Measuring the attenuation length of seawater in KM3NeT/ORCA with atmospheric muons

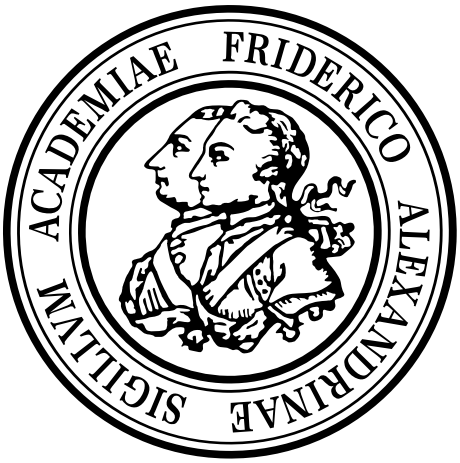
Master's Thesis in Physics

submitted by

Martin Schneider

Date of Distribution: 02. May 2022

Friedrich-Alexander-Universität Erlangen-Nürnberg



Supervisor: PD Dr. Thomas Eberl

Abstract

The KM3NeT neutrino detectors are currently under construction in the deep Mediterranean Sea. KM3NeT-ORCA, the low-energy part of KM3NeT, is an underwater Cherenkov neutrino detector featuring a dense configuration of optical modules designed for the measurement of atmospheric neutrinos down to the few GeV energy regime. As it is located in a deep-sea environment, the detector performance depends on the optical water properties. In this thesis, a study of the optical properties of the sea water surrounding the ORCA detector is presented using a sample of atmospheric muon events obtained with 4 detection units of the KM3NeT-ORCA detector. A dataset with a livetime of approximately 120 hours was used. For this, Monte Carlo simulations obtained by varying the light attenuation length through a change in absorption and scattering length are compared to atmospheric muon data. Using a simple model a fit is performed in an attempt to estimate the attenuation length for the different simulations. Tentatively, a small increase in absorption length does compare to the data favourably.

Contents

Chapter 1

Introduction

Neutrinos were first detected in an experiment in 1956 by Clyde Cowan and Frederick Reines, a feat that was later acknowledged with the Nobel prize in physics in 1995. Soon after this discovery several projects attempted to build a neutrino telescope or wanted to provide a proof of concept. The goal of a neutrino telescope is to search for neutrino sources in the universe. As neutrinos, like photons, do not carry a charge, they are not deflected by magnetic fields on their way to earth and point directly at their origin. Neutrinos only interact via the weak force and can as a result easily pass through interstellar gas clouds. In fact, they may even be used to give early warnings about impending supernovae, since powerful neutrino emission occurs well before the light of a supernova becomes visible. Because of the fact that neutrinos interact only via the weak interaction, a large target mass is needed. Therefore experiments are built deep underground, using the deep sea water as detector medium. This is the case for the ANTARES experiment, while the IceCube experiment located at the South Pole deploys its optical sensors about two kilometers deep into the ice. The Earth can act as a shield in this case, filtering out all particles except neutrinos. The detection of upward-going leptons via Cherenkov light is then used as a signature of neutrino interactions in the matter below the detector.

The ANTARES detector, which once was the largest neutrino detector in the Northern Hemisphere, will be succeeded by the KM3NeT project, also located in the Mediterranean Sea. It is planned to cover about one cubic kilometer of volume and consists of two components ARCA (Astroparticle Research with Cosmics in the Abyss) and ORCA (Oscillations Research with Cosmics in the Abyss) which are designed with different science goals in mind.

The main objectives of the KM3NeT experiment include the discovery and observation of high-energy neutrino sources in the universe and the determination of the mass hierarchy of neutrinos [1]. For this purpose, understanding the absorption and scattering of light in the detector medium is fundamental to be able to optimize the performance of the detector. While atmospheric muons mostly represent unwanted background when observing neutrino sources in the universe, they may be used as a frequent and steady source to study the detector medium.

The seawater optical and oceanographic properties of the KM3NeT-It site were first investigated in August of 2002 [2]. In this study the absorption and scattering

of light in seawater was measured with an accuracy of approximately 10% using standard oceanographic probes and a transmissometer. Whereas the scattering and absorption from pure seawater (salt and water) is well-determined [3], the main contribution to the uncertainty lies in the presence of fine, dust-like particles. With the KM3NeT calibration system in place that consists of monitoring tools, probes installed on the detector elements and dedicated calibration units [4], there already exists a powerful tool to further improve on this knowledge. A variation of the so-called NEMO water model is in use in the KM3NeT project today. Given that dedicated sea operations to measure the optical water properties at the detector site are costly and additionally require a significant amount of labour, researchers explore additional opportunities to obtain measurements for these quantities.

This master thesis aims to use atmospheric muon data to try to measure the attenuation length of seawater in the KM3NeT/ORCA detector. This thesis is structured as follows: First the KM3NeT detector is introduced. Then atmospheric muons and their interaction in the detector will be briefly explained. After that the different steps to create a Monte Carlo simulation are explained, followed by the description of the Run and Event selection of the data. Finally measurements of the attenuation length will be compared in data and Monte Carlo simulations. For that multiple Monte Carlo simulations with changes in absorption and scattering length have been produced. At last the results are summarized and a conclusion will be given.

Chapter 2

The KM3NeT Detector

KM3NeT is a large-scale underwater instrument for the detection of neutrinos of cosmic and atmospheric origin that has been under construction since 2015. The name of this research project - KM3NeT - is an abbreviation for Cubic Kilometre Neutrino Telescope. Located in the Mediterranean Sea this neutrino telescope will consist of a water Cherenkov detector with an instrumented volume on the order of one cubic kilometer. Currently the KM3NeT project has deployed research equipment at multiple locations in the Mediterranean Sea. The KM3NeT site in Italy (near Portopalo di Capo Passero) features the ARCA (Astroparticle Research with Cosmics) detector which is installed at a depth of 3400 m. It will succeed the ANTARES neutrino telescope which stopped data taking in February of 2022 and continue the search for high-energy cosmic neutrino sources in the universe. The KM3NeT site in France (near Toulon) features the ORCA (Oscillations Research with Cosmics in the Abyss) detector which is installed at a depth of 2475 m. It has a different scientific goal - namely to measure the fundamental properties of neutrinos. As its purpose is not that of a classical neutrino telescope but rather that of a particle physics detector to study neutrino properties, it has been optimised accordingly. In the following, the detector design of ARCA and ORCA is described. The content of this section is largely based on the 'Letter of Intent' [1].

2.1 The KM3NeT building block

The KM3NeT detectors are designed with their respective scientific goal in mind. The conceptual design report [5] showed that configurations with densely instrumented regions are preferable for low neutrino energies, whereas lower density homogeneous designs yield better efficiency for an intermediate energy range (1 TeV to 100 TeV), which is preferable for point source searches. The ARCA and ORCA detector use the same general structure, so-called building blocks. Each building block consists of 115 strings or detection units (DUs) that are anchored to the sea floor and supported by buoys. Each detection unit in turn comprises 18 sensor spheres called digital optical modules (DOMs) and each module is outfitted with 31 photomultiplier tubes (PMTs). The DOMs have a diameter of roughly 43 cm to fit the 7.6 cm PMTs and additional supporting electronics. The KM3NeT building block is used for both ARCA and ORCA but with different inter-string and inter-DOM distances that are suitable for their respective physics case. An artists impression

of a building block is shown in Figure 2.1.

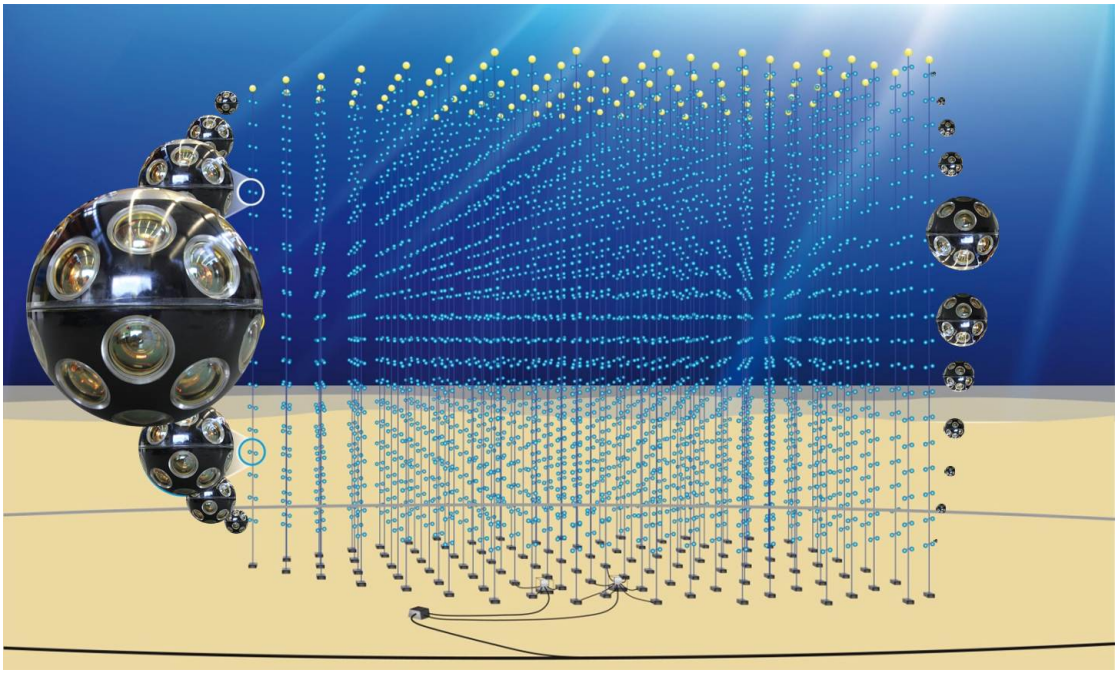


Figure 2.1: A building block of KM3NeT with 115 strings. Each string holds 18 DOMs that contain 31 PMTs. (courtesy Marc de Boer/Ori Ginale) [6]

2.2 The ARCA detector

As the successor to ANTARES, the purpose of the ARCA detector is to measure high-energy cosmic neutrino sources in the universe. A broad spacing of strings and DOMs is preferable to capture Cherenkov light induced by high-energy leptons over larger distances.

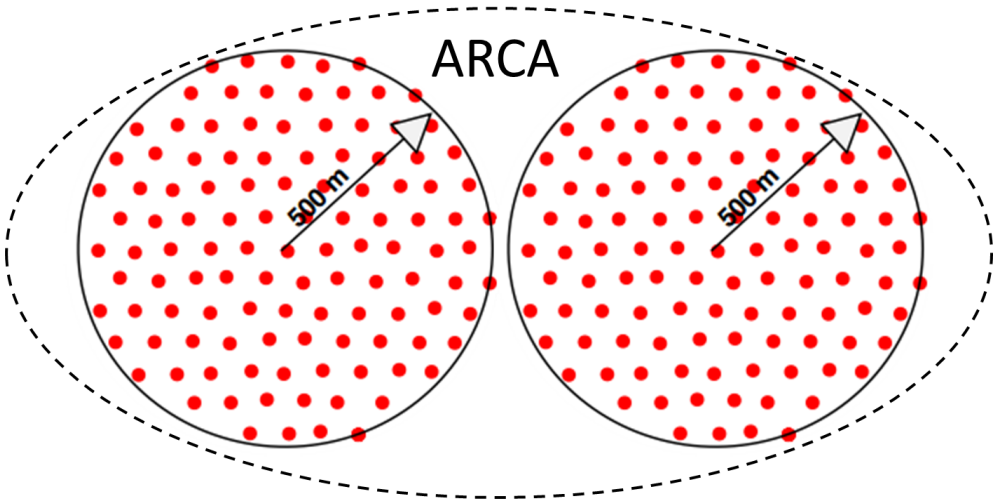


Figure 2.2: Footprint of the finished ARCA detector. Each building block has 115 strings (90 m average spacing) with 18 DOMs each (36 m spacing). The instrumented volume is $0.48\,\mathrm{km}^3$ (R=500 m, z=612 m) [7].

In case of the ARCA detector the 115 strings are spaced 90 m apart on average and the DOMs on the string itself are distributed with a distance of 36 m to each other. This results in an instrumented volume of 0.48 km^3. ARCA is planned to feature two building blocks, a footprint is shown in Figure 2.2. At the time of writing nine detection units have already been deployed, seven of those are currently in operation.

2.3 The ORCA detector

Compared to ARCA, the ORCA detector is much more compact and has only one building block. It is optimized to measure the properties of neutrinos and features a denser configuration which is better suited to lower neutrino energies. The footprint of the ORCA detector is shown in Figure 2.3. The ORCA building block has smaller

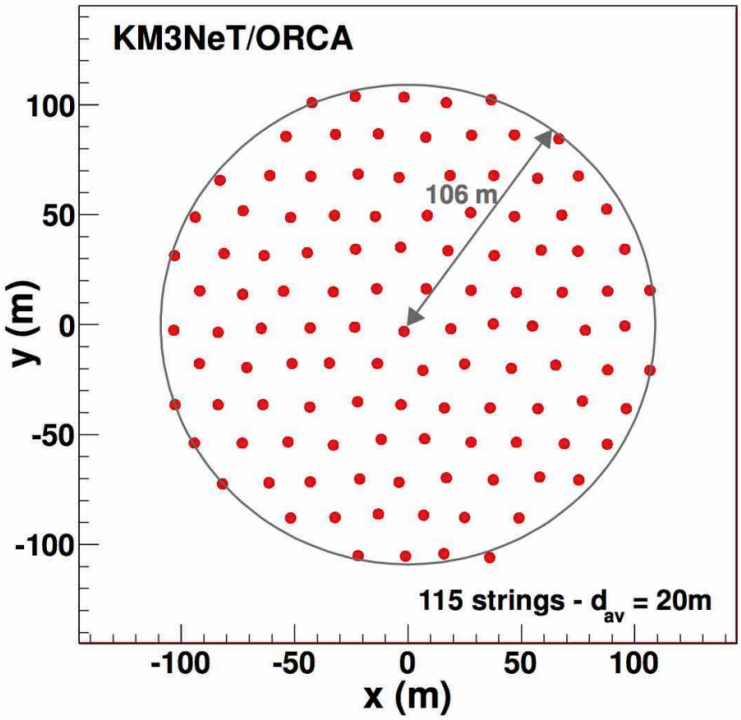


Figure 2.3: Footprint of the ORCA detector, with 115 strings (20 m spacing) with and 18 DOMs each (9 m spacing). The instrumented volume is 5.4 m^3 to 6 m^3 (cylinder: R=106 m, z=153 m) [1]

inter-string and inter-DOM distances of 20 m and 9 m respectively. Currently the ORCA detector features ten detection units, eight of them are actively taking data.

For the purpose of this thesis, namely measuring the attenuation length of sea water, the ORCA detector is well suited. The compact configuration allows for an ample signal of atmospheric muons, especially for a small number of DUs. Additionally with several DOMs in close vicinity, light attenuation effects are expected to be clearly visible.

Chapter 3

Atmospheric Muons

Cosmic radiation reaches the earth from the sun, the milky way and other galaxies in the universe. A major part of this radiation are cosmic ray nuclei (high-energy massive particle like protons and atomic nuclei that move through space at nearly the speed of light). When they interact with air molecules in our atmosphere, they produce a slew of new particles. These secondary cosmic rays often have enough energy to create particles themselves through further collisions and decays, thus creating a cascade called air shower. Muons are a main component of secondary cosmic rays. As muons are unstable particles and decay soon after they are created, they cannot reach us from interstellar space. Hence the name atmospheric muons. The descriptions in this chapter follow, unless stated otherwise, the reference [8] and references therein.

As the cascade in the air shower develops further towards the ground, the particles are less and less energetic since the energy of the incoming cosmic ray is split and

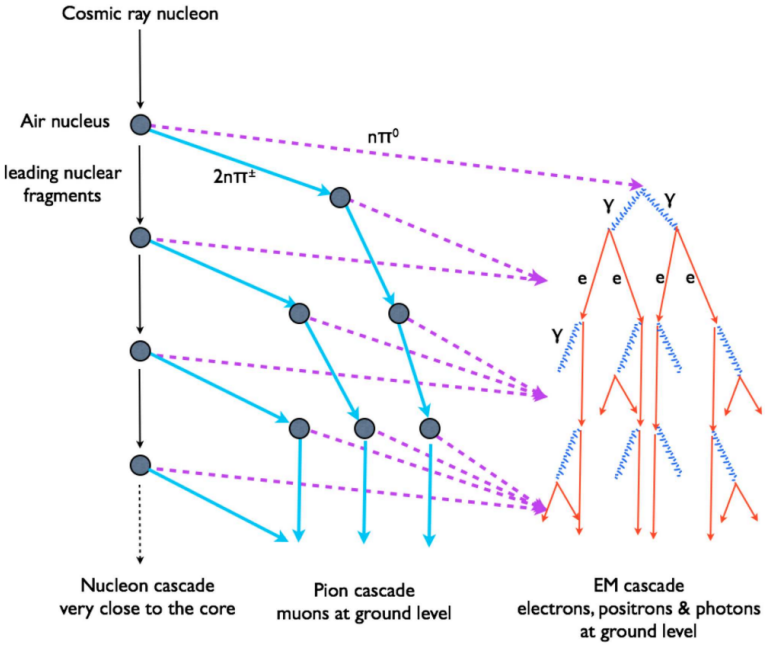


Figure 3.1: Schematic evolution of a cascade initiated by a cosmic ray nucleon. At each step, roughly 1/3 of the energy is transferred from the hadronic cascade to the electromagnetic one. Sketch taken from [9]

redistributed among more and more particles. Additionally, the secondary particles tend to pick up more and more lateral momentum causing the air shower to spread over a large area. A sketch of the temporal evolution of an air shower with the hadronic and electromagnetic component is shown in Figure 3.1. Muons predominantly originate from the decay of secondary charged pions and kaons. The most important decay channels and their respective branching ratios are listed below.

$$\pi^{\pm} \rightarrow \mu^{\pm} \overset{(-)}{\nu}_{\mu} \quad (\sim 100\%) \tag{3.1}$$

$$K^{\pm} \rightarrow \mu^{\pm} \overset{(-)}{\nu}_{\mu} \quad (\sim 63.5\%) \tag{3.2}$$

Atmospheric muons are the most abundant charged particles arriving at sea level and the only ones able to reach deeply underground. The reason is their small energy loss in the whole atmosphere ($\sim 2\,\mathrm{GeV}$), the relatively long lifetime and the fairly small interaction cross section compared to other secondary cosmic rays.

At the ground the muon flux with $E_{\mu} > 1\,\mathrm{GeV}$ through a horizontal area is roughly equal to one particle per cm^2 per minute. The angular distribution of muons also depends on their energy. While atmospheric muons have on average an energy of

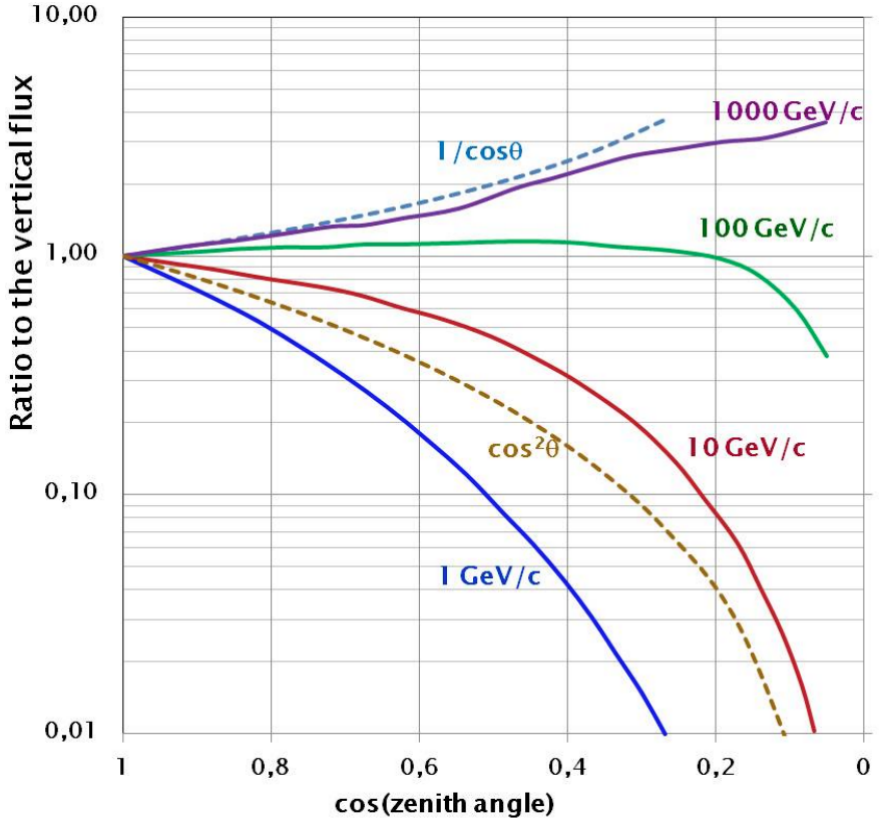


Figure 3.2: Monte Carlo simulation of the angular distribution of muons at the ground for different muon energies. The measured angular distribution at sea level is $\propto \cos^2\theta$. The angular distribution becomes increasingly steep at lower energies and flattens at higher energies. Figure taken from [8].

approximately 4 GeV when reaching sea level, at large angles low energy muons decay before reaching the surface. Additionally high energy pions decay before they

interact, resulting in an increase in average muon energy. The angular distribution for different simulated muon energies is shown in Figure 3.2. At ground level the total distribution is proportional to $\cos^2\theta$. The abundance of muons decreases more rapidly with the zenith angle, the less energy they have. For very high energies it even increases slightly.

3.1 Muon energy loss

While the mean muon energy at the ground may only be about 4 GeV, muon energies measured at the ORCA detector, 2500 m below sea level, will certainly be higher. This is caused by the increase in path length as well as the increased density of sea water compared to air. For most muons below $\sim$ 500 GeV the energy loss is dominated by ionization. Ionization is a continuous process and only depends weakly on the muon energy. For relativistic particles it can be considered constant. Discrete energy losses such as electromagnetic interactions with nuclei, bremsstrahlung and pair production become important for even higher energies. These radiative losses are called discrete because energy is lost in a burst along the muon path. The total muon energy loss can be written as

$$\frac{dE_\mu}{dX} = -\alpha - \beta E_\mu \tag{3.3}$$

with X given by the thickness of the material that is to be traversed, α denoting losses through ionization and β denoting the sum of possible radiative losses. The factors α and β depend also on the medium, specifically on the proton number Z and the mass number A : $\alpha \propto Z/A$ and $\beta \propto Z^2/A$. The energy loss of muons in copper as a function of kinetic energy T is shown in Figure 3.3. For copper radiative losses dominate above $E_{\mu c} = \alpha/\beta = 315$ GeV [10]. For seawater this threshold is

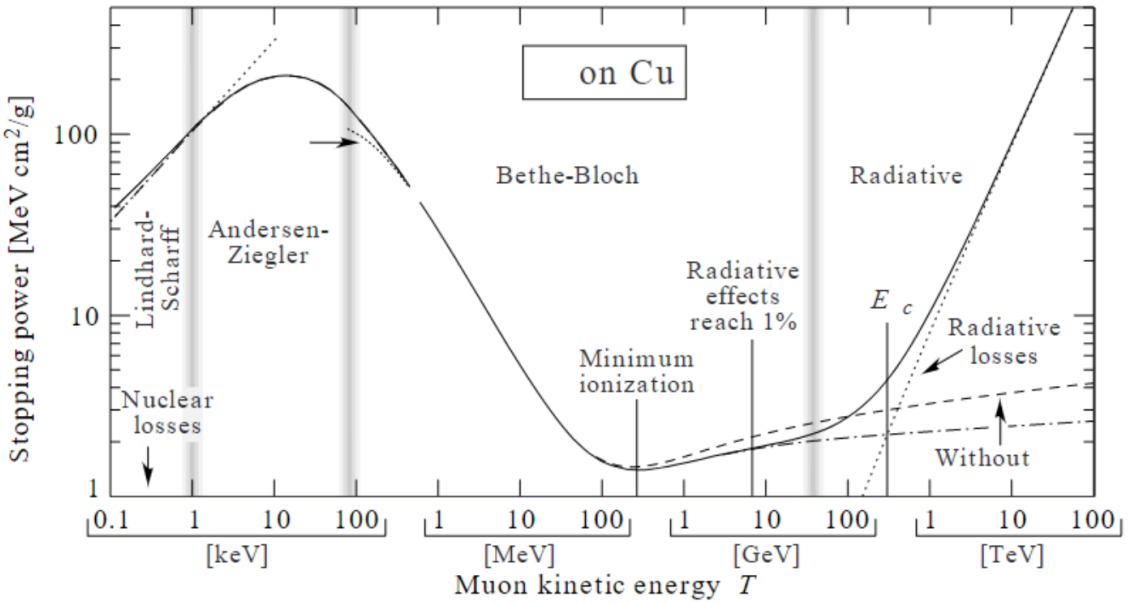


Figure 3.3: Stopping power ($-dE_\mu/dX$) for positive muons in copper as a function of kinetic energy T. The total stopping power is indicated by the solid curve [10]

expected to be higher since the significantly lower proton number Z compared to copper ($Z = 29$). For measurements of muons with the ORCA detector burst-like radiative losses should be avoided as they would interfere with the goal of measuring the attenuation length, see section 3.2. If necessary one could put a restriction on the allowed zenith angle of measured muons to filter out horizontal muons as motivated by Figure 3.2. This would remove mostly higher energy muons as at a shallow incident angle low energy muons decay before reaching the detector due to the increase in path length.

3.2 Interaction in the detector

Under water neutrino detectors use a large volume of seawater as detector medium and the surrounding sea as target material. Neutrinos can be detected indirectly, for example through charged current interactions of muon neutrinos that result in high-energy muons, see Figure 3.5. In the case of very high-energy neutrinos, a large fraction of the muon energy consists of the original neutrino's energy. Due to momentum conservation the muon direction is then almost identical to the direction of the neutrino.

When a muon passes through the ORCA detector characteristic electromagnetic radiation called Cherenkov radiation is emitted by the medium. This happens for all charged particles if they move in a medium with refractive index n with a velocity v that exceeds the velocity of light $c_n = c/n$ in that very medium. The condition reads as follows:

$$\beta = \frac{v}{c} \geq \frac{1}{n} \tag{3.4}$$

Cherenkov radiation is emitted at an angle of

$$\theta_{\mathrm{C}} = \arccos \frac{1}{n\beta} \tag{3.5}$$

relative to direction of the muon momentum. One can calculate the number of photons a particle of charge z would produce in the visible spectral range ($\lambda_1 = 400\,\mathrm{nm}$ to $\lambda_1 = 700\,\mathrm{nm}$) due to the Cherenkov effect with the following equation [11]

$$\frac{dN}{dx} = 2\pi\alpha z^2 \frac{\lambda_2 - \lambda_1}{\lambda_1 \lambda_1} \sin^2\theta_{\mathrm{C}} \approx 490 z^2 \sin^2\theta_{\mathrm{C}}\,\mathrm{cm}^{-1} \tag{3.6}$$

where $\alpha = 1/137$ is the fine-structure constant. As the muons passing through the detector are relativistic particles ($\beta \approx 1$), the Cherenkov angle in water is about 42°. Thus in water, around 220 photons per centimeter are produced by a single relativistic muon. These photons are then detected by the PMTs in the detector and used to reconstruct the direction of the muon. A sketch of the Cherenkov effect is shown in Figure 3.4.

Because the Cherenkov light will be subject to attenuation effects within the ORCA detector, one can attempt to model the decrease in detected photons as a function of distance travelled by the Cherenkov photons. As mentioned earlier, very high energy muons that experience radiative energy losses in the detector need to be

avoided. For example secondary particles created through these discrete energy losses can themselves be subject to the Cherenkov effect and thus interfere with the measurement.

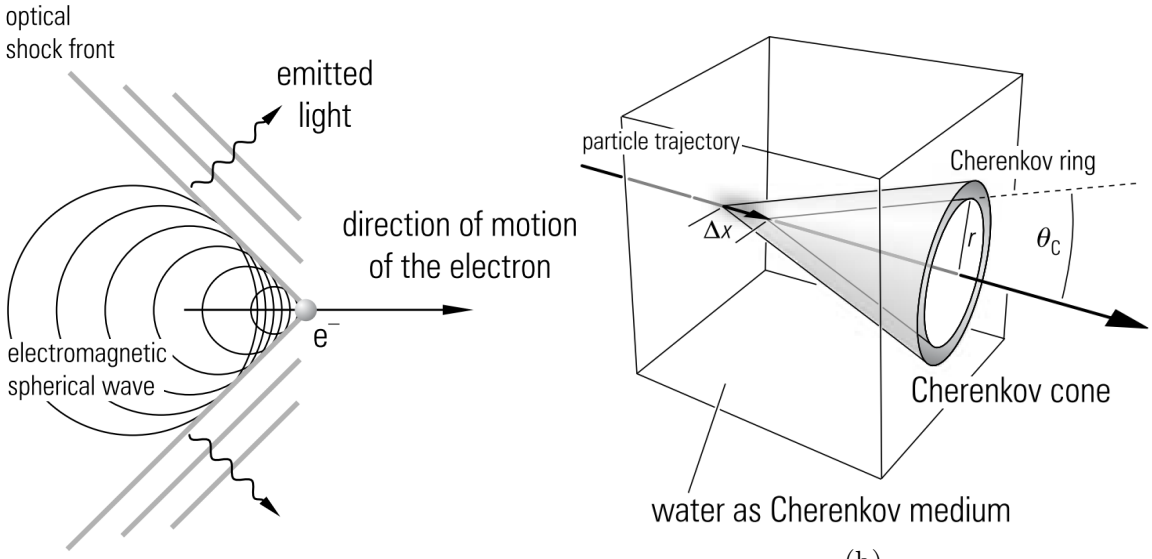


Figure 3.4: **a,** Emission of Cherenkov radiation in an optical shock wave by an electron. The production of Cherenkov radiation is the optical analogue to sound shock waves created by an aircraft when it exceeds the velocity of sound. **b,** Production of a Cherenkov ring under water. Both sketches taken from [11]

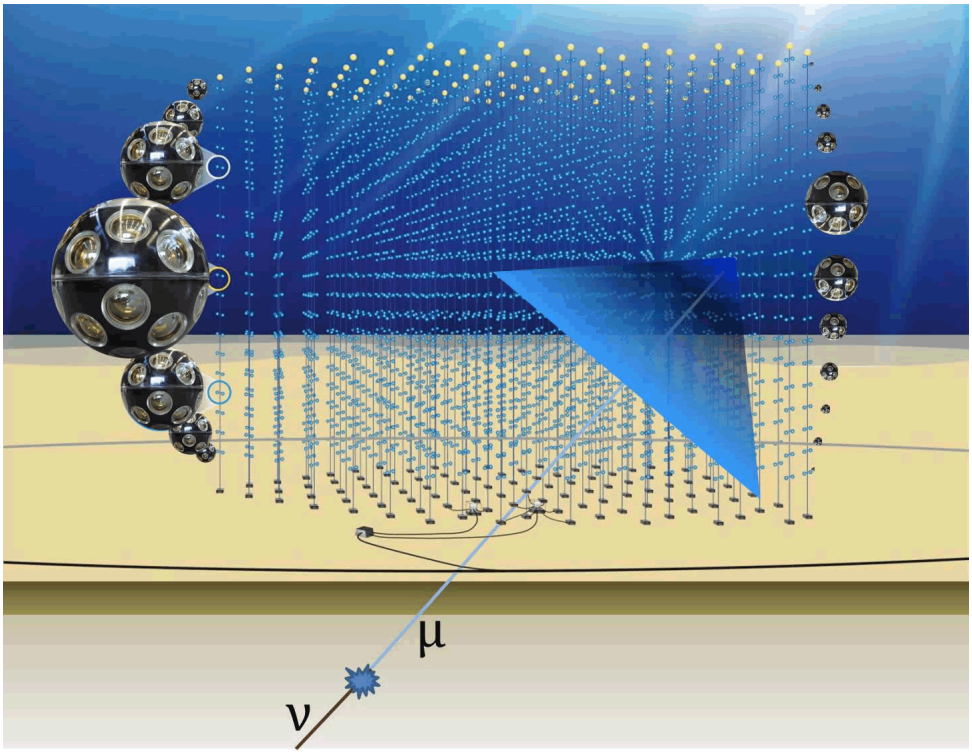


Figure 3.5: An artists impression of a neutrino interaction resulting in a muon producing Cherenkov light in the detector (courtesy Marc de Boer/Ori Ginale) [12]

Chapter 4

Detector Operation

In the previous section the main process by which high energy muons cause light emission in the detector, Cherenkov radiation, was described. Before the complete process of simulating atmospheric muons passing through the detector is outlined, it is briefly explained how data acquisition works in practice. This will be summarized in the following for the ORCA detector. Furthermore the purpose of this chapter is also to introduce the so-called detector files, which are heavily used during MC simulations.

4.1 From PMT signals to events

The KM3NeT detectors operate under the "All data to shore" principle. Every time one or multiple photons are detected by a PMT (meaning the analogue pulse of the PMT rises above a certain threshold) a hit is registered. This information is collected over a period of 100 ms (called a time slice) and then sent to the server on shore. A hit contains only very basic information like the PMT identifier, the time of the analogue pulse crossing the threshold and the time the pulse stays above the threshold (time over threshold). As the PMTs produce a signal at a rate on the order of kHz, saving only a few bytes is more practical than recording the full waveform for each PMT. Each hit is first calibrated, meaning it is associated with a time offset, as well as the position and orientation of the PMT. A hit with this information is called L0 hit.

As a significant amount of hits is produced by natural radioactivity (^{40}K decays) and bioluminescence, the data filter uses trigger algorithms to parse through all the hits recorded in the detector to find causally related hits (e.g. hits that stem from the Cherenkov light of a muon). The trigger algorithms use the time coincidence of hits and constraints on the PMT angle to check if a set of hits is compatible with e.g. the Cherenkov light of a muon or a shower event. If a set of hits satisfies the trigger conditions a (DAQ) "Event" is recorded. It consists of the triggered hits and a snapshot of all hits in the detector in a certain timeframe around the trigger time. The detector operates with several different active triggers, only one of them needs to be satisfied in order for an Event to be recorded. The Event information can then be used to fit specific models to the data in the reconstruction process. In the case of atmospheric muons the main properties to be reconstructed are the muon direction and its energy.

4.2 Detector calibration

Reconstructed properties like direction and energy are obtained by using calibrated hit information such as the positions of the PMTs, the PMT timings, PMT efficiency and the PMT gain. The accuracy of those measurements can be vastly improved by using a precise calibration of the detector. For this purpose several devices are installed in the detector. Acoustic emitters / receivers can be used for positional calibration and nanobeacons / lasers are used in the calibration of PMT timings. Natural phenomena are also used in calibration, namely natural radioactivity through ^{40}K decays and atmospheric muons. The light emitted from ^{40}K decays and atmospheric muons can be used in PMT efficiency and gain measurements and to calibrate PMT timings. An example how this done in practice can be found in reference [13].

The calibration information can be accessed through detector files which can be identified by their filename extension `.detx`. They are an essential tool in order to carry out high level data analysis, but are also heavily used in the production of new MC simulations. As the detector calibration can not be carried out in real time the detector operates using an unoptimized calibration file, also referred to as trigger detector file. This is not a problem for recording events since the trigger algorithms in use are quite robust. For each run (measurement period of $\sim 6\,\mathrm{h}$) one detector file is created. During each week of detector operation several short runs are taken using the calibration equipment to measure the PMT positions and timings. With this additional information new detector files are produced for each run that contain the improved calibration information. They are also referred to as calibrated detector files. As the MC simulation in general need a very accurate detector configuration, they often make use of calibrated detector files. Trigger detector files are then used to convert MC truth into raw data, to emulate the operation of the real detector.

Chapter 5

Simulating the Detector

Before analyzing any data generated by the detector it is first important to understand how the process of detecting atmospheric muons works - and how it can be simulated. In this chapter the general procedure of simulating atmospheric muons and the process of detecting them is explained. The following sections provide an overview over the different programs used to generate a Monte Carlo (MC) simulation of atmospheric muon events for the ORCA detector. By extension, this also covers some aspects of data taking, as once a detector signal is simulated, it is to be treated in the same way as actual data. As a detailed explanation is beyond the scope of this thesis, it is recommended to read through the references mentioned and to consult the source code for additional information.

5.1 The simulation chain

In general the simulation chain, which consists of the steps to generate a complete MC simulation, can differ depending on what kind of events are to be simulated and whether speed or precision is preferred. In order to simulate an atmospheric muon that passes through the ORCA detector, several different algorithms need to be applied in series.

$$\texttt{MUPAGE} \rightarrow \texttt{JSirene} \rightarrow \texttt{JTriggerEfficiency} \rightarrow \texttt{JORCAMuonReconstruction} \quad (5.1)$$

At first the `MUPAGE` program is used to calculate the kinematics of muons arriving at the surface of an imaginary cylinder surrounding the active volume of the ORCA detector. In the next step `JSirene` is used to simulate the detection probability of photons emitted along the path of muons. This includes signal produced through the emission of Cherenkov light but also signal induced by processes responsible for the energy loss of a muon in water. After that `JTriggerEfficiency` is used to convert the MC truth produced by `JSirene` into raw data i.e. by applying PMT inefficiencies, a background signal, etc. and then calibrating the data. It also applies initial trigger algorithms to the data that assign a cluster of hits to an event (e.g. a muon or a shower inside the detector). Finally the output is passed to the `JORCAMuonReconstruction` script together with a calibrated detector file. This will apply reconstruction algorithms for every event. In this thesis it will be used to reconstruct the properties of muons passing through the detector. Thus the final output are events that contain a cluster of hits and reconstructed properties of a muon that fits best to this cluster of hits.

5.2 Generating atmospheric muon events with MUPAGE

In this section a short description of the MUPAGE program is given. It heavily relies on the reference [14].

To generate Atmospheric events for the ORCA detector one possibility is to use a full Monte Carlo (MC) simulation, that starts with the generation of atmospheric showers induced by cosmic rays. These kind of simulations are able to accurately reproduce the main features of atmospheric muons reaching the detector. However it consumes a large amount of CPU time. The simulations in this thesis were created using MUPAGE, a fast simulation of atmospheric muon bundles (muons produced in the same cosmic ray interaction) for underwater neutrino telescopes. MUPAGE uses a different approach in that atmospheric muon events are generated with the help of parametric formulas [15]. These were derived from a full MC simulation (HEMAS) of primary cosmic ray interactions and shower propagation in the atmosphere. The parametric formulas provide the flux, the multiplicity, the radial distribution and the energy spectrum.

A MUPAGE event is a bundle of muons with multiplicity m_c on the *can*, an imaginary cylinder surrounding the active volume of a generic detector (see Figure 5.1). The purpose of the *can* is that it defines the volume in which light from muons can reach the detector. Muons outside the *can* do not need to be simulated as they do not produce a signal in the detector. Muons in the bundle are assumed to be parallel to the shower axis, and to reach a plane perpendicular to said axis at the same time. At first, the bundle multiplicity, direction and impact point of the shower axis on the *can surface* are generated. After that, for each muon in the bundle (multiplicity), the distance from the shower axis, the energy and the coordinates of the impact point on the *can surface* are calculated. The input parameters for the MUPAGE program refer to the the detector geometry, the medium and to the range of simulation parameters (multiplicity, zenith angle and muon energy). The output the MUPAGE program generates is a `.evt` file. In the native format, it is an ASCII table containing the kinematics of events at the surface of the *can* mentioned earlier. This table can be used as input for a detector-dependent MC simulation. In this thesis, it is used for the ORCA detector with 4 strings. The next step is to use this information to produce the Cherenkov light that is emitted along the path of muons passing through the *can*.

5.3 Light generation with JSirene

As JSirene is a part of the software package JPP, the information presented in the following can be found in the JPP documentation, mainly in references [16] and [17].

Photons emitted along the path of muons and the probability of their detection by the DOMs are simulated with JSirene (included in the JPP package) using multidimensional interpolation tables. Apart from that, JSirene can also be used to simulate the detector response to taus or shower like events. JSirene mainly

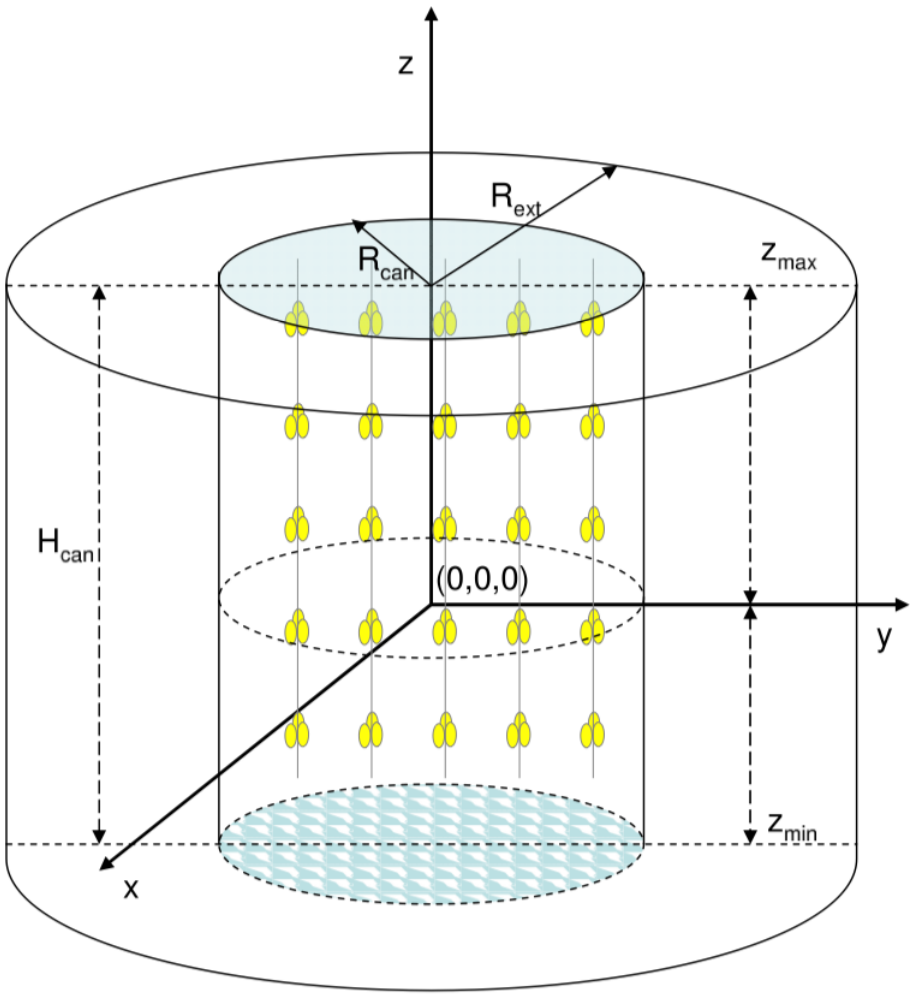


Figure 5.1: Sketch of the cylinder (R_{can}, H_{can}) surrounding the instrumented volume of 5 lines of the ANTARES detector. This cylinder is the (*can*). The events are generated on an extended *can*. [14]

relies on tabulated values of the probability density functions (PDF) of the arrival time of light in combination with interpolation methods instead of simulating single photons. The muons are propagated from a starting position for example a neutrino interaction vertex. In our case this starting position is provided by `MUPAGE`. At first a margin is introduced between the surface of the cylinder (*can*) and the positions of the PMTs. This is the maximal distance used to create the light generation volume. Inside the cylinder the muon is propagated, respecting energy loss. The light produced by the muon includes the emitted Cherenkov radiation and light emitted because of energy loss. Specifically, that includes the light emitted by a minimal ionizing particle (MIP), δ rays and bremsstrahlung induced showers. The term MIP refers to particles whose mean energy loss through matter is close to the minimum, while δ rays are secondary electrons that have enough energy to themselves produce further ionization a significant distance away from the primary radiation. The emission of photons from δ rays is assumed to be isotropic.

Due to these various sources of light, multiple PDFs of the arrival times of photons need to be calculated. That includes the simple case of direct light from a muon,

but also for example indirect light due to the energy loss of a muon. The PDFs incorporate characteristics of the PMT e.g. the photo-cathode area, the quantum efficiency (ratio of detected electrons to incident photons expressed as a percentage) and the angular acceptance. Also they consider the angular distribution of the light scattering, the scattering length and the absorption length.

In order to simulate different absorption lengths for the detector medium, separate PDFs for each absorption lengths need to be produced. For this thesis an increase / decrease of 5% / 10% in absorption lengths was simulated. A snippet of the PDF plotted over the expected arrival time of Cherenkov light is shown in Figure 5.2. As expected the PDF is sharply peaked around the expected arrival time (0 ns). Also an increase / decrease in absorption length leads to an increase / decrease in the number of photo electrons detected. As input `JSirene` requires, in addition

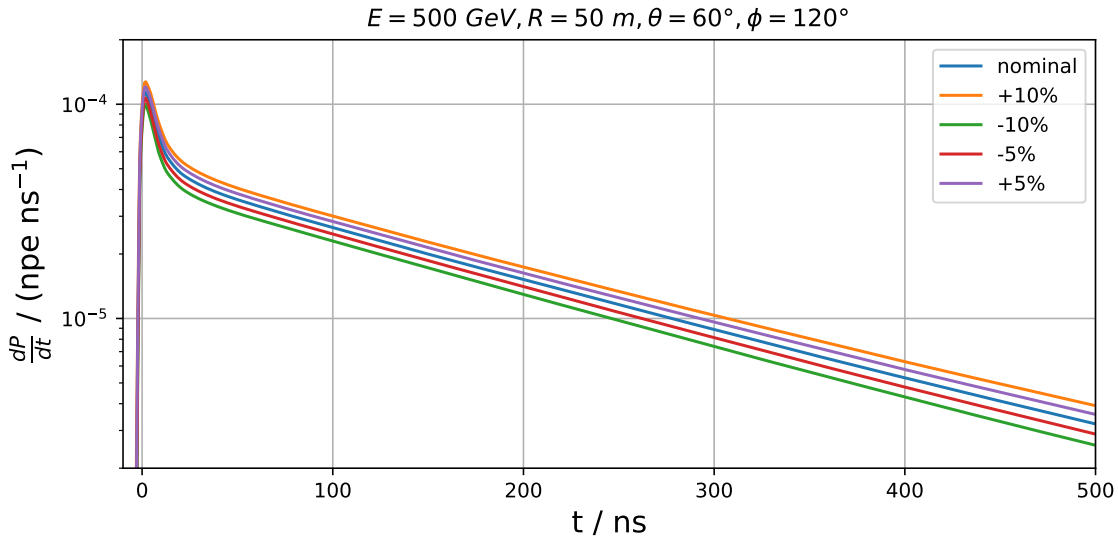


Figure 5.2: Snippet of the PDF simulated for different absorption lengths, with fixed energy and angles. The ordinate measures the expected arrival time of Cherenkov light at the PMT. As expected an increase in absorption length leads to more signal

to the `MUPAGE` output, a calibrated detector file as the position of the PMTs needs to be known as precise as possible. The output will be a `ROOT` [18] formatted `.aanet` file. In summary, `JSirene` is designed for speed and uses PDFs to implement the physics of the detection principle.

5.4 Triggering

In the next step another program from the JPP package named `JTriggerEfficiency` is used. This section only provides information about the purpose and some of the requirements of this algorithm. The source code can be found in the `JPP` documentation.

At this point the results produced with the simulation are to be treated in the same way as the data. `JTriggerEfficiency` is an application that applies the same trig-

ger software to simulated events as `JDataFilter` to real data. It is a place holder for all detector related effects like PMT efficiencies, discriminator thresholds, etc. Also it applies a background signal (mostly caused by ^{40}K decays and bioluminescence in the real detector). Like with real data the trigger software searches for coincidences in a timeslice of 100 ms of data. From that it generates multiple clusters of hits in time that satisfy the coincidence conditions of the trigger algorithms. This happens in real time during the detector operation. In this thesis `JTriggerEfficiency` is operated in run-by-run mode (i.e. it uses summary data taken with real detector) as the full MC simulation will be compared to real data consisting of 20 runs later.

As input `JTriggerEfficiency` needs two detector files. The first one is a trigger detector file that is used in order to convert MC truth to raw data. The second one is a calibrated detector file that helps to convert raw data to calibrated data. The MC truth here is the output of `JSirene`. The output is a `ROOT` formatted DAQ file that is then used to reconstruct events such as muons passing through the detector.

5.5 Reconstruction

This section relies on the brief introduction to the reconstruction process from the `JPP` documentation [19]. If more information is needed, it is recommended to consult this reference or the source code itself.

At the reconstruction step a model is fitted to the data. As atmospheric muons are simulated, the model assumes that a muon passes through the detector and emits Cherenkov light in a cone as it propagates. The model attempts to reconstruct two important properties of the muon. The energy and the direction of the muon. The track length of the muon is correlated with the energy of the muon itself. The reconstruction is done through multiple steps, commonly referred to as the "reconstruction chain". It consists of multiple algorithms applied in sequence

$$\texttt{JPrefit} \rightarrow \texttt{JGandalf} \rightarrow \texttt{JStart} \rightarrow \texttt{JEnergy} \tag{5.2}$$

These are applied to an event file, in this case the output of `JTriggerEfficiency`.

First an initial maximum likelihood fit is performed by `JPrefit`. It starts by assuming a number of track direction hypotheses in all directions in 5° steps, to scan across the whole solid angle of the sky. Thus the only free parameters to be reconstructed are the position of the muon and the time where it crosses a certain reference plane perpendicular to its direction. The parameters are then obtained by minimizing the χ^2 value, which is based on the time residuals (difference between detected and expected arrival time of the photons) and is normalised to an assumed time resolution. Only a smaller cluster of causally related hits is selected from data (in 1D), in order to minimize the linear fit being affected by the optical background. Additionally some outliers are removed. The directions that produce the best fits are then stored and passed to the next reconstruction step.

`JGandalf` uses these fits as a starting point and performs a scan around these fits and then executes the full fit. 3D clustering of causally related hits takes place. The fit uses a more elaborate method called the Levenburg-Marquardt method for the

likelihood scan. It uses a set of PDFs to describe the PMT response similar to the PDFs in `JSirene`. However these PDFs do not account for direct or scattered light due to delta-rays. `JGandalf` also has a quality parameter $Q = -\chi^2$ associated with it. The larger the quality, the better the fit.

`JStart` determines the starting position of the muon trajectory. This is achieved by back-projecting the hits onto the track under the Cherenkov angle. The first emission point associated with the track that exceeds the random background level is selected as the starting position.

`JEnergy` determines the energy. It uses the spatial distribution of hit and non-hit PMTs. All PMTs in a pref-defined width around the muon trajectory are used. The hit probability for a PMT is calculated and then compared to the occurrence of a hit. This procedure involves a 1-parameter likelihood fit of $\log E$. At last `JEvt` can be used to convert the `.root` formatted files produced by `JEnergy` to `.aanet.root` formatted files.

Chapter 6

A Model for the Attenuation Length

The goal of this thesis is to measure the attenuation length of seawater using atmospheric muons with the ORCA detector. Atmospheric muons can be considered a frequent and steady calibration source as they have low seasonal variation in their rate [20]. For neutrino telescopes they are mostly background that needs to be filtered out, or they are used for calibration purposes. In the deep Mediterranean sea the ORCA detector benefits from excellent water properties, close to optically pure sea water with an absorption length $l_{\text{abs}} \approx 50\,\text{m}$ and a scattering length of about $l_{\text{scat}} \approx 50\,\text{m}$ in the 400 nm to 500 nm regime (Figure 9.6). With the ORCA detector it is not possible to measure pure absorption, as scattering is always present in sea water, too. However one can measure the attenuation, which encompasses scattering and absorption effects. The attenuation length l_{att} is related to the absorption length l_{abs} and the scattering length l_{scat} as follows:

$$\frac{1}{l_{\text{att}}} = \frac{1}{l_{\text{abs}}} + \frac{1}{l_{\text{scat}}}. \tag{6.1}$$

The main idea behind measuring the attenuation length is that the hit probability p_{hit} is proportional to the attenuation law.

$$p_{\text{hit}} \sim \exp\left(-d/l_{\text{att}}\right) \tag{6.2}$$

In the detector signal we see the attenuation effects when we look at the number of photons as a function of distance. More precisely, the number of photons detected by PMTs in the detector should decrease with the distance to the muon track because of light attenuation. To describe this effect we start with a simple model. The muon path, along which Cherenkov light is emitted, is treated as a line-like light source. The number of photons as a function of the distance is then approximately given by

$$N_{\text{photons}} \approx \exp\left(-d/l_{\text{att}}\right) \cdot d^{-1} \tag{6.3}$$

where the first term describes light attenuation due to absorption and scattering and the second term is the phase space factor. It describes the density of radiation at a given distance. For a point-like light source the proportionality factor is d^{-2}, the well-known inverse square law. In principle the phase space factor can depend on the detector configuration and the exponent might be fractional. This is neglected here, but should be considered for a more extensive analysis.

To obtain the number of photons at a given distance to the muon track, the hit probability is used. As the photon detection process is a counting experiment, Poisson statistics apply.

$$P_k(\lambda) = \frac{\lambda^k}{k!} \exp(-\lambda) \tag{6.4}$$

In this case k is the number of photons and λ is the expected value. The probability of a hit p_{hit} is equal to the probability of one or more photons hitting a PMT.

$$p_{\text{hit}} = p_{k>0} \tag{6.5}$$

To obtain the expected number of photons one can use the relation

$$P_0(\lambda) = 1 - p_{k>0} = \exp(-\lambda) \tag{6.6}$$

and solve for the expected value λ. This yields

$$N_{\text{photons}} = \lambda = -\ln(1 - p_{\text{hit}}(d)) \tag{6.7}$$

for the expected number of photons. Finally one can calculate the hit probability p_{hit} as a function of distance d by dividing the number of hits by the number of PMTs as a function of distance. This will depend on the detector geometry. After plotting Equation 6.7 as a function of distance, one can fit the result using the model described in Equation 6.3 using the attenuation length as a fit parameter. This fit will be performed for the data and the MC simulation.

For the ORCA detector the PMT charge information is not stored directly, rather the number of photons in a single hit has to be inferred from the time over threshold (section 4.1). One can avoid relying on the hit-charge information by using larger distances. In section 8.3, it will be shown that one hit is roughly equal to one photon at about 60 m of travel distance for Cherenkov photons (Figure 8.6).

Chapter 7

Run and Event Selection

7.1 Run selection

For this thesis 20 runs were analysed, each corresponding to a measurement time of about six hours, in total a livetime of about 120 hours. The runs were selected using the silver selection criteria. That means they are physics runs that have a lifetime of more than an hour and a non-zero trigger rate. Additionally there is a restriction on the time the detector spends idle.

$$|(\text{UTCmax} - \text{UTCmin}) - \text{Livetime}| < 100\,\text{s} \tag{7.1}$$

The 20 runs chosen have a very similar livetime. Thus, the same number of events was generated with `MUPAGE` for each run ($\sim$ 8.6 million). Meaning the MC simulations feature the full livetime to a good approximation. The run list used can be found in Appendix B.

7.2 Event selection

For each event registered by the detector, several muon tracks are being reconstructed using `JGandalf` (see section 5.5). Out of those tracks only the one with the highest quality parameter (likelihood) is selected. This track represents our best guess for the muon track and any significant deviations of the estimated track direction will heavily influence the expected direction and time of arrival of Cherenkov photons. Furthermore, only down-going tracks are considered and some initial quality cuts are applied. They are listed in Table 7.1. These cuts are used to set a baseline, both in terms of the number of hits used in the reconstruction and the overall track quality.

Table 7.1: Quality cuts applied, T_x and T_y are the estimated uncertainties in the x and y direction cosines provided by `JGandalf`

likelihood	$Q = -\chi^2$	≥ 75
$\beta_0 = \sqrt{(T_x^2 + T_y^2)}$	β_0	≤ 0.06
$\beta_1 = \sqrt{(T_x \cdot T_y)}$	β_1	≤ 0.04
hits used in reconstruction	n_{reco}	≥ 35

The parameters are extracted from `JGandalf` (section 5.5). In addition to that only muon tracks are used that pass through the very center of the detector in order to have good coverage of the Cherenkov light with most of the PMTs in the detector. Such tracks should provide ample signal and a good reconstruction of the track, meaning lower uncertainties with respect to the track direction compared to tracks that pass the detection units to the side. As mentioned before, in this thesis the ORCA detector with a detector configuration of 4 detection units is used. For this layout, a cylinder with a diameter of 10 m and a height of 40 m in the center of the detector was defined. This is illustrated in Figure 7.1. Muon tracks that do not

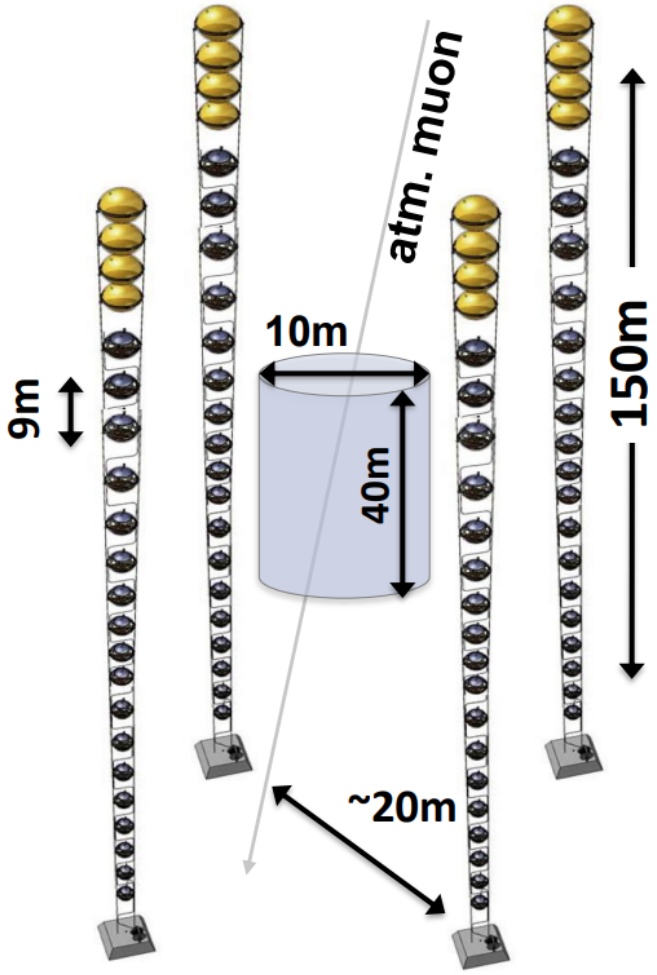


Figure 7.1: Sketch of the detector layout with 4 detection units. Only reconstructed muon tracks that pass through a cylinder with a diameter of 10 m and a height of 40 m in the center of the detector are considered

intersect this cylinder are discarded. This is a severe restriction with only about 5% of the events surviving this cut. However, once the detector grows in size and more detection units are installed this volume can of course be increased.

A special type of event that one needs to consider are multi-muon events. As multiple atmospheric muons can stem from the same initial cosmic ray interaction, they can reach the instrumented volume of the detector almost simultaneously. In the MC simulation this is handled by simulating muon bundles, see section 5.2.

This can interfere with the measurement of the attenuation length, mainly because of two reasons. Firstly the reconstruction only tries to reconstruct a single muon and secondly, light emitted from a second muon is not included in the model that is used to estimate the attenuation length (see Equation 6.3). Thus, multi-muon events should be avoided. Using 4 detection units, the instrumented volume is small enough that most events only feature a single muon and for most multi-muon events a single muon is responsible for the bulk of the hits. From MC simulations for this detector configuration the number of multi-muon events is estimated to be about $\sim$20% based on a single run. Nonetheless some additional cuts are applied in an attempt to filter out these events.

For this it is useful to study the shape of the time residuals of the hits. The time residuals are the time difference between the expected time of a hit based on the reconstructed muon track and the actual detected hit time. Using the quality cuts from Table 7.1 and the cylinder cut, the time residuals of single muon and multi-muon events are compared. This is shown in Figure 7.2 using $\sim$ 15000 events (before cuts) from the MC simulation of a single run. Multi-muon events tend to produce

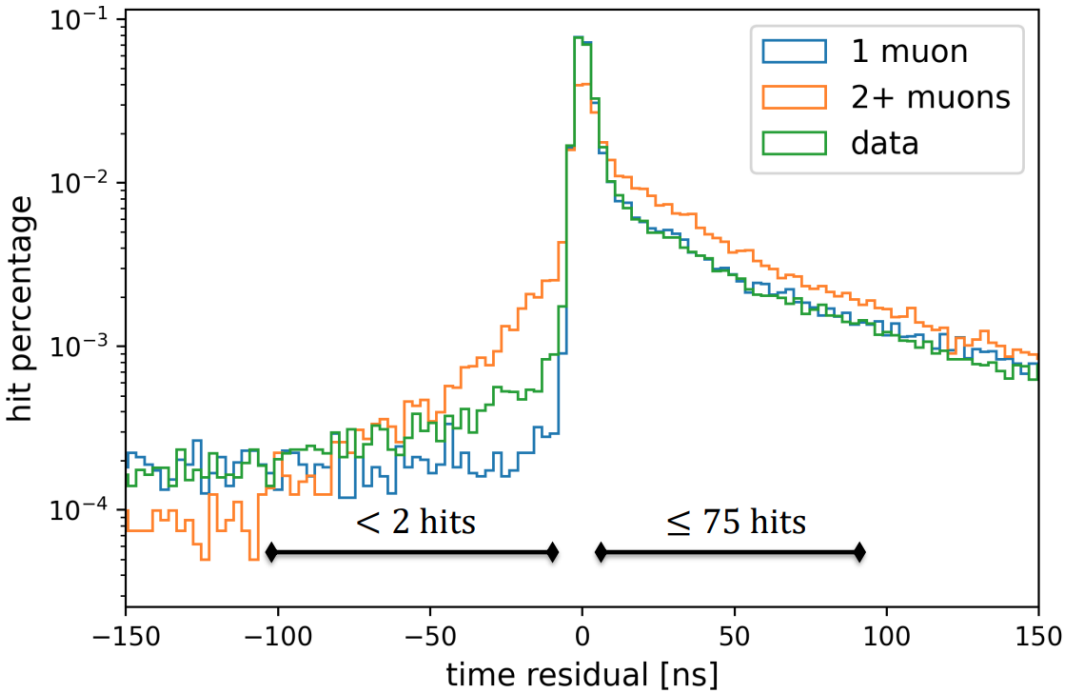


Figure 7.2: Time residuals of single muon events and multi-muon events that feature a second muon that produces at least 5 hits in the detector. The histograms are normalized to 1 for a comparison of shapes

more hits that arrive too early or too late compared to single muon events. The cause for this might be that some hits from the second muon are used for the track reconstruction process, resulting in a muon track that does not represent either muon too well. If the number of hits contributed by a second muon is less than 5, it is not counted as a multi-muon event. This is done as for a small number of hits it is nearly impossible to filter them out, since such events look very similar to ones with slightly more noise (e.g. due to ^{40}K decays or bioluminescence). As such, they

are likely removed by additional cuts that will be introduced in section 8.2 which are aimed at reducing noise hits.

By removing events that have 2 or more hits within -100 ns to -10 ns and 75 or more hits between 10 ns to 100 ns the number of multi-muon events can be reduced. After these cuts are applied about 3% of multi-muon events are left in the MC simulation. The cuts could in principle be improved further as they are only an initial guess, chosen "by eye". For this analysis it yields diminishing returns as a relatively small data set was used, meaning the efficiency of the cut is more important than the purity of the data. A detector configuration with more strings will be more prone to multi-muon events and thus a more elaborate way of filtering out multi-muon events (e.g. with machine learning techniques) will be required.

Chapter 8

First Results and Additional Cuts

8.1 Zenith angle and muon energy distribution

After the cuts motivated in chapter 7 are applied, about 25000 events are remaining. The zenith angle distribution of the muon tracks for both the data and MC simulation are shown in Figure 8.1. The zenith angle is measured between the direction

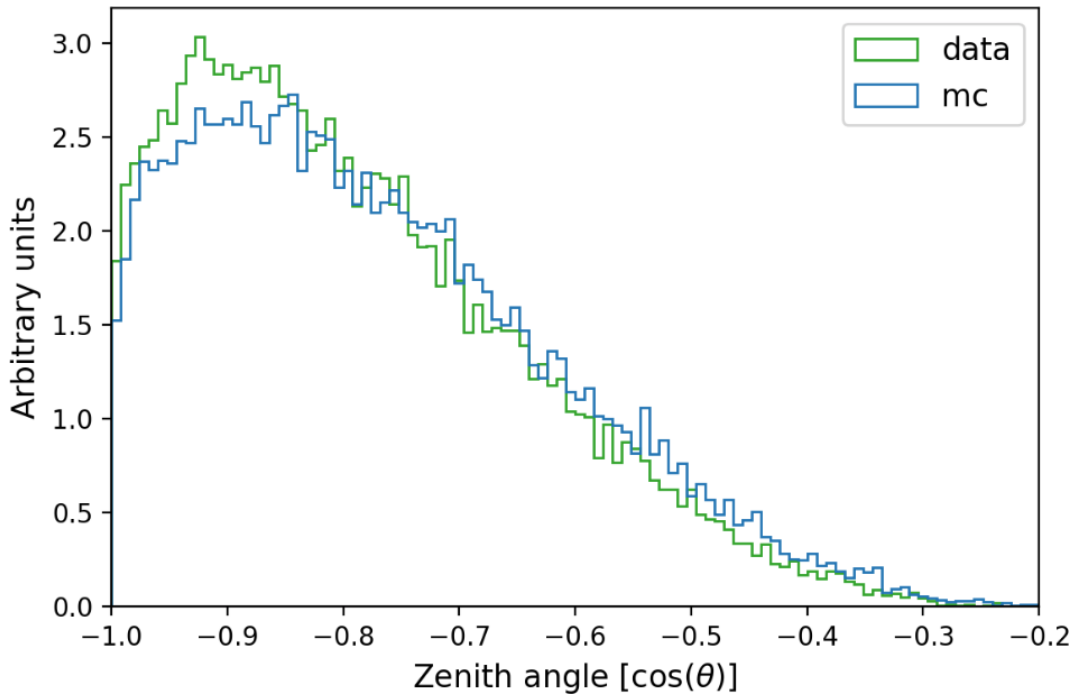


Figure 8.1: Normalized histograms of the zenith angle distribution for data and MC simulation

of the muon track and the vertical (e.g. a muon track from directly above has a zenith angle of 180°). The distribution peaks around 150° to 160° with respect to the vertical. Comparing the MC simulation and the data, slightly more vertical muon tracks are observed in data compared to the simulation. The overall shape of the distribution is mainly influenced by the cylinder cut introduced in section 7.2. Increasing the diameter tends to allow for more vertical tracks, while on the other hand increasing the height of the cylinder would favor more horizontal tracks on average.

The muon energy does also depend on the zenith angle. A 2D histogram of the muon energy and the zenith angle is shown in Figure 8.2. This plot was obtained by extracting the simulated muon energy for the first muon in each bundle and its corresponding zenith angle from the MC simulation for each event. About 75% of

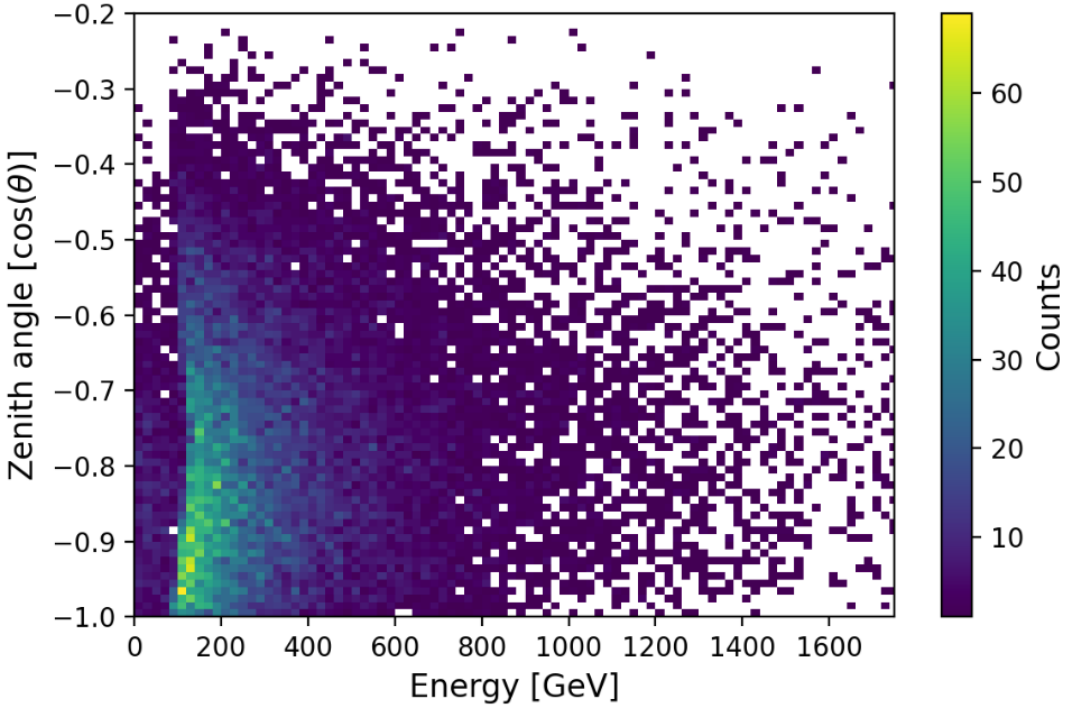


Figure 8.2: Histogram of the muon energy and the zenith angle for the first muon in each bundle. The MC simulation predicts the bulk of muons with energies around 100 GeV to 400 GeV

.

muons have an energy below 500 GeV and lose energy dominantly through ionization. While the percentage of high energy muons above 1 TeV estimated by the MC simulation is about 6% the probability of radiative losses in the detector volume is assumed to be low and an impact on the measurement is not expected.

8.2 Additional cuts

In order to filter out noise hits, additional cuts are applied. The first cut applies to the the time residuals mentioned in section 7.2. Even though events are created if a set of causally connected hits prompts a trigger algorithm, all hits in a certain time frame around the triggered hits are included in the event. This time frame also includes noise hits due to ^{40}K decays or bioluminescence. In order to filter out some of this noise only hits that are "on time" (within ±6 ns) are considered. Furthermore, one can use constraints based on the expected impact angle of the photon. For this the angular acceptance of the PMTs needs to be examined. It is depicted in Figure 8.3 for different PMT configurations. The configuration used in the ORCA detector features a reflector ring and the PMTs can detect a photon with an angle of incidence up to $\sim 90°$. To reduce contribution through scattering and noise, PMT hits are only considered if they face the expected Cherenkov light

front. This can be achieved by comparing the direction of the PMT to the expected direction of a Cherenkov photon based on the reconstructed track of the atmospheric muon. Initially only hits are selected that differ from the ideal impact angle of 180° by at most 90°. This artificially limits the angular acceptance of the PMT.

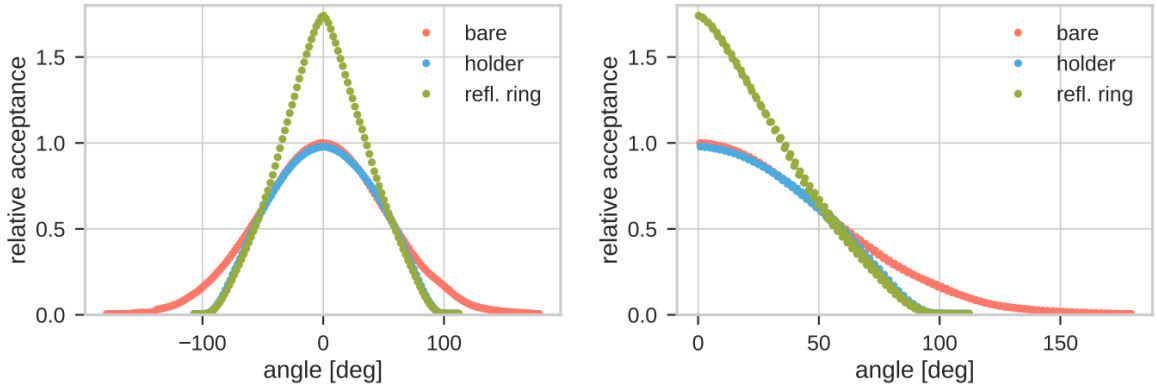


Figure 8.3: Measured relative acceptance as a function of the angle of incidence of the parallel photons (left) and for absolute values of the angle (right). The reflector-equipped PMT used in the ORCA detector is shown in green. Taken from [21].

Additionally not all the PMTs on the DOMs are used in the analysis. Based on the measured zenith angle distribution (see Figure 8.1) almost no signal is expected for PMTs that look towards the sea floor. A picture of a DOM is shown in Figure 8.4. There are 6 rings with PMTs on a DOM, each ring indicated by a letter A to F. The

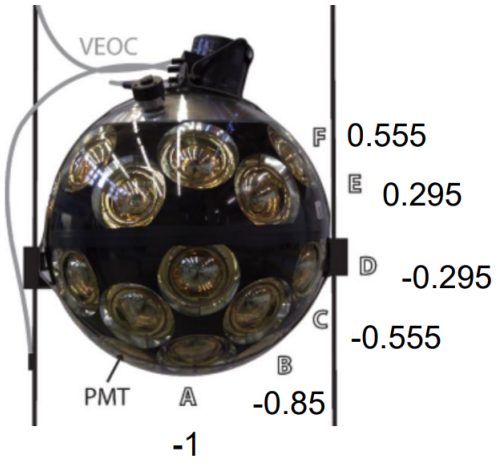
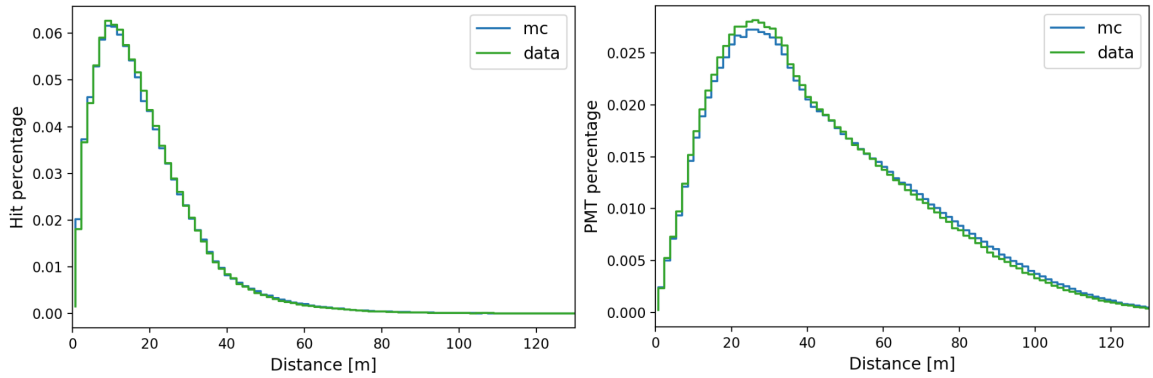


Figure 8.4: Depiction of DOM used in the ORCA detector. It consists of 31 PMTs on 6 rings denoted with letters A to F. The direction of the PMT in each ring is given by the cosine of its angle

direction of the PMT is described the cosine of its angle next to the letters. The PMTs at position A and B can be safely neglected and are ignored in the analysis due to their direction.

Using these additional cuts distance histograms for the hits and PMTs are created. This was done by calculating the distance from the reconstructed track to the posi-

tion of a hit / PMT, assuming the trajectory of a Cherenkov photon, see Figure 8.5. The maximum of the distance distribution is at about 10 m for the hits and about 27 m for the PMTs. Both histograms are normalized to compare the shape of the data and MC distribution. Slight differences are observed, in particular the distance to PMTs are somewhat shorter in data compared to the MC simulation. This is a result of the different zenith angle distributions, see Figure 8.1, as vertical muon tracks lead to shorter distances compared to horizontal ones.



(a) (b)

Figure 8.5: **a,** Distance histogram of the hits in the detector. **b,** Distance histogram of the PMTs in the detector. Both histograms are normalized to have an area equal to 1 to allow for a comparison of shapes. The distances are calculated based on the path of a hypothetical Cherenkov photon from the reconstructed track to the PMT. Cuts on the time residuals (±6 ns) and the angular acceptance $< 90°$.

8.3 Fitting the attenuation length

Using the additional cuts introduced in the previous section, a histogram of the number of photons over the distance can be plotted using Equation 6.7. For this a histogram of the number of hits is divided by a histogram of the number of PMTs (Figure 8.5) to obtain the hit probability. As both distributions are essentially derived via counting, a Poisson error is assumed. The statistical error is then obtained through simple error propagation neglecting correlations. The error on the distance is set to half a bin width.

$$N_{\text{photons}} = -\ln\left(1 - \frac{N_{\text{hits}}}{N_{\text{pmts}}}\right) \tag{8.1}$$

$$\begin{aligned}
\Delta N_{\text{photons}} &= \sqrt{\left(\frac{dN_{\text{photons}}}{dN_{\text{hits}}} \cdot \Delta N_{\text{hits}}\right)^2 + \left(\frac{dN_{\text{photons}}}{dN_{\text{pmts}}} \cdot \Delta N_{\text{pmts}}\right)^2} \\
&= \frac{1}{1 - \frac{N_{\text{hits}}}{N_{\text{pmts}}}} \sqrt{\left(-\frac{1}{N_{\text{pmts}}} \cdot \sqrt{N_{\text{hits}}}\right)^2 + \left(\frac{N_{\text{hits}}}{N_{\text{pmts}}^2} \cdot \sqrt{N_{\text{pmts}}}\right)^2} \\
&= \frac{1}{1 - \frac{N_{\text{hits}}}{N_{\text{pmts}}}} \sqrt{\frac{N_{\text{hits}}}{N_{\text{pmts}}^2} + \frac{N_{\text{hits}}^2}{N_{\text{pmts}}^3}}
\end{aligned} \tag{8.2}$$

The model (Equation 6.3) is fitted using the method of orthogonal distance regression, meaning that the sum of the squared weighted orthogonal distances from the curve to the data points is to be minimized. The weights are determined by the covariances. For this `scipy.odr` is used, a `Python` [22] package, which uses a modified trust-region Levenberg-Marquardt algorithm [23] to estimate the function parameters. The fit was performed at distances between 60 m and 80 m. In this regime almost all hits consist of a single photon. To verify this, one can plot the time over threshold as a function of distance (Figure 8.6). For a single photon the time over threshold is about 25 ns. In Figure 8.7 the number of photons over the

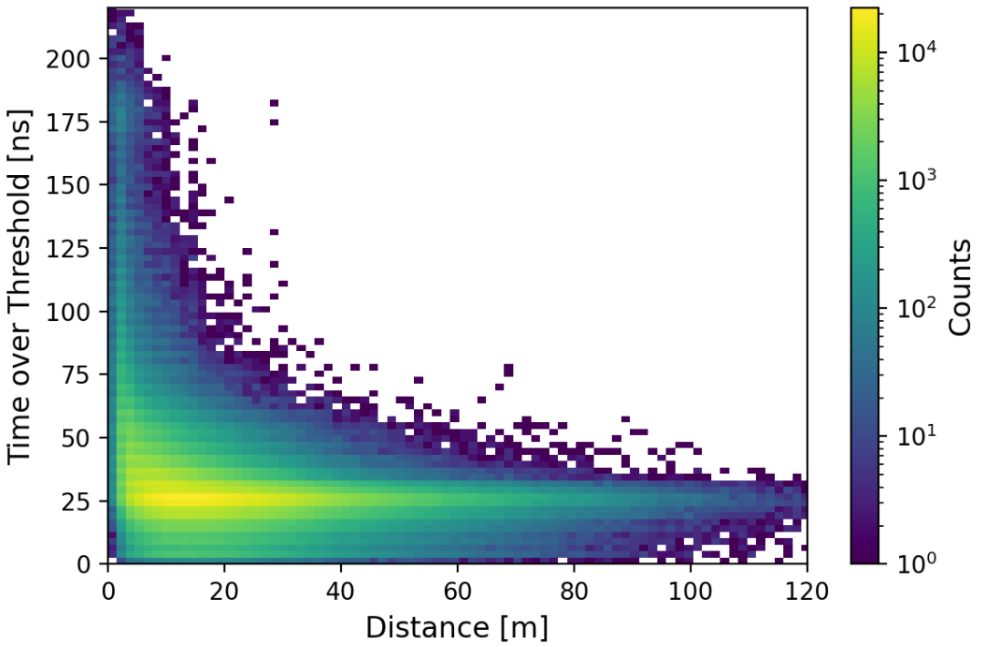


Figure 8.6: Time over threshold distribution of the data with respect to the distance. The time over threshold is the time the analogue pulse of the PMT stays above the threshold. It can function as a proxy for the PMT charge. A single incident photon is known to cause a time over threshold of about 25 ns

distance is plotted and a fit of the attenuation length was performed. The result was obtained using hits with time residuals within ±6 ns and a direction that differs from the ideal impact angle of 180° by less than 90°. It seems that the data is not well described by the model with the chosen cuts. The reasons for this mismatch are unclear. It is possible that it is caused by noise due to ^{40}K decays and bioluminescence or that there is still a significant contribution by scattered photons. In order to decrease these effects the cut on the hit direction can be increased, which in a sense artificially limits the angular acceptance of the PMTs. As the measurement of the attenuation length depends on the slope and not on the overall normalization, one might expected that it would not affect the attenuation length too much. However, fits of the attenuation length using cuts on the angular acceptance of 60° and 50° respectively (shown in Figure 8.8) show a significantly lower attenuation length. On the one hand, the efficiency of the PMTs could vary too much across the whole detector and the PMTs still capture a lot of noise. On the other hand, the cut could also have a non-negligible effect on the exponent of the phase space factor (Equation 6.3). Why exactly this cut affects the attenuation length considerably

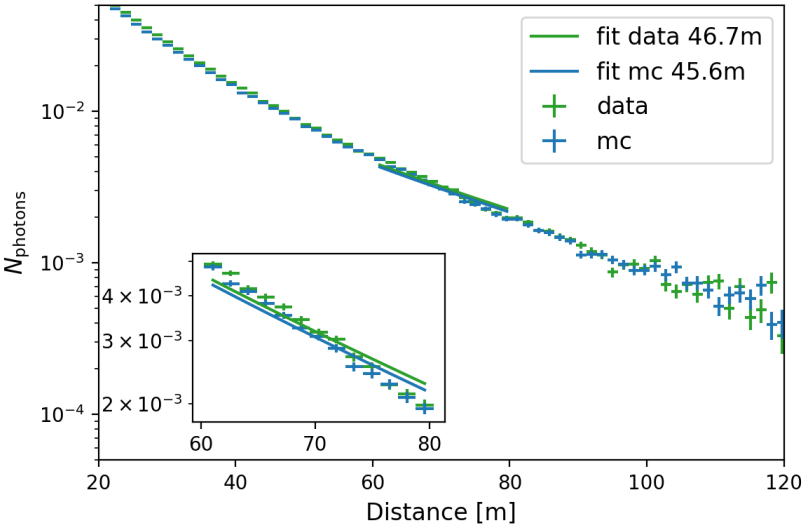
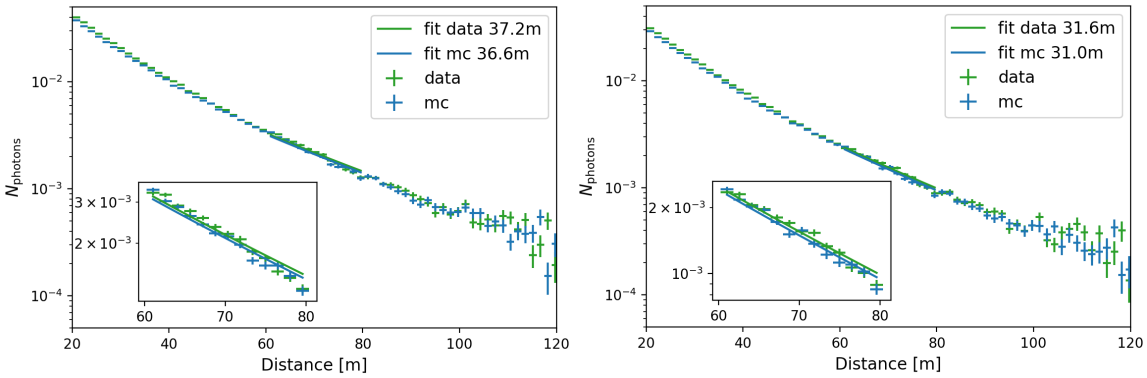


Figure 8.7: Fit of the attenuation length using the model described in Eq. 6.3. Using a cut on angular acceptance of 90°, the data is not well represented by the fit



(a) (b)

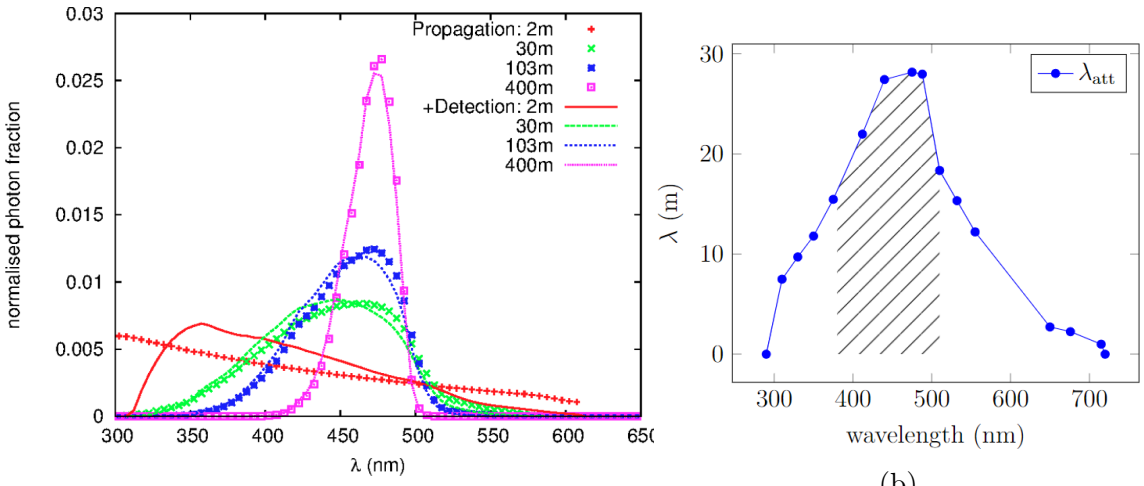
Figure 8.8: Fit of the attenuation length using an angular acceptance cut of **a,** 60° and **b,** 50°. Limiting the angular acceptance decreases the attenuation length and increases the quality of the fit

still remains an open question. For further comparison, the cut of 50° will be used in the following as the quality of the fit of the attenuation length is optimized by this cut. The fit parameter and fit quality ($\chi^2/n.d.f$) for the fits are listed in Table 8.1 Using a cut on the angular acceptance of 50° yields an attenuation length of ~ 31 m

Table 8.1: Fit parameter and fit quality

angular acceptance cut	$\chi^2/n.d.f$	attenuation length [m]
90° data	4.78	46.7 ± 0.71
90° mc	3.53	45.6 ± 0.60
60° data	2.23	37.2 ± 0.36
60° mc	1.76	36.6 ± 0.32
50° data	1.03	31.6 ± 0.21
50° mc	0.97	31.0 ± 0.20

for the MC simulation. To put this into perspective, one can compare this to the attenuation length that was used as an input for the MC simulation. For that is also important to consider the distance at which the attenuation length was measured as the attenuation length depends on the wavelength of the photons. This in turn means that the spectral composition of the Cherenkov photons also depends on the distance. To have a rough estimate of the contributing wavelengths, one can look at the photon fraction after different distances of photon propagation (Figure 8.9). As the attenuation length is fitted in the regime of 60 m to 80 m the majority of contributing wavelength are within 380 nm to 500 nm. In combination with the input of the MC simulation the estimated attenuation length would be about 24 m to 27 m. This could not be recovered with the cuts used in this analysis. Although a stricter cut on the angular acceptance can decrease the attenuation length further, the effects of this cut and to what extent this affects the validity of the model are not yet well understood.



(a) (b)

Figure 8.9: **a,** Normalised photon fraction for different propagation lengths, provided by C. W. James [24]. **b,** Input attenuation length for different wavelengths used in the MC simulation. The shaded area highlights the part of the spectrum that has a major contribution for a measurement between 60 m and 80 m. The data is provided in [25].

Chapter 9

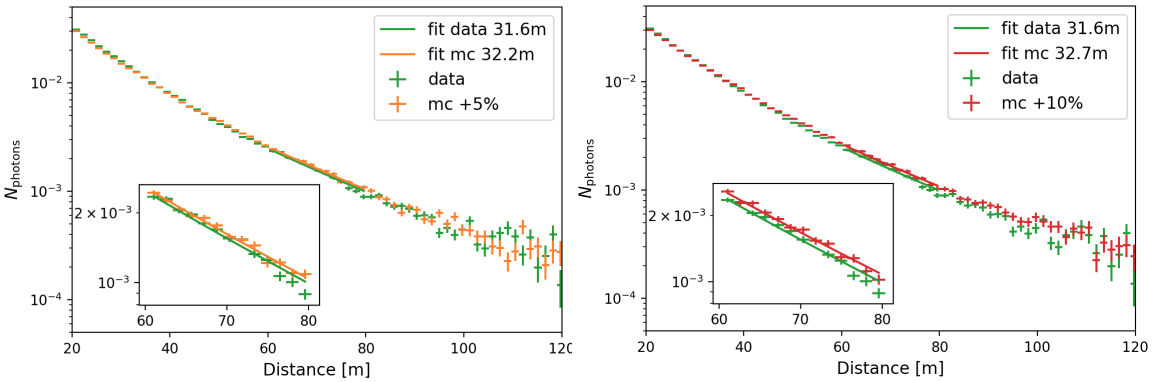
Comparing MC Simulations with Different Attenuation Lengths

A `Nextflow` [26] project was created together with R. Gracia to produce multiple MC simulations with different absorption and scattering lengths. The code is available on the KM3NeT Git server. `Nextflow` is a workflow management system that enables scalable and reproducible scientific workflows using software containers and data-driven computational pipelines. The `Nextflow` project is used to produce multiple MC simulations with different absorption / scattering lengths by applying the algorithms mentioned in chapter 5 in sequence. As an input two factors can be specified to separately increase the absorption length and scattering length with respect to the nominal value. The MC simulations were generated for the 20 runs selected in section 7.1 with the full livetime of 120 hours. More information about the `Nextflow` project along with the run list can be found in Appendix A and B.

Initially MC simulations with an increase / decrease in absorption length and have been produced. As the absorption and scattering length were initially measured to an accuracy of about 10% [2] an increase / decrease of 5% and 10% was chosen. They will be compared to the data in the following chapter. For the scattering length only two MC simulations with an increase / decrease of 5% were produced due to time constraints. The results are shown in section 9.2.

9.1 Varying the absorption length

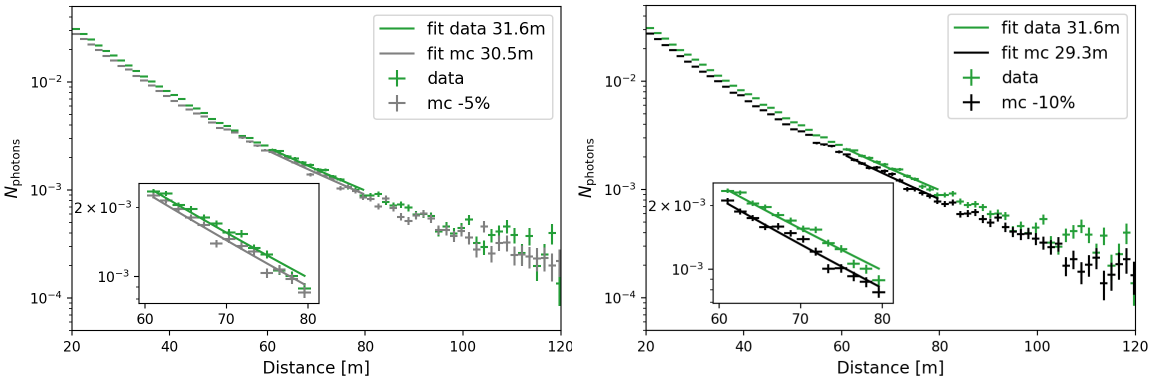
In this section the effect of different absorption lengths on the MC simulation is explored. Additionally the attenuation length is fitted in the same way as in chapter 8 to verify the change in absorption length. In total four additional MC simulations are analyzed that feature changes in absorption length of +5%, +10%, -5% and -10%. The cut on the time residuals is unchanged and a cut on the angular acceptance of 50° was used as this optimizes the quality of the fit (see Table 8.1). The MC simulations with increased absorption are shown in Figure 9.1 and with decreased absorption in Figure 9.2. The fits and error calculation were performed in the same way as described in section 8.3. The associated fit parameters are provided in Table 9.1 at the end of the section.



(a) (b)

Figure 9.1: Comparison between data and a MC simulation with **a,** 5% and **b,** 10% increased absorption length. A slight increase in absorption length of 5% leads to an overestimation of the attenuation length. By eye this MC simulation is still a good representation of the data similar to the original MC simulation (compare to Figure 8.8). Further increasing the absorption length does not yield an improvement in that regard

Comparing the different simulated absorption lengths it seems that the data prefers a slight increase in absorption length. The attenuation length of the nominal MC simulation is about 0.6 m too low according to the fit, while it is the opposite for the MC simulation with a 5% increase in absorption length. Also a decrease in

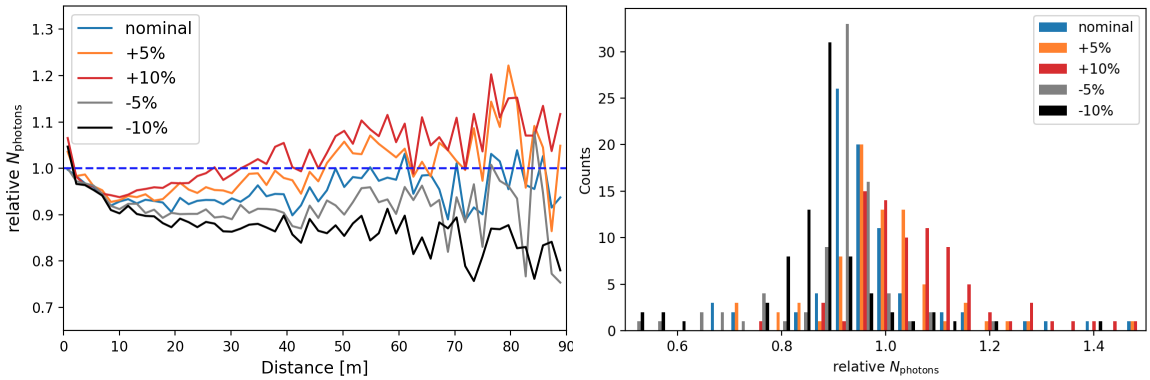


(a) (b)

Figure 9.2: Comparison between data and a MC simulation with **a,** 5% and **b,** 10% decreased absorption length. As the original MC simulation already features a slightly lower attenuation length according to the fit, further decreasing the absorption length does not improve the MC simulation as expected.

absorption length does lead to a decrease in the fitted attenuation length as expected. However, the fit only covers distances between 60 m and 80 m.

The differences of the MC simulations with respect to the data over a bigger distance interval is shown in Figure 9.3. For that purpose, the histograms of the MC simulation were divided by the histogram of the data. This comparison only covers distances up to 90 m because of a lack of statistics beyond that. The MC simulations



(a) (b)

Figure 9.3: **a,** Histogram of MC simulations divided by data for different absorption lengths, the dashed blue line indicates ideal representation **b,** Projection of the histogram onto the y-axis. The more counts close to a ratio of 1.0 the better the representation of the data

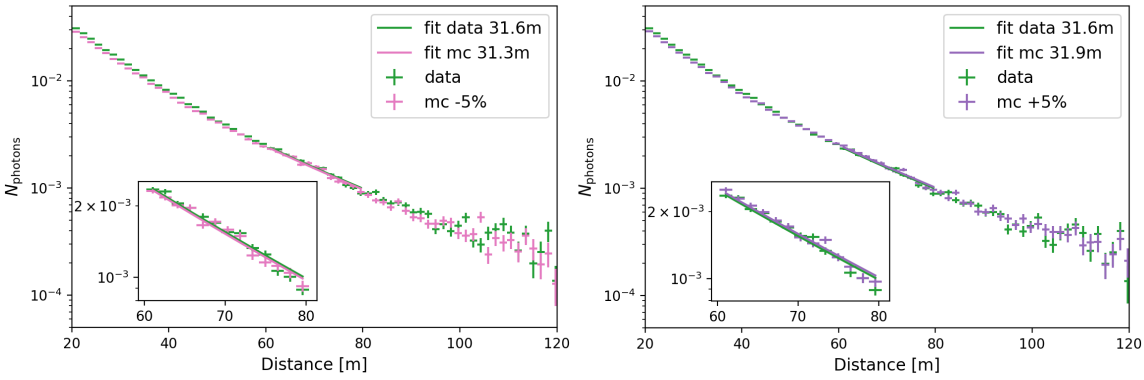
can now be judged based on how far away the ratio is from the ideal representation, represented by the dashed blue line. As expected the increase / decrease in absorption leads to an increase / decrease in the number of photons with respect to the nominal MC simulation and the effect increases with distance. Up to about 45 m an increase in absorption length improves the MC simulation. In the regime of 45 m to 75 m there is no clear favorite between the nominal MC simulation and the one with a 5% increase in absorption. For even larger distances the statistics are not good enough to draw a meaningful conclusion. In summary, the preliminary conclusion is that a small increase in absorption length can benefit the comparison of data to the MC simulation.

Table 9.1: Fit parameter and fit quality for different simulated absorption lengths

	$\chi^2/n.d.f$	attenuation length [m]
data	1.03	31.6 ± 0.21
mc nominal	0.97	31.0 ± 0.20
mc +5%	0.35	32.2 ± 0.13
mc +10%	0.47	32.7 ± 0.15
mc -5%	0.88	30.5 ± 0.19
mc -10%	0.80	29.3 ± 0.17

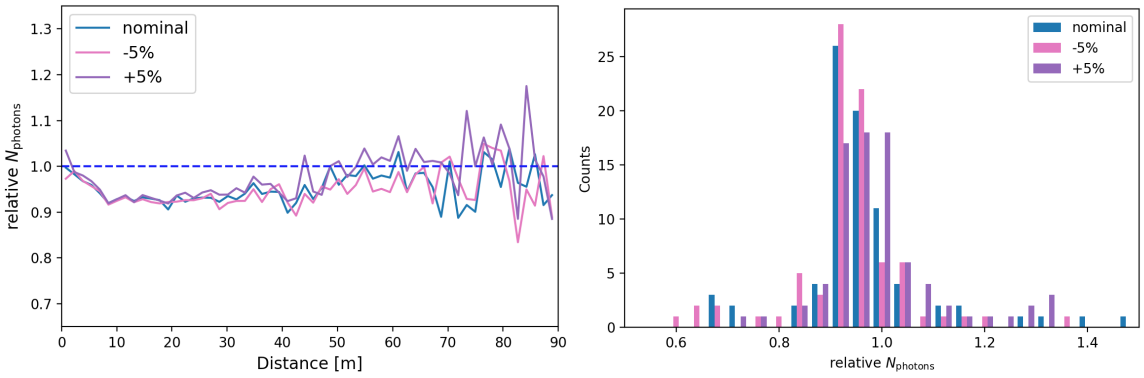
9.2 Varying the scattering length

Instead of changing the attenuation length through a change in absorption length, one can also change the scattering length instead. For this only an increase / decrease of 5% of the scattering length was simulated. The same cuts on the time residuals and angular acceptance were maintained to allow for comparisons to the previous section. The MC simulations are again compared to the data in Figure 9.4 and Figure 9.5. Fits and associated fit parameters are provided at the end of the section (Table 9.2).



(a) (b)

Figure 9.4: Comparison between data and a MC simulation with a 5% decrease (**a,**) and a 5% increase (**b,**). This increase has a smaller effect on the attenuation length compared to a change in absorption length. Furthermore compared to Figure 8.8, the fit yields an increase in attenuation length, at least in the 60 m to 80 m regime



(a) (b)

Figure 9.5: **a,** Histogram of MC simulations divided by data for different attenuation lengths, the dashed blue line indicates ideal representation **b,** Projection of the histogram onto the y-axis. The more counts close to a ratio of 1.0 the better the representation. The impact of the change in scattering length is mainly visible from 20 m to 60 m. For larger distances it is overshadowed by the statistical uncertainties

For the same change percentage-wise, a change of the absorption length does have a slightly larger impact on the attenuation length than a change in scattering length

(see Figure 9.3). On the one hand the scattering length is slightly lower in the 400 nm to 500 nm regime (see Figure 9.6). On the other hand some scattered photons may still partly contribute to the measured attenuation length in the plots should they survive the cuts on the time residuals and the angular acceptance. It should also

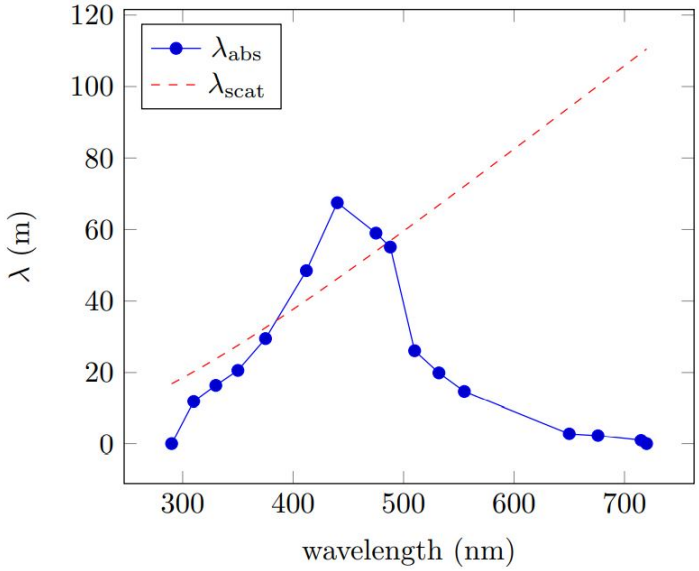


Figure 9.6: Absorption length and scattering length as a function of the wavelength of the photon used in the MC simulation. Plot taken from [25]

be noted that the measurement of the scattering length in [2] was performed by independently measuring the absorption length and attenuation length and inferring the scattering from both measurements. Thus it is subject to a larger error. In sum this complicates studies of the scattering length. Given the available data, one can not draw a conclusion on whether an increase or decrease in scattering length would be beneficial to the comparison of data to the MC simulation.

Table 9.2: Fit parameter and fit quality for different simulated scattering lengths

	$\chi^2/n.d.f$	attenuation length [m]
data	1.03	31.6 ± 0.21
mc nominal	0.97	31.0 ± 0.20
mc -5%	0.77	31.3 ± 0.18
mc +5%	0.75	31.9 ± 0.18

Chapter 10

Summary and Conclusions

In this thesis, the KM3NeT-ORCA neutrino detector was used to attempt to measure the attenuation length of seawater using atmospheric muon events. For this purpose, a detector layout of 4 detection units and a dataset with a livetime of approximately 120 hours was used. Only atmospheric muons passing through the center of the detector were used with the help of a cylinder cut. Furthermore, the time residuals were used to limit the number of multi-muon events. Using hits that arrived on time (± 6 ns) and a cut on the angular acceptance ($< 50°$) the hit probability was calculated. Using a simple model, the attenuation length was then fitted to the data yielding a result of $31.6\,\mathrm{m} \pm 0.21\,\mathrm{m}$. This does not include systematic uncertainties. Additionally new MC simulations were produced with changes in absorption length of +5%, +10%, -5%, -10% and changes in scattering length of +5% and -5%. For these MC simulation the attenuation length was fitted to verify the change in absorption/scattering and they were subsequently compared to the data.

Tentatively the results indicate that a small increase in absorption length can improve the comparison of data to the MC simulation. While the fit of the attenuation length can not recover the expected attenuation length, which should be in the realm of 24 m to 27 m, the fit can be used to verify the change in absorption length introduced by the different MC simulations. As an open question remains how exactly the cut on the angular acceptance influences the measurement of the attenuation length and if a different exponent in the phase space factor (e.g. a fractional one) would be more appropriate.

At the time of writing the KM3NeT-ORCA detector consists of 10 detection units, and more strings are planned to be added over time. This kind of analysis can benefit from a larger volume in two ways. Firstly, the cylindrical volume cut to use events passing through the center of the detector can be increased further, which will increase the amount of data significantly as this cut has the biggest impact on statistics. Secondly, more detection units enables fitting at larger distances. This can be useful to limit the spectrum of wavelengths contributing to the measurement of the attenuation length. On the other hand, a larger instrumented volume does increase the importance of filtering out multi-muon events. For this purpose using new machine learning techniques that allow for reconstructing new observables like the muon multiplicity might be useful [28].

For a future analysis with more detection units it will be crucial to understand how a

different phase space factor would influence the measurement and if the phase space factor depends strongly on the detector layout. Furthermore, a more elaborate way of choosing a cut for the angular acceptance would be helpful. The `Nextflow` project used in this thesis can also be used to produce MC simulations for a different detector layout in the future with only small modifications. Additionally it supports simulating a change in scattering length in addition to a change in absorption length. Eventually in a few years measurements of the attenuation length might come in useful as part of weekly calibration, since only 120 hours of livetime were used in this analysis, and a larger detector will provide enough statistics.

Appendix A

The `Nextflow` Project

The `Nextflow` workflow is separated into two main workflows. The first one is used to create the PDFs and CDFs that are needed for `JSirene`. A code snippet is shown below. In this workflow, the factors for the scattering and absorption length have to be provided. The factor scales the nominal values of the absorption length and scattering length for the MC simulation. The MakePDF process demands most of the computation time ~ 1.5 days to create the PDFs for one pair of absorption and scattering factor (see Figure A.1). The workflow executes multiple processes in parallel for each step according to the number of combinations to be simulated.

```
workflow {

  absorption = channel.of(0.9, 0.95, 1.0, 1.05, 1.1);
  scattering = channel.of(1.0);

  absorption.combine(scattering).set{ tmp };
  tmp.multiMap { it ->
              abs: it[0]
              scat: it[1]
              }.set { combinations };

  MakePdf(combinations.scat, combinations.abs);

  SaveMakePdf(MakePdf.out...);

  MergePdf(MakePdf.out...);

  SaveMergePdf(MergePdf.out...);

  MakeCdf(MakePdf.out...);

  SaveMakeCdf(MakeCdf.out...);
}
```

For the next step a JSON file containing the runs to be processed along with trigger and calibrated detector files is created. This is done to save the information which runs have already been processed. Additionally, the total number of events to be generated is split into multiple smaller jobs (in this case 20 per run) to save time through parallelization, as the jobs are executed on the computing cluster. The muon events provided by MUPAGE are then passed to JSIRENE, which simulates the light propagation for every set of PDFs for the different absorption and scattering length factors produced with the previous workflow. The output is then propagated further through the simulation chain to JTriggerEfficiency and eventually JORCAMuonReconstruction. Finally the result is stored and the JSON file is updated once a run was processed successfully. A simplified snippet of the workflow is shown below. Using the RRZE Woody cluster, running 100 processes in parallel this workflow takes about 9 hours to complete for two MC simulations (Figure A.2). This can be further improved by accessing more nodes in parallel, as the individual processes require only up to 15 minutes of computation time. The workflows also stores the intermediate files for each process and thus parts of the workflow can be easily rerun should a few jobs fail during the execution of the workflow.

```
workflow {
  EXTRACT_MUPAGE_JOBS_FROM_JSON(...);
  DEFINE_MUPAGE_JOBS(inputs.n_jobs, inputs.n_events, ...);
  runMupage(jobs);

  light_pdfs.combine(runMupage.out.mupage_evt).set{ tmp }

  tmp.multiMap { it ->
                 pdf_path: it[0]
                 mupage_out: it[1]
               }.set { combinations };

  runJSirene(...
             );

  runJTriggerEfficiencyRbr(...);

  JOrcaMuonReconstruction(...);

  StoreOutput(...)

  updateConfigurationStatus(...);
}
```

The process scripts are provided in the modules sub-folder and the workflow inputs are specified in the parameters sub-folder. The profiles.config file contains specifies whether processes are executed locally or on the computing cluster and defines the computing resources needed. It also provides a path to a singularity container which is used for the necessary software packages.

The source code is available on the KM3NeT Git server.

```
N E X T F L O W  ~  version 21.10.6
Launching `make_pdf.nf` [mad_majorana] - revision: fce7382700
WARN: Nextflow version 21.10.6 does not match workflow required version: 21.04.3 -- Execution will continue, but things may break!
executor >  pbs (12)
[c3/74c8eb] process > MakePdf (2)      [100%] 2 of 2 ✔
[32/de3af0] process > SaveMakePdf (2)  [100%] 2 of 2 ✔
[0e/8c8507] process > MergePdf (2)     [100%] 2 of 2 ✔
[f5/50f905] process > SaveMergePdf (2) [100%] 2 of 2 ✔
[2f/8d8c96] process > MakeCdf (2)      [100%] 2 of 2 ✔
[7c/73b9d3] process > SaveMakeCdf (2)  [100%] 2 of 2 ✔
Pipeline complete
Command line: nextflow run make_pdf.nf -c profiles.config -profile woody_batch -with-timeline -with-report -w work_pdf
Completed at: 24-Apr-2022 22:49:38
Duration    : 1d 11h 29m 26s
CPU hours   : 70.9
Succeeded   : 12
```

Figure A.1: Command line output of the `Nextflow` workflow to create the PDFs for two different scattering lengths

```
N E X T F L O W  ~  version 21.10.6
Launching `run_mupage_chain.nf` [condescending_brown] - revision: f966b5fbea
WARN: Nextflow version 21.10.6 does not match workflow required version: 21.04.3 -- Execution will continue, but things may break!
executor >  local (822), pbs (2800)
executor >  local (822), pbs (2800)
[77/02287e] process > EXTRACT_MUPAGE_JOBS_FROM_JSON  [100%] 1 of 1 ✔
[f1/29fff7] process > DEFINE_MUPAGE_JOBS (19)        [100%] 20 of 20 ✔
[af/eb31b6] process > runMupage (400)                [100%] 400 of 400 ✔
[6a/50ef93] process > runJSirene (799)               [100%] 800 of 800 ✔
[6b/9eb10e] process > runJTriggerEfficiencyRbr (800) [100%] 800 of 800 ✔
[30/08373a] process > JOrcaMuonReconstruction (799)  [100%] 800 of 800 ✔
[af/c1788b] process > StoreOutput (800)              [100%] 800 of 800 ✔
[e9/0d5b82] process > updateConfigurationStatus      [100%] 1 of 1 ✔
Pipeline complete
Command line: nextflow run run_mupage_chain.nf -c profiles.config -profile woody_batch -with-timeline -with-report -w /home/saturn/capn/mppi110h/work
Completed at: 25-Apr-2022 19:39:16
Duration    : 8h 31m 59s
CPU hours   : 285.3
Succeeded   : 3'622
```

Figure A.2: Command line output `Nextflow` workflow to create two MC simulations with different scattering lengths for 20 runs (full livetime)

Appendix B

Run list

The runs used for the analysis in this thesis are shown in Table B.1. The were accessed using `xrootd` from the CC-IN2P3 computing centre in Lyon via the folder `/group/km3net/data/KM3NeT_00000044/v6.0/reco`. In total, they amount to a livetime of 117.73 hours. The MC simulation were created for the same runs with a total livetime of 120.55 hours. They are based on the v6.1 production that is available via the folder `/group/km3net/mc/atm_muon/KM3NeT_00000044/v6.1` in Lyon.

Table B.1: Run list

5521 5522 5523 5527 5529 5530 5531 5614 5628 5629
5668 5671 5672 5673 5685 5686 5687 5691 5698 5717

Bibliography

[1] S. Adrian-Martinez et al. "Letter of intent for KM3NeT 2.0". In: *J. Phys. G* 43.8 (2016), p. 084001. DOI: 10.1088/0954-3899/43/8/084001. arXiv: 1601.07459 [astro-ph.IM].

[2] G. Riccobene and A. Capone. "Deep seawater inherent optical properties in the Southern Ionian Sea". In: *Astropart. Phys.* 27 (2007), pp. 1–9. DOI: 10.1016/j.astropartphys.2006.08.006. arXiv: astro-ph/0603701.

[3] Raymond C. Smith and Karen S. Baker. "Optical properties of the clearest natural waters (200–800 nm)". In: *Appl. Opt.* 20.2 (Jan. 1981), pp. 177–184. DOI: 10.1364/AO.20.000177. URL: http://opg.optica.org/ao/abstract.cfm?URI=ao-20-2-177.

[4] Veronique Van Elewyck, P Kellerb, and M Lindsey Clarka. "The Calibration Units of the KM3NeT neutrino telescope". In: *Proceedings, 34th International Cosmic Ray Conference, PoS (ICRC2015)*. Vol. 1160. 2015.

[5] P. Bagley et al. (KM3NeT Collaboration). *Conceptual Design Report*. 2006. URL: https://www.km3net.org/wp-content/uploads/2015/07/CDR-KM3NeT.pdf (visited on 04/01/2022).

[6] KM3NeT web pages. 2006. URL: https://www.km3net.org/wp-content/uploads/2016/02/KM3NeT-Telescope-1.jpg (visited on 04/01/2022).

[7] KM3NeT web pages. 2015. URL: https://www.km3net.org/wp-content/uploads/2015/05/KM3NeT-ARCA-and-ORCA.png (visited on 04/01/2022).

[8] S. Cecchini and M. Spurio. "Atmospheric muons: experimental aspects". In: *Geoscientific Instrumentation, Methods and Data Systems* 1.2 (Nov. 2012), pp. 185–196. DOI: 10.5194/gi-1-185-2012. arXiv: 1208.1171 [astro-ph.EP].

[9] Antoine Letessier-Selvon and Todor Stanev. "Ultrahigh energy cosmic rays". In: *Reviews of Modern Physics* 83.3 (Sept. 2011), pp. 907–942. DOI: 10.1103/revmodphys.83.907. URL: https://doi.org/10.1103%2Frevmodphys.83.907.

[10] Donald E. Groom, Nikolai V. Mokhov, and Sergei I. Striganov. "Muon stopping power and range tables 10 MeV–100 TeV". In: *Atomic Data and Nuclear Data Tables* 78.2 (2001), pp. 183–356. ISSN: 0092-640X. DOI: https://doi.org/10.1006/adnd.2001.0861. URL: https://www.sciencedirect.com/science/article/pii/S0092640X01908617.

[11] Claus Grupen. *Astroparticle Physics*. 1st ed. Springer, Berlin, Heidelberg, 2005. ISBN: 978-3-540-27670-8. DOI: https://doi.org/10.1007/3-540-27670-X.

[12] KM3NeT web pages. 2006. URL: https://www.km3net.org/wp-content/uploads/2016/01/KM3NeT-NeutrinoToMuon-e1453717877680.png (visited on 04/01/2022).

[13] KW Melis. "In-situ calibration of KM3NeT". In: *PoS (ICRC2017)* 1059 (2017), p. 123.

[14] G. Carminati, A. Margiotta, and M. Spurio. "Atmospheric MUons from PArametric formulas: A Fast GEnerator for neutrino telescopes (MUPAGE)". In: *Comput. Phys. Commun.* 179 (2008), pp. 915–923. DOI: 10.1016/j.cpc.2008.07.014. arXiv: 0802.0562 [physics.ins-det].

[15] Y. Becherini et al. "A parameterisation of single and multiple muons in the deep water or ice". In: *Astroparticle Physics* 25.1 (2006), pp. 1–13. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2005.10.005. URL: https://www.sciencedirect.com/science/article/pii/S092765050500157X.

[16] M. de Jong. *JSirene: A program to simulate the detector response*. Jpp documentation. 2021. URL: https://sftp.km3net.de/documentation/Jpp/v15.0.0-rc.2/JSirene.PDF (visited on 04/05/2022).

[17] M. de Jong. *The probability density function of the arrival time of light*. Jpp documentation. 2021. URL: https://sftp.km3net.de/documentation/Jpp/v15.0.0-rc.2/JPDF.PDF (visited on 04/05/2022).

[18] Fons Rademakers et al. *root-project/root: v6.20/06*. Version v6-20-06. June 2020. DOI: 10.5281/zenodo.3895852. URL: https://doi.org/10.5281/zenodo.3895852.

[19] Brìan Ò Fearraigh. *Track Reconstruction in KM3NeT*. Jpp documentation. 2021. URL: https://sftp.km3net.de/documentation/Jpp/v15.0.0-rc.2/TrackReconstruction.PDF (visited on 04/05/2022).

[20] Adam Bouchta. "Seasonal variation of the muon flux seen by AMANDA". In: *26th International Cosmic Ray Conference (ICRC26), Volume 2*. Vol. 2. 1999, p. 108.

[21] Jonas Reubelt. "Hardware studies, in-situ prototype calibration and data analysis of the novel multi-PMT digital optical module for the KM3NeT neutrino telescope". PhD thesis. Erlangen - Nuremberg U., 2019.

[22] Python Software Foundation. *Python*. Version 3.7.5. Feb. 19, 2010. URL: https://www.python.org.

[23] P. T. Boggs and J. E. Rogers. "Orthogonal Distance Regression," in "Statistical analysis of measurement error models and applications: proceedings of the AMS-IMS-SIAM joint summer research conference held June 10-16, 1989," in: *Contemporary Mathematics* 112 (1990), p. 186.

[24] C. W. James. Personal communication. Apr. 21, 2022.

[25] Simulations and HE Astrophysics Working Groups. *Detector simulations for KM3NeT*. p. 20. October 21, 2021. URL: https://simulation.pages.km3net.de/input_tables/Simulations_Description.pdf (visited on 04/25/2022).

[26] Paolo Di Tommaso et al. "Nextflow enables reproducible computational workflows". In: *Nature biotechnology* 35.4 (2017), pp. 316–319.

[27] R. Gracia. Personal communication. 2022.

[28] S. Reck et al. “Muon bundle reconstruction with KM3NeT/ORCA using graph convolutional networks”. In: *37th International Cosmic Ray Conference. 12-23 July 2021. Berlin*. Mar. 2022, 1048, p. 1048.

Acknowledgements

This part is dedicated to the people who supported me during this thesis. I am grateful for the opportunity to work on a master's thesis at ECAP, even if I was not able to participate in everyday life at the institute to the extent I had hoped due to the pandemic. Nevertheless, I would like to thank the people who made my work at the institute as pleasant as possible. A special thanks goes to

- Thomas Eberl, for the opportunity to work on this project and the helpful weekly discussions and valuable feedback,
- Rodrigo Gracia-Ruiz and Jannik Hofestädt for helping out with supervision and always being approachable,
- Johannes Schumann and Robert Lahmann for undertaking shifting duties with me,
- My office colleagues Maximilian and Jutta for the pleasant and relaxed atmosphere
- and my brother Oliver Schneider for proof reading my thesis

Thanks again!

Erklärung / Statement of Authorship

Hiermit bestätige ich, dass ich diese Arbeit selbstständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe.

I hereby certify that this thesis has been composed by me and is based on my own work, unless stated otherwise.

Erlangen, 02. May 2022

Martin Schneider