

CIDeR-ML Workshop

Patrick “*Ka Vang*” Tsang (SLAC)

Jul 29, 2024



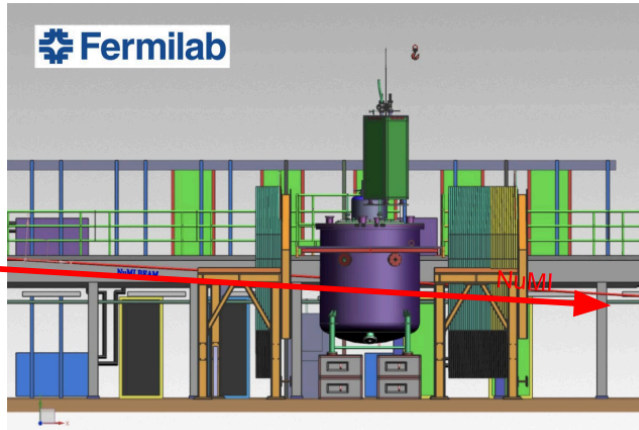
U.S. DEPARTMENT OF
ENERGY

Stanford
University

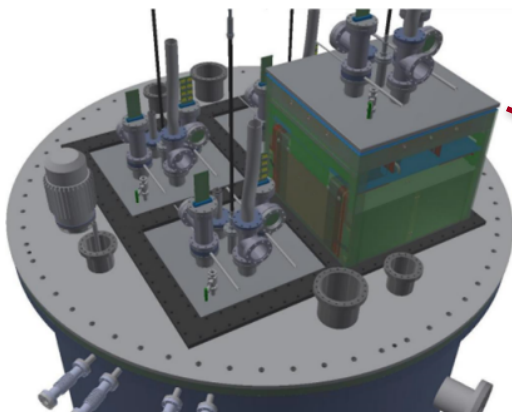
SLAC NATIONAL
ACCELERATOR
LABORATORY

Introduction: ND-LAr 2x2

ND-LAr 2x2 at FNAL/NuMI



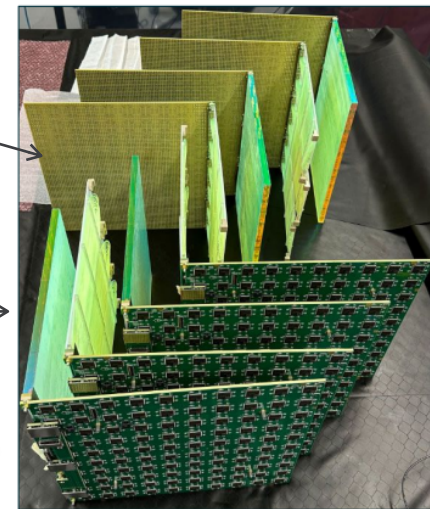
4 Modules



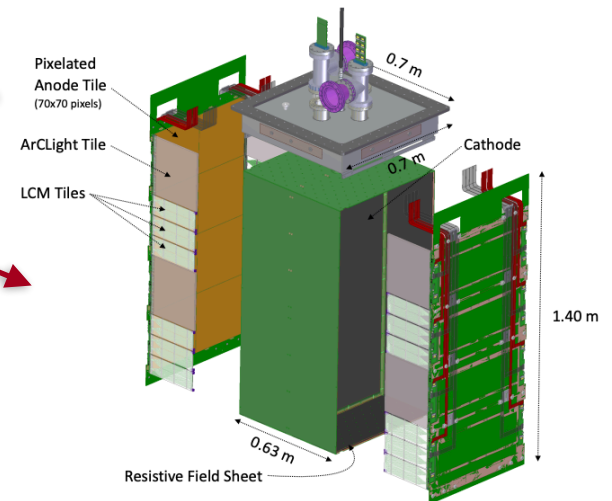
Tiles for Charge & Light Detection

Pixels

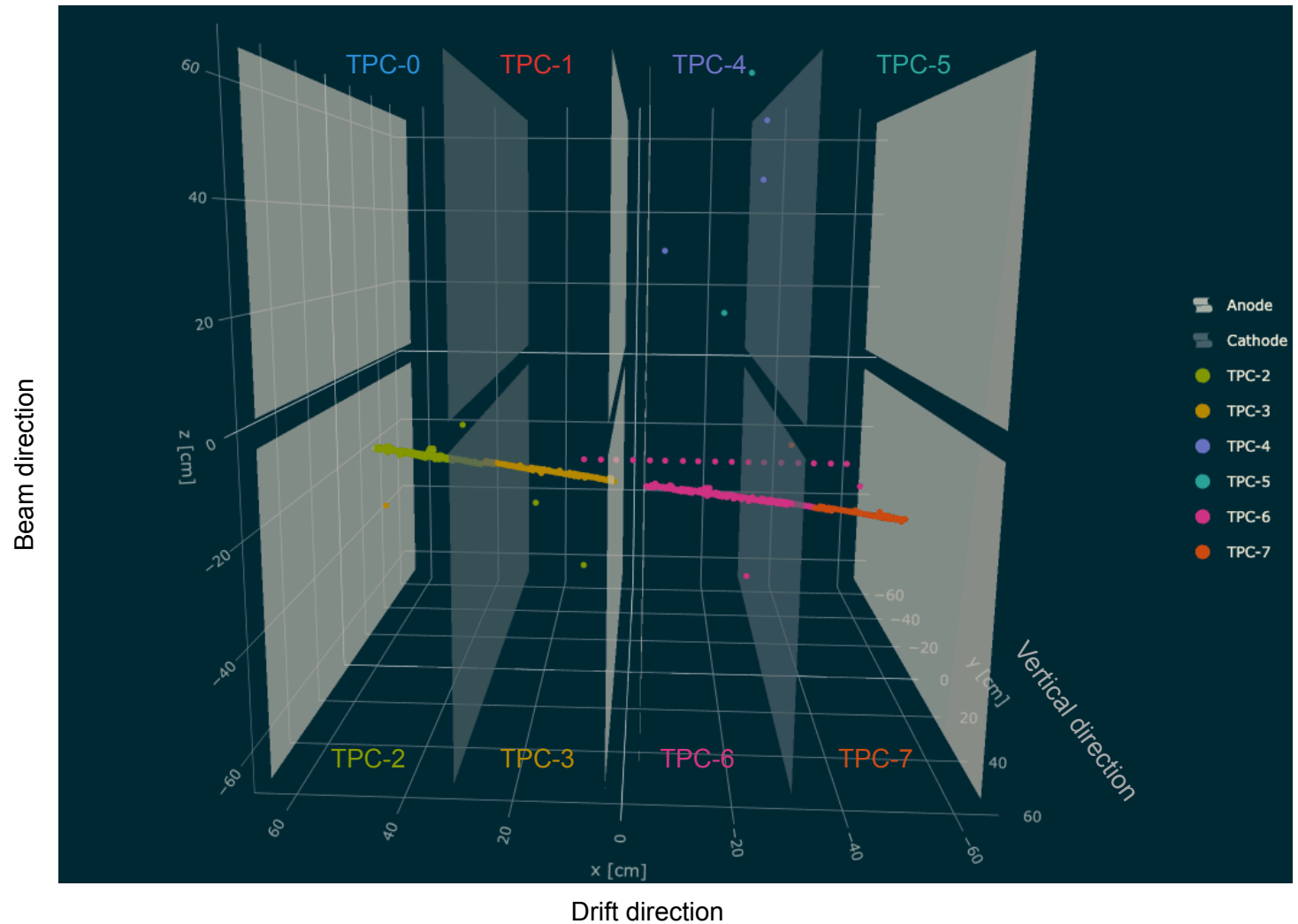
ArCLight (panel)
LCM (fibers)
=> SiPMs



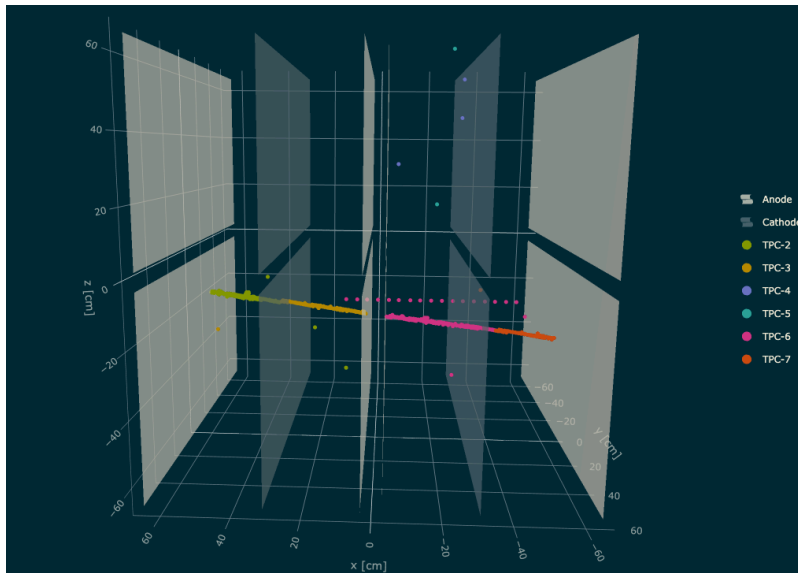
Breakdown of one module (2 TPCs)



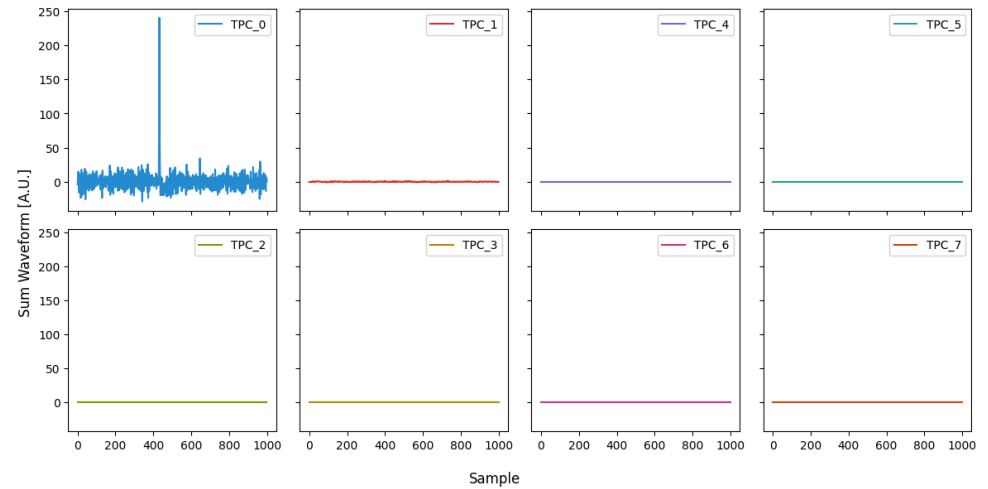
Example: Non-Beam Event at 500 V/cm



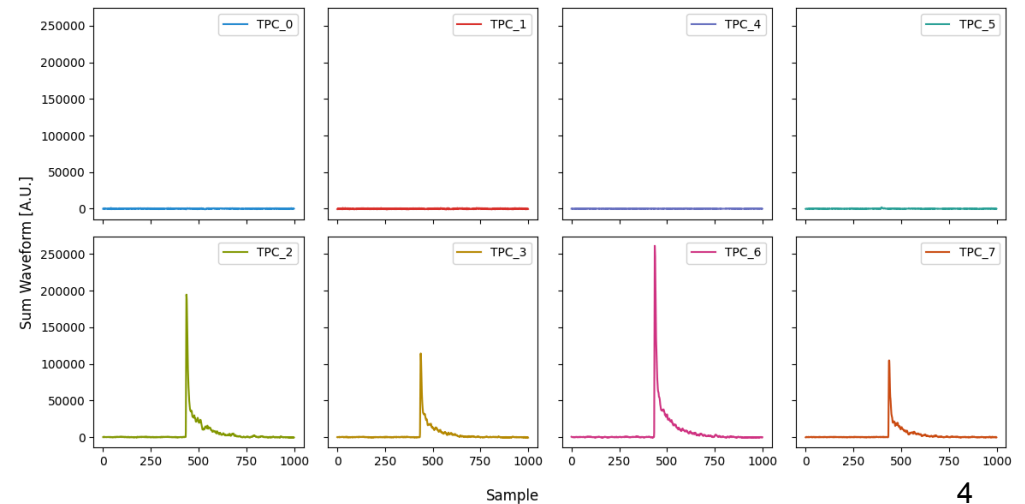
Example: Charge-Light Trigger Matching



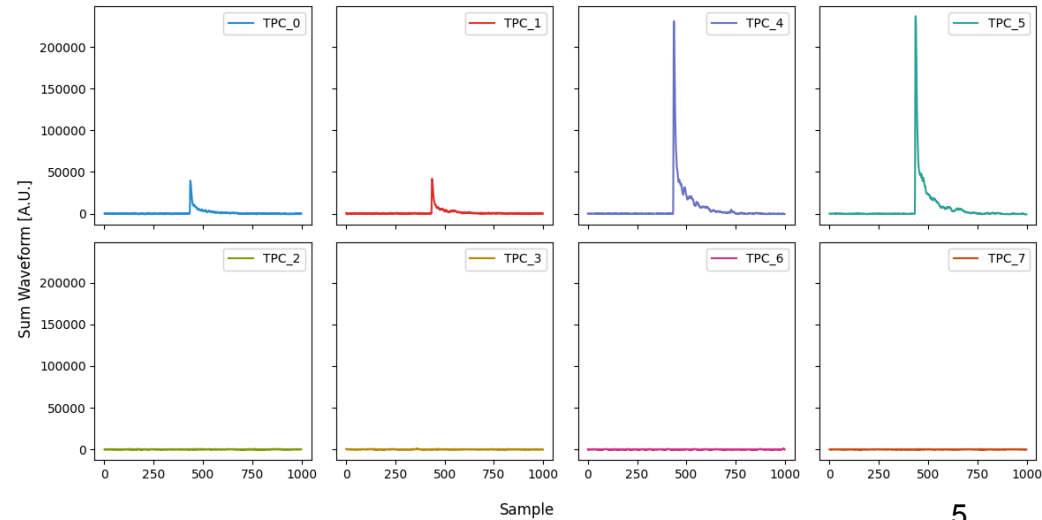
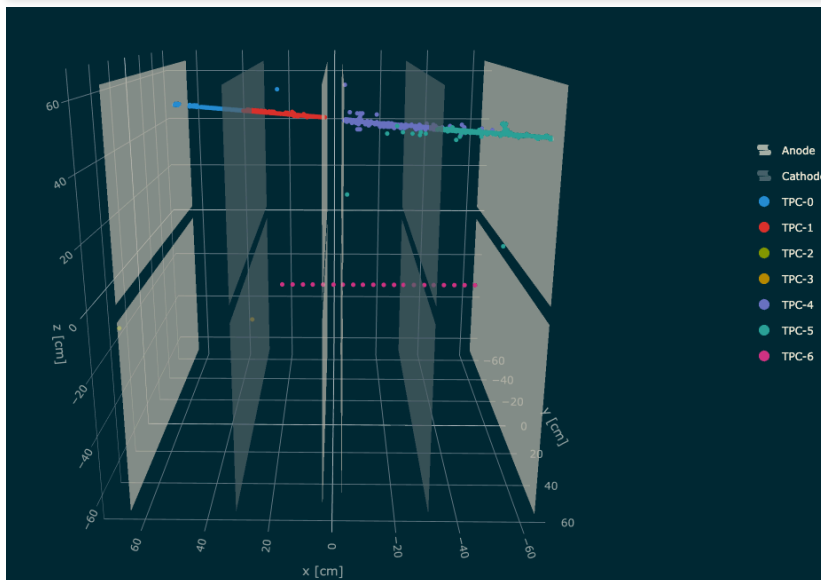
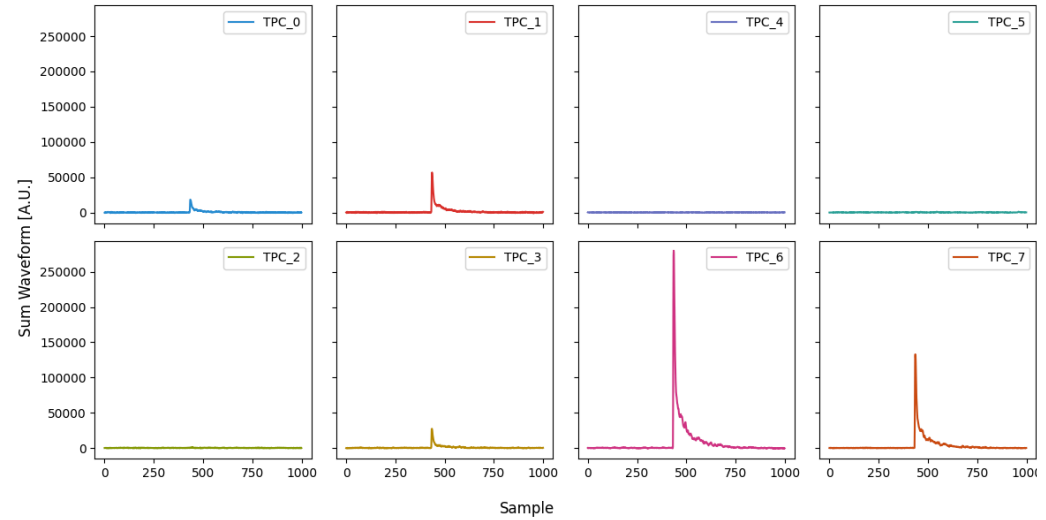
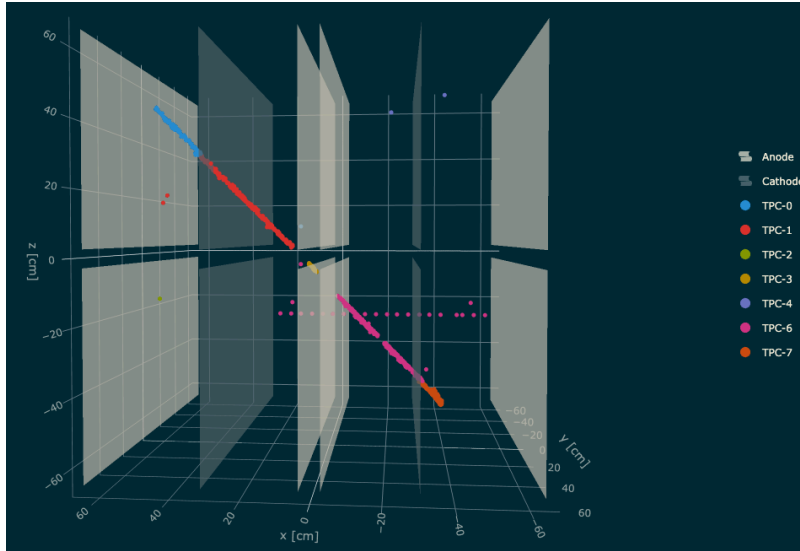
Default reco. w/ signal processing (NOT GOOD)



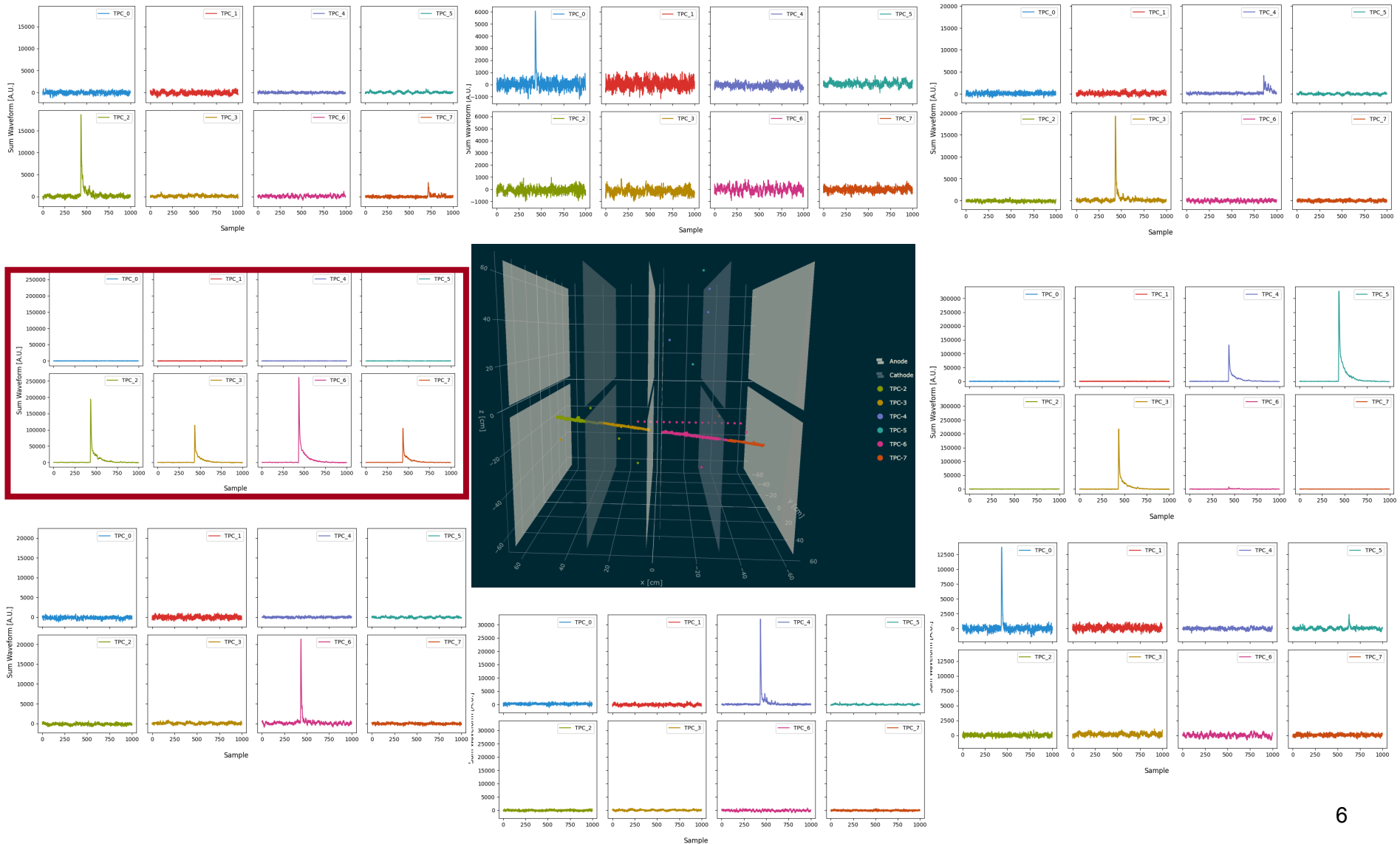
My educated guess using raw waveforms (OK?)



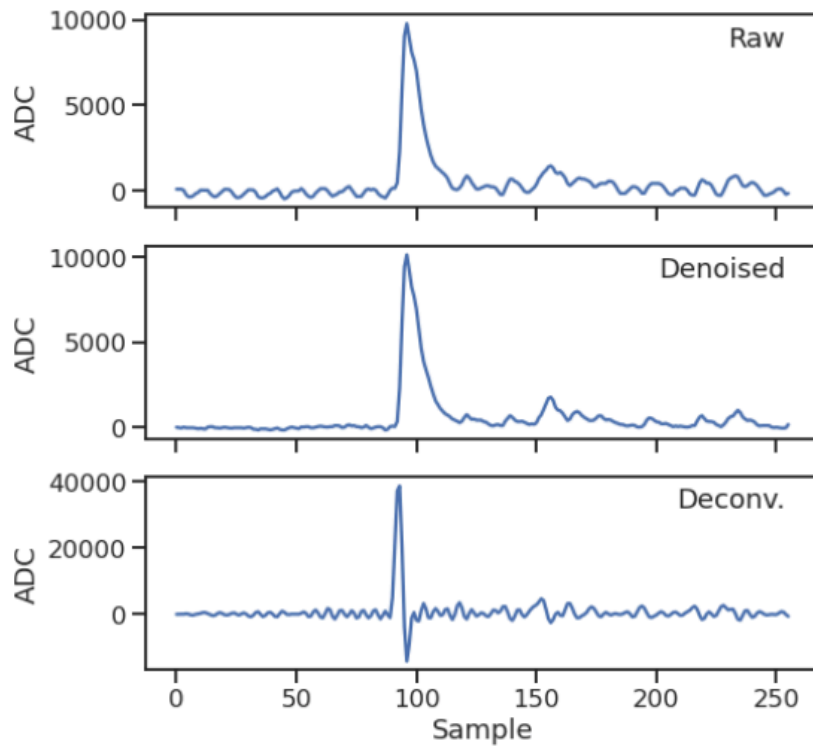
More Example (My Guess Matching)



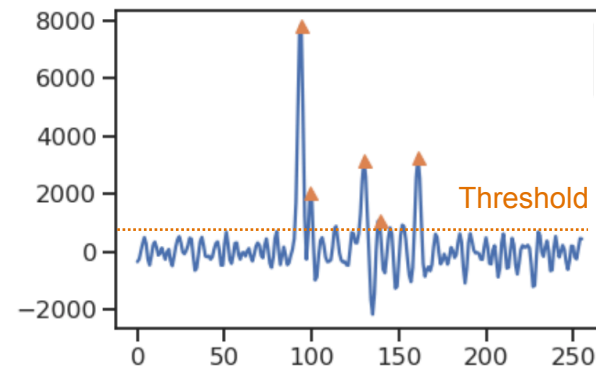
Problem: Multiple Light Triggers



Single Processing of SiPM Waveform



Peak Finding of SiPM Signal



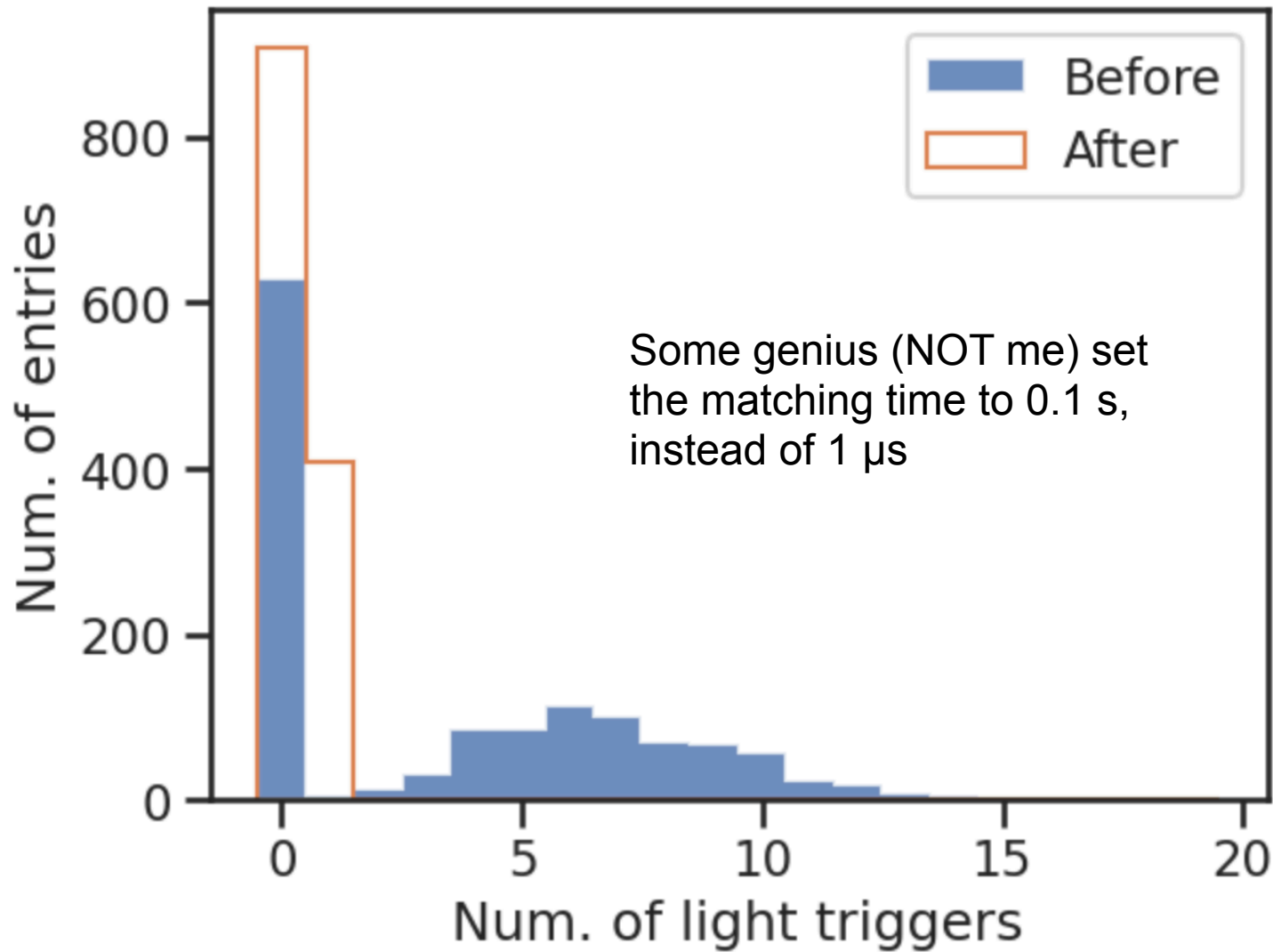
Calibration ADC=>p.e.

Goals: 2x2 Data Preparation

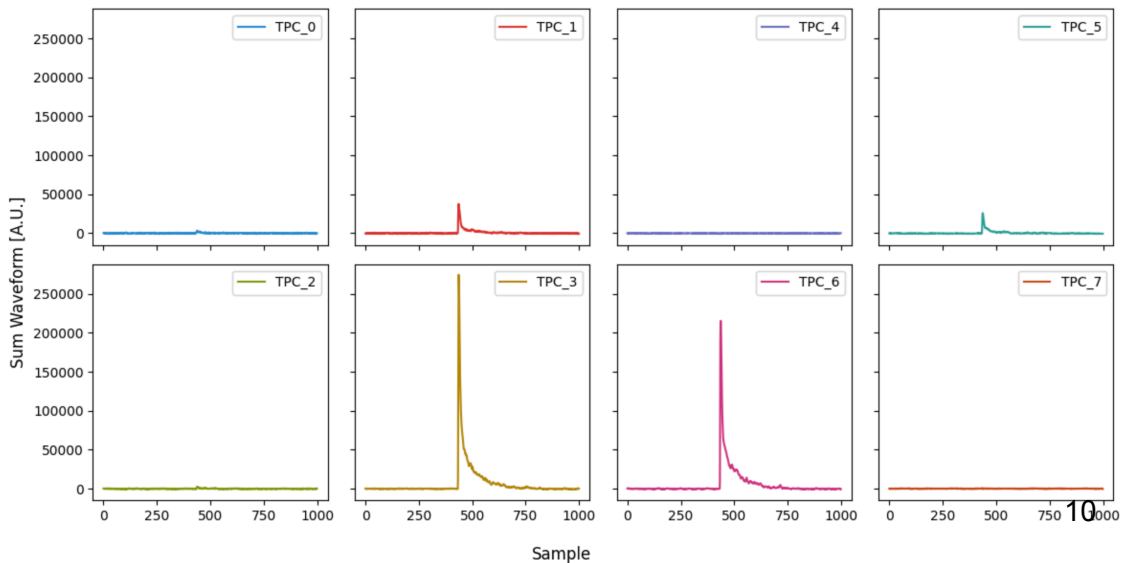
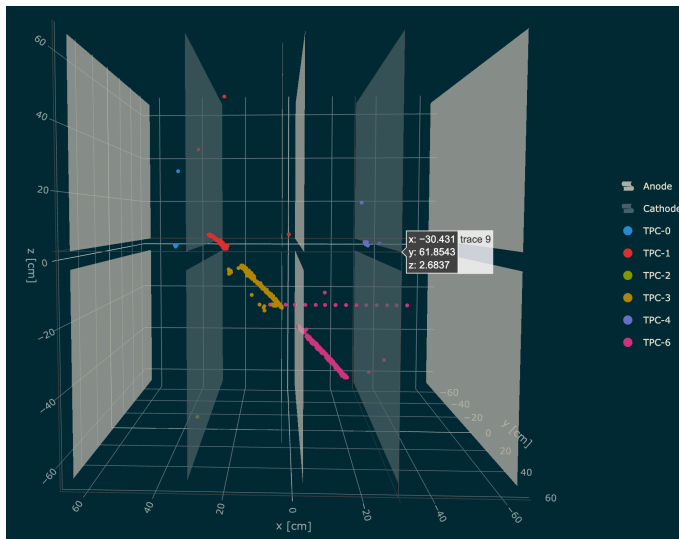
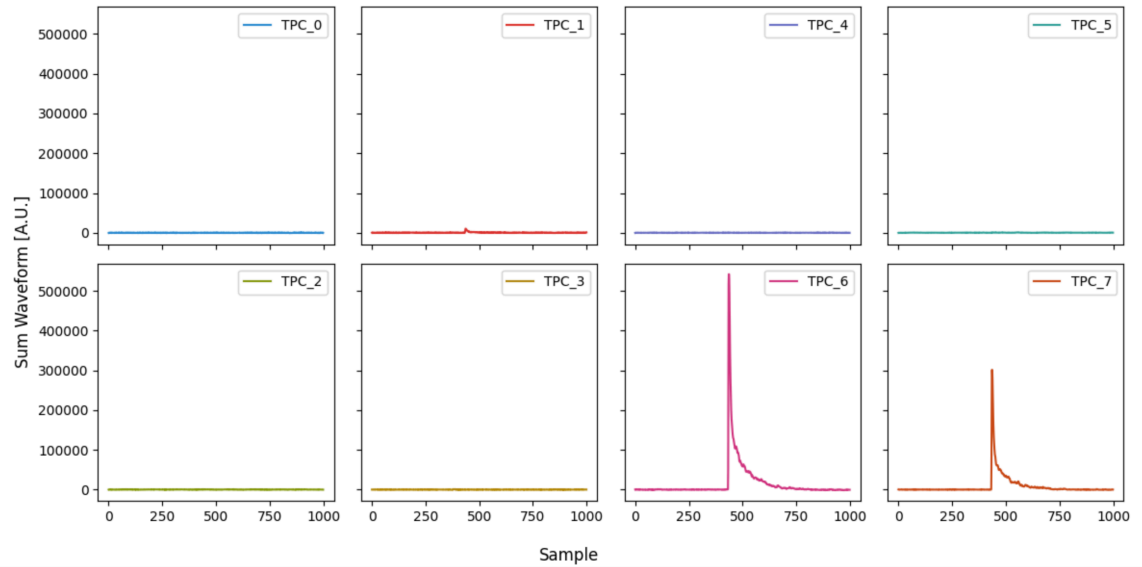
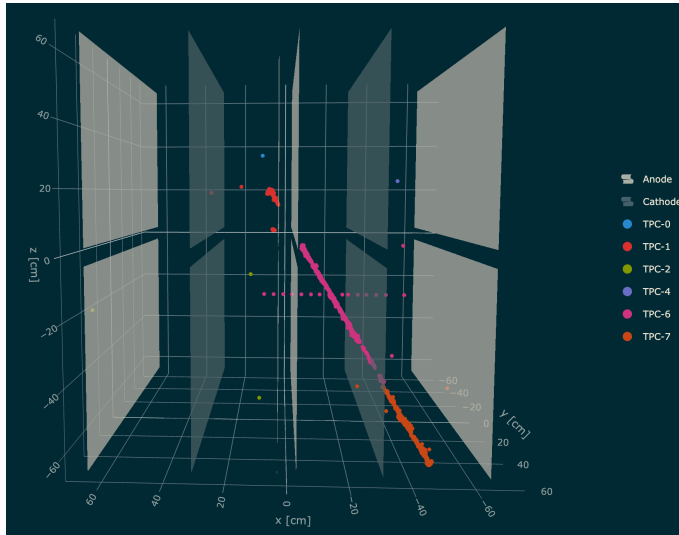
- Check trigger matching [~~1 day~~]
 - Reduce ambiguous triggers?
- Validate channel mapping [~~0.5 – 1 day~~]
 - Raw ADC (8x64) => TPC/SiPM (8x48)
 - Cross-check w/ LUT (~close to data)
- Signal Processing [maybe? 1-2 days]
 - Noise filtering
 - Deconvolution
 - Hit finding (determine threshold)
 - Calibration ADC => p.e.
- Event selection [~~1 day~~]
 - Clean sample of isolated long track + light signal
- Gather above codes [1 day]
- Documentation [1 day]

Prepare a workflow to process 2x2 data for SIREN track calibration.

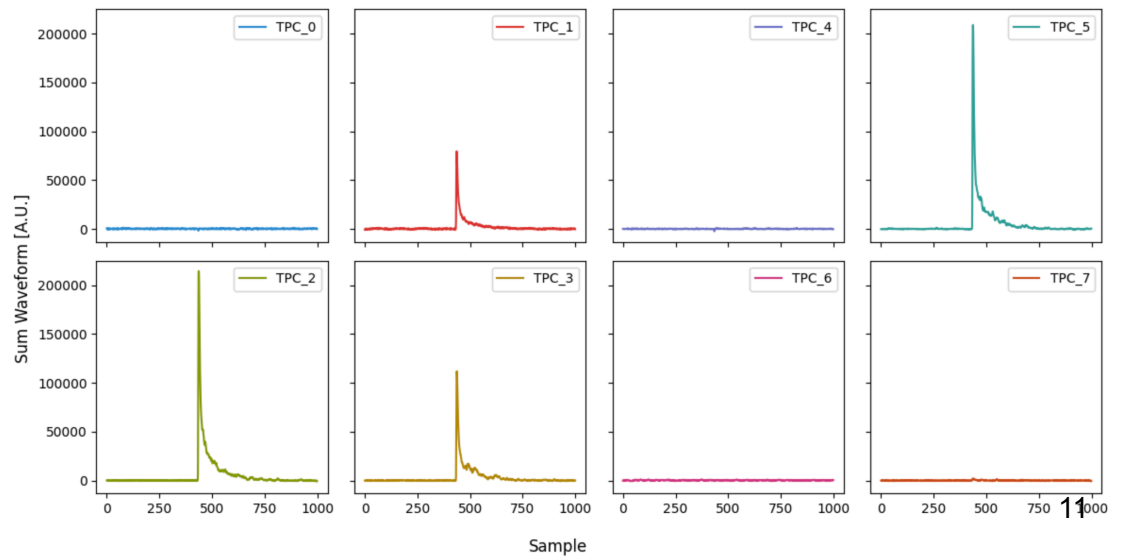
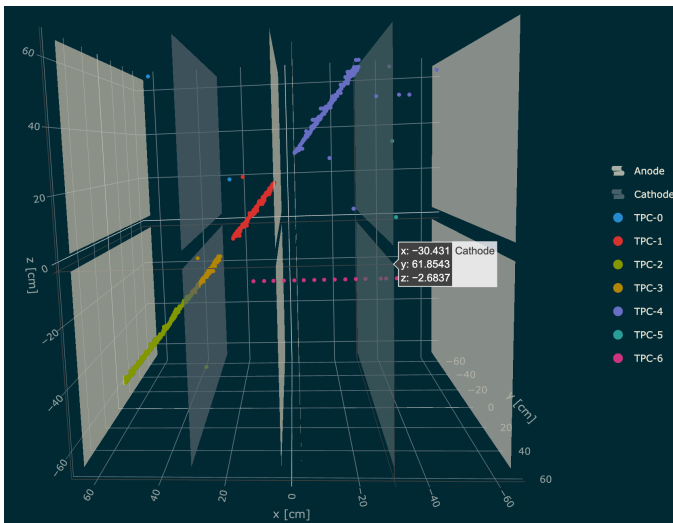
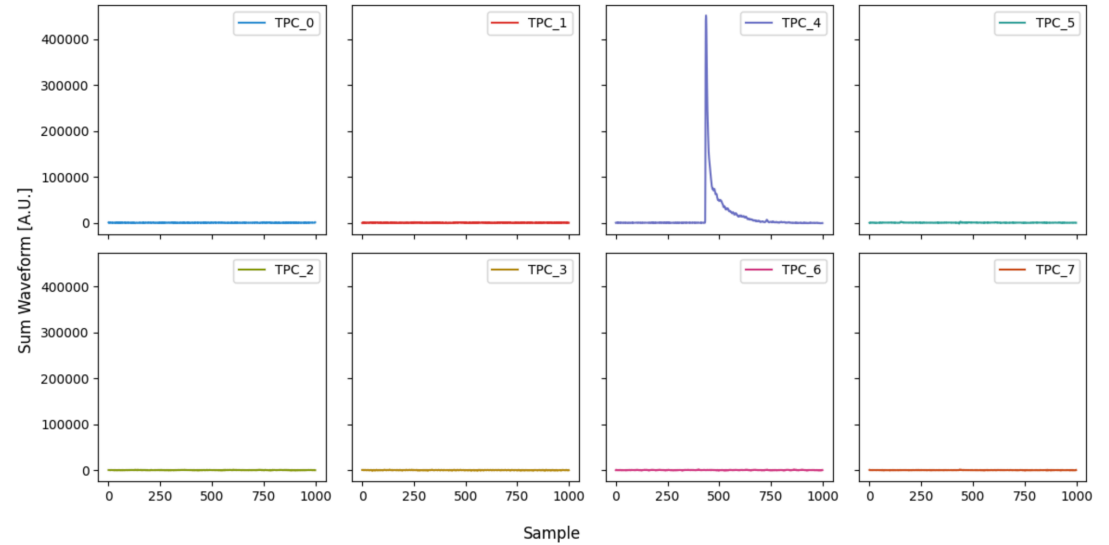
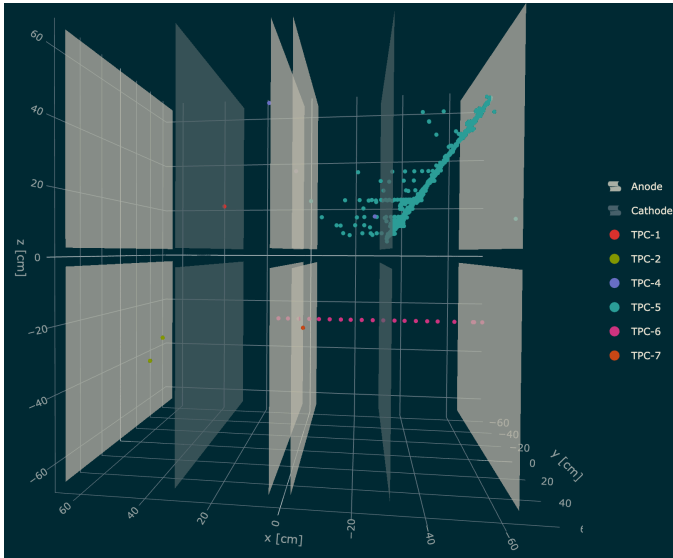
FIXED: Trigger Matching



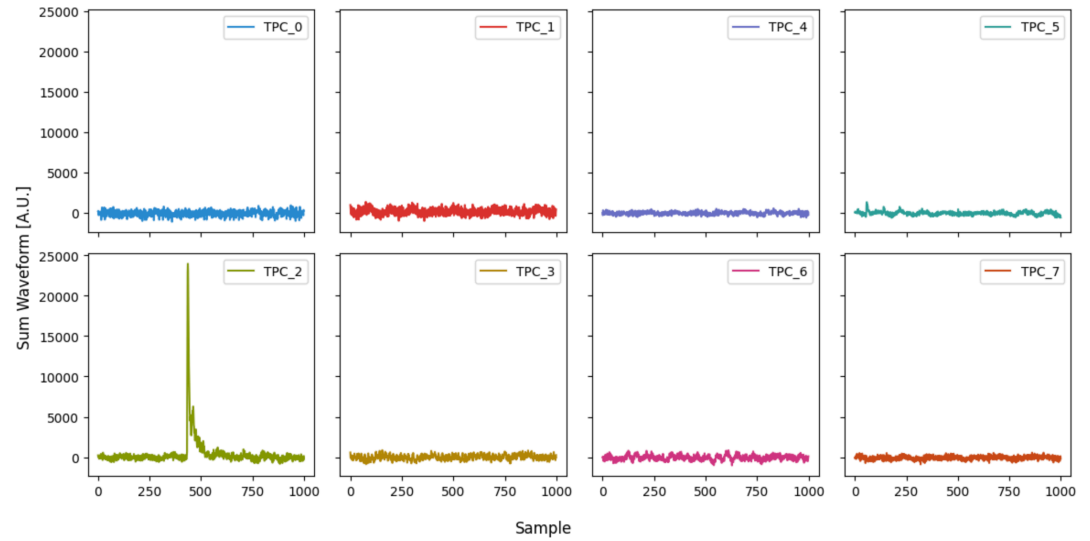
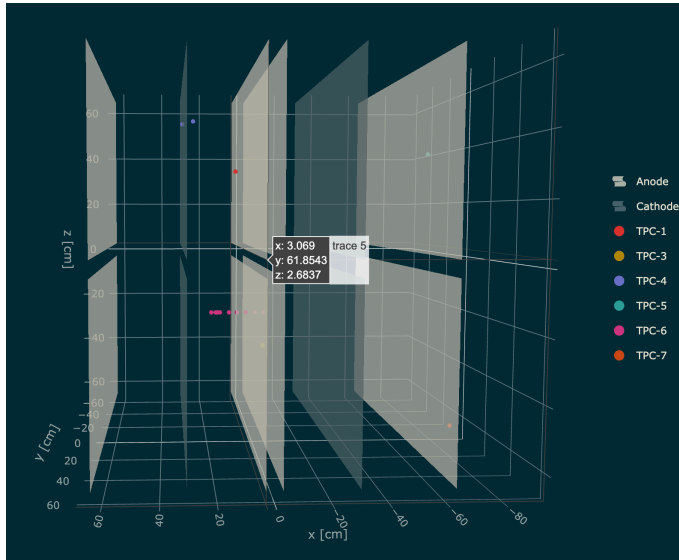
FIXED: Examples



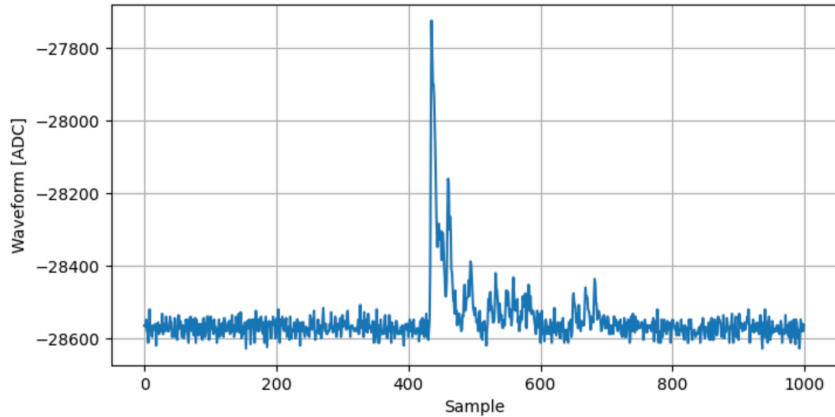
Problem: Swapping of TPC 4 & 5 (for Light)



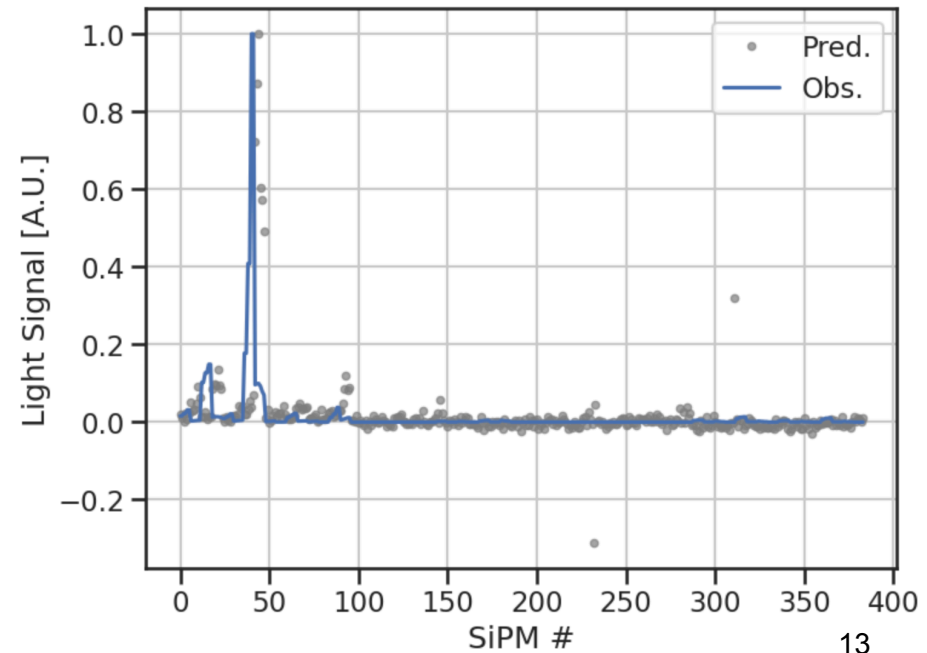
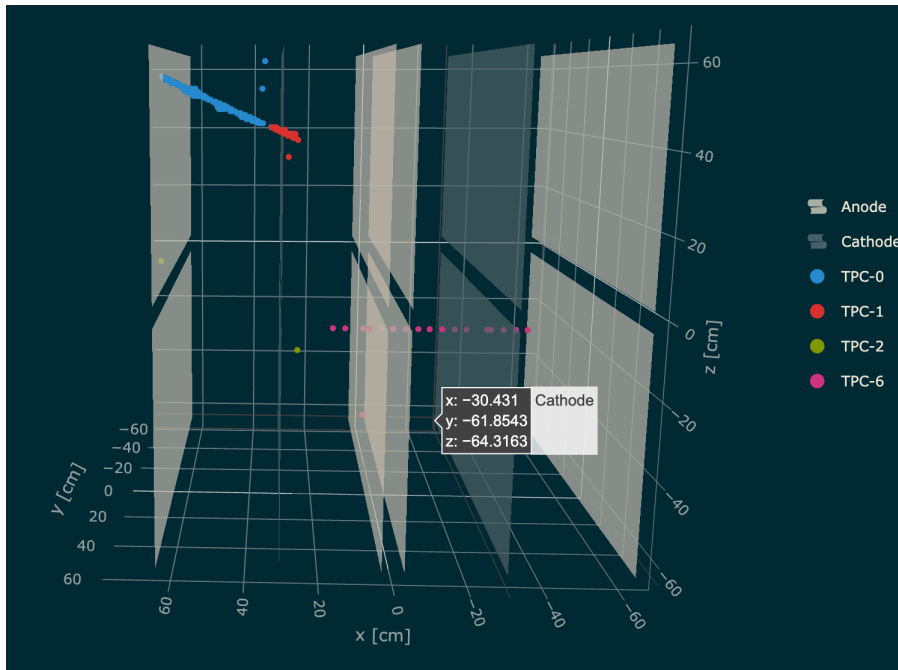
Problem: Light Trigger but No Charge Activity



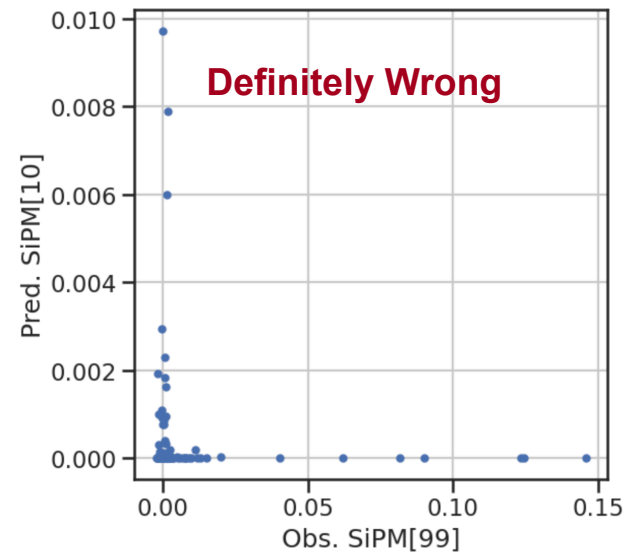
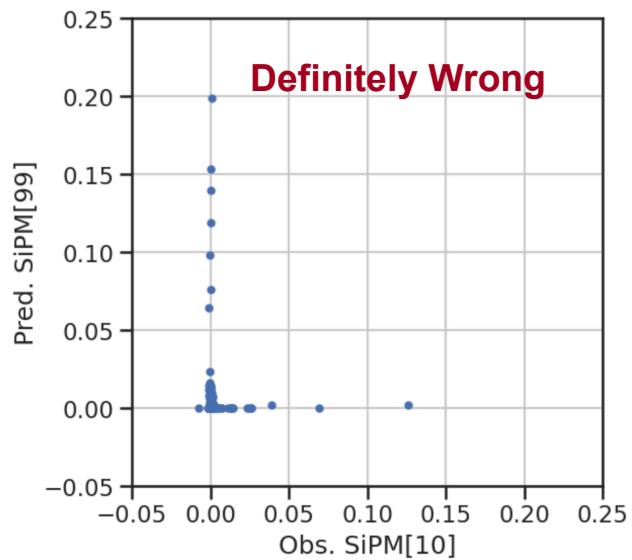
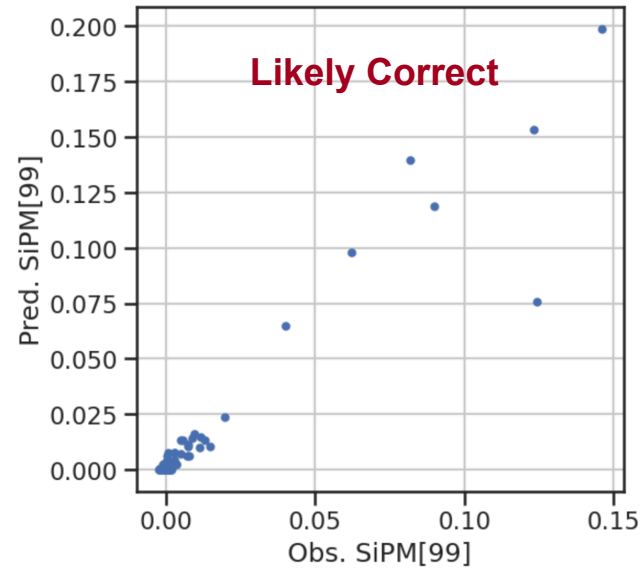
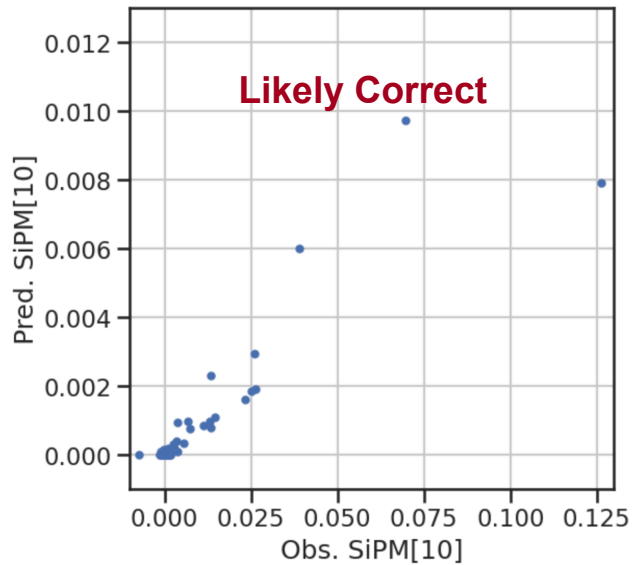
Charge-to-Light Prediction



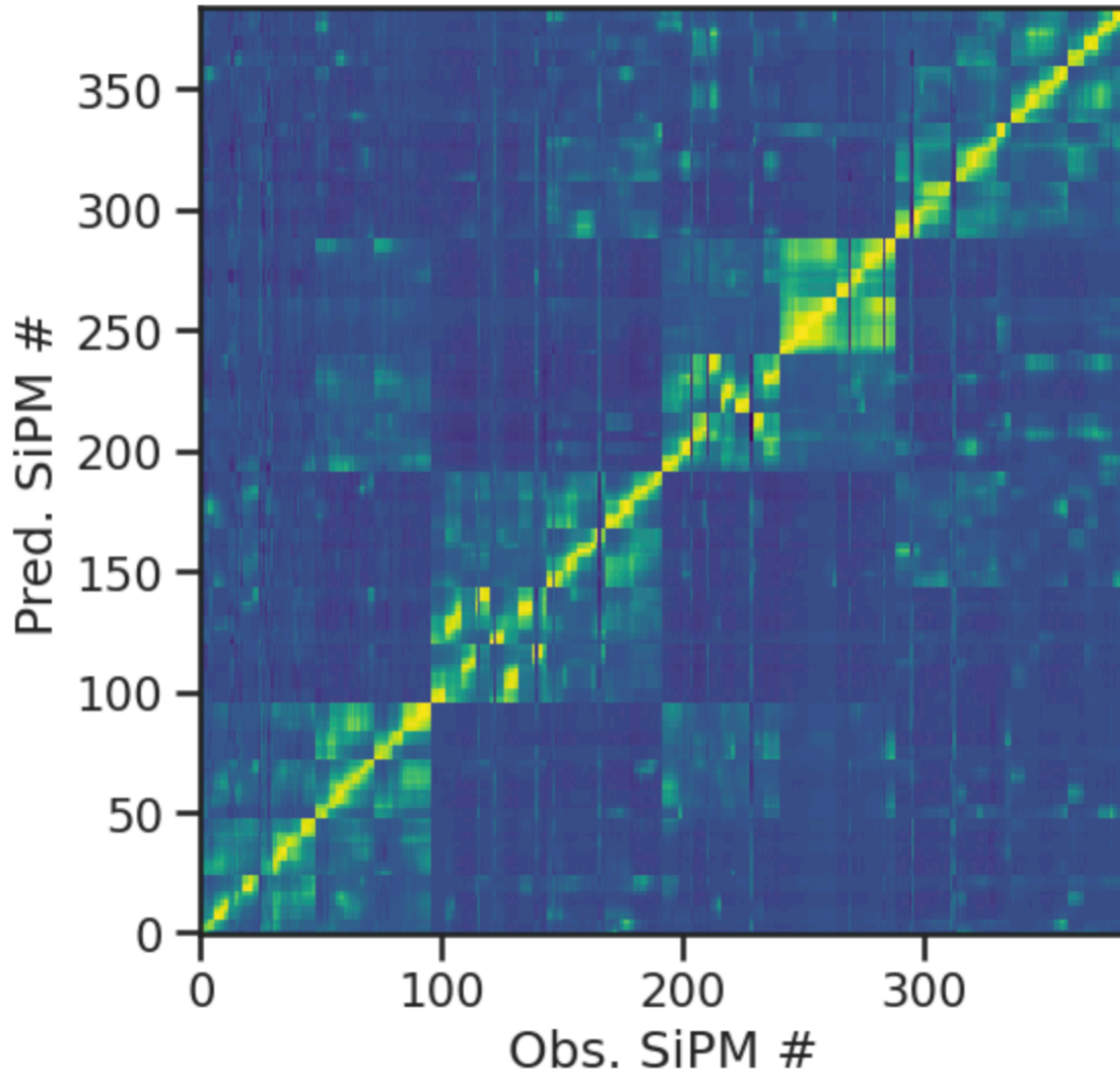
Obs.: Area of the waveform after pedestal subtraction
Pred: Prediction of light signal using charge image and visibility LUT



Validation Channel Map w/ Correlation Plots



Cross Correlation

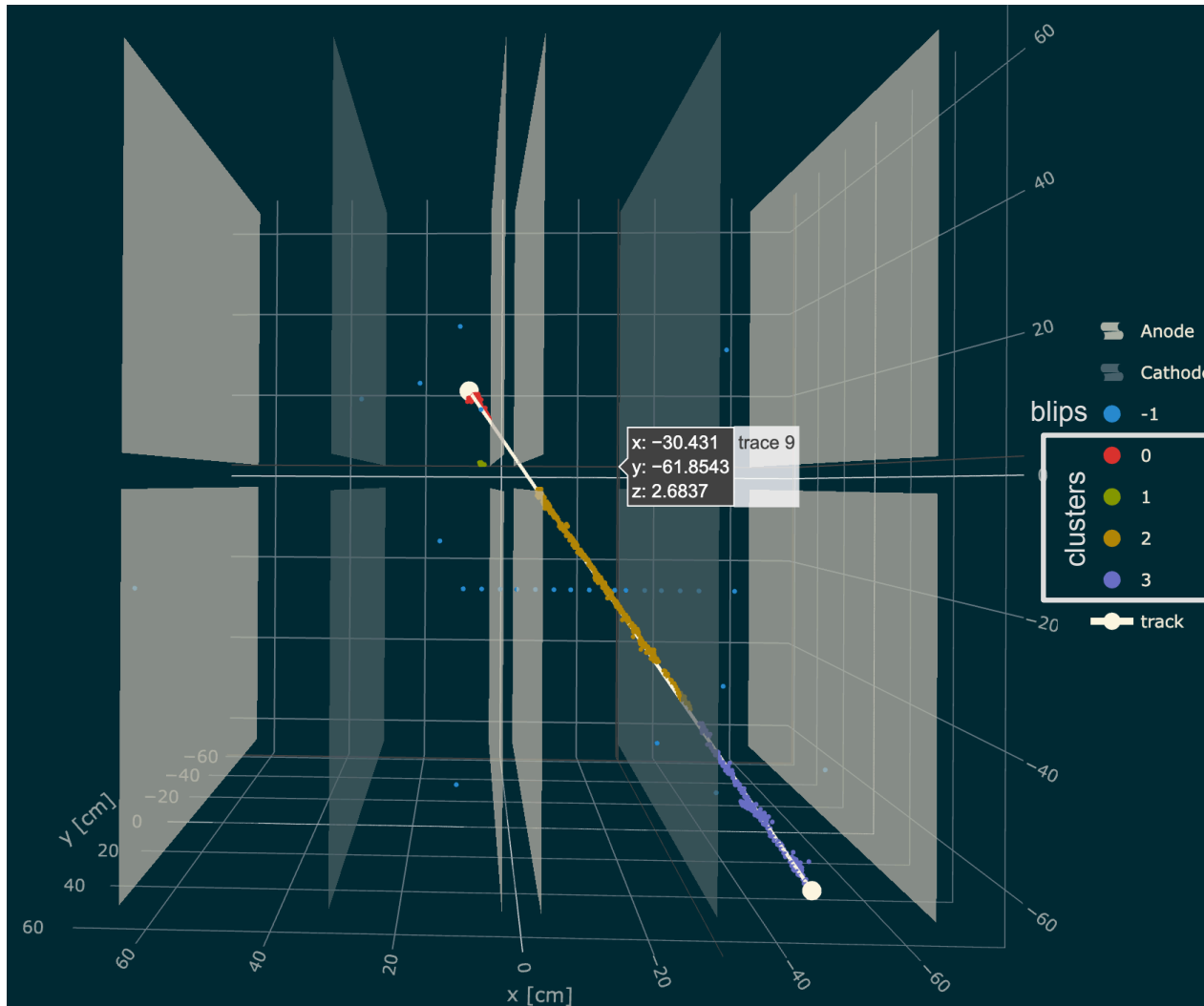


Ideally we expect high correlation on the diagonal if the channel map matched to the LUT.

Sometimes it's difficult to differential SiPMs in the same optical module (group of every 6 SiPMs), in particular adjacent pair, $2n$ v.s. $2n+1$

Definitely some parts are INCORRECT.

Event Selection



- Align each TPC w/ trigger time (T0)
- Run dbscan on charge image (min_samples=5, eps 2.5=cm)
- Remove small blips
- Fit the big clusters with a straight line and locate end points
- Possible to cut on goodness-of-fit to remove non-track like events (shower, multiple objects ...)

Current Status (Aug 7)

- One-to-one matching for charge and light event
- T0 correction based on trigger time
- Raw waveform processing
- SiPM channel maps
 - No show stopper as long as we pick all active channels
- [Nice to have] Light signal process
 - Noise filtering on baseline
 - Gain calibration table, ADC => p.e.
 - Workable in ADC unit
- Crossing track selection

Next Task: Gather all above code pieces and form a data processing workflow

Work in Progress ...

The screenshot shows a GitHub repository page for '2x2-dataprep' under the organization 'CIDeR-ML'. The repository is private. The main branch is 'main'. The repository contains three files: '.gitignore', 'README.md', and 'README'. The README file is highlighted, showing the title '2x2-dataprep' and the description 'Data preparation tools for CIDeR-ML'. The right sidebar shows the 'About' section with the description 'Data preparation tools for CIDeR-ML', '0 stars', '4 watching', and '0 forks'. The 'Releases' section shows 'No releases published' with a link to 'Create a new release'. The 'Packages' section shows 'No packages published' with a link to 'Publish your first package'. The 'Code' button is highlighted with a red box.

Navigation: Code (highlighted), Issues, Pull requests, Actions, Projects, Security, Insights, Settings

Repository: 2x2-dataprep (Private)

Branch: main

Search: Go to file

Commit: kvtsang Initial commit (dfaa5c8 · 4 hours ago) 1 Commit

File	Commit	Time
.gitignore	Initial commit	4 hours ago
README.md	Initial commit	4 hours ago

README

2x2-dataprep

Data preparation tools for CIDeR-ML

About

Data preparation tools for CIDeR-ML

- Readme
- Activity
- Custom properties
- 0 stars
- 4 watching
- 0 forks

Releases

No releases published
[Create a new release](#)

Packages

No packages published
[Publish your first package](#)

Un-Flow 2x2 Data File

Matched Charge-Light events in plain HDF5 format

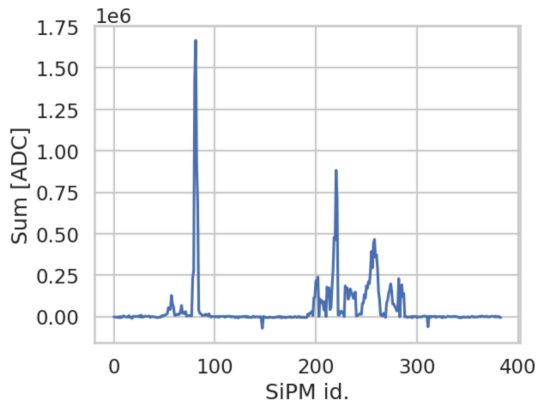
```
(nersc-flow) login27:2x2_beam$ h5ls -r test.h5
/
  Group
/qpts          Dataset {90052/Inf, 4}
/qpts__idx     Dataset {148/Inf}
/sum_adc      Dataset {147/Inf, 384}
/sum_adc_idx  Dataset {148/Inf}
/tpcs         Dataset {90052/Inf}
/tpcs_idx     Dataset {148/Inf}
/wvfm         Dataset {147/Inf, 8, 64, 1000}
/wvfm_idx     Dataset {148/Inf}
```

Charge dataset: qpts

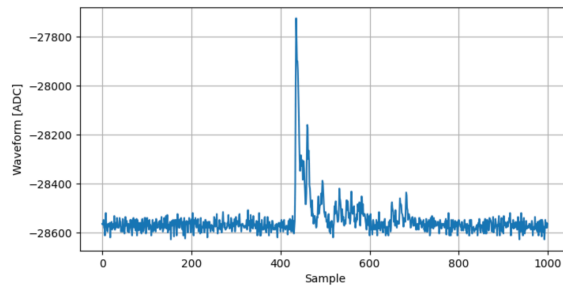
```
qpts__idx[:8]
array([ 0, 232, 483, 1012, 1909, 3042, 3333, 4047], dtype=uint64)

qpts
array([[ 10.149627 , -39.2409 , -25.2971 ,  7.5200195 ],
       [ 44.757607 , -46.7787 ,  23.0801 , 12.841797 ],
       [ 35.609932 , -49.8825 ,  19.0895 ,  5.391309 ],
       ...,
       [-39.569134 ,  51.40669 ,  20.890812 , 20.292286 ],
       [-29.873438 ,  39.767437 ,  8.475613 ,  9.64873 ],
       [-33.433525 ,  40.931362 ,  9.639538 , 10.713086 ]],
      dtype=float32)
```

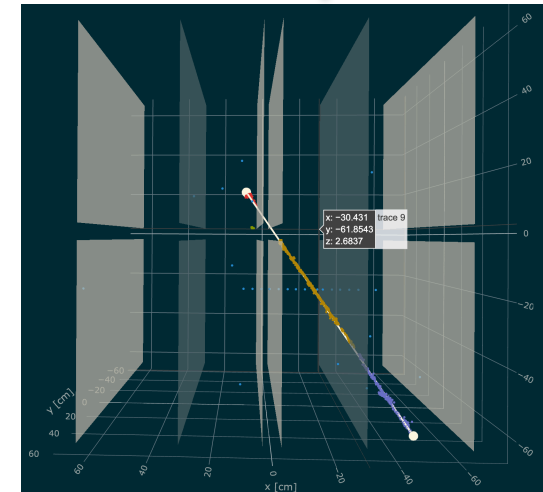
Light Signal (w/ channel map)
Baseline subtracted, sum area



Raw waveform (no channel map)



1-to-1
matching



Goals: 2x2 Data Preparation

At the end of the workshop

- Check trigger matching [~~1 day~~]
 - Reduce ambiguous triggers?
- Validate channel mapping [~~0.5 – 1 day~~]
 - Raw ADC (8x64) => TPC/SiPM (8x48)
 - Cross-check w/ LUT (~close to data)
- Signal Processing [maybe? 1-2 days]
 - Noise filtering
 - Deconvolution
 - Hit finding (determine threshold)
 - Calibration ADC => p.e.
- Event selection [~~1 day~~]
 - Clean sample of isolated long track + light signal
- Gather above codes [~~1 day~~]
- Documentation [1 day]

Prepare a workflow to process 2x2 data for SIREN track calibration.