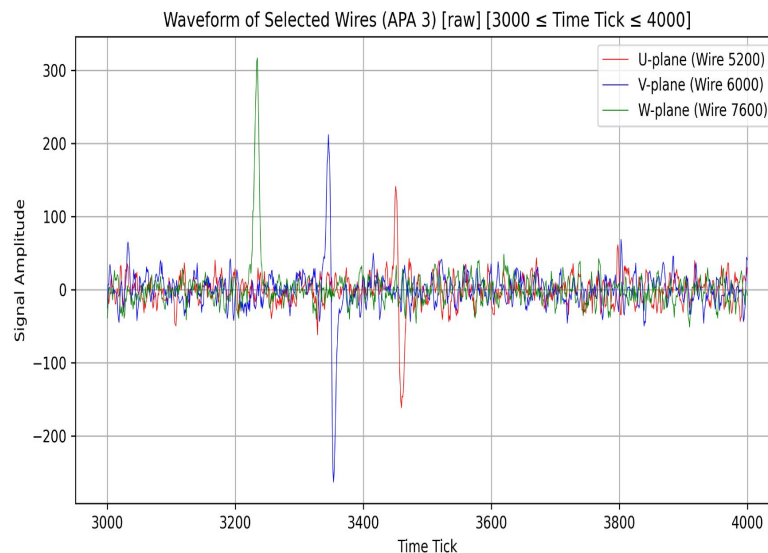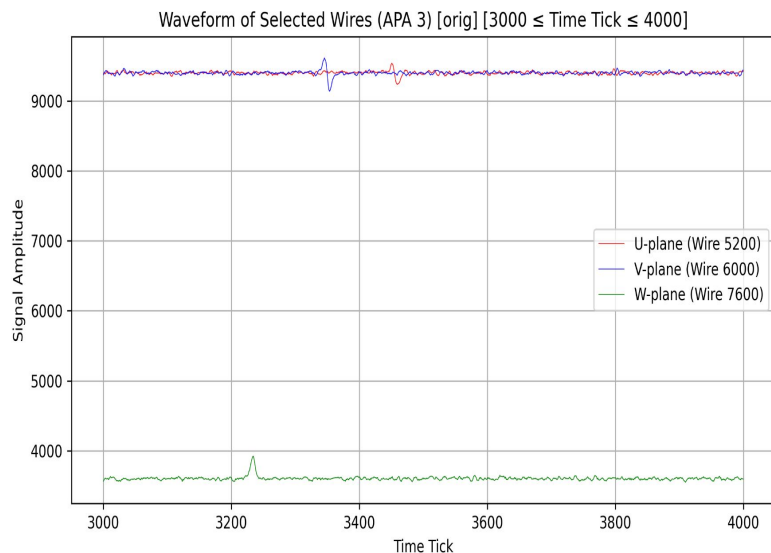# Status report on
WireCell Toolkit

Hokyeong Nam
Chung-Ang University

Outline

- PDHD Standalone Signal Processing & Data Processing

  - Waveform in each Sigproc. step

  - 2D Histogram - time tick vs wire channel

  - APA1 W plane event with Xuyang's field response

- DNN ROI training of the UNet on the WC cluster

  - Model training with different wire channel (U, V, U + V, overlap)

  - Loss trends

  - Implementation on Sigproc. chain

- Summary & Plan
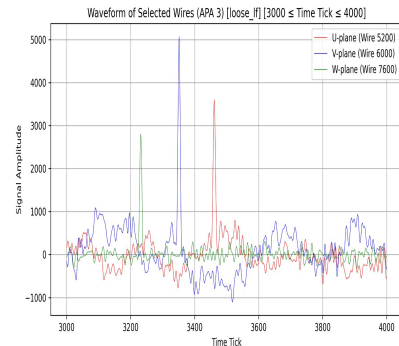
PDHD Signal processing (Sim.) - NF

Waveform of Selected Wires (APA 3) [orig] [3000 ≤ Time Tick ≤ 4000]

NF

Waveform of Selected Wires (APA 3) [raw] [3000 ≤ Time Tick ≤ 4000]

- PDHD simulation data
- comparison between original waveform and after noise-filtering

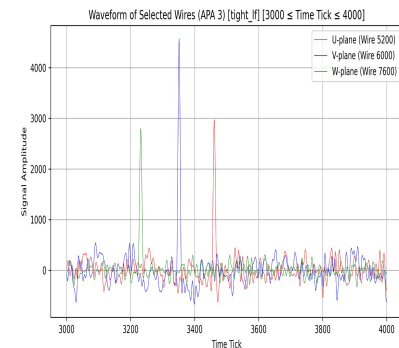# PDHD Signal processing (Sim.) – Wiener Filter & Low-Frequency Filter
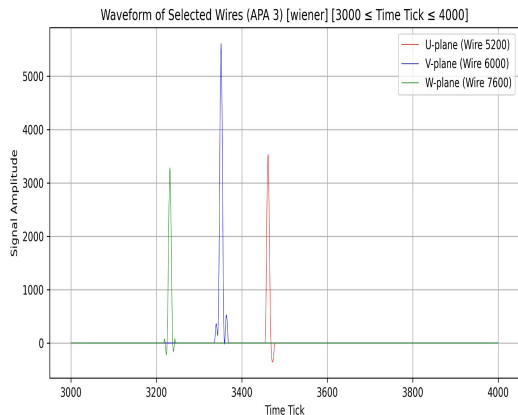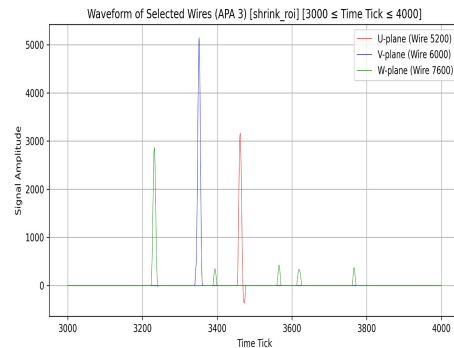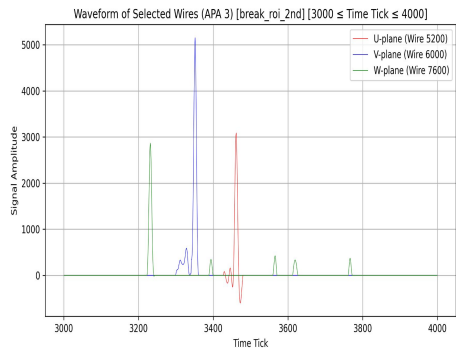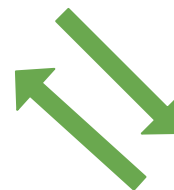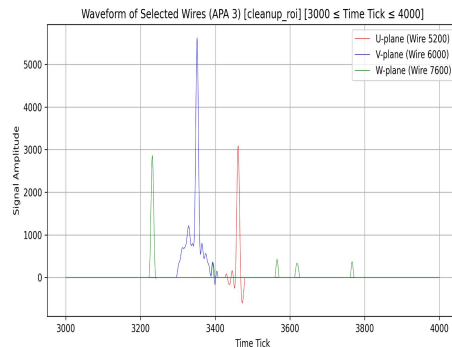

WF applied

loose

tight

- To find ROI,

 apply Wiener-Filter and define both loose and tight Low-Frequency filter

PDHD Signal processing (Sim.) – Finding ROI in classical way

- To find ROI, repeat
  - Break ROI
  - Shrink ROI
  - Cleanup ROI

Waveform of Selected Wires (APA 3) [cleanup_roi] [3000 ≤ Time Tick ≤ 4000]

Waveform of Selected Wires (APA 3) [break_roi_2nd] [3000 ≤ Time Tick ≤ 4000]

Waveform of Selected Wires (APA 3) [shrink_roi] [3000 ≤ Time Tick ≤ 4000]

PDHD Signal processing (Sim.) – Deconvolved waveform

- Gaussian filter is the result of 2D deconvolution
  - ROI is applied to the waveform after NF & before WF
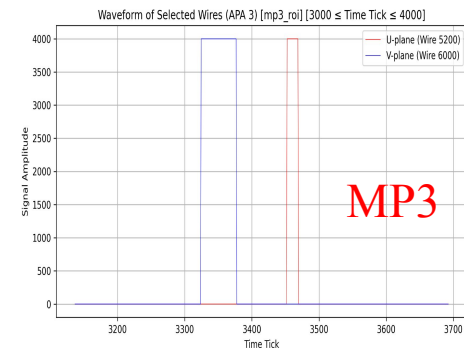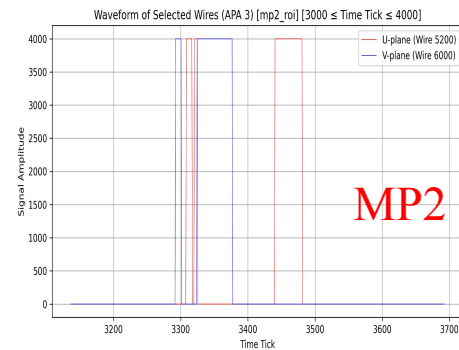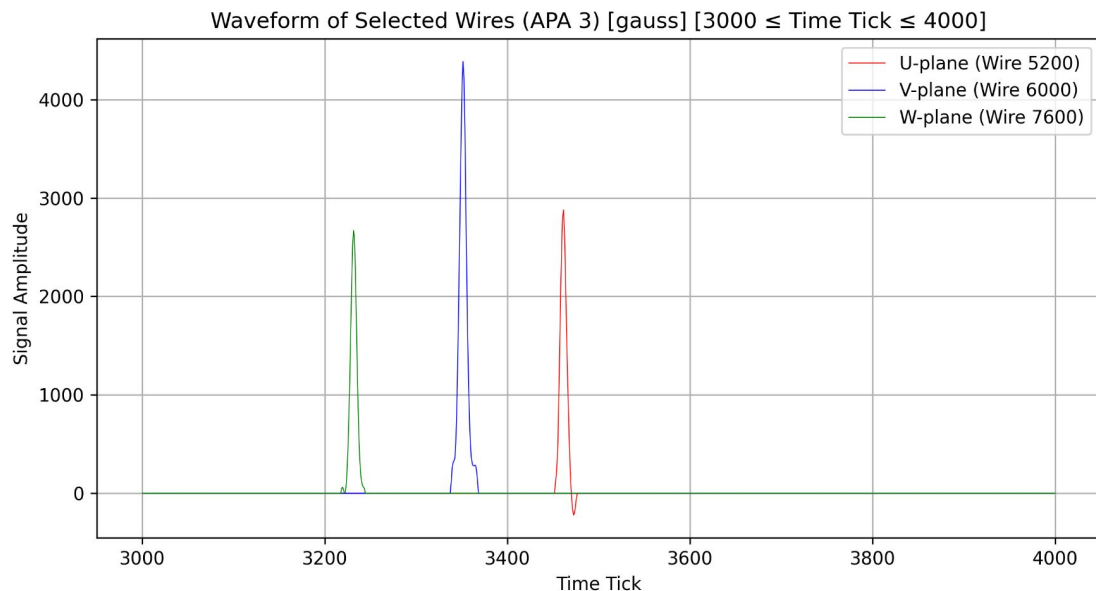
2D Histogram (Sim.) – time tick vs wire channel
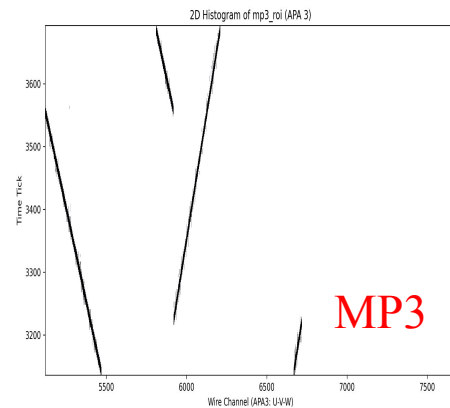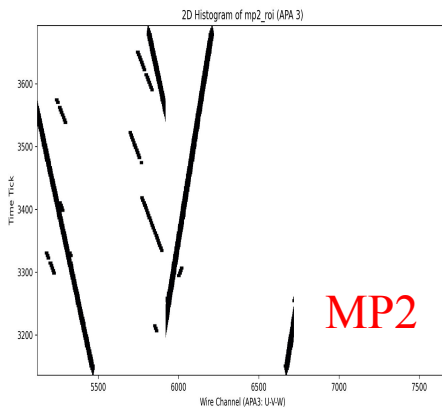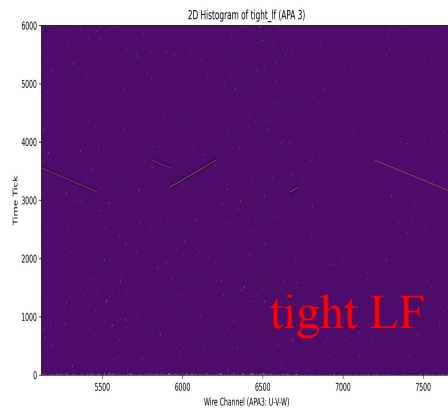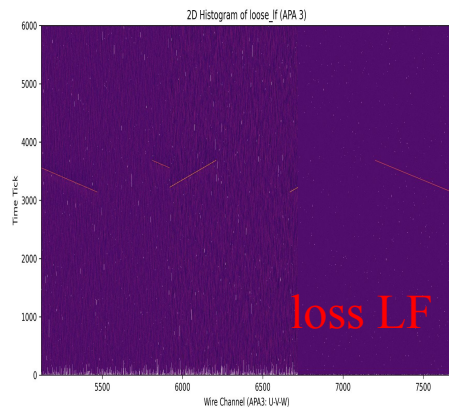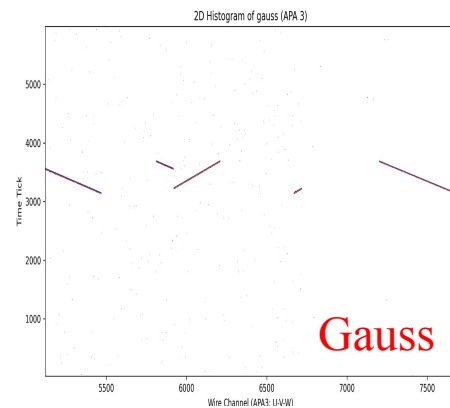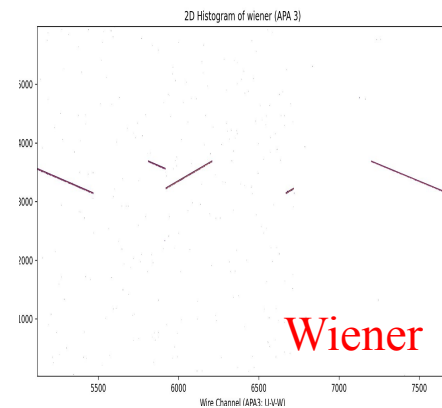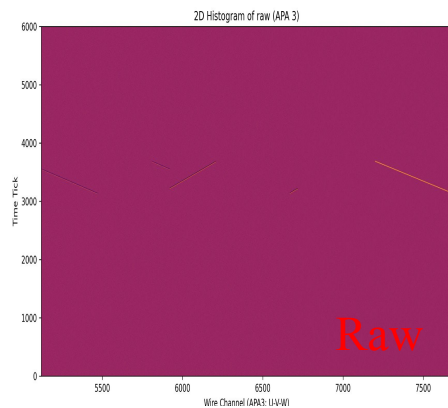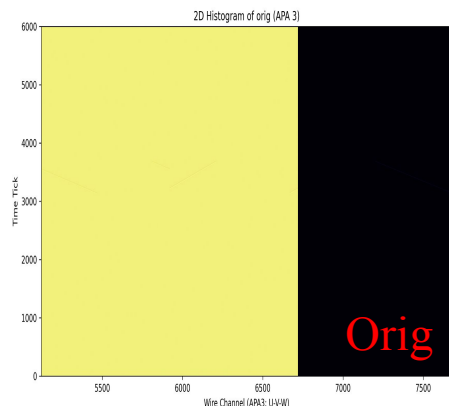
PDHD Signal processing (Data) - Overview
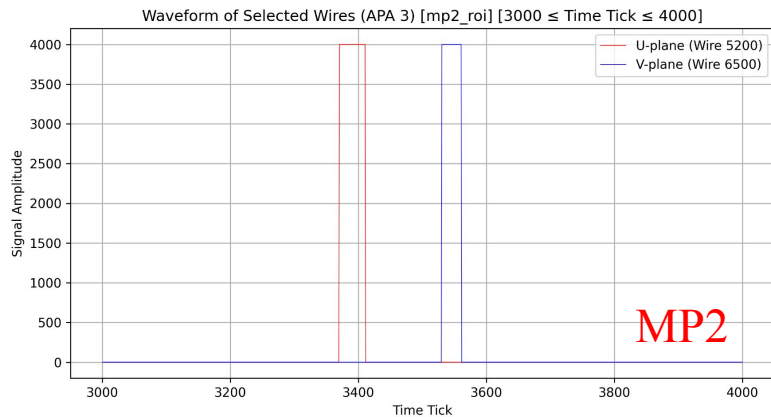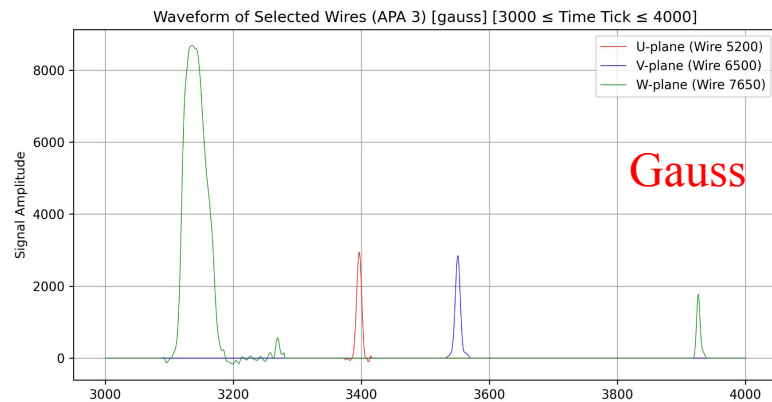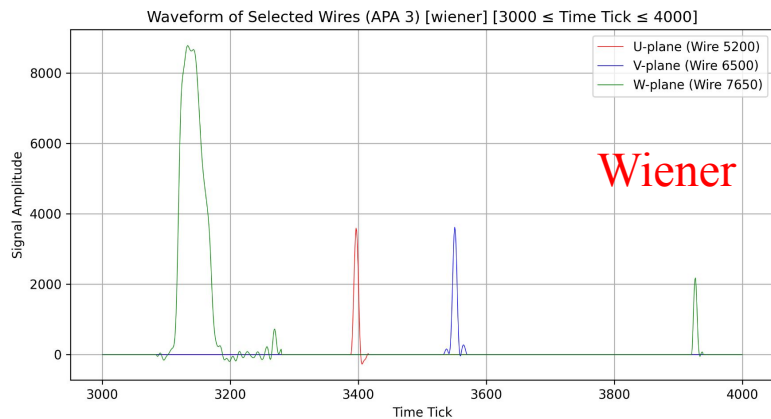
2D Histogram (Data) - time tick vs wire channel

Orig

Raw

Wiener

Gauss
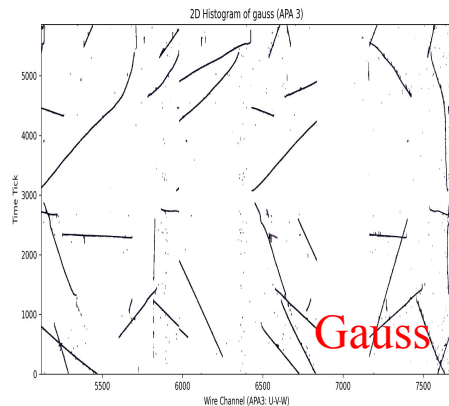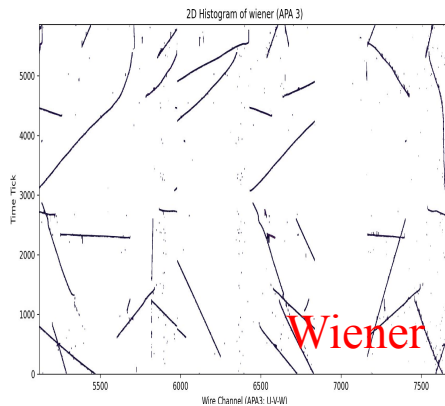
loss LF

tight LF

MP2

MP3
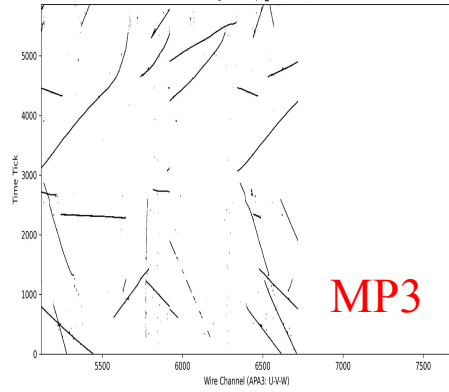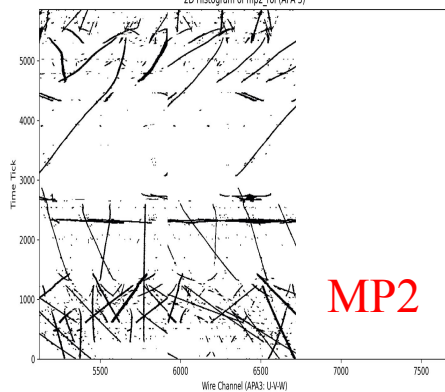
APA1 W plane event with Xuyang's field response
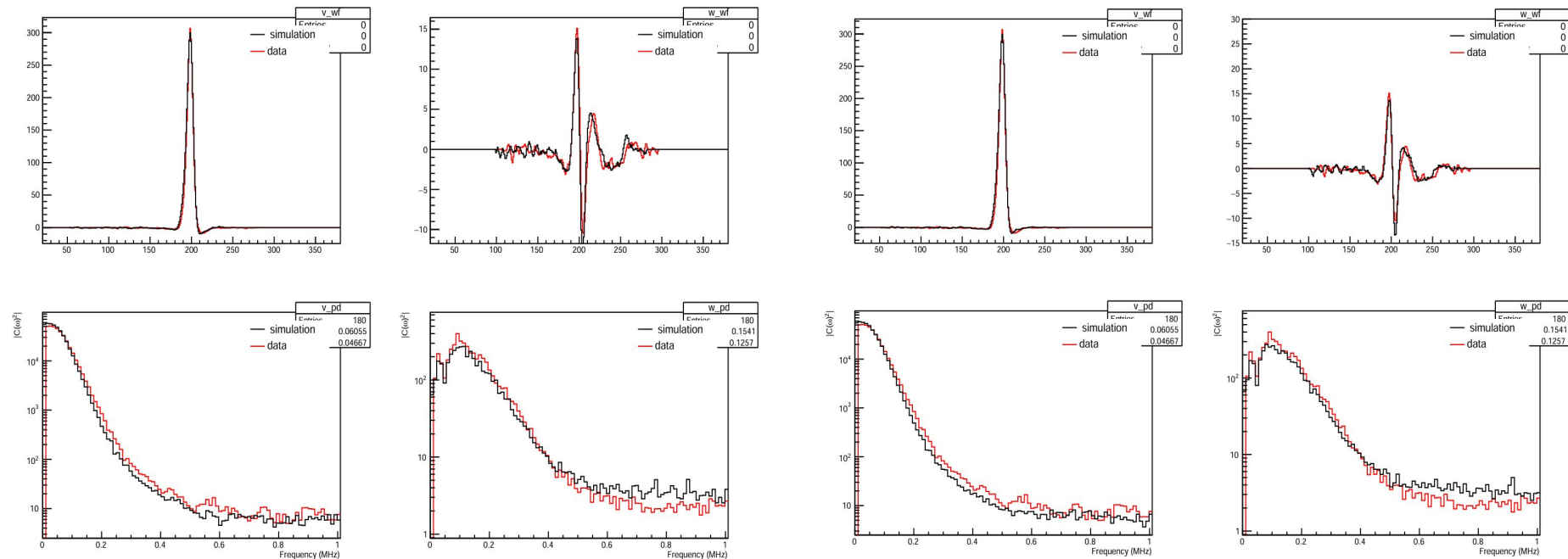
- My result

- Xuyang's result

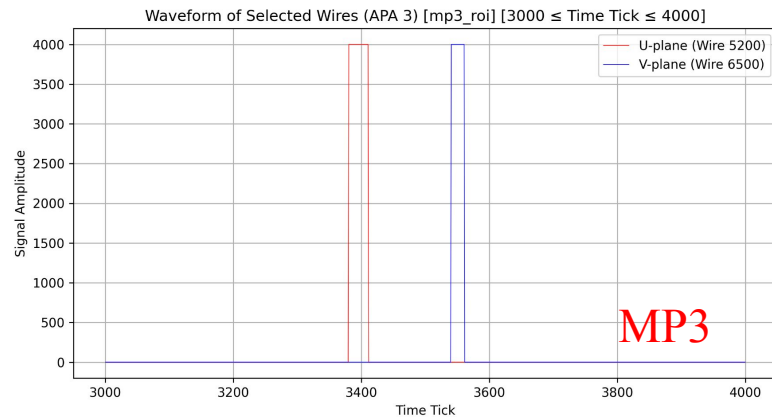# DNN training on the WC cluster - Choose channel

```
123        # x_range = [800, 1600] # PDSP, V, left-closed right-open interval
124        x_range = [476, 952] # PDVD, V
125        y_range = [0, 600]
126        z_scale = 4000
```

```
132        #x_range = [0, 800] # PDHD, U, left-closed right-open interval
133        #x_range = [800, 1600] # PDHD, V, left-closed right-open interval
134        x_range = [0, 1600] # PDHD, Induction, left-closed right-open interval
135        # x_range = [476, 952] # PDVD, V
```

- Trained with PDHD data
  - /nfs/data/1/renney/train_data_PDHD_fixedbug_separateWC

DNN training on the WC cluster - time consumption

```
Epoch finished ! Loss: 0.022248
Checkpoint e49 saved !
Validation Loss: 49.0000, 0.033984


real    25m3.591s
user    20m18.836s
sys     4m44.028s                U & V
```

```
49 : 0.9889 --- loss: 0.016800
Epoch finished ! Loss: 0.018984
Checkpoint e49 saved !
Validation Loss: 49.0000, 0.034819


real    23m13.408s
user    18m27.681s
sys     4m44.835s                U
```

```
Epoch finished ! Loss: 0.020041
Checkpoint e49 saved !
Validation Loss: 49.0000, 0.034918


real    23m21.587s
user    18m33.740s
sys     4m46.965s                V
```

- Average training time ~ 5 min.
  - /nfs/data/1/renney/train_data_PDHD_fixedbug_separateWC

DNN training on the WC cluster – Loss with wrong x_range
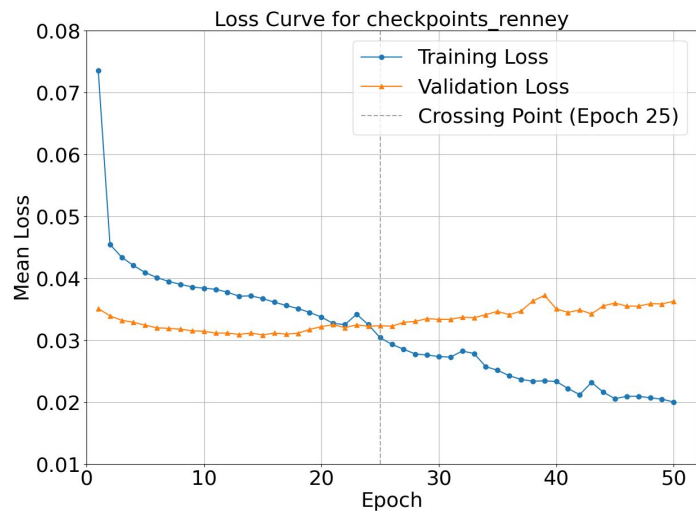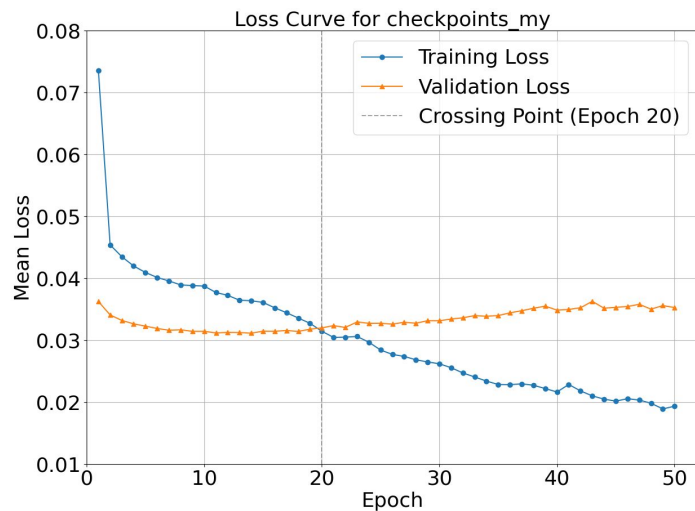
- Checkpoint: my - First crossing at Epoch 20.0, Training Loss = 0.031452, Validation Loss = 0.031995
- Checkpoint: my - Final Train Loss: 0.019340, Final Val Loss: 0.035267, Generalization Gap: 0.015927
- Checkpoint: renney - First crossing at Epoch 25.0, Training Loss = 0.030412, Validation Loss = 0.032341
- Checkpoint: renney - Final Train Loss: 0.020019, Final Val Loss: 0.036281, Generalization Gap: 0.016262

DNN training on the WC cluster – Loss with wrong U, V, and both of them
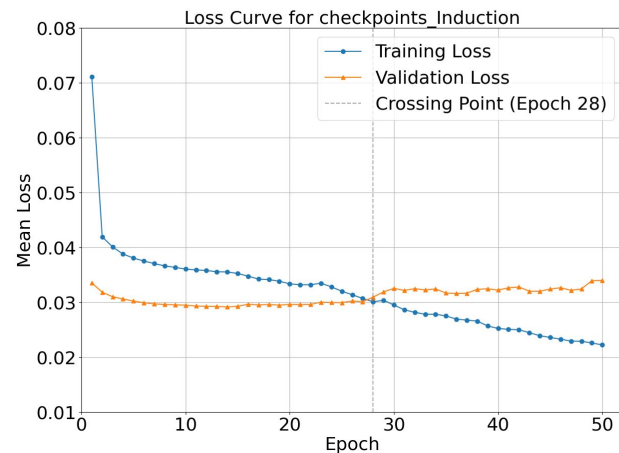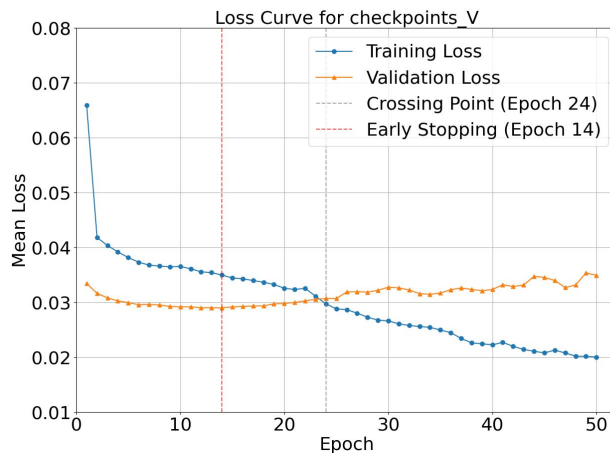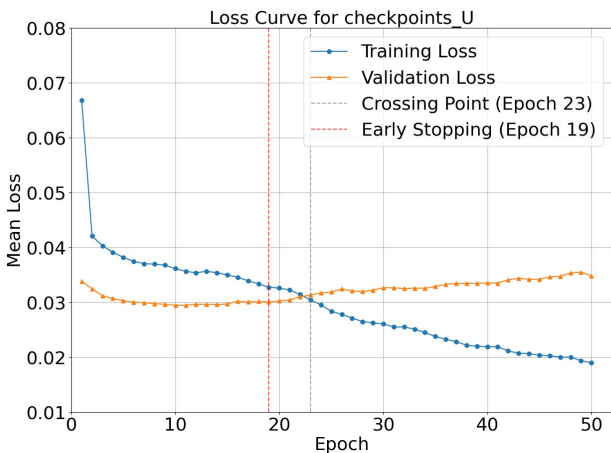
- Checkpoint: Induction - First crossing at Epoch 28.0, Training Loss = 0.030104, Validation Loss = 0.030948
- Checkpoint: Induction - Final Train Loss: 0.022248, Final Val Loss: 0.033984, Generalization Gap: 0.011736
- Checkpoint: U - First crossing at Epoch 23.0, Training Loss = 0.030450, Validation Loss = 0.031340
- Checkpoint: U - Early Stopping Recommended at Epoch 19.0, Training Loss = 0.032752, Validation Loss = 0.030063
- Checkpoint: U - Final Train Loss: 0.018984, Final Val Loss: 0.034819, Generalization Gap: 0.015835
- Checkpoint: V - First crossing at Epoch 24.0, Training Loss = 0.029711, Validation Loss = 0.030694
- Checkpoint: V - Early Stopping Recommended at Epoch 14.0, Training Loss = 0.034972, Validation Loss = 0.028990
- Checkpoint: V - Final Train Loss: 0.020041, Final Val Loss: 0.034918, Generalization Gap: 0.014877

Summary & Plan

- ➤ Followed up Xuyang's instruction on Sigproc. & Data processing
  - ○ Plotted 1D waveform and 2D histogram with PDHD simulation data
  - ○ Reproduced the APA 1 W plane event with the updated field response

- ➤ ML model training for DNN Sigproc.
  - ○ Trained the UNet with different wire channels
  - ○ Checked train vs val loss to find the file has reasonable weight
  - ○ Revised cfg. file to implement the trained model

- ➤ Next steps
  - ○ Run the full Sigproc. chain with DNN ROI
  - ○ Train the model with wirecell-dnn → ~1 min. is expected for training.
  - ○ Check the time consumption for the ML inference in the Sigproc. chain
  - ○ Try other ML techniques (e.g. transformer) to explore room for improvement
  - ○ Code review: cfg file, shell scripts, source file, etc.

Back Up