Technische Universität München

Abschlussarbeit im Bachelorstudiengang Physik

Simulation of the Signal Chain for STRAW

Simulation der Signalkette von STRAW

Maximilian Hoch

Erstgutachter (Themensteller): Prof. Dr. Elisa Resconi

Zweitgutachter: Prof. Dr. Lothar Oberauer

24.09.2018

Abstract

The aim of this Bachelor thesis is to simulate the signal chain of the STrings for Absorption length in Water (STRAW) experiment. The project was deployed to evaluate a potential site for a neutrino telescope. In order to do that the conditions like absorption length for emitted light and scattering length as well as the background mostly caused by bioluminescence have to be determined.

In order to evaluate the site LEDs are flashed and recorded. The photons emitted by the flasher unit are not emitted simultaneously but with a time distribution of a few nanoseconds. Also the Photomultiplier used for detection suffers from a transit time spread on the same order of magnitude. An amplifier, the PADIWA, then modulates the signal which has to be simulated. The signal then gets read out by the Trigger-Readout-Board which records time stamps from points where the signal crosses a threshold, rising or falling.

The simulation starts with FADC recordings of Photomultiplier pulses of singe photon events. After reproducing all the different effects caused by the flasher, PMT, amplifier and discriminator the output has the same form as STRAW itself.

The simulated data contains information of how many photons were detected and what time shifts they have. The results of this simulation will be used to analyze the data collected with STRAW in the future.

Contents

1   Introduction

1.1   Neutrino Physics

The Standard Model of Particle Physics includes 3 neutrinos which are electron neutrinos, muon neutrinos and tau neutrinos. These three are neutral partners of electrons, muons and taus. Therefore, neutrinos do not carry electric charge or colour charge. The neutrino only reacts with matter due to the weak interaction with a very small cross section. With a typical neutrino flux for solar neutrinos of $10^{10}cm^{-2}s^{-1}$ and a typical interaction cross section of about $10^{-45}cm^2$ the neutrino at low energy, 2 MeV, have an absorption length of 1600 ly in water. The flux from solar neutrinos measured on earth is about 10^{11} solar neutrinos per second and square centimetre. Using a large enough target for the neutrinos to interact with one measures a few events per day. Early neutrino detection experiments used the inverse β-decay [1] as seen in equation 1.

$$\bar{\nu}_e + p \rightarrow n + e^+ \tag{1}$$

The annihilation of the created neutrino and positron produces gammas which are detected with scintillator. The coincidentally detecting two gammas can be used to largely suppress background form cosmic rays and radioactivity which proved the existence of neutrinos.

Astrophysical neutrinos and atmospheric secondary particles formed in cosmic ray showers have a much lower flux and higher energies. Again, these are detected via secondary leptons and their reactions created in the first vertex (e, mu, tau). If the neutrino had enough energy the produced lepton is faster than light in the medium and sends out Cherenkov radiation. The Cherenkov light is emitted in a conic shape like it is illustrated in figure 1. The figure illustrates the geometry of the emitted Cherenkov light.
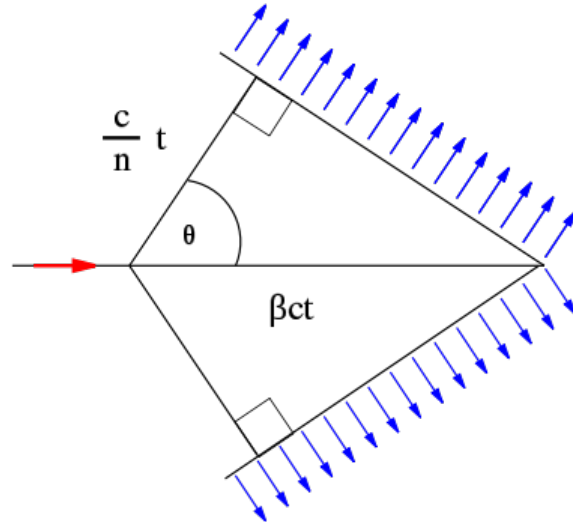

Figure 1: The Cerenkov light is radiating outward in a cone through the medium.

From the measured light the energy and the direction of the neutrino can be reconstructed [8]. A large-scale neutrino detector in water or ice can be used to reconstruct the direction and energies of neutrinos and at the same time be used for suppressing background from atmospheric muons.

1.2   Neutrino Sources

Neutrino sources measured on earth can be separated into two groups, astrophysical neutrinos and atmospheric neutrinos.

Astrophysical neutrinos are created by mechanisms that are not yet fully understood and are created far away from our atmosphere. These neutrinos created in the universe can have very high energies. The IceCube Neutrino Observatory measures neutrinos up to the PeV range [9]. The sources of these neutrinos are suspected to also be the sources of cosmic rays, which could produce those neutrinos in deep inelastic scattering of CRs with other nuclei [14].

Neutrinos which are created by energetic particles such as nuclei hitting the atmosphere are called atmospheric neutrinos. The cosmic rays which create those neutrinos consists of protons (79%) and alpha particles (15%) and other nuclei (6%) [13]. The energies of these atmospheric neutrinos are stretching over ten orders of magnitude.

The cosmic radiation is mostly isotropic. Solar winds and the magnetic field can influence particles with Energies below 10 GeV [2]. The cosmic rays are producing secondary particles, mainly hadrons, which decay into charged leptons and neutrinos [14, 13]. These cascades distribute over a large area on the ground. The neutrinos formed by such air showers are divided into conventional neutrinos from pion and kaon decays and prompt neutrinos coming from the decay of hadrons which have different energy spectra.

1.3   Neutrino Detection in Ice

The IceCube Neutrino Observatory is a large-scale neutrino detector. It is located on the geographic south pole. Therefore the transformation from horizontal into equatorial coordinates is simplified. The sensor units are distributed over a cubic kilometer of Antarctic ice. The Cherenkov light in the ice is measured to reconstruct the trajectory and energy of point sources in the TeV range. IceCube's last construction phase ended on December the 18th of 2010.

IceCube consists of spherical optical Modules, DOMs, deployed on strings. Each DOM has a photomultiplier tube, PMT, to detect the emitted light. The single board data aquisition computer sends the digital data to the counting house on the surface above an array [10]. Each string holds about 60 DOMs spread out in depths of 1.450 to 2.459 meters.

The construction of such a detector is very complex and labour intensive. But the construction can only take place during the Antarctic austral summer from November to February. The permanent sunlight during summer allows for 24-hour drilling. Drilling the holes was achieved with hot water which was heated on sight. To keep a drilled hole from collapsing the water had to stay in the hole. That is why deploying a string with the DOMs mounted into the hole was a time sensitive action.

The contraction took place over 5 years. In 2005 the first string was deployed. Each summer season 13 to 20 strings could be set up leading to a total of now 86 strings [11, 15]. The total cost of the project was 279 million USD [6] which was founded by numerous universities and research institutions worldwide [5].

Measuring neutrinos in ice is a very efficient way but the construction of such a detector takes a long time and takes place at a remote working area.

1.4   Neutrino Detection in Water

The "Strings for Absorption Length in Water" project, STRAW, uses a different medium in which Cherenkov light is emitted, water. The physics behind Cherenkov light produced in ice is the same as in water. The purpose of this detector is to investigate the area off the coast of Victoria in the so called "Cascadia basin" as feasible location of a future large volume neutrino telescope. The optical properties of water over the seasonal variations will be measured over the course of two years. The scattering properties of the sea is also affected by the background from radioactive materials such as K^{40} and bioluminescence. The emitted light from sea life is currently the largest contribution to the background.

The working principle of STRAW is to emit short but intense flashes of light deep under the surface and test the absorption length. These flashes will be detected by fast and sensitive detector units, which also allow to study background light. The ocean research institute that acts as collaboration partner for STRAW, Ocean Networks Canada (ONC), has a very good infrastructure in the deep sea.

2   STRAW

STRAW consists of two strings each with a length of 146.5 m on which a total of 3 light sources and 5 light sensors are mounted in a depth of about 2600 m. The geometry of the two mooring lines is very important in order to avoid blocking light paths from the light source to the light sensors. This is why importance was given to keep twisting of the strings to a minimum. In order to keep the strings up right even with the current pushing the string, floaters were added for high buoyancy.

The distance between the strings is 50m. The modules were mounted 30m, 50m, 70m and 110m above the sea floor. The strings are hold to the seafloor with an anchor. The light source module called Precision Optical Calibration Module, POCAM, is shown blue in figure 2, and the light sensor module called Straw Digital Optical Module, sDOM, shown red.
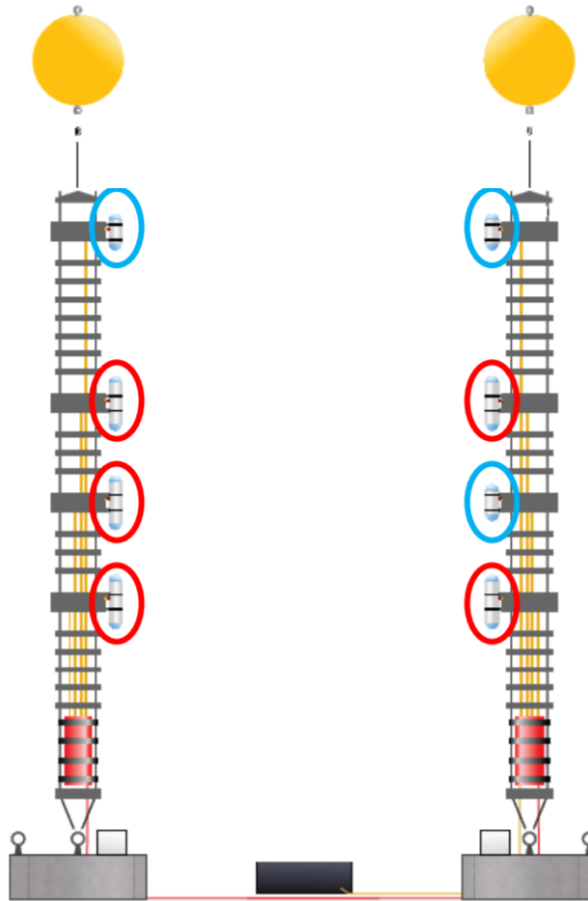

Figure 2: STRAW consists of two strings with 3 POCAMs (blue) and 5 sDOMs (red). The strings are fixed to the seafloor with an anchor and are held up right by floaters with a buoyancy of negative 420 kg . The schematic diagram was kindly provided by Christian Spannfellner.

One string without the modules consists of two carrying ropes with spacers each 5 m and data cables. Each module is mounted to a special spacer at the specific heights.

2.1   Flasher and Sensor Modules

Both modules share the same base design but with different electronics and instruments suited for their specific purposes. The detector is running Linux and is accessible via SSH. The modules are connected via a 100 Mbit connection which is the bottleneck in pulling the data from the detector. The flasher units use micro controllers and communicate via RS485.

The STRAW project includes three POCAMs which are used for isotropic (within one half sphere), nanosecond light flashes in the range of 380 to 600 nm produced by four different LEDs. One POCAM has one set of LEDs pointing upward and one set pointing downward as seen in figure 3.

Figure 3: This is a 3D model of a POCAM. The POCAM contains 2 LED arrays, one at each end. The 3D render was kindly provided by Kilian Holzapfel.

Five sDOMs are mounted on the strings which contain a PMT facing upwards and a PMT facing downward and all the needed electronics to record single photon events in order to determine the properties of water as previously mentioned. A cross section of a sDOM is shown in figure 4.

Figure 4: This is a cross section of an sDOM. The sDOM contains 2 PMTs, one at each end. The electronics are capable of handling both at the same time. The 3D render was kindly provided by Felix Henningsen.

As seen in figure 4 the ends are covered with a pressure resistant BK7 glass hemisphere with a diameter of 114 mm and a thickness of 7 mm. It is mounted with a certified deep sea adhesive to the housing. The glass lets photons pass through onto the three-inch PMT which is a Hamamatsu R12199. The window is providing better transmission in the UV light regime because it is made out of borosilicate glass. The 80 mm PMT has a high gain of up to $\sim 10^7$. It is responsive in wavelength between 300 to 650 nm which is a perfect range for the measurement. The electric pulses from detected light flashes are split, amplified and discriminated through four channels per PMT using the so called PADIWA module. The time information is then digitized with 100 ps precision using the Timing Readout Board (TRB).

The housing is made out of high-pressure-resistant titanium with a thickness of 12 mm. It is capable to withstand 600 bar. On the right side of figure 4 is a penetrator for the cable connection and to pump out the air replacing it with Nitrogen at ~ 300 mbar. This prevents possible corrosion from condensing water in the air surrounding the electronics. The cable which is used to connect the modules to the junction box is a FALMAT Xtreme-Net FM022208-12 that has been provided by ONC.

The read-out electronics are supplied by the junction box with 48 V. A constant voltage regulator steps down the voltage to 5 V producing a fair amount of heat which is distributed through an Aluminum link to the housing. The three main DAQ parts are the Odorid C2 which is running all the protocols needed, the PADIWA and the Triggered-Readout-Board, TRB.

Shaped pulses from the PADIWA are transmitted to the TRB. The PADIWA has voltage thresholds defined for four channels. Once the pulse exceeds the set threshold the TRB which hosts a time-to-digital converter, TDC, records the timing of rising and falling edges at the thresholds. The data is temporarily stored onto the single board Odroid.

2.2 Photon Propagation

2.2.1 Absorption

If an isotropic source sends out light with an intensity of I_0 the intensity I can be calculated at a specific distance x with

$$I(x) \sim I_0 \cdot \frac{1}{x^2} \cdot \exp\left\{\frac{-x}{l_{abs}}\right\} \tag{2}$$

In order to determine the absorption length l_{abs} which is a wavelength dependent both intensities and the distance has to be known. The absorption length is defined as the distance a photon can travel without the probability of absorption falling under e^{-1}. The absorption length for pure water is shown in figure 5.
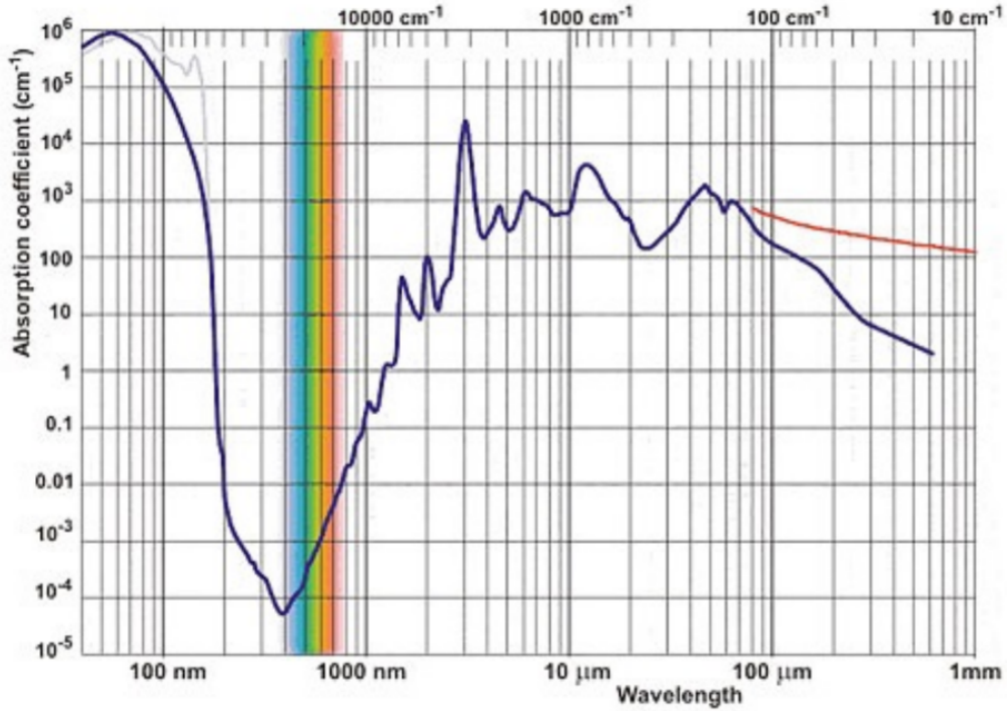

Figure 5: The graph shows the absorption coefficient depended on the traveled length for electromagnetic radiation [3]. The important wavelengths for the experiment are ultraviolet and green.

As seen in the graph the absorption length depends on the wavelength which means that blue light can travel up to 100 m without being absorbed. The absorption coefficient of water for wavelengths above 600 nm rapidly increases meaning that it will be absorbed far quicker.

The Cerenkov light which is relevant for neutrino detection emitted is dominantly in the regime between ultraviolet and green. This is why the used PMTs in STRAW are tailored to the regime of blue to green colors and should therefore yield a better resolution.

11

This is also true for the salt water in Cascadia basin. The absorption in sea water is also limited by floating objects like bio-mass. The bio-mass around the detector is blocking light and limiting the theoretical possible distance at which neutrinos could be measured.

The four different LEDs the absorption length of the sea water can be measured very precise because the intensity from the LED and the measured intensity detected by a module is known as well as the distance. But in order to get precise results the measured intensity has to be reconstructed from only a few points for which the simulations could be used. With equation 2 the absorption length can be calculated. Testing all four LEDs a absorption coefficient graph can be estimated which is very important to determined the feasibility for neutrino detection.

2.2.2 Scattering

In analogy to the absorption length the scattering length is defined. The scattering length is the length at which the probability that a photon has been scattered is equal to $1 - e^{-1}$.

The absorption length will be determined by flashing an LED which points directly to a PMT. In order to measure the scattering length an LED, facing away from the PMT, has to be flashed. Because if the PMT detects the flash pointed away from it then that means the photon was scattered. A scattered photon takes a longer time to arrive at the PMT because it had to travel a longer distance.

Both lengths will be calculated to gain insight into the path a photon takes. By comparing both length it determines if the average photon which has been detected has been scattered or not. In order to get accurate results a precise timing measurement in the order of nanoseconds is taken.

2.3 Background

2.3.1 Luminescence

Sea water provides a high natural luminescence which results in a high background noise. This could greatly impact the possible neutrino signals that could be measured.

A part of this luminescence in the water is caused by K^{40} decay. Potassium is part of the salt solved in the water. The decay rate is proportional to the salinity.

2.3.2 Bioluminescence

Unlike the well understood radioactive decay causing the background, bioluminescence has to be predicted. The bacteria and algae which are capable of producing light are dependend on the region. Over the course of seasonal variation which determines the temperature and the present bio-mass the background could change drastically. Furthermore the currents are also having an impact on the sea life and are changing constantly which can trigger the bioluminescence of certain algae or let other light emitting sea creatures drift away.

One of the first under water creatures that where encountered on the strings were Pyrosomes. These are free-floating colonial tunicates that live in the area. The Pyrosomes are

cylindrical colonies made up of hundreds of zooids which are all capable of emitting light. They are starting to emit light if there is danger or if they them self detect light [4].

Therefore the water has to be monitored over a period of two years to determined the impact on the project. This requires the ability to measure single photons at a rate of up to 100 kHz regime.

2.3.3   Water sample

In order to determined the salinity, in specific the amount of K^{40} presence in the water, was taken by the Remotely Operating Vehicle (ROV) during deployment. This probe will be analyzed at the University of Alberta. This will provide a better understanding of the background caused by the radioactive luminescence but it will only provide a temporary accuracy because changes during the seasons will change the amount of potassium.

2.4   Data Analysis for STRAW

STRAW is already taking data in the deep-sea. The remotely triggered modules are taking data and temporarily store it on an SD card. Loading the data over the ONC network is currently limited. The data which is taken from the background is investigated especially the change over one day. The strings were deployed in June 2018, this is why the investigation for changing conditions during the season has not yet taken place.

Measuring the absorption length and scattering length is under way. Flashing the LED with one of the four colours. The sDOMs are then recording photons which are proportional to the intensity they receive.

The intensity for each LED can be set by the supply voltage to the Kapustinsky circuit. The LEDs have been tested in the medium to high light regime which means a voltage starting from 10 up to 20 V. The light output and time frame in which the photons are send out is linear dependent on the bias voltage [7]. During the testing of the simulation a linear expansion to 5 V has been assumed. This assumption is mostly confirmed by the results. Figure 6 shows the data that has been measured by Felix Henningsen which suggest the constant linear dependency.
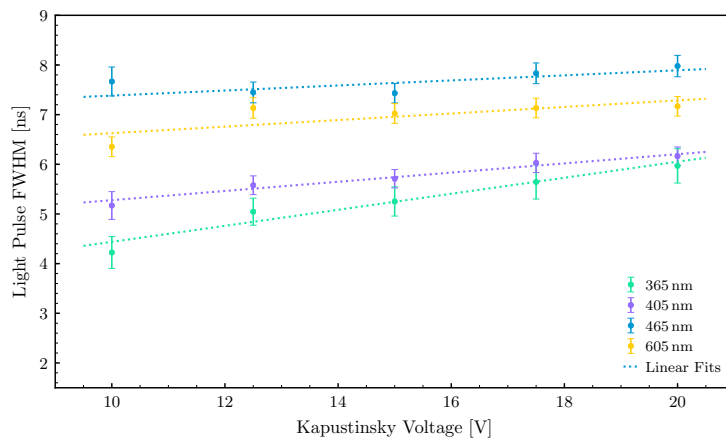

Figure 6: The graph shows the linear dependency of the supplied voltage to the light pulse FWHM that is sent out [7].

The data recorded by the sDOMs are time stamps. These time stamps correspond to the signal crossing a voltage threshold. Once the input surpasses or falls below a threshold the time is written down. Every threshold is twice as high as the one before with a total of four thresholds. These points in time which represent one of four specific voltages are used to reconstruct the signal but at the moment there is no procedure to do that. In order to achieve that a basic understanding of how these curves look like is needed. Furthermore, gaining insight of how many photons the signal was composed of results in a better calibration for detecting signals.

With a high background and the main part of that caused by a not easily understood source, the bioluminescence, the processes inside the detector becomes very important. Better understanding the signal leads to a better understanding of the area.

STRAW was deployed to investigate the area of Cascadia basin. If building a large-scale neutrino telescope would be an option in the area, used for vetoing the background of the atmospheric muon flux. In order to answer this question many things have to be done. One important part is understanding the detector.

3 Simulation of the signal chain

3.1 LED Flasher

The simulation starts at an LED array from a POCAM. The Kapustinsky flasher provides nanosecond light pulses with a good light yield. Within a flash of about 10 ns, 10^9 photons can be emitted. Sending out many photons over a short period of time is very important for the calculation of the absorption length. If too few photons would be emitted the biomass in the water could absorb all photons and the test run would yield no results. The time needed to send out the photons should be as short as possible to get more accurate results.

Each LED sends out a different Wavelength and is named with a number. The LEDs send out light with the wavelengths of $605, 365, 405, 465$ nm. The LEDs are shown in figure 7.
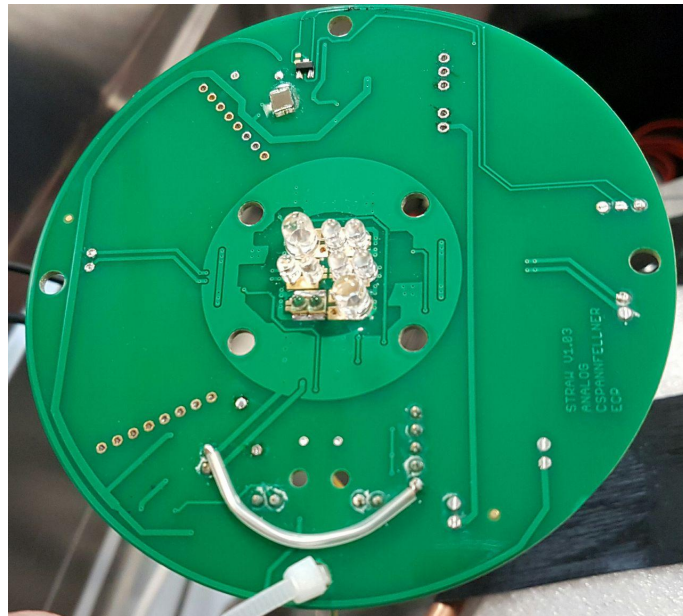
Figure 7: This is the PCB used in STRAW with the LED array.

The LEDs are not sending out isotropic light. This is why a semi-transparent PTFE sphere is mounted onto the LEDs to assure isotropic light transmission. This piece was not important for single photon event data that enters the simulation because only a few photons were recorded with a direct path to the PMT.

In order to run the simulation one LED has to be chosen and the voltage at which it is powered. From these two parameters the first time shift is drawn. Another parameter which has to be set is the average number of photons which are expected to reach the PMT. For each simulated signal the parameter is taken and a Poisson distributed random number is drawn which represents the number of photons detected by the PMT for the simulated signal.

The first time shift according to the point in time a photon has been emitted, which has been detected, is generated with standard deviation fitted to the LED pulse. How many photons are reaching the PMT is set with the Poisson random number according to

the expected number of photons. With these parameters the random Gauss numbers are drawn with the center being at zero. Because only the absolute different arrival times are important negative numbers are taken as well.

If the Poisson random number drawn according to the average photons is zero then the simulation adds zero peaks to the list for the number of peaks for a histogram of number of peaks. After that it skips the simulation of zero photons detected and runs the next signal simulation.

3.2   Photomultiplier Tube

Each photon entering the PMT has a time shift according to the point in time it has been emitted. In reality the photon hits the photo-cathode and causes a photo-electron to be sent out. This electron then gets multiplied by dynodes until the electrons get absorbed by the anode which causes a current pulse being sent to the PADIWA.

The time from converting the photon to an electron and the jitter of time an electron needs to pass through the PMT is called transit time spread, TTS. This time shift has to be added to each simulated photon entering the PMT in form of a random number.

The TTS is given by the spec sheet, it is used as the deviation of the normal distribution. A random Gauss number is drawn for each photon to represent the TTS.

The time shift from the LED and the PMT is added together for each particle. This time shift determines the recorded arrival of the signal.

3.3   Single Photon Peaks

Each photon that is recorded by the PMT produces a signal which is known as a single photon event. Single photon peaks have been recorded which are stored in sets of 10000. How they were recorded will be discussed later. Running the simulation determines which set of peaks will be used and does not change every time a signal is simulated. This is done because the time needed to open one file is rather long and causes the simulation to be slowed down because a file has to be opened. Using one set of peaks is not an issue for the simulation because usually 3 to 6 photons were expected to arrive at the PMT which means there are enough recorded peaks to run the simulation multiple times with one set.

If more photons are expected and the random Poisson number becomes too large for a set of 10000 peaks, multiple files could be added together. This would result in more peaks to choose from but the loading time into the memory accrues only once, at the beginning.

Before the simulation chooses peaks from the loaded file a *numpy*[1] array with the dimensions number of photons and data points is created in which later on the peaks will be written for further processing.

Each photon which is recorded by the PMT becomes a recorded single photon signal, chosen from the single photon file. The peak is chosen by the number of the row it is stored in. The row number is generated by a random number generator. An example for a single photon event recorded by the PMT is shown in figure 8.
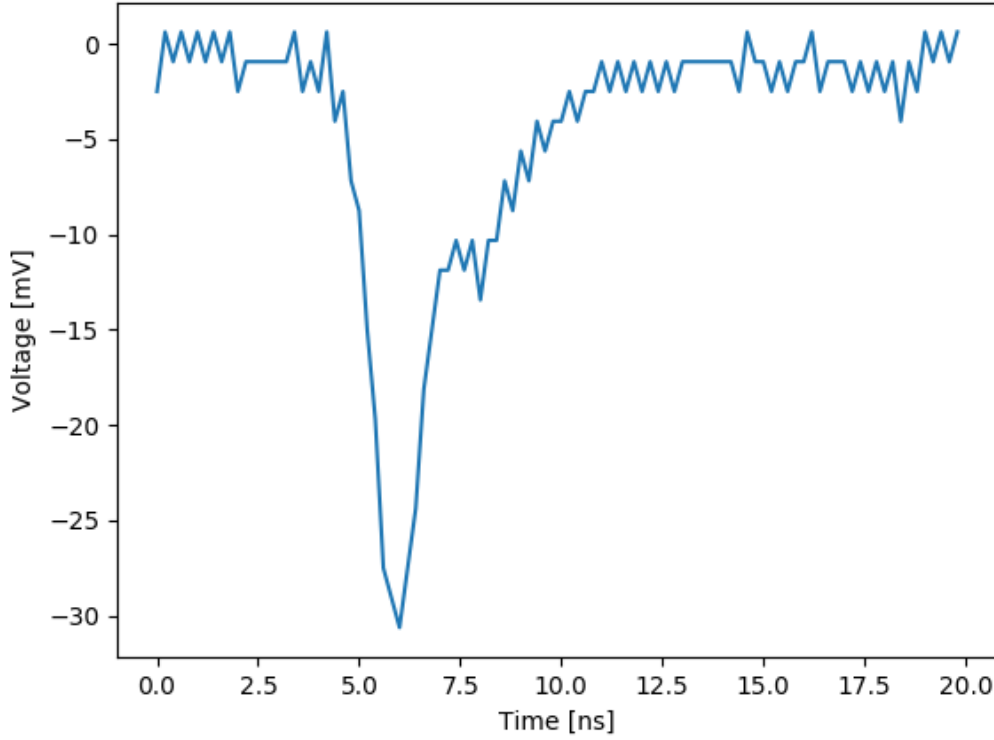
---

[1]www.numpy.org

Figure 8: The single photon peak has a fast rising edge to the first peak and a bumpier falling edge to the second peak which is caused by the electronics.

Once a row has been chosen the data points s which are stored as integer number in FADC counts is saved to a disk space. For the calculations the data has to be converted into mV which is done with equation 3, in which the variables $VRange$ and $Offset$ are used from the picoscope file.

$$s[mV] = s \cdot 2 \cdot \frac{VRange}{2^{16}} - Offset \tag{3}$$

After the data is converted into physical units the time shift has to be added. In order to set the peak at the exact time shift, which is given very precise from the generators, the data has to be interpolated.

The interpolation is done by a $scipy^2$ library. In order to get better results later on more points are interpolated than previously taken.

The interpolation takes the data points given and produces a graph. Now a new range of the axis has to be chosen which includes the maximum time shift that has been generated. The graph then gets asked for points from the new axis. At points where the signal has not yet been detected it results in 0 mV and at points where the signal has already finished it results in 0 mV as well.

---

^{2}www.scipy.org

Every photon goes through this process with its specific time shift. Once the the new data points are generated from the signal with the time shift the data is written to the *numpy* array. Each signal is stored in its own row and they all have the same number of data point according to the new axis that has been given.

When all peaks are written to the *numpy* array the simulation adds the voltages for the same time step together. This results into the real signal one would have measured. The photons are not recorded after each other but coming at different times in which others are already going through the PMT or have not yet arrived. This is why the data points have to be summed up to simulate the real output. An example for a simulated multi photon signal recorded by a PMT is shown in figure 9.
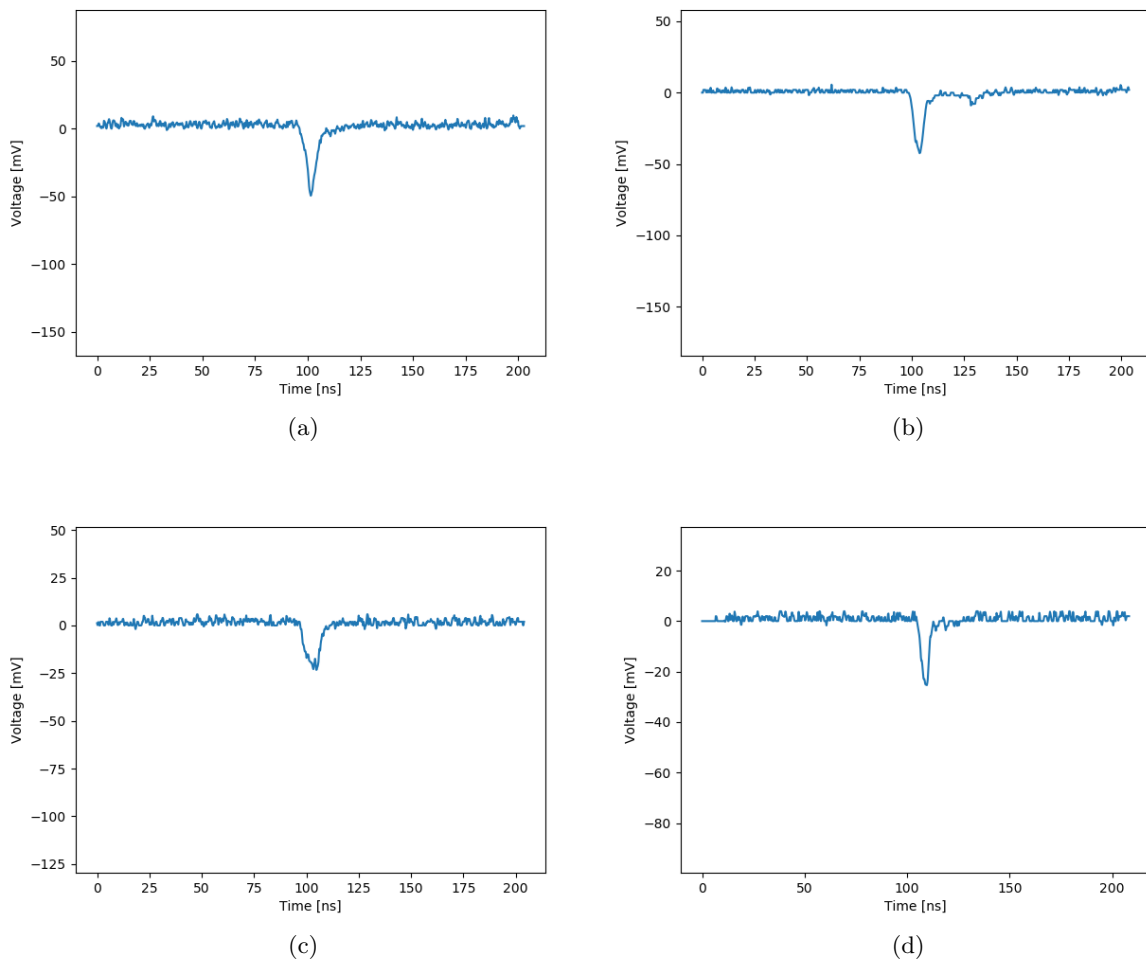

Figure 9: The simulated output from the PMT contains multiple photons added up to a multi photon event.

3.4   PADIWA

The PADIWA is an amplifier. It amplifies the signal coming from the PMT and reshapes it. Theoretically it is a low pass filter combined with a high pass filter. The simulation has to account for the physical elements used such as resistors and capacitors. Figure 10 shows the electric circuit of the PADIWA.
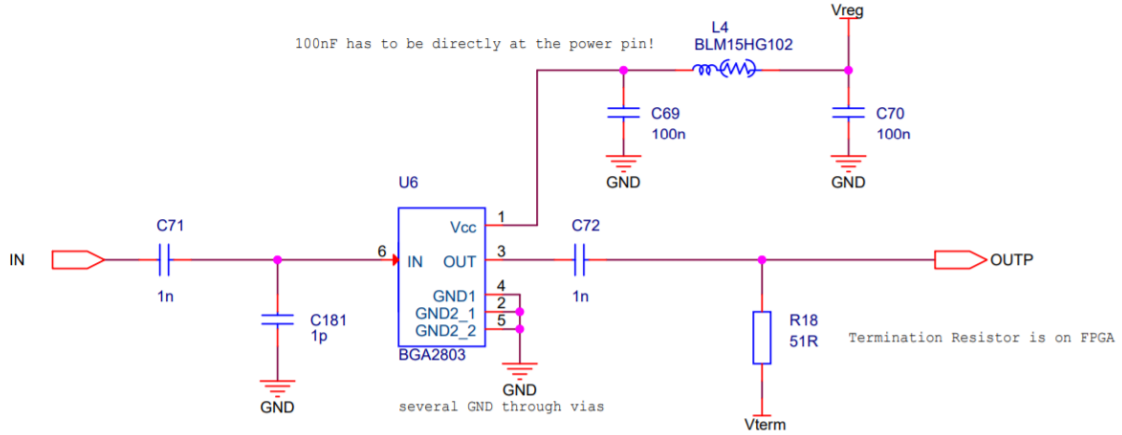

Figure 10: This is the electric circuit of a single input channel of the PADIWA [12] which partly had to be simulated. The capacitors where switched out, C71 and C72 is a 220 nF C181 is a 1 pF.

The signal enters on the left and the first capacitor has to be loaded. After that the current loads the second capacitor connected to ground. The electric component BGA2803 which is the wideband amplifier with internal matching circuit in a 6-pin SOT363 plastic SMD package modifies the signal. Leaving the amplifier another capacitor has to be loaded.

The simulation has to calculate all these electric components. The effect of the capacitors was simulated by numerically solving the differential equation. Additional effects, probably caused by the amplifier itself had to be compensated by an additional smearing kernel in order to produce the measured real PADIWA amplifier output. The amplifier itself has to be simulated by convolving a kernel consisting of variables corresponding to frequency cuts with the data coming from the calculation.

The output from the PADIWA is determined in three equations representing figure 11. Equation 4 calculates the current I_1, starting with the potential difference U_0 passing through the C181 capacitor witch a capacitance C_1 up to the amplifier BGA2803 with a resistance R_1.

$$I_1 = \frac{U_1}{R_1} = \frac{U_0}{R_1} - \frac{\int_0^t I_1 dt'}{C_1 \cdot R_1} \tag{4}$$

Equation 5 is representing the loss of current $I^{'}$ over capacitor C181 with a small parasite resistor $R^{'}$ connected to the ground.

$$I_2 = I_1 - I^{'} = I_1 - \left(\frac{U_1}{R^{'}} - \frac{\int_0^t I_1 - I_2 dt^{'}}{C_2 \cdot R^{'}} \right) \tag{5}$$

The signal is leaving the PADIWA over the last capacitor, C72, and a termination resistor, calculated in equation 6.

$$I_1 = \frac{I_2}{R_3} - \frac{\int_0^t I_3 dt^{'}}{C_3 \cdot R_3} \tag{6}$$

The simulation has to be scaled by a constant factor which has been determined to be about 8. More explanation on how the scaling factor was measured will be given later. The simulations output for the PADIWA is shown in figure 11.
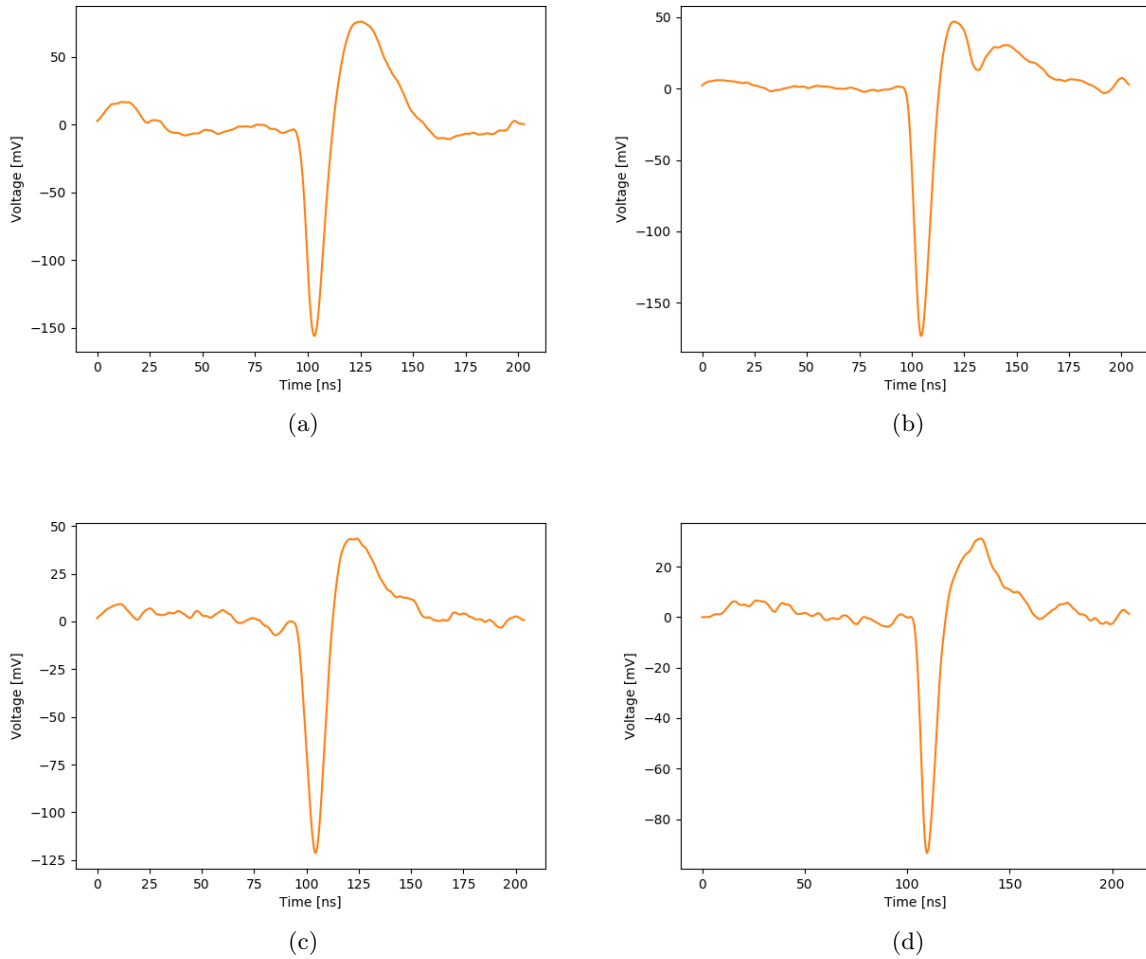

(a)

(b)

(c)

(d)

Figure 11: The multi photon event from the PMT is modified by the simulated PADIWA and has a characteristic overswing after the pulse which is caused by the capacitors.

3.5   Trigger-Readout-Board

The TRB has a very exact absolute clock which is used to record the data. The PADIWA has 4 thresholds triggering the TRB. These thresholds are triggered once the voltage of the signal surpasses it and when the signal falls below these thresholds. Every threshold is twice as high as the one before.

The simulation goes through all data points which are above the threshold and stores the edges in an array. The information that is stored is the time stamps at which the threshold is crossed.

These points correspond to the real measurement STRAW does. The simulations output at the end is in the same form but not like the real measurement more parameters are known. Such as how many photons were detected which energies do they correspond to and what was the time spread that lead to its shape with the whole graph.

After that the simulation plots the results. The thresholds and the simulated signal coming from the PADIWA with the time stamps from the TRB are shown in figure 12.
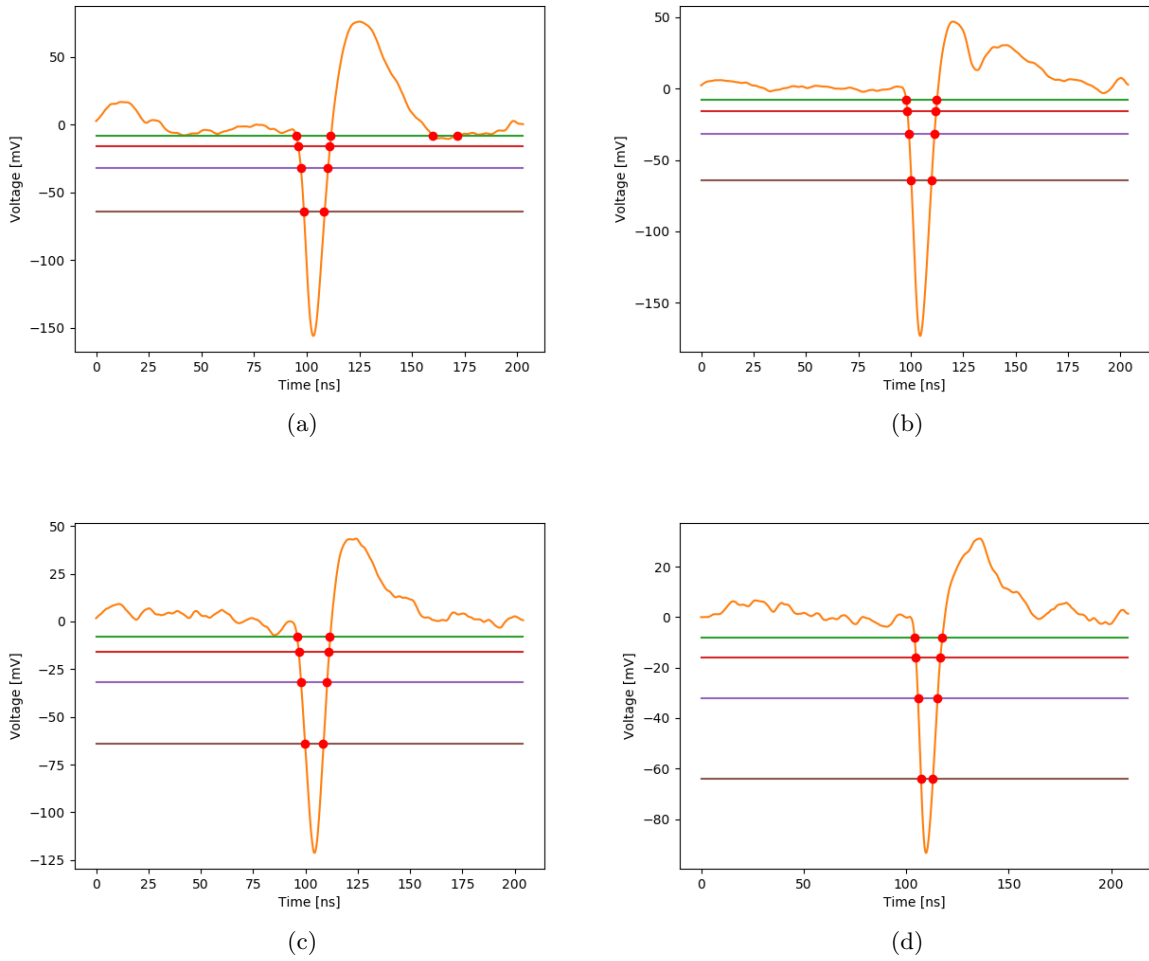

Figure 12: The TRBs threshold are symbolized by the green, red, violet and brown line which are each twice as high as the one before. The red dotes are the data points at which the time stamp is taken and represents the output from STRAW.

In order to show the accuracy of simulating the PADIWA the PMT output and the PADIWA output was measured simultaneously. The simulation then took the PMT output and simulated the PADIWA output which is shown in figure 13. The overshoot is higher in the simulation but STRAW does not measure these regions and is only interested in the negative peak caused by the photons.
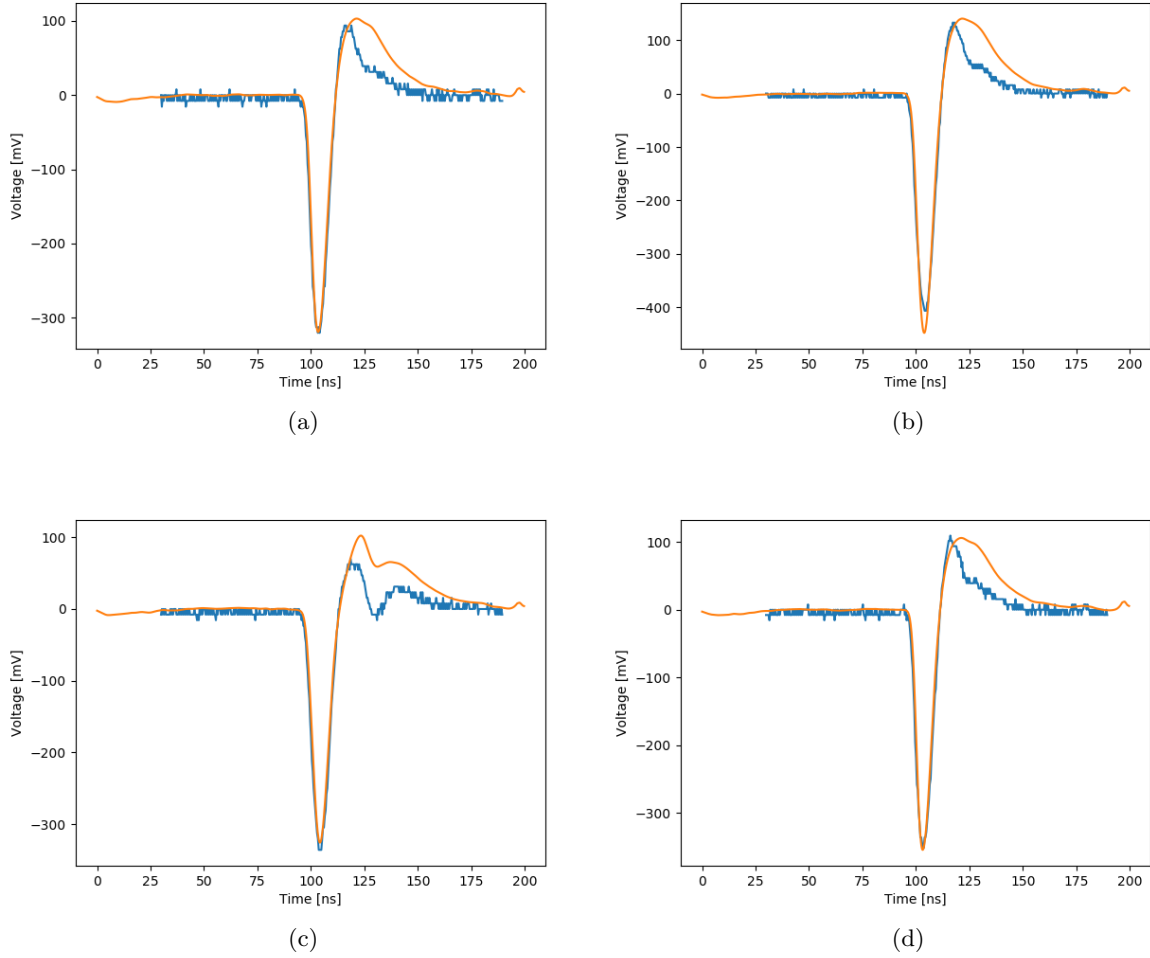

Figure 13: The PADIWA output (blue) was recorded with a picoscope. The simulated PADIWA output (orange) was generated using the measured PMT output from the same event as the PADIWA output.

4   Producing Simulated Pulses Step by Step

The first step in order to record the Test Sample was to write a python protocol that runs the PicoScope 6403D with which everything was recorded. This was achieved with the python library *picoscope*[3] from GitHub. The picoscope was connected to a PC via a USB cable.

The sensitive PMT was put into a metal box with black cloth inlay to keep light from entering. Thermally induced electrons and scintillation light produced by glass hemisphere were used to record single photon events.

4.1   Dark Rate and Gain

The PMT which was the same model as the ones build into STRAW was connected to the electronics of an sDOM, which powered the PMT and set the voltage with which the gain was determined. Channel A from the picoscope was connected to the output from the PMT.

The first set of data was recorded without a specific trigger coupled with $DC50$. The recording was 200 ms long with a sample interval of $0,2$ ns. The data was written into a python dictionary with important settings. The recorded data itself was written into a *numpy* array and the offset, range, sample interval and duration were written as integers. This resulted in a huge file size of over 1 Gb.

From the recorded data the peaks had to be detected and cut out. The background noise was dominant but with the right parameters the peaks were detected and each peak was stored into a new file.

In order to detect these peaks and cut only the peaks without the background a python script was written. The script loaded the recorded file and looped through the recorded data points which where higher than a set threshold. After the threshold was surpassed the script looked for peaks which are made out of one single data point and for peaks which are flat for a couple of data points. The peaks have a very smooth and steep rising edge therefore only triggering at a real peak was not a problem. After the first peak the signal has a characteristic second smaller bump which is caused by the PMT. To avoid triggering at the smaller peaks and the bumpy falling edge a cover of several nanoseconds had to be laid over the data starting with the fist peak it found. Under the cover the script did not look for peaks.

Once the highest point was found the data points around it that represent the whole peak were stored into a new file as mentioned. These files were significantly smaller, every recorded data set produced about 200 to 300 peaks.

Some of the peaks that where stored are after-pulses from of the PMT and not single photon events. Another python script was written to sort out data which had extremely high edges or to many peaks in one file. This reduced one recorded set by 10 to 15 files.

These methods were used to record the first 10000 peaks. The scripts used were automated with a bash script that repeatedly called each python script in order.

---

[3]https://github.com/colinoflynn/pico-python

The build in PMTs in STRAW were previously tested and the gain was set to $15 \cdot 10^6$. The mean multiplier gain, is the ratio between the anode charge, produced by a single photo electron and the electron charge. The PMT that was tested had not been previously calibrated regarding the gain.

In order to determine the gain the recorded single photon events were used. At first the voltage U for an event was calculated by summing up all data points and equation 7 with the resistor R at 50 kOhm.

$$U = \frac{R}{I} \tag{7}$$

A script was written that calculated the charges of all peaks. As with the PMTs from STRAW the balance point of all values was ascertained and the lowest 20 percent were cut away as well as everything bigger than 2,5 times the balance point. The gain G was calculated according to equation 8 for one peak where e is the charge of one electron.

$$G = \frac{U}{e} \tag{8}$$

In order to set that gain to $15 \cdot 10^6$ the voltage had to be adjusted. After trying out a couple of settings, the gain was achieved at $1,280$ V.

The power supply from the sDOM caused too much background noise for recording clean data for a simulation. This is why a different more reliable power supply was used to power the PMT and later on the LEDs. With the new power supply the screening script had to be updated because the parameters for the background changed. This increased the number of peaks found per minute to a point where recording a large amount of peaks was possible. With the gain set to $15 \cdot 10^6$ another 10000 peaks were recorded.

4.2   First Simulation Steps

The first step in order to simulate the signal chain was to understand the LEDs which send out photons. Measurements for the voltage corresponding to the intensity have already been done. The voltage in range of 10 to 20 V have been tested and recorded. The results showed a linear dependency between voltage and intensity. For STRAW which is designed for the deep sea these middle to high intensities are important because of the absorption from the water and bio-mass.

During the measurement for the simulation the PMT and the flasher were in a closed box. The box was big enough to fit the equipment which was also covered in black cloth to keep it as dark as possible. This is why a high light intensity was not necessary for our experiment because the absorption was so little that in order to only let a few photons reach the PMT a partially transmitting foil had to be used. The foil was taped on top of the LEDs. The LEDs pointed toward the PMT. The voltages tested ranged from 5 V to about 8 V.

The LEDs used were mounted onto a PCB using the same flasher circuit that is also used by STRAW, the same used in STRAW. In order to power one LED at a time a power supply was connected that supplied the specific voltage. A wave-generator was

then connected to the LED That set the pulse with which the LED is set to flash. The frequency of the pulses was set to 1 kHz with a pulse width of 20 μs.

The first multi photon peaks were measured with the same scripts as the single photon peaks. The data recorded looked as expected. The shape resembled photons arriving with different time shifts adding up to a signal representing the sum of the photons.

After analyzing these larger peaks the first step of the simulation was written in python. This included simulating which photons were reaching the PMT. An average number of photons reaching the PMT was set as a parameter. This parameter is then given to a random Poisson number generator which sets the number of photons arriving at the PMT in the simulation.

The first parameter causing the first time shift for each photon individually is determined by the LED used and the voltage set. With the linear dependency for the specific LED the deviation for a random Gauss number generator is set.

The second time shift for each individual photon is corresponding to the time the signal needs to travel through the PMT and get detected. This time spread is generated with the already known travel time as deviation for a random Gauss number generator.

The previously measured single photon events were used as source for a random choice for each photon. The photons in the simulation are known when they are detected and what signal shape each photon caused.

The next step in the simulation was summing up the signals at the point of their arrival. This was done by interpolating each photon peak at the specific point on the time axis which is the point of detection.

With the time shift simulated the signals produced by the simulation were able to reproduce data recorded from the experiment. In order to verify this step histograms where made. Two parameters were investigated, the full with at half maximum, FWHM, and the number of peaks.

The FWHM was achieved by determining the highest point from the signal and checking how many points were above this value. Each point is followed by another point which represents a time step of $0, 2$ ns. Adding the number of points up and multiplying it with the time step results in the FWHM value for the peak.

Counting the peaks for each event was not possible with the scripts written for cutting out peaks from raw data because these scripts relayed on a smooth rising edge and after finding one peak they are not able to find another one in the same signal.

The solution for this problem was found on GitHub. The python script *detect_peaks*[4] is able to find peaks in the data, an example is shown in figure 14. The Amplitude was turned upside down because the script is finding positive peaks.

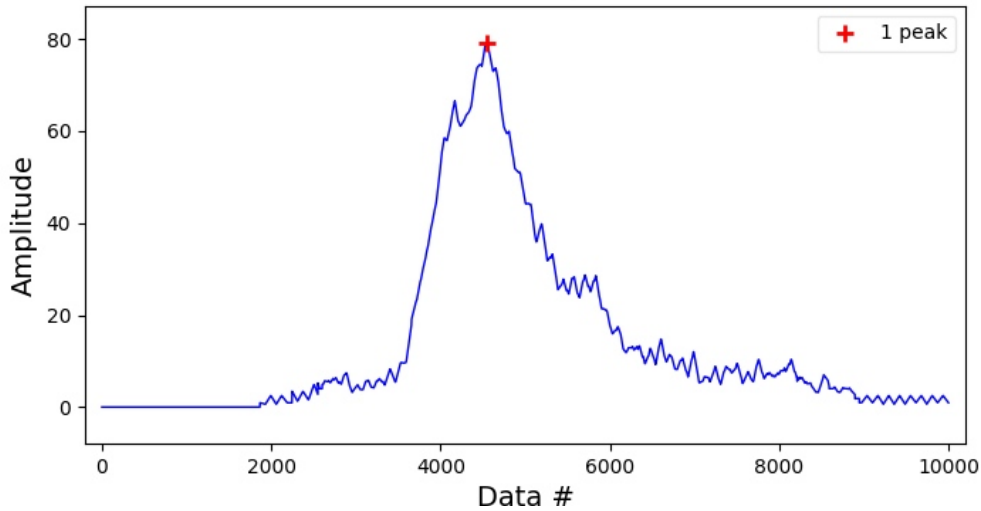[4]https://github.com/MonsieurV/py-findpeaks

Figure 14: This is a plotted representation of a multi photon event which has been turned upside down with one peak (blue) and the found peak is marked with a red plus. The plot is generated by the modified script and its output would be one peak.

The script has a few parameters to determine what is a peak but for the peaks that have been recorded the possible settings were not enough because the falling edge of the peaks are very bumpy. Furthermore peaks which were flat for a few data points were also not detected but theses are valid peaks for our experiment. This is why the script had to be modified.

After modifying the code according to the data that has to be analyzed the results were accurate. Another histogram shows that the simulated peaks are similar to the peaks that were recorded.

The recorded multi photon peaks so far are made up of photons from the flasher and single photon events. The flasher is powered by the wave-generator which then was used to record these multi photon events.

The previously used script to record data from the picoscope was modified using a trigger. The wave-generator was set as the trigger for the picoscope. The recorded signal was therefore only photons which came from the LED. The wave-generation signal was split and the second line was connected to channel C from the picoscope. Channel C was used because using B caused the possible sample interval to drop to only $0, 8$ ns. By using C the interval of $0, 4$ ns was achieved.

The coupling was also changed to AC. The picoscope did not take the time to load the capacitors for alternative current. This has to be done by setting up a dummy run in the python script which is long enough to load the capacitors in order to overcome this bug in the python picoscope library.

Another important change that was possible by using a trigger was that the peaks were now stored in one single file. With these changes recording data became less of an time consuming task and screening for falsely recorded background was unnecessary.

4.3   Simulation of the PADIWA

The PADIWA combines a wideband input amplifier with an electronically adjustable discriminator. The first attempt of simulating this electric component was theoretical. By transforming the data from the PMT into the Fourier room a high and low pass filter was simulated.

Using a python library the one-dimensional signal was transformed into the Fourier room. The Fast Fourier Transformation algorithm is a very efficient way of transforming and requires very little computing time.

The data from the transformation represents frequencies. A low pass filter is cutting low frequencies. This was simulated by multiplying the data points with a rising curve from 0 to 1. The high pass filter cuts high frequencies. This was achieved the same way with a curve falling from 1 to 0. After cutting the data it was transformed back into the real room.

Recording the data coming from the PADIWA was achieved by soldering cables directly onto the board. Figure 10 shows the electric circuit. Using the arbitrary function generator as trigger the output from the PADIWA was measured.

The data simulated with this method was not looking like the signal that was expected. It did not have the characteristic trends. This is because the amplifier has resistors and capacitors which have to be processed in a practical way.

By calculating the signal going through the electric components and convolving the result with a cutting curve the signal looked like expected. The only parameter needed for this calculation was the amount of how much the amplifier raises the values. This value was acquired by measuring the PMT output and the PADIWA output at the same time. Comparing the two highest points, results in a constant value which does not have to be changed anymore. The signal is about 8 times higher.

4.4   Simulation of the Trigger-Readout-Board

The last step in order to complete the simulation is to simulate the thresholds recording the rising and falling edges. This is done by finding the first value exceeding a threshold. This triggers a function that stores the time stamp from this point. This results into the simulation having the same output like STRAW.

This was achieved by finding the first value larger than the threshold and writing down the time stamp from that data point. At first a *scypy* library was used to find the root. But this method already results in a very accurate time stamp because the interpolations that have been done previously produced a very high resolution. The same is done with the falling edges, once the signal falls below the threshold the time stamp from that data point is taken.

5   Data vs Simulation

In order to test the simulation a specific set of data has to be taken. The aim for this test is to prove that the simulation reflects the real world results. For each data analyses plot 10000 peaks are analyzed. This is done to keep the run times in a reasonable time frame and it is enough to compare the results.

The first run includes the PMT output and the PADIWA output without the flasher being active. This is done to record single photon events from both devices.

Knowing these two, the amplification from the PADIWA can be calculated by comparing them to each other. This value is only relevant once but has to be determined. The result is that the PADIWA is amplifying by a value of about 8.

The second run takes the data that will be simulated, the PADIWA output. It is recorded using the trigger from the LED which is set to flash. This means that the LED sends out photons which are detected by the PMT and then are modified by the PADIWA.

The third run includes the PMT output and the PADIWA output with the LED active. The trigger was set onto the PMT output. This results in data which contains the multi event signal from both devices.

After all three runs the average number of photons coming from the LED can be calculated by dividing the multi peak signals from run three with the single photon events from run one. This has to be done every time because if the foil or the array itself is moved the number of photons arriving could change.

Knowing the number of photons and having the single photon events from the first run the simulation can predict the PADIWA output from run two. Run two is used to conclude if the simulation is correct.

5.1   Photomultiplier Tube

The fist parameter is the full with at half maximum, FWHM. This value resembles the duration in time at which the measured voltage is above half the maximum value. It is used to determine the shape of the peak.

Figure 15 on the left shows the histogram containing the real world data analyses. Most of the peaks have a FWHM value of 10 ns or more. The tallest peaks have the largest FWHM value because it takes a longer time for the signal to rise and fall. In the shorter FWHM area the number of vales drops and then rises again. This is caused by noise in the signal which can be observed in all histograms. The high end also shows a very smooth slope of up to 23 ns which shows that the cut off in the higher region was not set to low and all signals were detected.

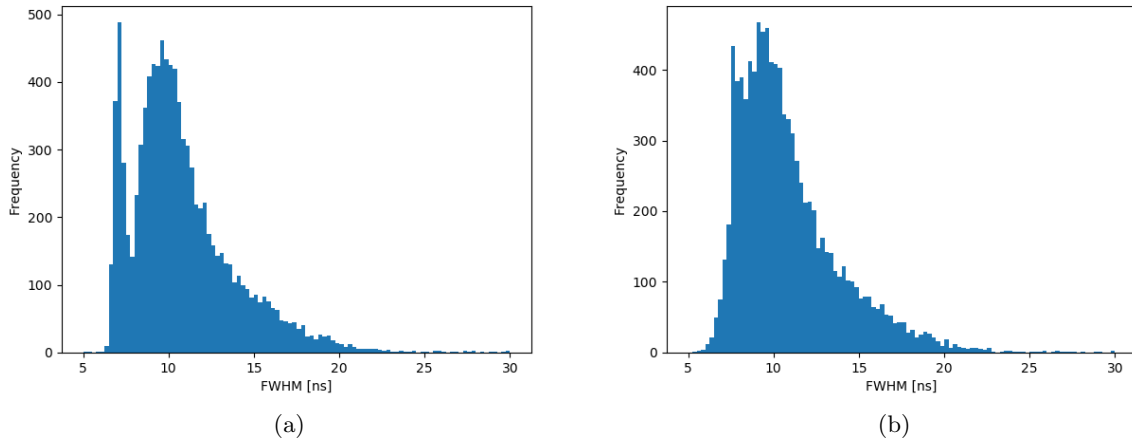(a)   (b)

Figure 15: The histograms are showing the FWHM values for the data taken from the experiment (a) and the simulated signal (b) coming from the PMT.

The simulated data is shown in figure 15 on the right. The important peaks have a FWHM of 10 ns or higher. The characteristic of the values in the low end are similar to the real data but the recorded single photon peaks are less noisy. The high end goes up to 23.5 ns with a steady slopee.

The first parameter suggests that the simulation of the PMT is accurate. This parameter only shows that the time shape of the signal is correctly simulated but not that the shape is correct. This is why the number of peaks for each results were counted.

The number of peaks in the PMT output shows if the shape is the same as the simulated data and not just if the width is the same. The data taken from the PMT does not account for zero photon events that were not recorded because the trigger only recorded events. The simulation sometimes draws zero single photon peaks from the pool and therefore the peaks from the simulated PMT are also zero.

The results from counting the peaks are at first not significant. The data from the PMT in both cases show that only one peak is detected. It can be concluded that if all signals have only one peak and the width is correct the signal itself is simulated very accurate to the real world.

The assumption that the LED has a linear dependency to the voltage down to 5 V is mostly correct because if it was not there would be a discrepancy in the number of peaks recorded.

5.2 PADIWA

Simulating the PADIWA was very time consuming because the multiple integrations over the interpolated data with a lot of data points is very CPU intensive. Nevertheless, to get the best results the resolution was set high.

Because the PADIWA is the last part in the chain that has to be simulated three parameters were investigated. The first parameter is the FWHM, the second is the number of peaks and the third is the area under the curve of the peak.

The full width at half maximum of the recorded data has the most values at 22.5 ns shown in figure 16 on the left. The same trend for smaller values is seen like before. The higher region is falling very rapidly after 24.5 ns. All events were recorded because the edges have a very steady trend.

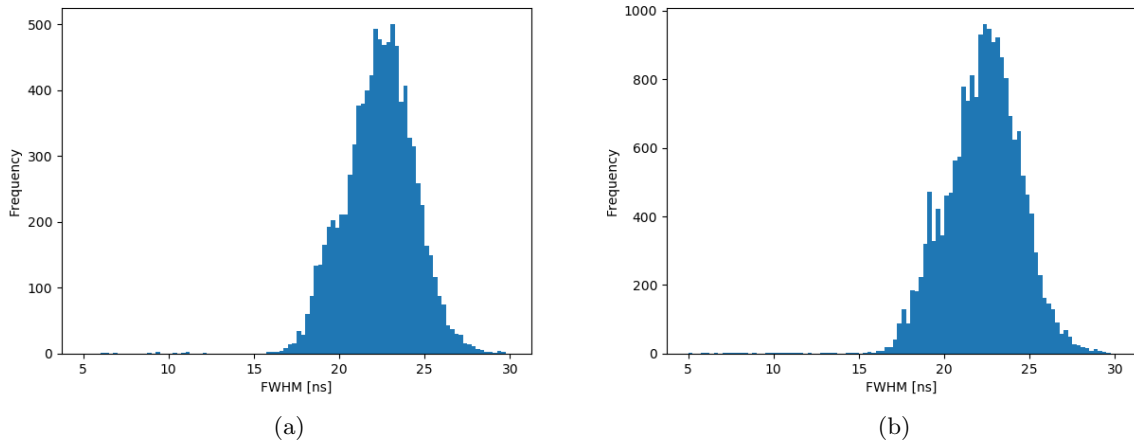(a)   (b)

Figure 16: The histograms are showing the FWHM values for the data recorded from the experiment (a) and the simulated signal (b) coming from the PADIWA.

The FWHM values for the simulated data is shown in figure 16 on the right. Most values are at 24 ns and above. The lower end shows that the values are starting with a width of 16 ns like the recorded data. The higher end goes up to 30 ns with a steady slopee.

The second parameter that has been investigated is the number of peaks in the output form the recorded data and the simulated data. Signals coming from the PMT have only one peak. After the PADIWA the real world data does not change the number of peaks. The recorded data from the PADIWA has one peak. The simulated data has the same characteristic as the recorded data it only simulates data with one peak.

The PADIWA output has the same number of peaks, one, like the simulated data. The FWHM value distribution form the recorded data is showing the same trends and values as the simulated data.

In order to give more proof that the simulation is representing the detector another value is investigated. The area under the generated curve is the sum of the recorded millivolt and represents a charge. The taken data from the experiment shows again the characteristic trend for smaller values which here corresponds to higher numbers seen in figure 17. The balance point of the histogram is at about -7550 mV. The edges for the middle part are falling smoothly with a few values higher than -15000 mV which are after-pulses.

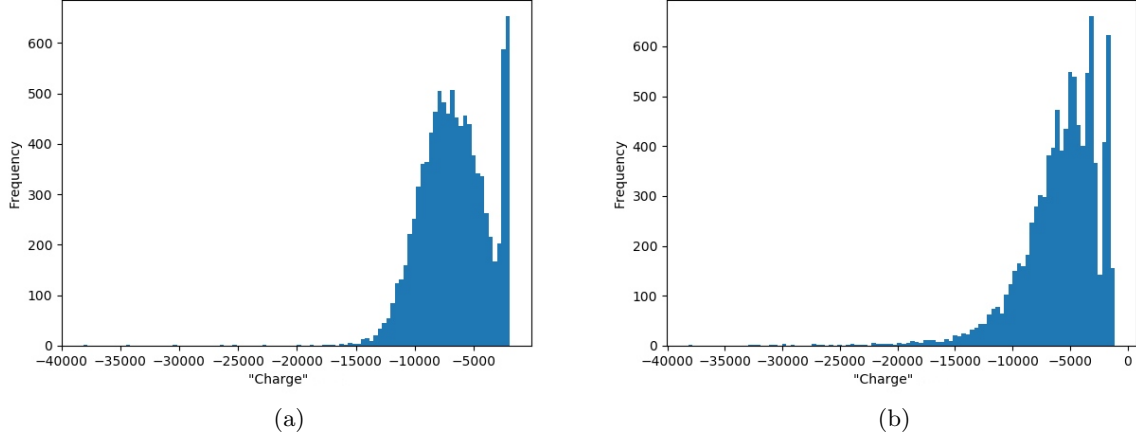(a)   (b)

Figure 17: The histogram shows the charge for the taken data (a) and the simulated data (b) after the whole signal chain.

The histogram for the area under the curve for simulated data is shown in figure 17. The histogram for the simulated data has its balance point at -7500 mV. This is close to the measured results and the highest values for the recorded data and the simulation are both going up to -15000 mV. The simulation shows a few results with higher vales but not in significant amount. The area for the smaller region has a balance point at -2000 mV which is the same as the recorded data. This concludes that the area under the curve is the same for the real world data and the simulated data.

The analyzed values for the PADIWA simulation is concordant with the measured values. The number of peaks are the same, one, the FWHM are of the same trend and value as the recorded data as well as the simulated charge.

6   Conclusion

The simulation procedure presented in this thesis can effectively reproduce the behaviour of the real signal chain of the STRAW sDOM. Number of peaks and the full width at the half maximum are nearly identical. Simulating the effects on the signal induced by the amplifier of the PADIWA has been challenging as simple theoretical (bandpass filter) models failed and the output had to be reproduced by a numerical procedure. But the simulation is able to produce signals with the same shape as the real PADIWA and the signal represents the same energy.

Furthermore the simulation is able to give an output which is the same format as the data taken from STRAW. This was the main point for further applications.

The simulation gives insight into how the electronic signal observed and recorded by the PADIWA and TRB is influenced by the different components along the signal chain. In the future, MC simulations of the whole detector are required. An example for the whole simulation for multi photon peaks is given in figure 18 which includes the output from the PMT, the output from the PADIWA and the threshold time stamps.
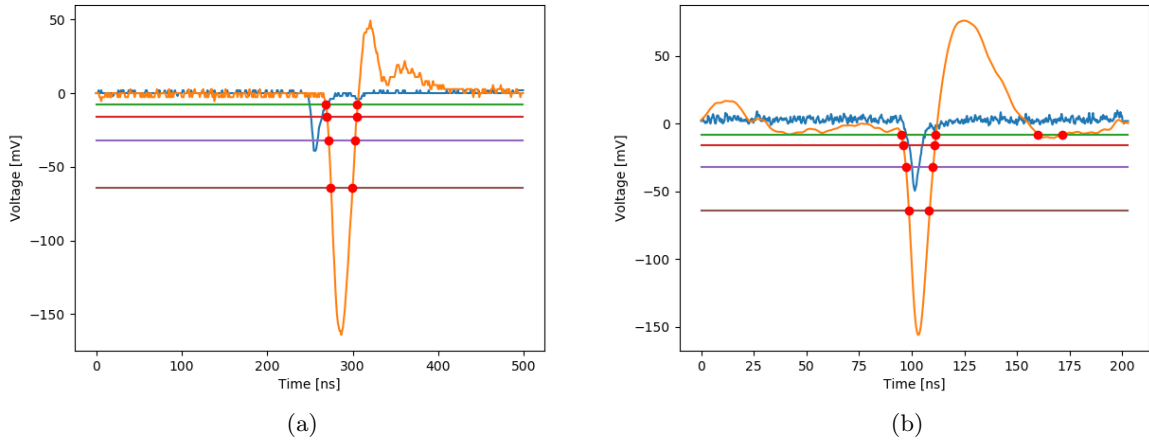

Figure 18: The left figure shows the recorded data from the PMT (blue) and the PADIWA (orange) as well as the symbolized thresholds at which the time stamps (red dots) are taken. The right side shows a simulated output from the PMT (blue) and the PADIWA (orange) with the thresholds and time stamps.

Due to the limited amount of time that could be invested in the frame of a Bachelors thesis, further tuning of the parameters or other improvements might still be necessary.

7   Outlook and open Issues

At the moment there is no way of reconstructing the data taken from STRAW to determine the number of photons that have been detected.

The output from the simulation is the same as the data output form STRAW. STRAW returns the time stamps from the rising and falling edges at the thresholds. A feature that would take these points from the real experiment and tries to fit a simulated signal to these points could be very useful. This would give very good insight of how many photons were measured and how much time the photons needed to arrive at the detector. And from just a couple of points the whole graph could be simulated.

By fitting the simulation to recorded data the height and trend of the measured event can be simulated. With these information the intensity recorded by the sDOM can be determined. By calculating the recorded intensity and knowing the intensity the LED was flashed the absorption length and scattering length can be determined.

Due to the limited amount of time there are also things that should be perfected. The time needed to simulate a signal could be reduced by leaving out the histograms which are only used to analyze the simulation.

Measurements with the LED set to higher voltages would be interesting too. In order to take data in the range from 10 to 20 V more transparent layers of foil should be attached to the LED in order to keep the number of photons hitting the PMT low.

Acknowledgements

First I would like to thank Prof. Elisa RESCONI for offering me a place in here group to write my bachelor thesis.

I also want to thank Dr. Christian FRUCK for coming up with such a great topic and taking the time to supervise me and in particular providing me with corrections and suggestions during the writing process for this thesis.

Furthermore I want to thank Michael ENGELHARD for handing me his LaTeX template and Alexander SZAMBOR for helping me find most of the spelling mistakes.

References

[1]   C. L. Cowan et al. "Detection of the Free Neutrino: a Confirmation." In: *Science* 124.3212 (1956), pp. 103–104.

[2]   R. Abbasi et al. "Measurement of the anisotropy of cosmic-ray arrival directions with IceCube." In: *The Astrophysical Journal Letters* 718.2 ().

[3]   Davide Anguita et al. "Smart Plankton: a Nature Inspired Underwater Wireless Sensor Network." In: (2008).

[4]   M.R.; E.A. Widder; J.F. Case Bowlby. "Patterns of stimulated bioluminescence in two pyrosomes (Tunicata: Pyrosomatidae)." In: *Biological Bulletin* 179.3212 (1990), pp. 1340–350.

[5]   *Collaborating Organizations.* https://icecube.wisc.edu/collaboration/institutions, 2017.

[6]   *Frequently Asked Questions.* https://icecube.wisc.edu/about/faqcost.

[7]   Felix Henningsen. "Optical Characterization of the Deep Pacific Ocean: Development of an Optical Sensor Array for a Future Neutrino Telescope." MA thesis. Munich, Tech. U., 2019.

[8]   M. G. Aartsen et al. (IceCube Collaboration). "Energy Reconstruction Methods in the IceCube Neutrino Telescope." In: *JINST* 9.P03009 (2014).

[9]   M. G. Aartsen et al. (IceCube Collaboration). "Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector." In: *Science* 342.1242856 (2013).

[10] R. Abbasi; et al. (IceCube Collaboration). "The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping." In: *Nuclear Instruments and Methods* 601.3 (2009).

[11] *IceCube Neutrino Observatory Construction Completed.* https://web.archive.org/web/20101225234422/http://icecube.wisc.edu/complete.php, 2010.

[12] M.Traxler. *PADIWA3.* GSI. https://jspc29.x-matter.uni-frankfurt.de/trb/schematics/PADIWA3-all.pdf, 2012.

[13] K. A. Olive et al. (Particle Data Group). "Review of Particle Physics." In: *Chinese Physics C* 38.9 (2014).

[14] E. Resconi T. K. Gaisser R. Engel. "Cosmic rays and particle physics." In: *Cambridge University Press* (2016).

[15] *World's largest neutrino observatory completed at South Pole.* https://news.wisc.edu/worlds-largest-neutrino-observatory-completed-at-south-pole/, 2010.

List of Figures