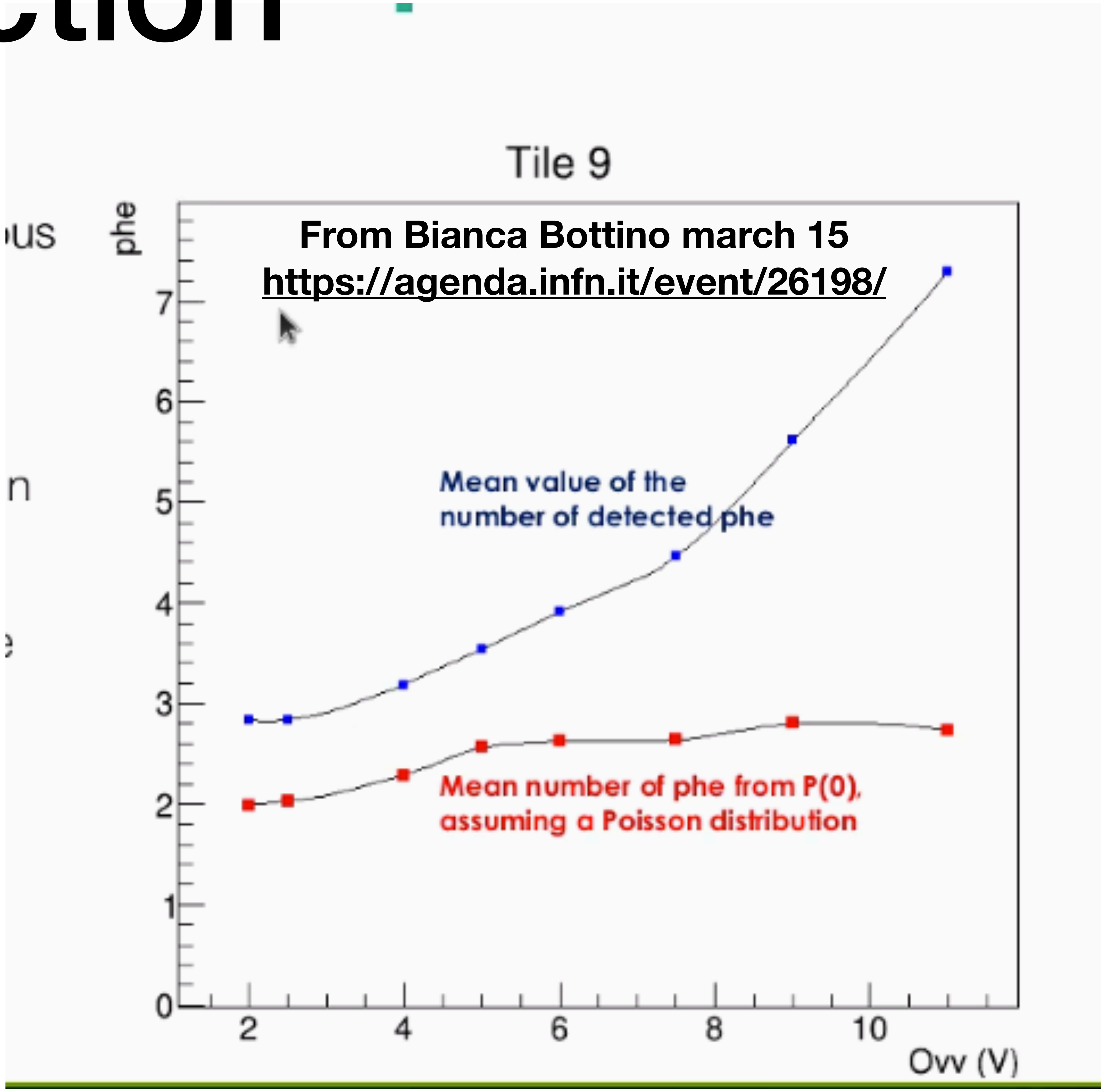
Correlated noise in LFoundry tile 21 with laser liquid nitrogen data

2021-04-20

Giacomo Petrillo
University of Pisa
info@giacomopetrillo.com

Introduction

- We want to model the pe distribution for simulation and VETO studies.
- On an LFoundry tile, with a TPC FEB.
- To disentangle the primary photon distribution from cross talk and afterpulsing.



Data

- http://ds50tb.lngs.infn.it:2180/SiPM/Tiles/LFOUNDRY/pre-production-test/TILE_21/ (collected by Lucia Consiglio using the LNGS cryogenic teststand).

- Sampling frequency 1 GSa/s.

- Each event lasts 15 μs with 9 μs before the laser.

- Three overvoltages: 5.5 V, 7.5 V, 9.5 V (read from file name).

- 200k events (3 seconds) in 10 files, per overvoltage.

- Data collected at the end of 2020.

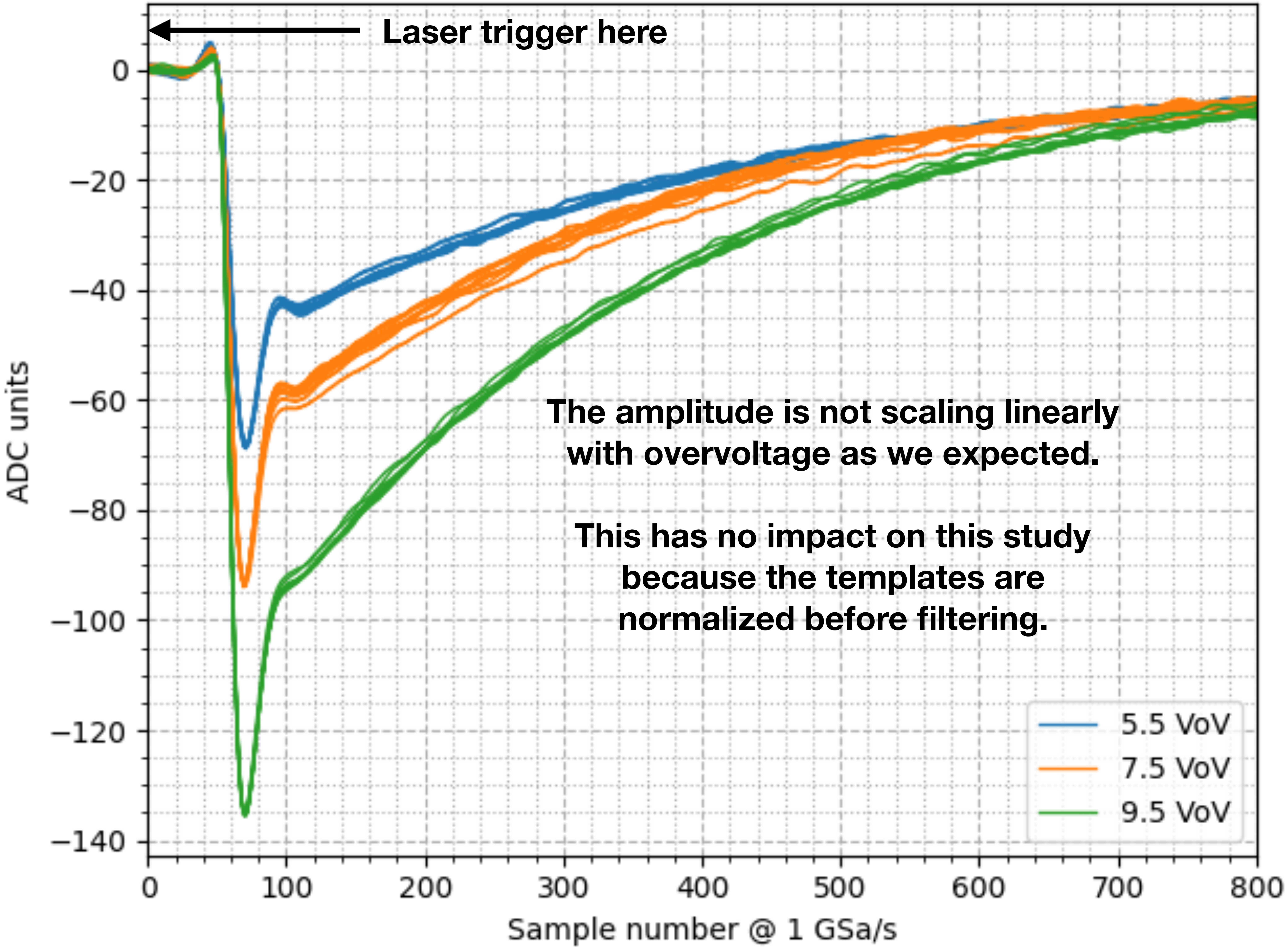
- Caveat: measurement not specifically designed for estimating correlated noises (there are laser reflections).

Filter (1/2)

We use a cross correlation filter.
We do the template this way:

1. Compute the baseline with the pretrigger 8 μs.
2. Select 1 pe signals with a 1.5 μs integration.
3. Average over these waveforms the 3.5 μs segment starting from the laser pulse, to get a first template.
4. Use the first template to filter the 1 pe signals and align them, then average again.

We do the template separately for each of the 10 files.

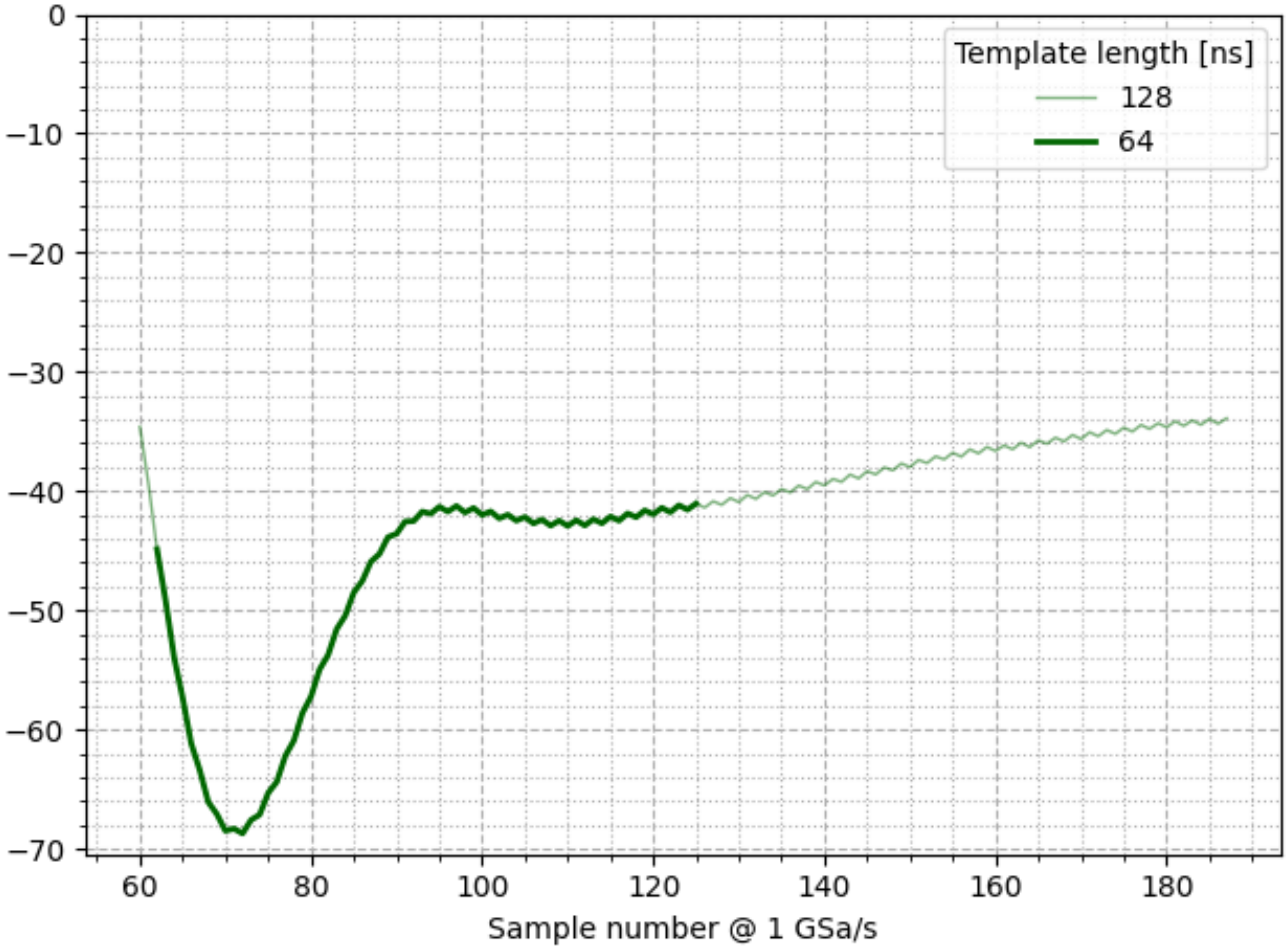


Filter (2/2)

We will use filters with different lengths.

We truncate the template by keeping the fixed length subtemplate that maximizes the sum of squares of the samples.

We will use the shortest possible filters to separate well close peaks (128 ns @ 5.5 VoV, 64 ns others).



Peak finding (1/3)

After filtering, to find the laser pulse:

1. Take a window of ±30 samples around the expected signal position (obtained from the trigger).

2. Take the minimum in the window, excluding minima occurring at the edges.

Peak finding (2/3)

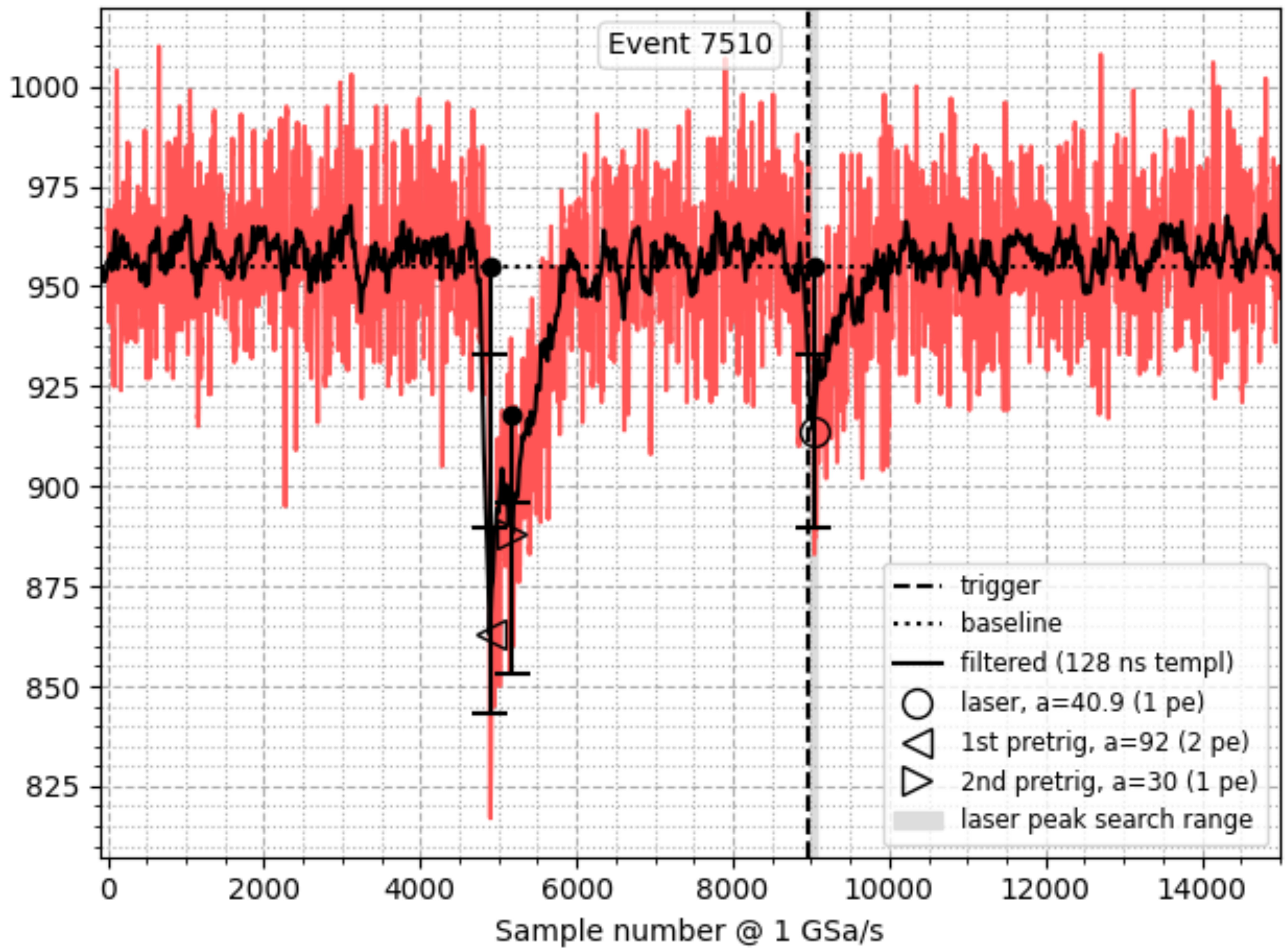
For peaks other than the laser one we use a prominence-based peak finder.

- We use the filtered waveform.
- We divide the event in pre- and post-laser regions and search separately.
- We keep the two highest-prominence peaks in each region.
- It is not necessary to require a minimum distance between peaks because the filtered waveform is smooth (on the scale of the filter length).

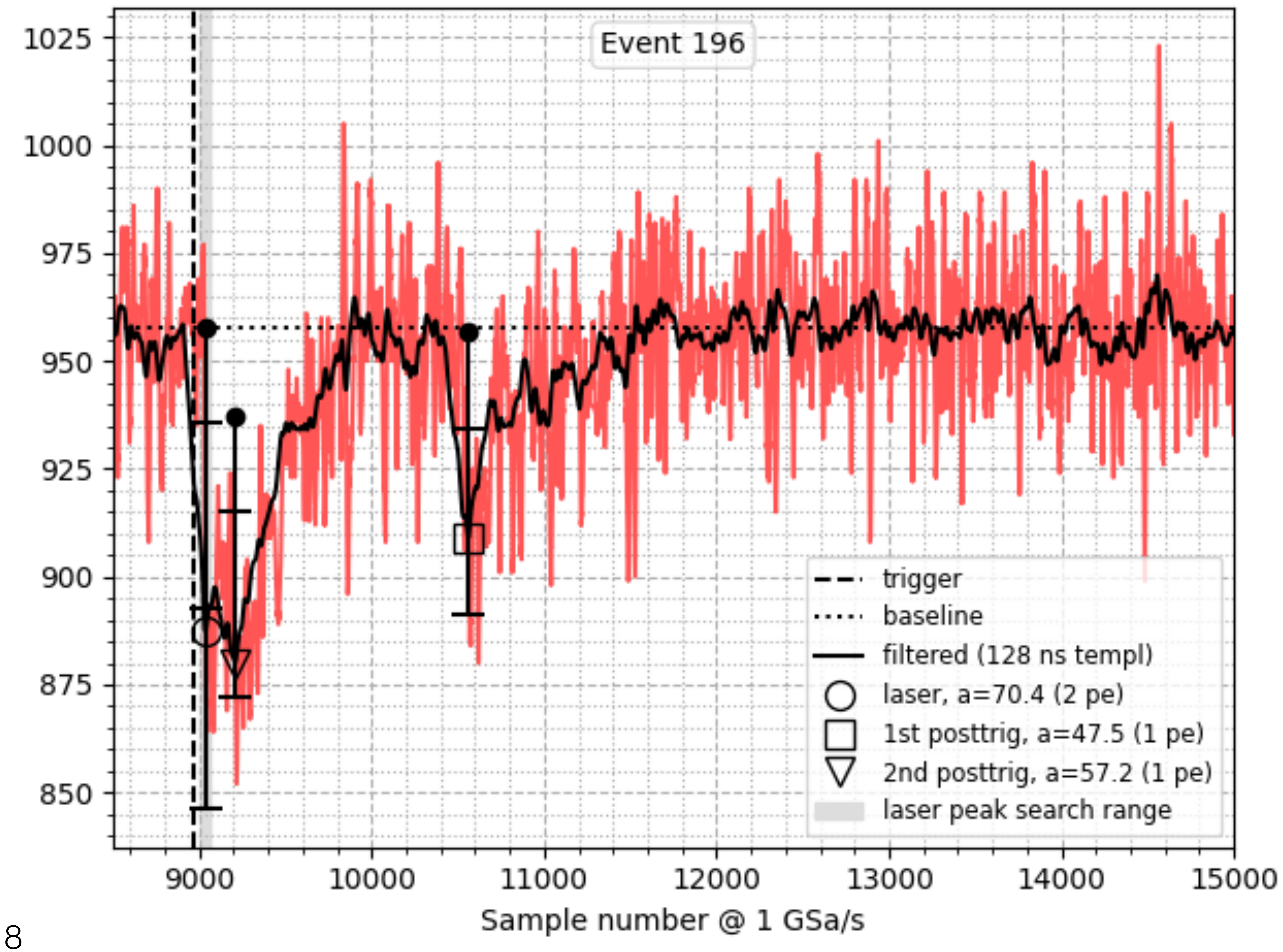
Peak finding (3/3)

Examples.

Pre-trigger



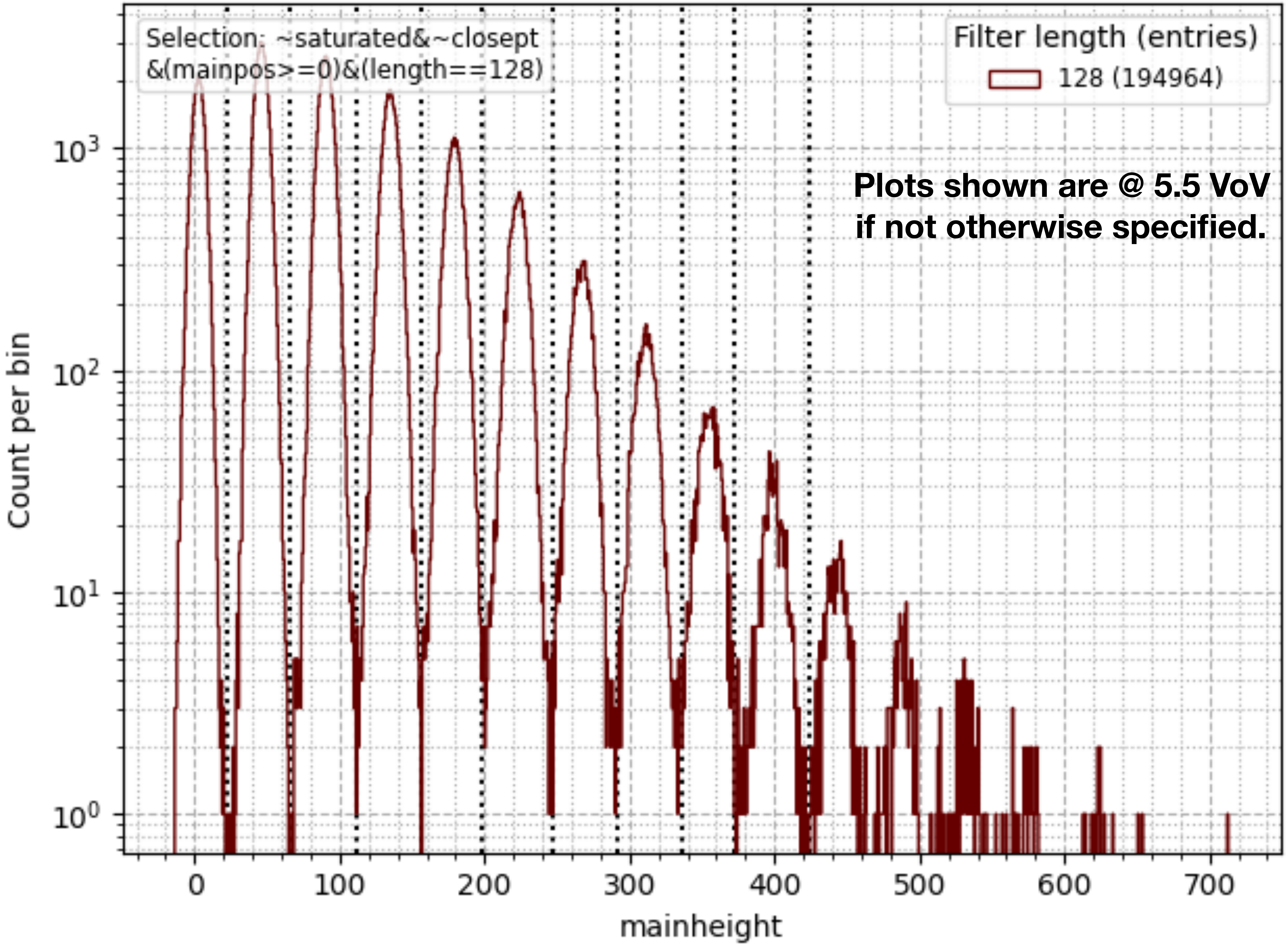
Post-trigger



Converting height to pe

We select events where the laser peak height should be reliable (isolated, not saturated).

We histogram the laser peak height, detect peaks in the histogram, and place boundaries at the midpoint between the most distant consecutive samples between each peak.



Amplitude for close peaks

Peaks close to each other influence each other's peak height in the filtered waveform, we want to correct this.

We make this approximation: the position of the peak is not changed by the presence of other peaks (because the signal has a sharp peak).

To make it fast, instead of fitting the waveform, we use only the filtered peak heights. (For an untruncated noise-corrected matched filter this would be already optimal, it isn't for our truncated cross correlation filter).

If y_i is the filtered signal waveform (computed from the template), t_j and h_j the position and height of peak j in the filter output, a_j the unknown amplitude of signal j, we have the linear system

$$h_k = \sum_j y_{t_k - t_j} a_j$$

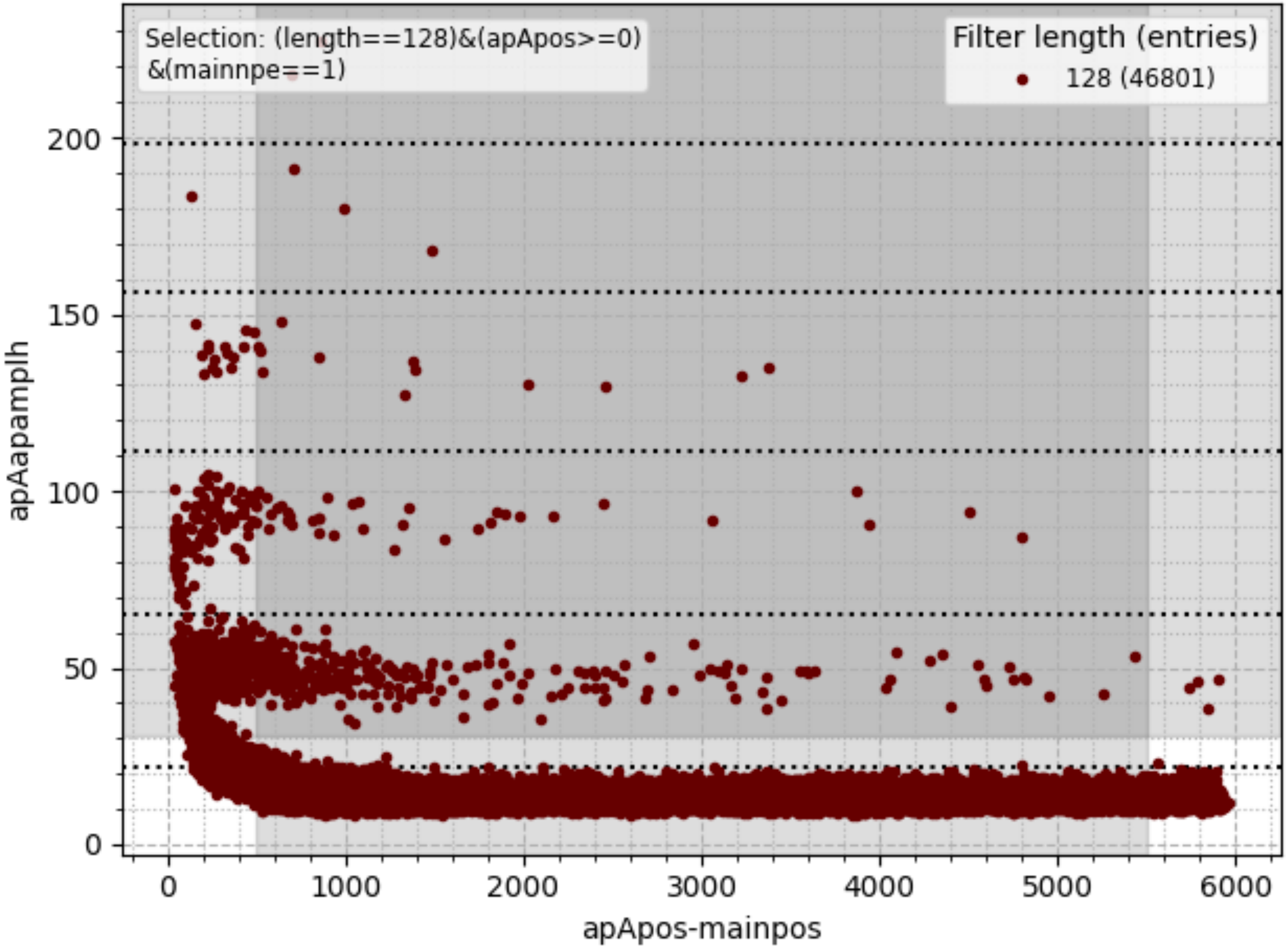
which we solve to find the amplitudes a_j. The y is normalized to have peak height 1 such that a has the same units of h, and shifted such that $y_0 = 1$ is the peak.

Afterpulses (1/3)

Scatter plot amplitude vs. delay from laser peak of post-trigger pulses for 1 pe laser peaks.

Afterpulses close to the laser are too small to be detected. Also, there could be peak finding artifacts.

So we take a window from 500 ns to 5500 ns after the laser peak, then choose the cut on the amplitude.

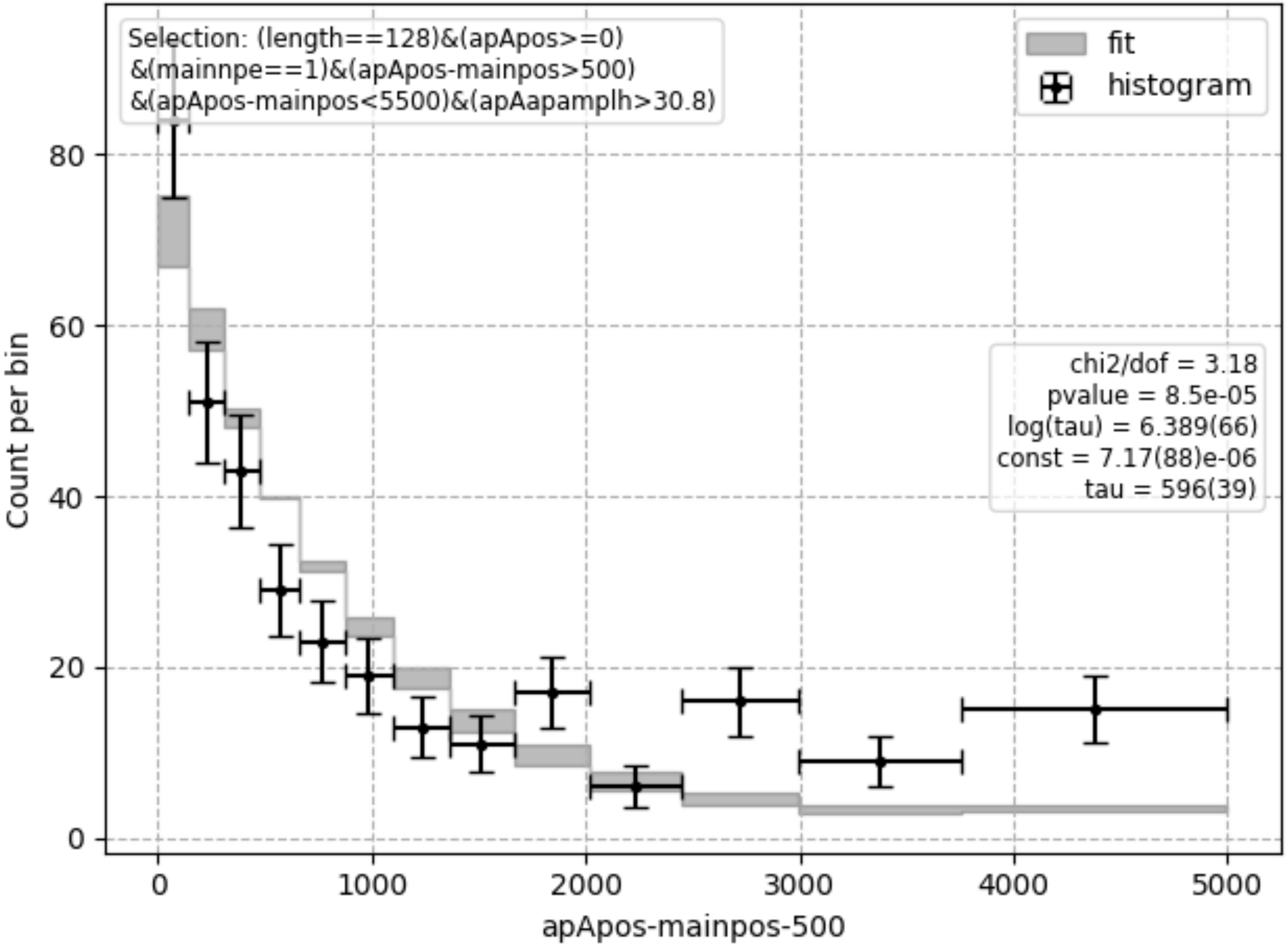


Afterpulses (2/3)

We fit with exponential + uniform where the uniform has a datapoint from the random pulses rate (measured separately, not shown).

The fit is not very good. For other overvoltages the shape of the residuals (data – fit) is similar.

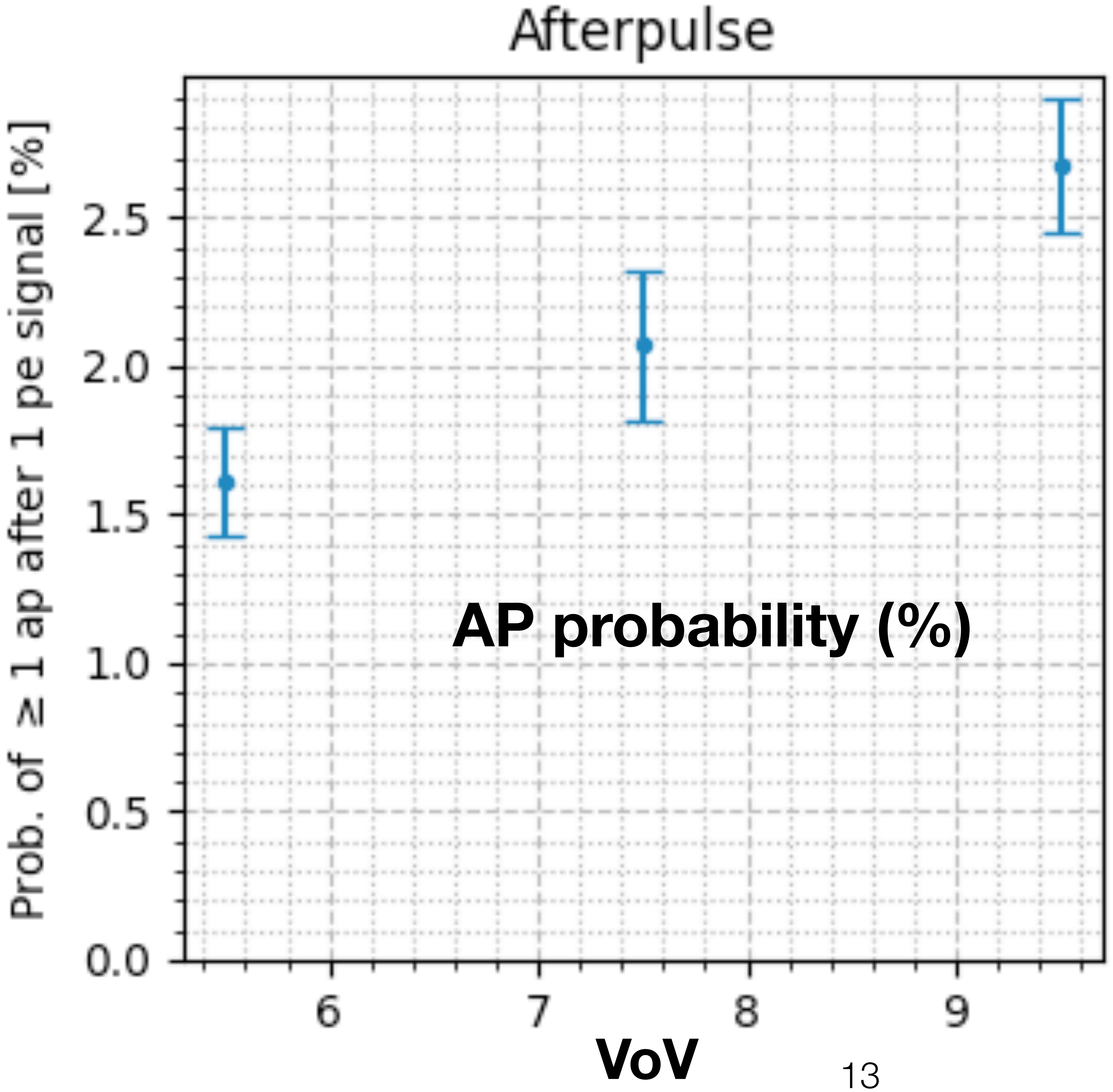
We rescale the tau error with sqrt(chi2/dof) to add a systematic.

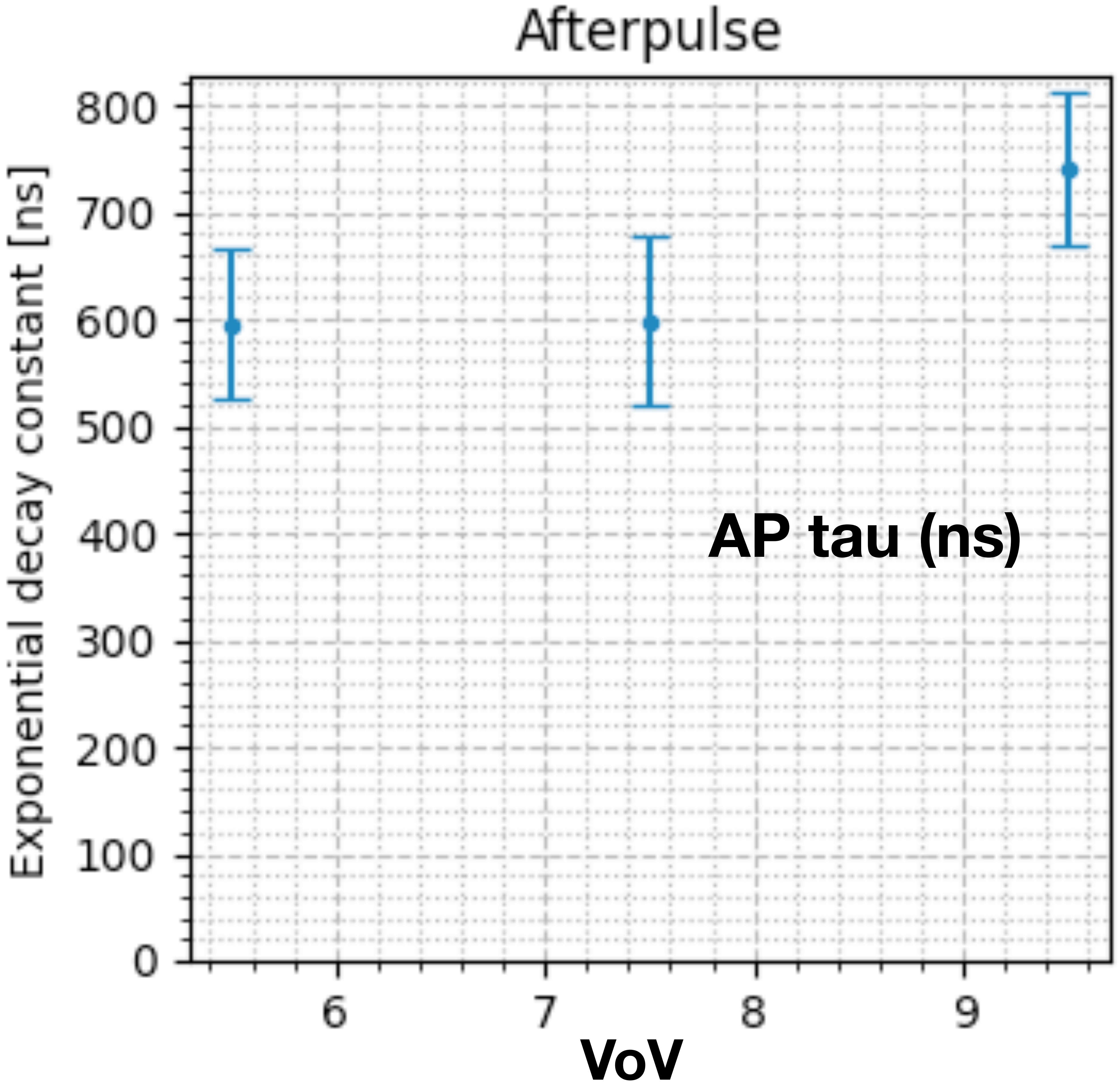


Afterpulses (3/3)

Correcting the afterpulse count for the temporal cut with the fitted model (exponential+background), we obtain this.

Although the model is not very good, if a simulation was run with these parameters, it would reproduce the correct afterpulse count with delay above 500 ns.





Direct cross talk model

Poisson branching process: each pulse generates a poisson count of child pulses with mean μ_B. The total number of pulses (root + descendants) is Borel distributed:

$$P(n; \mu_B) = \exp(-\mu_B n)\frac{(\mu_B n)^{n-1}}{n!}$$

DCT for afterpulses and random pulses

If the initial number of pulses is poisson-distributed with mean μ_P, the total with cross talk is:

$$P(n; \mu_P, \mu_B) = \exp\big(-(\mu_P + n\mu_B)\big)\frac{\mu_P(\mu_P + n\mu_B)^{n-1}}{n!}$$

DCT for laser

References: arXiv 1109.2014, 1609.01181.

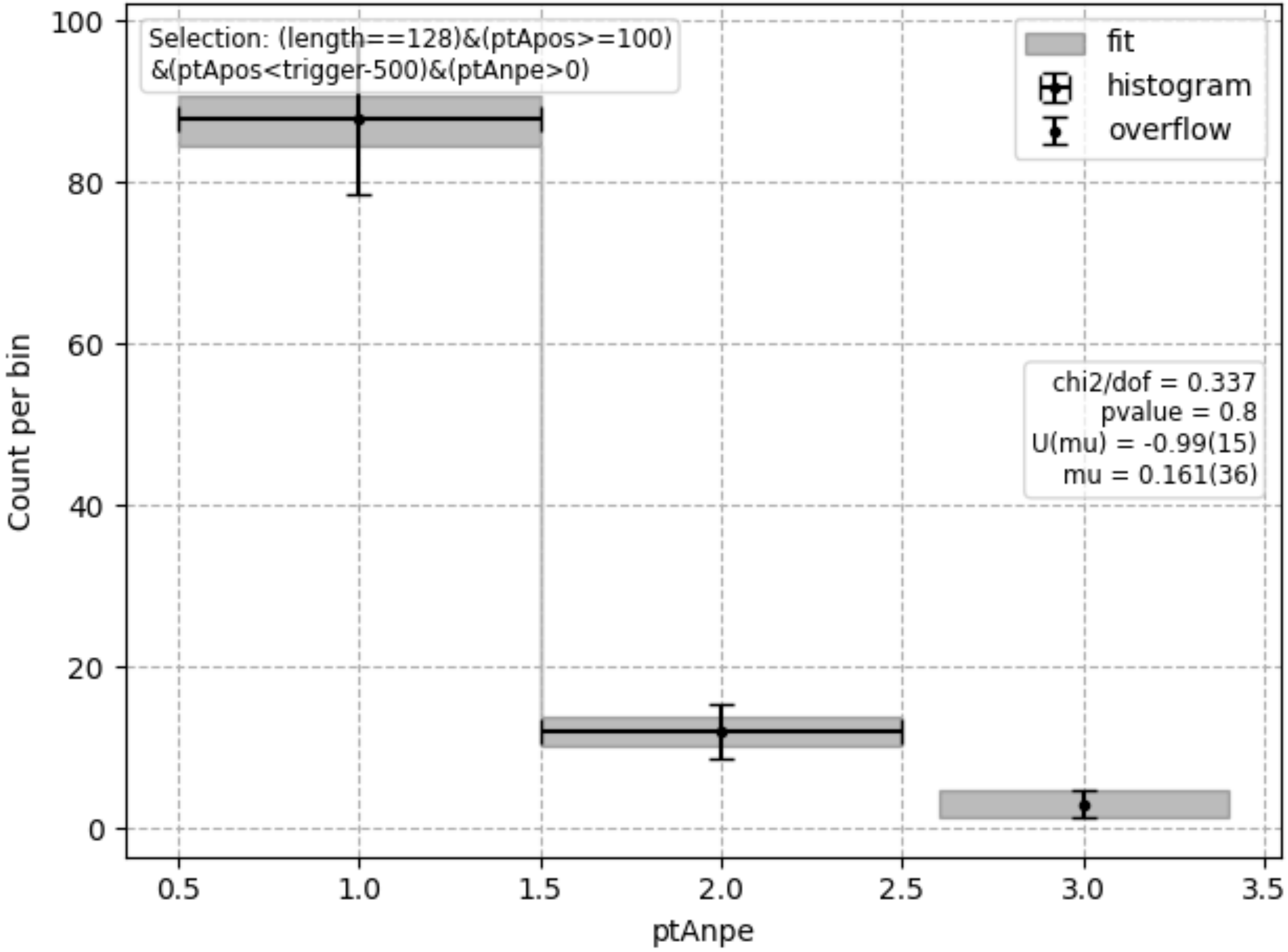
Cross talk in pre-trigger pulses

We count the pe of appropriately selected pre-trigger pulses, using the pe boundaries.

We fit with the Borel distribution.

The last bin is an overflow bin, fitted with the integral to infinity.

These fits work well at all overvoltages.



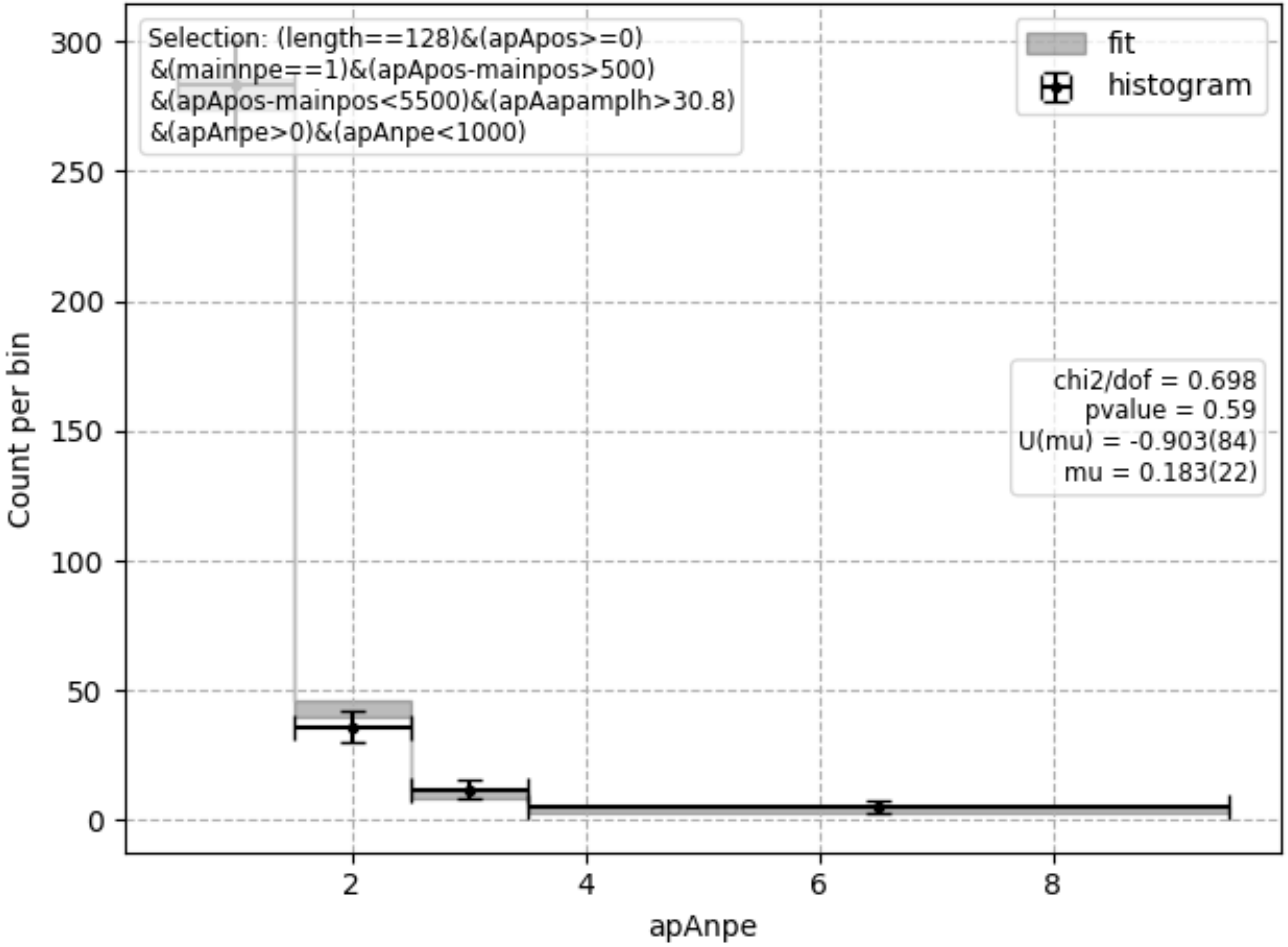
Cross talk in afterpulses

We redo it similarly for laser afterpulses.

This time we don't fit the overflow because it gives problems at higher overvoltages.

Still need to investigate if this is due to wrong model or problems in the analysis.

Without overflow, all fine.



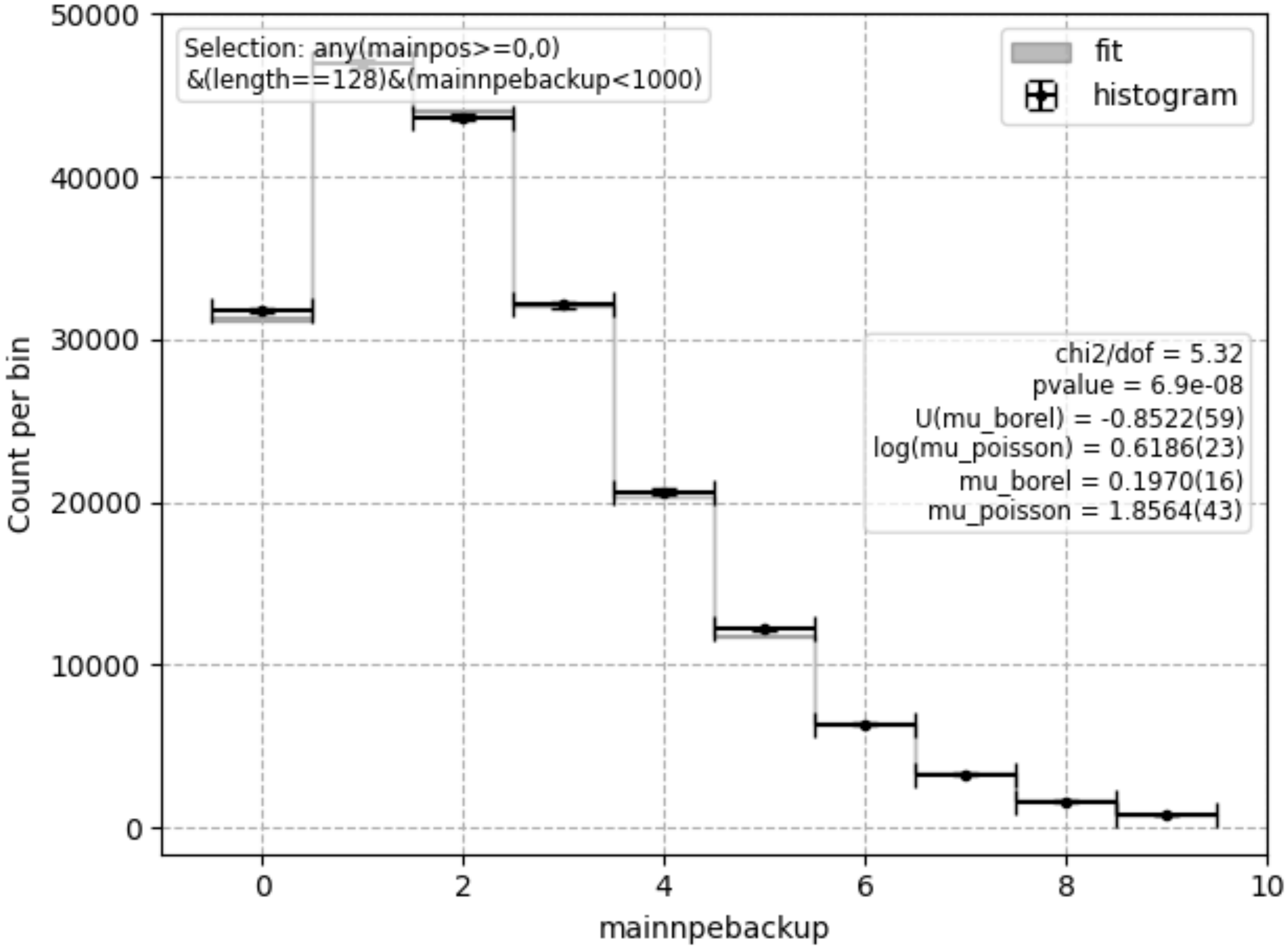
Cross talk in laser pulses

We fit using the Poisson+Borel distribution.

We don't fit the overflow as for afterpulses.

The chisquares are always bad but here the sample size is very high, the agreement is satisfactory.

We add systematics multiplying the error by sqrt(chi2/dof).

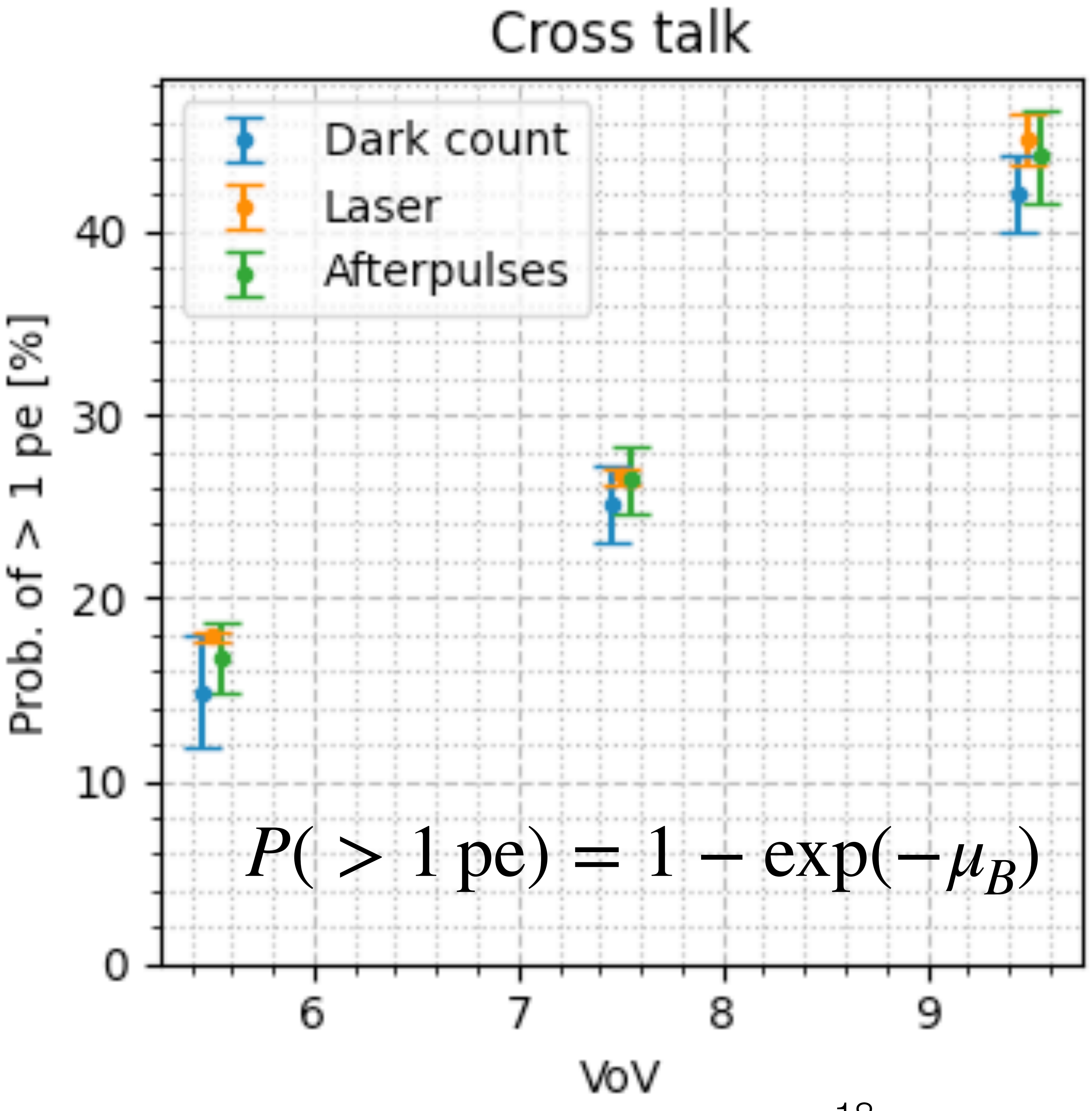


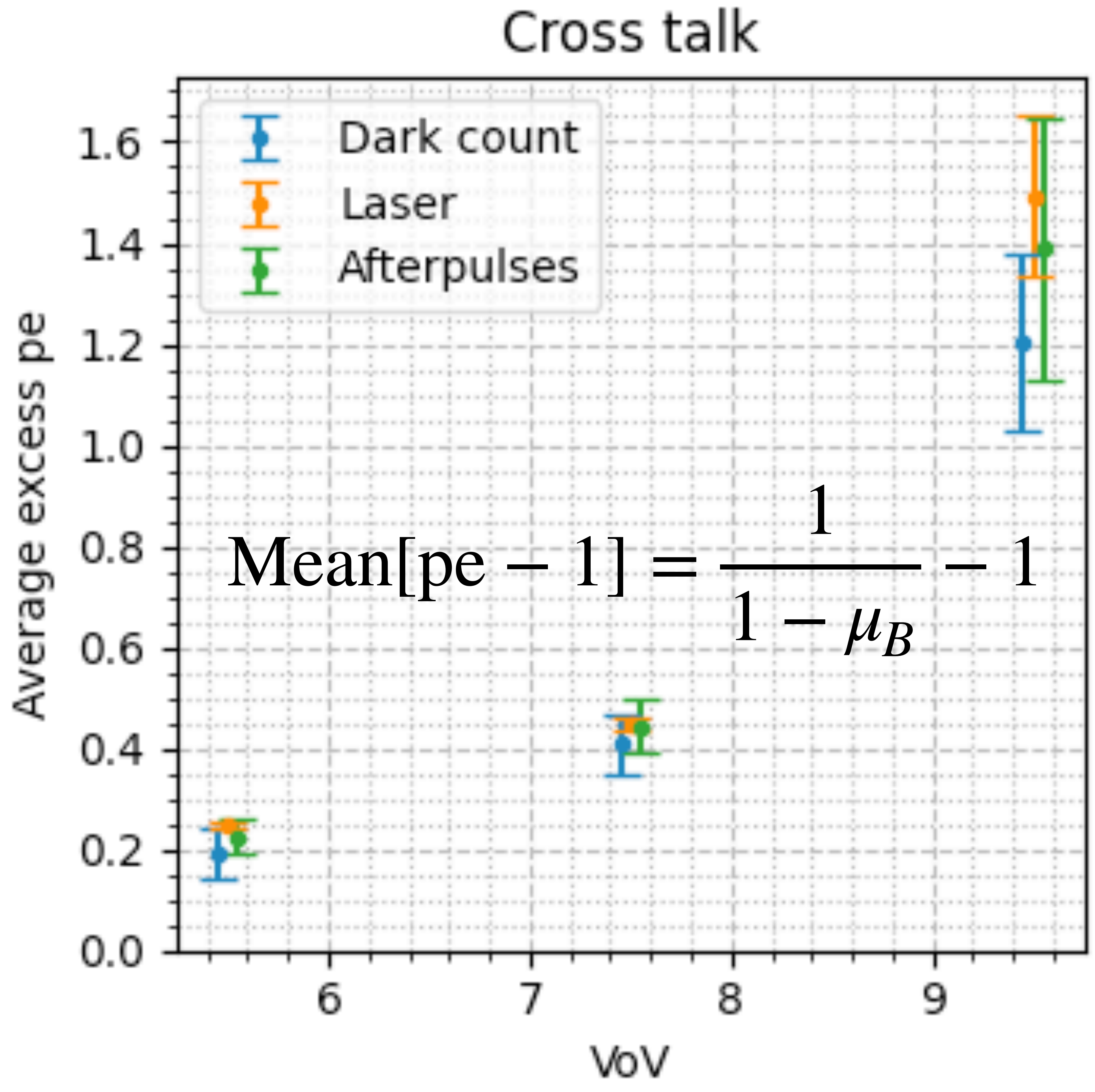
Cross talk results (1/2)

We express the amount of cross talk in two different ways: A. the probability of having cross talk, and B. the average number of excess pe.

Both computed from the fitted model.

The different measurements agree.



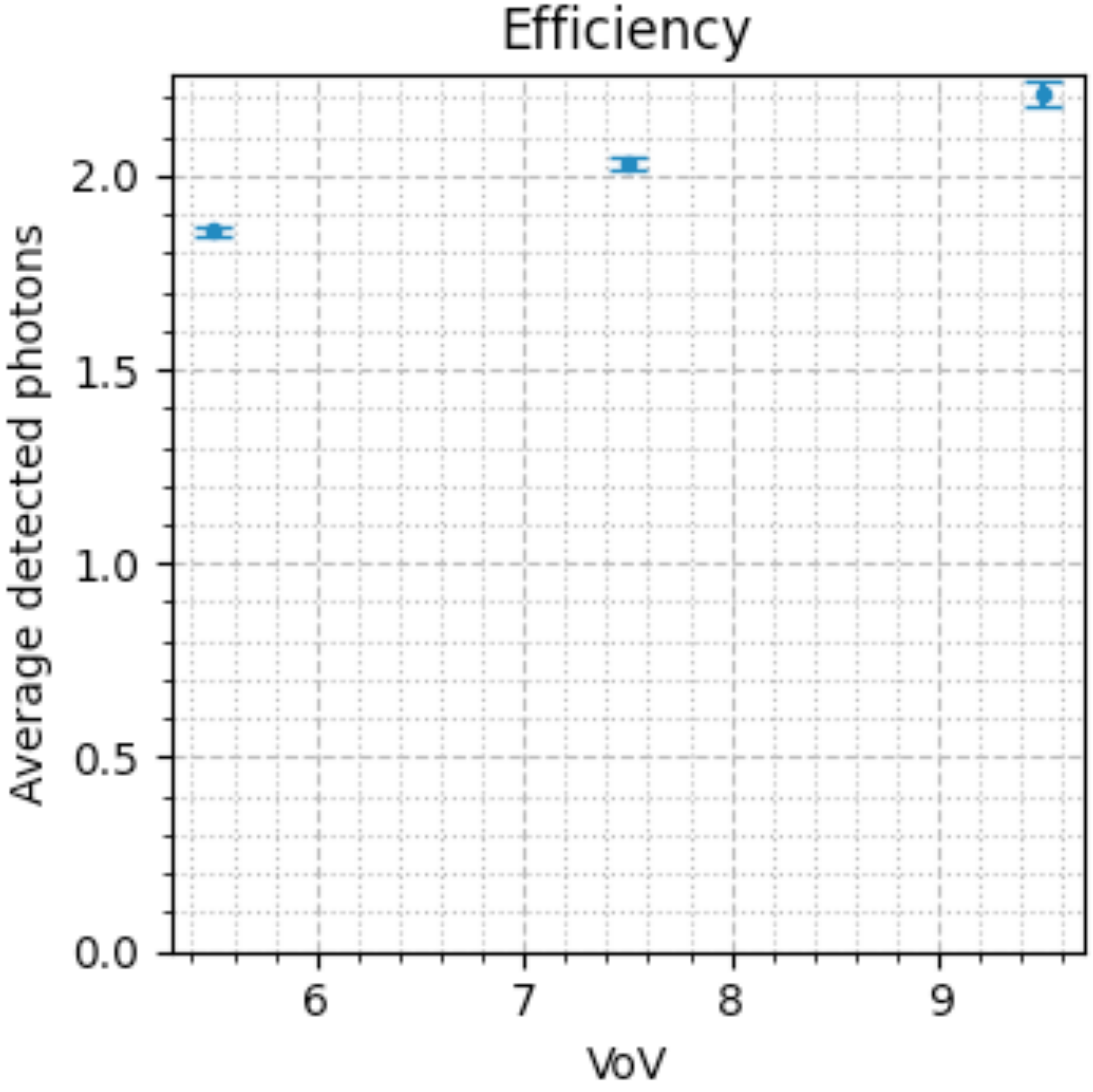


Cross talk results (2/2)

We check the other parameter in the model: the mean laser pe before cross talk μ_P.

Expected to be proportional to PDE, if laser is stable and illumination conditions constant.

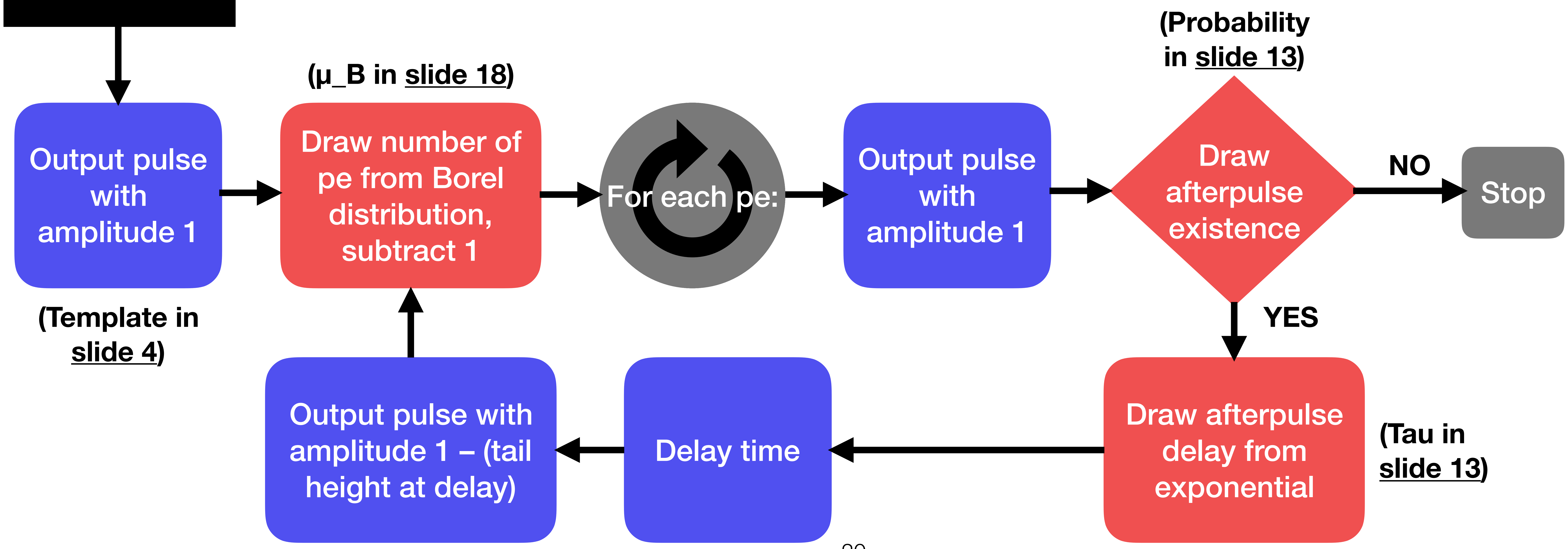
Increasing, as expected.



Mapping to simulation

We express the model we used in the analysis as the equivalent simulation algorithm. In other words, the analysis is coherent with this simulation.

Note that it is not very physically accurate, there are evident approximations.



Conclusions

We have analyzed correlated noises using cross correlation filter and peak finding (instead of charge integration methods).

The Poisson branching process model for the direct cross-talk seems good, as reported in the literature.

The laser pe distribution seems to be Poissonian in LNGS.

This model can be used to interpret VETO data.

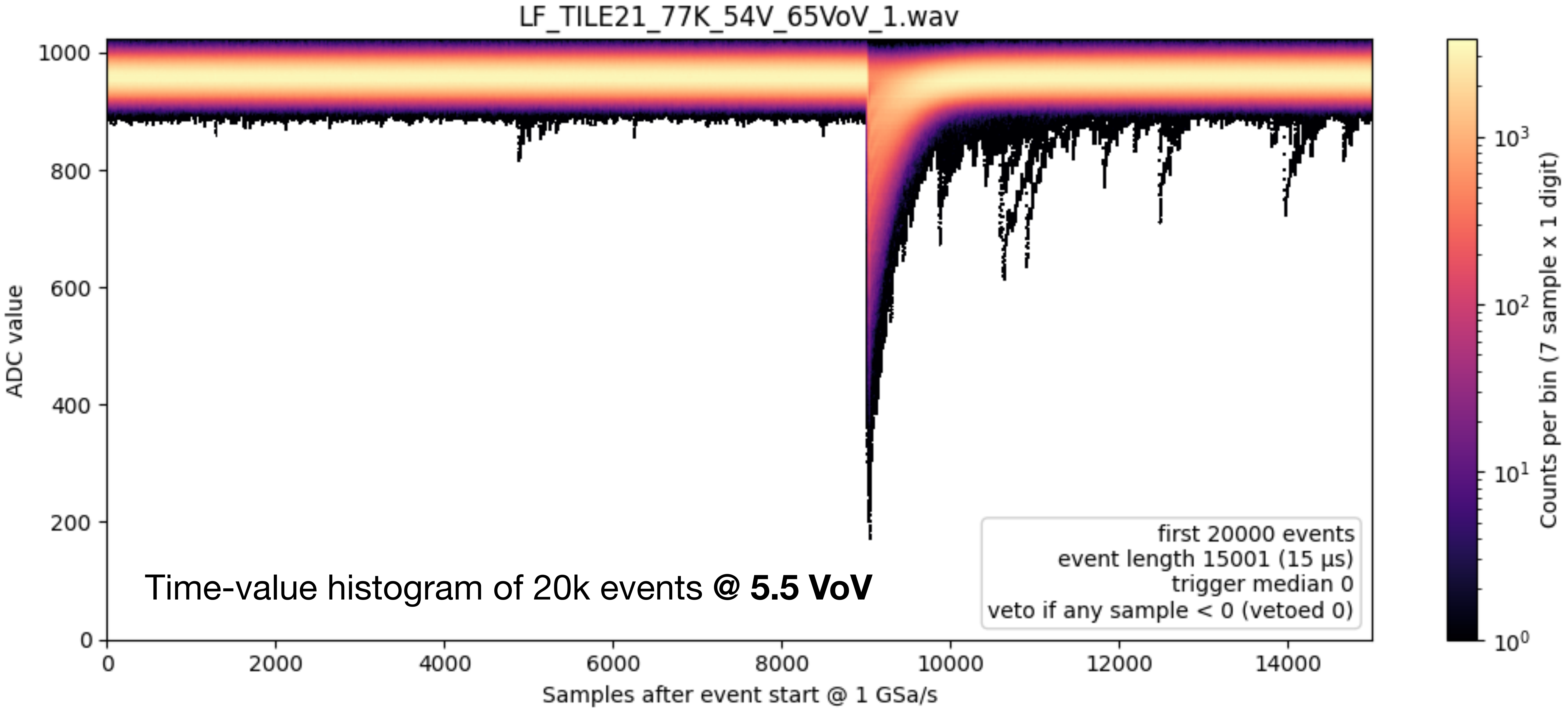
The afterpulse decay distribution appears different from the expected exponential. Fitting the exponential, tau = (650 ± 100) ns.

If it is not a bug in the analysis, maybe there are two populations of afterpulses with two taus, due to different energy levels?

Or different afterpulse tau for each SiPM in the PDM?

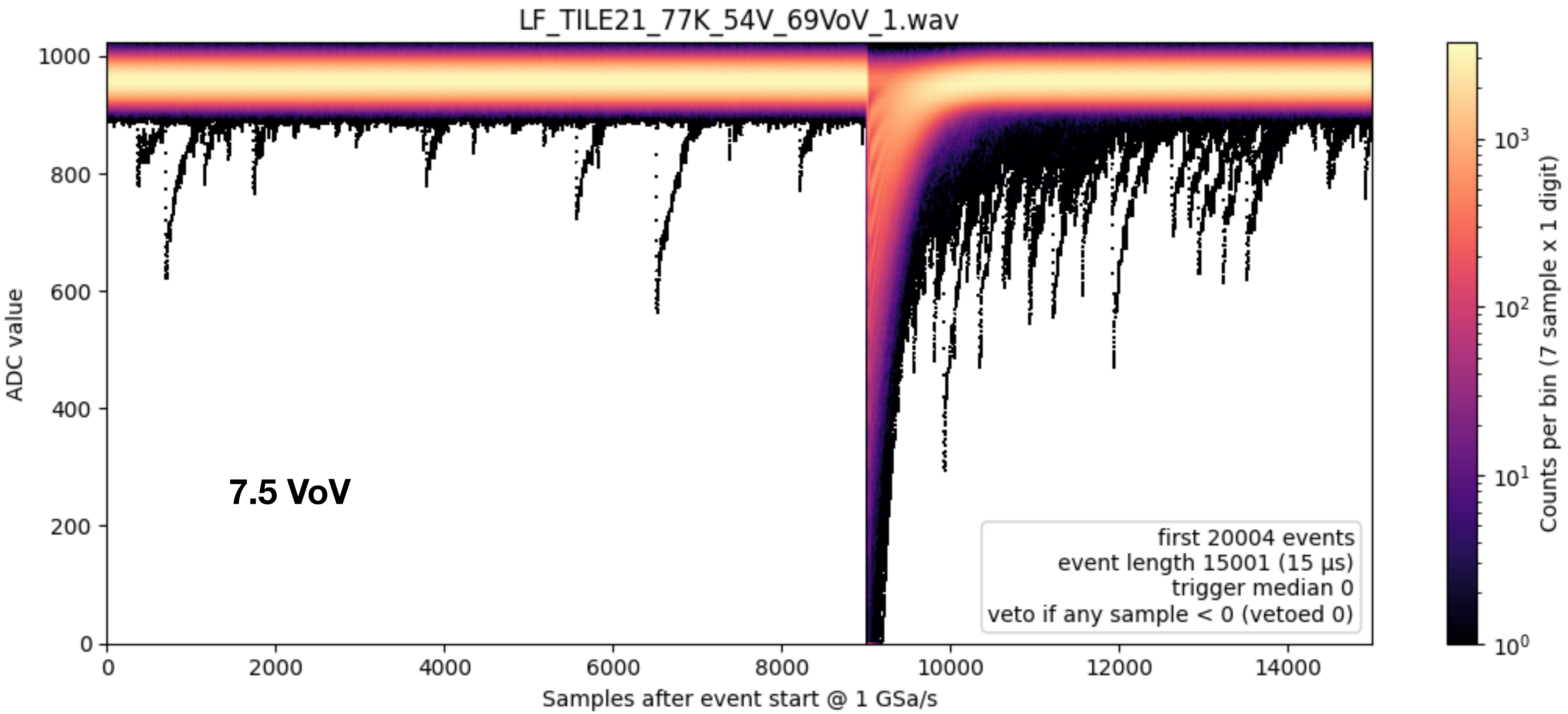
Backup slides

Data (1/3)



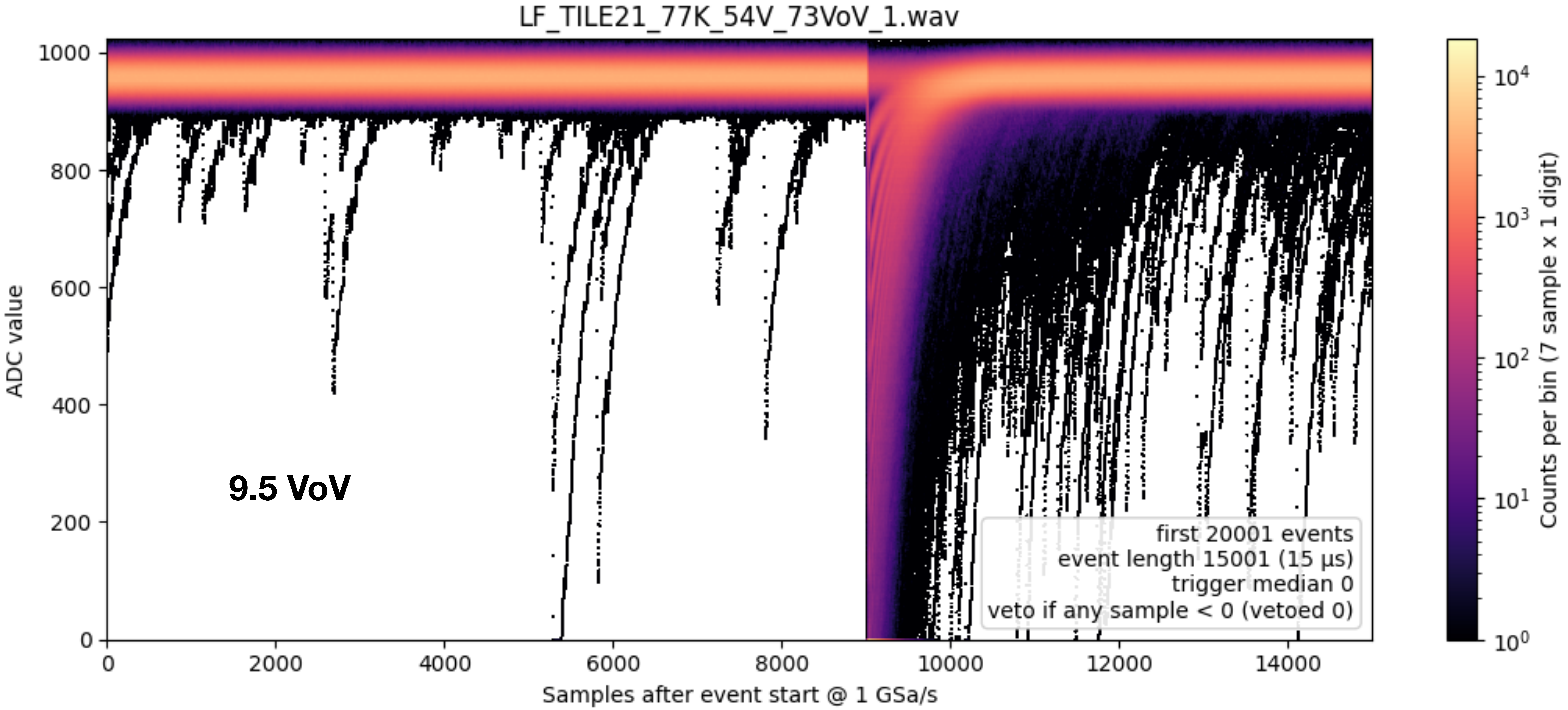
Time-value histogram of 20k events @ **5.5 VoV**

Data (2/3)



LF_TILE21_77K_54V_69VoV_1.wav
1000
800
600
400
200
0
ADC value
0
2000
4000
6000
8000
10000
12000
14000
Samples after event start @ 1 GSa/s
7.5 VoV
first 20004 events
event length 15001 (15 µs)
trigger median 0
veto if any sample < 0 (vetoed 0)
10^3
10^2
10^1
10^0
Counts per bin (7 sample x 1 digit)

Data (3/3)

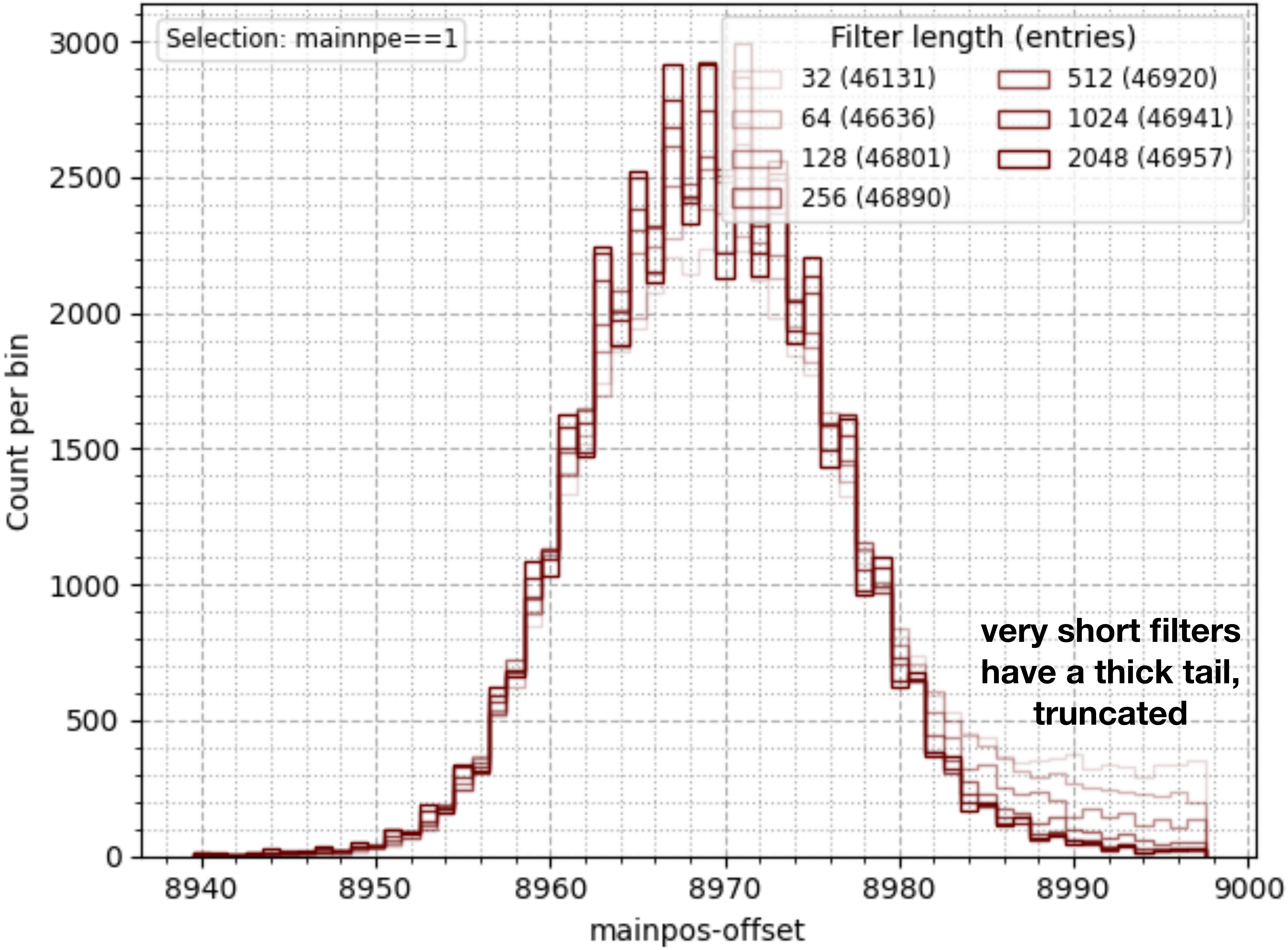


LF_TILE21_77K_54V_73VoV_1.wav
ADC value
Samples after event start @ 1 GSa/s
Counts per bin (7 sample x 1 digit)
9.5 VoV
first 20001 events
event length 15001 (15 µs)
trigger median 0
veto if any sample < 0 (vetoed 0)

Laser peak finding (1/4)

Distribution of the laser peak position for 1 pe pulses.

The distribution is skewed to the right and **truncated**. We have to **make sure this is not due to afterpulsing**, because it would bias the afterpulse count.



Laser peak finding (2/4)

I found the tail is due to random fluctuation, not afterpulses.

- It appears with short filters because they have lower SNR.

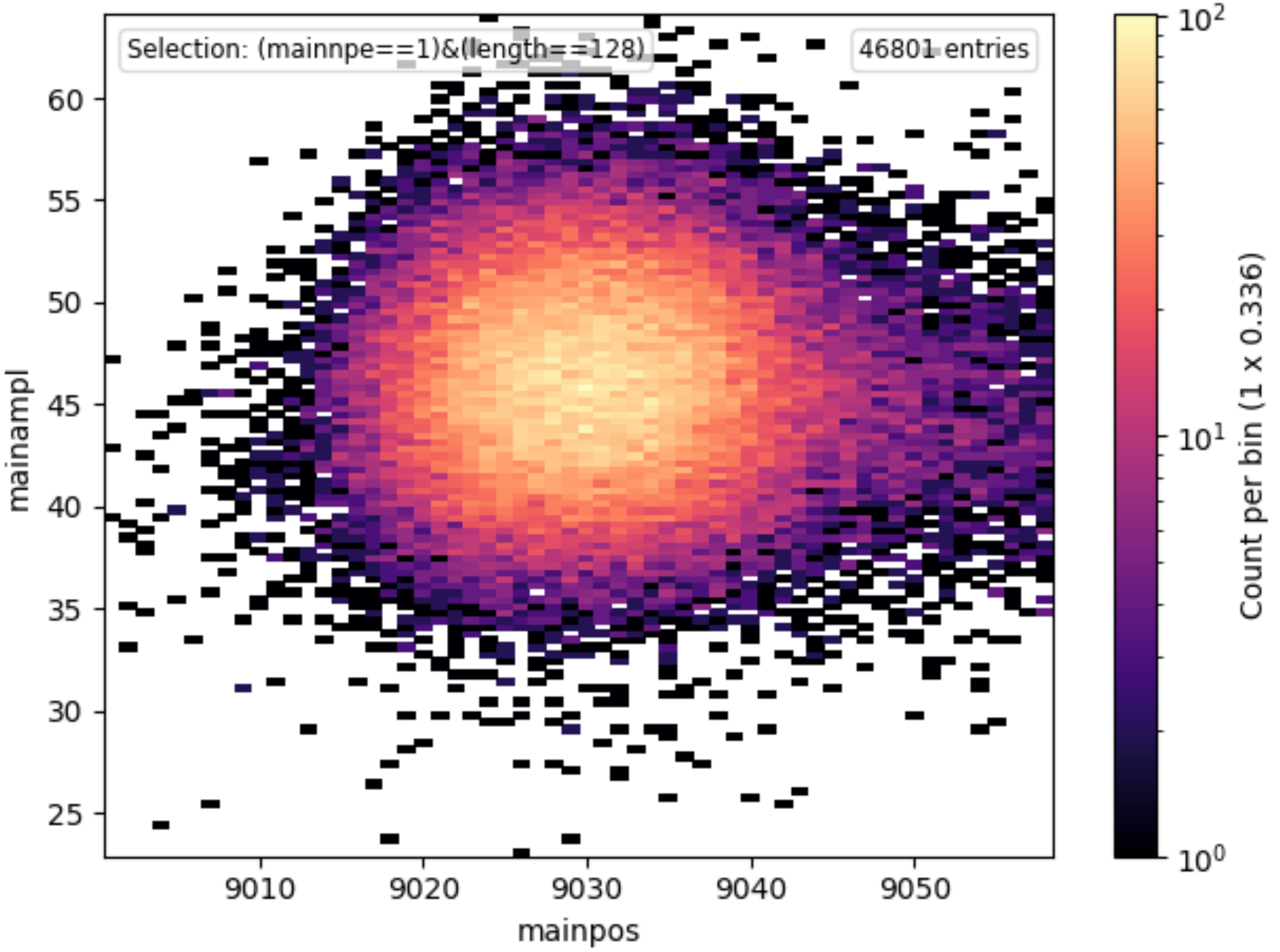
- It is to the right because of the asymmetric signal shape.

Evidence in the next slides.

Laser peak finding (3/4)

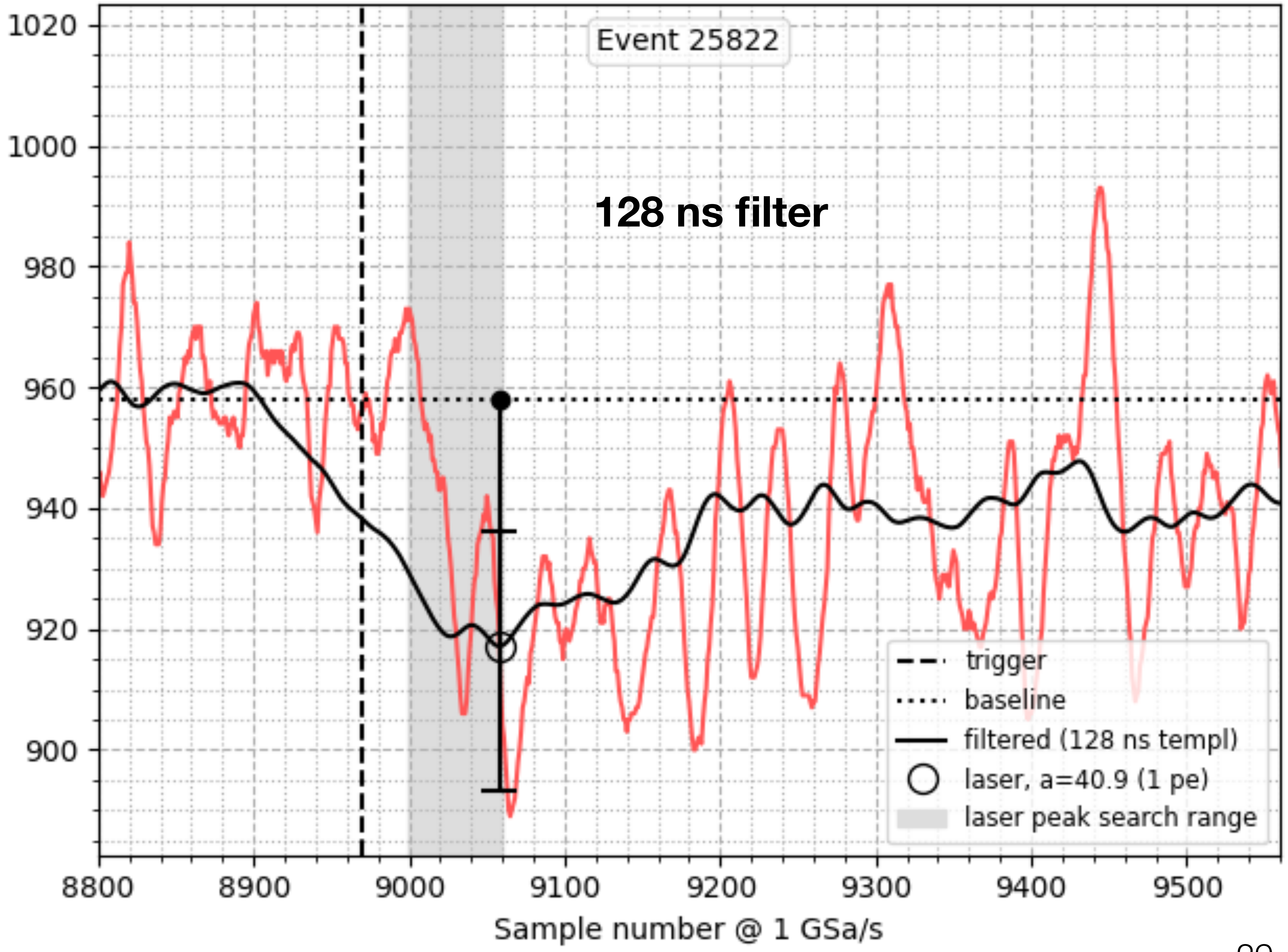
Position-amplitude histogram with short filter.

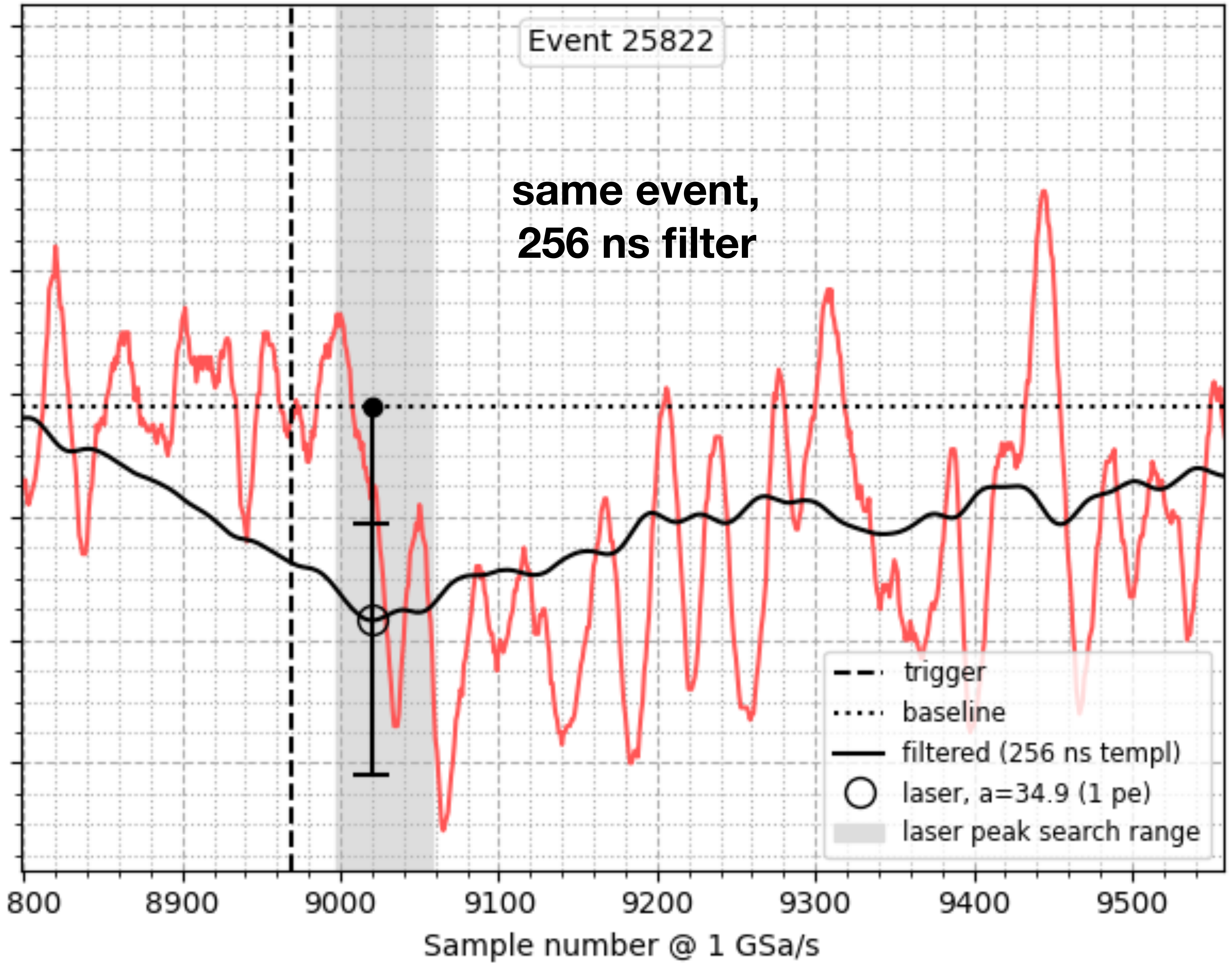
If the tail was due to afterpulses, I would expect a positive correlation. Instead it seems slightly negative.



Laser peak finding (4/4)

I also looked at many events in the tail. They are mostly like this: a bit longer filter is sufficient to cancel the noise and bring back the position to the center.



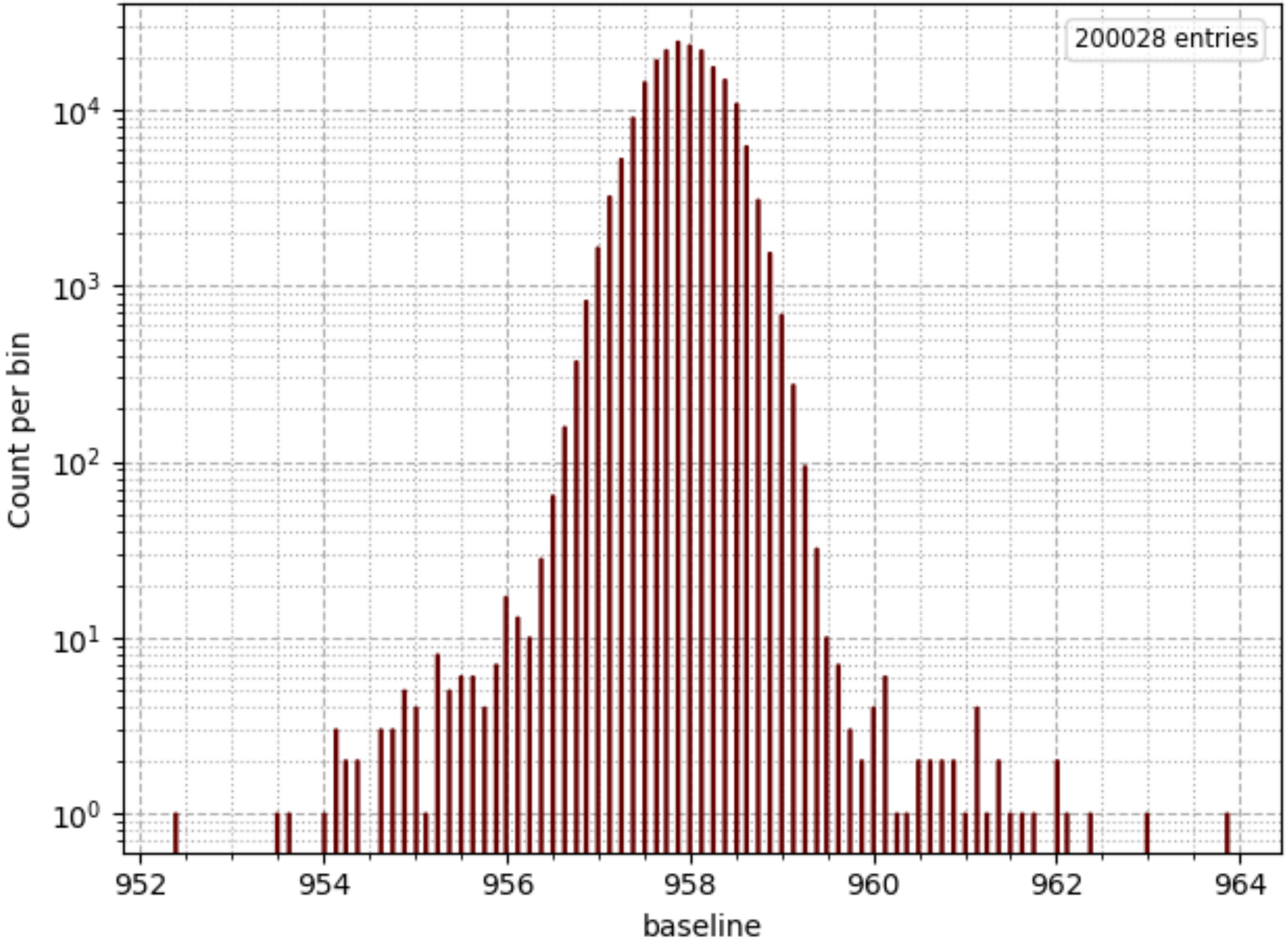


Baseline (1/4)

- I compute the baseline as the median of the pre-trigger region.
- To avoid discretizing too much (the median is discrete), I average the medians of 8 interleaved subarrays.
- If any pre-trigger sample is less than 700, I reuse the baseline from a previous event.

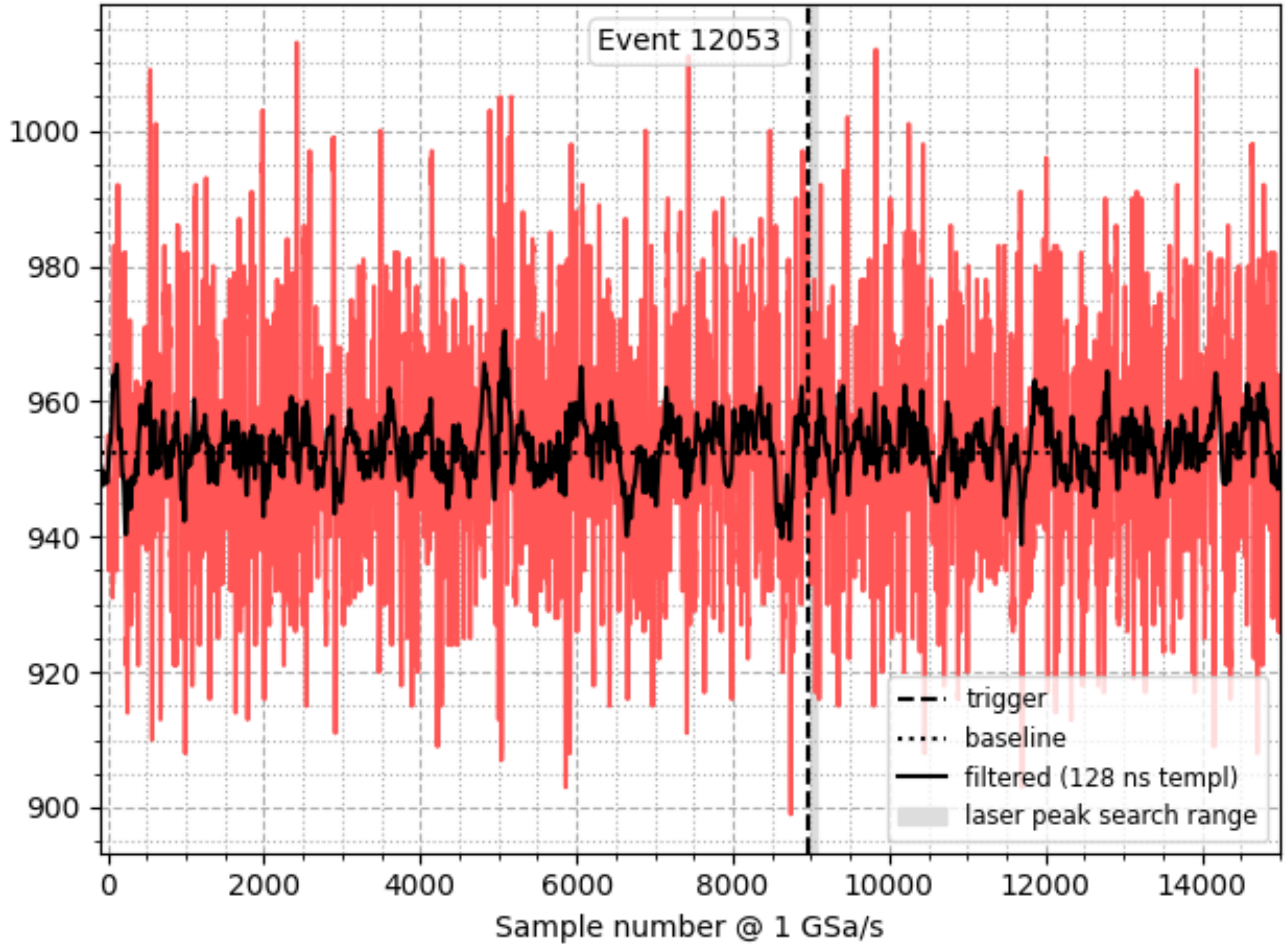
Baseline (2/4)

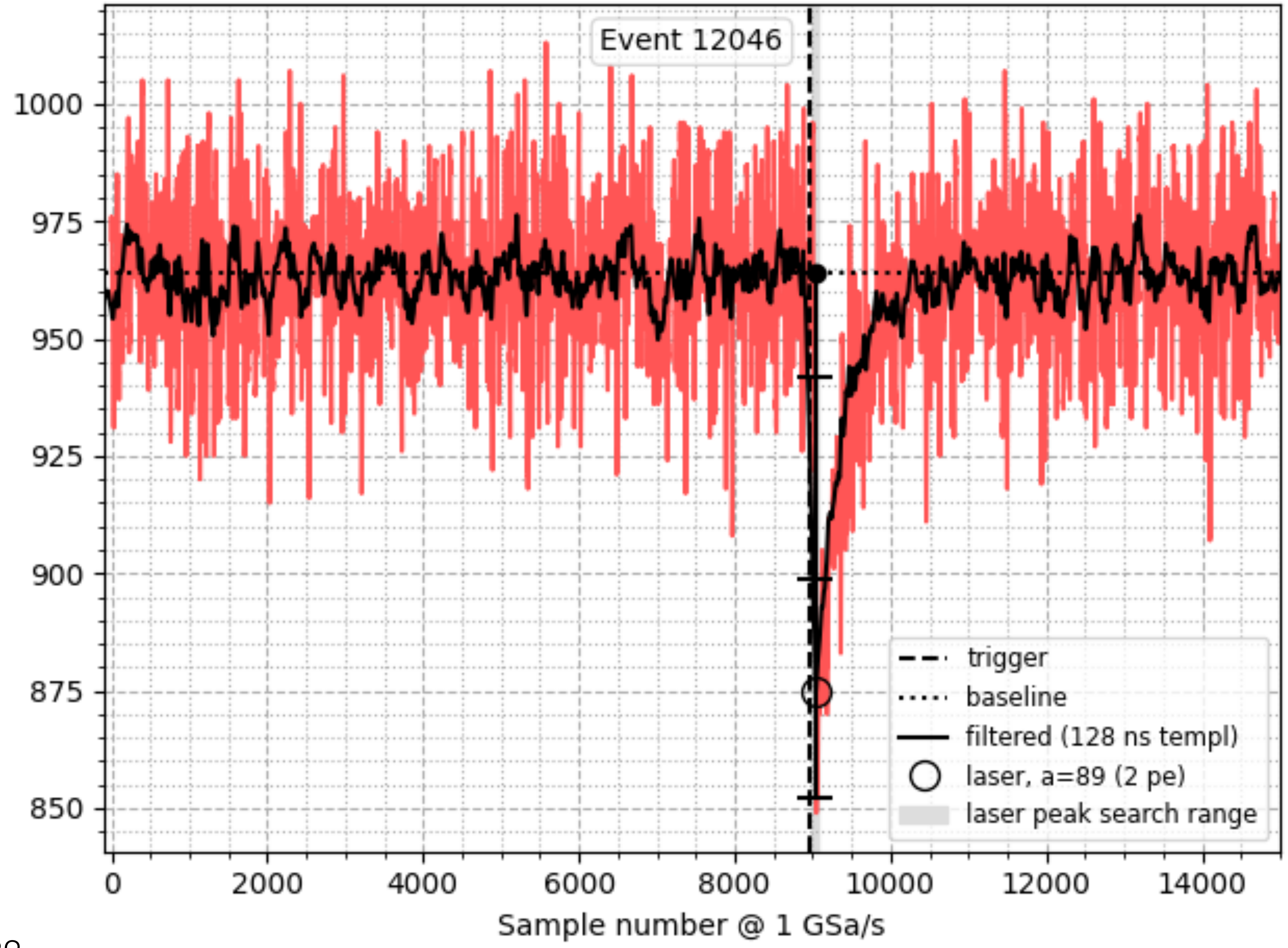
Distribution of the baseline (logscale). There's a small tail to the left.



Baseline (3/4)

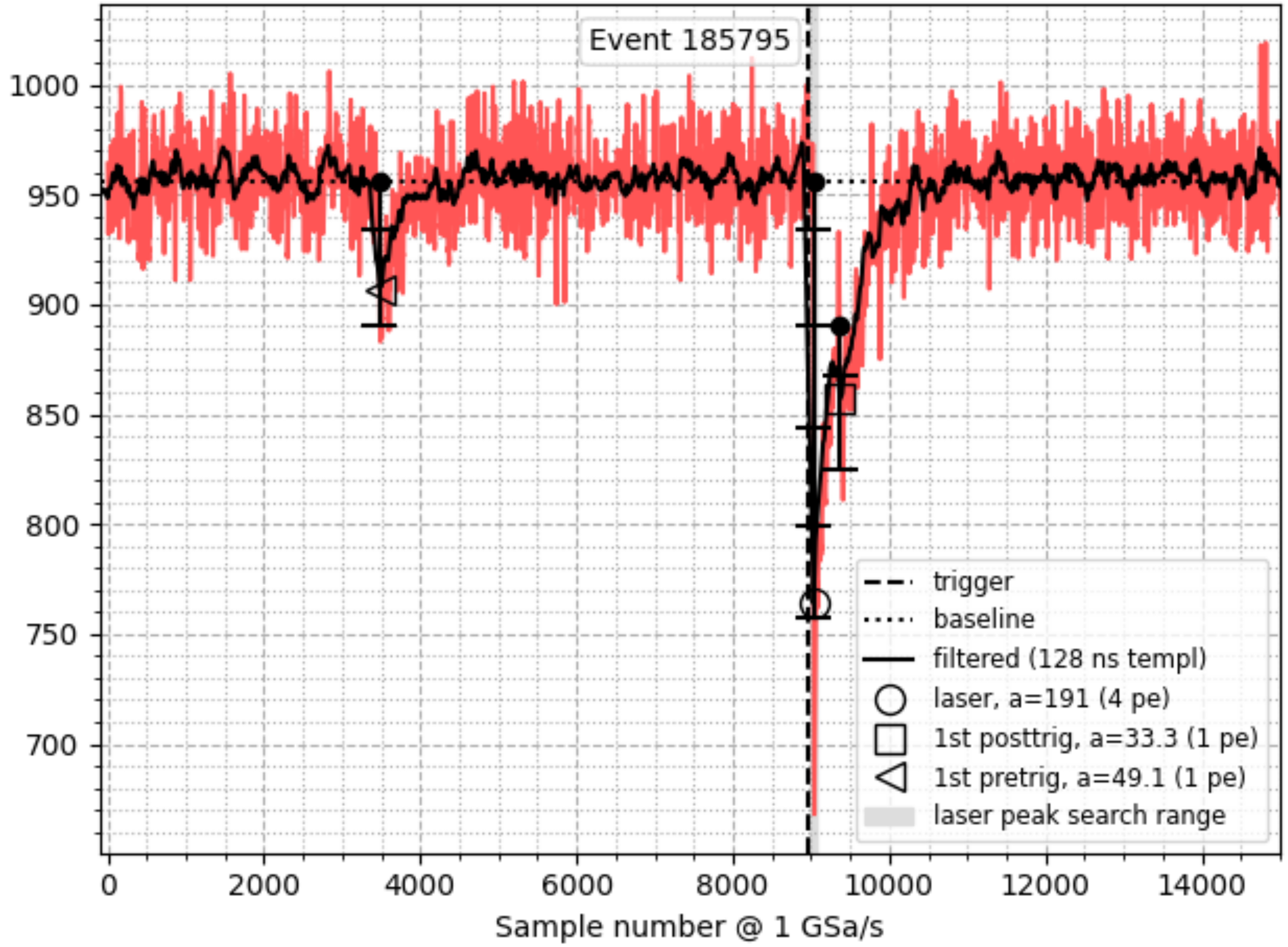
The leftmost and rightmost events in the previous histogram. They are genuinely extreme baselines.

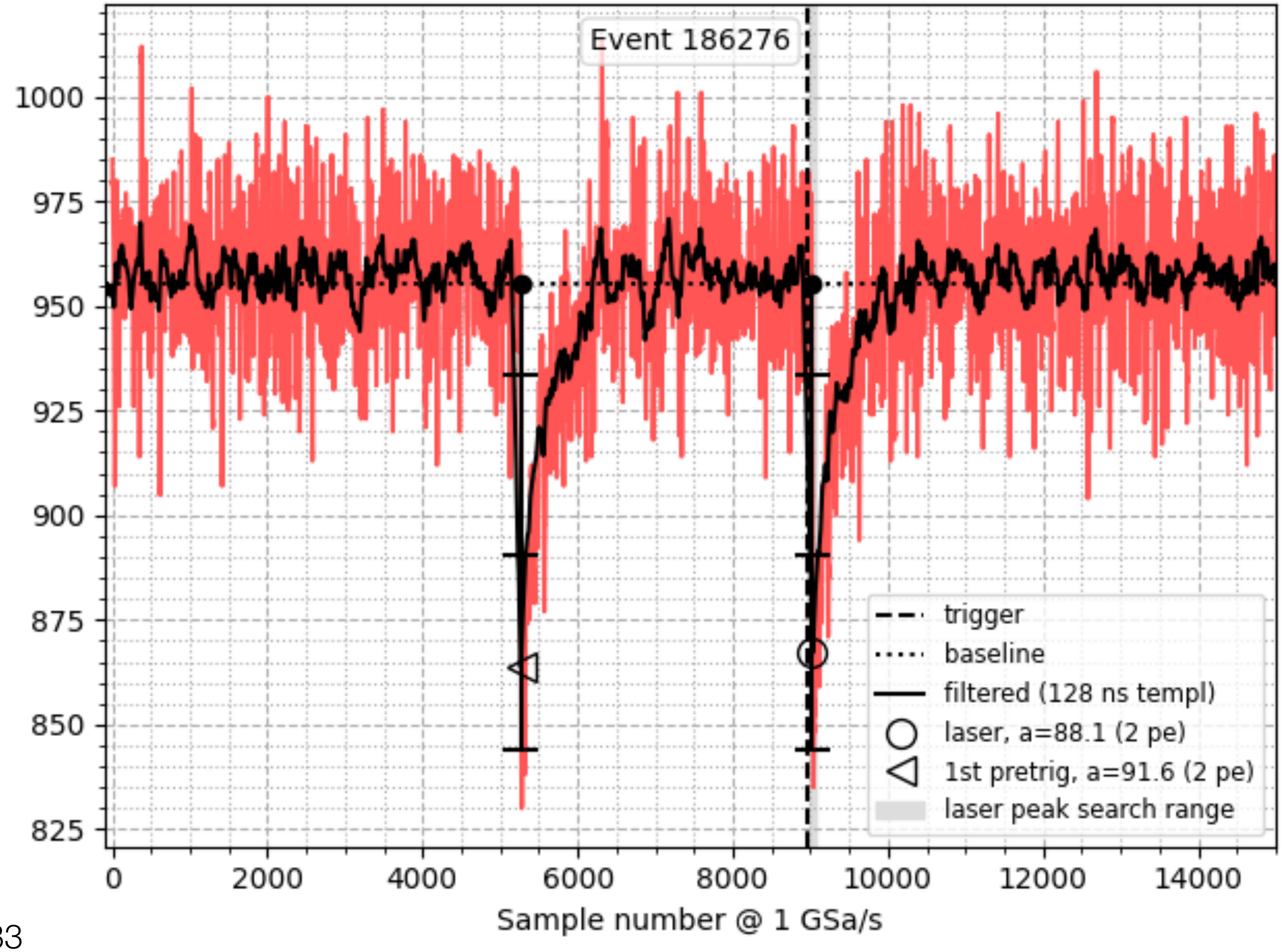




Baseline (4/4)

The events in the left tail instead are mostly events with dark pulses. This means that dark pulses will have their amplitude underestimated, by at most 4. This is not a problem so I did not fix it.





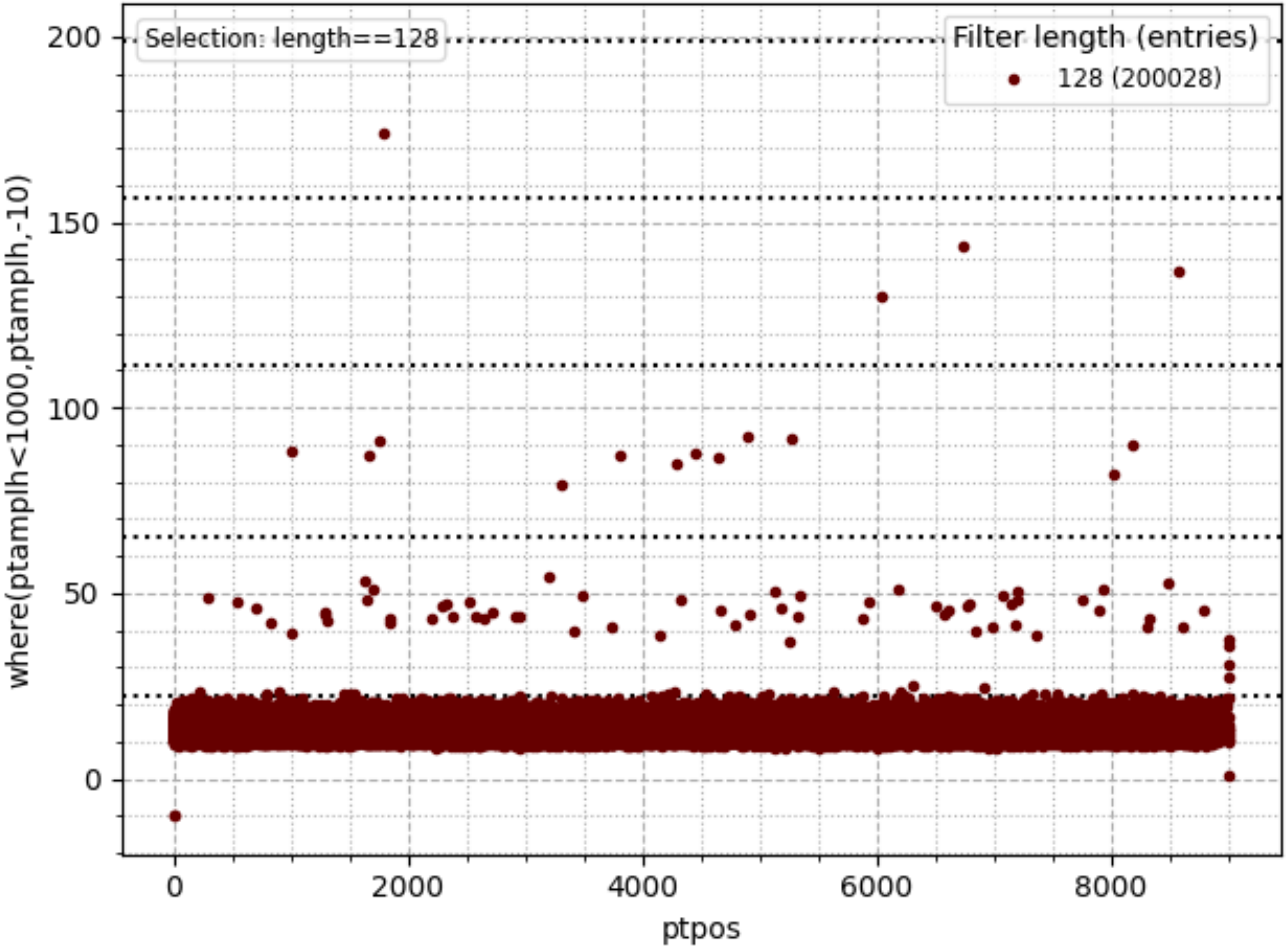
Random pulses rate (1/6)

We need the random pulses rate to subtract the background when counting afterpulses.

We plot the most prominent pre-trigger peak amplitude vs. position.

(The probability of multiple pulses in 9 μs is negligible.)

We notice that there are boundary effects.

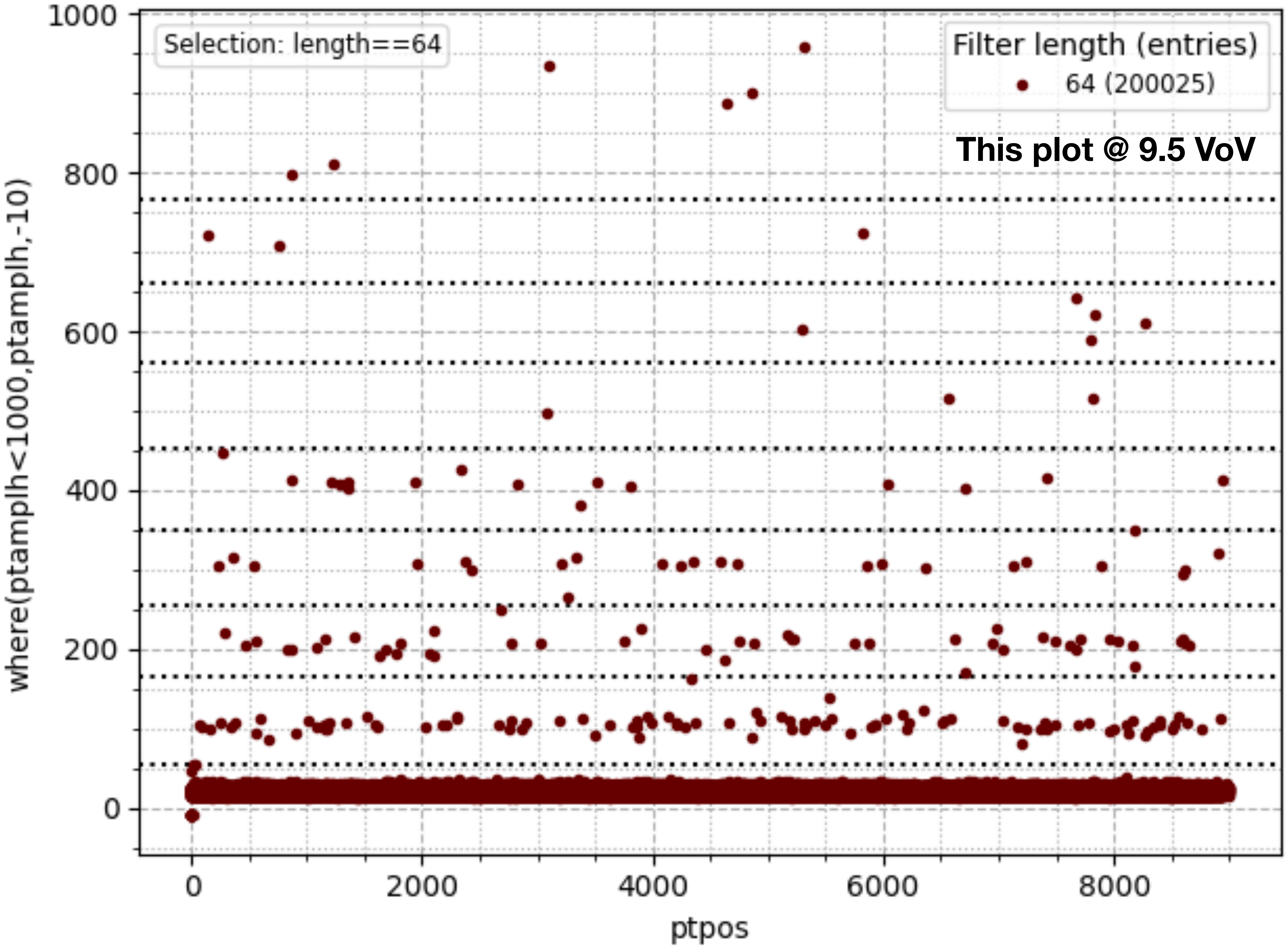


Random pulses rate (2/6)

However, at 9.5 VoV there is another problem: there are too many points very close to a pe bin boundary.

Given the few events and the clean fingerplot, this is unlikely.

(This is not a problem for counting them, it is when measuring cross-talk.)



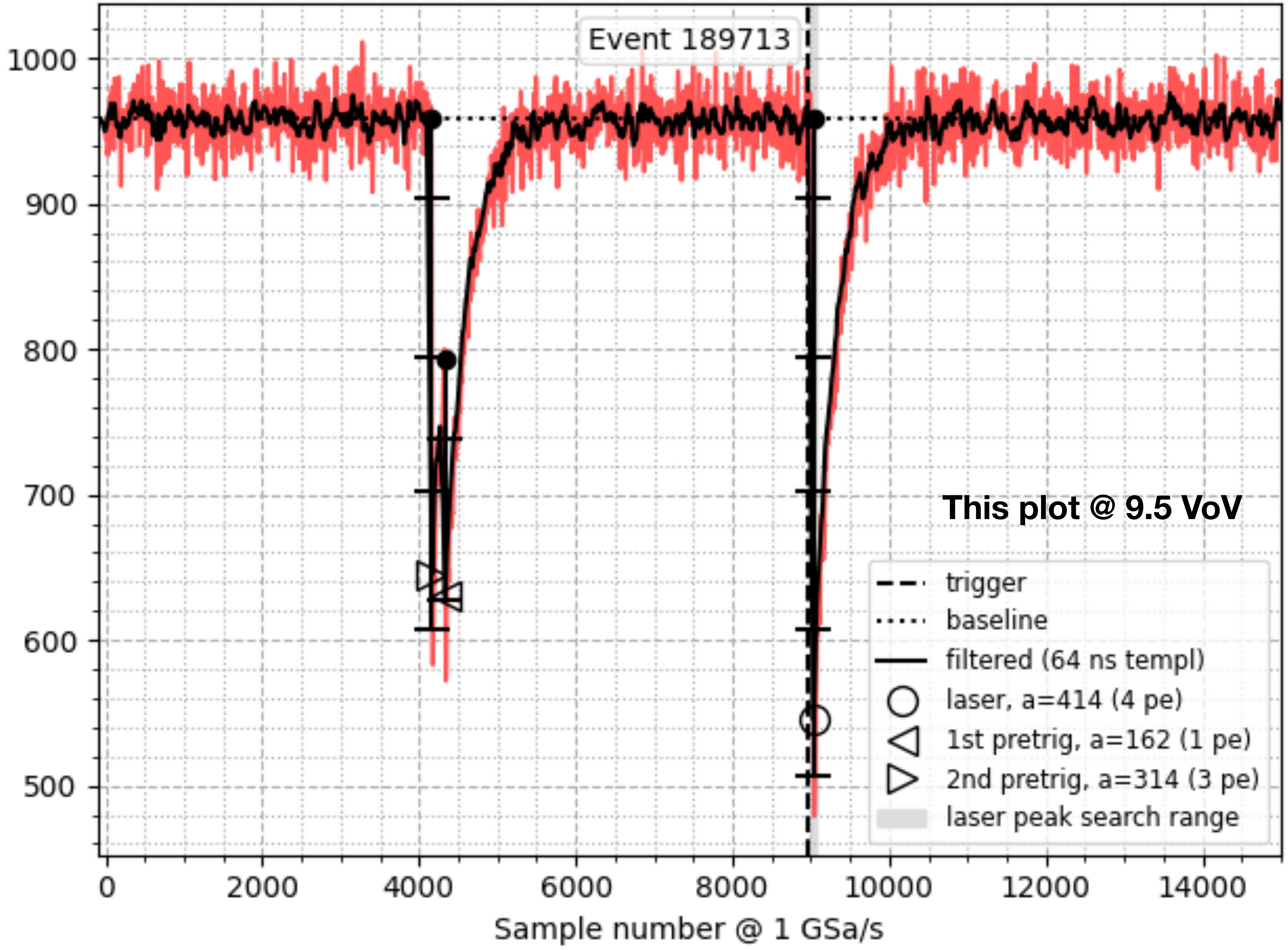
Random pulses rate (3/6)

Looking at these events, we see they are afterpulses.

The afterpulse has the highest prominence and height.

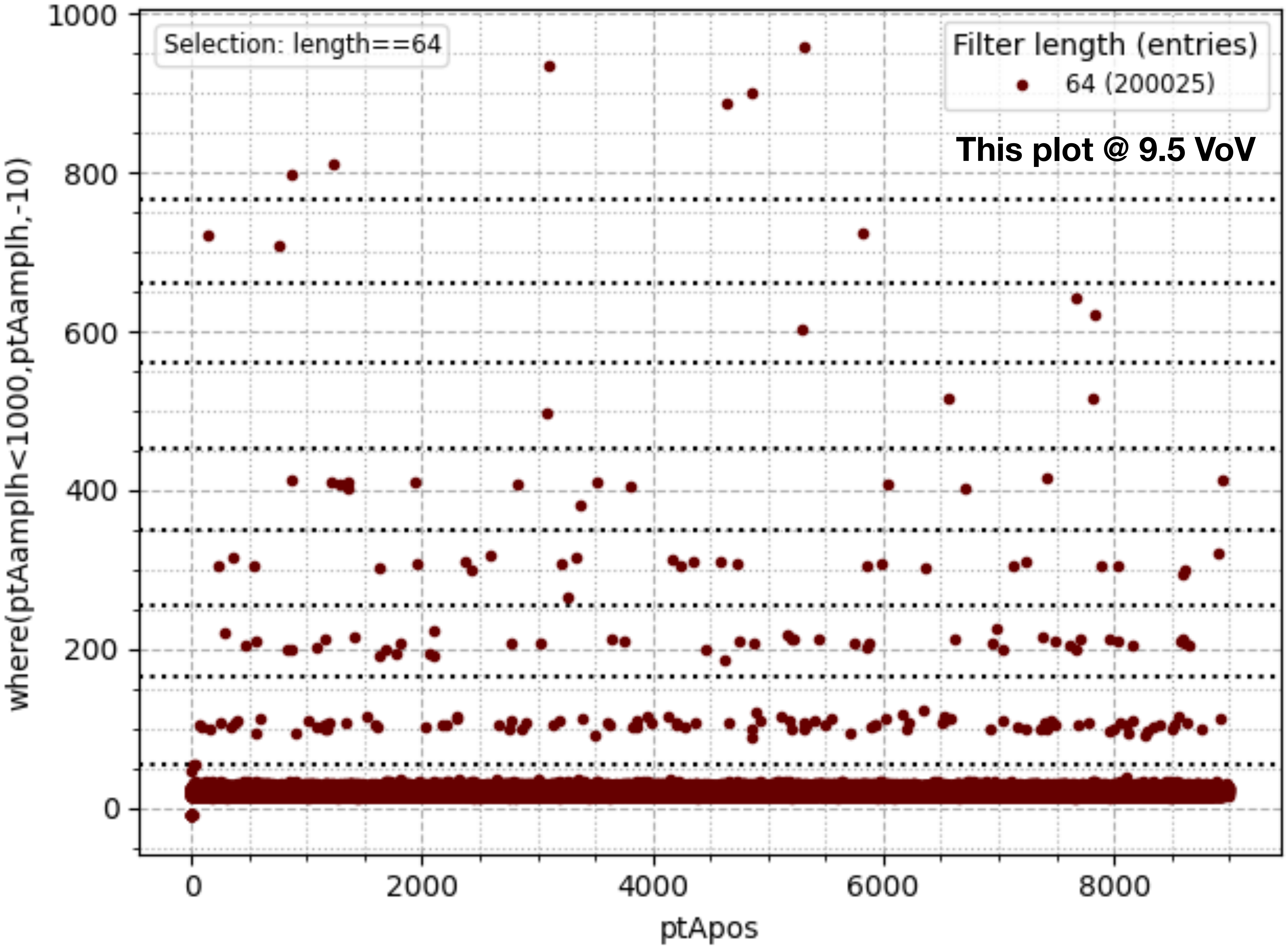
We could select by amplitude instead, but an afterpulse could still be higher.

So when both peaks are above threshold we take the leftmost one, this avoids afterpulses.



Random pulses rate (4/6)

The separation is a bit cleaner now.

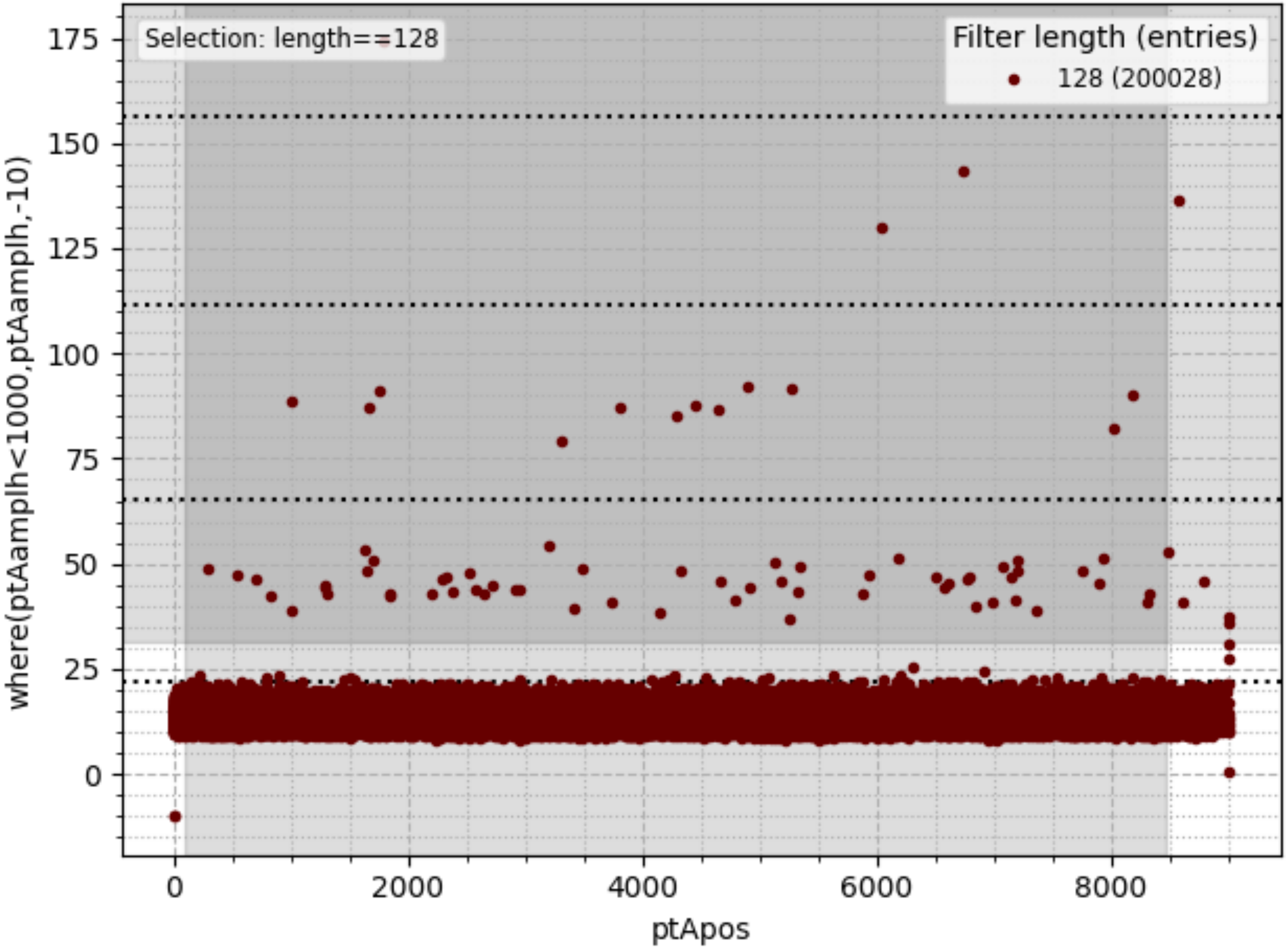


Random pulses rate (5/6)

To count pulses we use a threshold a bit higher than the boundary from the fingerplot, since here the 0 pe count is overwhelming.

We cut 0.1 μs on the left, and 0.5 μs from the right just to be safe since the laser peak could have an effect.

(Same selection used for cross-talk fit.)

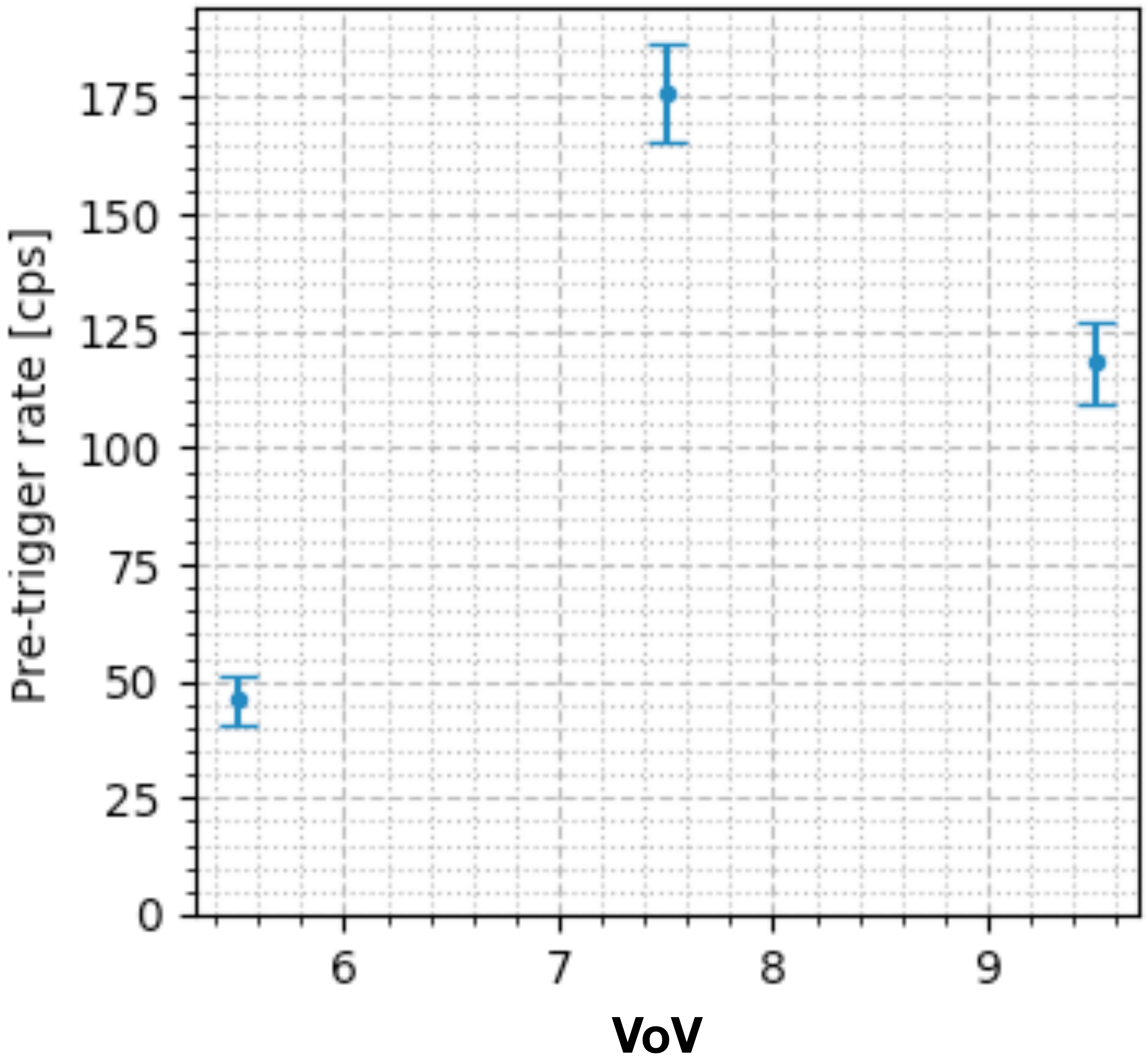


Random pulses rate (6/6)

We count pulses, take the poisson error, divide by the total time with the cuts.

The higher rate at 7.5 VoV would be unlikely if whis was only dark count rate, so we suppose the setup is not light-tight, **take these as upper bounds for DCR**.

However, as correction to the afterpulses count, we do not care if these are not true dark pulses.



Amplitude for close peaks

A detail on the computation of the amplitude (slide 10).

When solving the linear system, we have to decide what peaks to input in the equations.

Most peaks will be random fluctuation and they would introduce a bias since they are selected as the most prominent peaks in the event.

We take the boundary between 0 and 1 pe from the fingerplot, peaks lower than that are ignored.

Important: this does not miss afterpulses because we cut on the absolute height relative to the baseline, not the prominence.

In the plots, when using the amplitude, we actually plot the raw height for peaks which did not pass this preselection.

Amplitude for afterpulses

The afterpulses get shorter as they get closer to the originating pulse.

To select the afterpulses with straight cuts, we add a correction to their amplitude.

The correction we use is the height of a 1 pe signal, evaluated with a delay from the peak, where the delay is the delay of the afterpulse to correct.

This has no particular physical meaning, it is just a tweak sufficient for this analysis.

In fact looking at the afterpulse scatterplot (slide 11) it is evident the dots are still not on a straight line.

Multiple post-trigger pulses

Similarly as we do for the pre-trigger pulses, when there are **two** post-trigger **peaks** above threshold, we **take the leftmost one**.

A. If **one of them is not an afterpulse**, this will strongly select afterpulses since most of them are close to the laser peak.

B. If they are both afterpulses and the second is an afterpulse of the first (**"series" afterpulses**), then this correctly chooses the afterpulse of the laser pulse.

C. If they are **"parallel" afterpulses**, i.e. both originate from the laser pulse, this is a problem (that could be corrected statistically). However each discharge can only generate one afterpulse at a time (because an afterpulse resets other nesting afterpulses), so for 1 pe laser pulses (our selection) it's ok.