

Directed tree clustering algorithm and PFA status

Dhiman Chakraborty, Guilherme Lima,
Rob McIntosh, Vishnu Zutshi



NORTHERN ILLINOIS
UNIVERSITY

ILC Simulations Workshop
Univ. Colorado, January 9-11, 2006

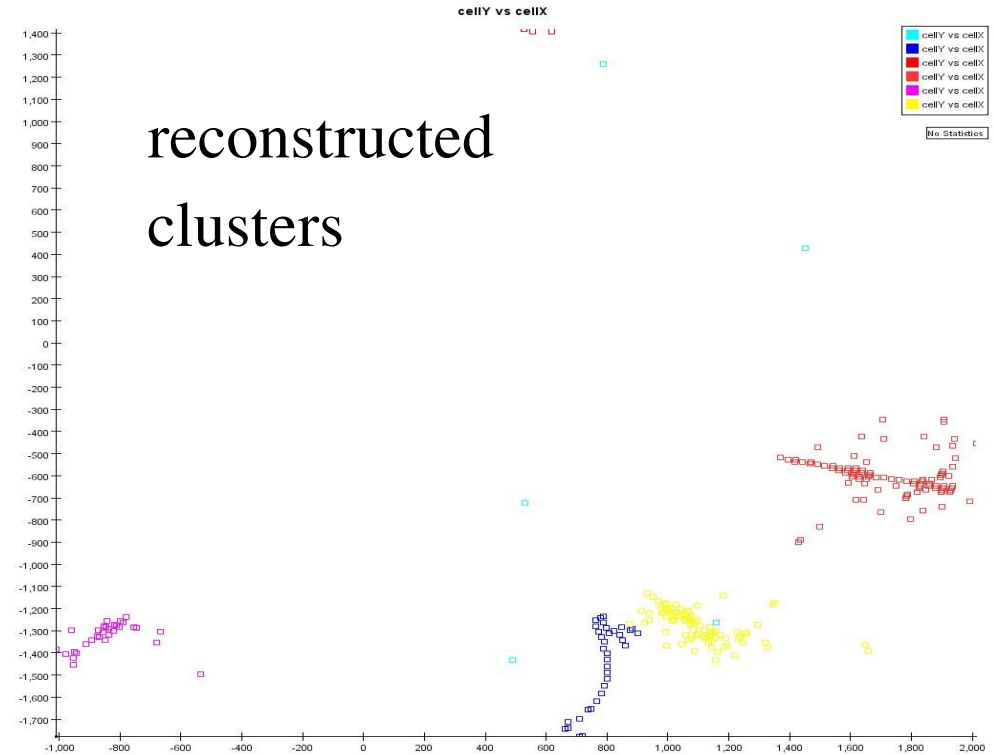
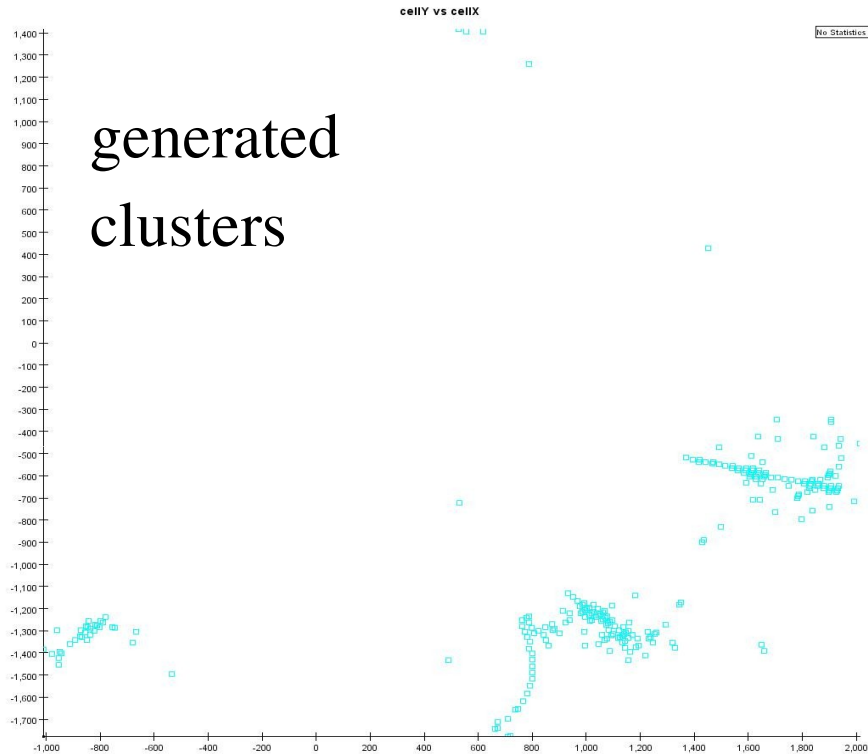
Clustering and PFA

- Goal: develop a Particle Flow Algorithm, based on the flexible Directed Tree clustering algorithm, implemented by Vishnu Zutshi.
- Participate in the detector optimization effort
- Development has been based on SiD and non-projective geometries, but algorithm is in no way restricted to these characteristics
[SDNPHOct04](#) is based on [SDJan03](#), [steel/scintillator](#), with non-projective HCal barrel and replacing 34 1cm-thick with [41 0.5cm-thick](#) scintillator layers, simulated by LCDG4.
A SLIC model has been prepared for SLIC, to be used for certification
- Digitization effects to be studied for detector optimization (DigiSim)

Directed Tree Algorithm

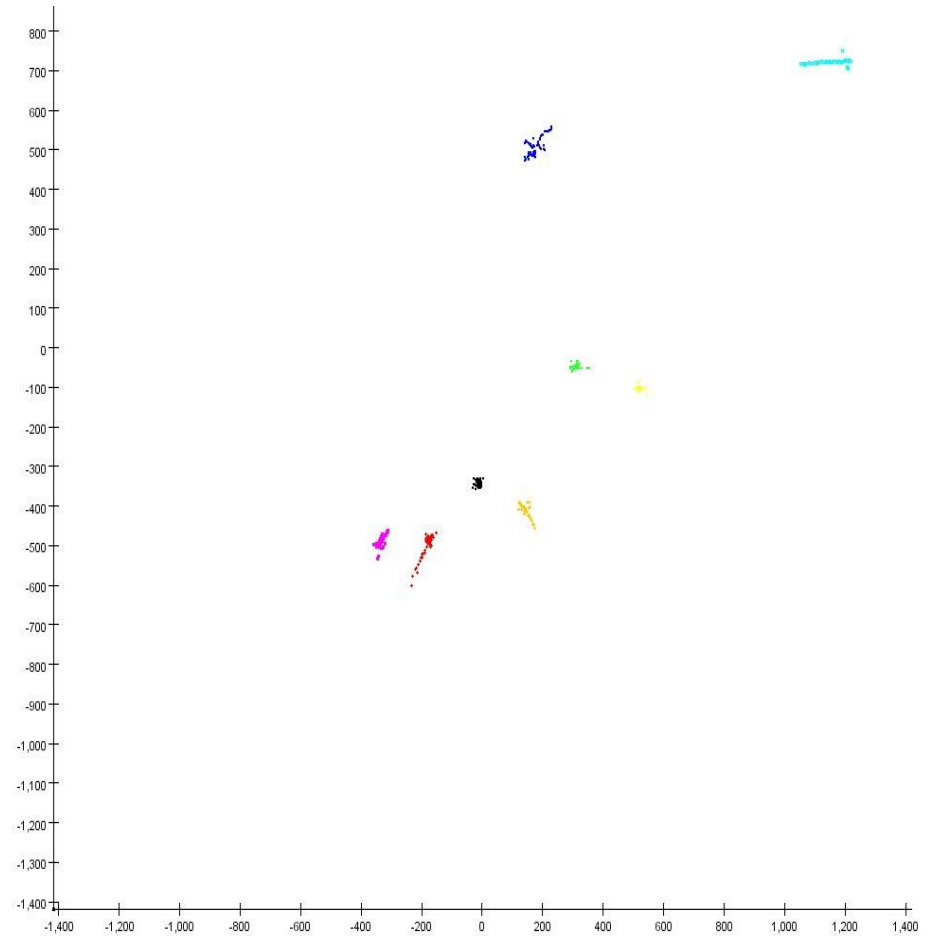
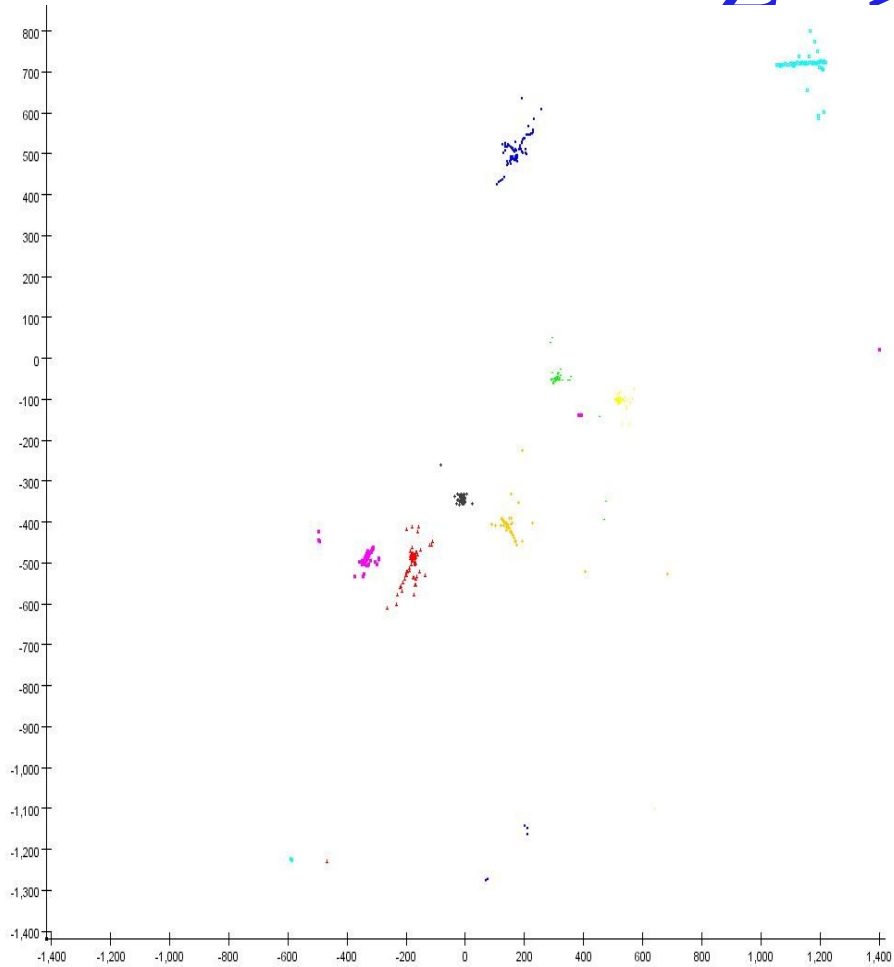
- Define a cell neighborhood (e.g. “5,5,9” for +/-5 in phi, +/-5 in z, +/- 9 layers around the reference cell)
- Discard low-energy hits ($\frac{1}{4}$ MIP cut)
- Calculate density for each hit, based on number of hits on neighborhood
- calculate Distance-Weighted Density Differences, $(D_j - D_i) / d_{ij}$, for all hits j in the i 's neighborhood
- Find i,j pair of hits (i,j) with maximum DDWD
 - negative max: i is the seed of a new cluster
 - positive max: i is attached to j , which may be a seed or become attached to some other seed

Single particle events



No problem to reconstruct the clusters from single hadrons.
Some fragments are clear though.

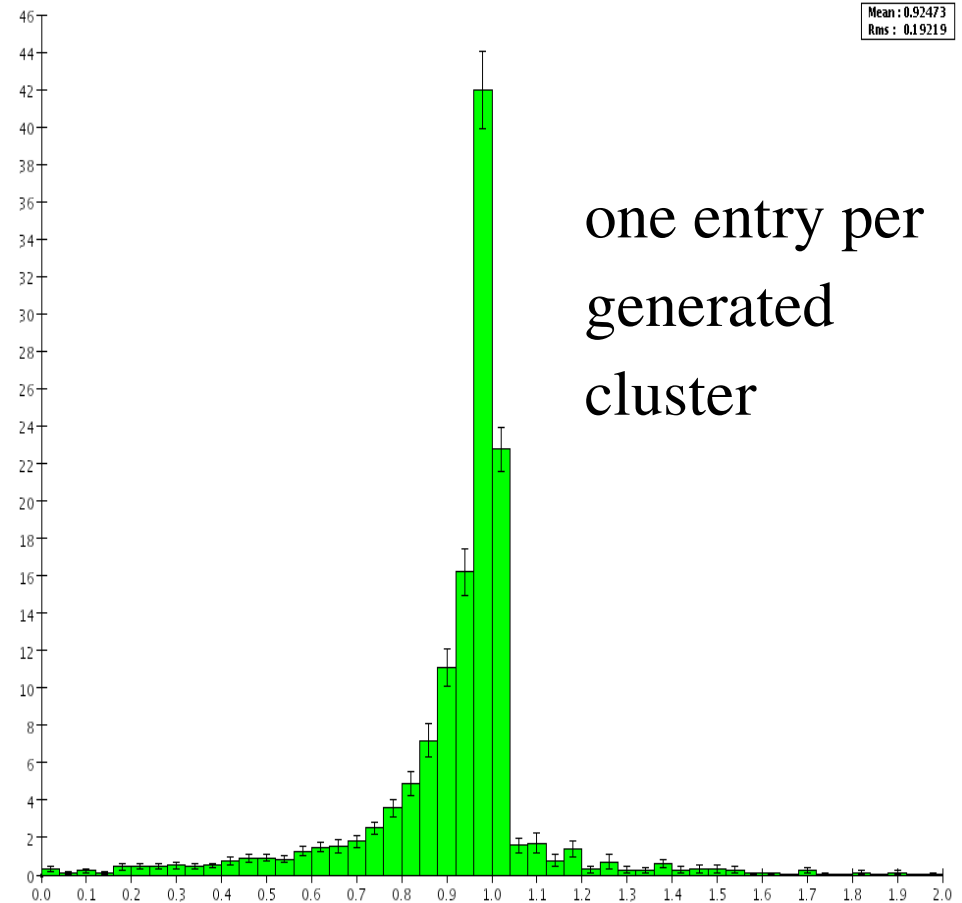
Z --> hadrons



Seems to be doing a good job, but one needs to quantify the performance

Algorithm performance

- Calorimeter only
- Compare reconstructed to generated cluster energies after cluster matching
- Calculate $E_{\text{rec}} / E_{\text{gen}}$ for each generated cluster
- Fill histogram with one entry per cluster, weighted by $E_{\text{gen}} / E_{\text{total}}$
- Ideal: spike at 1

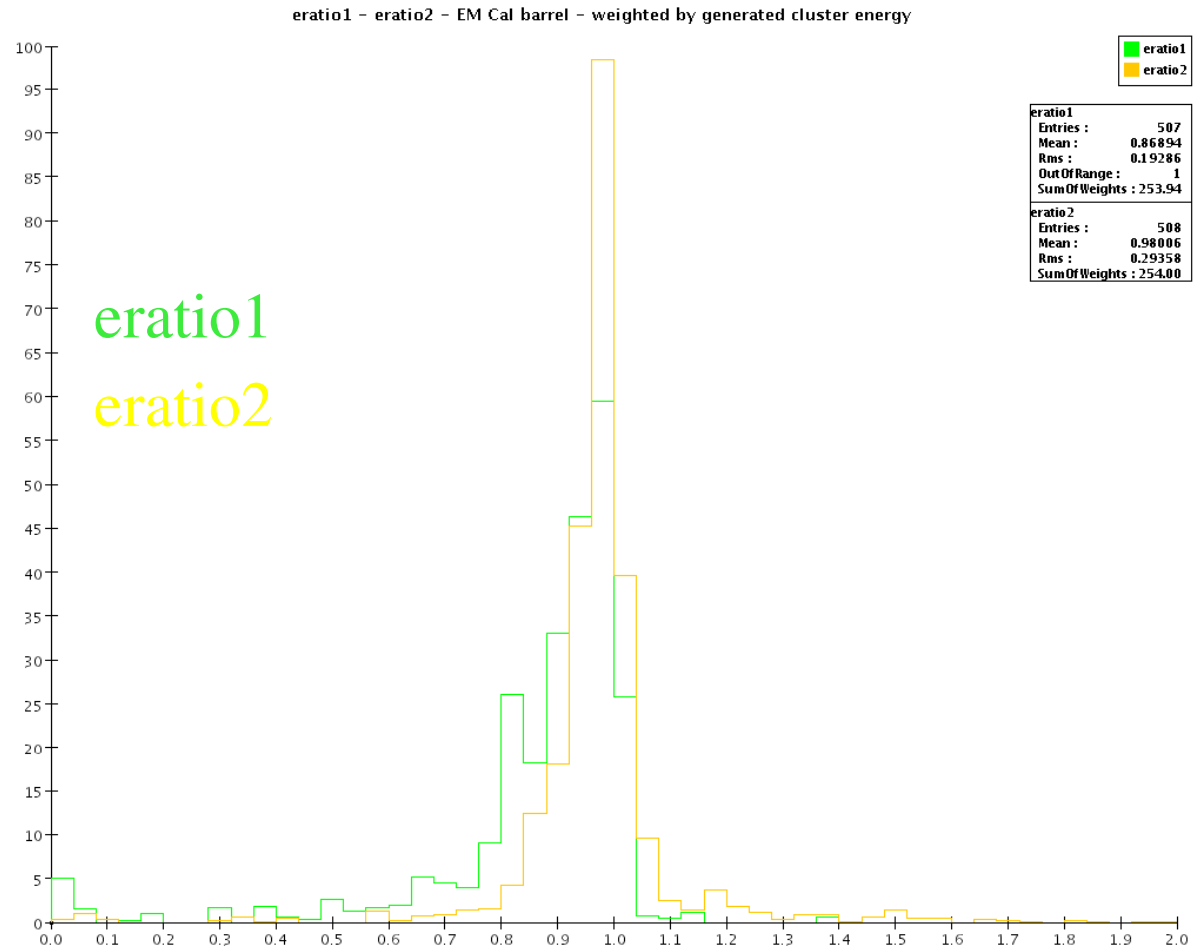


Cluster matching and merging algorithms

- Stage 1: one-to-one gen-reco matching, based on distances (3D or angular)
--> several remaining clusters (“fragments”)
- Stage 2: attach satellites to reco clusters, based on angular distances
possible cuts on angular separation, fragment energies/# hits and distance to primary clusters

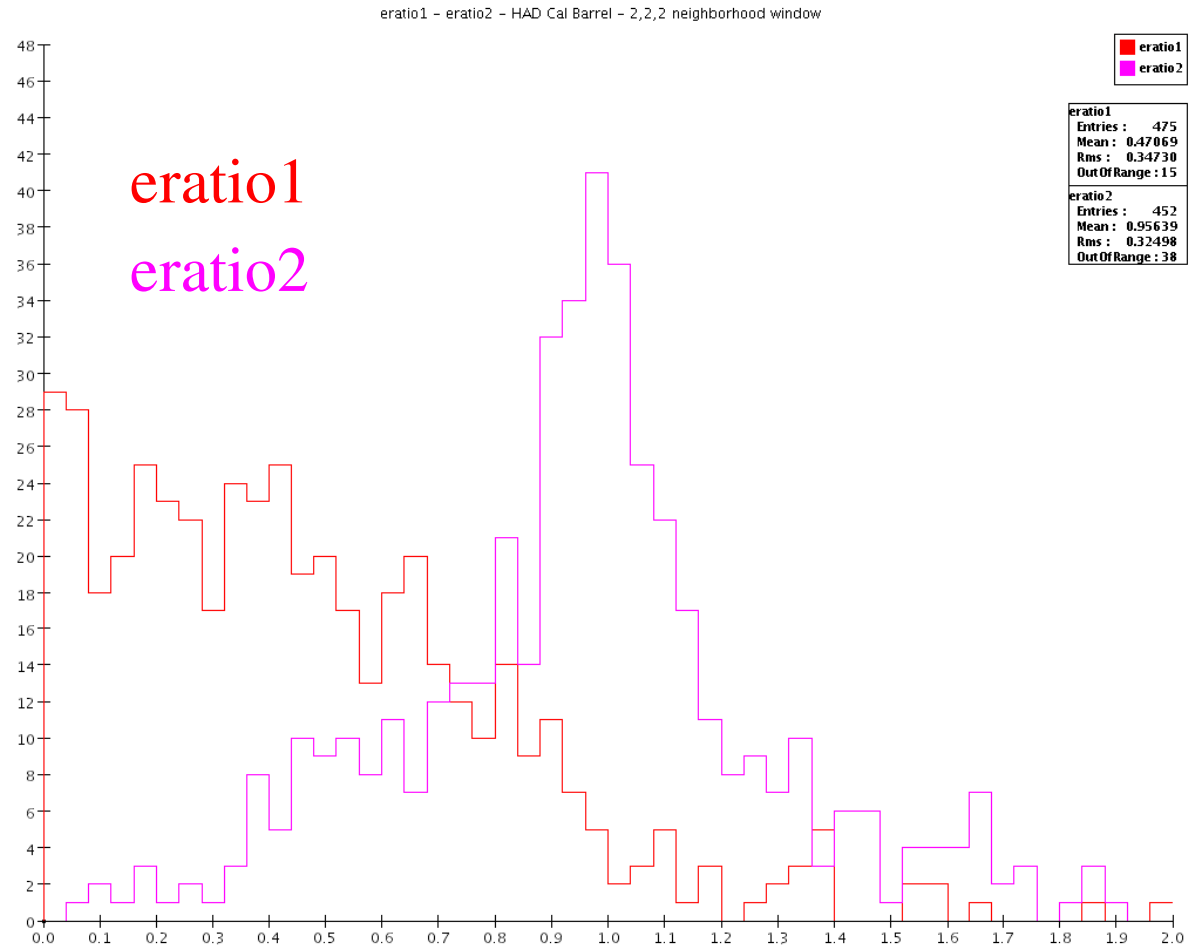
Preliminary ECal analysis

- 500 events, with 2-pions
10cm apart at Ecal face,
using SDNPHOct04 detector
- neighborhood definition:
(dphi=5, dZ=5, dlayer=9)
- discard events with decays or
interactions before Ecal
- Look at:
 - eratio1: Erec/Egen after
stage 1 (matching)
 - eratio2: Erec/Egen after
stage 2 (merge satellites)

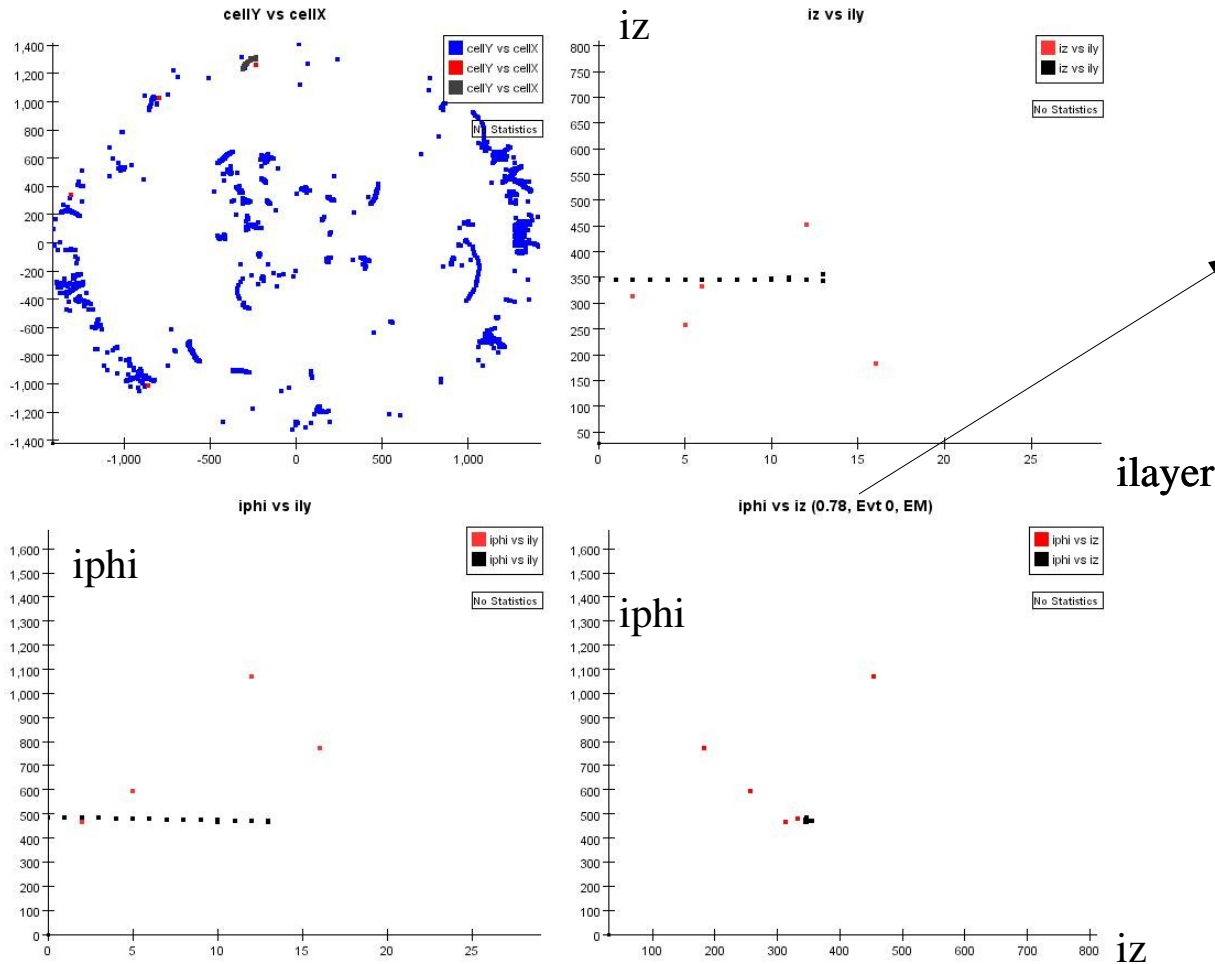


Preliminary HCal analysis

- 500 events, with 2-pions
10cm apart at Ecal face,
using SDNPHOct04 detector
- neighborhood definition:
($d\phi=2$, $dZ=2$, $dlayer=2$)
- discard events with decays or
interactions before Ecal
- Look at:
 - eratio1: Erec/Egen after
stage 1 (matching)
 - eratio2: Erec/Egen after
stage 2 (merge satellites)

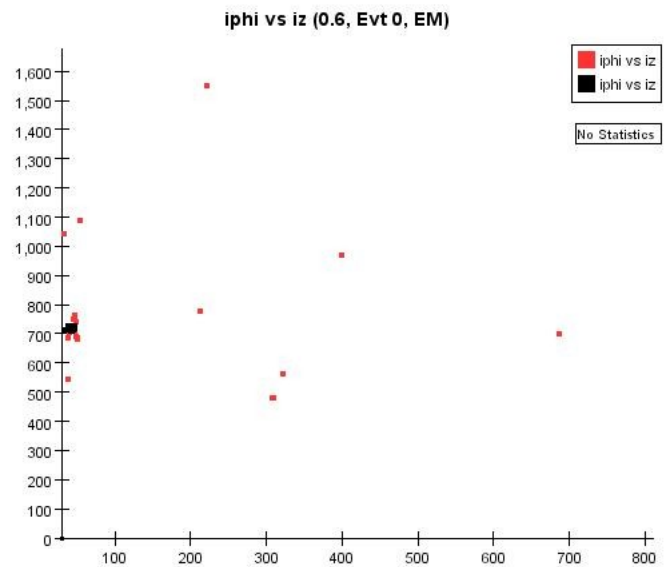
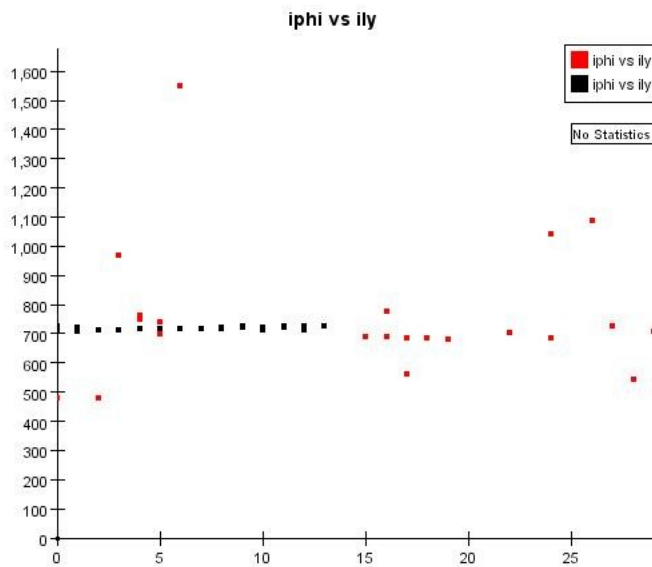
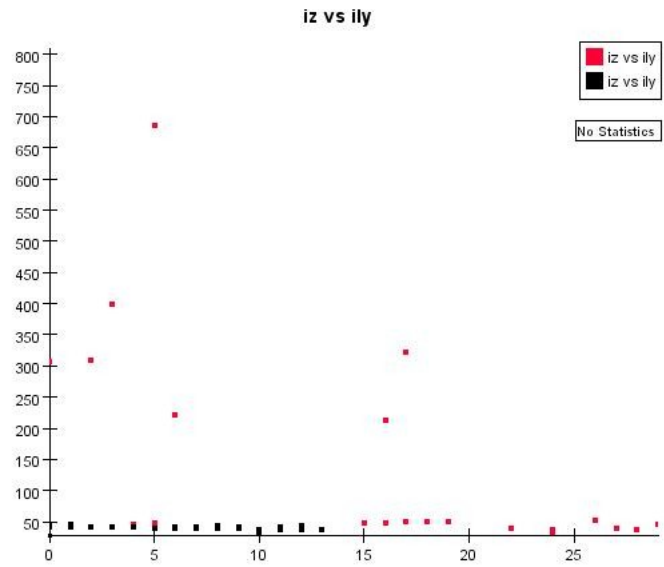
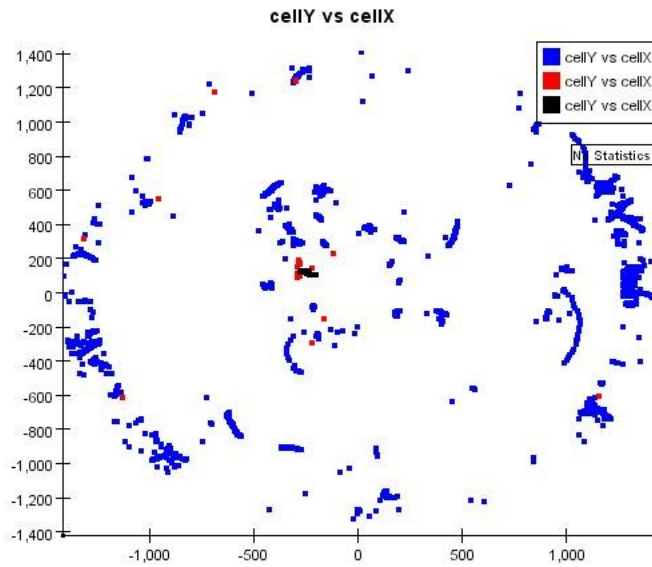


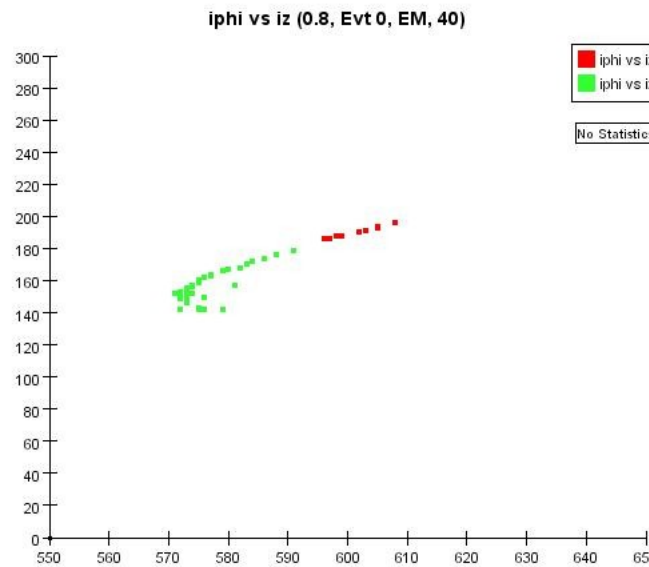
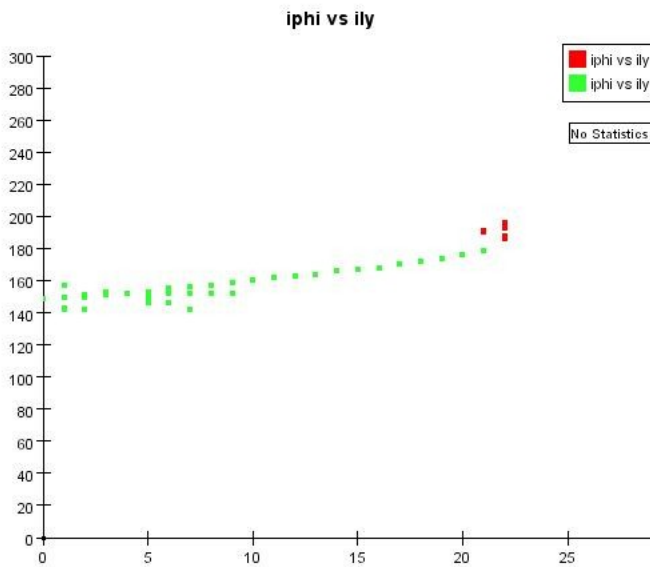
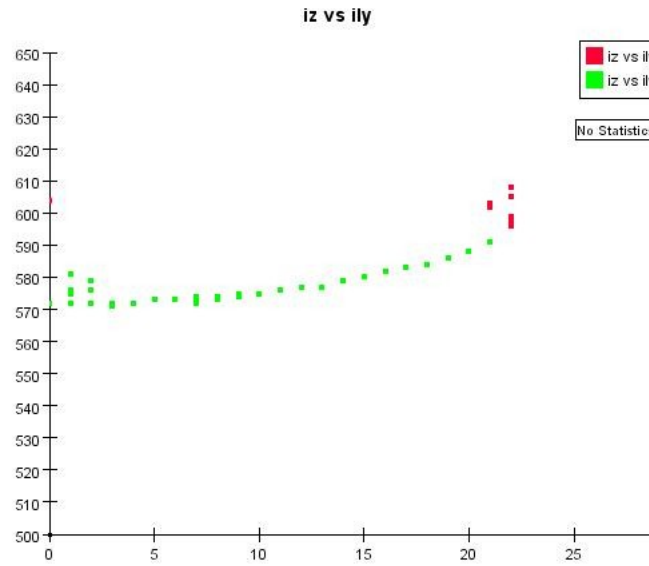
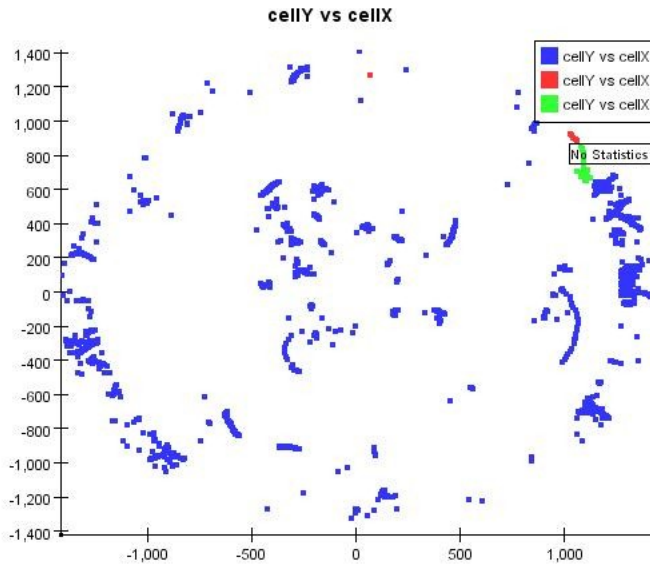
More on fragment association



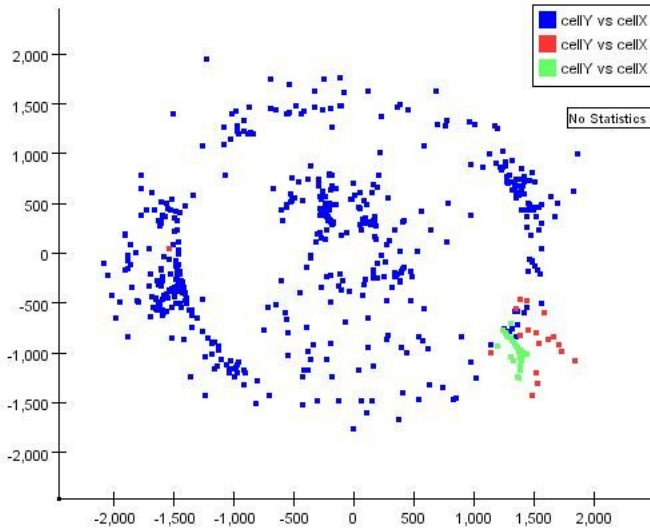
Scan through a few tracks with low efficiency for hit clustering

All hits in event
All hits from a given MCParticle
All hits clustered

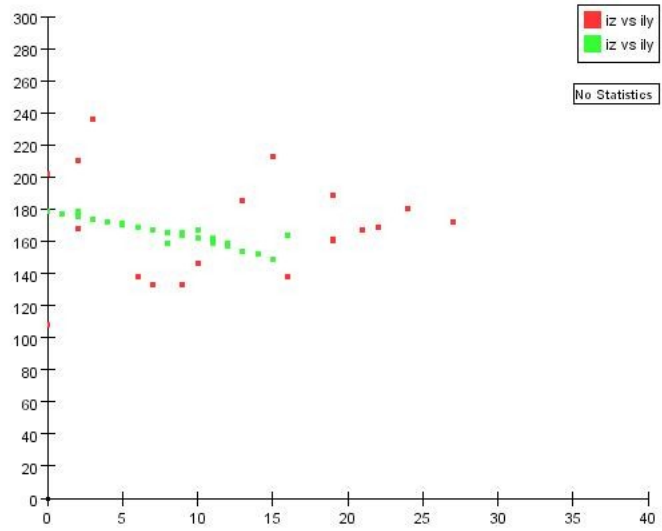




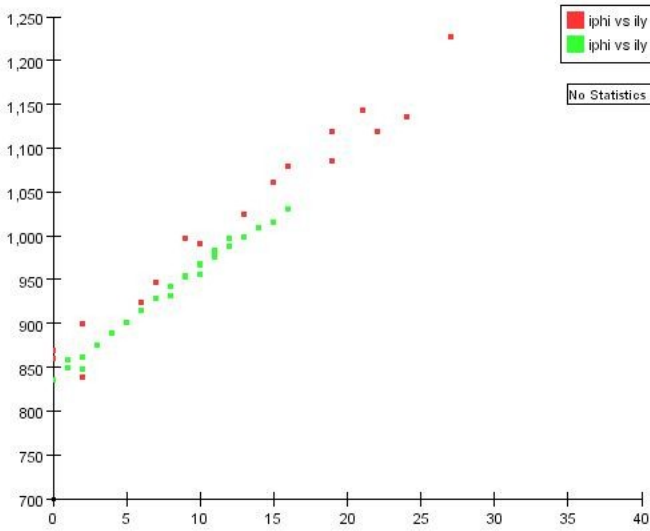
cellY vs cellX - cellY vs cellX - cellY vs cellX - cellY vs cellX - cellY ...



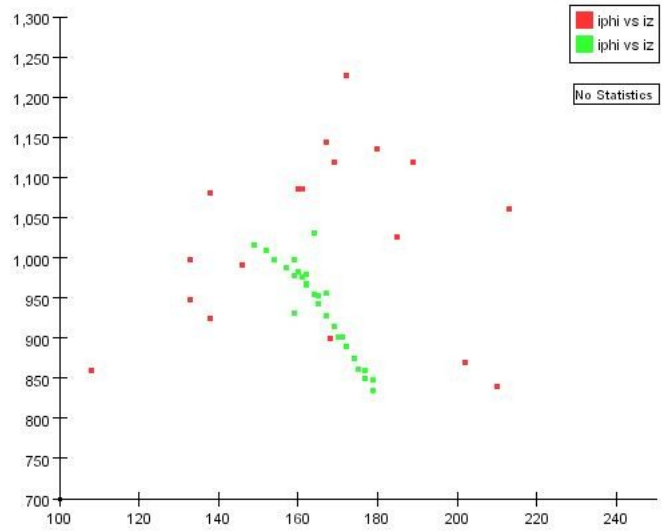
iz vs ily

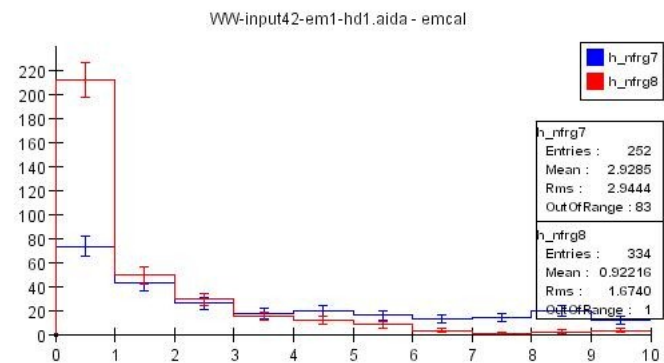
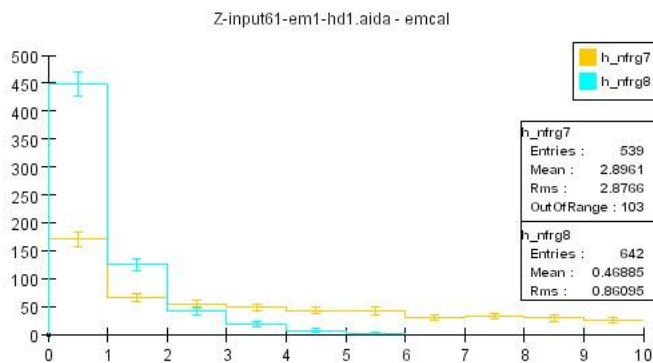
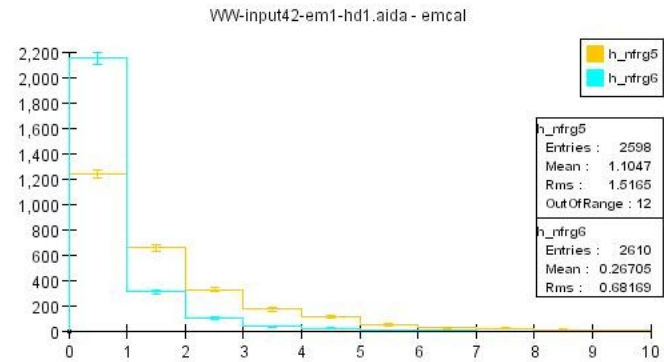
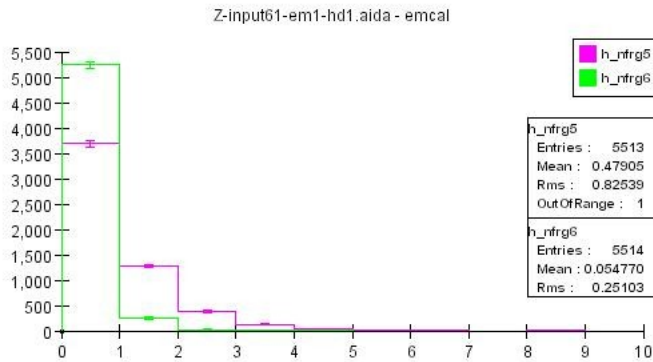
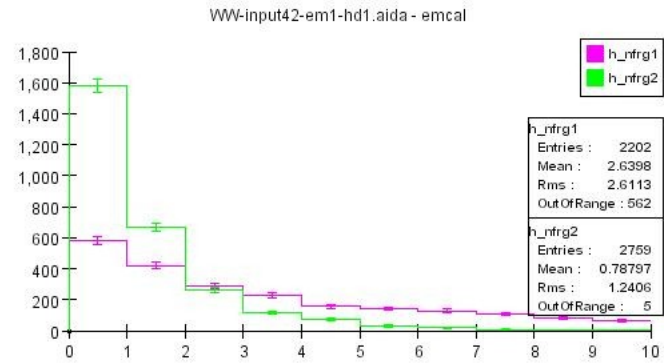
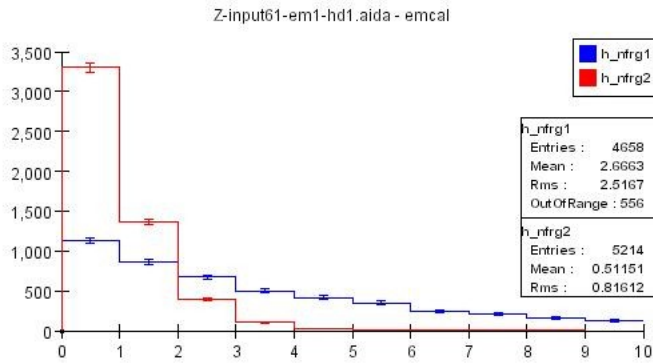


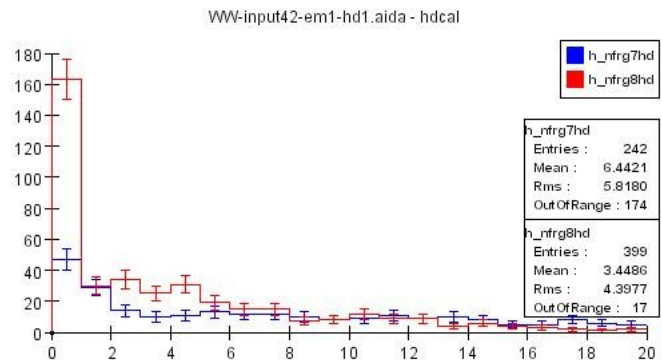
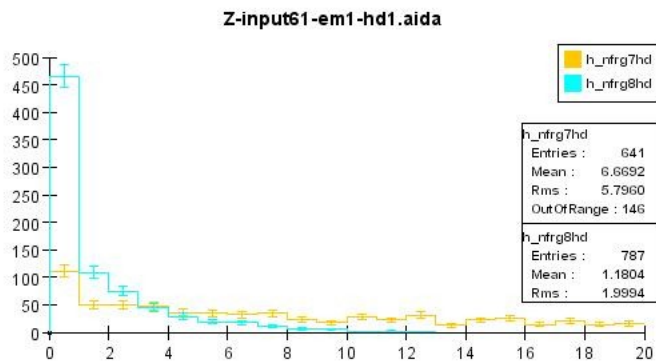
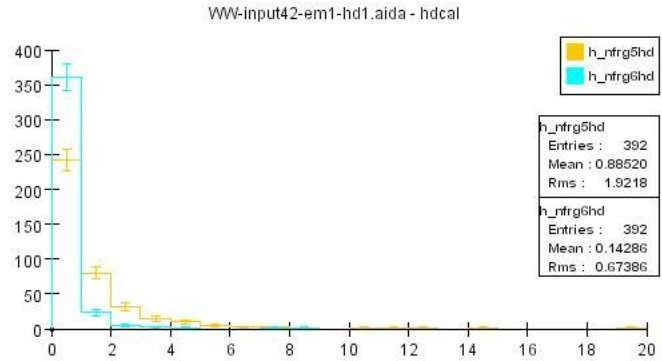
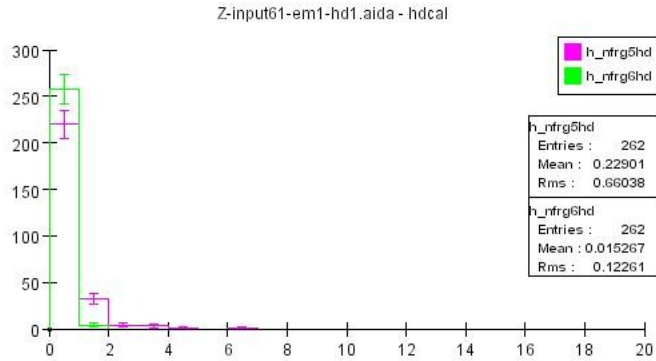
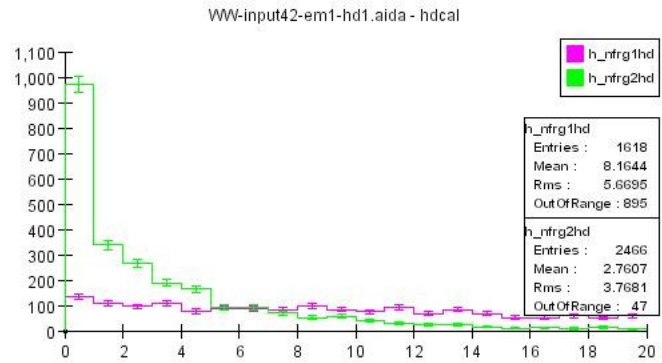
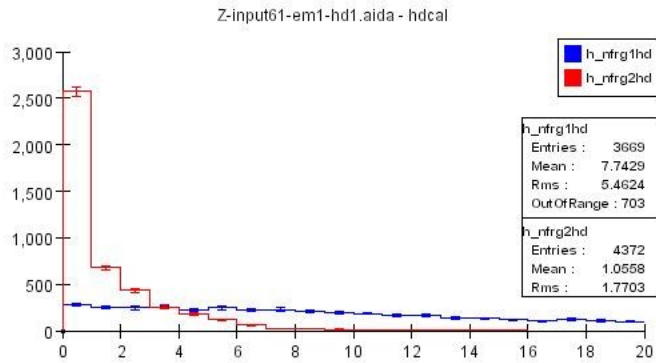
iphi vs ily



iphi vs iz (0.6, Evt 0, HD, 27)





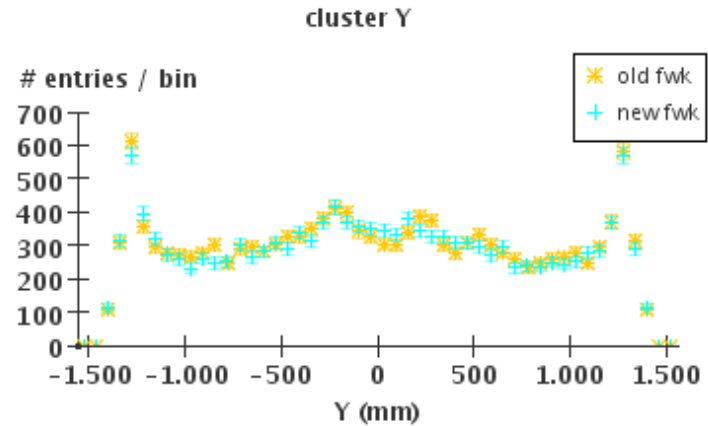
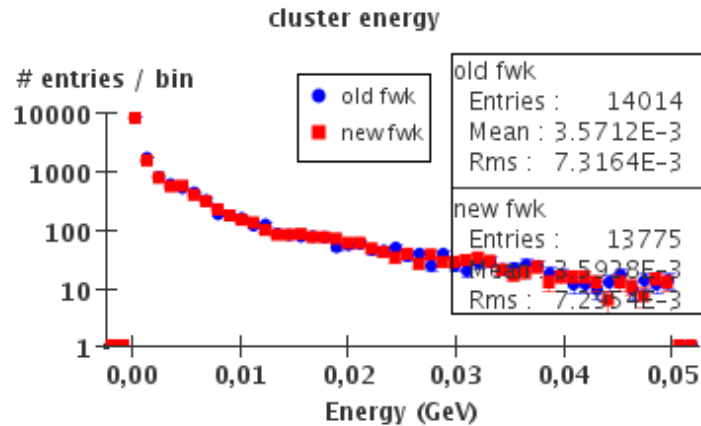


Certification studies

- Algorithm was developed using LCDG4 + LCIO + hep.lcd (old java framework)
- Ported to SLIC + org.lcsim, certification being finalized
- Common geometry: SDNPHOct04
- Very good agreement in ECal, more fragments in HCal

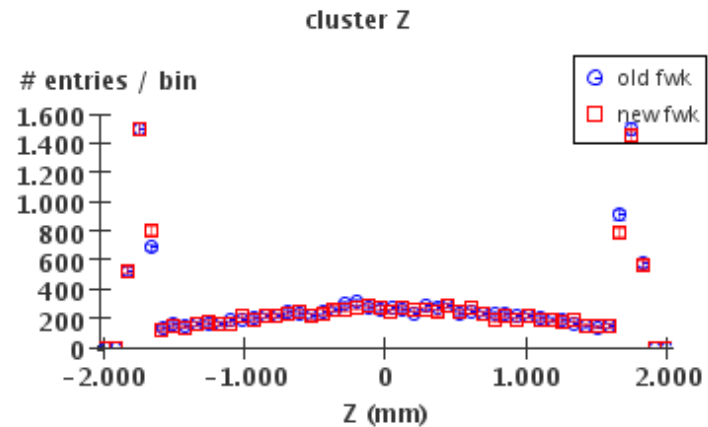
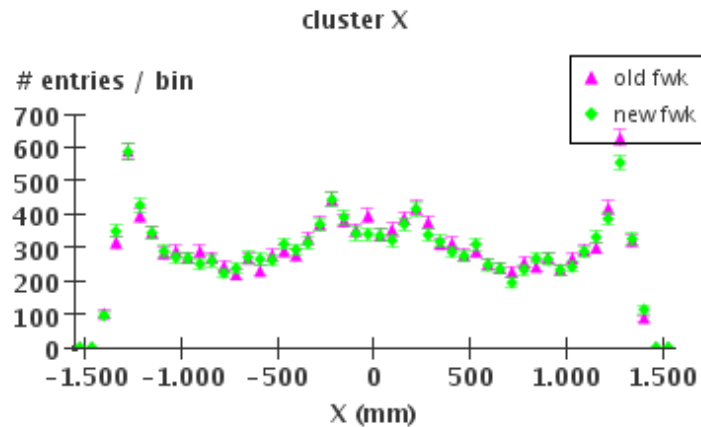
ECal barrel+EC: directed tree clustering

E



Y

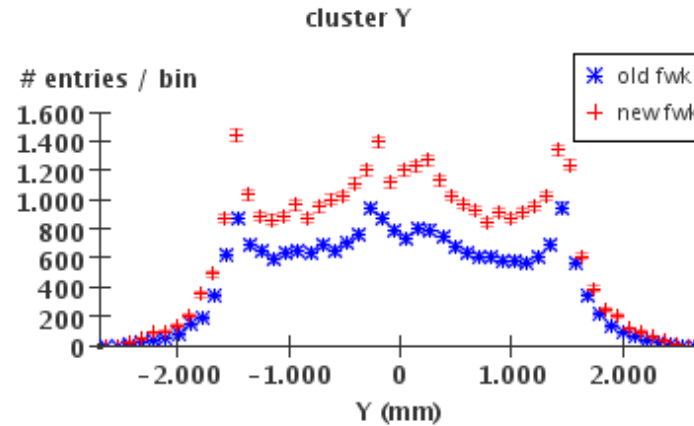
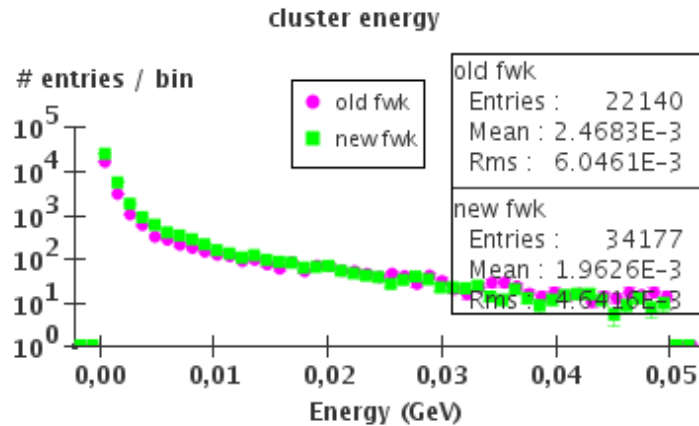
X



Z

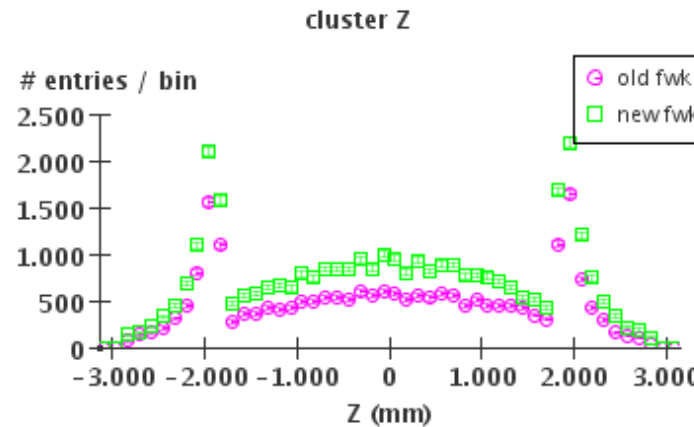
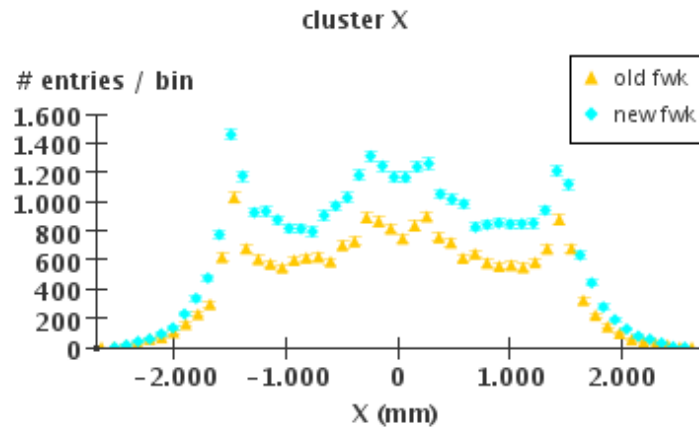
HCal barrel+EC: directed tree clustering

E



Y

X



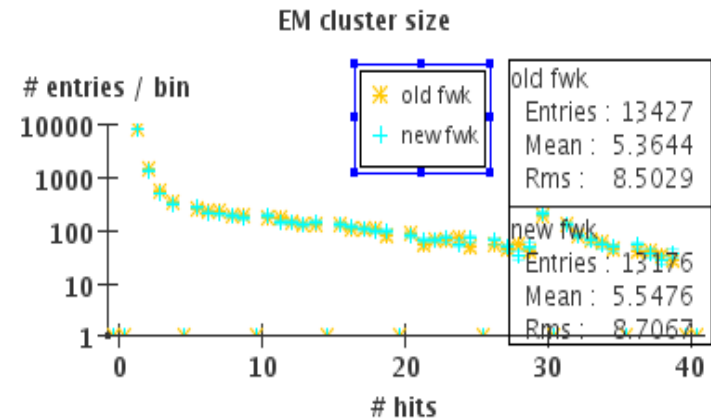
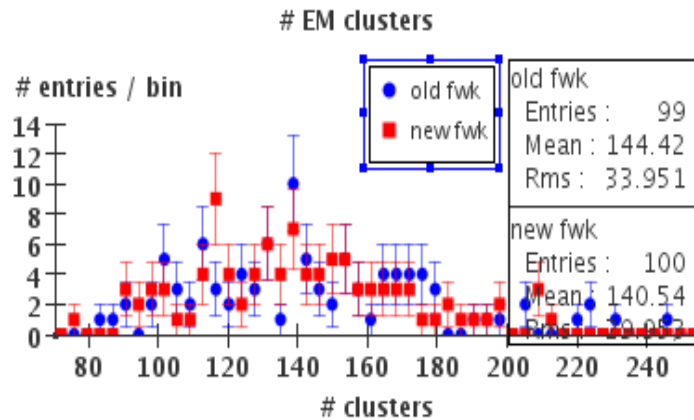
Z

directed tree clustering

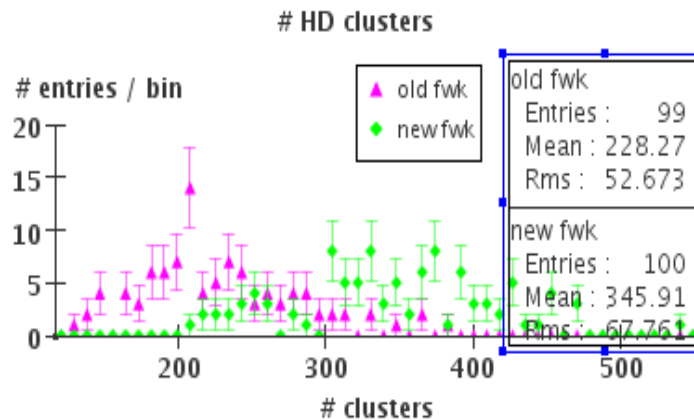
clusters

Cluster sizes (log scale)

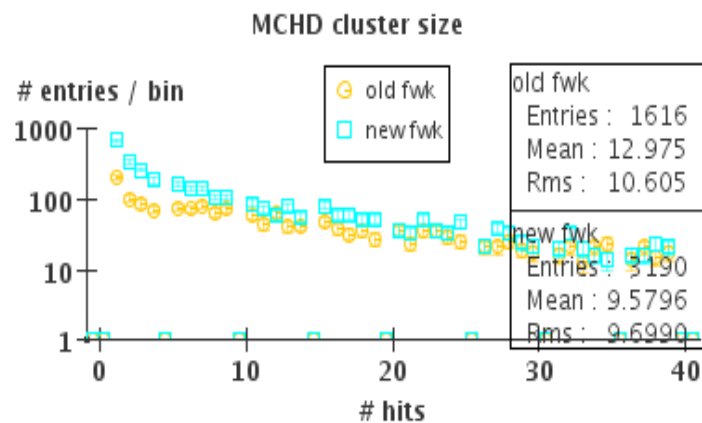
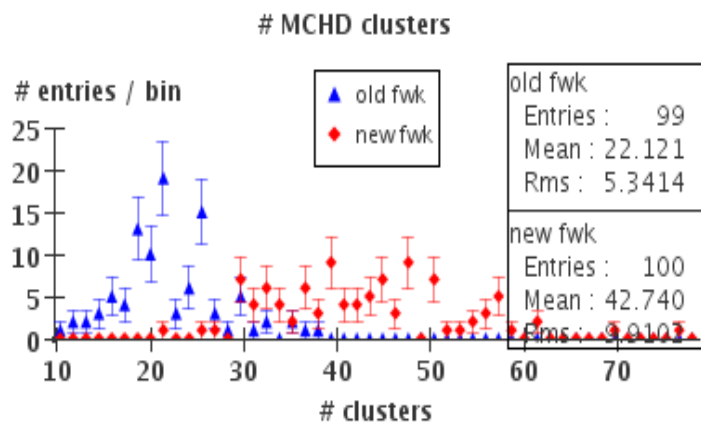
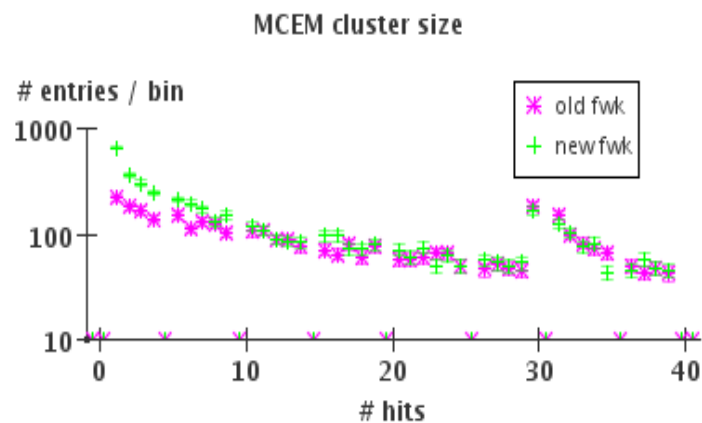
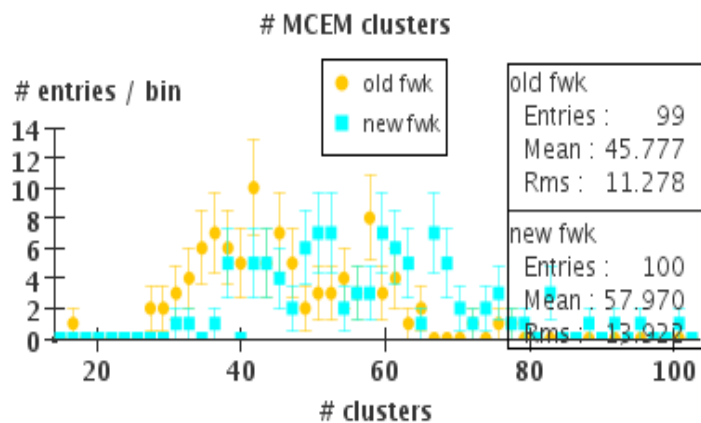
EM



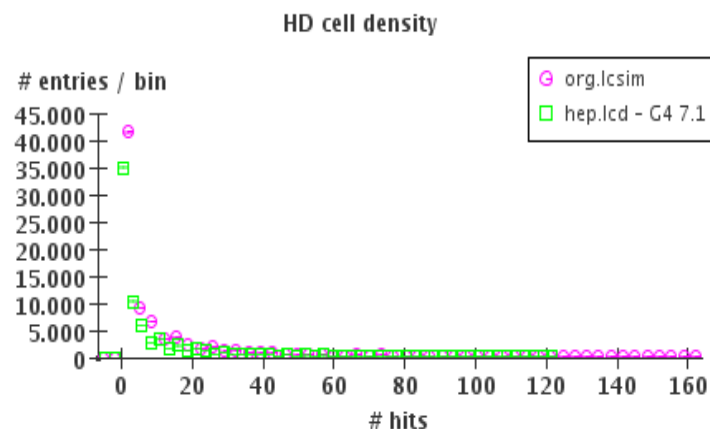
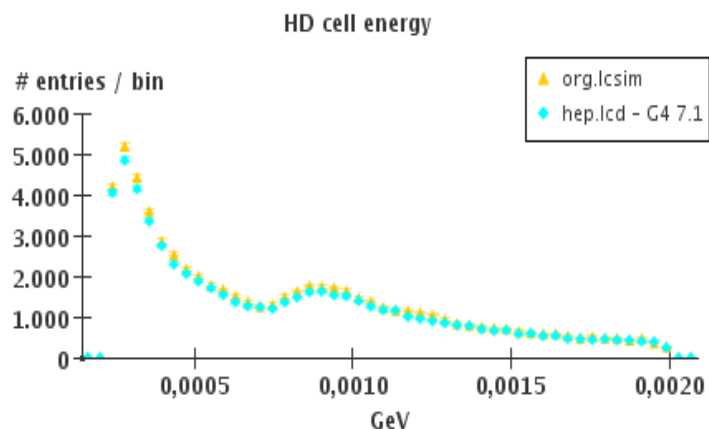
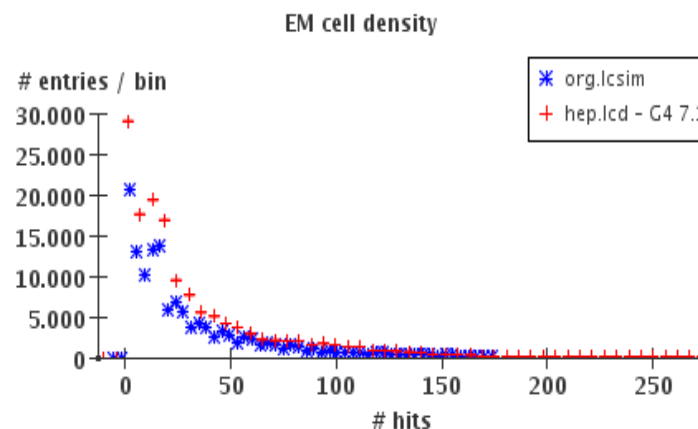
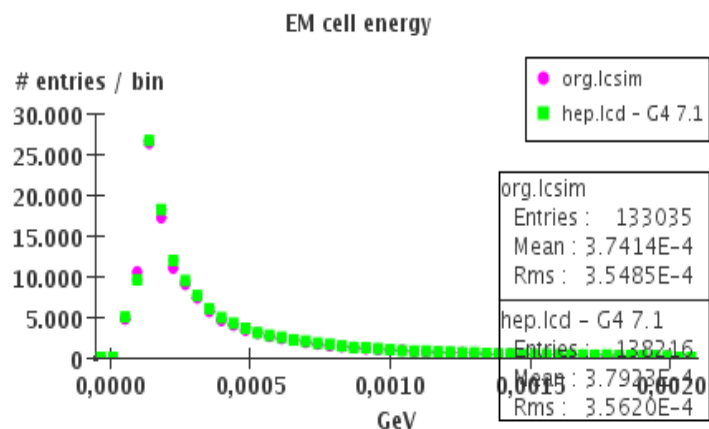
HAD



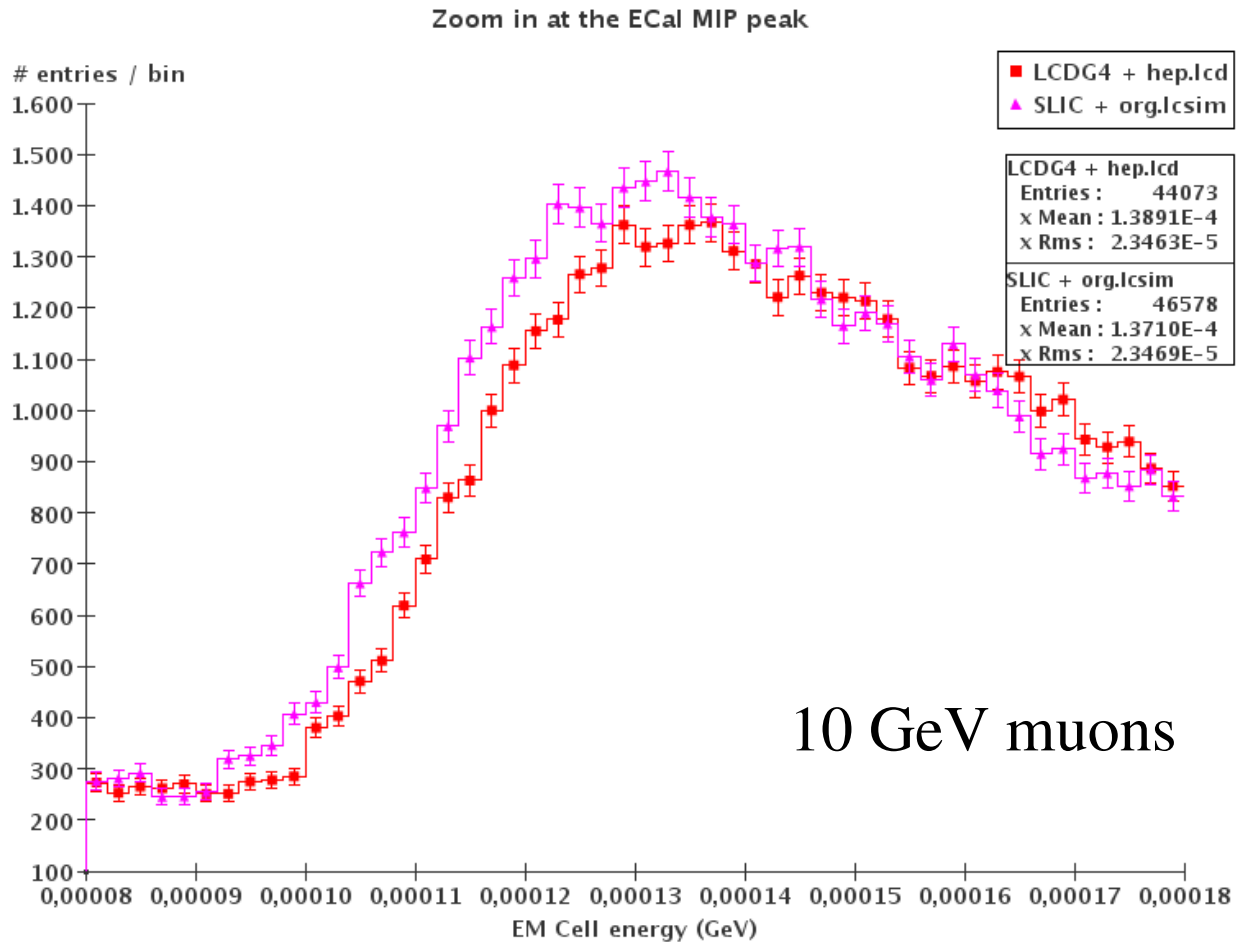
Perfect clustering (cheating)



Energies and densities



Zoom in at the ECal mip peak



Current status

- Directed tree clustering: neighborhood tuning is under way
- Correct assignment of fragments is challenging
discarding low multiplicity satellites and get hints from the track extrapolation seem to give good improvements, need to check if effect can be compensated by calibration without degrading resolution
- Clustering algorithm converted and well integrated to org.lcsim, certification to be finalized soon.
- More focusing on the PFA development effort
- This is work in progress, there is a lot of work to do!...

Things to do

- Some more parameter optimization (Hcal) (RM)
- Fragment assignment under study (VZ)
- Add other tools for a complete PFA algorithm
 - track matching (done for hits, clusters under way)
 - photon ID / cluster shape (H-Matrix)
 - More sophisticated calibration schemes
 - digitization effects