

# Gauss Hit Finder Proposed Changes

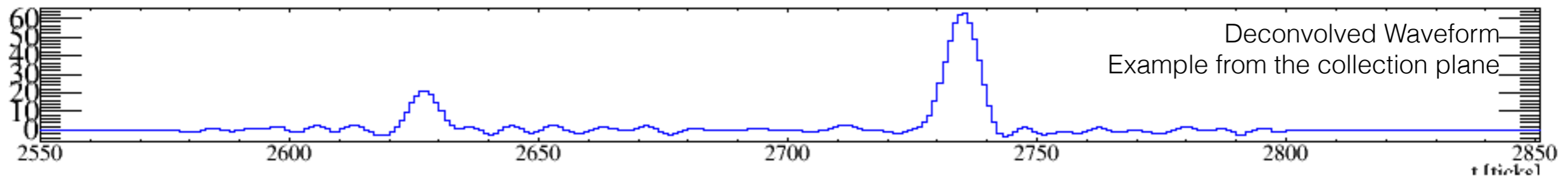
Tracy Usher  
February 2, 2016

# Overview

- Reminder of the strategy for finding/fitting pulses
- Proposed modifications to the GausHitFinder\_module
- A few example waveforms for illustration
- Summary

# Gaushit Finder - Reminder

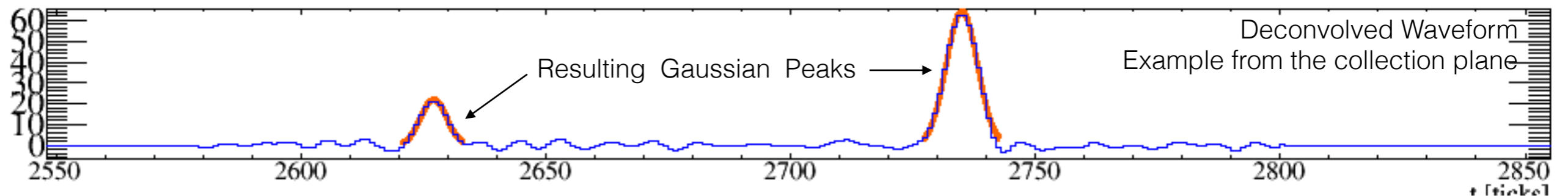
- Hit finder takes as input an “ROI” waveform from the output of the deconvolution process:



- The hit finder assumes that pulses to be found will be (positive) unipolar gaussian shaped pulses
- It also assumes that multiple pulses that are merged together can be modeled as individual gaussians
- This latter assumption is problematic when tracks start to run along the drift direction or are run parallel to a wire

# Gaushit Finder - Reminder

- The hit finder has three main elements:
  - For a given ROI, search the waveform for maxima points over a fixed threshold and then the corresponding minima points on either side
    - I'll call these "snippets"
  - Loop over these "snippets" to perform the following two tasks:
    - If two adjacent snippets are overlapping then merge them together
    - Fit the candidate snippet(s) to a Gaussian shape



- Return parameters describing the fit pulse
  - Peak time, width, integrated area, etc.

# A Few Details...

- The fit stage actually consists of two parts:
  - The initial gaussian fit to the candidate peak(s)
  - If the initial fit chi-square is “poor” then add another peak and refit.
    - Extra peak is added near the end of the current fit range
      - this can lead to negative amplitude pulses!
    - Take the second fit solution if the chi-square is better than first fit
- Fit initialization:
  - Peak positions are constrained to lie within the (merged) snippet
  - Peak width is initially set to an expected (per plane) value

# A Few More Details...

- Current code limits “merging” of snippets to 2 peaks
  - Long waveforms consisting of many peaks will be processed in a series of fits to merged snippets containing 2 peaks at a time - plus refits if poor chi-square
    - Source of observation that area under waveform does not equal sum of fit peak areas
- Fits which return a “bad” chi-square, even after adding an extra peak and re-fitting, will be rejected
  - There is no record of these rejections in the output
  - Sometimes the fit can fail even for “good” hits...
- There is an opportunity to address various output issues
  - Hit number is on snippet, not on ROI
  - “sumADC” is sum over entire snippet, not just the returned fit peak
  - etc.

# Proposed Modifications to Hit Finder

- Restructure
  - Three main elements of hit finder broken into independent functions
    - hopefully easier to maintain/update in future
  - Single function to do gaussian fits to handle both first and second fits
    - Hopefully less complexity to future developers
- Tweaks
  - Allow multi-peak fits to long pulses
    - Up to 10 peaks fit simultaneously
      - Utilizing an existing fhicl parameter which was already set to 10
      - Value is arbitrary (?) but observation is that going past ~15 begins to really impact algorithm time

# Proposed Modifications to Hit Finder

- Tweaks (continued)
  - Waveforms with more than 10 (see above `fhicl` parameter) candidate peaks are NOT fit, a single hit will be returned for this pulse
    - Time is center of pulse, width is taken from length of pulse, height from average of ADC values
    - Similar to `cchitfinder`
  - Initializing the fit
    - Candidate peak positions constrained within peak range, not full snippet
    - Initial peak widths from width of candidate peak, not fixed value
  - Limit poor chi-square refits
    - Still add extra peak near end but ONLY if amplitude will be positive
  - Chi-square failures now returned as single hit as for long pulses

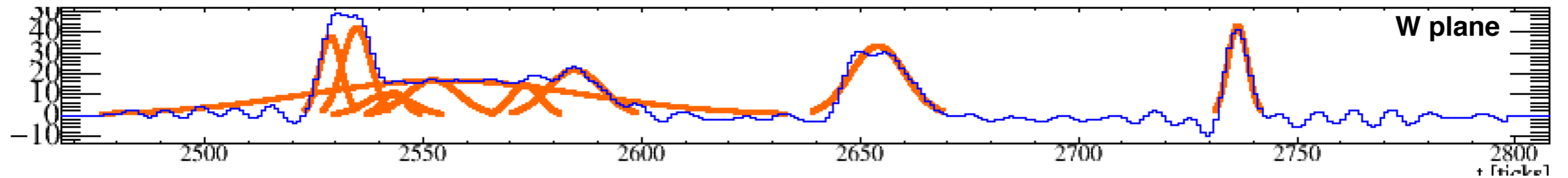


# Proposed Modifications to Hit Finder

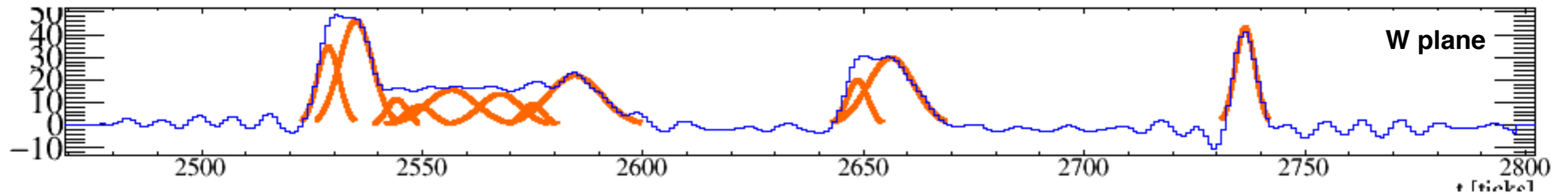
- Goals:
  - Make module simpler
    - Primarily for ease of future maintainability
    - ~920 lines of code down to ~760
  - No change to results for single gaussian fits
  - Improve fits for potentially true multi hit waveforms
  - Provide alternative output for complex waveforms which probably don't have gaussian peak structures
  - Output for all fit attempts
    - e.g. let user decide to ignore poor chi-square fits
  - Improve information going into the output hits

# Some Comparisons

## Original Hit Finder

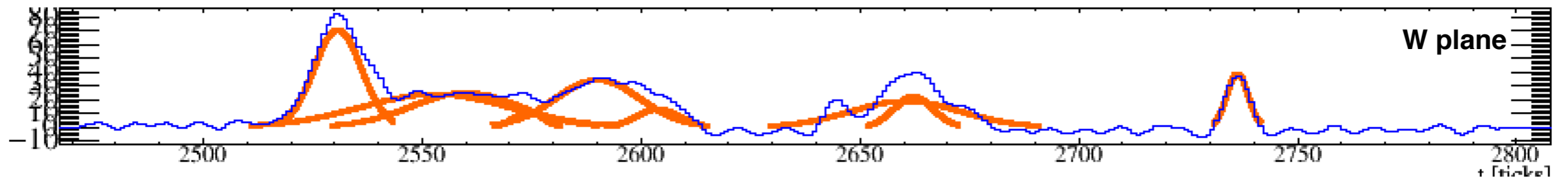


## Updated Hit Finder

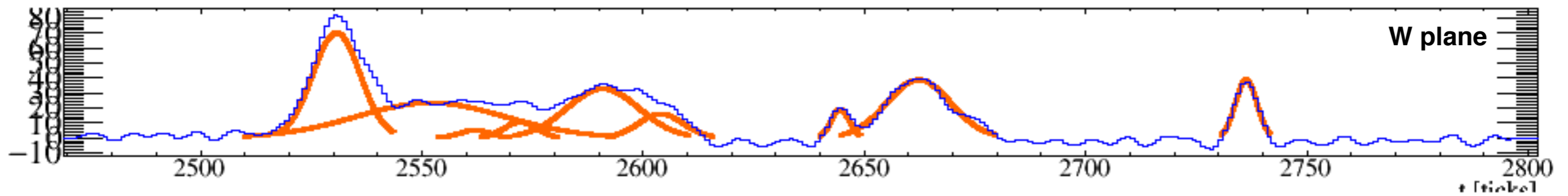


# Some Comparisons

## Original Hit Finder

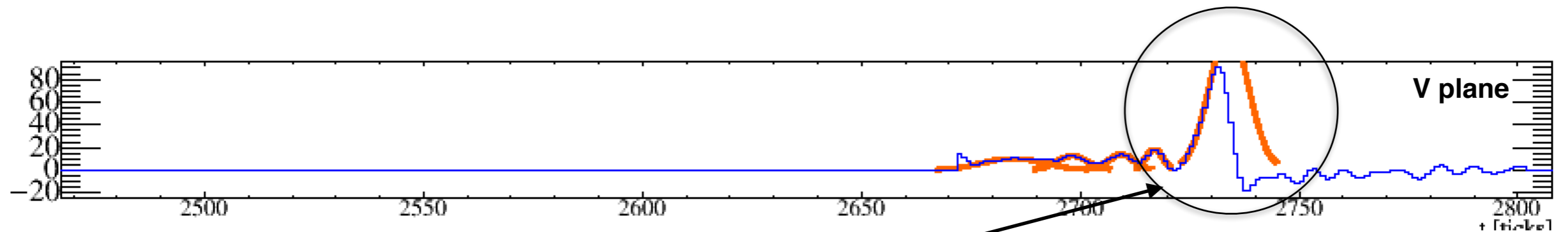


## Updated Hit Finder



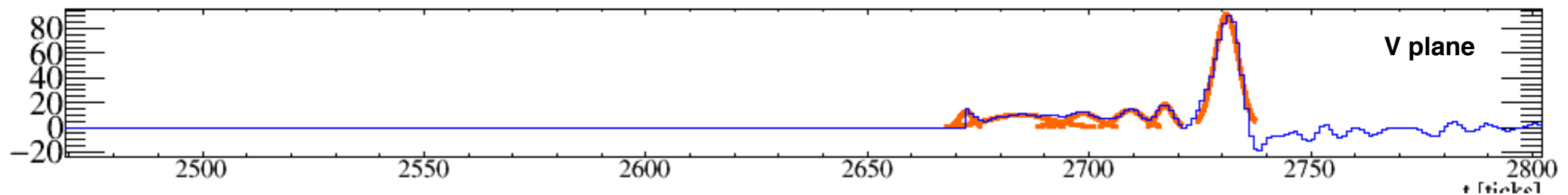
# Some Comparisons

## Original Hit Finder

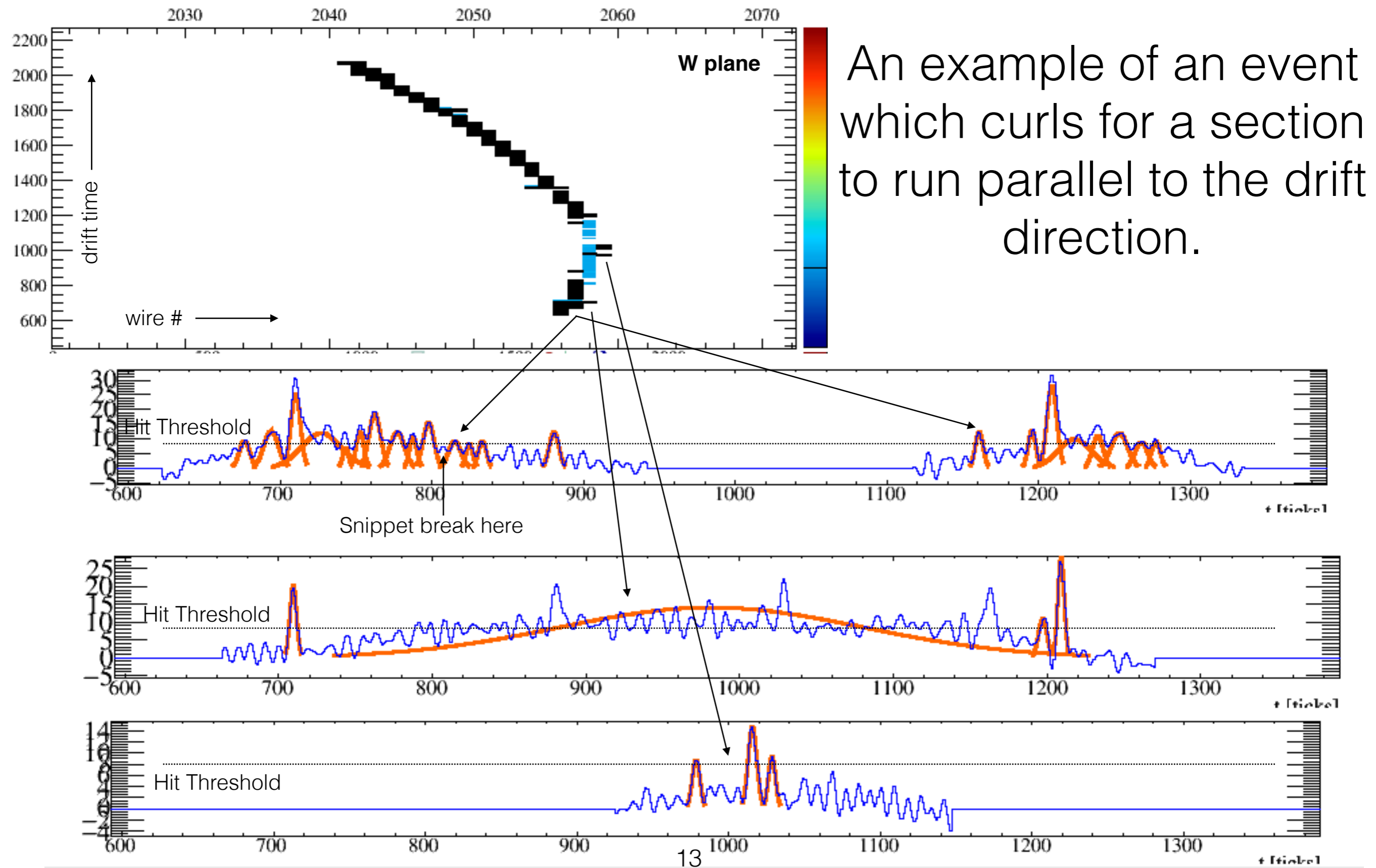


This happens on a relatively frequent basis

## Updated Hit Finder



# Long Pulses



# Summary

- Propose an updated version of the GausHitFinder\_module
  - Same strategy and logic of the initial peak finding
  - Some tweaks and fixes to improve performance in special cases
    - Better initialization of fits
    - Allow multi hit fits of longer pulse trains
    - Recovery for fit failures, mechanism to summarize overly long pulse trains
  - Essentially the same timing as the original algorithm
    - Depends on number of gaussians allowed in pulse trains
- GausHitFinder\_module is shared code - will impact other users
  - Tingjun has run a trial and sees no adverse impact for dune/lariat
- On larreco feature branch: feature/usher\_gaushitrestructure