Charged Pions in MAIA

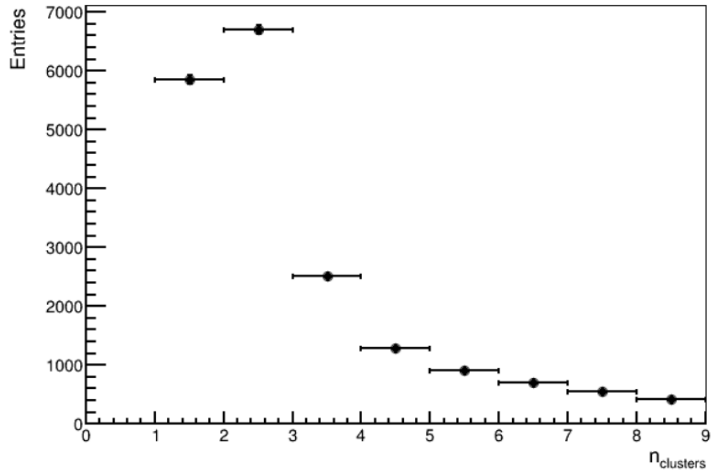
Sarah Demers, Ethan Martinez, and Gregory Penn
Yale University

Quick Reminder

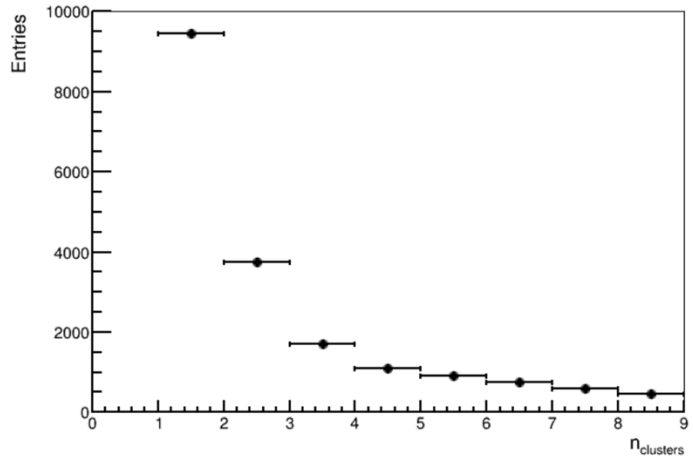
- w/ pion gun: found drastically different cluster “qualities” (link)

SiTrack

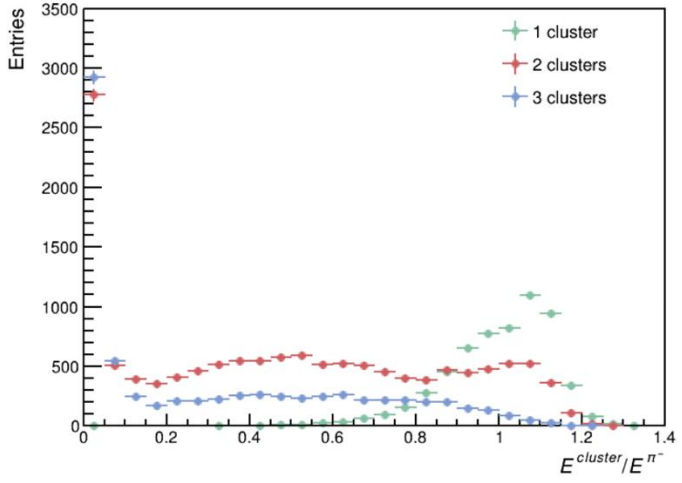
clusters less split up

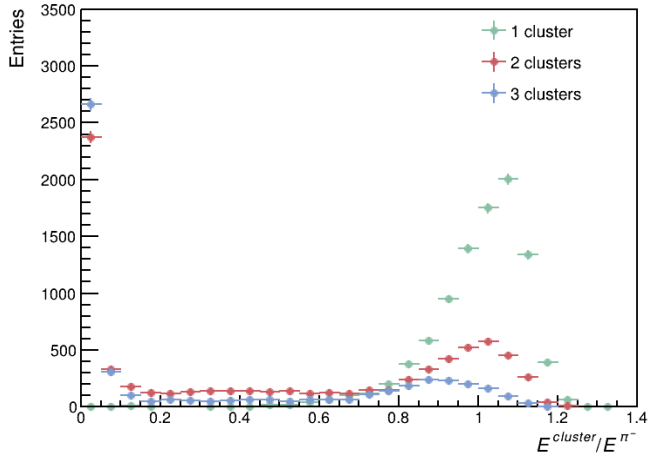


SiTrack_Refitted



one cluster more often captures all $\pi^{\pm}$ E

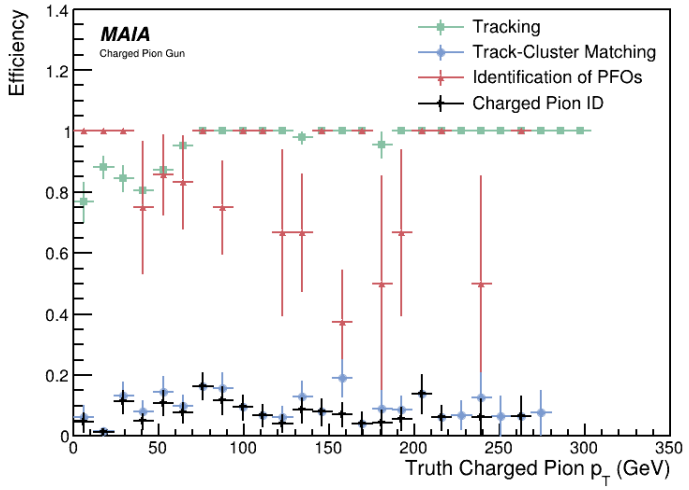


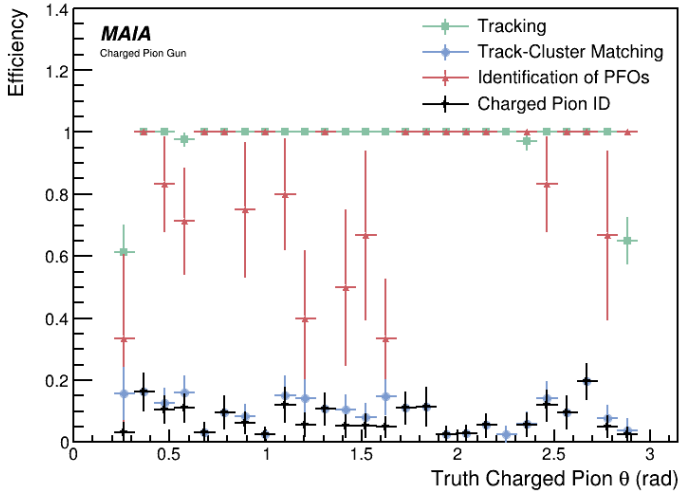


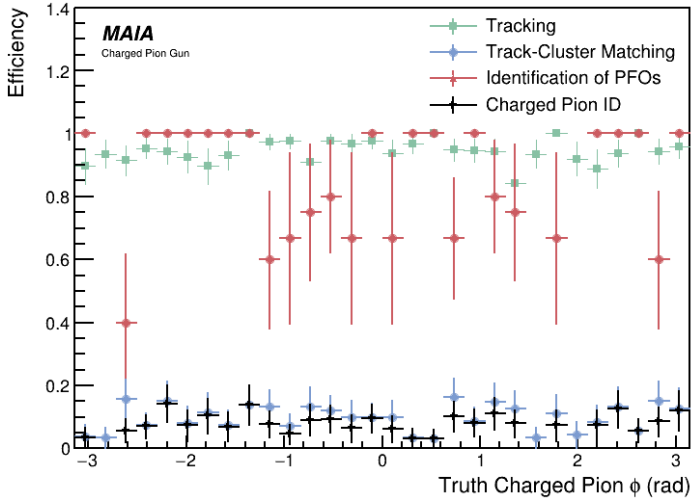
Quick Reminder

- This propagates down reconstruction in the form of improved track-cluster matching efficiency:

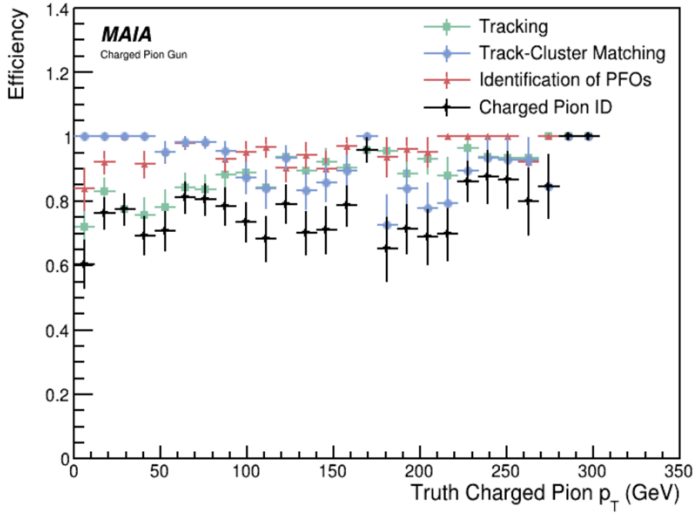
SiTrack

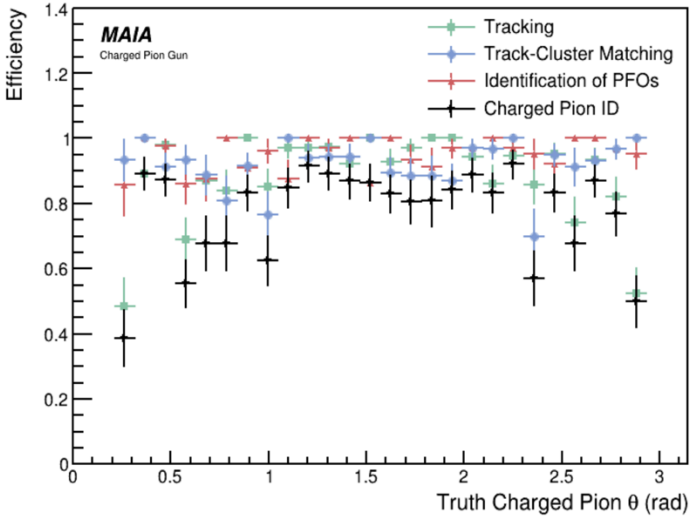


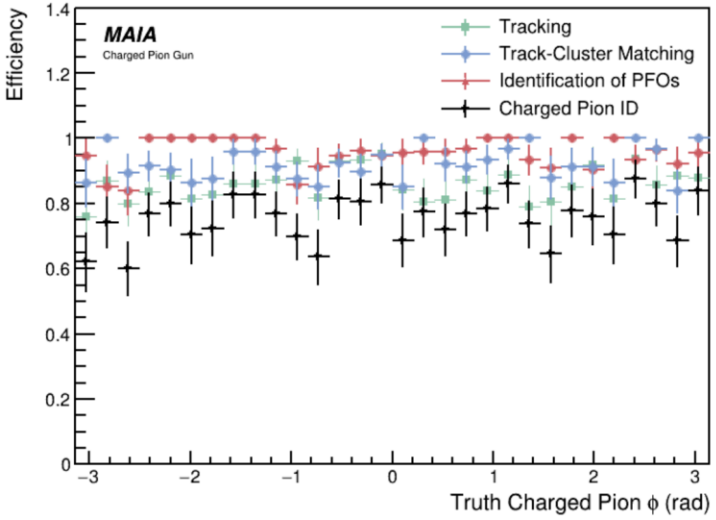




SiTrack_Refitted

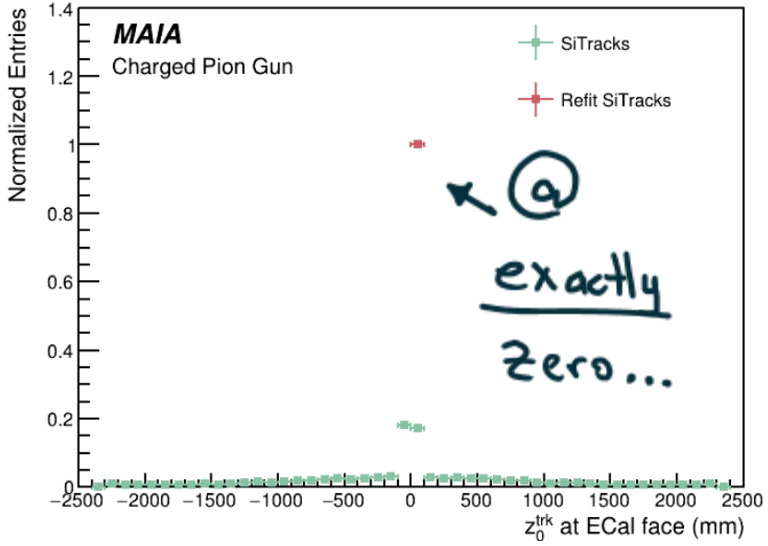


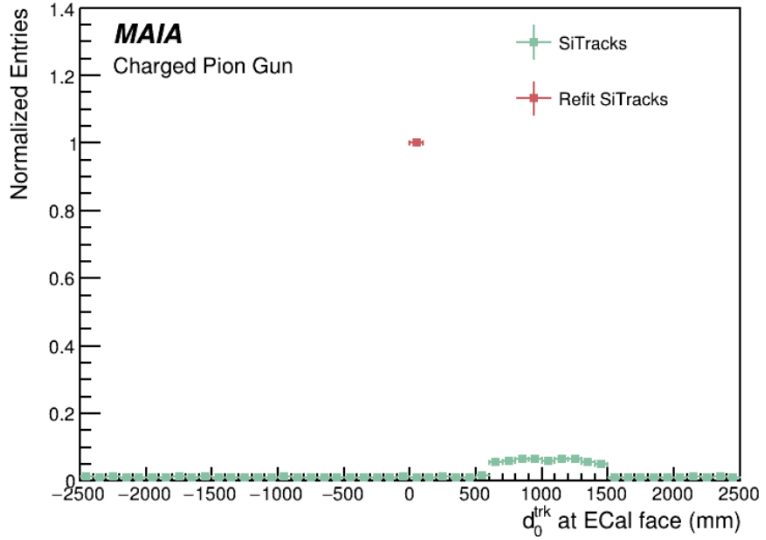


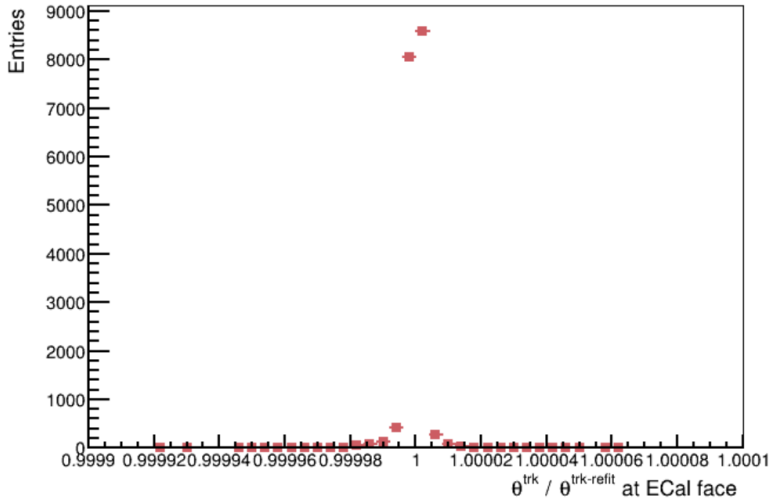


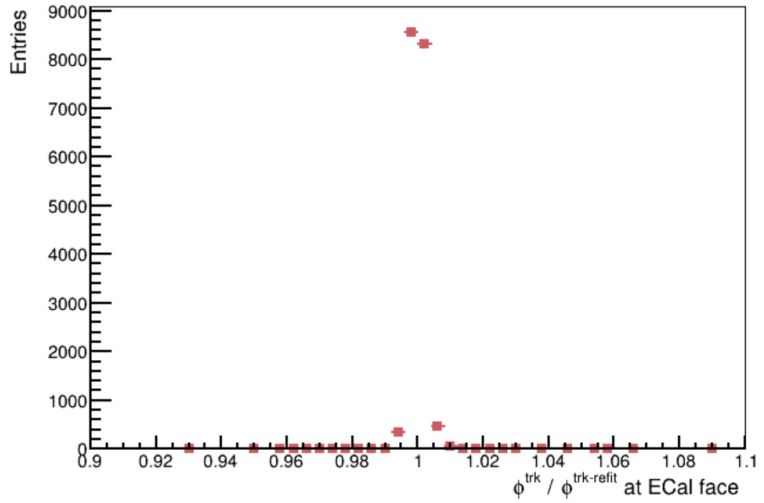
Quick Reminder

- Track states @ ECal face were not particularly enlightening (link)



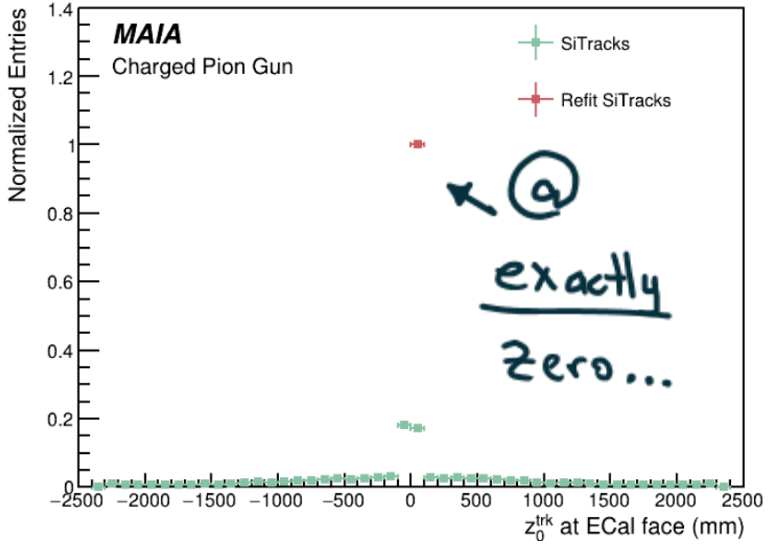


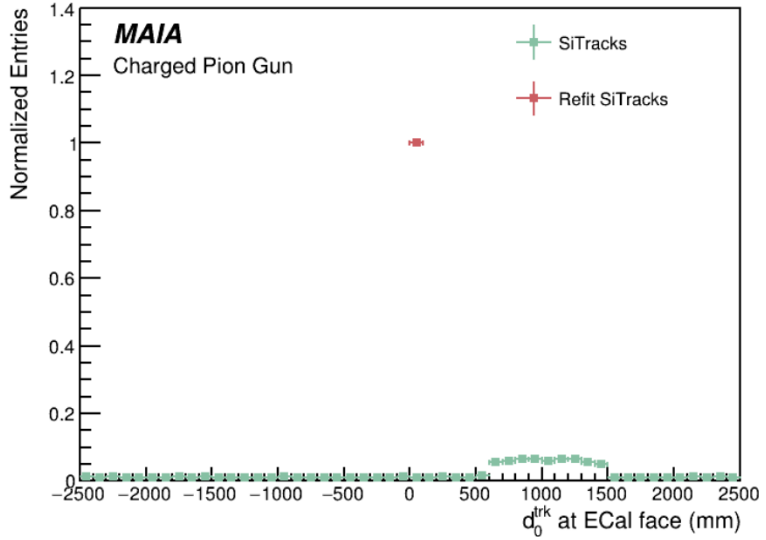


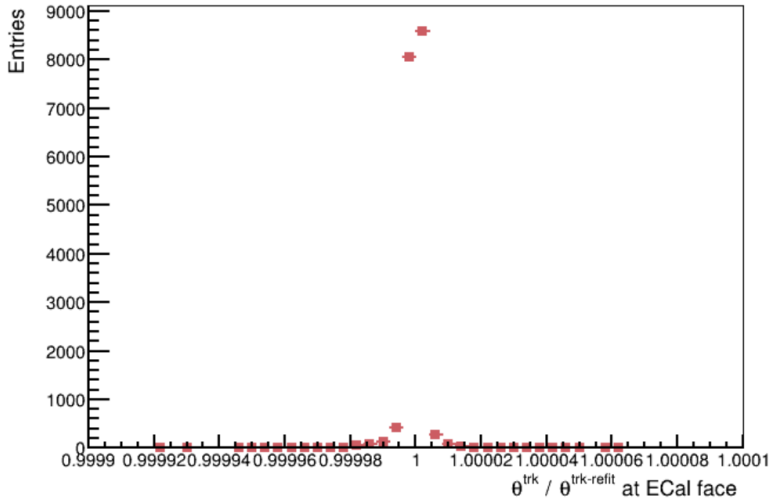


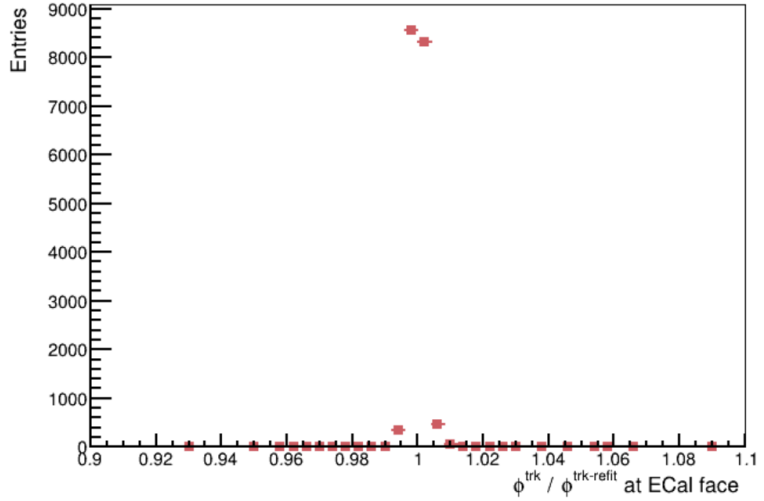
Quick Reminder

- Track states @ ECal face were not particularly enlightening (link)









motivates a deep look into Pandora!

this week.

Running Pandora Locally

- Given these bugs and previous challenges attempting to alter Pandora from the outside, I figured it was worth the time to check pandora out locally
- Thanks to help from Thomas Madlener, I was able to do so
- Instructions can be found here
- It's not difficult to replicate
  - Following the instructions takes ~ 5 minutes of work

Pandora Algorithm in a Nutshell

Link to Pandora paper (thanks Tova!)

- Pandora is the combination of many algorithms:
  1. Track selection (section 4.1 of paper) – details in previous presentations (1, 2)
     1. Not the culprit of poor efficiency
  2. Clustering
     1. Photon clustering
     2. Fast photon identification
     3. Cone clustering
     4. Topological cluster merging
        1. This involves ***9(!)*** edge cases that recluster, targeting various phenomena (some involving tracks)
     5. Re-clustering
     6. Photon identification and recovery
     7. Fragment removal
     8. Formation of particle flow objects
- It's clear that most, if not all (I haven't checked everywhere yet) of the algorithms are run somewhere in the code
- What's not as clear from the paper is the extent to which they all talk to each other

Track-Cluster matching all over!

- Previously I thought that track-cluster matching (code) was just done once, after clustering
  - The paper implies this
- The Pandora config I took from the tutorial has ***24*** instances of the track-cluster matching algorithm

```
<algorithm type = "TopologicalAssociationParent" description = "ClusterAssociation" instance = "reclusterAssociation">
    <associationAlgorithms>
        <algorithm type = "LoopingTracks"/>
        <algorithm type = "BrokenTracks"/>
        <algorithm type = "ShowerMipMerging"/>
        <algorithm type = "ShowerMipMerging2"/>
        <algorithm type = "BackscatteredTracks"/>
        <algorithm type = "BackscatteredTracks2"/>
        <algorithm type = "ShowerMipMerging3"/>
        <algorithm type = "ShowerMipMerging4"/>
        <algorithm type = "ProximityBasedMerging">
            <algorithm type = "TrackClusterAssociation"/>
        </algorithm>
        <algorithm type = "ConeBasedMerging">
            <algorithm type = "TrackClusterAssociation"/>
        </algorithm>
        <algorithm type = "MipPhotonSeparation">
            <algorithm type = "TrackClusterAssociation"/>
        </algorithm>
        <algorithm type = "SoftClusterMerging">
            <algorithm type = "TrackClusterAssociation"/>
        </algorithm>
        <algorithm type = "IsolatedHitMerging"/>
    </associationAlgorithms>
</algorithm>
<algorithm type = "TrackClusterAssociation" description = "TrackClusterAssociation"></algorithm>
```

for example: called 4 times in reclustering

- This is in addition to several instances of track-cluster distances being computed within individual clustering algorithms (example: GetDistanceToTrackSeed in Cone clustering)

The temporary fix

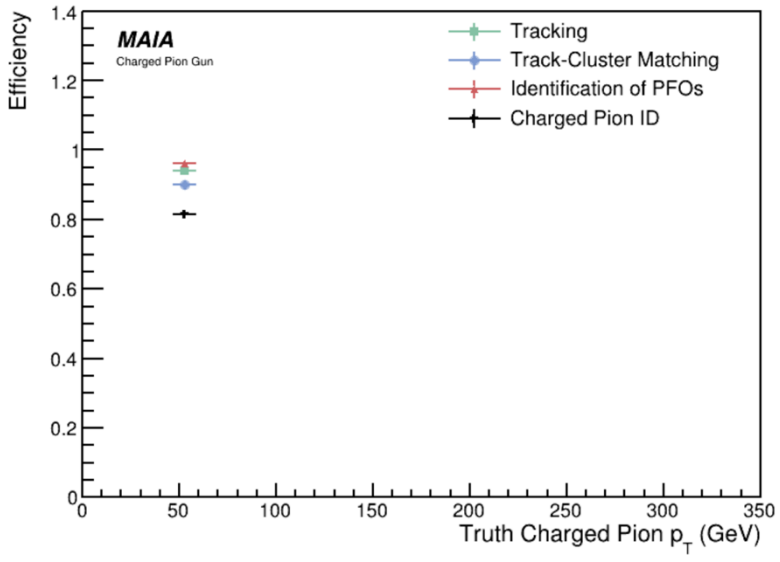
- The paper makes it clear that any final track-cluster association must be made by this track-cluster matching (code)
- Rose has investigated this (link)
  - Found very small gains when fully relaxing cuts
- I also replicated this, assuming I only need to change the parameters right when PFOs are made
- But now I realize that we call it many times, and with my local code I can hard-code the numbers in the C++ file:

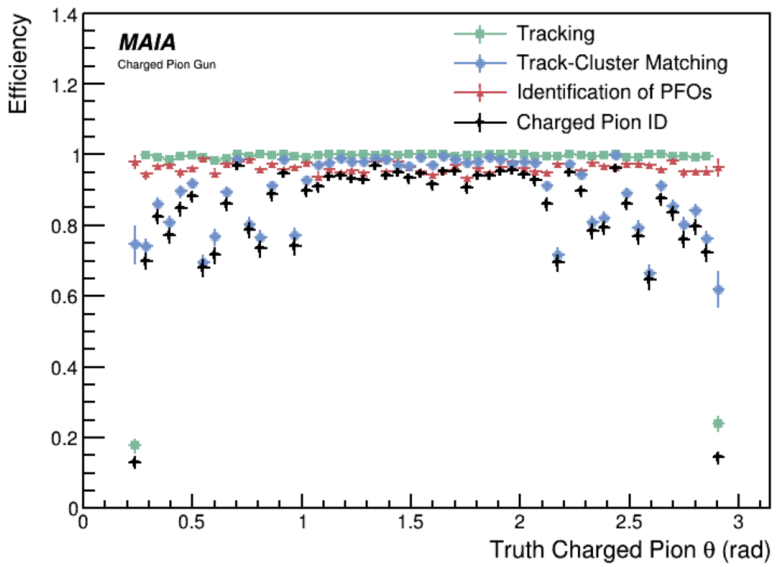
```
43  TrackClusterAssociationAlgorithm::TrackClusterAssociationAlgorithm() :
44      m_lowEnergyCut(0.2f),            cluster energy threshold
45      m_maxTrackClusterDistance(10.f), perpendicular distance
46      m_maxSearchLayer(9),             layers of calorimeter to search for track
47      m_parallelDistanceCut(100.f),    parallel distance, track and cluster
48      m_minTrackClusterCosAngle(0.f)   angle between track and cluster
49  {
50  }
```

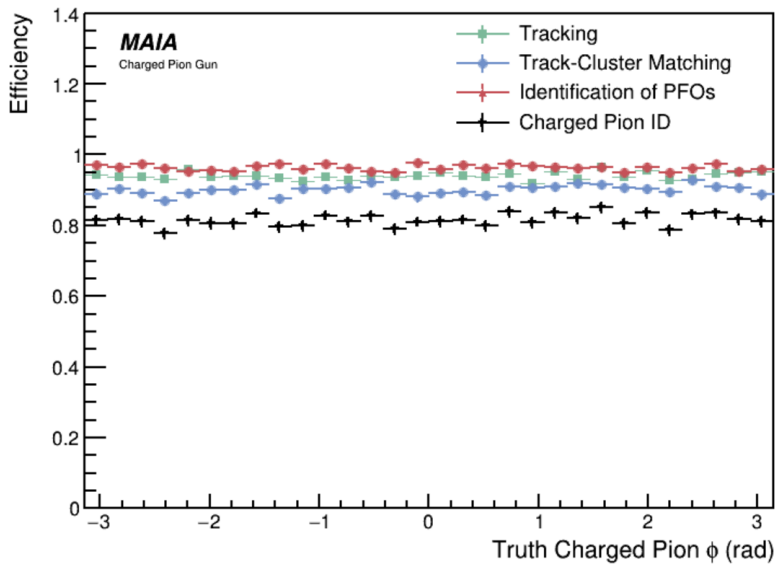
```
43  TrackClusterAssociationAlgorithm::TrackClusterAssociationAlgorithm() :
44      m_lowEnergyCut(0.f),
45      m_maxTrackClusterDistance(100000.f),
46      m_maxSearchLayer(100),
47      m_parallelDistanceCut(1000000.f),
48      m_minTrackClusterCosAngle(0.f)
49  {
```

The temporary fix

- Track-cluster matching efficiency looking much better
- Using SiTrack container:
  - Good tracking efficiency
  - Tracking studies in paper used this
- Charged pion gun, flat in theta and phi, pT = 50 GeV
- A few questions to investigate:
  - What's driving the segmentation in theta?
  - Why not perfect track-cluster matching efficiency?

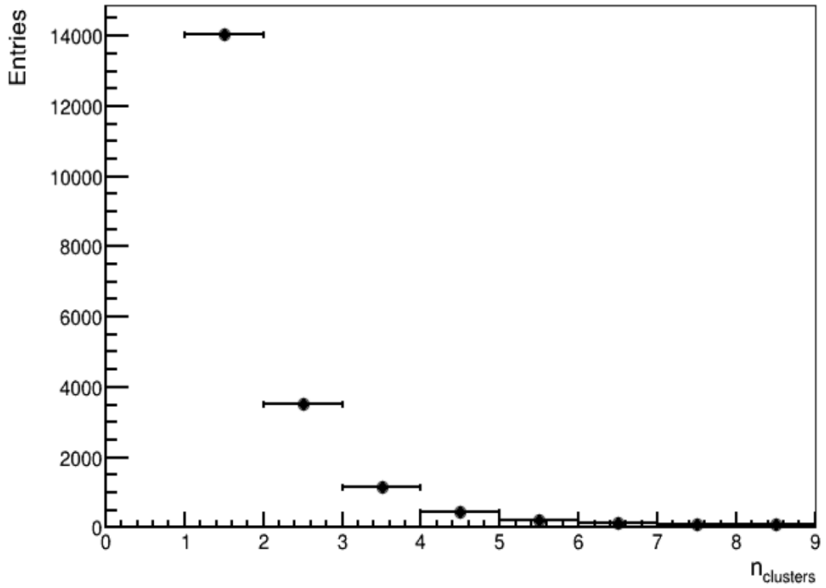


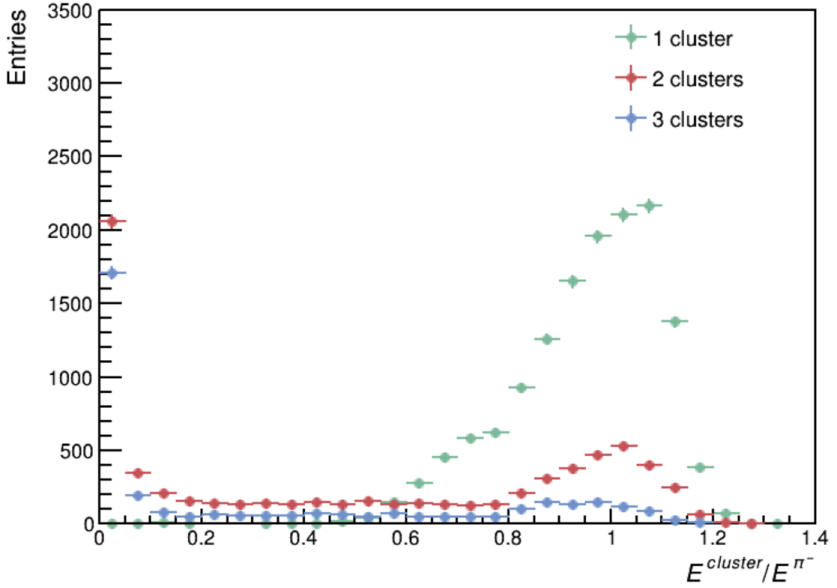




The temporary fix

- Track-cluster matching efficiency looking much better
- Using SiTrack container:
  - Good tracking efficiency
  - Tracking studies in paper used this
- Charged pion gun, flat in theta and phi, pT = 50 GeV
- A few questions to investigate:
  - What's driving the segmentation in theta?
  - Why not perfect track-cluster matching efficiency?
- Clustering also looking much better
  - Confirming that track-cluster matching is called many (24!) times during Pandora





Thoughts and next steps

- Pandora is a complicated algorithm
- The results I showed today are hard to interpret
  - Where exactly did this matter?
- Next steps should be to *simplify* the algorithm to make it interpretable, then expand as needed
- What I have in mind:
  - Remove all but cone clustering and PFO creation in the config
    - Understand the results – plot track-cluster distances when it is only being called in a few places
    - Add BIB?
  - Add sub-algorithms back in one at a time, understanding their impact w/ and w/o BIB individually
    - How does photon clustering impact charged pion reconstruction?
    - How impactful are the 9 “edge case” topological clustering algorithms in a Muon Collider environment?
    - Do we need to recluster to see anything reasonable?
- Very happy to hear other thoughts and opinions…