Content

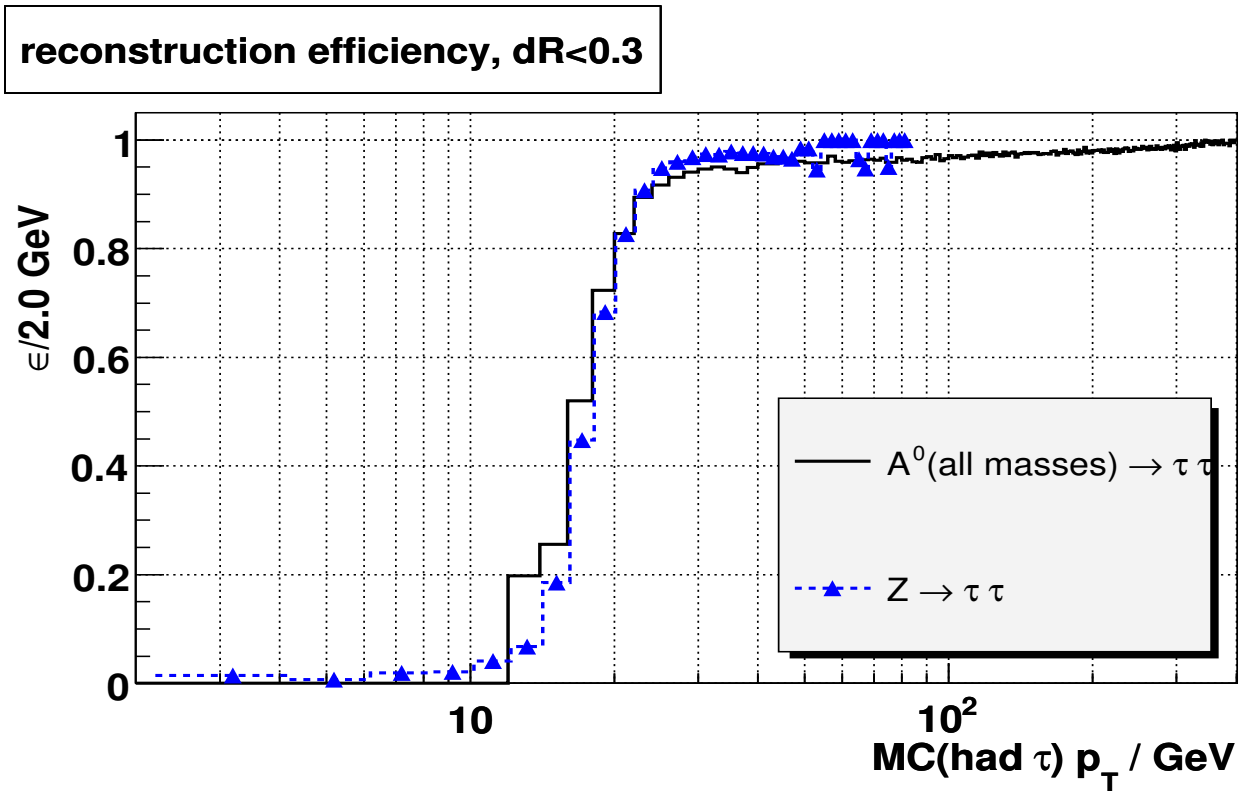
TauRec

Michael Heldmann
Universität Freiburg

- **Reconstruction**
- **Calibration**
- **Identification**
- **Running with TopoClusters**

TauRec (reconstruction)

- TauRec starts from CombinedClusters
- CombinedClusters are clusters from a sliding window algorithm
- Current settings for seeds give the following reconstruction efficiency (e.g. the probability for a mc true hadronic tau to have a reconstructed cluster within $\Delta R<0.3$)
- Because of this turn-on already on reconstruction level tauRec is (if seeded by CombinedClusters) only suitable for tau with $p_T>20GeV$
- TauRec can be seeded (only a change in the jobO needed) by any collection of vectors, e.g. Jets, tracks, or topoclusters

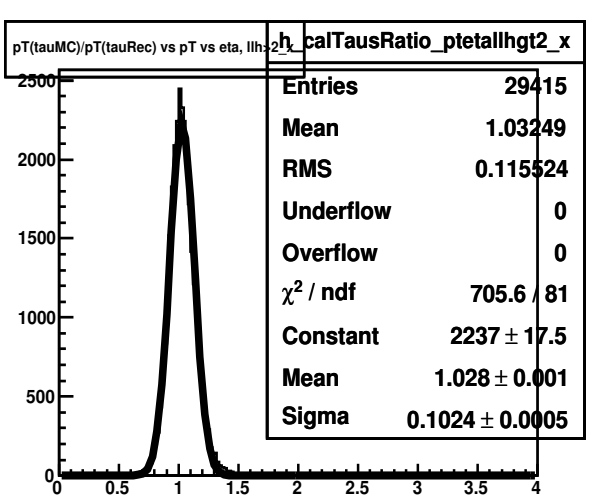


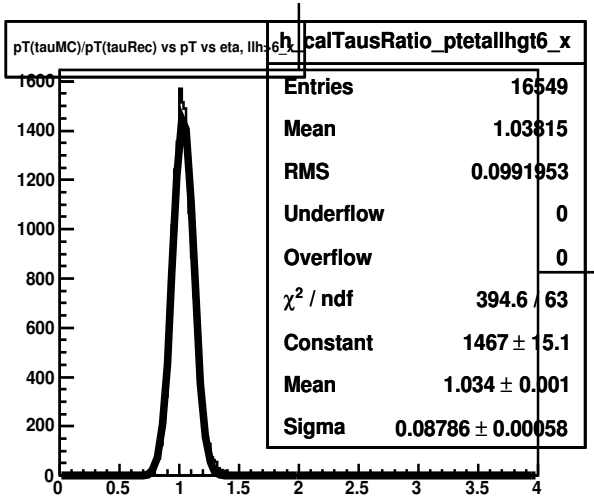
TauRec (reco open questions)

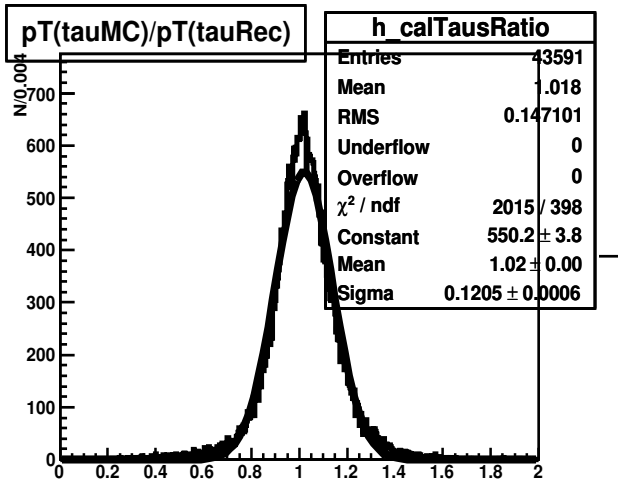
- Other seeds have never been thoroughly investigated for tauRec
- Taus seeded by jets (probably cone?) from TopoClusters could be interesting
- The TopoCluster algorithm could act as a good noise suppression, both for the calibration and for identification
- One may also look for sub-TopoClusters in such a jet (mainly for very low p_T taus), using the categorization into EM and HAD Clusters to separate the Pi^0 from Pi^+ clusters
- This may also allow new ID variables as e.g. Inv. Mass between track(s) and EM subcluster(s)
- It may also improve the reconstruction efficiency for low p_T taus in general, allowing for potentially more spread out energy

TauRec (calibration, current)

- Noise cut: take only cells which are above $2 \cdot \sigma$ (electronic noise for the cell)
- This cut is applied to cells used for identification and for cells used in calibration
- Calibration is done using an H1 style algorithm, weights are the ones obtained for jets
- A p_T and η dependant correction factor, specific for taus, is applied after the cell level H1 calibration
- For CSC with old weights the calibration is still OK (after fixing the sampling fraction difference in the tile), but not perfect







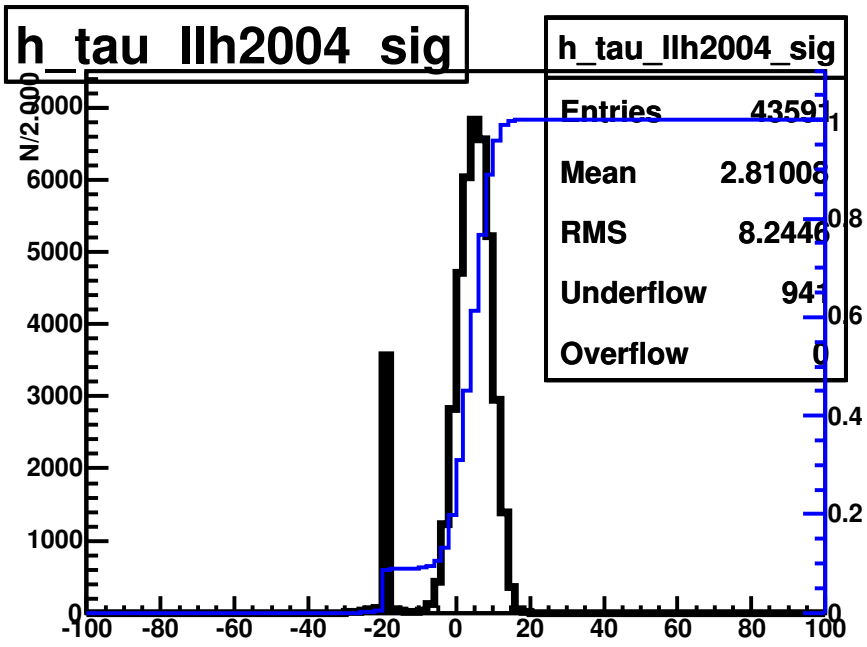
TauRec (calibration, open questions)

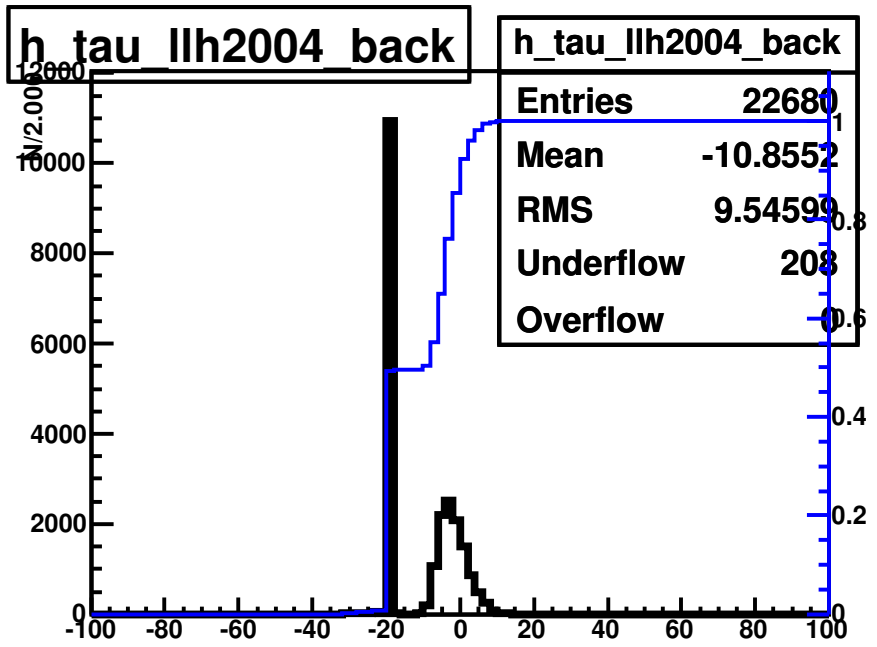
- H1 style Calibration should have weights trained for taus
- They should be optimized separately for different tau types (e.g. Pi^{+}, Pi^{+} + Pi^{0}, $3Pi^{c}$, ...) as far as these can be separated
- It depends on the ID cuts chosen on the level of a few percent change in the mean, because one selects e.g. cleaner tau candidates
- It obviously depends on p_T and η, and the weights might require a smaller η binning than for jets, because one could potentially resolve smaller details given the narrower energy deposit
- At the moment it generalises between samples only at the few % level (e.g. If you go from one sample, perfectly calibrated with a fudge factor in p_T and η to another sample you still might see ~ 2% percent offset, before ID cuts, less after ID)
- The local hadron calibration provided with the TopoClusters should certainly be tried out, expecting to make a physics object calibration on top of that
- The last point might improve automatically with the points before

- **Energy flow type of treatment is covered by Elzbieta + tau1P3P**

TauRec (identification)

- ID variables show some funny features if plotted directly, one has to remove the bad cluster seeds by hand, e.g. tau_et != nan, tau_et > 15 GeV
- ID dropped somewhat on CSC because of problems with the seeding clusters. We suspect that a splitting and merging was introduced which is not correctly treated in tauRec. This has to be understood and fixed.
- Conclusion: taus are very sensitive to seeding, so using TopoClusters could yield a more stable seed for the taus





TauRec (identification, open questions)

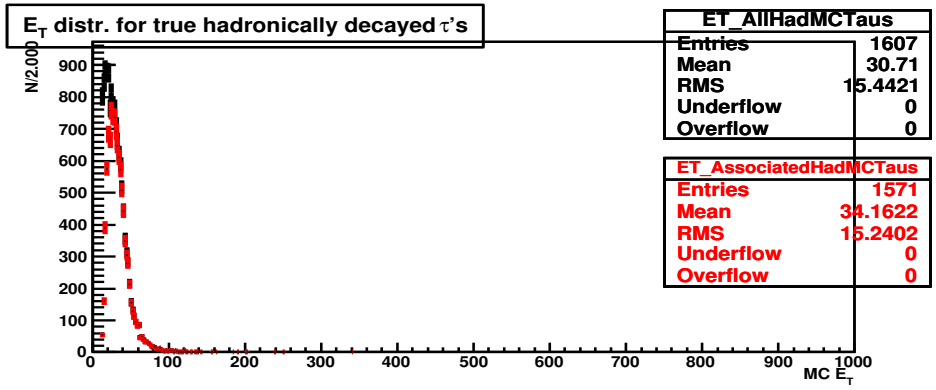
- Should the identification variables also be calculated using only cells from TopoClusters ?
- Some noise suppression is certainly needed (at the moment simple cut $E_{cell}>2*\sigma$)

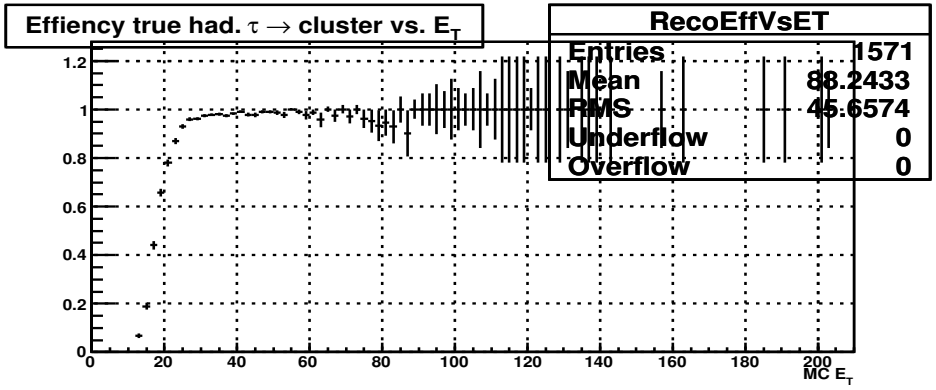
- More sophisticated treatment of taus in the calorimeter could be promising like
- Searching for EM subclusters could be done separating HAD and EM subcluster
- One could use TopoCluster in EM only to find pi^0
- and in HAD+EM (for the hadronic subshower)
- or use even more specialized information e.g. in the eta-strip layer only
- also for very low p_T taus the strategy may be based more on finding individual components of the tau, for these also in combination with tracks

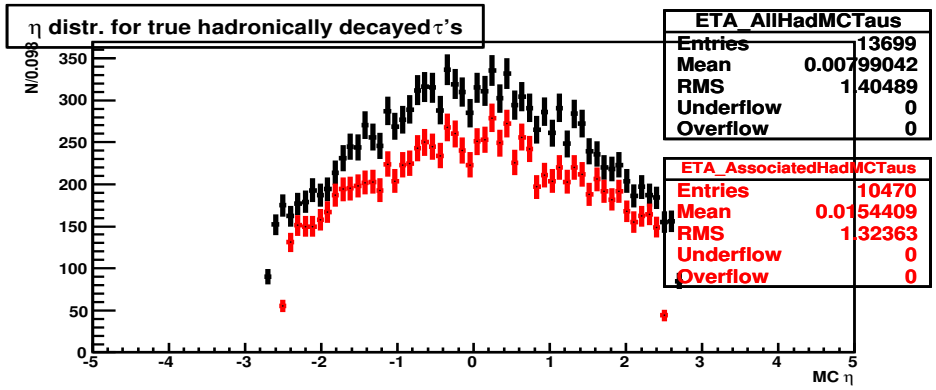
TauRec with TopoCluster

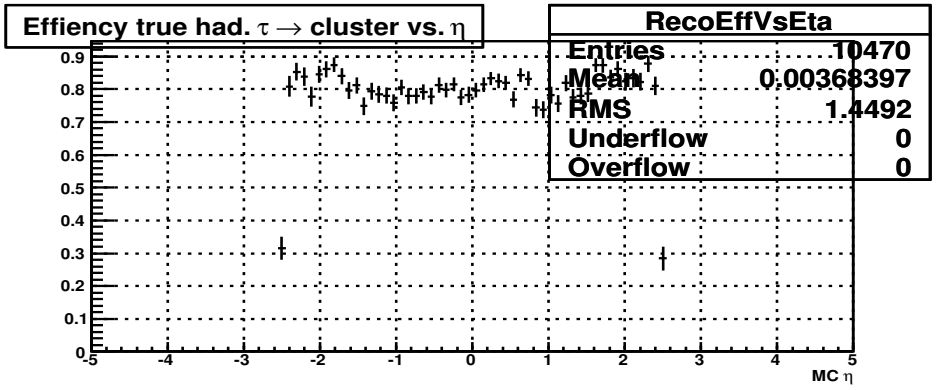
- In tauRec-02-03-05-08 and on the tauRec Wiki page I provided some example jobO for running the existing ConeClusterJet and using these jets as seeds for taus. The resulting CBNT has both normal tauRec and tauRecTopoCluster variables.
- Cells are already taken from TopoClusters, for Calibration and for ID variables
- missing:
  - The correct treatment of cell overlaps between TopoClusters (in tauRec)
  - Calibration tool to make use of the local hadron calibration missing
  - TopoCluster specific tools are missing completly (monitoring, additional variables, ...)
- This is just a first try to get things going ! Hubert's group (Susanne et al) showed interest to contribute to this topic. Maybe already results for Barcelona ?

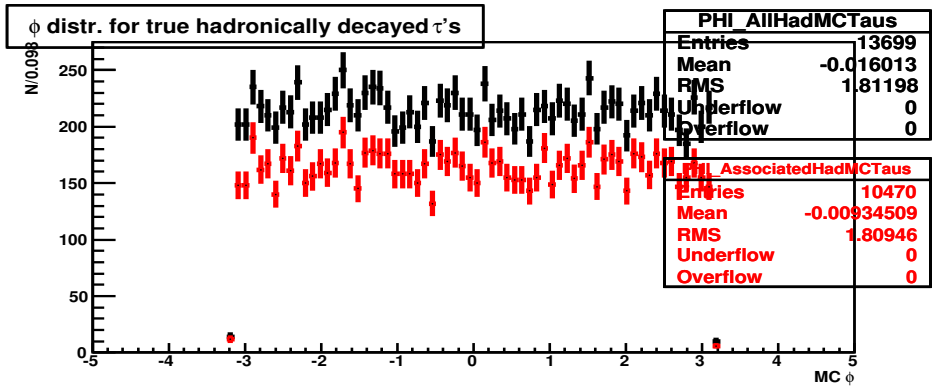
TauRec with CombinedCluster

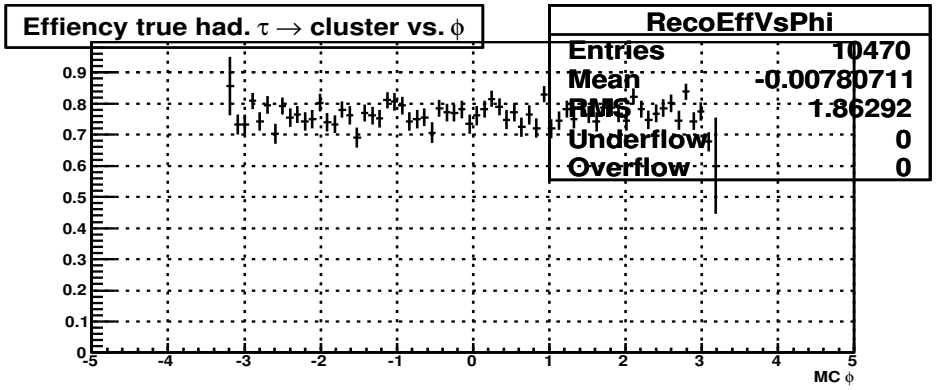




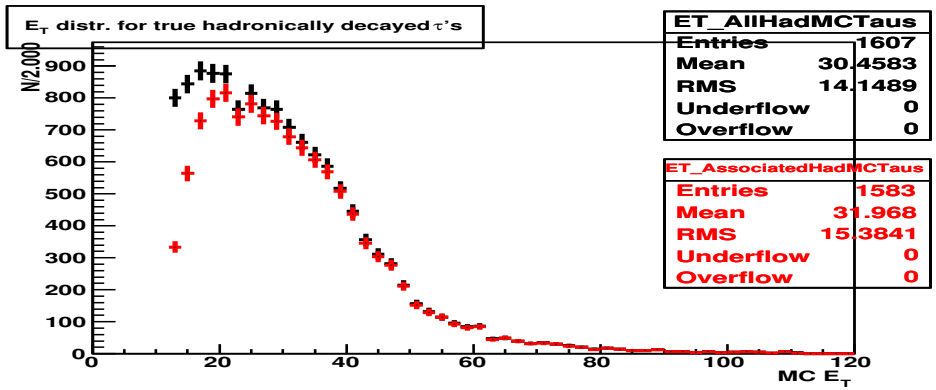


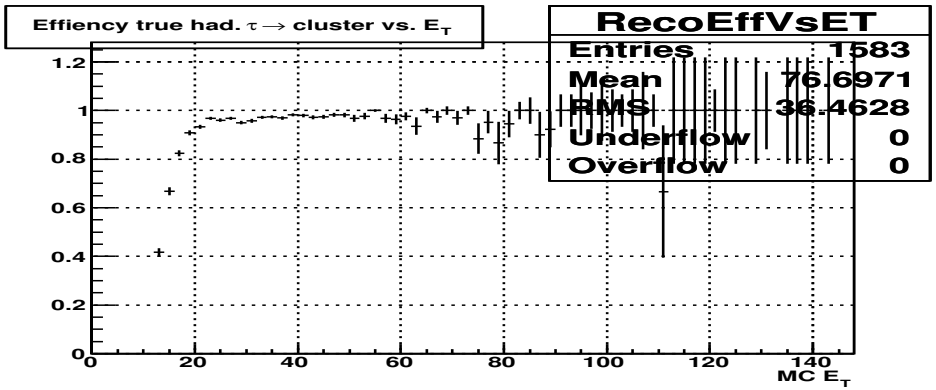


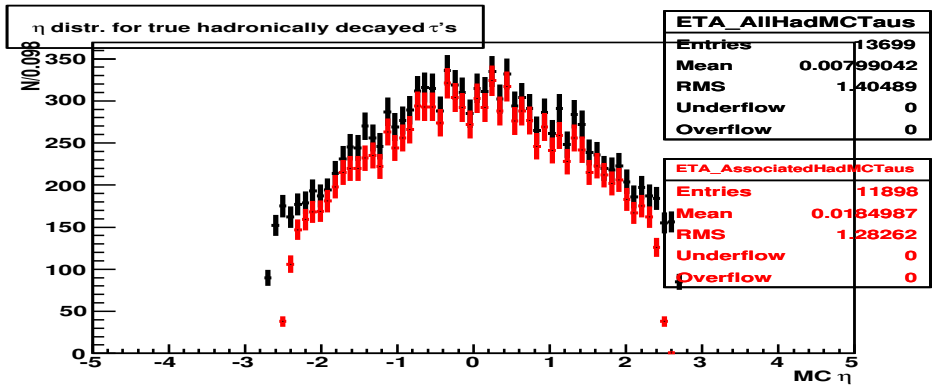


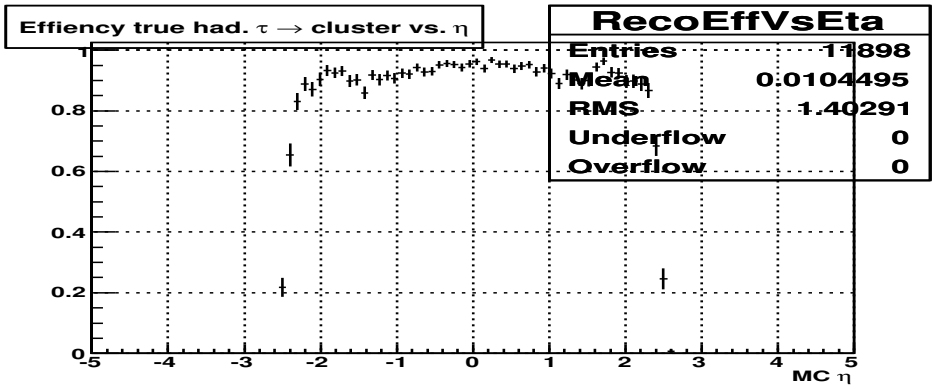


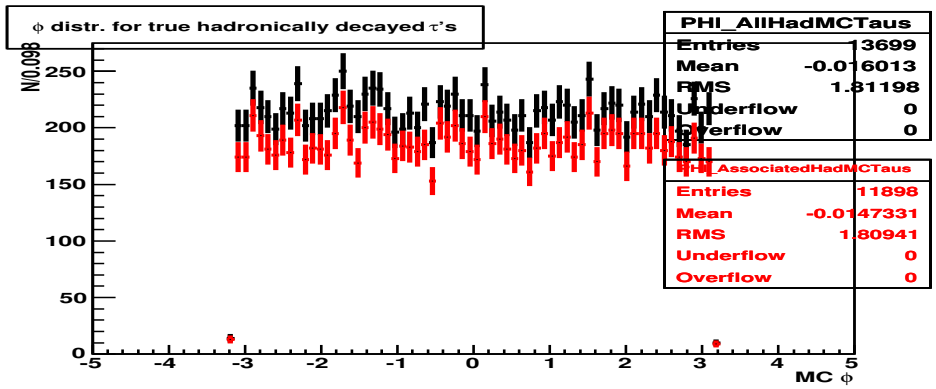
TauRec with TopoCluster

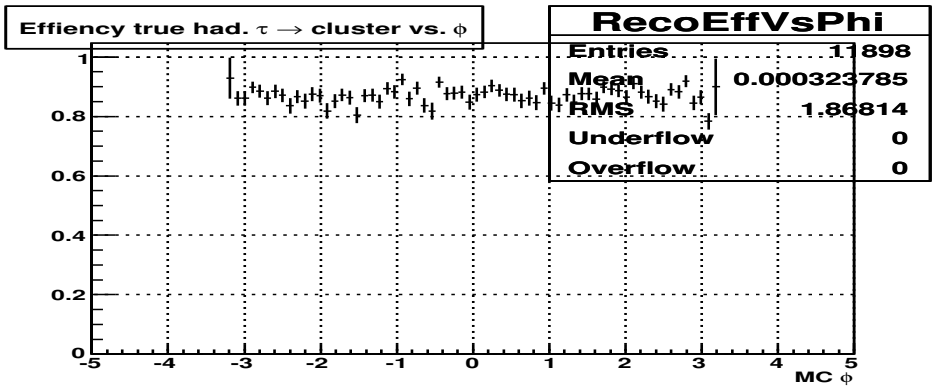




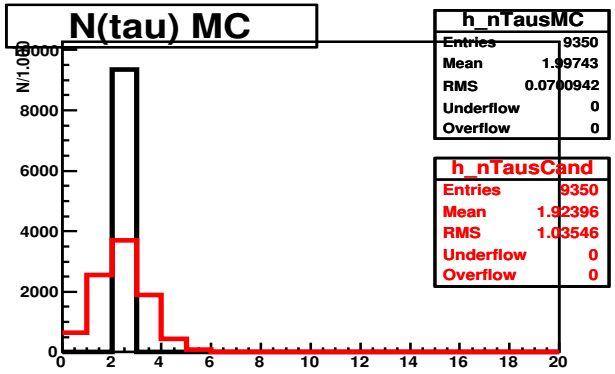


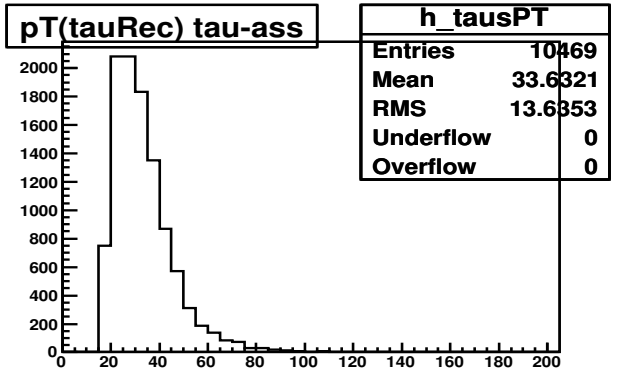


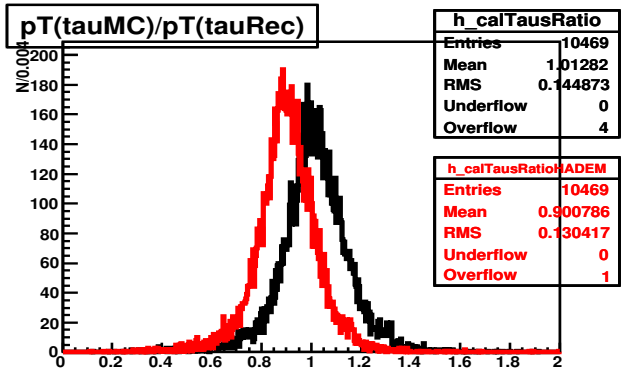


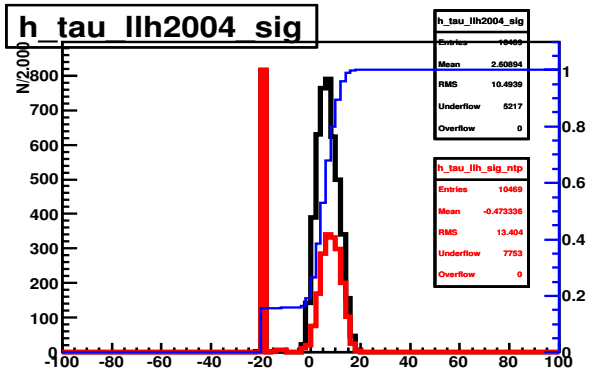


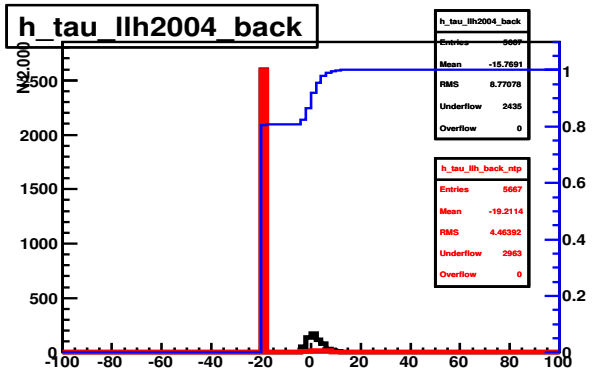
TauRec with CombinedCluster

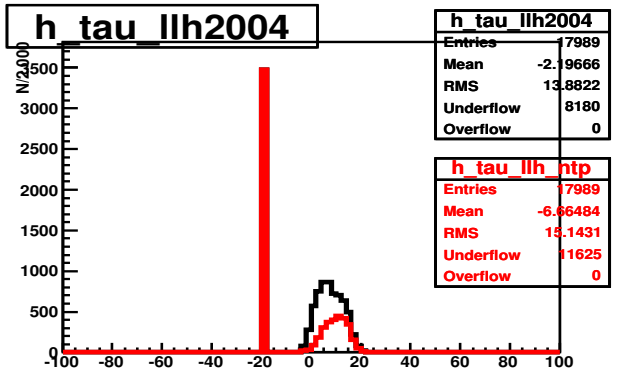


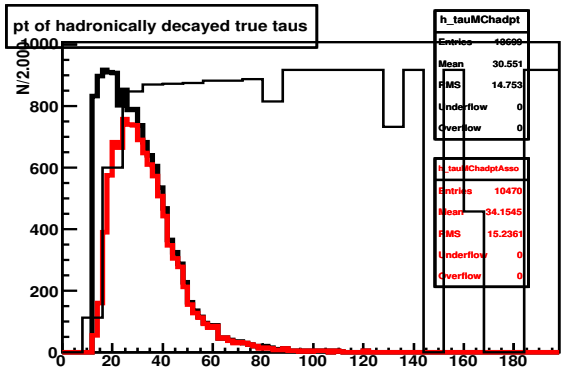


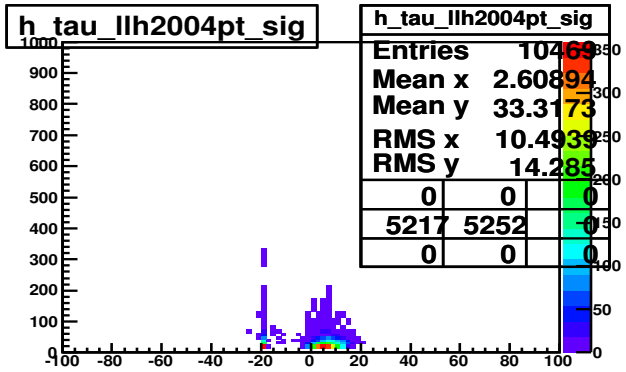


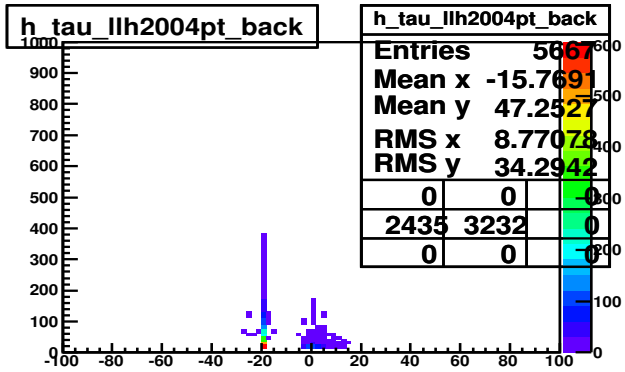




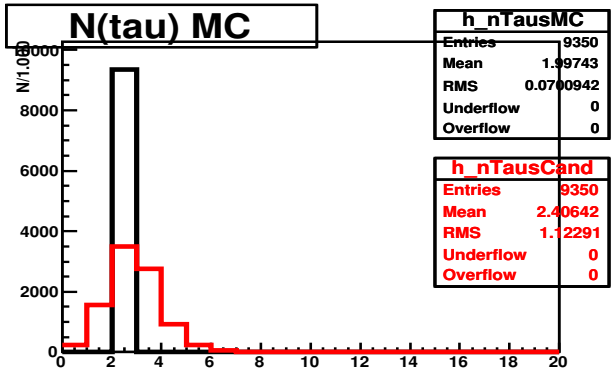


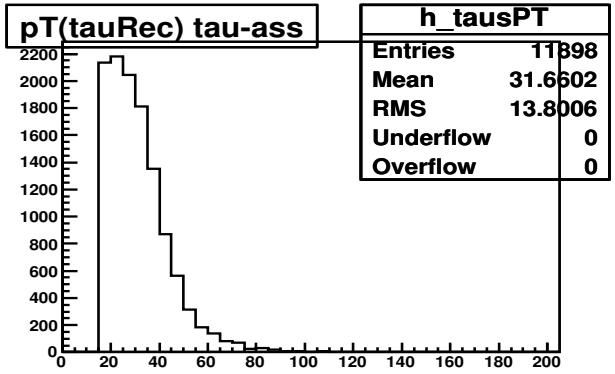


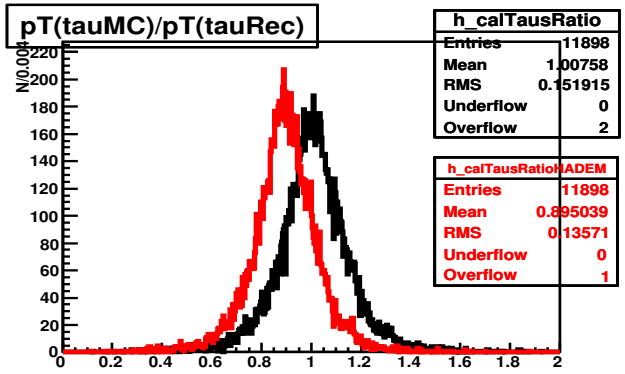


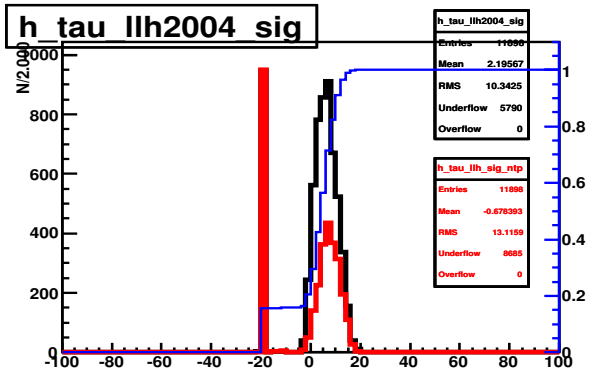


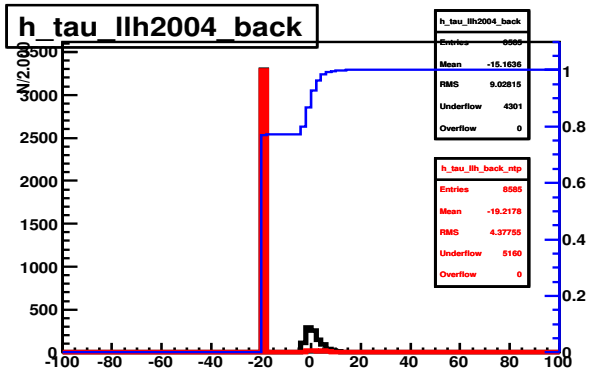
TauRec with TopoCluster

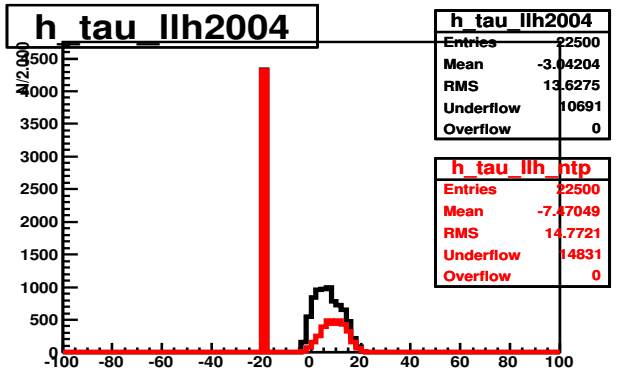


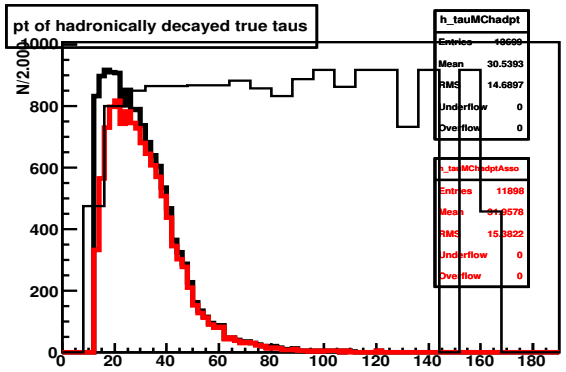


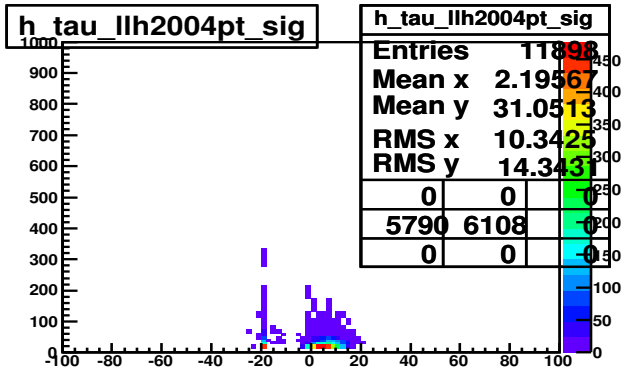


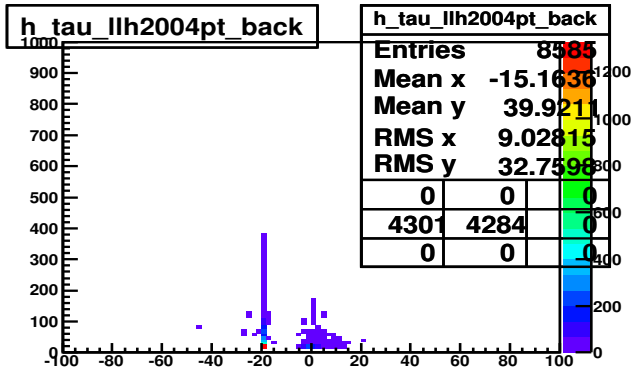




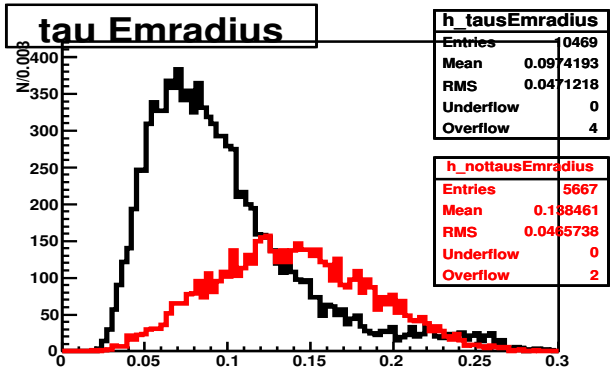


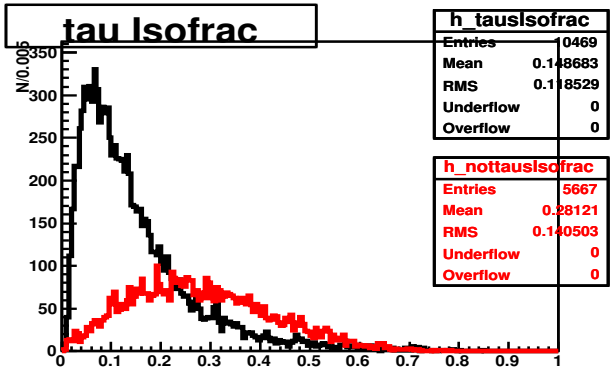


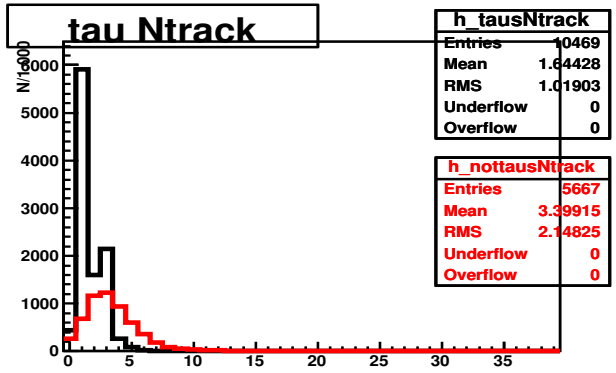


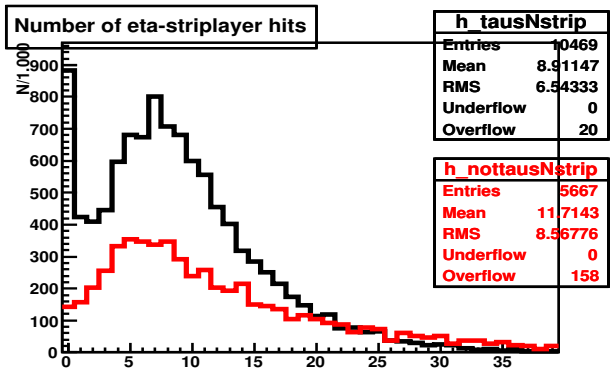


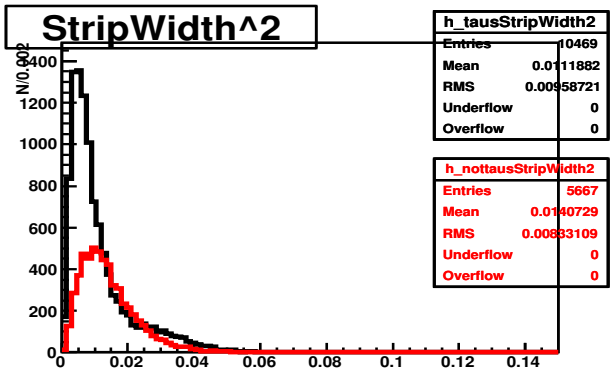
TauRec with CombinedCluster

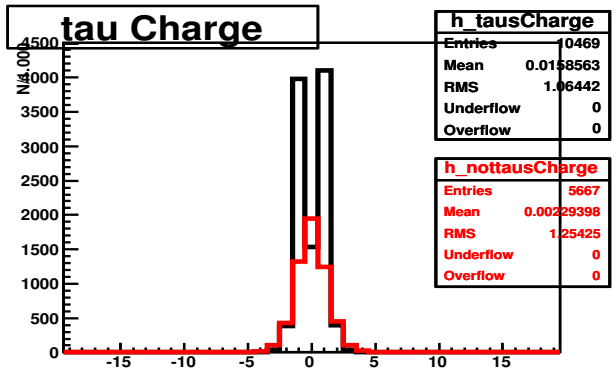


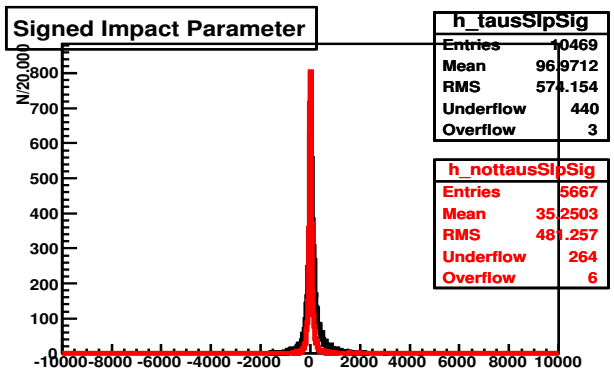


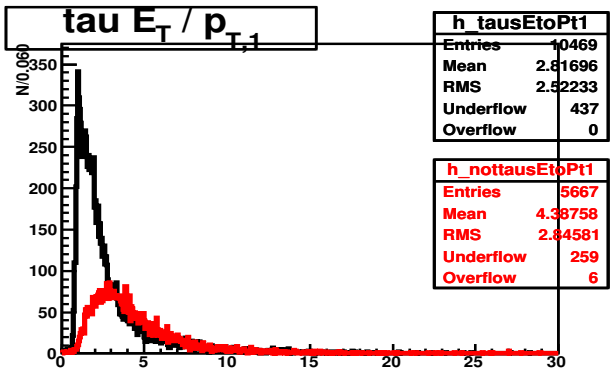


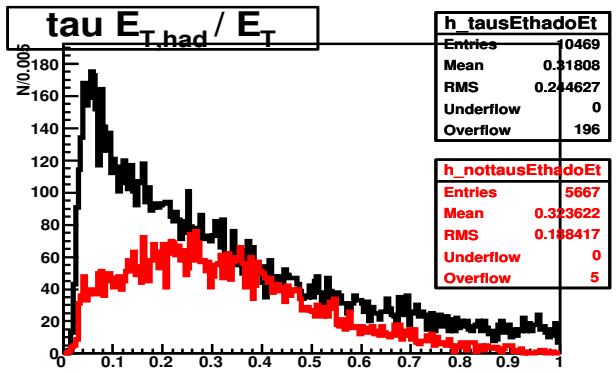




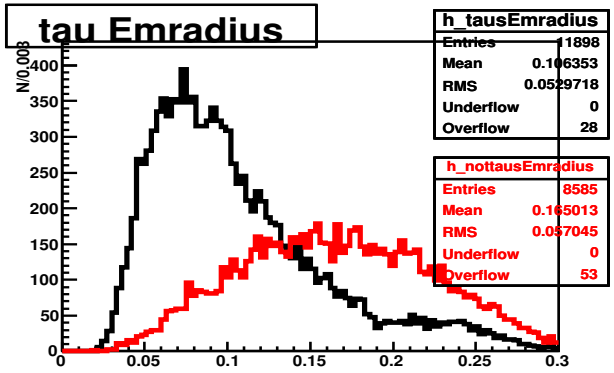


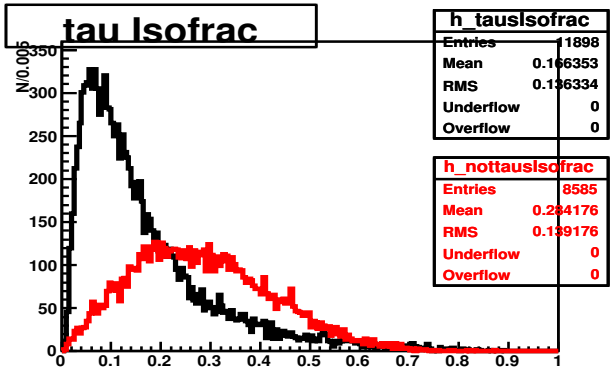


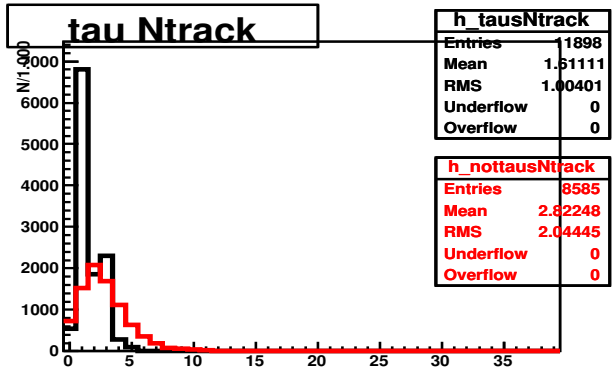


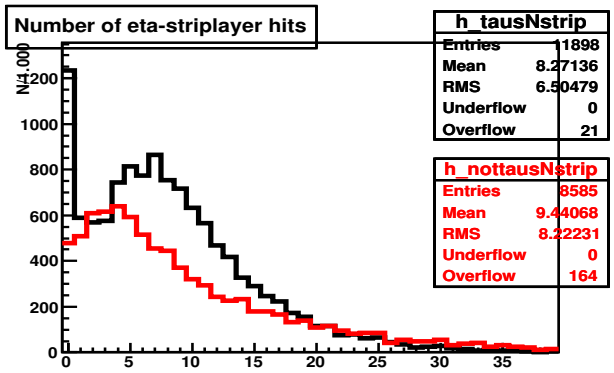


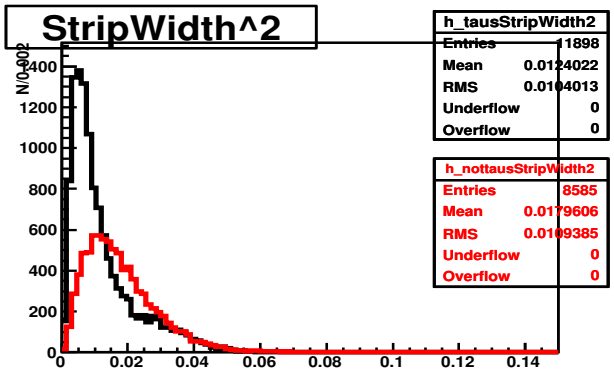
TauRec with TopoCluster

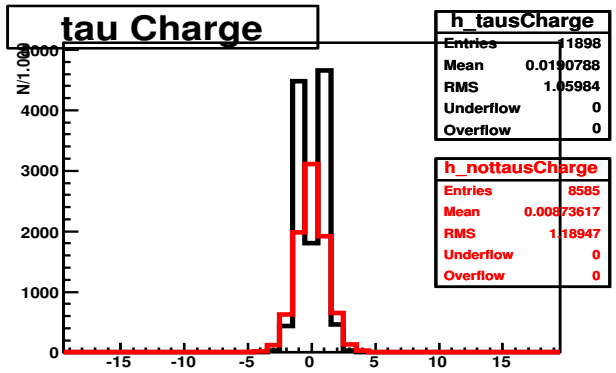


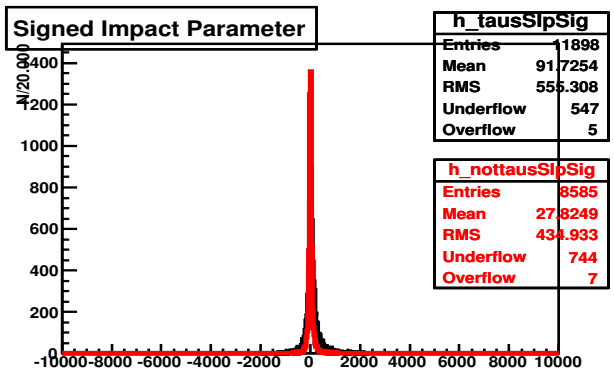


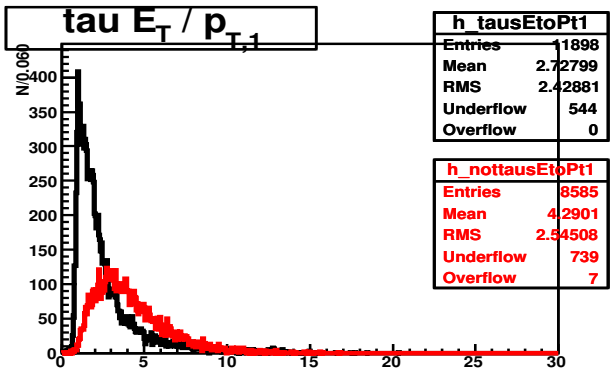


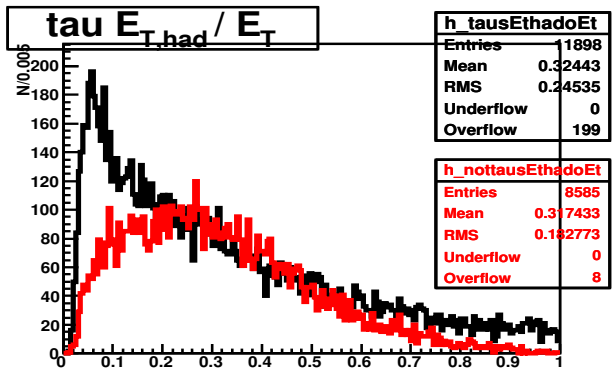










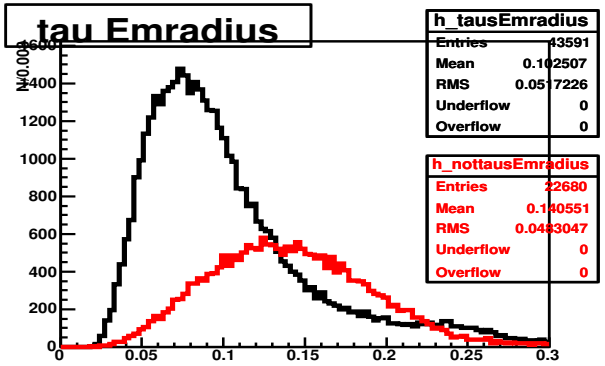
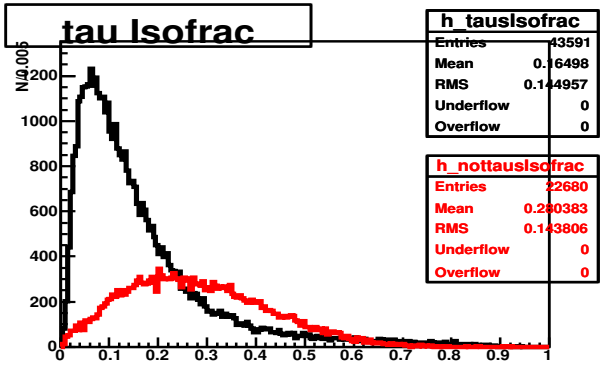
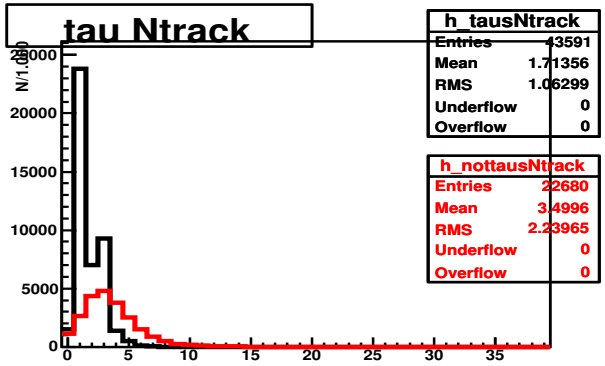
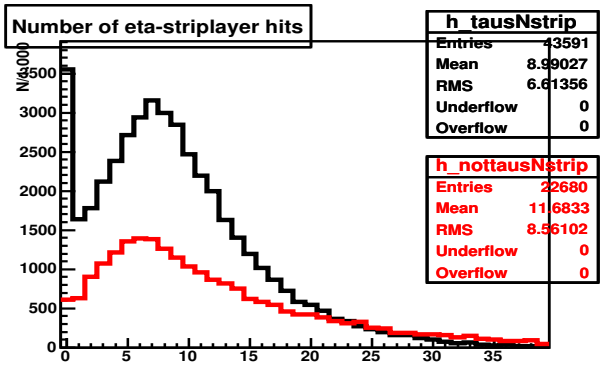
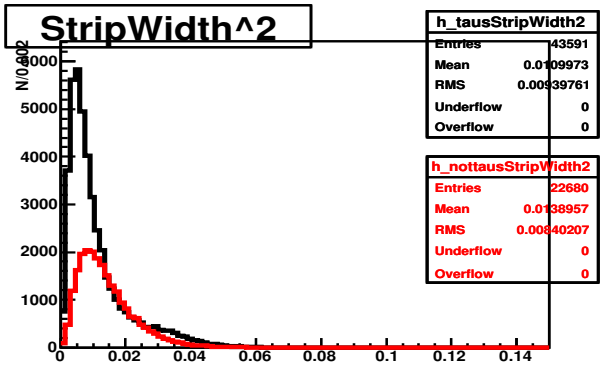
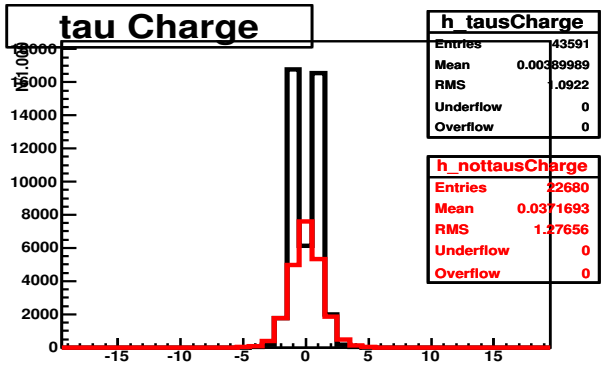
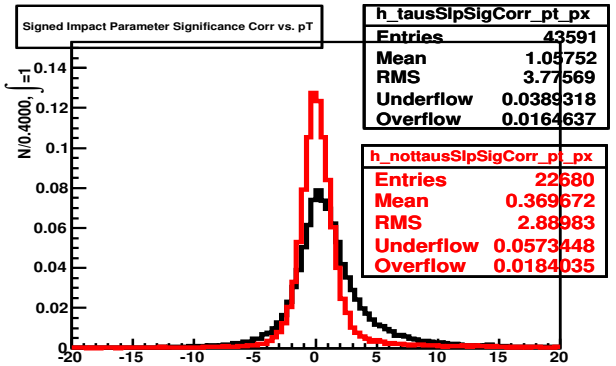
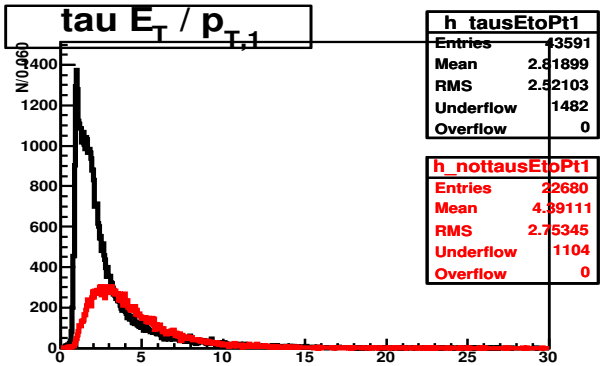
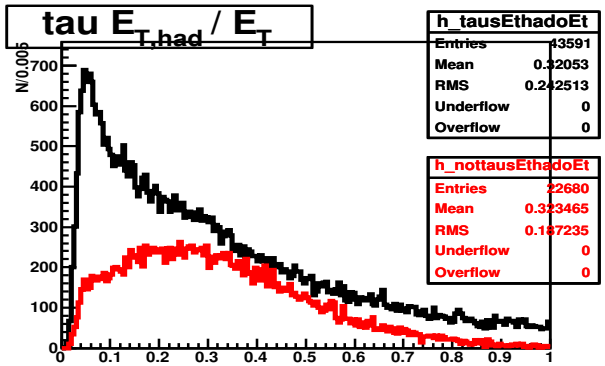


BACKUP SLIDES

General open questions

- Events with pileup (not even low lumi) have never been studied
- Z→ tau tau not only to get PDFs is still not thoroughly studied in fullsim (Asen Christov is going to do this for his Ph.D.)
- A clear strategy for how to validate ("certify" as in D0) a certain algorithm for tau ID is missing
- A systematic monitoring of the input quantities with e.g. RTT also with respect to data taking should be established
- Current strategies for selecting "calibration" events would require a rather good understanding of MET, maybe we should try to decouple that (e.g. no collinear approximation for Z → tau tau if possible)
- Maybe later we want to improve our understanding of 3 prong decays. Questions of tracking (e.g. For high pT where it's difficult to separate tracks) or secondary vertexing for taus are still not thoroughly understood (a simple sec. Vertexing is done in tauRec but not fully studied)
- Very low pT taus are quite difficult (< 10 GeV), what use cases are there ? What would be the best strategy to reconstruct and identify them.

TauRec (identification, open questions)

- Retraining the old likelihood with ROME data improved performance on ROME data by 5-20% in rejection (depending on pT)
- A retraining on CSC might be feasible (after figuring out the problem with the seeds)
- A 2 dimensional likelihood might be interesting, given a good smoothing is used
- Retraining likelihood for different tau type separately might be interesting, but first attempt did not yield a significant improvement (also the available statistics might have been the problem (less pT bins possible))
- Asen is looking into electron ↔ tau separation
- A neural network is also available in tauRec from Kamal, but only for a small pT range (17-35GeV)
- Many more ID variables than the usual ones are implemented in tauRec, mostly copied from D0
- Some are promising but of course highly correlated with the normal variables

TauRec (identification, CSC)