Status report on

DNNROI sigproc

Hokyeong Nam
Chung-Ang University

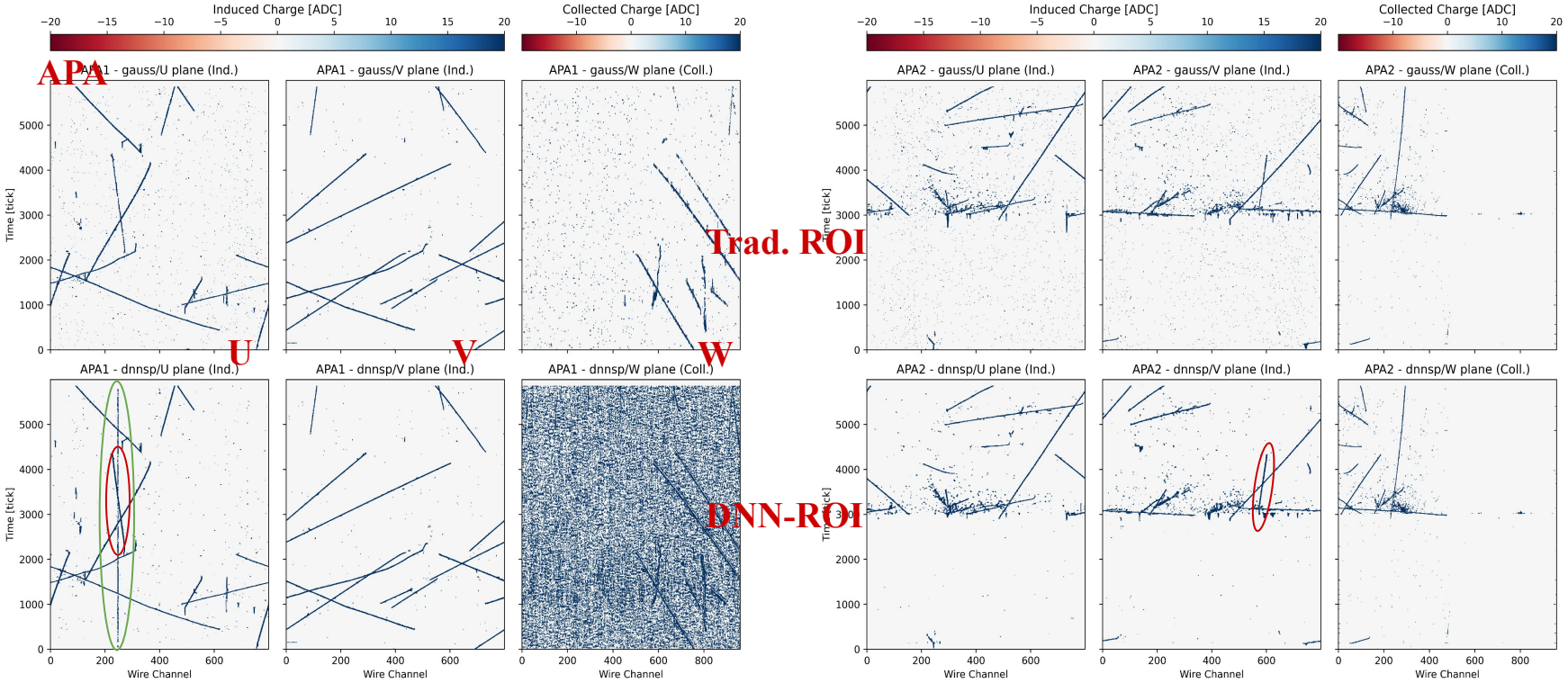
Outline

- **PDHD**
  - **Real Data (w/ MobileUNetV3, rebin = 10, Truth_th = 100)**
  - **Single Track Evaluation (WC standalone)**
  - **1D Waveform Study (WC standalone)**

- PDVD
  - Parameter List for Simulation

- Memory Profiling

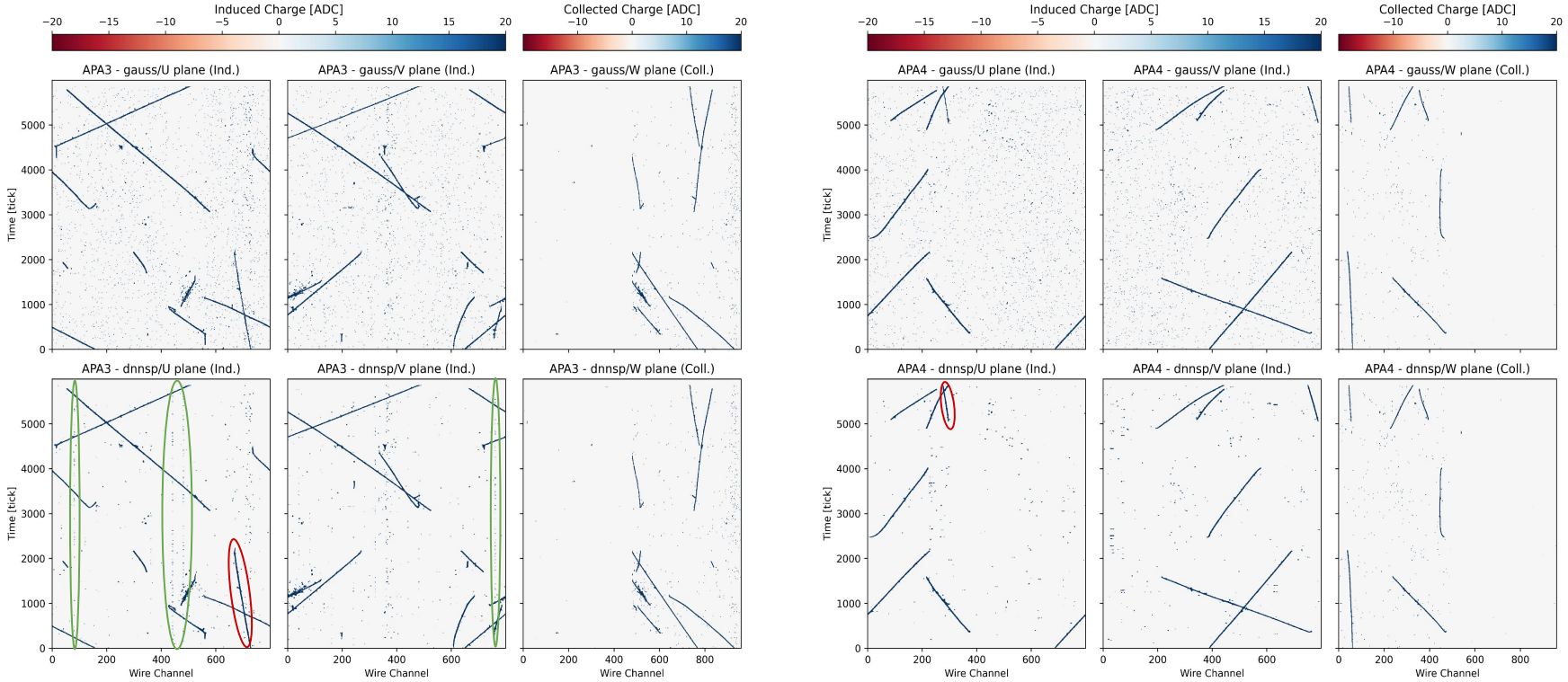
PDHD Real Data - Results w/ baseline model



Run 026763-225 (Better, Worse, Similar, Discuss later)

- DNN-ROI doesn't work for the W plane on APA 1

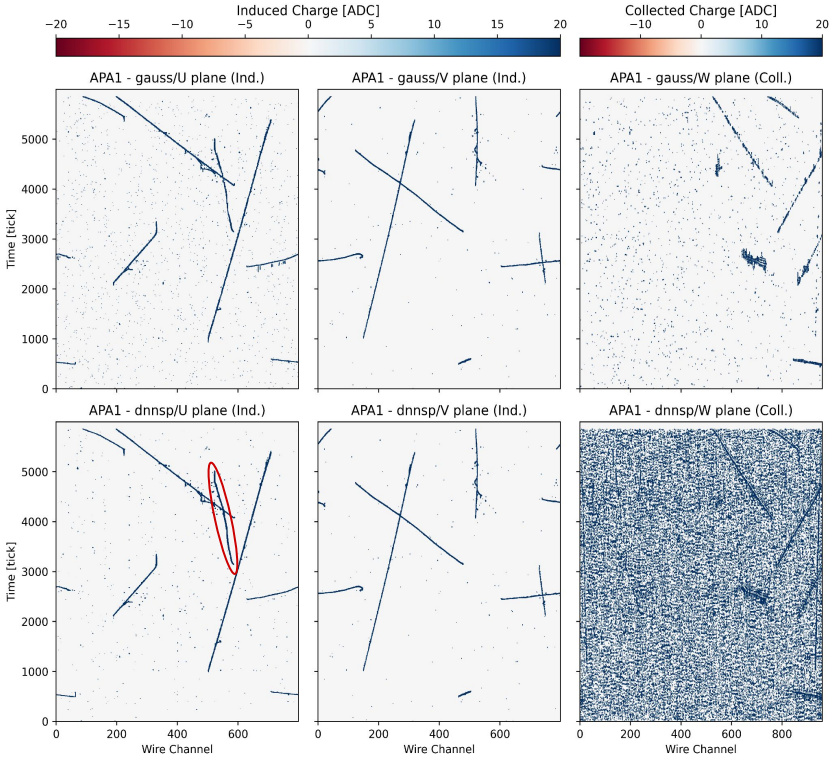
PDHD Real Data - Results w/ baseline model

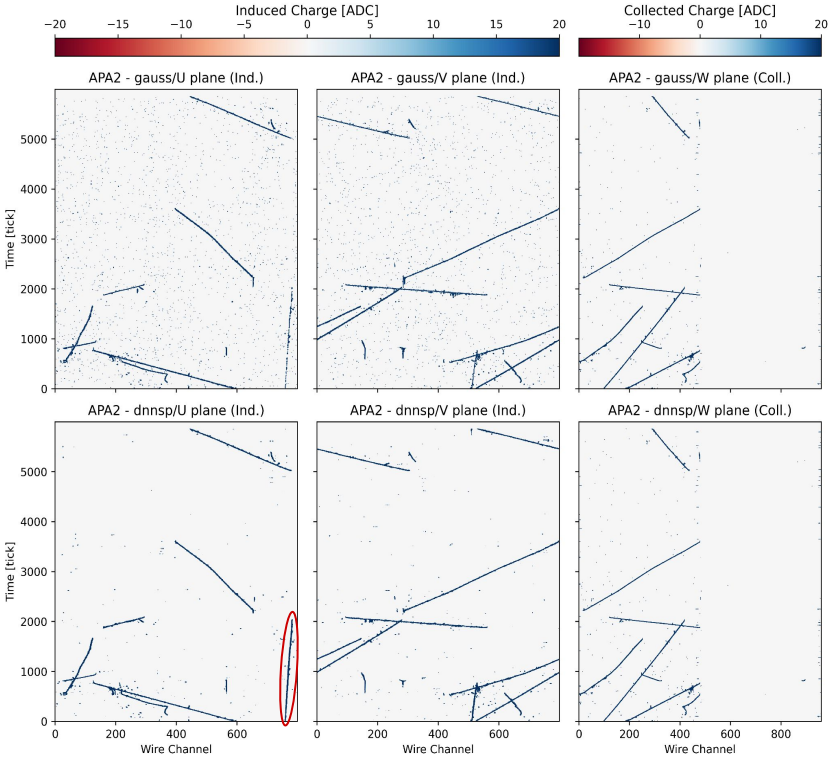


Run 026763-225 (Better, Worse, Similar, Discuss later)

- For APA3, discrete vertical noise for 3 channels consistently observed

PDHD Real Data - Results w/ baseline model

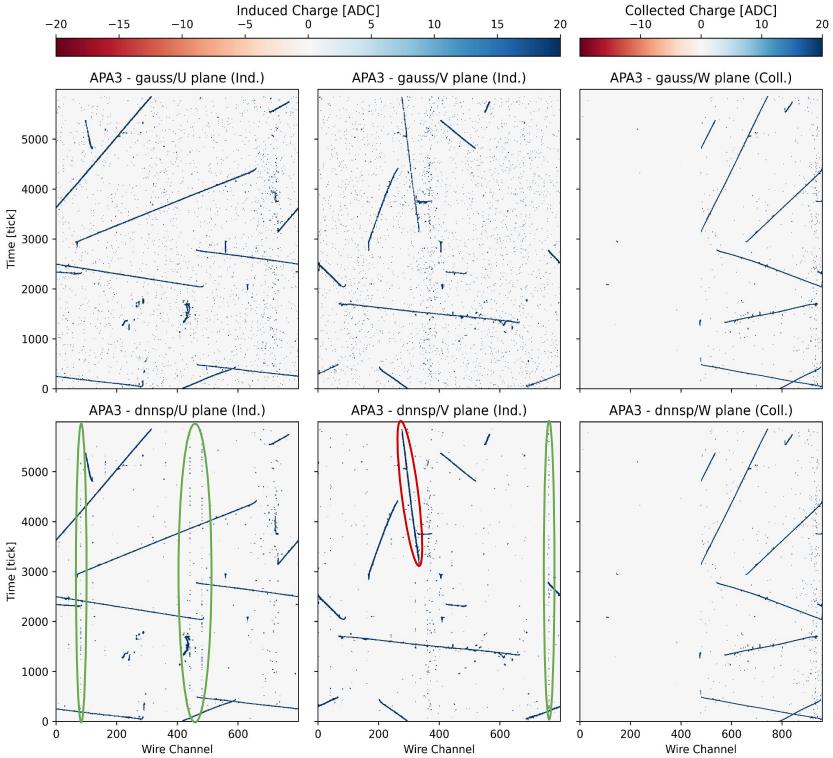


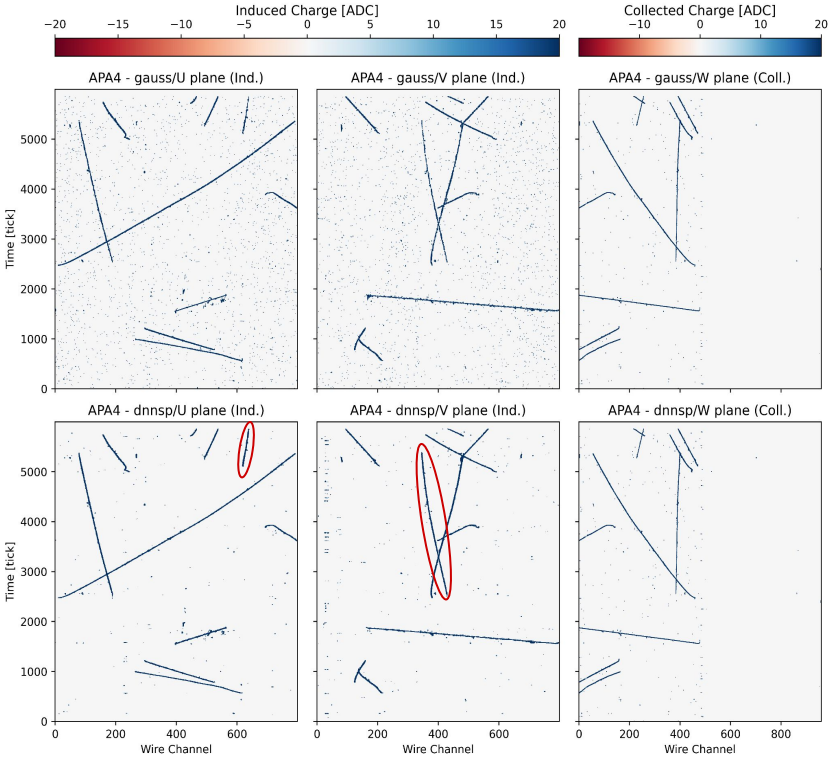


Run 027673-1 (Better, Worse, Similar, Discuss later)

- TBD

PDHD Real Data - Results w/ baseline model

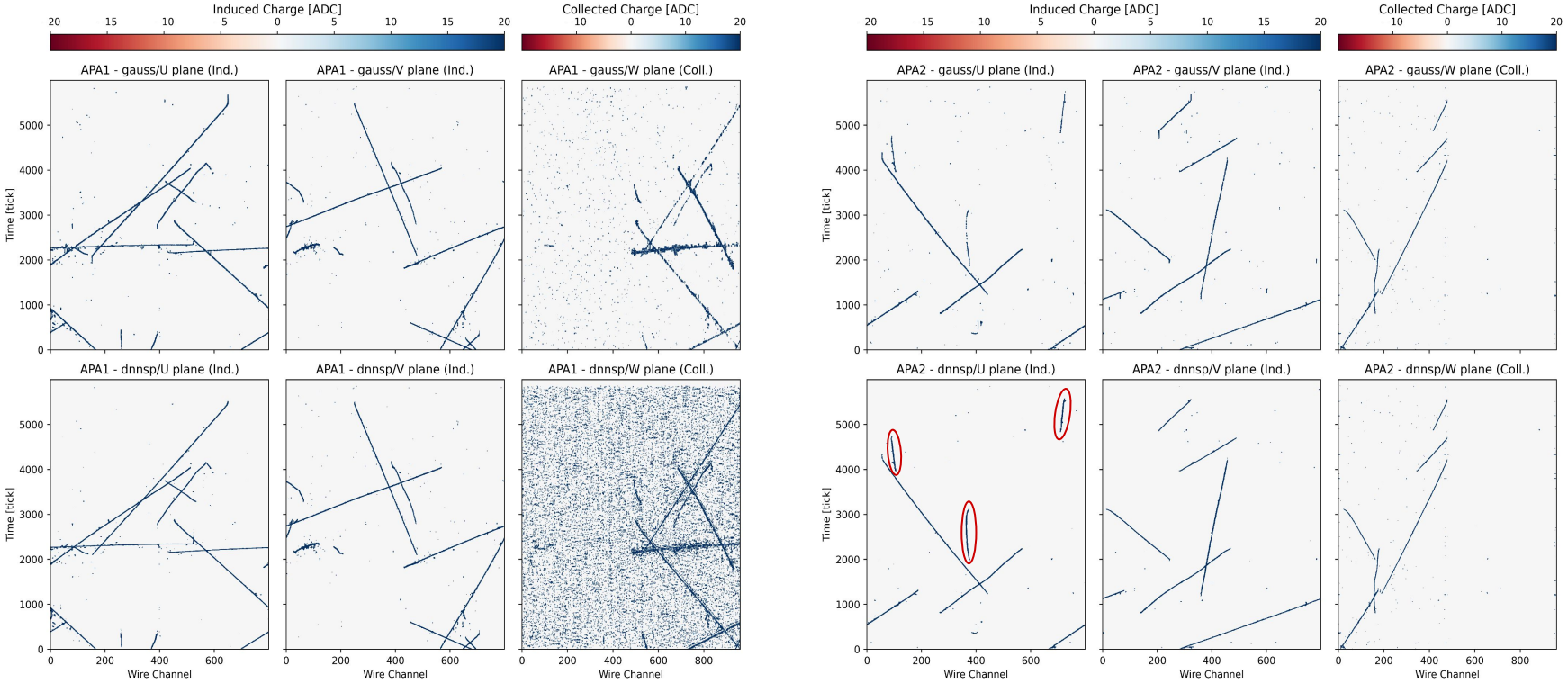




Run 027673-1 (Better, Worse, Similar, Discuss later)

- TBD

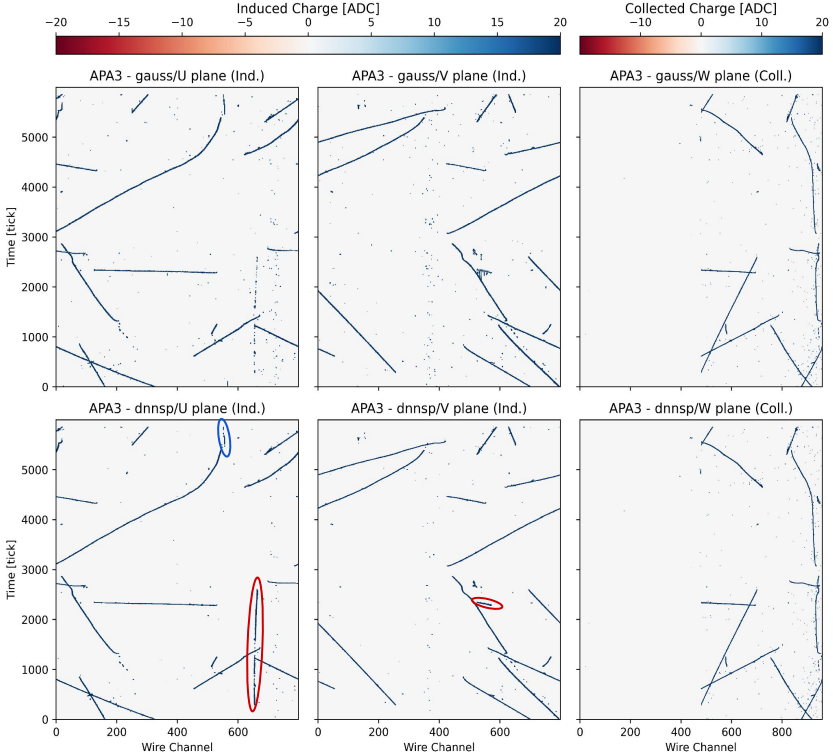
PDHD Real Data - Results w/ baseline model

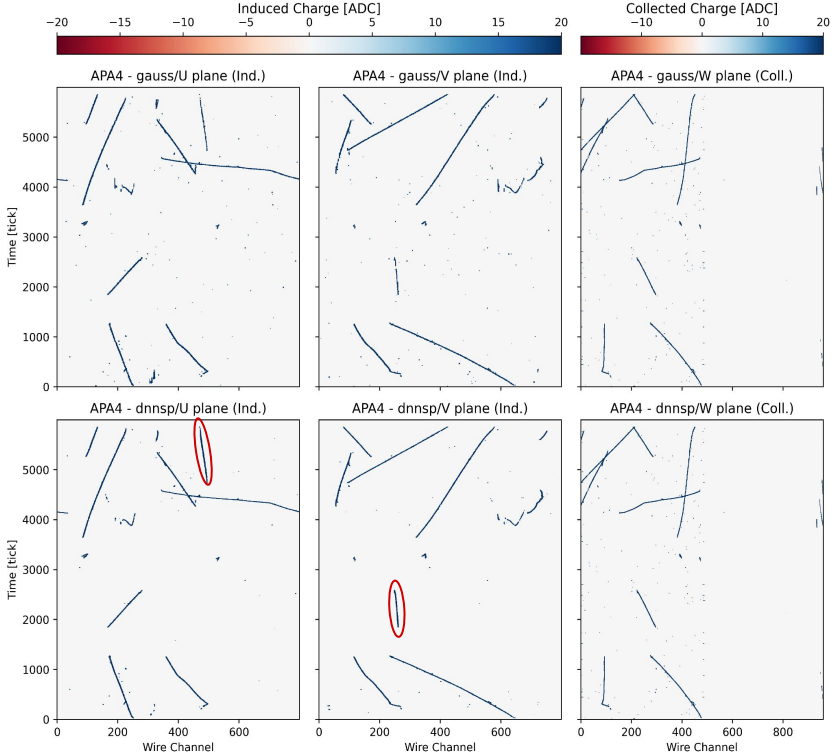


Run 028588-4562 (Better, Worse, Similar, Discuss later)

- TBD

PDHD Real Data - Results w/ baseline model

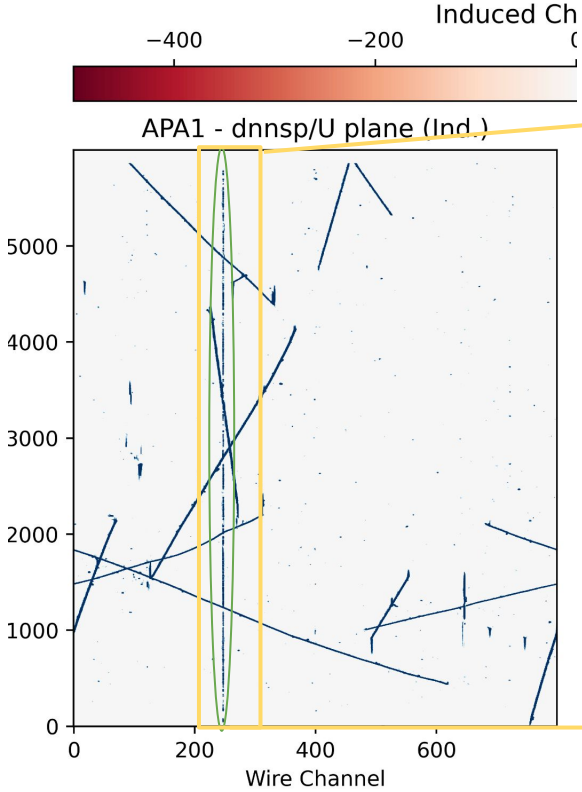




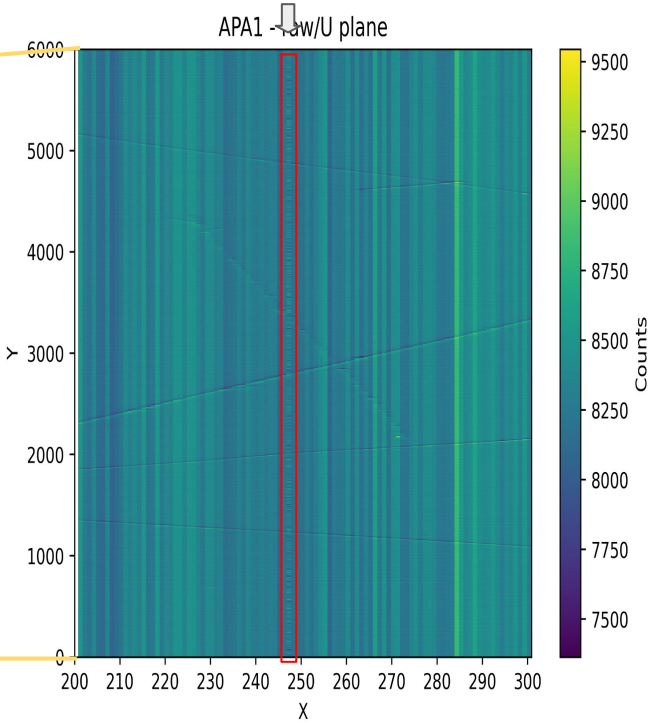
Run 028588-4562 (Better, Worse, Similar, Discuss later)

- For this event, noises are disappeared

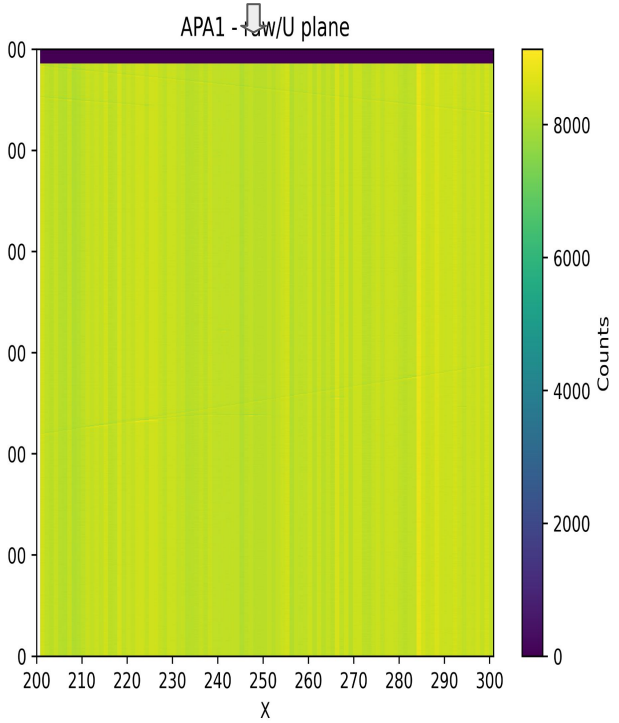
APA 1 noisy channel



Noisy



Not Noisy

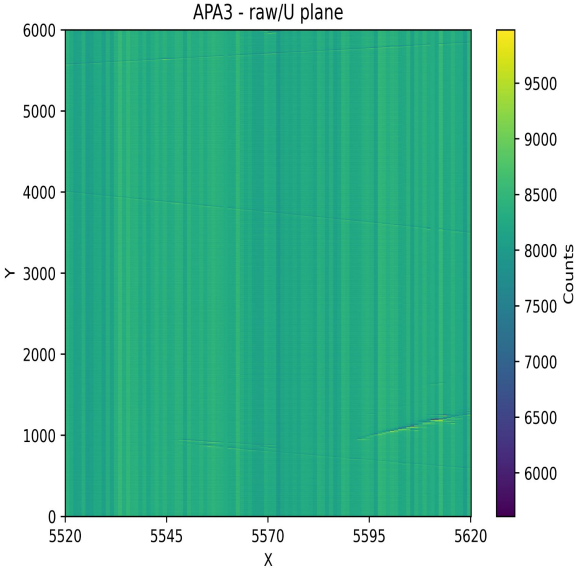


❑ Run 026763-225

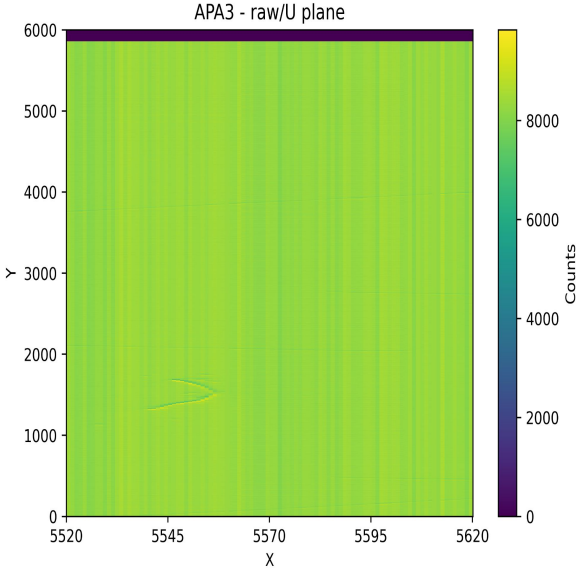
❑ Run 027673-1

- For the APA 1, a ghost track appeared on a particular noisy channel

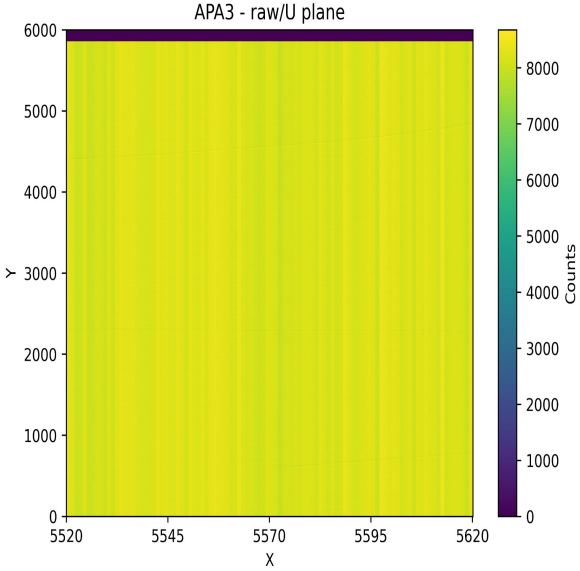
APA 3 noisy channel



❑ Run 026763-225

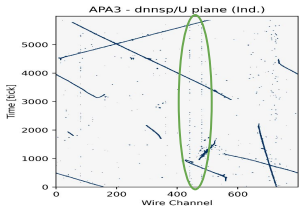


❑ Run 027673-1



❑ Run 028588-4562

- Not like the APA 1, no noisy channel is observed for the APA 3
- Reconstructed false signals in this case seems just “noise” rather than ghost tracks



Single Track Evaluation - Phase Space

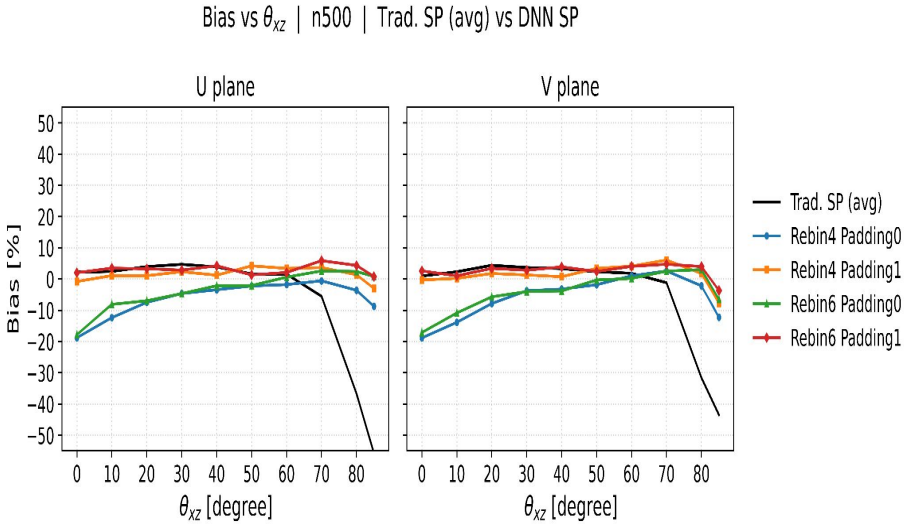
ML training

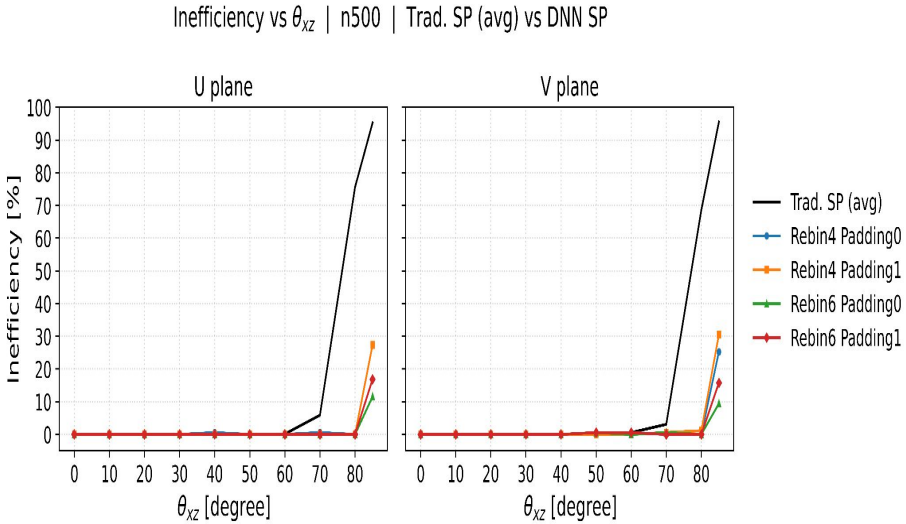
- Rebining: 4 or 6
- Truth_threshold: 100 or higher
- Padding: factor depends on the options above & which side we want to do?
- MP_tick_resolution: 10
- Model: per plane or same model

ML inferencing

- tick_per_slice: corresponds to Rebining in ML training
- MP_tick_resolution: 10
- Padding: for future?
- Chunking: not used for MobileUNetV3 model

Single Track Evaluation - WC standalone

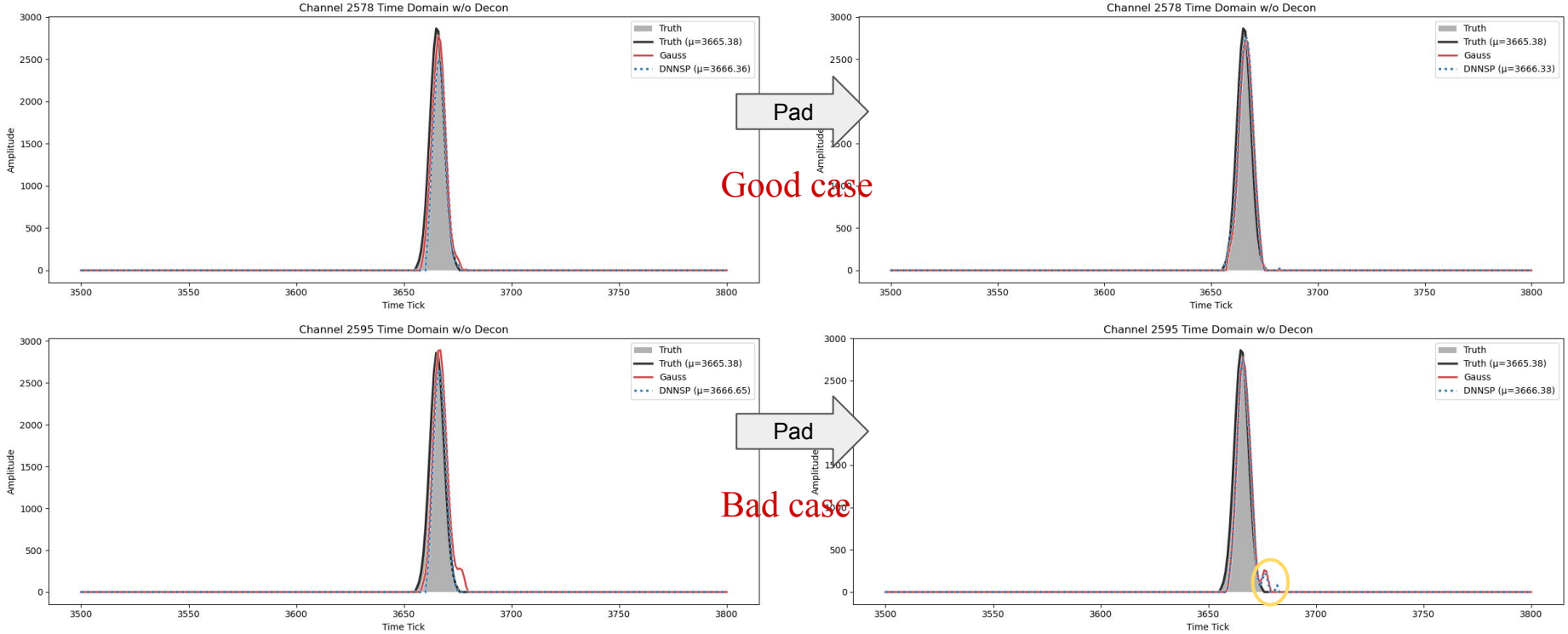




- Implemented padding strategy successfully reduced bias in isochronous cases
- The best combination from this evaluation:

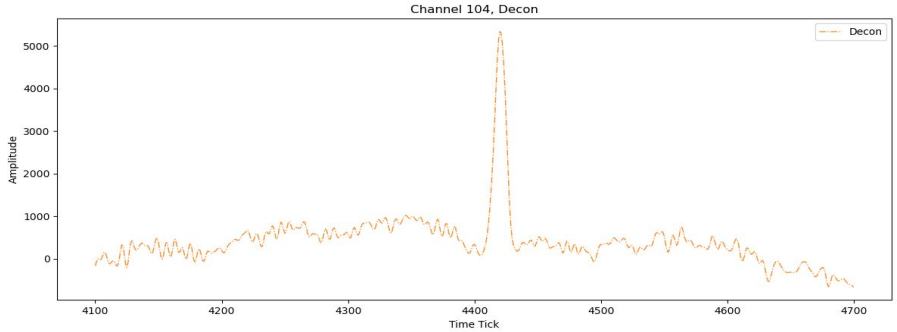
Rebin	Truth_th	Padding
6	100	both side, 1

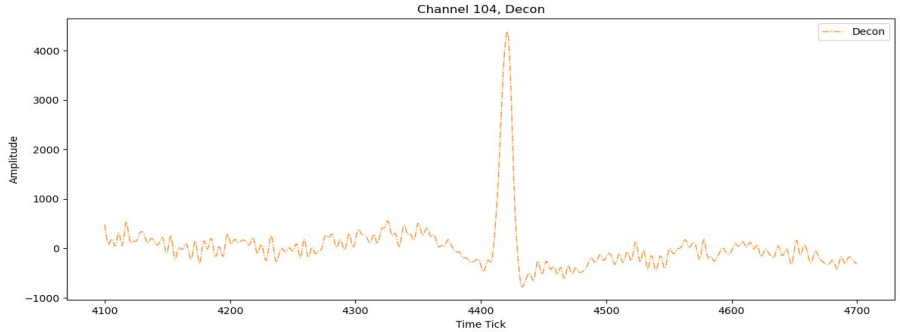
1D Waveform Study - WC standalone

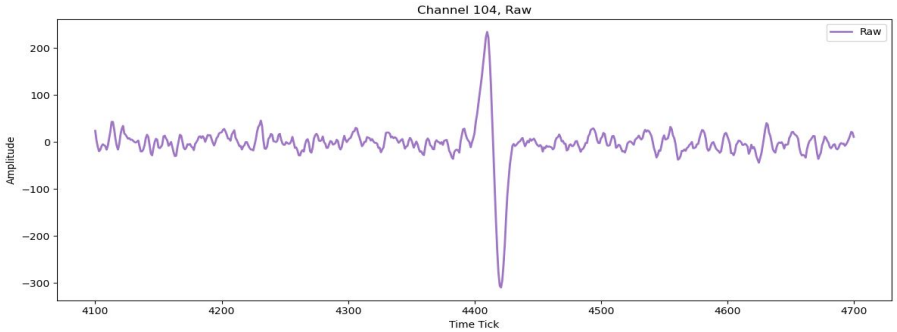


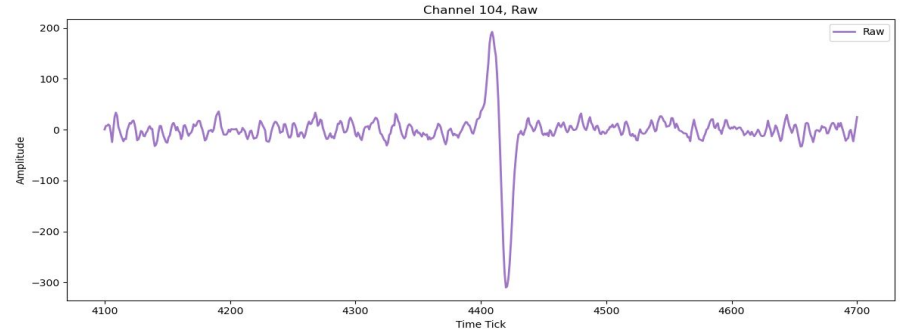
- Studied 1D waveform of isochronous tracks: rebin 6, truth_th 100, (before/after) padding 1
- Padding is able to recover subtracted ROI window, but sometimes recover the window too much
- Padding 1 in this case will add 6 time ticks to both side → need refinement?

1D Waveform Study - Randomness Issues









- Executed PDHD simulation twice, w/ NO changes
- The simulated raw waveform is differ from run by run

  → 1D waveform is less deterministic since we feed different decon for each model

Outline

- PDHD
  - Real Data (w/ MobileUNetV3, rebin = 10, Truth_th = 100)
  - Single Track Evaluation (WC standalone)
  - 1D Waveform Study (WC standalone)

- **PDVD**
  - **Parameter List for Simulation**

- Memory Profiling

Parameter List for Simulation (Jsonnet)

- For the main Jsonnet file (wcls-sim-drift-deposplat.jsonnet), a Google sheet is ready: https://docs.google.com/spreadsheets/d/1WyMjH4deXoRYhCQUaNqwk-U6DGCiAg1ozsu5Ce2ZACo/edit?usp=sharing
- Key parameters:

Type	Name	Parameter	Value
FieldResponse	field0	filename	protodunevd_FR_3view_speed1d55.json.bz2
WireSchemaFile		filename	protodunevd-wires-larsoft-v3.json.bz2
ColdElecResponse	elecresp0	gain/nticks/postgain/shaping/tick	See the link above
JsonElecResponse	elecresp1	filename	dunevd-coldbox-elecresp-top-psnorm_400.json.bz2
EmpiricalNoiseModel	empericalnoise-anode0	spectra_file	pdvd-bottom-noise-spectra-v1.json.bz2

- For the main FHiCL file (pdvd_wirecell_sim_deposplat.jsonnet), a Google sheet is under preparation
- Key parameters: protodunevd_v4_refactored.gdml, dune35t_response_v1.0.root, uboone_filters_v0.1.root, protoDUNETPCChannelMap_RCE_v4.txt, SCE_DataDriven_180kV_v4.root, Pedestal, Gain Cal. etc.

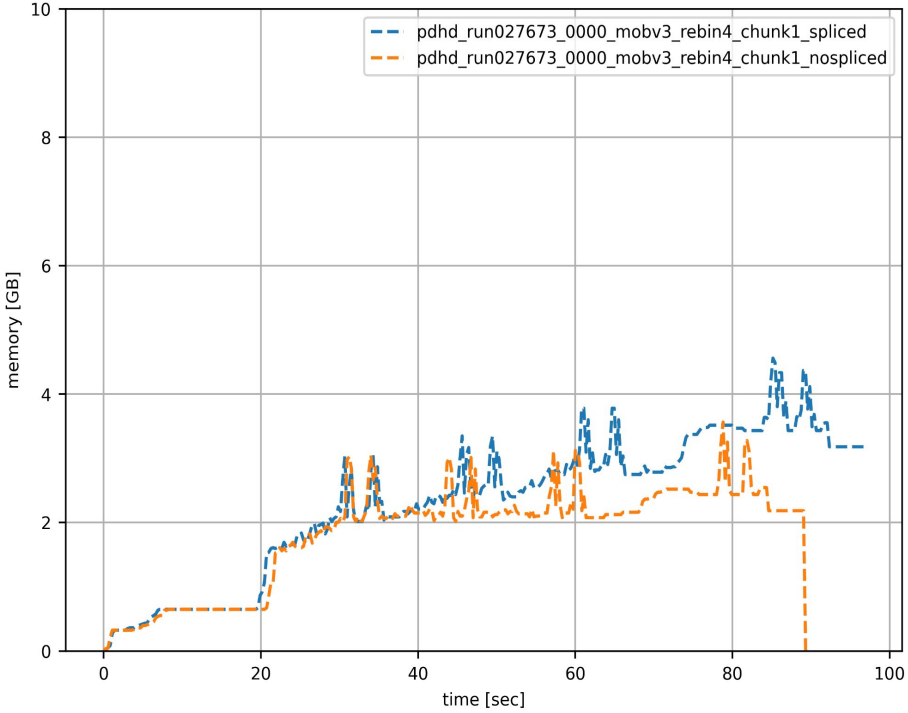
Outline

- PDHD
  - Real Data (w/ MobileUNetV3, rebin = 10, Truth_th = 100)
  - Single Track Evaluation (WC standalone)
  - 1D Waveform Study (WC standalone)

- PDVD
  - Parameter List for Simulation

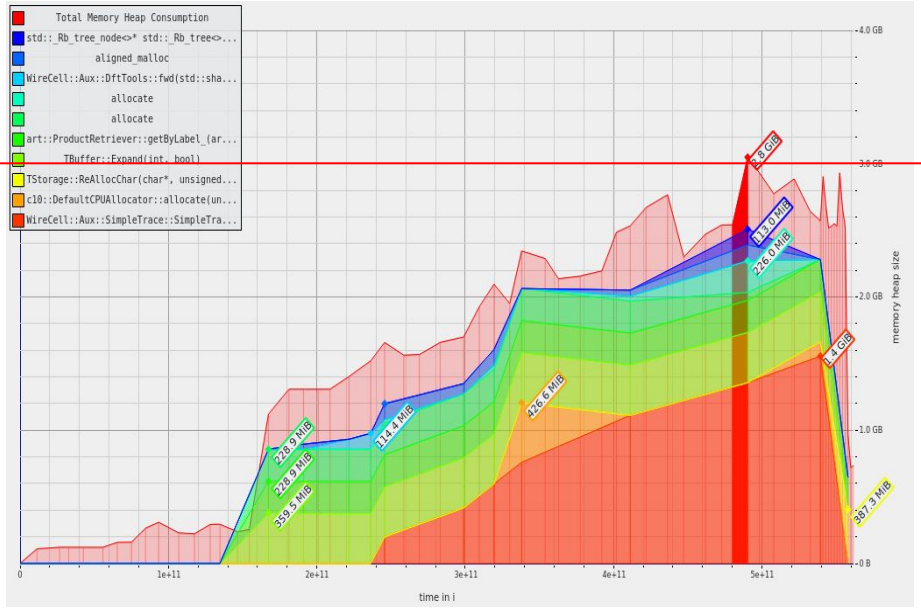
- **Memory Profiling**

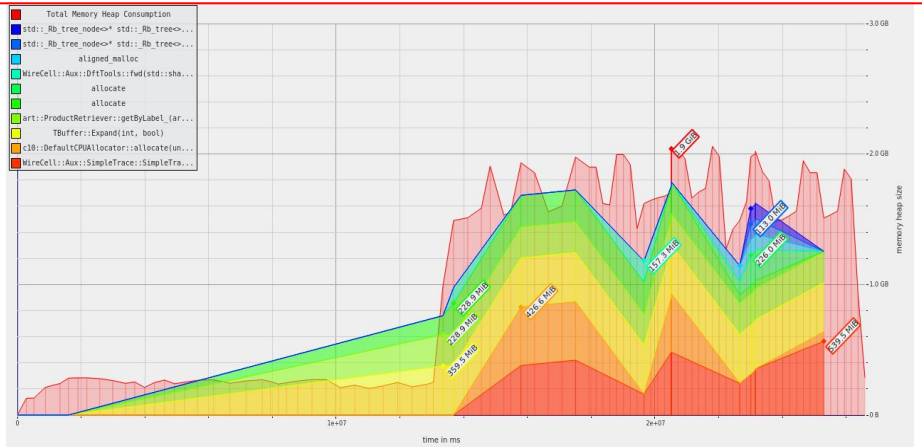
Memory Profiling - Activity Logger



- Tested with PDHD data 027673-0000
- Orange - no spliced vs. Blue spliced
- Accumulated memory is solved
  → Pedestal is decreased about 1 GB
- What is the “spliced” option?

Memory Profiling - Valgrind Massif



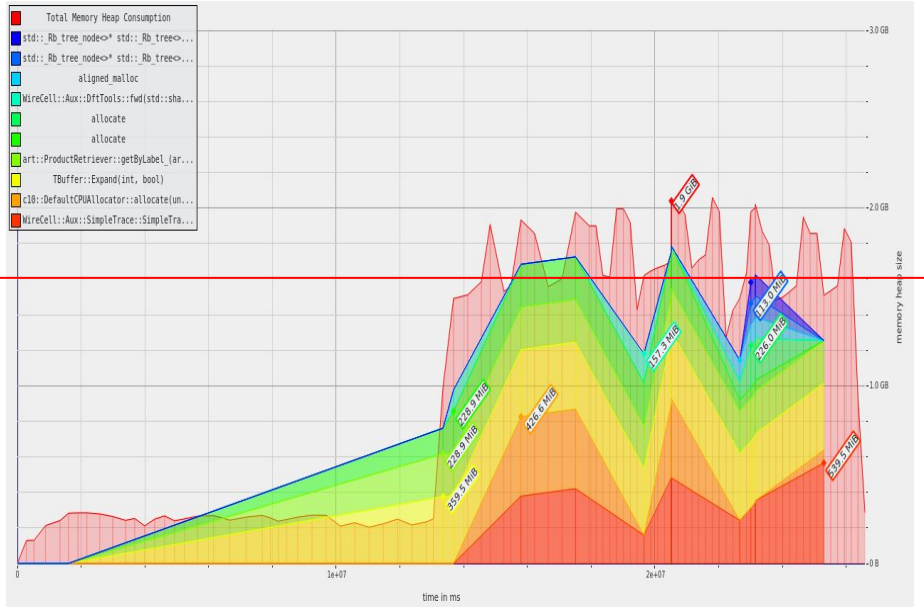


- ❑ PDHD real data, splices = true
- ❑ PDHD real data, splices = false

- Heap profiling confirmed accumulated memory is due to “spliced” option

Memory Profiling - Valgrind Massif

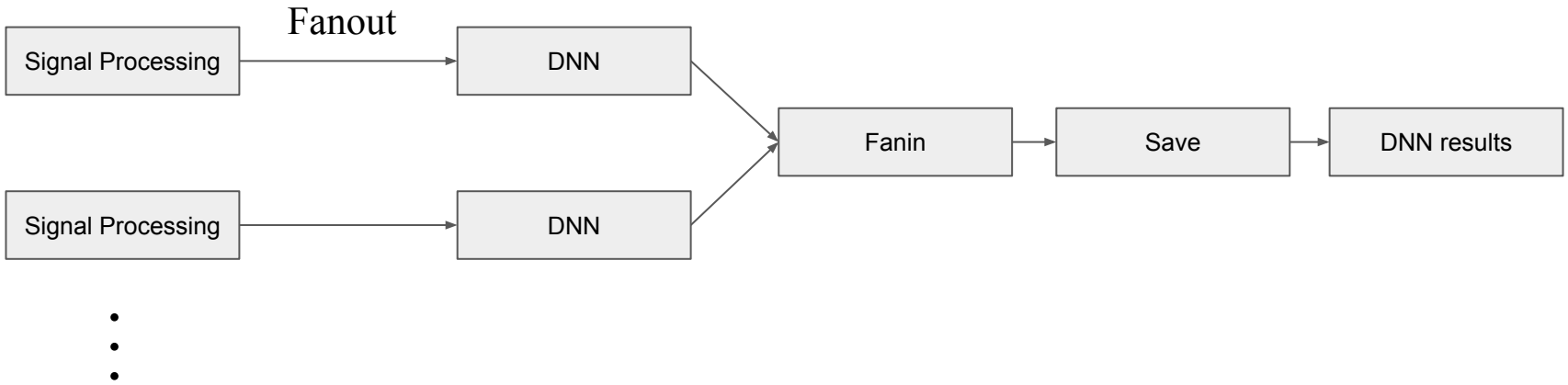
❑ PDVD real data, splices = false



❑ PDHD real data, splices = false

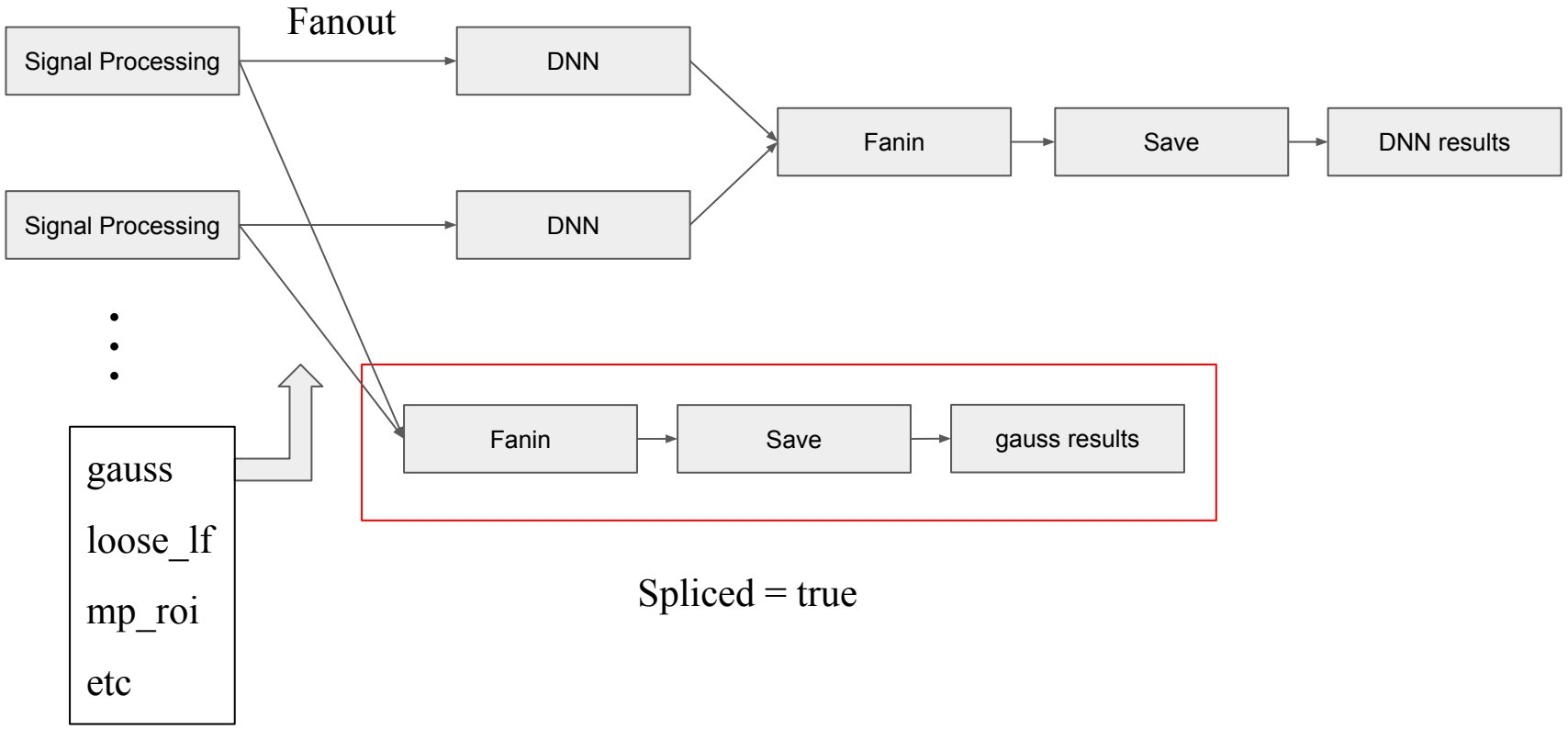
- In this case, the spliced option is set to false for both PDVD and PDHD data processing
  → Now, the overall trend is much similar to each other

Memory Profiling - Spliced Explained

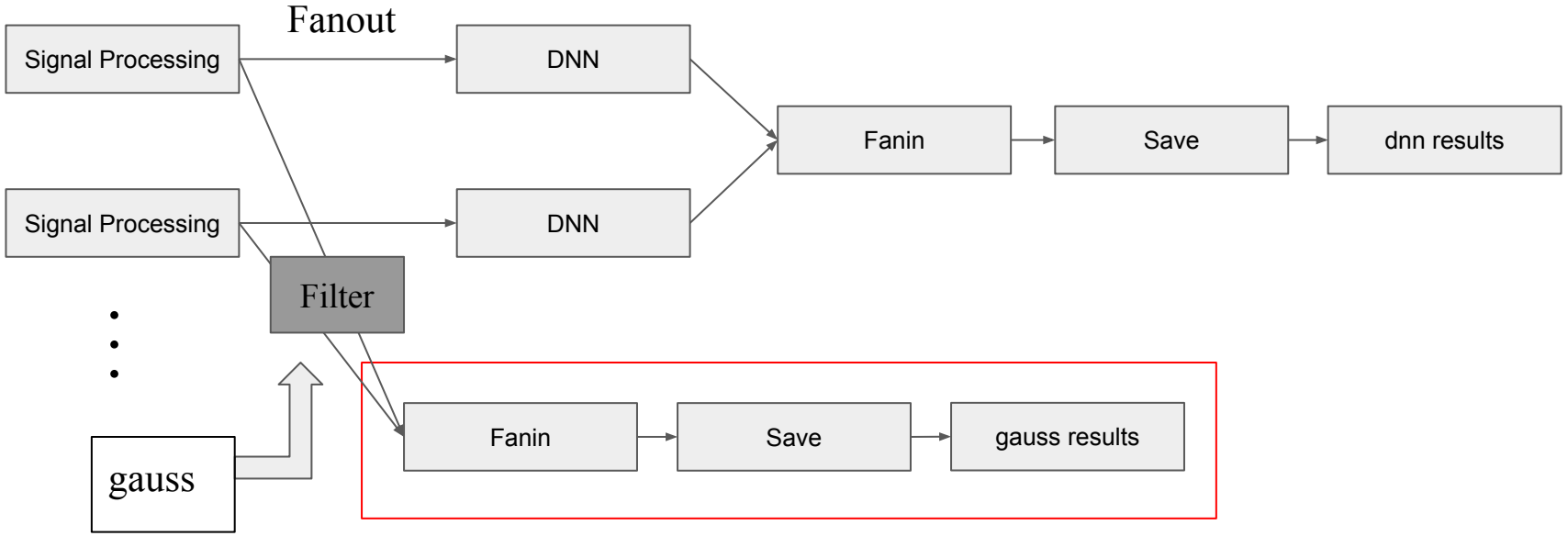


Spliced = false

Memory Profiling - Spliced Explained



Memory Profiling - Spliced Explained



Expected memory ~ few MB

Spliced = true

→ Memory would be still accumulated per APA processing, but not that big as before

Many thanks to Brett and Haiwang

Memory Profiling - History w/ PDHD 027673-0000

Model change

- UNet + rebin 10 + chunk 1 + spliced: 6.7 GB
- MobileUNetV3 + rebin 10 + chunk 1 + spliced: 4.1 GB
- MobileUNetV3 + rebin 4 + chunk 1 + spliced: 4.9 GB

Chunking implementation → Lower the ML **inference peak** but not the **pedestal**

- MobileUNetV3 + rebin 10 + chunk 4 + spliced: 3.8 GB
- MobileUNetV3 + rebin 4 + chunk 4 + spliced: 4.0 GB

Accumulated memory issue → Lower the **pedestal** but not the ML **inference peak**

- MobileUNetV3 + rebin 10 + chunk 1 + no spliced: 2.9 GB
- MobileUNetV3 + rebin 4 + chunk 1 + no spliced: 3.6 GB

Combined

- UNet + rebin 4 + chunk 4 + no spliced: 4.3 GB (300 s)
- **MobileUNetV3 + rebin 4 + chunk 4 + no spliced: 2.7 GB (102 s) - Latest!**

Summary

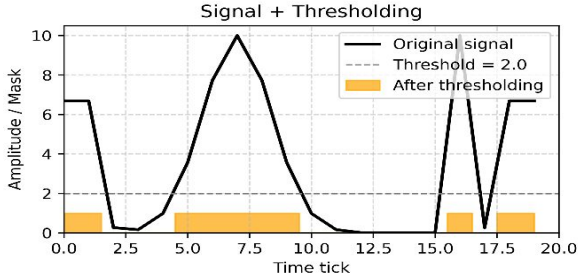
- General
  - The cause of the accumulated memory issue is finally identified
  - Discussed with Moon about how to handle bad channels for DNN-ROI

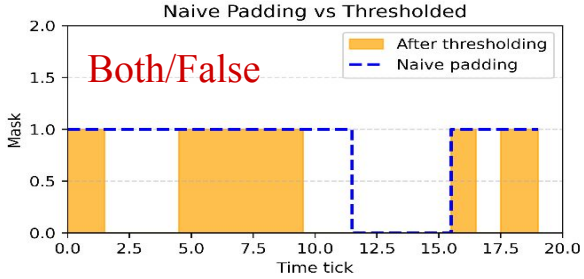
- PD-HD
  - Performed visual inspection on real PDHD data
  - Padding works on isochronous shower and reduced bias

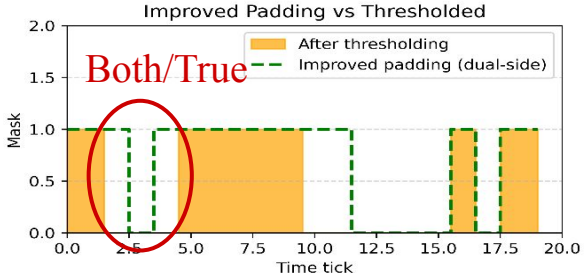
- PD-VD
  - Parameter lists are under preparation and should be validated

Back Up

Padding Strategy





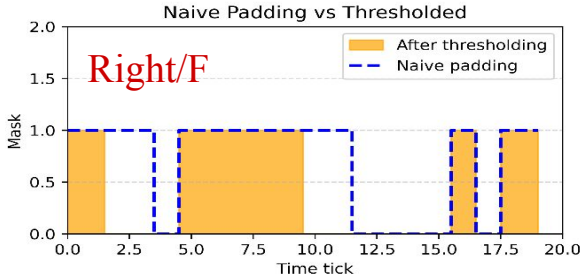


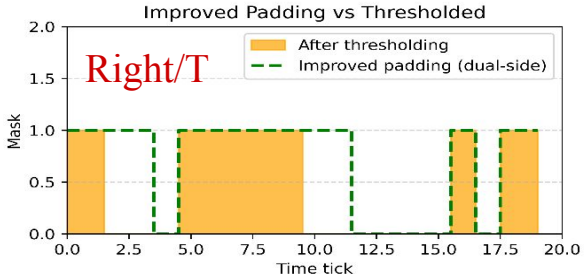
What is changed?

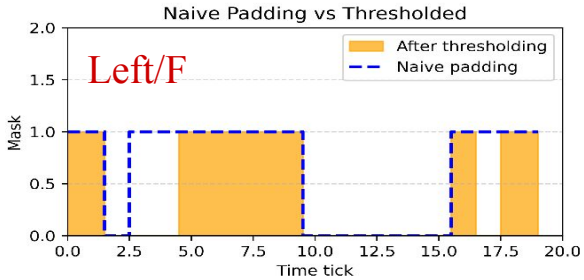
- Can choose which sides to pad
- Can avoid merging

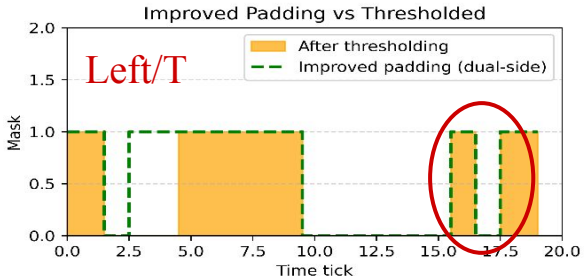
Parameters for this example

- padding = 2
- min_run = 2
- min_gap = 1
- padding_side = both/right/left
- avoid_merge = False/True









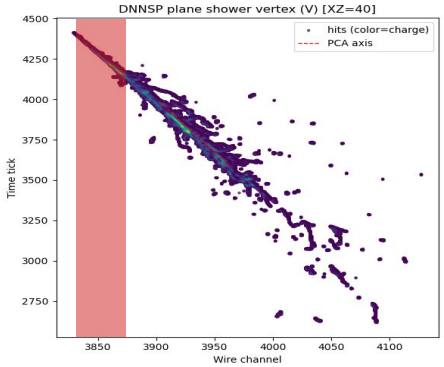
DNN-ROI Performance Evaluation

- For track events, three metrics are used: Bias, Resolution, and Inefficiency

$$Bias = 100 \times \left(\left\langle \frac{Q_{reco}}{Q_{truth}} \right\rangle - 1 \right) \qquad Resolution = 100 \times \frac{RMS\left(\frac{Q_{reco}}{Q_{truth}}\right)}{\left\langle \frac{Q_{reco}}{Q_{truth}} \right\rangle} \qquad Inefficiency = 100 \times \frac{Number\ of\ bad\ channels}{Number\ of\ valid\ truth\ channels}$$

- For shower events, a charge profile based on vertex information was added as the fourth metric:
  - Sum the charge along the shower direction up to 42 wire channels (≈ 1–2 radiation lengths)
  - Compare the reconstructed-to-truth ratio charge ratio

$$Q_{method} = \sum_{w \in W} Q_{method}(w) \qquad R_{cprofile} = \frac{Q_{reco}}{Q_{truth}}$$



- Samples were generated with
  - Detector configuration: ProtoDUNE - Horizontal Drift (PD-HD)
  - XZ angle: 0°, 10°, 20°, 30°, 40°, 50°, 60°, 70°, 80°, 85°
  - Shower energies: 100 MeV, 500 MeV, 1 GeV, 2 GeV, 3 GeV, 5 GeV
  - Software: WCT standalone (Tracks), LAr-WCT (Showers)