Wire-Cell TPC Responses, Simulation, Signal Processing and Implications for Vertical Drift Designs

Brett Viren and Andrea Scarpelli

January 21, 2021

Topics

- Concepts in **responses** for **simulation** and **signal processing**.
- Requirements and progression of LArTPC detector response models.
- Review of performance of Wire-Cell Toolkit's implementations for wire detectors.
- Challenges for responses for **strips+holes** anodes and performance of Wire-Cell Toolkit with 50-L detector data.

$\mathcal{R}$: modeling LArTPC ionization response

For $x \in$ "*det*" (real detector) or "*sim*" (detector simulation):

- $\mathcal{S}_x$ drifted **ionization charge distribution** ("signal"),
- $\mathcal{R}_x$ **detector response** in anode to drifting electrons,
- $\mathcal{N}_x$ non-signal related "**noise**",
- $\mathcal{M}_x$ a **measurement** (eg ADC waveforms on channels).

Simulation is a **convolution** (with $\mathcal{R}_{det}$ or $\mathcal{R}_{sim}$):

$$\mathcal{M}_x = \mathcal{N}_x + \mathcal{R}_x \circledast \mathcal{S}_x$$

Signal processing is (mostly) a **deconvolution** with $\mathcal{R}_{sp}$ to get **reconstructed signal**:

$$\mathcal{S}'_x = F_{sp} \circledast \mathcal{R}_{sp}^{-1} \circledast \mathcal{M}_x$$

(more details in backups)

Requirements on responses

We then must **choose** $\mathcal{R}_{sim}$ and $\mathcal{R}_{sp}$ and wish to minimize the **per-event difference** between reconstructed and “true” ionization signal in the *sim*:

$$|\mathcal{S}'_{sim} - \mathcal{S}_{sim}|$$

and simultaneously minimize an **ensemble difference** between reconstructed signal over **similar event samples** from *det* and *sim*:

$$|\langle \mathcal{S}'_{det} \rangle - \langle \mathcal{S}'_{sim} \rangle|$$

This obviously implies we want:

$$\mathcal{R}_{det} \approx \mathcal{R}_{sim} \sim \mathcal{R}_{sp}$$

IOW, we want $\mathcal{R}_{sim}$ as **close to reality as computational power allows** and $\mathcal{R}_{sp}$ as **close to reality tempered by our limited basis of measurement** $\mathcal{M}$ (ie, channel-level info).

Historical Progress of Response Sophistication

$$\mathcal{R}^{1D} \rightarrow \mathcal{R}^{2D} \rightarrow \mathcal{R}^{2.5D} \xrightarrow{?} \mathcal{R}^{3D}$$

LArSoft → Wire-Cell Toolkit (with wires) → WCT (with strips+holes) → ???

1D response model: $\mathcal{R}^{1D}(x)$

- Response depends on 1D coordinate (drift direction).
  - Sim and SP assume current only in wire nearest to drifting electron.

Pros:

- Computationally fast and algorithmically easy.
- Still available for use in LArSoft SP and sim.
- For sim, strongly minimizes: $|\mathcal{S}'_{sim} - \mathcal{S}_{sim}| \approx 0$

Cons:

- For sim, the $|\mathcal{S}'_{sim} - \mathcal{S}_{sim}| \approx 0$ minimum is “too perfect”.
- For data, avg reco signal differences $|\langle\mathcal{S}'_{det}\rangle - \langle\mathcal{S}'_{sim}\rangle|$ are large.
  - The “long range induction” effects can not be ignored.
  - MicroBooNE data demonstrated this model is too simplistic.

Wire-Cell 2D response model: $\mathcal{R}^{2D}(x, \rho)$

- Response depends on 2D coords (drift + pitch directions).
  - May induce current on range of nearby wires (1 ± 10) or strips (1 ± 5).
  - $\mathcal{R}_{sim}$ varies w/in one *wire region* (10 sub-pitch bins).
  - Average over each wire region: $\langle \mathcal{R}^{2D}_{sim}(x, \rho) \rangle |_{\rho} \rightarrow \mathcal{R}^{2D}_{sp}(x)$

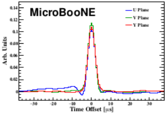
Pros:

- Well minimizes both $|\mathcal{S}'_{sim} - \mathcal{S}_{sim}|$ and $|\langle \mathcal{S}'_{det} \rangle - \langle \mathcal{S}'_{sim} \rangle|$.
- Validated, optimized implementation in Wire-Cell Toolkit.
  - Now established as default in most LArSoft uses.
- On average, works well on **some** non-2D geometries (eg wires).

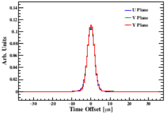
Cons:

- Field response calculations more difficult than 1D, but reasonable.
- Sim and sigproc algorithms more complex, somewhat slower than 1D.
- The more 3D the geometry ⇒ more imperfect is a 2D model.
  - Particularly, **strips + holes** stress the model.

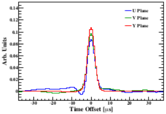
1D SP *vs* Wire-Cell 2D SP on MicroBooNE *det* data



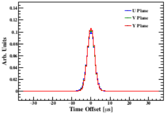
(a) 1D deconvolution, $5^\circ < \theta_{xz} < 15^\circ$.



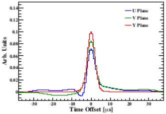
(b) 2D deconvolution, $5^\circ < \theta_{xz} < 15^\circ$.



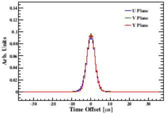
(c) 1D deconvolution, $15^\circ < \theta_{xz} < 30^\circ$.



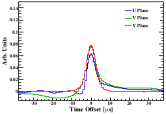
(d) 2D deconvolution, $15^\circ < \theta_{xz} < 30^\circ$.



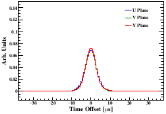
(e) 1D deconvolution, $30^\circ < \theta_{xz} < 50^\circ$.



(f) 2D deconvolution, $30^\circ < \theta_{xz} < 50^\circ$.



(g) 1D deconvolution, $50^\circ < \theta_{xz} < 70^\circ$.



(h) 2D deconvolution, $50^\circ < \theta_{xz} < 70^\circ$.

Plotted: average **reconstructed ionization signal** $\mathcal{S}'^{1D}_{det}$ and $\mathcal{S}'^{2D}_{det}$ *vs* sample time for ensemble of tracks in four angle bins.

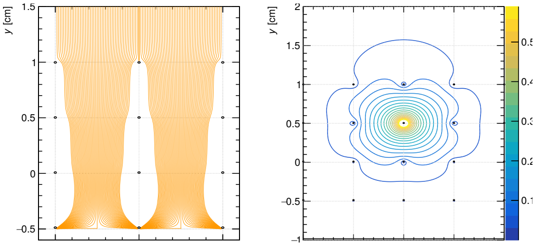
2D Wire-Cell signal processing is able to correctly recover identical **average** track reco signal $\mathcal{S}'^{2D}_{det}$ **independently from each wire plane**.

Exploit LArTPC technology for tomography!

Ionization Electron Signal Processing in Single Phsae LArTPCs II. Data/ Simulation Comparison and Performance in MicroBooNE
MicroBooNE Collaboration arXiv:1804.02583, JINST 13, P07007 (2018).

2D Field Response Calculations - Wires

- 2D **slice** across 3D geometry
  - plus some fictional alignment!
- Wires are **infinite, parallel and uniform**, and there are **no edge effects**.
- Drift paths in applied $\vec{E}$-field.
- Per-conductor *weighting* field.
- 126 drift paths per plane: 1 ± 10 wires, 6 “impact positions” per wire at $\frac{1}{10}^{th}$ pitch, exploit translation and mirror symmetries.



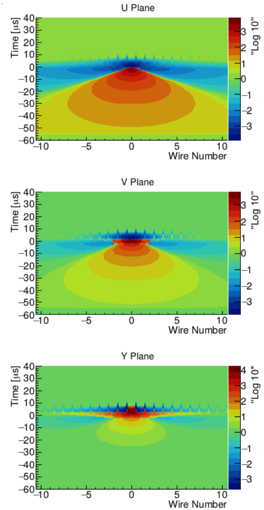
2D (x, ρ) drift paths and example U-wire weighting field for ProtoDUNE-SP calculated by drifires/Garfield++ (we are migrating from the venerable GARFIELD).

Instantaneous induced current: $I(t_i) = q\vec{W}(\vec{r}_i) \cdot \vec{v}(\vec{r}_i); \vec{r}_i = \vec{r}(t_i); t_i = t_0 + i\Delta t$

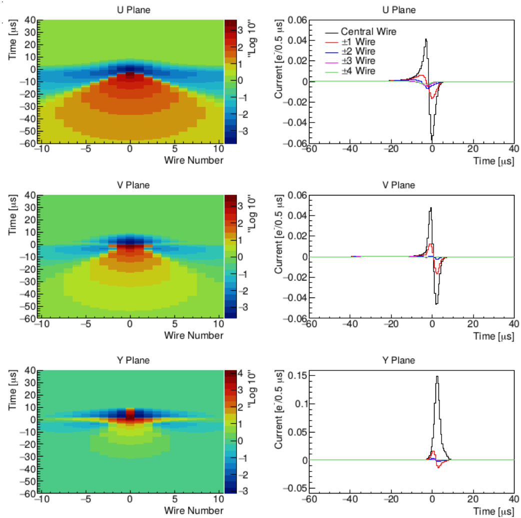
- $\vec{W}$: *weighting field* is $\vec{E}$-field with conductor-of-interest at 1V, all else at 0V.
- $\vec{v}$: drift velocity along path $\vec{r}$ calculated from LAr physics and solving for applied $\vec{E}$-field.
- q: an infinitesimal element from the distribution of drifted ionization electrons at the start of a drift path.

Wire-Cell 2D response for MicroBooNE

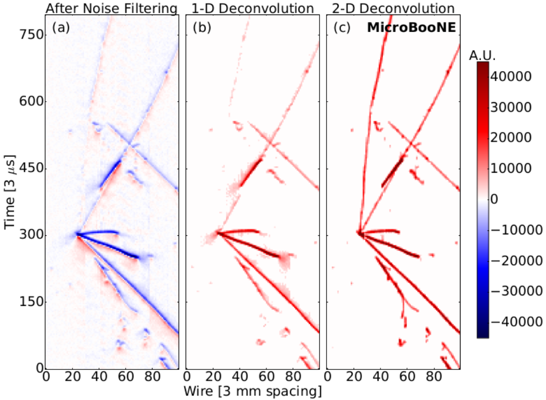
$\mathcal{R}^{2D}_{sim}(x, \rho)$
(drift vs impact position)



$\mathcal{R}^{2D}_{sp}(x)$
(drift vs wire region + select wire regions)



More Wire-Cell NF/SP performance on MicroBooNE



Signal processing on MicroBooNE **detector data** event:

- (a) WCT noise filtered,
- (b) 1D SP and
- (c) Wire-Cell 2D SP

Similar performance on ProtoDUNE-SP.

Ionization Electron Signal Processing in Single Phase LArTPCs I. Algorithm Description and Quantitative Evaluation with MicroBooNE Simulation MicroBooNE Collaboration arXiv:1802.08709, JINST 13, P07006 (2018).

Challenges of strips+holes for $\mathcal{R}$

Strips+holes strongly violate 2D model

- **non-parallel strips between the planes**
  - (a “feature” shared with wires)
- **non-uniform along their length**.

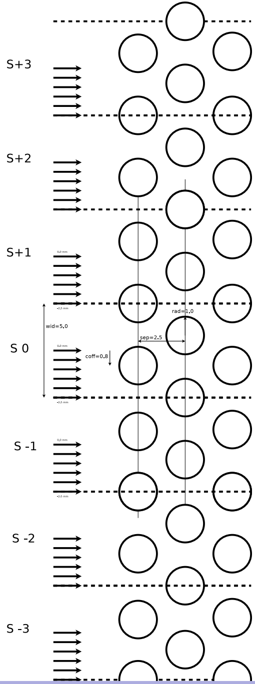
Extra challenges:

- ind/col hole patterns differ between strips-in-plane and strips-across-planes
- hole-pattern has some finite *repetition distance*
  - 50-L has a 2-hole repetition,
  - longer repetition for some 3-view designs.

Small bonus: fields drop faster with strips than wires

- 1 ± 10 wires $\rightarrow$ 1 ± 5 strips.

50-L detector collection stips+holes ⟶



2.5D model for strips+holes $\mathcal{R}^{2.5D} \rightarrow \mathcal{R}^{2D}_{sim,sp}$

“2.5D” trick

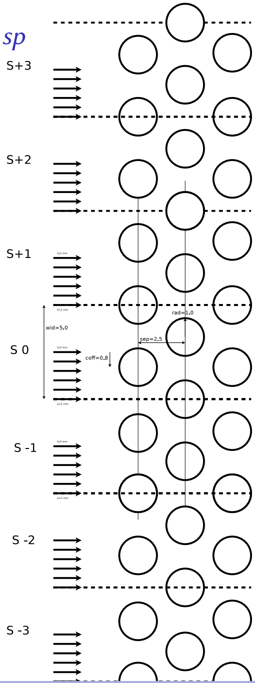
- Construct **slices** across strips spanning the *repetition distance*.
- Each slice defines one $\mathcal{R}^{2D}$ problem domain.
- Calculate per-slice $\mathcal{R}^{2D}_{slice}$,
- Take **average** over slices to get $\mathcal{R}^{2D}_{sim,sp}$

New problems for calculating $\mathcal{R}^{2D}_{slice}$:

- How best to define and combine slices?
- How many slices are needed?
- How to exploit symmetry to reduce calculation?
- **How wrong** is this on **average** and in **detail**?
  - tests ongoing

Chosen slices shown as vertical lines ⟶

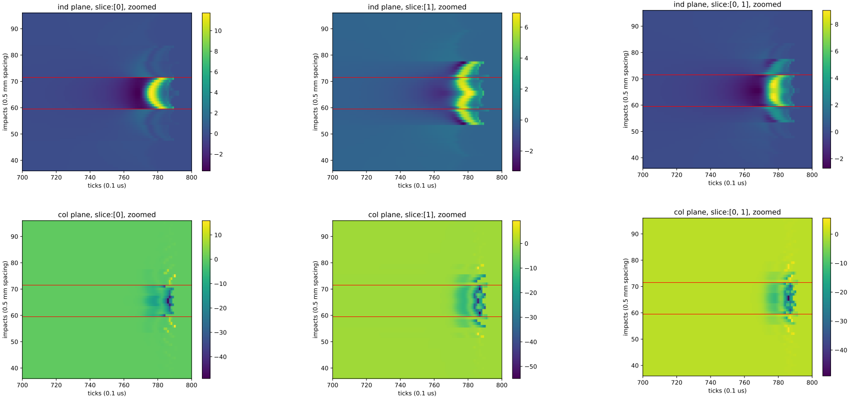
(details how we use GARFIELD to perform calculation in backups)



50-L test detector: $\mathcal{R}^{2.5D} \rightarrow \mathcal{R}^{2D}_{sim}$

The PCBro (PCB anode readout) package processes GARFIELD output to form the per-slice responses to produce WCT-compatible $\mathcal{R}^{2D}_{sim,sp}$ as linear color scale.

(induction plane, slices 0, 1 and average)
(collection plane, slices 0, 1 and average)



Slice 0. Slice 1. Slices $0.5 * (0 + 1)$.

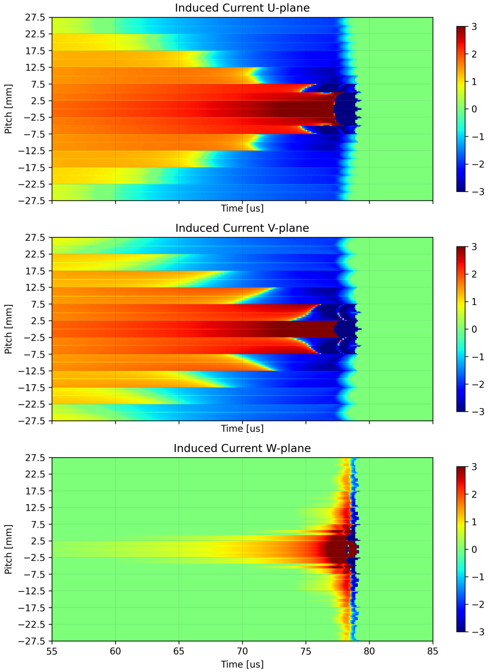
Next 3 slides

tl;dr: focus on bottom two plots.

To better view tails we use “±log 10” scale for 50-L’s $\mathcal{R}^{2D}_{sim}$.
Each slide shows a slice-average or a specific slice:

1. highlight slice 0
2. highlight slice 1
3. average over both slices

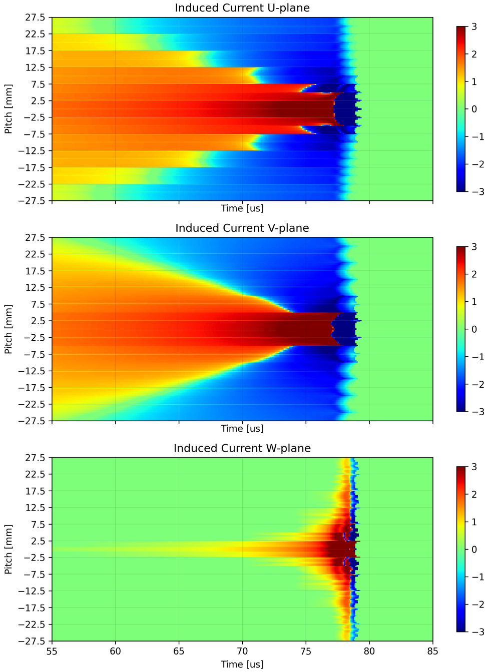
50-L detector only has 2-views: induction + collection.
The “U” and “V” labels indicate different forms for the **induction** plane info.
The “W” is always the **collection** plane info.



Slice 0

50-L $\mathcal{R}^{2D}_{sim}$

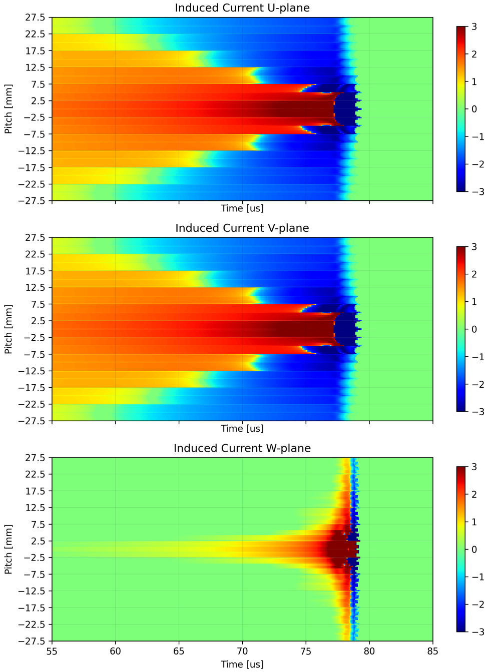
1. U is average over both induction slices
2. V is induction slice 0
3. W is collection slice 0



Slice 1

50-L $\mathcal{R}^{2D}_{sim}$

1. U is average over both induction slices
2. V is induction slice 1
3. W is collection slice 1



Average

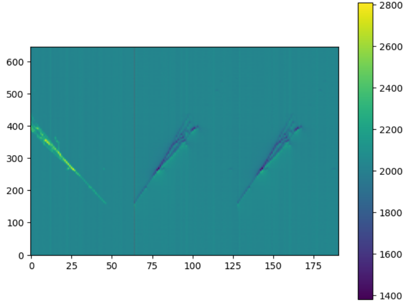
$50\text{-L}\ \mathcal{R}^{2D}_{sim}$

1. U is average over the two induction slices.
2. V is average over the two induction slices.
   - (the two are identical)
3. W is average over the two collection slices.

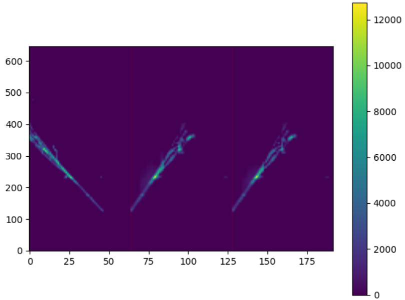
50-L raw data and WCT sigproc with $\mathcal{R}_{sp}^{2.5D}$

The PCBro package also provides a 50-L raw data decoder and hooks to run **Wire-Cell Toolkit signal processing** on 50-L data.

50-L raw data event
PCBro data decoder



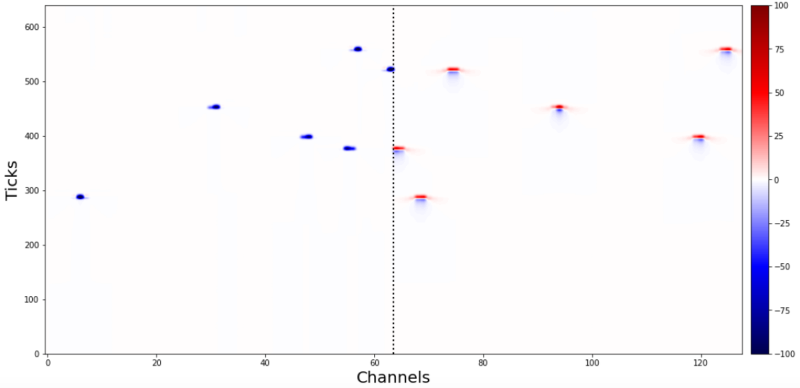
Same event after **Wire-Cell signal processing**



Note, the "**double induction**" plane is merely duplicate to fit nominal WCT assumptions of 3 planes. The PCBro package uses this "extra" plane to test different field responses in the same job. In production processing, we need not waste CPU on the duplication.

50-L WCT simulation of ^{39}Ar "blips"

Andrea Scarpelli
collection plane | induction plane



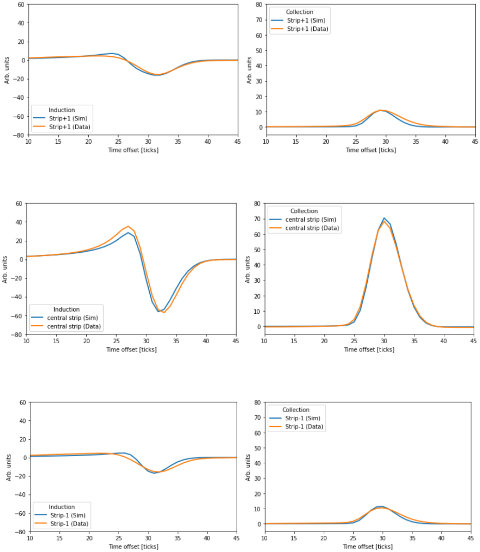
Samples ^{39}Ar energy spectrum scanned from arXiv:1705.05726v1.

50-L data / WCT sim comparison - ^{39}Ar “blips”

Andrea Scarpelli and Serhan Tufanli

average raw waveforms

50-L data (orange) and Wire-Cell 2D sim (blue) using a 2-slice $\mathcal{R}^{2.5D}$. Separated by induction (left) and collection (right) planes and for central strip (middle) and central ±1 strips (top/bottom).



3-view strips+holes

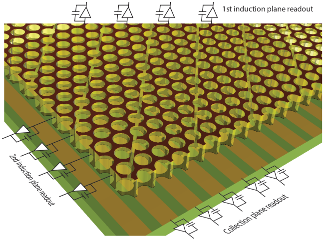
At least 3 views are needed to well exploit the tomographic power of LArTPC!

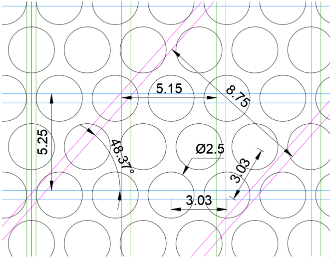
Various design, installation and **response** challenges for a 3-view **strips+holes** detector.

One possible design adds “diagonal” strip with “skewed” hole pattern. ⟶

- $\mathcal{R}$ may require $N > 2$ slice average: “2.7D … 2.9D”
  - Have scheme to produce **data-driven optimal weighting** of slices. Non-trivial, but there if we need it.
- Strip angle is such that **no slice** goes only through hole diameters.
  - Further biasing of the “2.x D” approach?
- Even with 3D, prefer to maximize regularity of patterns.

Design wish: 3-view isosceles or hexagonal strip+hole pattern. A novel hexagonal design by Bo looks nice!





From Bo’s presentation at the recent DUNE collab call

Summary and some next steps

- Wire-Cell's use of $\mathcal{R}^{2D}_{sim,sp}$ improves over $\mathcal{R}^{1D}$ as demonstrated with MicroBooNE and ProtoDUNE.
- The $\mathcal{R}^{2.5D}$ trick brings Wire-Cell support to 50-L strips+holes det.
  - Eyeball SP event display and average raw waveform tests look okay.
  - These **average** test metrics may hide some variational problems.
- Unclear (yet) "2.5D" trick is enough esp. for 3-view designs.
  - Require WCT SP/sim and test detector data to confirm.
- The more regular the strip+hole pattern the better!
  - **isosceles/hexagonal 3-plane** designs for vert. drift det, please!
  - a very new design already in this direction!
- **Precision tests** to check for correct **variations** *vs* position even more important for strips+holes.

In general: we will continue to improve support in Wire-Cell Toolkit for strip+hole test detectors and for the eventual DUNE VD module(s)!

$\mathcal{FIN}$

backups

Conceptual LArTPC Simulation

Real detector (and its simulation) produces an event via a **convolution** of:

$\mathcal{S}$ an **ionization charge distribution** (“signal”) with

$\mathcal{R}$ a **detector response** to drifting electrons, plus

$\mathcal{N}$ all non-signal related “**noise**”, producing

$\mathcal{M}$ a **measurement** (eg ADC waveforms on channels).

$$\mathcal{M}_x = \mathcal{N}_x + \mathcal{R}_x \circledast \mathcal{S}_x$$

With $x =$ “*det*” (real detector) or “*sim*” (detector simulation).

knowns $\mathcal{M}_{det}$, $\mathcal{M}_{sim}$, $\mathcal{S}_{sim}$ and $\mathcal{N}_{sim}$ (modeled), $\mathcal{R}_{sim}$ (but imperfect).

unknowns $\mathcal{R}_{det}$, $\mathcal{S}_{det}$ and $\mathcal{N}_{det}$.

Conceptual LArTPC Noise Filtering

Noise filtering is a transformation F_{nf} on the measurement:

$$\mathcal{M}_x \rightarrow \mathcal{M}'_x = F_{nf}(\mathcal{M}_x)$$

designed to strongly remove **excess** or **external** noise and potentially reduce **inherent noise** leaving residual noise n_x:

$$F_{nf}(\mathcal{N}_x) \rightarrow n_x \ll \mathcal{N}_x$$

while attempting to leave the signal term approximately invariant:

$$F_{nf}(\mathcal{R}_x \circledast \mathcal{S}_x) \approx \mathcal{R}_x \circledast \mathcal{S}_x$$

Conceptual LArTPC Signal Processing

Signal processing attempts to recover a good approximation of $\mathcal{S}_x$ from $\mathcal{M}_x$.

It uses a **deconvolution** of the measure with a response and with added filters F_{sp} to further suppress the residual noise term.

$$\mathcal{S}'_x = F_{sp} \circledast \mathcal{R}_{sp}^{-1} \circledast \mathcal{M}_x, x \in \{sim, det\}$$

We may not use the detailed $\mathcal{R}_{sim}$ here as it is in terms of the detailed, inaccessible "true" signal coordinates so we use an **average** $\mathcal{R}_{sp}$ (ie, per-channel vs sample time matching $\mathcal{M}_x$).

Note: for **induction channels**, $\mathcal{R}_{sp}^{-1}$ diverges at DC, thus **amplifies residual, low-frequency noise**. To counter, F_{sp} includes special algorithmic high-pass "filters" called *signal region of interest* (signal-ROI) and *local baseline correction*.

Conceptual LArTPC Responses

We then must **choose** $\mathcal{R}_{sim}$ and $\mathcal{R}_{sp}$ and wish to minimize the **per-event difference** between SP and “true” ionization signal in the *sim*:

$$|\mathcal{S}'_{sim} - \mathcal{S}_{sim}|$$

and simultaneously minimize an **ensemble difference** between SP signal over similar samples from *det* and *sim*:

$$|\langle \mathcal{S}'_{det} \rangle - \langle \mathcal{S}'_{sim} \rangle|$$

This obviously implies we want:

$$\mathcal{R}_{det} \approx \mathcal{R}_{sim} \sim \mathcal{R}_{sp}$$

IOW, we want $\mathcal{R}_{sim}$ as **close to reality as computational power allows** and $\mathcal{R}_{sp}$ as **close to reality tempered by our limited basis of measurement** $\mathcal{M}$ (ie, channel-level info).

The Test Metric Caveat

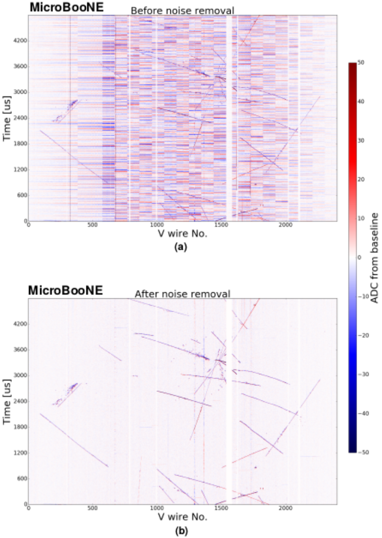
So far we use metrics sensitive to **response averages over space**.

- Demonstrates WCT $\mathcal{R}^{2D}$ **good on average** for SP(*sim*) $\approx$ SP(*det*).
  - Does control for **3D direction**, important variation for SP efficiency!
- May not be sensitive to **imperfect detailed variations** in $\mathcal{R}^{2D}_{sim,sp}$
  - Eg, is there SP(*sim*) $\neq$ SP(*det*) bias/resolution at specific locations?
  - Particularly strip+holes have large variations along strip direction.
    - ★ (will show)
- Examples of more **precise metrics** to apply in future *det* vs *sim*:
  - Signal matching between planes with ^{39}Ar or other “blips”.
  - Detailed comparison of dE/dX with tracks from full 3D reco.

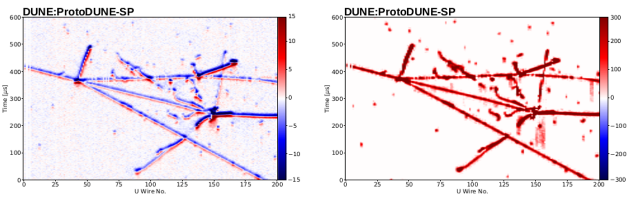
More Wire-Cell performance on MicroBooNE: NF

Wire-Cell software **noise filter** applied to MicroBooNE data event.

Noise Characterization and Filtering in the MicroBooNE Liquid Argon TPC MicroBooNE Collaboration arXiv:1705.07341, JINST 12 P08003 (2017).



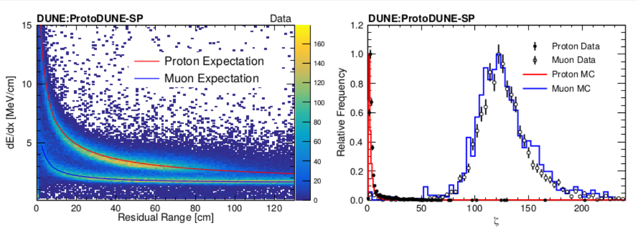
Wire-Cell signal processing on ProtoDUNE-SP data



(a) After Noise Filtering

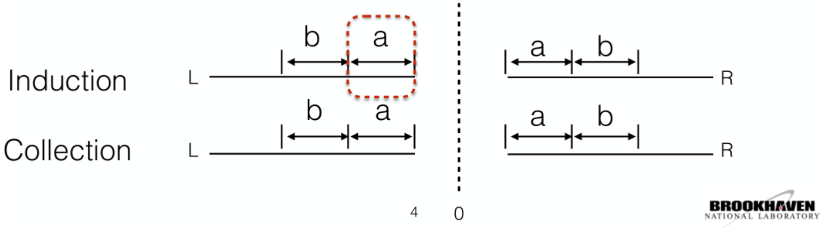
(b) After Deconvolution

Wire-Cell simulation/data comparison dE/dX for μ and P on PDSP



First results on ProtoDUNE-SP liquid argon time projection chamber performance from a beam test at the CERN Neutrino Platform. arXiv:2007.06722, JINST 15 (2020) P12004.

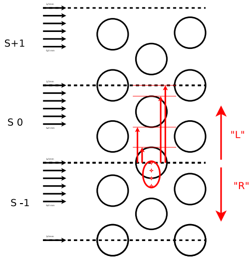
The 2.5D trick applied to GARFIELD



GARFIELD setup for 2D single-hole (Yichen Li)

Clever work-around to limitations of 2D and GARFIELD:

- **Single-hole**, 2D geometry with strips constructed as array of hundreds of "**micro wire**" sensing conductors
  - the L/R and A/B blocks work around some GARFIELD technical limits.
- Manual labor intensive post processing
  - Catalog maps of drift paths on slice to single-hole geometry.
  - Catalog micro wire selection criteria for each drift path.
  - Longer the hole pattern repetition distance ⇒ more the effort.



50-L data/sim checks t.b.d.

Other sources:

- 1 MeV e- Bismuth source “blips”.
- MIP tracks in different direction bins.

Address “Test Metric Caveat” with more precise *det vs sim* comparisons:

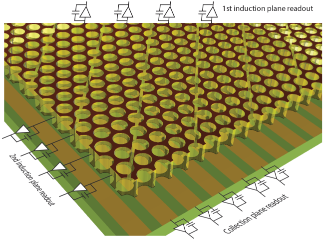
- SP ind/col ratio for “blips”, ideally $= 1.0$
- Invariant values (eg, raw ind waveform integral $= 0.0$)
- Raw waveforms from “blips” as $f(\rho)$?

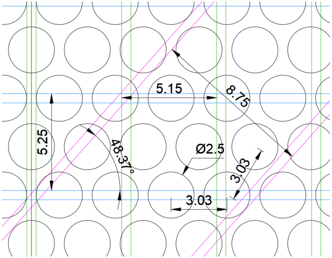
Skewed hole pattern: 2.5D? 2.7D?

50-L: **average** tests look good for $\mathcal{R}_{sim}^{2.5D}$, **eyeballing** SP event display looks good for $\mathcal{R}_{sp}^{2.5D}$.
But, 50-L has fairly **regular hole pattern**.
New worry: 3-view's "diagonal" strip with "skewed" hole pattern. $\longrightarrow$

- May need $N > 2$ slice average: "2.7D … 2.9D"
  - Have scheme to produce **data-driven optimal weighting** of slices. Non-trivial, but there if we need it.
- Strip angle such that **no slice** goes only through hole diameters.
  - Further biasing of the "2.x D" approach?
- Even with 3D, prefer to maximize regularity of patterns.

∴ **Design wish**: rectangular 2-plane or isosceles/hexagonal 3-plane!





From Bo's presentation at the recent DUNE collab call

So, why not $\mathcal{R}^{3D}$?

Some work in 3D exists.

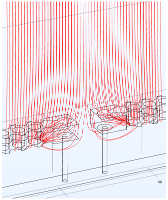
- Mostly for "near by" fields or "far fields" but for **wires**.
- Far field 2D calculation takes **minutes-hours**, 3D calculations take **hours-days** (using BEM, effectively impossible with FEM).
- Strips+holes need finer meshing ⇒ more processing (than wires).
- Longer hole repetition distance ⇒ more processing (than 50-L).

And, given $\mathcal{R}^{3D}$ sim must contend with an explosion of data.

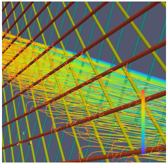
- $\mathcal{R}_{sim}$ **drift paths per plane** (some estimate/guesses)
  - 2D wires: 126.
  - 2.5D strips: O(100) for 50L, O(1,000) for skewed hole pattern
    - but at least results in "standard" $\mathcal{R}^{2D}_{sim,sp}$!
  - 3D: O(50,000), and worse: **old simulation must be thrown out**.

3D simulation

- Same concepts as 2D sim, but need all new algorithms/code.
- 2D sim exploits 10-way interlacing across common, wire-relative impact positions in order to use 2D FFT + 10-way sum for fast convolution.
  - How to even apply this trick in 3D? More variety along the strip direction will increase interlacing ⇒ harder/slower calculation.



From Bo's talk again



LARF