

# BACoN analysis

Carmen Romo Luque

March 13, 2024

## 1 Introduction

The BACoN detector is designed to investigate the scintillation properties of liquid argon when doped with xenon. The setup consists of a 100-liter cylindrical cryostat filled with liquid argon, featuring a single upward-facing PMT mounted at the bottom and four rows of three equidistantly spaced SiPMs. Each SiPM row is separated by 10 cm from the next one. On top of the cylinder, an Americium source is placed, emitting gammas and alphas that interact with the liquid, producing scintillation light. The first row of SiPMs is positioned close to the decay source and serves as the trigger system when all of them detect at least one photoelectron. These three SiPMs are collimated to prevent potential saturation effects.

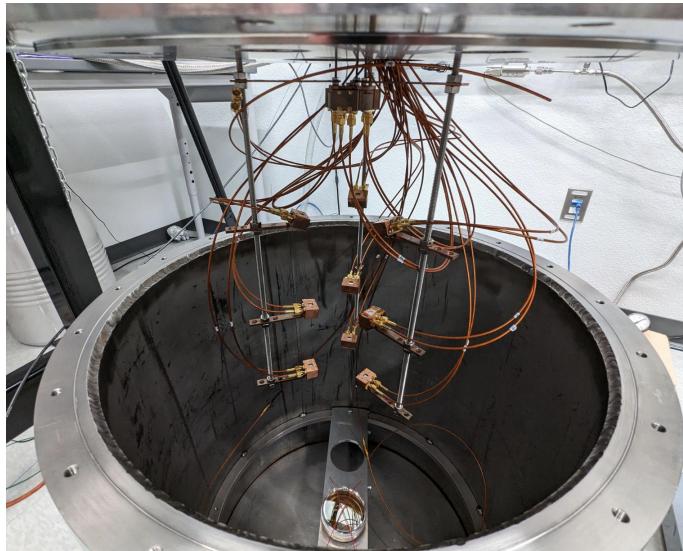


Figure 1: New setup of the BACoN detector. The PMT is placed at the bottom of the chamber and 4 rows of SiPMs arranged around 10 cm apart, facing the americium source.

The cryostat's liquid argon level is continuously monitored by measuring its weight. Additionally, the argon undergoes purification via a SAES PS4-MT3/15-R getter, which removes impurities such as H<sub>2</sub>O, CO, CO<sub>2</sub>, N<sub>2</sub>, H<sub>2</sub>, CH<sub>4</sub>, reducing their concentrations to less than 1 part per billion (PPB) prior to liquefaction.

## 2 SiPM calibration in a dark box

Maintaining the optimal performance of sensors in the BACoN system is crucial for accurate data collection, enabling the detection of variations in light levels within the detector. For

this, periodic calibrations of the SiPMs are conducted to verify their proper functioning and stability. Despite employing the same sensor model across all SiPMs, variations in analog-to-digital conversion (ADC) counts in response to identical light inputs are inevitable. Hence, regular calibration of the SiPMs at the single photon level is imperative to standardize their responses and assess their long-term stability.

Prior to installation within the chamber, the SiPMs were calibrated in a controlled environment using a dark box illuminated by a pulsed LED. The amplitude of the LED pulse utilized during calibration runs depended on various factors including the LED characteristics, the electronics chain, and the sensor gain. The light intensity was adjusted using an oscilloscope to achieve a few ADCs, guaranteeing operation at the single photoelectron level. Subsequently, pulse height spectra were generated by extracting the pulse heights. Given the low amplitude of pulses at the photoelectron level, they occasionally merge with dark counts, rendering them indistinguishable. Figure 2 presents an illustrative example of a waveform captured during a calibration run.

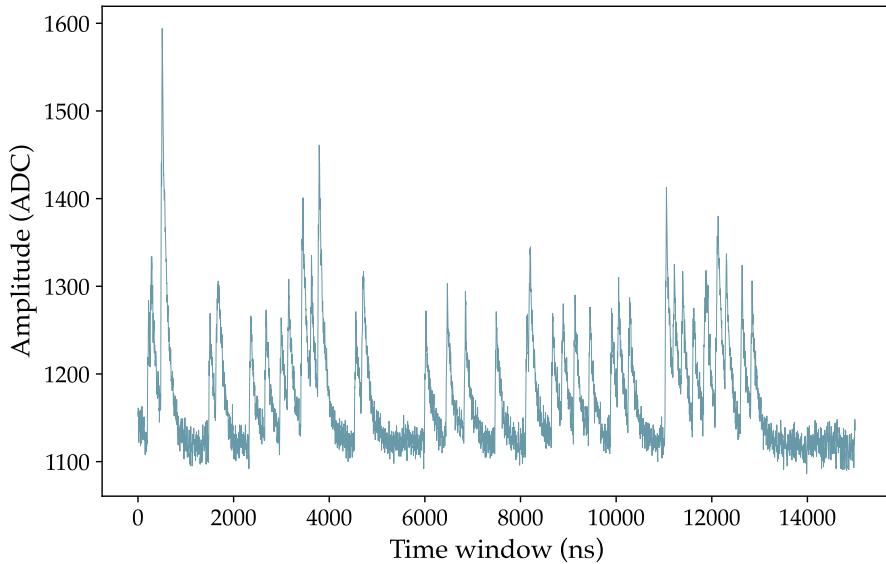


Figure 2: Example of a signal waveform of a SiPM tested in the dark box.

First, we subtract the baseline of the waveform, which allows us to obtain the absolute height of the peaks. The baseline subtraction procedure can be performed by computing either the mean value or the mode of the entire waveform or just a certain range. Figure 3 shows an example of a waveform before and after subtracting the baseline.

Then, we compute the single photon spectra by getting the height of the peaks at the trigger region. Figure 4 shows the sum of all waveforms for one channel with the peak in the trigger region colored in grey.

Subsequently, a histogram is generated displaying the heights of the peaks in the trigger window, allowing for the identification of individual single photoelectron peaks (refer to Figure 5). First, a poisson distribution of a certain number of gaussians is performed in order to extract the mean position of each photoelectron peak. The gain is then calculated from the spectrum using the distance between the pedestal peak and those for 1, 2, ... photoelectrons.

Then, the linear regression of the mean achieved enables the extraction of the conversion factor between ADC and photoelectrons. With these values, the response across all channels can be standardize.

Figure 6 shows the gain resulting from the 12 channels tested in the black box at two different bias voltages.

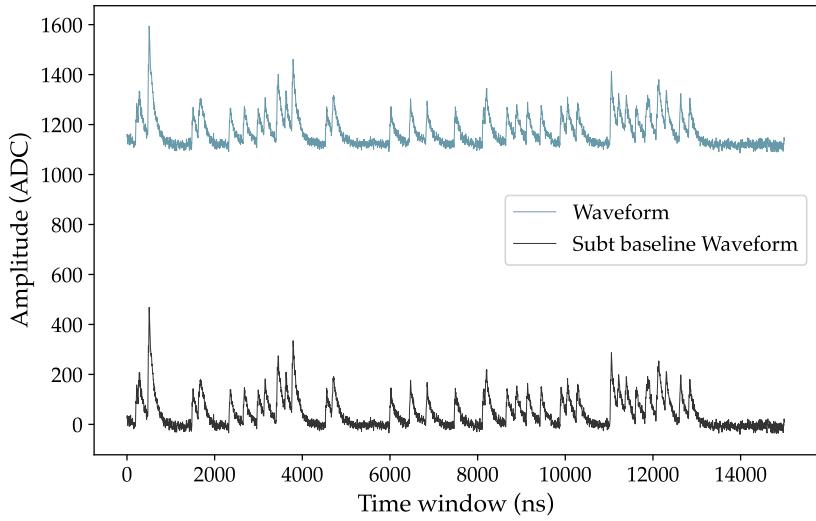


Figure 3: Previous signal waveform and the waveform after subtracting its baseline.

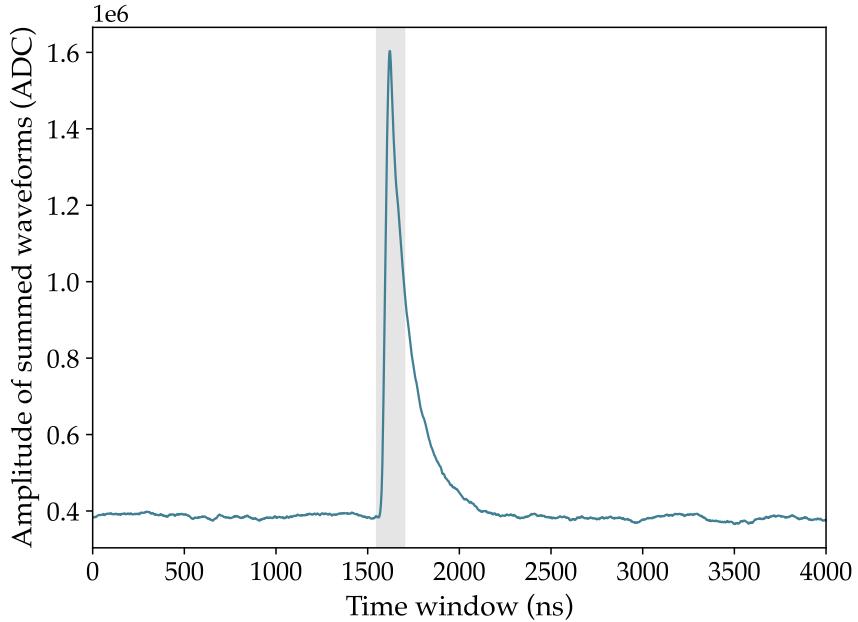


Figure 4: Sum of all subtracted waveforms for one file for a chosen channel. The shadowed range corresponds to the trigger region.

### 3 SiPM calibration in the chamber

Ideally, a chamber should be equipped with LEDs for performing regular calibrations during data collection. That would allow us to monitor gain variations over time at a controlled light intensity, particularly important for long-term data acquisitions lasting several months or few years. Additionally, since we operate at low temperatures, the gains of the SiPMs won't match those obtained during black box calibration. However, in the absence of LEDs, we can still utilize low-light data for similar calibrations, though it's important to note that the values may differ.

Due to the low temperature required to maintain argon in its liquid state within the BACoN

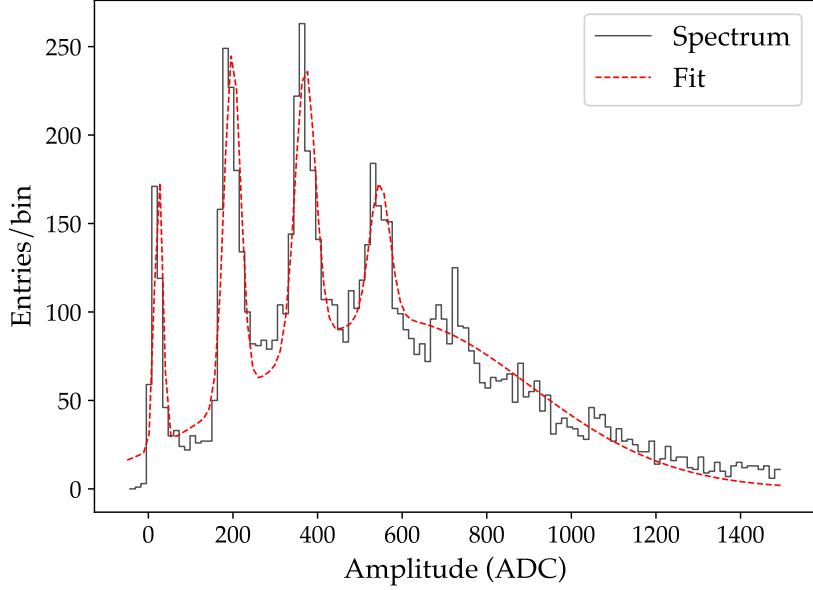


Figure 5: Single pe spectrum for one channel showing the individual peaks. The distribution has been fitted of multiple Gaussian distributions, collectively following a Poisson distribution.

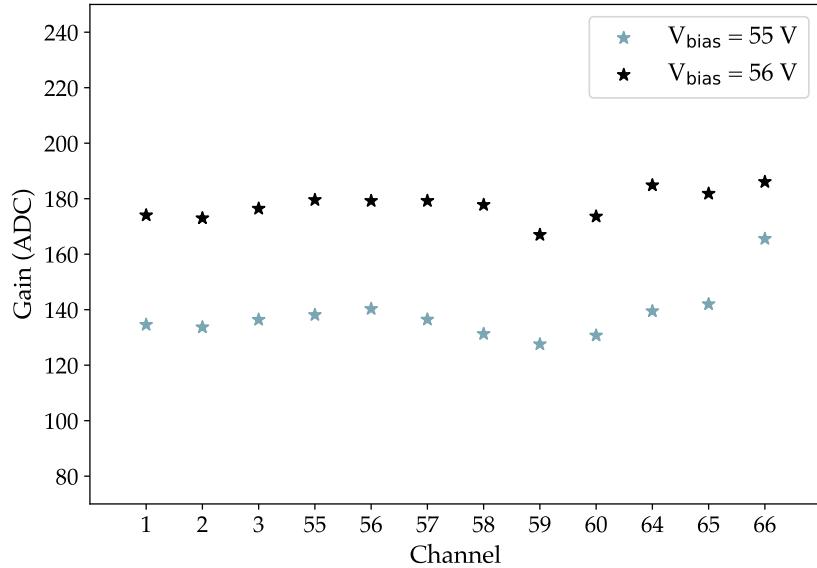


Figure 6: Gain obtained for each channel at 55 V (blue) and 56 V (black) bias voltages.

detector, the bias voltage of the SiPMs needs to be reduced. This compensates for the decrease in breakdown voltage of the silicon diodes within the SiPM structure under reduced thermal conditions. At lower temperatures, the breakdown voltage tends to increase, meaning that the same bias voltage applied at room temperature could potentially cause the SiPMs to operate in a regime where the breakdown voltage is exceeded. This can lead to excess dark count rate and decreased performance. Therefore, reducing the bias voltage when operating SiPMs at lower temperatures helps ensure that they remain within their optimal operating conditions and maintain stable performance. That's the reason why the gains of the SiPMs inside the chamber are higher even at a lower bias voltage, as can be seen in Figure 7.

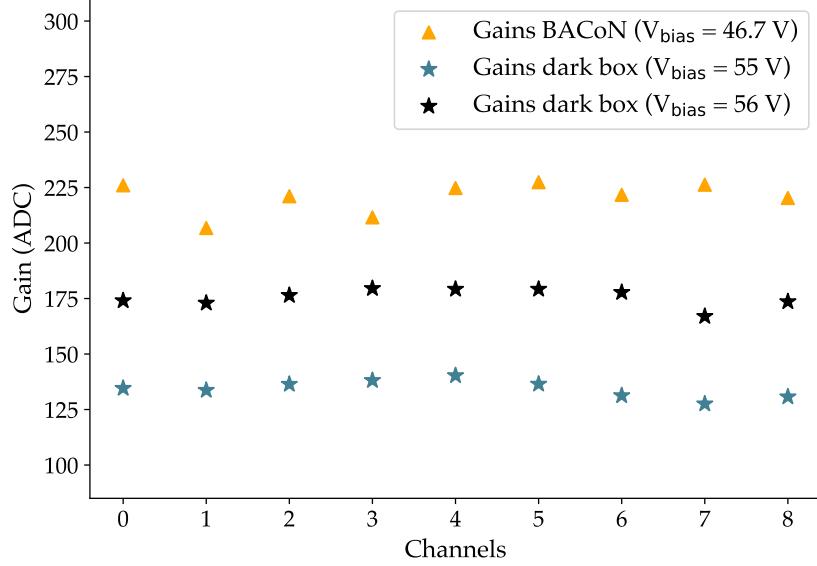


Figure 7: Gain obtained for each normal channel at 55 V (blue) and 56 V (black) bias voltages in the dark box and at 46.7 V inside the BACoN detector.

## 4 Data Run 2 analysis

The second period of data collection in BACoN began in mid-November 2023. Figure 8 shows the waveform for each channel for a particular event. Once the three trigger SiPMs detect one photon each, the event starts, and the rest of the channels (from channel 0 to 8) trigger as well. During the initial days, only argon was inside the chamber, and the gain of the trigger SiPMs was adjusted twice after several days (first on November 21st and then on December 2nd, as shown in Figure 9). The PMT malfunctioned and stopped working, so the data for these runs do not include that information. Additionally, in December 2023, channel 5 of the normal SiPMs started to have issues and was disconnected.

Figure 8 depicts the appearance of the data. In this specific event, besides the signal in the trigger SiPMs, which is required to store the event, other channels detect light, evident in the form of single peaks. Conversely, there are channels where no light is detected, and only the baseline originating from the electronics is visible in the waveform. In our setup, the signal from the trigger SiPMs is inverted due to the specific electronic chain for those channels. Therefore, an additional step in the analysis chain will be required to flip their signal.

### 4.1 First look at the peaks

For the data analysis is crucial to fully understand the shape of data and the peaks, since their amplitud (either height or integral) is related to the photons detected by each sensor. Figure 10 shows different peaks zoomed in using a window of 1500 ns registered by the same channel. In most cases, an isolated single peak (that can correspond to few photoelectrons) is registered, but in many cases, before the peak goes down, more photons are detected, producing a second peak that is convoluted with the first one. Our analysis algorithm should take this into account and try to deconvolute both.

### 4.2 Baseline subtraction

As mentioned in Section 2, it is necessary to subtract the baseline of each independent waveform to quantify the amount of light of the waveform either by using the height of the peaks

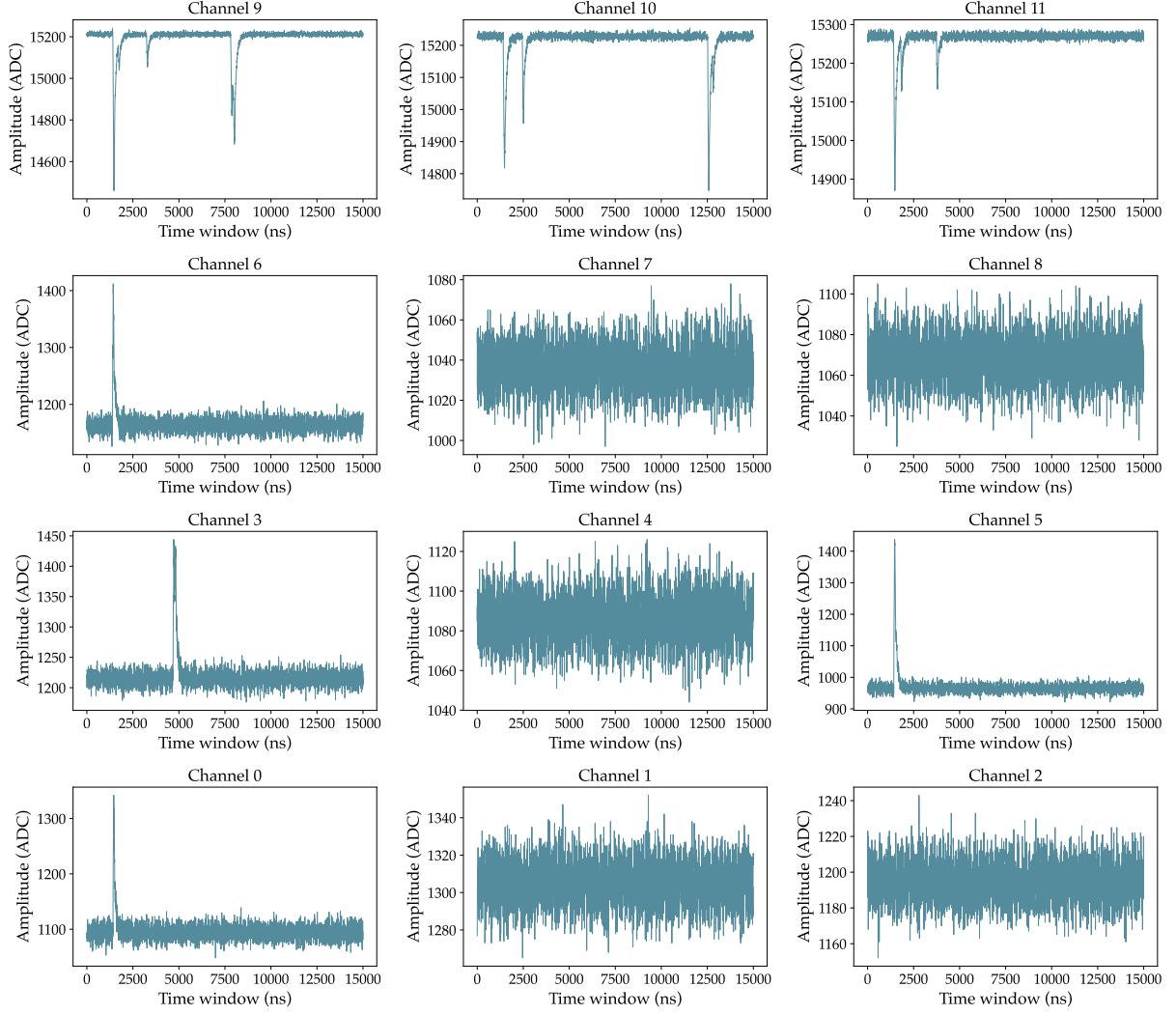


Figure 8: Raw waveforms of all channels for one particular event. Channels are sorted in closer proximity to the source, with channels 9, 10, and 11 being the closest to the source (in fact, they are the trigger SiPMs), while channels 0, 1, and 2 are the farthest from the source, closer to the bottom of the vessel.

or the integral. The baseline subtraction can be achieved by utilizing either the mean value or the mode of the entire waveform or by focusing on a specific range. In the BACoN data, we employ the range of the waveform just before the trigger region to minimize the impact of detected light on baseline computation, which could inflate its value. Furthermore, we opt for the mode as it has been determined to provide better accuracy in case a background event falls in the region before the trigger time.

Figure 11 shows the waveform before (left) and after (right) baseline subtraction. In the plot on the left, the pink region indicates the area used to compute the baseline. Note that in the plot on the right, the baseline is centered at 0 ADC.

### 4.3 Analysis methods

After baseline subtraction, there are two ways of analyzing waveforms, which can be complementary. The first one consists of summing all the waveforms for a channel (some cuts can be

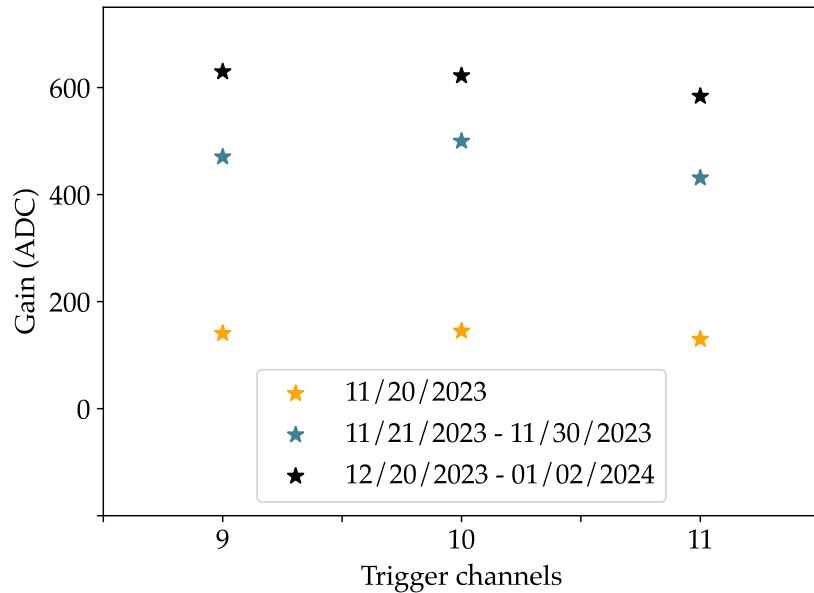


Figure 9: Gain obtained for the trigger channels inside the BACoN detector for the different periods of data taken.

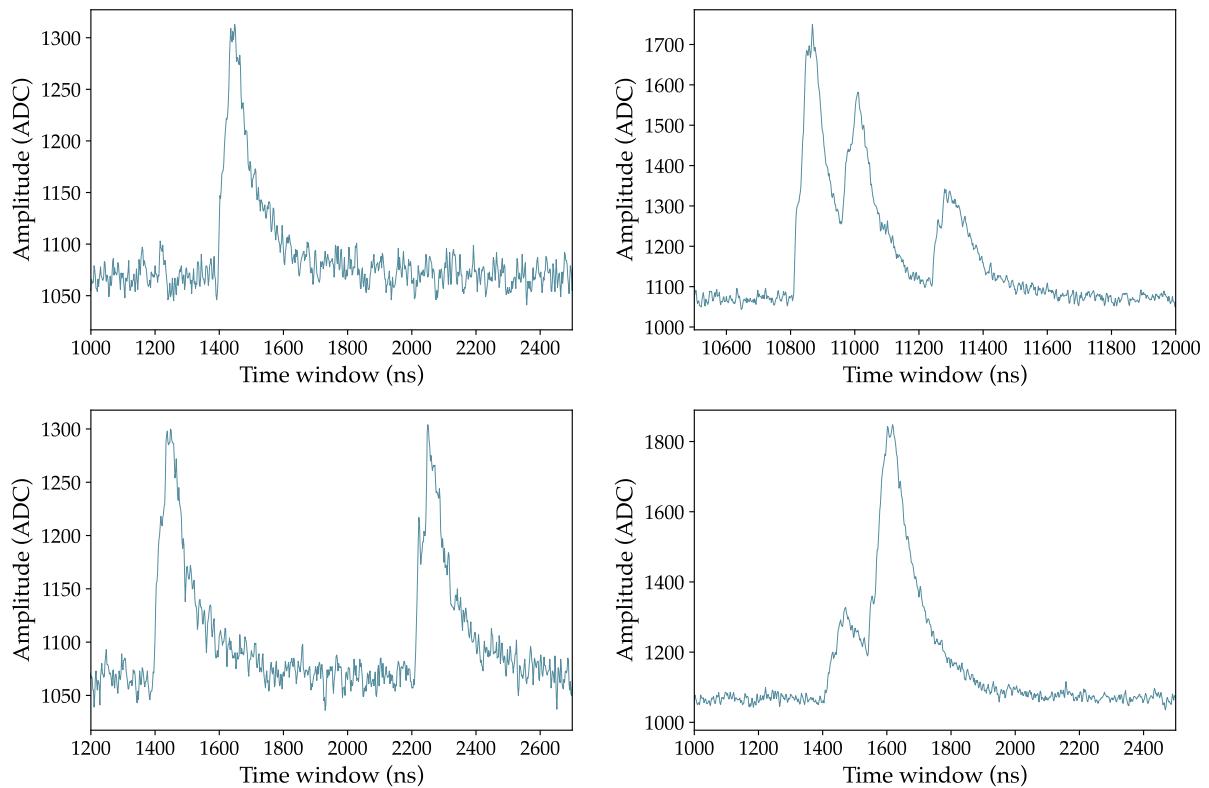


Figure 10: Different kind of peaks selected for channel 8.

applied beforehand, such as cosmic cuts), and the second one consists of analyzing the hits (the peaks) and the times when the peaks occur.

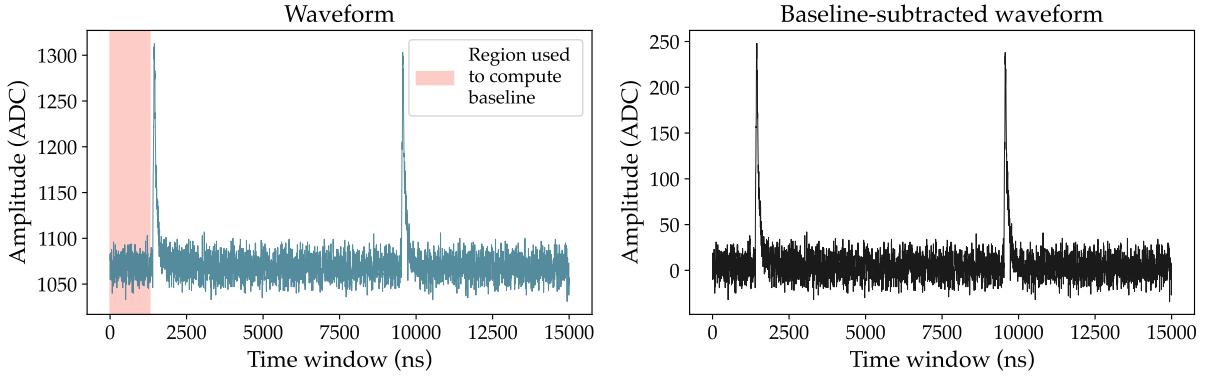


Figure 11: Left: original waveform showing in pink the time window used to compute the baseline using the mean or the mode. Right: Waveform after baseline subtraction.

#### 4.3.1 Summing waveforms

This method is the simplest one in order to study the scintillation light model and extract the decay constants from the fit to understand the difference in the light yield and times after adding xenon. After summing the selected waveforms, the result must be normalized with the number of events summed and the gain extracted from the calibration procedure (Figure 12).

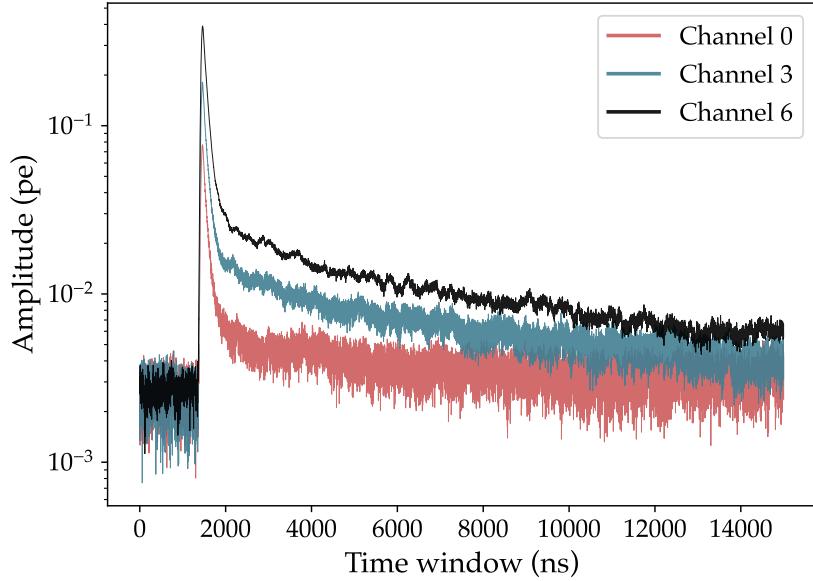


Figure 12: Average calibrated waveforms for channels 0, 3 and 6.

To reduce background events, we can remove waveforms that have peaks above a certain threshold before the trigger region. This is because the signal in the trigger region marks the start of the event. Therefore, the three trigger channels should record signals above 1 pe. If an event has a signal before the trigger time, it indicates that it does not come from the decay source but rather from the environment and should be rejected. Two examples of such waveforms are illustrated in Figure 13. The number of waveforms with a signal before the trigger time depends on the channel and the file, but ranges from 50 to 200 waveforms per data file.

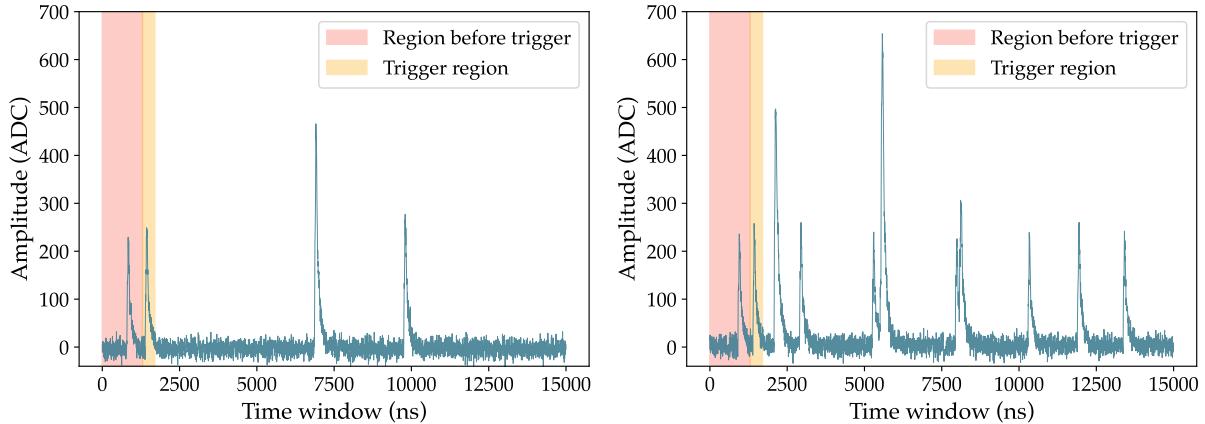


Figure 13: Example of two background events (both present a peak before the trigger time).

#### 4.3.2 Analysis of hits

To identify valid hits in each channel waveform, the following steps are carried out:

##### **Smoothing waveforms using the Savitzky-Golay filter**

First, it is necessary to apply a threshold to reject noise. However, noise counts could be mistaken for signal if they exceed the selected threshold. To mitigate this issue, we employ the Savitzky-Golay (SG) filter. This filter effectively removes noise from signals while preserving important features such as peaks and valleys. It is particularly useful when dealing with noisy data, enabling clearer analysis of the underlying signal.

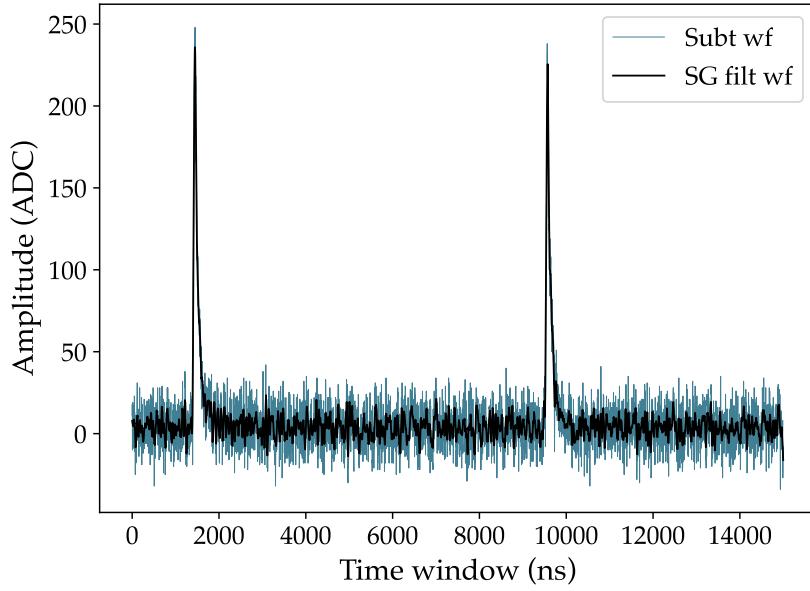


Figure 14: Subtracted waveform (black) and waveform (salmon) after applying the Savitzky-Golay filter.

Figure 14 shows a subtracted waveform and the result after filtering using the Savitzky-Golay (SG) filter. As depicted in Figure 15, with this filter, we can ensure that no counts from

the baseline exceed the desired threshold.

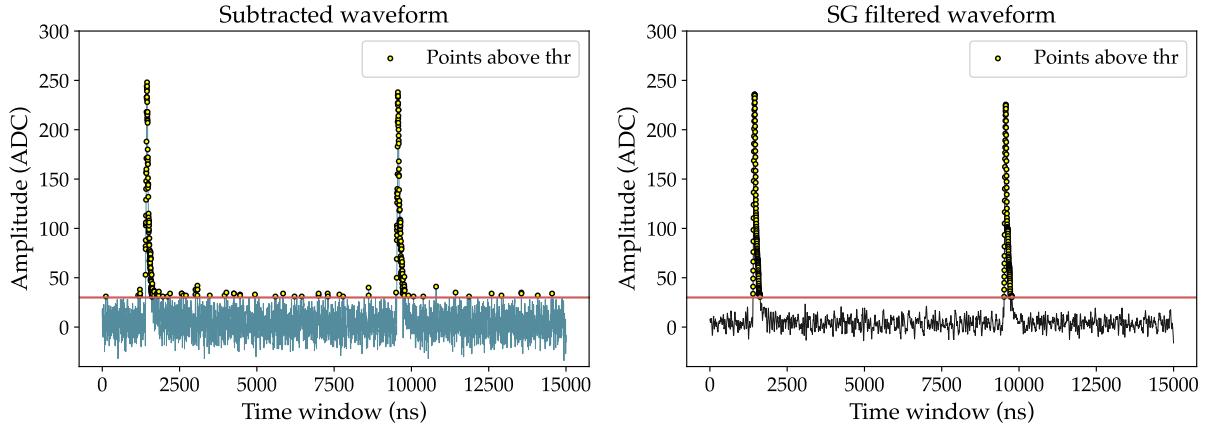


Figure 15: Left: subtracted waveform; right: subtracted and filtered waveform. The points above the chosen threshold (30 ADC, for example) have been plotted to illustrate that after filtering, no points from the baseline are present that could lead to a misinterpretation of the signal.

One consideration when employing this method is that the heights of the peaks after filtering will be slightly reduced. Therefore, to address such variations, SiPM calibration should also be conducted using this method.

### **Set a certain threshold and compute zero suppressed waveforms**

An intermediate step between filtering and obtaining the peaks is to compute the zero-suppressed waveforms. In this process, all parts of the waveform below the chosen threshold, which removes the noise, will be converted into zeros (see Figure 16). This makes it easier for the algorithm (which will be used and explained in the next step) to identify the different peaks.

### **Search for peaks above the threshold**

In order to get the peaks from our zero suppressed waveforms, we are using the python package `PeakUtils`. As written in the documentation: "this package provides utilities related to the detection of peaks on 1D data. Includes functions to estimate baselines, finding the indexes of peaks in the data and performing Gaussian fitting or centroid computation to further increase the resolution of the peak detection". We should provide this algorithm the minimum distance between peaks and the absolute threshold in ADC. For the moment we are using 50 ADC both for the threshold the minimum distance between peaks.

The outcome of the algorithm is the peak position in the waveform and using that we can extract the peak height, shown in Figure 17.

Using this method, we can get all the peaks from each channel and compute a hit map like the one shown in Figure 18. In the plot, defined horizontal lines can be observed. They are due to the individual photons registered by the SiPM.

### **Additional ways to compute the hit map**

In the previous method, we were taking the timestamp of the peak as the one corresponding to the maximum in the peak. However, another option is to take the time when the peak exceeds

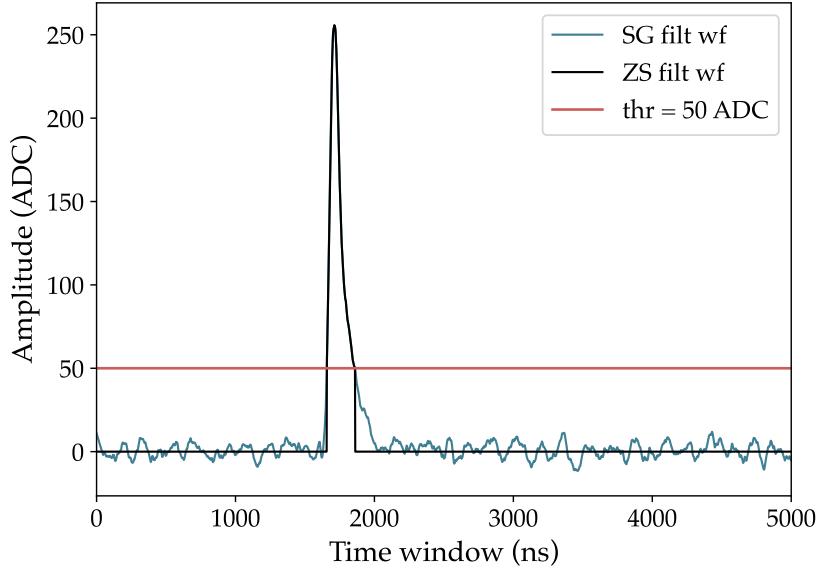


Figure 16: Peak of a selected waveform zoomed in. The blue curve corresponds to the filtered waveform while the black one is the zero suppressed waveform, where the values below 50 ADC have been set to zero.

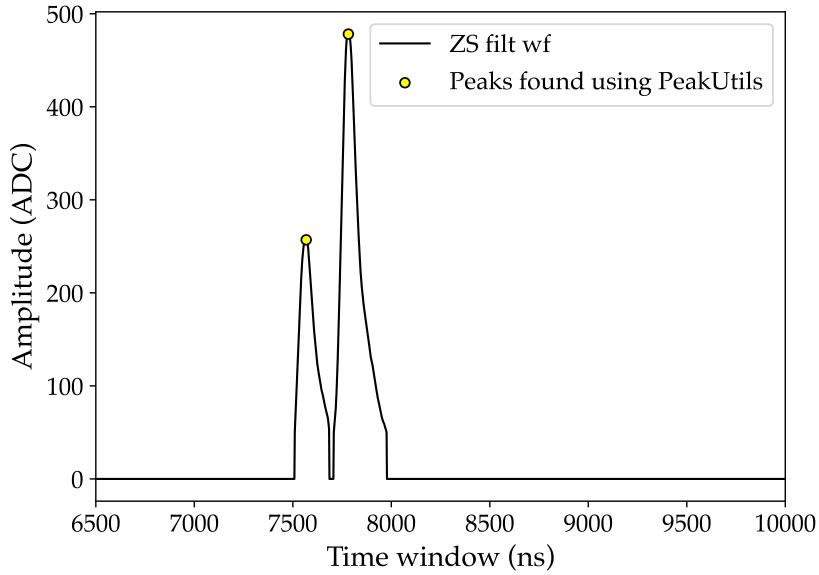


Figure 17: Example of the two peaks found by PeakUtils in a zero suppressed waveform.

the threshold value. This way the effect of the rise time of each peak would be reduced (Figure 19).

Another factor that we should take into account is the height of the peaks when two peaks overlap. If the second peak starts before the first peak reaches the baseline, the height of the second peak should start from the intersection with the first peak, as shown in Figure 20.

After combining the calibrated hits from all channels of a single file, the resulting hit map is depicted in Figure 21 left. Subsequently, when the hits for each number of photoelectrons (represented by different colors up to 4 SPE) are aggregated into a histogram, Figure 21 right is generated. It is observed that few number of photoelectrons are more common in the wave-

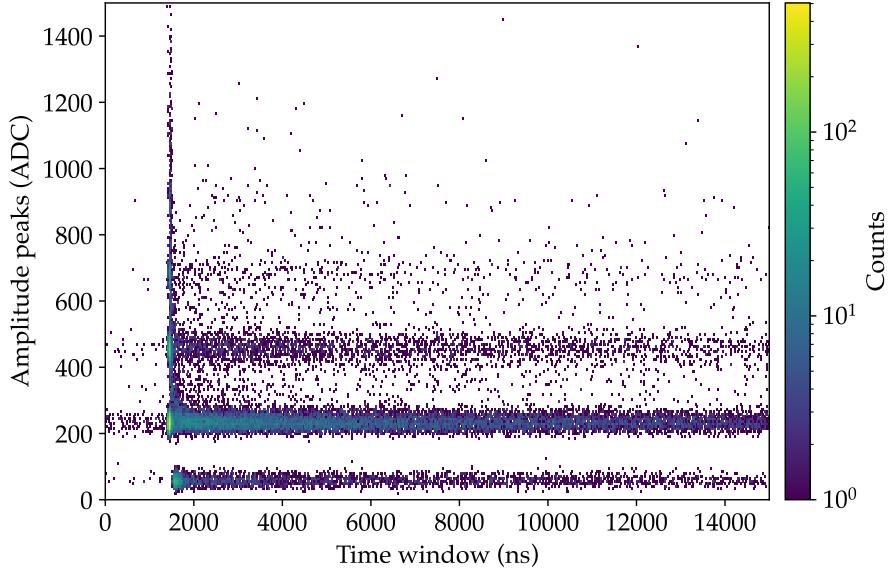


Figure 18: Hit map for one channel. Each bin represents a peak with its timestamp (position in the waveform) and height (amplitude). Each horizontal line corresponds to a integer number of photoelectrons.

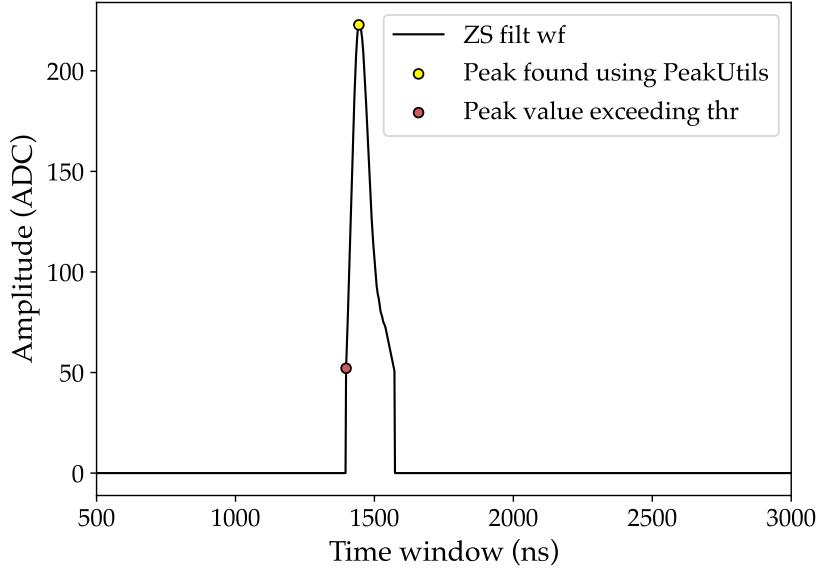


Figure 19: Example of a peak where the timestamp is taken from the maximum of the peak (yellow) or from the point where the threshold is exceeded (red).

forms.

#### 4.4 Comparison between hit map and waveforms sum

Figure 22 shows the comparison between the different analyses. The red line corresponds to the sum of the hits for each timestamp using the time of the maximum of each peak. The blue line has been computed also from the hit map but computing the time of the waveform when it crosses the threshold. Finally, the black line corresponds to the sum of all the waveforms in

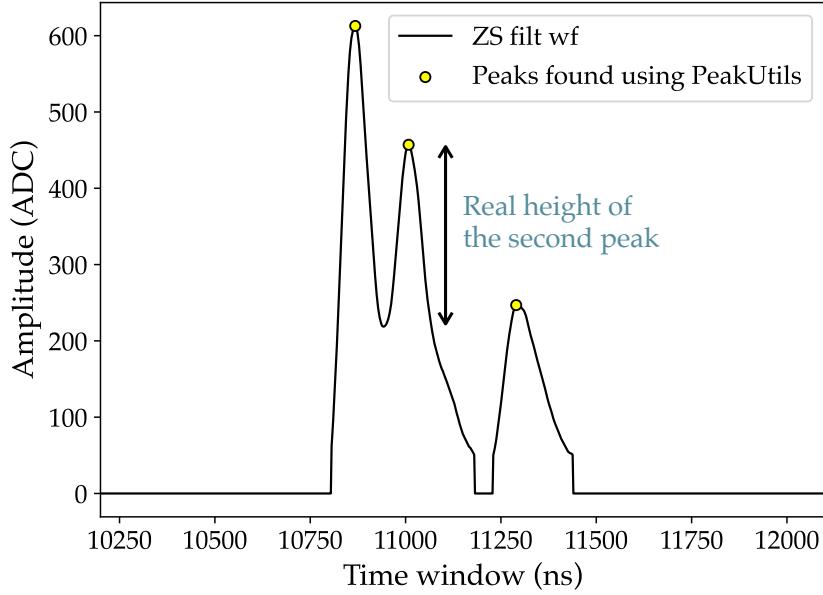


Figure 20: Example of two peaks overlapping. As it is written in the figure, the real height of the second peak corresponds to the difference between its height and the height of the intersection with the first peak.

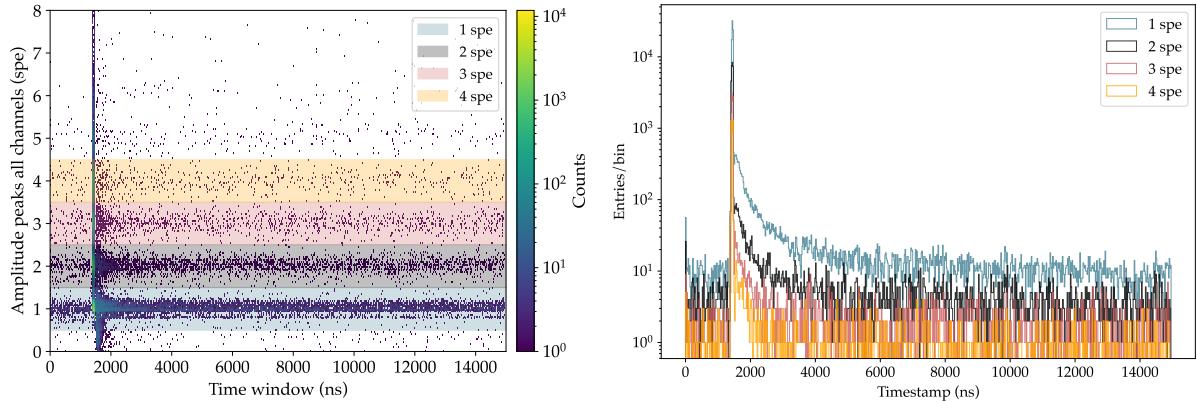


Figure 21: Left: calibrated hit map for all channels of a file with the colored bands corresponding to a different photon number. Right: histograms corresponding for each photoelectron number until 4 SPE.

the selected file. All cases consider all the events with no signal before the trigger rate.

## 5 Single photoelectron shape

Peak shape for the different number of photoelectrons.

Fit of the 1 spe shape.

## 6 Trigger time

We define the trigger time as the moment when the peak exceeds the threshold of 50 ADC in the trigger region of a waveform. Creating a histogram with the values for each channel results in

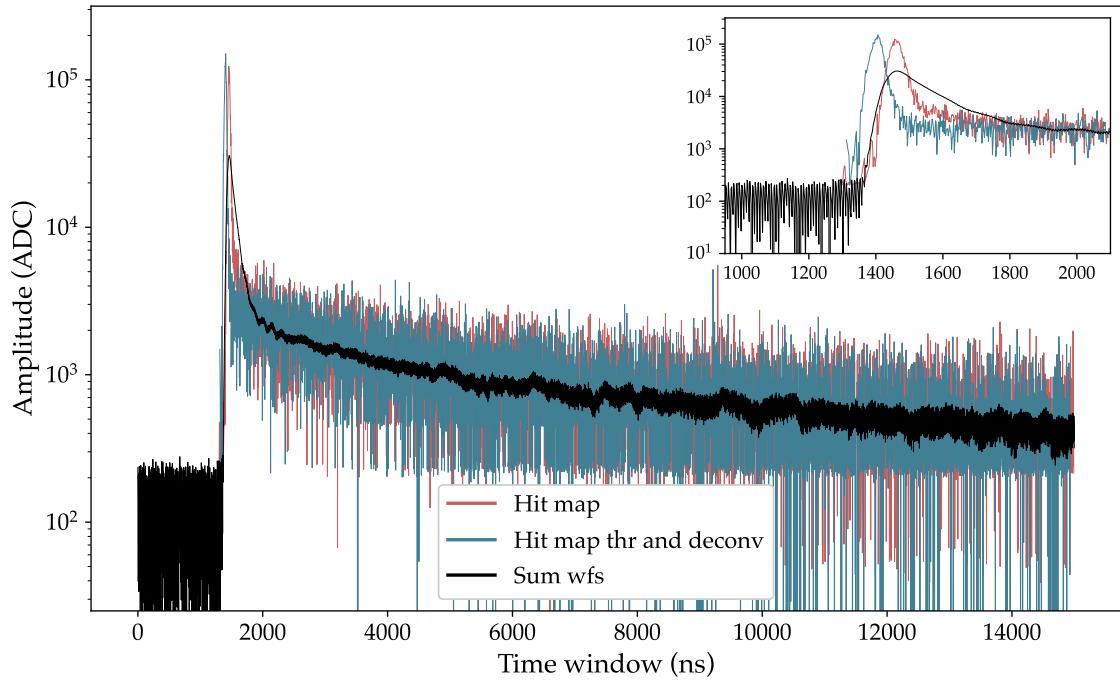


Figure 22: Comparison between the analysis using the hit map with the time of the maximum of the peak (red) and using the time where the waveform exceeds the threshold and deconvolving the second peak from the first peak (blue) and the analysis using the sum of the waveforms (black), all three cases after removing events with peaks occurring before the trigger time. The sum of the waveforms has been normalized to achieve a similar order of magnitude as the other cases, enabling a comparison of the shapes of the three analyses. The plot on the upper right corner zooms in the peak region.

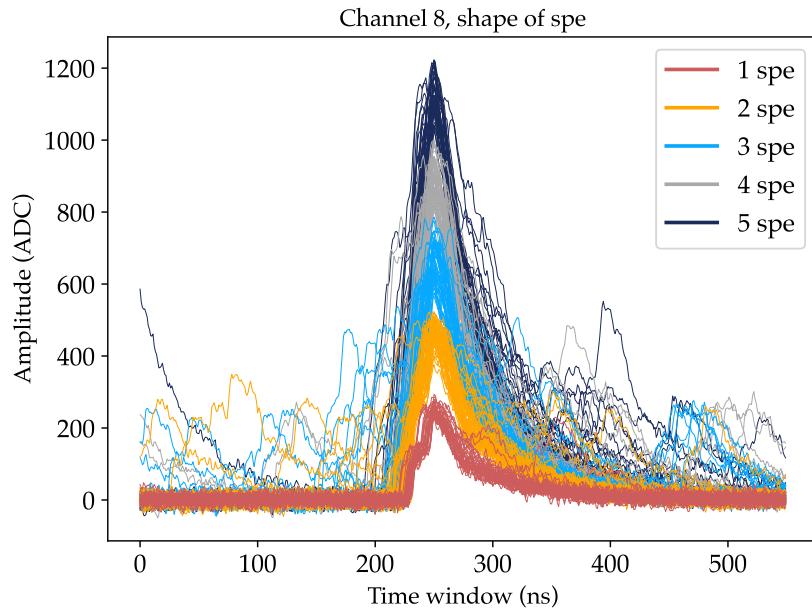


Figure 23:

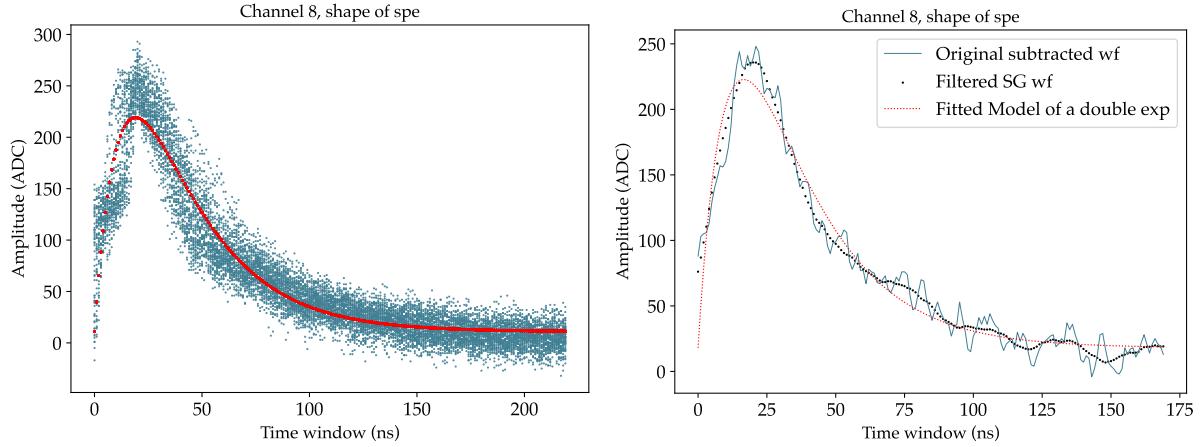


Figure 24:

the plots depicted in Figure 25. On the left, the trigger time for the trigger SiPMs is displayed, while the plot on the right illustrates the same parameter for the regular SiPMs. In both cases, the peaks are centered around 710 ns, but the FWHM is narrower in the trigger SiPMs (close to 40 ns, corresponding to the trigger window). This result was expected, as the trigger SiPMs are closer to the source.

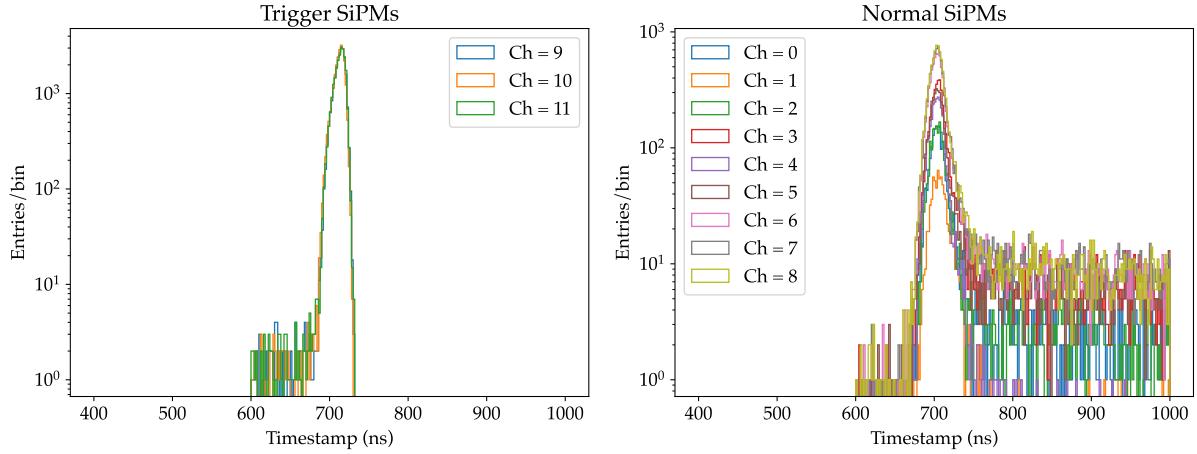


Figure 25: Trigger time for the trigger SiPMs (left) and for the regular SiPMs (right).

Plot 26 depicts a fit for the trigger time combined for the source SiPMs (left) and the regular SiPMs (right). The histogram shape of the trigger SiPMs is non-gaussian, whereas it appears more gaussian for the regular SiPMs. As mentioned, the width of the distribution is narrower for the source SiPMs.

## 7 Rise time

We are computing the rise time of the peaks as the difference between the maximum of the peak and the point that surpass the selected threshold.

Figure 28 shows the rise time vs amplitude in number of photoelectrons for all SiPMs. As expected, the rise time of a peak increases with the number of photoelectrons.

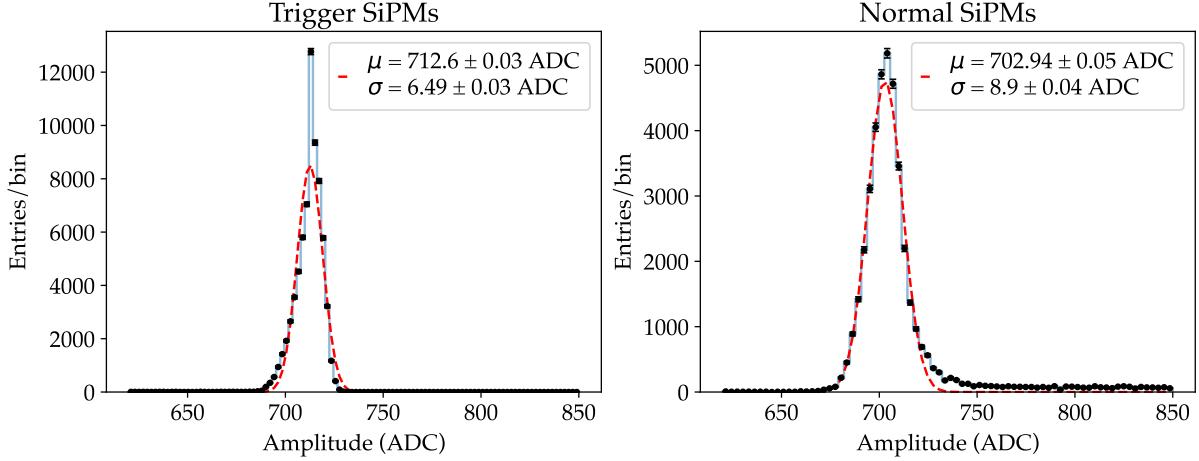


Figure 26: Fit of the trigger time combined for the trigger SiPMs (left) and the normal SiPMs (right).

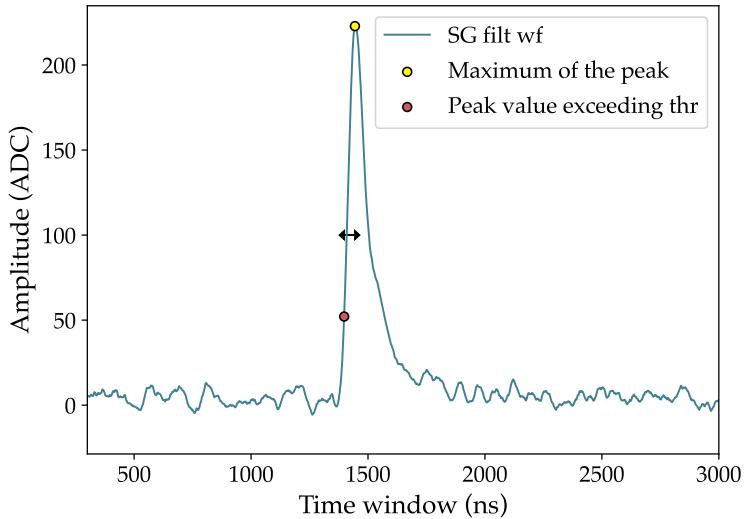


Figure 27: Waveform with the maximum in yellow and the point where the threshold of 50 ADC is exceeded in red. From the difference of the two times, the rise time is computed.

## 8 Energy spectrum

The energy spectrum obtained with our setup can be computed using the amount of light detected by the three trigger SiPMs per event. Various scenarios are illustrated in Figure 29: the sum of the peaks in the entire window versus the trigger window, and all events versus only centered events. Centered events are selected by summing only the charges of each channel within a difference of 2 pe with the other 2 channels. This way, we can ensure that the event occurred at a similar distance from all sensors, that is, at their center.

In Figure 29, individual photoelectrons (starting at 3, since the sum is represented) are observable on the left. Near energy levels corresponding to several photoelectrons, a subtle bump is noticeable in all cases. This bump is likely attributed to the alpha particles emitted by the source. Unlike the LLAMA detector, which exhibits a distinct peak resulting from gamma rays at around 40 photoelectrons and no discernible peak for alphas, our setup displays a minor

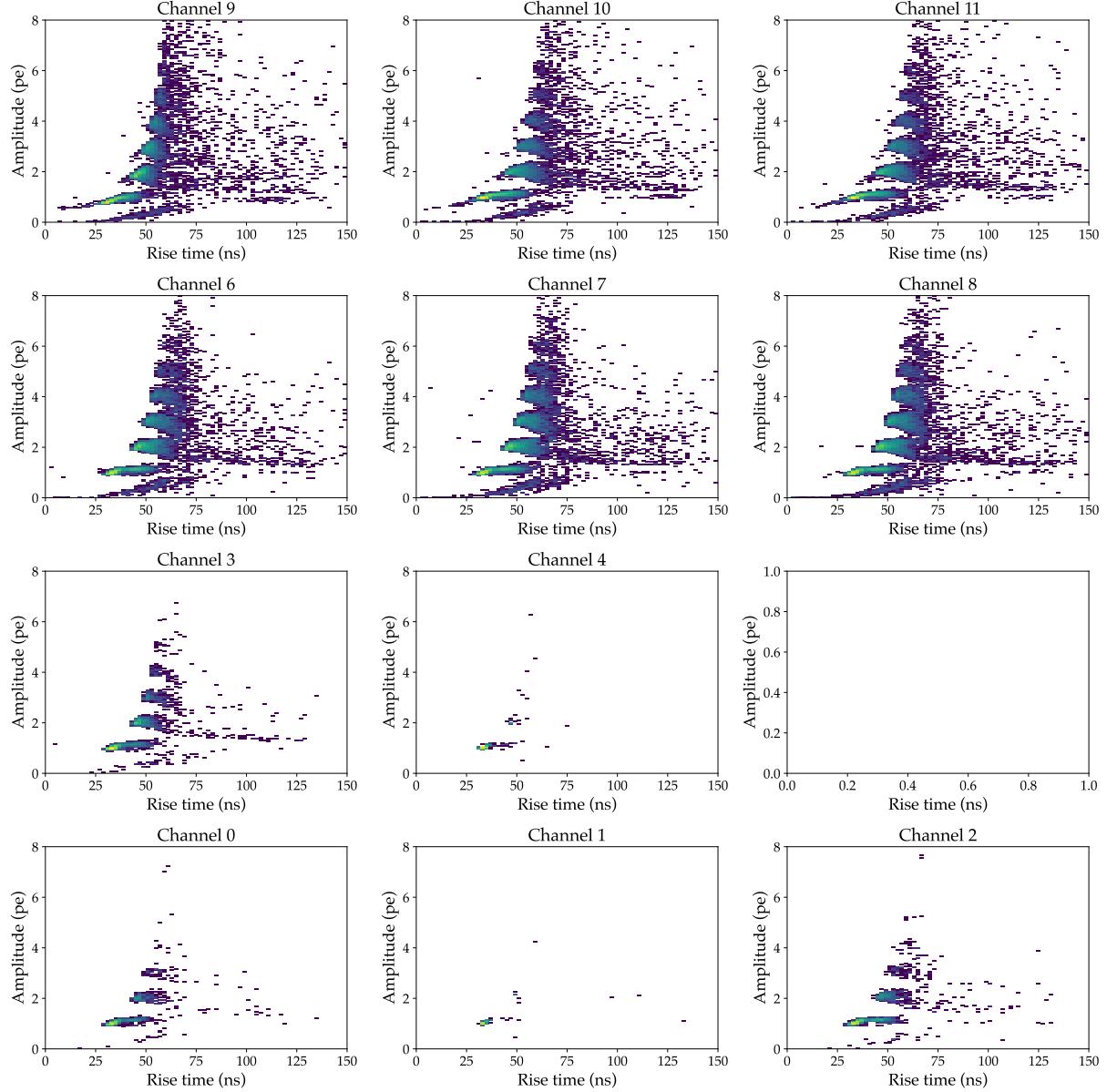


Figure 28: Rise time versus number of photoelectrons for each channel for a Dec 28th run (first file). Channels 9, 10 and 11 are the trigger SiPMs, close to the source. Channel 5 was not working correctly.

bump indicative of alphas, but an indistinguishable gamma peak. Several disparities exist between the two detectors: LLAMA's underground location results in fewer background interferences. Additionally, they employ a higher threshold in the electronics of their trigger SiPMs, and their source is sealed, precluding alpha particle detection. Furthermore, their SiPMs lack collimation, resulting in a higher acceptance, approximately 2 orders of magnitude more than in BACoN. Therefore, it is reasonable that in our setup, the alpha peak appears at a lower energy region0.

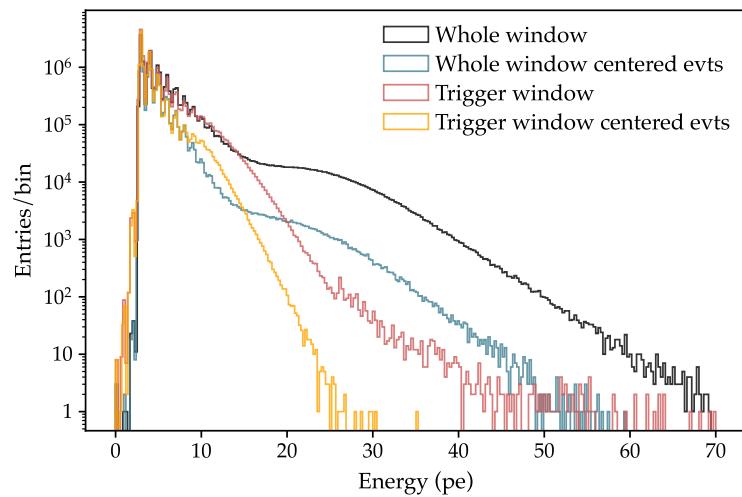


Figure 29: Summed energy of the three trigger SiPMs for 4 different cases: taking the whole window of the event, taking only the trigger window, and the two same cases but selecting those events with  $\pm 2$  pe difference between the three channels.