Towards generic nu selection in SBND

Haiwang Yu

2025-01-30

Outline 1

- **Clustering validation + tool development – Ewerton, Hanyu**
  - goal:
    - develop a useful tool to evaluate img/clus results
    - look for cluster failures and tune/improve the clus parameters/alg.
  - clustering code not crashing **in all cases**
  - save out SimEnergyDepo
  - group depos based on time
  - match Cluster points with depo points
  - metric calculation

- **QL-Matching – Lynn, Haiwang**
  - goal:
    - adopt the many-to-many uboone ql-matching alg. to
      - determine cluster t0 (tag in-beam cluster)
      - improve clustering
  - Implement data objects to provide charge/light information
  - Bring in the semi-analytical photon library
  - Initial implementation of matching alg.
  - ...

Outline 2

- **DL-Clustering – Moon, Avinay, Haiwang**
  - goal:
    - improve neutrino clustering especially pi0 related
    - prioritize efficiency more
  - Data preparation
    - NCpi0 + cosmic sample with current clustering alg.
      - https://github.com/HaiwangYu/sbnd-dl-clustering/blob/main/README.md#clustering-input-data
    - Save out reco. info
    - Save out truth info
  - Algorithm development
    - information to use
    - network architecture
    - ...

as_pctree crashing with empty root nodes

https://github.com/WireCell/wire-cell-toolkit/issues/364

aux/src/TensorDMpointtree.cxx +4 -1

```
@@ -161,7 +161,10 @@ WireCell::Aux::TensorDM::as_pctree(const TensorIndex& ti,
161                                                    161
162         // Loop cross product of (PC name,node)    162         // Loop cross product of (PC name,node)
163         for (const auto& [pcname, pcds] : pointclouds) {  163         for (const auto& [pcname, pcds] : pointclouds) {
164 -                                                  164 +           if (!lpcmaps_ds.get(pcname)) {
                                                       165 +               /// TODO: @BV what does this mean?
                                                       166 +               continue;
                                                       167 +           }
165             // Get local PC map vector as a span on DS array.  168             // Get local PC map vector as a span on DS array.
166             auto lpcmap = lpcmaps_ds.get(pcname)->elements<lpcmaps_index_type>();  169             auto lpcmap = lpcmaps_ds.get(pcname)->elements<lpcmaps_index_type>();
167                                                    170
```

80b80a3

Another crash for DUNE-FD-VD

Can be reproduced using standalone configuration
working on some instructions to capture that
this can be fixed using: 2692f

aux/src/TensorDMpointtree.cxx +4

```
@@ -124,6 +124,10 @@ std::unique_ptr<WireCell::PointCloud::Tree::Points::node_t>
124  WireCell::Aux::TensorDM::as_pctree(const ITensor::vector& tens,
125                                     const std::string& datapath)
126  {
127  +     /// TODO: FIXME: this works, but need to understand why
128  +     if (tens.size() <= 2) {
129  +         return std::make_unique<WireCell::PointCloud::Tree::Points::node_t>();
130  +     }
127  124      TensorIndex ti(tens);
128  125      return as_pctree(ti, datapath);
129  126  }
130  127
131  128
132  129
133  130
```

Brookhaven National Laboratory

QL-Matching, q/l information providing

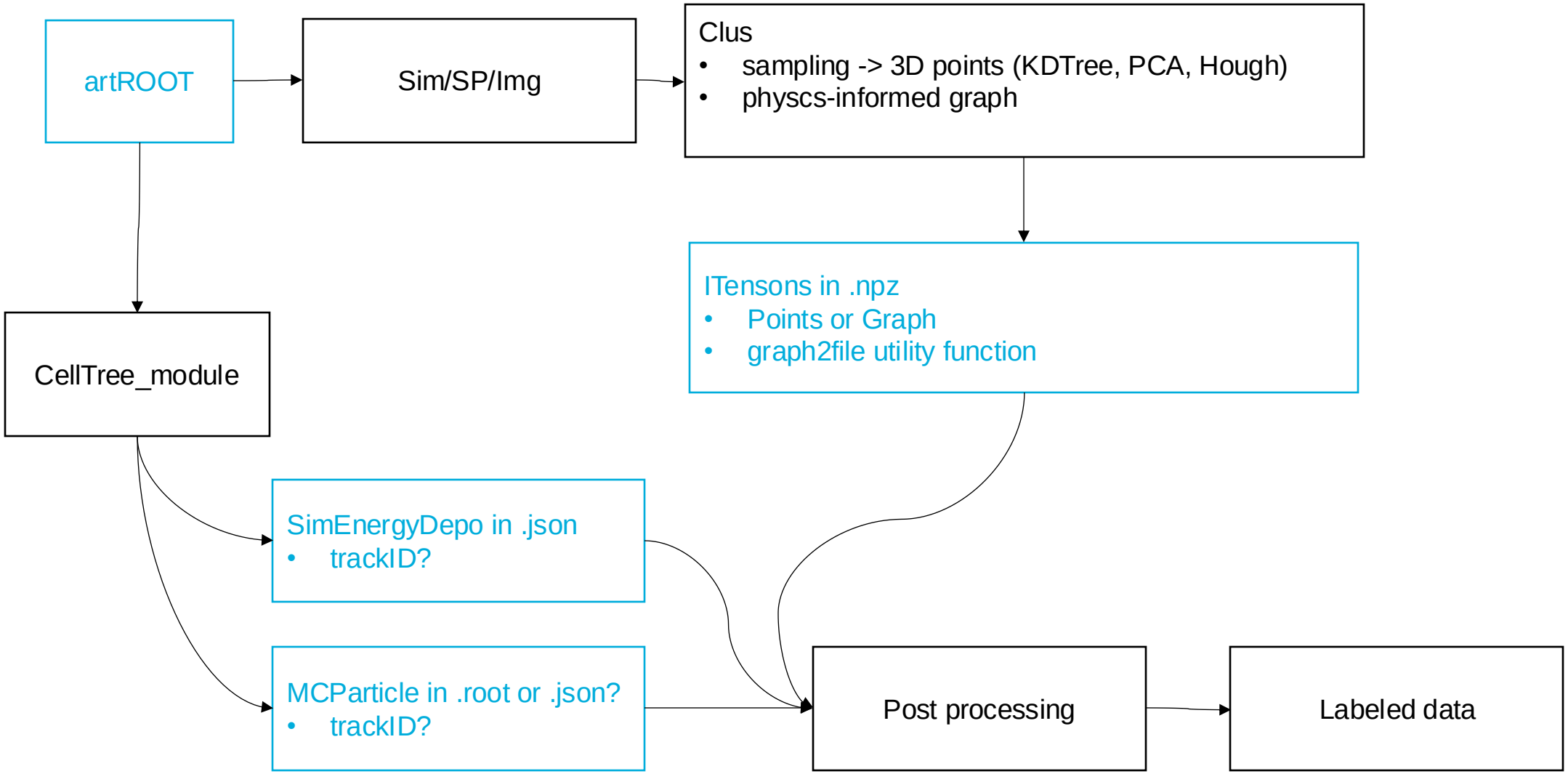
- The src code is here: https://github.com/HaiwangYu/larwirecell/tree/qlmatch-ls991
  - The demo part is here: https://github.com/HaiwangYu/larwirecell/blob/qlmatch-ls991/larwirecell/qlmatch/QLMatching.cxx#L93-L152
- The cfg and a simple instruction is here: https://github.com/HaiwangYu/sbnd-op/tree/main

https://www.phy.bnl.gov/twister/bee/set/69cb6967-b68a-4f63-9717-f20e24bf8475/event/2/
lynn-sim.root, event 1, eventNum 2

```
593 [19:59:37.051] D [matching] <QLMatching:matching0> checking TimingTPCBundle object
594 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 0 icluster0 time -0.6394203060764309 total_PE 88611.48962499975 ks_dis 0 chi2 0 ndf 60 consistent_flag true
595 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 0 icluster1 time -0.6394203060764309 total_PE 88611.48962499975 ks_dis 1 chi2 144587046.05169508 ndf 60 consistent_flag false
596 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 0 icluster2 time -0.6394203060764309 total_PE 88611.48962499975 ks_dis 1 chi2 144587046.05169508 ndf 60 consistent_flag false
597 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 0 icluster3 time -0.6394203060764309 total_PE 88611.48962499975 ks_dis 1 chi2 144587046.05169508 ndf 60 consistent_flag false
598 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 0 icluster4 time -0.6394203060764309 total_PE 88611.48962499975 ks_dis 1 chi2 144587046.05169508 ndf 60 consistent_flag false
599 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 0 icluster5 time -0.6394203060764309 total_PE 88611.48962499975 ks_dis 1 chi2 144587046.05169508 ndf 60 consistent_flag false
600 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 0 icluster6 time -0.6394203060764309 total_PE 88611.48962499975 ks_dis 1 chi2 144587046.05169508 ndf 60 consistent_flag false
601 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 0 icluster7 time -0.6394203060764309 total_PE 88611.48962499975 ks_dis 1 chi2 144587046.05169508 ndf 60 consistent_flag false
602 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 0 icluster8 time -0.6394203060764309 total_PE 88611.48962499975 ks_dis 1 chi2 144587046.05169508 ndf 60 consistent_flag false
603 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 0 icluster9 time -0.6394203060764309 total_PE 88611.48962499975 ks_dis 1 chi2 144587046.05169508 ndf 60 consistent_flag false
604 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 0 icluster10 time -0.6394203060764309 total_PE 88611.48962499975 ks_dis 1 chi2 144587046.05169508 ndf 60 consistent_flag false
605 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 1 icluster0 time -0.24027122833153763 total_PE 88234.34492500071 ks_dis 0 chi2 0 ndf 58 consistent_flag true
606 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 1 icluster1 time -0.24027122833153763 total_PE 88234.34492500071 ks_dis 1 chi2 283001048.94328725 ndf 58 consistent_flag false
607 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 1 icluster2 time -0.24027122833153763 total_PE 88234.34492500071 ks_dis 1 chi2 283001048.94328725 ndf 58 consistent_flag false
608 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 1 icluster3 time -0.24027122833153763 total_PE 88234.34492500071 ks_dis 1 chi2 283001048.94328725 ndf 58 consistent_flag false
609 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 1 icluster4 time -0.24027122833153763 total_PE 88234.34492500071 ks_dis 1 chi2 283001048.94328725 ndf 58 consistent_flag false
610 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 1 icluster5 time -0.24027122833153763 total_PE 88234.34492500071 ks_dis 1 chi2 283001048.94328725 ndf 58 consistent_flag false
611 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 1 icluster6 time -0.24027122833153763 total_PE 88234.34492500071 ks_dis 1 chi2 283001048.94328725 ndf 58 consistent_flag false
612 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 1 icluster7 time -0.24027122833153763 total_PE 88234.34492500071 ks_dis 1 chi2 283001048.94328725 ndf 58 consistent_flag false
613 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 1 icluster8 time -0.24027122833153763 total_PE 88234.34492500071 ks_dis 1 chi2 283001048.94328725 ndf 58 consistent_flag false
614 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 1 icluster9 time -0.24027122833153763 total_PE 88234.34492500071 ks_dis 1 chi2 283001048.94328725 ndf 58 consistent_flag false
615 [19:59:37.052] D [matching] <QLMatching:matching0> bundle flash 1 icluster10 time -0.24027122833153763 total_PE 88234.34492500071 ks_dis 1 chi2 283001048.94328725 ndf 58 consistent_flag false
```

Brookhaven National Laboratory

DL-Clustering data preparation



Save out reco. info

blob: 2D tensor
- [blob_idx], q, qp0 (uq), dqp0 (duq) ..., wind_min_dp0 (wind_u_min) ..., corner0_x, corner0_y, corner0_z, ... corner5_x,

points: 2D tensor
- [pt_idx], blob_idx, **x, y, z,** q, qp0 (uq), dqp0 (duq) ...

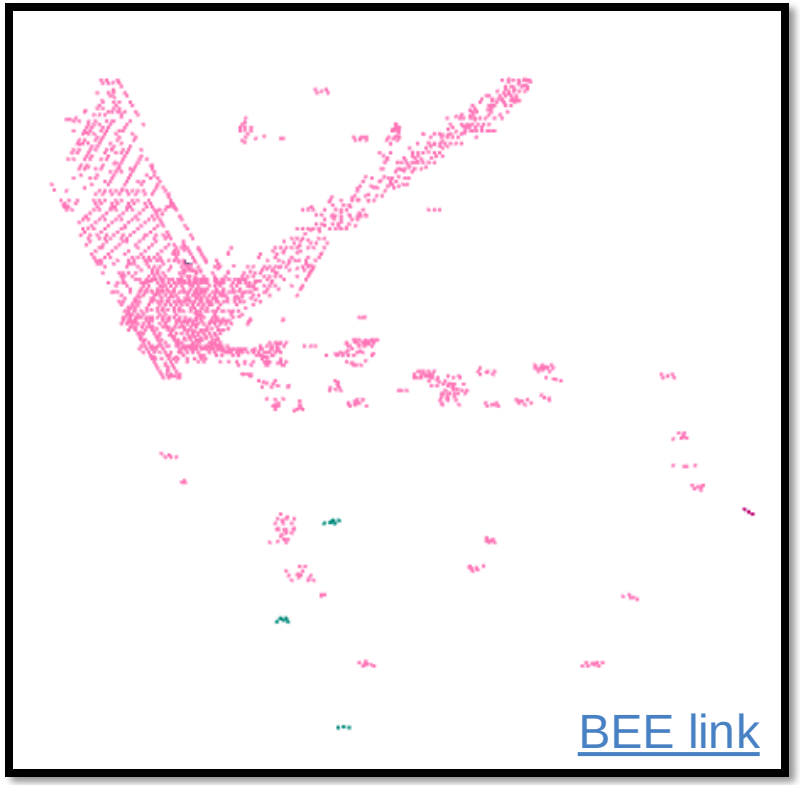
point_graph: 2D tensor
head (pt_idx), tail, weight

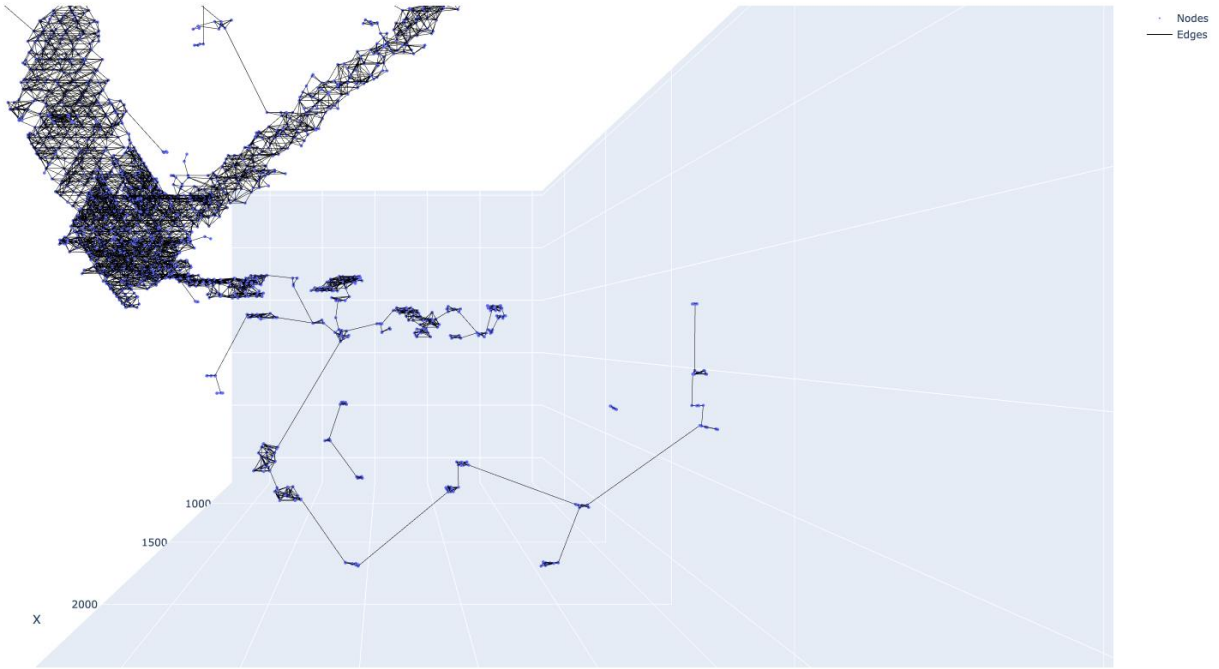
light: ?

Save out reco. info

- torch_geometric.data
- visualization: plotly
- https://github.com/HaiwangYu/sbnd-dl-clustering/blob/main/explore.py

A DUNE-FD-VD numu event in one APA





Brookhaven National Laboratory

Save out truth info

https://github.com/LArSoft/larreco/blob/develop/larreco/WireCell/CellTree_module.cc

https://github.com/LArSoft/lardataobj/blob/develop/lardataobj/Simulation/SimEnergyDeposit.h
https://github.com/NuSoftHEP/nusimdata/blob/develop/nusimdata/SimulationBase/MCParticle.h

Are these trackIDs syncronized?

`SimEnergyDeposit::trackID` or `SimEnergyDeposit::origTrackID`

?

`MCParticle::TrackID()`