Clustering and PFA plans

Dhiman Chakraborty, Guilherme Lima, Rob McIntosh, Vishnu Zutshi

Snowmass Workshop
August 14-27, 2005
Clustering and PFA

- Goal: develop yet another complete Particle Flow Algorithm based on a calorimeter-only clustering algorithm (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
- Plans to consider digitization effects in the detector optimization (DigiSim)
Directed Tree Algorithm

- Define a cell neighborhood
- Discard low-energy hits (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
- Enter into histogram with weight $E_{\text{gen}} / E_{\text{total}}$
- Ideal: spike at 1

one entry per generated cluster
Cluster matching and merging algorithms

- Stage 1: one-to-one gen-reco matching, based on distances (3D or angular)
  --> several remaining clusters (“satellites”)

- Stage 2: attach satellites to reco clusters, based on angular distances
  possible cuts on angular separation, satellite energies, #hits
Preliminary ECal analysis

- 500 events, with 2-pions 10cm apart at Ecal face, using SDNPHOct04 detector
- neighborhood definition: $(d\phi=5, dZ=5, \text{dlayer}=9)$
- discard events with decays or interactions before Ecal
- Look at:
  - $\text{eratio1}$: $E_{\text{rec}}/E_{\text{gen}}$ after stage 1 (matching)
  - $\text{eratio2}$: $E_{\text{rec}}/E_{\text{gen}}$ 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:
  - $eratio1$: $E_{\text{rec}}/E_{\text{gen}}$ after stage 1 (matching)
  - $eratio2$: $E_{\text{rec}}/E_{\text{gen}}$ after stage 2 (merge satellites)
Current status

• Analysis of complex events shows some problems with too many isolated satellites
  satellites are isolated reconstructed clusters, too far from the main shower (how to connect them?)

• Clustering algorithm converted to org.lcsim, to be certified. Committed to LCSim CVS repository

• More manpower for 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)
- Add other tools for a complete PFA algorithm
  - track matching
  - photon ID
  - cluster shape
  - MIP tracking in calorimeters (to help connecting satellites?)
  - digitization effects