Noise & cross-talk: Isolating spurious signals

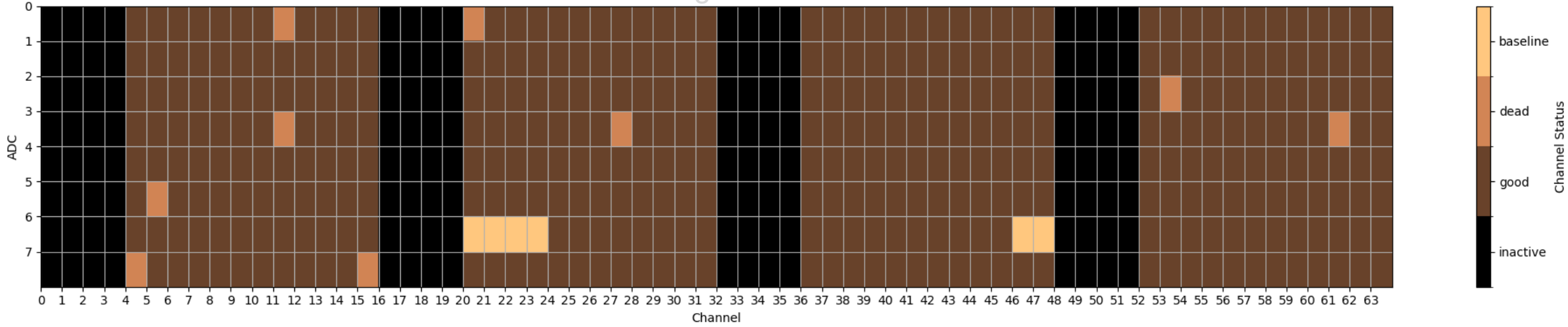
James Vincent Mead

Data sample

- ***File being used on all subsequent slides:***
  - mpd_run_hvramp_rctl_104_p123.FLOW.hdf5

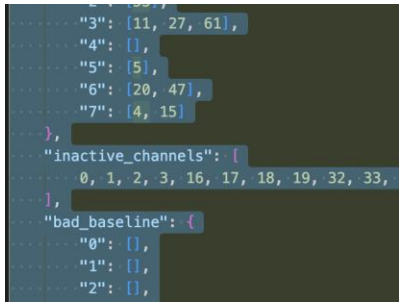
2x2 channels

Excuse the choice of colour scheme – I will change this



- ***Would like to standardise:***
  - Channel status format
  - Channel map format
  - Their interface

Human readable



Some channels have more than one status

&

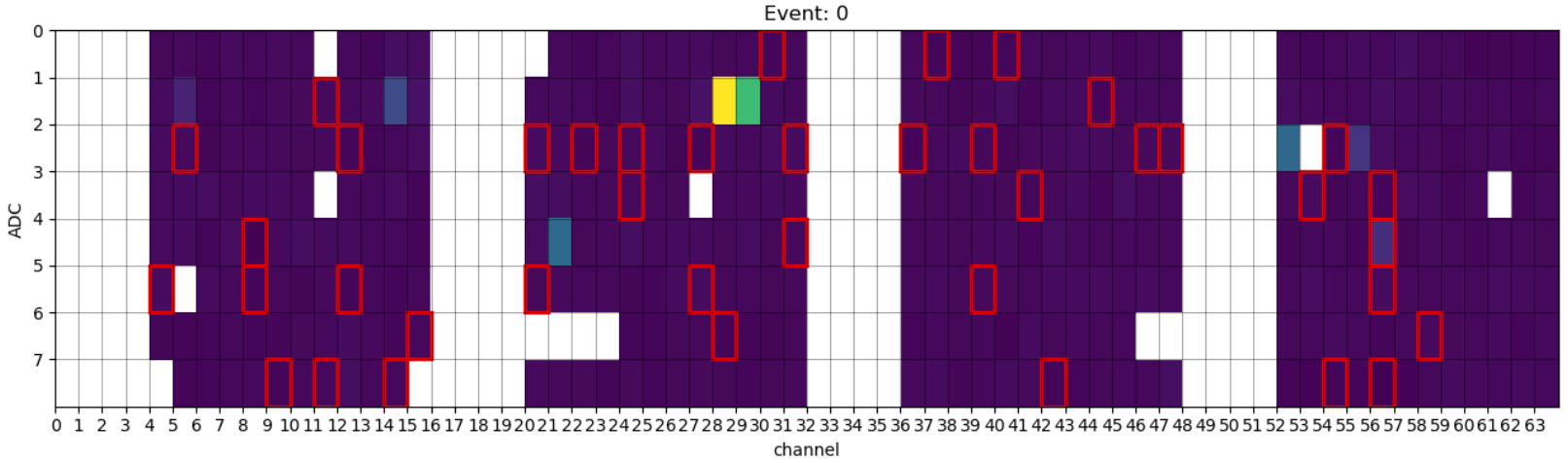
numpy readable

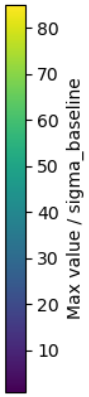
```
-1.0,-1.0,-1.0,-1.0,0.0,0.0,0.0,0.0,0.0,
-1.0,-1.0,-1.0,-1.0,0.0,0.0,0.0,0.0,0.0,
-1.0,-1.0,-1.0,-1.0,0.0,0.0,0.0,0.0,0.0,
-1.0,-1.0,-1.0,-1.0,0.0,0.0,0.0,0.0,0.0,
-1.0,-1.0,-1.0,-1.0,0.0,0.0,0.0,0.0,0.0,
-1.0,-1.0,-1.0,-1.0,0.0,1.0,0.0,0.0,0.0,
-1.0,-1.0,-1.0,-1.0,0.0,0.0,0.0,0.0,0.0,
-1.0,-1.0,-1.0,-1.0,1.0,0.0,0.0,0.0,0.0,
```

Single 2D array of [-1:2] => status
OR
Four 2D arrays of binary values

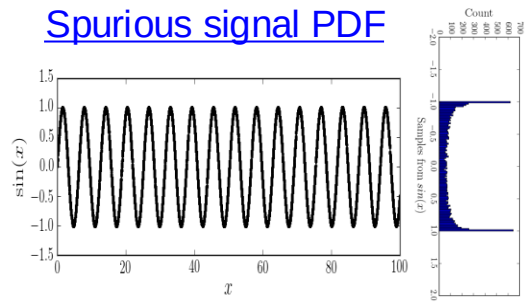
Current baseline

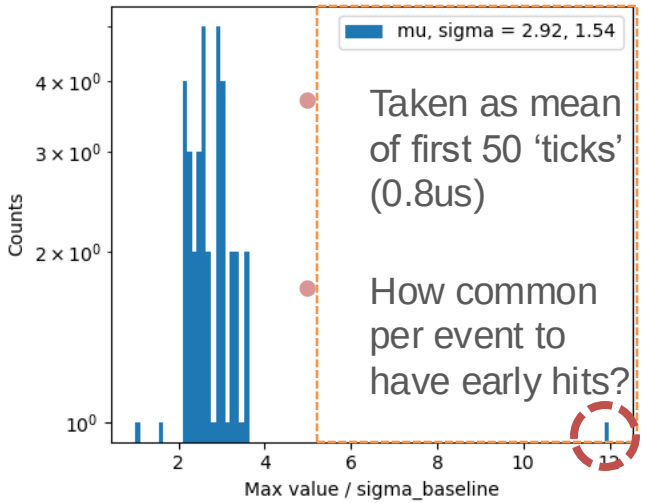
Channel with max bin in first 50 ticks





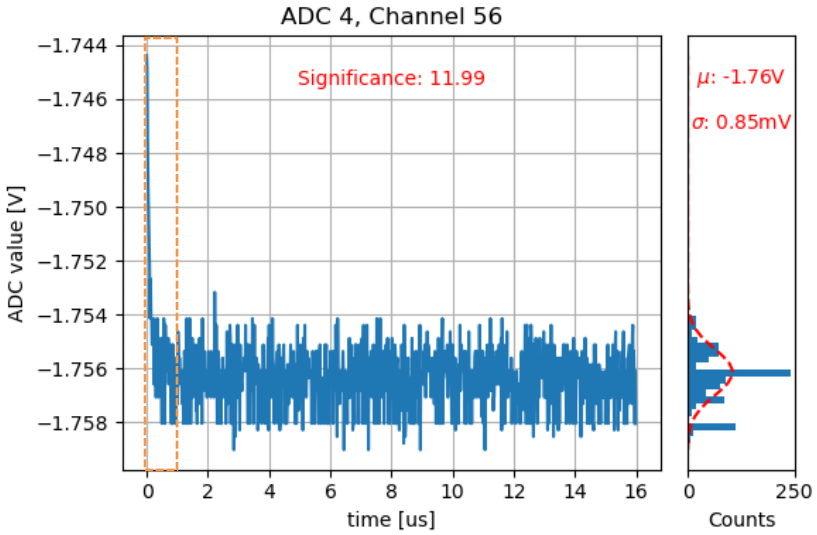
Spurious signal PDF





Taken as mean of first 50 ‘ticks’ (0.8us)

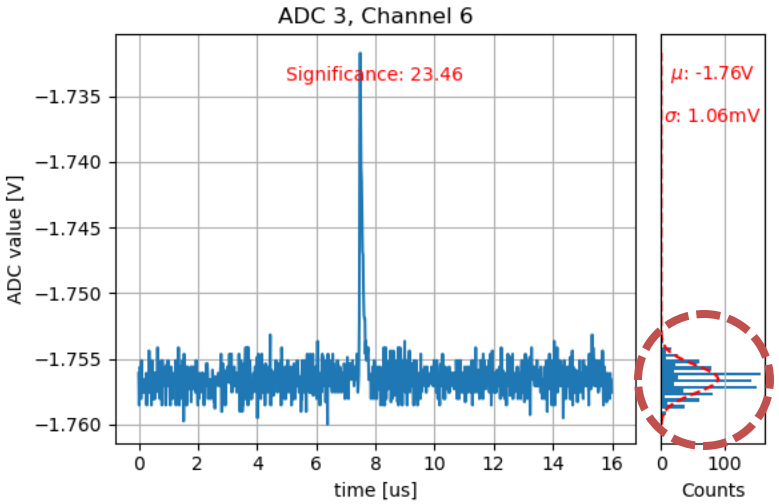
How common per event to have early hits?

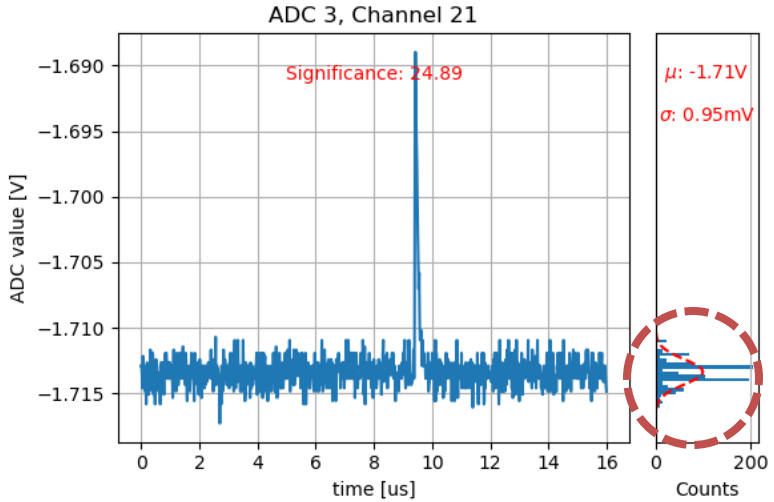


- Can baseline bias from early hits be ignored?
- Is the median unstable / susceptible to small biases from spurious signals?
- Is fitting a gaussian too slow? Can it be sped up?

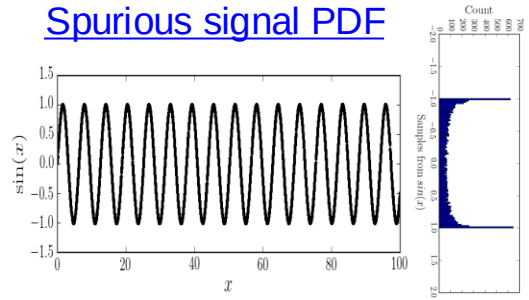
Need more robust method – max bin might be later in waveform but still include coincidence in 1st 50 ticks

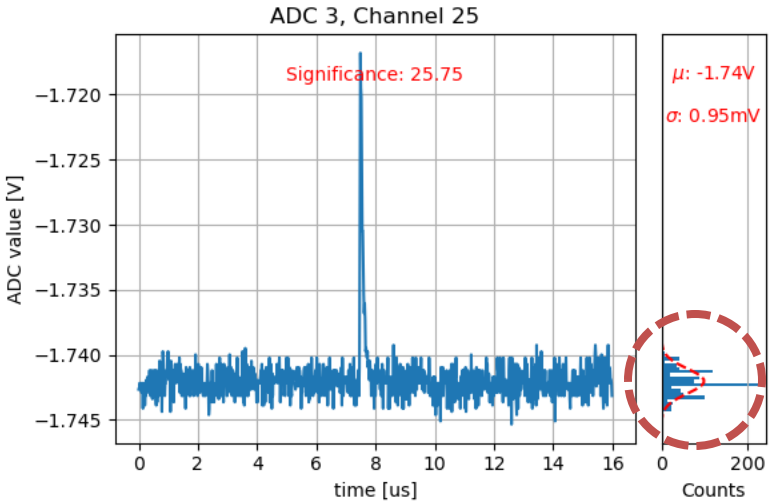
Partially discretised baseline?

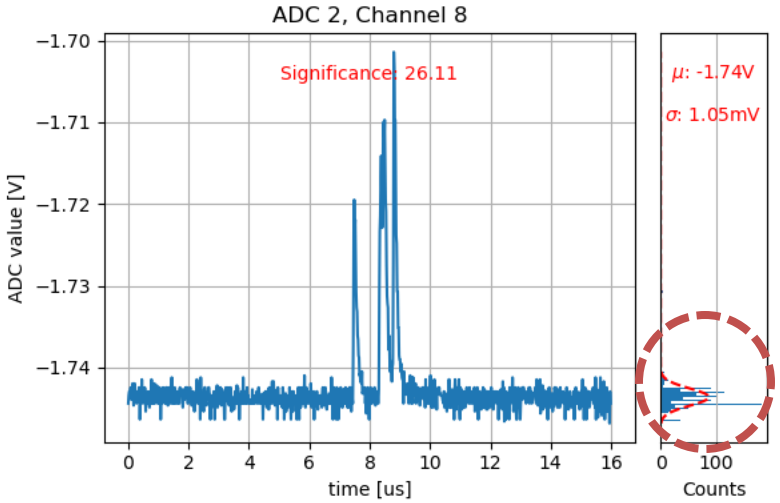


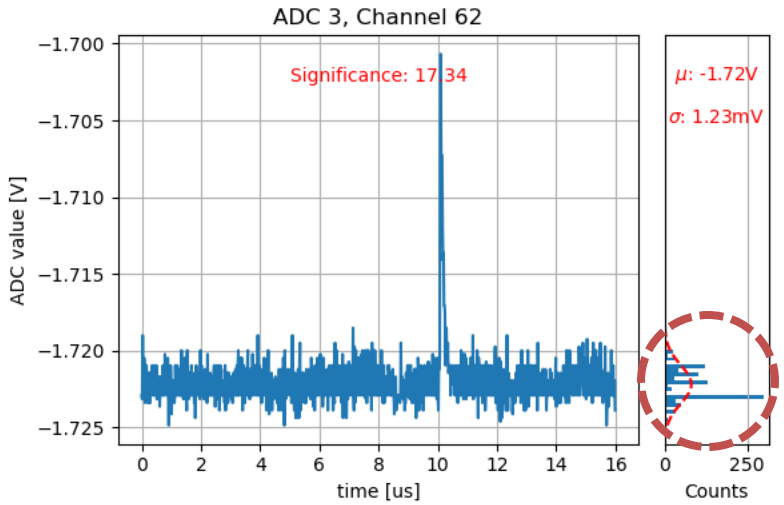


[Spurious signal PDF]



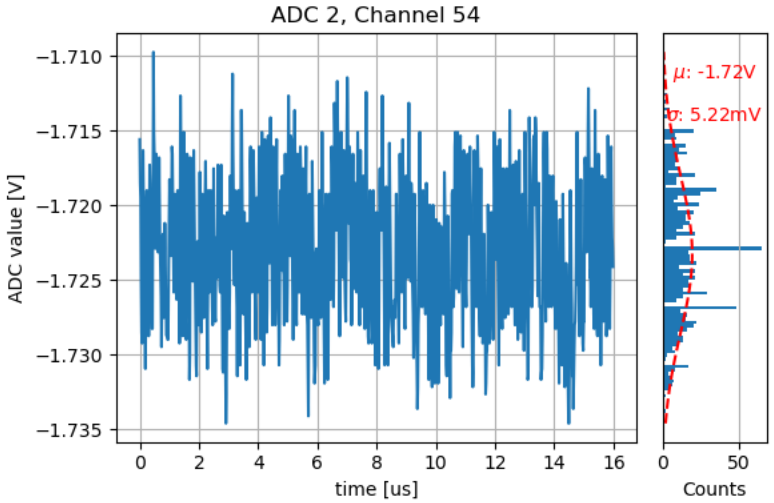


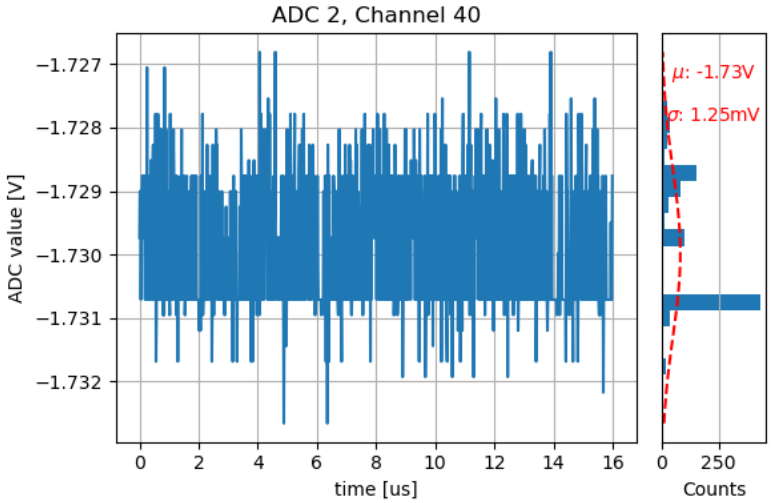


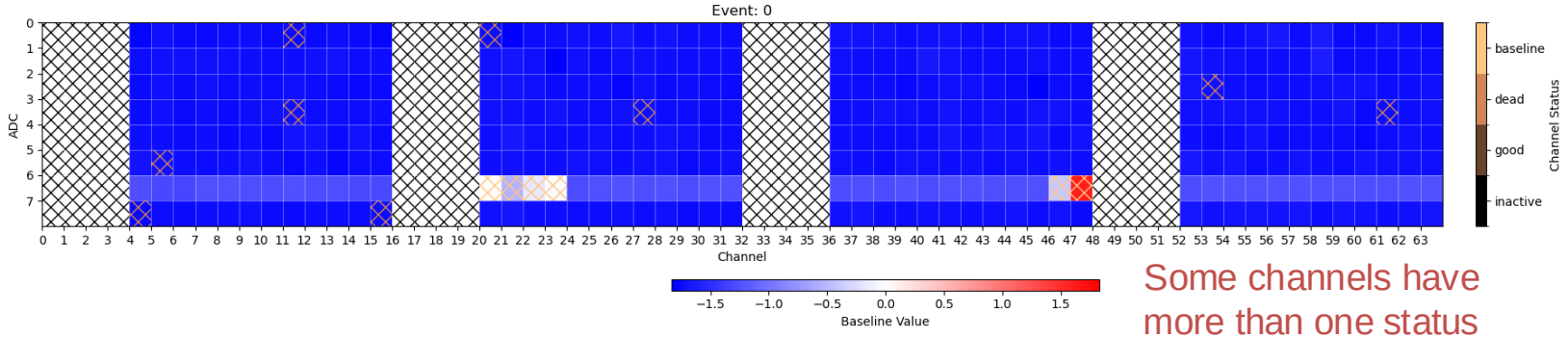


Channel status sanity check

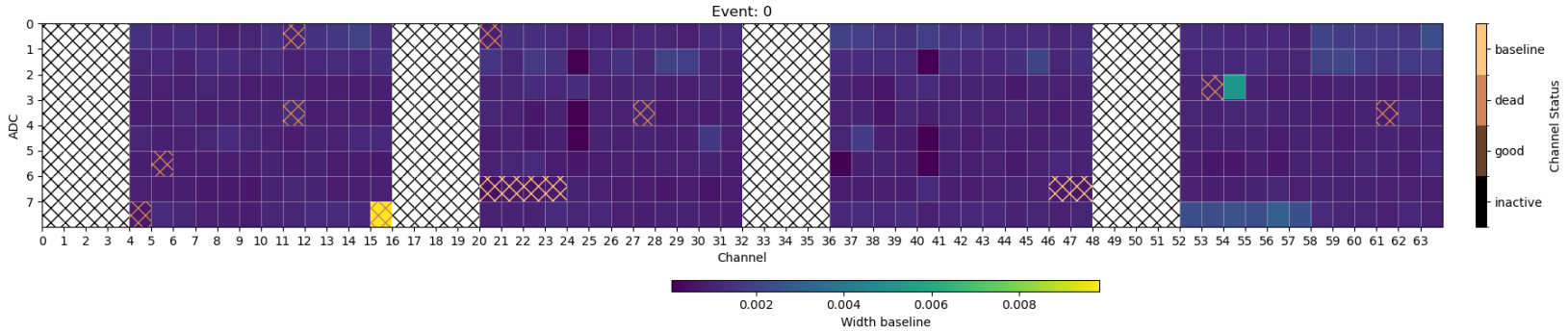
Matches expectation from status map and all fits converged

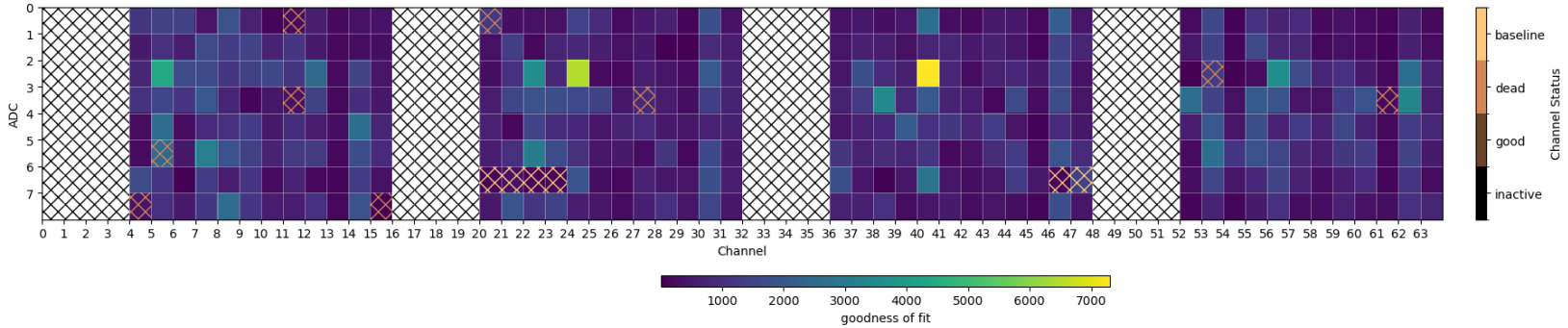




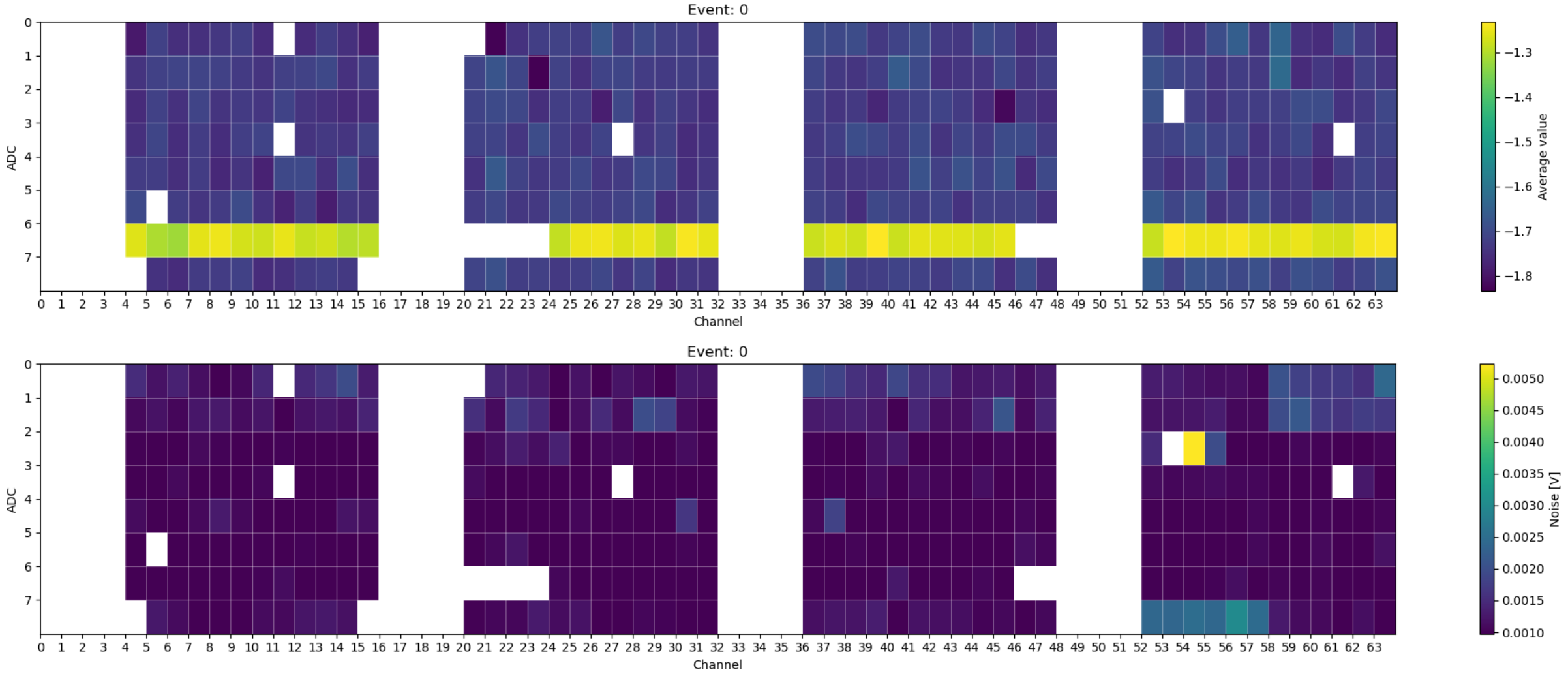


Some channels have more than one status

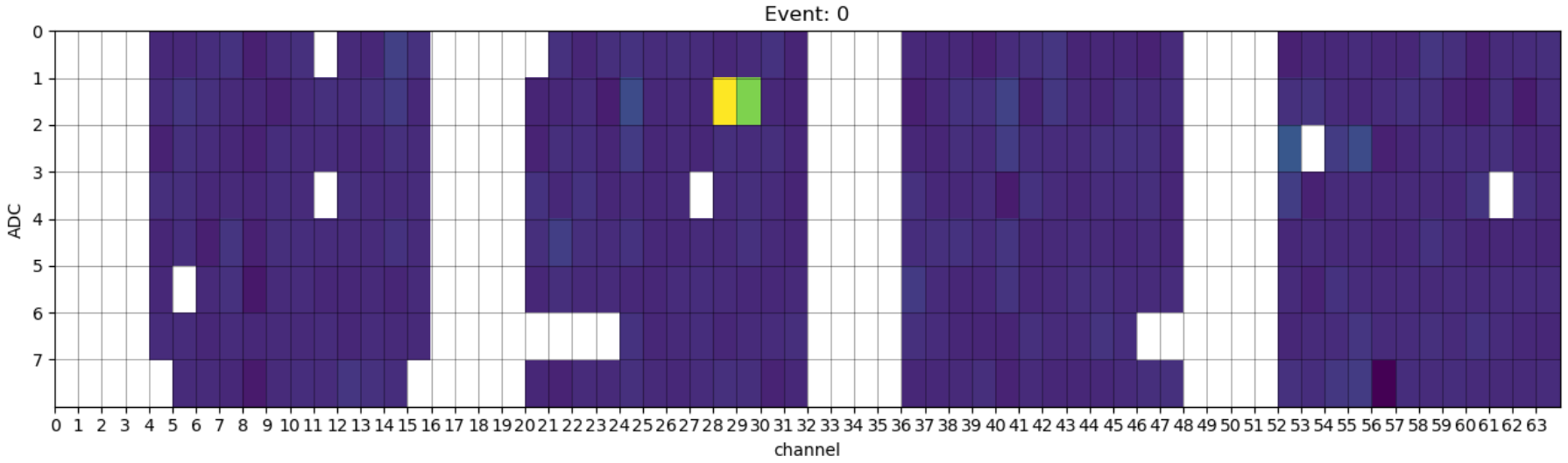


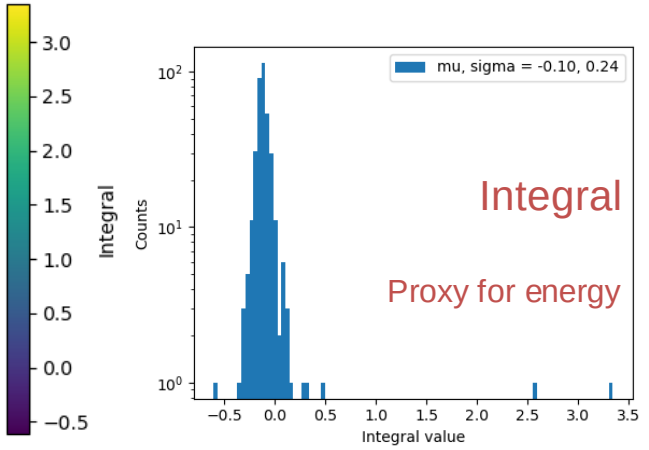


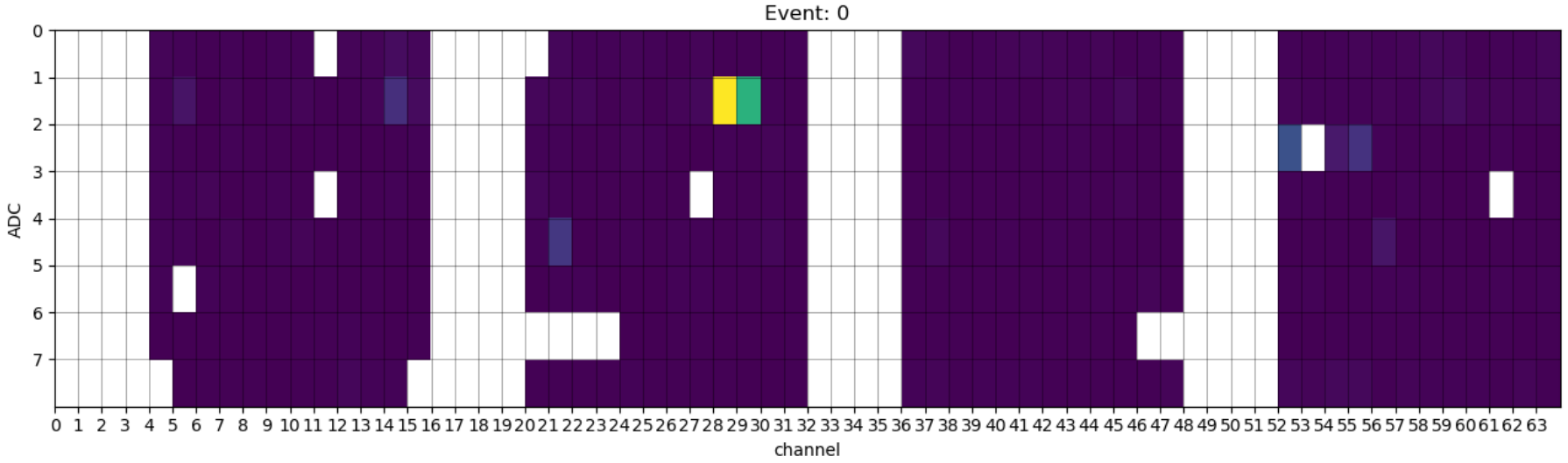
Channel baseline & noise

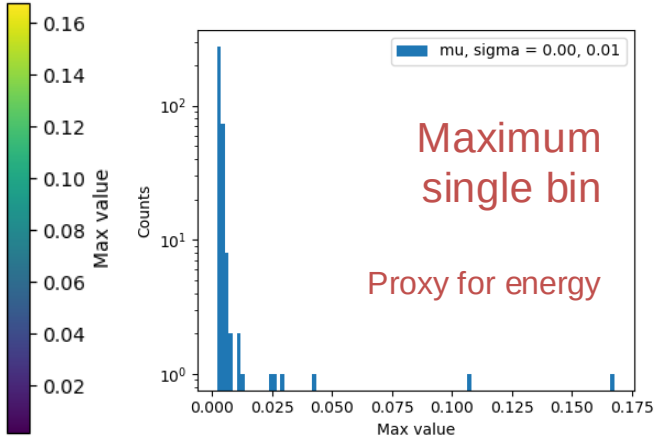


Absolute thresholds

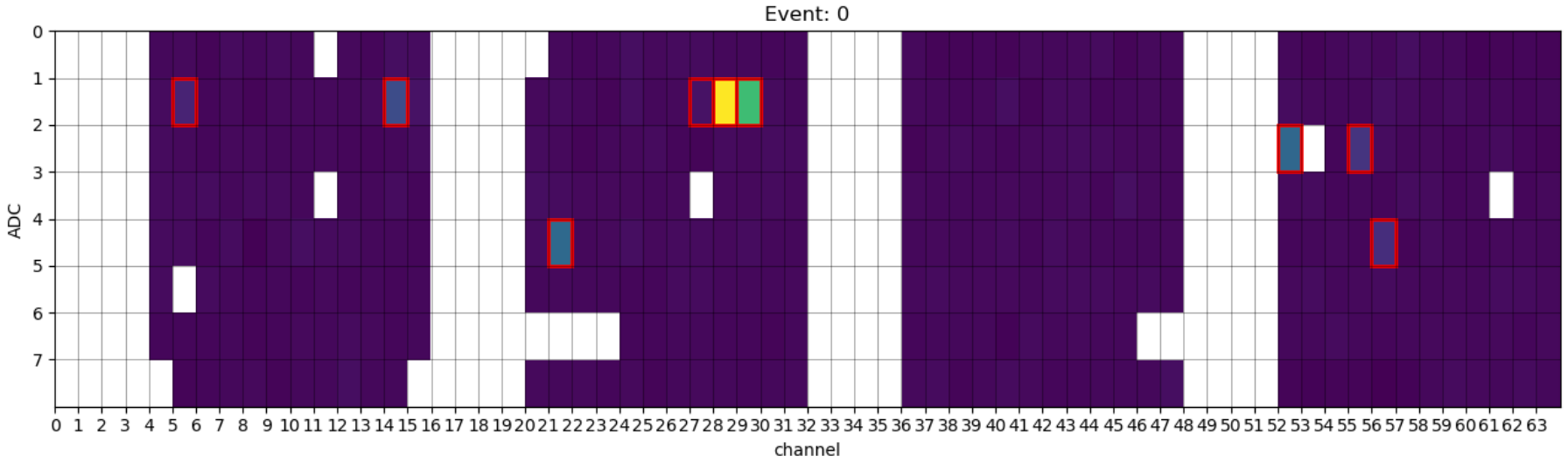


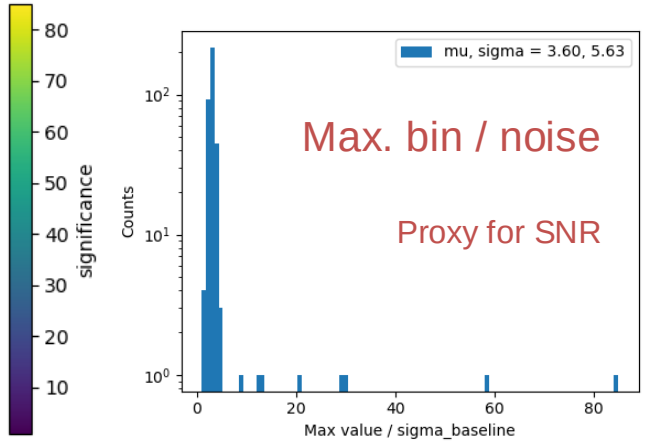


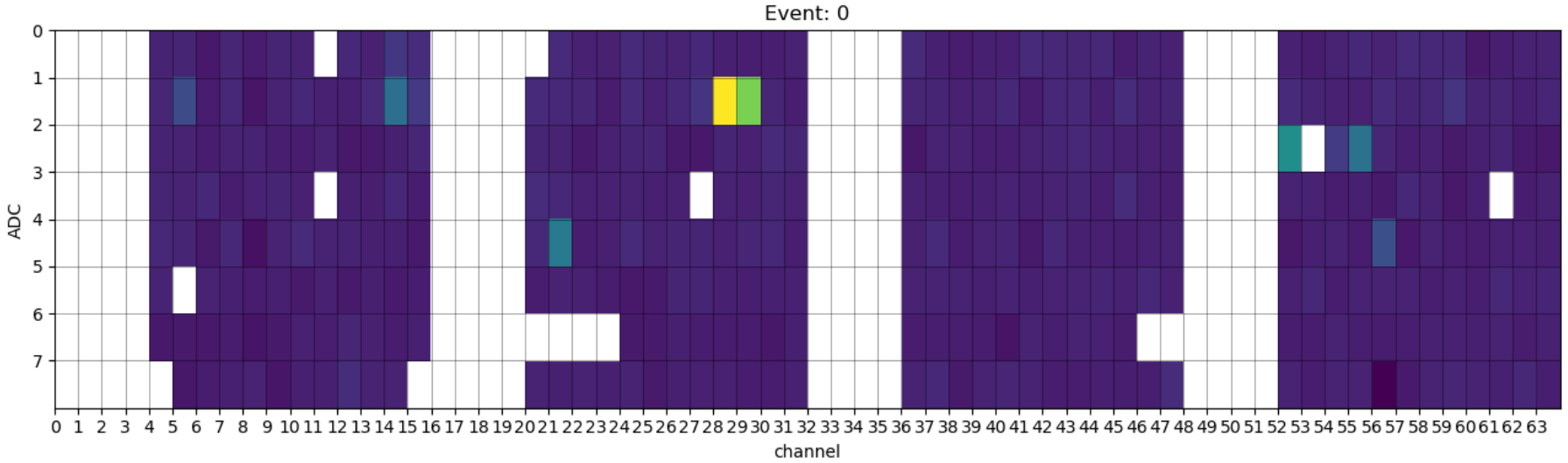


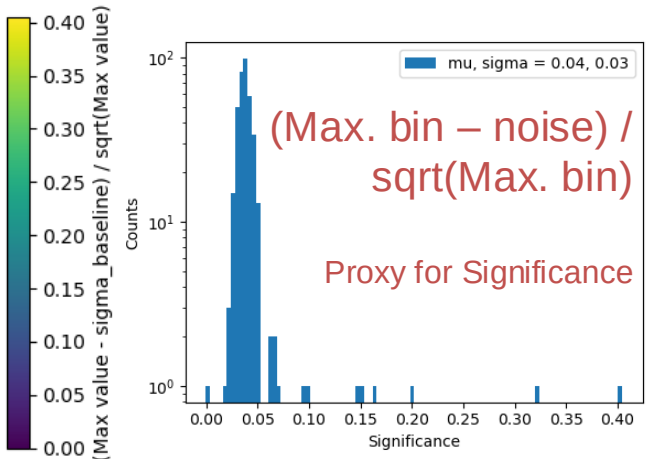


Relative thresholds

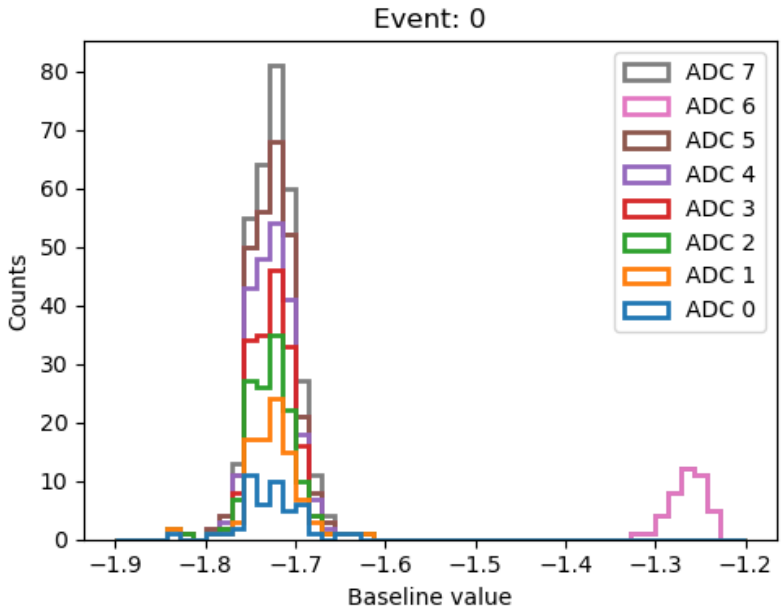


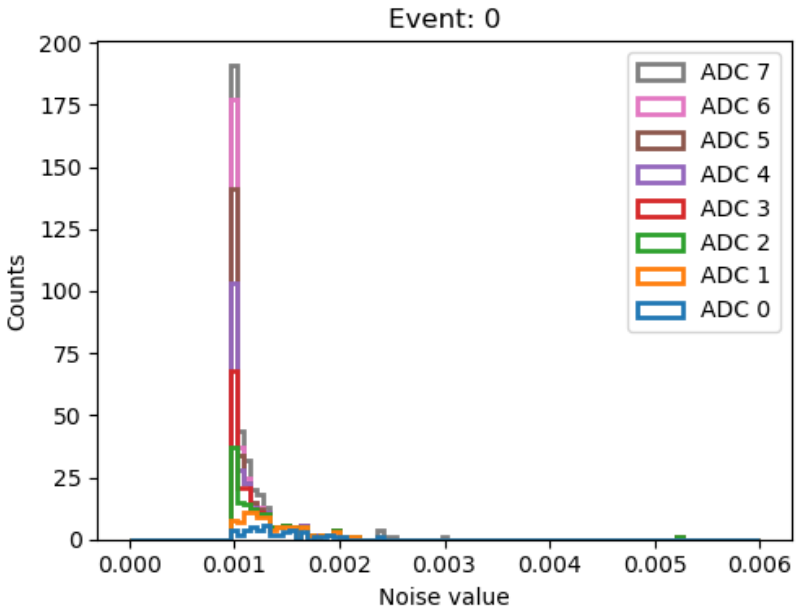


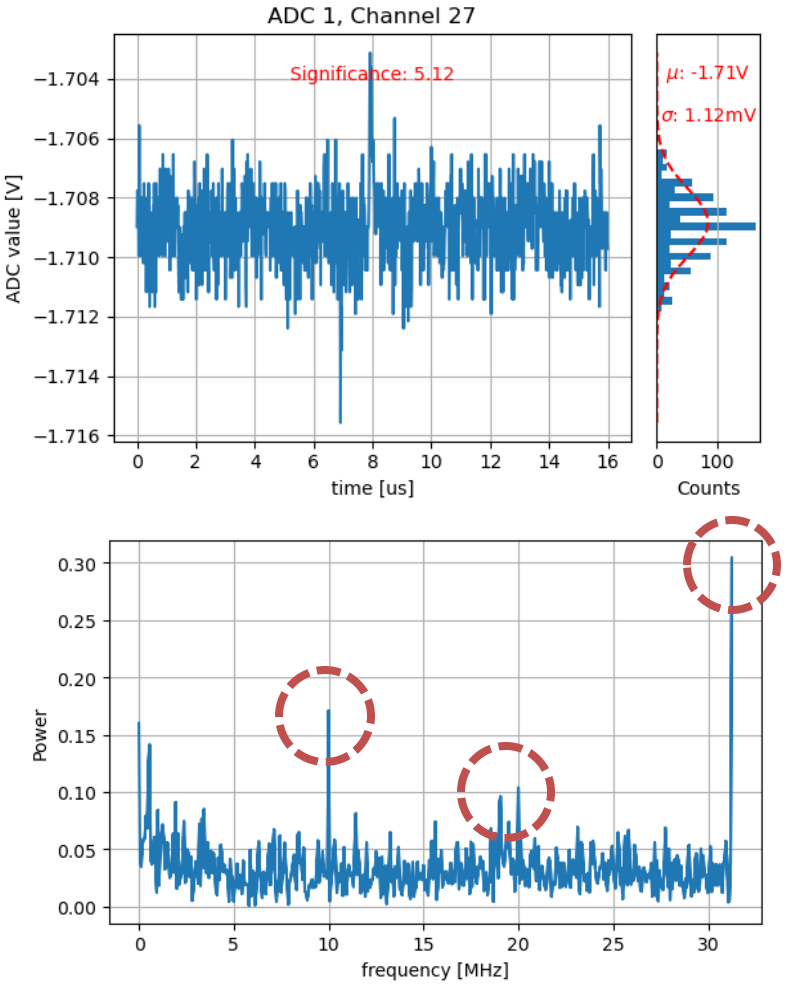




Baseline & noise by ADC

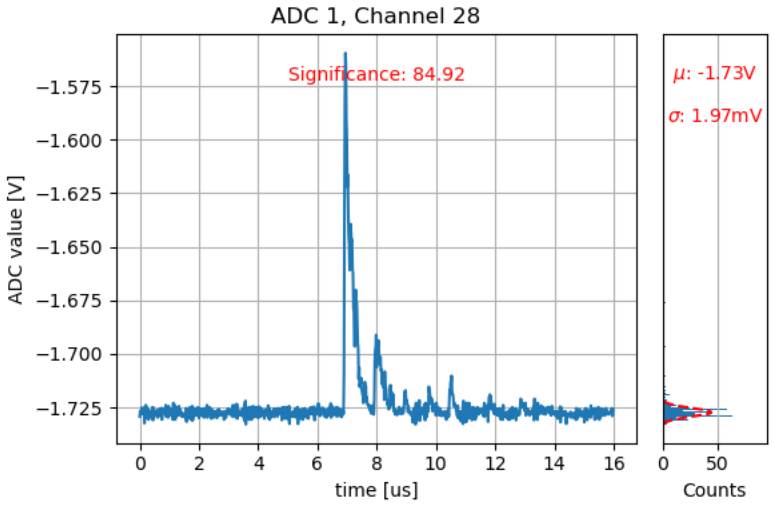


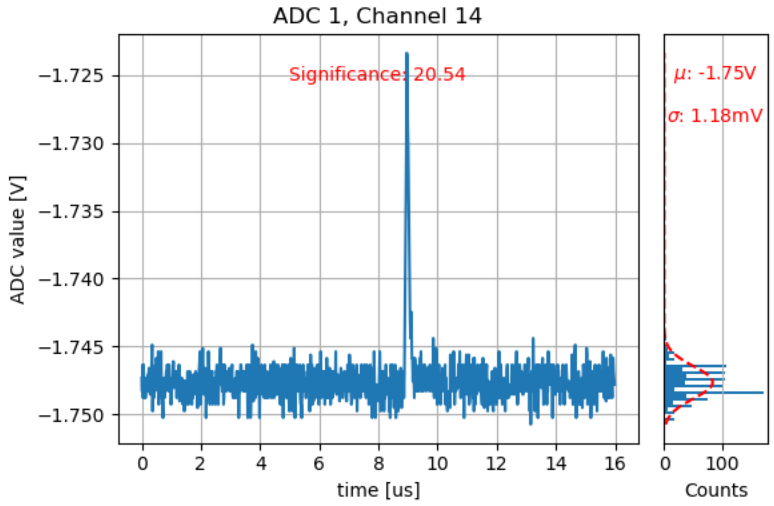


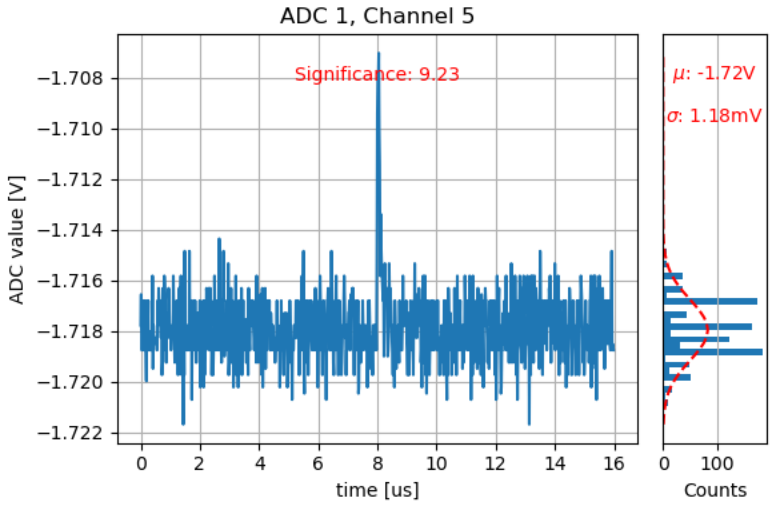


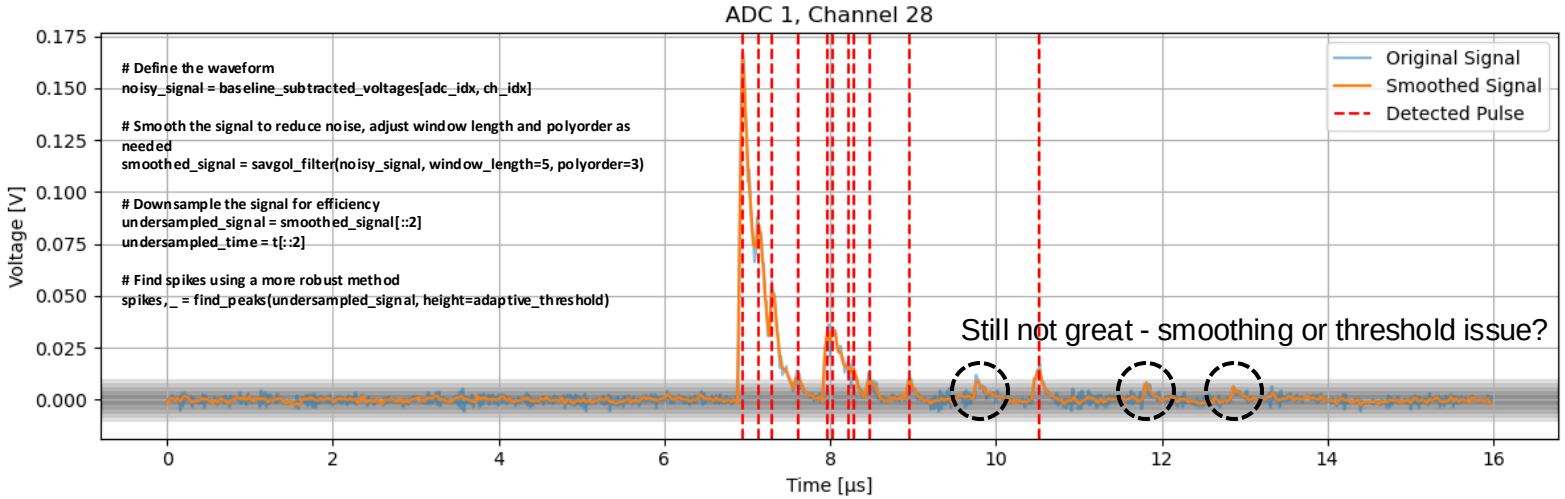
- Small pulses would be lost without channel specific thresholds
- Stick to uniform threshold and throw away small pulses?
- Even relatively quiet channels have spurious signals

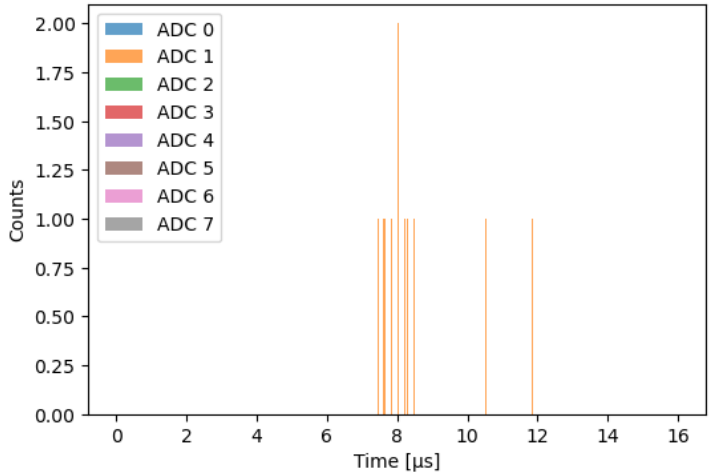
Rudimentary hit finder



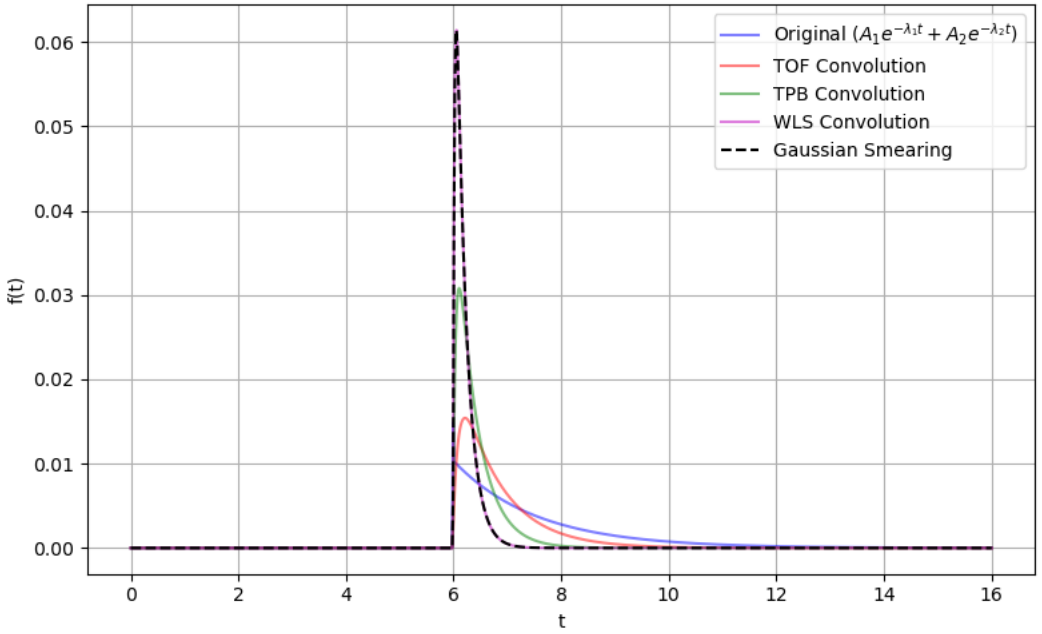


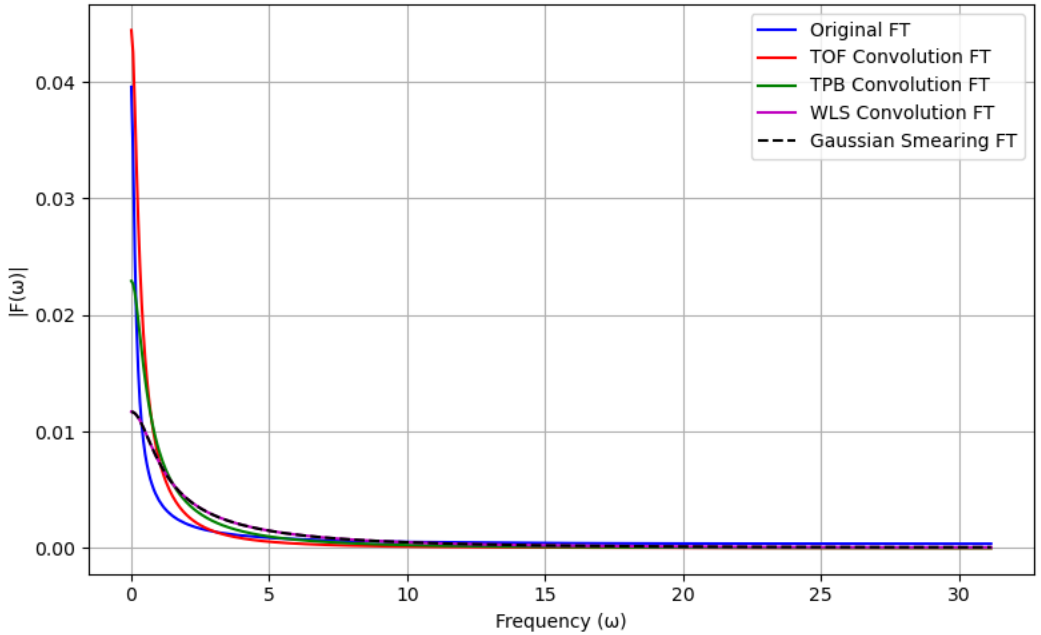


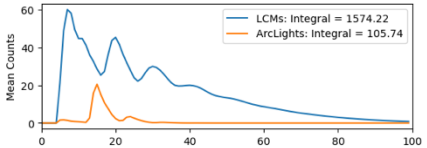




Pulse shape in Fourier space

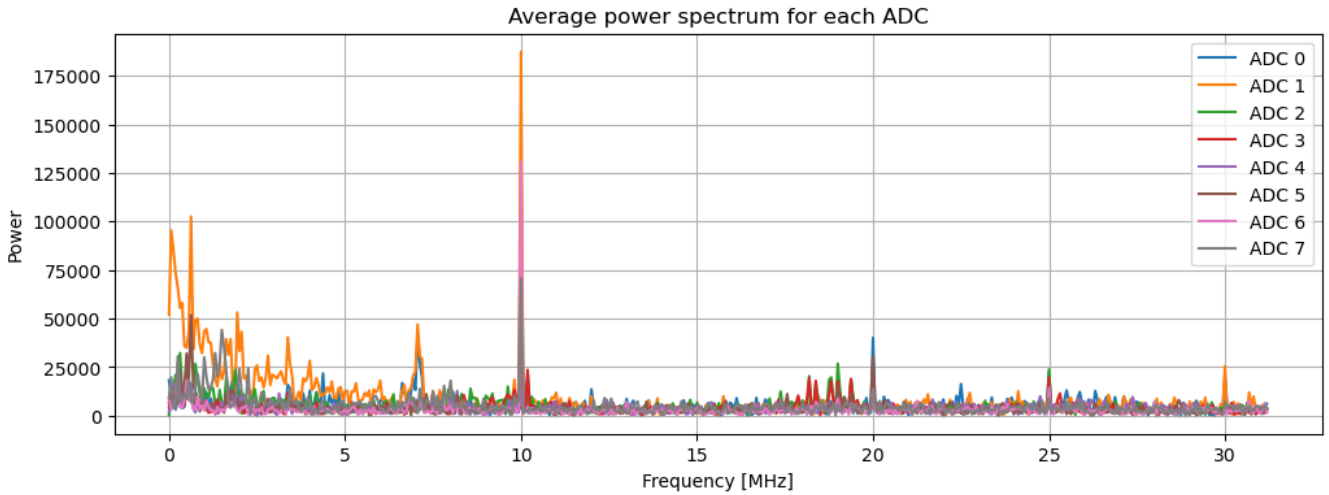


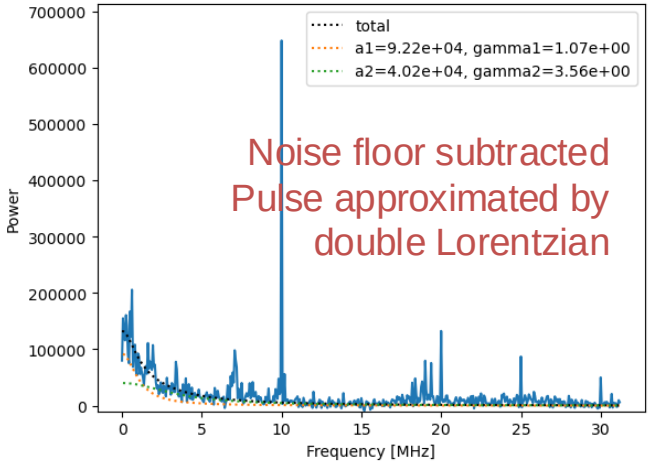




ToF approximated with an exp placeholder

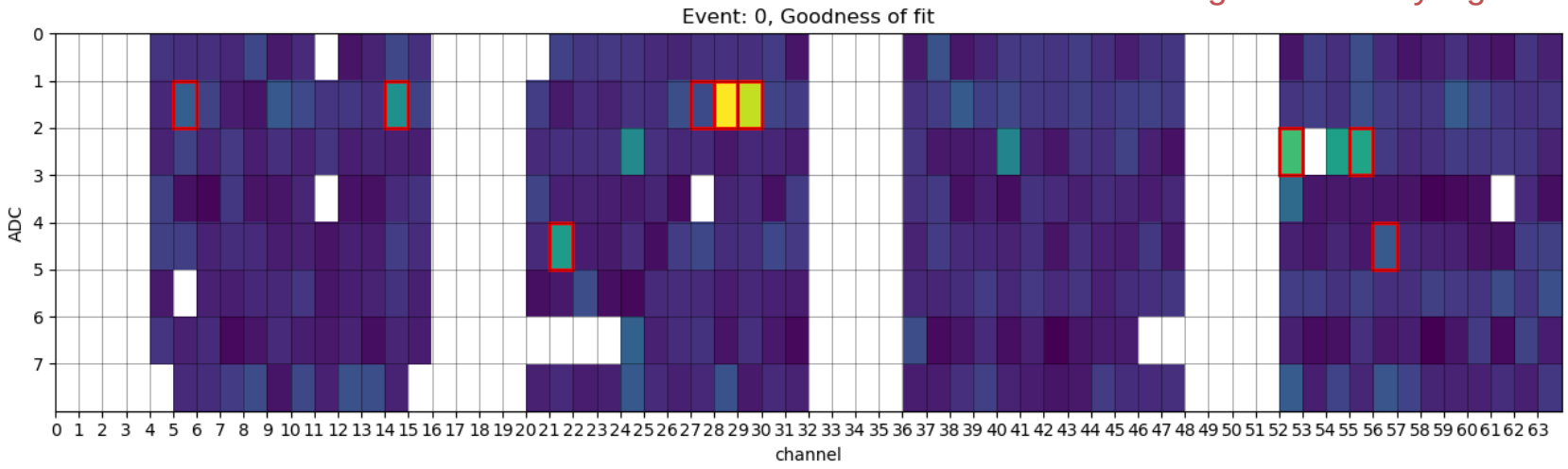
Isolating spurious signal candidates

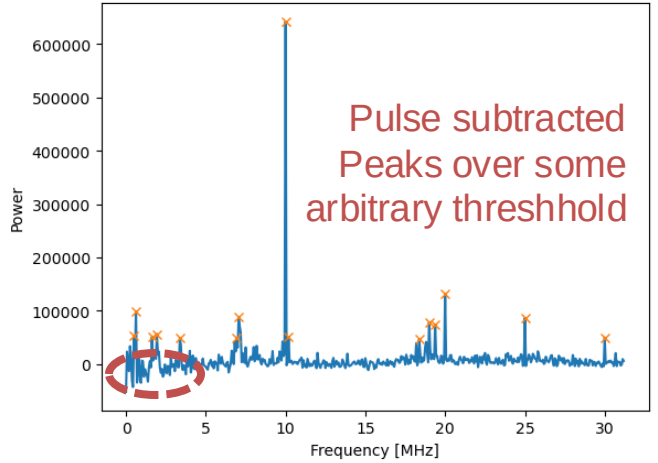




Noise floor subtracted
Pulse approximated by double Lorentzian

Badness of fit correlates with highest intensity signals

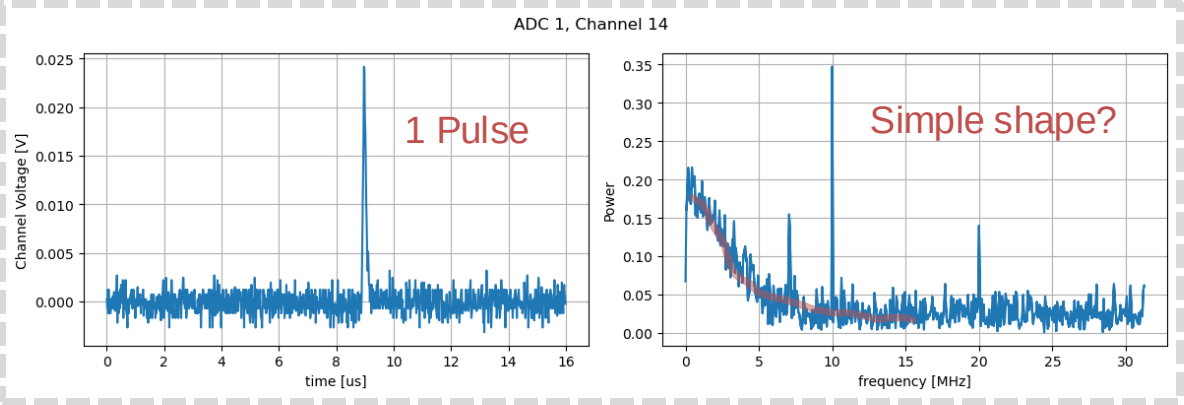


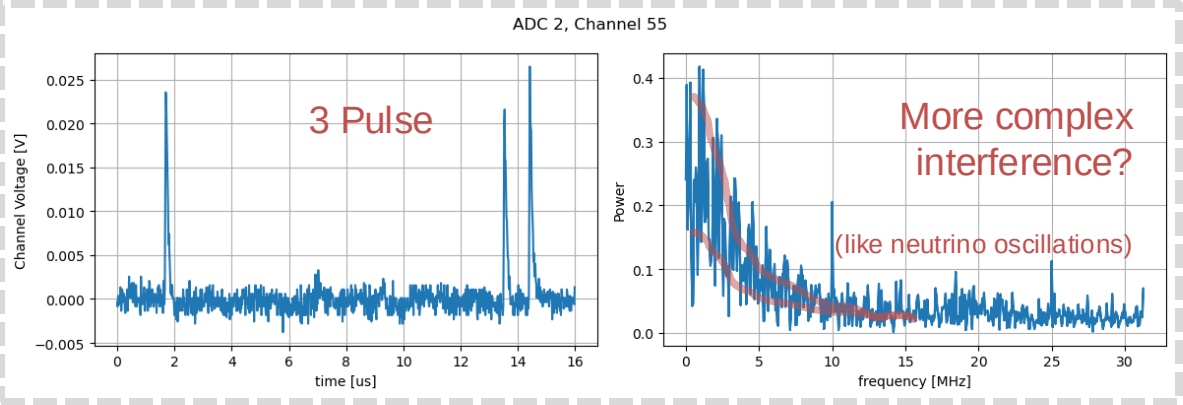


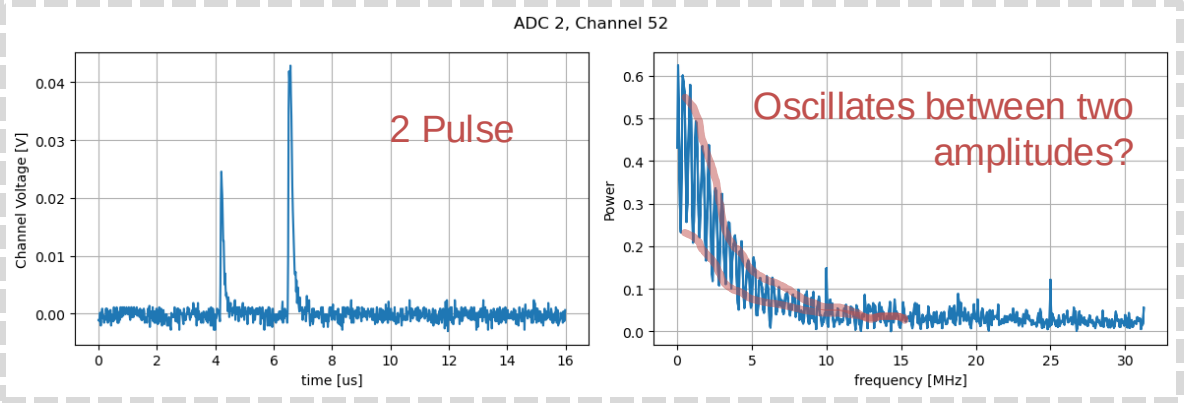
Pulse subtracted
Peaks over some arbitrary threshhold

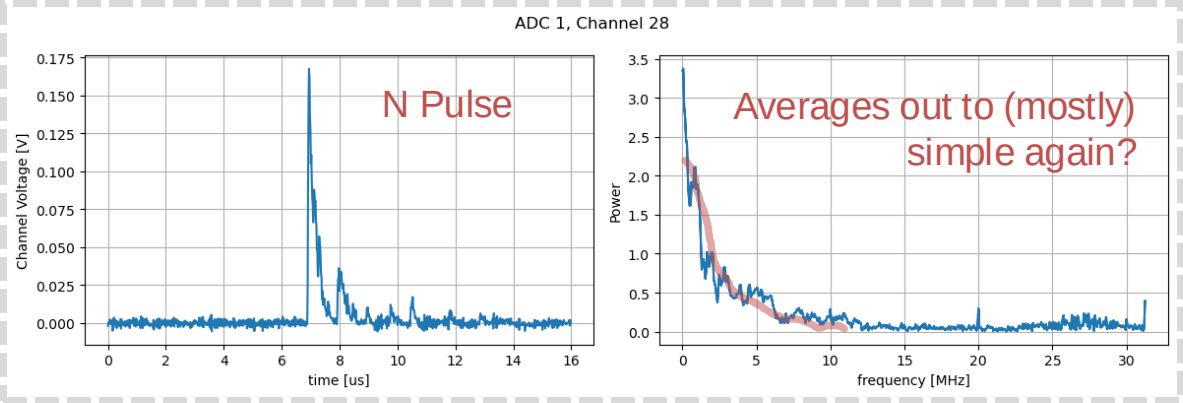
N pulse examples in freq. space

Highest yield waveforms dominate average – more likely to have multiple peaks



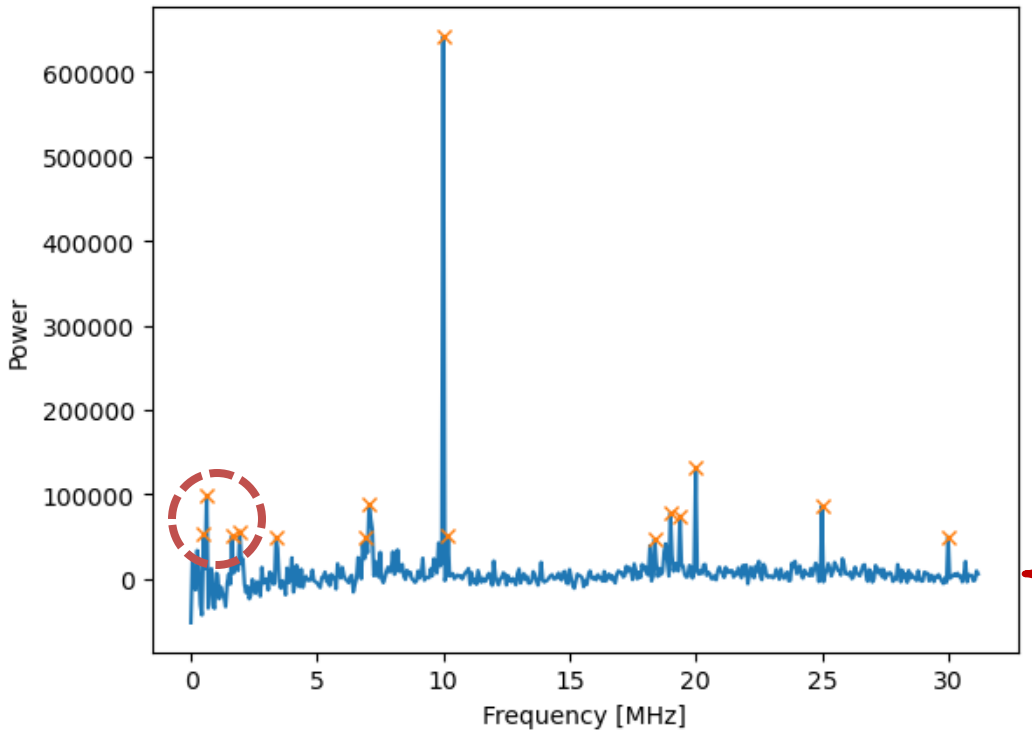


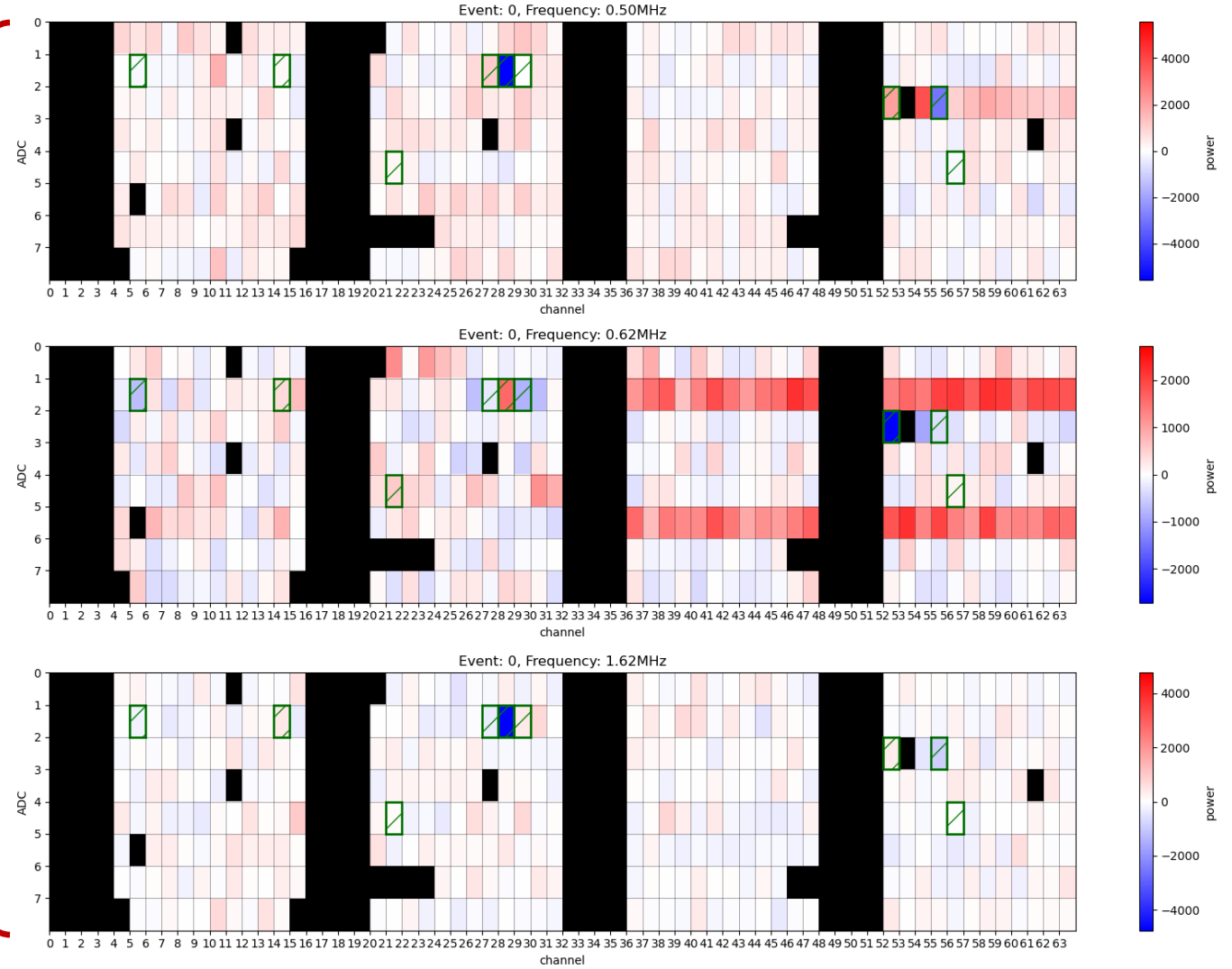




Subtracting the fit results in negative and positive fluctuations

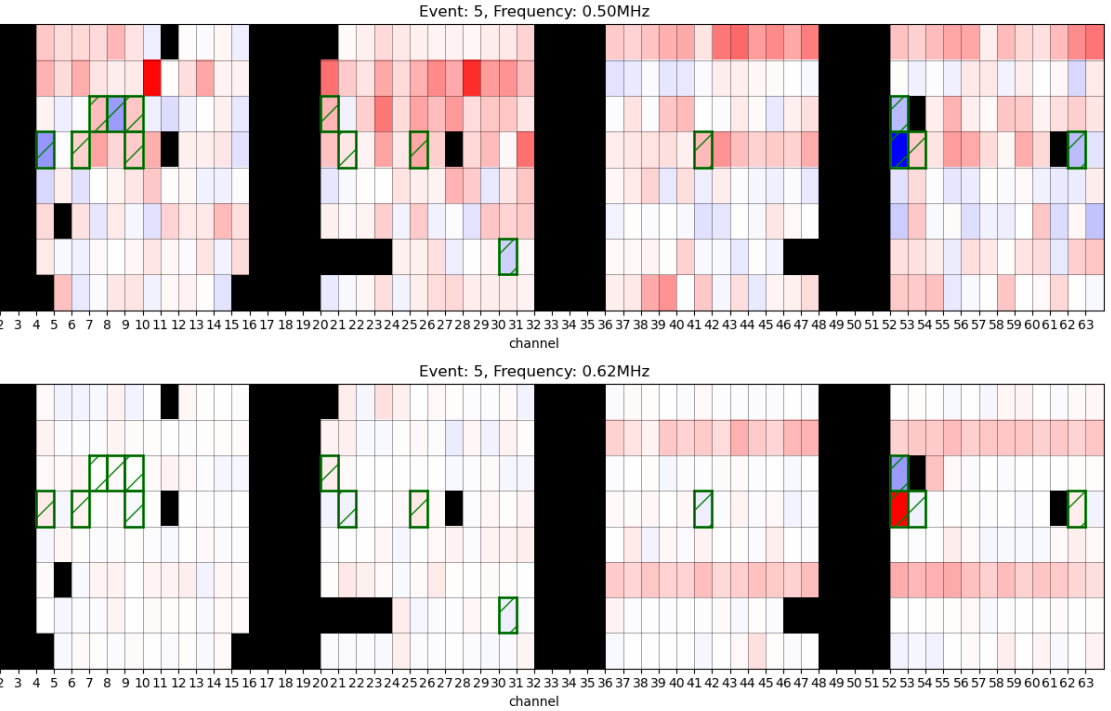
Localising the signals





Localising the signals

EVENT 5

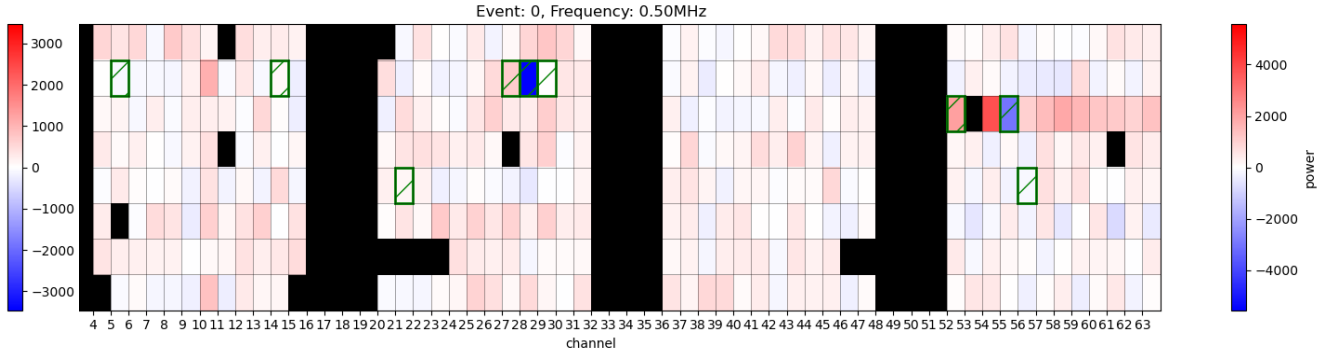


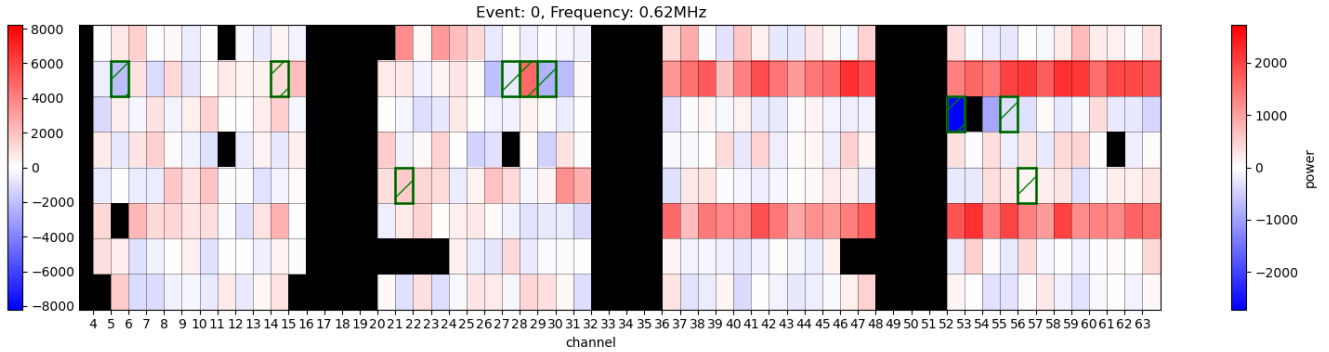
0.5 MHz
- Unclear, partially dominated by light of event
- Could be artifact of ill-fitting subtraction

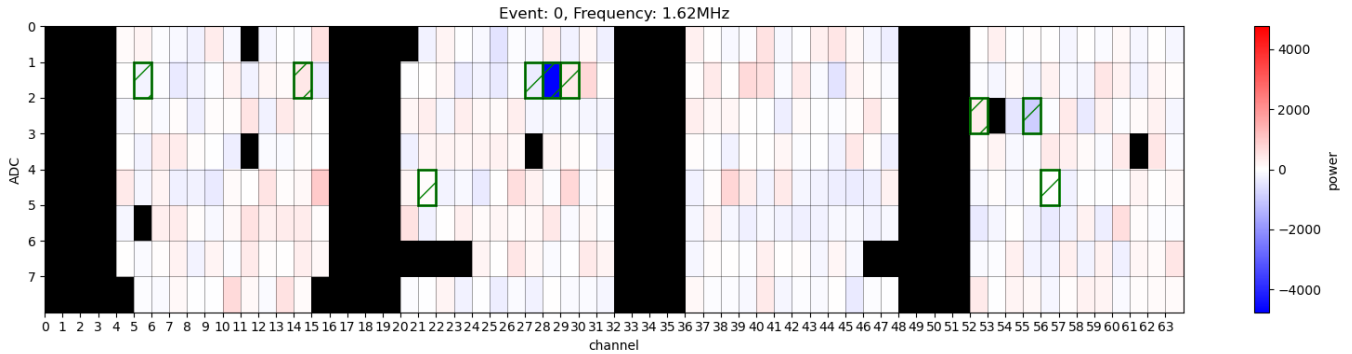
0.62 MHz
- Appears consistent across events
- No clear correlation with presence of hits

1.62 MHz
- No clear correlation across detector or with hits
- Not tagged as a significant peak in i_evt_lrs=5

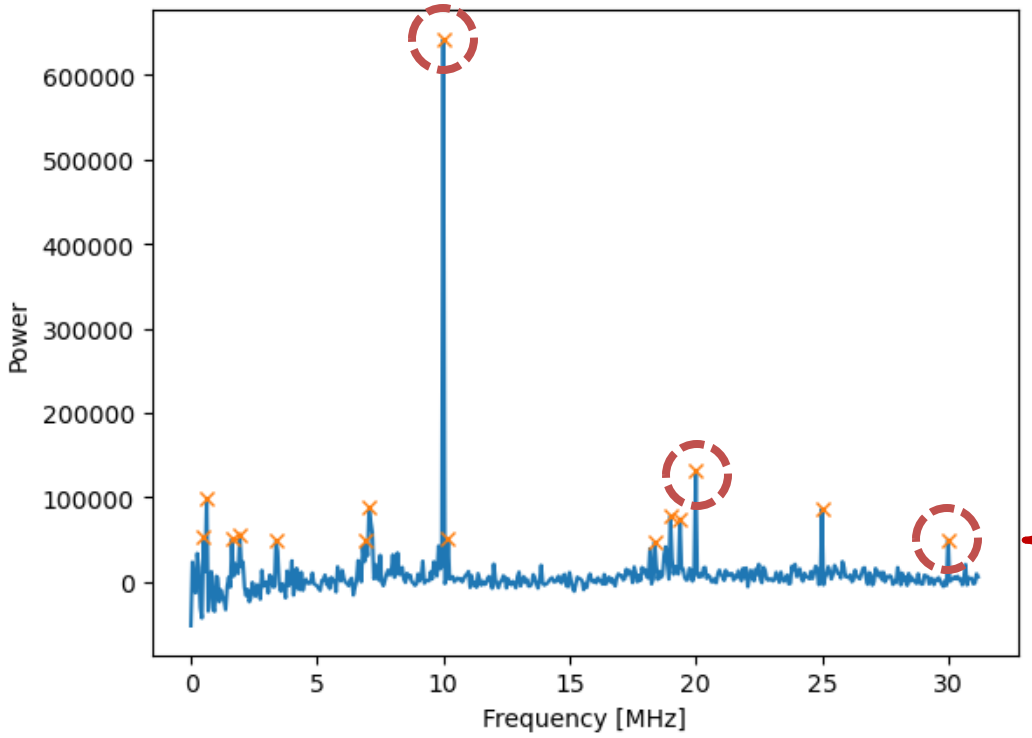
EVENT 1

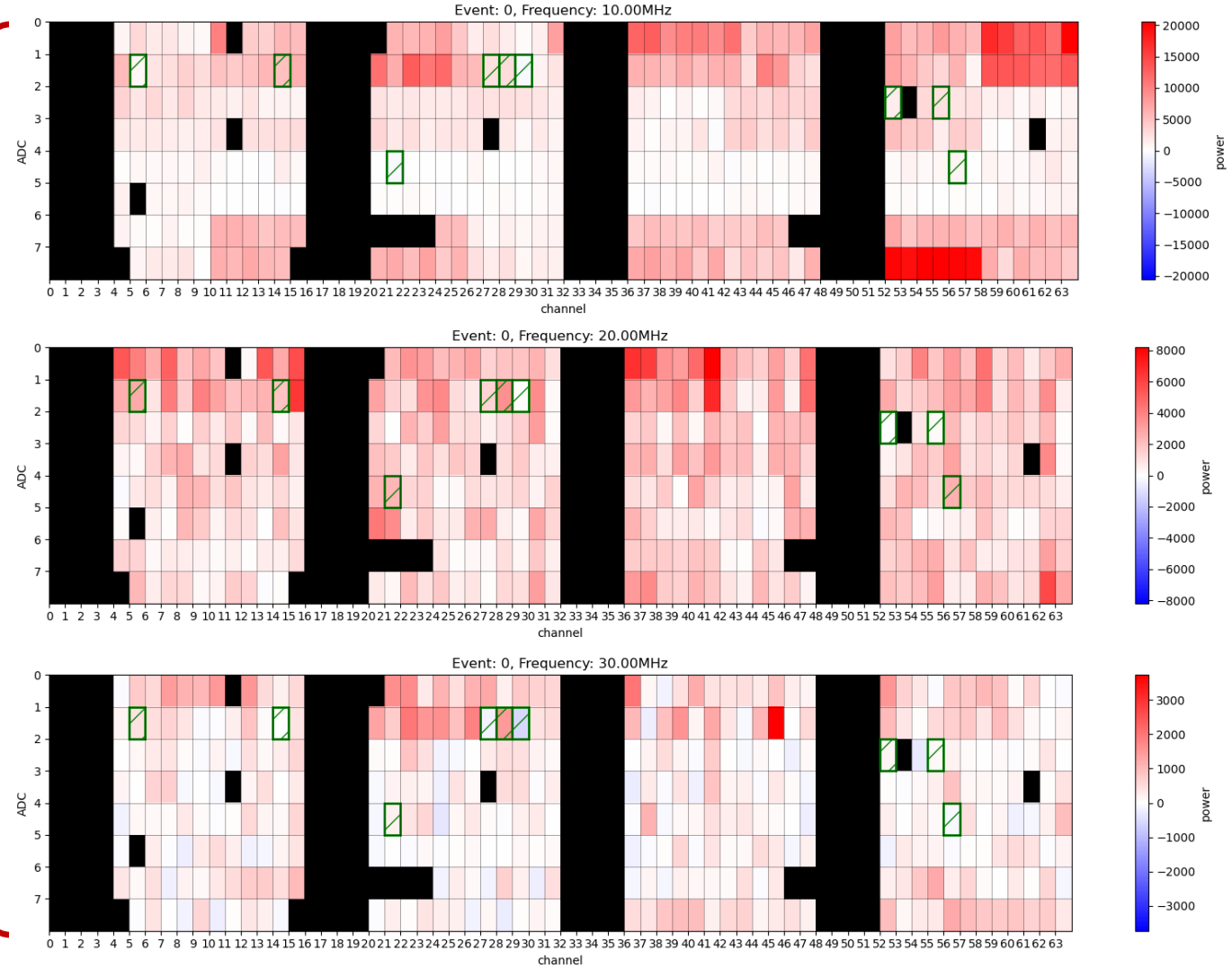






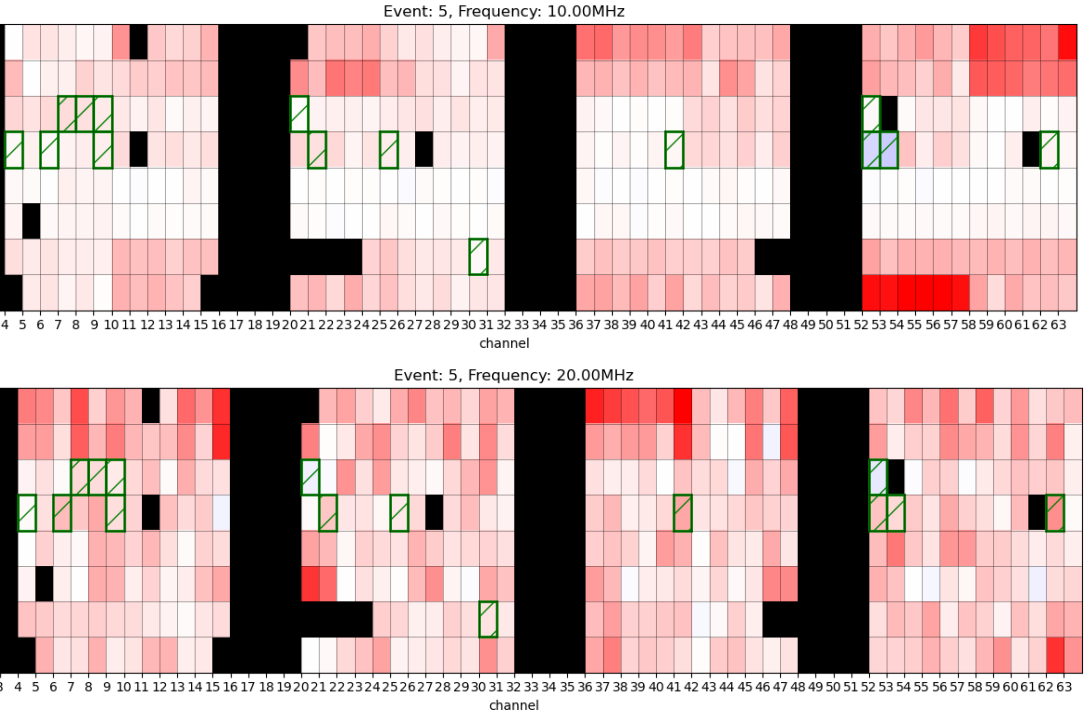
Localising the signals





Localising the signals

EVENT 5

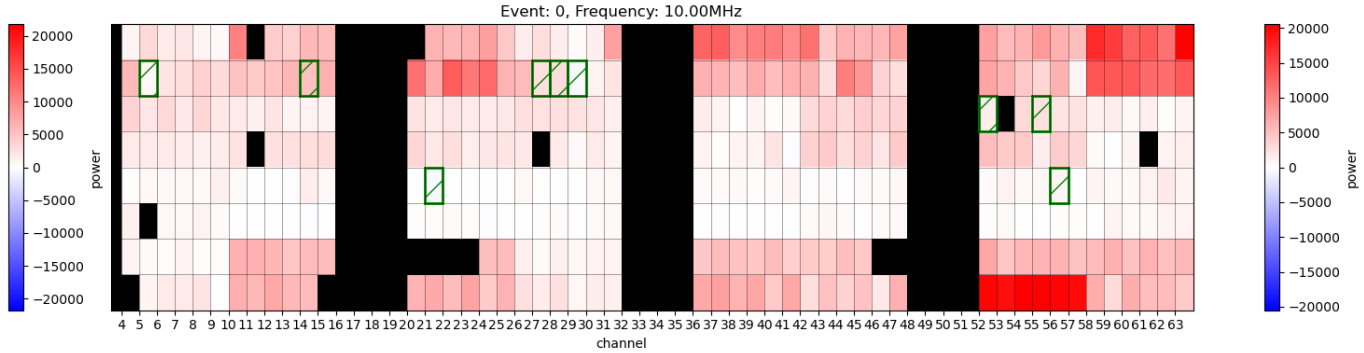


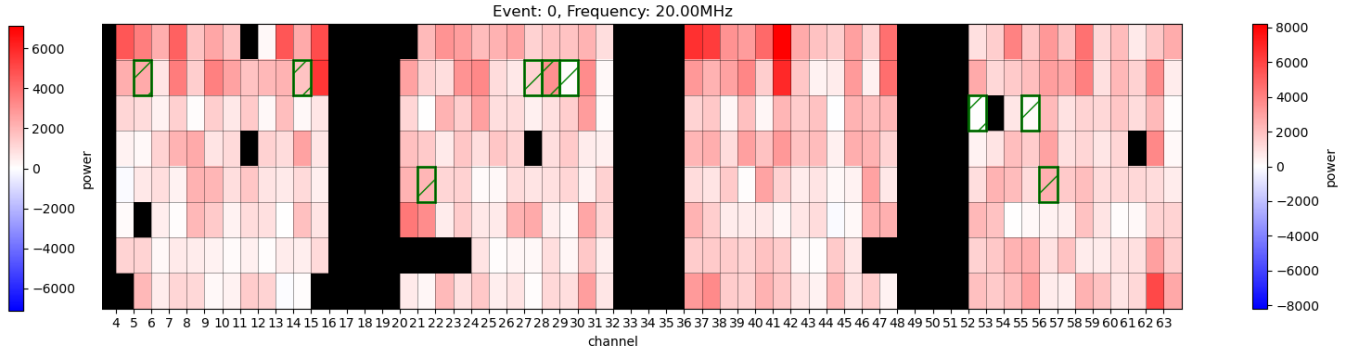
10 MHz
- Appears consistent across events
- No clear correlation with presence of hits

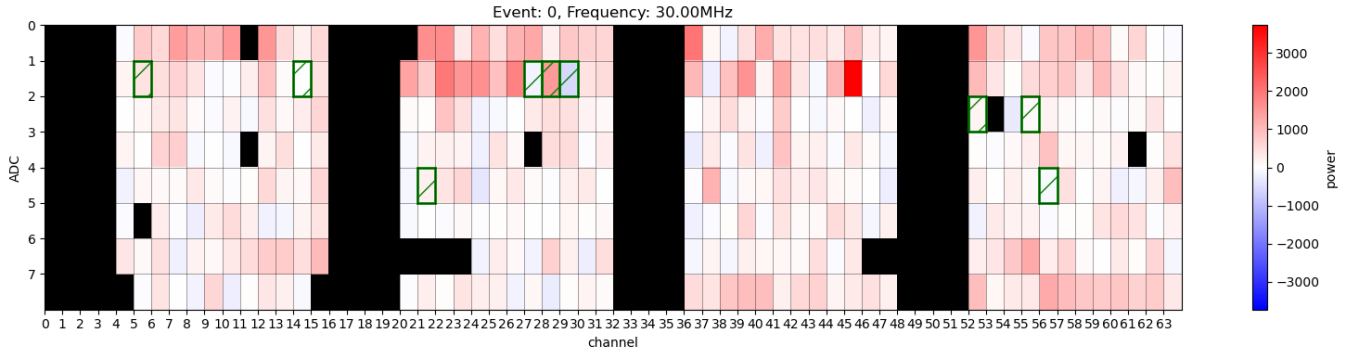
20 MHz
- Some correlation with 10MHz peak
- No clear correlation with presence of hits

30 MHz
- Some correlation with 10MHz peak
- Not tagged as a significant peak in i_evt_lrs=5

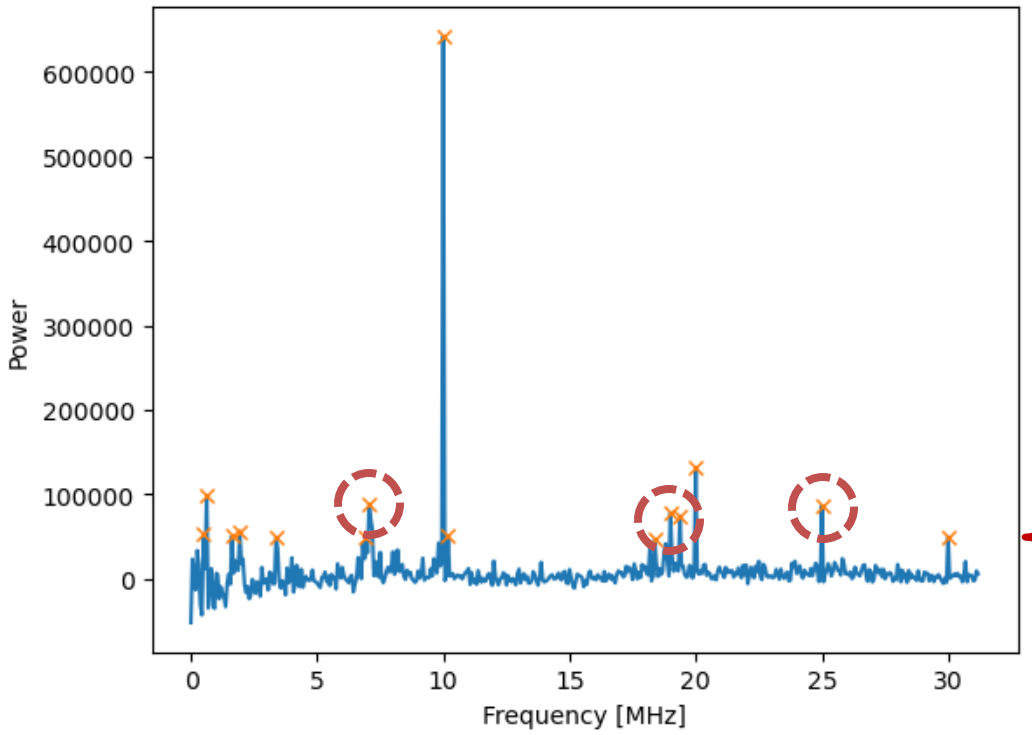
EVENT 1

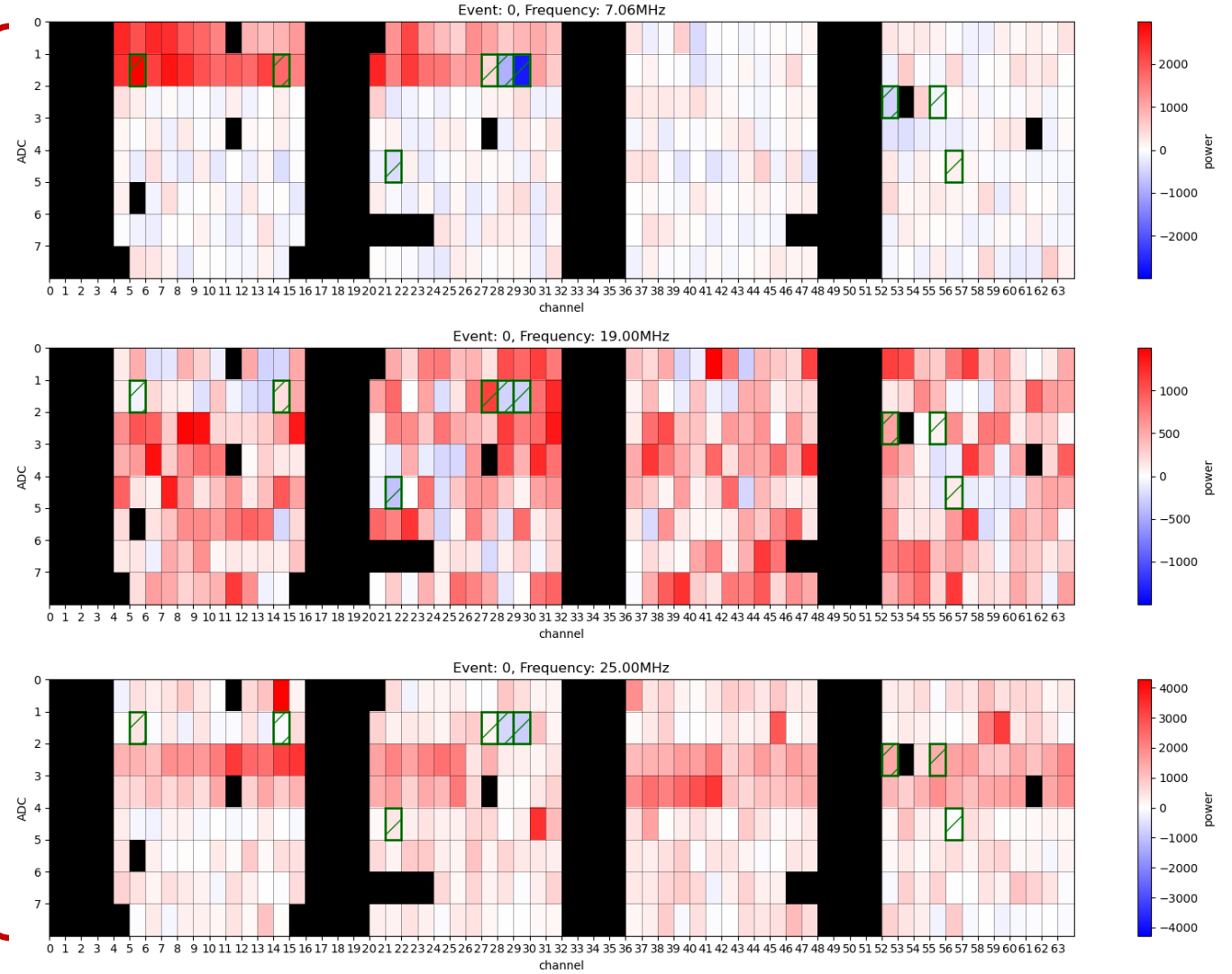






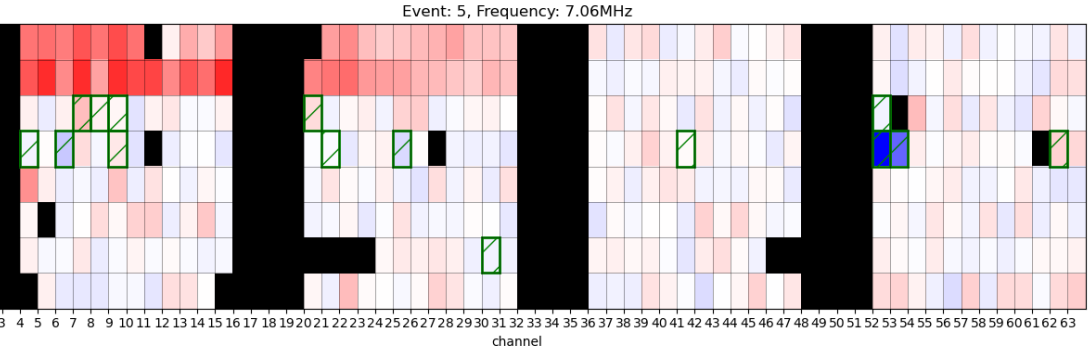
Localising the signals





Localising the signals

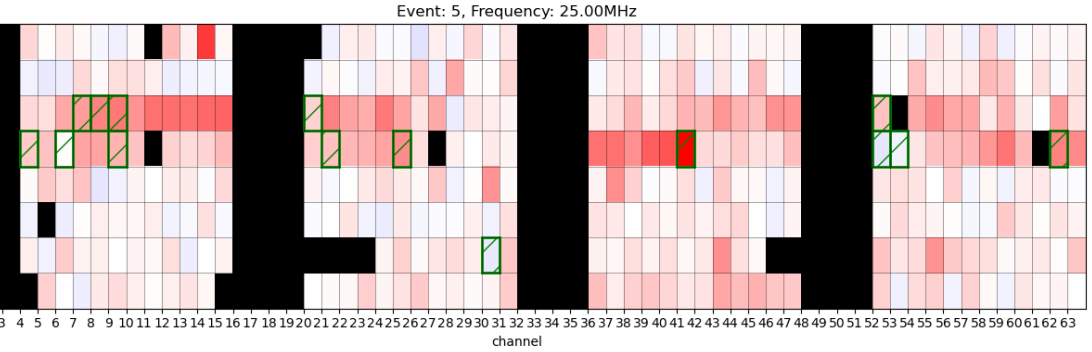
EVENT 5



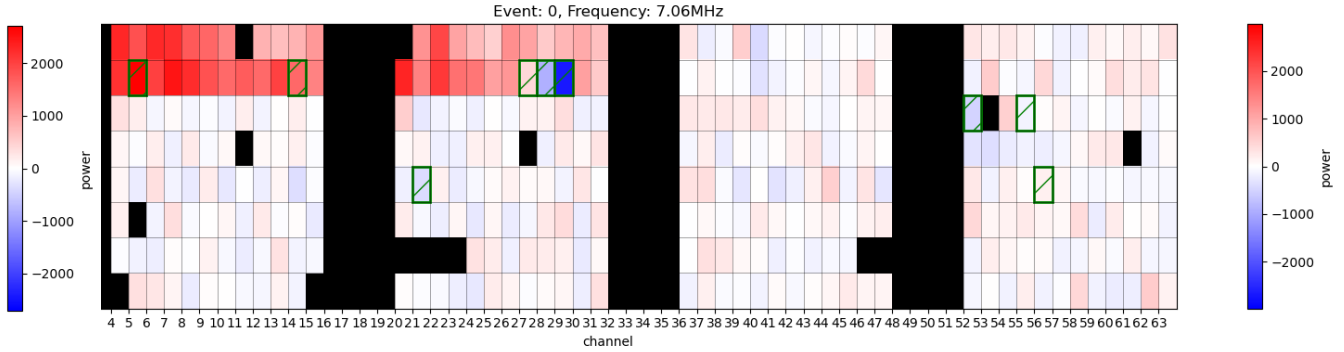
7.06 MHz
- Appears consistent across events
- No clear correlation with presence of hits

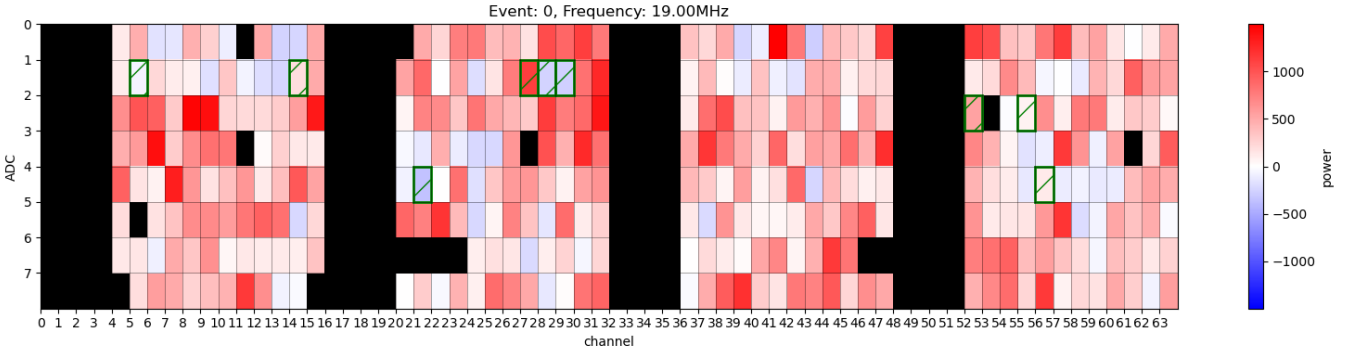
19 MHz
- No clear correlation across detector or with hits
- Not tagged as a significant peak in i_evt_lrs=5

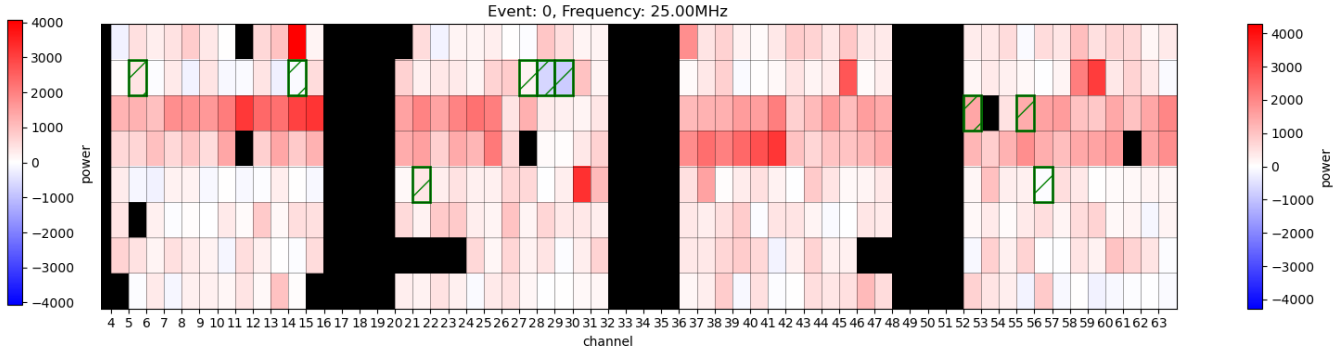
25 MHz
- Appears consistent across events
- Some correlation with presence of hits (check more events!)



EVENT 1

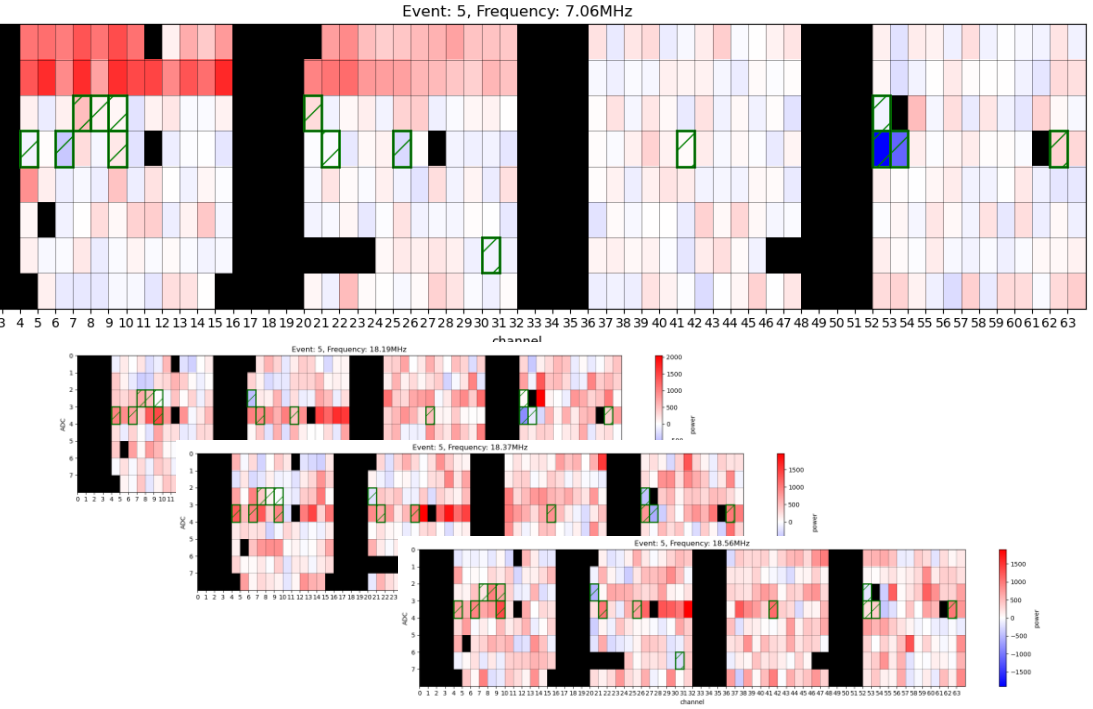


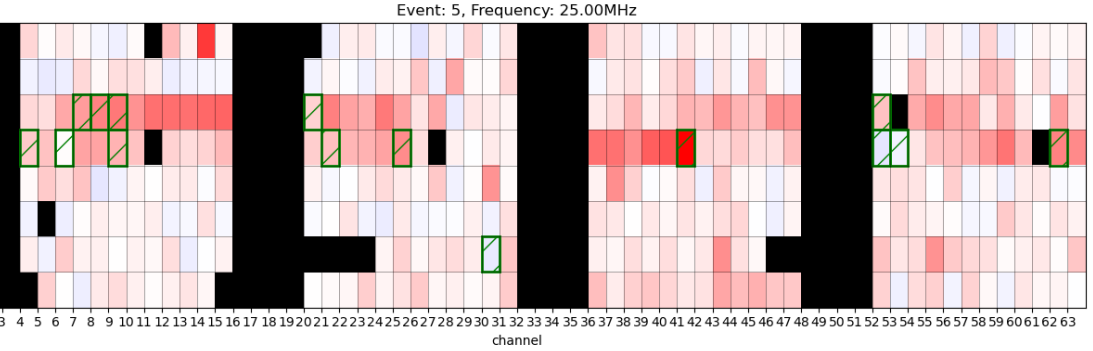




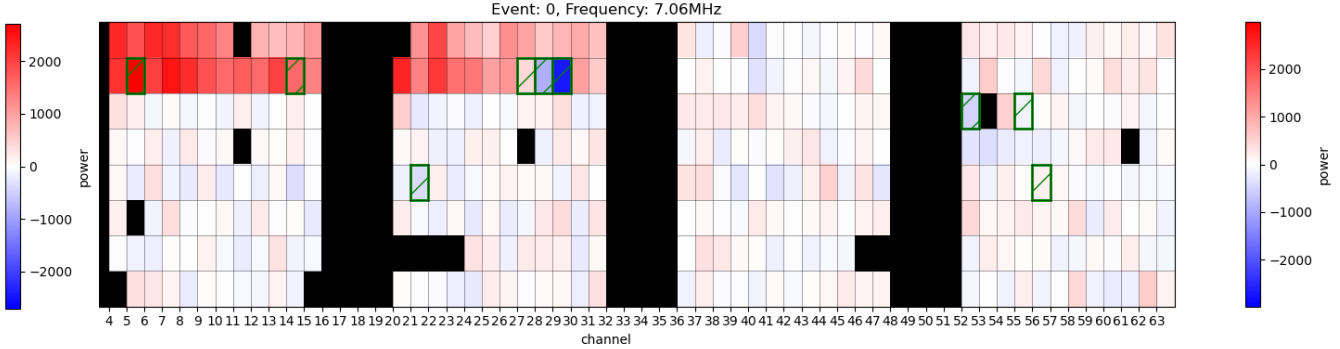
Localising the signals

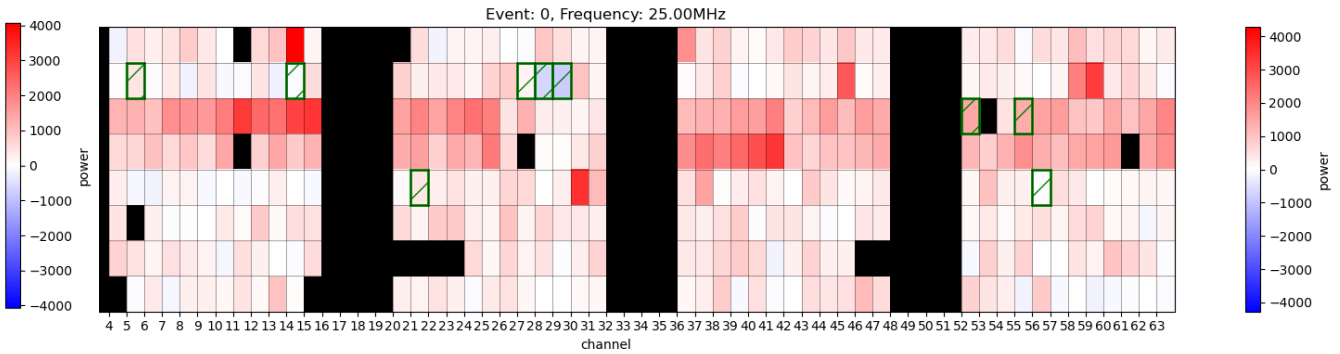
EVENT 5





EVENT 1





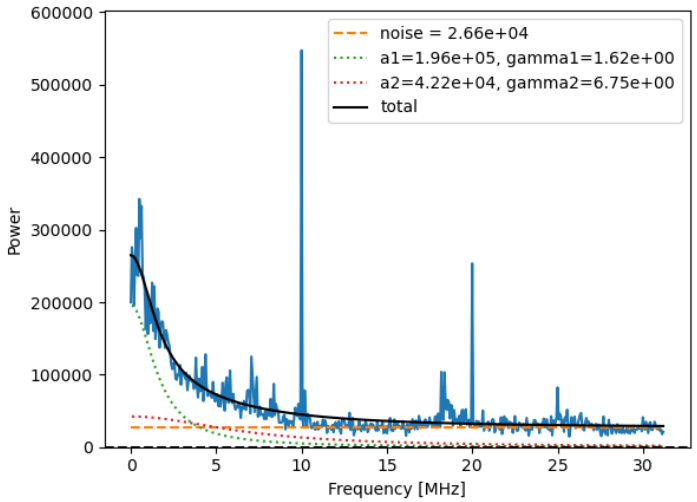
18-19 MHz

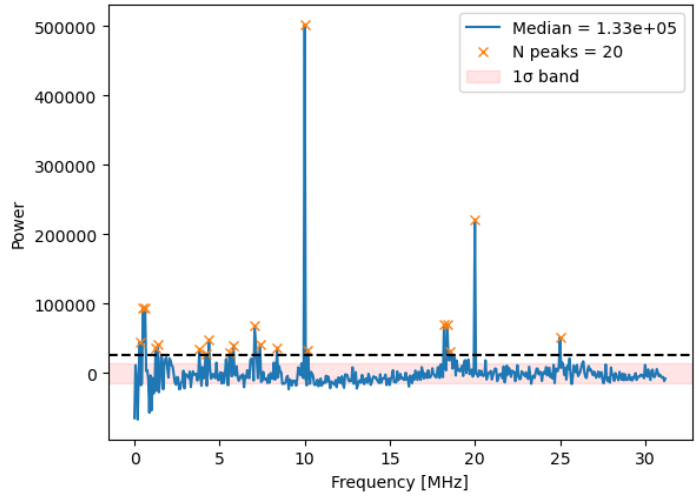
- Cluster of localised peaks
- Some correlation with presence of hits
- Check across more events
- Candidate for cross-talk or just poorly modelled signal?

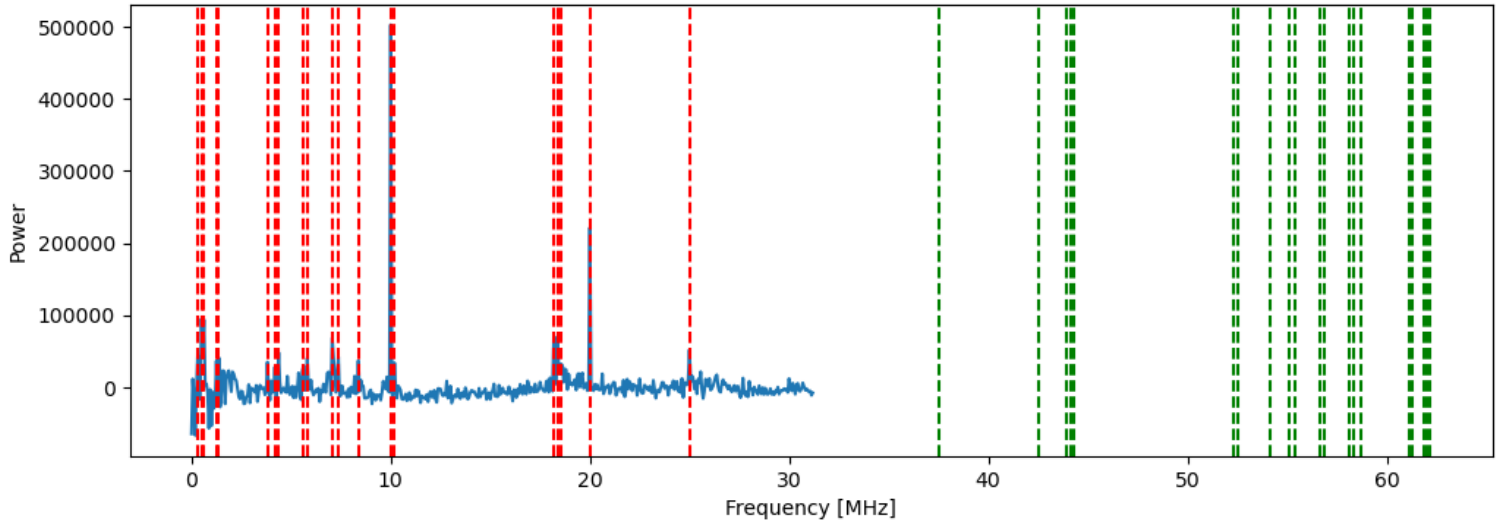
Aliasing: spurious signals

Fitting double Lorentzian + noise floor

(WIP)







Channel map & geom. files

Next step:

Map properly to optically isolated channels i.e. per TPC
https://github.com/DUNE/ndlar_flow/blob/feature_run_on_data/data/proto_nd_flow/light_module_desc-5.0.0.yaml

Conclusion

- ***En-route to:***
  - Isolating spurious signals
    - Overlay TPCs on channel vs ADC to illustrate optical isolation
    - Could these “non-Gaussian noise” sources be affecting the hit finder deconv?
    - If that were the case, they shouldn’t impact MC – so maybe not the culprit
  - Could they be cross-talk?
    - Looks like external sources – i.e. intersystem cross-talk?
    - For interchannel cross-talk, we would need a better test (WIP @ Nikhef)
  - Standardising some tools for diagnostics / analysis
    - numpyfying channel maps – status, TPC, light-trap, (x,y,z) ?
    - bits -> voltage conversion [2+]14bits starting {11}
    - Baseline and noise estimate -> avoiding biases but needs to be fast
    - Channel specific adaptive thresholds -> any meaningful gains to be made?