# Correlated noise in Lfoundry tile 21 with laser liquid nitrogen data

2021-04-12

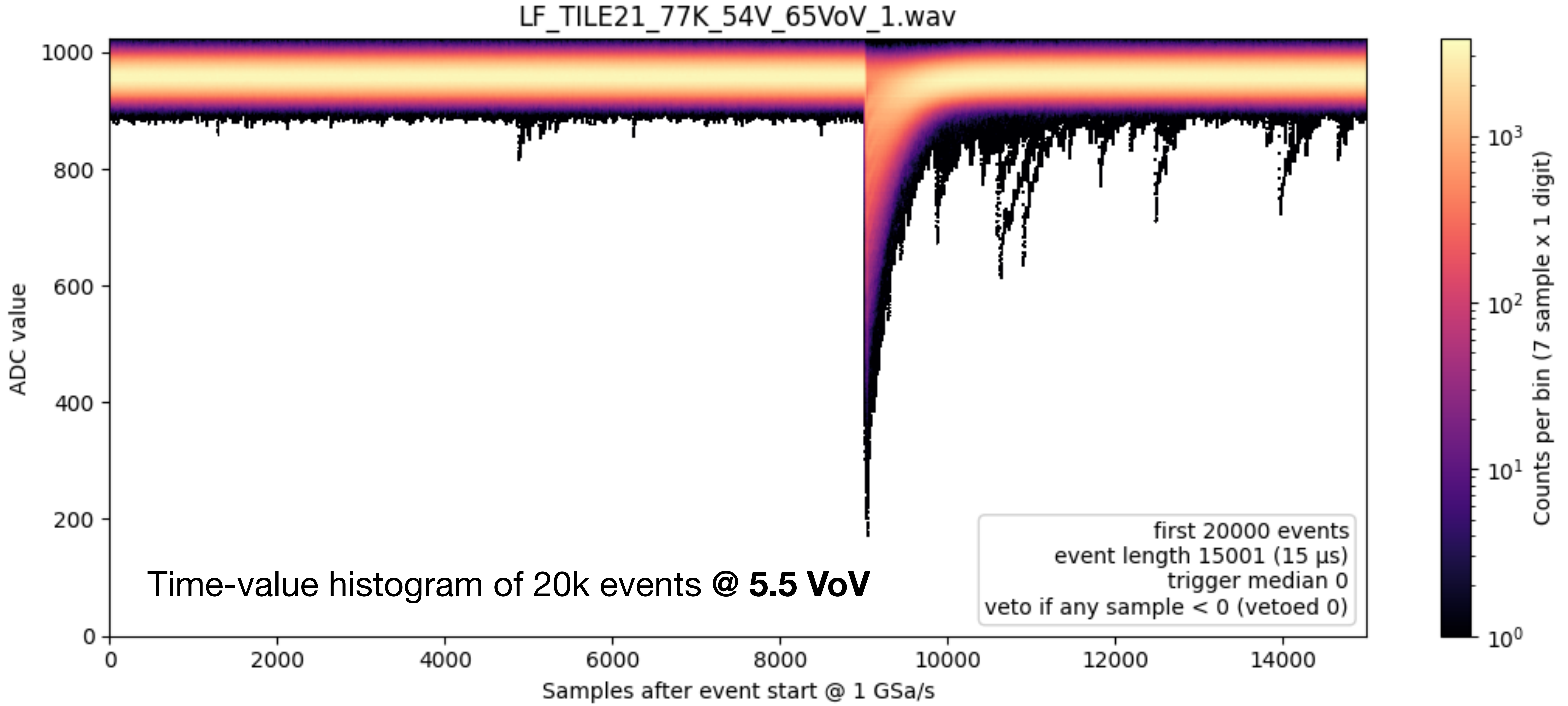
Giacomo Petrillo
University of Pisa
info@giacomopetrillo.com

https://github.com/Gattocrucco/sipmfilter/afterpulse_tile21.py

# Data (1/5)

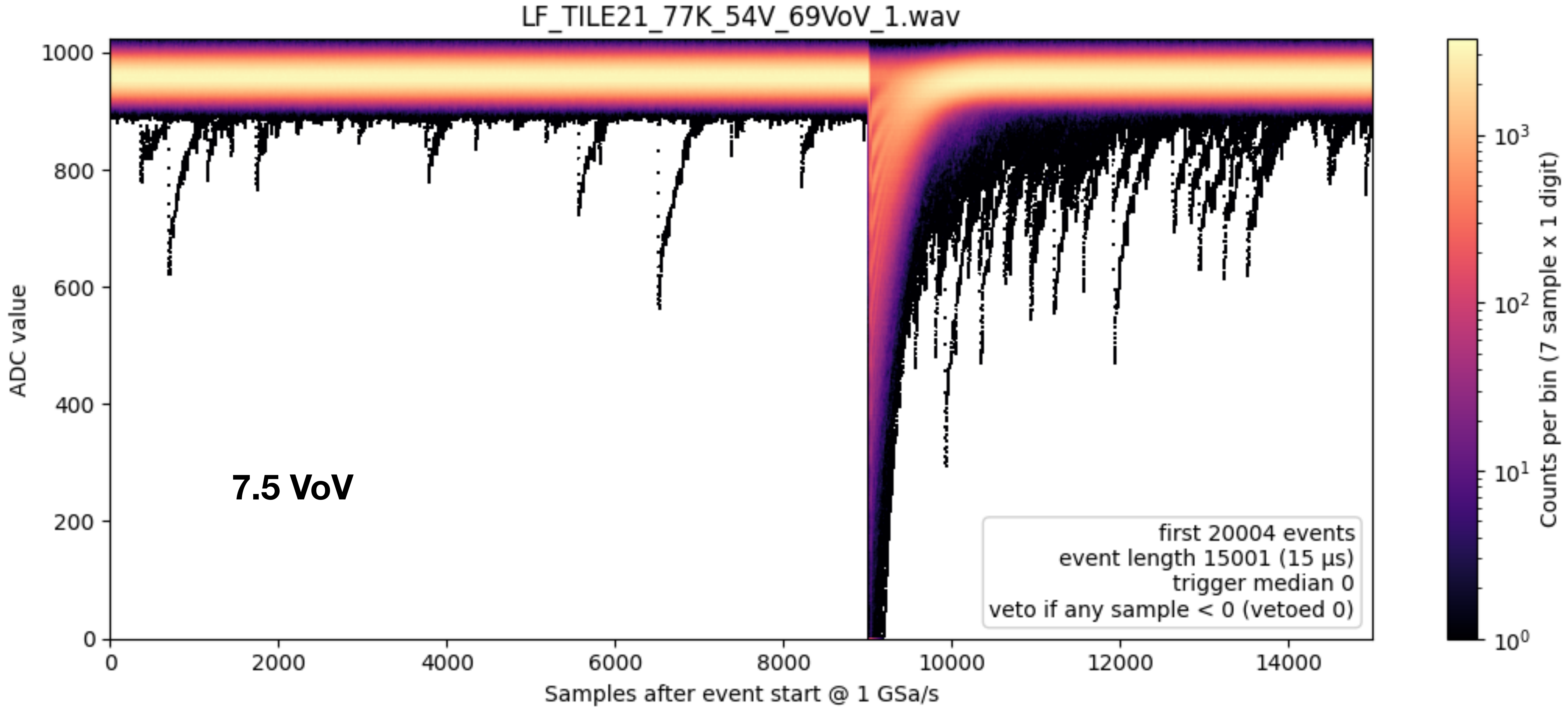
- http://ds50tb.lngs.infn.it:2180/SiPM/Tiles/LFOUNDRY/pre-production-test/TILE_21/
- Sampling frequency 1 GSa/s.
- Each event lasts 15 μs with 9 μs before the laser.
- Three overvoltages: 5.5 V, 7.5 V, 9.5 V.
- 200k events (3 seconds) in 10 files, per overvoltage.
- The laser trigger is not digitalized, however in many other LNGS files it is so I am sure the trigger occurs at sample 8969 ± 21.

# Data (2/5)

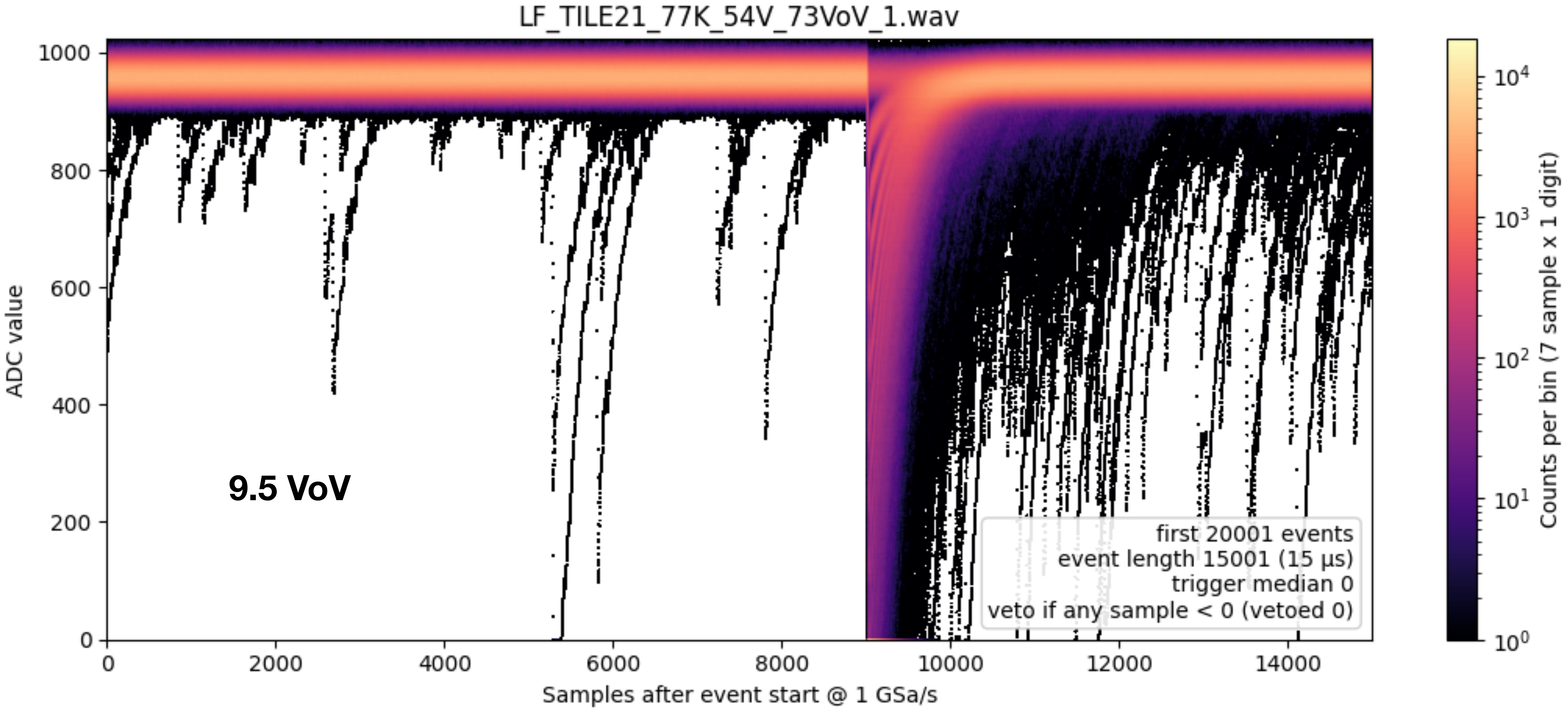


LF_TILE21_77K_54V_65VoV_1.wav
ADC value
1000
800
600
400
200
0
0
2000
4000
6000
8000
10000
12000
14000
Samples after event start @ 1 GSa/s
first 20000 events
event length 15001 (15 µs)
trigger median 0
veto if any sample < 0 (vetoed 0)
Time-value histogram of 20k events @ **5.5 VoV**
Counts per bin (7 sample x 1 digit)
$10^3$
$10^2$
$10^1$
$10^0$

# Data (3/5)

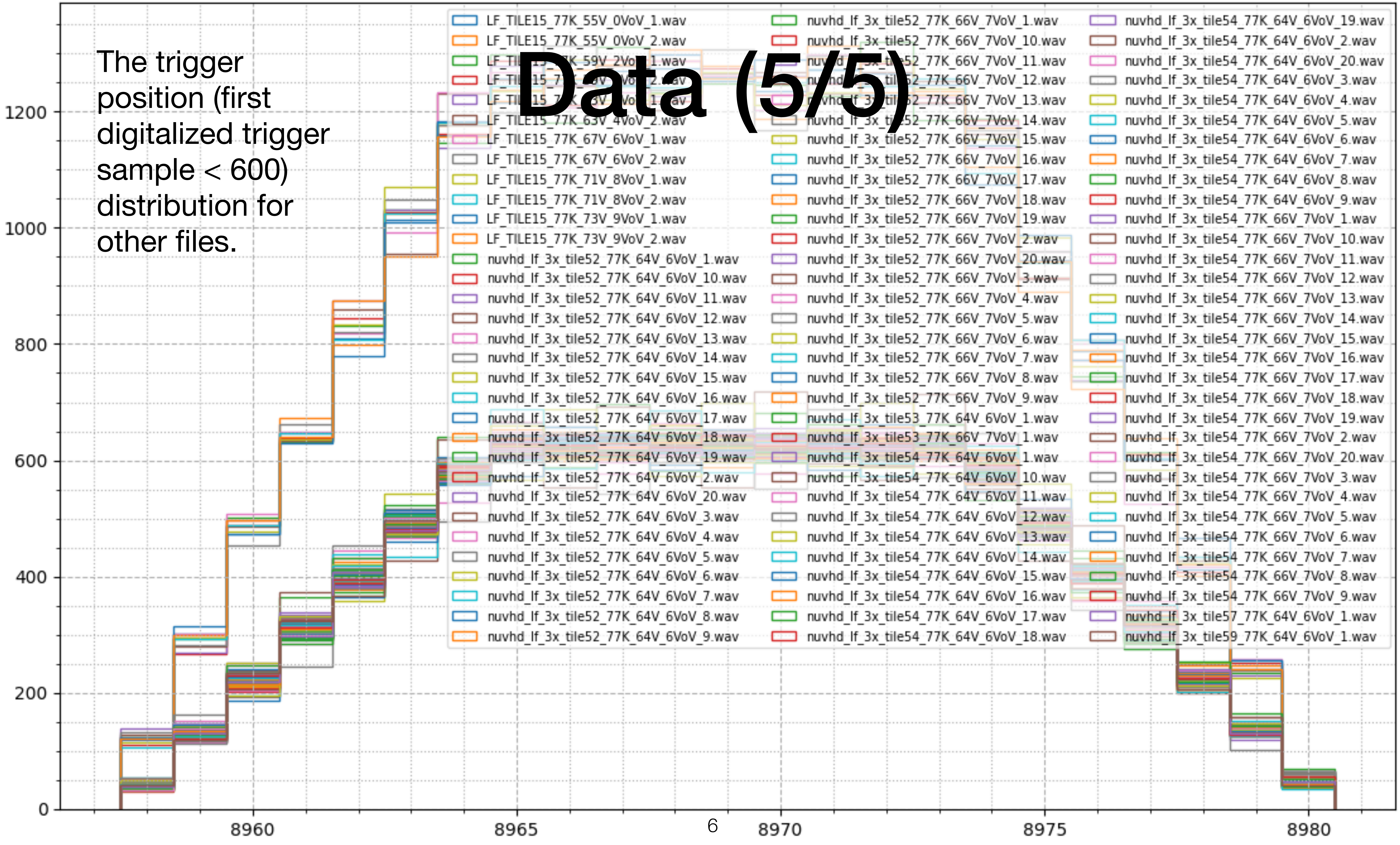

LF_TILE21_77K_54V_69VoV_1.wav
ADC value
Samples after event start @ 1 GSa/s
Counts per bin (7 sample x 1 digit)
7.5 VoV
first 20004 events
event length 15001 (15 µs)
trigger median 0
veto if any sample < 0 (vetoed 0)

# Data (4/5)


LF_TILE21_77K_54V_73VoV_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
9.5 VoV
first 20001 events
event length 15001 (15 µs)
trigger median 0
veto if any sample < 0 (vetoed 0)
10^4
10^3
10^2
10^1
10^0
Counts per bin (7 sample x 1 digit)

# Data (5/5)

The trigger position (first digitalized trigger sample < 600) distribution for other files.



# Filter (1/3)

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

1. Compute the baseline with the pretrigger 9 μs.

2. Select 1 pe signals with a 1.5 μs integration (we use 1 pe because it permits a cleaner selection than multiple pe).

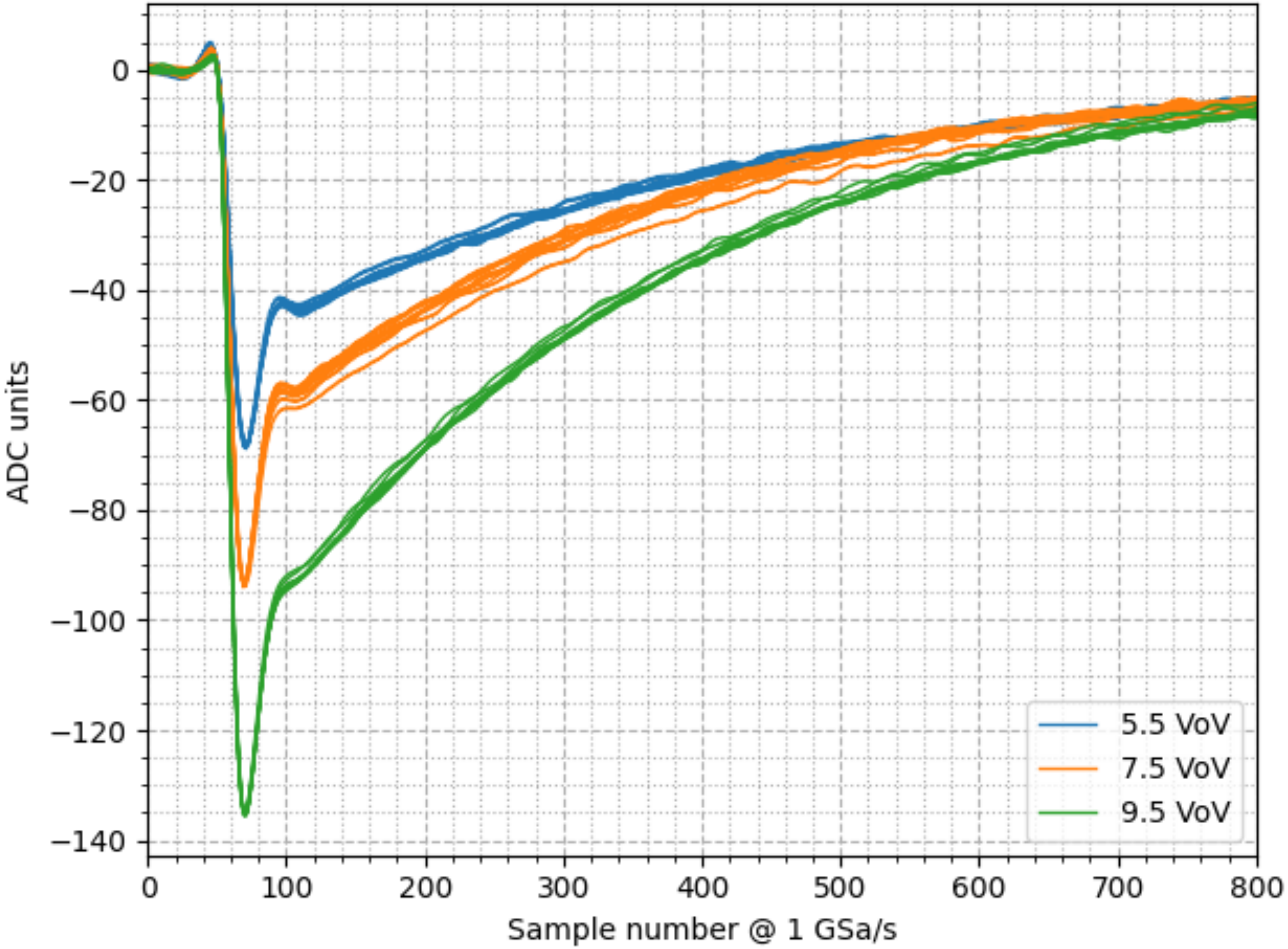
3. Average 3.5 μs starting from the (assumed) trigger position to get a first template.

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 (2/3)

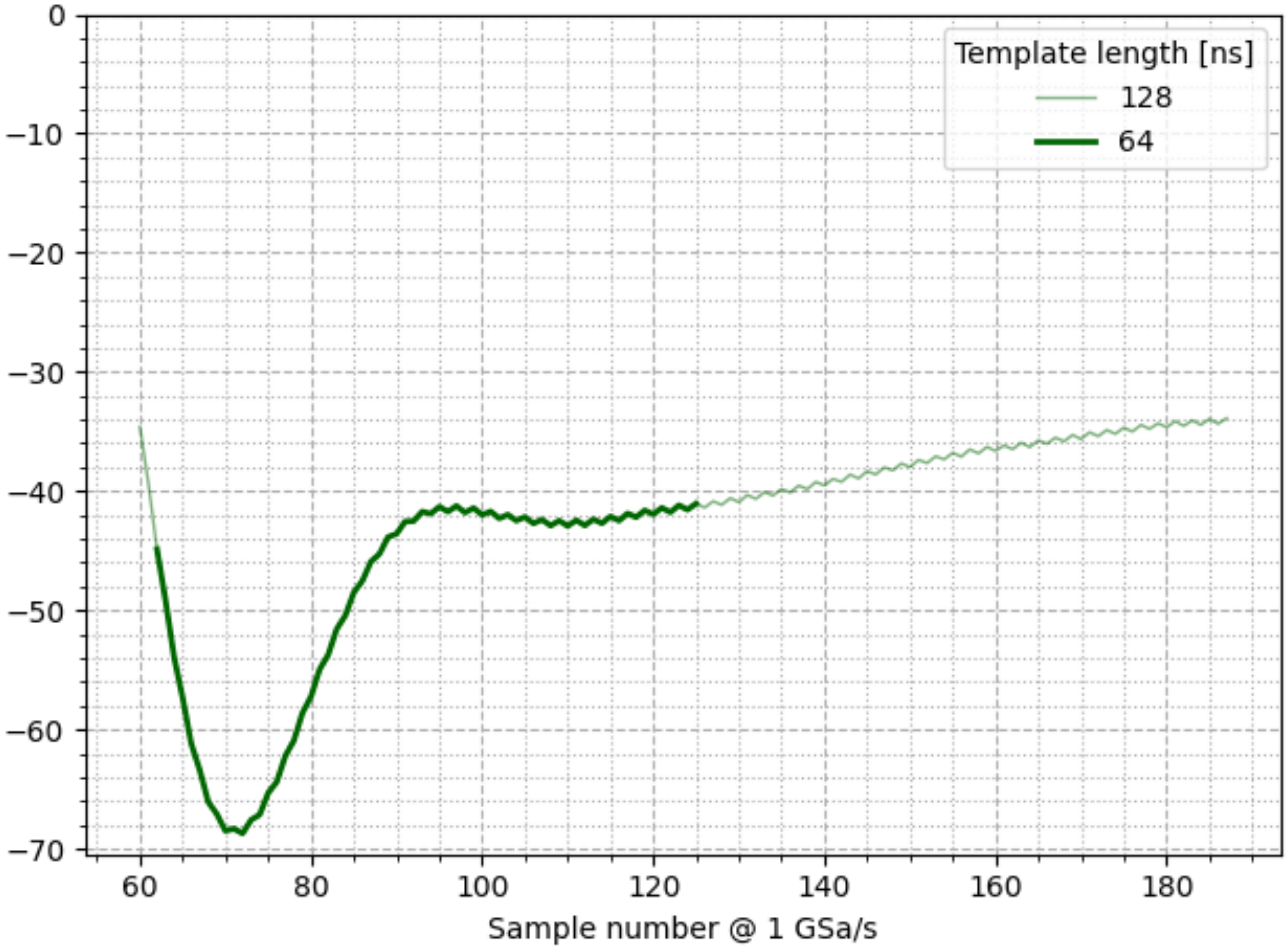
The templates obtained.



# Filter (3/3)

We will use filters with different lengths. We truncate the template by keeping the highest euclidean norm part.

(The wiggles are noise at the Nyquist frequency that does not get cancelled by the average.)



# Laser peak finding (1/5)

After filtering, to find the laser pulse:

1. Take a window of ±30 samples around the assumed mean trigger position 8969 (plus an offset to take into account filter template truncation).

2. Take the minimum relative minimum in the window.

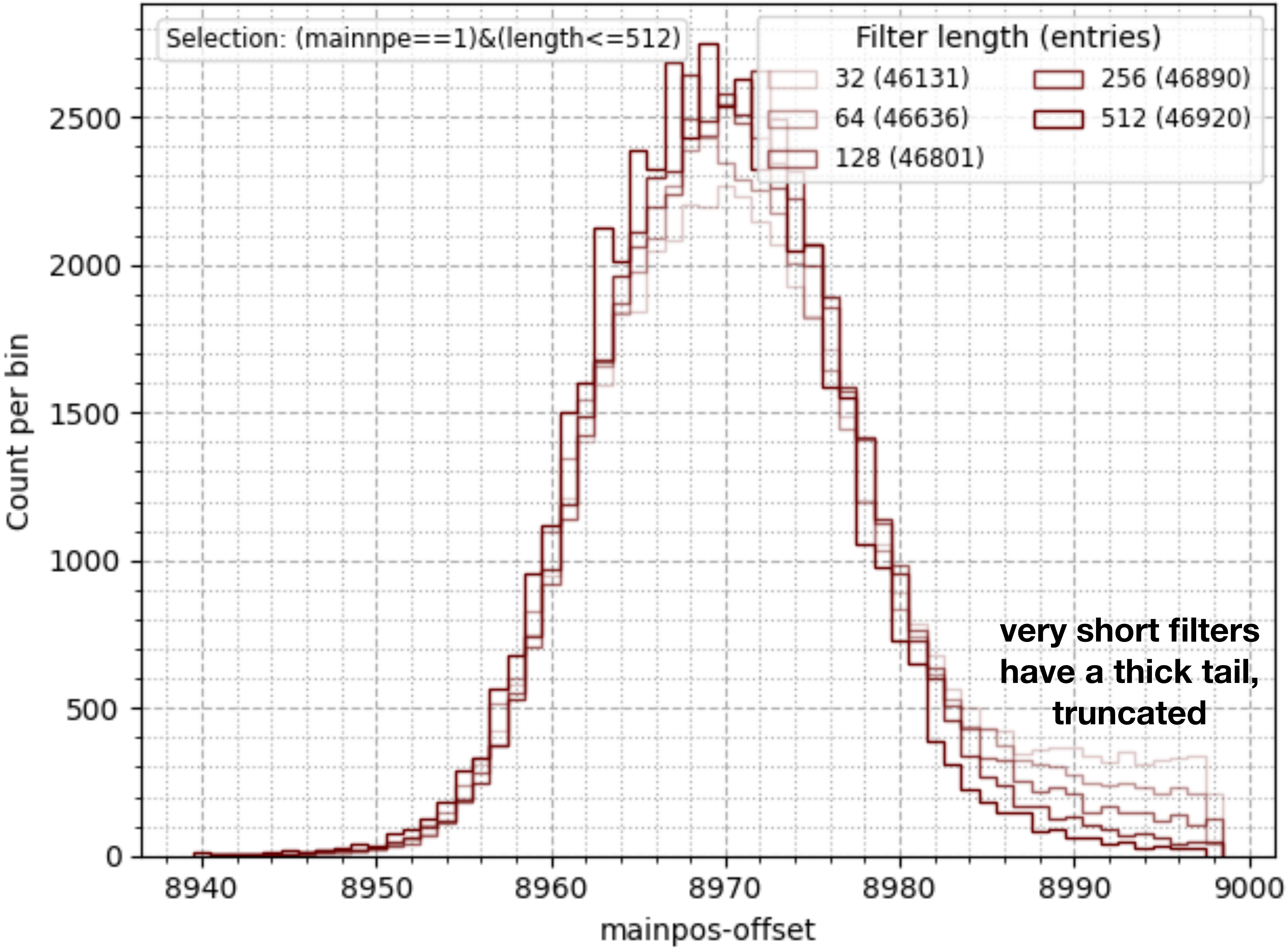
3. If there's no relative minimum, no peak found.

# Laser peak finding (2/5)

Distribution of the laser peak position. (When not otherwise stated, the plots are at **5.5 VoV**.)

**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.

(This plot is selected for 1 pe, I will describe later how this is done.)



# Laser peak finding (3/5)

I give this interpretation to the tail:

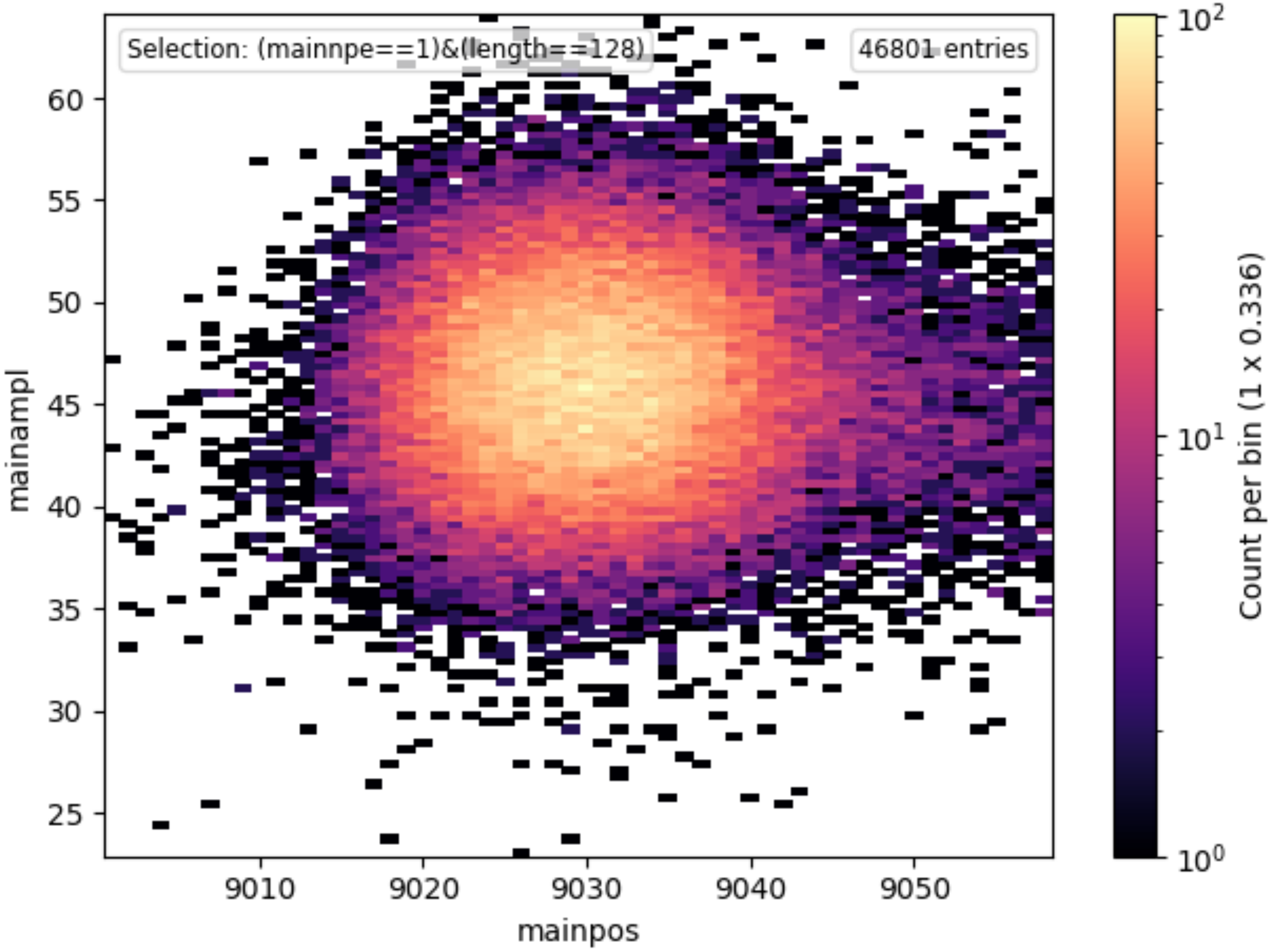
- The tail is random fluctuation.
- It appears with short filters because they have lower SNR.
- It is to the right because of the asymmetric signal shape.

Further evidence in the next two slides.

# Laser peak finding (4/5)

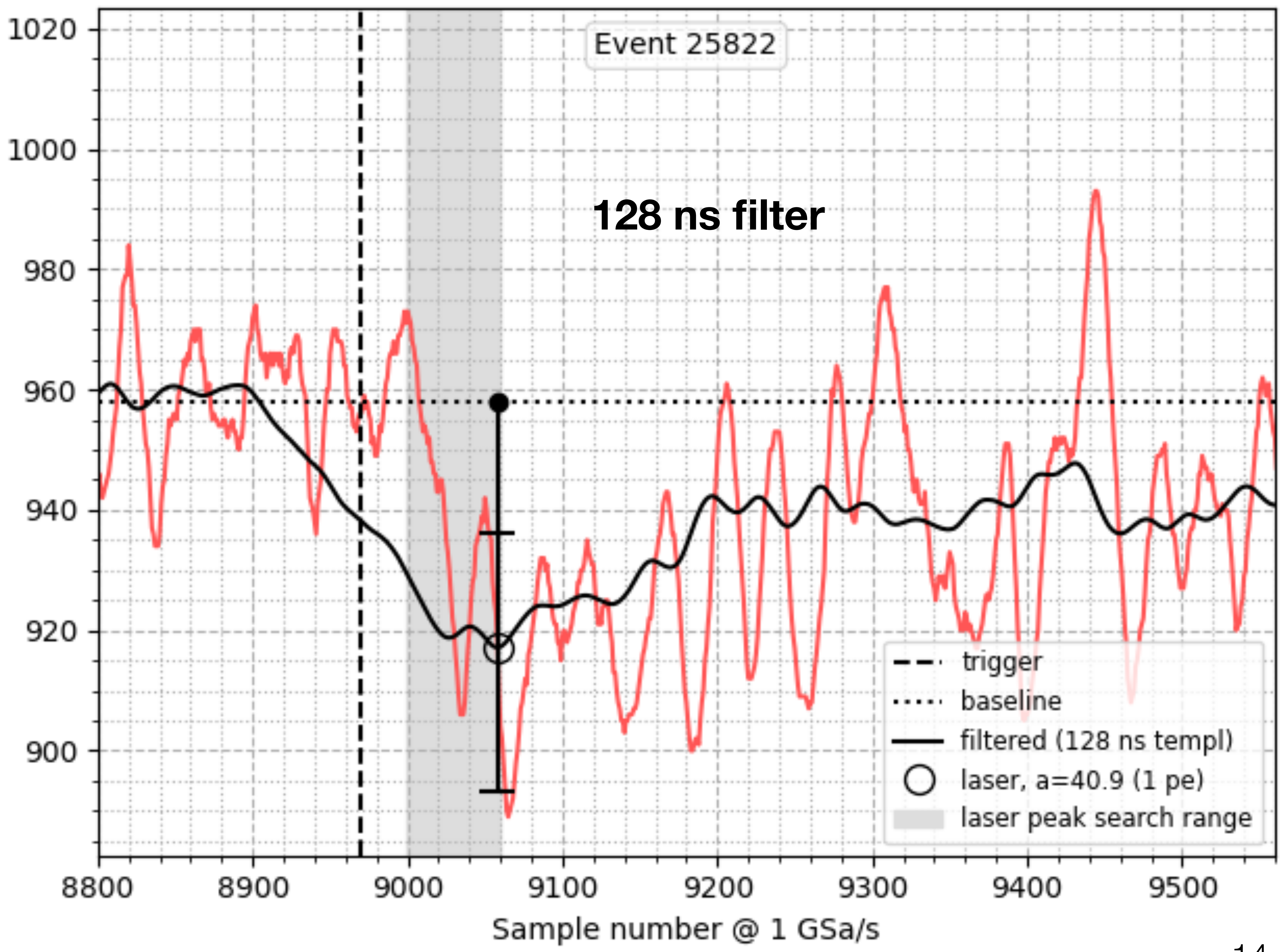
Position-amplitude histogram with short filter.

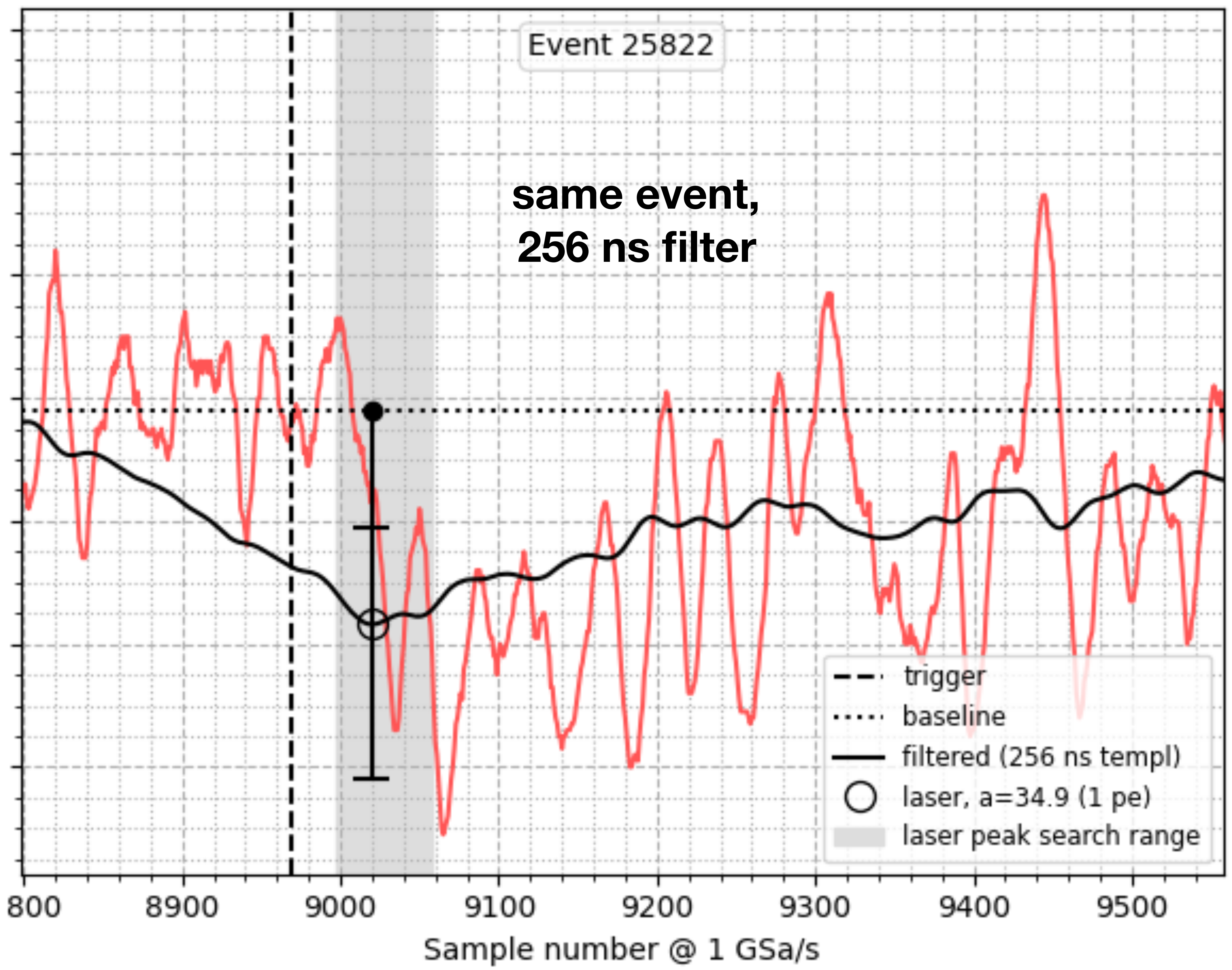
If the tail was due to afterpulses, I would expect a positive correlation. Instead it seems slightly negative.



# Laser peak finding (5/5)

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.



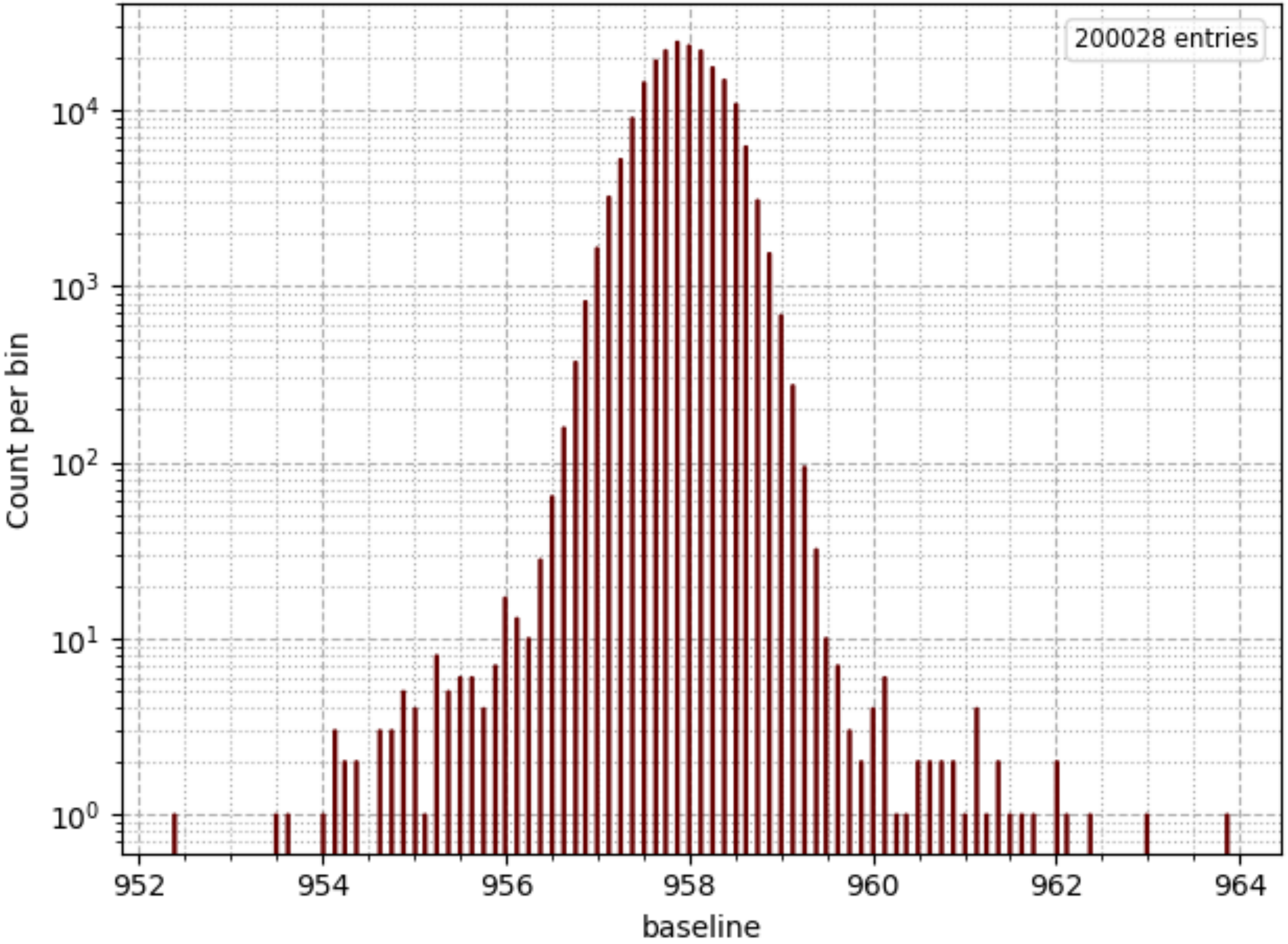



# Baseline (1/4)

- 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.

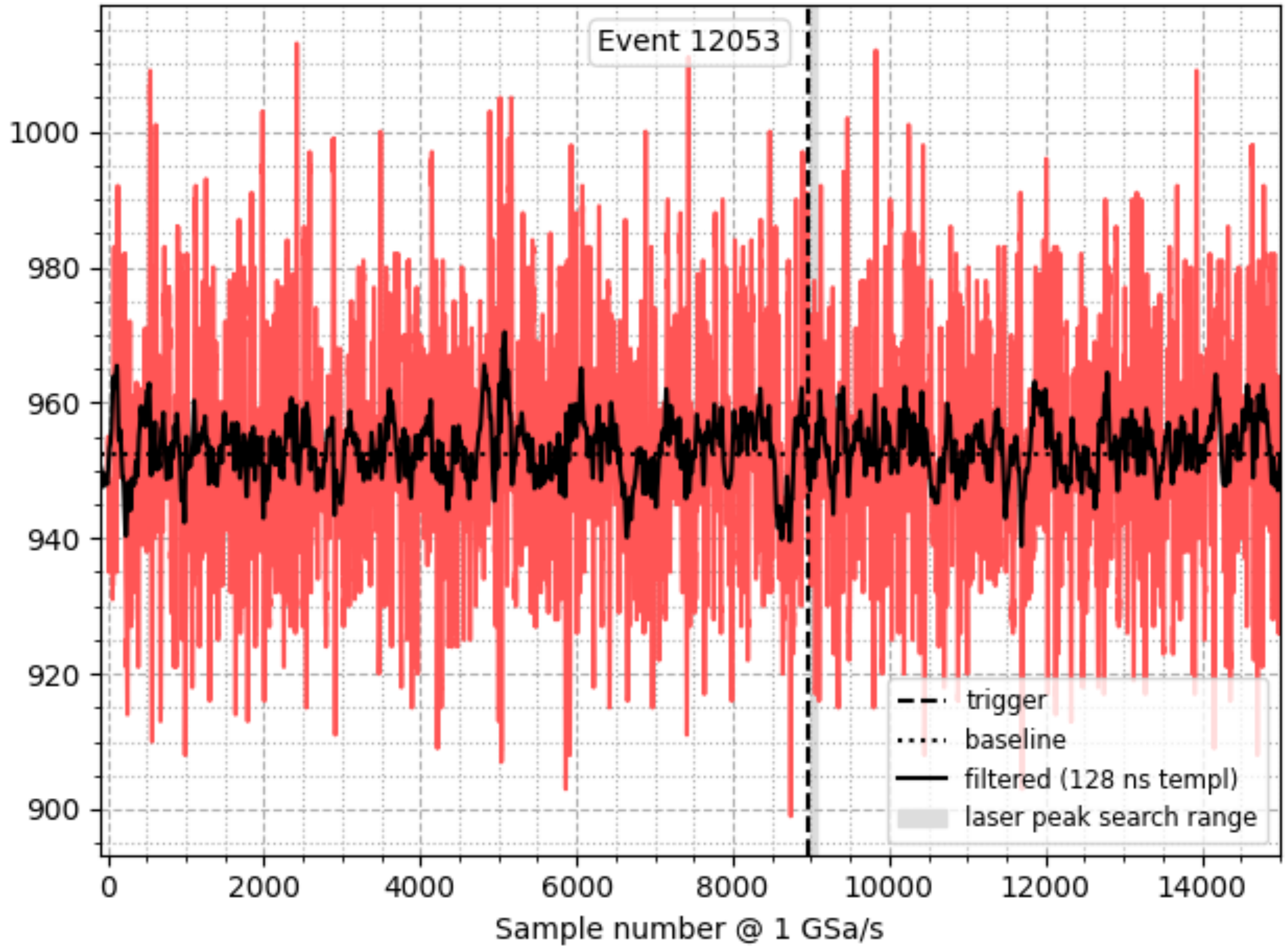
# Baseline (2/4)

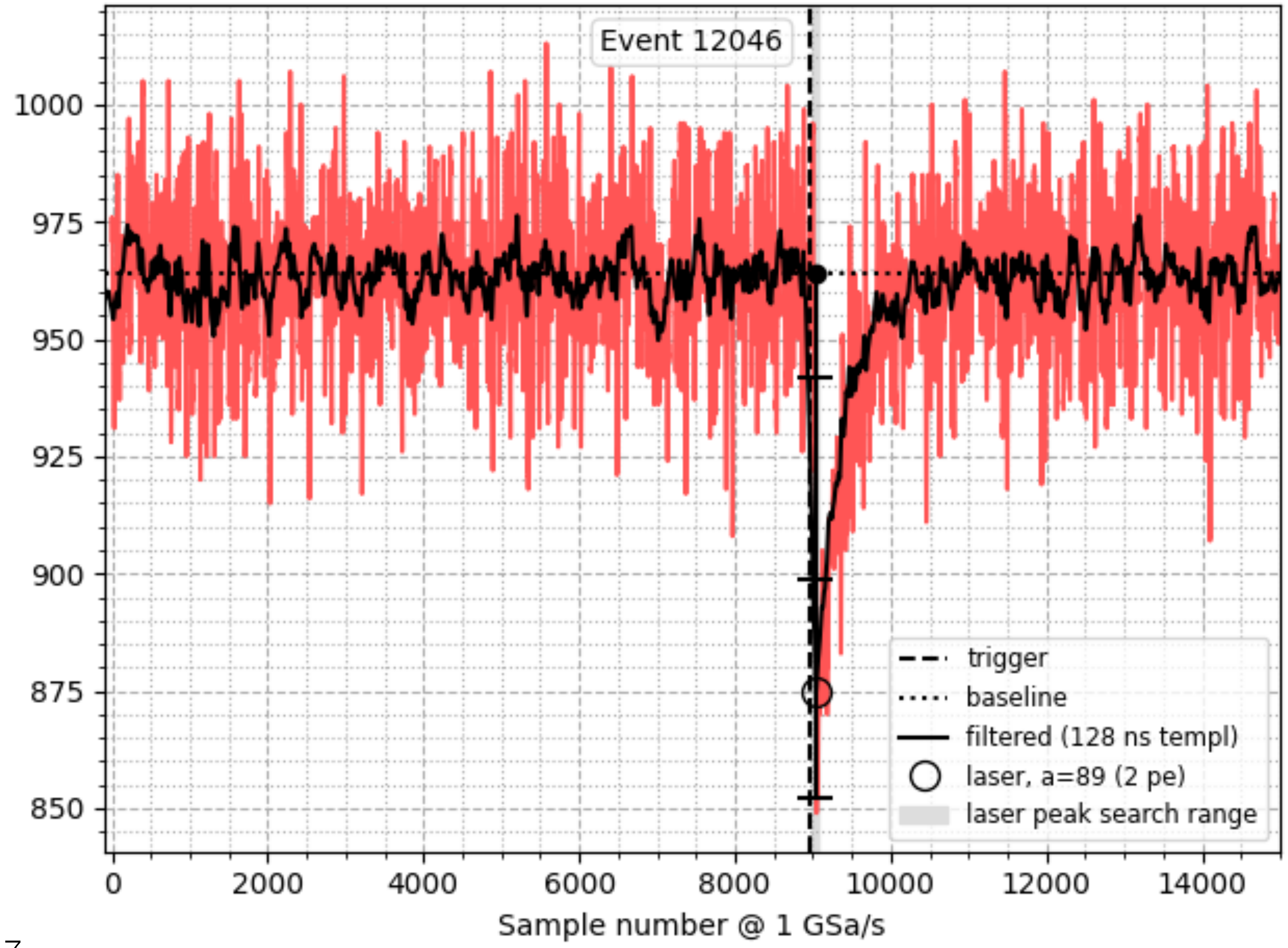
Distribution of the baseline (logscale). There's a small tail to the left.



# Baseline (3/4)

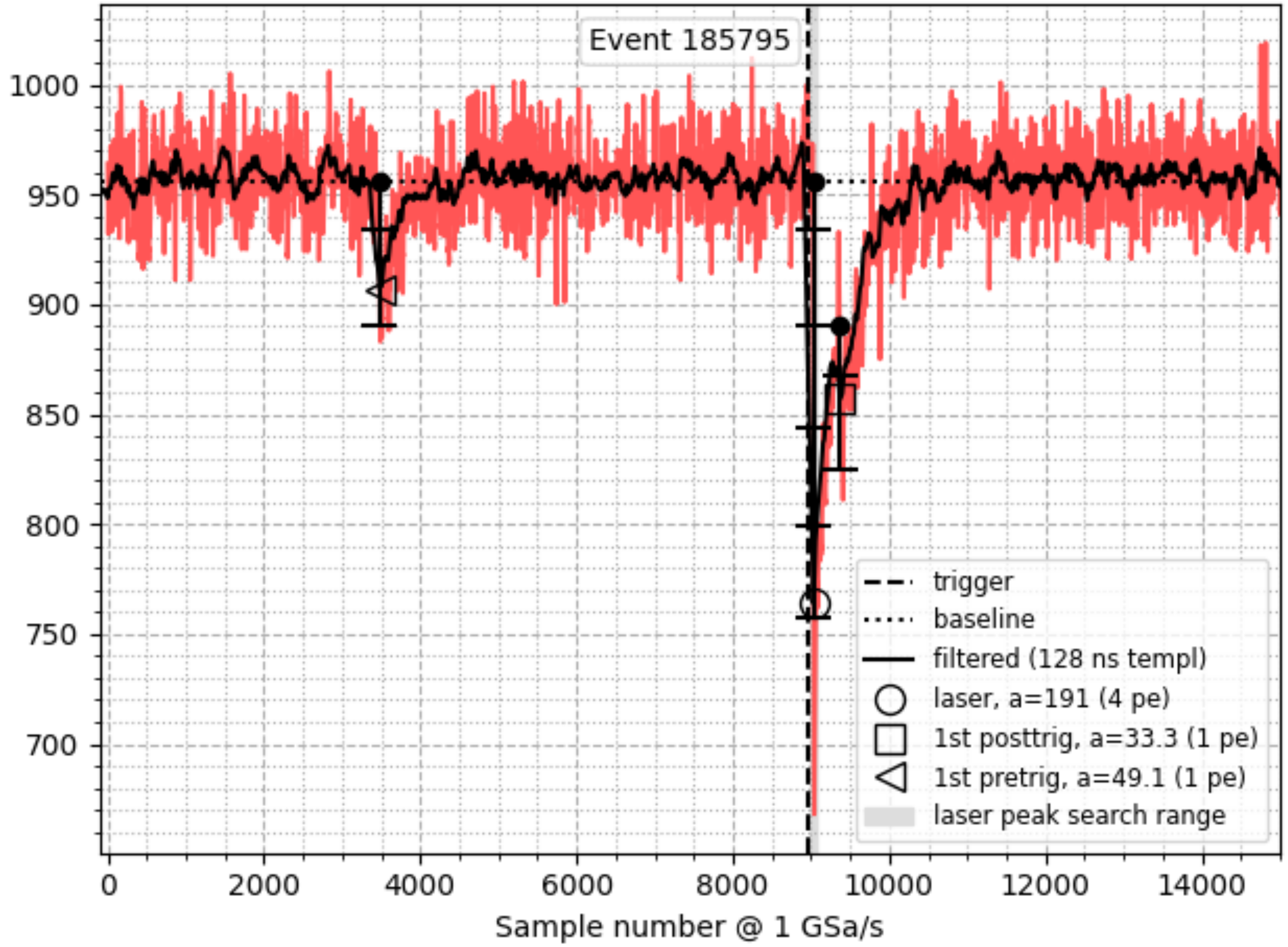
The leftmost and rightmost events in the previous histogram. They are genuinely extreme baselines.

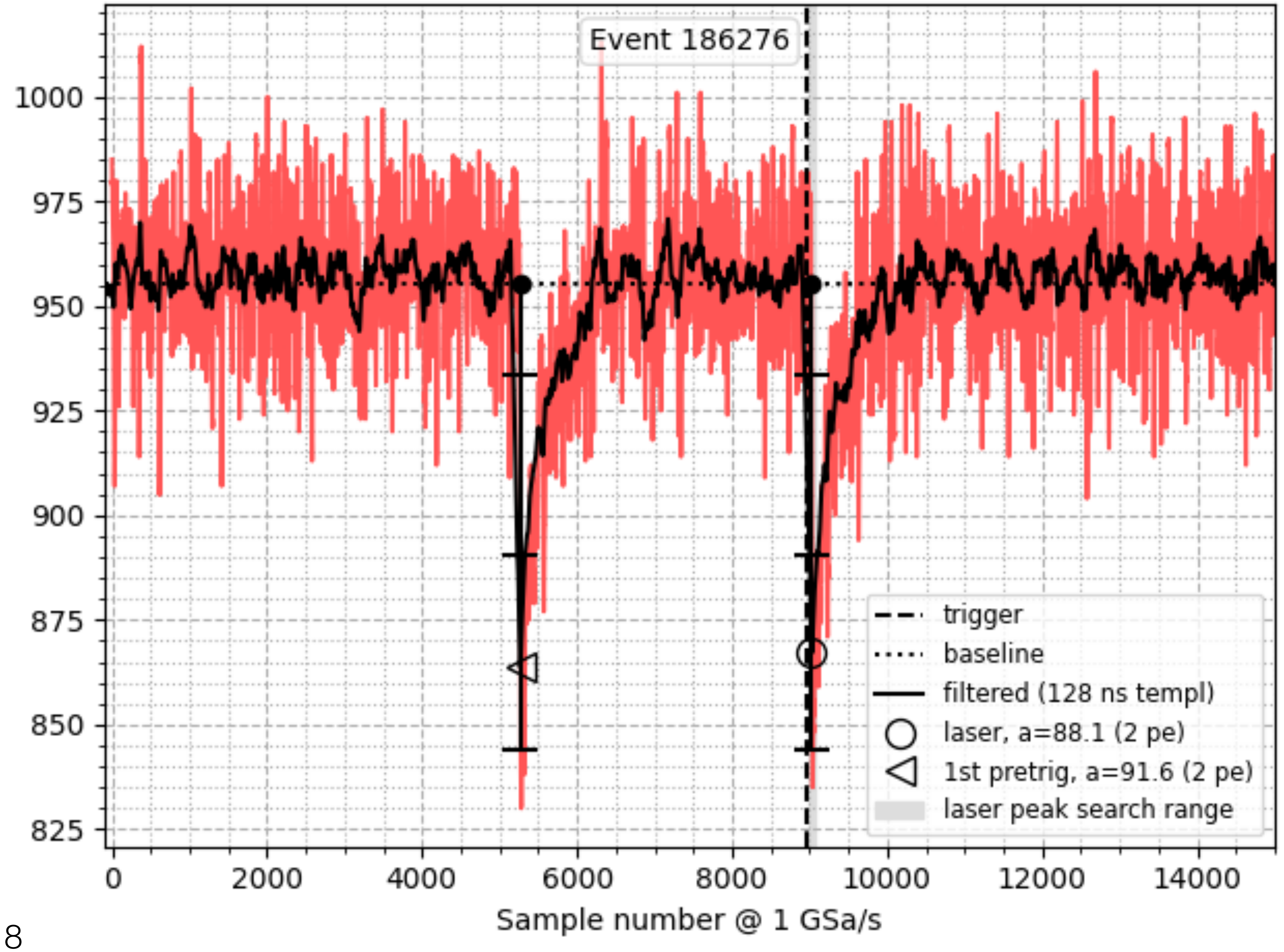


# Baseline (4/4)

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. We will see this is not a problem so I did not fix it.



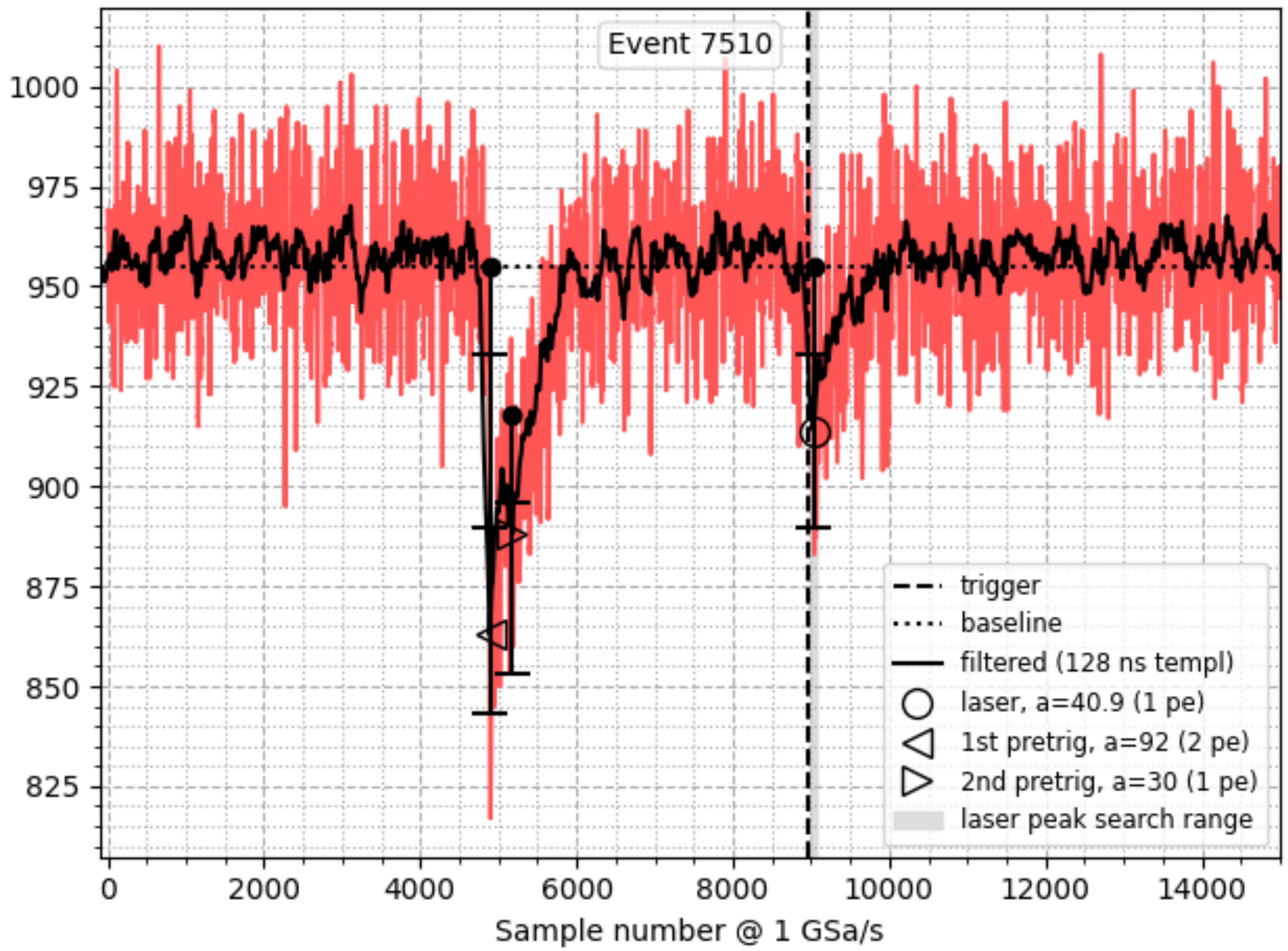
# Peak finding (1/2)

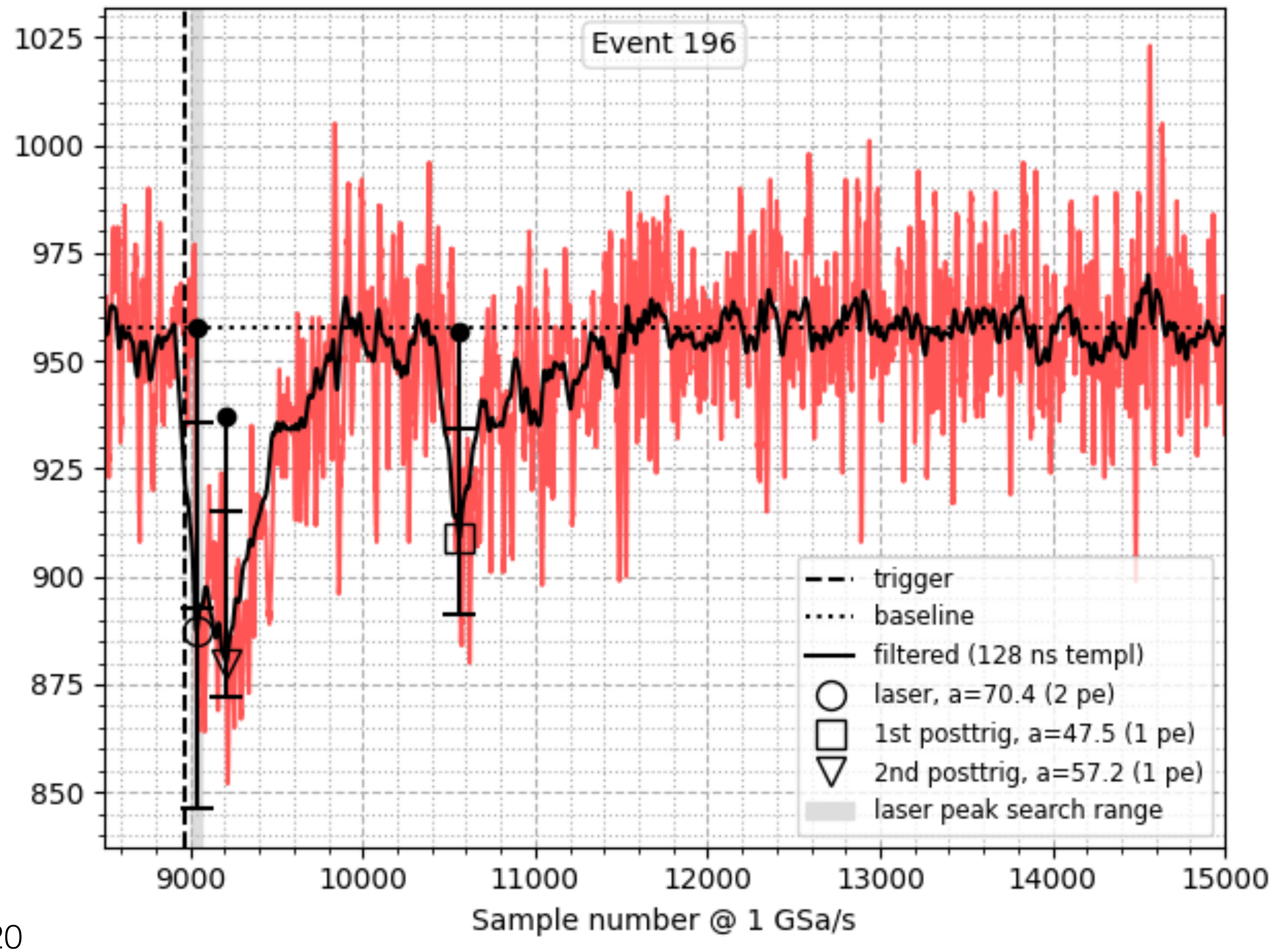
For non-laser peaks I use a prominence-based peak finder.

- I divide the event in pre- and post-laser regions and search separately.
- For computing the prominence, I cap the (inverted) valleys to the baseline, and I ignore the valley floor if it occurs at the edge of the search region (needed for pulses near the boundary), unless both the left and right floors occur at edges.
- I 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.

# Peak finding (2/2)

Examples.



# Counting pe (1/2)

In this data the SNR is high enough that we can classify the pulses by number of pe neglecting the overlap.

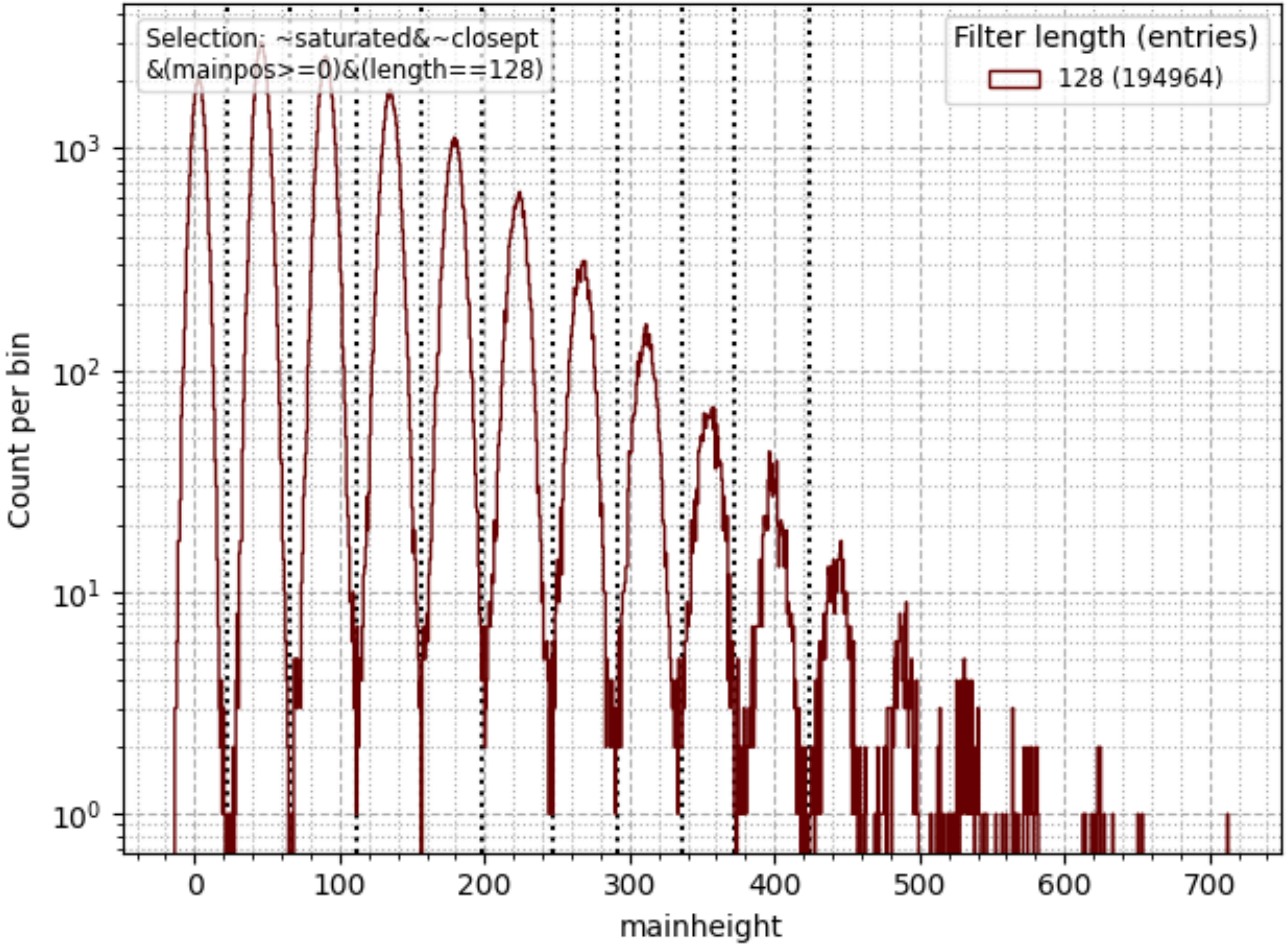
To choose the filter length, there's a compromise between noise (needs long filter) and afterpulses (needs short filter).

We use 128 ns for 5.5 VoV and 64 ns for higher overvoltages.

# Counting pe (2/2)

We select events where the digitizer output is not saturated anywhere and where there's no peak higher than 8 within 2.5 μs before the laser peak.

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



# Peak amplitude (1/2)

Close peaks influence each other's peak height in the filtered waveform. We make these approximations:

A. The position of the peak is not changed by the presence of other peaks (because the signal has a sharp peak).

B. All the information on the signal amplitude is contained in the peak height (because the matched filter has the highest SNR, but we are not correcting the noise spectrum).

So if $y_i$ is the filtered signal waveform, $t_j$ and $h_j$ the position and height of peak $j$, $a_j$ the unknown amplitude of the 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. The $y$ is normalized to have peak height 1 such that $a$ has the same units of $h$.

# Peak amplitude (2/2)

We have to decide wether to input a peak in the equations, because 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.

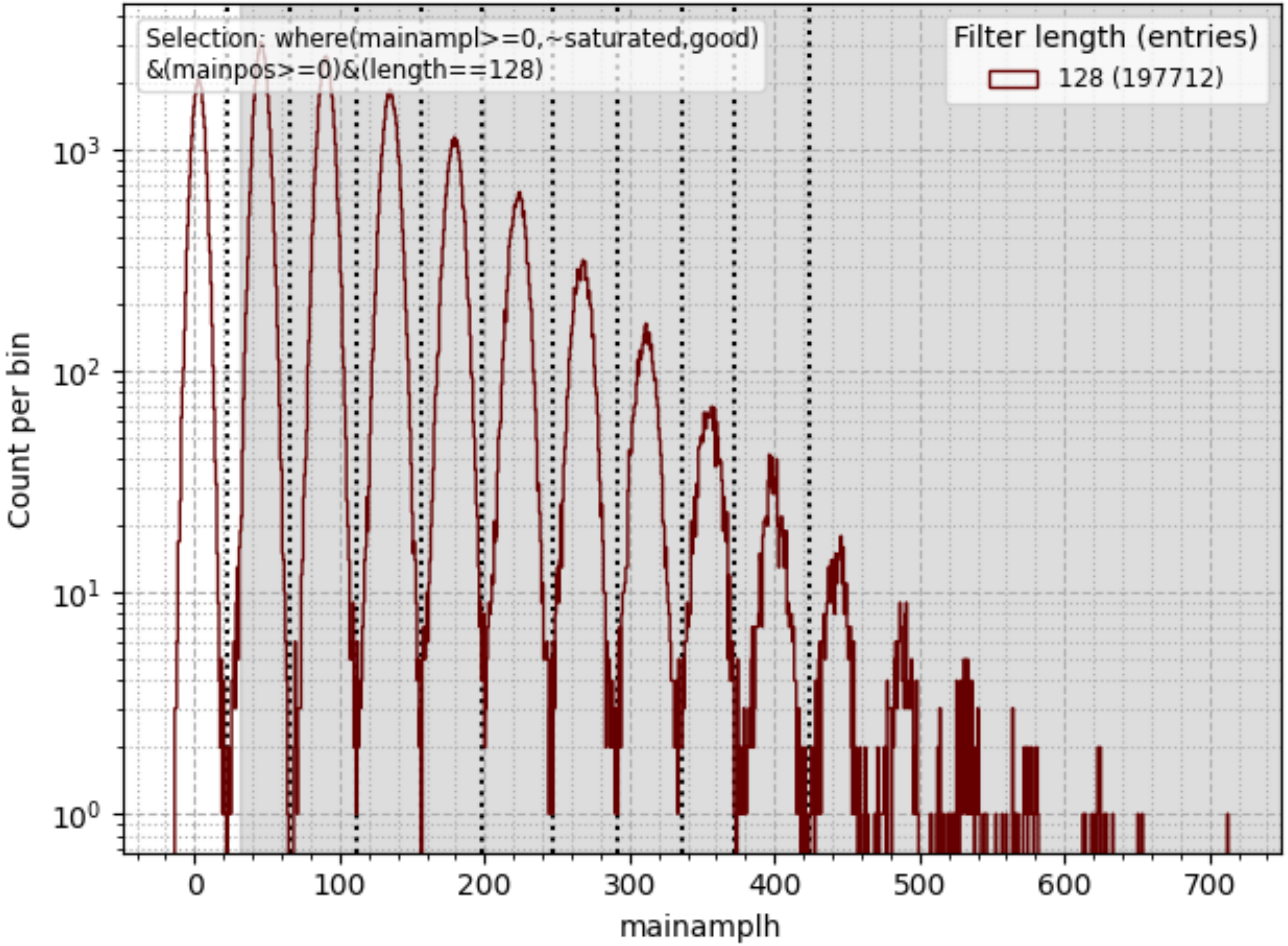
This does not miss afterpulses because we cut on the absolute height relative to the baseline, not the prominence.

# Counting pe again

We redo the fingerplot with the amplitude instead of the height. This time we exclude only saturated events.

The pe bins turn out to be practically the same.

We use the height instead of the amplitude for peaks that were ignored to have the 0 pe peak.



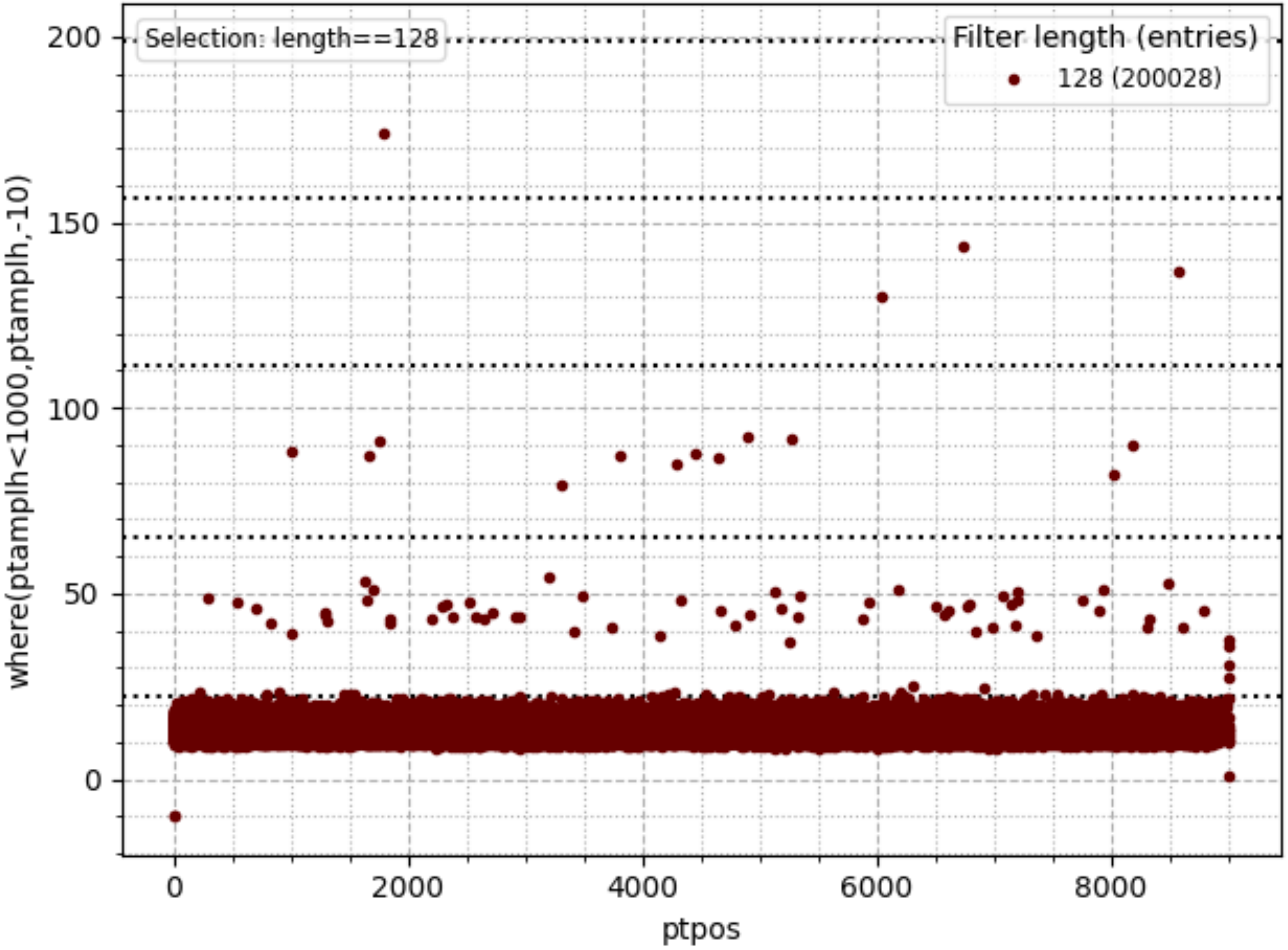
# Dark count rate (1/6)

We need the dark count rate to subtract the background when counting afterpulses.

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

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

We notice that there are boundary effects.

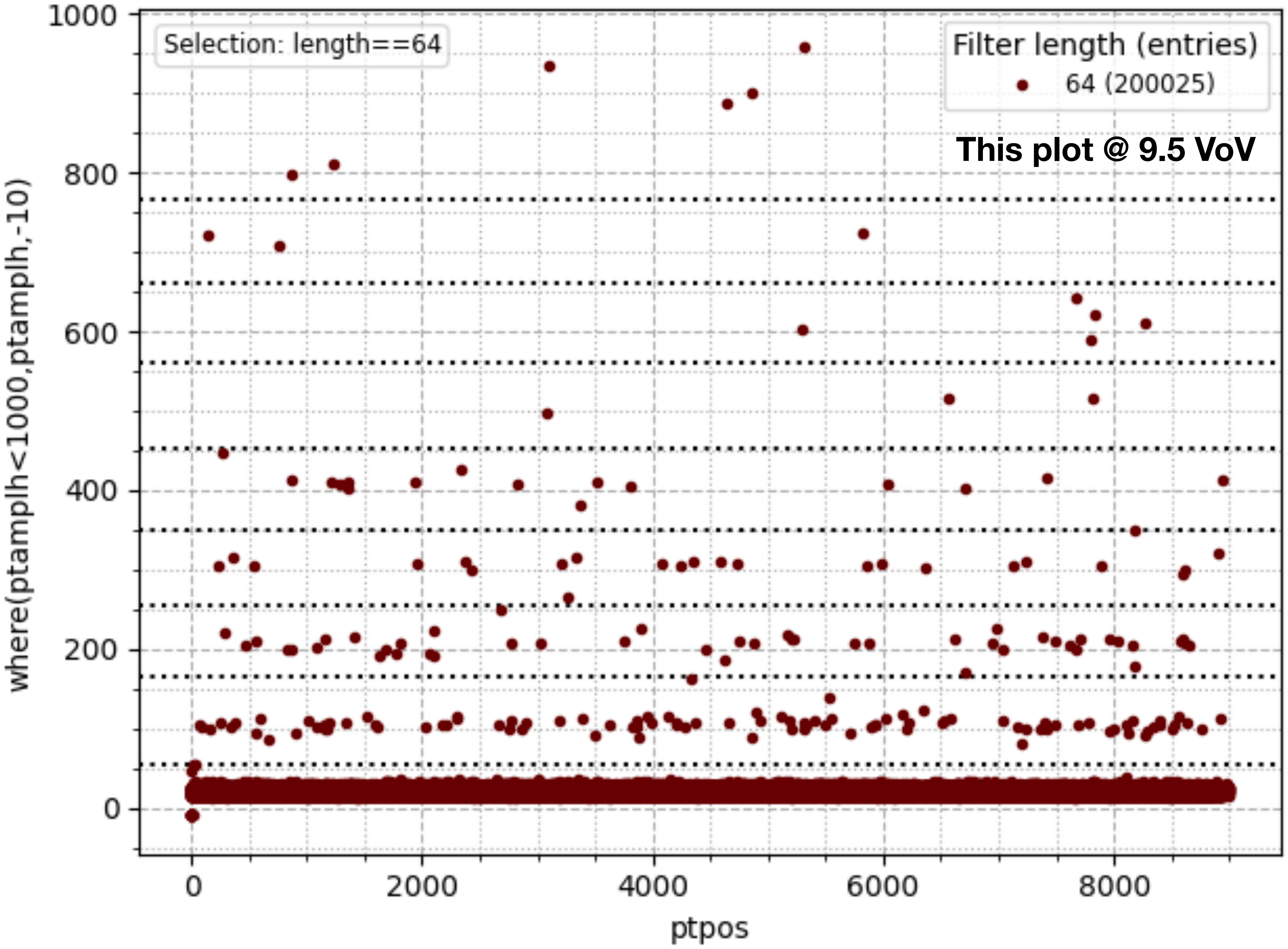


# Dark count rate (2/6)

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 will be when measuring cross-talk.)



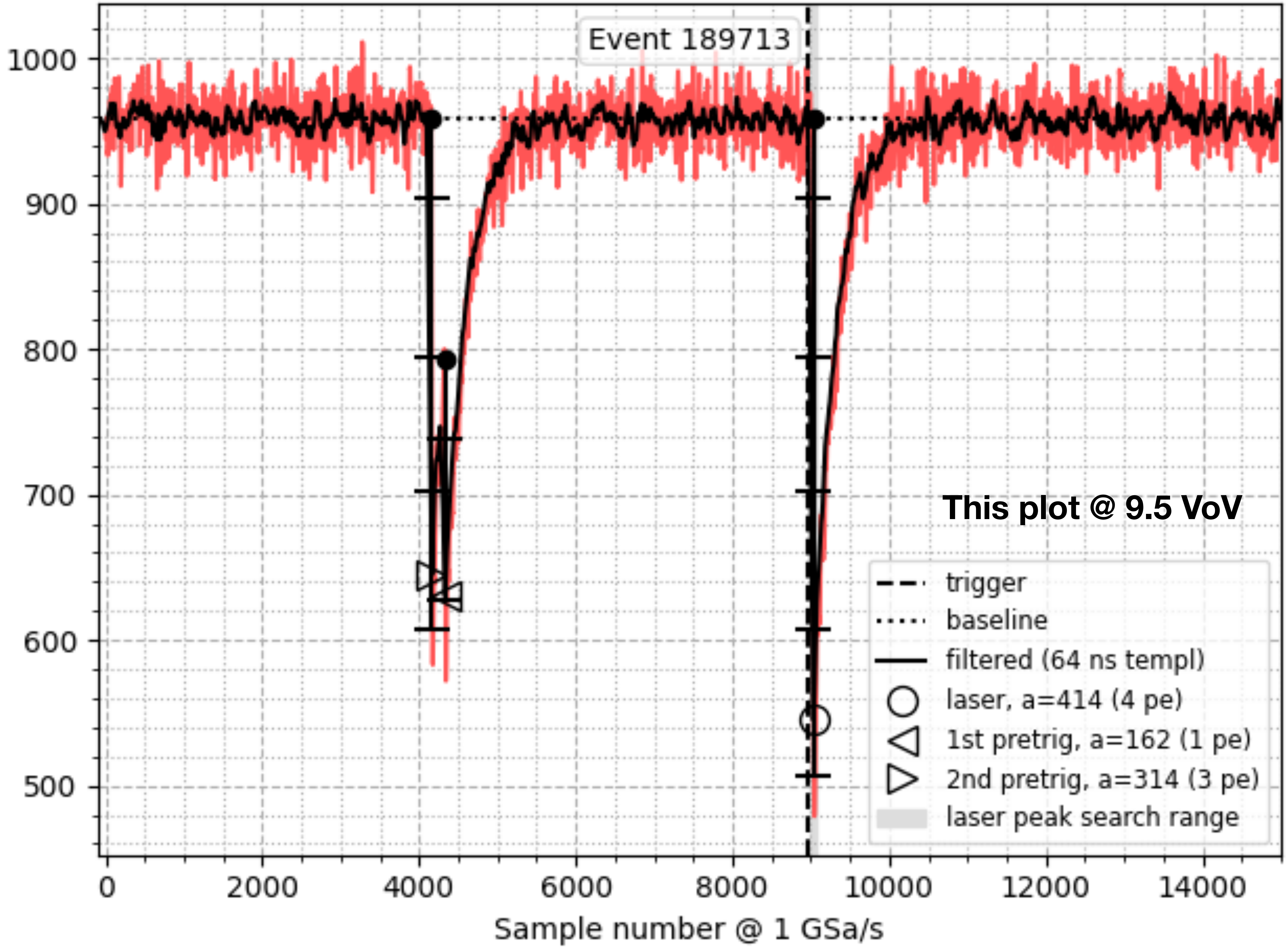
# Dark count rate (3/6)

Looking at these events, we see they are dark signals with afterpulses.

The afterpulse has the highest prominence and height.

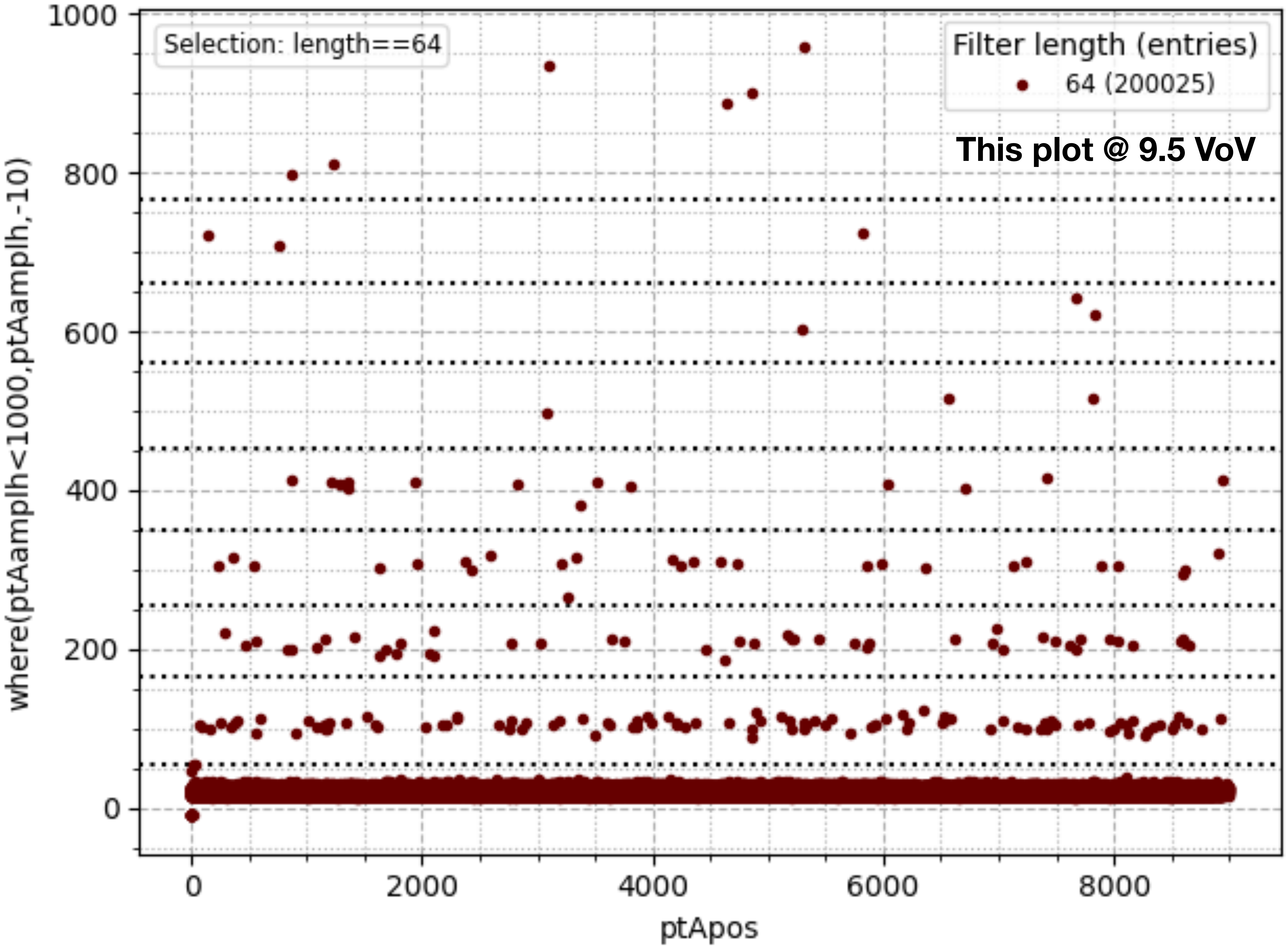
We could select by amplitude instead, but an afterpulse could still be higher.

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



# Dark count rate (4/6)

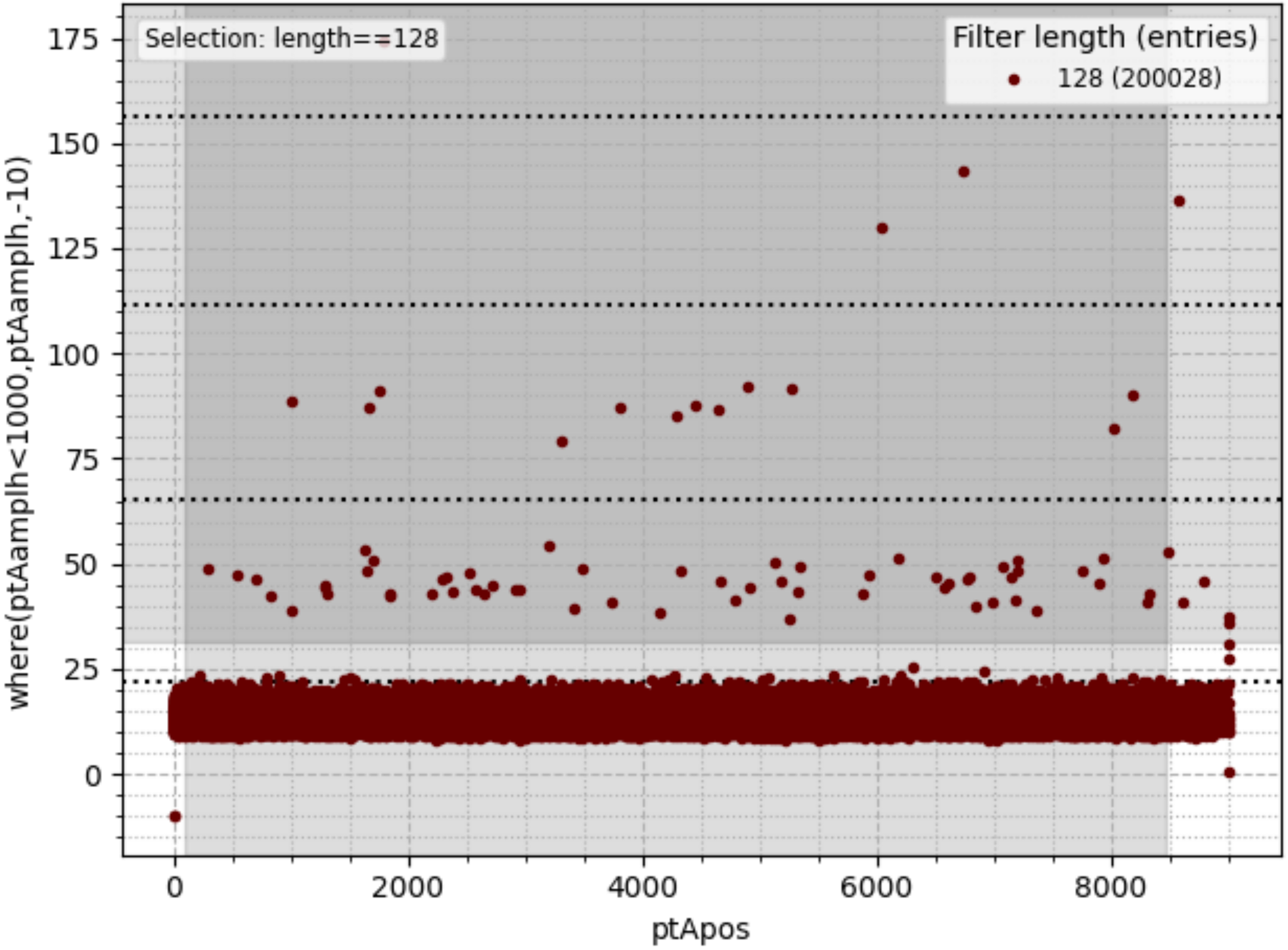
The separation is a bit cleaner now.



# Dark count 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 a bit on the left and 0.5 µs from the right just to be safe since the laser peak could have an effect.

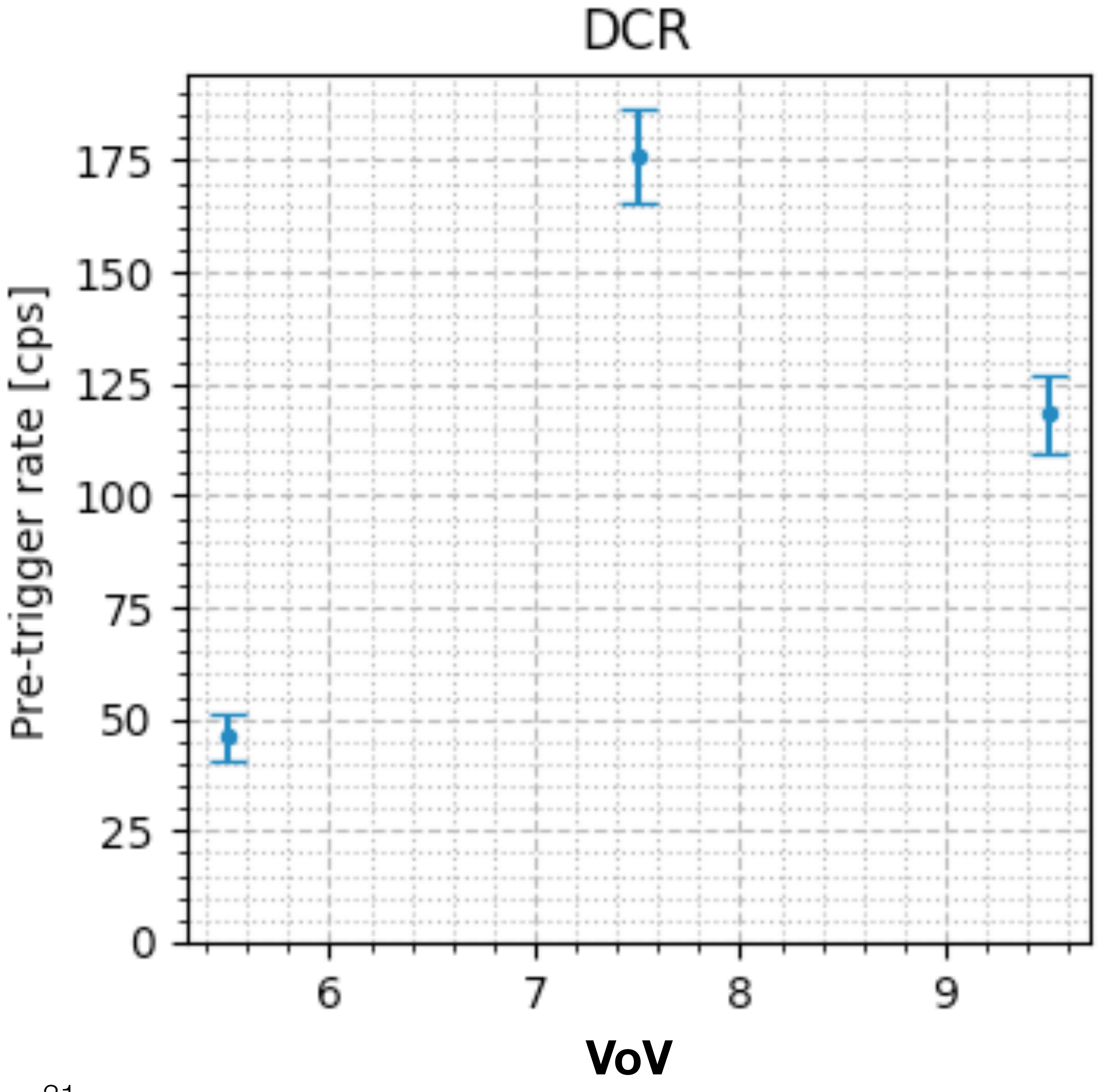


# Dark count rate (6/6)

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

The higher DCR at 7.5 VoV is highly unlikely, so we suppose the setup is not light-tight, **take these as upper bounds**.

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



# Afterpulses (1/6)

We first **select 1 pe laser pulses**, since the afterpulse probability should depend on the multiplicity.

Similarly as we did for the dark 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 it's ok.

# Afterpulses (2/6)

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

However, their absolute height is such that when added to the tail of a 1 pe pulse they reach a round number of pe, because they use up only and exactly the recovered charge from their cell.

So, to count the afterpulses with straight cuts, we add this quantity to the amplitude.

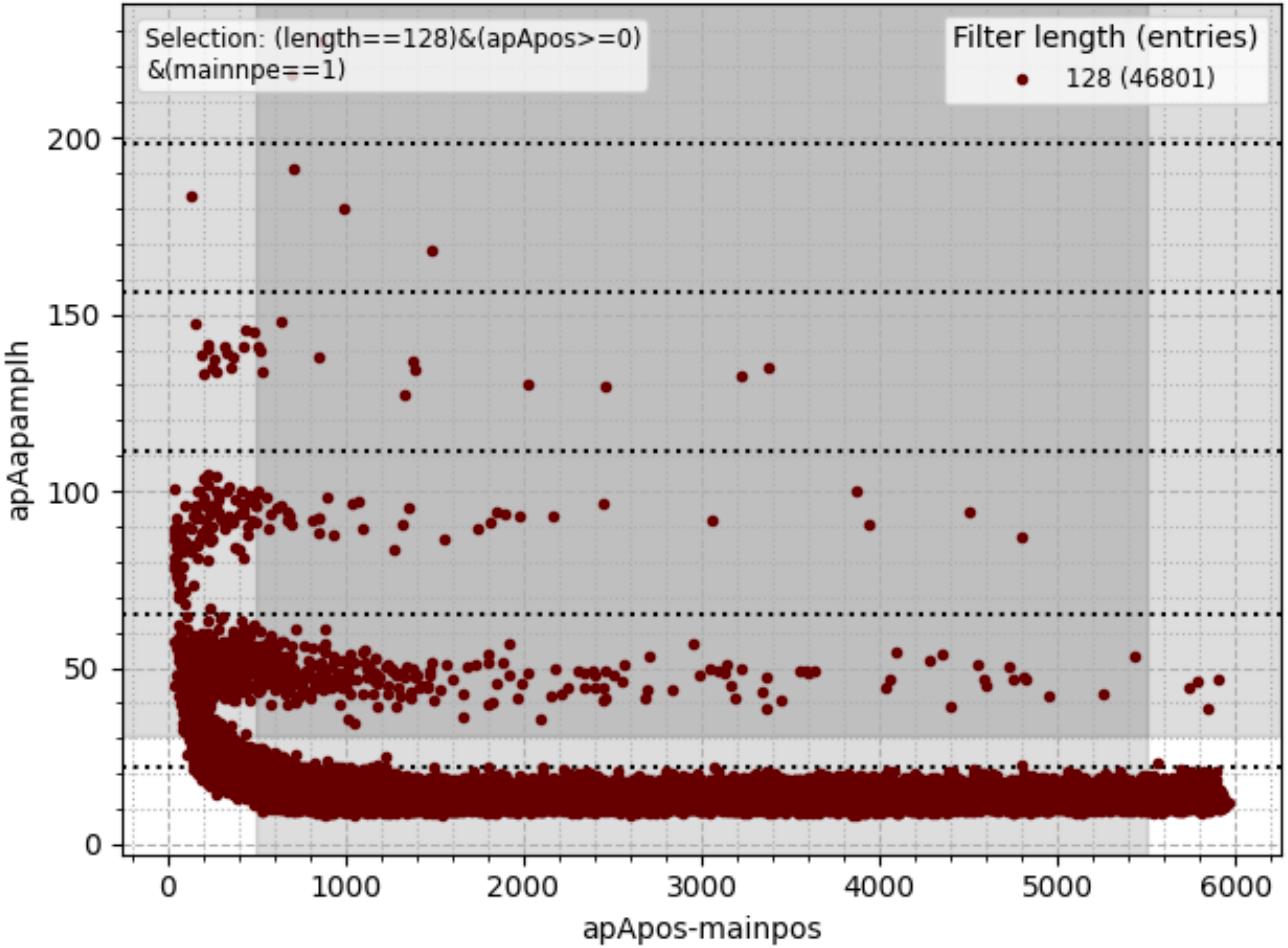
(Note this is different from just taking the height because we first subtract the filtered laser peak effect, and then add back the tail height as if it was just under the afterpulse, without filter smearing. And when selecting higher pe laser peaks the bare height would be higher.)

# Afterpulses (3/6)

Scatter plot corrected amplitude vs. position (relative to laser peak) of post-trigger pulses for 1 pe laser peaks.

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

We could still get a clean cut going closer to the laser peak, however we are afraid the distribution would be modeled by peak finding effects.

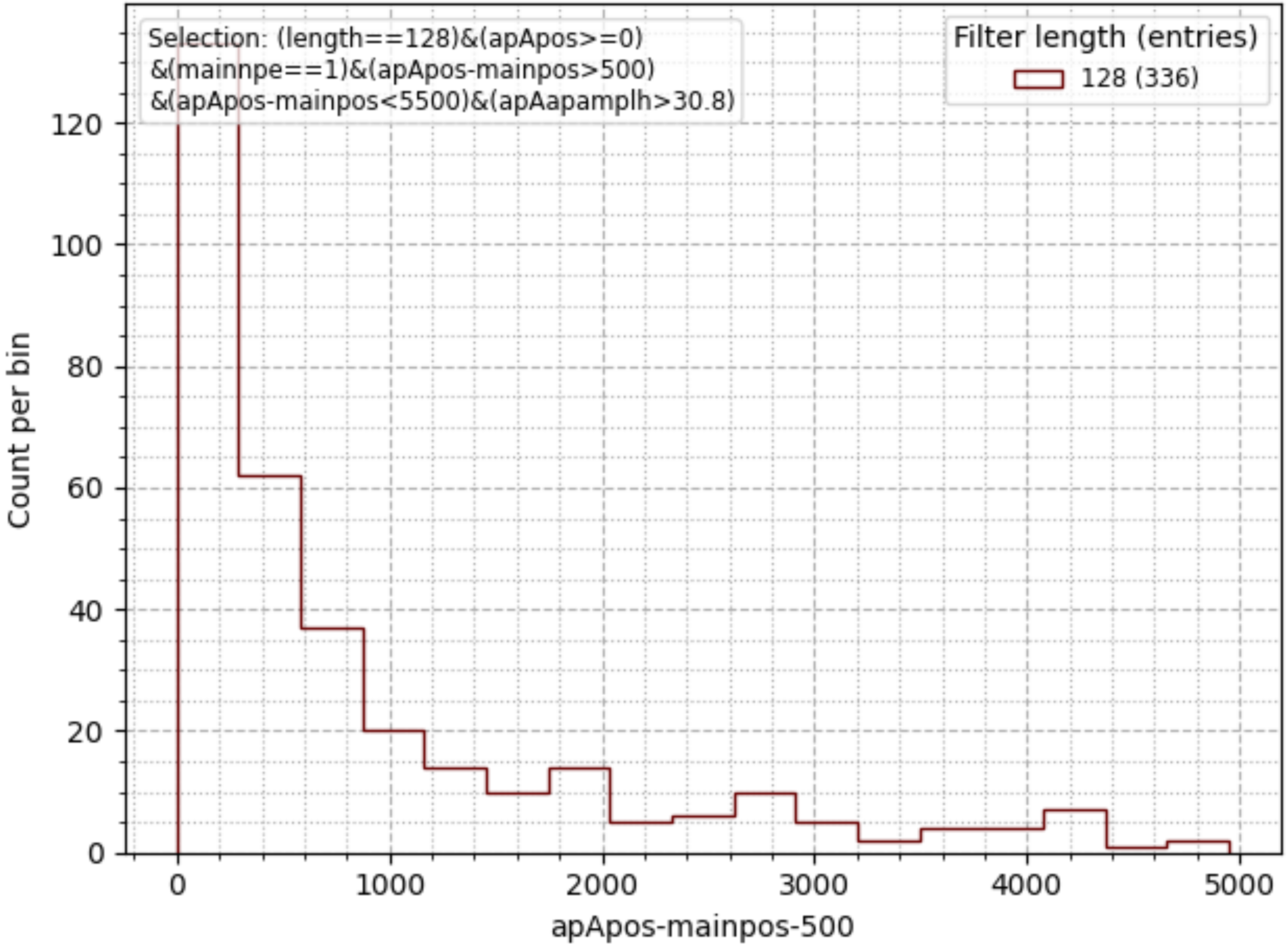


# Afterpulses (4/6)

Since we are cutting away afterpulses close to the main peak, we need to fit the distribution and add them statistically.

This is the delay histogram of selected afterpulses (starting 500 ns after the laser peak).

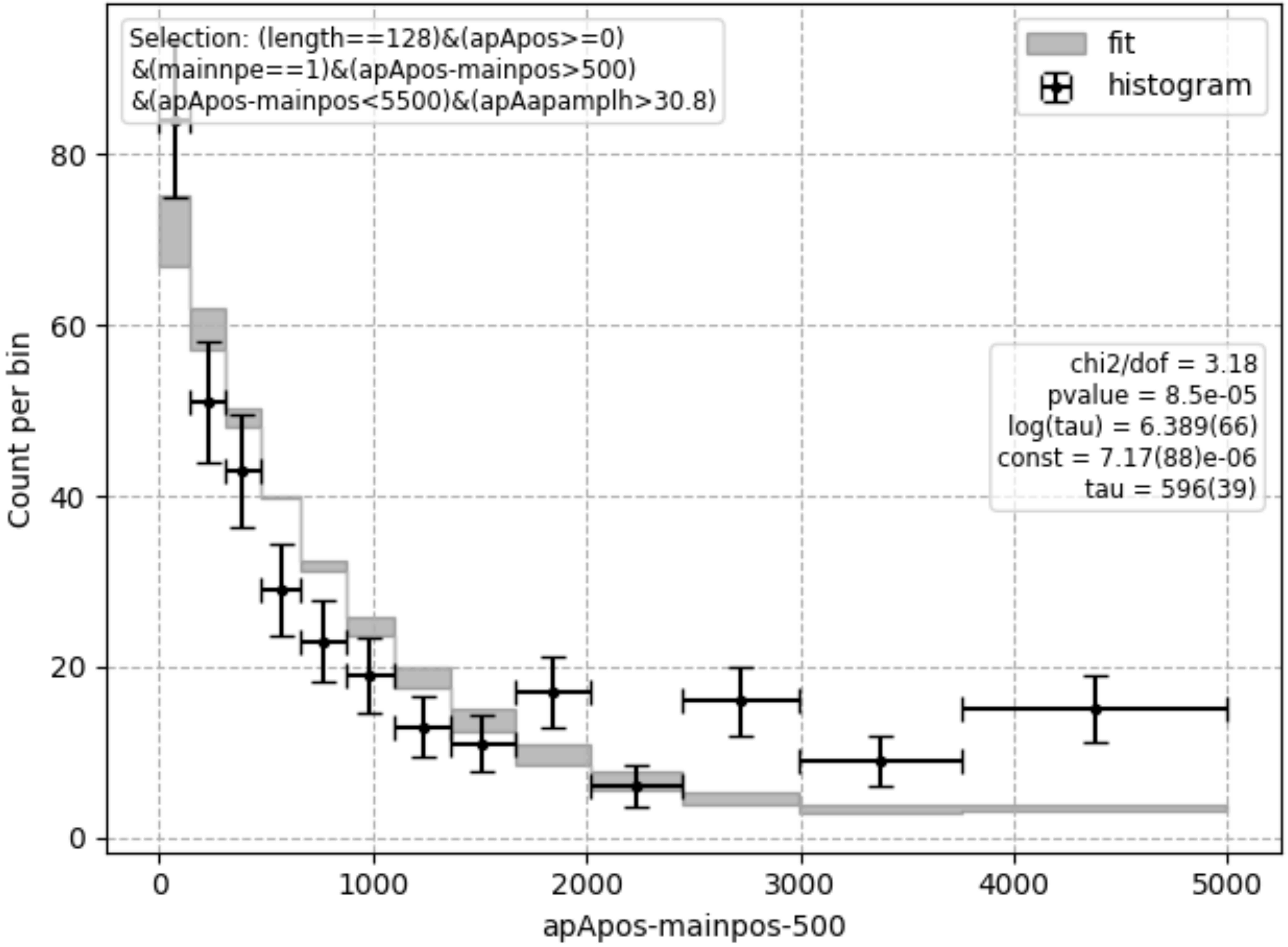
We fit an exponential plus uniform background, where the background is computed from the dark count rate.



# Afterpulses (5/6)

The fit is not very good. For other overvoltages the shape of the residuals is similar: up, down, up.

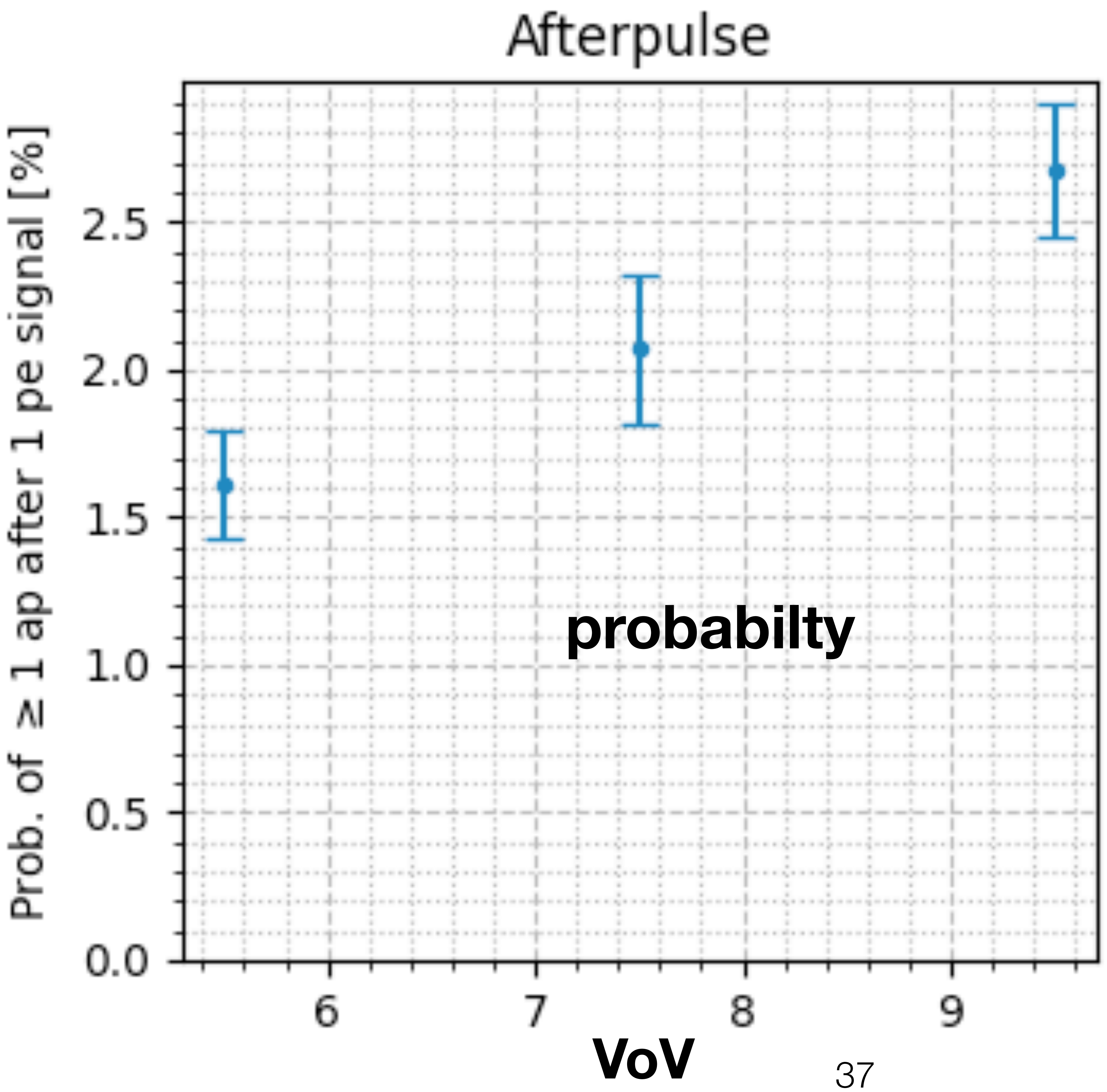
We rescale the tau error with sqrt(chi2/dof) to add a systematic.

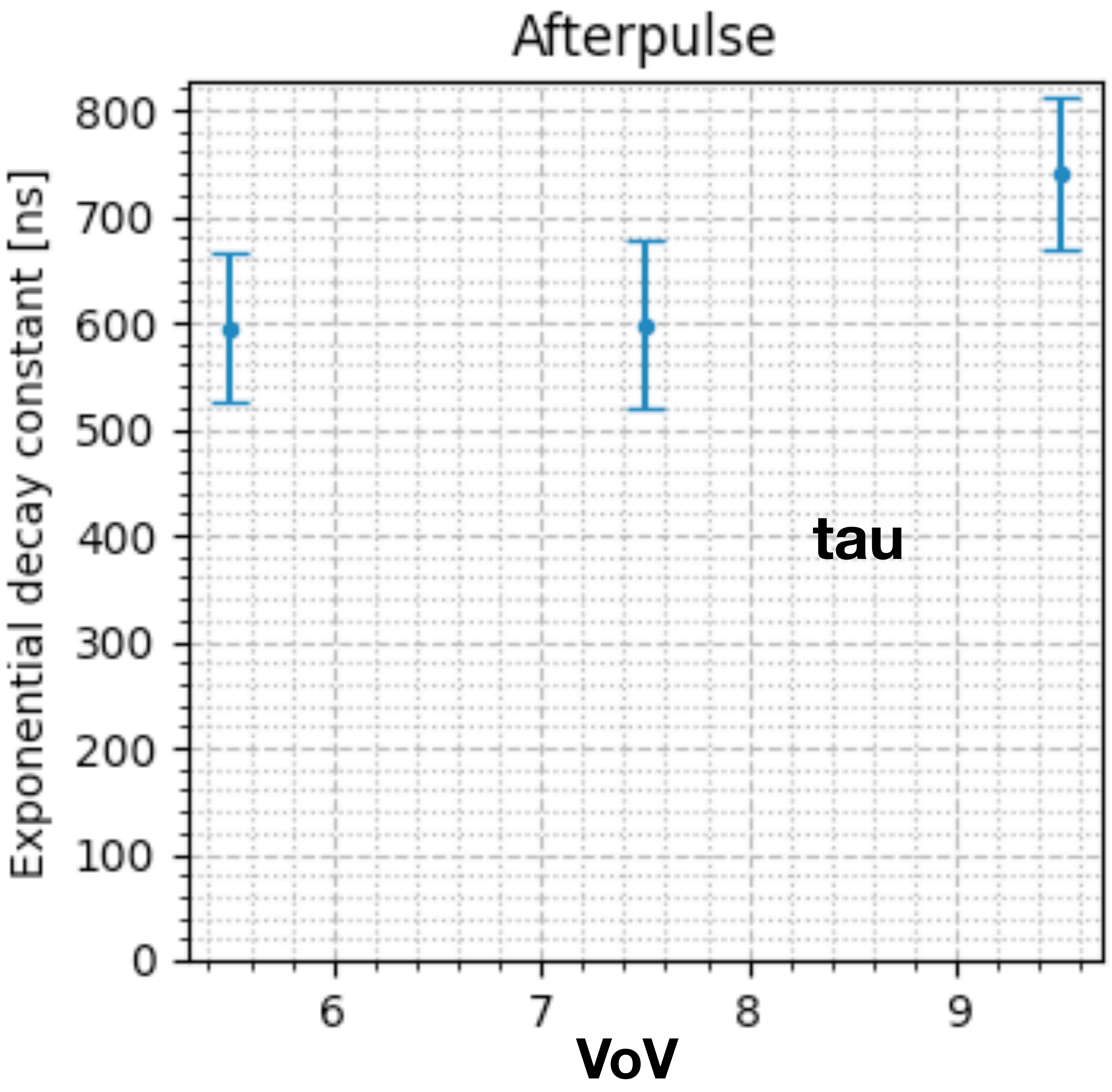


# Afterpulses (6/6)

Anyway, correcting with the DCR and tau, we obtain this. Thing missing that should be done: counting afterpulses for higher laser pe.

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 $\mu_B$. The total number of pulses (root + descendants) is Borel distributed:

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

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

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

References: arXiv 1109.2014, 1609.01181.

# Cross talk in dark count

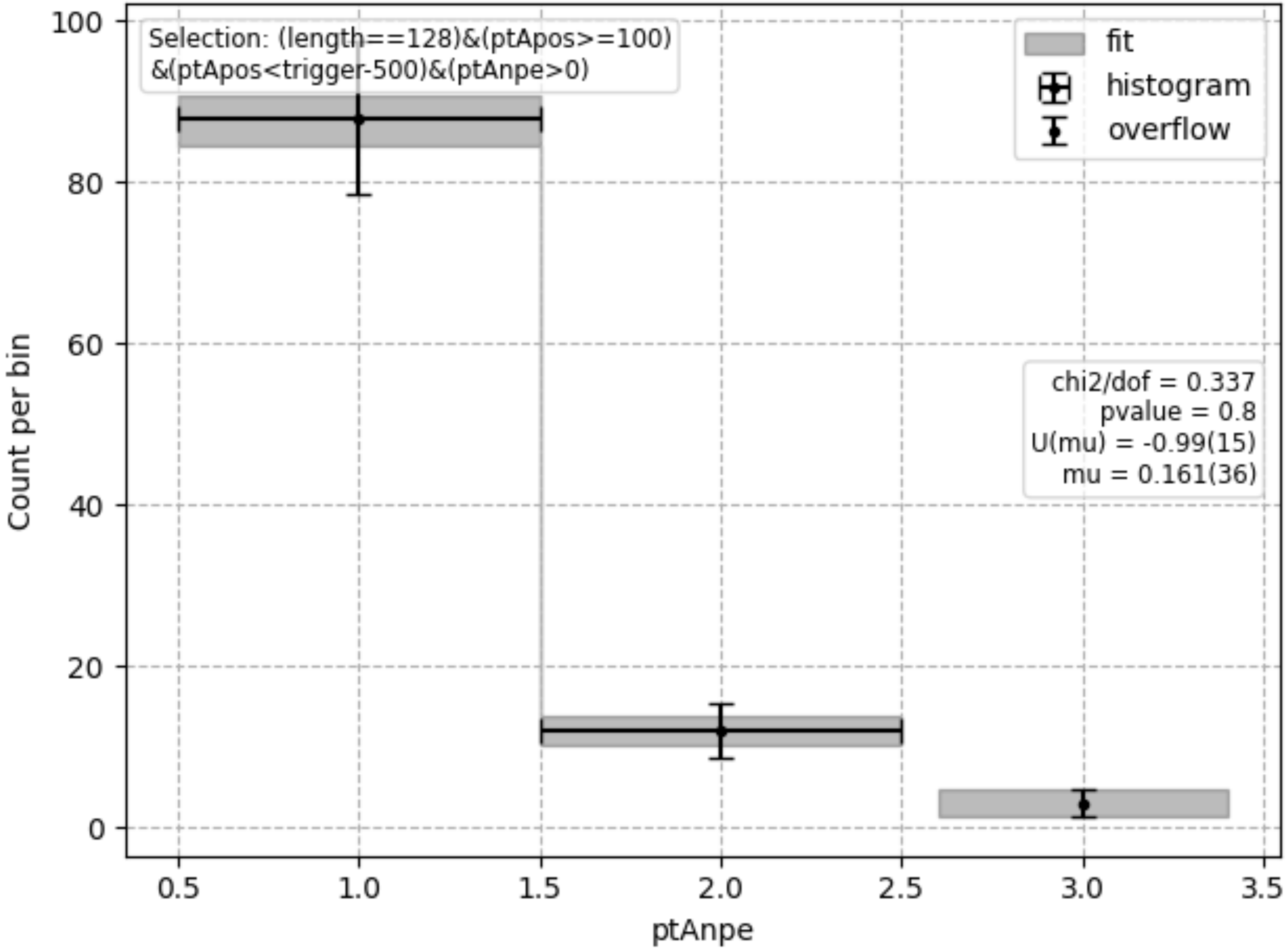
We count the pe of the pre-trigger pulses selected before, using the pe boundaries.

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.

(If you look closely at the fit, the chisquare may seem suspect. There's an additional squared term not plotted to put a uniform prior on μ_B.)



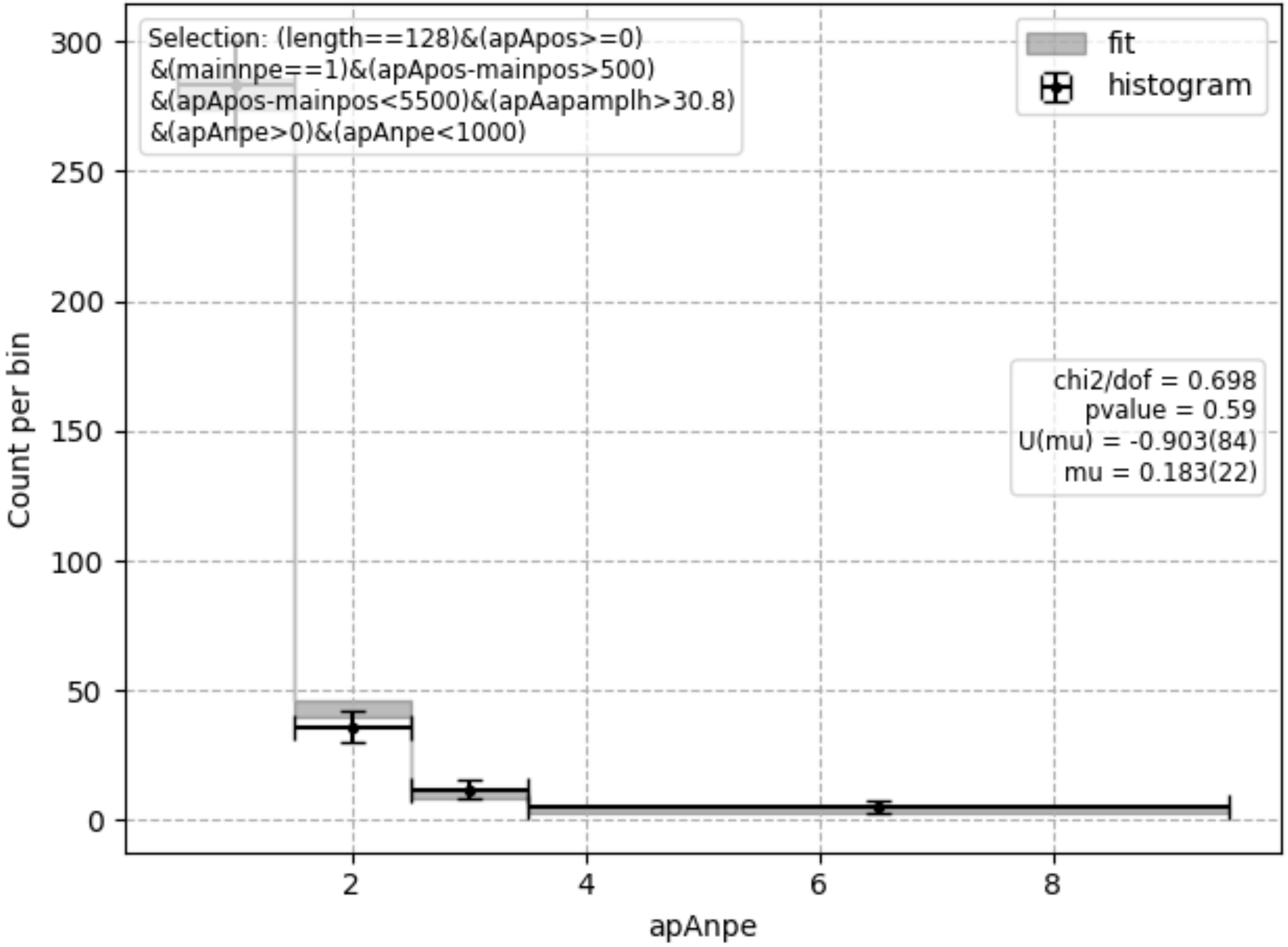
# Cross talk in afterpulses

We redo it similarly for 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 (1/2)

We said before that if there is no local minimum in the laser peak search range, we do not "find" it. Turns out about 1.5 % of events misses a laser peak for a fixed filter length.

These are mostly 0 pe, however there's also some contamination from very high afterpulses that "flatten" the main peak.

All this is highly sensitive to filter length, so for those events we use the shortest filter that finds a laser peak.

Only 0.1 % of events do not satisfy even this requirement, and we ignore them.

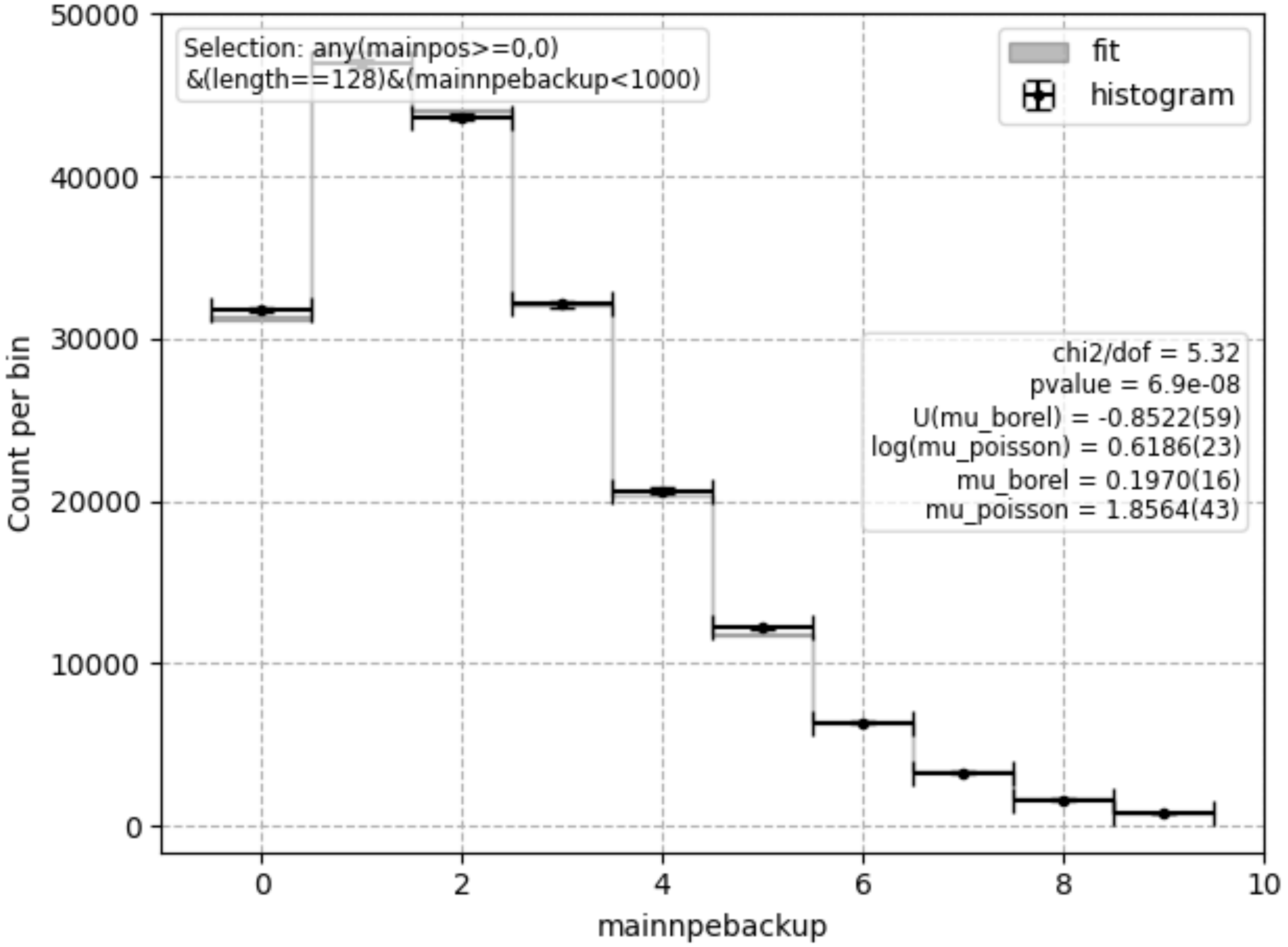
# Cross talk in laser pulses (2/2)

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, so whatever.

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

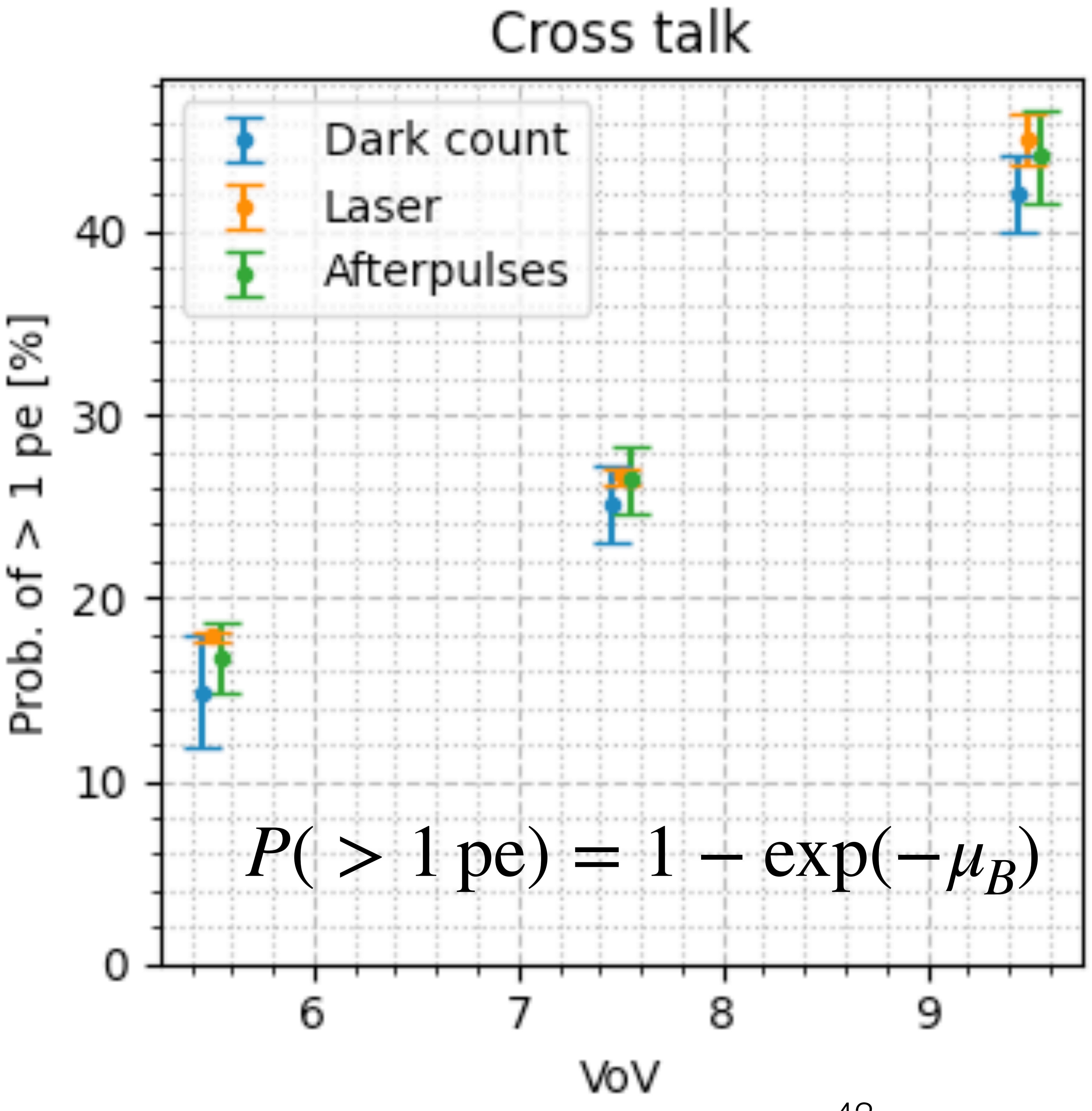


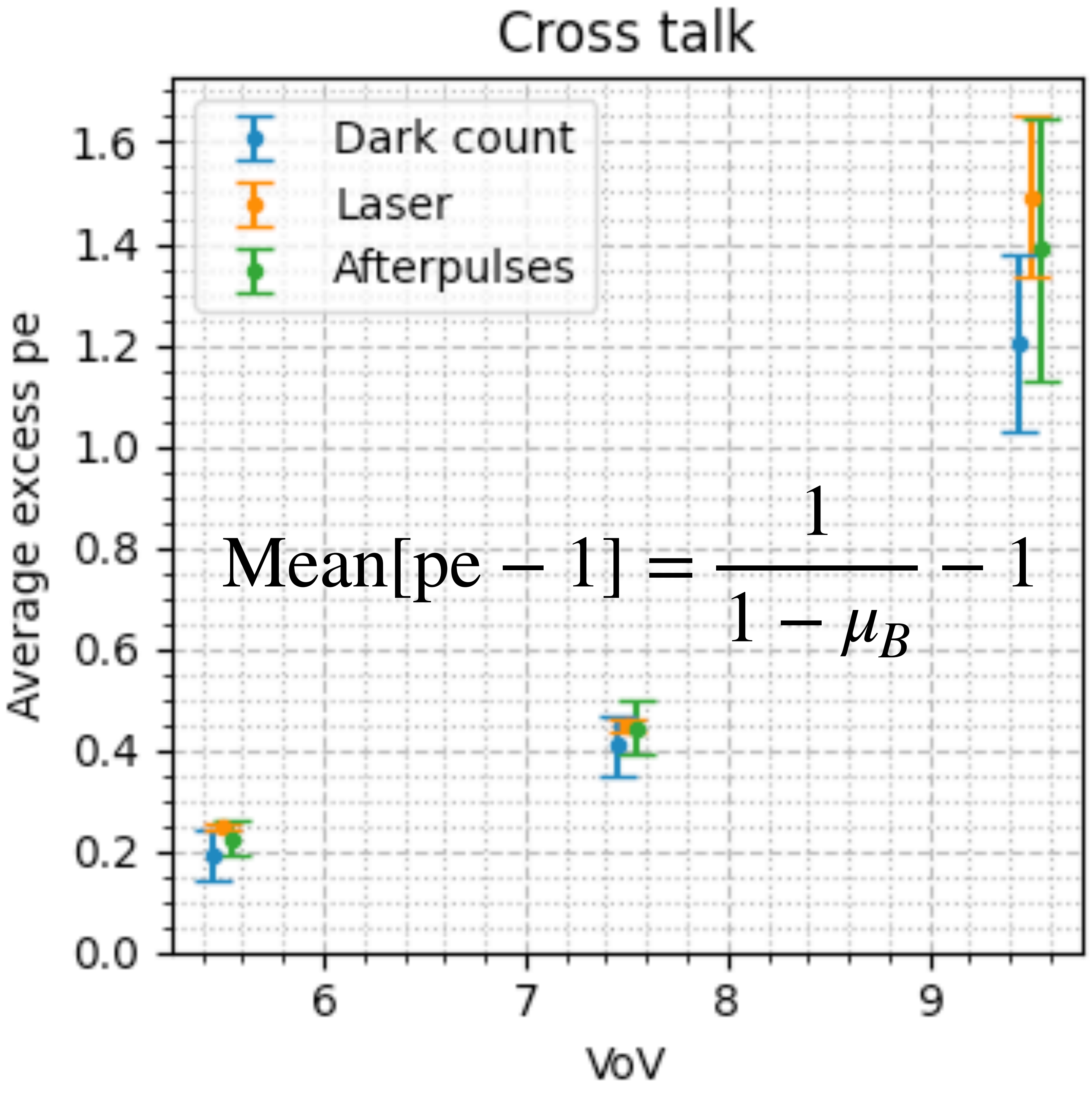
# Cross talk results (1/2)

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 methods agree.



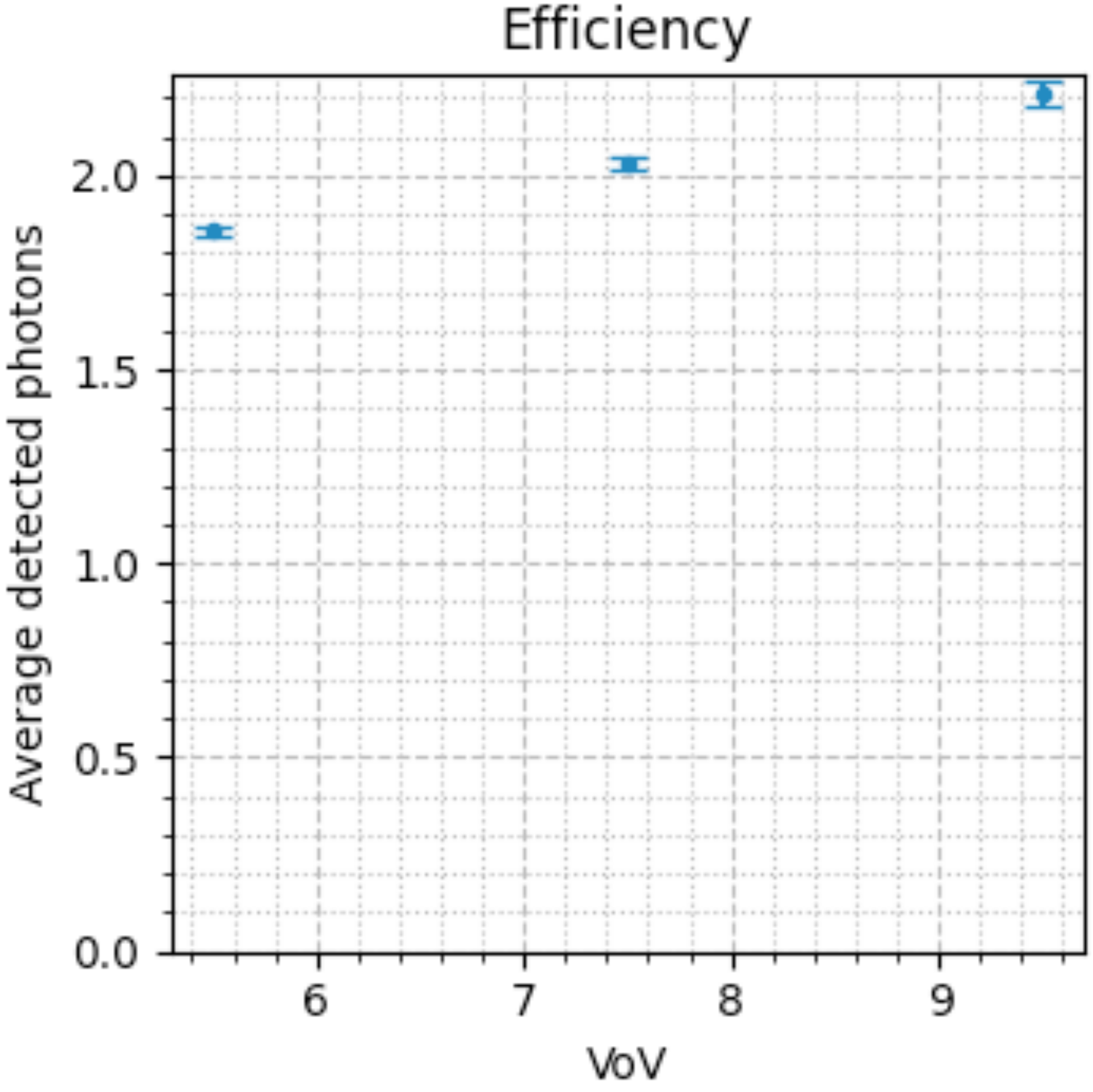



# Cross talk results (2/2)

The mean laser pe before cross talk, $\mu_P$.

It is proportional to the photodetection efficiency.

Increasing, as expected.



# Conclusions

We have analyzed correlated noises using peak finding.

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

The afterpulse decay distribution appears different from the expected exponential, with a tau shorter than what I had heard (1 μs).

This needs to be investigated (may be a problem in my analysis).

If it is not a bug, 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?