Track-Cluster Matching

Tristan Protzman

Lehigh University

April 17th, 2025

Intro

- I was asked to implement a track-cluster matching factory into EICRecon
  - Used as input to electron identification, particle flow algorithms
  - We need to move away from using truth information in these algorithms
- Based on stand-alone implementation from Derek Anderson
  - MatchProjectionsAndClusters.cxx - github.com
- This is a first pass to enable us to start developing algorithms without truth data, I'm sure there's plenty to improve
- Related issue: EICRecon Issue #1648
- Related PR: EICRecon PR #1768

- Definition: $\Delta R = \sqrt{\Delta\phi(\text{track},\ \text{cluster}) + \Delta\eta(\text{track},\ \text{cluster})}$

Algorithm

- Inputs: Track Segment Collection, Cluster Collection
- Output: TrackClusterMatchCollection
  - Association between edm4eic::Track and edm4eic::Cluster
- For each cluster, loop over all track segments
- Check if each point in the segment is at the surface of either the EMCal or HCal
- Find the closest point in $\eta - \phi$ space to the cluster
- Create an association

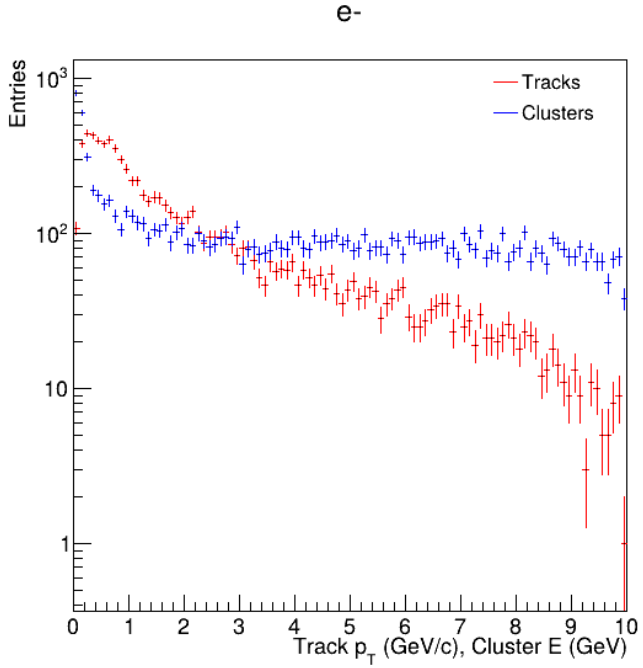
Pseudocode of TrackClusterMatch.cxx

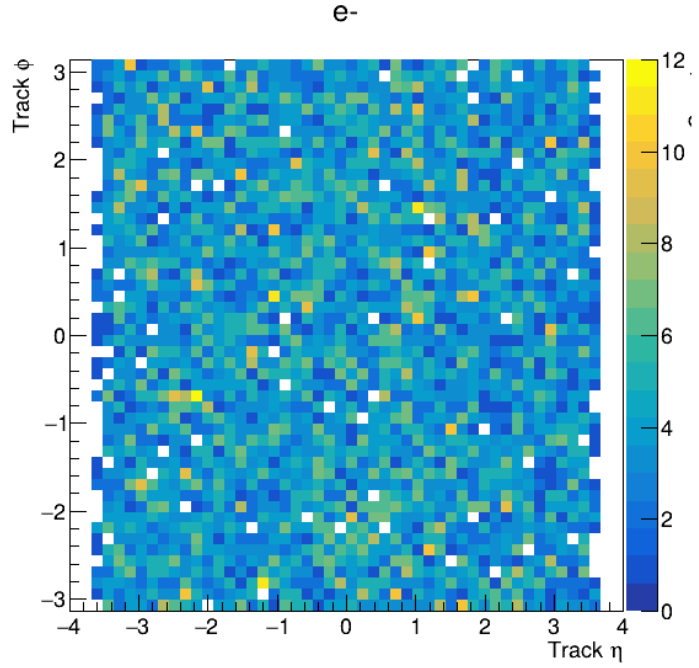
```
for cluster in clusters:
  best_match = none
  for track in tracks:
    if track is used:
      continue
    for point in track:
      if point is not at calo surface:
        continue:
      if ΔR(track, cluster) < best_match:
        best_match = track
```

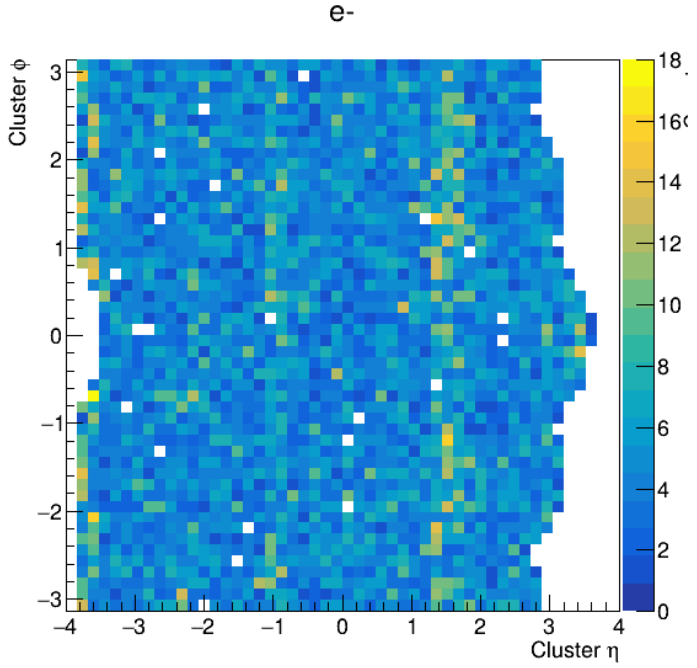
Performance studies - input

- Using single particle simulations
  - 0.1 to 10 GeV/c momentum
  - 2 to 178 degrees theta, flat in eta
  - Electrons, muons, pions, kaons, protons, and their antiparticle

Generated electron sample (other species here)

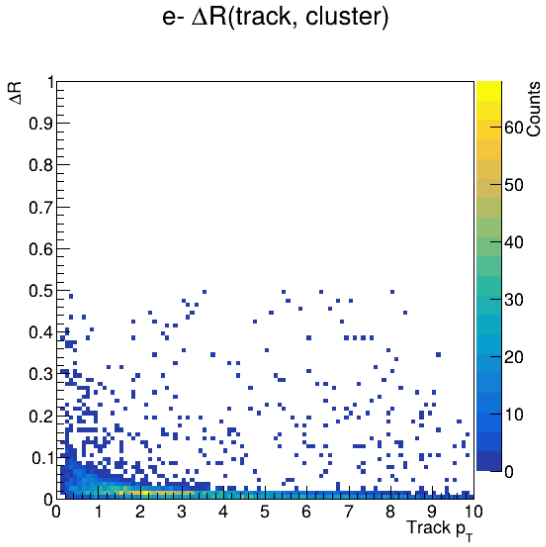


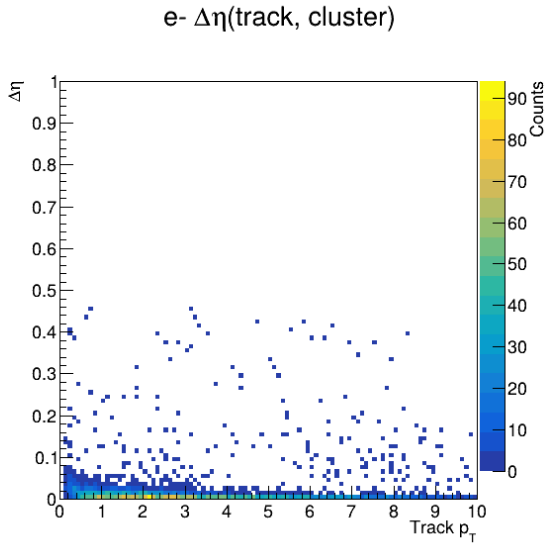


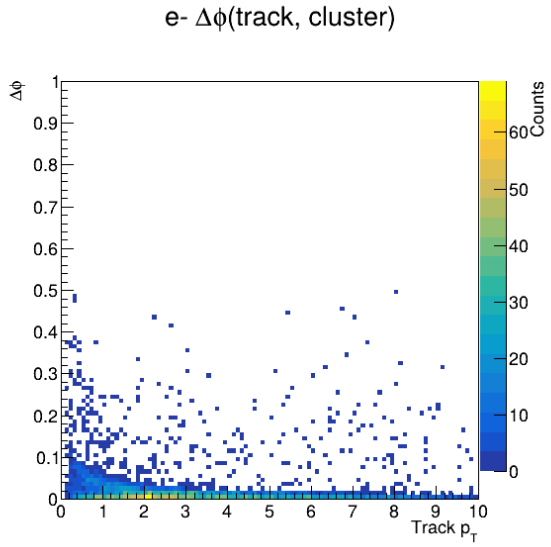


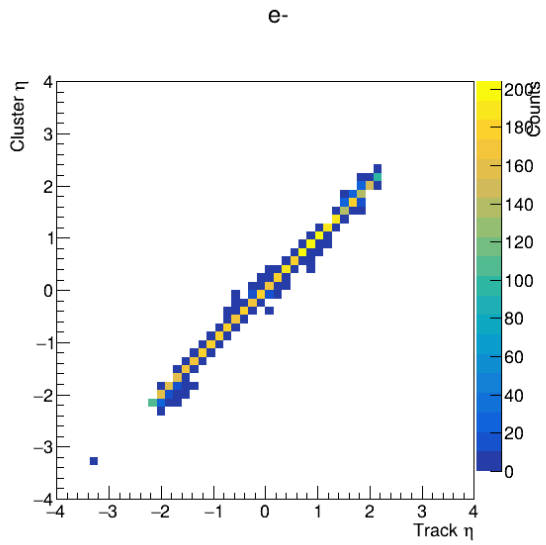
Performance studies - matching

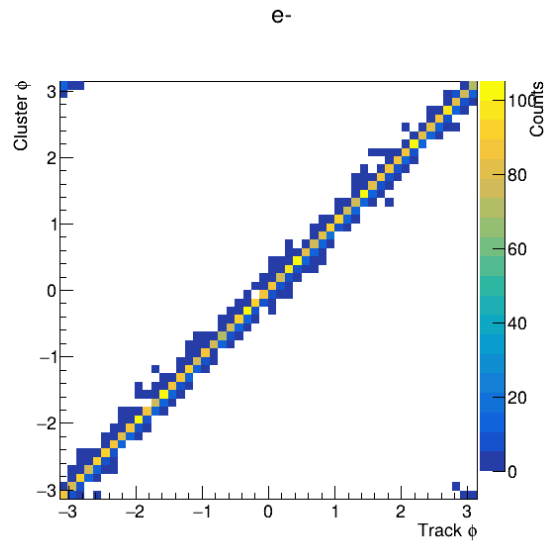
- Focus on electrons for input to electron ID algorithm
- Some structure in ΔR and $\Delta\phi$ at low p_T
- No matches outside ~|2| in η, not sure where the actual limits are
  - Tracks and clusters are generated beyond this range
- Other species here

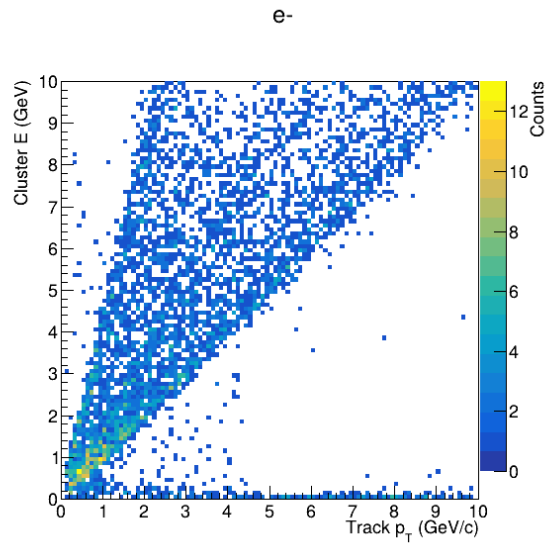






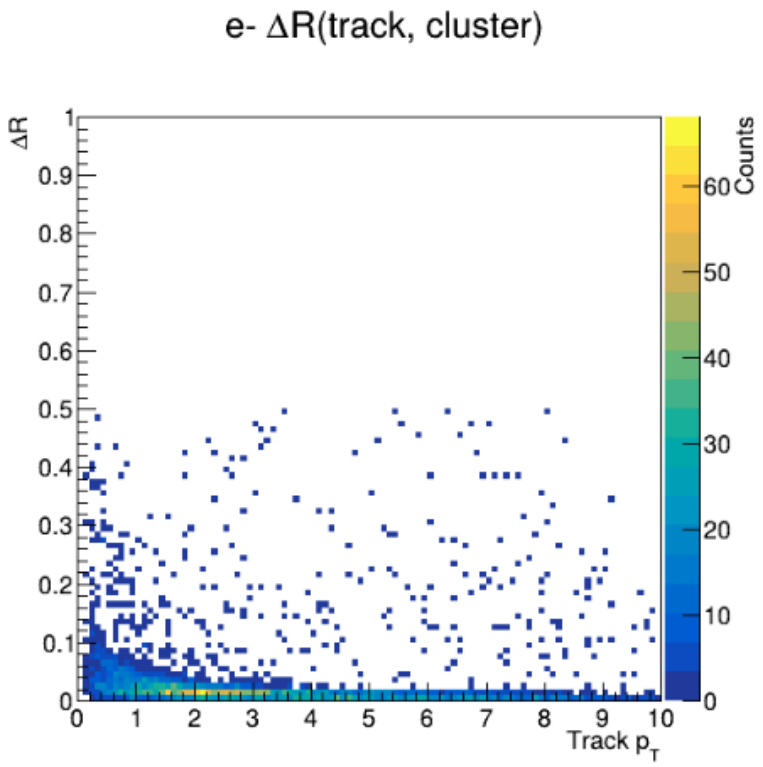


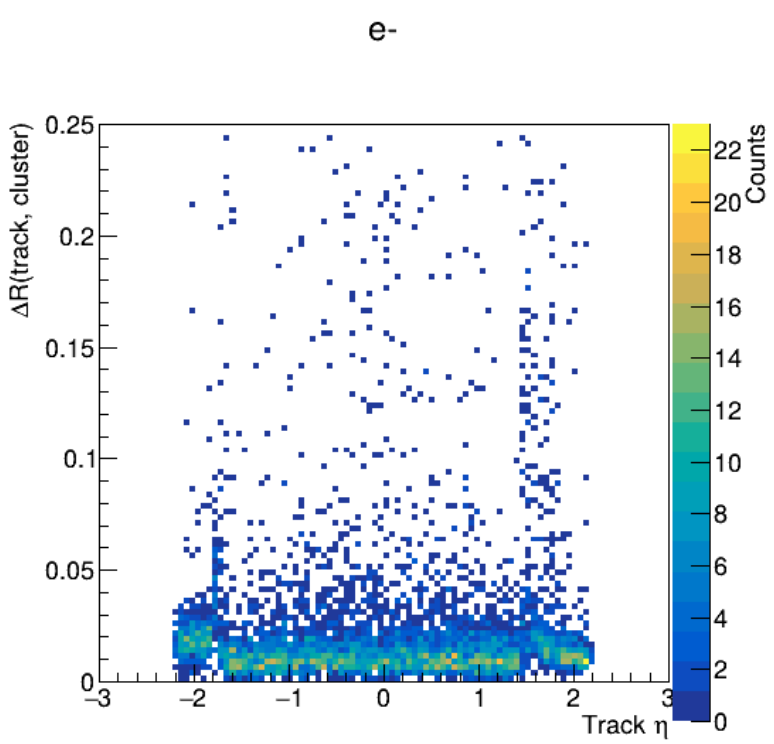


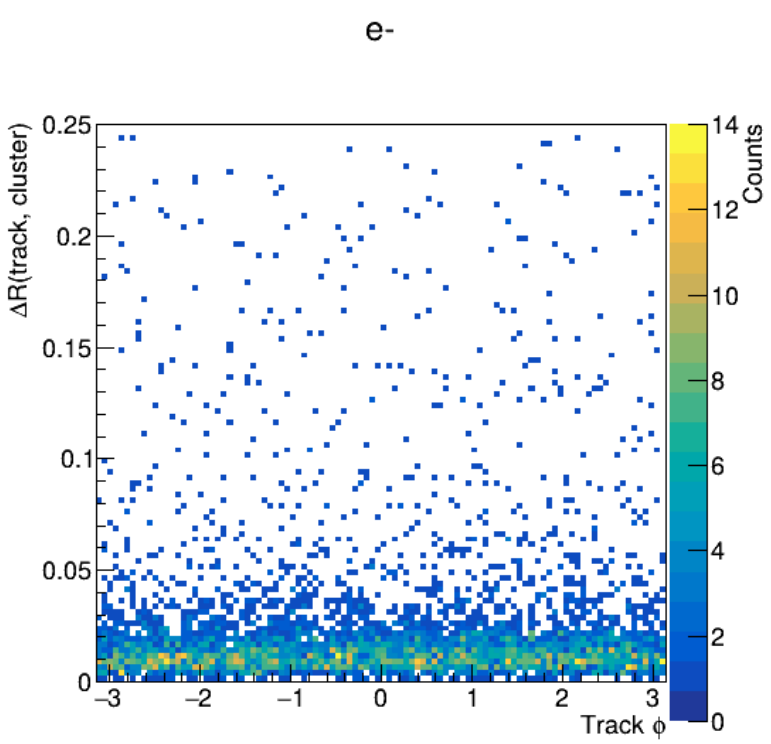


Performance studies - uniformity

- Looking at ΔR as a function of η and ϕ shows discontinuities at what (I think) are detector boundaries

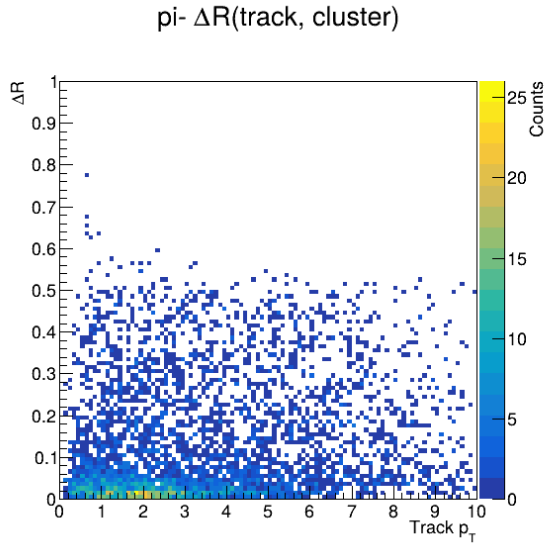


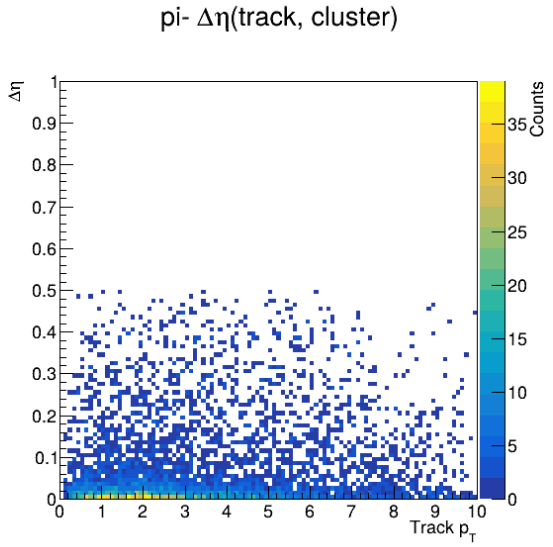


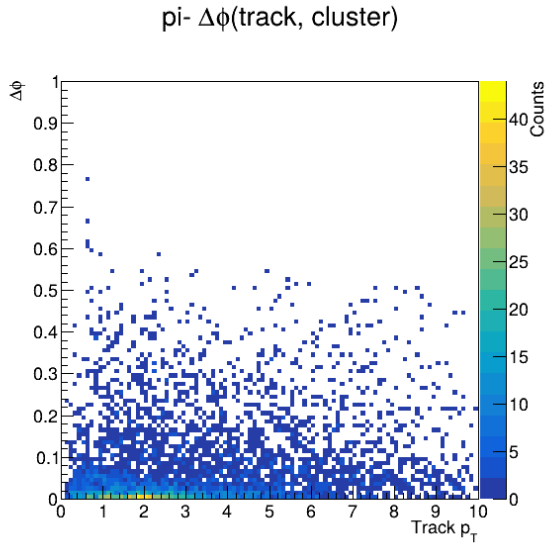


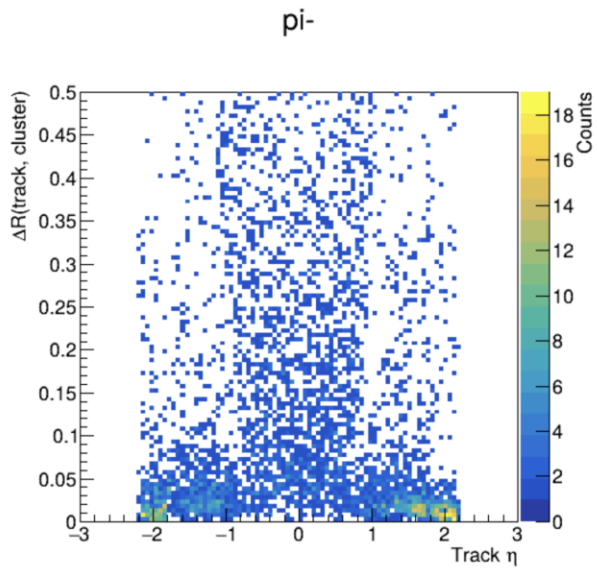
Performance studies - hadrons

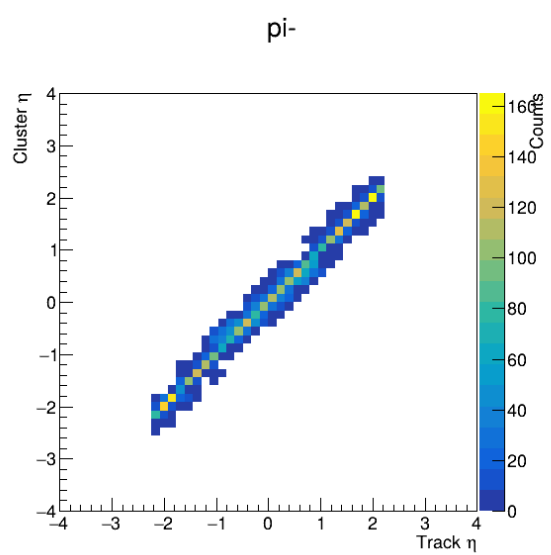
- Hadrons have similar characteristics with wider distribution
- Probably not too surprising, hadronic showers will be larger than EM showers
- Strong η dependence

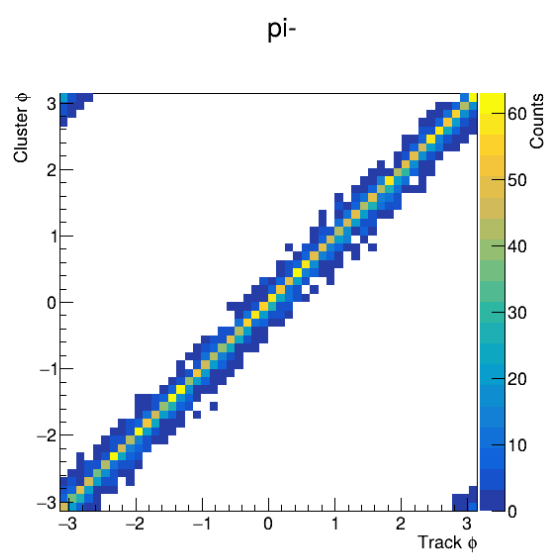


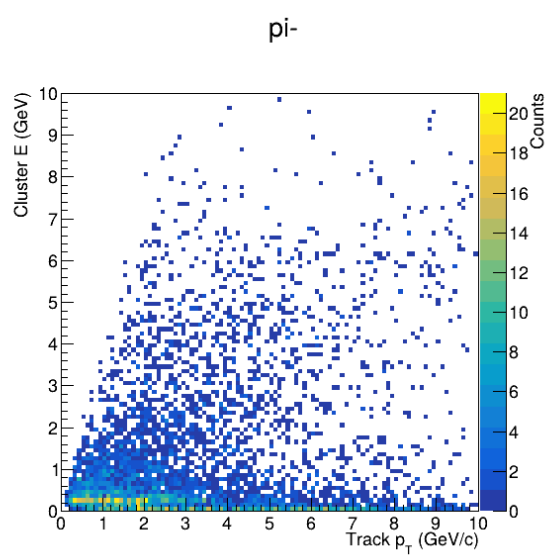










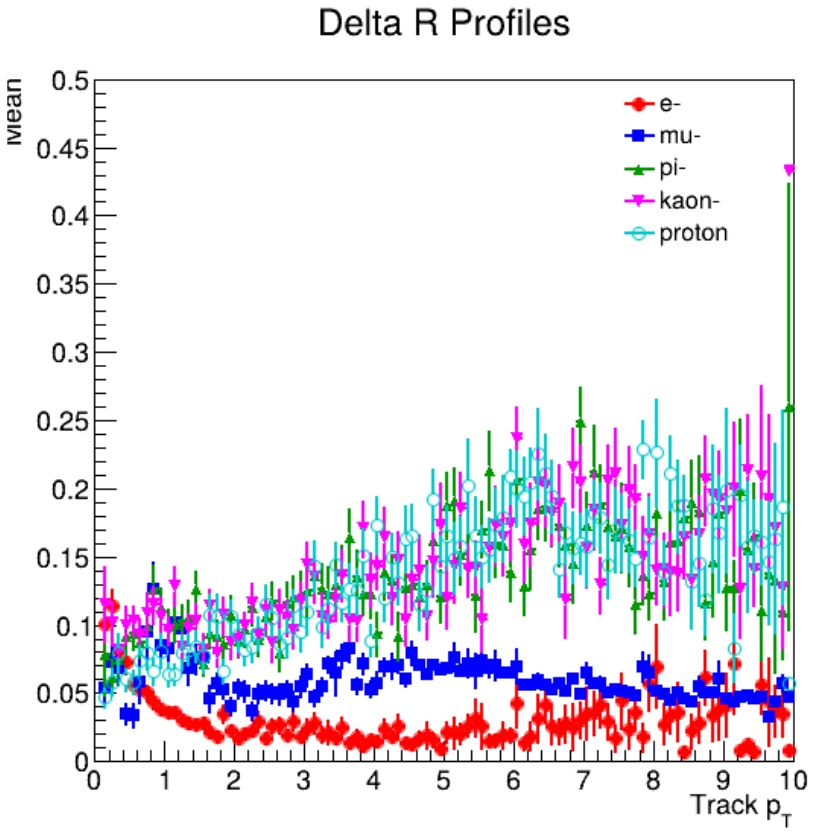


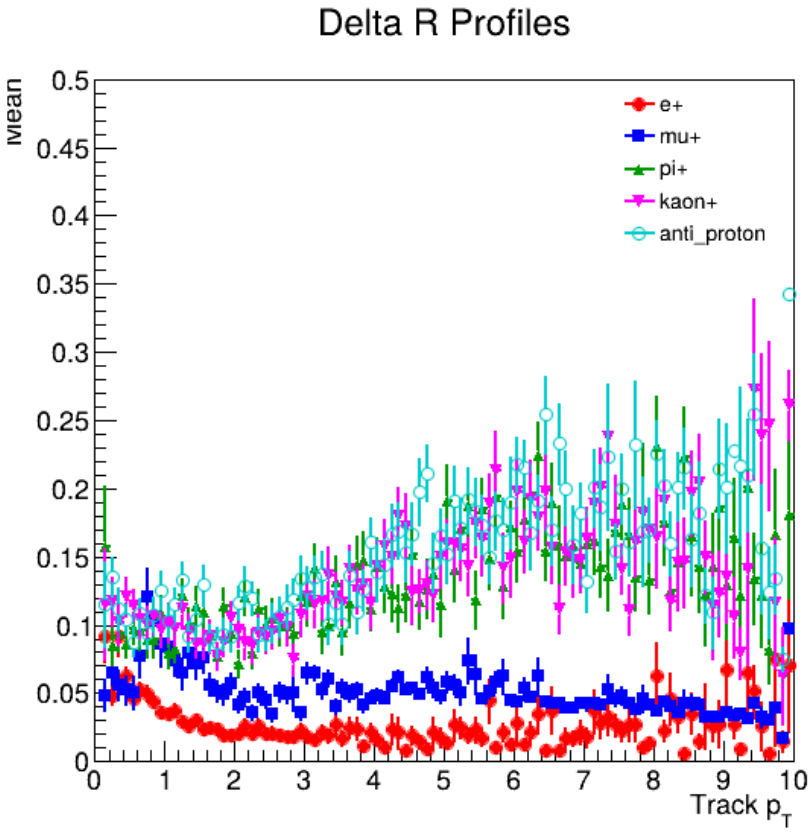
Summary

- Profile of ΔR for each particle type
- I've been told the calorimetry clustering doesn't deal well with muons yet

Next steps

- Adapting electron ID code to use this rather than truth
- Package into a reconstruction benchmark to run for each PR
- Evaluation of full events





What would be helpful from the tracking side?