Augmented Signal Processing in Liquid Argon Time Projection Chambers with a Deep Neural Network

- Software Effort to Improve LArTPC Reconstruction
- *JINST* 16 (2021) 01, P01036

Haiwang Yu (BNL)

CPAD Instrumentation Frontier Workshop 2021

18-22 March 2021

Signal Processing (SP) of Liquid Argon TPC (LArTPC)

LArTPC is the central detector in many next-gen neutrino experiments, e.g. DUNE, SBN etc.

- Calorimetry and rich topology information

LArTPC wire-readout measures induced charge $\otimes$ response

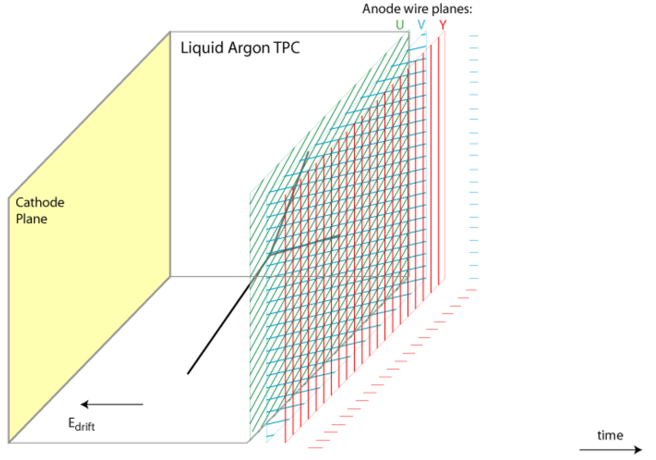
Signal Processing (SP) of LArTPC resolves charge from the original measurement:

$$M(t',x') = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R(t,t',x,x')\cdot S(t,x)dtdx + N(t',x')$$

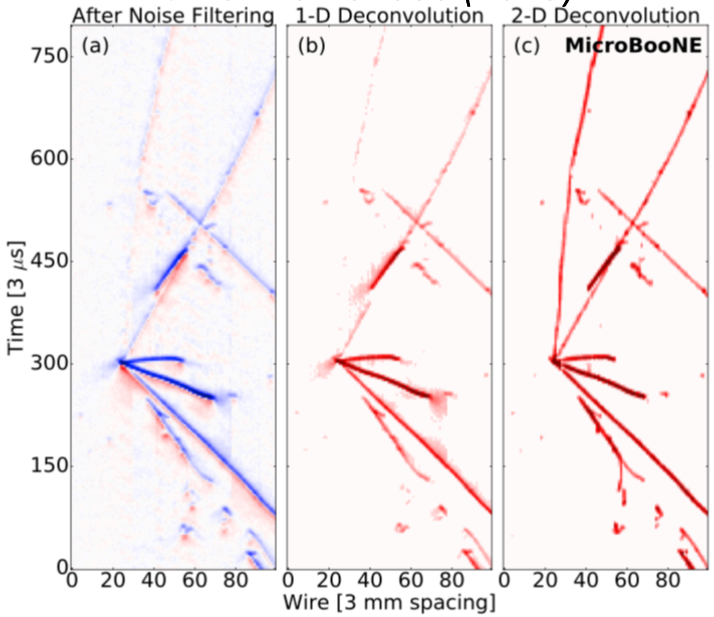
$$S(\omega_t,\omega_x) \sim \frac{F(\omega_t,\omega_x)\cdot M(\omega_t,\omega_x)}{R(\omega_t,\omega_x)} \quad \xrightarrow{IFT} \quad S(t,x)$$

Current state of the art SP algorithm based on 2D deconvolution: *JINST 13 P07006 (2018)*

Signal formation in a 3 plane LArTPC "Single-Phase" wire-readout, from Bo Yu (BNL), https://lar.bnl.gov/wire-cell/

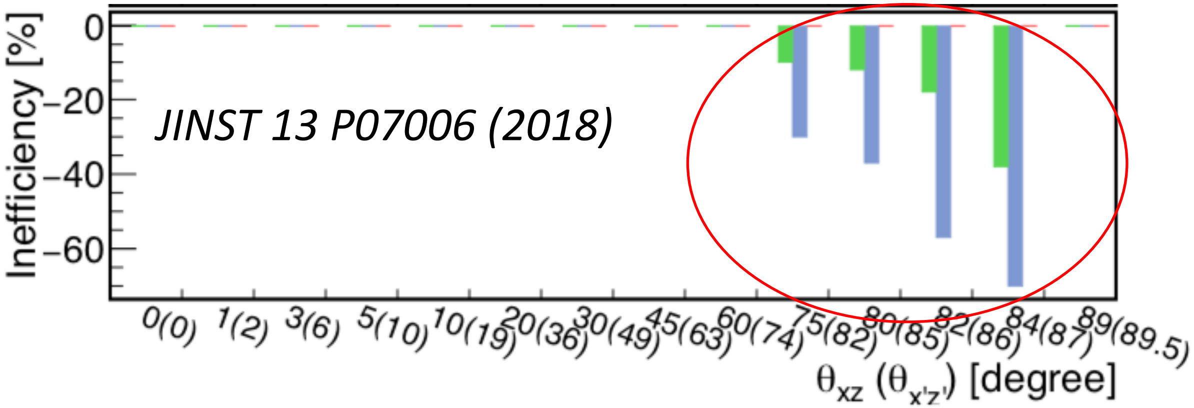


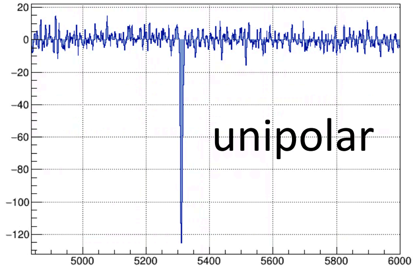
JINST 13 P07006 (2018)

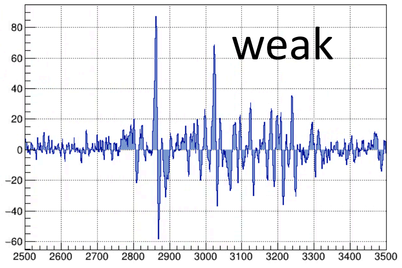


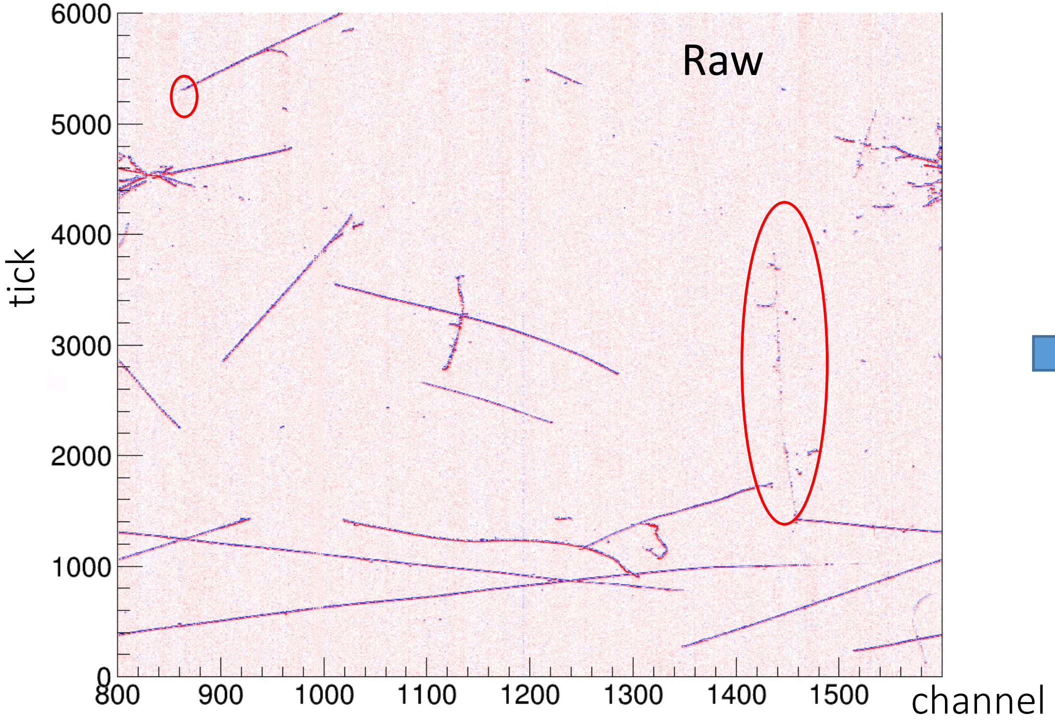
Remaining Challenges in Current Signal Processing (SP)

- "Prolonged Track" – weak signal
- "Tear Drop" - distorted waveform
- Noisy dots - noise

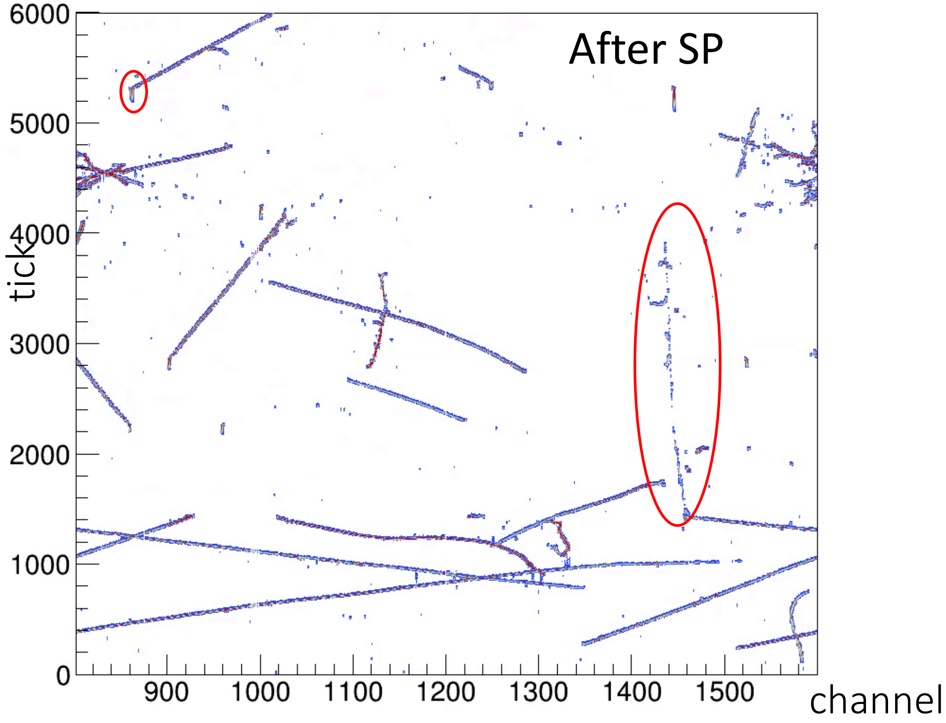






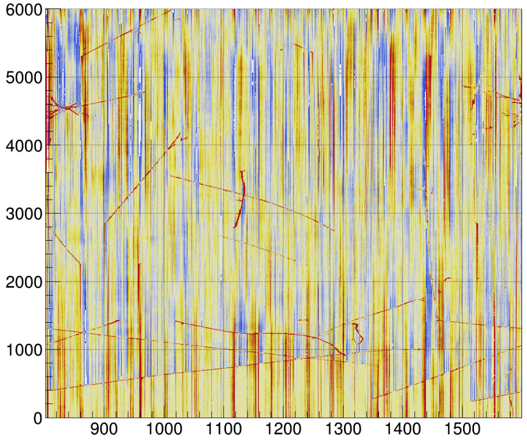


SP

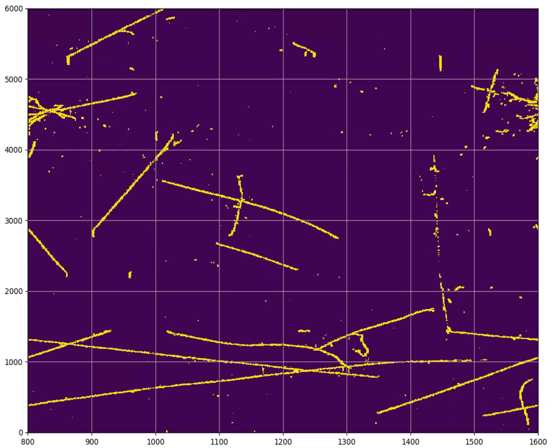


Need to improve “ROI” (Region of Interest) finding

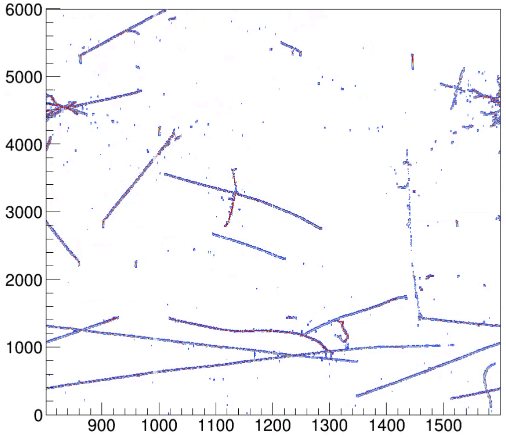
Decon for charge
Waveform ⟶ charge, dense

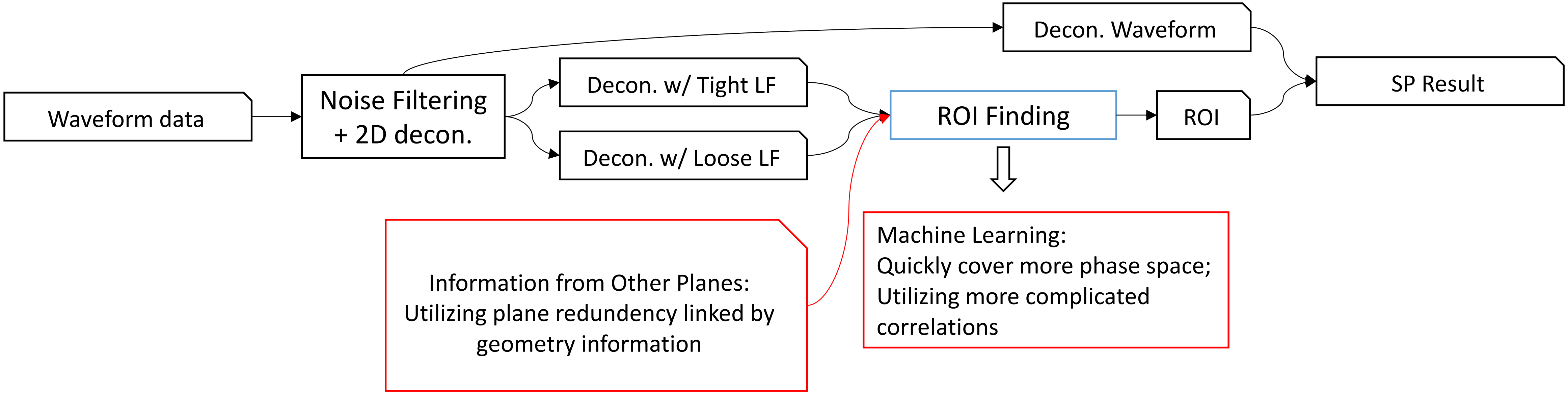


ROI:
Hit finding, sparsify



SP result:
Sparse, charge



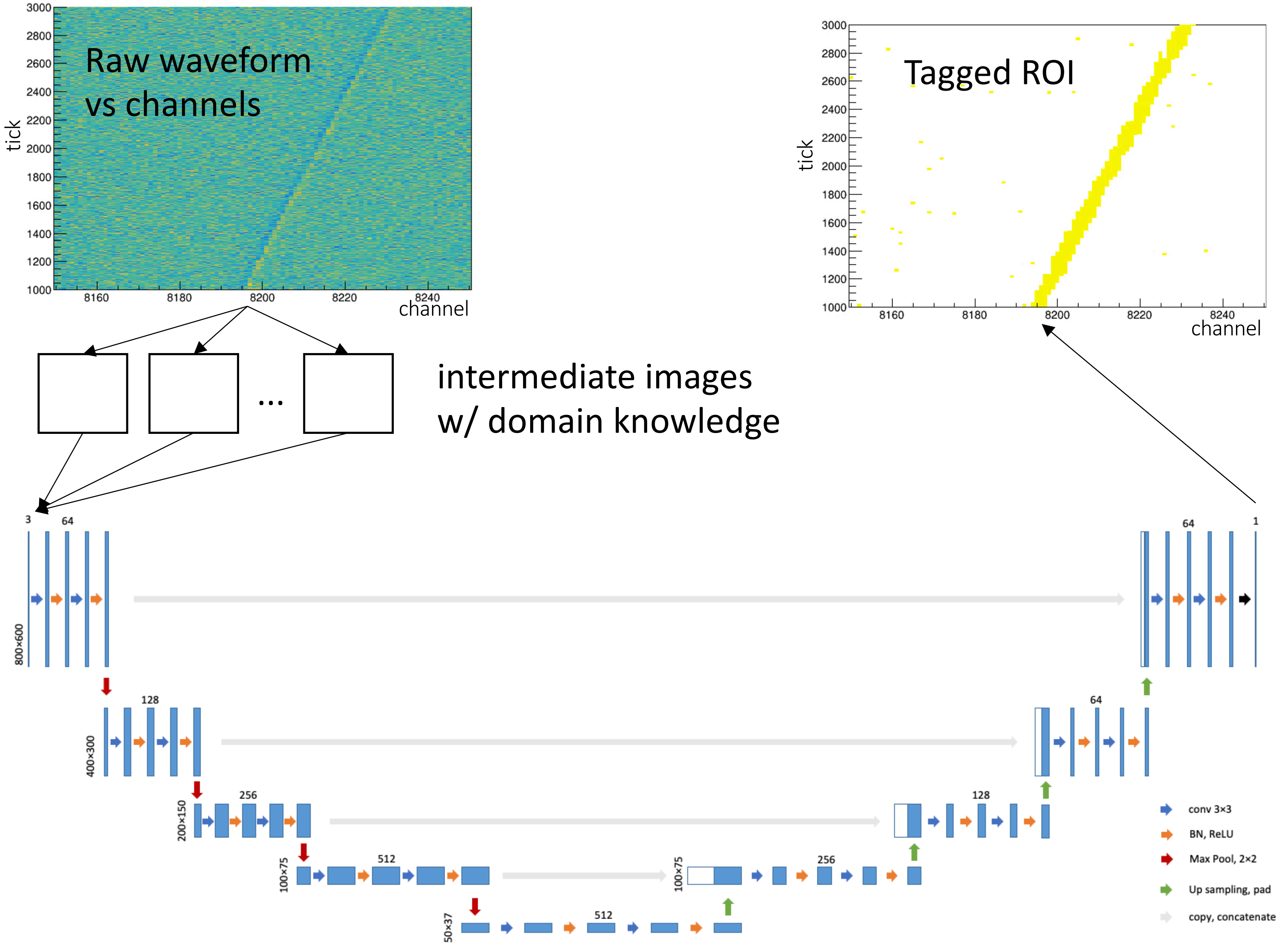


ROI finding as semantic segmentation

2D semantic segmentation – UNet
- arXiv:1505.04597

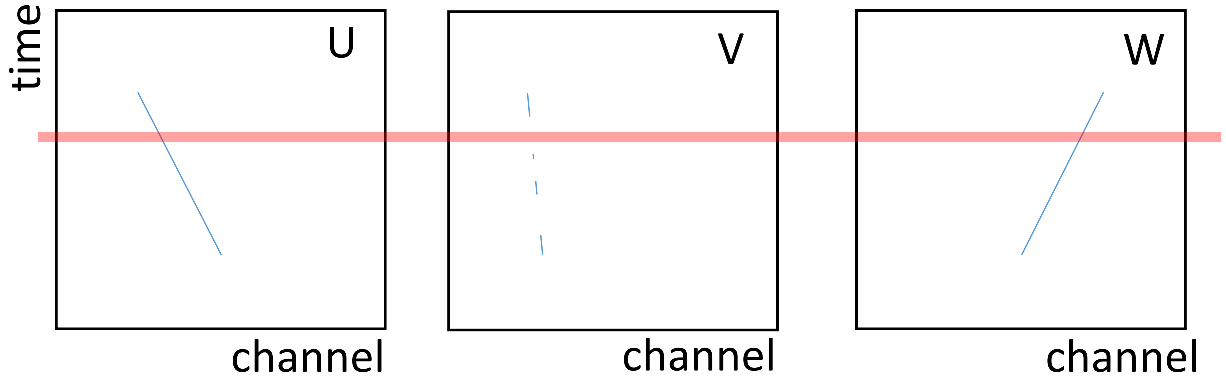
UNet: auto encoder-decoder + skip connections
- Output is sparsely connected components
- Input and output are similar at leading order

Forked a Pytorch implementation:
https://github.com/HaiwangYu/Pytorch-UNet

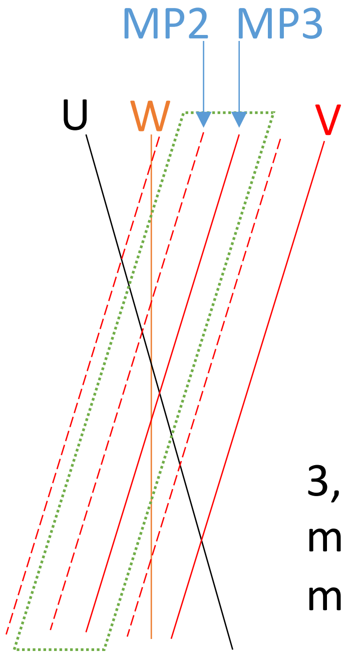


Muti-Plane Information – Geometry

1, make time slices



2, Matching active wires (with initial ROI) from multiple planes



active wire in the time-slice : ref. plane, target plane

in-active wire in the time-slice

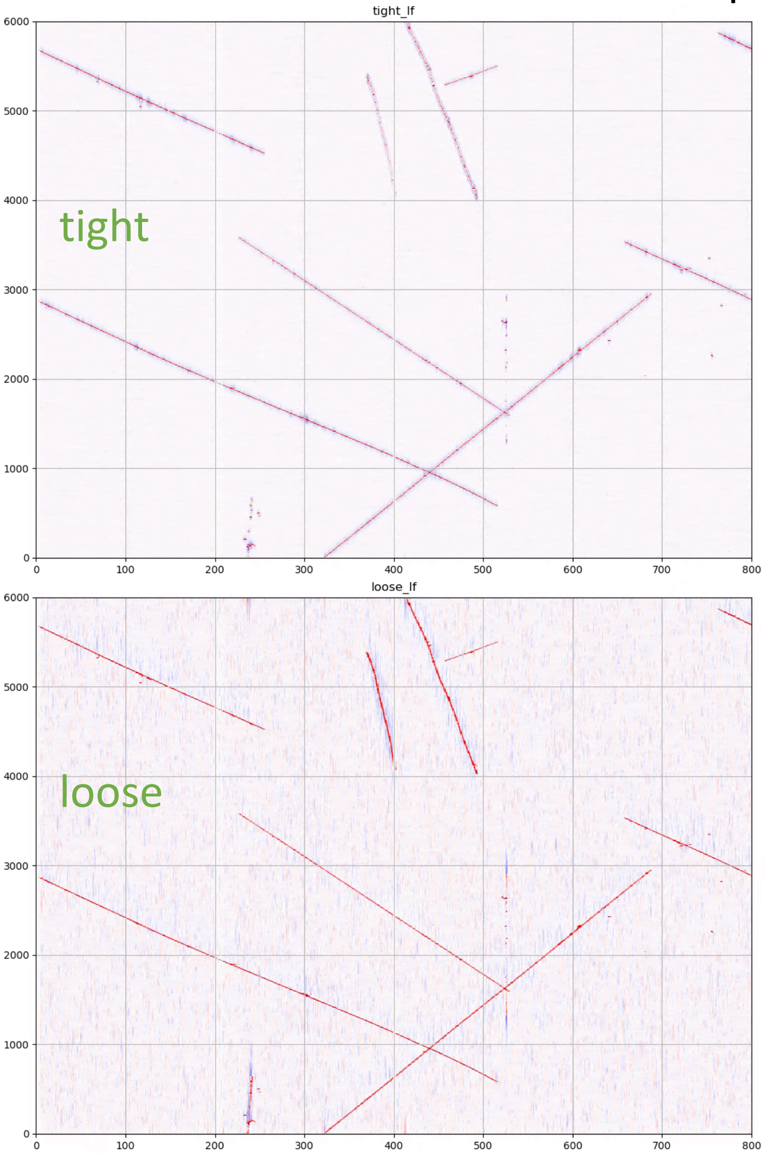
3, On target plane, tag 3-Plane matched ROIs (MP3) or 2-Plane matched ROIs (MP2)

Using other plane information projected to the target plane: utilizing plane redundancy

- Inspired by Wire-Cell '3D Imaging'.
  - *arxiv:1803.04850*
- Key algorithm "Fast projection" realized using "RayGrid" tools: https://github.com/WireCell/wire-cell-docs/blob/master/presentations/updates/20190321/latexmk-out/img.pdf

Input and label for DNN - ProtoDUNE Single-Phase simulation

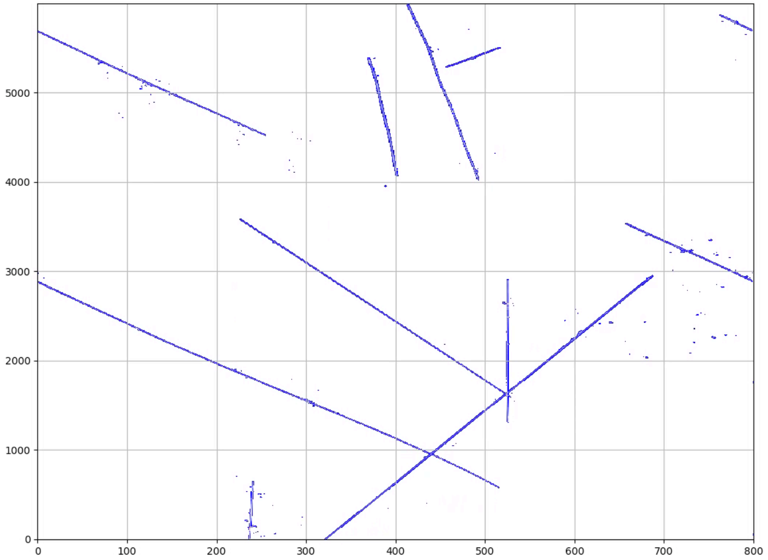
Input candidates



tight/loose: deconvolution with tight/loose low frequency filter

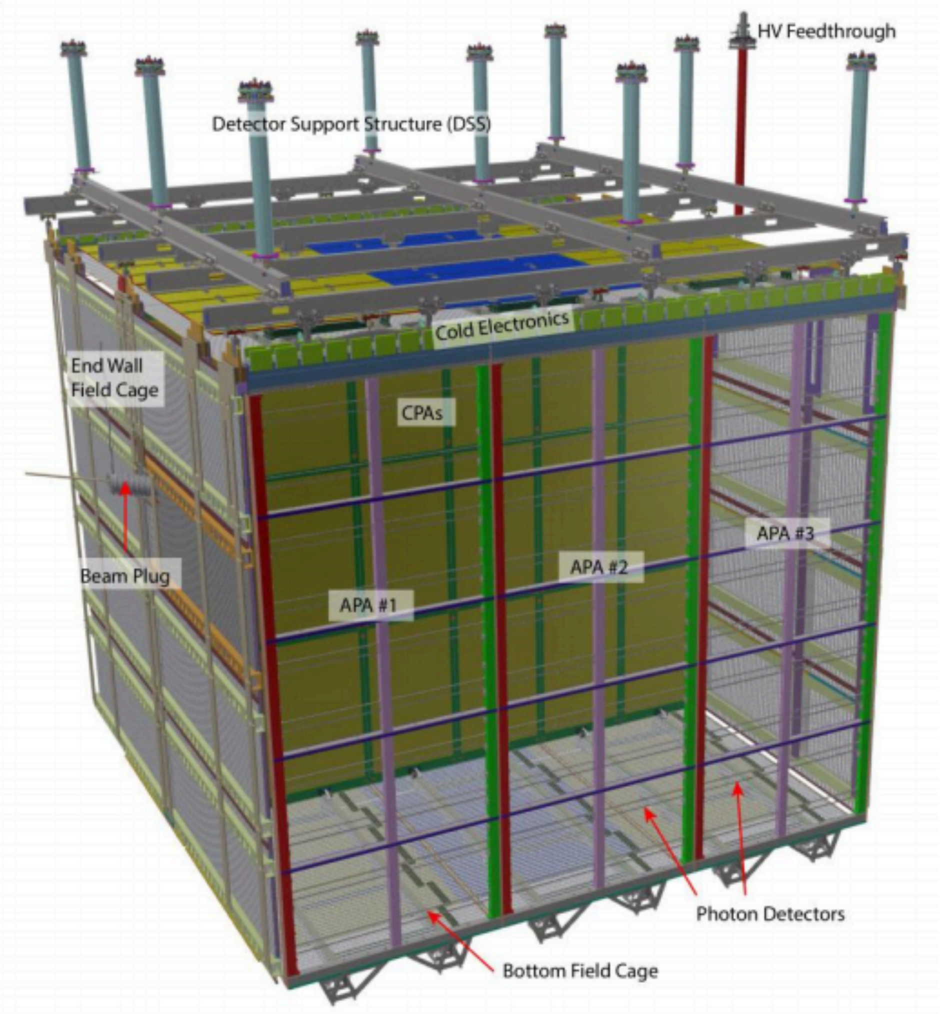
label was made by rasterization of charge depos – 2D gaussian response

Label



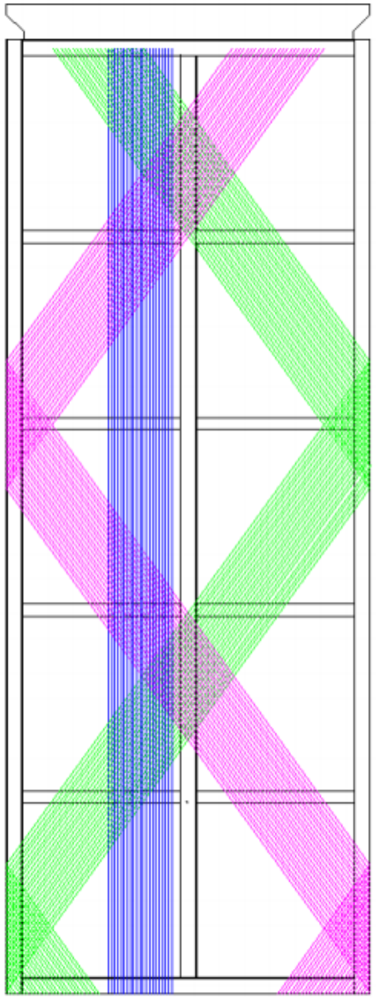
ProtoDUNE Single-Phase

ProtoDUNE Single-Phase (the NP04 experiment at CERN) is designed as a test bed and full-scale prototype for the elements of the first far detector module of DUNE (next-generation, long-baseline neutrino oscillation experiment).



major components of the ProtoDUNE-SP TPC from arxiv:1706.07081

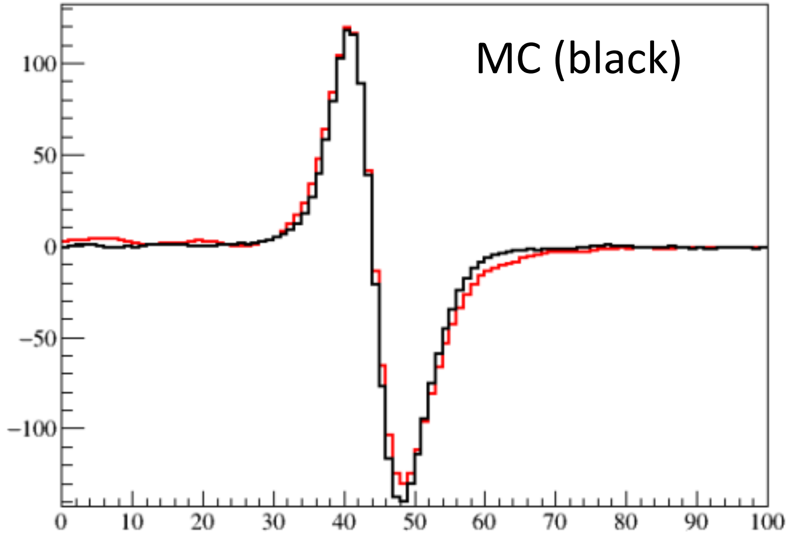
Sketch of a ProtoDUNE-SP APA from arxiv:1706.07081



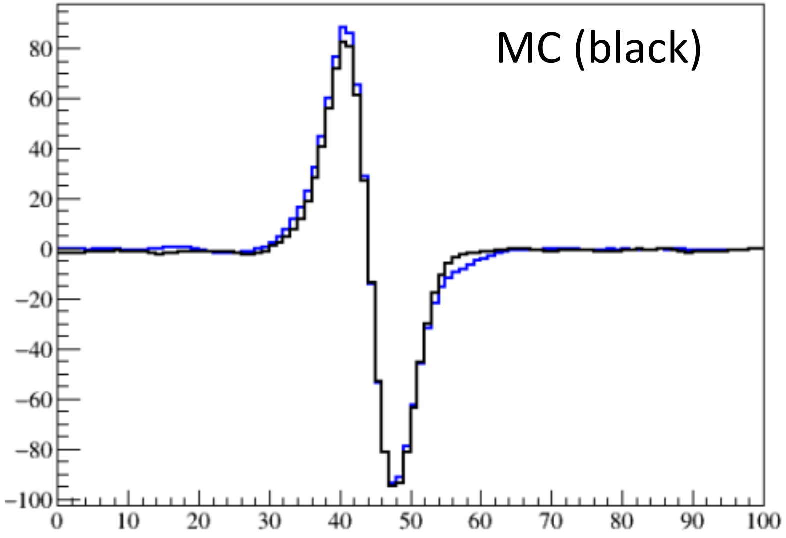
Signal Simulation Validation

2D convolution based simulation for LArTPC
- *JINST 13 P07006 (2018)*
- Software stack:
  - LArSoft for the generator and Geant4 simulation
  - Wire-Cell Toolkit for the detector response

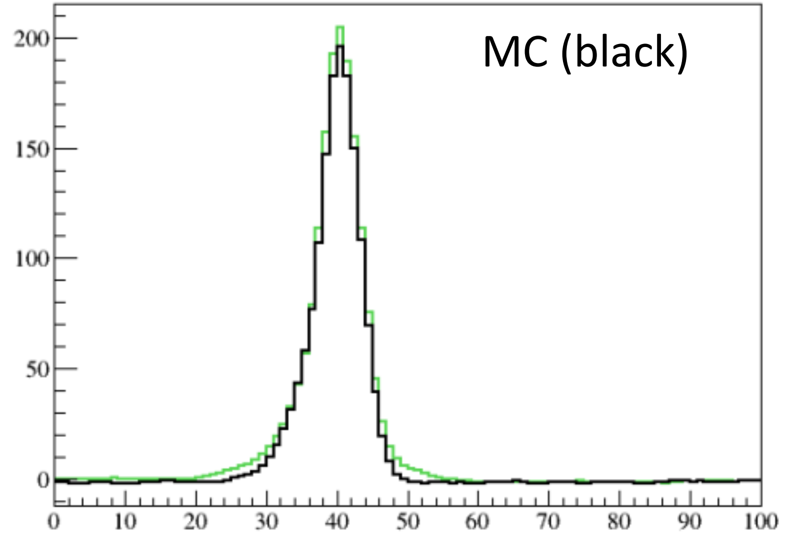
Average waveform: data vs. MC



U plane
MC scaling: 1.6
Tick: [570, 970]
Ch: [8050, 8200]



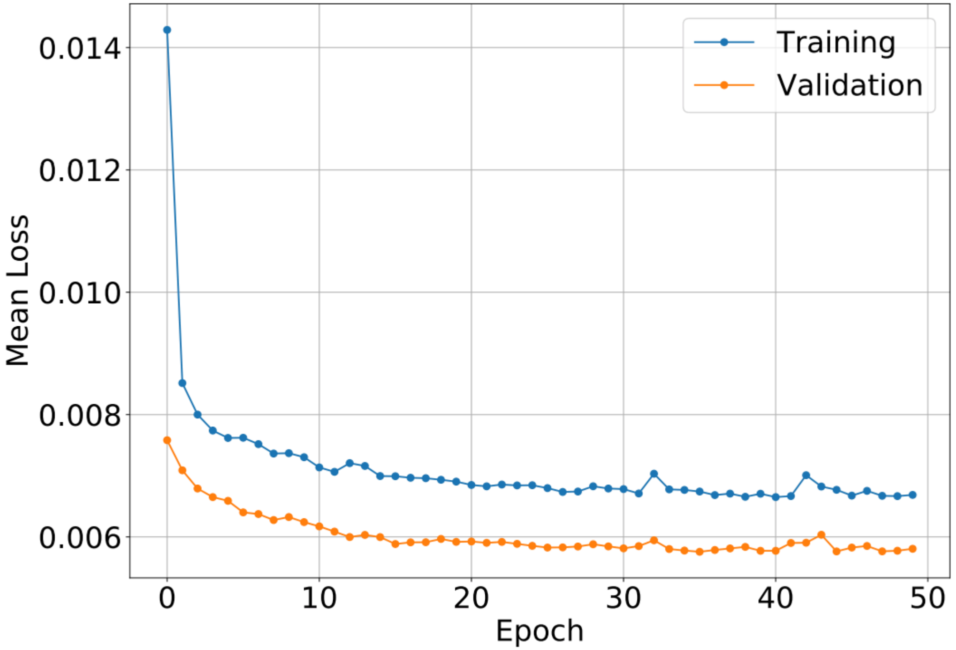
V plane
MC scaling: 1.6
Tick: [570, 970]
Ch: [8950, 9200]



W plane
MC scaling: 1.6
Tick: [570, 970]
Ch: [9500, 9750]

Efficient network – 500 samples already provide good performance – 6 min/epoch

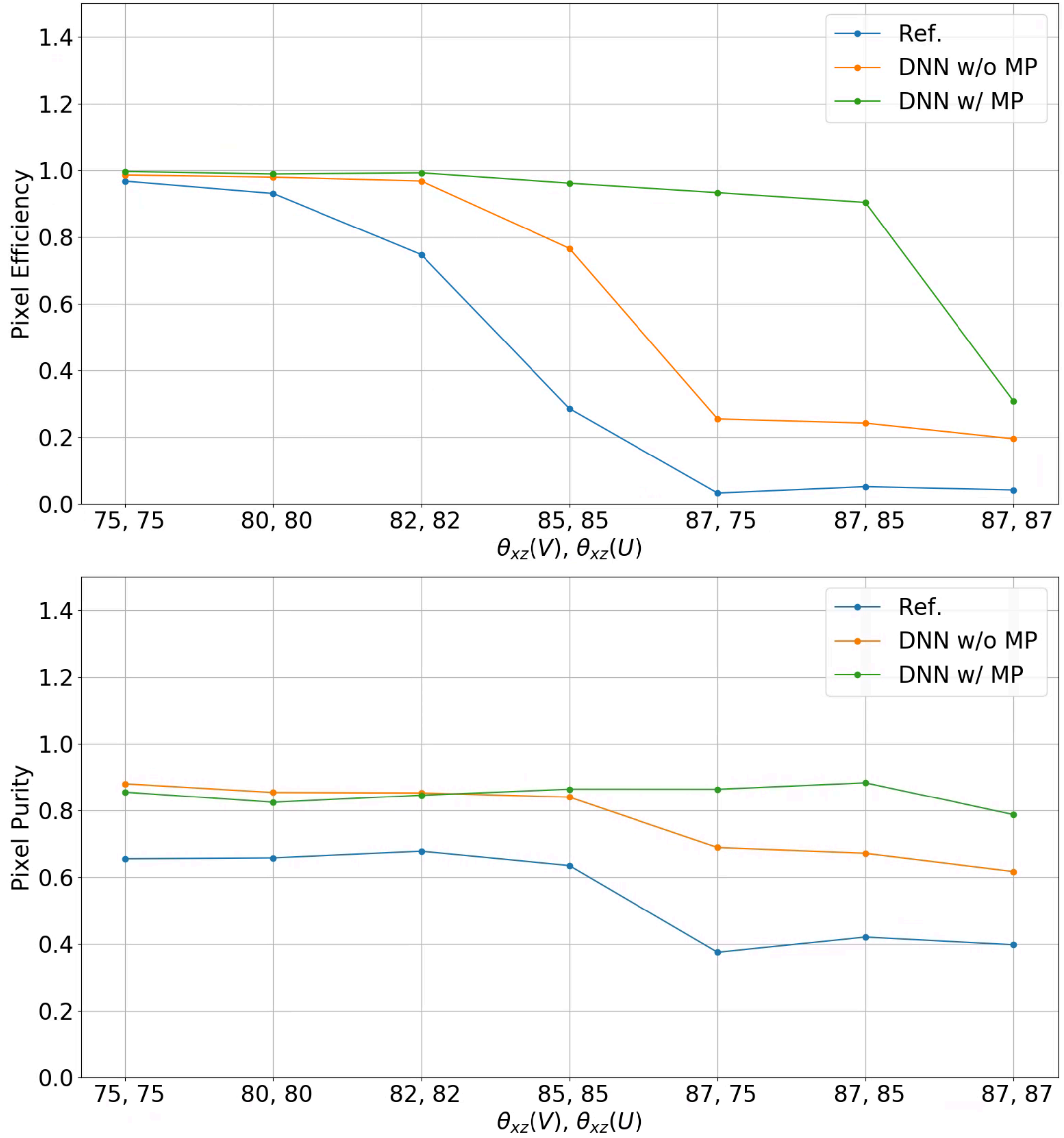
- Platform: I9-9900K, 32 GB memory, Nvidia GTX 2080 Ti 11GB, Samsung 970 500GB NVMe SSD
- Sample: 500 APA Planes using cosmic generator (450 training, 50 validation)
- Loss: cross entropy
- Optimizer: Stochastic Gradient Descent with momentum



Evaluation with simulation

- DNN ROI finding has higher efficiency and purity than the reference ROI finding especially at large angles.
- More effective on tracks have asymmetric angles induction planes
- DNN w/o MP information performs slightly better than Ref. but much worse than DNN w/ MP information

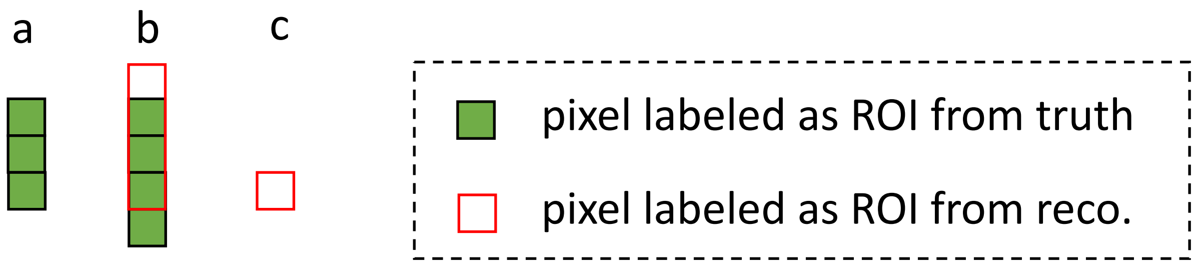
ROI finding on V plane



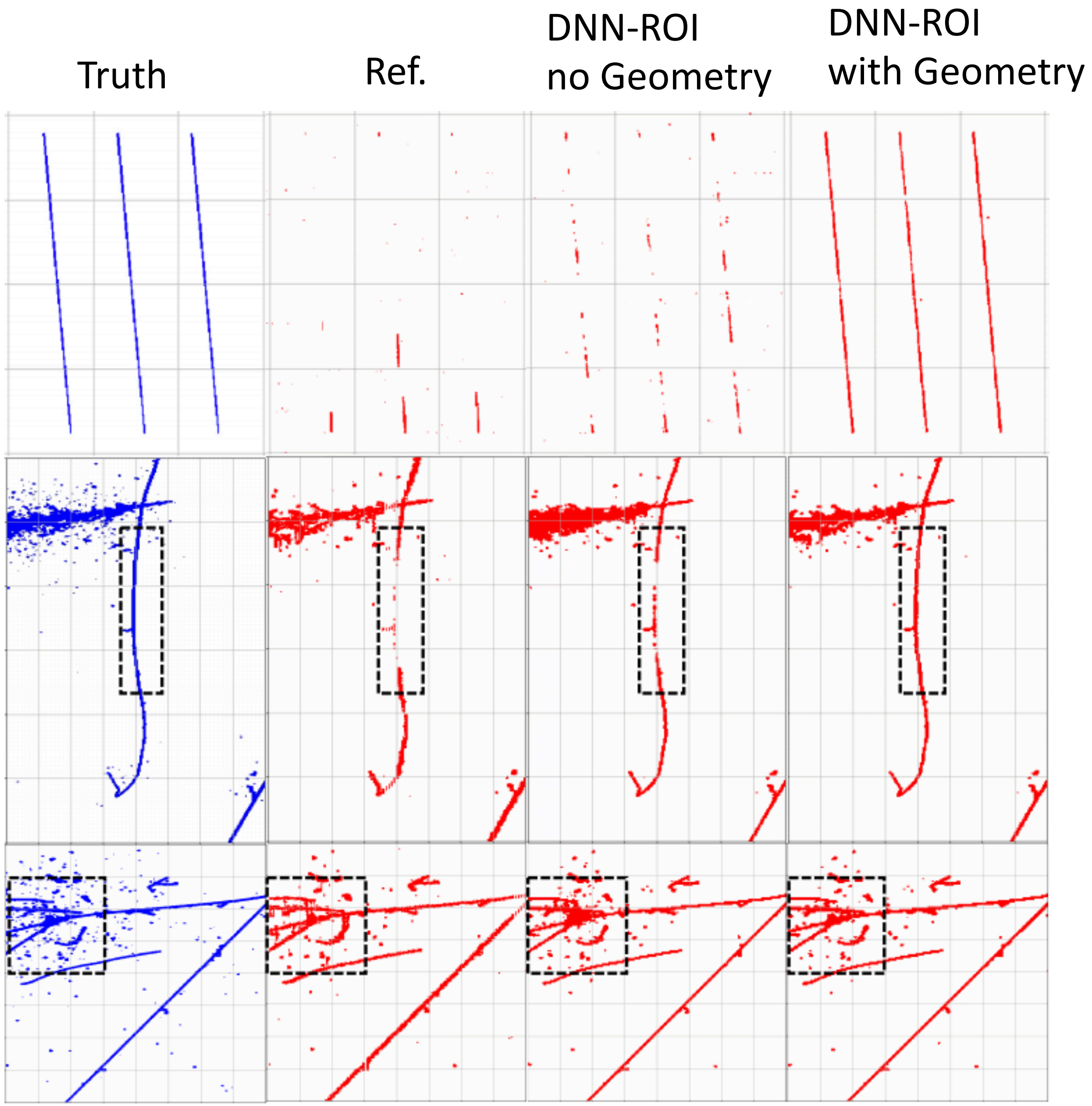
pixel based evaluation:

eff = (0 + 3 + 0)/(0+3+4) = 3/7

purity = (0 + 3 + 0)/(0+4+1) = 3/5



Example ROI finding event displays on simulated data



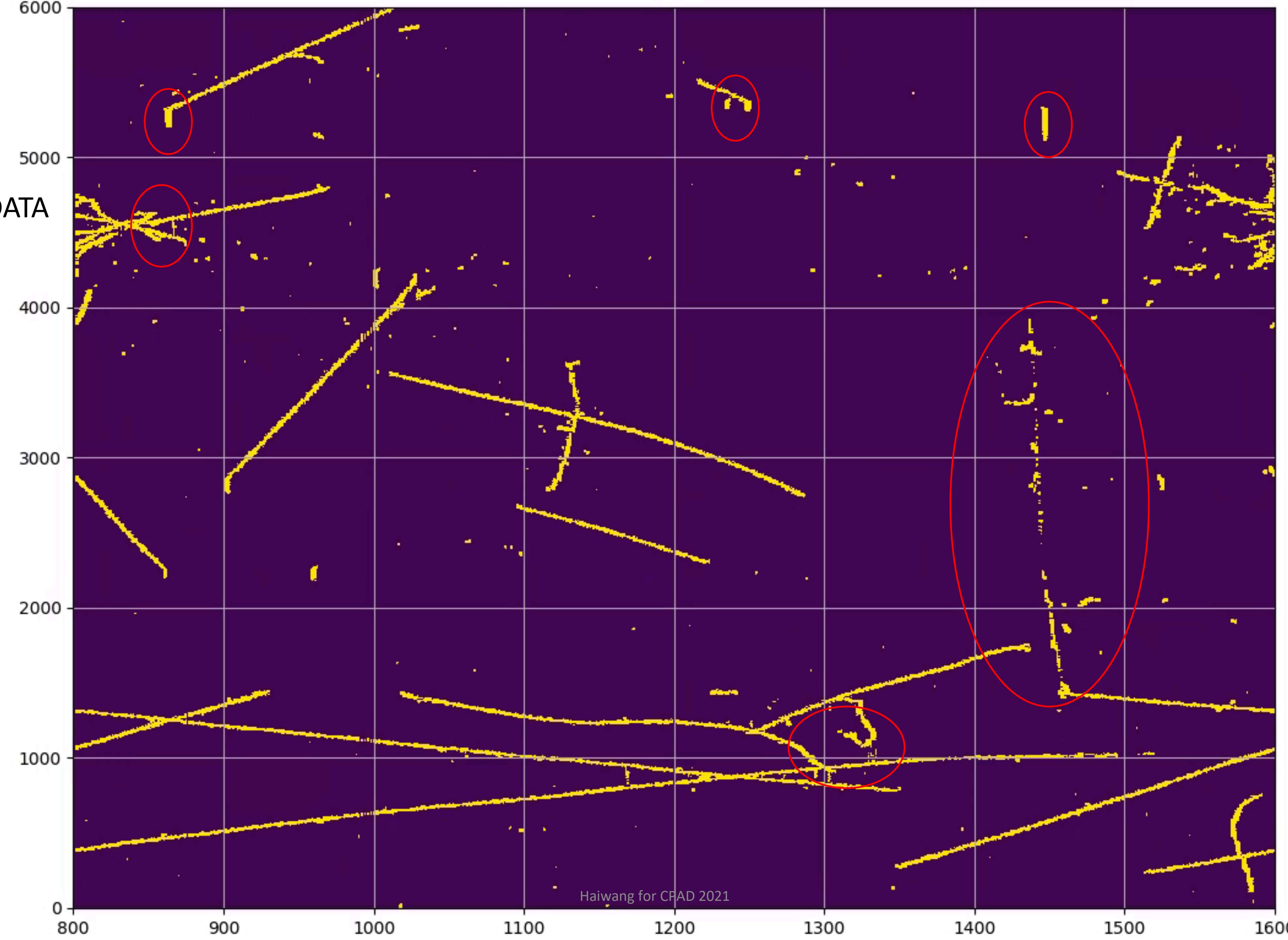
straight-line prolonged tracks

cosmic track with a section featuring a large projection angle

an interaction vertex of a charged pion with many activities

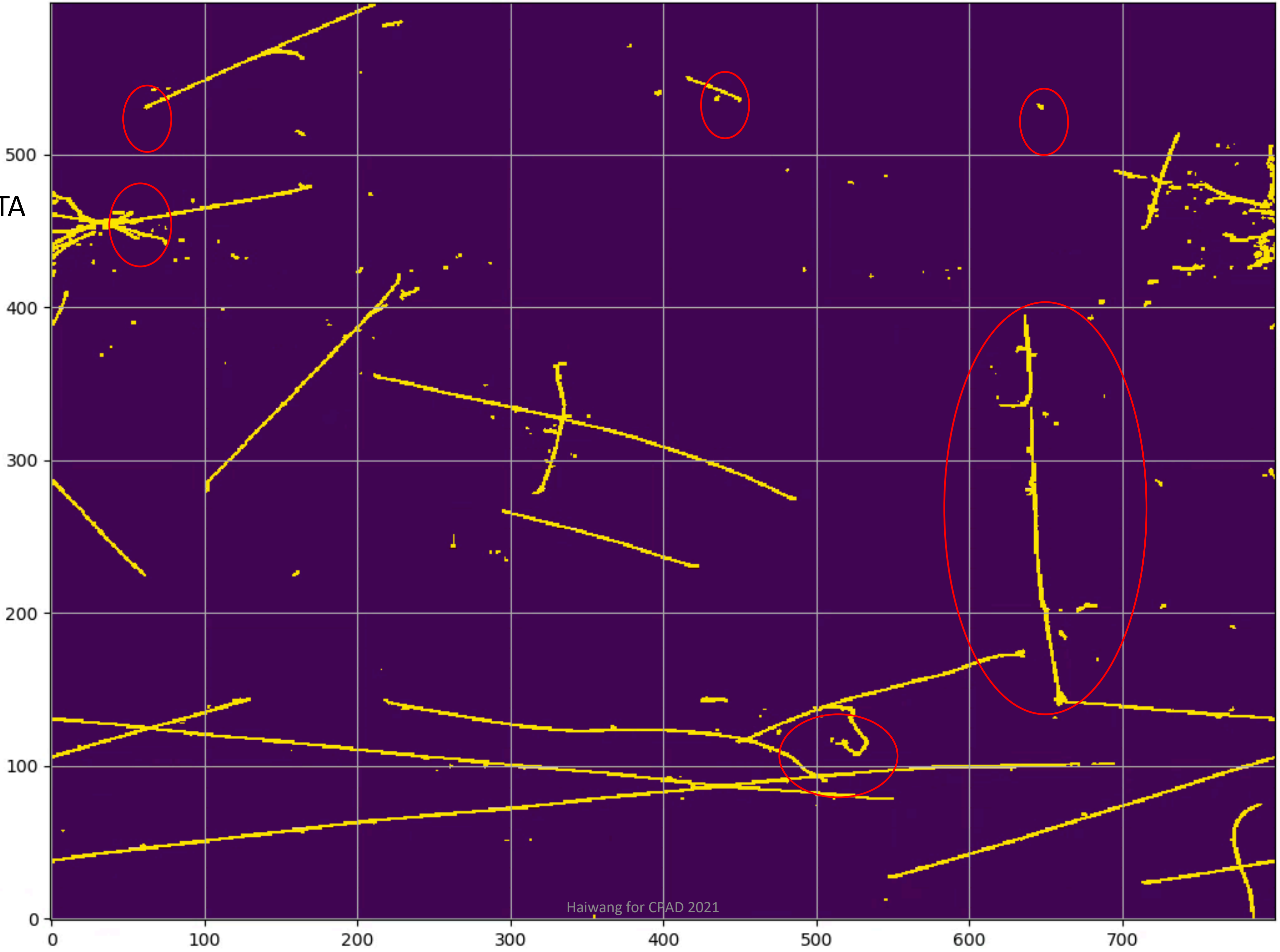
Reference SP

ProtoDUNE DATA
run: 5145
subRun: 1
event: 26918
V plane



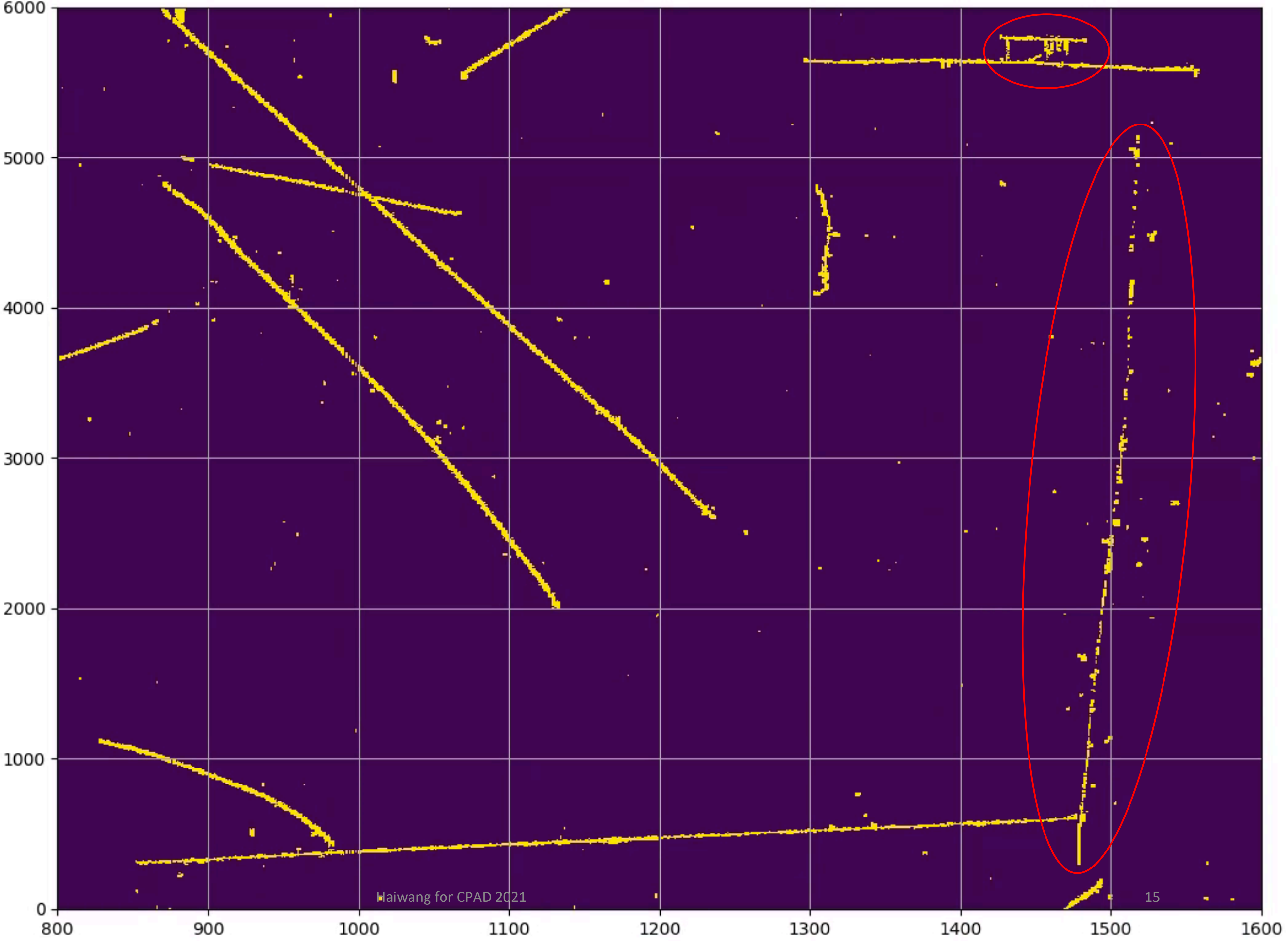
DNN-SP

ProtoDUNE DATA
run: 5145
subRun: 1
event: 26918
V plane



Reference SP

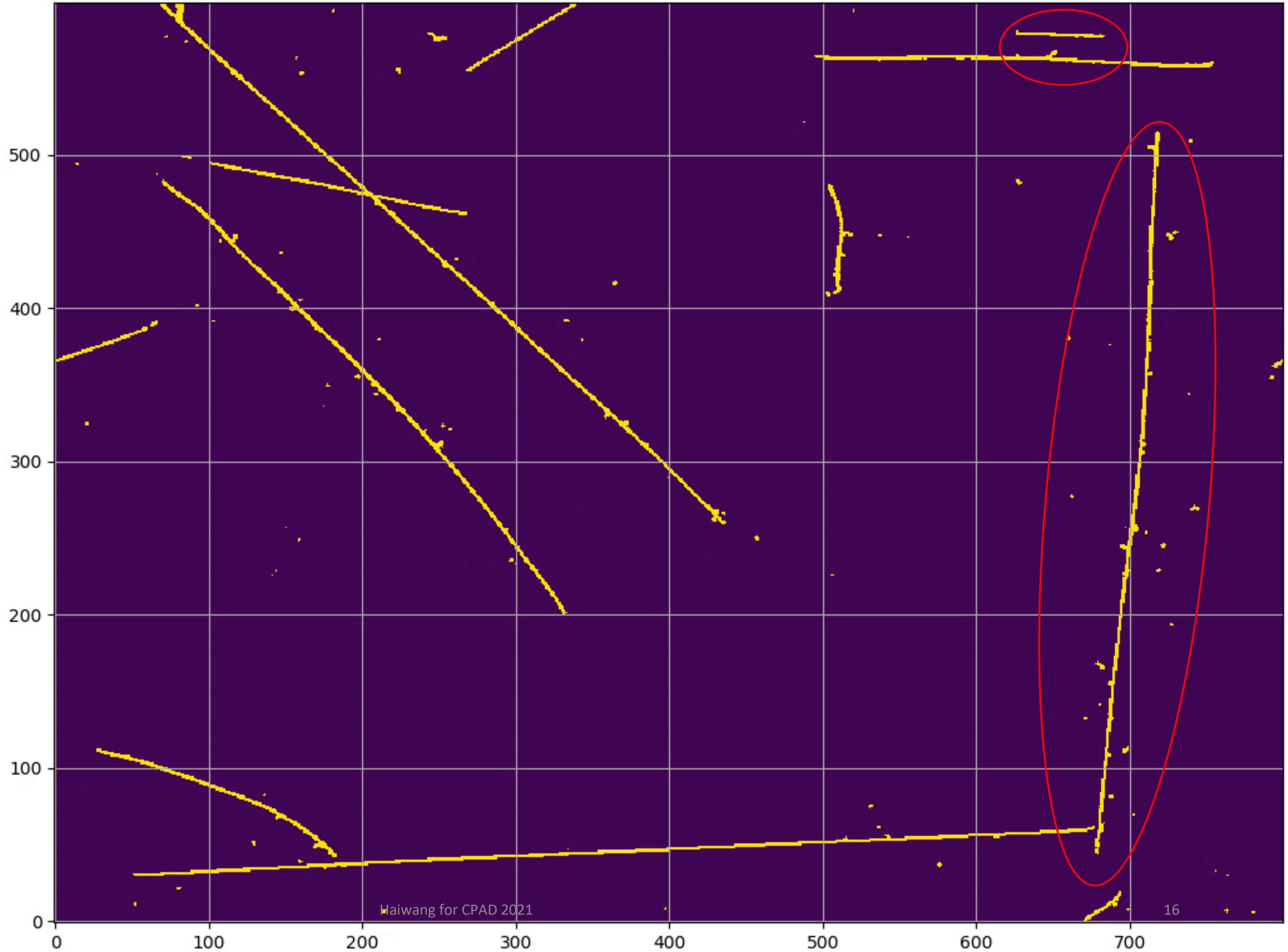
ProtoDUNE-SP Data
run: 5145
subRun: 1
event: 26945
V plane



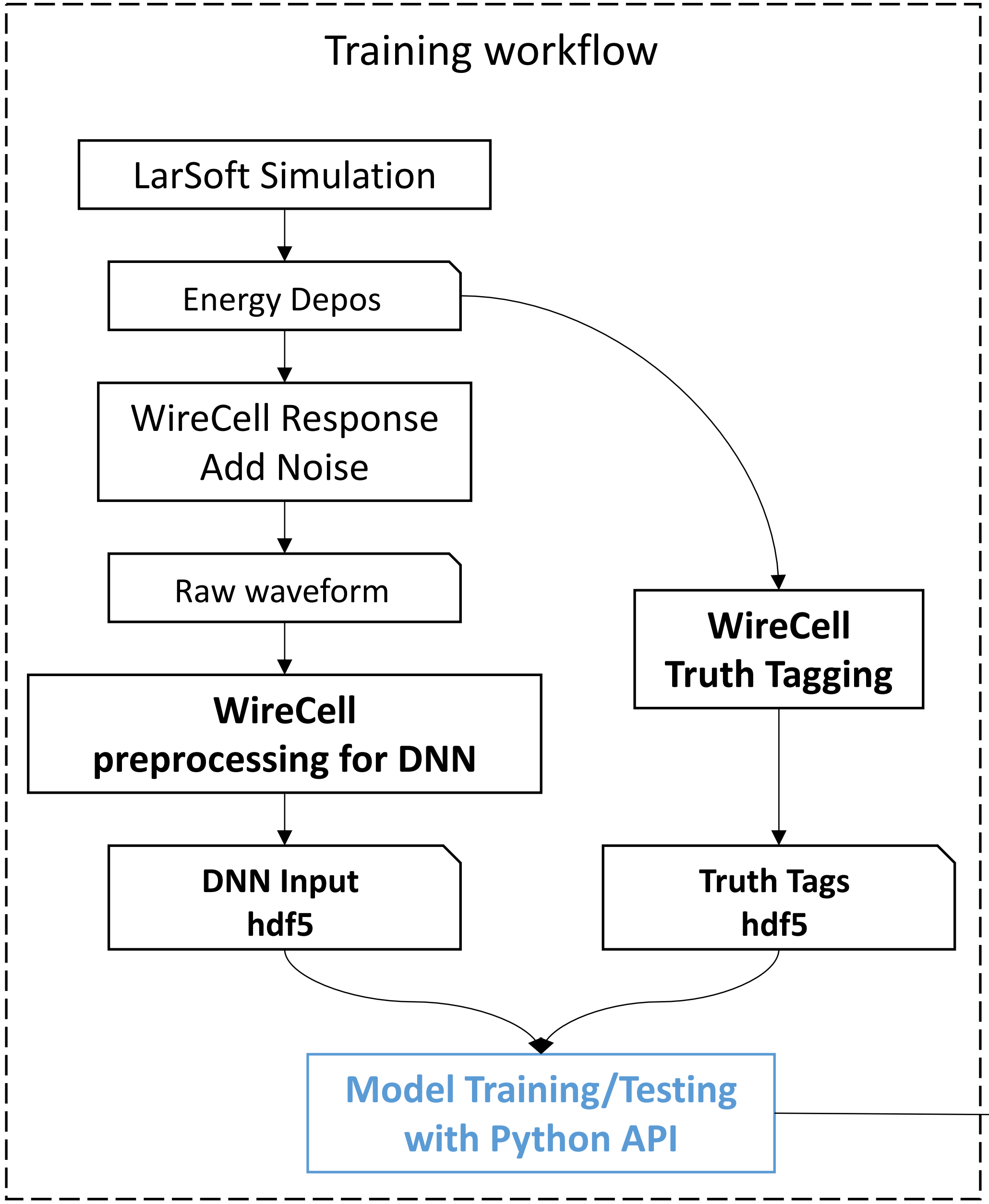
ProtoDUNE: how and why
L. Manenti: X02.00001
and other ProtoDUNE Talks

DNN-SP

ProtoDUNE-SP Data
run: 5145
subRun: 1
event: 26945
V plane

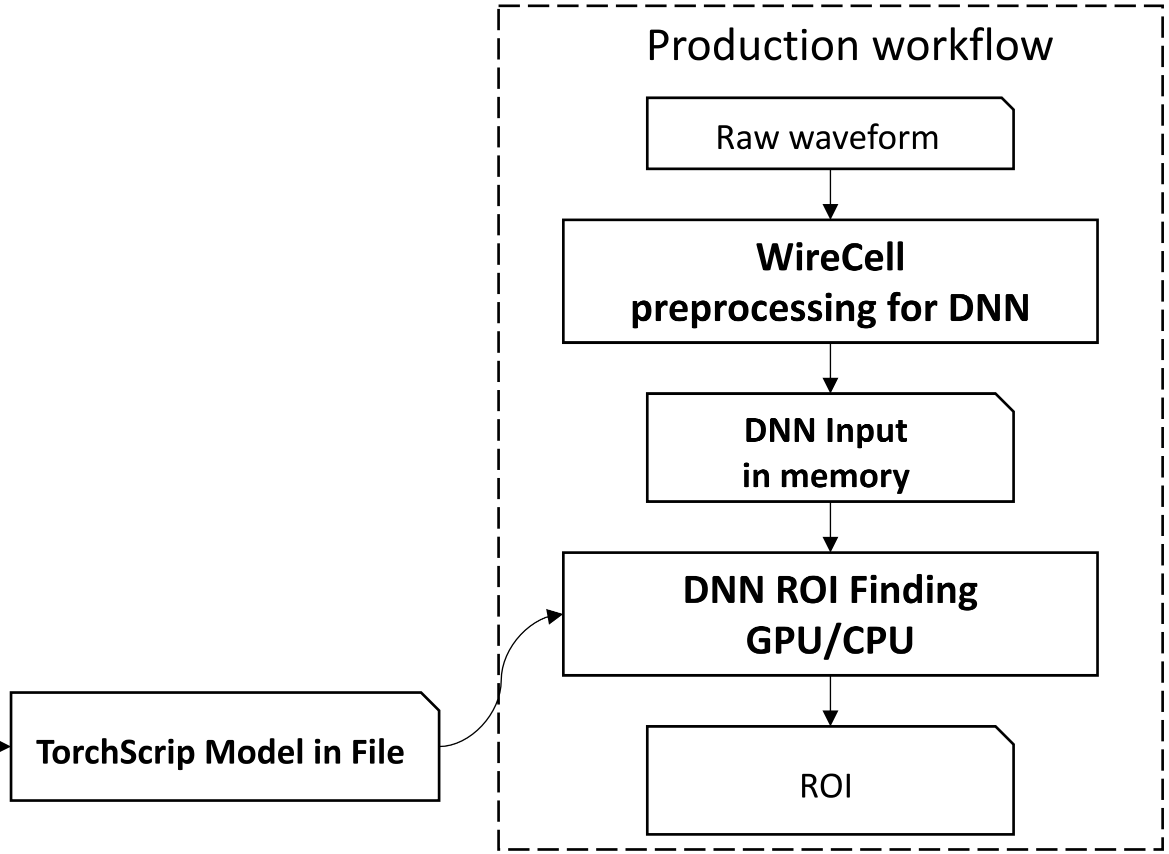


Deep-Learning in Wire-Cell Toolkit



Established initial machinery for Deep-Learning in Wire-Cell Toolkit
https://github.com/WireCell/wire-cell-toolkit

- Data preparing in LarSoft/Wire-Cell
- Training with python
- Production with C++



Summary

DNN ROI finding algorithm provides a software improvement on LArTPC reconstruction

- based on the the hardware feature - redundant wire plane
- combines domain knowledge with machine learning
- significant improvement on both simulation and ProtoDUNE-SP data

Ref. *JINST 16 (2021) 01, P01036* for more.

Wire-Cell: https://lar.bnl.gov/wire-cell/

Backup

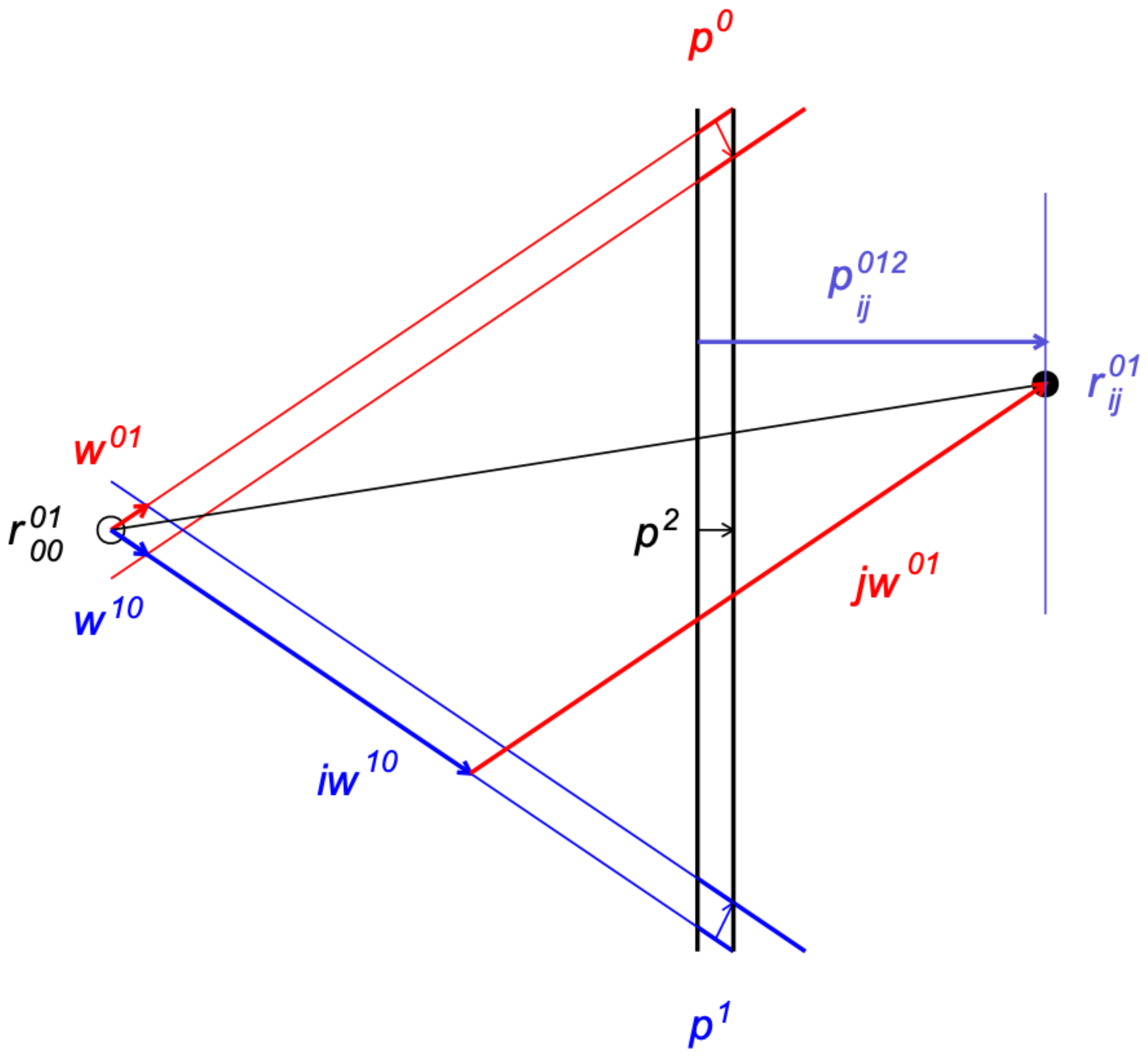
RayGrid fast projection

Non-orthogonal coordinate system

$$r_{ij}^{lm} = r_{00}^{lm} + jw^{lm} + iw^{ml}$$

Projection on the "n" plane

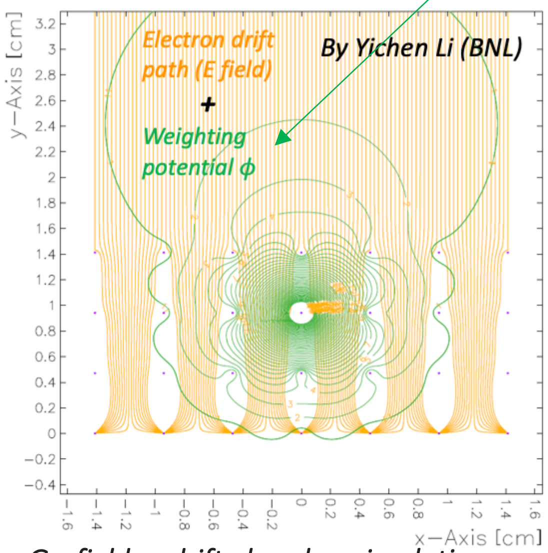
$$p_{ij}^{lmn} = (r_{ij}^{lm} - c^n) \cdot \hat{p}^n$$



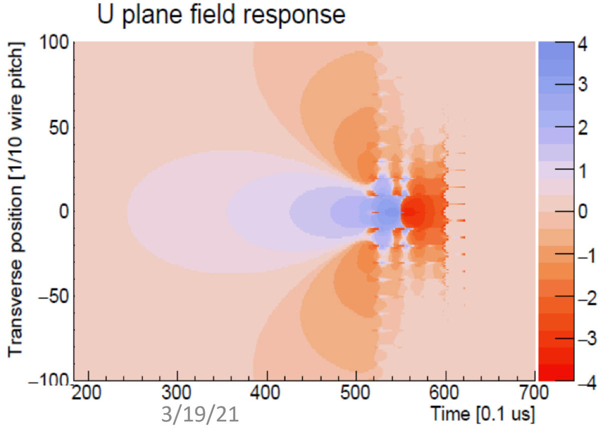
WireCell Detector Response Simulation

Ramo's theorem

$$i = e\vec{v} \cdot \vec{E}_v = e\vec{v} \cdot (-\nabla\phi)$$



Garfield, a drift-chamber simulation program User's Guide, Veenhof, R. (1993)



Hanyu Wei' talk, DUNE collab. meeting May 2019

Response matrix:

- Field response from Garfield simulation
- Electronic response

2D convolution – realized by FFT and IFFT

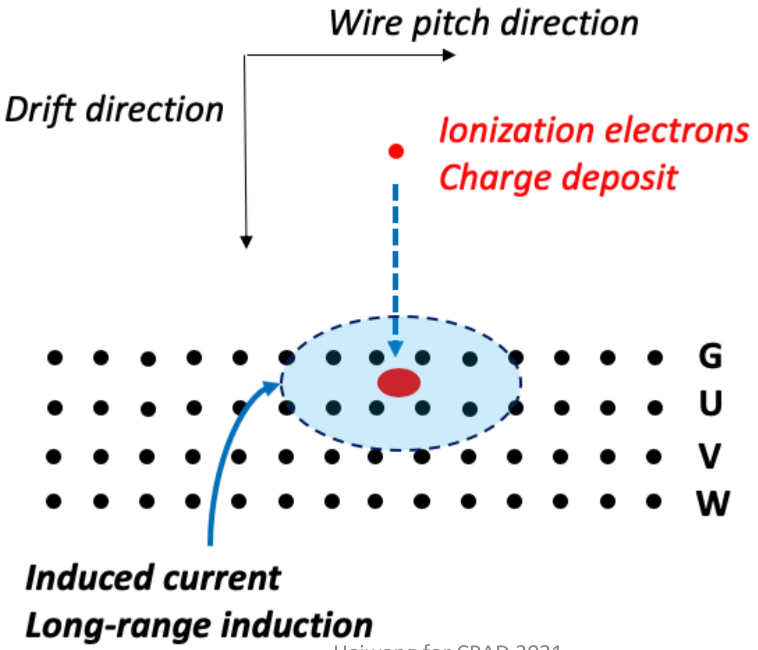
Add inherent electronics noise

- M. Diwan. *Basic mathematics of random noise part ½*

Refer: *JINST 13 P07006 (2018)*

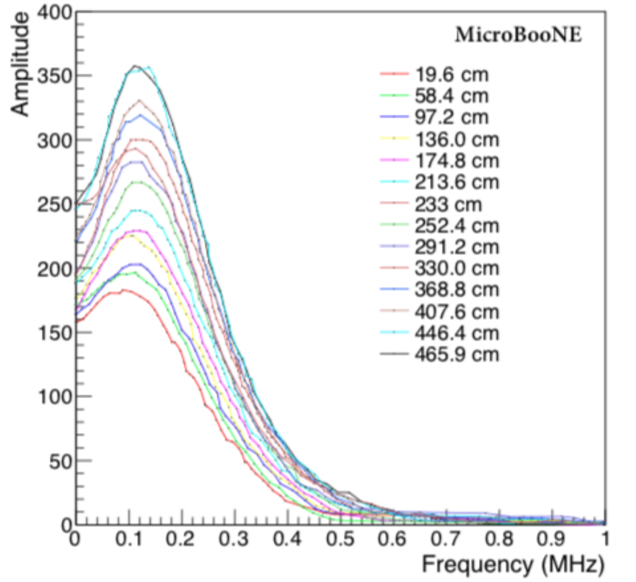
2D Conv.

$$M(t',x') = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R(t,t',x,x')\cdot S(t,x)dtdx + N(t',x')$$



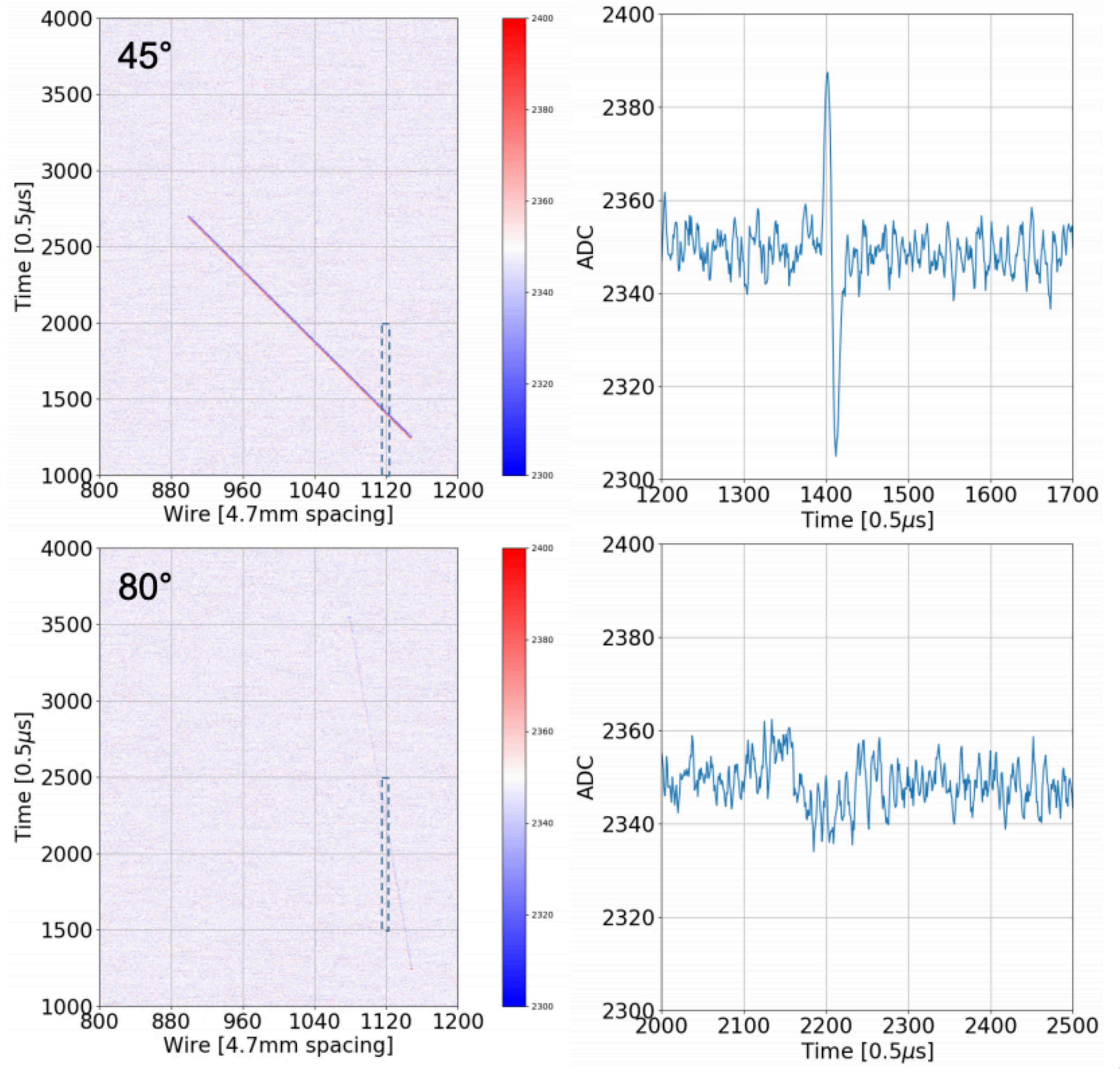
Hanyu Wei' talk, DUNE collab. meeting May 2019

Inherent electronic noise spectrum for different lengths of wires, *JINST 13 P07006 (2018)*

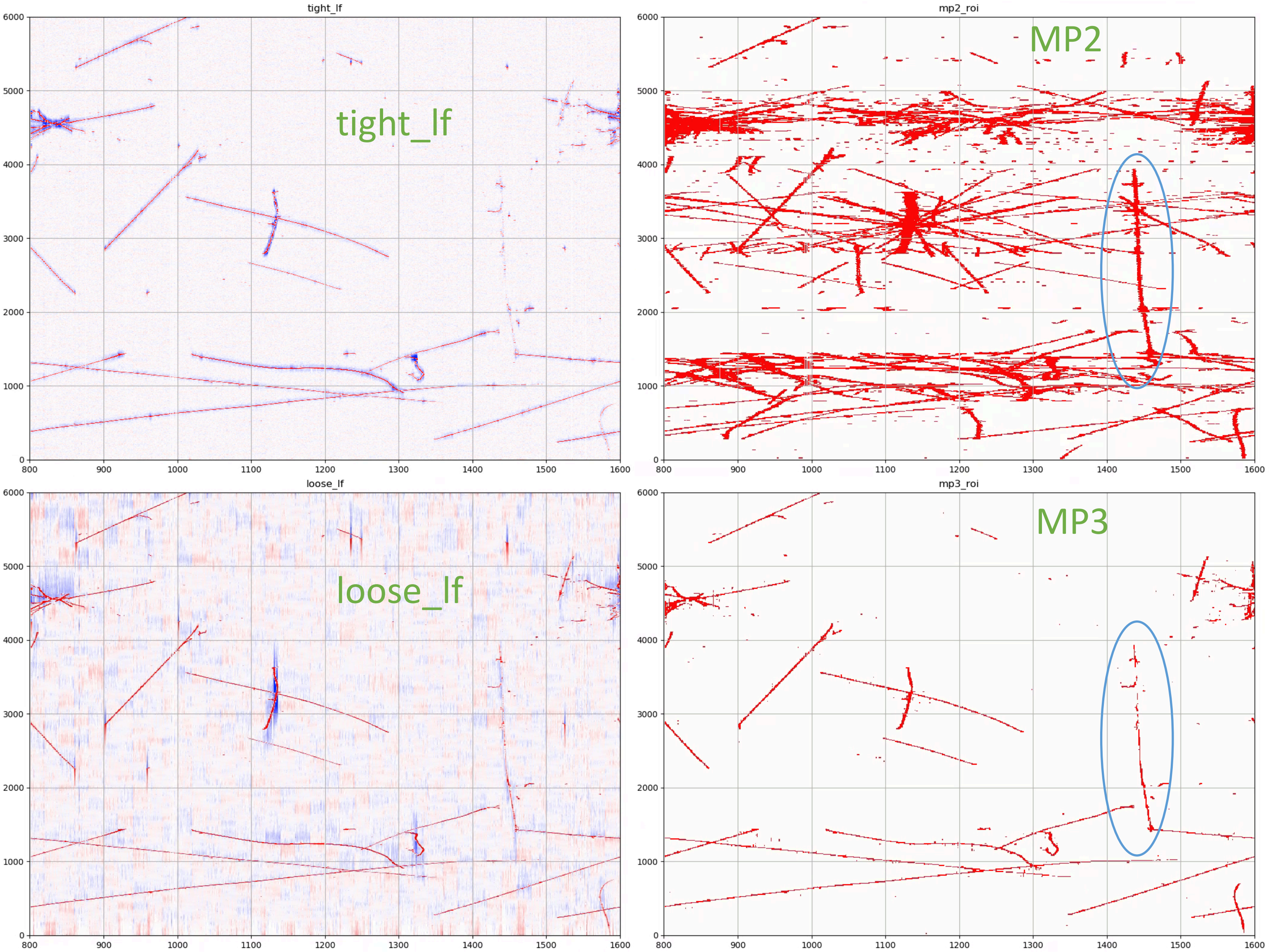


Signal vs. track angle

projection of a minimum ionizing particle (MIP) track on one induction plane



Input Candidates for DNN - ProtoDUNE data

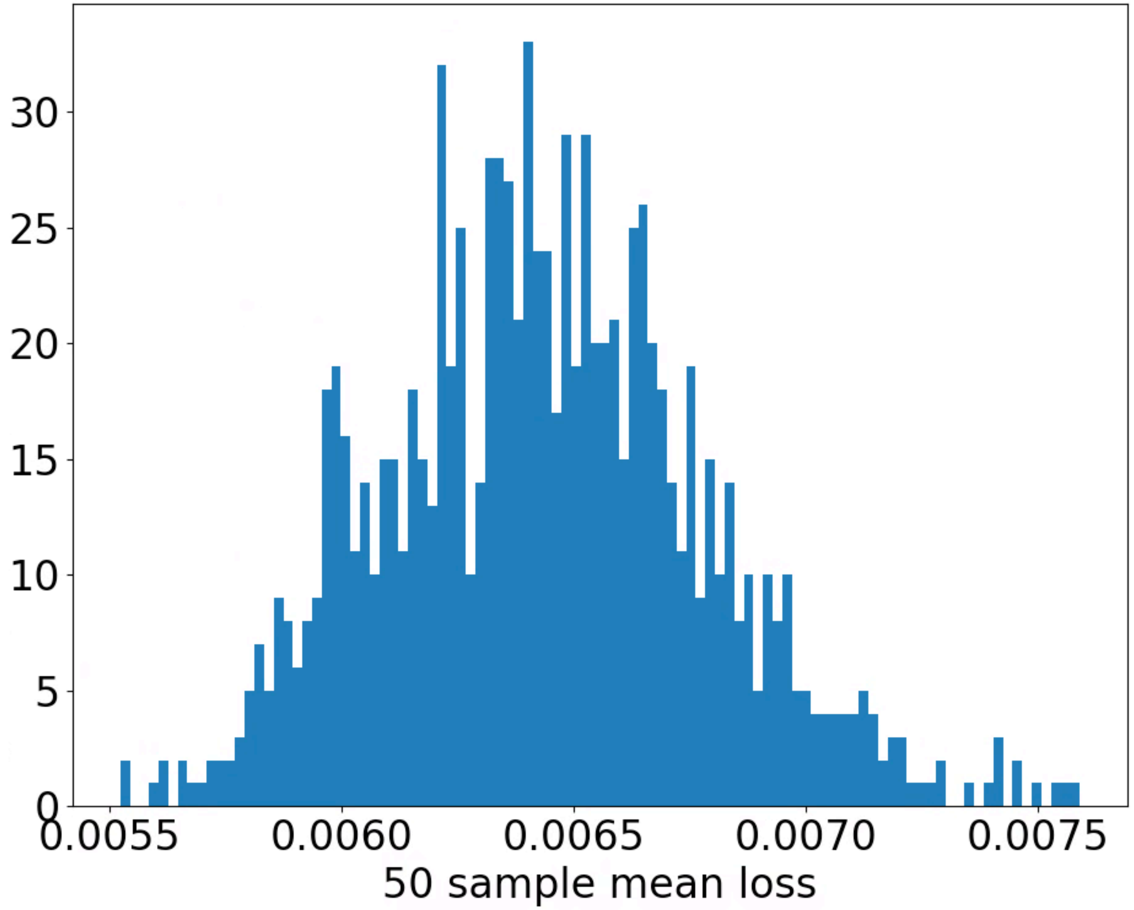


Training

	Unit Time	Total Time
Generator	2 sec/event	0.3 hour/500events
G4	23 sec/event	3.2 hour/500events
detector response, truth tagging and waveform preprocessing	68 sec/APA	9.4 hour/500 APA
Network training	6 min/epoch (1 epoch: 1 iteration of 500 APA sample)	5 hour/50epoch
Sum		17.9 hour

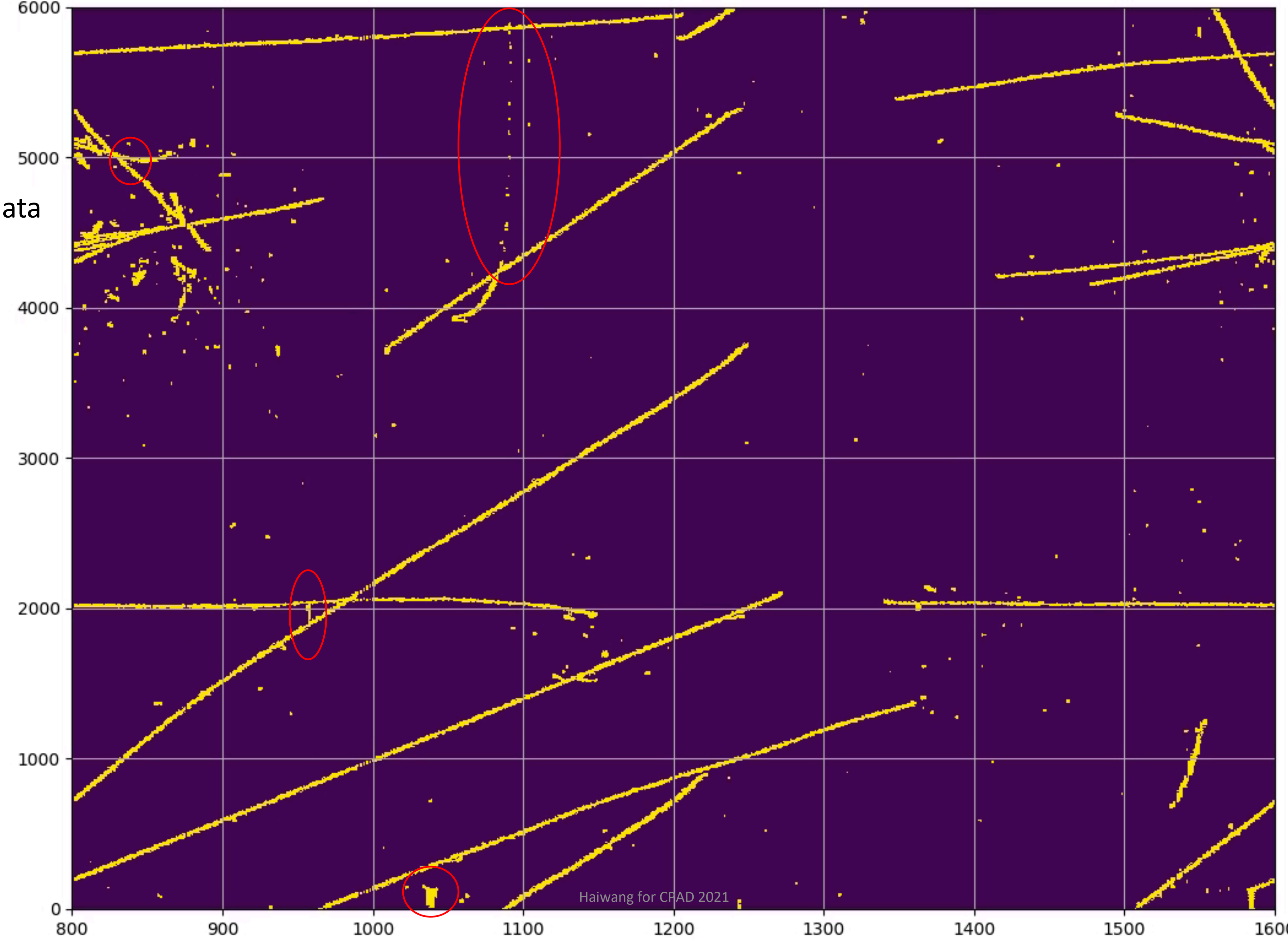
Mean loss of image samples are also expected to vary

We validated this hypothesis using a bootstrapping method. 1000 test samples (each sample contains 50 images) were random chosen from a test data set. Mean loss value was calculated for each sample and the distribution is shown below. We can see that the sample mean loss has a Gaussian like distribution that varies from 0.0055 to 0.0075.



Reference SP

ProtoDUNE Data
run: 5145
subRun: 1
event: 26925
V plane



DNN-SP

ProtoDUNE Data
run: 5145
subRun: 1
event: 26925
V plane