Cluster Reconstruction Studies [Updated] [Short Version]

Victor Genty

Nevis Labs

June 28, 2013

Outline

1. Clustering in LArSoft (Omitted)
2. Fuzzy Clustering (Omitted)
3. Purity and Efficiency Evaluation
   - Single Electrons
   - Single Muons (Omitted)
   - $1e^- + 1p$ Final states
   - Future (Omitted)

(Omitted)→see longer talk

My Evaluation

1. Generate single electron, muon and uniform flux CC ν_e events with singles.fcl and GENIE. Filter for $1e^- + 1p$ final states
2. Reconstruct clusters with modified uboone offline .fcl script
3. Feed to a module I created to calculate purity and efficiency of reconstructed clusters
4. Compare DBscan, FuzzyCluster

$$\text{Purity} = \frac{\#\text{ of hits from trackID in cluster}}{\text{total }\#\text{ of hits in cluster}}$$

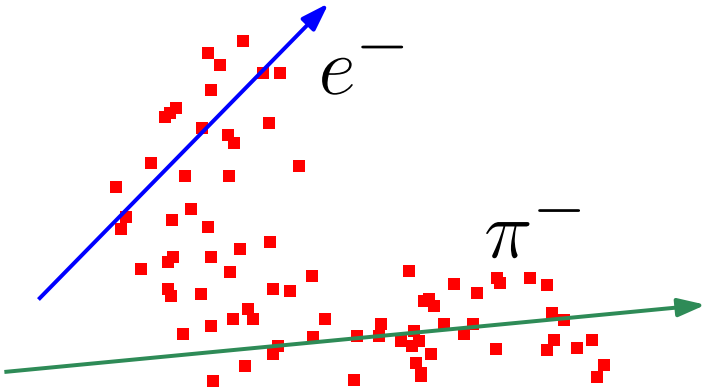
$$\text{Efficiency} = \frac{\#\text{ of hits from trackID in cluster}}{\text{total }\#\text{ of hits for that trackID}}$$

Purity

Formula

$$\text{Purity} = \frac{\#\text{ of hits from trackID in cluster}}{\text{total }\#\text{ of hits in cluster}}$$

Example



Hit Count

Recon: Total = 50

Truth: e^- = 15 $\rightarrow$ Purity=15/50=0.3

Truth: π^- = 35 $\rightarrow$ Purity=35/50=0.7

Measures

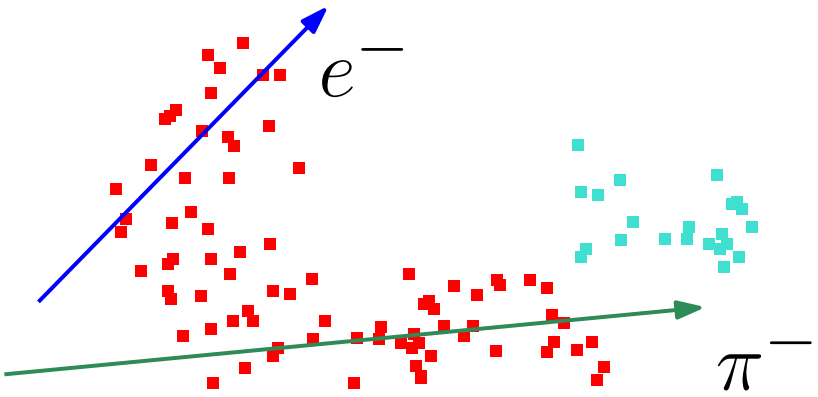
- How much of a cluster is composed of a each true particle
- If less than 1: either true tracks were on top of each other or clustering algorithm failed

Efficiency

Formula

$$\text{Efficiency} = \frac{\#\text{ of hits from trackID in cluster}}{\text{total }\#\text{ of hits for that trackID}}$$

Example



Track Hit Count

Truth: Total $\pi^- = 100$

Truth: Cluster 1 $= 75 \rightarrow$ Eff=0.75

Truth: Cluster 2 $= 25 \rightarrow$ Eff=0.25

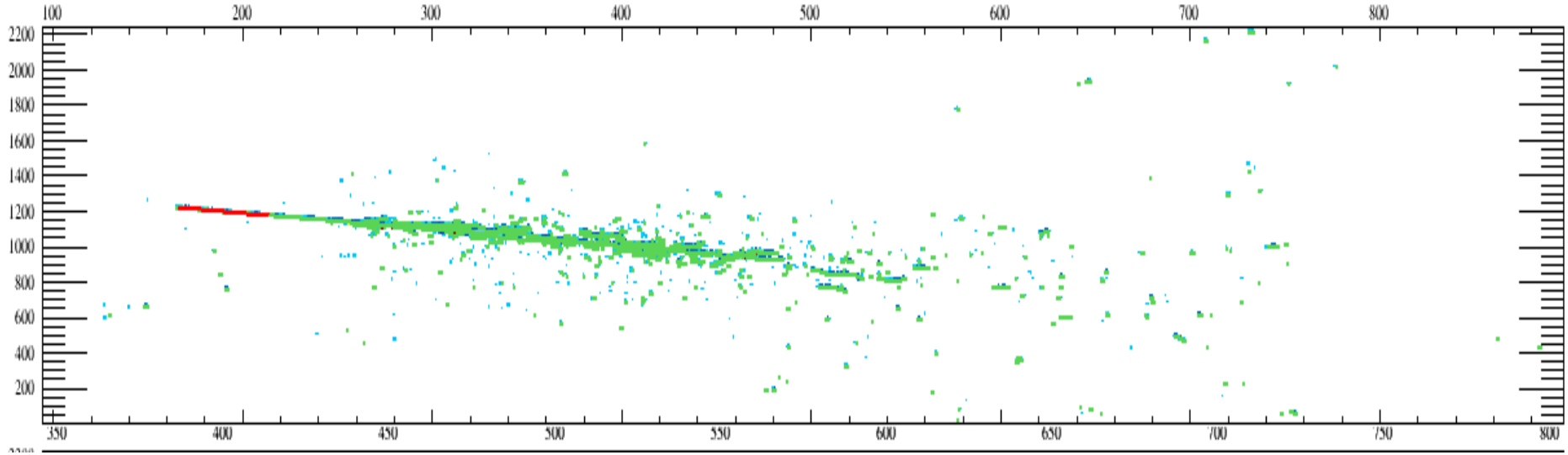
Recon: Grouping of hits

Measures

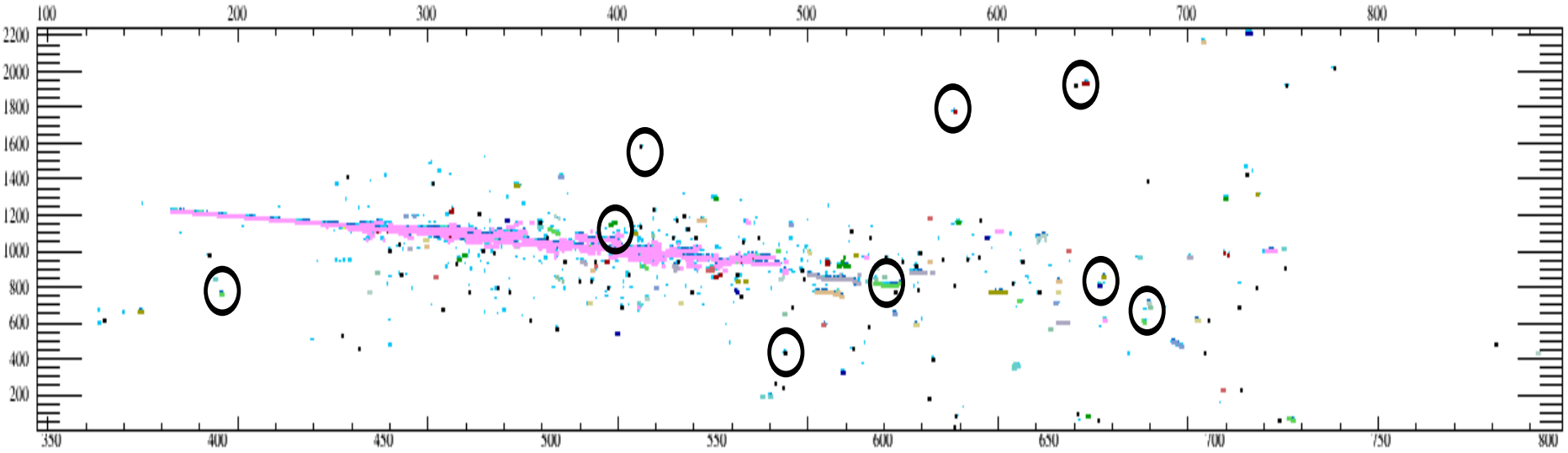
- How many of all hits the particle generated are in a specific cluster
- If less than 1: Algorithm failed to group the hits created by the particle into a single cluster

Event Display - Electron

Fuzzy

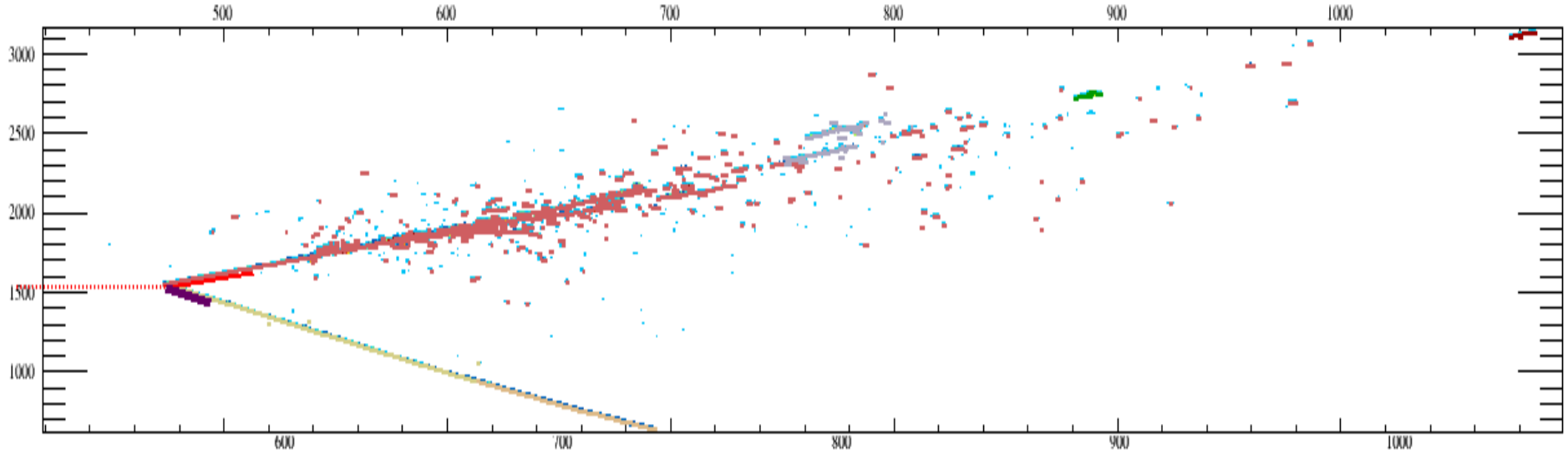


DB

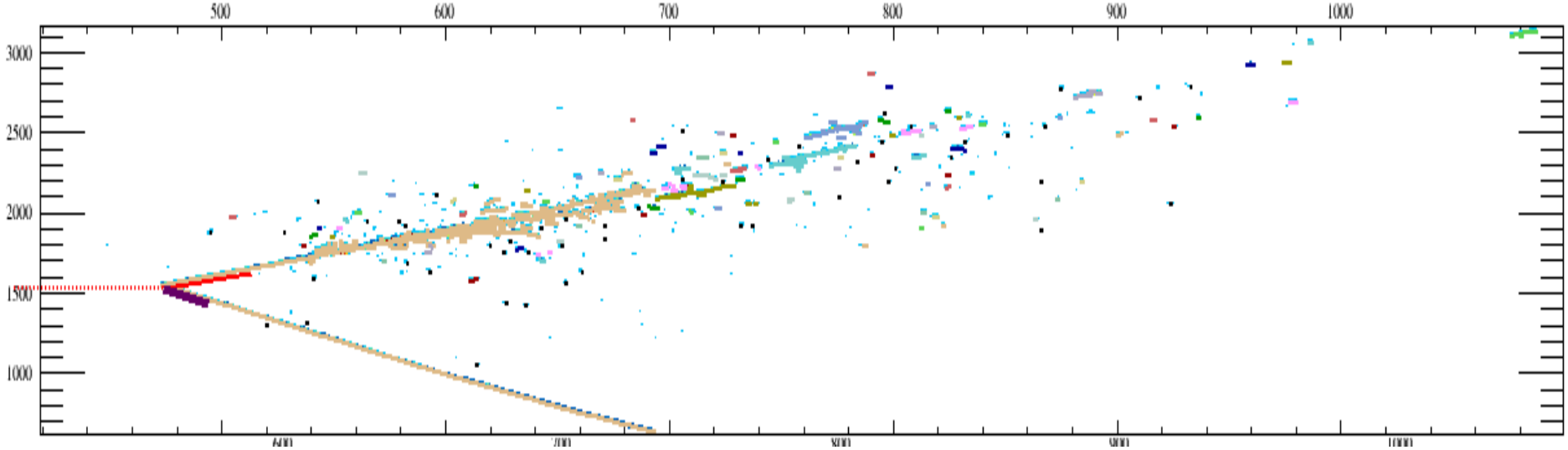


Event Display - $1e^- + 1p$

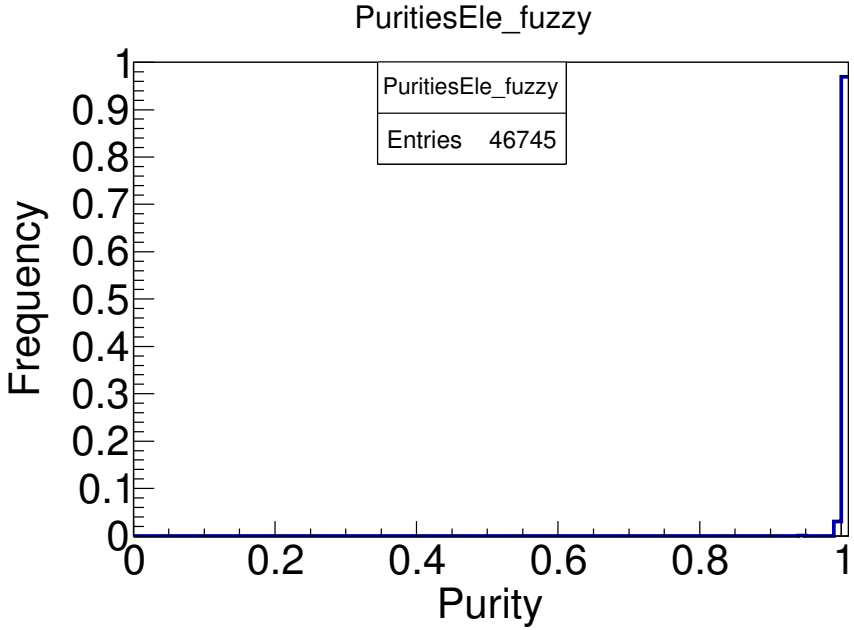
Fuzzy

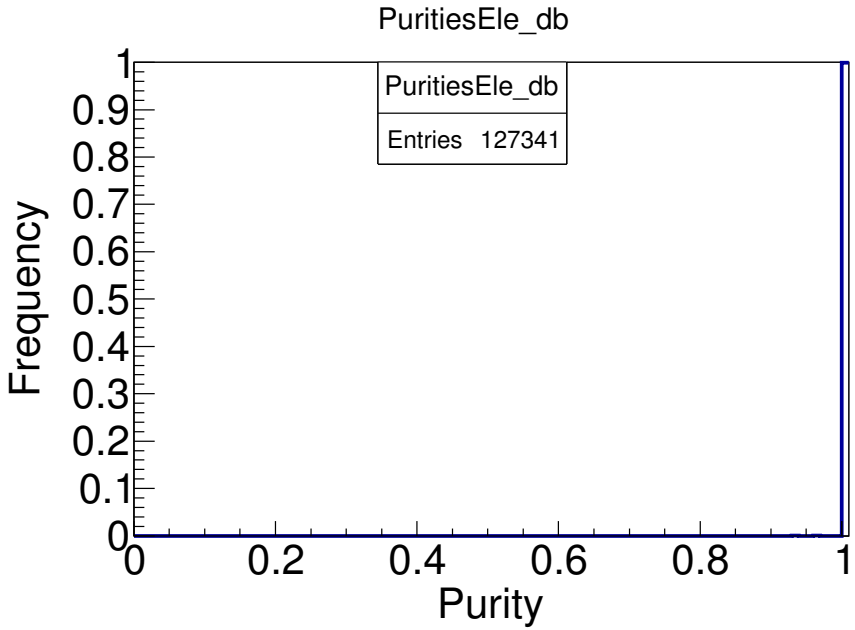


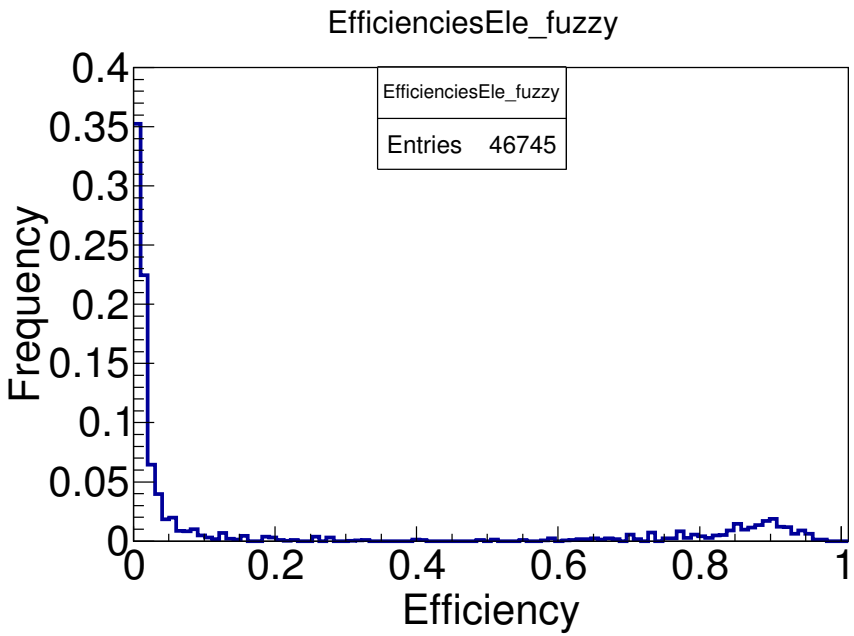
DB

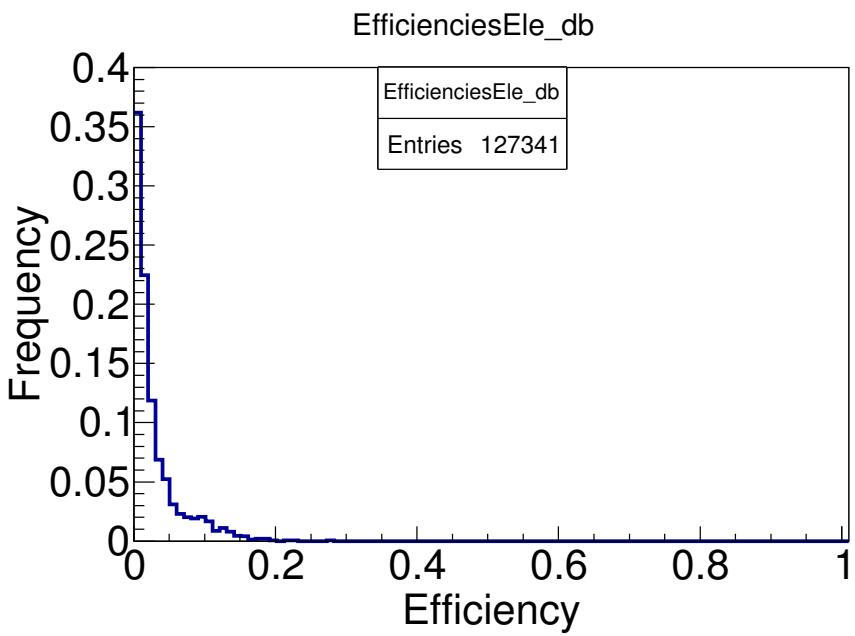


Single Electrons - Fuzzy left, DBscan right

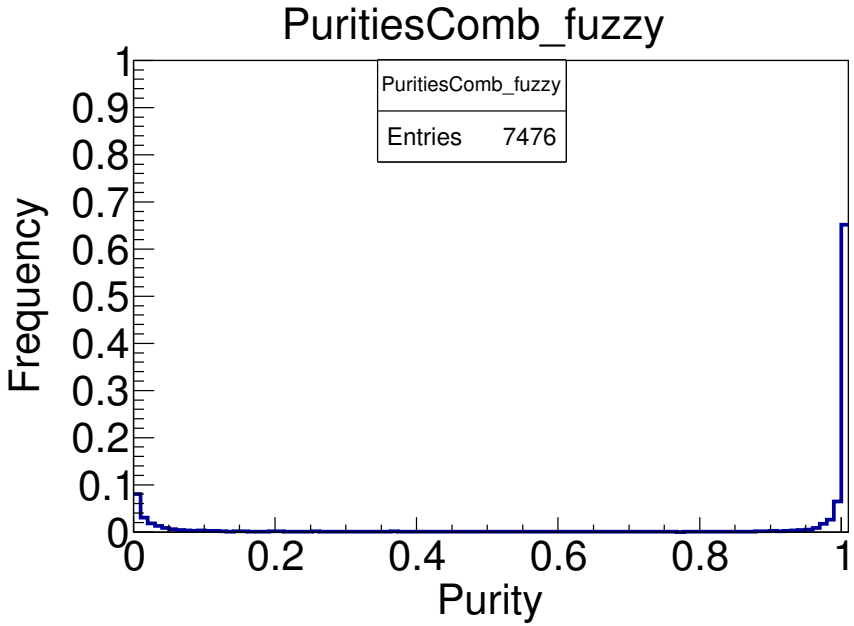


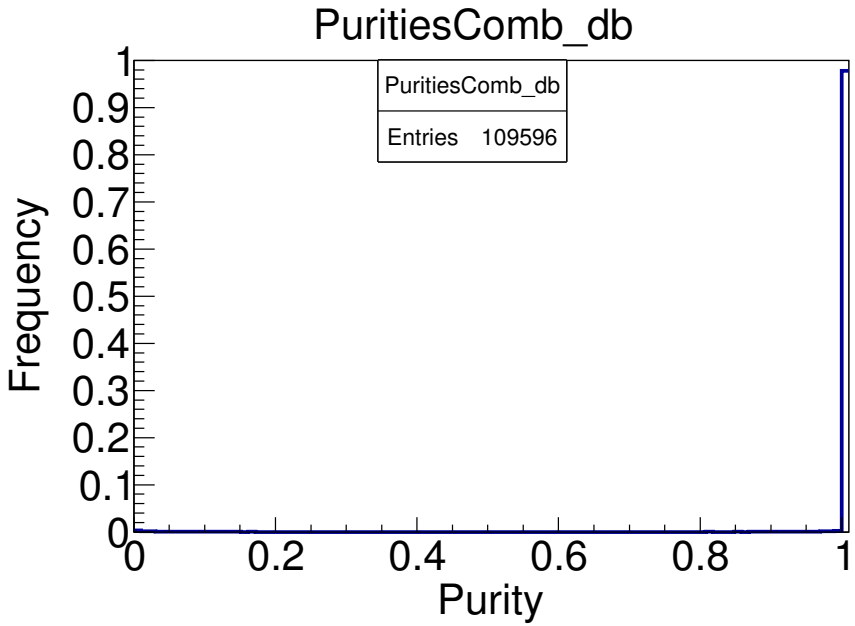


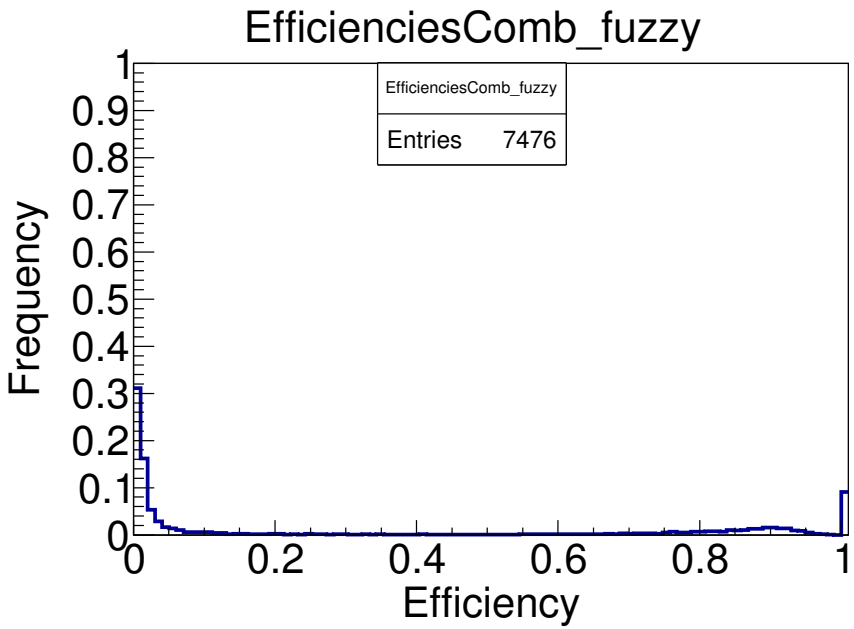


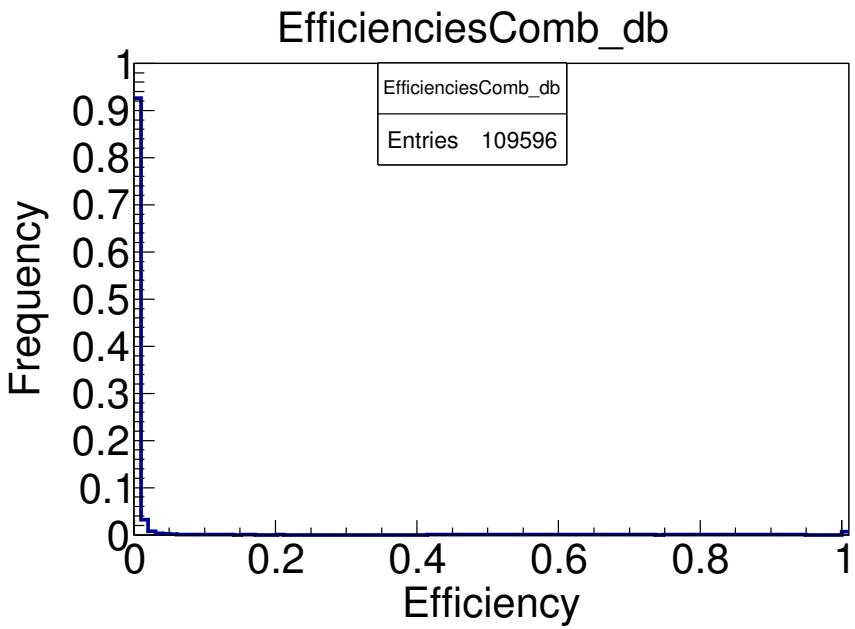


$1e^- + 1p$ - Combined

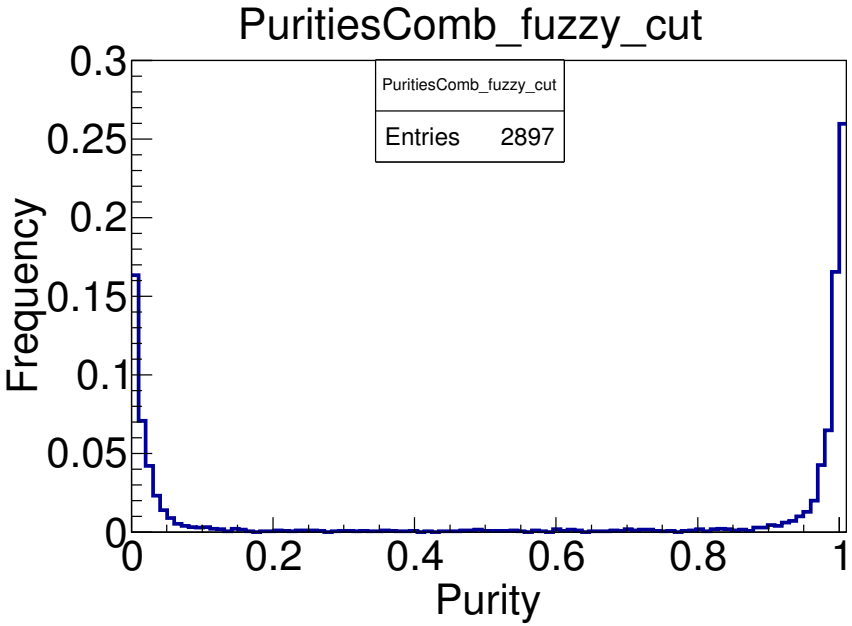


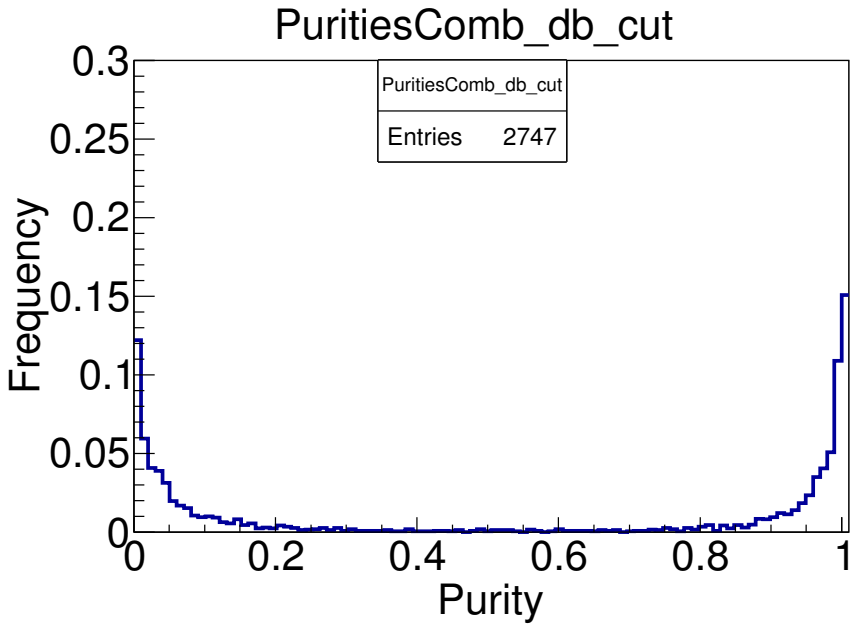


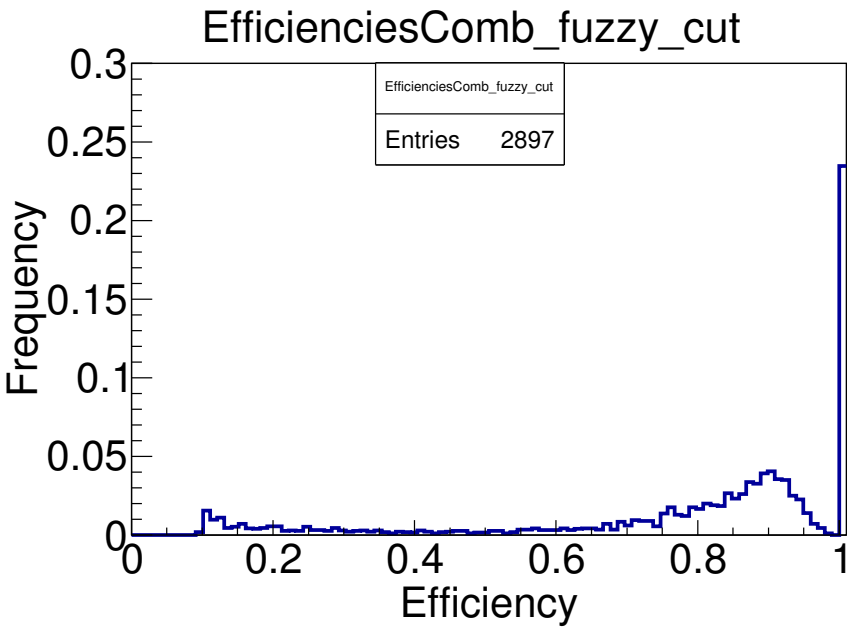


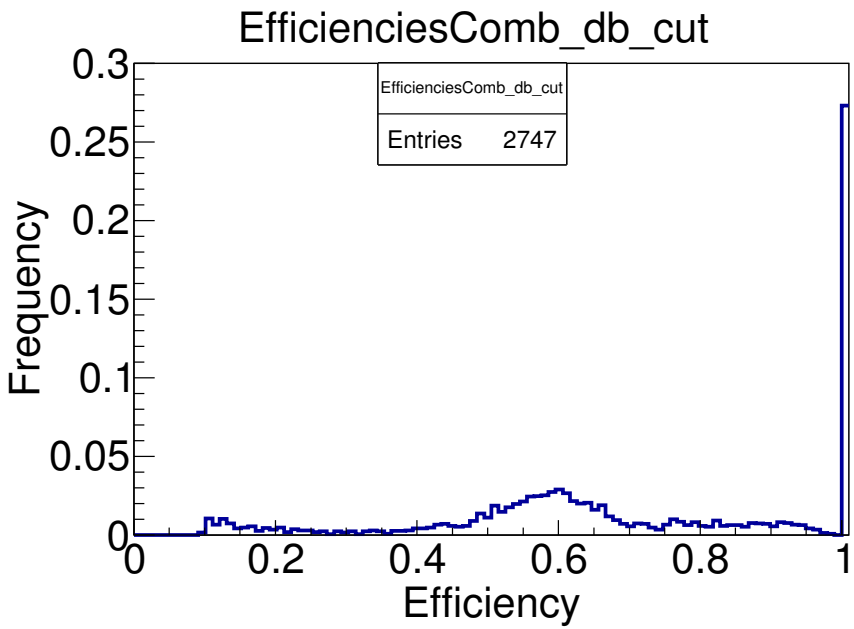


$1e^- + 1p$ - Combined Cut









Conclusions

1. Fuzzy cluster shows improved purity and efficiency values for filtered $1e^- + 1p$ events
2. Clustering also looks improved based on event display output
3. The main issue I've come accross is that the algorithm will fail to cluster proton hits separate from electron showers, especially when proton track is underneith the electron shower
4. Low efficiency clusters with high purity containing 1-5 hits still plague the algorithm