Use of calorimetric profiles for reclustering in Pandora

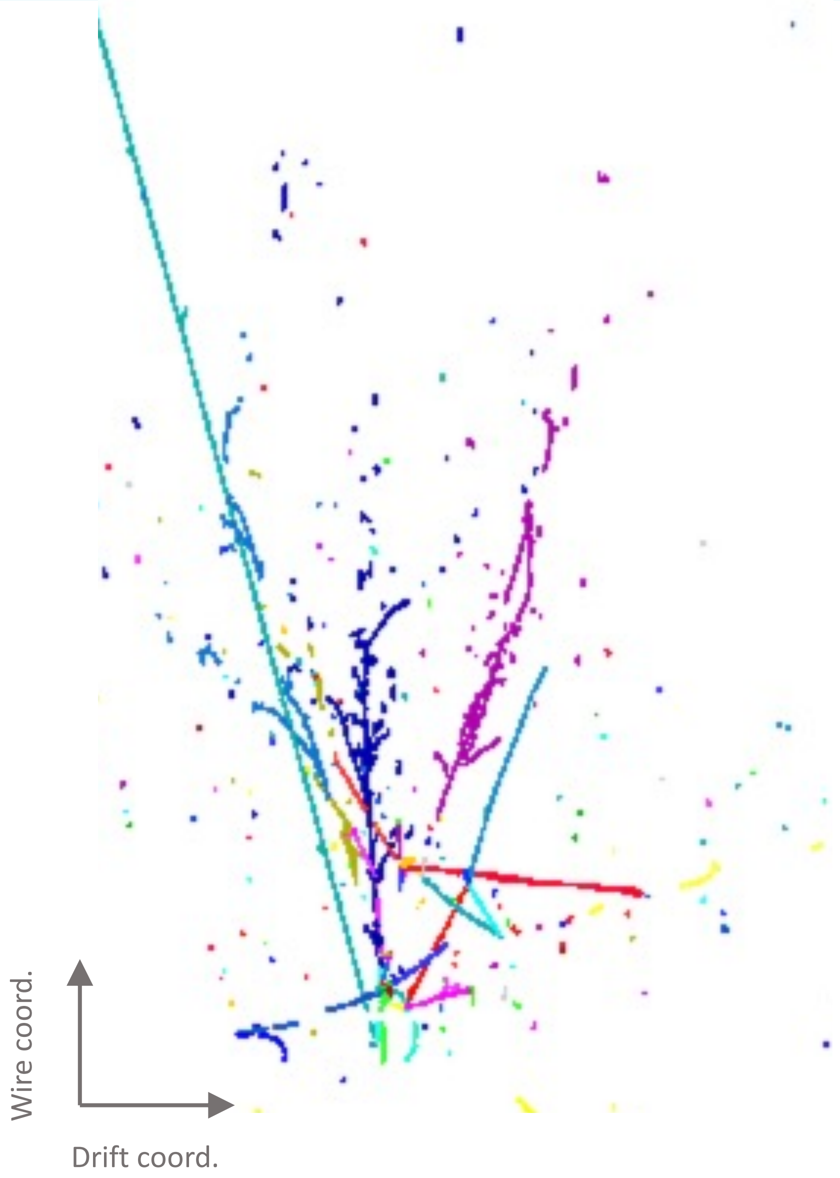
Maria Brigida Brunetti

DUNE UK – Physics and computing – 5 July 2022

WARWICK
THE UNIVERSITY OF WARWICK

LArTPC reconstruction challenges

- High-energy interactions in LArTPCs produce complex images with multiple overlapping tracks and showers
- This creates a reconstruction challenge: disentangle merged tracks and showers
- Merging of two or more showers: a relevant problem affecting many different environments and analyses



6.7 GeV DUNE FD neutrino interaction
cheated pattern recognition

π^0 mass reconstruction

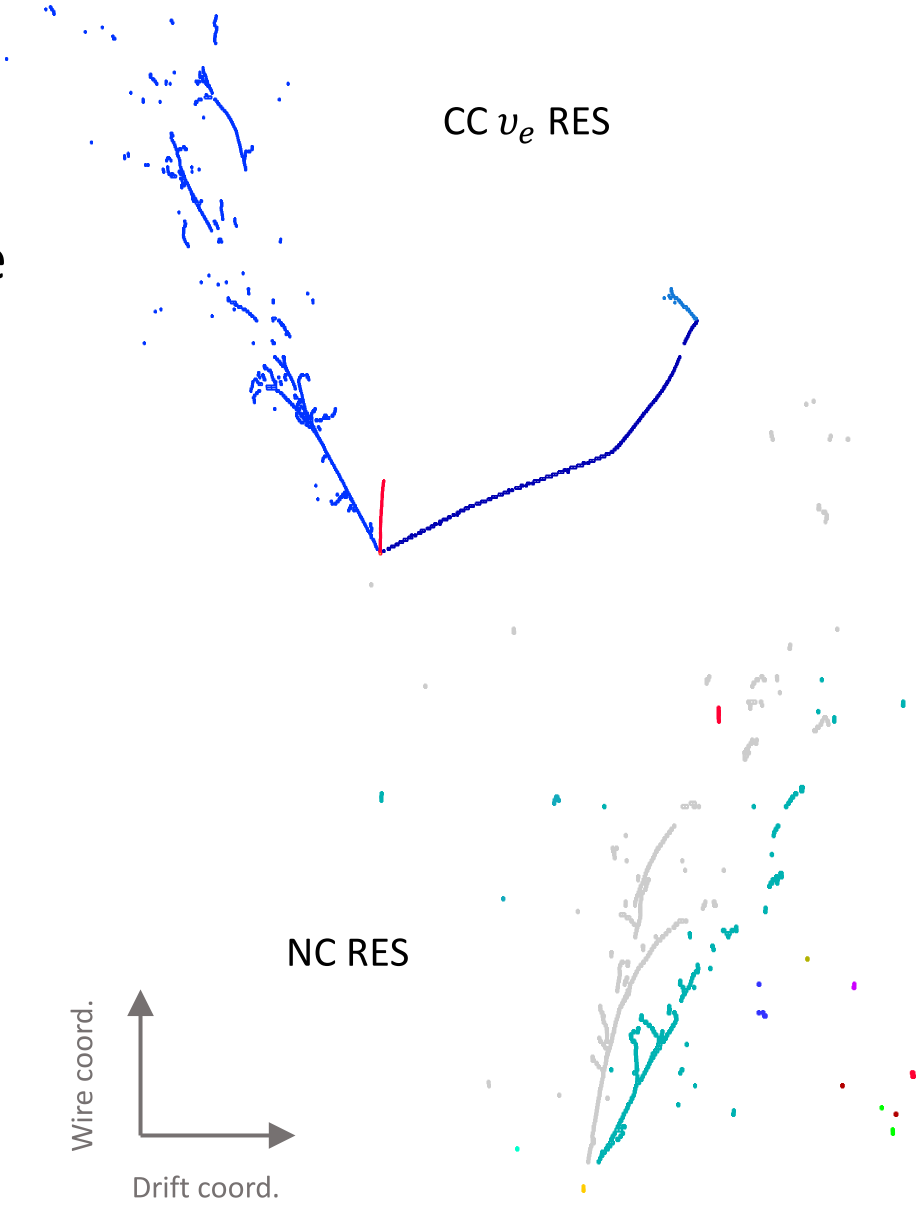
- In υ_e appearance analyses, merging of photon showers from π^0 decays can lead to mis-tagging the shower as electron-like and mis-labelling the event

- Leading causes of performance loss

  https://indico.fnal.gov/event/53402/

- I am working on optimising Pandora using π^0-mass reconstruction-related metrics as a guiding point

  (Neutrino interactions in the 1x2x6 HD FD)



Analysis-driven Pandora optimisation

1) Choose analysis metrics

2) Identify main issues and develop new targeted algorithms
3) Assess metrics improvements

- This approach complements pattern recognition-driven metrics

  MicroBooNE Pandora paper: arXiv:1708.03135v1

  Andy C.'s new patrec metrics: https://indico.fnal.gov/event/50215/contributions/232770/

- The use of analysis-driven metrics as a guide for Pandora optimisation has been successfully demonstrated by I. Mawby in deltaCP sensitivity studies

Splitting merged showers: reclustering

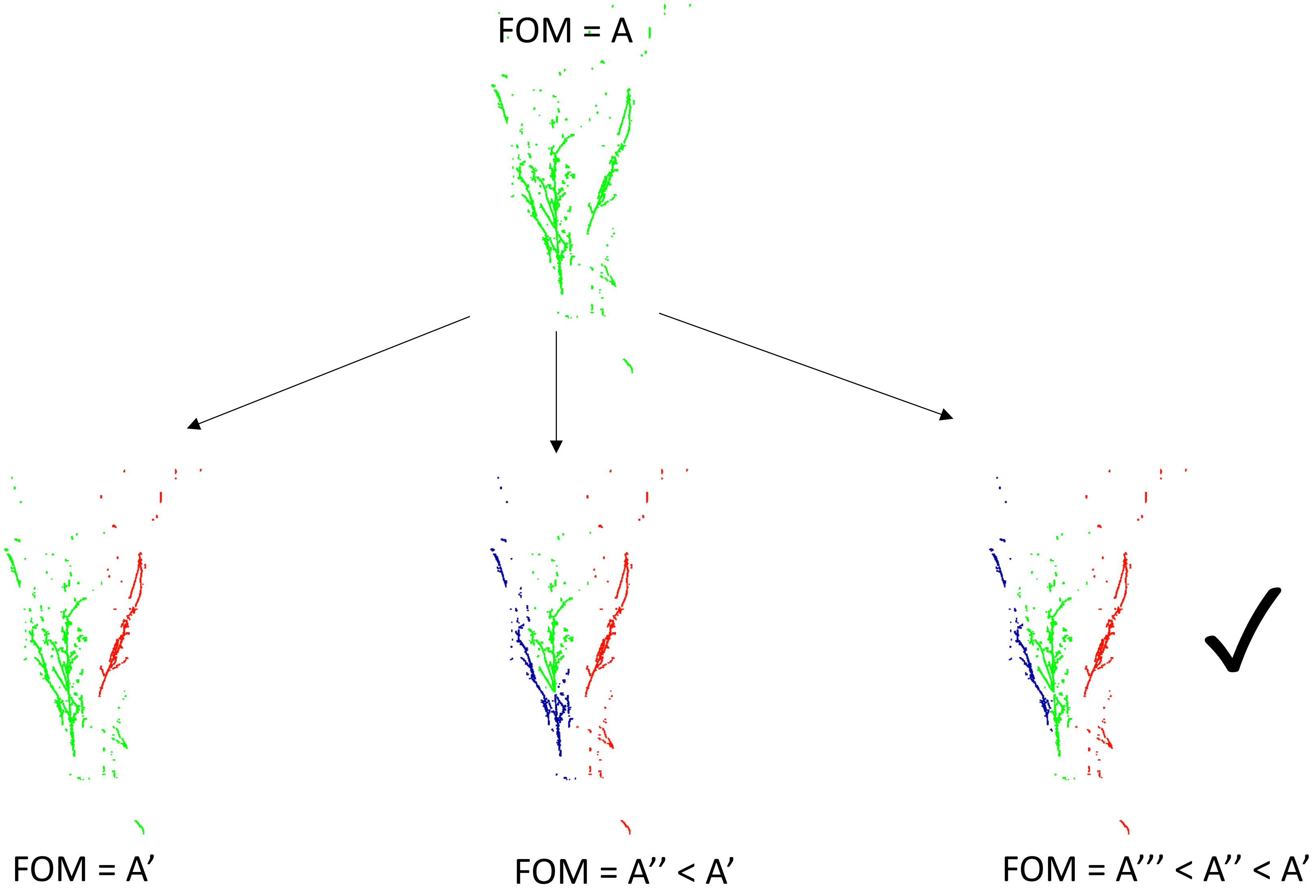
- We can implement **reclustering** in Pandora (main novel design feature!)

  (Following ILC's strategy https://arxiv.org/abs/1506.05348)

1. **Choose figures of merit (FOMs)** to assess whether a shower is merged
2. Start from a list of shower particles, and calculate initial figures of merit
3. Remove the 3D clusters from the parent particle, and **make 3D hits available**
4. Iterate over many encapsulated clustering algorithms. Each will produce a **new list of candidate 3D clusters**
5. Calculate FOMs for all outcomes, and pick best one (or keep original)

→ Integrate with ML/DL approach

Splitting merged showers: reclustering (2)



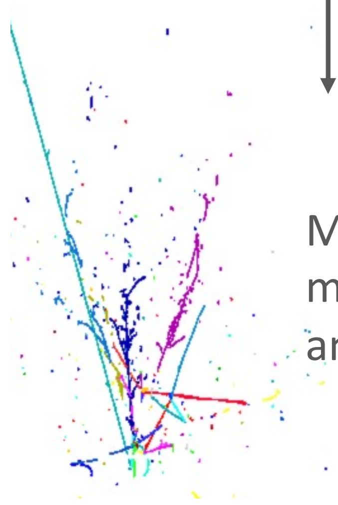
Reclustering example

- The main algorithm runs many encapsulated clustering algorithms
- Each produces a list of clusters and has an associated FOM
- The main algorithm picks cluster outcome with best FOM

Shower energy profiles

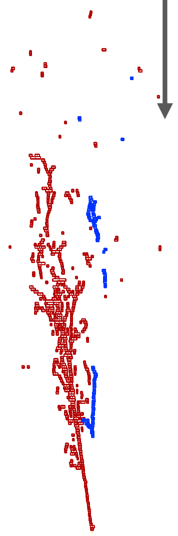
- Use hit calorimetric information to drive reclustering decision via a FOM
- Explore use of shower energy profiles: https://indico.fnal.gov/event/54472/
- Start from **transverse profiles**, look at **longitudinal profiles** at a later stage

(*)



Most useful when merged showers have angular separation

(*)

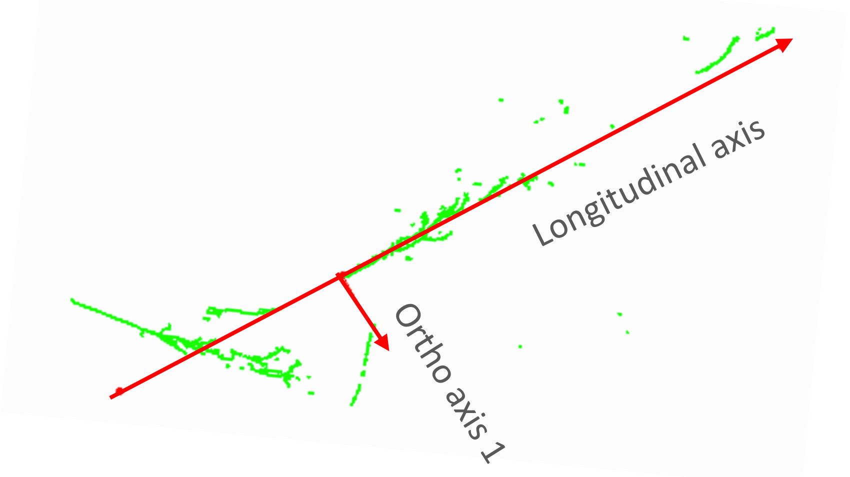


Most useful when two merged showers start at different depths along the principal axis

*cheated patrec

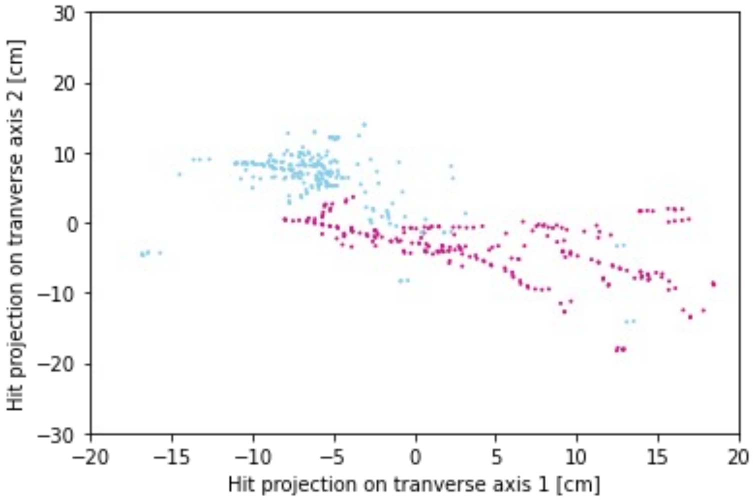
Transverse shower energy profiles

- Find shower principal axis
- Project all hits onto transverse plane
- Produce 2D binned profile summing energy depositions for hits that fall in same bin

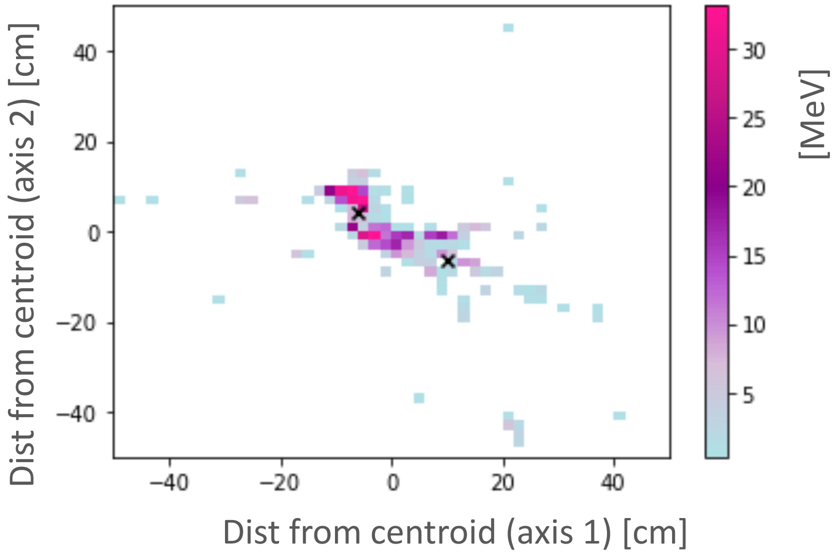


- A possible clustering algorithm uses k-Means clustering to predict cluster centers and distributions for for N = 2 or more underlying clusters → can employ in reclustering

Hit projection on transverse plane
Displaying most contributing MC particle

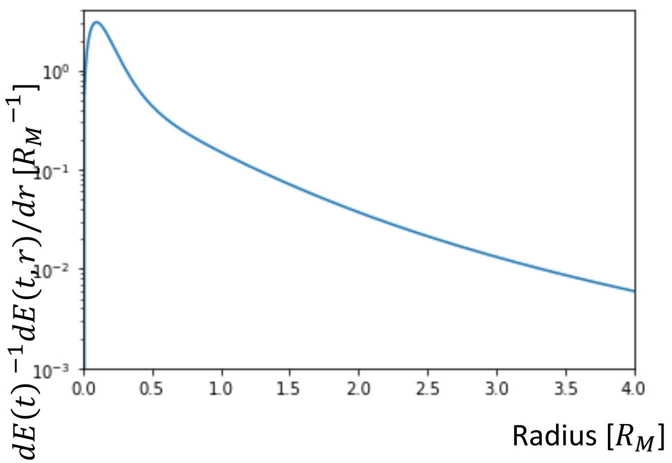


Normalised binned observed energy profiile

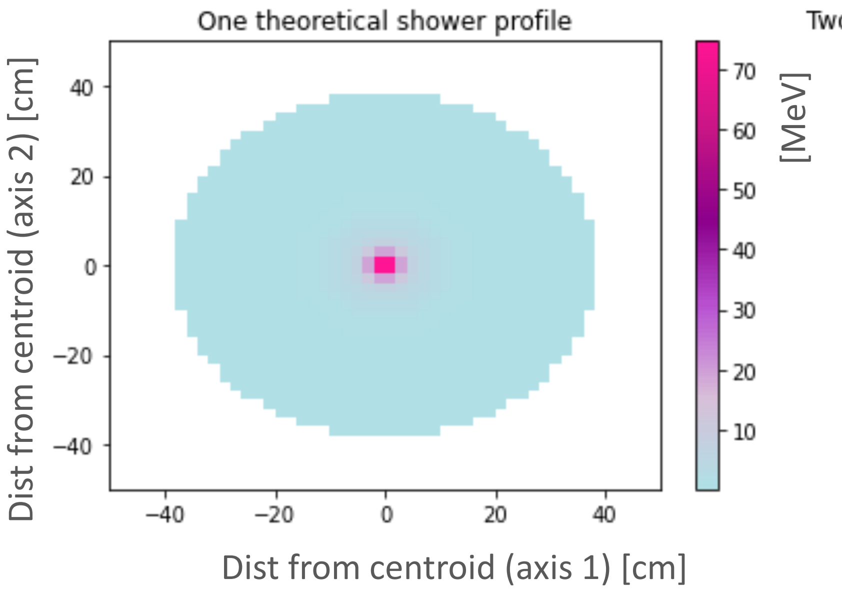


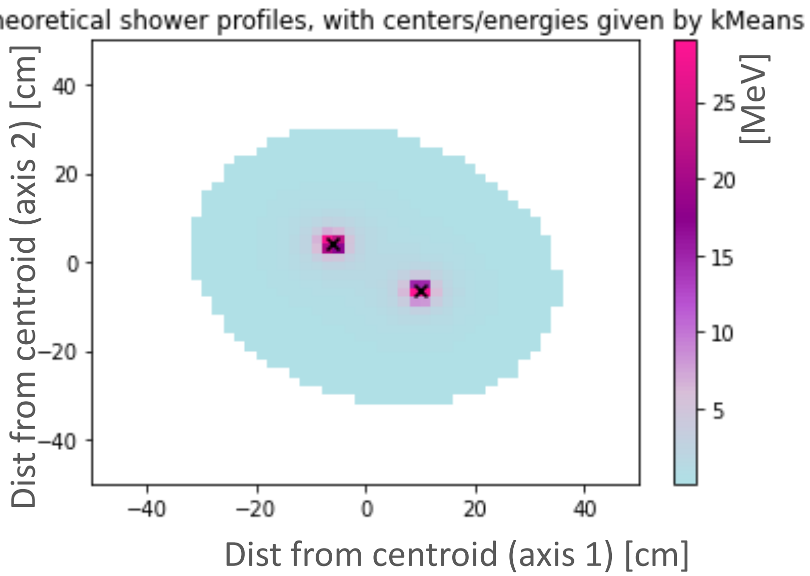
Transverse shower energy profiles (2)

- Transverse shower profile parametrization from arXiv:hep-ex/0001020v1 (Grindhammer)
- Derive 2D binned expected profile
- Can combine many profiles together to create N overlapping showers predictions
- Need external input for center positions and relative energy (e.g. for kMeans)



Example expected radial profile for a single 500 MeV shower (normalized)





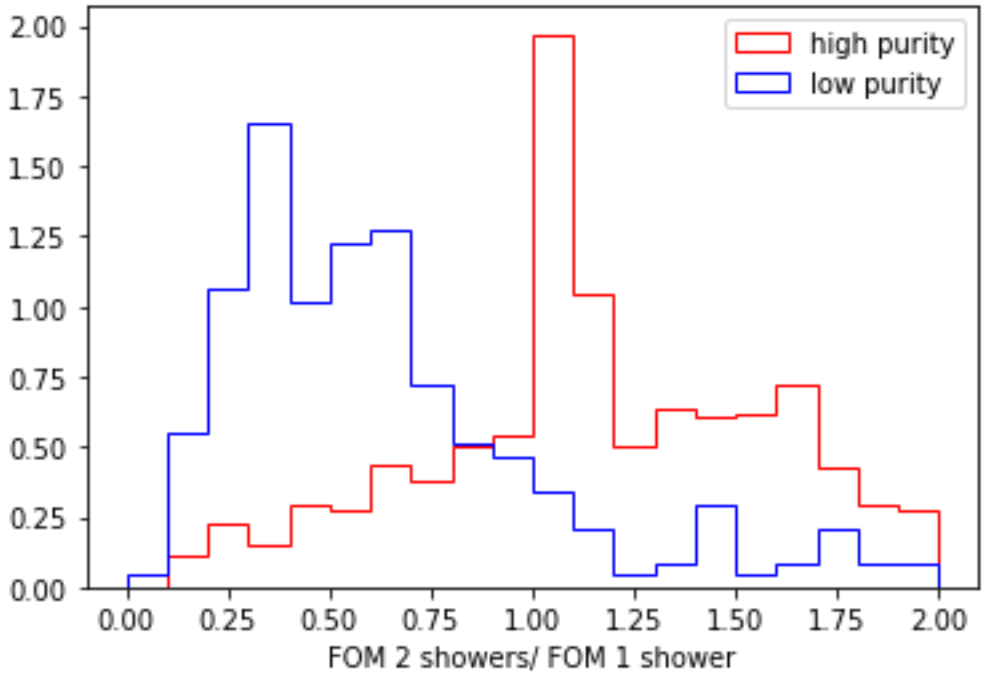
Splitting decision procedure

- Does the observed profile agree better with a 1 shower, or a 2 showers prediction?
- Compare bin-by-bin the observed profile to the predictions for different numbers of showers
- Define a FOM:

$$FOM = \frac{\sum_{Bins} \frac{(N_{OBS} - N_{EXP})^2}{N_{EXP}}}{total\ Energy}$$

- If the FOM is smaller for a 2-shower prediction than for a 1-shower prediction, can decide to split cluster
- Can combine different FOMs, both calorimetric and topological, in multivariate approach

1000 showers sample
Ratio of FOM(2 showers) and FOM(1 shower) shows good separation



Conclusions

- Reclustering strategy implementation in Pandora is under development and proceeding well (foreseen completion on the timescale of a few months)
- Use reclustering to split merged showers, benefitting many analyses
- First test: π^0 invariant mass reconstruction in neutrino interactions at the DUNE FD
- Calorimetric shower profiles are a useful tool in this context
- Can integrate into a multivariate approach, including ML/DL

Transverse profiles - parametrization

- from arXiv:hep-ex/0001020v1

$$f(r) = \frac{1}{dE(t)} \frac{dE(t,r)}{dr},$$

$$\begin{aligned} f(r) &= p f_C(r) + (1-p) f_T(r) \\ &= p \frac{2 r R_C^2}{(r^2 + R_C^2)^2} + (1-p) \frac{2 r R_T^2}{(r^2 + R_T^2)^2} \end{aligned}$$

$$\begin{aligned} R_{C,hom}(\tau) &= z_1 + z_2 \tau \\ R_{T,hom}(\tau) &= k_1 \{\exp(k_3(\tau - k_2)) + \exp(k_4(\tau - k_2))\} \\ p_{hom}(\tau) &= p_1 \exp\left\{\frac{p_2 - \tau}{p_3} - \exp\left(\frac{p_2 - \tau}{p_3}\right)\right\} \end{aligned}$$

with

$$\begin{aligned} z_1 &= 0.0251 + 0.00319 \ln E \\ z_2 &= 0.1162 + -0.000381 Z \\ k_1 &= 0.659 + -0.00309 Z \\ k_2 &= 0.645 \\ k_3 &= -2.59 \\ k_4 &= 0.3585 + 0.0421 \ln E \\ p_1 &= 2.632 + -0.00094 Z \\ p_2 &= 0.401 + 0.00187 Z \\ p_3 &= 1.313 + -0.0686 \ln E \end{aligned}$$

t = longitudinal shower depth in units of radiation length
τ = t/T = shower depth in units of the depth of the shower maximum
r = radial distance from the shower axis in Moliere radius units
E = shower energy in units of critical energy

ADC->MeV conversion factor=0.0075 MeV/ADC
Argon properties
https://pdg.lbl.gov/2014/AtomicNuclearProperties/HTML/liquid_argon.html