Vertex finding with Pandora Deep Learning

Andy Chappell

05/07/2022

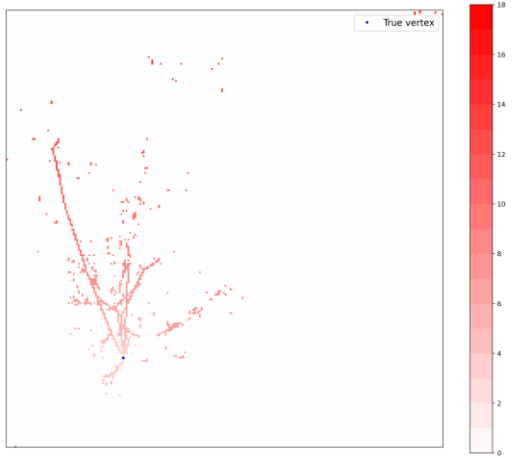
DUNE UK Collaboration Meeting

DEEP UNDERGROUND NEUTRINO EXPERIMENT

WARWICK

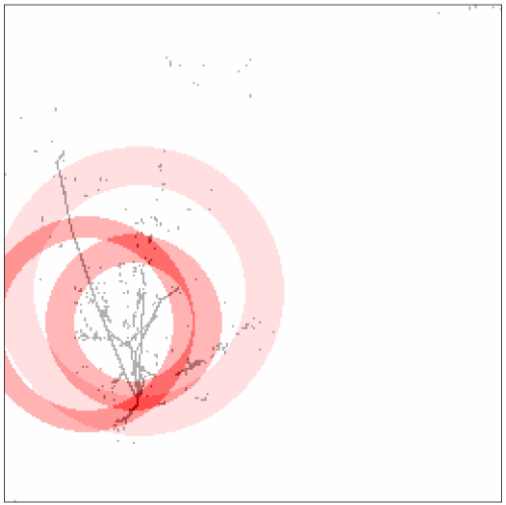
Reminder

- Identifying the neutrino interaction vertex seems like something a CNN should be good at
- But defining a single point/region as the signal means you need a huge number of events to train the network



True pixel classification

- Instead encode the truth in all hits in each event, by describing the distance each hit is from the interaction vertex
- Gives the network a direct handle on how all of the information in a single event relates back to the vertex location
- Post-process the distance classification to resolve the interaction vertex



Heat map generated from 3 classified pixels

Evolving the network

WARWICK

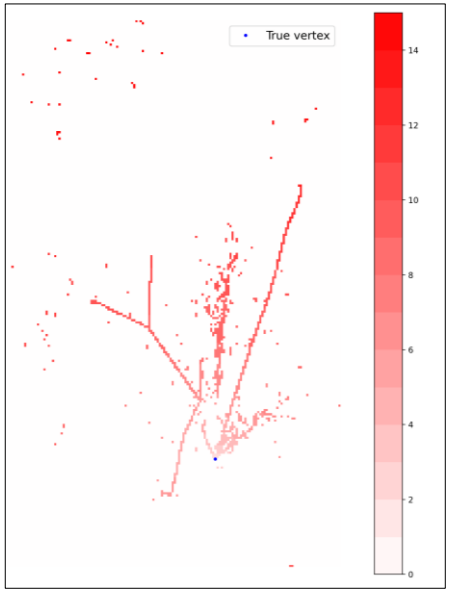
- Preliminary results were quite reasonable
  - 68% of events yielded a reconstructed vertex within ~2.6 cm of the true vertex
  - Compared to ~2.1 cm for the standard BDT-based reconstruction
  - (No fiducial volume cuts applied)
- Problem: Resolution
  - Events are sampled to a 256 x 256 pixel image per view
  - If the event spans more than a couple of metres the pixels begin to represent large regions, limiting the network resolution
- Solution: Add a second pass
  - To address this problem, take the result of the first pass and zoom in to this region to identify the vertex at higher precision

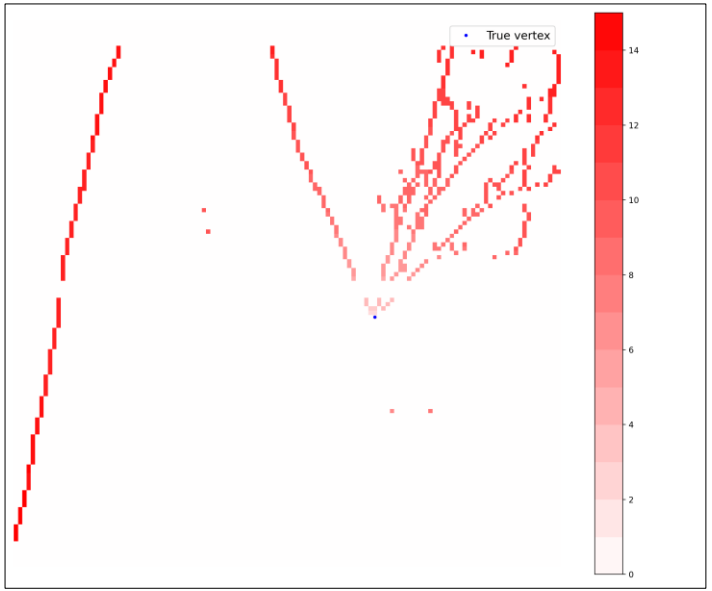
Training the second pass

WARWICK

- We want this to be quick, so we compromise
  - Smaller images at 128 x 128 pixels
  - 64 cm x 64 cm event region, allowing for 0.5 cm resolution
- If the first pass vertex is off by much more than 32 cm, pass 2 probably won't help
  - Will comment on potential mitigation later
- To define the training dataset I take a perturbed version of the true vertex
  - Gaussian (0 cm, 15 cm) perturbation in X and Z
  - Treat this as the centre of the image
  - Our first pass reconstruction will be imperfect and we want to ensure the network doesn't simply learn to pick the centre of the image in the second pass
  - A check is applied to ensure that the region contains hits
  - No check on true vertex containment because this technique can, in principle, find uncontained vertices, so we should let it try

Training the second pass

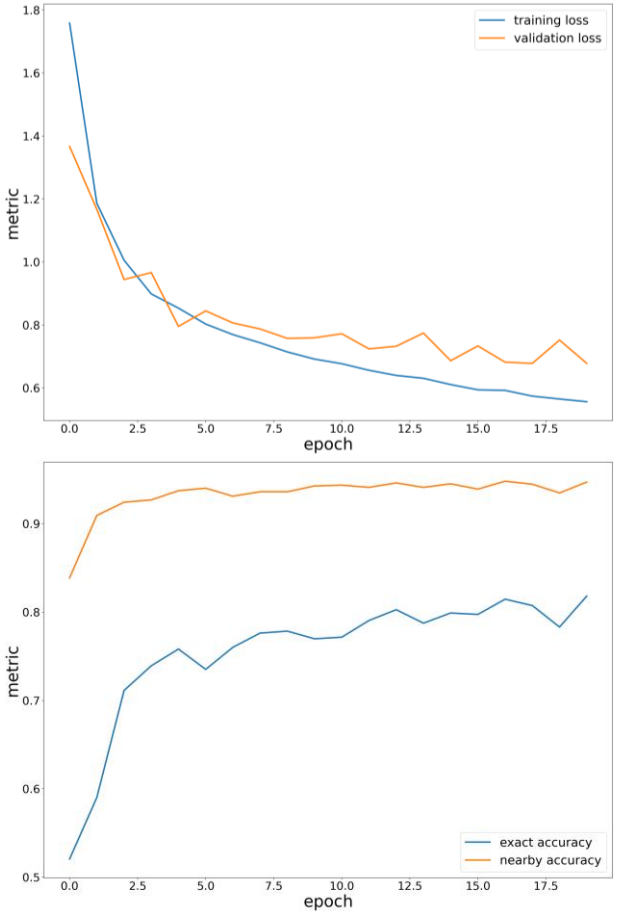




- Left: Pass 1, full event (cropped here), event aspect ratio not (necessarily) respected
- Right: Pass 2, 64x64 cm region of interest, note non-central true vertex

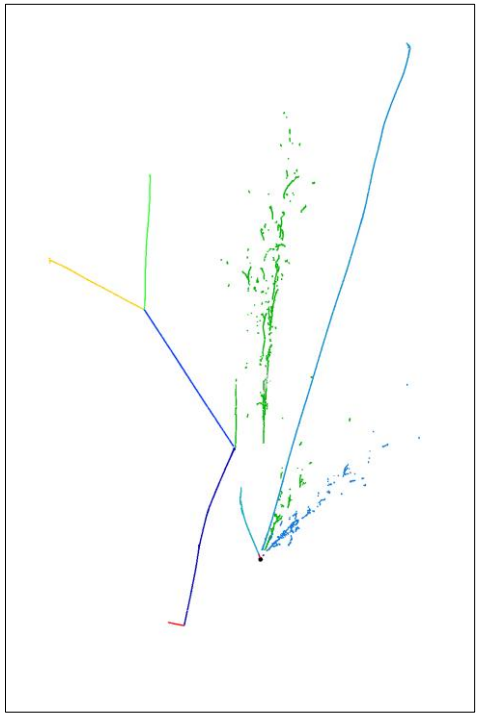
Training the second pass

- Second pass network appears to train well
- W view indicates there may be scope for further improvement with tweaks to the training procedure
- Given 19 distance classes to choose from, classification accuracy is quite good
  - ~80% of pixels correctly classified
  - ~94% of pixels either correctly classified or placed in one of the classes adjacent to the correct class



Example NC DIS Event Reconstruction

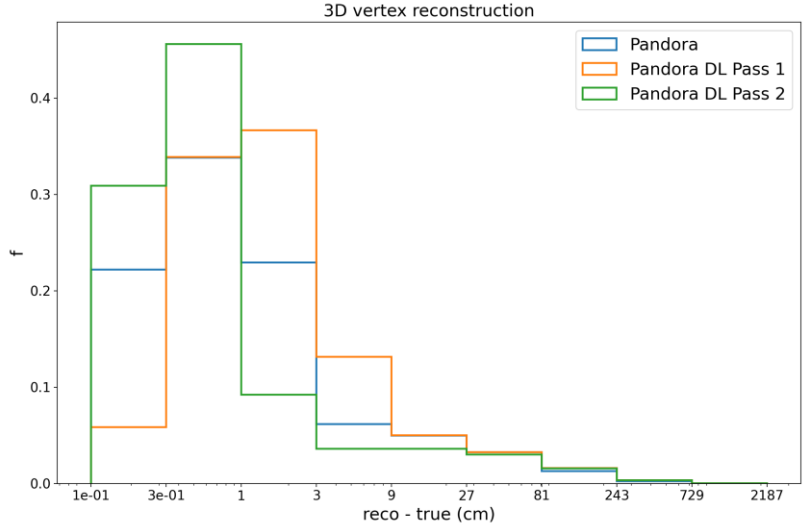
WARWICK

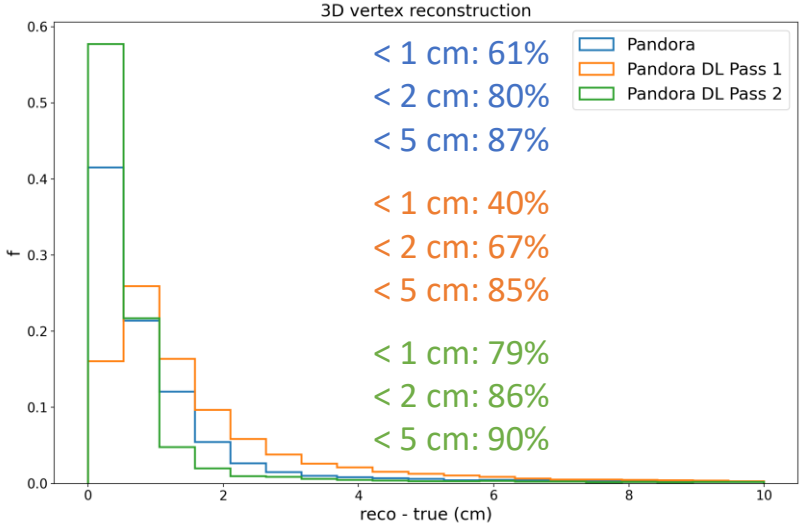


- 3D hits projected into the W view
- A well reconstructed vertex yields good reconstruction of final state particles

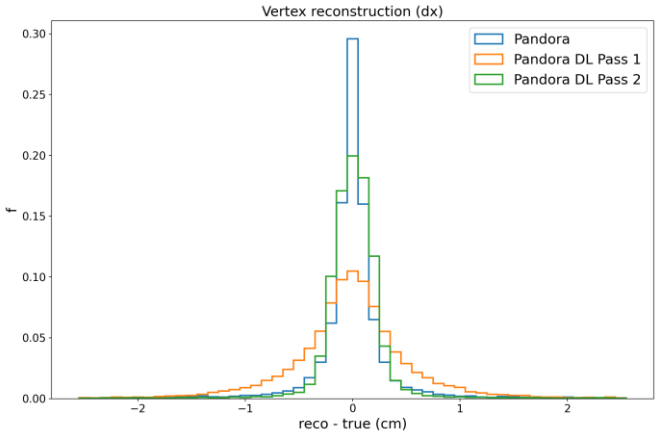
Reco – True Vertex Deltas

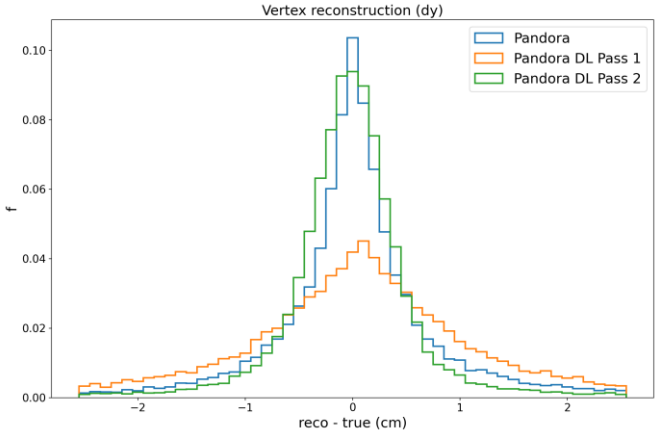
- Compared Pandora, and Pandora's two DL passes on a 50,000 event sample
- Even split ν_μ and ν_e MCC 11 1x2x6
- True vertex must be within the fiducial volume:
  x: (-310, 310) cm; y: (-550, 550) cm; z: (50, 1244) cm;
- Unsurprisingly, second pass improvement limited when first pass error is large

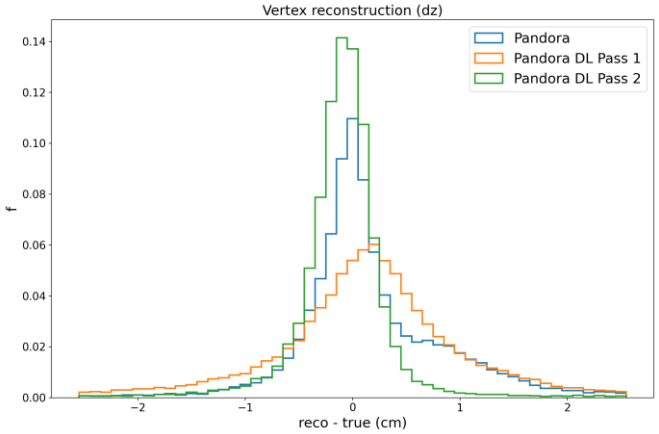




Reco – True Vertex Deltas



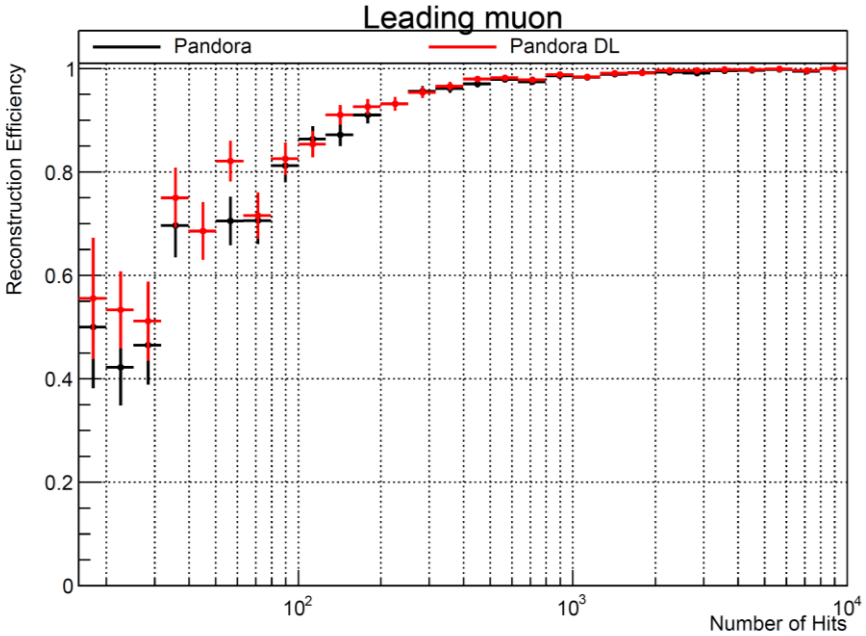


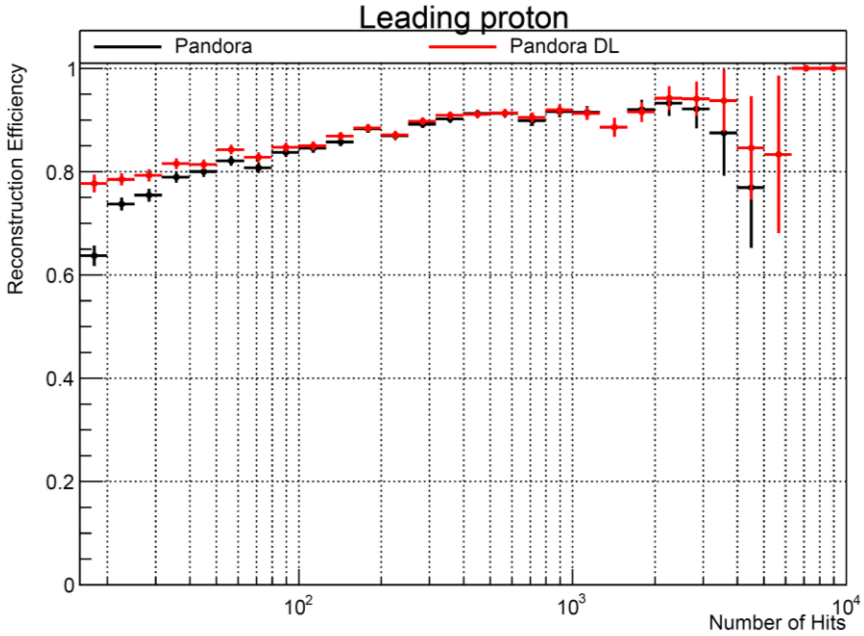


- DL vertexing shows less bias and less population of tails
  - Peaks at zero in X and Y
  - Peaks at -0.1 cm in Z
- Standard Pandora has a sharper peak in x
  - Suspect this could be driven by lower network resolution in x

Leading particle reconstruction efficiency

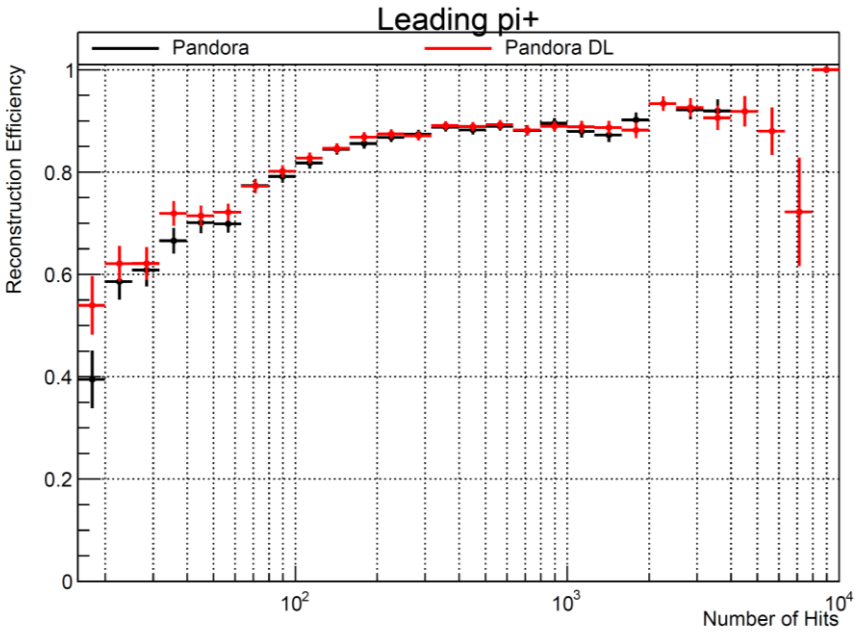
- Fraction of true leading particles with a matching PFP
- No cuts applied

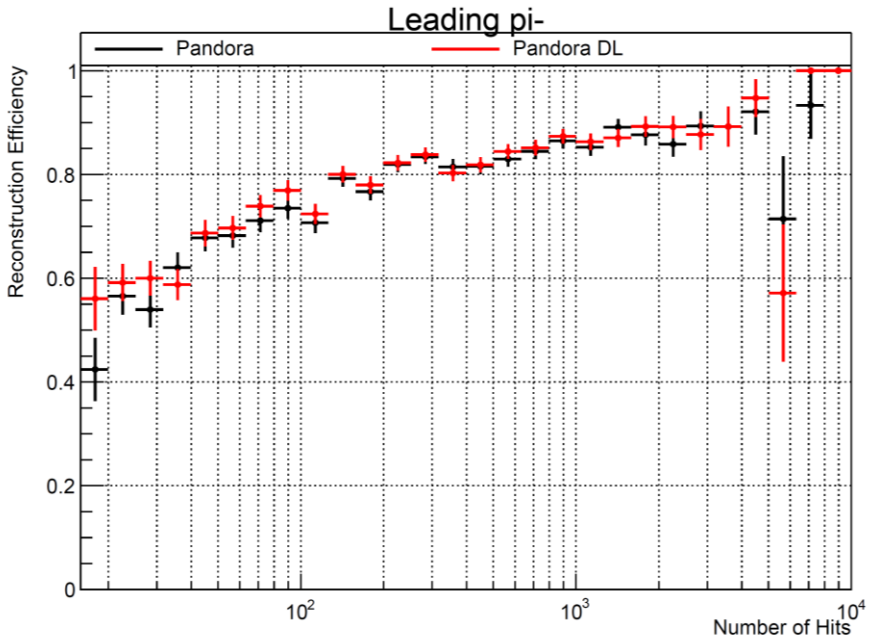




Leading particle reconstruction efficiency

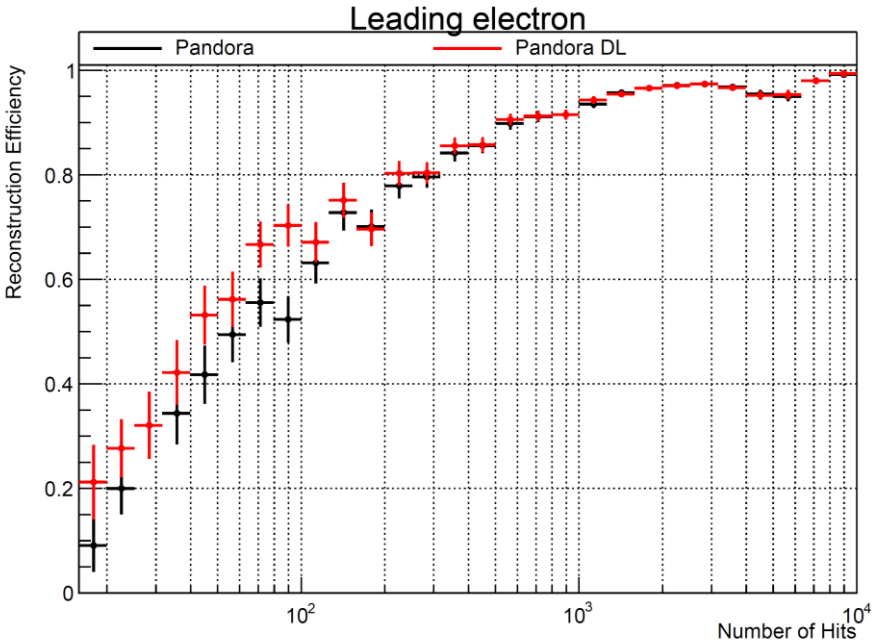
- Fraction of true leading particles with a matching PFP
- No cuts applied

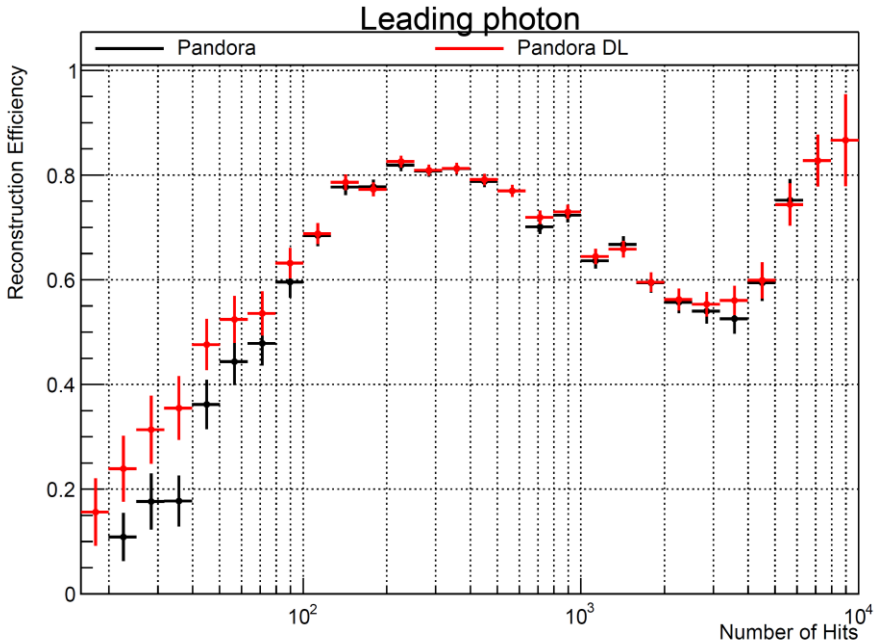




Leading particle reconstruction efficiency

- Fraction of true leading particles with a matching PFP
- No cuts applied





Next steps for vertexing

WARWICK

- Pull request in preparation
  - Aim to release initial version next week
  - Further refinement in future releases
- Dealing with large pass 1 errors
  - Picking entirely the wrong region in pass 1 means pass 2 isn't too helpful
  - Identify a few candidates in pass 1 and zoom in on each
- Multi-view/pass refinement
  - Check consistency of positioning between passes
  - Can use detector geometry information to correlate heat maps and refine 3D position
- Vertex finding in atmospheric samples in testing
- Vertex finding in Vertical Drift geometry
- Secondary vertexing
  - It may be possible to extend this method to find secondary vertices as well