Basic Wire Readout LArTPC
Signal Processing

Overview

- Generally, signal processing is the procedure for returning the basic input to event reconstruction from the raw data provided by the detector
  - For LArTPCs the TPC data are typically waveforms from each wire which can be used to determine the drift time for track ionization in the argon and to provide a measure of the charge deposited
  - The basic building blocks for reconstruction are “Hits” - objects describing the measured drift time and charge for individual charge deposits seen by each wire
- Many challenges to Signal Processing
  - Handling detector noise
  - Signal-to-Noise can make signal finding difficult
  - Signal decimation at steeper track angles…
- ICARUS, in particular, is very challenged
  - Significant coherent noise and very low Signal-to-Noise are particularly troublesome
- Aim here to convince that much work needed to help ICARUS, hence SBN program, achieve ultimate science goals

Goals of Signal Processing

- Simulation
  - Provide an accurate representation of the signal response of the ICARUS detector to charge depositions in the liquid argon
  - Output waveforms per wire channel in the same format as presented in the data

- Reconstruction
  - LArTPC Reconstruction Dual Mandate:
    - Given waveform readout of sense wires in ICARUS provide a robust estimate of the charge deposited by the ionizing track/shower in the liquid argon
    - As well, estimate the time of the charge deposit = the distance of the charge deposit from the sense wire
  - The ultimate goal of the signal processing chain is to return 3D positions of charge deposits as well an estimate of the deposited charge

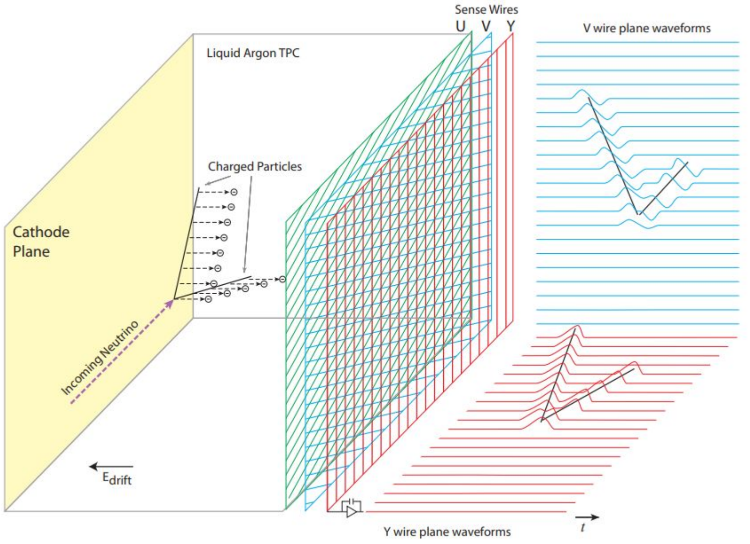
- To date the two goals have always shared a common methodology

Outline

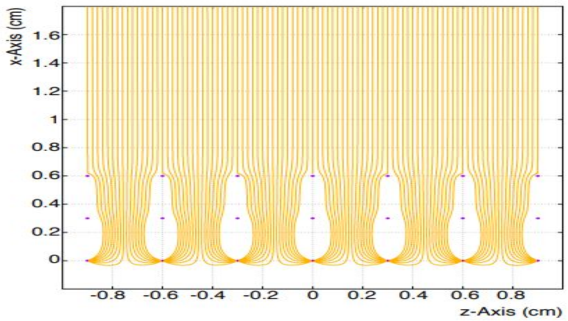
- Signal Simulation and signal reconstruction overview
- Outline of steps in Signal Processing Chain
- Noise handling for ICARUS
- Mention ROI/Hit Finding
- What is the current situation? Can we do better?
- Turn over to Francois for some ideas on improving

Basic Overview of Signal Simulation

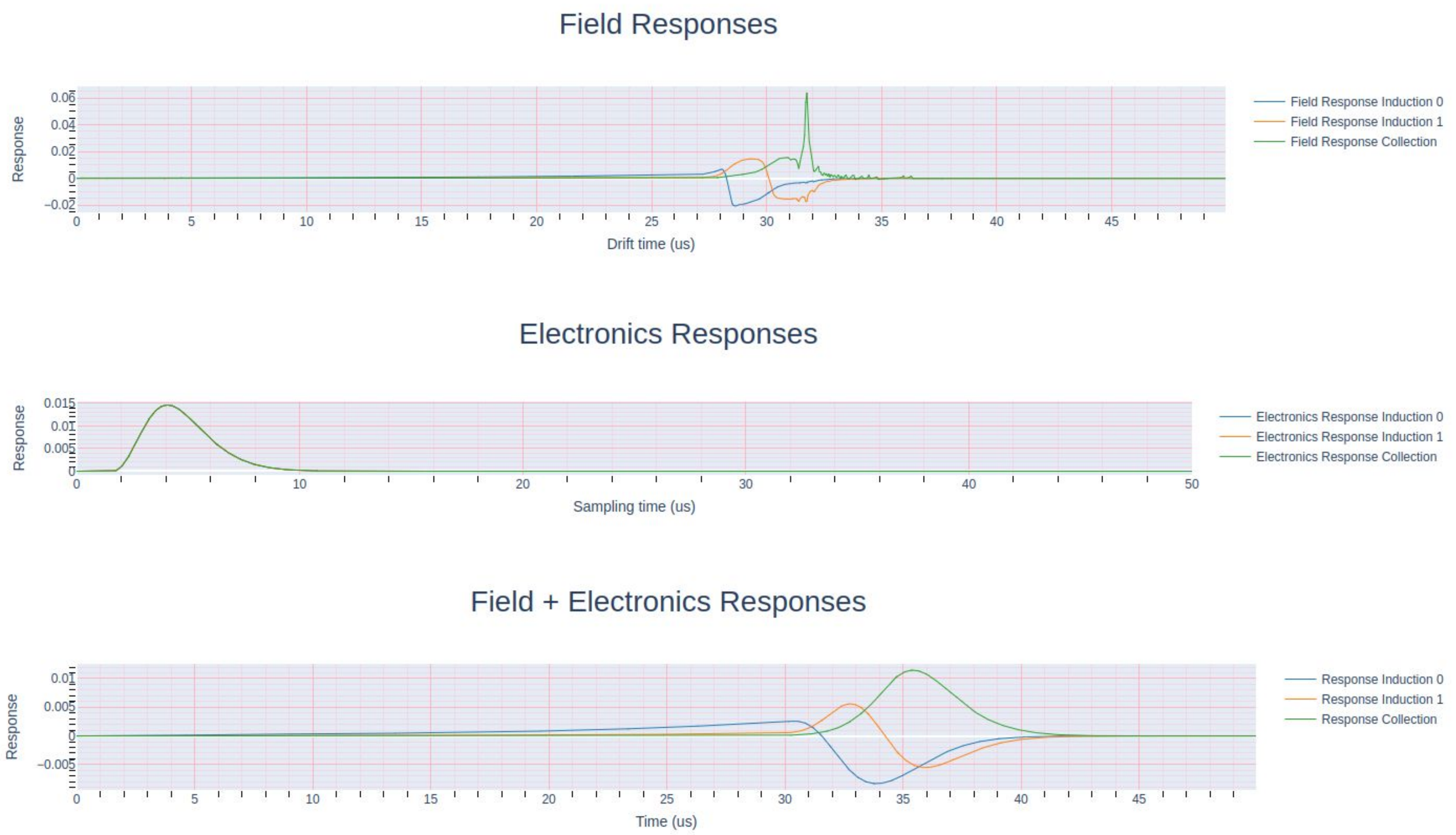
- Within the bulk region of the TPC drift transport electrons in “bunches”
  - Assume uniform electric field
    - Drift velocity constant, determined by liquid argon parameters
  - Group electrons within track spacing of ~1/10 sense wire spacing to form “bunches”
  - Provide corrections for recombination, lifetime, diffusion, field distortions, space charge, etc.
  - Transport to a plane some distance above the wire planes



- Granularity of wires becomes apparent as first wire plane is approached
  - Switch from bulk region drift to using field responses to determine charge induced or deposited on wires
    - Field responses are determined using a 1D model from the Garfield electrostatics program

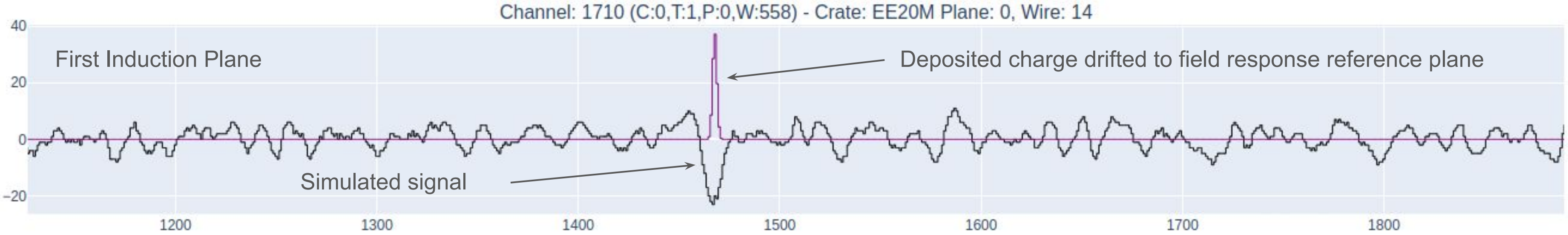


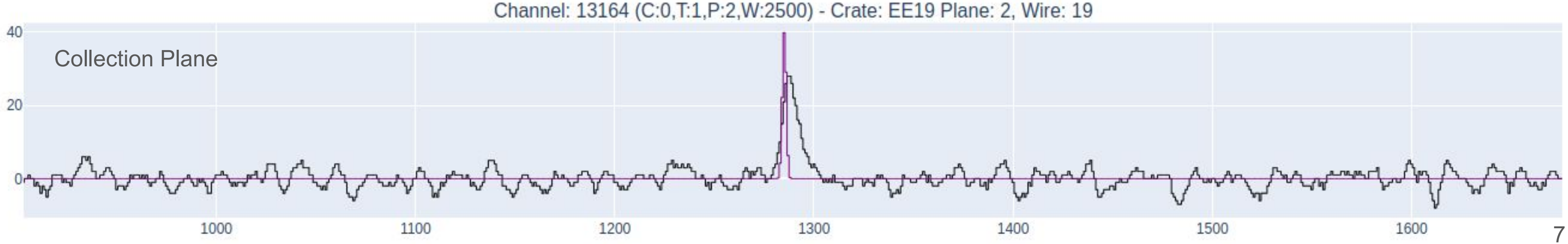
Basic Overview of Signal Simulation - Time Domain



Basic Overview of Signal Simulation - Final Result

- Simulation of the waveforms now straightforward
  - Simulate signal by convolving simulated charge at field response reference plane with the response functions from previous page
  - Add simulated noise from a data driven model





Basic Overview of Signal Reconstruction

- Our goal is to estimate the charge deposited from the measured waveforms
  - Signal Simulation gives us the measured charge:

$$M(t') = \int_{-\infty}^{\infty} R(t, t') \cdot S(t) dt.$$

Measured signal

Response Function

Original Signal

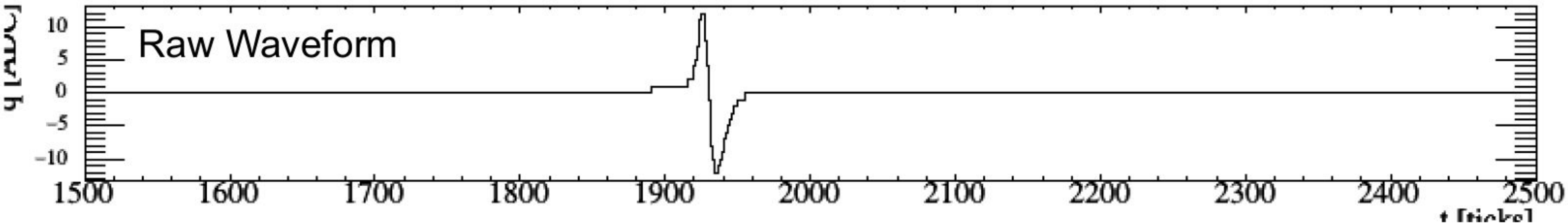
  - Convert to the frequency domain and run the above in reverse:

$$S(\omega) = \frac{M(\omega)}{R(\omega)}.$$

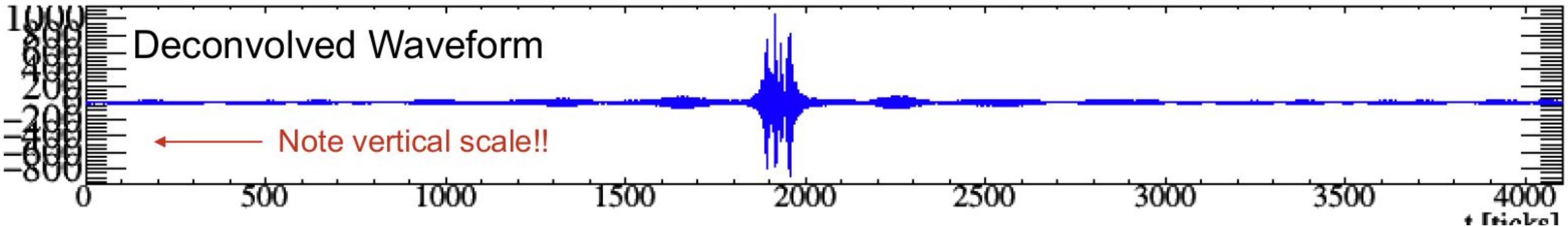
- Issues:
  - Note that our real measured signal will contain noise
  - We are doing numerical computations with discrete values
  - Induction plane responses tend to zero approaching zero frequency

Effect of Discrete Numbers

Test case of signal on waveform with no noise:



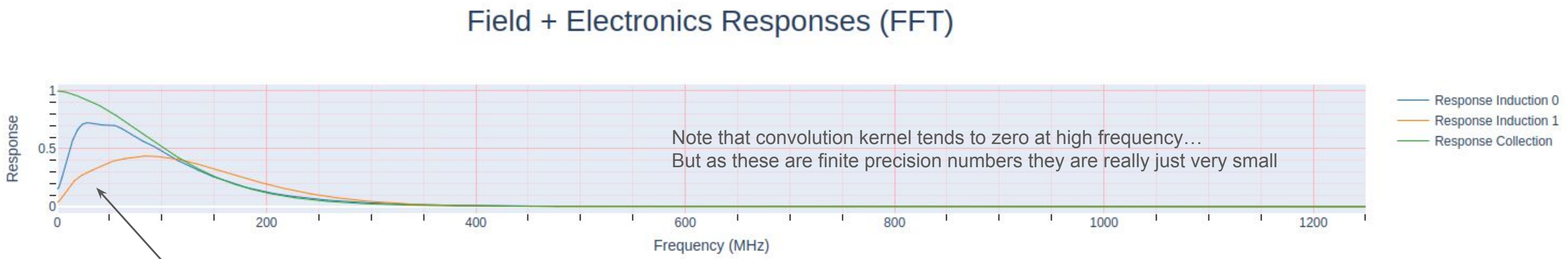
Result of using equation on previous page to deconvolve:



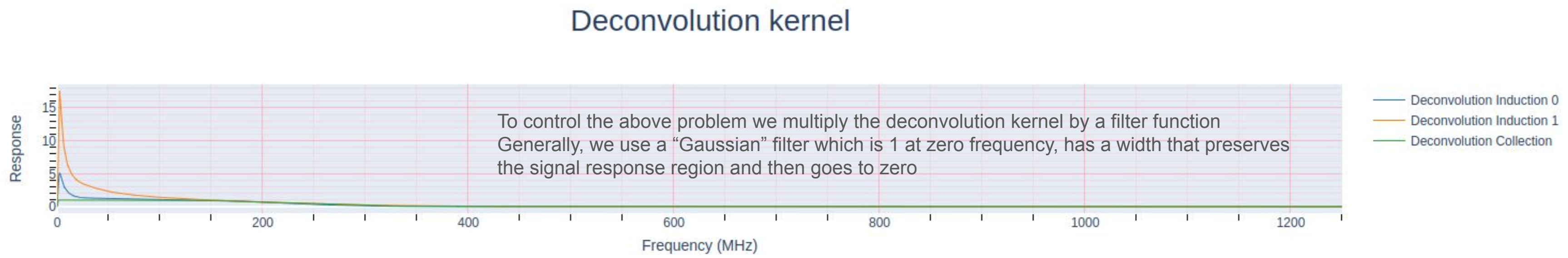
This madness results from dividing signal waveform by what can be very tiny numbers in the convolution kernel, primarily at high frequency

Convolution and Deconvolution Kernels

The convolution and deconvolution kernels for the three sense planes:



Note induction planes have responses that go to zero at zero frequency!



Problem: Induction plane responses go to zero at zero frequency… but applying a high pass filter component for this can impact signal reconstruction… Current filter compromises but not the enhancement at low frequency!

Basic Overview of Signal Reconstruction

- To estimate the original signal we use:

$$S(\omega) = \frac{M(\omega)}{R(\omega)} \cdot F(\omega).$$

Estimated Signal — $S(\omega)$; Measured signal — $M(\omega)$; Response Function — $R(\omega)$; Filter Function — $F(\omega)$

- Choice of Filter Function is key:
  - Best filter for charge reconstruction is “Gaussian” where width needs to be carefully selected
  - It is also important to include a high pass component on the induction plane to prevent amplification of low frequency noise… but must be careful to minimize impact on signal!

- Not discussing today:
  - 2D Deconvolution - all of above still holds, just more computationally complicated
  - Does the above model really work correctly for long pulse trains?
    - When does this break down?

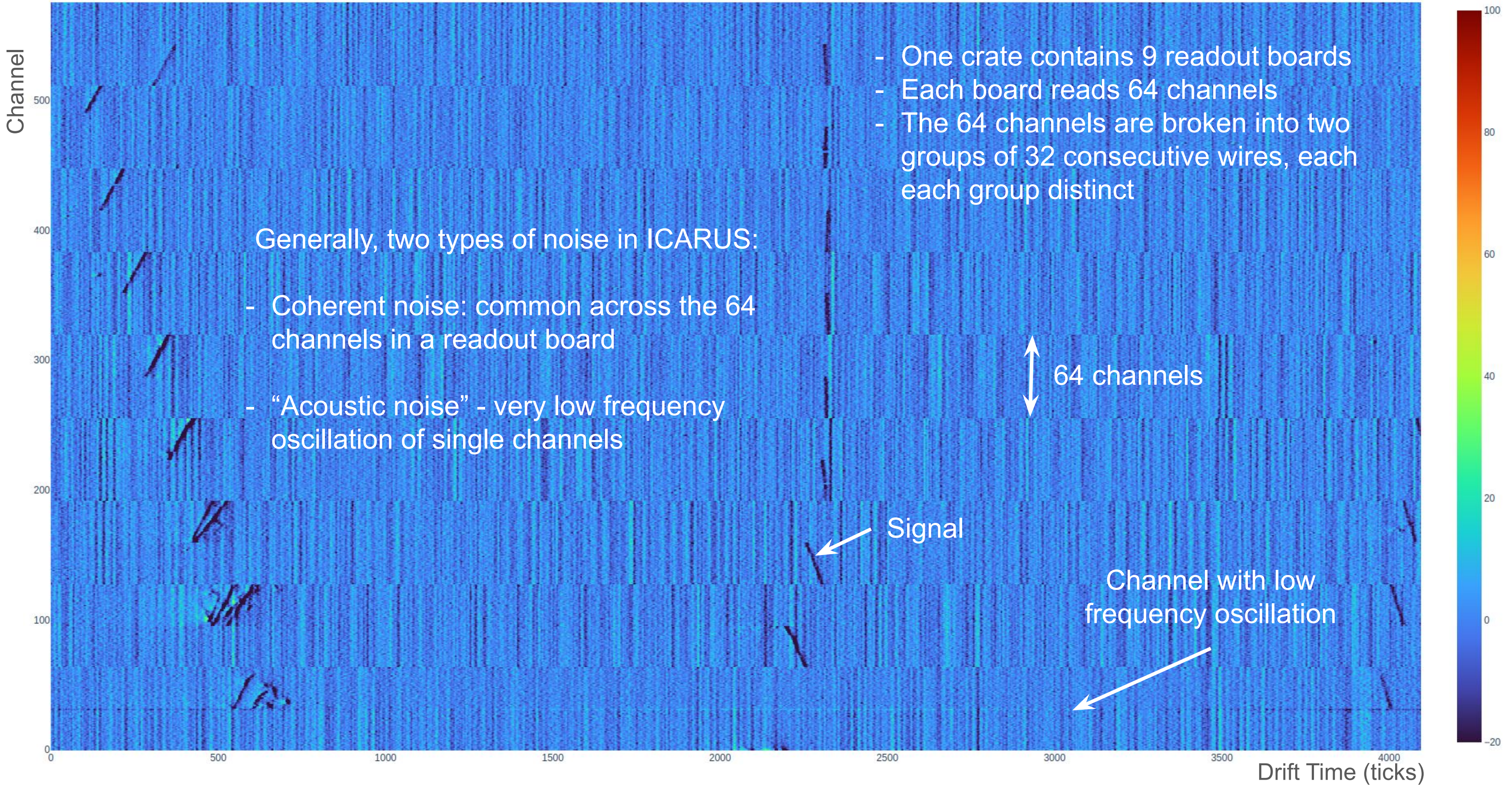
- Can we do better? Do we actually need deconvolution to find signals?

Outline of the Signal Processing Chain

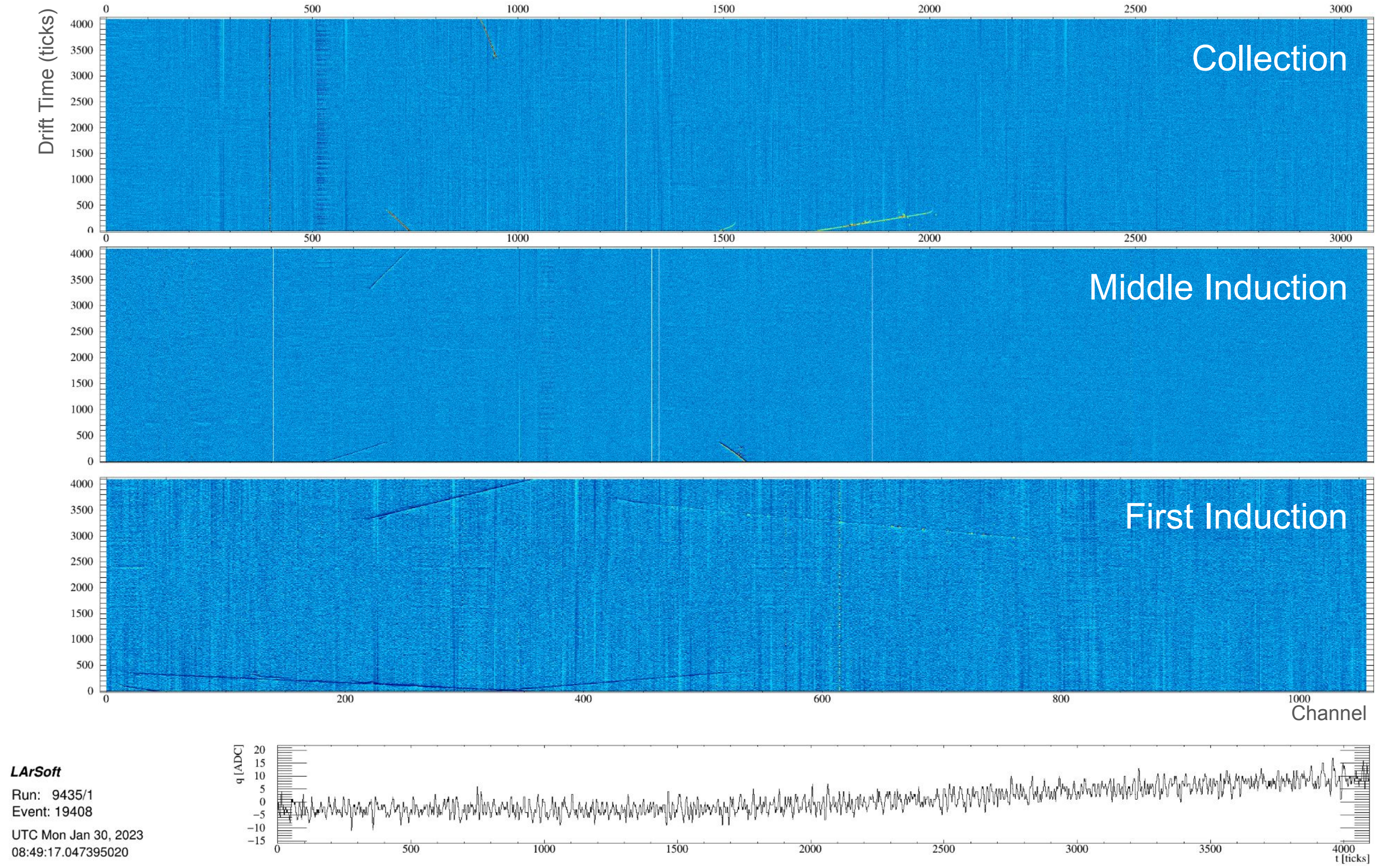
- Generally, there are four steps to signal processing
    - Data unpacking and noise handling
        - Convert from compressed daq format to something usable
        - Take advantage of this stage to mitigate noise issues
    - Deconvolution
        - Convert induction bipolar signals to unipolar
    - Region of Interest (ROI) finding
        - Regions in waveforms which are likely to have signal
    - Hit Finding
        - Search the ROIs for individual “hits” representing distinct charge deposits

Noise Handling for ICARUS

Image for a single crate
576 channels

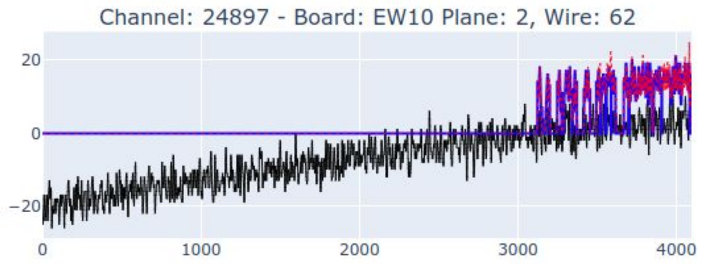


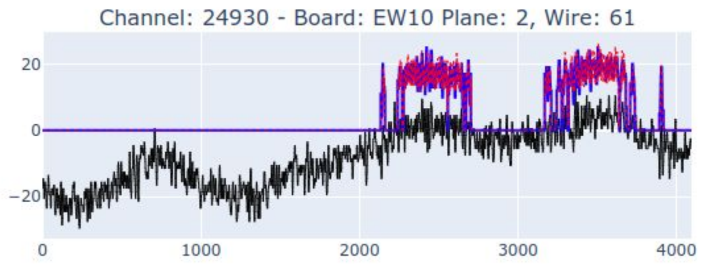
Low Frequency Oscillation



Low Frequency Oscillation - Issues

- Generally oscillation wavelength is large - longer than readout time of waveforms
    - Though can have shorter wavelength also
- This can adversely impact the 1D deconvolution on the collection plane
    - Collection plane, unlike induction planes, does not filter low frequency components
    - This appear to distort the baseline and/or amplify the low frequency oscillation
    - This can result in a large number of ROIs and then reconstructed hits
- Obviously should try to remove this
    - Be nice if it was not there to start with!



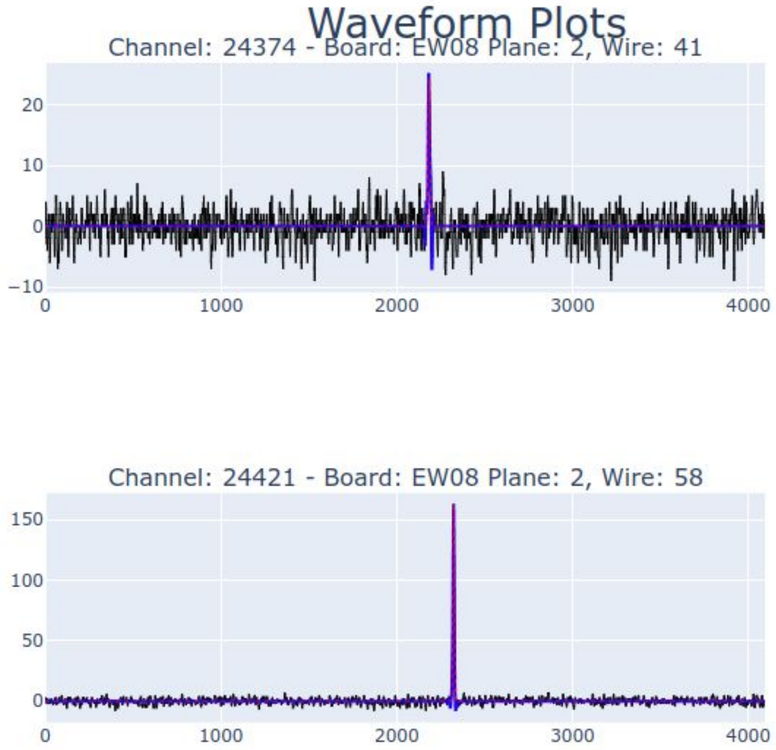


Low Frequency Oscillation - Mitigation

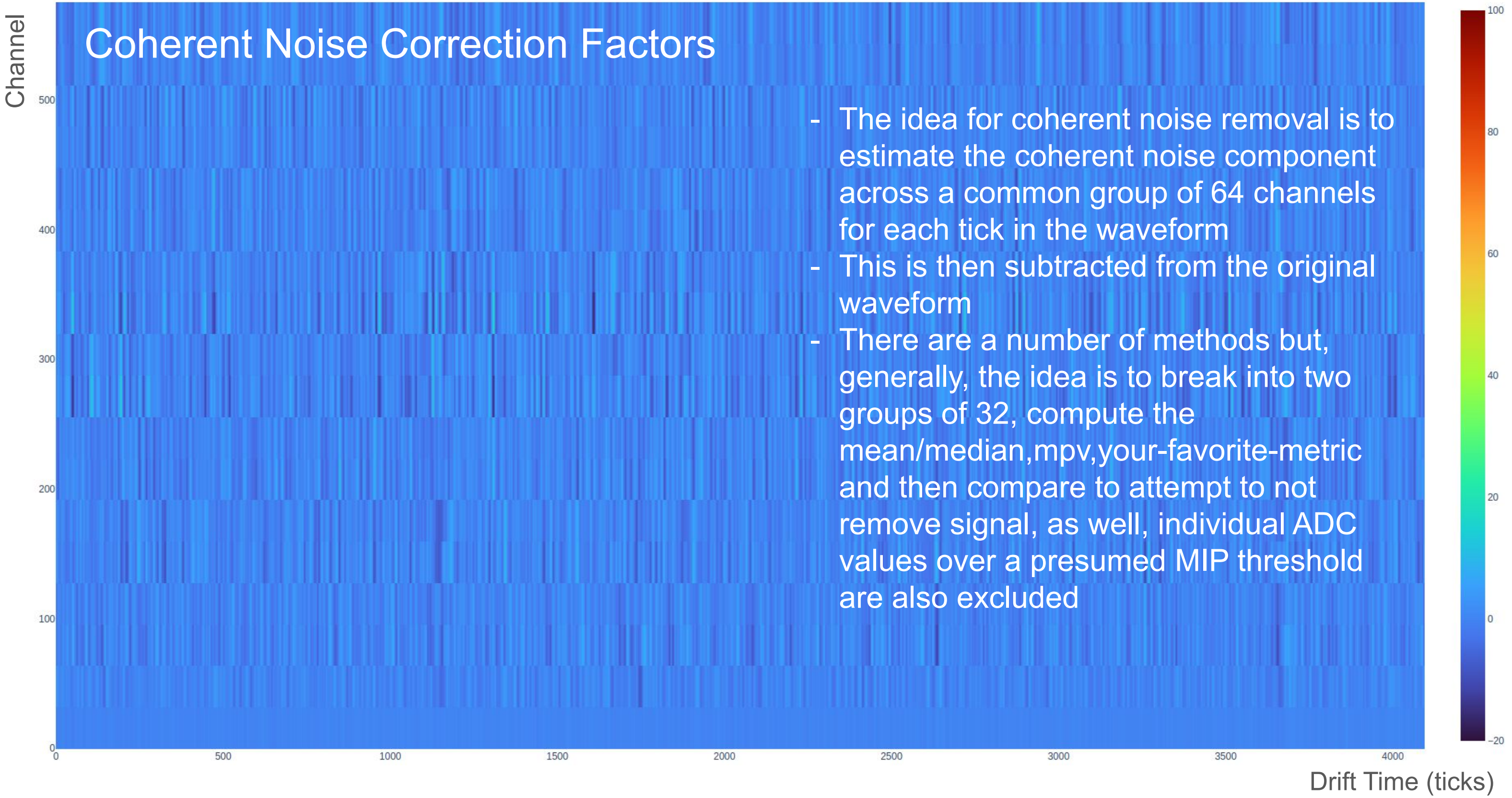
- Problem: “obvious” methods for removal (high pass filters, running averages, etc.) can impact the frequency response of signal, particularly for longer charge deposits.

- Fortunately, oscillation wavelength generally longer than readout time, most “oscillations” essentially linear

- Can use PCA to “fit” to slow oscillation and subtract from waveform
    - Working to extend this for remaining cases were there is a “roll”



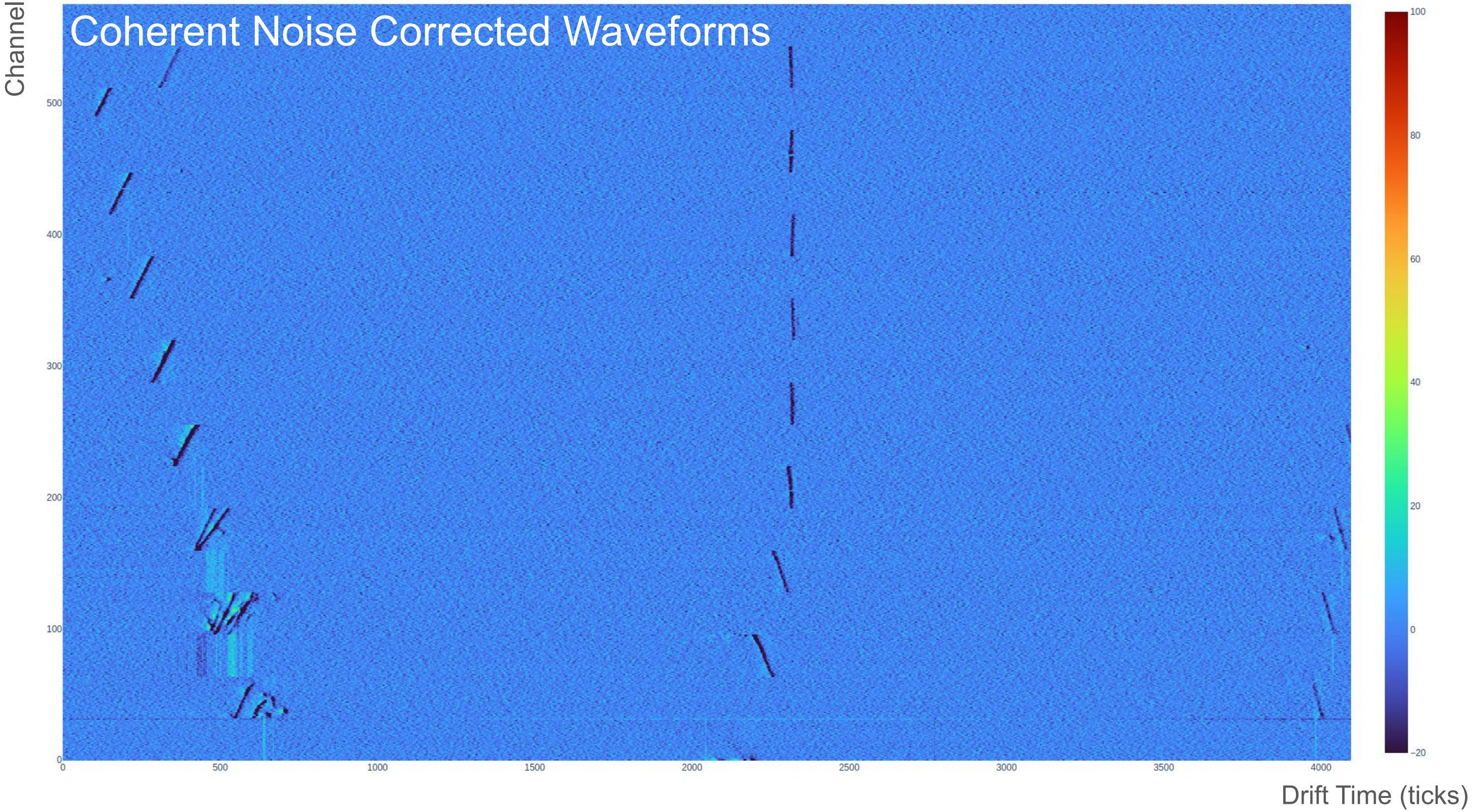
Coherent Noise Removal



Coherent Noise Correction Factors

- The idea for coherent noise removal is to estimate the coherent noise component across a common group of 64 channels for each tick in the waveform
- This is then subtracted from the original waveform
- There are a number of methods but, generally, the idea is to break into two groups of 32, compute the mean/median,mpv,your-favorite-metric and then compare to attempt to not remove signal, as well, individual ADC values over a presumed MIP threshold are also excluded

Coherent Noise Removal



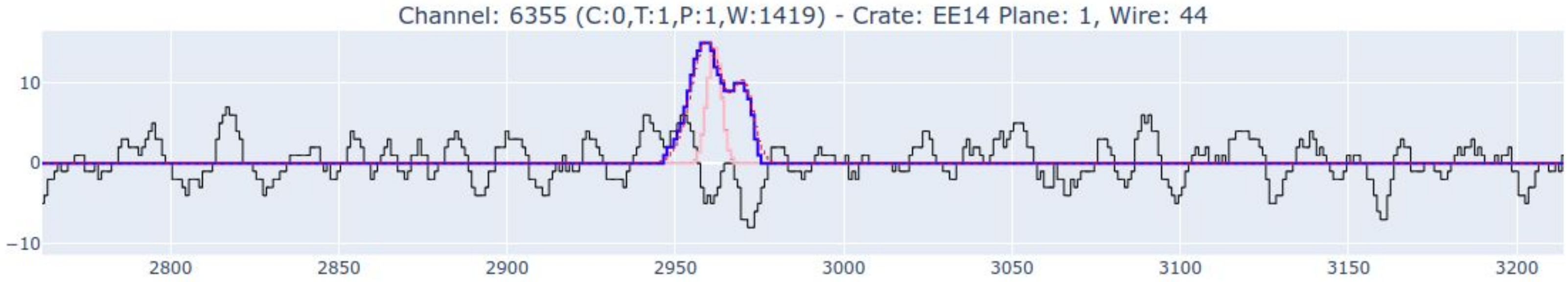
Regions-of-Interest and Hit Finding

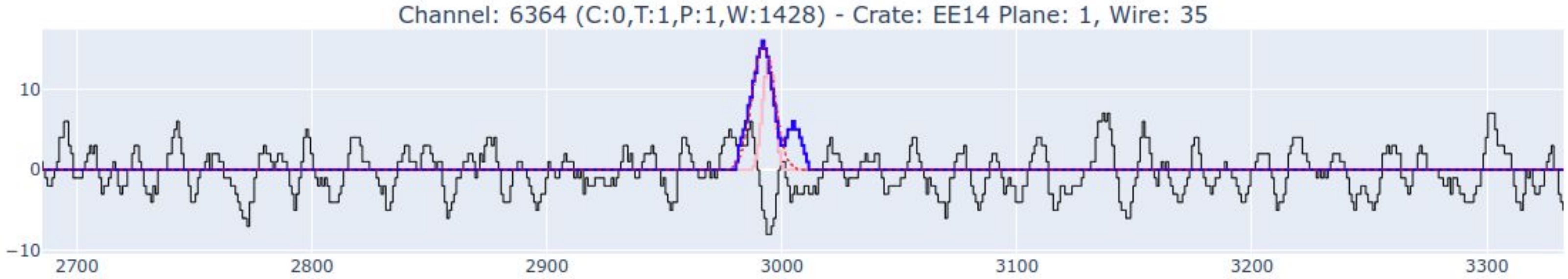
- Regions-of-Interest (ROIs)
  - Idea is to identify candidate regions on waveforms which are likely to contain signal
    - Should be fast and high efficiency at expense of purity
  - Originally developed to reduce overall data volume and speed up hit finding
  - To increase efficiency ICARUS uses a 2D Morphological Filter approach (Dae-Heun Koh)

- Hits
  - Hits are meant to be individual pulses in the waveform and return both the drift time for the hit and an estimate of the charge deposited (in ADC counts)
  - In ICARUS candidate pulses are found using a simple threshold over baseline method
    - Assumes baseline is at zero!
  - Fit a gaussian, or multiple gaussians for connected pulse trains, to return time/charge
    - If pulse train length too long then don’t fit, convert to “long hits”
  - Use the LArSoft supplied “GausHit Finder” module

Example Waveforms

Black: RawDigit waveform
Blue: ROI from Deconvolved waveform
Red Dash: Hit from hit finding
Pink: SimChannel deposit (ADC counts/3)





The Current Situation

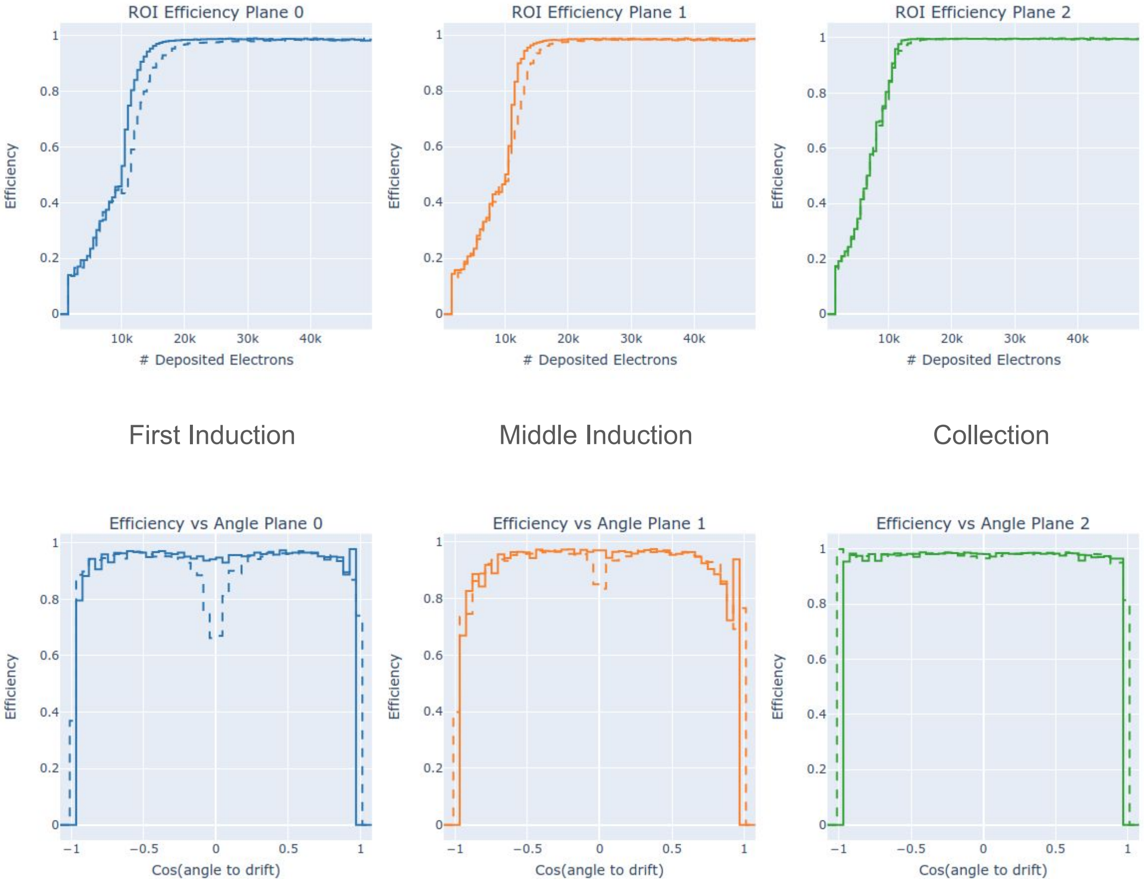
ROI Finding Efficiency

- As a function of number electrons deposited
- As a function of cosine of track ange to wire planes

Dashed line is simulation with the coherent noise

Solid line is simulation with NO coherent noise

Addressing signal loss during coherent noise subtraction has potential to massively improve efficiency for both ROI and Hit finding



Can We Do Better?

- Important to achieving its ultimate physics goals ICARUS needs to address its hit finding efficiency problem
    - Particularly for tracks running parallel to the sense wire planes
    - Also for tracks nearly perpendicular to sense wire planes

- Fixing the parallel track problem means
    - Fixing the coherent noise subtraction to not remove signal (is this possible?)
    - Developing a new method to identify signals
        - Using the raw input waveforms?

- Fixing the perpendicular track problem
    - What is the correct strategy for returning “hits” for long pulse trains
    - Is deconvolution the correct way to estimate deposited charge for these pulses?

Handoff to Francois

To explain that we can do better and how to do so…