Available on CMS information server CMS NOTE 2008/XXX

The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

28 May 2008

Studies of Simulated QCD Events with Anomalous Missing Transverse Energy

J. Temple

University of Maryland, College Park, Maryland

Abstract

This note describes studies of QCD multijet events with anomalously large missing transverse energy (MET). Events from the CSA07 METHIGH_SKIM with more than 150 GeV of reconstructed MET and less than 50 GeV of generated MET were investigated. A significant fraction of these events were affected by GEANT bugs regarding long-lived heavy and strange hadrons. The remaining events appear to be due to expected jet resolution effects.

1 Introduction

Missing transverse energy (or MET) is expected to be a useful variable both for identifying the presence of known particles (such as neutrinos) and for detecting previously undiscovered particles (e.g., neutral supersymmetic particles) at CMS. Correct measurement of the MET requires a fully functioning detector with high hermiticity as well as a good understanding of the responses and corrections of the detector components. In order to using MET as a discriminating variable, its properties must also be understood in the simulated events used to mimic physics signals. This note details a study of MET within simulated QCD events, concentrating on events with large missing transverse energy. As QCD events should contain no intrinsic missing energy, large MET values may indicate a problem with the generation or simulation of these signals. Large MET values may also give insight into the causes of tails in the MET resolution of the detector.

2 Event Sample and Selection

Events were chosen from the CSA07 "Gumbo" dataset, consisting to QCD, minimum bias, and photon+jet events, in the METHIGH_SKIM (/CSA07JetMET/CMSSW_1_6_7-CSA07-Gumbo-B1-PDJetMET-ReReco-100pb-Skims0-metHigh_FEVTSIM/USER) generated with CMSSW_1_4 and reconstructed with CMSSW_1_6_7. This is a skim of 1.1 million events which pass the HLT1MET trigger (requiring 65 GeV of reconstructed MET). Further cuts were then applied in order to select events with MET that could not be ascribed to jet resolution effects. These cuts are described in detail below.

2.1 Reconstructed and Generated MET Cuts

The reconstructed missing transverse energy in each event was corrected for jet energy scale effects using the iterativeCone5CaloJets corrections in "JetMETCorrections/Type1MET/data/MetType1Corrections.cff". The resulting MET, $corMetType1Icone5$, was required to have a value of at least 150 GeV. In addition, the generator-level MET ($genMet$) was required to be less than 50 GeV, in order to eliminate events containing significant intrinsic MET.

2.2 Jet-MET Cuts

Jets were reconstructed using the "iterative Cone" algorithm, forming jets based on a $\Delta R = 0.5$ cone. Jet energy scale corrections were applied to the jets using "JetMETCorrections/MCJet/data/ MCJetCorrections152.cff", to form $MCJetCorJetIcone5$ jets. Only jets with $E_T > 20$ GeV were included in this study. Unless otherwise noted, all further references to "reconstructed" jets or MET in this note refer to the energy scale-corrected objects.

Due to inherent jet resolution and reconstruction uncertainties, the energies of the corrected and generated jets are not expected to exactly match. In an event containing high energy jets, small relative differences between the generator-level and reconstructed jets can lead to significant missing transverse energy. This behavior is expected, and such an event should not be considered as anomalous.

For events in which MET arises solely due to jet resolution effects, the MET is expected to point either along or against the jet axis. In order to exclude such events from this study, a cut was placed on the azimuthal angle between the MET and each of the leading two jets in the event. For each jet j_i (i=1 or 2), events were excluded if $\Delta\phi(j_i, MET) < 0.349$ (20 degrees) or $(\pi - \Delta\phi(j_i, MET)) < 0.349$.

3 GEANT Bug with $b-$ and $c-$ Hadrons

Even after the cuts described in sections 2.1 and 2.2 were applied, a large number of events were seen in which the MET was aligned against the direction of one of the leading two jets (see Figure 2). Event displays indicated that such events contained reconstructed jets with energies far less than their generator-level counterparts, as shown in Figure 3.

Further investigations by Seema Sharma and Shuichi Kunori led to the discovery of a bug regarding the tracking of long-lived decay particles in GEANT. (Here, "long-lived" refers to the decay length as defined through a runtime parameter, and was set to 100 mm for all events in this study.) GEANT4 provides a mechanism to decay such particles and then make use of the daughters of the decays. However, an incorrect implementation of code in CMSSW caused the kinematic quantities of the daughters to be computed in the rest frame of the mother particle rather than the laboratory frame. For instance, in the decay $B \to D \to K + \pi$, the energies of the final kaon and pion were computed in the rest frame of the D, and thus ignore the energy of the initial B.

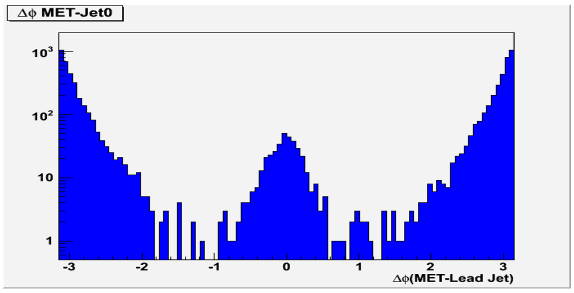


Figure 1: Δϕ(lead jet,MET) prior to applying Δϕ cut.

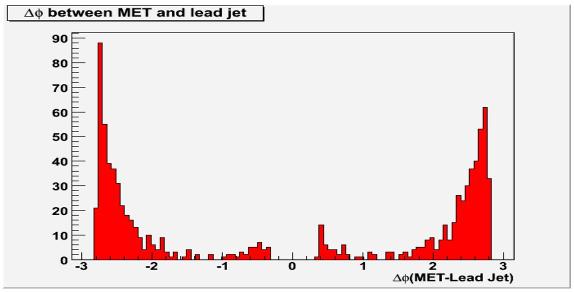


Figure 2: Δϕ(lead jet,MET) after applying cut.

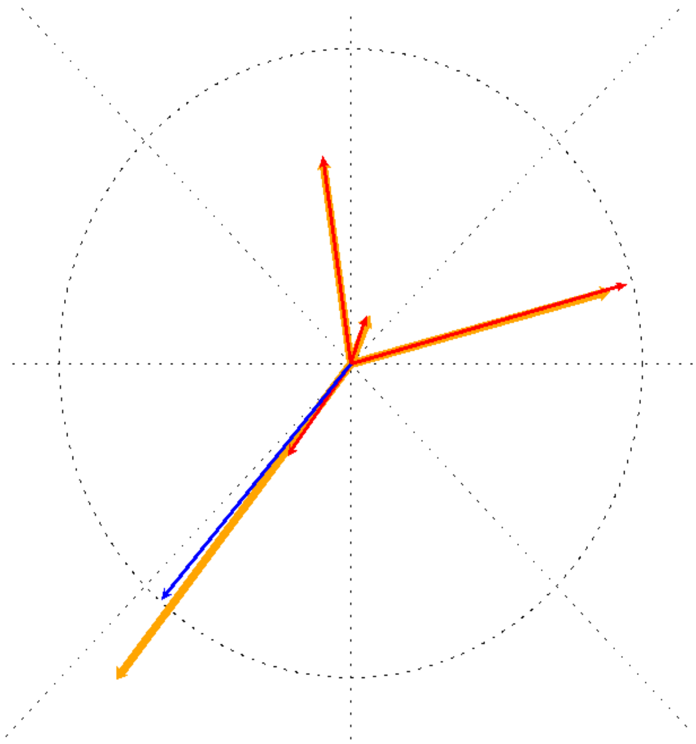


Figure 3: $x-y$ schematic view of an event containing a jet with reconstructed energy much less than its generator-level energy. Reconstructed jets appear in red, generator jets in orange, and the reconstructed MET in blue. The dashed circle represents 500 GeV of transverse energy.

In events with long-lived heavy particles (particularly b- and c- hadrons), this bug caused severe violations of energy conservation, leading to significant MET. A fix for this bug was added to V00-02-16-01 of the SimG4Core/Generators package. However, that fix obviously could not be applied to the existing CSA07 samples. For those samples, Sunanda Banerjee created a simple test to determine whether an event could be affected by the bug:

```
bool badEvent = false;
edm::Handle<SimTrackContainer> simTracks;
iEvent.getByLabel("g4SimHits",simTracks);
SimTrackContainer::const_iterator simTrack;
for (simTrack = simTracks->begin(); simTrack != simTracks->end(); ++simTrack){
  if ((*simTrack).genpartIndex() > 0) {
    if ((*simTrack).momentum().mag() > 1.8) badEvent = true;
  }
```

This code checks for GEANT simTracks from particles with masses greater than 1.8 GeV, and sets the boolean *badEvent* flag to "true" is any such particles are found. In this analysis, it was found that roughly half the events passing the cuts detailed in Section 2 were flagged as bad events. Such events were discarded from the analysis.

It should be noted that simply discarding such events may not be the best procedure for all analyses. Removing these events can have a large effect on the resulting kinematic distributions of the sample, with the size of the effect depending on the fraction of events flagged as bad. For CSA07 QCD samples with p_T-hat between 380-600 GeV, roughly 1.5 % of all events have a "true" *badEvent* flag. However, for QCD events in the METHIGH_SKIM sample, that fraction rises to approximately 5%, and reaches 8% in a CSA07 sample of Z' decays. The effect of this bug must be independently determined on an analysis-by-analysis basis.

4 Revised selection cuts

After the GEANT bug was discovered, events in the METHIGH_SKIM with a "true" boolean flag were discarded, and the cuts from section 2 were reapplied. The resulting number of events passing those cuts is shown in Table 1 and Figure 4. As expected, events with the GEANT bug have tend to have a larger MET significance than events without the bug. This is due to the erroneously small reconstruction energy of at least one jet in each such event, which simultaneously lowers the $\sum E_T$ of the event and increases MET.

Table 1: Number of METHIGH_SKIM events passing each selection cut

Selection Cut	Events Remaining
No cut	1136158
GEANT *badevent*=false	1078664
Generator MET $<$ 50 GeV	980204
Reconstructed MET $>$ 150 GeV	65107
$\Delta\phi(lead\,jet, MET) > 0.20, < \pi - 0.20$	6007
$\Delta\phi(second\,jet, MET)$¿0.20,¡$pi−0.20	3506

4.1 MET Significance Cuts

Further investigations indicated that many events passing the above cuts contained two highly energetic jets, with energies of 1 TeV or more. Because the energy of such jets was so large, small differences between the angle of the reconstructed and generated jet led to significant MET that did not point along the jet axis. A sample event displaying this behavior is shown in Figure 4.1. In such events, the MET angle fell outside the $\Delta\Phi$ cut described in Section 2.2, and these events were thus not excluded by the default event selection. However, the MET in these events was produced via expected fluctuations in jet reconstruction, and these events should thus not be considered "anomalous".

It was observed that, although the MET of such events was large, the MET significance, defined as the ratio of the MET to the scalar sum of transverse energy in the event:

$$\sigma_{MET} = |MET| / \sum E_T, \quad (1)$$

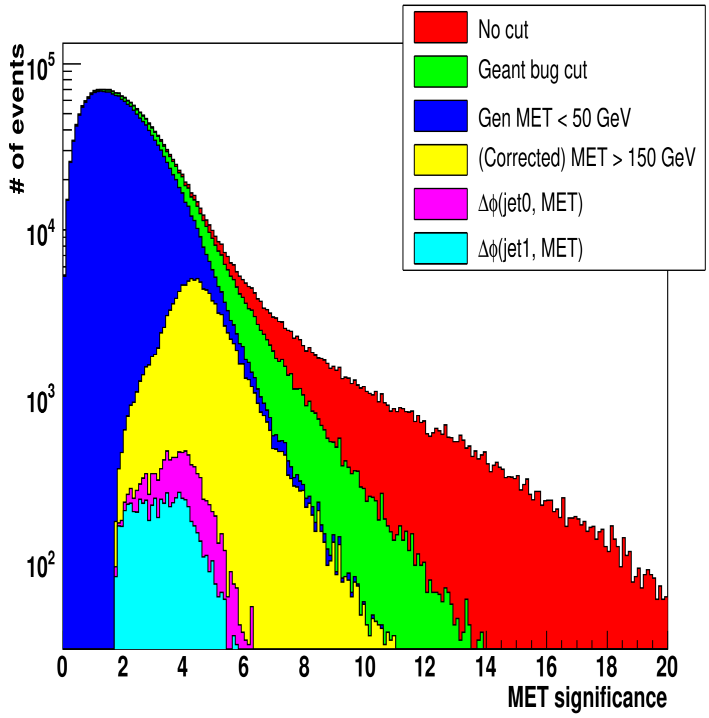


Figure 4: Number of events passing each selection cut as a function of MET significance.

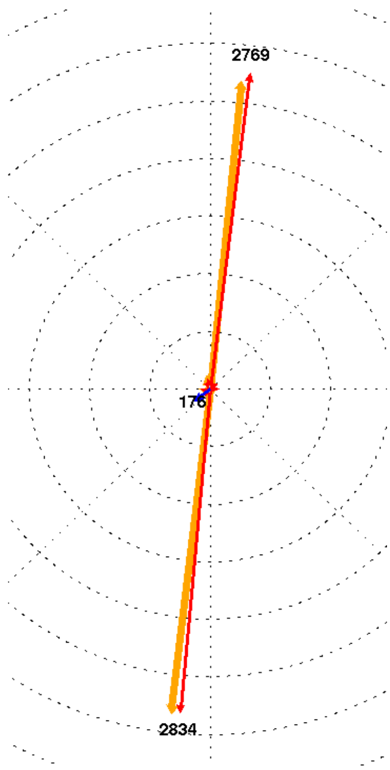


Figure 5: Sample event with MET$>$ 150 GeV produced by small reconstruction errors in two high-p_T jets.

was relatively small. In the largest p_T^{HAT} bin (.ge.3500 GeV, the MET in all observed events passing the selection cuts was due to small differences between generated and reconstructed jets. The significance of most of these events was found to be under 4, as shown in Figure 4.1. A cut requiring MET significance > 4 was introduced in order to eliminate such events from the sample. After this cut was applied to the 3506 events passing the original selection cuts, 1041 events remained.

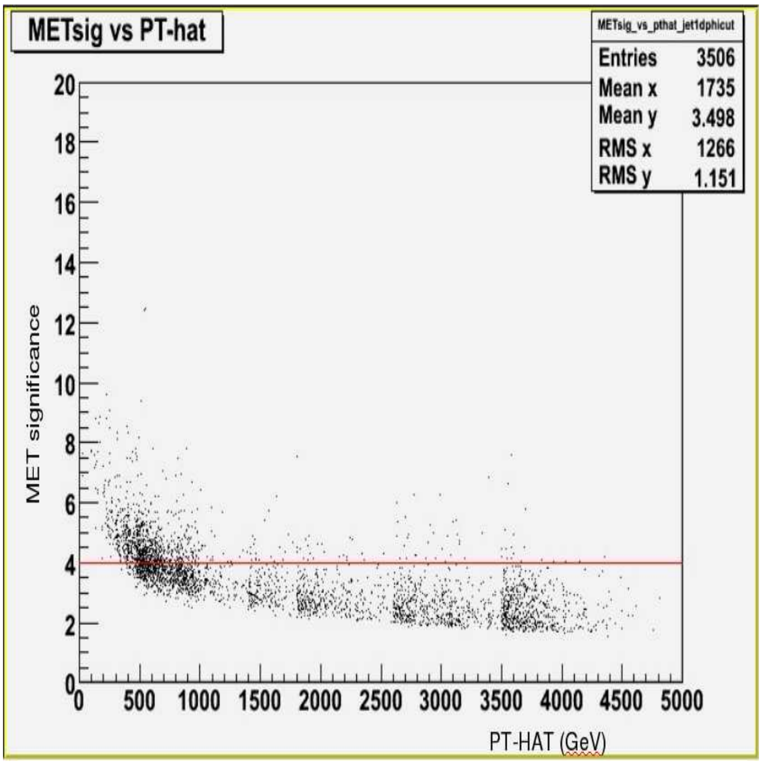


Figure 6: MET significance vs. p_T^{HAT} for events passing basic selection cuts. The line in red indicates an additional applied cut requiring significance $>$4.

5 Categorization of Anomalous Events

Examination of the 1041 events passing all cuts indicated that the majority of events fell into one of three "failure mode" categories:

- Events with an extra reconstructed jet not associated with a generator-level jet;
- Events with an extra generator-level jet;
- Events with significant differences between generated and reconstructed jet energies.

Each of these categories will be discussed in turn below, and the cause of each failure mode will be explained.

5.1 Events with Extra Reconstructed Jets

Of the 1041 selected anomalous events, 87 contained an extra reconstructed jet that did not appear to be correlated with any generated jet in the event. Extra jets are identified as reconstructed jets with E_T greater than 50 GeV for which there no associated generator jets greater than 20 GeV within a ΔR of 0.3.

One such event is shown in Figure 5.1. In such an event, the presence of an extra jet resulted in large reconstructed MET opposite the direction of the jet. Studies of generator-level particles associated with these jets found that dominant energy contributor to the jets tended to be strange particles (mainly Λ, Ξ, and Σ). Further investigations found that there was a large difference in angular positions between the generator particles and the jets, as shown in Figure 5.1.

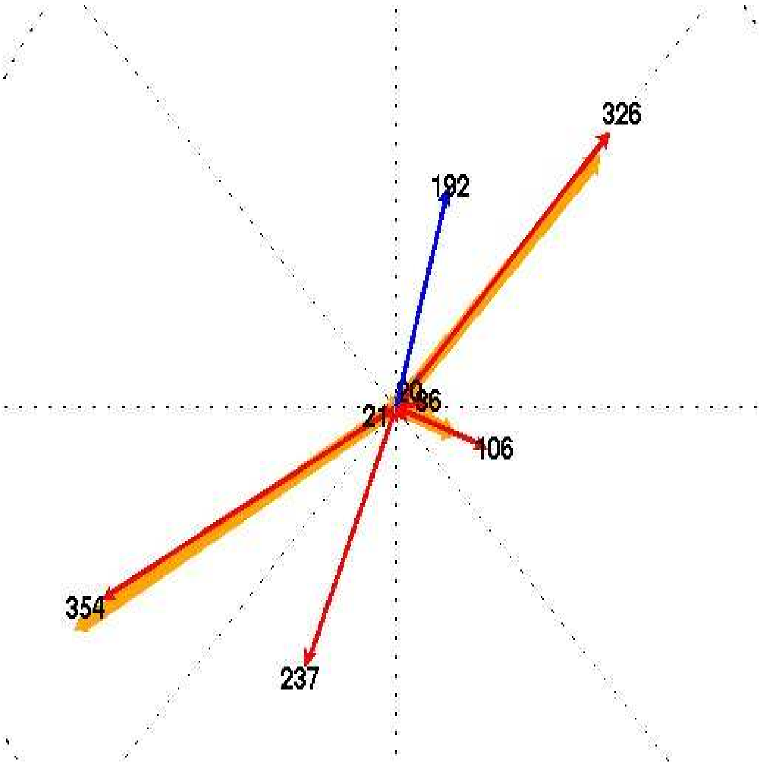


Figure 7: An event with an “extra” 237 GeV reconstructed jet.

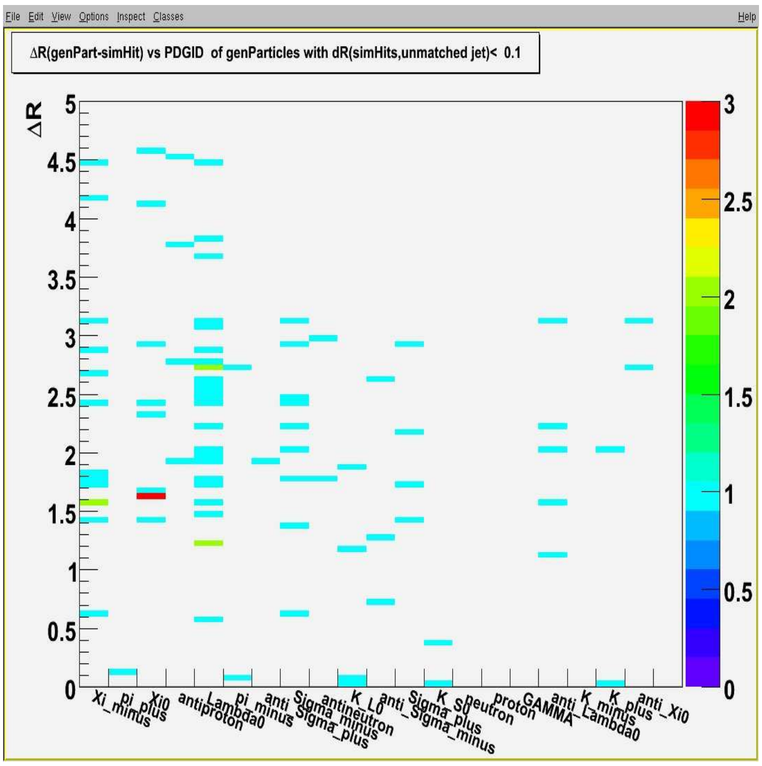


Figure 8: ΔR between “extra” reconstructed jets and the largest-energy generator-level particle in the jet. The events are dominated by strange particles with $\Delta R > 1$ from the reconstructed jet.

This behavior was traced to a limitation of GEANT4's handling of long-lived strange baryons. For strange particles interacting in the detector, energy-momentum conservation may be violated, resulting in a decayed daughter particle having far more energy that its parent particle in the rest frame of the parent. This seems related to the reference frame bug detailed in Section 3. At the current time (17 June 2008), two proposals have been offered for dealing with this problem:

- provide a fix inside GEANT4 for this issue;
- provide a momentum check within the CMS simulation that will skip events with large momentum imbalances.

The timescale for implementing either of these solutions has not yet been provided. An alternate fix of ignoring events containing strange baryons with decay lengths $\beta\gamma c\tau > 10$ cm has also been suggested. However, it is not yet known how such a cut would affect overall sample distributions, or whether that cut is sufficient to eliminate all instances of this bug. Further investigations into this problem are currently being conducted.

5.2 Events with Extra Generator-Level Jets

Conversely, a small number (13 out of 1041) events contain a high-p_T generator jet that is not matched to a reconstructed jet. These events were found to be the result of two generator-level jets being merged into a single reconstructed jet. In 5 of these events, the errors in reconstructed energy due to the jet merging provide a negligible contribution to the MET, as in Figure 9. For the remaining 8 events, though, the direction and magnitude of the MET is related to the merged jets, as in Figure 10. However, it cannot be determined whether the large missing energy is due solely to the merging of the jets or to simple jet resolution effects. It is known that the jet energy scale corrections do not scale linearly with jet energy, so treating two jets as one can result in an overcorrection that can affect the MET.

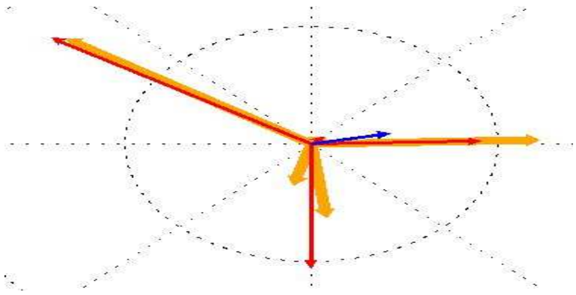


Figure 9: Event display for an event in which two generator-level jets (pointing downward) are merged into one reconstructed jet.

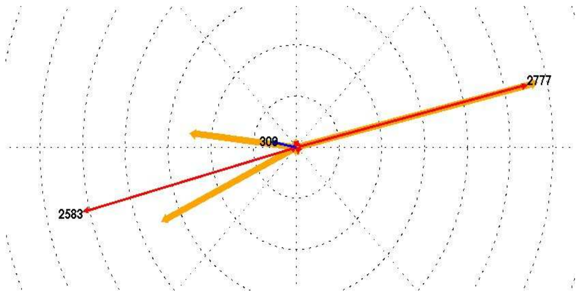


Figure 10: Event display for an event in which jet merging results in large MET.

5.3 Events with Generator-Reconstructed Energy Disagreements

The remaining events in the selected sample have all reconstructed jets > 20 GeV matched to a corresponding generator jet. MET arises in these events purely from differences in reconstructed and generated jet momenta. In these events, the E_T difference $\vec{\Delta}(ET)$ between each reconstructed jet and its associated generator-level jet is calculated. For reconstructed jets in which this difference is greater than 2σ, where σ is estimated as:

$$\sigma = 125\% \sqrt{(}E_T), \tag{2}$$

a categorization variable α is defined. This variable is a comparison between the projection of the jet energy difference onto the MET axis and the overall MET:

$$\alpha = [|\vec{\Delta}(ET)|/MET] cos(\phi_{\vec{\Delta}(ET)} - \phi_{MET}). \tag{3}$$

Each reconstructed jet is assigned a label based on its α value, according to the categories defined in Table 2. The jet is further classified by whether its energy is larger or smaller than that of its generator counterpart. If larger, the jet is labeled as "over-reco"; otherwise, it is labeled "under-reco".

Events are classified according to the jets with the largest α value. For example, an event with one "large over-reco" and two "medium under-reco" is given an overall label of "large over-reco". If an event contains both an over-reco and under-reco jet with the same α category, the event is given a separate label "both". A list of the number of events in each category is given in Table 3.

Table 2: Categories of jets in high-MET events

Category Name	α value
Biggie	$\alpha > 0.5$
Large	$0.5 \geq \alpha > 0.375$
Medium	$0.375 \geq \alpha > 0.25$
Small	$0.25 \geq \alpha > 0.10$

Table 3: Number of events in each MET category

Category Name	Number of Events
Biggie over-reco	268
Biggie under-reco	246
Biggie both	8
Large over-reco	154
Large under-reco	61
Large both	26
Medium over-reco	36
Medium under-reco	52
Medium both	42
Small over-reco	15
Small under-reco	8
Small both	5

As indicated in Table 3, the remaining high-MET events are dominated by events with either "biggie" or "large" discrepancies between reconstructed and generated jets. Additional investigations show that most of these events contain only a single poorly reconstructed jet (see Figures 11 and 12). Furthermore, the reconstructed energies are not egregiously different from the generated values, as Figures 13 and 14 demonstrate.

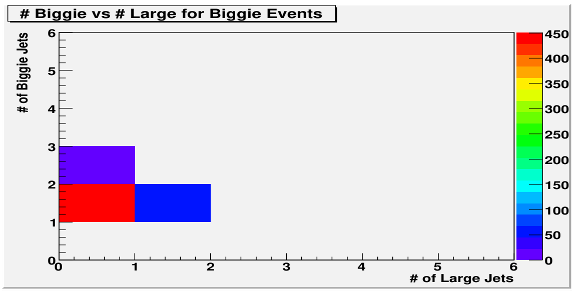
Figure 5.3 illustrates a sample Biggie event. In this event, the major contribution to the MET is due to the generator-reco differences of the leading jet. This jet has an energy of 1261 GeV, with a corresponding resolution of $1.25 * \sqrt{(1261)} = 35.5$ GeV. Its corresponding generator jet has an energy of 1132 GeV, and thus differs from the reco jet by more than 2σ but less than 3σ. Likewise, the overall MET significance in this event is 4.01, indicating that the event nearly failed the significance cut.

Other events show similar behavior. There are no obvious bugs or oddities causing the large MET. Instead, the MET is the result of expected resolution effects in high-energy events. The number of such events may be reduced by making a better calculation of jet energy resolution than is done in Equation 2. It has also been suggested that the definition of MET significance should be altered to better describe the MET resolution at large E_T. Either change would affect the number of events in the α categories, but neither would affect the overall conclusion that these events appear to be the results of normal fluctuations in jet energies.

5.4 Unknown Failure Modes

There remain 20 events that do not fall into any of the categories above. Event displays for a few of these events are shown in Figures 5.4 - 5.4. The majority of these events have identifiable failure modes, even though they do not meet the failure criteria described above.

The event shown in Figure 5.4 contains two very high-E_T jets, with energies of roughly 4 TeV each. The associated generator jets are each within 2σ of their reco partners, and so no jet is assigned to one of the bad jet categories. Although there is good agreement between generated and reco jets, the energy scale of these objects is so large that very slight differences lead to nearly 400 GeV of missing transverse energy. The MET significance cut is designed to remove such events, but this one just passes the cut, with a significance of 4.10.



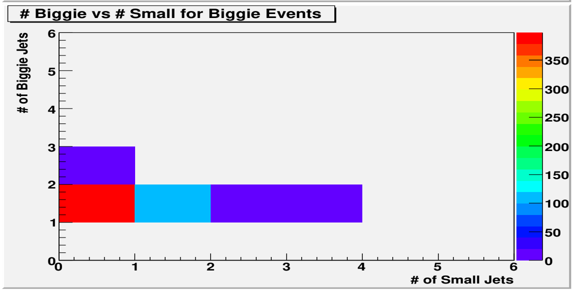
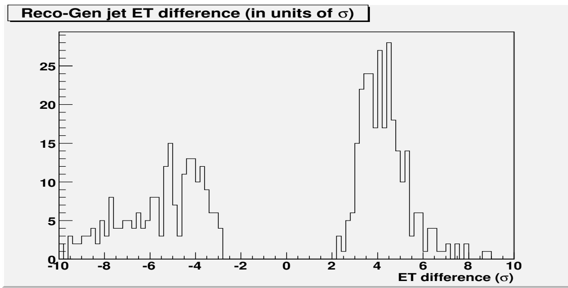


Figure 11: Number of "Biggie" jets vs. Number of "Large" jets in "Biggie" events.

Figure 12: Number of "Biggie" jets vs. Number of "Small" jets in "Biggie" events.



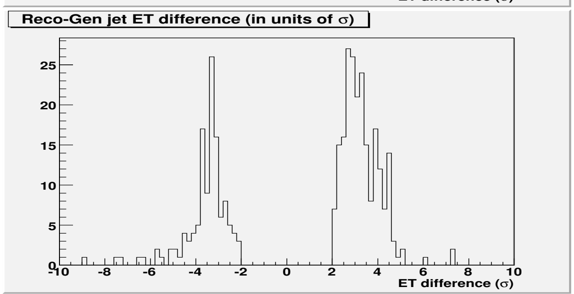


Figure 13: Difference between reco and generator "Biggie" jets. Jets must have difference $> 2\sigma$ to be plotted here. Biggie jet energies tend to be overestimated relative to the generator jets.

Figure 14: Difference between reco and generator "Large" jets. Jets must have difference $> 2\sigma$ to be plotted here.

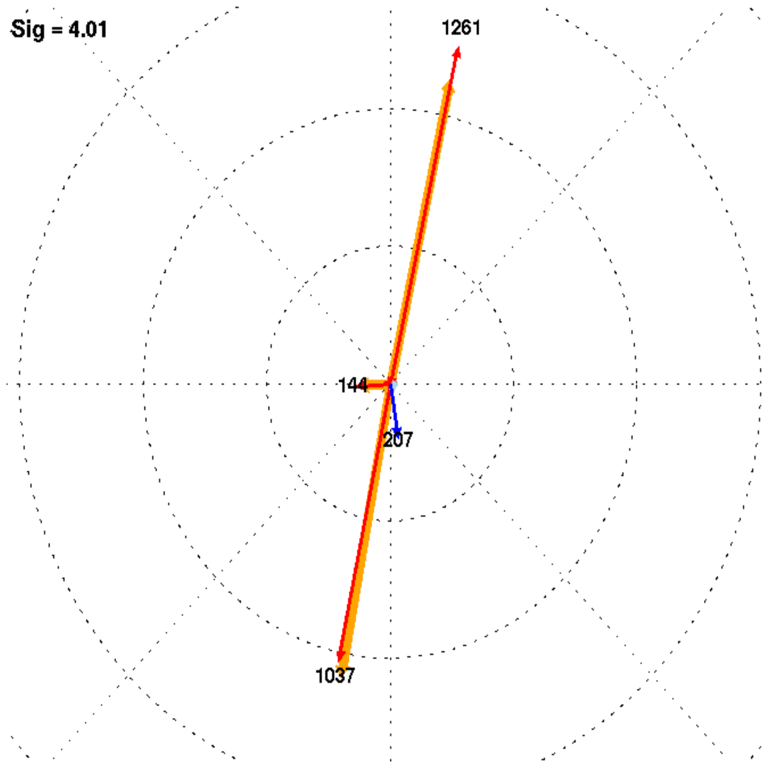


Figure 15: A sample "Biggie" event.

The event in Figure 5.4 has lower-energy jets, with the largest having a value of 680 GeV. It is evident from this event display that there are multiple small differences between generated and reco jets. None of these differences is large enough to trigger a bad jet categorization. However, the net result of these small differences leads to a large missing transverse energy. Again, the MET significance in this event is just above the cutoff value.

The event in Figure 5.4 remains a bit of a mystery. The missing ET from all corrected reconstructed jets greater than 20 GeV in this event is only 69 GeV, far less than the calculated corrected MET of 155 GeV. However, the MET from generated jets in the event is 25 GeV, while the genMet value is only 0.7 GeV. It has been verified that there are no neutrinos or other invisible particles in this event which could be contributing to the missing energy. It is possible that there are a large number of low-p_T jets in this event whose sum significantly changes the overall MET. It may also be that this effect is a result of the GEANT bugs described above. This effect remains under investigation, but it appears to affect no more than 5 of the "unknown failure" events.

6 Muons and Electrons

The effect of muons and electrons on the MET in the QCD events studied here was found to be negligible. Muons were reconstructed using the *globalMuon* objects, and were required to have $p_T > 20$ GeV. The muons were also required to be isolated both the tracked and calorimeter, with a track $\chi^2 < 10$. No such muons were found in the 1041 tested events.

Electrons were reconstructed as *pixelMatchGsfElectrons*, and were also required to have $p_T > 20$ GeV. Loose track and tower isolation requirements were also placed on electrons. There were 104 events containing electrons in the final sample, with the high number due in part to the looser requirements on electrons than muons. However, most of these events contained electrons with p_T just above the cutoff threshold, as in Figure 6. In no events was the electron a significant contributor to the overall missing energy.

7 Conclusions

Over 1 million QCD events in the CSA07 METHIGH_SKIM were studied. Two GEANT bugs were discovered which contributed heavily to the number of events with large missing transverse energy. Simple patches now exist to ignore events containing these bugs, although a true fix does not yet exist. The remaining events were found

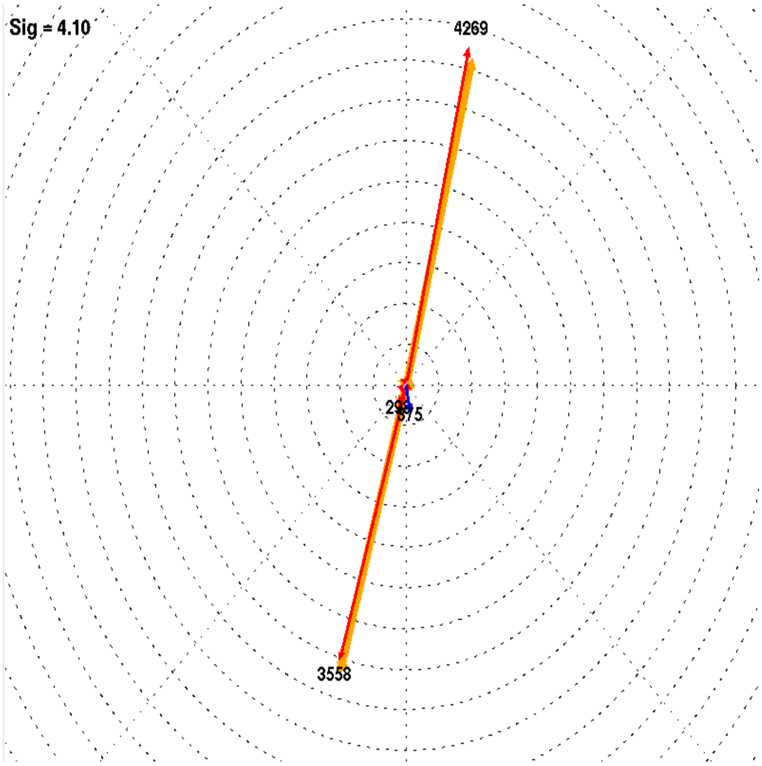


Figure 16: A sample event which does not fall into any of the failure mode categories.

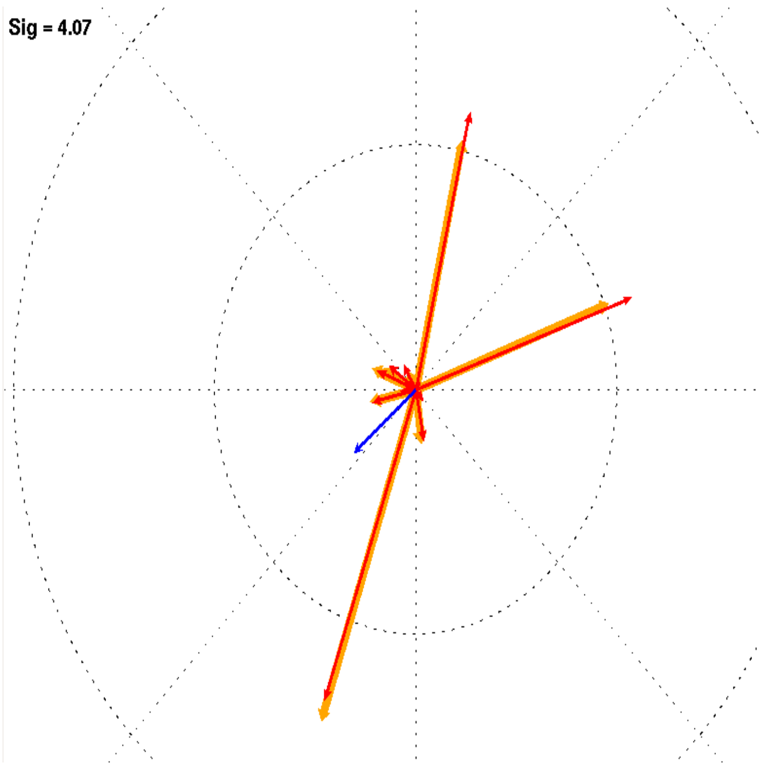


Figure 17: A sample event which does not fall into any of the failure mode categories.

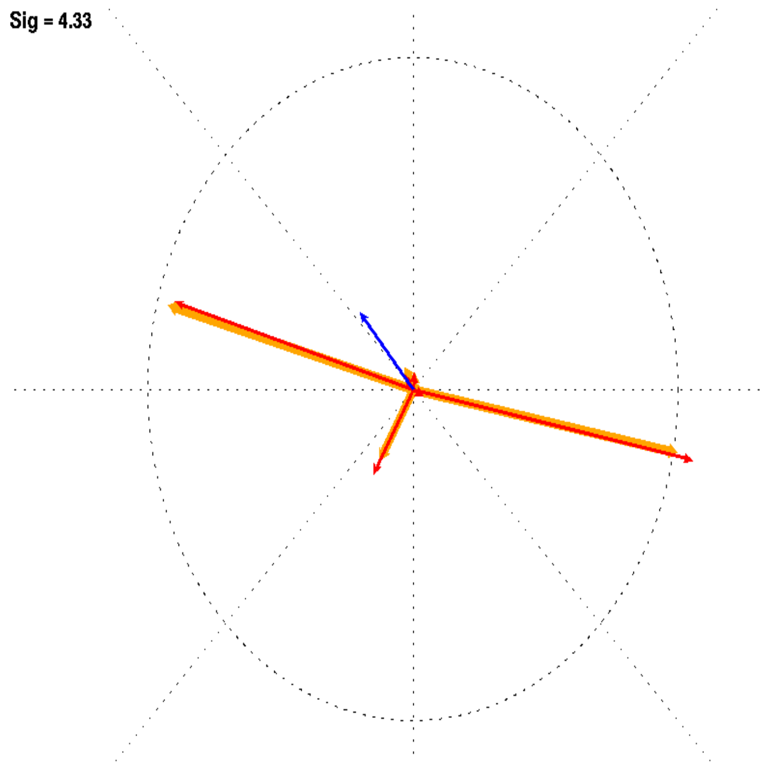


Figure 18: A sample event which does not fall into any of the failure mode categories.

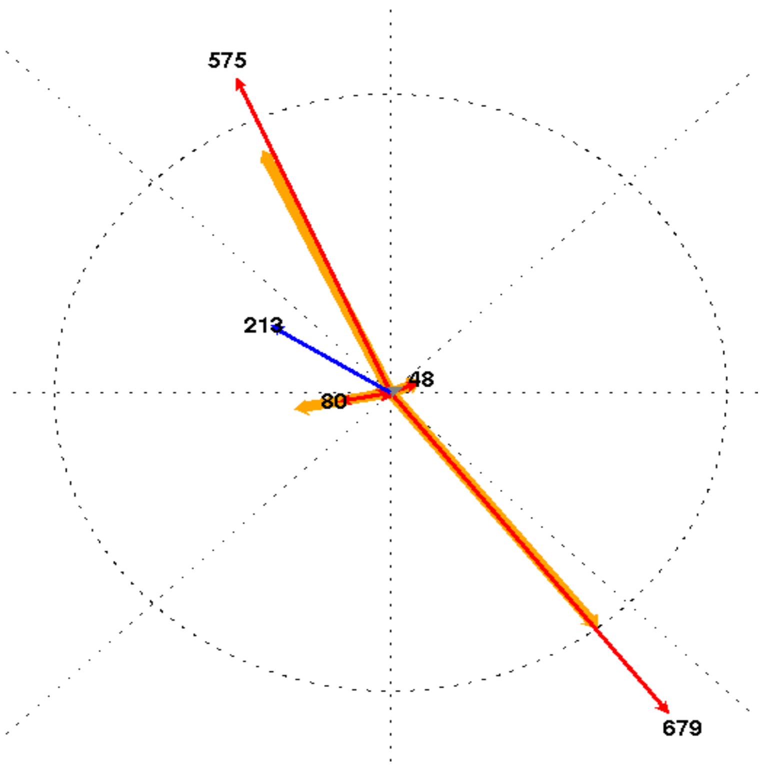


Figure 19: Sample event containing an electron. The electron is just visible as a gray arrow, with an energy just over 20 GeV.

to have significant but not unexpected fluctuations in reconstructed jet energies. These fluctuations led to large missing transverse energy within the events. A handful of events show behaviors which may require further study.