MPD Simulation and Reconstruction with GArSoft

Tom Junk DUNE Near Detector Discussion Meeting Tata Institute of Fundamental Research, Mumbai, India 28 February, 2020



The GArSoft Team

• Fermilab:

People have many roles, I may be missing some! Apologies if I am misrepresenting something.

- Faculty: Tom Junk: Picked up Brian Rebel's GArSoft in 2018 and adapted it to the ALICE chambers. Wrote TPC tracking code.
 Maintains software releases. Liaison to computing people.
- Faculty: Leo Bellantoni: Analysis trees, Coherent-pion analysis, dE/dx, Track-cluster associations, vertexing, backtracking (MC truth matching) organizes the to-do list.
- Postdoc: Tanaz Mohayai: G4 simulation and parameterized reconstruction. electron analysis.
- Faculty: Alan Bross and Jen Raaf: Group leaders who consult on issues



The GArSoft Team

- DESY
 - Postdoc Fellow: Eldwan Brianne: Calorimeter simulation and reconstruction. Geometry. Event Display.
- Minnesota
 - Faculty: Andy Furmanski: Parameterized reco and physics.
 Overlay samples.
- University of Colorado
 - Faculty: Alysia Marino
 - Undergrad: Susan Born -- did a great job with a Kshort analysis
 - Postdoc: Thomas Campbell machine-learning short track ID at the primary vertex. Left for industry in 2019.

🛠 Fermilab 🖂 💦

The GArSoft Team

- Indiana University
 - Faculty: Mark Messier and Jon Urheim
 - Postdoc: Erica Smith interested in pulse-shape and noise modeling
 - Postdoc: Ashley Back interested in vertexing and low-threshold tracking (picking up where Thomas Campbell left off)
- Kyiv National University
 - Faculty: Vladimir Aushev and Yuriy Onishchuk
 - Postdoc: Olga Gogota
 - Students: Igor Neporozhnii, Sasha Ostapenko, Valiriia Ivanets



Near Detector Software People

- Wichita State University
 - Faculty: Mathew Muether: ND Software Integration. Samples, computing, CDR prep, discussion of how to build the final product
- LArTPC (ArgonCube)
 - Kazu Terao and Gianluca Petrillo (SLAC Faculty)
 - James Sinclair (Bern Postdoc)
- SAND
 - Davide Sgalaberna and Guang Yang
 - Contact Sergio Bertolucci and Chang Kee Jung to learn more



Motivation for GArSoft

- Detailed detector modeling is important for physics sensitivity calculations.
- Edep-sim models use GEANT4 but list true energy deposits. Parameterized reco may be enough for a CDR
- We and our review committees want to ensure that the detector design meets the requirements and doesn't have significant blind spots induced by detector features.
- We have to show that projected performance is achievable
- For the future, we have to have software we can run to simulate and process real detector data and work within the full DUNE experiment.

Software Structure

- The DUNE Far Detector software is based on LArSoft <u>https://www.larsoft.org</u>, which uses the *art* framework <u>https://art.fnal.gov</u> and <u>https://cdcvs.fnal.gov/redmine/projects/art/wiki</u>
- DUNE Software and Computing goals are to share frameworks, dependent products, build systems and computing resources between the Near and Far detectors so that collaborators can easily work on both projects without facing separate learning curves. And joint analyses are made easier with shared tools.
- Much of the ND conceptual work is standalone however devleoping standalone is faster than developing in a framework



What's Available Now in GArSoft

- Detailed geometry: SPY magnet or 5-coil system, TPC+ECAL
- GEANT4 interface similar to LArSoft's
- Generators
 - GENIE for neutrino interactions and some BSM. Interfaces to LBNF and custom flux models.
 - Particle Gun
 - Text-File
 - Radiologicals (not tested)
 - Cosmic Rays (CRY) (tested long ago). We probably want CORSKA too but don't have it integrated yet.
 - Would like more (NEUT, NuWro, GiBUU, ...). There was a generator workshop at FNAL in January 2020 to help make generators easier to interface to simulation software stacks.



What's Available Now in GArSoft

- Constant magnetic field (help needed! slides later)
- Drift and diffusion of individual electrons
- IROC, OROC pad geometry from ALICE. Placeholder hole-filler geometry
- Spatial pad response functions parameterized from functions in the ALICE TDR



TPC channel positions

- 18 Sectors of IROC and OROC channels now using ALICE nominal geometry
- Rectangular array of pixels in a disk in the center
- Pixel size: 6mm x 6mm
 c.f. 4 mm x 7.5mm for inner
 pad rows.
- Total channels per side is now 339068.
- Total on both sides: 678136
- About 18% of channels are in the disk.





Figure 4.12: Pad layout of the ALICE TPC readout chambers. Distances from the beam axis are in mm.

Straw-person hole-filler pads: 6 x 6 mm²

🗱 Fermilab 🖂 💦

Readout Geometry



J. Alme et al., NIM A 622 (2010), 316-367

Fig. 9. Cross-section through a readout chamber showing the pad plane, the wire planes and the cover electrode.



Fig. 10. Wire geometries of the outer (left) and inner (right) readout chambers.



Pad Response Function

From the ALICE TDR: Charge Induction Response and IROC pad response





Q(x,y) function

Width ~1.2 mm, with a small tail



Figure 7.7: Pad response function for rectangular 4×7.5 mm² pads. Normalization is to unity.

Anode-plane wires are spaced 2.5 mm apart.

$$\mathsf{PRF}(x,y) = \int_{S} Q(x',y') \mathrm{d}S,$$



My Reproduction of the IROC PRF

From the ALICE TDR: Charge Induction Response and IROC pad response



Figure 7.6: Induced-charge distribution according to Ref. [11]. Normalization is arbitrary.



Figure 7.7: Pad response function for rectangular 4×7.5 mm² pads. Normalization is to unity. IROC Pad Response





My Predictions of IROC, IOROC, OOROC PRF's





What's Available in GArSoft

- Waveform simulation: Time ticks with induced charge are filled with induction fraction numbers summed over arriving electrons.
- Raw digits are zero-suppressed on readout. No noise is simulated.
- First step in reconstruction: Hit finding
 - Hits are pulses on pads over threshold
 - Some simple logic to keep hits from getting too long in drift time
 - Optimization welcome
- Nearby hits are then clustered together, with charge-weighted centroids
 - improves resolution, reduces number of points to fit. Hits make a scattered cloud around the tracks.
 - An issue close tracks may have hits that get mis-clustered.



TPC Cluster Residuals vs. Angle: IROC



Wide hit clustering: 4 cm (transverse) x 2 cm (longitudinal) search window



TPC Cluster Residuals vs. Angle: IOROC



Wide hit clustering: 4 cm (transverse) x 2 cm (longitudinal) search window



TPC Cluster Residuals vs angle: OOROC





Point Resolution Summary





Widths of fits of Gaussians to the abs val. residuals on previous pages

Resolution of OOROC for tracks 90° from the pad axis is still good – 1.6 mm. Optimization of TPC Clustering algorithm expected to improve all resolutions. Fiducial volume cuts (already in the 1 ton definition) ensure enough hits are on tracks to make a good momentum measurement.



What's Available in GArSoft

- Pattern Recognition
 - TPC clusters grouped into "Vector Hits", with at least 3 TPC Clusters, and up to 20 cm long. ("tracklets")
 - Vector hits are grouped into track candidates
 - Initial track parameters (six, including starting point) are estimated from TPC clusters.
 - New cheated pattern recognition available
- Track Fitting
 - TPC Clusters on list belonging to the pattern-recognition tracks are fit twice, once forwards and once backwards
 - To do let the track fitter find TPC clusters instead of relying on pattern recognition alone



What's Available in GArSoft

- Vertex Finding
 - Searches for nearby track ends
 - Linearly extrapolates tracks from their ends to fit a vertex
 - Lots of room for improvement here.
- Stitching of tracks that cross the cathode (get t₀ from the two track segments that need to line up)
- Track-ECAL extrapolation
- Analysis tree making. It's a lot easier to analyze a flat TTree than a framework file.



Tracking Performance: π^{\pm} and μ^{\pm}

Estimated using Leo B's sample of ν_{μ} events with the optimized LBNF FHC spectrum



Charged pion and muon tracking efficiency

Electrons are similar, but including them produces a kink at 20 MeV (bigger than the one that's there).

Low-energy electrons curl around – only partial efficiency for them

Low-energy pions and muons stop – have a track length cut of 20 TPC Clusters

Protons with P<150 MeV have very little KE and thus stop quickly – plot their efficiency vs. KE



Tracking Performance: Protons

Estimated using Leo B's sample of ν_{μ} events with the optimized LBNF FHC spectrum



Very short track efficiency overestimated near a dense primary vertex due to combinatorics – fake matches.

Efficiency should go to zero at KE=0.

Work in Progress – Optimizations will improve this



Tracking Performance: Muon Angles and Momenta

Work in Progress – Optimizations will improve these

🛠 Fermilab 🖂 🖊



~1 Degree angular resolution, and ~4.2% momentum resolution At 0.4 T B field. Should be 3.6% resolution at 0.5 T.

Tracking Performance: 4π **Coverage**



n.b. Charge modeling on the pads is naive – induced signals will be less for trains of charge arriving on the same pad over lengths of time



What's Available in GArSoft

- TWO (!) Event Displays
 - One modified from a NOvA display using ROOT's TView3D. Shows MCParticles, raw digts, hits, TPC Clusters, and tracks very well
 - One based on TEve. Very nice and more modern.
 - Both could use some work however.



A Simulated and Reconstructed ν_{e} Charged Current Event in the HPGTPC



 $\nu_{\rm e}$ + Ar \rightarrow e- + π + + p + n

Neutron with p = 0.23 GeV/c at the P.V. not shown



The Same Event in the LAr FD



CVN selects it as a $\nu_{\rm e}$ CC. Nick and Tingjun's Ereco = 1.914 GeV (2.0 was true)

⇔Fermilab DUNE

The 3D TEve-based Event Display



🛠 Fermilab 🖂 🖊

Eldwan Brianne wrote this display.

Reconstructing Vertices

- Initial vertex reco algorithm requires at least two tracks with nearby endpoints. Extrapolated helices used to fit vertex
- Some neutral-current events have no tracks at all
- Some charged-current events have just one track that starts in the middle of the detector
 - We can declare bare track endpoints as vertices but do not do so at the moment
 - Neutrals entering from outside the TPC can scatter and make fake vertices
 - Photon conversions make their own vertices
- Scattering of throughgoing tracks makes fake vertices with 2 tracks



Vertex Resolution



Thomas Campbell

 $oldsymbol{
u}_\mu$ CC Events with beam pointing along Z



Finding Short Tracks Near the Primary Vertex

5000

4000

3000

2000

1000

1.0

0.8

efficiency 90

0.2

0.0

-10

-5

- Thomas Campbell's work
- RANSAC line finding + Neural Network for p/pi separation and energy estimation
- Energy is dominantly from range for short tracks. Longer tracks use curvature. We need an algorithm that uses both optimally

33



Results

Kinetic Energy Residual for 3-15 MeV Protons

ò

12

14

10

true KE (MeV)

KE_measured - KE_true (MeV)





Cathode-Crossing Track Stitching Example Unstitched Stitched





MC true trajectories not shown. TPC Clusters not moved (assume t = 0 when drawing) Primary vertex lines up with true neutrino interaction better.



ECAL Model in GArSoft

- The default design is 60 layers of 2 mm Copper, and 5 mm of plastic scintillator
- First eight layers are high-granularity tiles (2.5 x 2.5 cm²)
- Fits inside magnet coils for now, but space-saving measures may require coils to be further in. Muon system would have to fit between coils
- Remaining 52 layer are 4 cm strips, crossed at each layer
- Energy resolution: $\frac{\sigma_E}{E} = \frac{5.6\%}{\sqrt{E}} \oplus 4\%$
- Angular resolution: $\sigma_{\theta} = \frac{6.29 \text{ deg}}{\sqrt{E}} \oplus 3.92 \text{ deg}$ (CDR numbers)
- Axially symmetric for now but under discussion
- Timing resolution at 700 ps/strip but needs proof of achievability.



Near-Term Project: Realistic Waveform Simulation

- Currently have: Electron arrival times with longitudinal and transverse diffusion
- Each electron is simulated separately, though only diffusion is applied.
- Lorentz angles not accounted for (B field is uniform and along E currently)
- Pulse shapes at the ROCs are nontrivial Need a Garfield model
- We also have no noise in the simulation need to add some
- We encouraged the Indiana group to work on this, but they are short on time.

Near-Term Project: Apply Field Map

- Vladimir Kashikhin (FNAL) gave us a field map spreadsheet (2D, r vs x), for the five-coil model and I wrote a GArSoft service to read it in and interpolate it in 3D.
- Need to use it in two places: Simulation and Reconstruction
 - Reconstruction requires changing the track parameter curvature to momentum and changing the Kalman filter code to use the field at each point.
 - Reco also needs to test hypotheses of which tracks and clusters are associated. Moving a track in time moves it in space, which samples a different field.
 - Simulation GEANT4 steps need to know about the field.
 Examples exist, such as in G4LBNF
- We encouraged the Kyiv group to work on this



from Vladimir Kashikhin

Five-Coil Field Map



Vladimir's X and Z are swapped wrt ours



from Vladimir Kashikhin

BRadial

Five-Coil Field Map

BAxial 0.03 0.6 0.02 0.58 0.01 0.56 0 0.54 -0.010.52 -0.02-0.030.5 0.48 0 0.5 2 z (m) $2 \times (m) = 0.5 \times 1^{-1.5} \times 1.5 \times 1^{-2} \times 10^{-1.5} \times 10^{-1.5}$ 1 1.5 1.5 0 0.5 1 1.5 *¥(m)* 0 0.5

Vladimir's X and Z are swapped wrt ours



Near-Term Project: Study and Improve TPC Cluster Resolution

- Current algorithms are quick and dirty
 - TPC Hits are formed from the zero-suppressed raw digits just by looking at over-threshold time periods
 - Some logic added to keep hits from becoming very long when a track is pointed at a pad (or a looper corkscrews towards a pad)
 - TPC Clusters really just start with any hit and keep adding until they collect charge from too far away. Charge centroid calculated.
 - ALICE code provides examples of how to do this better (but it is tied to their geometry. We need something that works for us in 3D with the hole-filler too)
- We encouraged the Kyiv group to work on this



Near-Term Project – Study dE/dx

- For now parameterized for CAF files using PEP-4 data as a guide for resolution. ALICE resolution is similar between 5 and 8%.
- Need to produce this from GEANT4-simulated and fully reconstructed MC. Needs realistic waveform simulation





Near-Term Project: Vertex and Track-Finding Improvements

- Vertex finder tends to find many vertices in a crowded environment
- Tracks are split near the primary vertex
 - Tracklets steal hits from each other
 - Tuneable parameters are set to break tracks easily into pieces
 - Do not want to miss a real vertex
 - But a scatter can look like a vertex
- Split tracks make for many combinations of possible vertices at the primary.
- Project: Select the best vertex and constrain tracks to it. Refit tracks after constraint.



Near-Term Project: Track Pattern Recognition Improvements

- Tracks are found by grouping vector hits together
- Vector hits are like ALICE's "seeds".
- ALICE uses the track Kalman filter to search for hits to add after an unambiguous seed has been found
- ALICE's problem is "easier" in that the primary vertex position is known (it's in the beampipe). Tracks are fit "outside-in". We do not have either luxury.
- Don't want to continue tracks through the primary vertex. My first attempt at pattern recognition did just this.
- Adding hits consistent with an extrapolation of the track path will do this – need to look at activity nearby to indicate a vertex.



Near-Term Project: Identify and Reconstruct Low-Energy Loopers

- Target physics: Conversions and Compton scatters
- Currently apply a momentum cut on tracks which is effective at reducing Comptons and Conversions
- Need to be able to pair up the e⁺ and e⁻ tracks as conversion candidates
- One electron may have much lower energy than the other
- Collider tracking algorithms are notoriously inefficient for lowenergy looping tracks.



Near-Term Project: Short Stubs at the Primary Vertex

- Thomas Campbell started a project to use machine learning techniques to find short stubs (mainly protons, but we expect some pions down there too) at the primary vertex
- RANSAC-based pattern recognition and stub counting
- The low threshold is one of the selling points of a HPGTPC
- Thomas moved on to an industry job, and we need a replacement
- Ashley Back (Iowa State) has expressed interest in taking this project on. He has NOvA experience.



ECAL Software To-Do

- Downselect on magnet design can help focus the studies on fewer geometries
- Need a crossed-strip algorithm to recover 4 x 4 cm² granularity from the long strips. Use time to associated strip hits to each other.
- Study/Improve cluster resolution
 - Lots of fragmentation of clusters in current algorithm
- PID with track/ECAL matching
 - Leo has developed a track extrapolator to the ECAL
 - Need to develop an analysis using isolation, nhits, energy matching, position, shower shape, etc.



Tips on Getting Started

- Lots of documentation linked on the DUNE at Work web page: <u>https://atwork.dunescience.org/</u>
- The wiki: <u>https://wiki.dunescience.org/wiki/Main_Page</u>
- How-to's
 <u>https://wiki.dunescience.org/wiki/DUNE_Computing/Computing</u>

 <u>How-To_Documentation</u>
- Setting up a local SL7 (or CentOS 7) system to use CVMFS: <u>https://wiki.dunescience.org/wiki/Setting up your own local S</u> <u>L7 system</u>
- GArSoft wiki how to run pre-installed release, or build and develop you own code <u>https://cdcvs.fnal.gov/redmine/projects/garsoft/wiki</u>



Analysis Updates

- Coherent pion analysis finished for CDR Leo Bellantoni
 - pre-track-stitch reco with spill timing \rightarrow broken tracks



Also: Leo's done good work on the Track-ECAL matching algorithm, MC particle to vertex associations, DST-maker and anatree improvements

- · ~56 days of "data" (with downtime) ⇒ get 330 signal events
- Reco backgrounds clearly reducible
- Reverse |t| cut to 0.06 < |t| < 0.30 GeV² \Rightarrow get 2331 RES & DIS @ 77% purity "for free"
- Reverse dE/dx cut, & 0.14 < |t| < 0.40 GeV² \Rightarrow get 6683 CCQE @ 72% purity "for free"
- 1 page, 1 figure writeup for CDR ready in Overleaf

48



Summary

- Lots of good progress recently on pad response functions, cathode-crossing tracks, track-ecal matching (no plots for latter)
- Many opportunities to improve modeling and reconstruction algorithms
 - Waveform modeling noise and signal shape
 - dE/dx simulation and reco
 - TPC cluster improvements and characterization
 - Pattern recognition in dense environments and low-energy loopers
 - Vertexing improvements
 - ECAL clustering improvments
 - use TPC/ECAL associations for PID, event slicing, analysis







New Geometry Files: MPD-only World Volume

- MPD_ECalOctagon_60I_UniformMagnet.gdml

Baseline MPD with HPgTPC and Pressure Vessel (PV), an ECAL in Octagonal shape (8 sides) with 60 layers in Barrel and Endcap of 2 mm Cu, 5 mm Scint and a uniform cylinder magnet of Aluminium of around 100 turns

- MPD_ECalDodecagon_80I_UniformMagnet.gdml

MPD with HPgTPC and PV, an ECAL in Dodecagonal shape (12 sides) with 80 layers in Barrel and 60 layers in the Endcap of 2 mm Cu, 5 mm Scint and a uniform cylinder magnet of Aluminium of around 100 turns

E. Brianne



New Geometry Files: MPD with LAr in ND Hall

- MPD_LAr_Hall_2CoilsMagnet_NoYoke.gdml

LArTPC and Baseline MPD with HPgTPC and PV, an ECAL in Octagonal shape (8 sides) with 60 layers in Barrel and Endcap of 2 mm Cu, 5 mm Scint and a 2 Helmholtz coil configuration w/o PRY

- MPD_LAr_Hall_4CoilsMagnet.gdml

LArTPC and Baseline MPD with HPgTPC and PV, an ECAL in Octagonal shape (8 sides) with 60 layers in Barrel and Endcap of 2 mm Cu, 5 mm Scint and a 4 Helmholtz coil configuration (removed central coil)

- MPD_LAr_Hall_5CoilsMagnet.gdml

LArTPC and Baseline MPD with HPgTPC and PV, an ECAL in Octagonal shape (8 sides) with 60 layers in Barrel and Endcap of 2 mm Cu, 5 mm Scint and a 5 Helmholtz coil configuration

E. Brianne



New Geometry Files: MPD with LAr in ND Hall

continued from previous slide

- MPD_LAr_Hall_SPYMagnet.gdml

LArTPC and Baseline MPD with HPgTPC and PV, an ECAL in Octagonal shape (8 sides) with 60 layers in Barrel and Endcap of 2 mm Cu, 5 mm Scint and the SPY Magnet configuration and a PRY according to the ND WS at DESY (https://indico.fnal.gov/event/21340/session/5/contribution/40/material/slides/0.pdf)

- MPD_LAr_Hall_UniformMagnet.gdml

LArTPC and Baseline MPD with HPgTPC and PV, an ECAL in Octagonal shape (8 sides) with 60 layers in Barrel and Endcap of 2 mm Cu, 5 mm Scint and a uniform cylindrical magnet of Aluminium of around 100t





New Geometry Files: Full ND Complex

- ND_Baseline_Hall.gdml
- Full ND Complex baseline in the hall
- ND_onlyMPD_Hall.gdml Only MPD baseline in the hall





Status: Track Stitching across the Cathode

- The issue: spill duration is $10 \mu sec$
- $v_{\text{drift}} \approx 3 \text{ cm}/\mu \text{sec}$, so displacements are up to 30 cm
- Getting the interaction time wrong means tracks are shifted along *X*.
- Tracks won't match up with the ECAL if their timing is wrong.
- If the B field is inhomogeneous, it's more than just a shift. More on this in later slides.
- Tracks that pass through the cathode plane have segments that are shifted in equal amounts in opposite directions

An Example

Purple: MC proton track

Red and blue: reco tracks on either side of the cathode.

Yellow: reco muon.

Vertices are yellow dots

blue: mc muon

dashed cyan: MC neutrino



🛠 Fermilab 🖂 🖊 🗄

Complications

- Tracks are associated to vertices.
- Shift a track in *X*, and we need to shift its vertices too
- A track endpoint may be associated to multiple vertices.
- Both track endpoints may be associated to vertices. Primary, secondary, delta rays, etc.
- Other tracks are associated to these vertices too.
- If you pick up a track segment and move it in *X*, you have to move everything it is attached to vertices, tracks, vertices ...
- The stitcher makes new collections of Tracks, Vertices, Trackloniz, and associations between them and TPCClusters

🛠 Fermilab 🕞 🖯

Unstitched







Magenta and cyan got stitched. Vertex moved close to edge of the drift volume. Maximum track length in X is still limited to size of the TPC.



Inhomogeneous Magnetic Field

- The relationship between curvature and momentum depends on the B field.
- Kalman fit can follow approximately helical tracks just fine
- But the curvature is defined in the (*Y*, *Z*) plane. Need to generalize this as the field doesn't point along *X* everywhere.
- Tracks will curl tighter where B is stronger, in corners near the coils.
- Displaced tracks will still curl tighter. Fit follows tighter curls, but relationship between curvature and momentum will be wrong.
- Extrapolation to ECAL will be wrong if the curvature is changing locally.



Inhomogeneous Magnetic Field

- Leo pointed out that E x B effects no longer vanish when there's a radial component of B.
- Drift paths now have a Lorentz angle proportional to IE x BI and the electron mobility (drift velocity).
- In LAr, the Lorentz angles can be neglected because the drift velocity is so slow.
- Shifting tracks, vertices, and TPC clusters along X due to timing shifts is only approximate if we're being careful, it's insufficient
- Need to know where in space a cluster on the endplate in (Y,Z) came from as a function of t



The dunetpc Dependency Tree (v08_18_00)



