

UNIVERSITEIT VAN AMSTERDAM

FACULTEIT DER NATUURWETENSCHAPPEN,  
WISKUNDE EN INFORMATICA

---

MASTER THESIS

---

# Reconstruction of High-energy Neutrino-induced Particle Showers in KM3NeT.

---

BY

---

**K. W. Melis**

---

Amsterdam, November 2014

*Supervisors:*

Dr. R. Bruijn

Prof. Dr. P.M. Kooijman

## Abstract

Sources of high-energy cosmic rays can be identified by measuring the direction- and energy-flux of cosmic neutrinos. The KM3NeT neutrino detector is being constructed in the deep Mediterranean sea to detect this flux. Neutrinos of all flavors can induce a detectable particle shower in the detector. The direction and energy of this shower gives an estimate of the direction and energy of the neutrino. In this research, a method has been developed to reconstruct the parameters of neutrino-induced showers with energies  $> 10^4$  GeV in the KM3NeT detector. The performance of the reconstruction method has been tested using simulated shower events. The shower direction can be reconstructed with a median error in the direction of less than  $2^\circ$ . The energy resolution is  $\pm 15\%$ . By selecting a subset of the reconstructed events, a direction resolution better than  $1^\circ$  can be achieved. The estimated shower position may be sufficiently accurate to distinguish between electron-and tau-neutrinos. This could lead to a neutrino source analysis based on all properties of detected neutrinos.







# Contents

<b>1</b>	<b>Introduction</b>	<b>8</b>
1.1	Cosmic rays . . . . .	8
1.1.1	Cosmic ray spectrum . . . . .	9
1.1.2	Fermi acceleration . . . . .	10
1.1.3	Cosmic ray sources . . . . .	10
1.2	Neutrino sources . . . . .	12
1.2.1	Pion decay . . . . .	12
1.2.2	Cosmic neutrinos . . . . .	12
1.2.3	Cosmogenic neutrinos . . . . .	13
1.2.4	Atmospheric neutrinos . . . . .	13
1.3	Large volume neutrino detectors . . . . .	13
1.4	Motivation and research question . . . . .	14
<b>2</b>	<b>The KM3NeT neutrino detector</b>	<b>15</b>
2.1	Detection principle . . . . .	15
2.1.1	Neutrino interactions . . . . .	15
2.1.2	Event signatures . . . . .	16
2.1.3	Cherenkov light . . . . .	17
2.2	Theoretical bounds on estimated neutrino parameters . . . . .	17
2.3	The KM3NeT phase-1.5 detector . . . . .	18
2.3.1	Photo-multiplier tubes . . . . .	18
2.3.2	Digital optical modules . . . . .	19
2.3.3	Detector outline . . . . .	19
2.4	Optical background . . . . .	20
2.4.1	Atmospheric muons . . . . .	20
2.4.2	Potassium-decay and bio-luminescence . . . . .	21
2.5	Data acquisition and time-over-threshold technique . . . . .	21
2.5.1	Time-over-threshold technique . . . . .	21
<b>3</b>	<b>Monte Carlo simulations</b>	<b>23</b>
3.1	Neutrino source, propagation and interaction . . . . .	23
3.2	Shower simulation . . . . .	23
3.3	Photon emission by particle showers . . . . .	24
3.3.1	Total number of emitted photons . . . . .	24
3.3.2	Wavelength distribution . . . . .	24
3.3.3	Electromagnetic and hadronic showers as a line source of light . . . . .	24
3.3.4	Longitudinal emittance profile . . . . .	24
3.3.5	Angular emission distribution . . . . .	25

3.4	Photon propagation . . . . .	26
3.5	PMT angular acceptance and quantum efficiency . . . . .	27
3.6	Combined photon hit . . . . .	27
3.7	Optical background . . . . .	27
3.8	PMT electronics . . . . .	27
3.9	Time-over-threshold . . . . .	28
<b>4</b>	<b>Shower Reconstruction</b>	<b>30</b>
4.1	Reconstruction principle . . . . .	30
4.1.1	Likelihood principle . . . . .	30
4.1.2	Shower likelihood function . . . . .	30
4.1.3	Reconstruction principle . . . . .	31
4.2	Position fit . . . . .	31
4.2.1	Position fit principle . . . . .	31
4.2.2	Position shower likelihood function . . . . .	32
4.2.3	Hit selection criterion . . . . .	33
4.2.4	Performance . . . . .	34
4.2.5	Filter on outliers . . . . .	35
4.3	Direction and energy fit . . . . .	37
4.3.1	Principle . . . . .	37
4.3.2	No-timing shower likelihood function . . . . .	38
4.3.3	PMT selection . . . . .	44
4.3.4	Scanning method . . . . .	44
4.3.5	Performance . . . . .	46
<b>5</b>	<b>Results</b>	<b>48</b>
5.1	Main sample . . . . .	48
5.2	Comparison with <i>km3</i> -simulated events . . . . .	50
5.3	Comparison with <i>aashowerfit</i> . . . . .	51
5.4	Event selection . . . . .	53
5.4.1	Event selection criteria . . . . .	53
5.4.2	Error in estimated shower position . . . . .	54
5.4.3	Comparison . . . . .	54
<b>6</b>	<b>Discussion and conclusions</b>	<b>56</b>
6.1	Discussion . . . . .	56
6.2	Conclusions . . . . .	57
<b>7</b>	<b>Appendices</b>	<b>60</b>
	Appendix A: Maximization of the point source likelihood function . . . . .	60
	Appendix B: Best estimate of the shower energy . . . . .	61
	Appendix C: Isotropic hemisphere sampling . . . . .	63

## Note on the contents

This thesis is partly based on reference material. The chapters "Introduction", "The KM3NeT detector" and "Monte Carlo simulations" are completely based on literature. In the latter chapter, the code of the 'JPP' software package has also been used to describe the relevant processes.

The work presented in the remaining chapters, the "Shower Reconstruction", "Results", and "Discussion and conclusions", is based on the work of the author. At some points in these chapters, references are made to used literature.

The appendices are derived on general assumptions and common knowledge.



# 1 Introduction

Looking up at the night sky, the light of stars being emitted billions of years ago can be observed. With telescopes covering a broad range of the electromagnetic spectrum, even the radiation of the very early Universe can be observed. However, the range of view in the optical high-energy regime is limited due to interactions of photons with interstellar matter and radiation<sup>[1]</sup>. In figure 1, the visible horizon in the optical domain is plotted.

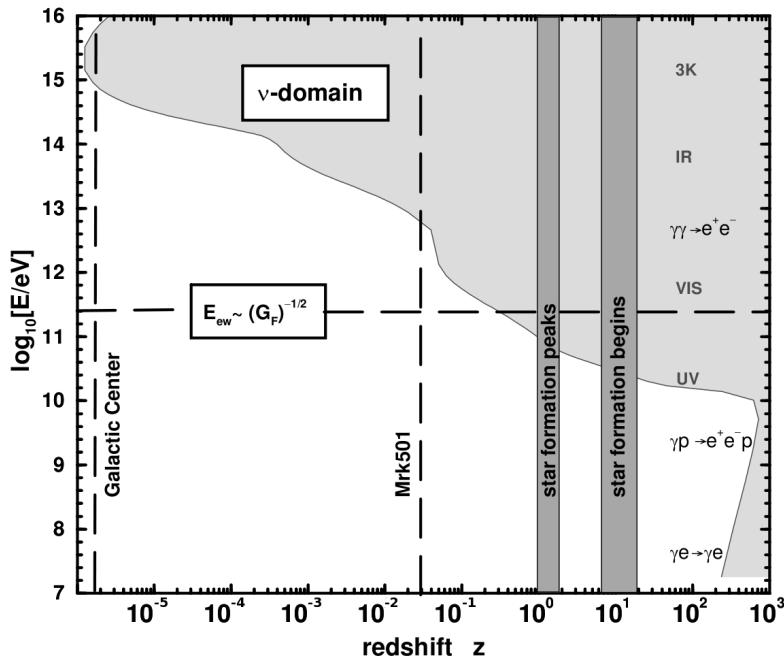


Figure 1: Visible range in the Universe, where the shaded area is invisible in the optical domain. The redshift on the horizontal axis gives a measure of the distance into the Universe. The vertical axis gives the photon energy. Figure taken from<sup>[1]</sup>

As can be seen, for photon energies about  $4 \cdot 10^6$  GeV, the horizon is limited to the distance from the Earth to the galactic center. If only photons could provide information about the Universe, many properties of the Universe would still be unknown.

## 1.1 Cosmic rays

In 1912, Hess detected high energetic particles penetrating the Earth's atmosphere, originating from the Universe<sup>[2]</sup>. At relatively low energies, these 'cosmic rays' are verified to constitute mainly of protons, helium and iron nuclei<sup>[3]</sup>. Since the discovery, the origin and the energy spectrum of cosmic rays have been of great interest in astroparticle physics.

### 1.1.1 Cosmic ray spectrum

Numerous experiments have investigated the flux of cosmic rays as function of energy. Figure 2 shows the combined results.

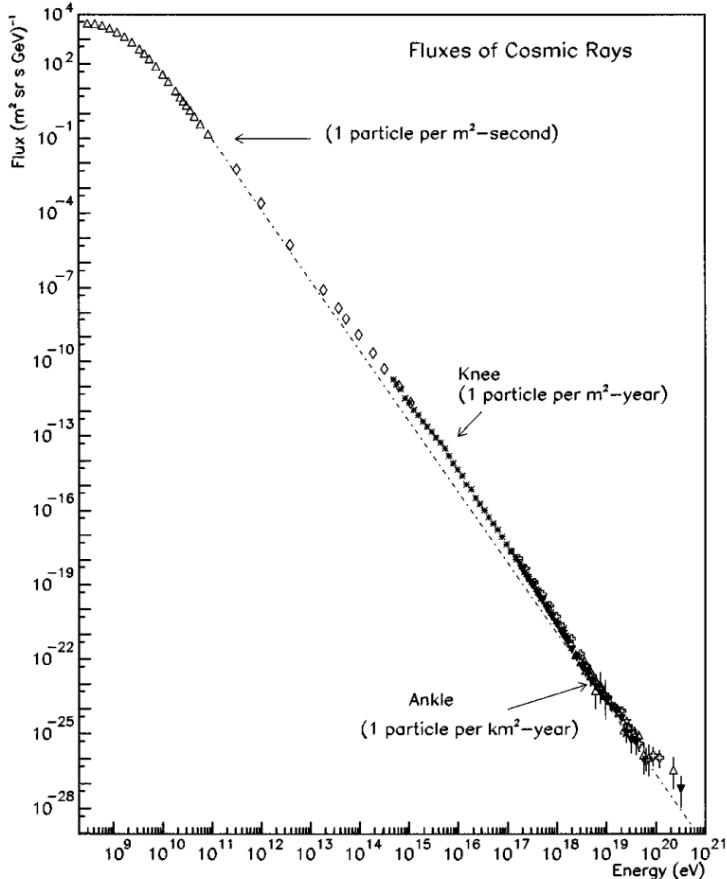


Figure 2: Cosmic ray energy spectrum. Data compiled by S. Swordy.

The dashed line represents a fit of the measured energy spectrum, and is given by  $dN_{CR}/dE_{CR} \propto E_{CR}^{-2.7}$ . From the cosmic ray energy spectrum, two questions arise.

1. What causes the discontinuous change in the spectrum at the 'knee' and the 'ankle'?
2. Which cosmological objects are capable of producing cosmic rays with energies up to  $3 \cdot 10^{20}$  eV<sup>[4]</sup>?

It is believed that neutrinos in the energy range above the knee originate from extragalactic sources<sup>[5]</sup>. A hypothetical mechanism that provides an answer to the second question is the Fermi acceleration mechanism<sup>[6]</sup>. In this mechanism, the very-high energy cosmic rays originate from lower energetic cosmic rays, which are accelerated in fast moving magnetic fields.

### 1.1.2 Fermi acceleration

Fermi acceleration is proposed as a candidate for accelerating charged particles in a fast moving magnetic field, as can be produced by the matter in a relativistic moving shock wave. Two types of Fermi acceleration have been suggested<sup>[7]</sup>. In the first order acceleration, cosmic rays are accelerated each time they transverse the shock front. In the second order Fermi acceleration, the particles 'bounce off' from the shock wave. In this section, only the first order Fermi acceleration will be discussed, since the second order mechanism fails to explain the observed cosmic ray energy spectrum.

Suppose a certain cosmological object emits matter and is producing the necessary fast moving shock wave. It can be shown that a charged particle transversing the shock wave from either side gains momentum<sup>[8]</sup>. If the magnetic field of the cosmological object is strong enough, the charged particles will follow a curved path. This may cause the particle to transverse the shock wave multiple times. Even if the momentum gain of a single pass is relatively low, this mechanism can lead to a very-high energy particle. It can be shown that the resulting energy spectrum is of the form:

$$\frac{dN_{CR}}{dE_{CR}} \propto E_{CR}^{-\alpha} \quad (1.1)$$

For the second order Fermi acceleration,  $\alpha$  is expected to equal  $\alpha = 2^{[7]}$ .

This mechanism could explain the observation of very-high energy cosmic rays. However, the question remains which cosmological objects are capable of accelerating particles this way.

### 1.1.3 Cosmic ray sources

The Fermi shock acceleration requires that a highly energetic charged particle crosses the wave front multiple times. The magnetic field of the cosmological object should account for this.

Suppose the acceleration region of a cosmological object has a radius  $R$  and the magnetic field has strength  $B$ , then only particles with the Larmor radius  $R_L \equiv \frac{E}{qB} \leq R$  ( $E$  and  $q$  are the energy resp. charge of the particle) are able to cross the shock wave multiple times. This criterion is known as the Hillas criterion. The maximal energy that a particle can obtain via Fermi acceleration from a source is then given by  $E_{max} = qBR$ .

In figure 3, the radius of the acceleration region and the magnetic field strength of several cosmic objects are plotted. For comparison, the maximum achievable energy  $E_{max}$  is plotted for protons ( $q = 1$ ) and iron nuclei ( $q = 26$ ) at  $E_{max} = 10^{11}$  GeV and  $E_{max} = 10^{12}$  GeV.

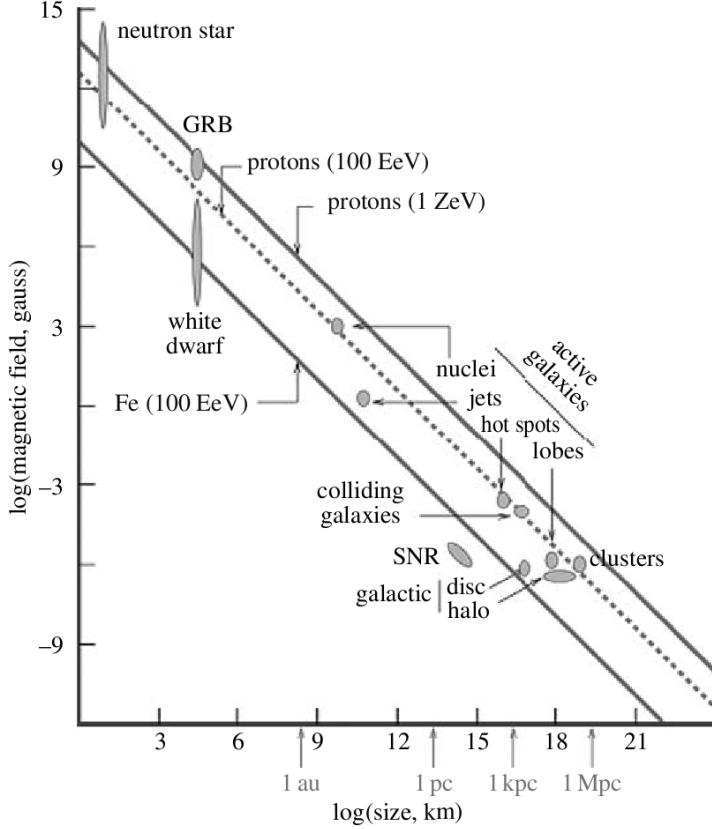


Figure 3: The strength and radius of the magnetic field for several cosmological objects. The lines correspond to the maximum achievable energy via Fermi acceleration for protons and iron nuclei as function of the strength and radius of the magnetic field. Figure taken from<sup>[6]</sup>

All cosmological objects plotted above the dashed line are theoretically capable of accelerating protons up to energies of  $10^{11}$  GeV. In the following sections, three of the drawn objects will be discussed.

**Active galactic nuclei.** Active galactic nuclei form the centers of galaxies. They are confirmed to consist of massive black holes<sup>[9]</sup>. As the surrounding matter is accelerated into the black hole, cosmic rays are emitted. When accelerated, these cosmic rays can reach energies up to  $10^{10}$  GeV<sup>[10]</sup>.

**Supernova remnants.** As a supernova explodes, electrons are accelerated to high energies<sup>[11]</sup>. As these electrons interact with surrounding interstellar matter, cosmic rays up to energies of  $10^5$  GeV can be produced<sup>[12]</sup>.

**Gamma-ray bursts.** The origin of Gamma-ray bursts has been a mystery since their discovery in 1973<sup>[13]</sup>. The emitted photons from Gamma-ray bursts form an expanding

'fireball'. In this fireball, as a result of the Fermi-acceleration mechanism, charged particles are accelerated<sup>[14]</sup>. The energy of these cosmic rays ( $> 10^{11}$  GeV) makes them the most highly energetic cosmic rays<sup>[15]</sup>.

Identifying the sources of high-energy cosmic rays would be a major achievement in astroparticle physics. The nonlinear path of cosmic rays from the source to the Earth, as a result of interactions with interstellar magnetic fields, complicates such an analysis.

Neutrinos are produced in the vicinity of sources of high-energy cosmic rays. Since neutrinos are chargeless and the interaction cross-section of neutrinos is small, these neutrinos are hardly influenced by interstellar matter, thus following a straight path from the source to the Earth. This makes them ideal probes for the origin of (high-energy) cosmic rays. However, distinguishing the flux of these 'cosmic neutrinos' from the flux of other neutrino sources forms a major challenge.

## 1.2 Neutrino sources

The observed neutrino flux at Earth mainly originates from three neutrino sources. For all these three sources, the majority of the produced neutrinos are the result of decaying pions.

### 1.2.1 Pion decay

Pions and kaons are produced when cosmic rays interact with matter or photons. In the decay of these pions, muons are produced, decaying to neutrinos:

$$\begin{aligned}\pi^+ &\rightarrow \mu^+ + \nu_\mu \rightarrow e^+ + \nu_e + \bar{\nu}_\mu + \nu_\mu \\ \pi^- &\rightarrow \mu^- + \bar{\nu}_\mu \rightarrow e^- + \bar{\nu}_e + \nu_\mu + \bar{\nu}_\mu\end{aligned}\tag{1.2}$$

The pion decay process predicts the production of neutrinos in ratio  $\nu_e : \nu_\mu : \nu_\tau = 1 : 2 : 0$  (not distinguishing between neutrinos and anti-neutrinos). A fraction of the cosmic ray energy will be transferred to the neutrino. As a result, the energy spectrum of the produced neutrinos will follow the same energy spectrum as the source of the cosmic rays.

The direction, energy and flavor distribution of neutrinos observed at Earth is dependent on the source. If the energy, direction and flavor of many detected neutrinos are reconstructed, the different fluxes should be distinguishable. A large contribution of detected neutrinos fulfilling the properties of a cosmic neutrino source could prove the existence of this source. In the following sections, the flux properties of the three main contributions to the neutrino flux at Earth will be discussed.

### 1.2.2 Cosmic neutrinos

Based on the matter density and cosmic ray flux, a measurable flux of neutrinos should be produced in the vicinity of a source of cosmic rays. As a result, the flux of cosmic

neutrinos is strongly correlated to the position of the source. Since the energy spectrum of the produced neutrinos follows the energy spectrum of the source, a  $dN_\nu/dE_\nu \propto (E_\nu)^{-2.0}$  cosmic neutrino energy spectrum is expected. As a result of neutrino oscillations during the path from the source to the Earth, the ratio of cosmic neutrinos arriving at the Earth equals approximately  $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$ .

### 1.2.3 Cosmogenic neutrinos

As a cosmic ray propagates through the Universe, it interacts with low-energy photons, such as those originating from the cosmic microwave background. In this interaction, pions are produced, decaying to neutrinos. The corresponding neutrinos ('cosmogenic' neutrinos) are predicted to carry a large fraction of the energy of the initial cosmic ray, ranging up to  $E_\nu = 10^{12}$  GeV<sup>[7]</sup>.

Since the production of cosmogenic neutrinos is not bounded to cosmological objects, the observed cosmogenic neutrino flux will be isotropically distributed. Due to neutrino oscillations, the ratio between cosmogenic neutrinos at arriving Earth will approximately equal  $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$ .

### 1.2.4 Atmospheric neutrinos

A cosmic ray interacting with molecules in the Earth's atmosphere can create a shower of particles. The pions and kaons in this shower can decay to neutrinos as described in section 1.2.1. These neutrinos are referred to as atmospheric neutrinos.

Atmospheric neutrinos are isotropically distributed. The energy spectrum of atmospheric neutrinos should follow the energy spectrum of cosmic rays ( $dN_{CR}/dE_{CR} \propto (E_{CR})^{-2.7}$ ). However, due to interactions of the produced pions with molecules in the atmosphere, the expected energy spectrum of atmospheric neutrinos is softer (approximately  $dN_\nu^{atm.}/dE_\nu^{atm.} \propto (E_\nu^{atm.})^{-3.7}$ )<sup>[16]</sup>. Neutrino oscillations play a minor role in the vicinity of the earth, resulting in an expected  $\nu_e : \nu_\mu : \nu_\tau = 1 : 2 : 0$  flavor ratio for atmospheric neutrinos.

The lifetime of pions increases with energy. At high energies, interactions of the pion with molecules in the atmosphere are likely to occur before the pion decays to neutrinos. although models of this process are not in agreement with each other, the flux of neutrinos at high energies is generally believed to be dominated by the decay of charmed mesons. The corresponding flux is called the prompt flux<sup>[17]</sup>. The energy spectrum of the prompt flux has not been determined yet.

## 1.3 Large volume neutrino detectors

In the previous chapter, measuring the neutrino flux has been suggested as a way to identify the sources of high-energy cosmic rays. However, due to the small interaction cross-section, detecting neutrinos is not trivial.

In 1961, Markov and Zhelenznykh introduced the idea of a neutrino detector covering a large volume of (sea)water or ice with optical sensors<sup>[18]</sup>. The optical sensors will detect the

Cherenkov light emitted by the charged particles produced in an interaction of a neutrino with the seawater. From the detected pattern of photon hits, the flavor, direction and the energy of the neutrino can be estimated.

## 1.4 Motivation and research question

Many studies using large volume neutrino detectors aim at identifying cosmic neutrino sources. Of key importance in such an analysis is the reconstruction of the neutrino parameters. The significance with which a source can be identified is positively related to the accuracy of the estimated direction, energy and flavor of the observed neutrinos.

In the past years, muons produced in charged current interactions of muon-neutrinos have been used as the main channel of many neutrino source analyses. Since a muon track has a long range, the direction can be determined with sub-degree accuracy. However, muon tracks form only a fraction of the total number of detectable events. In addition, if the muon track is not fully contained in the instrumented volume, only a lower limit on the muon energy can be determined.

The remaining fraction of the detected neutrino interactions is characterized by the production of a particle shower. At the start of this research, it was believed that the direction resolution of showers is insufficient to be used for neutrino source analyses. The recent observation of shower-like events with shower energies up to  $E_S \approx 2 \cdot 10^6$  GeV in the IceCube neutrino detector has revived the interest for shower-like events<sup>[19]</sup>. The energy resolution of showers can be expected to be satisfactory, since the full shower energy is dissipated in a short range.

In this research, an algorithm to reconstruct the parameters of high-energy shower-like events has been designed. For neutrino energies  $E_\nu > 1 \cdot 10^5$  GeV, the expected background from atmospheric neutrinos is below 1 event per year<sup>[19]</sup>. In this research, the shower energy range of interest holds  $E_S > 1 \cdot 10^4$  GeV. The question this research will answer is:

What resolution is achievable in reconstructing the direction and energy of high-energy showers in the KM3NeT neutrino detector?

## 2 The KM3NeT neutrino detector

Based on the design as proposed by Markov and Zheleznykh, several large-volume neutrino detectors have been constructed, such as IceCube, Super-Kamiokande and ANTARES<sup>[18]</sup>. Currently under construction in the deep Mediterranean Sea is the KM3NeT neutrino detector. When fully operational, it will consist of five or six neutrino detectors (the 'building blocks') at different locations, each covering a large volume of seawater with optical sensors.

The building blocks of the KM3NeT infrastructure will consist of thousands of optical sensors. In order to obtain physical results without constructing and testing the full detector, smaller subsets of detection blocks will be made operational first. The relatively small 'phase-1' detector is expected to be constructed during the year 2015. This detector will be followed by the 'phase-1.5' detector, which forms a complete building block. In this research, the phase-1.5 detector will be used as reference detector.

### 2.1 Detection principle

A neutrino can be detected if it interacts with matter in the vicinity of the KM3NeT detector. This interaction can lead to the production of relativistic charged particles, propagating through the seawater. The Cherenkov light emitted by these particles can be detected by the optical sensors. In the following sections, these three steps will be discussed in more detail.

#### 2.1.1 Neutrino interactions

A neutrino can interact via the deep inelastic scattering of a neutrino on a nucleon of a molecule. In this weak interaction, either a  $W^\pm$ - (charged-current interaction) or a  $Z^0$ -boson (neutral-current interaction) is exchanged between a parton and the neutrino. In figure 4, the Feynman diagrams of the neutrino interaction processes are drawn:

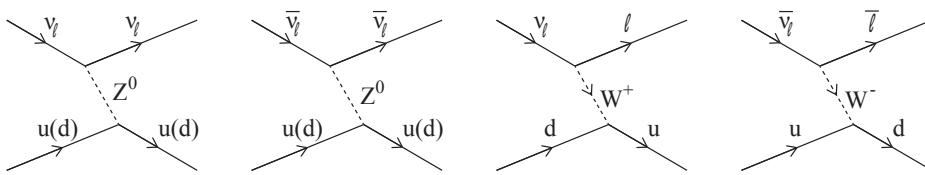


Figure 4: Deep inelastic scattering of a (anti-)neutrino on a nucleon. The two left diagrams give the neutral-current interaction, whereas the right two diagrams are charged-current interactions.

In a neutral-current interaction, a neutrino is produced, likely to leave the detector undetected. In the charged-current interaction, a lepton with the same flavor as the neu-

trino is produced. In all deep inelastic scattering processes, a fraction of the neutrinos four-momentum is transferred to the target nucleon.

### 2.1.2 Event signatures

A neutrino can be detected only if relativistic charged particles are produced in the weak interaction of a neutrino. These particles can be produced by the momentum transfer from the neutrino to the nucleon, or by the lepton produced in the charged-current interaction. The different event signatures are schematically summarized in figure 5:

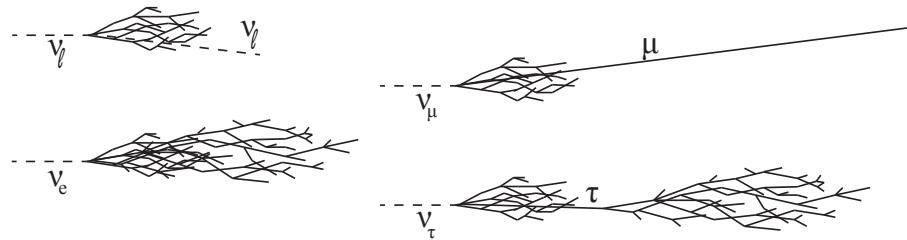


Figure 5: Possible event signatures of neutrinos interacting with seawater. Hadronic resp. electromagnetic showers are given by the small resp. large showers.

The transferred momentum from the neutrino to the parton is in general high enough for the production of hadrons. The decay of the produced hadrons (mainly pions) and interactions between the hadrons and the water lead to the production of secondary hadrons, photons and neutrinos. The photons (mainly created by the decay of neutral pions) can cause additional electromagnetic showers. The collection of the resulting shower is called a hadronic shower<sup>[20]</sup>.

**Neutrinos produced in neutral-current interactions are likely to leave the detector undetected.** On the other hand, leptons produced in charged-current interactions will leave a distinct event signature in the detector.

A muon produced via a charged-current muon-neutrino interaction has a relatively long lifetime. As a result, the muon leaves a track-like signature through the detector. For muon energies above  $10^4$  GeV, the length of the track exceeds 10 km<sup>[21]</sup>. Due to this long leverage arm, the direction of the muon can in general be reconstructed with sub-degree accuracy. If the muon emits high-energetic photons, electromagnetic showers can be produced along the track<sup>[20]</sup>.

An electron produced in a charged-current electron-neutrino interaction or via the decay of the tau produced in a charged-current tau-neutrino interaction will emit photons due to bremsstrahlung, scattering and ionization. The emitted photons posses a high fraction of the electron's energy, and may produce an electron-positron pair. These can emit photons again and again. All particles produced this way (the 'secondary' particles) form an electromagnetic shower<sup>[20]</sup>.

In case of a charged-current electron-neutrino interaction, the electromagnetic and hadronic shower overlap. The tau lepton produced in the charged-current interaction of a tau-neutrino has a lifetime long enough to produce a so-called 'double-bang' signature.

### 2.1.3 Cherenkov light

A charged particle propagating through a transparent medium with a velocity higher than the group velocity of light in this medium emits Cherenkov light at an angle

$$\cos(\theta_C) = \frac{1}{n_g \beta} \quad (2.1)$$

relative to the direction of the particle, where  $n_g$  denotes the group refractive index of the medium and  $\beta = v/c$  gives the ratio between the velocity of the charged particle and the speed of light in vacuum.

The number of emitted photons is dependent on the wavelength  $\lambda$  of the emitted photon and is described by the Frank-Tamm equation:

$$\frac{d^2N}{dx d\lambda} = \frac{2\pi}{\alpha \lambda^2} \left(1 - \frac{1}{n_g^2 \beta^2}\right) \quad (2.2)$$

where  $dx$  corresponds to one meter along the path of the charged particle,  $d\lambda$  gives a wavelength range of one meter and  $\alpha \approx 137$  is the electromagnetic coupling constant. A charged particle will dissipate all its energy in the form of Cherenkov radiation, until the point it becomes non-relativistic or decays.

## 2.2 Theoretical bounds on estimated neutrino parameters

A perfect reconstruction method can exactly reconstruct the true energy  $E_S^{tru.}$  and (average) direction  $\vec{d}_S^{tru.}$  of a shower (or muon track) induced by a neutrino interacting in the detector. Due to the deep inelastic scattering process, the reconstructed shower direction and energy will in general differ from the true neutrino direction and energy. This introduces a limit on the best achievable reconstruction accuracy.

In figure 2.2, the shower energy and direction is compared to the energy  $E_\nu^{tru.}$  and direction  $\vec{d}_\nu^{tru.}$  of the interaction neutrino ( $\theta_{\nu,S}$  gives the angle between  $\vec{d}_\nu^{tru.}$  and  $\vec{d}_S^{tru.}$ ):

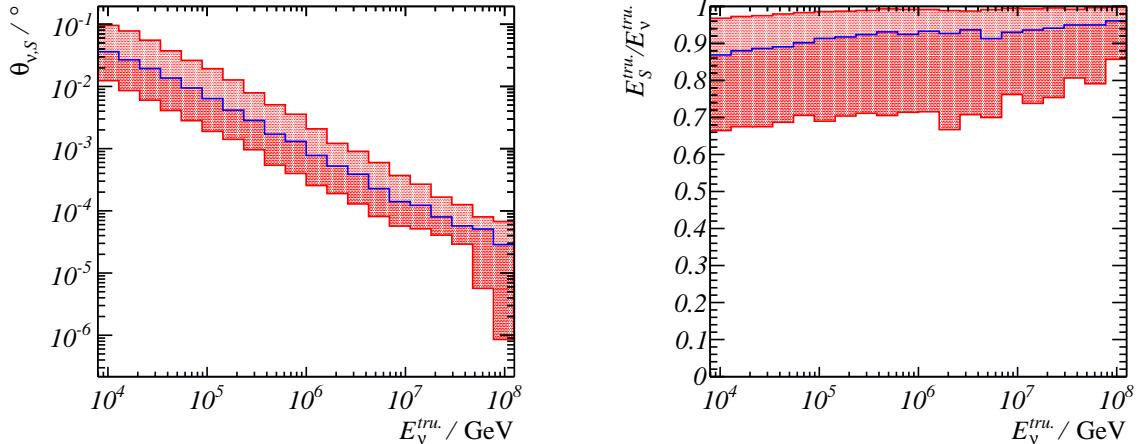


Figure 6: Median (blue) and 14% and 86% quantiles (red) of the angle between the neutrino- and shower direction (left) and of the fraction of the neutrino energy carried by the shower (right) as function of the true neutrino energy. This figure is obtained from Monte Carlo simulations.

As can be seen, the accuracy of the reconstructed neutrino direction and energy is limited. Especially at high energies, this limit is in general negligible with the resolution of the reconstruction method.

### 2.3 The KM3NeT phase-1.5 detector

The key detection units of the KM3NeT detector, the optical sensors, are photo-multiplier tubes (PMTs). 31 PMTs are grouped in a 'digital optical module', which are attached to vertical strings in the deep sea. In this section, the outline of the KM3NeT phase-1.5 detector will be discussed.

#### 2.3.1 Photo-multiplier tubes

A PMT consists of a cathode, anode and a series of dynodes in a vacuum. When a photon hits the first cathode, a photo-electron can be emitted as a result of the photo-electric effect. This photo-electron is accelerated by an electric field induced by the series of dynodes. Each time, a (photo-)electron hits a dynode, more electrons are emitted, resulting in a cascade of electrons hitting the anode. The resulting current on the anode produces a measurable voltage (the 'pulse'). From the amplitude of the pulse, the number of electrons hitting the anode can be estimated.

Not each photon hitting the PMT surface will be detected. The detection efficiency of a PMT is characterized by its angular acceptance (the probability that a photon incident at a certain angle hits the cathode) and its quantum efficiency (the probability that a photon hitting the cathode will emit a photo-electron).

The stochastic process from an emitted photo-electron leading to a cascade of electrons hitting the anode can be described by the time it takes for the cascade to propagate through the PMT (the 'transit time') and the number of electrons hitting the anode after a single emitted photo-electron is being emitted. For the used PMTs in the KM3NeT detector, both are approximately normally distributed. On average, the values of these standard deviations of the PMTs are approximately 2ns for the transit time and  $0.3 \cdot \sqrt{n^{e,abs}}$  for the number of electrons.

### 2.3.2 Digital optical modules

In order to withstand the pressure of the water in the deep sea, the optical sensors are grouped in pressure resistant glass spheres. The sphere is called a digital optical module, or DOM in short. Each DOM contains 31 PMTs, pointing outward through the exterior of the vessel. In order to increase the effective detection surface of the PMTs, a reflector ring around each PMT guides the light to the PMT. In figure 7, a schematic overview of such a DOM is given:

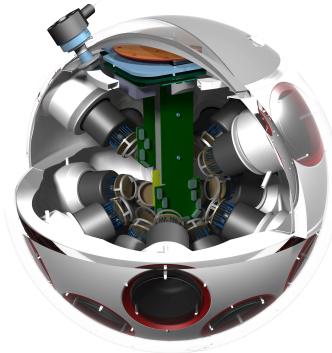


Figure 7: Schematic overview of a digital optical module (DOM) of the KM3NeT detector (image property of the KM3NeT collaboration)

In this figure, the PMTs are colored black, the reflector rings are drawn in red. In the center, the electronics board that collect and process the measured pulses can be seen (green).

### 2.3.3 Detector outline

The DOMs of the KM3NeT will be positioned along vertical strings. Eighteen DOMs are attached to each string, anchored to the seabed and kept vertical by buoys. The DOMs on a string are separated by 36m, starting of 100m from the anchor.

In figure 8, the positioning of the strings on the seabed of the phase-1.5 detector is given:

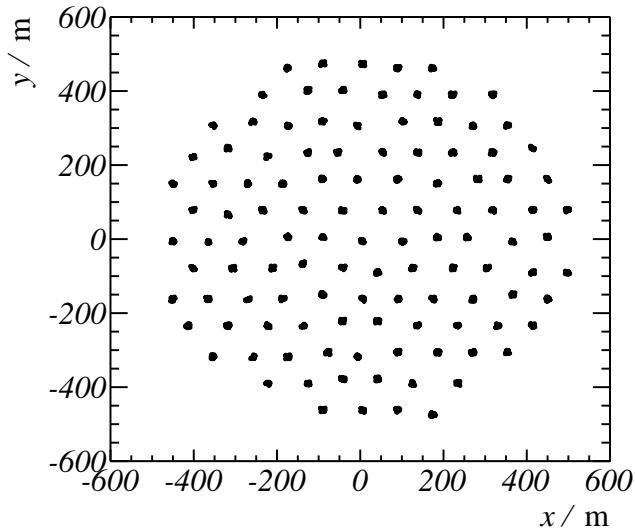


Figure 8: Positioning of the strings of the phase-1.5 detector on the seabed.

The optical modules of the phase-1.5 detector will instrument a cylindrical shape of radius  $\approx 500$  m and height  $\approx 600$  m, covering an instrumented volume of approximately  $0.5 \text{ km}^3$  seawater.

## 2.4 Optical background

The KM3NeT detector will be situated in the deep Mediterranean sea. As a result, optical background from natural processes and atmospheric muons is present. This optical background complicates the reconstruction of events, especially since the background photon hits are indistinguishable from signal photon hits. In this section, two sources of optical background will be discussed.

### 2.4.1 Atmospheric muons

When a cosmic ray interacts in the upper atmosphere, a particle shower can be created. Such a 'cosmic ray shower' can produce (bundles of) high energetic muons, with a lifetime long enough to reach the detector. The Cherenkov light emitted by these muons is indistinguishable from signal muon tracks with the interaction vertex outside the detector.

The background from atmospheric muons is dominant, although it is greatly reduced by the positioning of the detector in the deep sea. By doing this, the seawater covering the detector reduces the probability of a muon reaching the detector, whilst being of negligible influence on the propagation of a neutrino.

A special case of this atmospheric muon reduction is obtained by selecting only upward going muon tracks in the analysis. In this way, the probability that a atmospheric muon is being selected is greatly reduced.

### 2.4.2 Potassium-decay and bio-luminescence

Seawater contains the radioactive potassium isotope  $^{40}K$ . Photons are emitted by electrons produced in the  $\beta$ -decay of these isotopes. Potassium-decay forms a large fraction of the measured signal.

Some bacteria and organisms in the water are capable of producing light. The light emitted by these organisms is referred to as bio-luminescence.

The optical background of bio-luminescence and  $^{40}K$ -decay causes hits with a frequency of approximately 5 kHz. Typically, the intensity of the emitted light is not sufficient to be detected on multiple DOMs.

## 2.5 Data acquisition and time-over-threshold technique

The electronics board in each DOM collects the measured signal, processes it and sends it to the shore via a network of optical fibers.

A unique feature of the KM3NeT detector is that, at each time, all measured data is sent to shore. To suppress the dominating signal of optical background, dedicated filters in the shore station remove all data with a low probability that a neutrino interaction has occurred. For neutrino events in the highly-energetic regime this research is aiming at, it can be expected that all events in the vicinity of the detector pass at least one of these triggers.

### 2.5.1 Time-over-threshold technique

Mainly due to the significant optical background, the bandwidth of the optical fibers is not sufficient to continuously transmit the data of all PMTs. To solve this, the amount of transferred data is greatly reduced by introducing a time-over-threshold technique on the measured pulse of each PMT, before sending the data to shore.

In the PMT electronics, the output pulse of each PMT acts as the input voltage of a RC-circuit. In the time-over-threshold technique, the output pulse is compared to a certain threshold. In case the (absolute value of the) measured voltage exceeds this threshold, the moment of passing through threshold (the ' $t_0$ -time') and the time over threshold (the  $ToT$ -value) are recorded and send to shore. Each pair of the measured  $t_0$ -time and  $ToT$  value will from now on be referred to as a 'DAQ hit'.

In figure 9, the time-over-threshold technique is illustrated by plotting the output voltage of the comparator and a typical PMT output pulse:

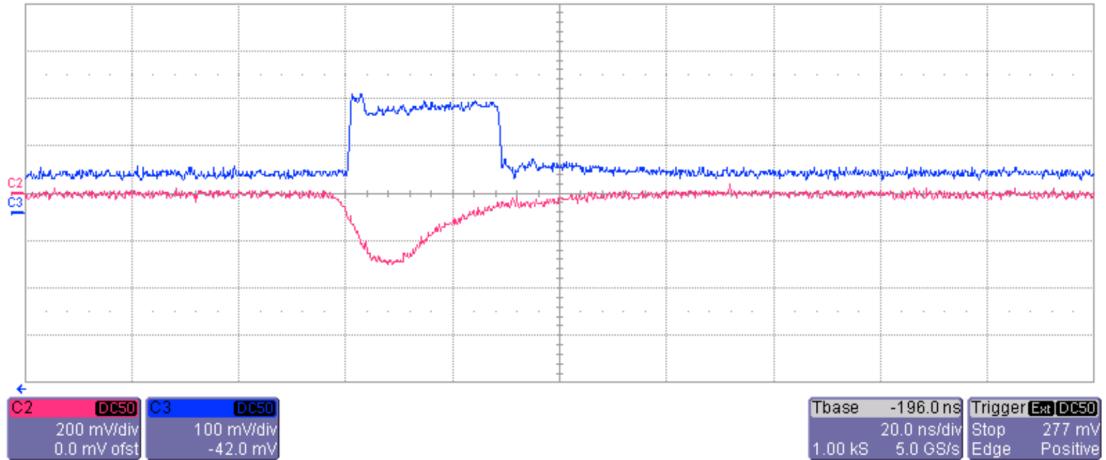


Figure 9: Measured PMT pulse (pink) and corresponding time-over-threshold output (blue) for a single photon hitting the PMT. Figure from<sup>[22]</sup>

In this figure, the output voltage was measured after a single photo-electron was emitted. If multiple photo-electrons are emitted shortly after each other, they contribute to the same DAQ hit. The *ToT*-value increases with the total charge. For a low total charge, this correlation is stronger than for high values. As a result, the resolution at low total charge is good, with a wide range of detectable pulse amplitudes.

## 3 Monte Carlo simulations

In this thesis, a method for estimating the parameters of neutrino induced showers in the KM3NeT detector will be presented. The shower parameters are not directly measurable. In addition, the hits originating from a shower are the result of a stochastic process. These two reasons, in combination with the fact that the KM3NeT phase-1.5 is far from completion, make it impossible to test the accuracy of the reconstruction routine using real data. For this reason, Monte Carlo simulated shower events will be used in this analysis.

The Monte Carlo simulation gives a detailed approximation of the reality. The full simulation consist of a chain of consequent processes. In this chapter, the different steps will be summarized.

### 3.1 Neutrino source, propagation and interaction

The Monte Carlo simulations start with the interaction of a neutrino with the matter in the vicinity of the detector. A set of neutrino interactions have been simulated in a container of seawater surrounding the detector (the generation volume). For these neutrino interactions the following conditions are set:

- Only charged-current electron-neutrino interactions.
- Neutrinos propagating through the generation volume from all directions
- A  $\frac{dN_\nu}{dE_\nu} \propto (E_\nu)^{-1.4}$  neutrino energy spectrum in a  $10^4 \text{ GeV} < E_\nu \leq 10^8 \text{ GeV}$  energy range.

Due to the first setting, a large fraction of the simulated events will contain an electromagnetic shower. The energy spectrum of the generated neutrinos is chosen such that a relatively large fraction of the neutrinos are high energetic. This spectrum doesn't represent the spectrum of a physical neutrino source.

In order to keep track of the cross-section of each of the simulated neutrino interactions, the probability that the neutrino is produced by a certain source, has transferred the Earth, propagates through the generation volume to finally interact, giving a triggered event with the above settings is calculated.

For this analysis, two sets of events were simulated. A relatively small sample (the 'test events') containing 6934 shower events has been used to design, test and optimize the reconstruction algorithm. In order to minimize biases, a larger sample of 69263 shower events (the 'main events') has been simulated to obtain the final results.

### 3.2 Shower simulation

From the simulated interaction products, electromagnetic and hadronic showers are simulated, using the GENHEN software package. The output of this simulation is a list of the parameters of all (secondary) particles in the event. It is beyond the scope of this research to cover these simulations in detail. Please see<sup>[23]</sup> for a detailed description of GENHEN.

### 3.3 Photon emission by particle showers

Each relativistic charged particle in a shower emits Cherenkov light. The simulated emission of photons is simplified by considering the collection of all relativistic charged particles, rather than the individual particles. As will be discussed, the photon emittance distribution of a shower is described by a total number of emitted photons, a wavelength-, longitudinal- and angular emission distribution.

#### 3.3.1 Total number of emitted photons

The dissipated energy in the form of photons of an electromagnetic shower is approximately equal to the shower energy<sup>[20]</sup>. For hadronic showers, a fraction of the shower energy is dissipated in the form of photons. From this it follows that the total number of emitted photons is proportional to the shower energy.

#### 3.3.2 Wavelength distribution

The number of emitted Cherenkov photons as function of the wavelength of the photon follows from the Frank-Tamm equation (eq. 2.2). The number of emitted photons as function of the wavelenght  $\lambda$  is approximately proportional to  $-\lambda^{-1}$ .

#### 3.3.3 Electromagnetic and hadronic showers as a line source of light

Using simulated shower events, the distribution of all particles in a shower event has been investigated<sup>[24]</sup>. In this study, the shower direction is defined as the average direction of all charged particles, weighted with their momentum.

One of the results is that 63% of all emitted light is produced in a cylinder with a radius less than 3 cm, concentric with the shower direction. Comparing this to the typical spacing between the DOMs ( $\approx 36\text{m}$ ) in the detector, the shower can be approximated as a line source of light.

#### 3.3.4 Longitudinal emittance profile

After the primary particle in a shower has decayed or interacted, more and more (charged) particles will be produced. As a result, the deposited energy increases. After reaching a maximum, the number of secondary particles and the energy of these particles decrease, resulting in a 'tail' of deposited energy. The longitudinal distribution of the energy deposition by a shower, and thus of the emitted light is well described by a gamma distribution<sup>[20]</sup>:

$$P_{(z|E_S)}^{\gamma,\text{emit}} = \frac{1}{\Gamma(a_{(E_S)})} \cdot \frac{z^{a_{(E_S)}-1} \cdot e^{-z/b}}{b^{a_{(E_S)}}} \quad (3.1)$$

Where  $z$  is the distance from the emission point to the interaction vertex position along the shower direction. The values of  $a$  and  $b$  depend on the shower energy via<sup>[21]</sup>:

$$a_{(E_S)} = 1.85 + 0.62 \cdot \ln \left( \frac{E_S}{GeV} \right) \quad (3.2)$$

$$b = 0.54$$

In figure 10, the longitudinal shower emission profile is drawn for several shower energies.

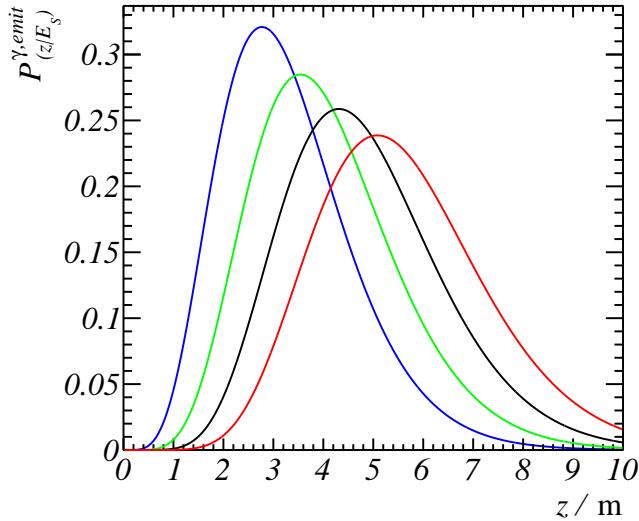


Figure 10: Longitudinal shower profile for shower energies  $E_S = 10^4$  GeV (blue),  $10^5$  GeV (green),  $10^6$  GeV (black) and  $10^7$  GeV (red).

### 3.3.5 Angular emission distribution

The majority of the secondary particles in a shower are aligned with the shower direction<sup>[24]</sup>. As a result, the angular photon emission distribution will be centered around the Cherenkov angle. This distribution is independent on the shower energy or the emission point on the shower axis, and is well described by<sup>[24]</sup>:

$$P_{(\theta^{emit})}^{\gamma,emit.} \propto e^{b \cdot (|\cos(\theta) - 1/n|)^a} \quad (3.3)$$

In the seawater surrounding the KM3NeT detector, the refractive index  $n$  equals approximately 1.38. The values of  $a$  and  $b$  are respectively  $a = 0.35$  and  $b = -5.40$ <sup>[24]</sup>.

In figure 11, the angular emission distribution is drawn:

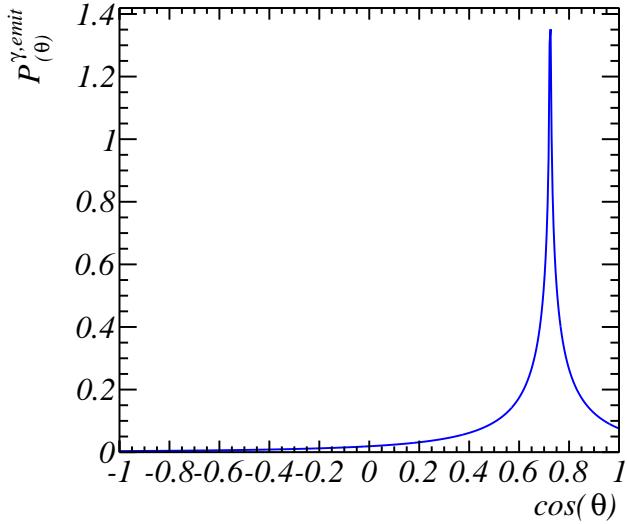


Figure 11: Angular photon emission distribution, for any value of the shower energy.

### 3.4 Photon propagation

The propagation of a photon in the seawater can be fully described by dispersion, scattering and absorption. All these processes are well-investigated. Despite of this, a generic simulation of the photon propagation through the water is not straightforward.

The photon paths of all emitted photons could be simulated using Monte Carlo techniques, resulting in a distribution of photon hits on the PMTs. Since the number of emitted photons increases linearly with the shower energy, a large number of photon paths should be simulated.

In order to save simulation time, the Monte Carlo simulations used in this analysis are based on the assumption that photons are scattered only once. As a result, the probability that a photon of wavelength  $\lambda$ , emitted from a certain position in a certain direction hits a PMT  $i$  can be approximated by a relatively simple integral.

This single scattering approximation can be expected to provide reliable probabilities for PMTs positioned close to the photon emission position, where the scattering length (in the order of 40 meters) is comparable to the distance from the PMT to the photon emission point. Based on the following reasoning, the single scattering approximation should also provide reliable results at longer distances:

- The total path length of a photon hitting a PMT at large distances will increase strongly with increasing scattering angles.
- The probability that a photon is absorbed increases with the total photon path length.
- The majority of detected photon hits at high distances will originate from single scattered photons.

The simulation of the photon propagation is based on the software package 'Sirene'. For this analysis, the method to simulate photon hits has been modified to include simulation of photon hits at large distances from the shower to the PMT.

### 3.5 PMT angular acceptance and quantum efficiency

The probability that a photon is detected by a PMT is characterized by the quantum efficiency and the angular acceptance of the PMTs. In the simulation, this probability is approximated by a linear interpolation of the measured angular acceptance and the quantum efficiency.

### 3.6 Combined photon hit

For the simulation of the photon hits, the JPP software has been used. Instead of simulating the emission, propagation and PMT-response of each individual photon, the simulated photon hits are drawn from a probability density function giving the probability to detect  $n_{i(t)}^{\gamma, sig.}$  signal photons on a PMT  $i$  in an infinitesimal time bin at time  $t$ , given the shower parameters  $S$ :

$$P_{(n_{i(t)}^{\gamma, sig.} | S)} \quad (3.4)$$

By assuming that the photon hits are Poisson distributed with some mean expected number of photon hits, where the expected number of photon hits is proportional to the shower energy, this function has been simplified such that it can be evaluated in a relatively small time.

### 3.7 Optical background

A set of hits has been added to the simulated signal photon hits in order to simulate background photon hits originating from  $^{40}K$ -decay and bio-luminescence. Hits from a single background photon are added with a frequency of  $5kHz$  per PMT. Background hits with a high enough amplitude to give multiple correlated hits on a single DOM are added with a frequency of  $100Hz$ . The algorithm used for this simulation is available in the JPP software package.

### 3.8 PMT electronics

As explained in section 2.3.1, the number of electrons hitting the anode as a result of a photo-electron being emitted is a stochastic quantity. It is convenient to give the number of electrons hitting the anode in units of 'effective' photo-electrons. With this definition, each single photon hit gives on average one effective emitted photo-electron.

In the Monte Carlo simulations, it is assumed that the effective number of emitted photo-electrons from a single photon hit is normally distributed with mean 1 (by definition) and standard deviation 0.3.

The statistical fluctuations in the transit time are also taken to be normally distributed, with a standard deviation of 2ns. The mean transit time will introduce a global shift of all hit times. Since hit times are only relevant relative to the other hits, the value of the mean transit time is arbitrary. A mean transit time of 0 nanoseconds has been chosen.

### 3.9 Time-over-threshold

At the moment of writing, a detailed simulation of the transition from a set of emitted photo-electrons to a set of DAQ hits is still under development. In this research, a simplified model of the time-over-threshold technique has been used, based on the simulation as available in the JPP software. although this model is still under development, it is thought to give a good approximation of the time-over-threshold technique.

In the time-over-threshold simulation, the pulse of  $n^{p.e.}$  effective photo-electron emitted at the same time increases linearly during the first 5 nanoseconds (the 'rise time', or  $t_{rise.}$ ) of the pulse, followed by a 'plateau' and an exponential decay with decay time  $t_{decay} = 4.5$  ns and decay rate  $R_{decay} = 0.02$ . As a result, the time-over-threshold of this set of emitted photo-electrons can be modeled by:

$$ToT_{(n^{p.e.})} = t_{rise} + t_{decay} \cdot \log(n^{p.e.}/R_{decay}) + C \cdot n^{p.e.} \quad (3.5)$$

The linear contribution to the time-over-threshold ( $C \cdot n^{p.e.}$ ) approximates the saturation effects in the PMT electronics. In the simulations, the proportionality constant  $C$  is set to  $C = 4.5$ . In figure 12, equation 3.5 is plotted:

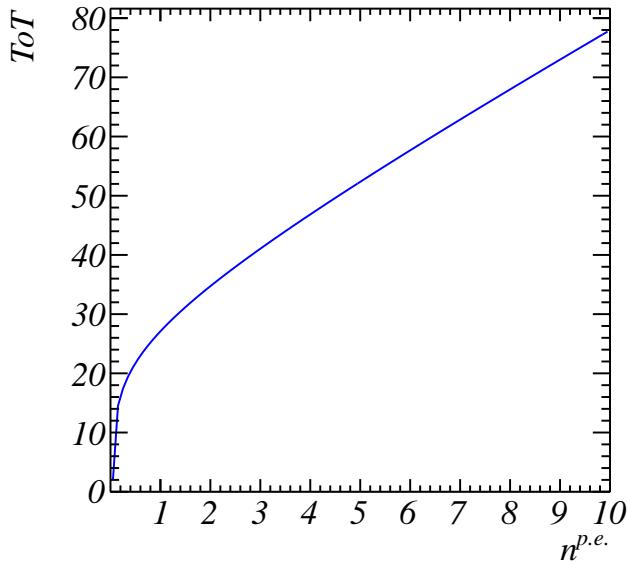


Figure 12: Relation between the time-over-threshold value as the result of  $n^{p.e.}$  effective photo-electrons emitted at approximately the same time (within the rise time of each other), as used in the simulation of the time-over-threshold process.

Suppose  $n_1^{p.e.}$  effective photo-electrons are emitted at  $t_1$ , and  $n_2^{p.e.}$  effective photo-electrons at  $t_2$ . If these two sets are emitted within a short time, then the individual contribution of the two pulses to the total pulse will be indistinguishable. If  $t_2^{p.e.}$  follows  $t_1^{p.e.}$  at a later time, while the pulse resulting from  $n_1^{p.e.}$  is still above the threshold, then the combined pulse will form a combination of the two pulses.

In the simulations, this effect is modeled by summing all emitted effective photo-electrons  $n_i^{p.e.}$  within the rise time of the previous emitted photo-electron ( $t_i - t_{i-1} < t_{rise.}$ ). The time-over-threshold value of the resulting pulse follows equation 3.5. If a second (summed) number of effective photo-electrons ( $n_i^{p.e.}$ ) is emitted at a time  $t_i$  within the time-over-threshold domain ( $t_j + ToT_{(n_j^{p.e.})}$ ) of a previous set  $j$ , then the total time-over-threshold value is taken to be the maximum of  $t_i + ToT_{(n_i^{p.e.})}$  and  $t_j + ToT_{(n_j^{p.e.})}$ .

## 4 Shower Reconstruction

This research aims at reconstructing the parameters of neutrino induced high-energy showers. It is not feasible to reconstruct the parameters of all the charged particles in a shower individually. Instead, the interaction vertex position ( $\vec{r}_S^{ver.}$ ), interaction time ( $t_S^{ver.}$ ), shower direction ( $\vec{d}_S$ ) and energy ( $E_S$ ) will be reconstructed. In this, the shower direction corresponds to the average direction of all charged particles in a shower, weighted with the momentum of each. The shower energy is the summed energy of all charged particles in the shower. In this chapter the developed shower reconstruction method will be discussed.

### 4.1 Reconstruction principle

#### 4.1.1 Likelihood principle

The criterion used to test each shower hypothesis ( $S = \{\vec{r}_S^{ver.}, t_S^{ver.}, \vec{d}_S, E_S\}$ ) is the likelihood principle. The likelihood principle states that the best estimate of the true shower parameters is the shower hypothesis that maximizes the shower likelihood function<sup>[25]</sup>.

The shower likelihood function  $L_{(S)}$  is proportional to the probability that a shower hypothesis  $S$  results in the detected set of DAQ hits with hit times  $\mathbf{t}^{DAQ}$  and  $ToT$ -values  $\mathbf{ToT}^{DAQ}$ :

$$L_{(S)} \propto P_{(\mathbf{t}^{DAQ}, \mathbf{ToT}^{DAQ}|S)} \quad (4.1)$$

As will be shown, the missing information on the distribution of individual photon hits complicates this shower likelihood function.

#### 4.1.2 Shower likelihood function

The sets of detected DAQ hits on different PMTs are uncorrelated to each other. From this it follows that the probability to measure the total set of DAQ hits ( $P_{(\mathbf{t}^{DAQ}, \mathbf{ToT}^{DAQ}|S)}$ ) on all PMTs of the detector can be factorized in the probabilities to measure the subset of DAQ hits ( $P_{(\mathbf{t}_i^{DAQ}, \mathbf{ToT}_i^{DAQ}|S)}$ ) on PMT  $i$ :

$$P_{(\mathbf{t}^{DAQ}, \mathbf{ToT}^{DAQ}|S)} = \prod_{i=1}^{N^{PMTs}} P_{(\mathbf{t}_i^{DAQ}, \mathbf{ToT}_i^{DAQ}|S)} \quad (4.2)$$

Three processes contribute to  $P_{(\mathbf{t}_i^{DAQ}, \mathbf{ToT}_i^{DAQ}|S)}$ ; the probability that a shower hypothesis  $S$  gives a certain set of photon hits on this PMT at times  $\mathbf{t}_i^\gamma$ , the probability that this set of photon hits results in a set of emitted effective photo-electrons  $\mathbf{n}_i^{p.e.}$  at times  $\mathbf{t}_i^{p.e.}$  and the probability that the set of effective photo-electrons produce the detected set of DAQ hits:

$$P_{(\mathbf{t}_i^{DAQ}, \mathbf{ToT}_i^{DAQ}|S)} = \iiint P_{(\mathbf{t}_i^{DAQ}, \mathbf{ToT}_i^{DAQ} | \mathbf{t}_i^{p.e.}, \mathbf{n}_i^{p.e.})} \cdot P_{(\mathbf{t}_i^{p.e.}, \mathbf{n}_i^{p.e.} | \mathbf{t}_i^\gamma)} \cdot P_{(\mathbf{t}_i^\gamma | S)} d\mathbf{t}_i^{p.e.} d\mathbf{n}_i^{p.e.} d\mathbf{t}_i^\gamma \quad (4.3)$$

Evaluation of this shower likelihood function requires an integral over multiple dimensions. Due to this, scanning the shower likelihood parameter space is not feasible. For this reason, the shower has been reconstructed using a different approach.

#### 4.1.3 Reconstruction principle

In this research, the shower parameters ( $S$ ) are estimated in two steps. In the first step, the shower position and interaction time are estimated. The direction and the energy of the shower are estimated in the second step. In both steps, simplified versions of the shower likelihood function are used.

### 4.2 Position fit

#### 4.2.1 Position fit principle

Suppose a shower emits light from a certain position. The number of PMTs present at a certain distance to the shower position increases with this distance. As a result, the majority of the detected signal photons from a shower will be detected at relatively large distances (see figure 13). Since the shower extends several meters, the shower can be approximated by a point source of light.

In figure 13, the hit times of signal and background DAQ hits ( $t_i^{DAQ}$ ) minus the interaction time ( $t_S^{ver.}$ ) are plotted as function of the distance from the hit PMT ( $\vec{r}_i^{DAQ}$ ) to the true interaction position  $\vec{r}_S^{ver.}$ :

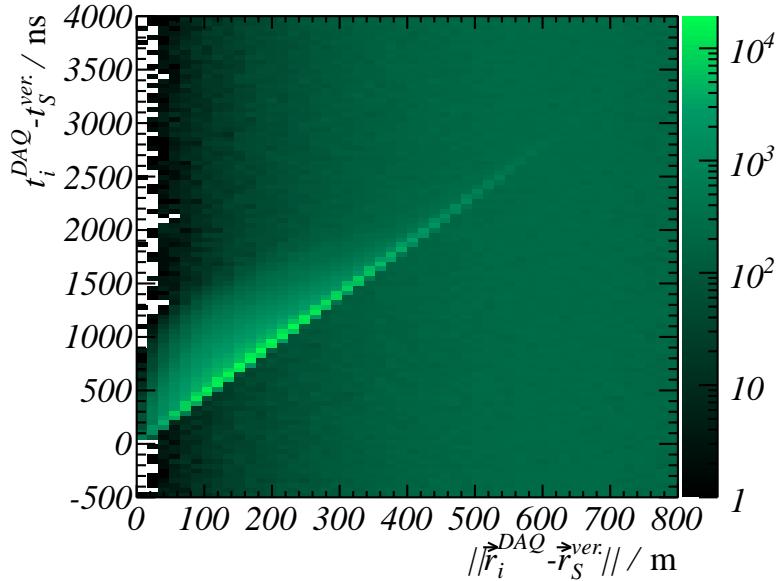


Figure 13: Number of signal and background DAQ hits, as function of the distance from the true interaction vertex position to the hit PMT and the hit time minus the neutrino interaction time  $t_S^{ver.}$ .

As can be seen, the hit times of a high number of DAQ hits are linearly related to the distance from the vertex position to the PMT. These 'early' hits origin from unscattered or softly scattered light. Also detected are 'late' hits (originating from scattered photons with long path lengths) and background hits (uncorrelated to the shower).

The linear relation between the hit times of the early hits and the distance from the PMT to the shower position is the key principle in the reconstruction of the shower position. By approximating the shower as a point source, the shower position and interaction time can be estimated analytically, as will be shown in the next section.

#### 4.2.2 Position shower likelihood function

In appendix A, an analytical function for the best estimated position  $\tilde{\vec{r}}_{PS}$  and time of emittance  $\tilde{t}_{PS}$  of a point source of light is derived, on basis of the following 'point source shower likelihood function':

$$L_{(\vec{r}_{PS}, t_{PS} | \vec{r}^{DAQ}, t^{DAQ})}^{PS} \propto \prod_{i=1}^{N^{DAQ,sel.}} P_{(\vec{r}_i^{DAQ}, t_i^{DAQ} | \vec{r}_{PS}, t_{PS})} \\ P_{(\vec{r}_i^{DAQ}, t_i^{DAQ} | \vec{r}_{PS}, t_{PS})} = \mathcal{N}_{(x_i | 0, \sigma^2)} \\ x_i \equiv \left( \vec{r}_i^{DAQ} - \vec{r}_{PS} \right)^2 - \bar{v}^2 \left( t_i^{DAQ} - t_{PS} \right)^2 \\ \mathcal{N}_{(x | \mu, \sigma^2)} \equiv \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2\cdot\sigma^2}} \quad (4.4)$$

Where  $\mathcal{N}_{(x | \mu, \sigma^2)}$  gives the normal distribution with mean  $\mu$  and standard deviation  $\sigma$  and the 'average speed'  $\bar{v}$  is approximately equal to the speed of light in water. The fitted point source parameters are independent of the value of the standard deviation  $\sigma$ .

The estimated point source position  $\tilde{\vec{r}}_{PS}$  will in general not estimate the vertex position  $\vec{r}_S^{ver.}$ , but rather the position of the 'average' of the emission distribution of the shower. In this analysis, the estimated point source position  $\tilde{\vec{r}}_{PS}$  is assumed to give an estimate of the 'shower maximum position'. The shower maximum position is defined as the position along the shower direction at which the longitudinal emission profile (equation 3.1) is at the maximum:

$$\tilde{\vec{r}}_{PS} \approx \vec{r}_S^{max.} \\ \vec{r}_S^{max.} \equiv \vec{r}_S^{ver.} + z_{(E_S)}^{max.} \cdot \vec{d}_S \\ z_{(E_S)}^{max.} \equiv 0.46 + 0.33 \cdot \ln \left( \frac{E_S}{GeV} \right) \quad (4.5)$$

The time of emittance of the point source  $\tilde{t}_{PS}$  will then correspond to the time of emittance from the shower maximum position:

$$\tilde{t}_{PS} \approx t_S^{max.} \equiv t_S^{ver.} + z_{(E_S)}^{max.}/c \quad (4.6)$$

The 'point source shower likelihood function' can be thought of as the average of the  $x_i$ -values of all selected hits. If all hits would be used in the position fit, background hits and late hits will distort the estimated shower maximum position. To avoid this, a hit selection is applied to select early signal photon hits (originating from unscattered- and softly scattered light).

#### 4.2.3 Hit selection criterion

In figure 14, the average number of photon hits on DOMs at several distances from the shower position are plotted as function of the hit time, minus the expected hit time of direct light ( $t_i^{exp.} \equiv t_S^{ver.} + \|\vec{r}_i - \vec{r}_S^{ver.}\| \cdot n^{water}/c$ ):

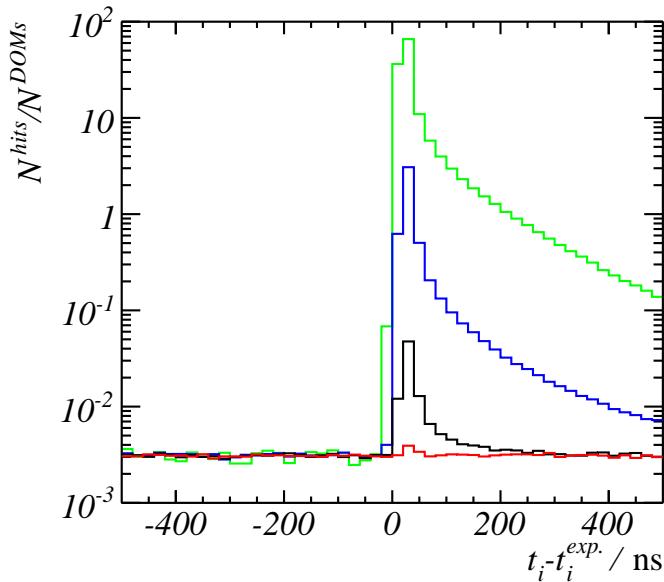


Figure 14: Average number of detected DAQ hits on each DOM as function of the hit time minus the expected hit time of direct light  $t_i^{exp.}$ . The distributions are shown for all DOMs positioned in 0-200 meter (green), 200-400 meter (blue) 400-600 meter (black) and 600-800 meter to the shower vertex position. The given DAQ hits originate from a single shower event of  $10^5$  GeV.

The PMTs on a DOM are positioned within a few decimeters of each other. Suppose a certain DAQ hit is the earliest signal DAQ hit on a DOM, then it is likely that this DAQ hit will be followed by DAQ hits on the other PMTs in the same DOM in a small time span. In figure 14, this can be observed as the sharp increase of detected hits at  $t_i \approx t_i^{exp.}$ . The hit selection criterion aims at selecting the earliest hit by the following hit selection criterion:

- A hit is selected if it is the first hit of a series of at least  $N$  hits on a DOM, and
- the hit times of two consecutive hits in this series do not differ more than  $\Delta_t$  nanoseconds.

Based on the number of selected hits and the spread in the  $x_i$ -values of the selected hits, the values  $N = 3$  and  $\Delta_t = 10\text{ns}$  have been chosen. Due to this rather strict criterion, a vast majority of all hits are not selected. However, this loss is acceptable in the view of the large number of hits that can be expected in a high-energy shower event.

#### 4.2.4 Performance

Reconstructing the shower position requires at least four selected hits. Applying the position fit algorithm to the test sample, consisting of 6934 shower events, results in 4867 reconstructable events. In figure 15, the number of unreconstructable events is plotted as function of the distance from the true shower maximum position to the center of the detector ( $R^{ev.}$ ) and as function of the true shower energy:

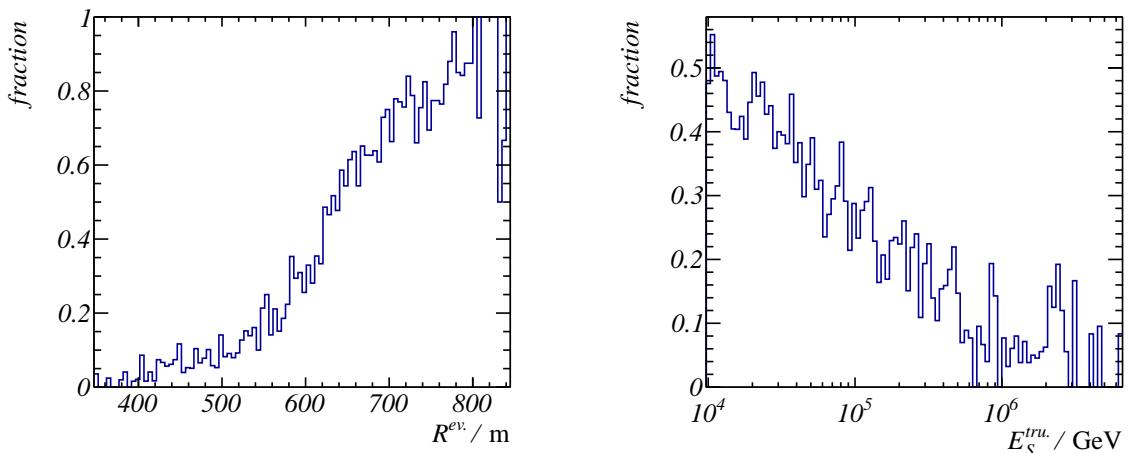


Figure 15: Fraction of the test events that were not reconstructed, as function of the distance from the interaction vertex to the center of the detector (left) and true shower energy (right).

As can be seen, the unreconstructable events constitute mainly of low energetic events at large distances to the detector center. It can be expected that these unreconstructable events, if they would have been reconstructed, will be badly reconstructed in the direction and energy fit.

Of all reconstructed events, the distance between the fitted and the true shower maximum position  $\delta_R \equiv \|\vec{r}_S^{max} - \tilde{\vec{r}}_S^{max}\|$  is plotted in figure 16 as function of the true shower energy ( $E_S^{tru.}$ ). The median of  $\delta_R$  for all events in each certain energy-bin is plotted in the right figure.

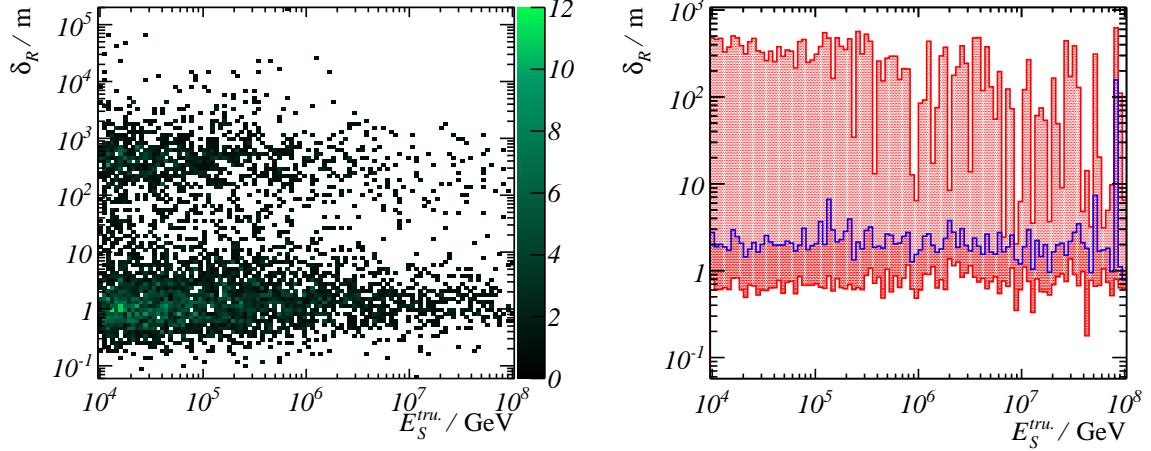


Figure 16: Difference between the reconstructed shower maximum position and the true shower maximum position as function of the true shower energy (left figure). In the right figure, the median (blue) and 84% and 16% quantiles (red) of all events in each energy-bin are given.

As can be seen, the majority of the events are reconstructed within an error of several meters. However, a fraction of the events are reconstructed with an error of more than 100 meters. To improve the position estimate, a second position fit is applied. For this fit, hits passing a second hit selection on basis of the estimated shower position are used.

#### 4.2.5 Filter on outliers

The position prefit principle as presented in section 4.2.2 is sensitive to selected DAQ hits originating from background- or late photon hits. In case such an 'outlying' hit is selected, the fitted shower position can be strongly distorted.

In figure 17, the reconstruction error  $\delta_R$  of each event is plotted as function of the maximal present outlier in the selected hits of that event, as calculated from the fitted position and expected hit times ( $t_i^{exp.} = \tilde{t}_S^{max.} + ||\vec{r}_i^{DAQ} - \vec{r}_S^{max}||/\bar{v}$ ):

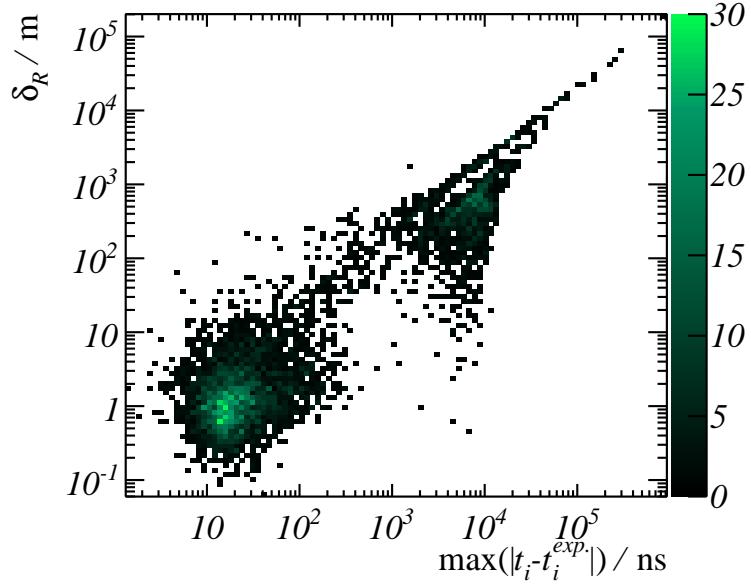


Figure 17: Difference between the reconstructed and true shower maximum position as function of the maximum time difference between the measured hit time and the expected hit time of unscattered light  $t_i^{exp}$  of all selected hits.

There is a clear relation between the reconstruction error and the maximal outlying hit. The reconstruction efficiency is improved by a second fit, using all selected hits with  $|t_i - t_i^{exp}| \leq 100 \text{ ns}$ . In figure 18, the reconstruction performance is shown after this second fit:

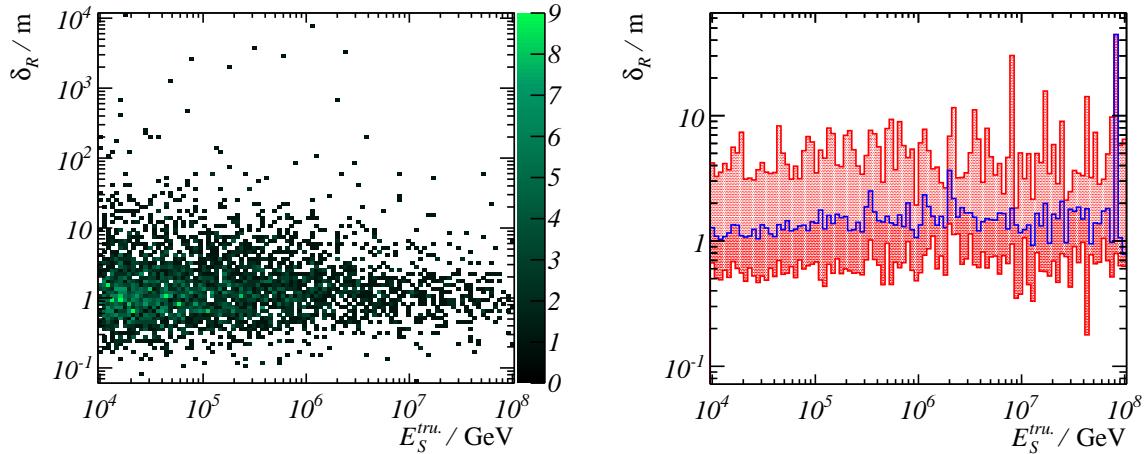


Figure 18: Difference between the reconstructed shower maximum position and the true shower maximum position as function of the true shower energy, after the filter on outlier hits has been applied. In the right figure, the median, 16 % and 84 % quantiles of all events in each bin are plotted.

As can be seen, only a few events with  $\delta_R > 100$  m remain. Some of the initially badly reconstructed events are reconstructed with a better accuracy, but the majority of these events are not reconstructable after removing the outlying hit. The unreconstructable events do not show a clear correlation with the position of the interaction vertex or the shower energy (figure 19).

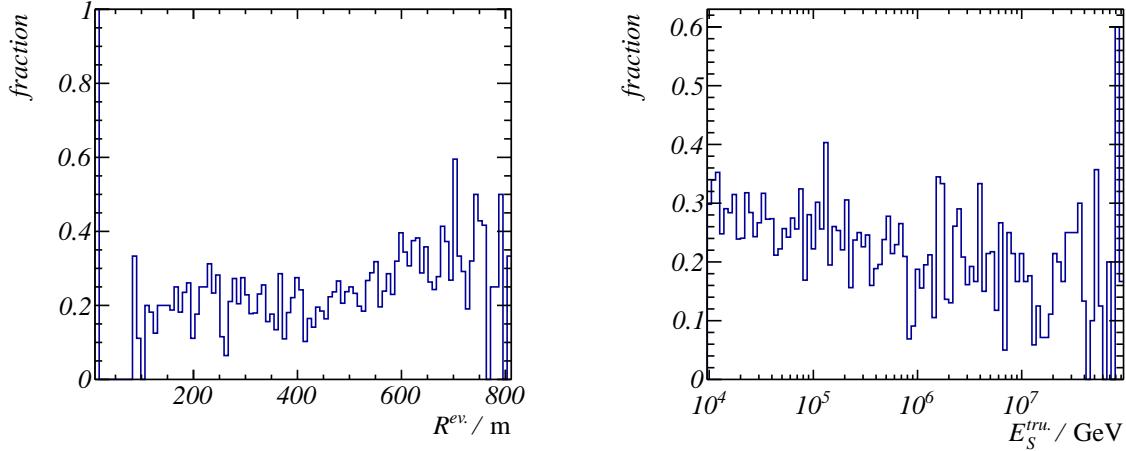


Figure 19: Fraction of reconstructed events that did not pass the second fitting procedure. The horizontal axis gives the distance from the interaction vertex to the center of the detector (left) and true shower energy (right)

### 4.3 Direction and energy fit

Estimating the shower (maximum) position and corresponding emittance time is the first step of the reconstruction method. The median error in the estimated shower position ( $\approx 1$ m) provides a solid fundament for the direction and energy fit. In this section, the method to estimate the shower direction and energy will be discussed.

#### 4.3.1 Principle

The subsequent production of secondary particles in a shower stops when the energy of each particle is less than a certain threshold<sup>[20]</sup>. As a result, especially at high energies, the number of emitted photons will be proportional to the total shower energy.

From the fact that the charged particles in a shower propagate along the shower direction, the angular distribution of the emitted photons will be centered at the Cherenkov angle. Although scattering, absorption and the PMT electronics influence the probability that an emitted photon is detected on a certain PMT, the distribution of detected photon hits is likely also to be maximal around the Cherenkov angle (with respect to the shower direction).

In figure 20, using the Monte Carlo truth information, the number of detected photon hits divided by the shower energy is plotted as function of the distance from the PMT to

the shower maximum position and the angle ( $\theta$ ) between the shower direction and the line from the shower maximum position to the PMT.

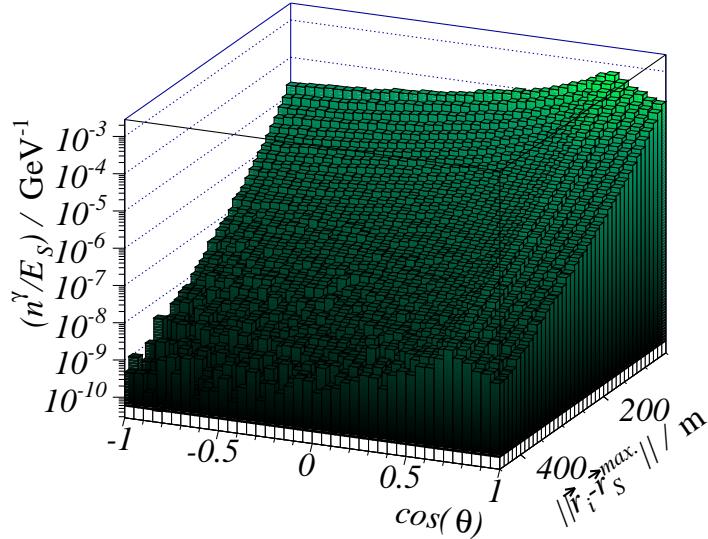


Figure 20: Detected number of photon hits divided by the energy of the shower, as function of the distance from the hit PMT to the shower maximum position and the angle ( $\theta$ ) between the line from the hit PMT to the shower maximum position and the shower direction.

As can be seen, the majority of the photon hits are detected at  $\cos(\theta) \approx 0.73$ . This angle corresponds to the Cherenkov angle in the seawater surrounding the KM3NeT detector.

The distribution of photon hits as function of the position and orientation of the PMT with respect to the shower direction, position and energy is the key in the 'no-timing shower likelihood function'. In this likelihood function, the information on the timing of the hits is disregarded, making it relatively easy to evaluate. This shower likelihood function will be derived in the following section.

The maximum of the no-timing shower likelihood function gives the best estimate of the shower direction and energy. The method to scan the no-timing shower likelihood function in the shower direction and energy parameter space will be given in section 4.3.4.

#### 4.3.2 No-timing shower likelihood function

Disregarding the timing information of the hits is equivalent to considering only the total number of photon hits and emitted effective photo-electrons:

$$\begin{aligned}
L_{(S)}^{no\ t} \propto P_{(\mathbf{T} \otimes \mathbf{T}^{DAQ}|S)}^{no\ t} &= \prod_{i=1}^{N^{PMTs}} P_{(\mathbf{T} \otimes \mathbf{T}_i^{DAQ}|S)}^{no\ t} \\
P_{(\mathbf{T} \otimes \mathbf{T}_i^{DAQ}|S)}^{no\ t} &= \sum_{n_i^{p.e.,tot.}=0}^{\infty} \sum_{n_i^{\gamma,tot.}=0}^{\infty} P_{(\mathbf{T} \otimes \mathbf{T}_i^{DAQ}|n_i^{p.e.,tot.})} \cdot P_{(n_i^{p.e.,tot.}|n_i^{\gamma,tot.})} \cdot P_{(n_i^{\gamma,tot.}|S)}
\end{aligned} \tag{4.7}$$

In this equation,  $P_{(n_i^{\gamma,tot.}|S)}$  gives the probability that a total number of  $n_i^{\gamma,tot.}$  photon hits hit a PMT  $i$ . The probability that this total number of photon hits results in a total number of emitted effective photo-electrons  $n_i^{p.e.,tot.}$  is given by  $P_{(n_i^{p.e.,tot.}|n_i^{\gamma,tot.})}$ . The probability  $P_{(\mathbf{T} \otimes \mathbf{T}_i^{DAQ}|n_i^{p.e.,tot.})}$  that the emitted photo-electrons will result in a set of measured set of DAQ hits does not depend the timing of the DAQ hits, as will be shown in section 4.3.2.

In the following sections, the probability density functions contributing to  $P_{(\mathbf{T} \otimes \mathbf{T}_i^{DAQ}|S)}^{no\ t}$  will be derived.

**Photon hit probability** Detected signal photon hits are uncorrelated to each other. As a result, the probability to detect a certain number of signal photon hits on a PMT is the outcome of a classical counting experiment. The corresponding probability density function is the Poisson distribution. If the probability to detect a certain number of background photon hits is taken to be Poisson distributed as well, then probability to detect a total number  $n_{i(t)}^{\gamma,tot}$  photon hits (background and signal) is also given by a Poisson distribution:

$$\begin{aligned}
P_{(n_i^{\gamma,tot.}|S)} &= \mathcal{P}_{(n_i^{\gamma,tot.}|\hat{n}_{i(S)}^{\gamma,tot.})} \\
\mathcal{P}_{(x|\mu)} &\equiv \frac{\mu^x \cdot e^{-\mu}}{x!}
\end{aligned} \tag{4.8}$$

The expected total number of photon hits ( $\hat{n}_{i(S)}^{\gamma,tot.}$ ) is the sum of the expected total number of signal photon hits ( $\hat{n}_{i(S)}^{\gamma,tot.,sig.}$ ) and background photon hits ( $\hat{n}_{i(S)}^{\gamma,tot.,bck.}$ ).

The expected total number of signal photon hits depends on the shower energy and the position and orientation of the PMT relative to the shower direction and position. This function has been evaluated directly from the probability density function used to simulate the events (see section 3.6). In this way, this probability density function describes the simulated events in the best possible way. In the used probability density function, the expected number of signal photons is proportional to the energy of the shower. The shower is again taken to be a point source of light, situated at the estimated shower maximum position.

Background photon hits originating from  ${}^{40}K$ -decay and bio-luminescence are expected on each PMT with a frequency of  $f^{bek.} \approx 5 \cdot 10^{-6} \text{ ns}^{-1}$ . Since background hits are present during the full event time, the total contribution of background hits is large compared to the contribution of signal photon hits. The reconstruction accuracy can be expected to increase with decreasing number of background hits.

Using the estimated shower maximum position and  $\tilde{t}_S^{max.}$ , the expected hit time of the earliest signal photon hits on a PMT can be approximated by  $t_i^{\gamma,exp.} = \tilde{t}_S^{max.} + \|\vec{r}_i - \tilde{\vec{r}}_S^{max.}\| \cdot n^{water}/c$ . In figure 21, the average number of measured signal and background photon hits on one PMTs in one event are plotted as function of  $t^\gamma - t^{\gamma,exp.}$ :

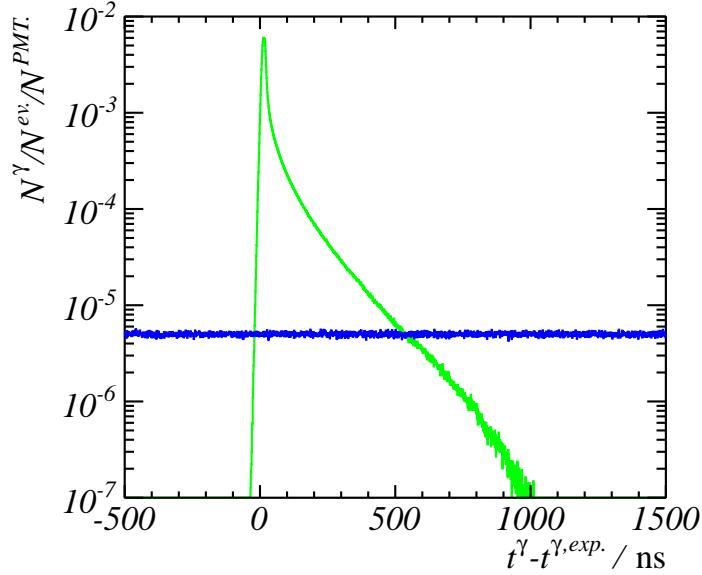


Figure 21: Hit time distributions for photon hits, normalized to the expected hit time as estimated from the fitted position. The distributions of signal photon hits (green) and background hits (blue) are plotted. The bin width is 1ns

The number of background hits can be reduced by selecting only DAQ hits in a time-window around  $t_i^{\gamma,exp.}$ . Based on figure 21, a time-window of  $-30 \leq t_i^{DAQ} - t_i^{\gamma,exp.} < 500$  ns has been chosen. Nearly all DAQ hits from signal photon hits are contained in this time-window, whereas the number of background hits is greatly reduced.

With this time-window, the expected number of photon hits from the shower and background is  $\hat{n}_i^{\gamma,tot.} = \hat{n}_{i(S)}^{\gamma,sig.,tot.} + f^{beg.} \cdot 530$ .

**Number of emitted photo-electrons** As discussed in section 2.3.1, each photon hitting the cathode of a PMT can result in a certain number of emitted effective photo-electrons. The probability that a single photon emits  $n^{p.e.}$  effective photo-electrons is taken as normally distributed, with a standard deviation of 0.3 and mean 1.

The probability that a total number of effective photo-electrons is emitted by a total number of photon hits  $n_i^{\gamma,tot.}$  follows from the normal distribution of the number of emitted effective photo-electrons from a single photon hit:

$$P_{(n_i^{p.e.,tot.}|n_i^{\gamma,tot.})} = \prod_{j=1}^{n_i^{\gamma,tot.}} \mathcal{N}_{(n_{i,j}^{p.e.}|1,0.3^2)} = \mathcal{N}_{(n_i^{p.e.,tot.}|n_i^{\gamma,tot.},0.3^2 \cdot n_i^{\gamma,tot.})} \quad (4.9)$$

**DAQ hit probability** In this section, the probability that a certain total number of emitted photo-electrons  $n_i^{p.e.,tot.}$  gives the detected set of DAQ hits will be derived. No conversion based on measurements from a set of DAQ hits to the total number of emitted photo-electrons is available. For this reason, this probability density function will be derived on basis of the approximations made in the ToT simulation method, as described in section 4.3.4.

With the approximations made for the simulation, each emitted photo-electron in  $n_i^{p.e.,tot.}$  contributes to the *ToT*-value of a detected DAQ hit in one of the following ways:

1. It forms the beginning of a pulse, or it is emitted within the rise time of a previous emitted photo-electron.
2. It is emitted within the time-over-threshold domain of a previous emitted photo-electron, but not within the rise time of the previous emitted photo-electron.

The photo-electrons of case one contribute to the ToT value of the corresponding *DAQ* hit via  $ToT_{(n^{mer})}$ , as given in equation 3.5. If all emitted photo-electrons on a PMT  $i$  would be of case one, then the number of emitted photo-electrons would equal:

$$n_i^{p.e.,tot.} = \sum_j \left[ ToT_{(ToT_{i,j}^{DAQ})}^{-1} \right] \\ ToT_{(x)}^{-1} \equiv W(e^{((x-5.0-4.5\cdot\log(50.0))/4.5)}) \quad (4.10)$$

Where  $ToT_{(x)}^{-1}$  is the inverse of equation 3.5, and  $W(z)$  denotes the Lambert *W*-function<sup>1</sup>.

Due to the presence of the 'case two' hits (approximately 11% of all DAQ hits), equation 4.10 does not hold. However, as will be shown next, it leads to a good approximation of the number of emitted photo-electrons resulting in a set of DAQ hits.

Using the simulated events, the sum of the inverse of all *ToT*-values of the measured DAQ hits on a PMT as function of the total number of effective photo-electrons on that PMT is plotted in figure 22. For comparison, the number of total effective photo-electrons for the same PMTs as function of the expected total number of photon hits is also given:

---

<sup>1</sup>The defining equation of the Lambert *W*-function is  $z = W(z) \cdot e^{W(z)}$

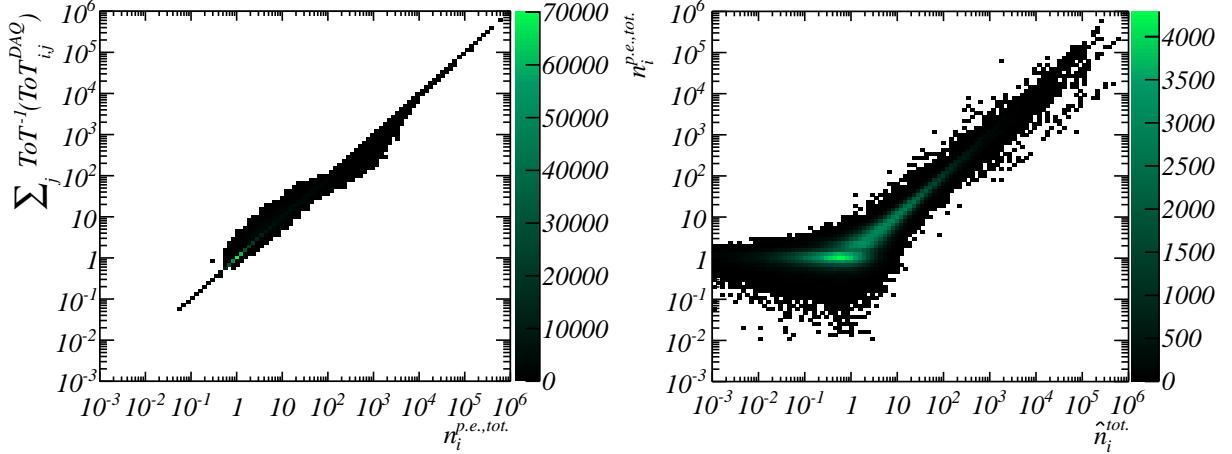


Figure 22: (Left) Sum of the inverse of all measured  $ToT$ -values on a PMT, as function of the total number of effective emitted photo-electrons. (Right) Total number of effective emitted photo-electrons as function of the expected total number of photon hits.

In the left figure, the contribution of 'case one' DAQ can clearly be observed. At lower values of the number of effective photo-electrons, the time-distribution of 'case two' hits is such that the sum of  $ToT$ -values is higher than that of a pulse during which all photo-electrons are emitted at the same moment. This indicates that a second photo-electron is emitted just before the pulse of a previous emitted photo-electron drops below the threshold. At higher  $n^{p.e.,tot.}$ , the estimated number of photo-electrons from the set of  $ToT$ -values is lower than the true number of emitted photo-electrons. This is the result of a relatively long pulse of multiple photo-electrons emitted within the rise time of each other, followed by some hits after this rise time, but with amplitudes so small that these do not contribute to the total pulse length.

The sum of the inverses of the  $ToT$ -values can differ from the total number of effective photo-electrons by orders of magnitude. However, compared to the spread in the total number of emitted effective photo-electrons with respect to the expected total number of photon hits (figure 22, right), the sum of the inverses gives a good estimate of the total number of emitted photo-electrons.

With the above approximation, the number of emitted effective photo-electrons can be estimated by:

$$\tilde{n}_{(ToT_i^{DAQ})}^{p.e.,tot.} \equiv \sum_{j=1}^{N_i^{DAQ,tot.}} \left[ ToT_{(ToT_{i,j}^{DAQ})}^{-1} \right] \quad (4.11)$$

The probability that a certain total number of effective photo-electrons  $n_i^{p.e.,tot.}$  gives the measured set of DAQ hits is then assumed to be one in case the estimated number of photo-electrons equals the number of emitted photo-electrons and zero otherwise:

$$P_{(\mathbf{ToT}_i^{DAQ} | n_i^{p.e.,tot.})} = \begin{cases} 1 & \tilde{n}_{(\mathbf{ToT}_i^{DAQ})}^{p.e.,tot.} = n_i^{p.e.,tot.} \\ 0 & \text{else} \end{cases} \quad (4.12)$$

**Recap** Having derived equations 4.8, 4.9 and 4.12, the no-timing likelihood is given by:

$$\begin{aligned} L_{(S|\mathbf{ToT}^{DAQ})}^{no\ t} &\propto \prod_i P_{(\mathbf{ToT}_i^{DAQ}|S)}^{no\ t} \\ P_{(\mathbf{ToT}_i^{DAQ}|S)}^{no\ t} &= \sum_{n_i^{p.e.,tot.}=0}^{\infty} \sum_{n_i^{\gamma,tot.}=0}^{\infty} P_{(\mathbf{ToT}_i^{DAQ}|n_i^{p.e.,tot.})} \cdot P_{(n_i^{p.e.,tot.}|n_i^{\gamma,tot.})} \cdot P_{(n_i^{\gamma,tot.}|\hat{n}_{i(S)}^{\gamma,tot.})} \\ P_{(\mathbf{ToT}_i^{DAQ}|S)}^{no\ t} &= \sum_{n_i^{\gamma,tot.}=0}^{\infty} \mathcal{N}_{(\tilde{n}_{(\mathbf{ToT}_i^{DAQ})}^{p.e.,tot.}|n_i^{\gamma,tot.}, 0.3^2 \cdot n_i^{\gamma,tot.})} \cdot \mathcal{P}_{(n_i^{\gamma,tot.}|\hat{n}_{i(S)}^{\gamma,tot.})} \end{aligned} \quad (4.13)$$

In figure 23, this distribution is plotted:

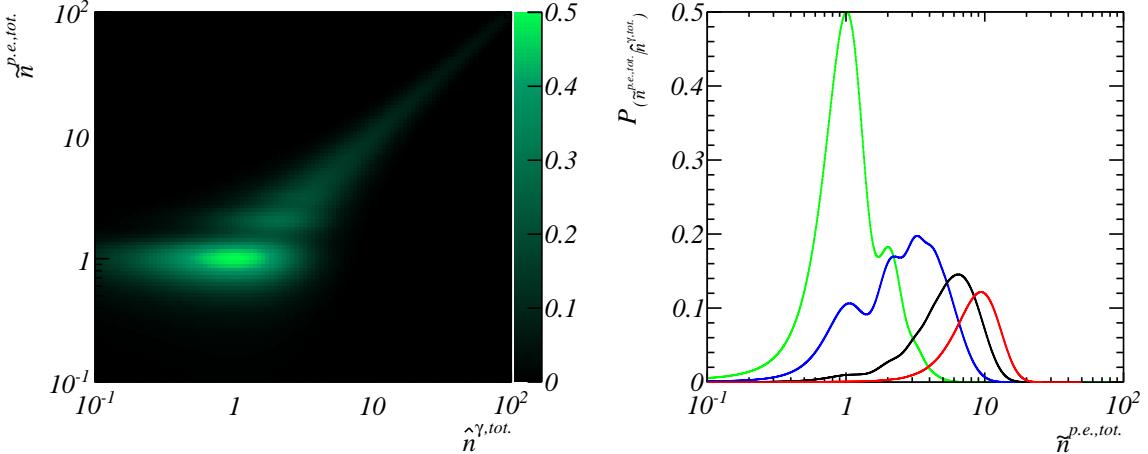


Figure 23: The distribution of estimated number of effective photo-electrons as function of the expected number of photons hitting a PMT (left). In the right figure, projections of this probability density function on the  $\tilde{n}^{p.e.,tot.}$ -axis at  $\hat{n}^{\gamma,tot.} = 1$  (green),  $\hat{n}^{\gamma,tot.} = 4$  (blue),  $\hat{n}^{\gamma,tot.} = 7$  (black) and  $\hat{n}^{\gamma,tot.} = 10$  (red) are drawn.

The distribution of  $P_{(\tilde{n}^{p.e.,tot.}|\hat{n}^{\gamma,tot.})}$  is dominated by the Poisson probability density function  $\mathcal{P}_{(n_i^{\gamma,tot.}|\hat{n}_{i(S)}^{\gamma,tot.})}$ . At a low number of expected photon hits ( $\hat{n}_{i(S)}^{\gamma,tot.} \leq 3$ ), the influence of the convolution of  $\mathcal{N}_{(\tilde{n}^{p.e.,tot.}|n^{\gamma,tot.}, 0.3^2 \cdot n^{\gamma,tot.})}$  at  $n^{\gamma,tot.} = \{0, 1, 2, 3, 4, 5\}$  can be observed. At higher number of photon hits, the Poisson distribution dominates.

### 4.3.3 PMT selection

If the expected number of photon hits on a PMT ( $\hat{n}_{i(S)}^{\gamma,tot.}$ ) differs significantly from the estimated number of emitted effective photo-electrons ( $\tilde{n}_i^{\gamma,tot.}$ ), then the best estimate of the shower direction and energy will depend mainly on 'matching' these two values. As a result, the estimated shower parameters can be distorted. For a set of ToT-values on a PMT with a small estimated total number of photo-electrons ( $\tilde{n}_i^{\gamma,tot.} < 3$ ), this may occur when a (background) photon hit is detected whereas the expected number of photon hits is very small. On the other hand, for  $\tilde{n}_i^{\gamma,tot.} > 100$ , the best fit of the energy will greatly depend on the expected number of photon hits on this single PMT.

In order to avoid this, all hit PMTs in an event with an estimated total number of emitted photo-electrons of  $\tilde{n}_i^{\gamma,tot.} < 3$  or  $\tilde{n}_i^{\gamma,tot.} > 100$  will be excluded from the direction and energy fit.

This PMT selection introduces a bias: PMTs with  $\tilde{n}_i^{\gamma,tot.} > 3$  and  $\hat{n}_{i(S)}^{\gamma,tot.} < 3$  will be selected, whereas PMTs with  $\tilde{n}_i^{\gamma,tot.} < 3$  and  $\hat{n}_{i(S)}^{\gamma,tot.} > 3$  won't. As a result, the estimated number of effective emitted photo-electrons will be systematically overestimated.

A positive side-effect of the PMT selection criterion is that the no-timing shower likelihood function can be simplified. Since the estimated number of emitted effective photo-electrons  $\tilde{n}_i^{\gamma,tot.}$  of a selected PMT must be larger than three, the number of expected photons that lead to this number is likely also to be higher than three (figure 23). In this regime, the Poisson distributed  $\mathcal{P}_{(n_i^{\gamma,tot.} | \hat{n}_{i(S)}^{\gamma,tot.})}$  dominates the no-timing shower likelihood function. As a result, the shower likelihood function can be approximated by:

$$\begin{aligned} L_{(S|\mathbf{T} \circ \mathbf{T}^{DAQ})}^{no\ t} &\propto \prod_{sel.PMT} P_{(\mathbf{T} \circ \mathbf{T}_i^{DAQ}|S)}^{no\ t} \\ P_{(\mathbf{T} \circ \mathbf{T}_i^{DAQ}|S)}^{no\ t} &= \sum_{n_i^{\gamma,tot.}=0}^{\infty} \mathcal{N}_{(\tilde{n}_{(\mathbf{T} \circ \mathbf{T}_i^{DAQ})}^{p.e.,tot.} | n_i^{\gamma,tot.}, 0.3^2 \cdot n_i^{\gamma,tot.})} \cdot \mathcal{P}_{(n_i^{\gamma,tot.} | \hat{n}_{i(S)}^{\gamma,tot.})} \\ &\approx \mathcal{P}_{(\tilde{n}_{(\mathbf{T} \circ \mathbf{T}_i^{DAQ})}^{p.e.,tot.} | \hat{n}_{i(S)}^{\gamma,tot.})} \end{aligned} \quad (4.14)$$

It is relatively straightforward to scan the three-dimensional parameter space of this shower likelihood function. In the following section, this procedure will be explained.

### 4.3.4 Scanning method

The shower direction and shower energy that maximize the no-timing shower likelihood function are found by scanning the likelihood space. For scanning, the estimated shower parameters  $\tilde{r}_S^{max}$  and  $\tilde{t}_S^{max}$  from the position fit will be used, leaving the three dimensional scanning over  $\vec{d}_S$  and  $E_S$ .

For a given shower direction, the shower energy that maximizes this shower likelihood function can be derived analytically. Based on the assumption that the expected number of photons hitting a PMT is proportional to  $E_S$ , this function is derived in appendix B.

This leaves a scan over the two-dimensional shower direction parameter space. As shown in figure 24, the shower likelihood function contains numerous local maxima.

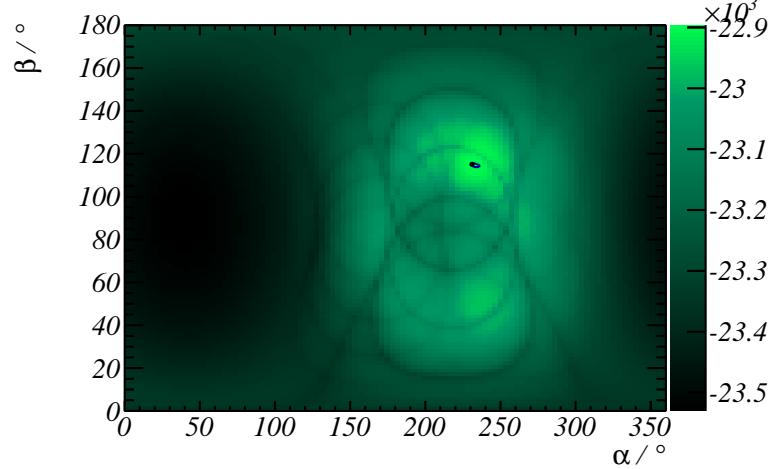


Figure 24: The logarithm of the likelihood function as function of the shower direction at the best estimate on the shower energy for a typical event.  $\alpha$  and  $\beta$  give the azimuth and zenith angle relative to the detector vertical axis. The small dot plotted corresponds to the true shower direction. The small ellipse surrounding this dot gives the circumference of the shower directions for which the likelihood value does not differ more than one from the maximum shower likelihood value.

In order to find the global shower likelihood maximum, the likelihood parameter space  $L_{(d_S | \tilde{r}_S^{max}, \tilde{r}_S^{max}, \tilde{E}_S)}^{no\ t}$  is 'roughly' sampled with 1932 test directions. The used sampling is described in appendix C. With this sampling, the probability that the global maximum is 'missed' is small. In figure 25, the test directions are plotted:

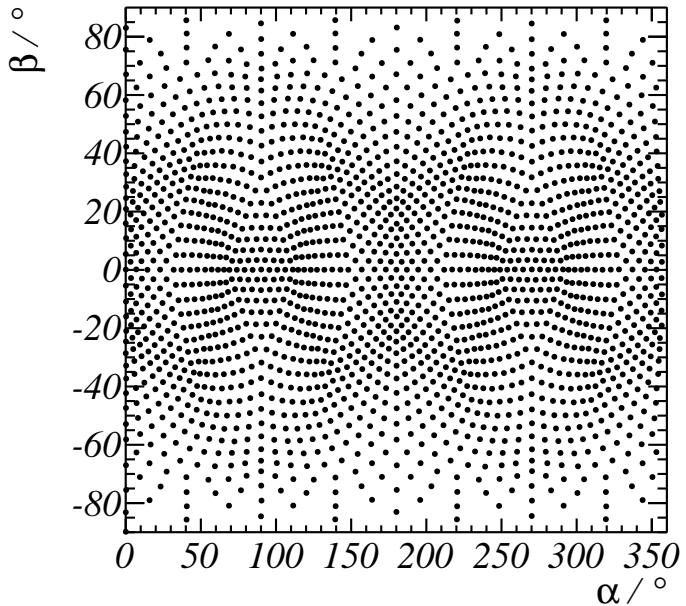


Figure 25: Distributions of the test directions for the 'rough' scanning method.

Since the number of sampling directions is finite, a median error in the estimated direction using this scanning method of approximately  $3^\circ$  can be expected. In order to get a better estimate, the 'rough' sampling is followed by a more precise sampling over directions close to the test shower direction with the highest likelihood value.

The asymptotic normality of the maximum likelihood estimator theorem states that shower likelihood maximum landscape close to the maximum is normally distributed. As a result, the log-likelihood landscape can be approximated by a parabola. The best estimate of  $\vec{d}_S$  is obtained by calculating the maximum of a fitted parabola through the log-likelihood landscape close to the best estimate from the rough sampling.

#### 4.3.5 Performance

To test the performance of the direction and energy fit procedure, the shower direction and energy of the events in the sample of 3674 test events that passed the position fit are estimated. The direction and energy fit requires 5 to 15 minutes per event (depending on the number of hits). In the following figures, the reconstruction error of the direction fit after applying the direction and energy fit procedure is given.

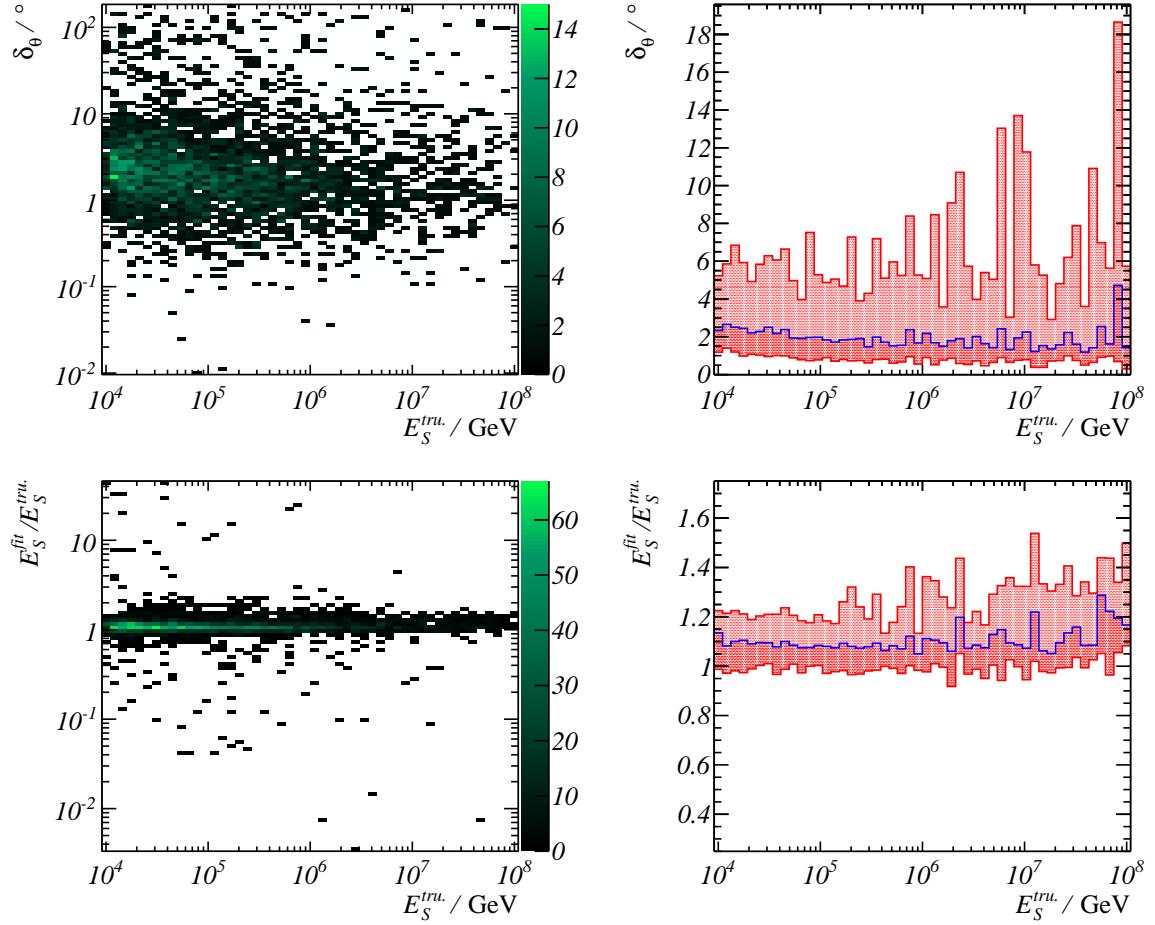


Figure 26: Angle between the reconstructed shower direction and the true shower direction ( $\delta_\theta$ ), and ratio between the fitted shower energy  $E_S^{fit}$  and the true shower energy  $E_S^{tru.}$  for all reconstructed events. The blue line resp. red lines in the right figures represent the median resp. 16 % and 84 % quantiles off the distribution in each energy-bin.

The median of the reconstructed direction of all events is approximately  $1.9^\circ$ . The shower energy is systematically overestimated. This is a direct result of the bias introduced by selecting only PMTs with an estimated total number of effective emitted photo-electrons larger than three and the corresponding approximations made for the shower likelihood function.

## 5 Results

In the preceding sections, all results were given for the 'test' set, consisting of 6934 events. The reconstruction method is tuned to give optimal results for these events. In order to avoid presenting biased results, the reconstruction algorithm has been applied to a large set of 69263 events (the 'main' sample). The results of this will be shown in section 5.1.

The method used to simulate the events is not the only one available. Main differences between different simulation methods are present in the emission of photons from showers and the propagation of photons through the seawater. The performance of the reconstruction algorithm on events simulated with the simulation package *km3* has been tested. The results will be shown in section 5.2.

Recently, a reconstruction algorithm ('*aashowerfit*') for high-energy showers in KM3NeT has been developed<sup>[26]</sup>. Since the aim of *aashowerfit* overlaps with the aim of this research, it forms an ideal reference for the reconstruction method developed in this study. In section 5.3, the reconstruction accuracy of *aashowerfit* is compared to the results of this analysis.

The angular error  $\delta_\theta$  between the true shower direction and the fitted shower direction and the ratio between the fitted shower energy  $E_S^{fit}$  and the true shower energy  $E_S^{tru.}$  will be used to quantify the performance of the reconstruction method. Since the position of the shower in the detector is in general of minor physical importance, it will not be presented in this section.

### 5.1 Main sample

Of the simulated 69263 shower-events in the main sample, 36866 have been reconstructed. All unreconstructed events did not pass the position fit. In figure 27, the accuracy of the reconstruction method obtained from the main sample is presented:

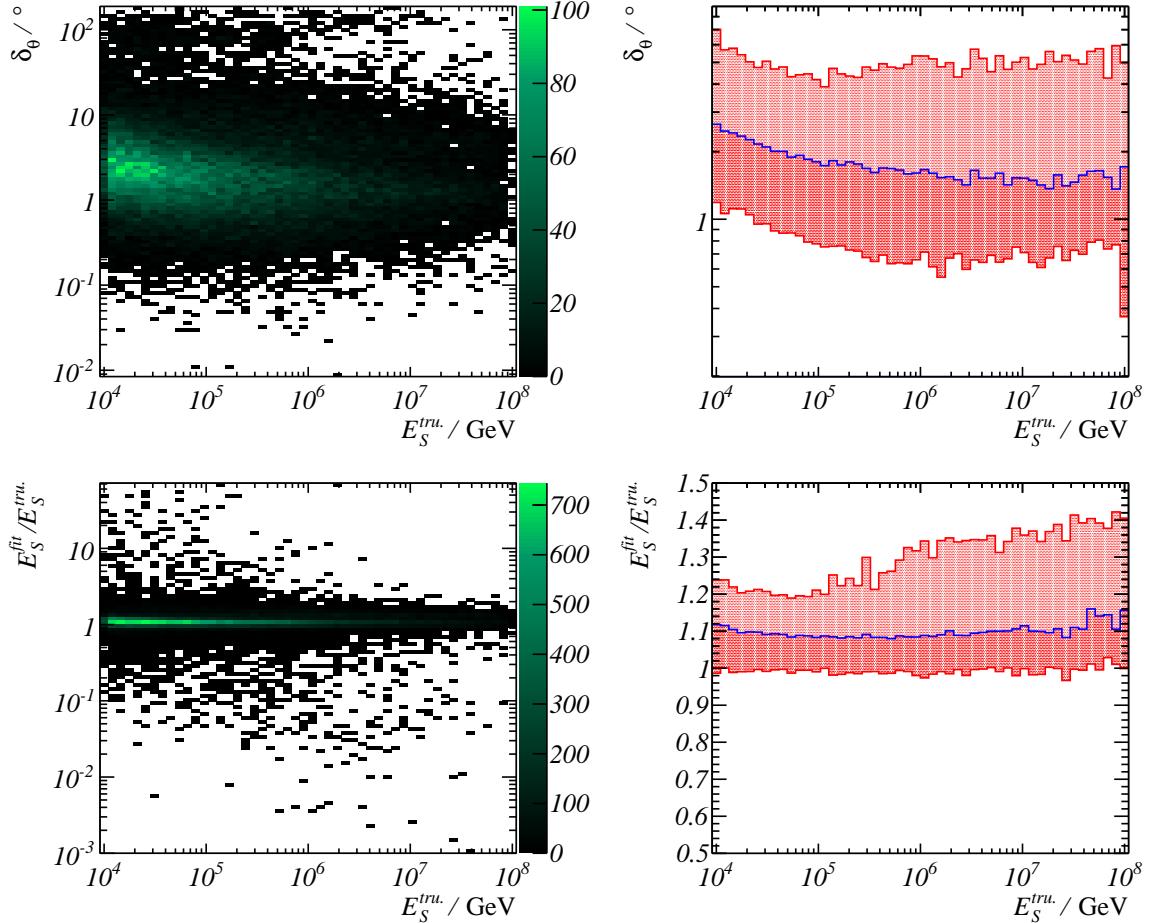


Figure 27: Angle between the reconstructed shower direction and the true shower direction ( $\delta_\theta$ ), and ratio between the fitted shower energy  $E_S^{fit}$  and the true shower energy  $E_S^{tru}$ . for all reconstructed events. The blue line resp. red lines in the right figures represent the median resp. 16 % and 84 % quantiles of the distribution in each energy-bin.

The results are in agreement with the results as obtained from the test sample, as presented in section 4.3.5.

The accuracy of the fitted direction increases with the shower energy. For some events, either a local maximum has been selected in the rough sampling, or the position is estimated with a large error. In figure 27, these events form the bulk of events at  $\delta_\theta \approx 90^\circ$ . Despite of the presence of these badly reconstructed events, the median angular error approaches approximately  $1.5^\circ$  at the highest shower energies. The median of the angular error is better than  $2^\circ$  at shower energies  $E_S^{tru.} \geq 4 \cdot 10^4$  GeV.

For a majority of the events, the fitted shower energy is approximately equal to the true shower energy. However, as can be seen clearly from the median of  $E_S^{fit}/E_S^{tru.}$ , the fitted shower energy systematically overestimates the true shower energy. The major cause of the overestimation comes from the bias introduced with the PMT selection.

By dividing the fitted energy by a function compensating for the systematic estimation,

the median of the fitted energy can be set at the true shower energy. If this method would be applied, then from the spread in the distribution of the fitted energies, the energy of the shower can be reconstructed with a resolution of approximately  $\pm 15\%$ .

## 5.2 Comparison with *km3*-simulated events

The reconstruction method can be used to give insights in differences between simulation methods. The *km3* simulation package is comparable to the used simulation method, but with different treatment of the emission and propagation of photons. The reconstruction method has been applied to a set of 6968 *km3*-simulated events. Of these, 3949 are reconstructed. In the following figures, the accuracy of the reconstructed parameters is given.

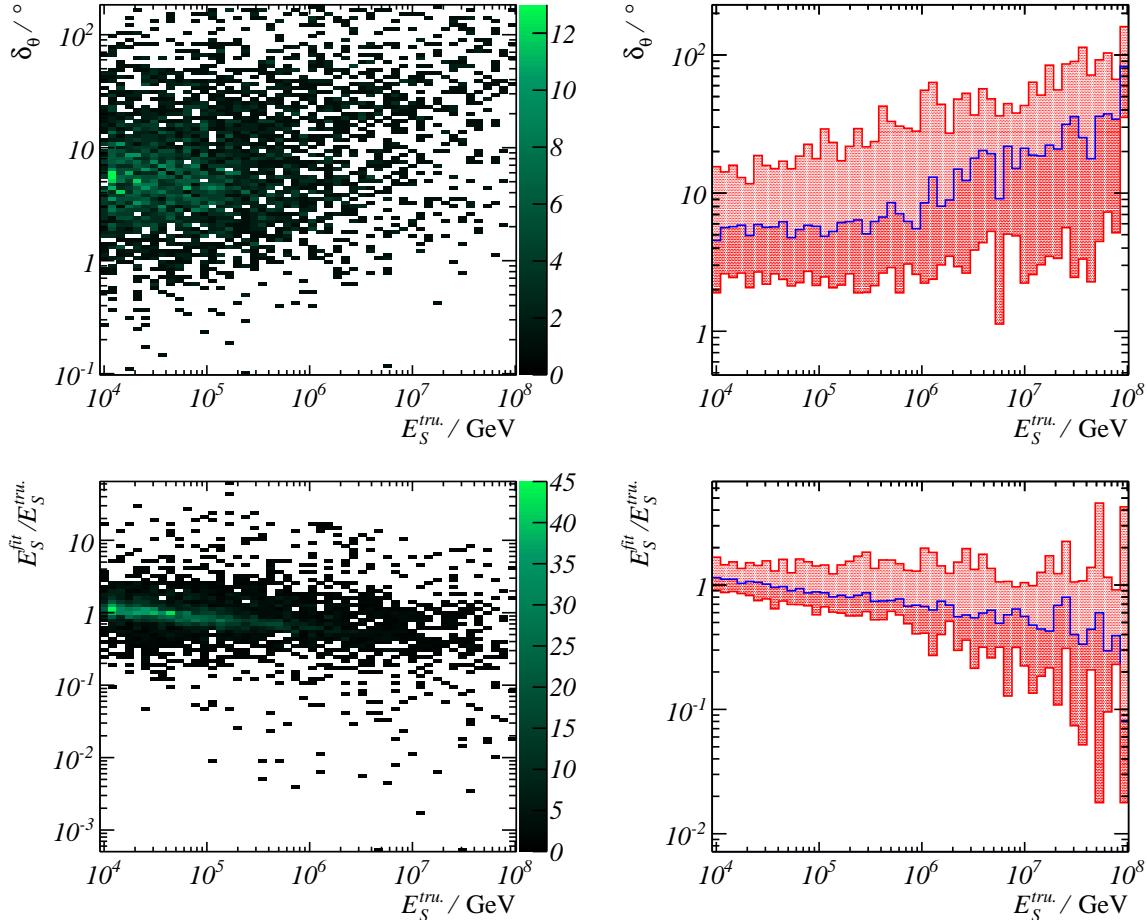


Figure 28: Angle between the reconstructed shower direction and the true shower direction ( $\delta_\theta$ ), and ratio between the fitted shower energy  $E_S^{fit}$  and the true shower energy  $E_S^{tru.}$  for all reconstructed events. The blue line resp. red lines in the right figures represent the median resp. 16 % and 84 % quantiles off the distribution in each energy-bin.

For true shower energies  $E_S^{tru.} < 3 \cdot 10^5$  GeV, the direction of the shower is reconstructed with a median error smaller than  $6^\circ$ . At higher energy, the accuracy of the estimated direction decreases. As can be seen, a systematic difference between the reconstructed shower energy and the true shower energy is present.

The accuracy of the estimated parameters of *km3*-simulated showers is significantly worse than the accuracy of the estimated parameters of showers simulated with simulation method described in section 3. This feature can be expected as the result of differences between the simulation methods, and from the fact that the reconstruction method is tuned for the used simulation method rather than for *km3*-simulated events.

These results show that a deeper understanding of the differences between the simulation methods should be achieved. Before applying the reconstruction method to real data, a more consistent simulation method based on both simulation methods should preferably be developed. After implementing minor changes to the reconstruction method, it can be expected that the reconstruction method will give results comparable with the results as presented in section 5.1.

### 5.3 Comparison with *aashowerfit*

The accuracy of the reconstruction method is compared to that of *aashowerfit*. The given performance of *aashowerfit* is based on events fulfilling a containment cut on the interaction vertex position. This simulation volume is given by  $\sqrt{x^2 + y^2} < 380$ m and  $|z| < 260$ m, where  $x, y, z$  represent the interaction vertex position relative to the center of the detector<sup>[26]</sup>. This relatively small volume guarantees the simulation of events which are expected to be well-reconstructable.

In order to compare the results of this analysis with those of *aashowerfit*, the same cuts are applied on the 'main' set of events. *Aashowerfit* compares the reconstructed shower direction and energy with the true neutrino direction and energy. Therefore the quantity  $\delta_\theta'$  is introduced to give the angle between the reconstructed shower direction and the true neutrino direction. In figure 29, the accuracy of both reconstruction methods on the contained events is plotted:

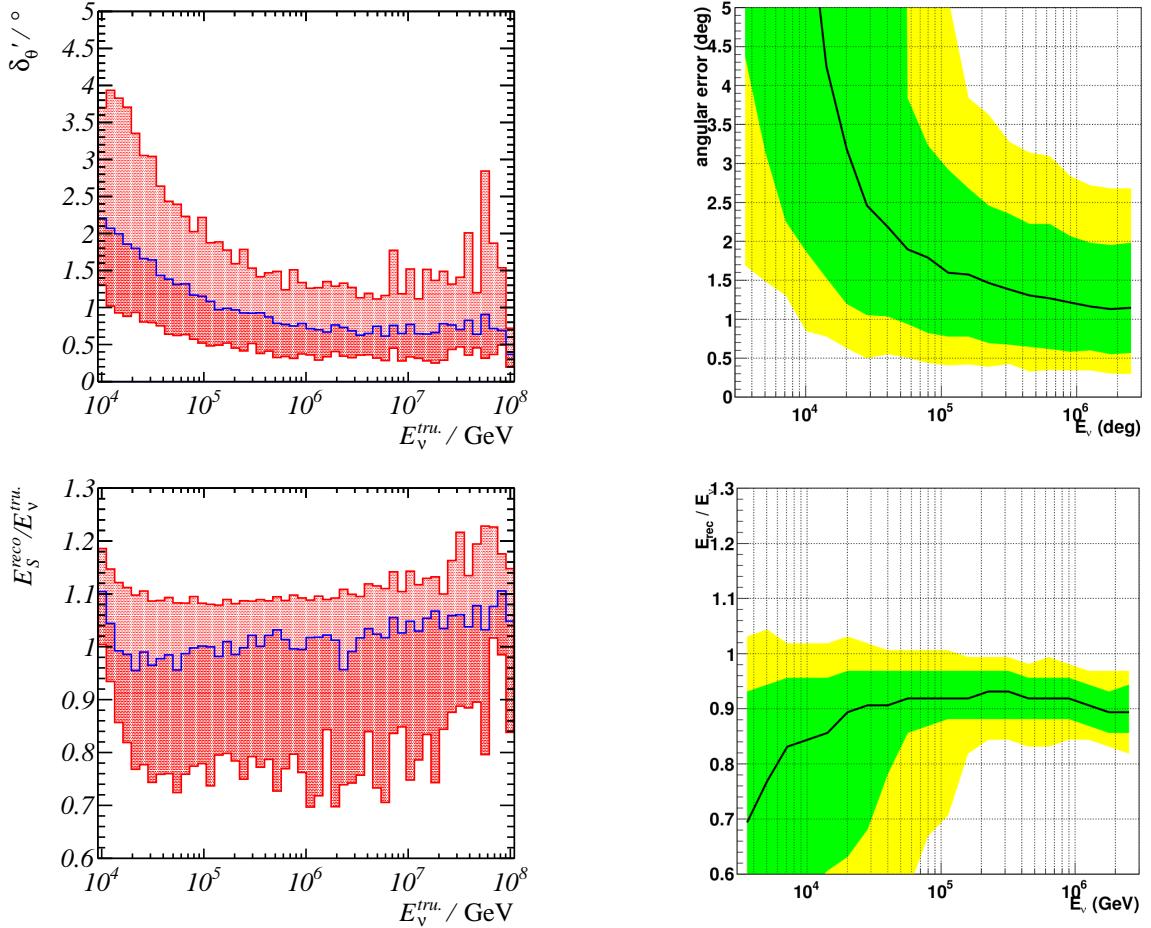


Figure 29: Angle between the reconstructed shower direction and the true neutrino direction ( $\delta'_\theta$ ), and ratio between the fitted shower energy  $E_S^{fit}$  and the true neutrino energy  $E_\nu^{tru.}$  for all reconstructed events. The figures left are obtained by the algorithm of this research. In these figures, the blue line resp. red lines represent the median resp. 16 % and 84 % quantiles off the distribution in each energy-bin. The right figures are taken from <sup>[26]</sup>. The green resp. yellow bands in these figures represent the 16 – 84 % resp. 5 – 95 % quantiles.

Especially in the reconstructed shower direction, the results of this analysis show a significantly better resolution than *aashowerfit*. For neutrino energies  $E_\nu^{tru.} \geq 10^5$  GeV, the direction of the neutrino using the reconstruction algorithm as developed in this research is estimated with subdegree resolution. Differences between the results can originate from the methods used to simulate the events and differences in the reconstruction methods. A major difference between the reconstruction methods is that *aashowerfit* only distinguishes between hit- and non-hit PMTs.

## 5.4 Event selection

The developed reconstruction method can be part of a more extensive analysis aiming at identifying neutrino sources. For such an analysis, it might be more efficient to select only events passing certain cuts on the reconstructed shower parameters. Although the number of selected events is then lower, the selected events are known to point to the source with a better accuracy.

The possible event selection criteria can be based on numerous (combinations of) measured and reconstructed quantities. In this chapter, the performance of several event selection criteria will be compared to each other.

### 5.4.1 Event selection criteria

In this section, five possible event selection criteria will be introduced, based on the estimated shower position, the estimated number of photon hits and the likelihood value.

**Distance to center of detector.** If a shower is contained within the instrumented volume of the detector, the shower is surrounded by PMTs. Due to the distinctive cone-shape at the Cherenkov angle of the emitted light, such an event can be expected to be well-reconstructable. On the other hand, if an event is positioned far from the instrumented volume, then only a few PMTs will detect the light and the cone-shape will produce a less distinctive hit pattern.

This leads to the event selection criterion  $R^{ev.} \leq R_{cut}^{ev.}$ , where  $R^{ev.}$  gives the distance from the estimated shower (maximum) position to the center of the detector.

**Cylindrical containment.** The instrumented volume of the KM3NeT detector can be approximated by a cylinder. Due to this shape, a cut on the distance to the center of the detector might not be the best cut on the estimated shower position.

A better event selection may be formed by selecting only events with estimated shower positions contained in a cylinder of radius  $\rho_{cut}$  and height  $z_{cut}$ .

**Estimated number of photo-electrons.** If a shower is high energetic and well-contained in the instrumented volume, then the number of photon hits is expected to be high. If this is the case, then also the error in the reconstructed direction can be expected to be low. The number of photon hits can thus provide an indication of the error in the reconstructed direction.

The number of photon hits is approximately equal to the (estimated) number of emitted photo-electrons. The corresponding event selection is formed by a lower limit on the total number of estimated effective photo-electrons  $\tilde{n}^{p.e.,tot}$ .

**Likelihood value** The probability that the estimated shower parameters can lead to the set of measured DAQ hits is proportional to the shower likelihood at these parameters.

The likelihood value thus gives an indication of the difference between the expected hits from the estimated shower parameters and the measured hits.

The shower likelihood function is given in equation 4.1. Since the probability to measure a certain set of DAQ hits on a PMT is always in the range zero to one, the likelihood value will decrease with the number of selected PMTs. A better measure for the event selection is given by the ratio between the logarithm of the likelihood value at the maximum ( $\log(L)$ ) and the number of selected PMTs ( $N^{PMT,sel.}$ ).

A high ratio between  $\log(L)$  and  $N^{PMT,sel.}$  can be the result of either a high likelihood value or a low number of selected PMTs. The likelihood value at the estimated shower parameters can thus be used in an event selection as  $l_{cut,1} < \log(L)/N^{PMT,sel.} \leq l_{cut,2}$ .

**Observed information** The quantity  $-d^2\log(L)/dS^2$  at the likelihood maximum is called the observed Fisher information. It gives a measure of the statistical error in the estimated shower direction.

The higher the observed information, the smaller the statistical error in the estimated shower direction. The event selection criterion  $-\log(L)'' > l_{cut}$  can thus be introduced.

#### 5.4.2 Error in estimated shower position

The reconstruction of the shower direction and energy takes an estimate of the shower position as input. The error in the reconstructed shower direction  $\delta_\theta$  can be related to the error in the estimated position  $\delta_R$ .

Suppose a measured or reconstructed quantity gives a perfect estimate of the error in the estimated position, then the hypothetical event selection criterion  $\delta_R < \delta_{R,cut}$  can be introduced. If the performance of this event selection is comparable to one of the 'real' event selection criteria, it will be assumed that the position fit performs well enough for the direction and energy fit. If, on the other hand, the performance of this hypothetical event selection is significantly better than other (real) event selection criteria, then methods to obtain a better estimate on the shower position should be investigated.

#### 5.4.3 Comparison

Each (set of) cut parameter(s) of a certain event selection criterion will result in a certain number of selected events  $N^{ev.,sel.}$ , with a median of the error  $\tilde{\delta}_\theta^{sel.}$  of the reconstructed direction of the selected events. For comparison of the presented criteria, the value(s) of the cut parameter(s) are not interesting.

The relation between  $\tilde{\delta}_\theta^{sel.}$  and  $N^{ev.,sel.}$  is plotted in figure 30 for all presented event selection criteria. In this figure, the median direction error  $\tilde{\delta}_\theta^{sel.}$  is calculated by weighting each event with a hypothetical atmospheric  $E_S^{-2}$  shower energy spectrum. For comparison, also the theoretically best possible event selection criterion ( $\delta_\theta < \delta_{cut}$ ) is plotted in gray. The event selection criteria with both a lower as an upper cut on a parameter can result in multiple values of  $\tilde{\delta}_\theta^{sel.}$  at a single  $N^{ev.,sel.}$  value.

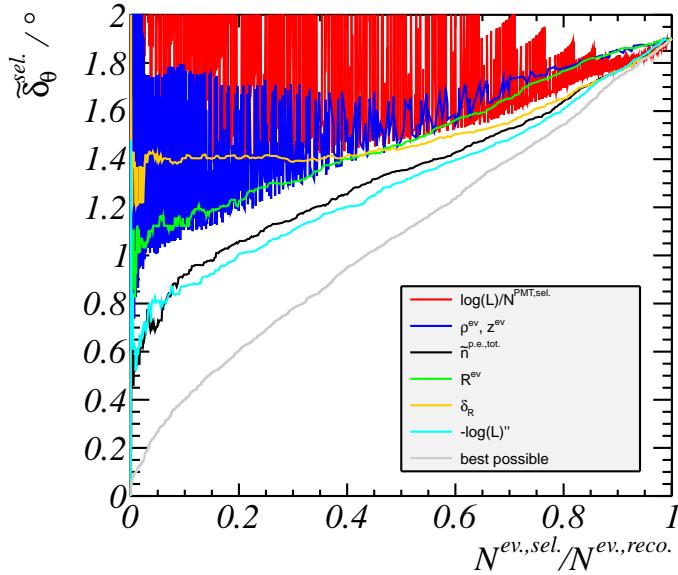


Figure 30: Performance of several event selection criteria. For the possible (sets of) cut parameter(s), the ratio of selected events is given on the horizontal axis. The median error in the estimated shower direction of the selected events is given on the vertical axis.

An event selection criterion should aim at selecting a high number of events  $N^{ev,sel.}$ , with a low median error  $\tilde{\delta}_\theta^{sel.}$  of the reconstructed direction of the selected events. The hypothetical event selection criterion on basis of  $\delta_R$  (in yellow) shows that, especially at small  $\delta_R$ , the error in the estimated direction is weakly correlated to the error in the estimated shower position. Other event selection criteria give a better measure of the error in the estimated shower direction. Based on this, the position fit can be assumed to perform well enough for the direction and energy fit. However, it should be noted that this doesn't imply that a better position fit will not lead to a better estimated shower direction.

Of all the presented criteria, the worst criterion for event selection is formed by the likelihood value (red). The event selection on basis of the observed Fisher information (cyan) gives the best event selection criterion. With the event selection criterion on basis of the observed Fisher information, subdegree median direction resolution of 18 % of all reconstructed events (weighted to the  $E_S^{-2}$  energy spectrum) can be achieved.

## 6 Discussion and conclusions

In this thesis, a new reconstruction method for high-energy showers in KM3NeT has been presented. In this section, the reconstruction method and results will be discussed, followed by conclusions on the performed study and obtained results.

### 6.1 Discussion

In this section, some features of the reconstruction method will be discussed. In most cases, possible solutions to the discussed issues will be given.

**Time over threshold conversion** In the reconstruction method, the number of emitted photo-electrons is estimated from the set of detected time-over-threshold values. The used relation between the number of emitted photo-electrons and the resulting time-over-threshold value is known to differ from measurements<sup>[27]</sup>. Comparing both relations, a slightly less accurate estimate can be expected in case the simulation and reconstruction method would be adapted to match the measured curve. This may also influence the performance of the reconstruction algorithm.

**Position fit** The reconstruction method starts with estimating the shower maximum position. Events are 'lost' due to the hit selection necessary for the position fit. The error in the estimated position is in general small enough not to influence the direction and energy fit significantly. However, in the rare occasion that the shower position is estimated with an error larger than 10 meters, the direction and energy fit can be distorted.

A different hit selection criterion for the position fit may increase the number of reconstructed events. A way to improve the accuracy of the position fit is to include the number of emitted photo-electrons. This amplitude information can be used not only to test a fitted position, but also to provide extra information via the position shower likelihood function.

An even better shower likelihood function would incorporate information on both the timing and the amplitude of both signal- and background photon hits. For maximizing this shower likelihood function, there would be no need for a hit selection.

**Direction and energy fit** The direction fit shows good accuracy. However, some events are reconstructed with large errors. This is explained by a poorly estimated shower position or from a local maximum selected by the 'rough' sampling.

A straightforward way to improve the direction reconstruction algorithm is by using more sampling points for the 'rough' sampling. This reduces the probability that a local maximum is selected, but the required number of sampling points will increase, at the cost of longer computation times.

The overall performance of the direction fit may be improved by including the timing of the detected hits in the used shower likelihood function. In this way, the shower likelihood

function becomes sensitive to the longitudinal distribution of emitted photons (several meters at high energies).

The systematical overestimation of the shower energy is the result of the bias introduced with the PMT selection and associating approximations on the shower likelihood function (equation 4.14). This bias can be removed by incorporating the probability that a small number of photon hits contributes to the number of estimated number of photo-electrons on a PMT (equation 7.9).

## 6.2 Conclusions

In this analysis, a reconstruction method for high-energy neutrino-induced particle showers with energies  $E_S > 10^4$  GeV has been developed. New methods and concepts have been developed, such as a method to estimate the number of emitted photo-electrons from the set of detected *ToT*-values on a PMT. Including this amplitude information in the shower likelihood function may explain the unparalleled performance.

The performance has been tested using simulated shower events. The direction of a shower can be reconstructed a median direction resolution of less than  $2^\circ$  and an energy resolution of  $\pm 15\%$ . By implementing event selection criteria, this resolution can be increased to a median direction resolution of less than  $1^\circ$ . This result makes the developed reconstruction algorithm the currently best performing one in KM3NeT.

The achieved direction and energy resolution indicates that (cosmic) neutrino sources can be identified by reconstructing shower-like events only. Combining this detection channel with the track-like detection channel will make the KM3NeT detector sensitive to all neutrino flavors. This can enhance the number of events in neutrino-source analyses by several factors.

The obtained median error in the estimated shower maximum position is in the order of one meter. This resolution could be sufficient to identify the individual showers contributing to the 'double-bang' signature after a tau-neutrino charged-current interaction. In this way, tau-neutrinos can be distinguished from electron-neutrinos. Methods exist to distinguish tracks from showers. Combining both principles can lead to an analysis in which also the ratio between the observed neutrino flavors helps to distinguish neutrino sources. With this, all characteristic properties that can be used to distinguish between different neutrino sources (section 1.2) can be probed.

## References

- [1] J. Learned and K. Mannheim, “High-energy neutrino astrophysics,” *Annu. Rev. Nucl. Part. Sci.*, vol. 50, pp. 679–749, 2000.
- [2] V. Hess, “Über beobachtungen der durchdringenden strahlung bei sieben freiballonfahrten,” *Phys. Z.*, vol. 13, pp. 1084–1091, 1912.
- [3] K. Nakamura *et al.*, “Review of particle physics,” *Journal of Physics G*, vol. 37, p. 075021, 2010.
- [4] J. Linsley, “Evidence for a primary cosmic-ray particle with energy  $10^{20}$  ev,” *Phys. Rev. Lett.*, vol. 10, pp. 146–148, 1963.
- [5] A. Hillas, “The origin of ultra-high-energy cosmic rays,” *Ann. Rev. Astron. Astrophys.*, vol. 22, pp. 425–444, 1984.
- [6] F. Fraschetti, “On the acceleration of ultra-high-energy cosmic rays,” *Phil. Trans. R. Soc. A*, vol. 366, pp. 4417–4428, 2008.
- [7] M. Bustamante *et al.*, “High-energy cosmic-ray acceleration,” *Proceedings of the 5th CERN-Latin American School of High-Energy Physics*, 2009.
- [8] D. Palioselitis, “Measurement of the atmospheric neutrino spectrum,” 2012. PhD thesis, Universiteit van Amsterdam.
- [9] B. Peterson and A. Wandel, “Evidence for supermassive black holes in active galactic nuclei from emission-line reverberation,” *The Astrophysical Journal*, vol. 540, no. 1, pp. L13–L16, 2000.
- [10] R. Protheroe and A. Szabo, “High energy cosmic rays from active galactic nuclei,” *Phys. Rev. Lett.*, vol. 69, no. 20, pp. 2885–2888, 1992.
- [11] K. Koyama *et al.*, “Evidence for shock acceleration of high-energy electrons in the supernova remnant sn1006,” *Nature*, vol. 378, pp. 255–258, 1995.
- [12] A. Hillas, “Can diffusive shock acceleration in supernova remnants account for high-energy galactic cosmic rays?,” *Journal of Physics G*, vol. 31, no. 5, p. R95, 2005.
- [13] R. Klebesadel *et al.*, “Observations of gamma-ray bursts of cosmic origin,” *Astr. Journ.*, vol. 182, pp. L85–L88, 1973.
- [14] T. Piran, “High energy neutrinos from cosmological gamma-ray burst fireballs,” *Phys. Rev. Lett.*, vol. 78, no. 12, pp. 2292–2295, 1997.
- [15] T. Piran, “Gamma-ray bursts and the fireball model,” *Phys. Rept.*, vol. 314, pp. 575–667, 1999.

- [16] M. Honda *et al.*, “Calculation of the flux of atmospheric neutrinos,” *Phys. Rev.*, vol. D52, p. 4985, 1995.
- [17] R. Enberg *et al.*, “Prompt neutrino fluxes from atmospheric charm,” *Phys. Rev.*, vol. D78, p. 043005, 2008.
- [18] M. Markov and I. Zheleznykh, “On high energy neutrino physics in cosmic rays,” *Nucl. Phys.*, vol. 27, pp. 385–394, 1961.
- [19] M. Aartsen *et al.*, “Observation of high-energy astrophysical neutrinos in three years of icecube data,” *Phys. Rev. Lett.*, vol. 113, p. 101101, 2014.
- [20] S. Eidelman *et al.*, “Passage of particles through matter,” *Physics Letters*, vol. B, no. 592, pp. 1–30, 2004.
- [21] C. Kopper, “Performance Studies for the KM3NeT Neutrino Telescope,” 2010. Doctoral thesis, Friedrich-Alexander-Universität, Erlangen-Nürnberg.
- [22] J. Reubelt, “Osci screenshots of base signals,” 2013. Presentation, ANTARES-KM3NeT Collaboration Meeting, Vilanova.
- [23] S. Agostinelli *et al.*, “Geant4a simulation toolkit,” *Nuclear Instruments and Methods in Physics Research*, vol. 506, no. 3, pp. 250 – 303, 2003.
- [24] R. Mirani, “Parametrisation of em-showers in the antares detector-volume,” 2002. Doctoral thesis, Universiteit van Amsterdam.
- [25] J. Berger and R. Wolpert, *The Likelihood Principle*. Hayward, California: Institute of Mathematical Statistics, 1984.
- [26] A. Heijboer, “Shower direction reconstruction: aashowerfit,” 2014. ANTARES-KM3NeT internal note.
- [27] J. Reubelt, “Pmt - base calibration,” 2014. Presentation, ANTARES-KM3NeT Collaboration Meeting, Leiden.

## 7 Appendices

### Appendix A: Maximization of the point source likelihood function

In this appendix, an analytical function for the parameters  $\vec{r}_{PS}, t_{PS}$  of a point source of light will be derived that maximize the likelihood function:

$$L_{(\vec{r}_{PS}, t_{PS})}^{PS} \propto \prod_{i=1}^{N^{DAQ,sel.}} P_{(\vec{r}_i, t_i | \vec{r}_{PS}, t_{PS})}$$

$$P_{(\vec{r}_i, t_i | \vec{r}_{PS}, t_{PS})} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{((\vec{r}_i - \vec{r}_{PS})^2 - \bar{v}^2(t_i - t_{PS})^2)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x_i^2}{2\sigma^2}}$$

$$x_i \equiv (\vec{r}_i - \vec{r}_{PS})^2 - \bar{v}^2(t_i - t_{PS})^2 \quad (7.1)$$

Where  $\vec{r}_i$  corresponds to the position of a selected hit at time  $t_i$ , and  $\bar{v}$  is the mean speed of the light.

Since  $x_i$  is normally distributed,  $x_{i,j} \equiv x_i - x_j$  is also normally distributed (with mean 0 and standard deviation  $\sqrt{2}\sigma$ ). Maximizing  $L_{(\vec{r}_{PS}, t_{PS})}^{PS}$  is then equivalent to maximizing:

$$L_{(\vec{r}_{PS}, t_{PS})}^{PS'} \equiv \prod_{i=1}^{N^{DAQ,sel.}} \prod_{j \neq i} e^{-\frac{x_{i,j}^2}{4\sigma^2}} \quad (7.2)$$

$$= e^{-\frac{(\sum_{i=1}^{N^{DAQ,sel.}} \sum_{j \neq i} x_{i,j}^2)}{4\sigma^2}}$$

Or, equivalently, to minimizing:

$$-\log(L_{(\vec{r}_{PS}, t_{PS})}^{PS'}) = \sum_{i=1}^{N^{DAQ,sel.}} \sum_{j \neq i} x_{i,j}^2 \quad (7.3)$$

As will be shown next, the minimum of this log-likelihood function can be found straightforward. It is convenient to write  $x_{i,j}$  as:

$$\begin{aligned} x_{i,j} &\equiv x_i - x_j \\ &= (\vec{r}_i - \vec{r}_{PS})^2 - \bar{v}^2 \cdot (t_i - t_{PS})^2 - (\vec{r}_j - \vec{r}_{PS})^2 + \bar{v}^2 \cdot (t_j - t_{PS})^2 \\ &= (\vec{r}_i^2 - \vec{r}_j^2) - \bar{v}^2(t_i^2 - t_j^2) - 2(\vec{r}_i - \vec{r}_j) \cdot \vec{r}_{PS} + 2\bar{v}^2(t_i - t_j) \cdot t_{PS} \quad (7.4) \\ &= (\vec{r}_i^2 - \vec{r}_j^2) - \bar{v}^2(t_i^2 - t_j^2) - 2 \cdot \begin{bmatrix} (\vec{r}_i - \vec{r}_j) & -\bar{v}^2(t_i - t_j) \end{bmatrix} \cdot \begin{bmatrix} \vec{r}_{PS} \\ t_{PS} \end{bmatrix} \end{aligned}$$

If  $\mathbf{r}_P \equiv \begin{bmatrix} \vec{r}_{PS} \\ t_{PS} \end{bmatrix}$  and the following vectors and matrix are introduced:

$$\mathbf{x} \equiv \begin{bmatrix} x_{1,2} \\ x_{1,3} \\ \vdots \end{bmatrix}, \quad \mathbf{b} \equiv \begin{bmatrix} (\vec{r}_1^2 - \vec{r}_2^2) - \bar{v}^2(t_1^2 - t_2^2) \\ (\vec{r}_1^2 - \vec{r}_3^2) - \bar{v}^2(t_1^2 - t_3^2) \\ \vdots \end{bmatrix}, \quad \mathbf{A} \equiv \begin{bmatrix} 2 \cdot (\vec{r}_1 - \vec{r}_2) & -2\bar{v}^2(t_1 - t_2) \\ 2 \cdot (\vec{r}_1 - \vec{r}_3) & -2\bar{v}^2(t_1 - t_3) \\ \vdots & \vdots \end{bmatrix} \quad (7.5)$$

then equation 7.3 can be written as:

$$\begin{aligned} -\log(L_{(\vec{r}_{PS}, t_{PS})}^{PS'}) &= \sum_{i=1}^{N^{DAQ,sel.}} \sum_{j \neq i} {x_{i,j}}^2 \\ &= \mathbf{x}' \mathbf{x} \\ &= (\mathbf{b} - \mathbf{A} \cdot \mathbf{r}_P)' (\mathbf{b} - \mathbf{A} \cdot \mathbf{r}_P) \end{aligned} \quad (7.6)$$

Where  $\mathbf{X}'$  denotes the transpose of a matrix  $\mathbf{X}$ . Setting the derivative of  $-\log(L_{(\vec{r}_{PS}, t_{PS})}^{PS'})$  with respect to  $\mathbf{r}_P$  to zero gives an analytical function for the best estimate  $(\tilde{\mathbf{r}}_P)$  of  $\vec{r}_P$  and  $t_{PS}$ :

$$\begin{aligned} 0 &= \frac{d \log(L^{PS'})}{d \mathbf{r}_P} \Big|_{\tilde{\mathbf{r}}_P} \\ 0 &= \frac{d (\mathbf{b} - \mathbf{A} \cdot \mathbf{r}_P)' (\mathbf{b} - \mathbf{A} \cdot \mathbf{r}_P)}{d \mathbf{r}_P} \Big|_{\tilde{\mathbf{r}}_P} \\ 0 &= \frac{d (\mathbf{b}' \mathbf{b} - 2\mathbf{b}' \mathbf{A} \mathbf{r}_P + \mathbf{r}_P' (\mathbf{A}' \mathbf{A}) \mathbf{r}_P)}{d \mathbf{r}_P} \Big|_{\tilde{\mathbf{r}}_P} \\ 0 &= \mathbf{b}' \mathbf{A} - (\mathbf{A}' \mathbf{A}) \tilde{\mathbf{r}}_P \\ \tilde{\mathbf{r}}_P &= (\mathbf{A}' \mathbf{A})^{-1} \mathbf{A}' \mathbf{b} \end{aligned} \quad (7.7)$$

## Appendix B: Best estimate of the shower energy

In this appendix, it will be shown that the shower energy that maximizes the no-timing shower likelihood  $L_{(S)}^{no,t}$  is given by:

$$\frac{\tilde{E}_S}{1 \text{ GeV}} = \frac{\sum_{sel.PMT} [\tilde{n}_{(ToT_i^{DAQ})}^{p.e.,tot.}]}{\sum_{sel.PMT} [\hat{n}_{i(S,E_S=1 \text{ GeV})}^{\gamma,tot.}]} \quad (7.8)$$

The no-timing shower likelihood function is derived in section 4.3 and is given by:

$$\begin{aligned}
L_{(S|\vec{\mathbf{T}} \cdot \mathbf{T}^{DAQ})}^{no\ t} &\propto \prod_{sel.PMT} P_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ}|S)}^{no\ t} \\
P_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ}|S)}^{no\ t} &= \sum_{n_i^{\gamma,tot.}=0}^{\infty} P_{(\tilde{n}_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ})}^{p.e.,tot.}|n_i^{\gamma,tot.})} \cdot P_{(n_i^{\gamma,tot.}|\hat{n}_{i(S)}^{\gamma,tot.})} \\
&\approx \mathcal{P}_{(\tilde{n}_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ})}^{p.e.,tot.}|\hat{n}_{i(S)}^{\gamma,tot.})} \\
\mathcal{P}_{(x|\mu)} &\equiv \frac{\mu^x \cdot e^{-\mu}}{x!}
\end{aligned} \tag{7.9}$$

The best estimate on the shower energy is given at the point where  $\frac{dL_{(S)}^{no\ t}}{dE_S} = 0$ . The derivative of the shower likelihood function is given by:

$$\begin{aligned}
\frac{d}{dE_S} L_{(S)}^{no\ t} &= \frac{d}{dE_S} \prod_{sel.PMT} \left[ P_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ}|S)}^{no\ t} \right] \\
&= \frac{d}{dE_S} \prod_{sel.PMT} \left[ \mathcal{P}_{(\tilde{n}_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ})}^{p.e.,tot.}|\hat{n}_{i(S)}^{\gamma,tot.})} \right] \\
&= \prod_{sel.PMT} \left[ \mathcal{P}_{(\tilde{n}_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ})}^{p.e.,tot.}|\hat{n}_{i(S)}^{\gamma,tot.})} \right] \cdot \sum \left[ \frac{1}{\mathcal{P}_{(\tilde{n}_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ})}^{p.e.,tot.}|\hat{n}_{i(S)}^{\gamma,tot.})}} \cdot \frac{d\mathcal{P}_{(\tilde{n}_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ})}^{p.e.,tot.}|\hat{n}_{i(S)}^{\gamma,tot.})}}{dE_S} \right] \\
&= L_{(S)}^{no\ t} \cdot \sum_{sel.PMT} \left[ \frac{1}{\mathcal{P}_{(\tilde{n}_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ})}^{p.e.,tot.}|\hat{n}_{i(S)}^{\gamma,tot.})}} \cdot \frac{d\mathcal{P}_{(\tilde{n}_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ})}^{p.e.,tot.}|\hat{n}_{i(S)}^{\gamma,tot.})}}{d\hat{n}_{i(S)}^{\gamma,tot.}} \cdot \frac{d\hat{n}_{i(S)}^{\gamma,tot.}}{dE_S} \right] \\
&= L_{(S)}^{no\ t} \cdot \sum_{sel.PMT} \left[ \left( \frac{\tilde{n}_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ})}^{p.e.,tot.}}{\hat{n}_{i(S)}^{\gamma,tot.}} - 1 \right) \cdot \frac{d\hat{n}_{i(S)}^{\gamma,tot.}}{dE_S} \right] \\
&= L_{(S)}^{no\ t} \cdot \sum_{sel.PMT} \left[ \frac{\tilde{n}_{(\mathbf{T} \cdot \mathbf{T}_i^{DAQ})}^{p.e.,tot.}}{\hat{n}_{i(S)}^{\gamma,tot.}} \cdot \frac{d\hat{n}_{i(S)}^{\gamma,tot.}}{dE_S} \right] - L_{(S)}^{no\ t} \cdot \sum_{sel.PMT} \left[ \frac{d\hat{n}_{i(S)}^{\gamma,tot.}}{dE_S} \right]
\end{aligned} \tag{7.10}$$

If it is assumed that the expected number of photons hitting a PMT  $i$  is proportional to the energy of the shower, then:

$$\begin{aligned}
\hat{n}_{i(S)}^{\gamma,tot.} &= \hat{n}_{i(S,E_S=1 \text{ GeV})}^{\gamma,tot.} \cdot E_S \\
\frac{d}{dE_S} L_{(S)}^{no,t} &= L_{(S)}^{no,t} \cdot \sum_{sel.PMT} \left[ \frac{\tilde{n}_{(\mathbf{T}o\mathbf{T}_i^{DAQ})}^{p.e.,tot.}}{\hat{n}_{i(S)}^{\gamma,tot.}} \cdot \hat{n}_{i(S,E_S=1 \text{ GeV})}^{\gamma,tot.} \right] - L_{(S)}^{no,t} \cdot \sum_{sel.PMT} \left[ \hat{n}_{i(S,E_S=1 \text{ GeV})}^{\gamma,tot.} \right] \\
\frac{d}{dE_S} L_{(S)}^{no,t} &= L_{(S)}^{no,t} \cdot \sum_{sel.PMT} \left[ \frac{\tilde{n}_{(\mathbf{T}o\mathbf{T}_i^{DAQ})}^{p.e.,tot.}}{E_S/1 \text{ GeV}} \right] - L_{(S)}^{no,t} \cdot \sum_{sel.PMT} \left[ \hat{n}_{i(S,E_S=1 \text{ GeV})}^{\gamma,tot.} \right]
\end{aligned} \tag{7.11}$$

Where  $\hat{n}_{i(S,E_S=1 \text{ GeV})}^{\gamma,tot.}$  is the estimated number of photon hits on PMT  $i$  from a shower of 1 GeV. Setting the derivative of the shower likelihood function to zero gives:

$$\begin{aligned}
0 &= \frac{d}{dE_S} L_{(S)}^{no,t} \Big|_{\tilde{E}_S} \\
0 &= \sum_{sel.PMT} \left[ \frac{\tilde{n}_{(\mathbf{T}o\mathbf{T}_i^{DAQ})}^{p.e.,tot.}}{\tilde{E}_S/1 \text{ GeV}} \right] - \sum_{sel.PMT} \left[ \hat{n}_{i(S,E_S=1 \text{ GeV})}^{\gamma,tot.} \right] \\
\frac{\tilde{E}_S}{1 \text{ GeV}} &= \frac{\sum_{sel.PMT} \left[ \tilde{n}_{(\mathbf{T}o\mathbf{T}_i^{DAQ})}^{p.e.,tot.} \right]}{\sum_{sel.PMT} \left[ \hat{n}_{i(S,E_S=1 \text{ GeV})}^{\gamma,tot.} \right]}
\end{aligned} \tag{7.12}$$

## Appendix C: Isotropic hemisphere sampling

In this appendix, a method for sampling the full hemisphere with a minimal number of sampling directions  $\vec{d}^{sam.}$  will be introduced.

Since the angles between the directions along the 30 edges of a dodecahedron (with 12 pentagonal faces) is constant, these directions form a convenient starting point.

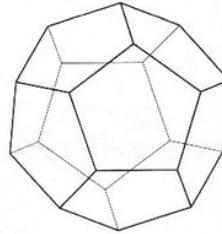


Figure 31: A dodecahedron

Adding the 12 directions pointing to the center of each face to the set of sampling directions is the second step in covering the full hemisphere. In this way, the 42 sampling directions form 60 equal triangles.

After the first two steps, extra sampling directions are added in the following way:

1. Division of each triangle in four equally shaped triangles.
2. All nodes of the sub-triangles give new directions.

By repeating these steps, an approximate isotropic sampling is achieved. As the final step, all sampling directions are normalized to one.

The total number of sampling directions ( $N^{sam.}$ ) depends on the number of times the subdivision is repeated ( $N^{steps}$ ). In table 1,  $N^{sam.}$  is given for some values of  $N^{steps}$ :

Table 1: Number of isotropic sampling directions

$N^{steps}$	$N^{sam.}$
0	30
1	42
2	132
3	492
4	1932
5	7692

