \cancel{E}_T in the direction opposite to the dijet opening angle. The type-II correction, however, calibrates the rest of the calorimeter energies, and makes the $\cancel{E}_{T||}$ distribution nearly symmetric again to the level of the $\cancel{E}_{T||}$ distribution before the type-I and type-II corrections are applied. This trend may be seen in Fig. 5(b) which shows the mean values of $\cancel{E}_{T||}$ as a function of $p_T^{\text{ave}}(\text{jet1},2) = (p_T^{\text{jet1}} + p_T^{\text{jet2}})/2$ for uncorrected, type-I corrected, and type-II corrected \cancel{E}_T .

4.4 \cancel{E}_T resolution versus $\sum E_T$

It is interesting to compare the \cancel{E}_T resolution for the three \cancel{E}_T reconstruction techniques. Because the \cancel{E}_T resolution has a strong dependence on the associated $\sum E_T$ it is presented as a function of $\sum E_T$. We characterize the \cancel{E}_T resolution based on the σ of a Gaussian fit to the $\cancel{E}_{x,y}$ distribution [2]. In order to make a meaningful comparison, we calibrate the measured \cancel{E}_T to the same scale. The calibration procedure employed to make this comparison possible is presented below.

The calibration for \cancel{E}_T is taken from a study of photon+jets events. Briefly, in this study the ratio $p_{T,||}^{\text{recoil}}/p_T^\gamma$ is measured as a function of p_T^γ , where p_T^γ is the p_T of the leading high p_T photon in the event, and $p_{T,||}^{\text{recoil}}$ is the p_T component of the “recoil” parallel to the leading photon direction.

The recoil is defined to be a vector sum of $\vec{\cancel{E}}_T$ and the photon momentum. When the \cancel{E}_T scale is well calibrated, we expect the $p_{T,||}^{\text{recoil}}/p_T^\gamma$ ratio to be unity. We use this ratio obtained from the photon+jets events as a \cancel{E}_T scale correction factor. From the calibrated \cancel{E}_T , new $\cancel{E}_{x,y}$ values are calculated and the \cancel{E}_T resolution is obtained. In the photon+jets study, the lowest p_T^γ value is 20 GeV; thus for any \cancel{E}_T below 20 GeV we use the calibration factor obtained at $p_T^\gamma = 20$ GeV. This was performed for the type-II corrected calo \cancel{E}_T , tc \cancel{E}_T , and pf \cancel{E}_T .

For $\sum E_T$, we use the pf $\sum E_T$ as measured by the particle-flow algorithm for all types of \cancel{E}_T . $\sum E_T$ is also measured using inputs to calo \cancel{E}_T and tc \cancel{E}_T (calo $\sum E_T$ and tc $\sum E_T$); however, pf $\sum E_T$

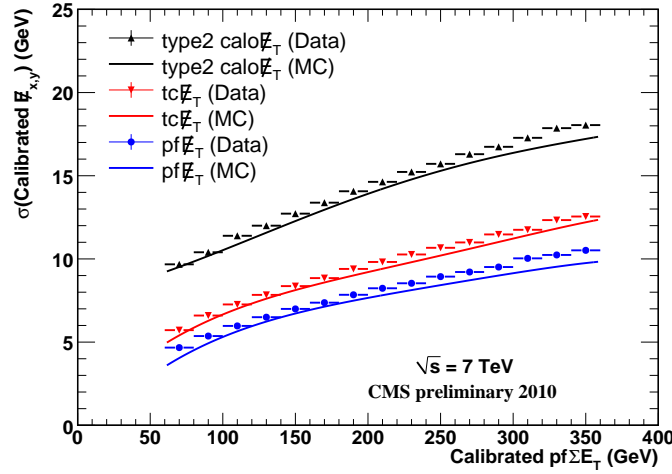


Figure 6: Calibrated E_T resolution versus calibrated $pf\sum E_T$ for the type-II corrected $caloE_T$, tcE_T , and pfE_T in data and Monte Carlo samples.

is closest to the actual particle-level $\sum E_T$. We calibrate $pf\sum E_T$ to the particle-level $\sum E_T$, on average, based on the correlation between the measured $pf\sum E_T$ and the particle-level $\sum E_T$ observed in PYTHIA 8 Monte Carlo.

Figure 6 shows the calibrated E_T Gaussian core resolution versus the calibrated $pf\sum E_T$ for different E_T reconstruction algorithms in events containing at least two jets with $p_T > 25$ GeV. Both tcE_T and pfE_T show improvements in the E_T resolution compared to the calorimeter only E_T in the resolution, and the pfE_T yields the smallest E_T resolution. The comparison between the type-II corrected $caloE_T$ and pfE_T is also presented in Ref. [12], which confirms the E_T resolution improvements in pfE_T compared to the type-II corrected $caloE_T$.

5 E_T Performance in Multi-Jet Events

Many searches for physics beyond the standard model critically depend on E_T in events with multiple (> 2) jets. However, such events are harder to simulate in Monte Carlo event generator programs. The limited understanding of multi-jet QCD production makes direct comparisons between data and Monte Carlo simulation challenging.

In this section, we present a study of the E_T performance in multi-jet events. In particular, we study the dependence of the E_T distribution on jet multiplicities. Figure 7 shows the E_T distribution in four different intervals of $\sum E_T$ for different jet multiplicities. Events in this figure are required to have at least two jets with $p_T > 20$ GeV and $|\eta| < 3$, and the jet multiplicity is defined to be the number of jets with the corrected $p_T > 20$ GeV.

The good agreement of the shape of the E_T distributions between different jet multiplicities in Fig. 7 indicates that the E_T performance is primarily driven by the total amount of calorimetric activity, parametrized by $\sum E_T$, and no significant contribution from jet multiplicities to E_T is visible at the presently available level of data statistics. This feature is useful for various physics analysis with multi-jets.

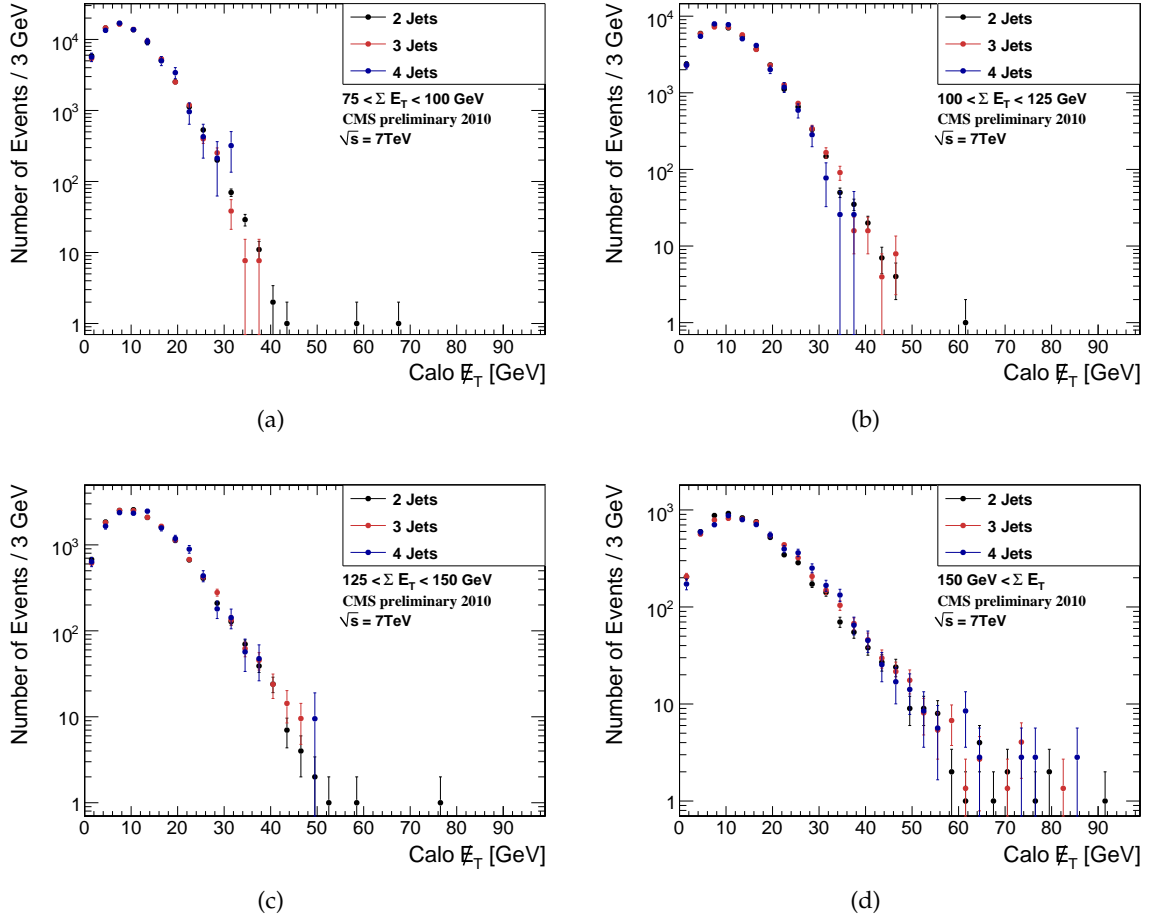


Figure 7: Calo \cancel{E}_T distributions in events with two, three, and four jets ($p_T > 20$ GeV and $|\eta| < 3$) in four $\sum E_T$ regions.

6 \cancel{E}_T in Pileup Events

The fraction of pile-up events, *i.e.*, a proton-proton beam bunch crossing that contains more than one pp collision, is quite low, $\sim 1\%$, in the data sample used in this note. However, as the instantaneous luminosity increases, so does pile-up. The fraction of pileup events is rapidly increasing, and it is therefore useful to study \cancel{E}_T in events with more than one reconstructed vertex. The quality requirements for additional primary vertices are the same as for the primary one, and the vertices are sorted according to their $\sum_{tracks} p_T^2$.

Figure 8 shows the \cancel{E}_T distributions for events with one or more than one reconstructed vertex. In pile-up events, higher $\sum E_T$ is expected. Re-weighting the \cancel{E}_T distributions based on $\sum E_T$ to match the $\sum E_T$ distribution in *non*-pile-up events aligns the shape of the \cancel{E}_T distribution in events with one and more than one vertex, as shown in Fig. 8. This shows that the widening of the \cancel{E}_T distributions in pile-up events is mainly the result of the higher event energy.

7 Performance of \cancel{H}_T

For some analyses, variables such as $H_T = \sum p_{T,jets}$ and $\cancel{H}_T = |\vec{\cancel{H}}_T|$ ($\vec{\cancel{H}}_T = -\sum \vec{p}_{T,jets}$), the calculation of which relies purely on clustered energy, present a more robust alternative to

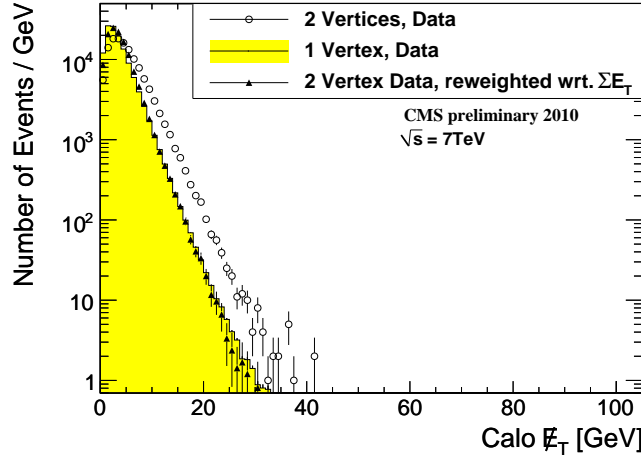


Figure 8: E_T distributions in 2-vertex events for $\text{calo}E_T$, with and without reweighting based on the $\sum E_T$ distribution to match the non-pile-up events. These distributions are compared to the E_T distributions for 1-vertex events. The distribution for 1-vertex events is normalized to the 2-vertex distribution.

$\sum E_T$ and E_T , which involve also unclustered energy and thus are more sensitive to detector noise and pile-up. In this section we present the behaviour of H_T and \cancel{H}_T and compare the resolution of \cancel{H}_T to that of E_T , as well as examining their deviations from Gaussian behaviour. For our studies we compare three different types of jet reconstruction that are currently in use within CMS: purely calorimeter-based jets (calo jets); Jet-Plus-Track (JPT) jets [16], and jets from Particle-Flow (PF).

Details of the jet selection criteria are described in Ref. [8]. In all three cases (calo, JPT, pf), the leading jet is required to have energy-scale corrected $p_T > 40$ GeV and uncorrected $p_T > 15$ GeV to ensure a high-level trigger efficiency $> 95\%$ for the selected event samples. Additional jets in the event are required to have transverse momentum $p_T > 20$ GeV and to be within pseudo-rapidity $|\eta| < 5$.

In order to reject any remaining noise from the calorimeters, calo and JPT jets are required to satisfy several jet identification criteria [8]. Furthermore, events are rejected if they contain a jet with corrected $p_T > 20$ GeV that does not pass the jet identification criteria. These criteria are designed to reject residual false energies from electronic noise and other artifacts of the ECAL and HCAL, which remain even after the noise cleanup procedure is applied, as discussed in Section 3.

Figure 9 shows the comparison of data to Monte Carlo simulated events for H_T on the left (a, c, and e) and \cancel{H}_T on the right (b, d, and f) for events with ≥ 2 jets. In all cases the simulation is scaled to the number of events observed in data. The distributions in the top (a and b), middle (c and d) and bottom (e and f) rows are for calo jets, JPT jets and particle-flow jets, respectively. In each plot we show the expectation from a QCD PYTHIA 8 Monte Carlo simulation [6] compared to data points (top window) and data divided by Monte Carlo to further examine the quality of the agreement in shape (bottom window). The data to Monte Carlo agreement is generally good.

The observed small difference in slope between data and Monte Carlo is mainly due to a faster falling H_T distribution in Monte Carlo simulation for dijet events. This is true for all three types of jet reconstruction methods.

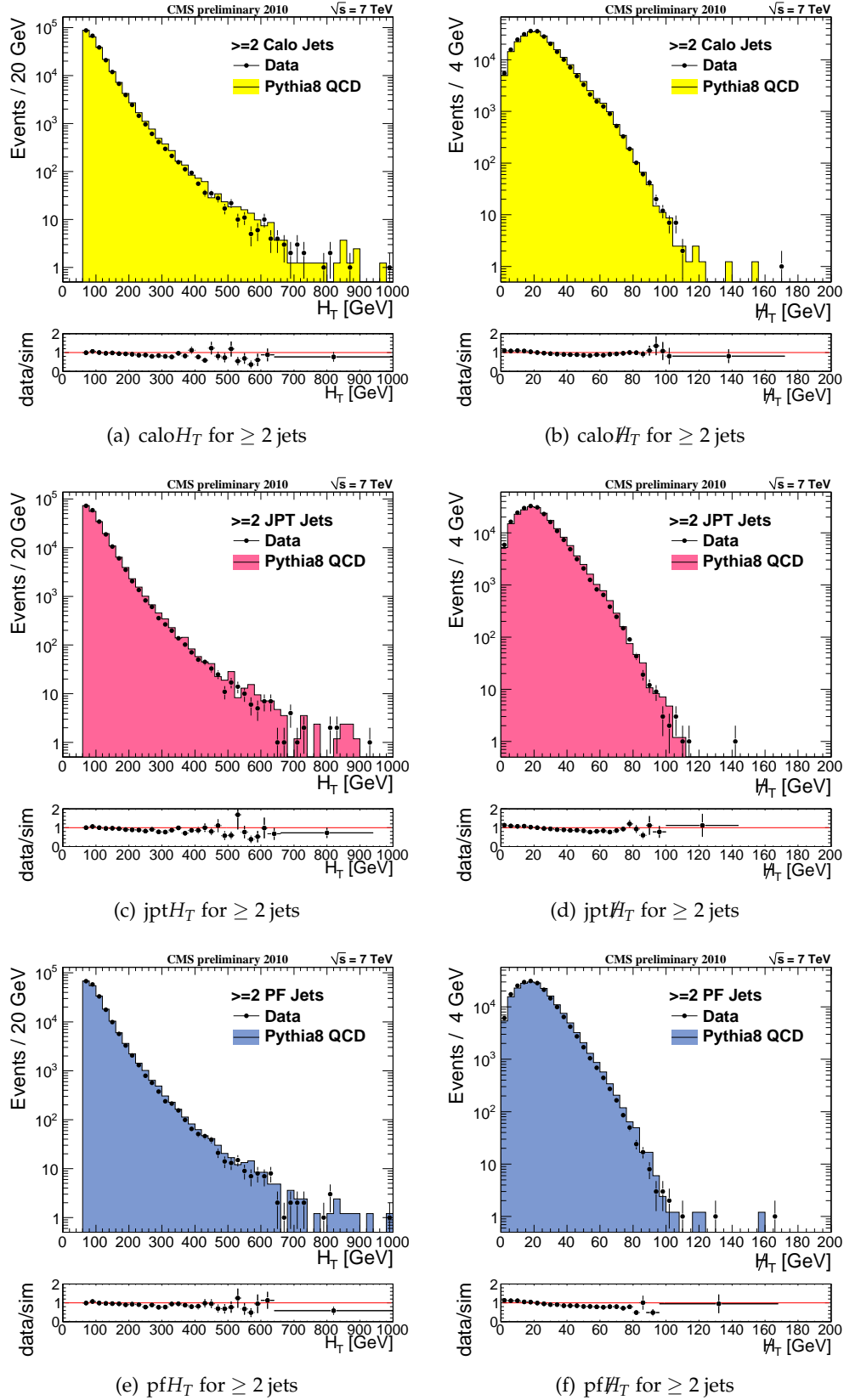


Figure 9: H_T (left) and $\#H_T$ (right) distributions in events with ≥ 2 jets, for calo jets (top), JPT jets (middle) and particle-flow jets (bottom).

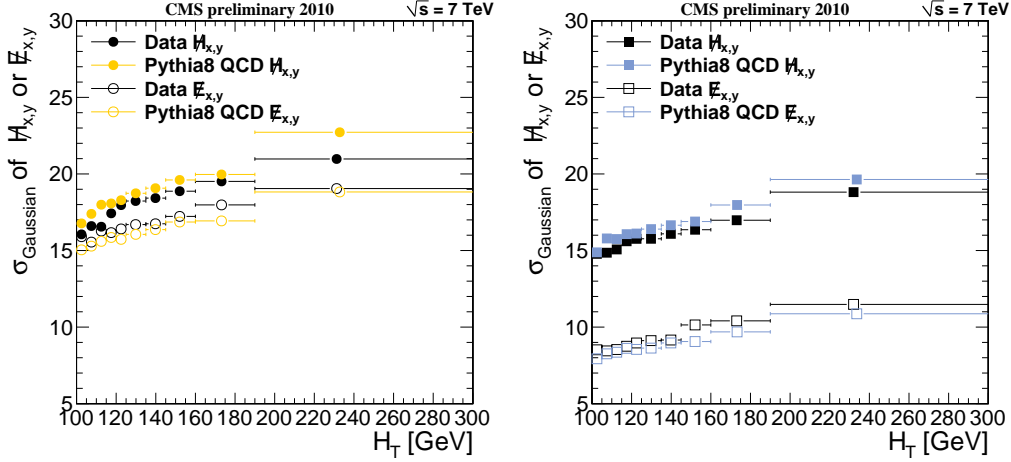


Figure 10: \cancel{H}_T and type II corrected \cancel{E}_T resolutions as a function of H_T in events with ≥ 2 jets for calo jets (left) and particle-flow jets (right).

In the \cancel{H}_T distributions no significant outliers are observed. The small shoulder at ~ 60 GeV stems from dijet events and is a result of the applied jet p_T thresholds of 40 GeV on the leading jet and 20 GeV on all other jets. For a small fraction of events both jet p_T 's are just above threshold and reconstructed in the same hemisphere while the balancing jets in the other hemisphere have fluctuated below the required jet p_T threshold.

To further evaluate the performance of the \cancel{H}_T approach we study the dependence of the resolution of \cancel{H}_T as a function of H_T . The resolution is defined to be the σ of a Gaussian fit to the $\cancel{H}_{x,y}$ distributions. In Fig. 10 we compare the resolutions of \cancel{H}_T in data and Monte Carlo. For comparison, we also plot the corresponding \cancel{E}_T resolutions. It can be seen that calo \cancel{H}_T has a slightly worse resolution than calo \cancel{E}_T , *i.e.* a wider core, due to the fact that \cancel{H}_T does not take into the account the unclustered energy in each event and jets that have fallen below the p_T threshold of 20 GeV. The effect is more pronounced when comparing to pf \cancel{E}_T , which uses a more sophisticated \cancel{E}_T reconstruction method.

We also assess the size of the tails in the \cancel{H}_T and \cancel{E}_T distributions, *i.e.* deviations from Gaussian behaviour. A good understanding of these tails is of particular importance to searches of new physics involving weakly interacting particles that escape the detector. To first approximation, disregarding detector noise, beam-halo events, pile-up and similar effects, $\cancel{E}_{x,y}$ ($\cancel{H}_{x,y}$) is expected to be Gaussian with a σ that depends on $\sum E_T$ (H_T). We test this through the following procedure. To reduce the dependence of the missing energy on the total energy in the event we divide $\cancel{E}_{x,y}$ by $\sqrt{\sum E_T}$ and $\cancel{H}_{x,y}$ by $\sqrt{H_T}$. We then determine the core width σ_{fit} of the $\cancel{E}_{x,y}/\sqrt{\sum E_T}$ and $\cancel{H}_{x,y}/\sqrt{H_T}$ distributions by iteratively fitting a Gaussian distribution to the events contained within $\pm 1.5\sigma$. In the case of perfect Gaussian behaviour, 68.3% of the events are contained within $\pm 1\sigma_{fit}$. We then find the actual range containing 68.3% of events in the observed distribution. Because of the non-Gaussian tails in the distribution this range deviates from σ_{fit} by some factor θ_1 . This procedure defines a point $(1, \theta_1)$ in Fig. 11. We then repeat this procedure for different multiples of σ , finding the actual $\cancel{H}_{x,y}/\sqrt{H_T}$ range, $\pm \theta_n$, expressed as a multiple of σ_{fit} that contains the corresponding fraction of events. We then plot θ_n versus n in Fig. 11. More rigorously θ_n can be defined as

$$\theta_n = F^{-1}(G(n))/\sigma_{fit} \quad (2)$$

where $G(n)$ is the cumulative distribution for a half-normal Gaussian with unit σ and F is the cumulative distribution of $|\cancel{H}_{x,y}|$. The solid line in Fig. 11 represents the diagonal $\theta_n = n$, which

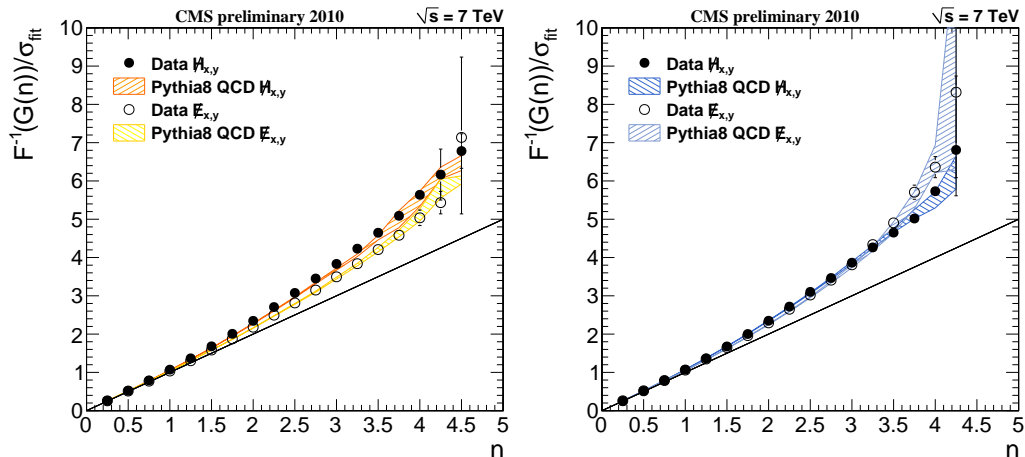


Figure 11: Characterizing the non-Gaussian behaviour of the tails in the $H_T/\sqrt{H_T}$ and $E_T/\sqrt{\sum E_T}$ distributions, constructed using calo jets (left) and particle-flow jets (right), for events with ≥ 2 jets.

corresponds to the perfect Gaussian behaviour. It is clear that above $\sim 2\sigma$ the distributions start deviating from Gaussian behaviour. The agreement between data and Monte Carlo is very good for both E_T and H_T cases, with the tails in data being only slightly larger in the case of calo jets. For calo jets, the tail in the calo E_T distribution is somewhat smaller than in the corresponding H_T distribution, while particle-flow jets, the H_T and pf E_T tails agree very well up to $n \sim 3.5$. In general, it is very encouraging to see that also the non-Gaussian tails in the E_T and H_T distributions are modeled well by the simulation.

8 Conclusion

We have presented an extensive test of the E_T reconstruction in CMS using the $\sqrt{s} = 7$ TeV data collected between the end of March 2010 and June 2010. The goal was to understand the performance of the E_T measurement using events where there is no significant intrinsic E_T .

We first studied the effect of instrumental anomalies and beam-induced backgrounds on the E_T measurement and demonstrated a cleaning procedure to identify and correct for these effects. After the cleaning procedure, the E_T and $\sum E_T$ distributions, as well as the E_T resolutions in the data are in general agreement with the Monte Carlo simulation predictions, although still some observed differences need to be investigated and worked out in future studies. Both track-corrected E_T and, especially, particle-flow E_T improve the resolution compared to the purely calorimetric E_T reconstruction.

For future searches for physics beyond the standard model (*e.g.*, Supersymmetry), we also studied the E_T performance in multi-jet events, pile-up events, and the H_T and H_T reconstruction (*i.e.*, the vector and scalar sums of jets). We found reasonable agreement between data and Monte Carlo simulation. In particular, it was shown that the E_T resolution depends primarily on $\sum E_T$ in multi-jet and pile-up events.

An excellent performance of the three different E_T and H_T reconstruction algorithms gives CMS the foundation of a robust physics programme by taking advantage of these largely independent measurement methods.

A Additional Figures for E_T Performance in Minimum-Bias and Jet Events

This section we present additional information to complement Sections 4.1 and 4.2. In Figs. 12 (a), (b) and (c), the x and y components of the E_T (E_x and E_y) distribution from three reconstruction techniques are shown. These distributions have two entries per event; one for the x component and the other for the y component. As expected, they are roughly symmetric with respect to zero, and general agreement is observed between data and Monte Carlo distributions, although the data distributions are slightly wider, indicating worse E_T resolution. This is consistent with the observation in the E_T distributions. This observed difference is primarily attributed to the imperfect response in the HCAL barrel and endcap regions. There is also a slight asymmetry in the $E_{x,y}$ distributions which is partially due to the non-uniform noise contributions in the ECAL endcap in the azimuthal angle. Nevertheless, the $E_{x,y}$ distributions are very close to being symmetric, and the sigma of the Gaussian fit to the $E_{x,y}$ distributions is used to evaluate the E_T resolution.

The $\text{tc}\Sigma E_T$ and $\text{pf}\Sigma E_T$ distributions in minimum-bias events are shown in Figs. 12 (d) and (e).

Figures 13 (a), (b) and (c) show the x and y components of the E_T (E_x and E_y) distribution in dijet events for three different reconstruction techniques. The $\text{tc}\Sigma E_T$ and $\text{pf}\Sigma E_T$ distributions in dijet events are shown in Figs. 13 (d) and (e).

Figures 13 (f), (g), and (h) show the azimuthal angle between the \vec{E}_T and the next-to-leading jet. In events containing hadronic jets, E_T is often introduced by jet mismeasurement and a

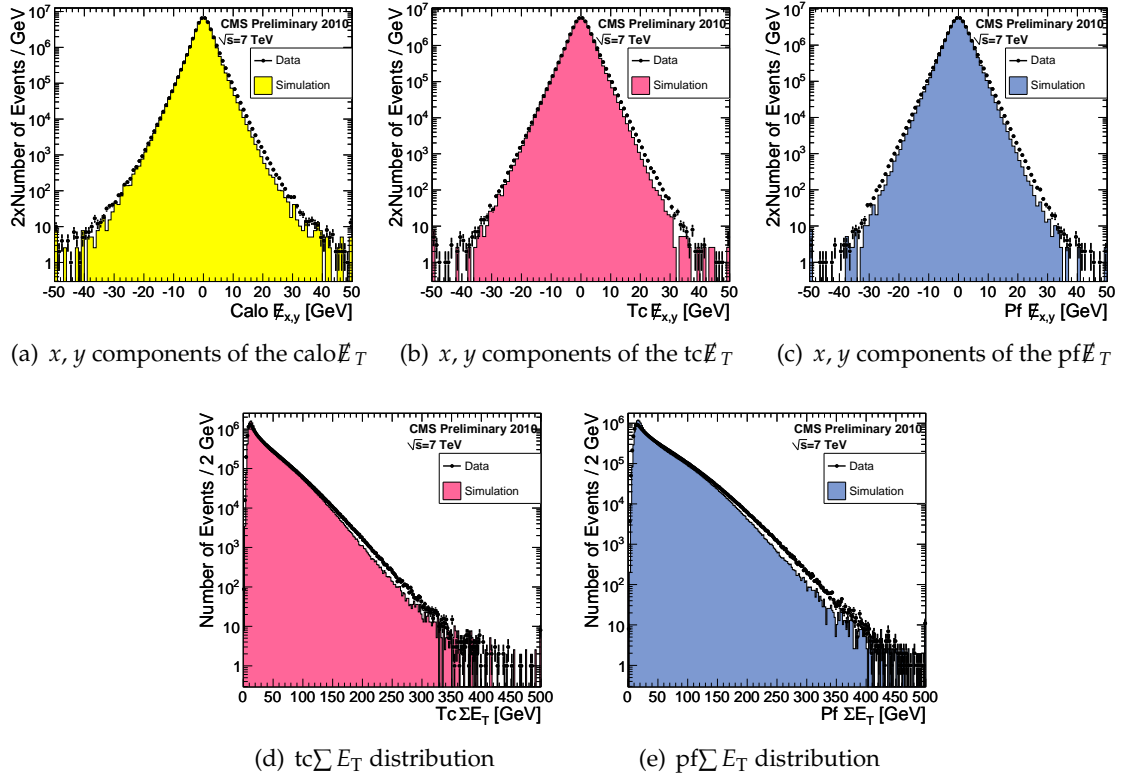


Figure 12: Distributions of x, y components of the $\text{calo}E_T$, $\text{tc}E_T$, $\text{pf}E_T$, respectively, and $\text{tc}\Sigma E_T$ and $\text{pf}\Sigma E_T$ in minimum-bias data compared with Monte Carlo simulation.

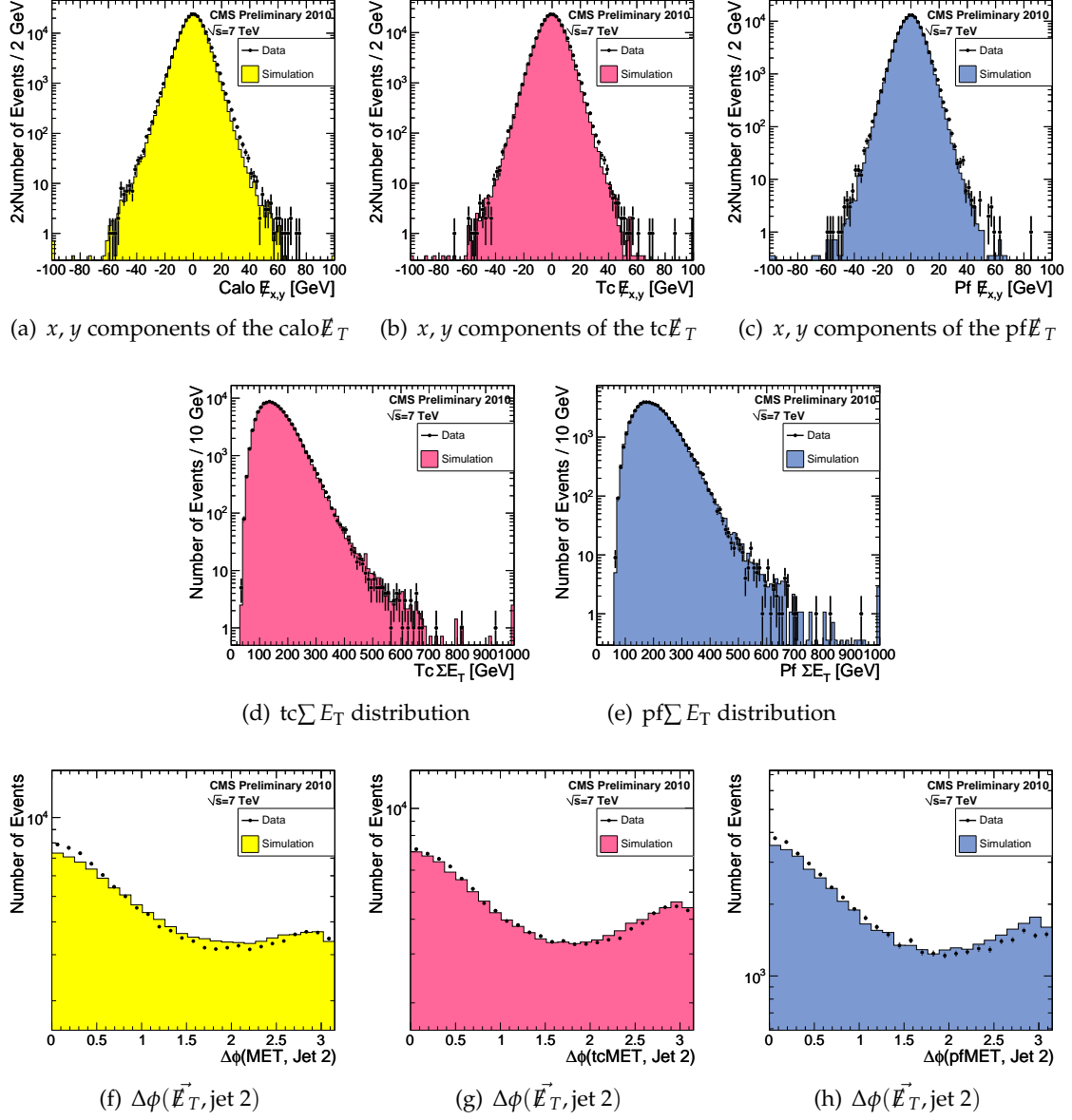


Figure 13: Distributions of x, y components of the $\text{calo}\vec{E}_T$, $\text{tc}\vec{E}_T$, $\text{pf}\vec{E}_T$, respectively (top row), $\text{tc}\Sigma E_T$ and $\text{pf}\Sigma E_T$ (middle row), and azimuthal angle between \vec{E}_T and the next-to-leading jet for $\text{calo}\vec{E}_T$, $\text{tc}\vec{E}_T$, $\text{pf}\vec{E}_T$, respectively (bottom row) in the inclusive dijet data ($p_T^{\text{jet1,2}} > 25$ GeV and $|\eta^{\text{jet1,2}}| < 3$) compared with Monte Carlo simulation.

clear angular correlation between \vec{E}_T and the next-to-leading jet is observed. $\Delta\phi(\vec{E}_T, \text{jet } 2)$ peaks toward low values as expected from the jet mismeasurement, and this behaviour is well modeled by the Monte Carlo simulation.

B Decomposition of \vec{E}_T using the Transverse Thrust Axis

In hadron-hadron collisions spherical symmetry is lost and event shape variables are re-casted in the transverse plane. For example, the Fox-Wolfram moments [17] when applied directly to hadron-hadron collisions interpret a minimum-bias event as a two-jet event.

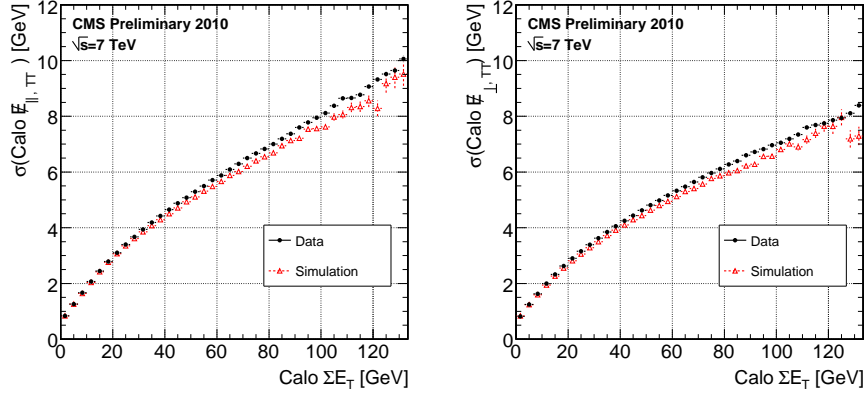


Figure 14: Resolution of calorimeter $E_{T\parallel}$ and $E_{T\perp}$ with respect to the event thrust axis as a function of the calorimeter $\sum E_T$. Data distributions are compared to the Monte Carlo simulation.

The Transverse Thrust (TT) axis is introduced here to decompose the canonical calorimetric missing transverse energy vector in a physics motivated reference frame. We calculate the TT axis in the transverse plane by maximizing the sum of the track momenta component along it, as follows:

$$TT = \max_{\phi_{TT}} \frac{\sum_i |p_T^i \cos(\phi_{TT} - \phi_i)|}{\sum_i p_T^i} \quad (3)$$

where ϕ denotes the azimuthal angle, p_T the transverse momentum and the index i runs over all high-purity tracks in the event. The E_T in the minimum-bias events can now be decomposed into two orthogonal components, denoted $E_{T\perp}$ and $E_{T\parallel}$ and corresponding to the E_T components perpendicular and parallel to the transverse thrust axis direction respectively:

$$E_{T\parallel} = \vec{E}_T \cdot \vec{p}_T^{TT} / |\vec{p}_T^{TT}|, \quad E_{T\perp} = \sqrt{|\vec{E}_T|^2 - |E_{T\parallel}|^2}. \quad (4)$$

An advantage of this decomposition is the use of the tracking in determining the axis of reference to project calorimetric measurements. Tracking provides a highly resolved direction of the momentum in the event independent of the calorimetry measurement. The parallel component of the calorimetric transverse energy is dominated by the calorimeter response to energy deposition while the perpendicular component is dominated by the ambient calorimeter noise.

The perpendicular and parallel components of E_T Gaussian core resolution are shown in Fig. 14 for a calorimeter-only E_T measurement. In both components the data-Monte Carlo agreement is satisfactory. There is a pure noise component of the resolution at low $\sum E_T$ and a resolution component encapsulating the response of the calorimeter to higher energy deposits at higher $\sum E_T$. Residual disagreements between the data and Monte Carlo at high energies are due to the simulation of the response of the calorimeter. The work needed to extract the calorimetric response from the data is underway. The agreement of the data at lower energies with the Monte Carlo simulation signifies a fair calorimetric noise model in the simulation.

C E_T in Heavy-Flavour-Enriched Events

In this section, results from a study of events containing b -quarks are presented. Events containing b -quarks can have different E_T distributions, because the B hadrons have unique fragmentation properties, and they sometimes lead to a final state containing neutrinos.

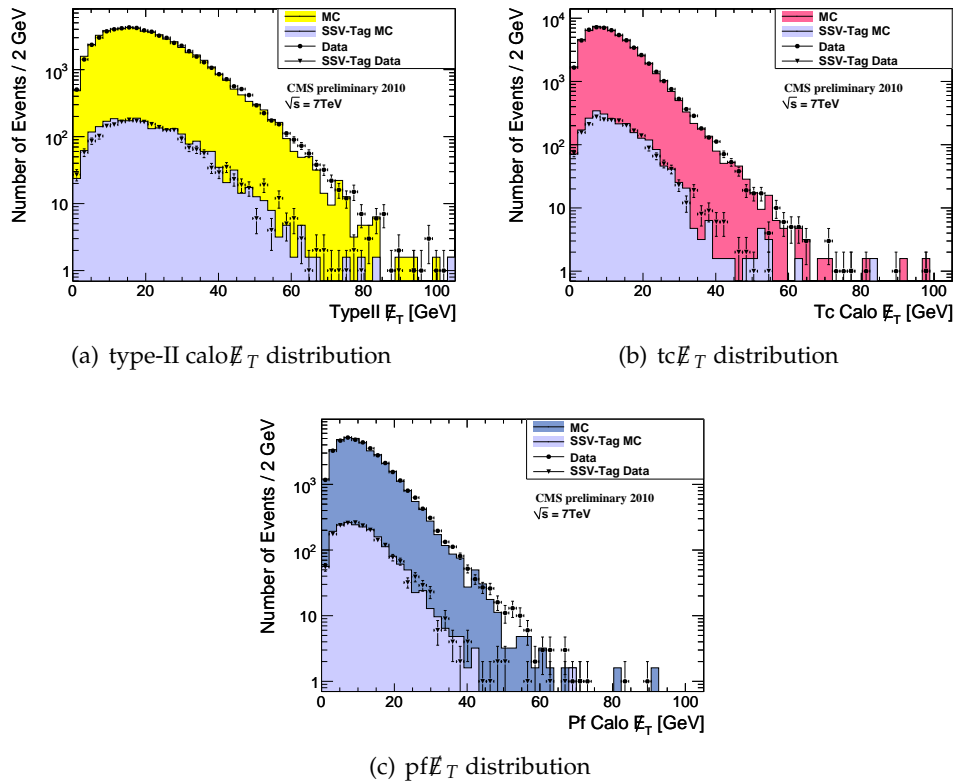


Figure 15: Distributions of (a) type-II corrected calo \cancel{E}_T , (b) tc \cancel{E}_T , and (c) pf \cancel{E}_T in dijet events with and without b -tagging.

In Fig. 15 we compare \cancel{E}_T in dijet events with and without a secondary vertex, *i.e.*, events where the leading jet has a positive SimpleSecondaryVertex (SSV) tag [18, 19]. Since this tagger relies on tracking information, we require that the leading jet has $|\eta| < 2.1$. Also, we require the leading two jets to have $p_T > 40$ GeV. Below 40 GeV, the b -quark tagging purity goes down significantly. Table 1 shows the mean values for the \cancel{E}_T distributions in Fig. 15 for comparison. The b -tagged dijet events tend to have a mean which is somewhat higher than the non b -tagged dijet events; however, the effect is small. Table 1 and Fig. 15 show that the Monte Carlo provides a reasonable description of the data distributions.

Table 1: Means and error on means of the \cancel{E}_T distributions for b -tagged dijet events of Fig. 15. All numbers in GeV, and the errors are statistical only.

	type-II calo \cancel{E}_T	tc \cancel{E}_T	pf \cancel{E}_T
no b -tag (data)	20.76 ± 0.05	12.08 ± 0.03	11.55 ± 0.04
no b -tag (MC)	20.47 ± 0.06	11.93 ± 0.04	11.33 ± 0.05
b -tagged (data)	21.19 ± 0.25	13.02 ± 0.17	11.73 ± 0.17
b -tagged (MC)	20.62 ± 0.30	12.60 ± 0.20	11.46 ± 0.20

References

- [1] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004.

- [2] CMS Collaboration, “Missing E_T Performance in CMS”, *CMS Physics Analysis Summary JME-07-001* (2007).
- [3] CMS Collaboration, “Track-corrected Missing Transverse Energy in CMS”, *CMS Physics Analysis Summary JME-09-010* (2009).
- [4] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS”, *CMS Physics Analysis Summary PFT-09-001* (2009).
- [5] CMS Collaboration, “CMS Tracking Performance Results from early LHC Operation”, arXiv:1007.1988.
- [6] T. Sjostrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1”, *Comput. Phys. Commun.* **178** (2008) 852–867, arXiv:0710.3820. doi:10.1016/j.cpc.2008.01.036.
- [7] Geant4 Collaboration, “Geant4: A simulation toolkit”, *Nucl. Instrum. and Methods A* **506** (2003) 250. doi:10.1016/S0168-9002(03)01368-8.
- [8] CMS Collaboration, “CMS Jet Performance in pp Collisions at $\sqrt{s} = 7$ TeV”, *CMS Physics Analysis Summary JME-10-003* (2010).
- [9] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-kt jet clustering algorithm”, *JHEP* **0804:063** (2008).
- [10] CMS Collaboration, “HCAL performance from first collisions data”, *CMS Detector Performance Summary DPS-2010/025* (2010).
- [11] CMS Collaboration, “Electromagnetic calorimeter commissioning and first results with 7 TeV data”, *CMS Note* **2010/012** (2010).
- [12] CMS Collaboration, “Commissioning of the Particle-Flow Event reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV”, *CMS Physics Analysis Summary PFT-10-002* (2010).
- [13] CMS Collaboration, “Single Particle Response in the CMS Calorimeters”, *CMS Physics Analysis Summary JME-10-008* (2010).
- [14] CMS Collaboration, “Jet and MET Commissioning Results from 7 TeV Collision Data”, *CMS Detector Performance Summary DPS-2010/014* (2010).
- [15] R. Field, Studying the Underlying Event at CDF and the LHC, Proc. of the 1st Int. Workshop on MPI at the LHC, Perugia, Italy, October 27-31, 2008, P. Bartalini and L. Fano, eds., arXiv:1003.422. DESY-PROC-2009-06.
- [16] CMS Collaboration, “The Jet Plus Tracks Algorithm”, *CMS Physics Analysis Summary JME-09-002* (2009).
- [17] G. C. Fox and S. Wolfram *Nucl. Phys.* **B143** (1979) 413.
- [18] CMS Collaboration, “Algorithms for b Jet Identification in CMS”, *CMS Physics Analysis Summary BTV-09-001* (2009).
- [19] CMS Collaboration, “Commissioning of b-jet identification with pp collisions at $\sqrt{s} = 7$ TeV”, *CMS Physics Analysis Summary BTV-10-001* (2010).