Reconstruction

LAr TPC R&D Workshop
Mar. 20, 2013

H. Greenlee

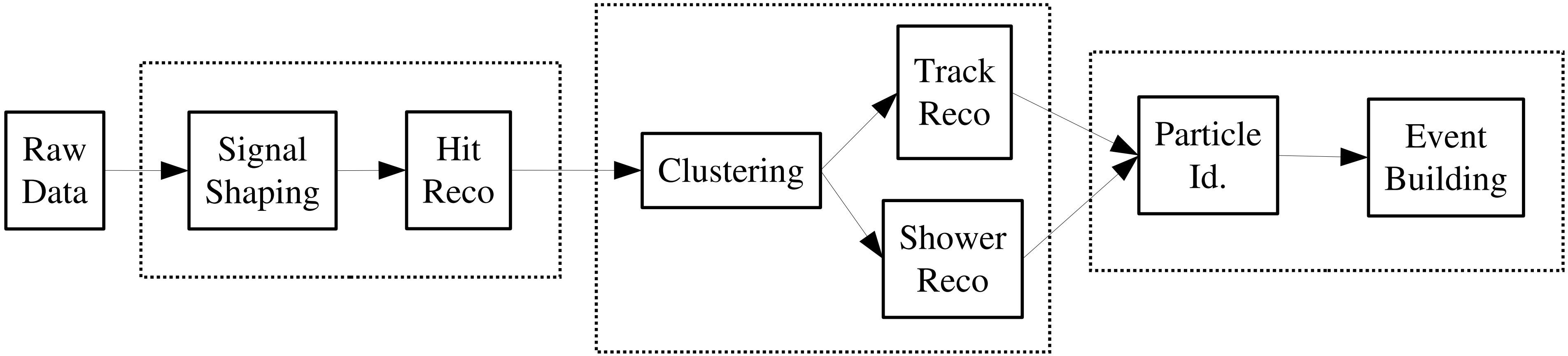
Outline

- Overview.
- Specific algorithms.
  - Signal shaping.
  - Hit finding.
  - Space points.
  - Pattern recognition.
    - Clustering.
    - Track reconstruction.
    - Shower reconstruction.
  - Calorimetry.
  - Optical reconstruction and cosmic rejection.

Overview

- I will give an overview of where things stand in larsoft, which is an umbrella group developing reconstruction algorithms for LAr TPCs at Fermilab.
- I will also give my opinion on what still needs to be done, and what the major problems are.

Reconstruction Chain



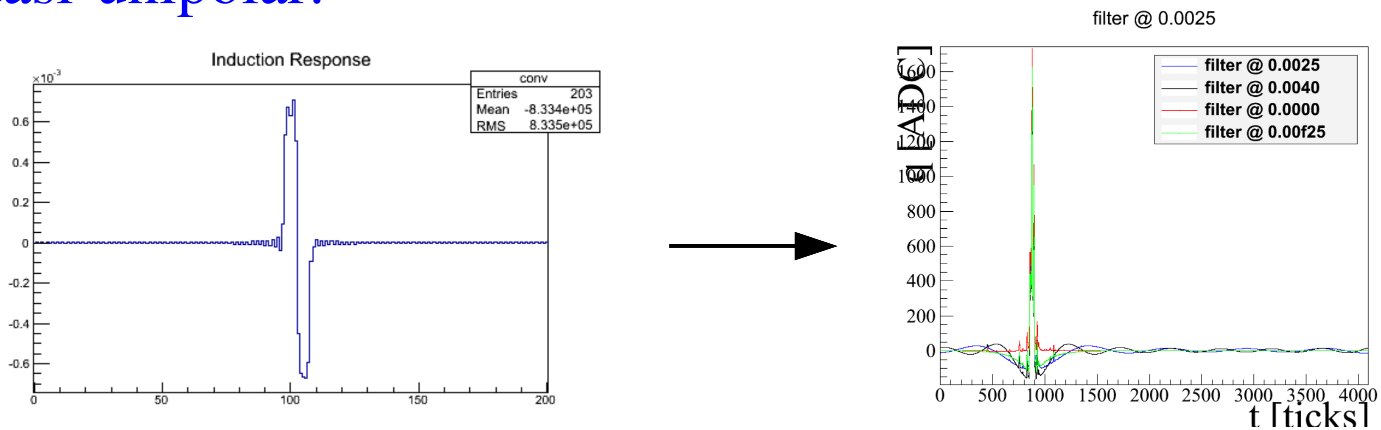
- Early stage reconstruction – Signal shaping and hit reconstruction.
  - Working algorithms exist. We know what the remaining problems are and what needs to be done.
- Pattern recognition – Clustering, track and shower reconstruction.
  - Hardest part of reconstruction problem. We don't yet have algorithms that can handle event reconstruction in the general case.
- Particle id.
  - Main challenge here is to understand calorimetry, including recombination effect. A lot of progress has been made.

Signal Shaping

- Raw ADC counts can be modeled (in simulation and real data) as the convolution of arriving charge with a response function, plus noise.
  - $A(f) = R(f)\, Q(f) + N(f)$.
- Above relation can be approximately inverted as
  - $Q_{\mathrm{reco}}(f) = A(f)\, F(f) \,/\, R(f)$.
  - Filter function $F(f)$ is necessary to limit noise arising from frequencies where the response $R(f)$ is zero or small.
  - For a given signal and noise, an optimal filter function is called the Wiener filter.
    - $F(f) = |R(f)\, Q(f)|^2 \,/\, (|R(f)\, Q(f)|^2 + |N(f)|^2$

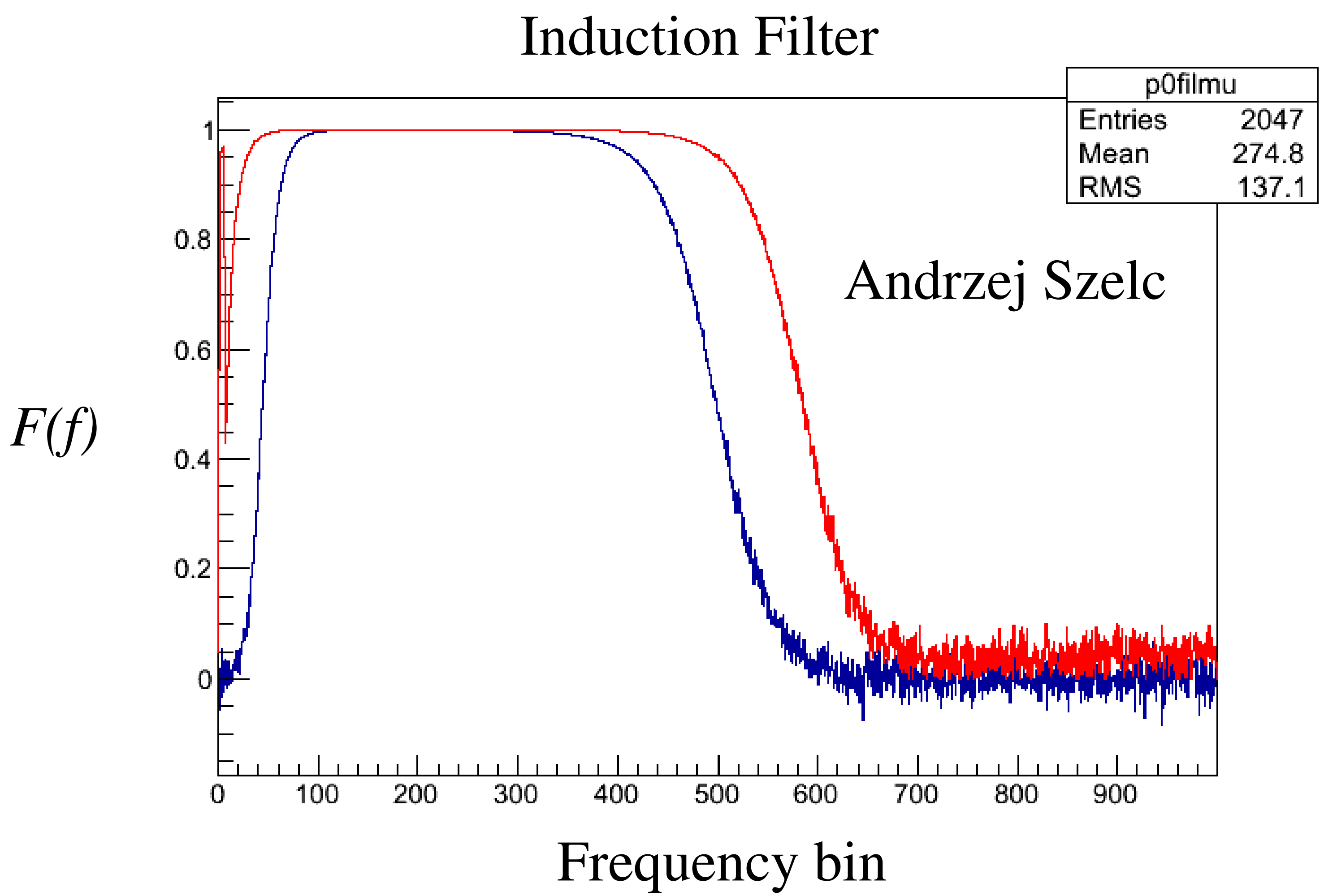
Signal Shaping (cont.)

- Signal shaping is used for both collection and induction planes, but its main purpose is to convert induction plane pulse from bipolar to quasi-unipolar.



- Actually, the shaped “unipolar” induction pulse is required to have zero area, because the response and filter are zero at zero frequency.
  - Shaped induction pulse consists of a narrow positive pulse superposed on a wide negative pulse.
  - The goal of induction signal shaping is to make the negative part of the pulse wider than the features you are trying to reconstruct, by pushing the low frequency filter cutoff as low as possible.

Wiener Filter



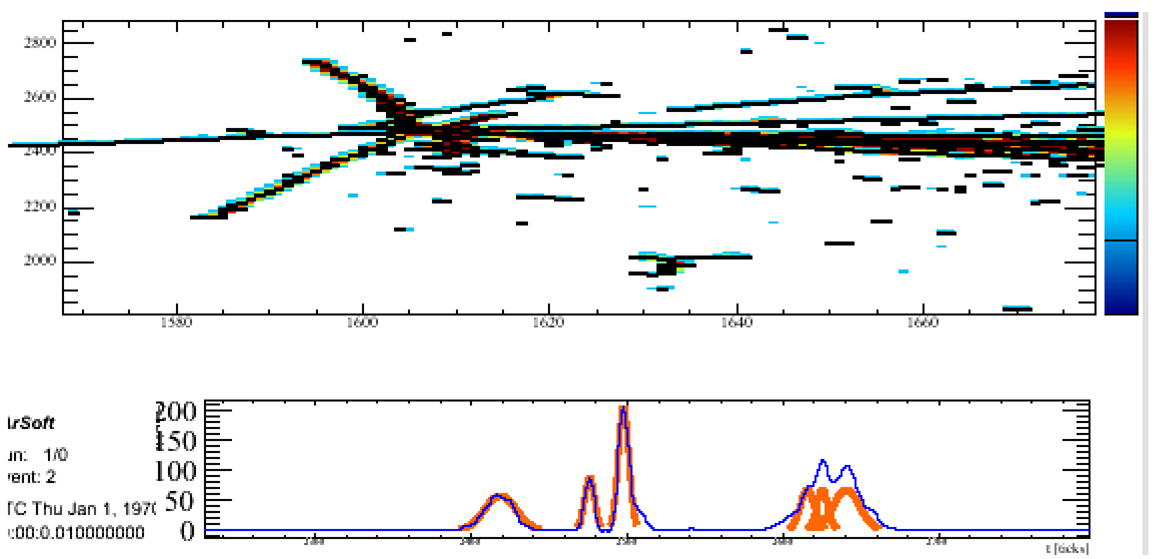
- Wiener filter depends on assumed signal size.

Signal Shaping Issues

- The most challenging issue regarding signal shaping reconstruction is knowing what the response function $R(f)$ and noise spectrum $N(f)$ are.
- Making the noise as small as possible is an important experiment design criterion.

Hit Finding

- Hit finding is the seemingly straightforward task of finding and characterizing peaks above baseline in the shaped readout waveform.



- Hits are characterized by 3-tuple: (position, width, area).

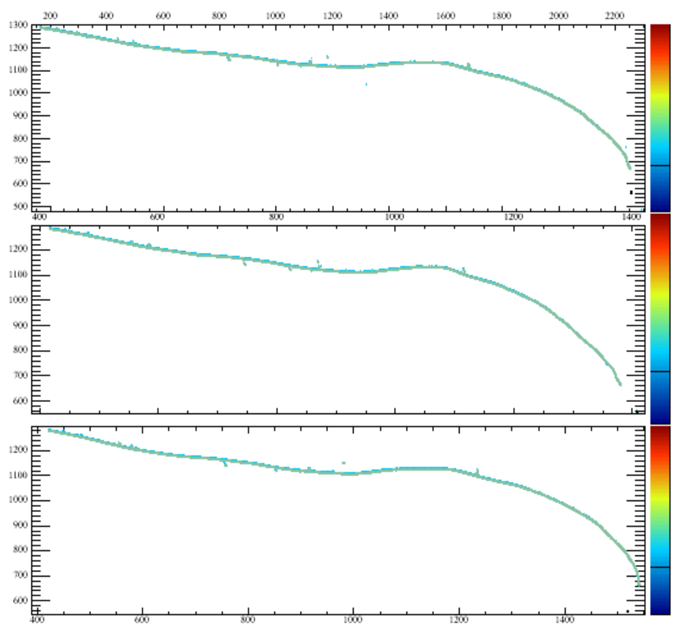
Hit Finding Issues

- Close hit separation – overlapping hits.
- Signal vs. noise hits.
- Error analysis – waveform is not a histogram.
- Using hit information.
  - Many simple pattern recognition algorithms view hits as pointlike (only use the hit position).
  - The second and third elements of the hit 3-tuple also give information that can also be used in pattern recognition.
    - Hit width is sensitive to “dip angle.”
    - Hit area is sensitive to particle type.

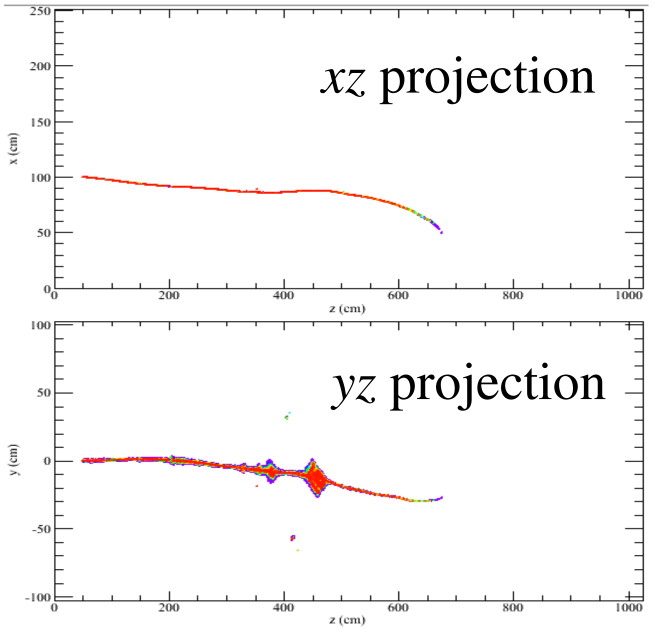
Space Points

- Hits from two or three views can be combined into (*x*,*y*,*z*) space points.
- Ab initio space point reconstruction from all hits (i.e. when not guided by some other kind of pattern recognition) is affected by a high fake rate, especially in case of particles moving parallel to readout wire planes.

Wire-Time View (Hits)

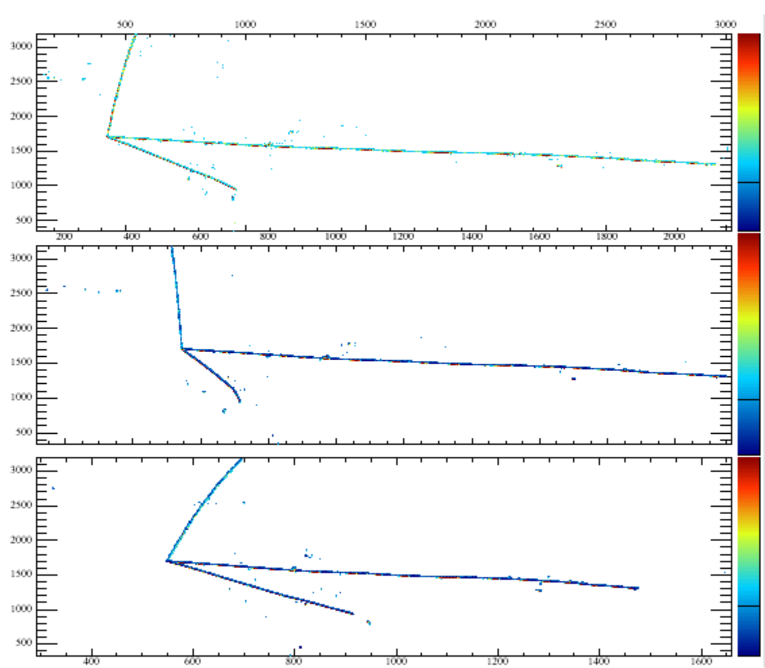


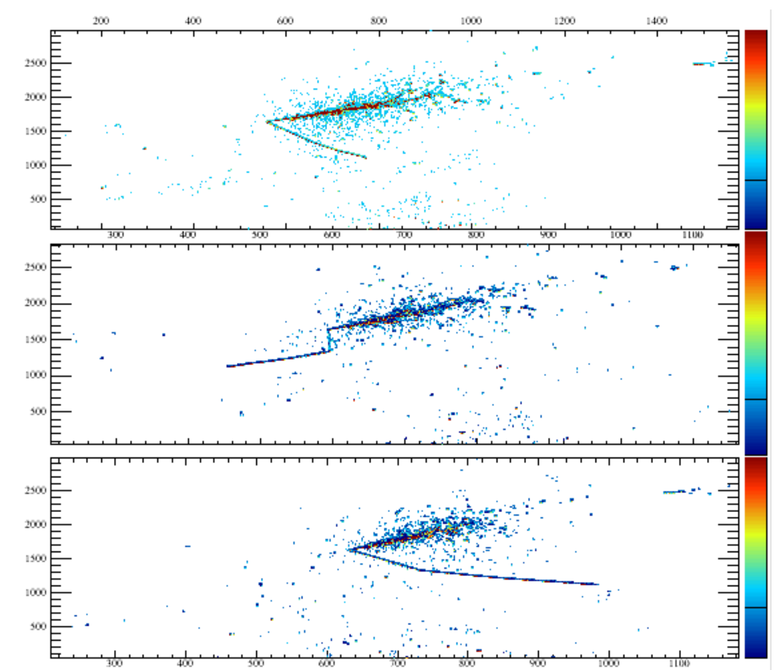
Ortho 3D View (SpacePoints)



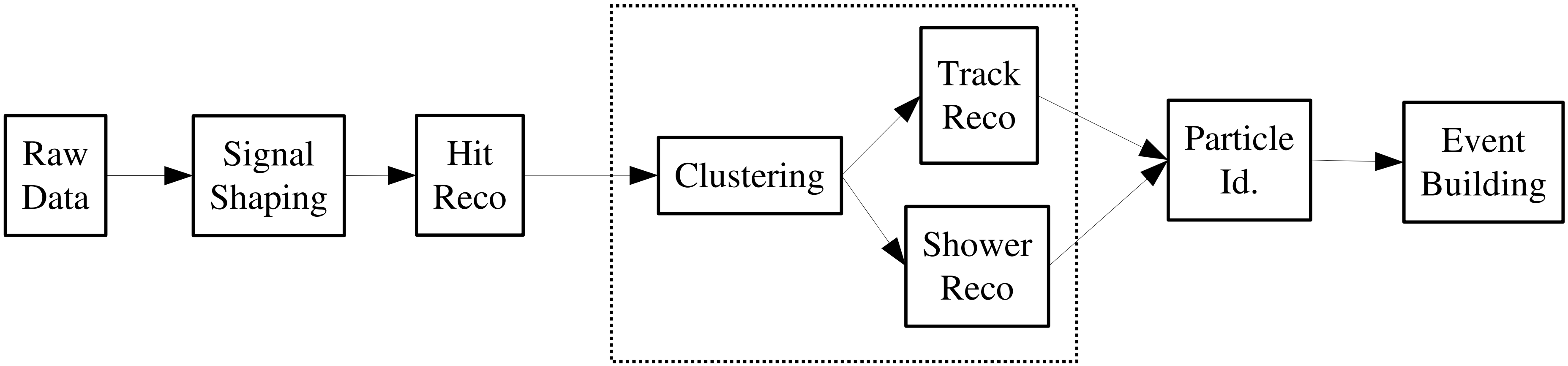
Pattern Recognition

- The job of pattern recognition algorithms is to identify and characterize related hits, to learn about the particle that caused the hits.
  - Hits caused by the same particle (within one view) are called clusters.
  - Clusters can be track-like or shower-like (electromagnetic showers).





Clustering



- Cluster reconstruction is the first stage of pattern recognition.
  - There is no sharp boundary between cluster reconstruction and either track or shower reconstruction.
  - Nevertheless, it can be useful to think about cluster reconstruction as a separate algorithm.
  - Cluster algorithms.
    - Dbscan.
    - Hough line finder.
    - Fuzzy clusters (Ben Carls).
    - Pandora (Andy Blake, Cambridge University group).

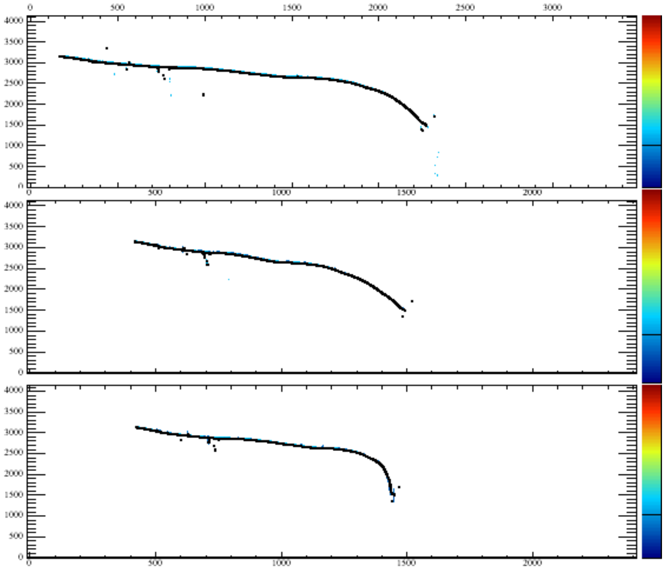
Dbscan

- Dbscan is a general purpose spatial clustering algorithm.
  - Not specific to high energy physics.
  - Clusters hits with an average density above some threshold.
  - Not aware of physical event features, such as vertices or kinks.
  - Dbscan is useful for rejecting isolated noise hits, and putting disjoint groups of hits into separate clusters.

Hough Line Finder

- Uses Hough transform to find straight track segments in one view (histogram method).
  - For small TPC, Hough lines can be used as tracks.
  - For large TPC, tracks are not straight enough to be found as single Hough lines.
  - Hough lines are still useful for finding parts of longer tracks, or track-like substructures in showers.

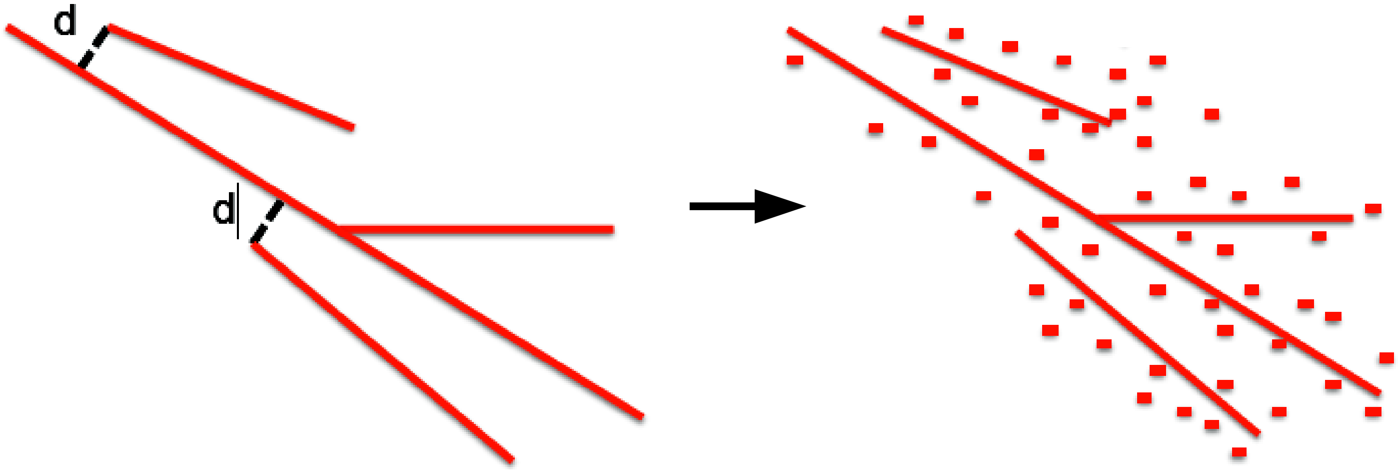
Single muon



Ben Carls

Fuzzy Clusters

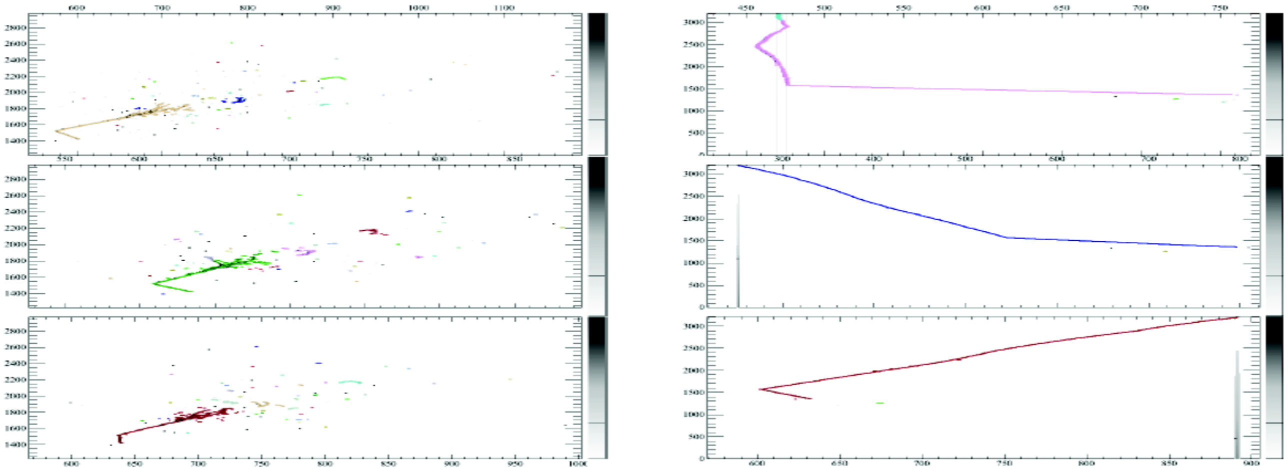
- Basic idea is to merge Hough lines based on proximity in space and angle (rather than just space, as with dbscan), then merge remaining nearby isolated hits.
  - Does a better job than dbscan in making “physics” clusters.



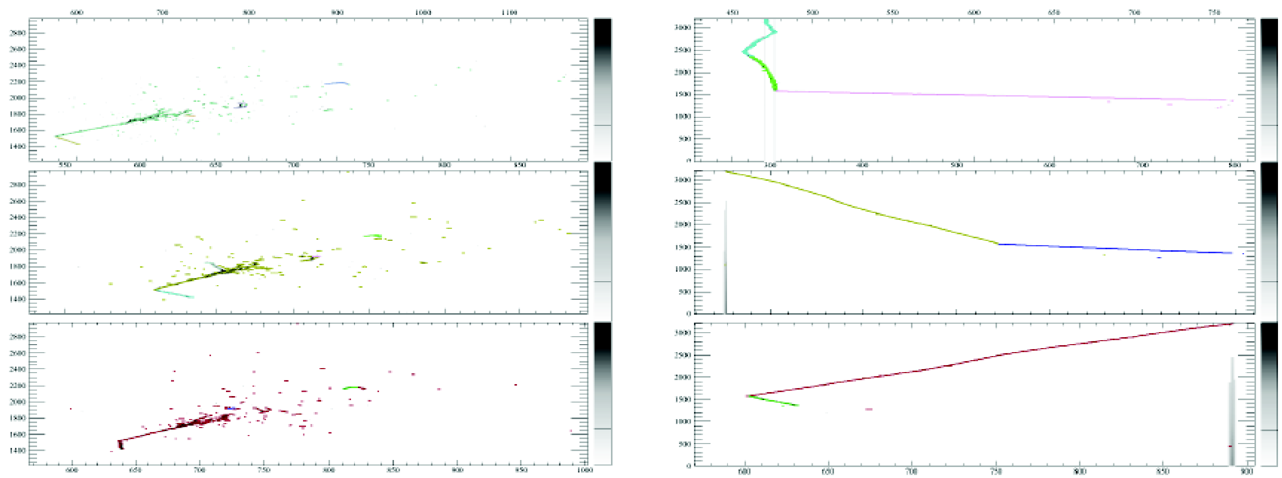
Ben Carls

Fuzzy Clusters

Dbscan



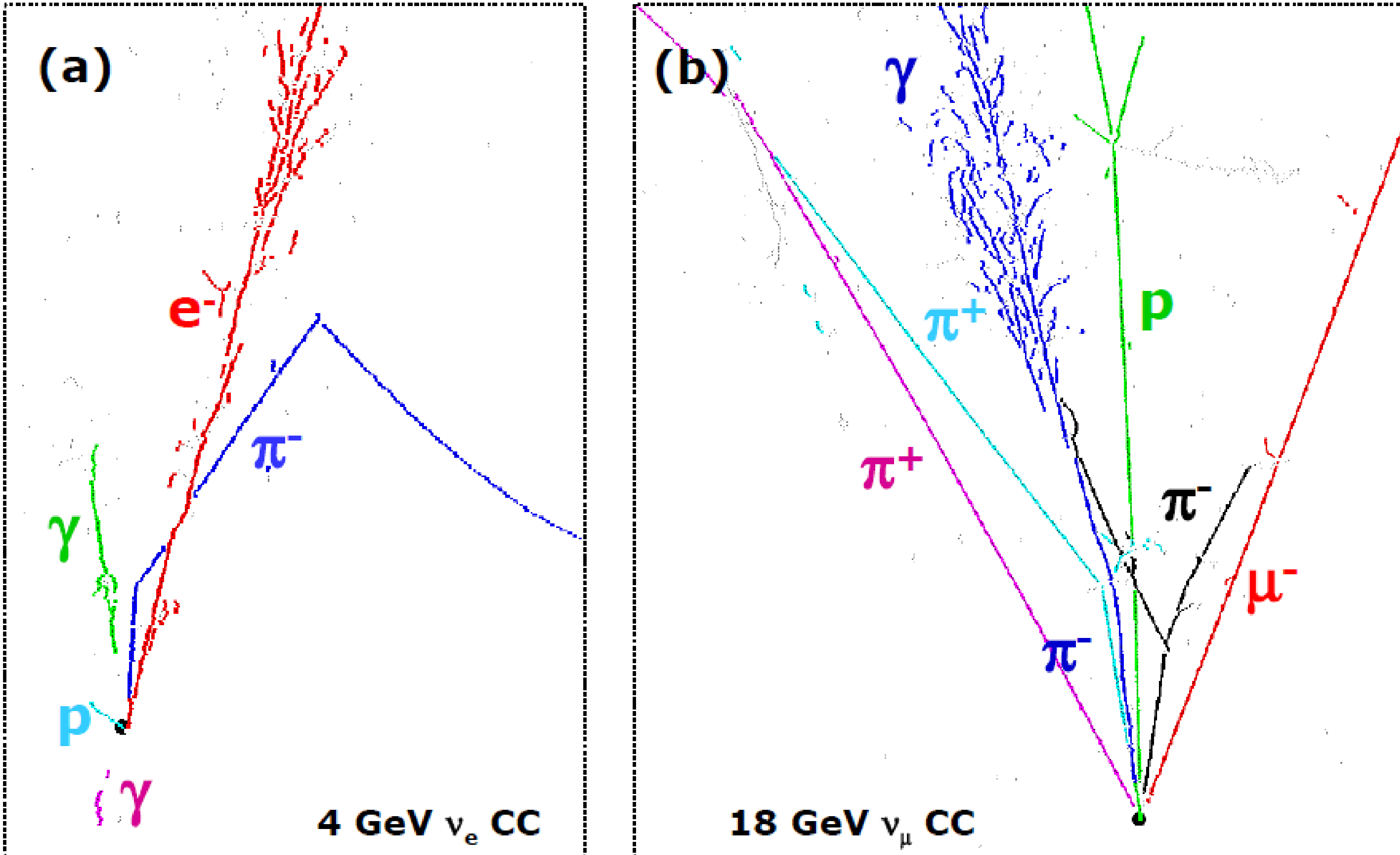
Fuzzy clusters



Pandora

- Presented by Andy Blake at November LBNE meeting in Houston. Also presented in larsoft.
- Starting from hits, pandora is a toolkit for doing all of the steps involved in pattern recognition, in 2D and (eventually) 3D.
- Pandora is installed as an external product at Fermilab, but not yet fully integrated into larsoft.

4. Final-State Particles



Coloured lines show reconstructed particles (labelled using MC truth)

Track Reconstruction

- Track reconstruction deliverables.
  - Track trajectory in 3D, start and end points, length.
  - Charge deposition *dQ/dx* at various points along track trajectory.
  - Track direction (i.e. forward vs. backward).
  - Momentum estimate.
    - Magnetic bend.
    - Range.
    - Multiple scattering (unique capability of LAr tpc).

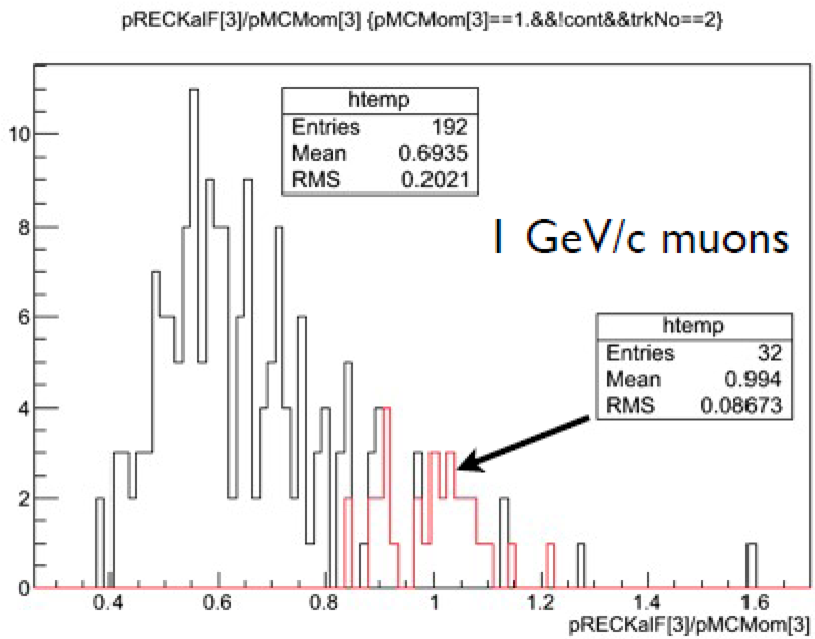
Track Reconstruction Algorithms in Larsoft

- 2D track reconstruction.
  - Merged hough lines.
    - Fuzzy clustering kind of already does this.
    - Larsoft includes a hough line merging algorithm, which has never been returned for large tpcs (but could be).
- 3D track reconstruction.
  - 3D Kalman filter using hits.
    - Algorithm exists in larsoft.
  - 3D Kalman filter using space points.
    - Algorithm exists in larsoft.
    - Probably not viable with ab initio space points, but potentially useful for use with space points obtained from pattern recognition.
  - 3D seed tracks.
  - Short track reconstruction.

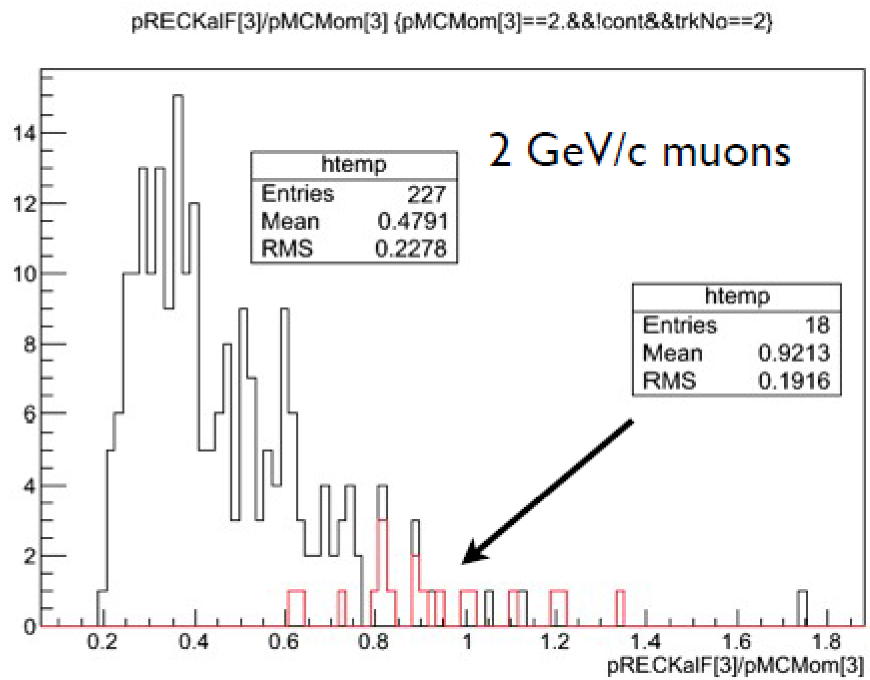
Multiple Scattering progress

Eric Church

17% (red) have 9% resolution
(require 20 **15 cm**-separated spacepts)



8% (red) have 19% resolution
(require 38 **15 cm**-separated spacepts)



These cuts effectively enforce seeing 50-60% of the track.

Can imagine doing better
with N separate decimations, using
all spacepoints. That's at top of list.

Shower Reconstruction

- In LAr TPC, showers always mean electromagnetic showers, not hadronic showers.
- Shower reconstruction deliverables.
  - Spatial reconstruction – 3D vertex and 3D shower axis.
  - Total shower energy.
  - Electron/photon separation (based on *dE/dx* in early part of shower).

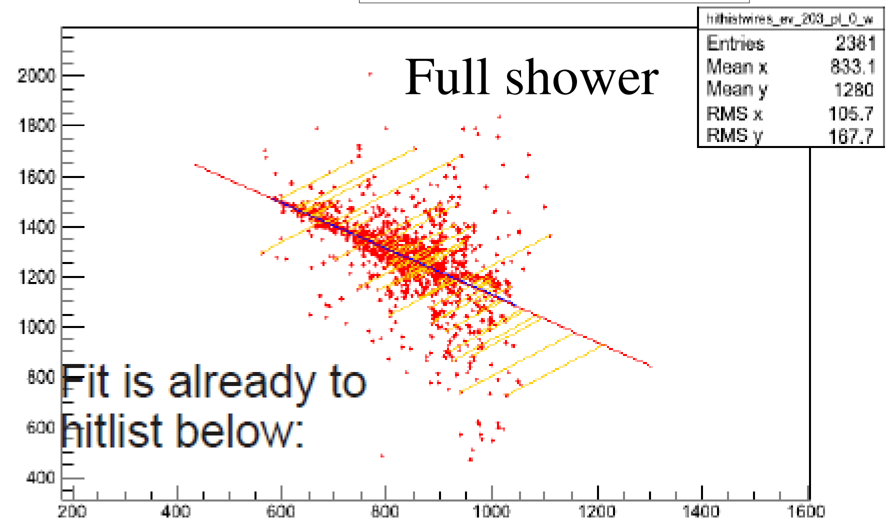
Shower Reconstruction in Larsoft

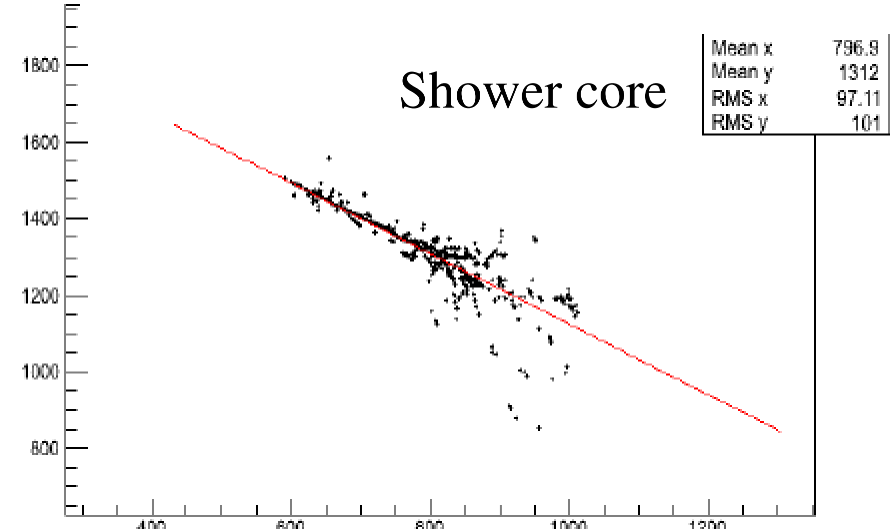
- Starting from clusters, shower reconstruction in larsoft takes place in two steps (modules).
  - ShowerAngleCluster – 2D shower reconstruction, produces refined 2D start point and 2D shower axis.
  - ShowerReco – Combines 2D showers in different views into 3D shower.
- Shower reconstruction challenges.
  - Showers are extended objects. Axis reconstruction can be pulled off by outliers.
  - Shower vertex reconstruction can be confused by backward-emitted photons.

Shower Reconstruction Recent Improvements and Ideas

Andrzej Szelc

- Use hits with higher than average charge to identify shower core.
  - Gives improved axis reconstruction.
- Improve vertex by rejecting hits (including backward hits) away from shower axis.
- Use track-like substructure (Hough lines) to help identify vertex.





Calorimetry

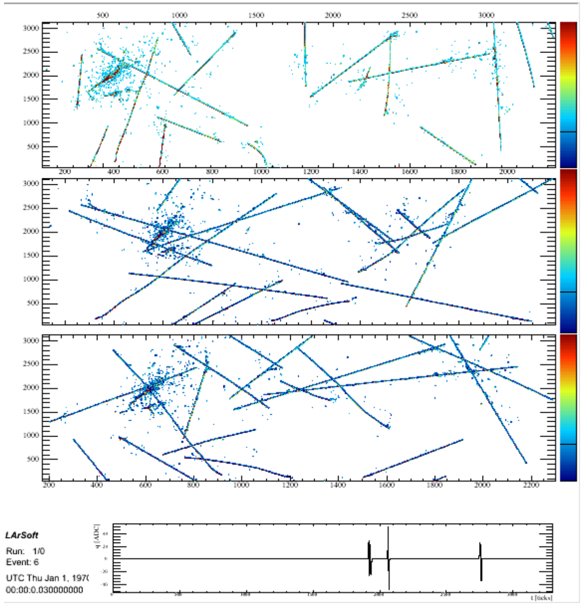
- Calorimetry has several applications in LAr tpc reconstruction.
  - Total energy of electromagnetic showers.
  - Electron/photon shower separation.
  - Track *dE/dx*.
- All of the above applications depend on the ability to convert collected charge Q to energy E.
  - Besides correcting for electron lifetime and particle angle, it is necessary to correct for charge quenching (recombination effect).
  - Our current understanding of the recombination effect is based on Birk's formula.
    - $dQ/dx = A\, dE/dx \,/\, [1 + (k/\varepsilon)\, dE/dx]$
    - where ε = electric field, and k = Birk's constant.

Birk's Formula Caveats

- Birk's formula only applies to charge depositions that are known to be from a single charged particle.
  - Not the case for electromagnetic showers.
- Accuracy of Birk's formula itself is ultimately in question.

Optical Reconstruction and Cosmic Rejection

- Because of their relatively long integration times, TPCs (especially TPCs on surface) have to deal with issue of cosmic rejection.
  - Optical systems.
  - Rejecting tracks that enter or exit (especially from top or bottom).
  - Rejecting tracks that cross spill (time) boundaries.
- Cosmic rejection has so far received relatively little attention.



Summary

- Early stage reconstruction (signal shaping and hit finding) are reasonably straightforward (but still needs work to get the details right).
- Pattern recognition is the hardest part of reconstruction.
  - Many unsolved problems.
  - It is not always obvious how to proceed (room for new ideas).
- Late stage reconstruction (esp. calorimetry).
  - Need to confront the issue of recombination effect, which is not that well understood theoretically.
  - Understand how to correct for recombination effect in showers.
- Cosmic rejection.