Cluster MCParticle Truth Matching

Alic Spellman 03/02/21







Intro

- Need reliable Cluster truth matching to finish evaluating performance of new KF+GBL compatible TrackClusterMatcher algorithm
- Find cases where a MCParticle is matched to multiple Clusters
 - Often these clusters are in completely opposite quadrants on the Ecal
 - How is it possible to truth match two clusters in opposite sides of Ecal to same MCParticle?
- May be issue with truth matching method
- May be issue with truth information in simulation...



Cluster Truth Matching

- Given EcalCluster, find "SeedHit"
- SeedHit is associated with Ecal Readout Hit
- EcalReadoutHit is associated with multiple Simulated Calorimeter Hits
- Each SimCalHit contributes some fraction of the EcalReadoutHit Energy
- Select SimCalHit with largest Energy contribution
- SimCalHits can have multiple MCParticles associated
- Select MCParticle that contributes most energy to SimCalHit
- Truth match MCParticle to Ecal Cluster

ecalReadoutHit **EcalCluster SeedHit** SimCalHits **MCParticles**



Checking Truth Match Energy

MCParticle Energy (GeV)

- After matching Clusters to MCParticles, 2d plot energy
- Top horizontal line shows FEE's

.

- Entries left of x=~2.3 represent edge hits
- **Central diagonal line** shows reasonable Cluster MCParticle matches
 - MCParticles with energy < ~400
 MeV should be outside Ecal
 acceptance
- Bottom horizontal line represents bad matches between Cluster and MCParticle

18 2.5 cluster vs electron mcp truth energy Entries 6704 16 Mean x 0.816 Mean y 1.141 14 Std Dev x 0.6272 Std Dev v 0.8586 12 cluster vs photon mcp truth energy 1.5 10 Entries 4947 Mean x 0.8021 Mean y 0.1638 8 Std Dev x 0.6265 Std Dev y 0.3398 6 cluster_vs_positron_mcp_truth_energy Entries 2284 4 0.5 Mean x 0.5823 Mean y 0.7348 2 Std Dev x 0.2924 Std Dev y 0.3467 0<u>`</u> 0.5 1.5 2.5 3.5 4.5 2 3 **Cluster Energy (GeV)**

cluster_vs_electron_mcp_truth_energy



- We find many cases (~12%) where single MCParticle is truth matched to multiple Clusters
 - Should this ever be possible?
- Maybe we can understand duplicate matches <u>if the Clusters are very close</u> <u>together in position</u>...
 - If they are close, should the Clusters be "clustered" themselves?
- Lets see how close together these duplicate
 Clusters are





- Given a MCParticle that gets truth matched to two different Clusters
- Check how far apart those two Clusters are
- We find cases of two Clusters that are supposedly products of the same MCParticle, on completely opposite halves of the Ecal...
- Top plot shows (Cluster1_y Cluster2_y)
- Bottom plot shows (Cluster1_x Cluster2_x)
- How can a single MCParticle make clusters in opposite halves of the Ecal?
 - Obvious mistake in truth matching







- Plots show position residuals in the xyplane between two Clusters that share a MCParticle
- Broken down into photons, electrons, positrons
- Seems like this issue isn't based on particle type...





cluster_positron_mcp_duplicates_dr





- Check the XY plane position residual between a Cluster's Simcalhit and the largest energy contributing MCParticle in that Simcalhit
- Plotted along MCParticle Endpoint Z, since it varies, where Simcalhit is always at Ecal face ~1330mm
- If MCParticle endpoint is in a tracking layer, why is it associated with a Simcalhit?
- We see cases where Simcalhit positions are very far (> 100 mm) from the largest contributing MCParticle on that hit....

simcalhit_truthmatched_MCP_z_v_dr





MCParticle Truth Matching Update Cluster energy: 1.1274969830000001 Ecal Cluster calhit raw energy: 1.1274969830000001 Seed Hit seedhit raw energy: 1.1274969830000001 seedhit cellID: 66317 Ecal Readout Hit readoutMatchHit cellID: 66317 Recalled from previous example of Cluster truth matching where Cluster with ~1 GeV of energy is simcalhit raw energy: 2.4397425295319408E-4 simcalhit MCParticleCount = 1 matched to MCParticle with ~10 MeV Potential issue (or misunderstanding on my part) simcalhit raw energy: 0.054429084062576294 Simulated Cal Hits simcalhit MCParticleCount = 1 with how the MCP simcalhit energy contributions simcalhit raw energy: 1.2112891674041748 are assigned in simulation... simcalhit MCParticleCount = 4 Appears that the contributed energy is simcalhit raw energy: 3.2516857027076185E-4 assigned to the wrong particle simcalhit MCParticleCount = 1 Or there are some particle tree complications simcalhit MCP Contributed energy: 0.18687056005001068 MCP energy: 2.3000081028758887 Solution is to match to largest energy MCParticle simcalhit MCP Contributed energy: 1.0177032947540283 in simcalhit *instead* of largest "contributed" energy MCP energy: 0.0018236056373091374 simcalhit MCP Contributed energy: 0.0065940129570662975 MCP energy: 6.695142248573957E-5 simcalhit MCP Contributed energy: 1.2121284089516848E-4 MCP energy: 6.694394932622357E-5 ratio of Cluster/MCP Energy: 618.2789523856177 🚤



MC Particles

Truth match MCP

w largest energy

contribution

MCParticle Truth Matching Update

- Solution is to match to largest energy MCParticle energy in simcalhit *instead of* largest "contributed" energy...
- Bad matches that formed the bottom horizontal line in previous plot significantly reduced
- Apply some additional cuts and cleaning....
- Require Cluster MCParticles origin at target
- Impose Py and cluster y position sign matching
 - MCPs with negative Py at target should hit bottom of Ecal, and vice versa





MCParticle Truth Matching Update

- We find number of MCParticles matched to multiple Clusters reduced from $\sim 12\%$ to $\sim 2\%$
- After those sanity cuts, check for duplicates and adjudicate each case if possible
- Compare MCParticle Endpoint position with each Cluster position
- If Cluster and MCParticle are in different quadrants, match can't (or shouldn't) be correct
- If duplicates remain...cut them all out





MCParticle Truth Matching Update

- In the end we see that the Cluster MCParticle truth matched pairs look much improved using this new method of matching
- Just a few stragglers remain, where the matching doesnt appear to be correct

•

 MCPs with P < 400 MeV shouldn't reach the Ecal Face

cluster_truth_stage_final_cluster_v_mcp_energy





Summary

- With the initial Cluster truth matching, we found sizeable instances of MCParticles being matched to multiple Clusters, even Clusters in different Ecal quandrants
- There may be some issue with the way Simcalhit MCP energy contributions are assigned in simulation
- Updating the truth matching to match a Cluster to the MCP with the largest energy has cleaned up most of the "bad" matches we had previously
- There remain a tiny fraction of weird cases
 - In principle, should these cases be impossible?
- The Cluster truth matching appears to be good enough to move forward

