Pandora for ND-LAr and 2x2

Steve Dennis
2x2 Analysis Workshop
20/01/2023

Introduction

- Today I'll be talking about a few things.

- A quick reminder of what Pandora is, and where it sits in our software chain.

- Current status of developments for ND-LAr.

- Future plans for both ND-LAr and the 2x2.

- In the next session, Leigh Whitehead will deliver an update on the current status of the 2x2 Simulation Challenge files.

Pandora

- Pandora is a general pattern recognition software framework.

  - Provides reconstruction for many of your favourite LArTPCs.

    - MicroBooNE

      https://github.com/PandoraPFA

    - ProtoDUNE

    - DUNE FD (HD and VD)

- Heavily developed and optimised for 3x2D readout TPCs

  - With low multiplicity.

- Now being developed for ND LAr.

 - With, of course, a specific focus on getting ready for 2x2 running.

UNIVERSITY OF CAMBRIDGE

Pandora in the Chain
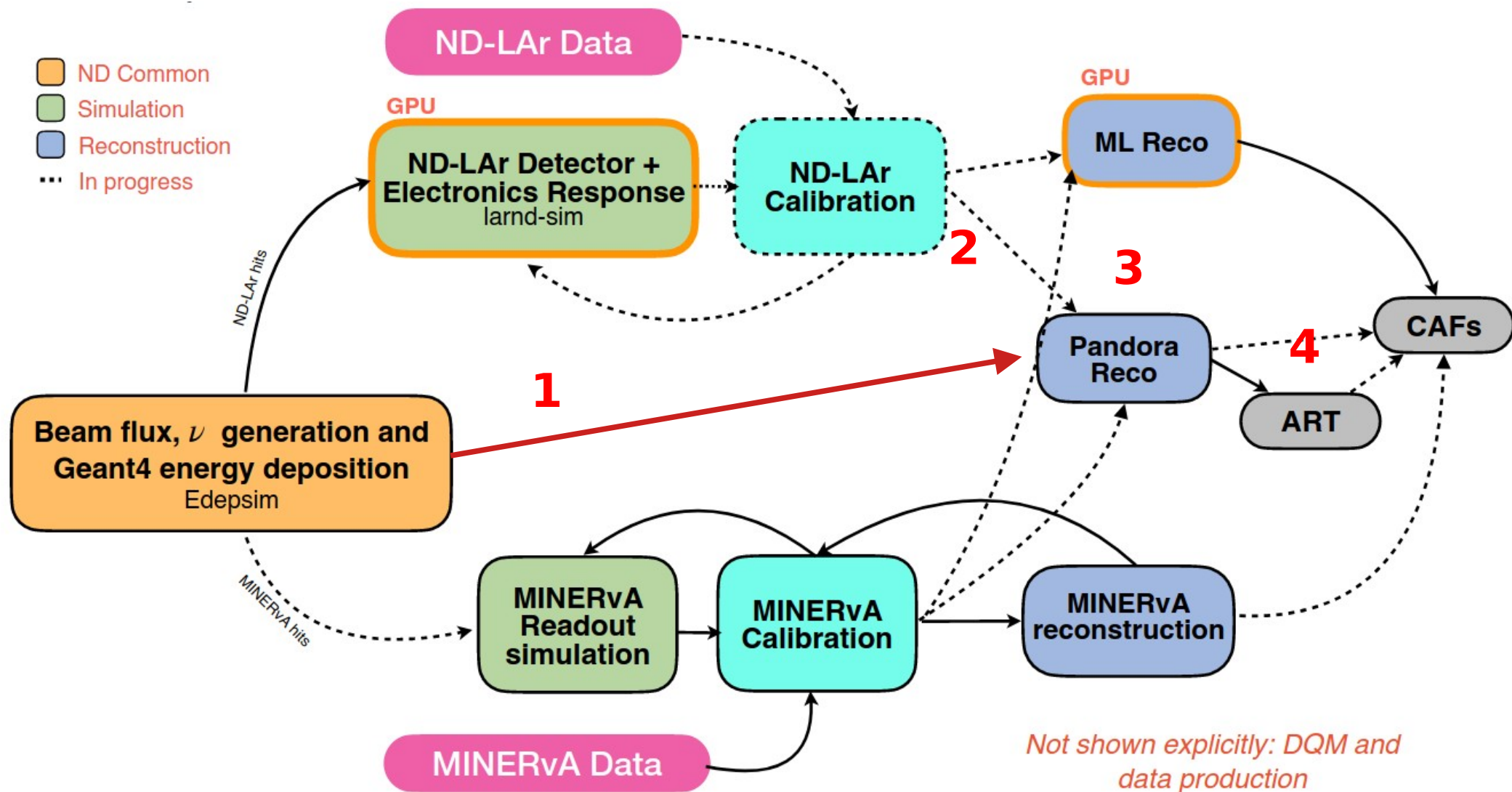
EdepSim/SED, very Briefly

- As you know, the simple way to get events from MC to reconstructable hits is using simulated energy depositions.

    - Eg EdepSim and LArSED (Simulated Energy Deposition, in LArSoft)

- Following our standard approach with LAr experiments, we offer two slightly different approaches to do this.

    - Going via LArSoft configuration/tools.

    - Running natively, independent of LArSoft.

- If you like using the LArSoft interfaces to everything, you can configure/run your generation and reconstruction using FHiCL files within LArSoft and output to ART.

    - Or you can just... not do that.

    - For example, when developing for the 2x2 I operate on my laptop that has never contained a trace of LArSoft.

Example of SimEnergyDeposit

SimEnergyDeposit
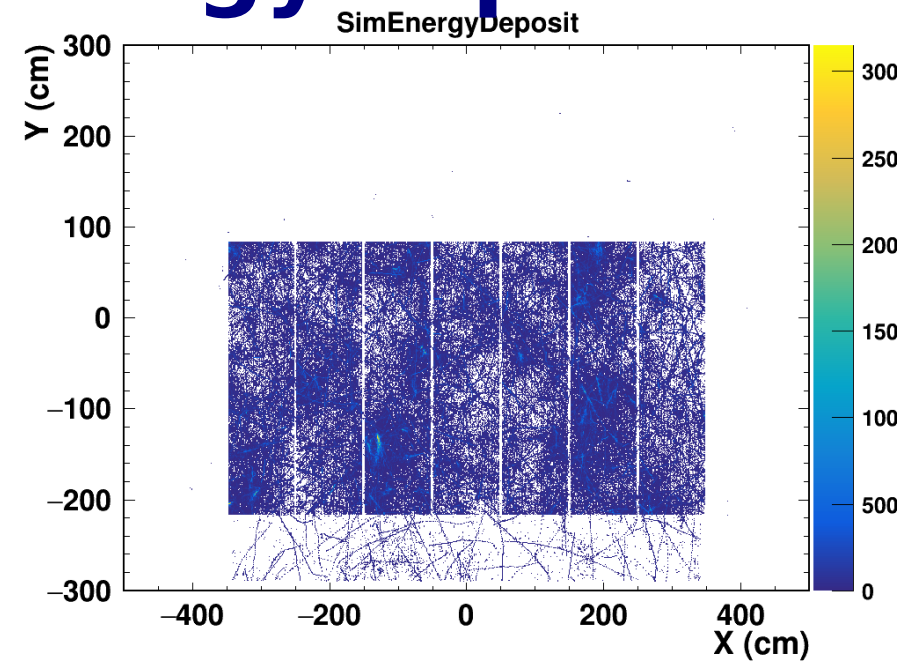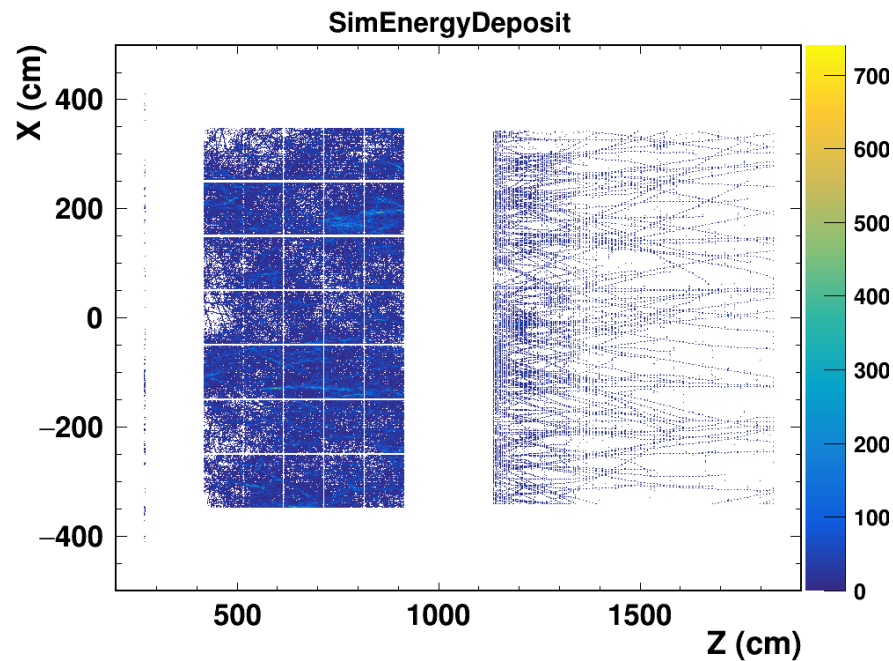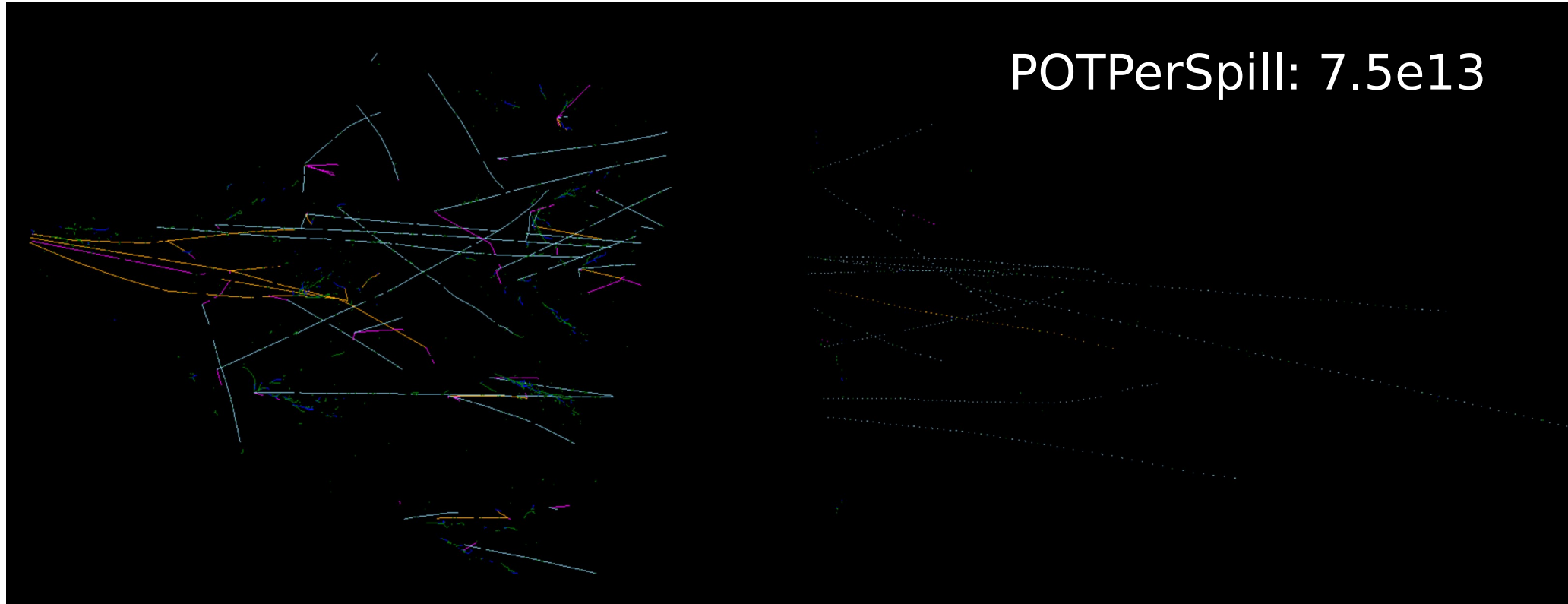

SimEnergyDeposit

- Energy depositions in three sensitive volumes.
- 7x5 LArTPC modules
  - Each has two drift volumes
- TMS is lower than LArTPC
  - Beam angle is -0.101 rad.

SimEnergyDeposit

UNIVERSITY OF
CAMBRIDGE
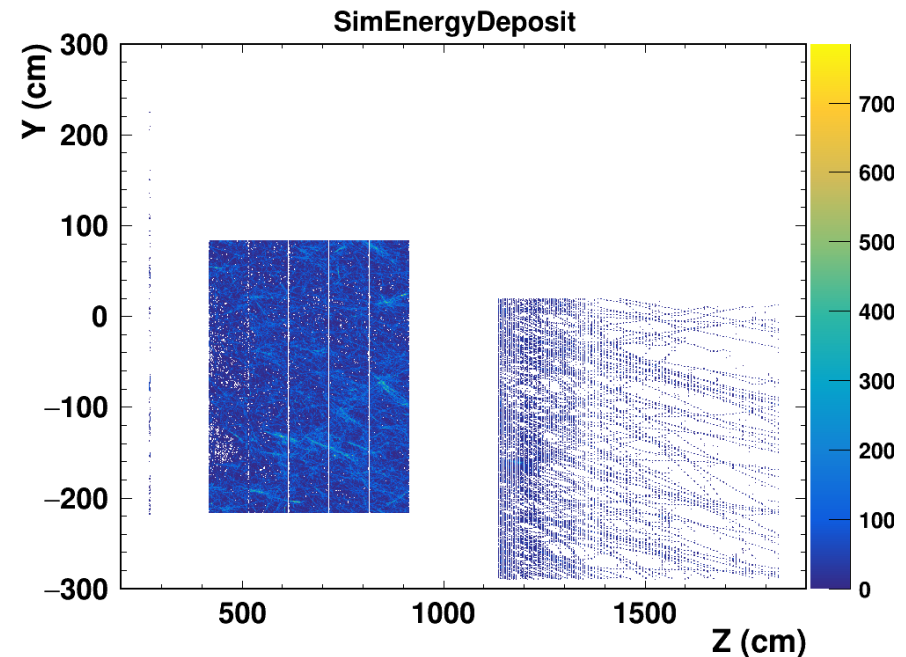
LArSoft Simulated Spill


POTPerSpill: 7.5e13

59 neutrinos!

UNIVERSITY OF CAMBRIDGE

Input from NDSim MC

- We are also capable of running from the ND MC files generated.

  - Leigh will be presenting the 2x2 Simulation Challenge events later, so I'm not going to say much here at all.

- This process involves Kevin making HDF5 files that resemble data.

- Richie decodes them into simple events with hits.

- We read these in and reconstruct them into particle hierarchies.

- This is all independent of LArSoft.

  - Unless you want LArSoft PFParticle objects as your output.

    – These were initially designed to represent Pandora outputs in LAr and closely resemble a Pandora internal structure.

The Pandora SDK

- The Pandora Software Development Kit is designed to handle generic reconstruction and clustering tasks.

  - From individual charge deposits.

  - Clustering.

  - Identifying particles.

  - Constructing hierarchies.

  - Making a "biography" of where a particle first interacted and everything it created.

- Currently most heavily developed for 3x2 Wire Plane LArTPCs.

  - SDK highly portable to new detector technologies.

  - New experimental designs always have some common needs and some unique ones.
   - Some work reusable, some must be fresh.
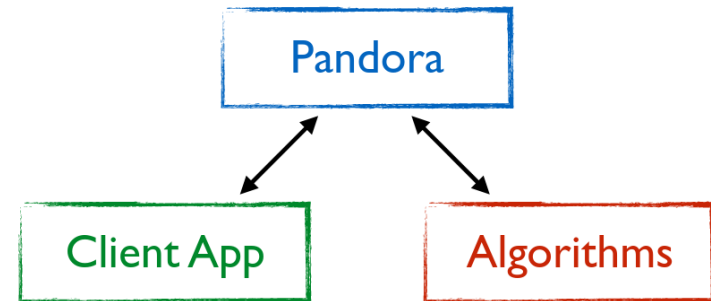
Integration with LArSoft

- Pandora is a standalone client application.

- It handles reconstruction in its own process, in its own data representation.

  - And small translation layers are constructed to interface with any relevant software stack.

- Think of Pandora as a tool in a way similar to GENIE or GEANT.

  - LArSoft can talk to it and invoke it, but it is not part of LArSoft.

- For example, the LArPandora module exists as an interface to Pandora.

  - But this is a part of LArSoft, not a part of Pandora.
    - https://github.com/LArSoft/larpandora
  - All actual reconstruction algorithms for LArTPCs are handled by LArContent, a package in Pandora that has no relation to LArSoft.
    - https://github.com/PandoraPFA/LArContent

- Plugging into an alternative toolchain means writing a new interface module, not a new reconstruction.

UNIVERSITY OF
CAMBRIDGE

Multiple Algorithm Approach

- The Pandora approach is to run many algorithms in sequence to steadily increase the amount of information known about the event.

  - Generally takes a "conservative" approach:
    - Each algorithm does things it is sure are correct, and leaves uncertain choices to be resolved further down the line.

- Algorithms can vary in complexity hugely:

  - Some are as simple as "ignore hits below this threshold".

  - Others can turn hits into identified particles using Deep Learning.

- Sequential algorithm use makes it trivial to do things "iteratively" with feedback between independent algorithms.

UNIVERSITY OF CAMBRIDGE
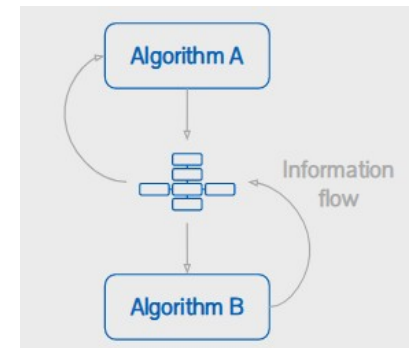
Actual Code Structure

- Pandora algorithms are written in C++, and have a simple structure: get input, do operations, write output.

  - All of this is controlled using XML files.

    - Actually running Pandora means taking in an XML file that contains a list of tasks to perform. These can be nested to perform multiple tasks in a single job, or sequentially in separate jobs.

      - For example, for doing an N-detector global reconstruction from scratch, the "master" file would spawn multiple worker processes, that reconstruct their individual detectors.

      - Then another instance that takes the output of these and performs global reconstruction.

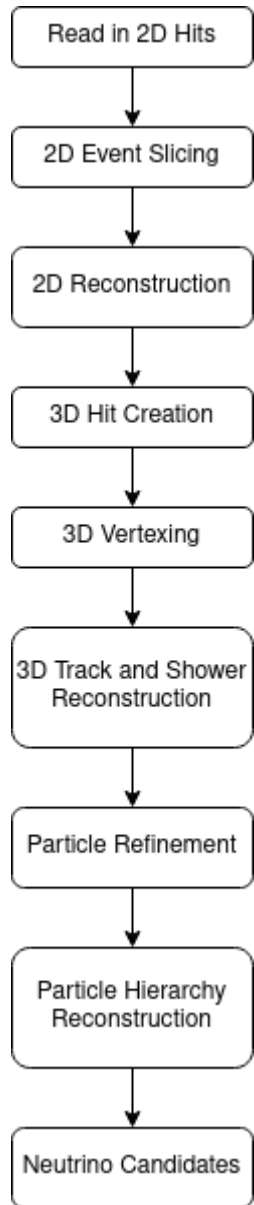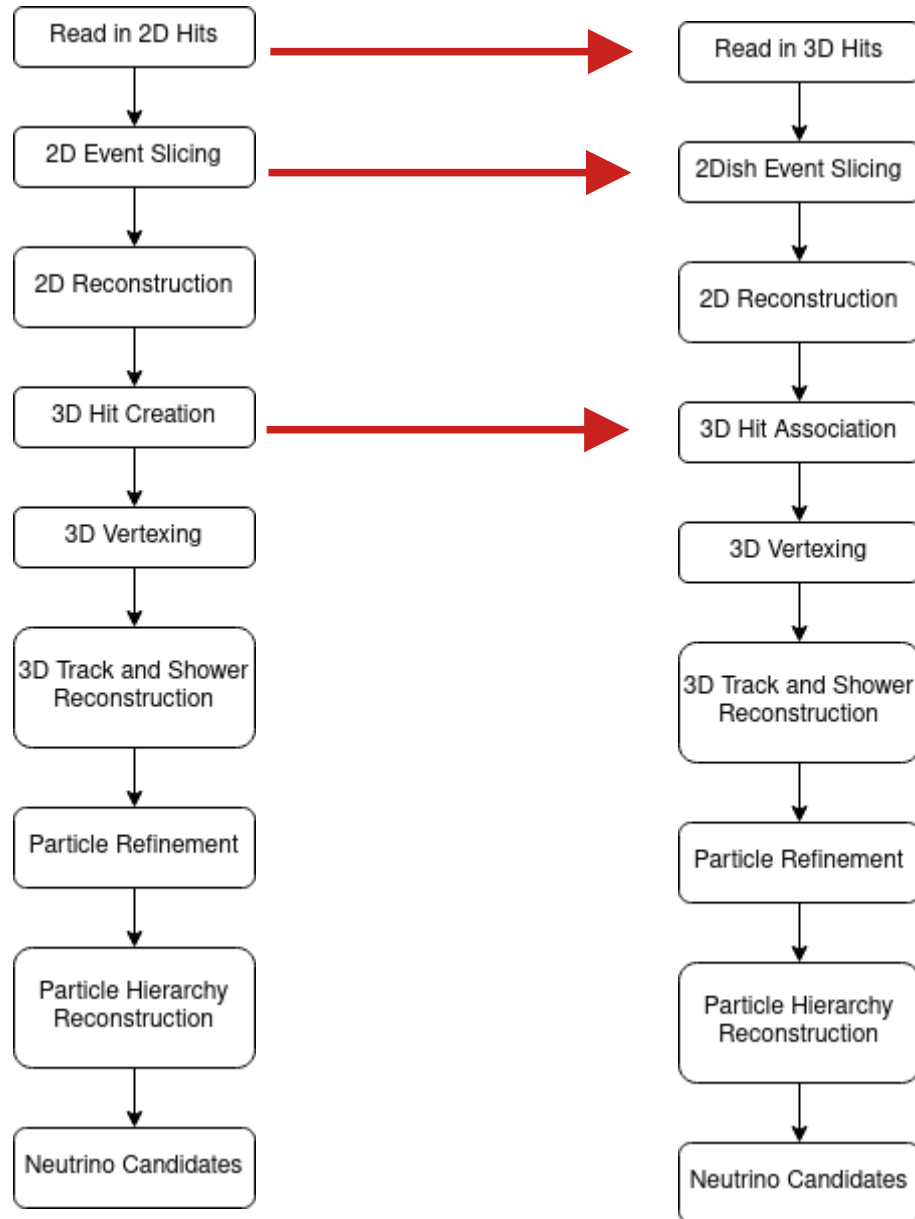| **Algorithm**	Pseudocode description of a client application for LAr TPC event reconstruction in a single drift volume
1:	**procedure** MAIN
2:	Create a Pandora instance
3:	Register Algorithms and Plugins
4:	Ask Pandora to parse XML settings file
5:	**for all** Events **do**
6:	Create CaloHit instances
7:	Create MCParticle instances
8:	Specify MCParticle-CaloHit relationships
9:	Ask Pandora to process the event
10:	Get output PFOs and write to file
11:	Reset Pandora before next event

Pandora Operations
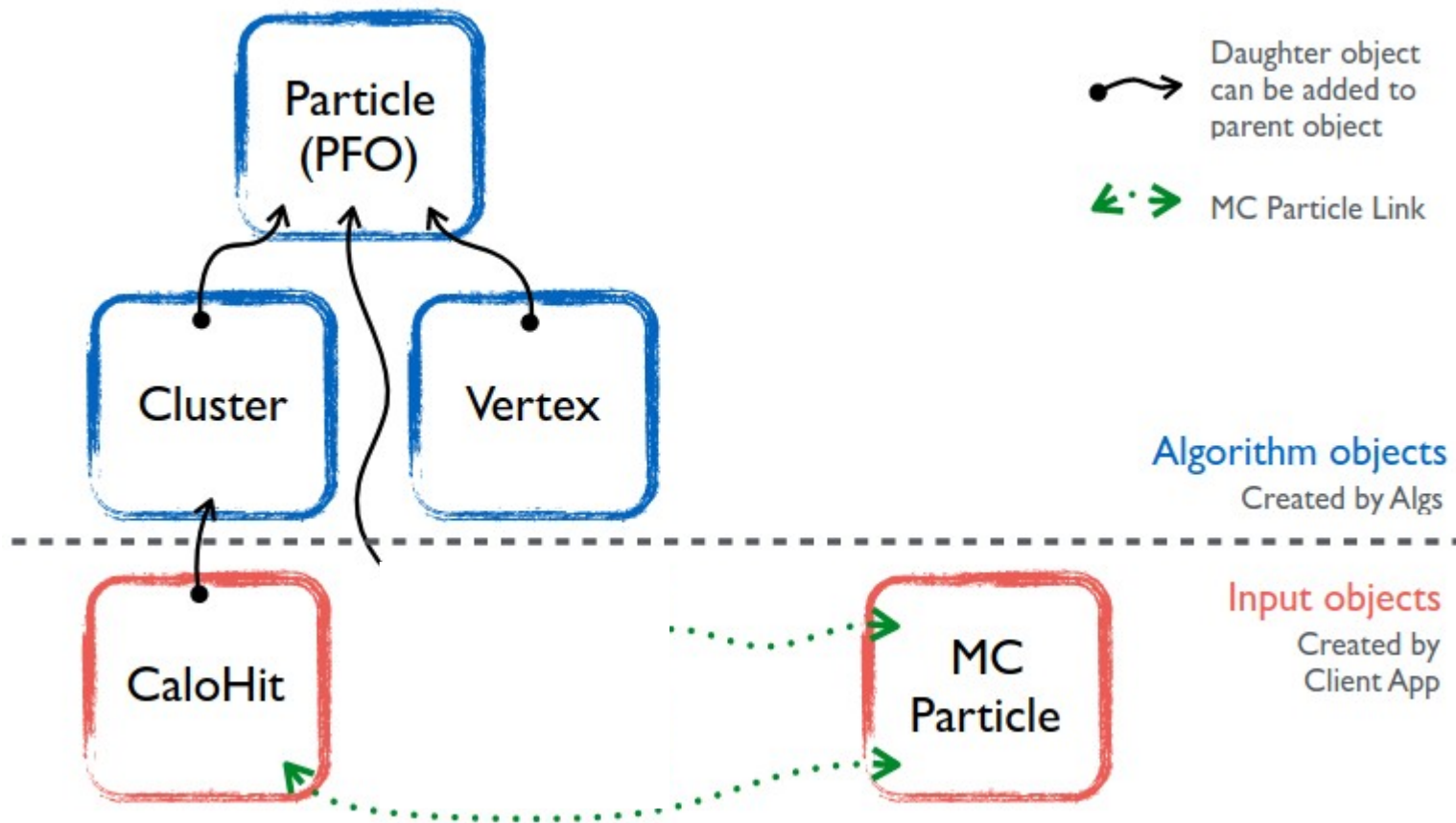
Pandora in the ND-Lar - History

- Initially began to use exactly the same reconstruction as the far detector, to within detector differences.

- This leverages the heavy development work that went into Wire Plane reconstruction.

    - And remains extremely very similar to the FD reconstruction.

    - But doesn't gain the full benefit of this detector's readout technology.

- So, currently upgrading our reconstruction to maximise reconstruction power at the ND using all 3D information.

    - Also investigating moving 3D hit matching earlier in our FD reconstruction so we can continue to maintain as much similarity as possible in later reconstruction stages.

Where we started (3x2D)

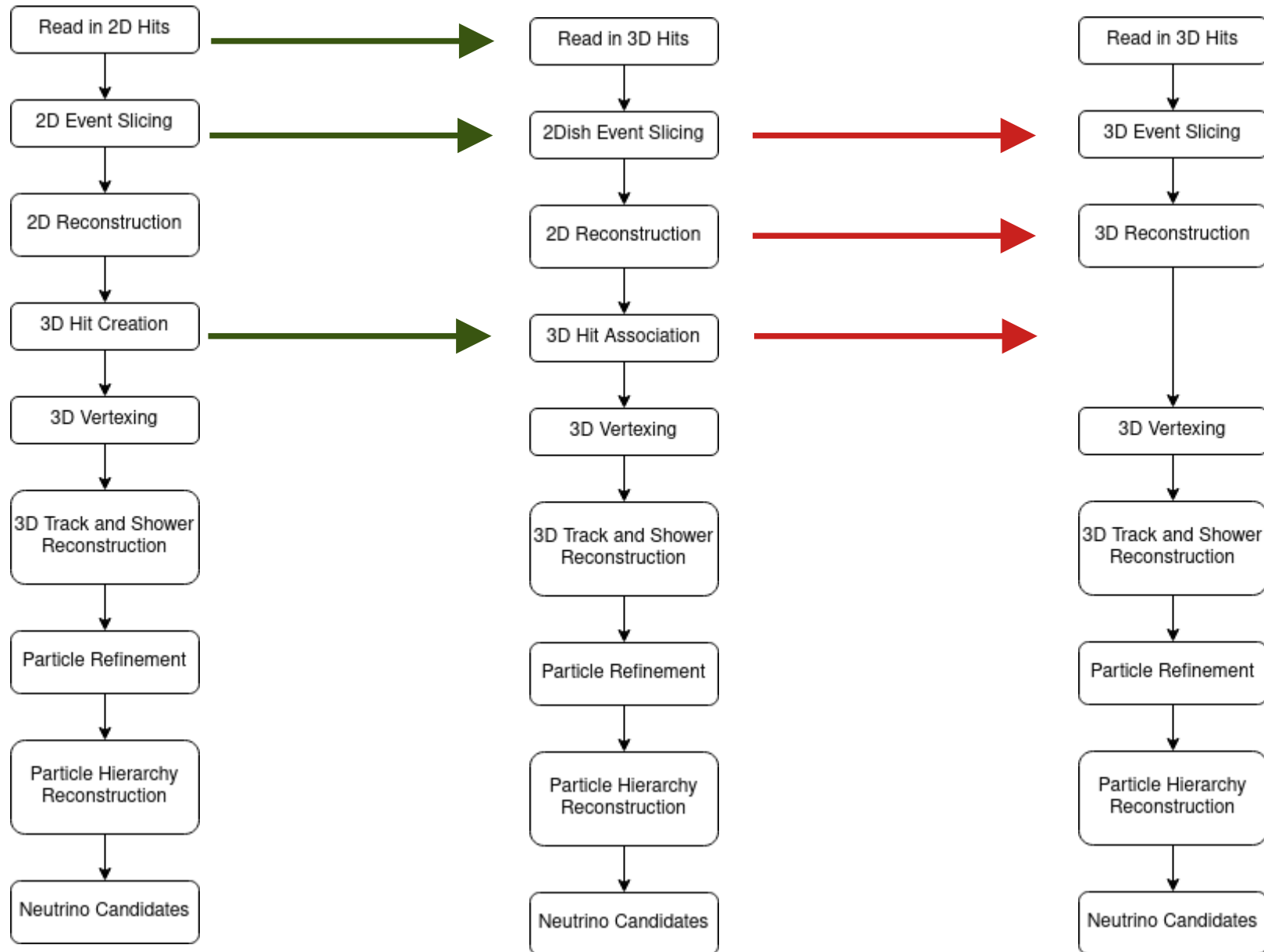Read in 2D Hits
↓
2D Event Slicing
↓
2D Reconstruction
↓
3D Hit Creation
↓
3D Vertexing
↓
3D Track and Shower Reconstruction
↓
Particle Refinement
↓
Particle Hierarchy Reconstruction
↓
Neutrino Candidates

Updates So Far

UNIVERSITY OF CAMBRIDGE

What we're developing

UNIVERSITY OF CAMBRIDGE

Outputs

- We're still primarily outputting our reconstruction to our own output format.
  - Which works great for us.
    - And can be almost transparently be turned into LArSofty events.
  - But clearly we need to produce something analysis-level.
    - Which is presumably CAFs.
    - And now there seems to be an appetite for this to be directly rather than via ART?
    - Opinions may vary.

- I've heard an awful lot of opinions on this, from several different perspectives.

  - Please speak up if you have one.

Summary

- Pandora is a well-established reconstruction package among liquid argon experiments.

  - Including as the primary reconstruction for DUNE FD physics reach proposals.

- Significant effort has been committed to maximise the output of the DUNE ND.

  - Using similar reconstruction to the FD for direct comparability.

  - And developing our techniques with the new capabilities of this detector.

  - A lot of progress has been made on the simulated detector in various simulations.

- I'll leave it to Leigh to discuss the results of our looking at the simulation challenge files.

- Thanks for listening.