# Studies on atmospheric neutrino oscillations with the ANTARES neutrino telescope

Studien zur Oszillation atmosphärischer Neutrinos
mit dem ANTARES Neutrinoteleskop

Der Naturwissenschaftlichen Fakultät
der Friedrich-Alexander-Universität Erlangen-Nürnberg
zur Erlangung des Doktorgrades Dr. rer. nat.

vorgelegt von
Friederike Schöck
aus Karlsruhe

Als Dissertation genehmigt von der Naturwissenschaftlichen Fakultät der Friedrich-Alexander-Universität Erlangen-Nürnberg

Tag der mündlichen Prüfung: 14. Februar 2011

| | |
|---|---|
| Vorsitzender der Promotionkommission: | Prof. Dr. Rainer Fink |
| Erstberichterstatter: | Prof. Dr. Uli Katz |
| Zweitberichterstatter: | Prof. Dr. Christian Stegmann |

# Contents

# List of Figures

# List of Tables

# 1 Introduction

*"Neutrino physics is largely an art of learning a great deal by observing nothing."*
Haim Harari

The night sky with its bright stars and the universe itself has been fascinating mankind since the beginning of time. Astronomy, which deals with the study of celestial objects as well as formation and development of the universe, is one of the oldest natural sciences. It is primarily based on optical observations, first with the naked eye, later on with instruments that became more and more advanced. Over the years, especially during the last decades, the observed photon energy spectrum was extended beyond the visible range towards higher (X-ray astronomy) and lower (radio astronomy) photon energies.

In 1911 though, when Victor Hess discovered the cosmic radiation (Nobel Prize 1936), a completely new window to our universe was opened up. Cosmic rays are messenger particles that convey precious information from outer space and can answer open astrophysical questions. They allow for a complementary insight to our universe, in addition to the conventional observation of electromagnetic radiation.

Cosmic radiation consists of charged particles that are permanently impinging on the Earth's atmosphere, where they interact with atmospheric gas molecules. New particles are generated in such interactions, which again interact and generate further particles. A cascade (so-called shower) develops, which is of hadronic or electromagnetic nature, depending on the primary particle type. This secondary or atmospheric radiation can be detected with dedicated instruments, and conclusions can be drawn about the original (primary) cosmic particles. Besides, it can also be employed for studying topics of particle physics, like neutrino flavour oscillations.

The energy spectrum of the primary cosmic radiation ranges over several orders of magnitude and reaches beyond $E = 10^{20}\,\mathrm{eV}$, which exceeds the energy that can be obtained at the currently most powerful particle accelerators on Earth by several orders of magnitude[1]. The flux is decreasing with a broken power law and ranges from about one particle per second per square meter at around $10^{12}\,\mathrm{eV}$ to less than one particle per century per square kilometer at about $10^{20}\,\mathrm{eV}$.

Even though a lot is already known about these cosmic messengers, there are fundamental questions yet to be answered, like their origin and their acceleration mechanisms. As charged particles are deflected in galactic and intergalactic magnetic fields,

[1] The LHC for example, is eventually expected to reach an energy of $E = 14\,\mathrm{TeV}$ in the centre of mass for proton-proton collisions.

they reach the Earth isotropically, not revealing their origin. Only at highest energies the deflection may be small enough for the particles to point back to their cosmic sources. It is generally assumed that e.g. active galactic nuclei are excellent source candidates for high-energy cosmic rays, but there is no definite observational proof yet.

Besides charged particles, also high-energy photons ($E > 100\,\mathrm{keV}$) and neutrinos are part of the cosmic radiation[2]. While $\gamma$-rays can be produced by different mechanisms in leptonic (e.g. synchrotron radiation) or hadronic processes (e.g. pion decay), neutrinos are exclusively generated in hadronic scenarios. The detection of high-energy neutrinos would directly imply the existence of relativistic hadrons. Furthermore, neutrinos are not deflected in magnetic fields and due to their very small interaction cross section, they propagate virtually undisturbedly through the universe. These characteristics render the neutrino an ideal messenger particle. Though one has to pay a certain price: their small interaction probability creates enormous difficulties to detect neutrinos on Earth and huge detectors are necessary to measure at least a few of them.

The IceCube experiment [1], located at the South Pole, is currently the largest experiment for the detection of these elusive cosmic messengers, with an instrumented volume of about one cubic kilometer. It employs the deep clear ice at the South Pole as a detection medium to measure high-energy neutrinos. This is possible via their charged interaction products that are emitting Cherenkov light when traversing through the ice. The light can propagate up to several 100 m and finally be recorded by photosensors. Typically, the Earth is employed as shielding against atmospheric charged particles (mainly muons) and therefore the IceCube field of view covers only the northern hemisphere[3].

On the basis of exactly the same detection principle, the ANTARES collaboration has designed and constructed a neutrino telescope in the depth of the Mediterranean Sea, employing sea water instead of ice as a detection medium. The aim of ANTARES, with an instrumented volume of about $0.01\,\mathrm{km}^3$, is the detection of cosmic neutrinos. Its field of view covers the complete Southern hemisphere, including the Galactic Centre. Like IceCube, ANTARES is designed for the detection of muon neutrinos that generate muons when they interact via charged current interactions. Due to their mass, muons with energies above some GeV can travel tens to several thousand meters in water. These long trajectories allow for a good directional reconstruction, which is essential for doing astronomy. Despite the fact that ANTARES is the largest water-based neutrino telescope worldwide, it is expected to be too small for the discovery of cosmic neutrinos. Therefore, a future multi-cubic-kilometer detector, KM3NeT [2], is currently in the planning phase.

The main hindrance in identifying cosmic signals for all cosmic neutrino detectors is the irreducible background of atmospheric neutrinos emerging from hadronic showers generated in the Earth's atmosphere by primary cosmic rays. Only by detecting an excess of neutrinos from a certain celestial direction or at particular high energy over

[2] In some literature, the term "cosmic radiation" only refers to the charged component.

[3] Lately analyses also include data from the southern hemisphere, some however at neutrino energies beyond $10^{15}\,\mathrm{eV}$.

the atmospheric background, a cosmic origin can be proven. On the other hand, even though the detected atmospheric neutrinos are only background for searches for cosmic neutrinos, they also provide great possibilities for studying particle physics aspects, in particular the phenomenon of neutrino oscillations.

Neutrino oscillations are a quantum mechanical effect, that was first proposed in the late 1950s and early 1960s [3, 4, 5, 6]. Definite experimental evidence for neutrino oscillations was first found by the Super-Kamiokande experiment [7, 8], employing a large water tank for measuring both solar and atmospheric neutrinos, also by making use of the Cherenkov effect. In short, the effect of neutrino oscillations is that neutrinos change their flavour identity while travelling from their generation to their detection point. For ANTARES, which is mainly designed for the detection of muon neutrinos, the consequence is a reduced flux of atmospheric muon neutrinos in an energy range of about 10 – 100 GeV. This is right at the low-energy sensitivity limit of ANTARES, which is given by the layout of the detector.

At such energies, neutrino oscillations have not yet been probed, and doing an oscillation analysis with the ANTARES data therefore allows for the investigation of neutrino oscillations in a new energy range. On the other hand, as the results of many different experiments are consistent with the theoretical description of neutrino oscillations, such an analysis can be employed to test the understanding of the ANTARES detector in the low-energy regime.

I have performed such a study of oscillations by means of atmospheric neutrinos detected with ANTARES. My thesis deals with this challenging analysis in the low-energy range of a neutrino telescope designed for high-energy neutrinos. For the analysis, I have recovered an old reconstruction program and have implemented it into the current ANTARES software framework. I have adjusted and improved the reconstruction algorithm at several levels, for example by including a new probability density function, using new hit selections and implementing a new fit procedure.

The construction of ANTARES, which is the first fully functional deep-sea neutrino telescope, was successfully completed in June 2008. However, data taking started already at the beginning of 2007 with the parts of the detector. The work presented in this thesis was done during the commissioning phase of the detector, and it must be seen in this context, that parts of the thesis focus on rather technical issues.

No official data production was available when I was working on that thesis and I had to process data and simulations by myself. This offered the possibility for testing an alternative concept concerning data selection and the simulation of experimental detection conditions. I have developed a concept for the selection of data, based on an event-by-event evaluation of environmental and detector conditions. It also incorporates the adjustment of simulations to the time varying data taking conditions. Due to several technical problems, for example with very slow access to the data base and a slow reconstruction speed, the processing itself took several months (February to July 2010). All developments and progress made by the ANTARES collaboration during that time could not be considered for the analysis anymore.

In Chapter 2 of this thesis, a brief overview of neutrino physics is given. The role of the neutrino within the Standard Model of particle physics is highlighted, different sources of neutrinos are discussed, as well as how neutrinos interact and how they can be detected. It closes with a short summary of neutrino oscillation experiments.

Chapter 3 is dedicated to the ANTARES experiment. The detector is described in detail, including the detection principle, the experimental setup, and the data acquisition. The ambient conditions are illustrated, including the optical background. This is followed by the event triggering and a discussion of signal and background events. The simulation processing chain in ANTARES is described and the software framework *SeaTray* is introduced, to which I made several contributions.

In Chapter 4, a detailed description and evaluation of the updated and improved low-energy reconstruction algorithm is given.

The details of the approach used for data evaluation and selection, as well as for simulating realistic detection conditions, and the respective software developments are discussed in Chapter 5. Furthermore, the processing and reconstruction chain of data and simulations is illustrated, and comparisons between data and simulation are shown in examples. Finally, quality selection cuts on the reconstructed events are discussed and the results are shown.

In Chapter 6, the relevant parameter distributions of simulation and data are discussed, and the oscillation analysis is introduced. The systematics and the conclusions that can be drawn from the analysis are also addressed within this chapter.

A summary of this thesis is given on pp. 155 – 157.

# 2 Neutrino physics

As the neutrino is the leading actor in this thesis, it is introduced in this chapter. Its role within the Standard Model of particle physics and beyond is illuminated in Section 2.1. A short introduction to the theoretical concepts of neutrino oscillations is given in Section 2.2. In Section 2.3 the main sources of neutrinos are presented and in Section 2.4 neutrino interactions and possible detection mechanisms for neutrinos are discussed. Finally, in Section 2.5, an overview of the experimental status of neutrino oscillation physics is given.

## 2.1 Neutrinos within the Standard Model and beyond

The Standard Model (SM) of particle physics is a basic theory of the elementary particles of matter and three of the four known fundamental interactions: electromagnetic, weak, and strong. All three interactions are explained by particles with spin 1 (bosons) that are exchanged between the interacting elementary particles. The probability amplitude of an interaction is determined by the transition matrix element. It incorporates the coupling strength of the interaction, which is given by the respective coupling constant, and a propagator term, which depends on the momentum transfer of the interaction and the mass of the exchanged boson.

There are twelve known elementary particles (and their twelve anti-partners), all of them being spin-1/2 particles (fermions). They are divided into two groups of six, the quarks and the leptons (see Fig. 2.1). Quarks (up $u$, down $d$, charm $c$, strange $s$, top $t$ and bottom $b$ quark) are massive particles with a charge of either minus one third or two thirds of the elementary charge $e$ and they are subject to all of the three interactions of the SM. They do not exist as isolated free particles but only in bound states.

Leptons comprise two types of particles: three massive and electrically charged particles (electron $e^-$, muon $\mu^-$ and tau $\tau^-$) and three corresponding electrically neutral neutrinos $\nu_e, \nu_\mu$, and $\nu_\tau$, that are very light (and are regarded massless within the SM). Each pair of a massive lepton and its corresponding neutrino partner are representing a so-called family or generation. Except for their mass, the characteristics and the interactions of the three generations are identical. Contrary to quarks, leptons do not couple to gluons, the force carriers of the strong interaction. Neutrinos, due to being electrically neutral, are only subject to the weak interaction. The force carriers of the weak interaction are the massive $W^+$, $W^-$ and $Z^0$ bosons. Both the $W^\pm$ and the $Z^0$ are heavy particles with masses of around 80 GeV and 91 GeV, respectively. As

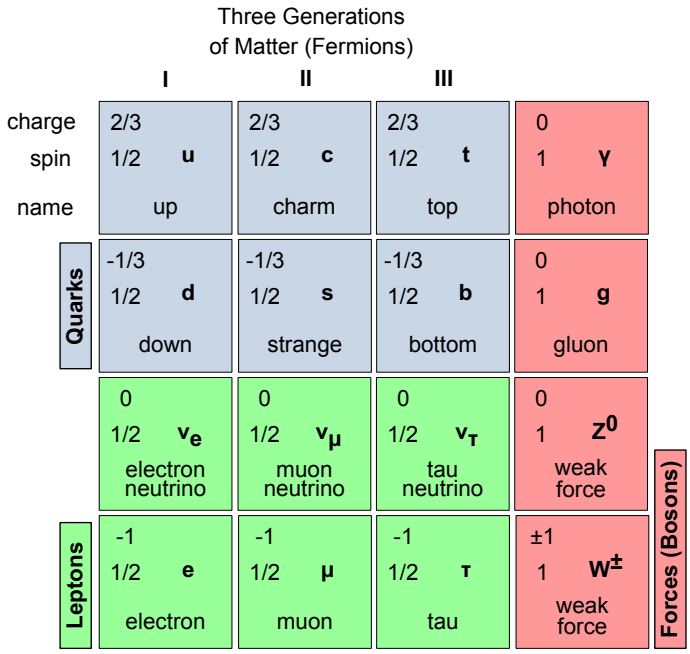



**Figure 2.1:** Schematic view of the Standard Model of particle physics. See text for details.

a consequence, the weak interaction is short-ranged and the interaction probability is extremely small compared to the electromagnetic or strong interaction[4].

This exclusive combination of attributes, being electrically neutral, massless and interacting strictly via the weak interaction, ensures a special role for the neutrino within the SM. Postulated by Wolfgang Pauli already in 1930 to explain the kinematics of the radioactive $\beta$-decay, neutrinos were first detected in 1956 by Cowan and Reines [10].

In the same year, 1956, parity violation of the weak interaction was observed for the first time [11]. Soon thereafter, it was found that all neutrinos are left handed [12], whereas all anti-neutrinos are right handed, implying that parity is maximally violated for these particles.

Measurements of the decay width of the $Z^0$ boson at LEP in 1989 showed that there are exactly three light-mass neutrinos [13]. As each of the three neutrinos has its corresponding electrically charged and massive lepton partner, this also indicates that there are only three lepton generations. The $\nu_\mu$ was detected for the first time in 1962 [14], six years after the first measurement of $\nu_e$. In 2000 finally, the observation of $\nu_\tau$ was reported [15, 16] and thus all three types of neutrinos have been directly detected.

In summary, the SM describes almost all observations made in the field of elementary particle physics amazingly well, and even various predictions basing on the SM could

[4]The exact value depends for example on the momentum transfer. Typically, an order of $10^{-13}$ for the weak interaction is given, while the strong interaction is of the order of 10 and the electromagnetic interaction is of the order of $10^{-2}$ [9].

experimentally be verified later on with extreme precision. However, though being remarkably predictive, it does not give a perfect description of all observable phenomena. It provides a good approach to the fundamental understanding of elementary particles and interactions and accordingly also to neutrino physics, but it has to be extended to include, for example, neutrino oscillations.

A deficit in the solar neutrino flux was already observed in 1968 by Davis et al., who performed measurements of the solar neutrino flux with with a radiochemical detector, the Homestake experiment [17, 18]. That this measured deficit was due to the oscillations of neutrino flavours (and not for example due to flaws in the solar models) could only be proven later on. The first definite experimental evidence for neutrino oscillations has been reported by the Japanese Super-Kamiokande group [8] in 1998. As explained in detail in the next section, neutrino oscillations require both neutrino flavour mixing and massive neutrinos. Neither of it is part of the SM theory. Neutrino oscillation experiments are not sensitive to the neutrino mass values themselves, but only to differences between squared neutrino masses. Information on neutrino masses can instead be obtained by three different methods, which are cosmological observations, search for neutrinoless double beta decay, and a direct determination of the neutrino mass by kinematics. Data from $\beta$-decay measurements [19] yield the most recent upper limit on the $\nu_e$ mass of about 2 eV [20].

Naturally, the question arises how neutrinos get their mass. The simplest assumption is that neutrinos (like charged fermions) are Dirac particles and couple to the Higgs boson. Dirac neutrinos would exist in four states (both left and right handed neutrinos and anti-neutrinos), but without having mass only left handed neutrinos and right handed anti-neutrinos are realised in nature. Mathematically this is explained by the Dirac mass term in the Lagrangian which mixes the left and right hand states of the fermions. Having only a single definite handedness like in the SM, the Dirac term vanishes and neutrinos become massless. On the other hand, if neutrinos are massive, right handed neutrinos and left handed anti-neutrinos must exist.

An alternative explanation for massive neutrinos is the Majorana scenario, where the neutrino is its own anti-particle and only one single particle with its two handedness states would exist for each lepton generation. This is theoretically possible because neutrinos have no electric charge. In the Lagrangian, a set of Majorana mass terms appears and, by invoking so-called "seesaw" models, small observable neutrino masses are motivated (see for example [20, 21, 22] and references therein). A direct consequence of the Majorana See-Saw Model is at least one heavy neutrino, which is essentially right handed. It is not yet observed because it is far too massive ($M_{\mathrm{R}} \leq 10^{15}$ GeV, [22]) . If neutrinos are Majorana particles, this would also impact on the theoretical description of neutrino flavour mixing and result in two additional CP violating phases (see Section 2.2). In any case, an extention of the SM is necessary and only if neutrinos are massive, the question about the Dirac- or Majorana-nature of neutrinos can eventually be answered.

## 2.2 Neutrino oscillations

The phenomenon of neutrino oscillations requires that neutrinos have mass. This is already beyond the Standard Model theory, as explained above. Furthermore neutrino mixing is required, which means that the observed weak eigenstates $|\nu_\alpha\rangle$ of the neutrinos that take part in weak interactions, are not eigenstates of mass. Instead they are linear combinations of three definite mass eigenstates $|\nu_i\rangle$. One eigenstate system is given in terms of the other via the unitary matrix $U_{\alpha i}$:

$$\text{Flavour eigenstates:} \qquad |\nu_\alpha\rangle, \qquad \alpha = e, \mu, \tau \tag{2.1}$$

$$\text{Mass eigenstates:} \qquad |\nu_i\rangle, \qquad i = e, \mu, \tau \tag{2.2}$$

$$|\nu_\alpha\rangle = \sum_i U_{\alpha i} |\nu_i\rangle, \qquad |\nu_i\rangle = \sum_\alpha U^*_{\alpha i} |\nu_\alpha\rangle. \tag{2.3}$$

The unitary transformation matrix $U_{\alpha i}$ is called MNS-matrix (according to Maki, Nakagawa, and Sakata, who first described flavour mixing in 1962 [5]) and is analogous to the CKM-matrix in the quark sector that describes the mixing between quark flavours. The MNS-matrix is commonly parametrised as

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \cdot \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \cdot \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} (\cdot V) \tag{2.4}$$

with $s_{ij} = \sin\Theta_{ij}$ and $c_{ij} = \cos\Theta_{ij}$. It depends on the three rotation or mixing angles $\Theta_{ij}$ and the complex CP violating phase $\delta$. The term $V = \mathrm{diag}(1, e^{i\Phi_1}, e^{i(\Phi_2+\delta)})$ contains two further CP violating phases $\Phi_1$ and $\Phi_2$ that arise as a direct consequence of a possible Majorana nature of neutrinos. If neutrinos are Dirac particles, $V$ equals the identity matrix. In any case, this term does not affect the calculation of oscillation probabilities and can be neglected for our purposes [21, 23].

Neutrinos are exclusively generated in weak interactions as pure flavour eigenstates and thus as a mixture of mass eigenstates. They are likewise detected via charged particles produced in weak interactions (Section 2.4), which again means measuring the flavour eigenstates of the neutrinos. The propagation through space time and therefore the time evolution of a neutrino is, in contrast to production and detection, described by its mass eigenstates. The time evolution of a general quantum state $|\nu_i\rangle$ is generated by the Hamiltonian $H_0$:

$$|\nu_i(t)\rangle = e^{-iH_0 t}|\nu_i\rangle = e^{-iE_i t}|\nu_i\rangle. \tag{2.5}$$

Provided that the mass eigenstates have different masses, they also have different energies and their associated waves will have different frequencies. Superposition of the waves leads to interference effects and hence the flavour composition changes with time. The quantum mechanical probability to measure one specific flavour eigenstate therefore depends on the energy of the neutrino, on the difference of the frequencies of the

mass eigenstates (i.e. on the difference in mass) and on the life time of the neutrino or, respectively, the distance between the points of production and detection (i.e. the path length of the neutrino). A pure flavour eigenstate $|\nu_\alpha\rangle$ at time $t = 0$ evolves as:

$$|\nu_\alpha(t)\rangle = \sum_i U_{\alpha i}|\nu_i(t)\rangle = \sum_i U_{\alpha i}e^{-iE_i t}|\nu_i\rangle = \sum_{i,\beta} U_{\alpha i}U^*_{\beta i}e^{-iE_i t}|\nu_\beta\rangle \tag{2.6}$$

The quantum mechanical probability of finding a flavour state $|\nu_\beta\rangle$ of a neutrino with energy $E$ after time $t$, that was produced in eigenstate $|\nu_\alpha\rangle$ is given by

$$P(\nu_\alpha \to \nu_\beta) = |\langle\nu_\beta|\nu_\alpha(t)\rangle|^2. \tag{2.7}$$

Assuming CP conservation (phase $\delta = 0$; current neutrino oscillation data have no sensitivity to CP violation) and replacing the time of flight $t$ by the path length $L$, the resulting probability is

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 4 \cdot \sum_{j>i} U_{\alpha i}U^*_{\alpha j}U^*_{\beta i}U_{\beta j} \sin^2 \frac{\Delta m^2_{ij} L}{4E}, \tag{2.8}$$

with the differences of the neutrino masses squared $\Delta m^2_{ij} = m^2_i - m^2_j$ (for the derivation of the formula see for example [24])[5]. The amplitudes of the individual terms in the oscillation probability depend on the entries $U_{\alpha i}$ of the MNS-matrix and therefore on the mixing angles $\Theta_{ij}$, whereas the frequency of the oscillation is given by $\Delta m^2_{ij}$ and the energy $E$ of the neutrino.

Altogether, there are six free parameters to be determined: the three mixing angles, the CP violating phase $\delta$ and two independent $\Delta m^2_{ij}$ parameters, $\Delta m^2_{12}$ and $\Delta m^2_{23}$; $\Delta m^2_{13}$ is given by:

$$\Delta m^2_{12} + \Delta m^2_{23} + \Delta m^2_{13} = 0. \tag{2.9}$$

The ordering of the masses, the so called mass hierarchy, is still unknown. The eigenstates can either follow the "normal" hierarchy ($m_1 < m_2 < m_3$) or the "inverted" hierarchy ($m_3 < m_1 < m_2$), for which $\Delta m^2_{23}$ changes its sign. Neither the absolute values of the masses nor the two Majorana CP violating phases $\Phi_1$ and $\Phi_2$ can be determined in oscillation experiments because the oscillation probability does not depend on these parameters. On the other hand, the three-flavour mixing scenario depends on the mass ordering being normal or inverted, although in subleading order due to the size of $\Theta_{13}$ (see below).

The values for $\Theta_{12}$ and $\Theta_{23}$, as well as for $\Delta m^2_{12}$ and $\Delta m^2_{23}$, are well known from several measurements (Section 2.5). Only few measurements exist for the third mixing angle $\Theta_{13}$ but they hint to a small value compatible with zero. Though the CP violating phase $\delta$ is in principal accessible through oscillations, the fact that it only appears in combination with $\sin\Theta_{13}$ in the MNS-matrix makes it difficult to extract its value.

The three-flavour mixing scenario can be simplified using the results of various experiments. Given that $\Delta m^2_{12} \ll \Delta m^2_{23}$ and in the limit of $\Theta_{13} \approx 0$, the oscillations

[5] This is for ultrarelativistic neutrinos. Due to the small neutrino mass this is already a good approximation for energies at some 10 eV.

$\nu_e \leftrightarrow \nu_\mu$ and $\nu_\mu \leftrightarrow \nu_\tau$ can be considered independently. Furthermore, the effect of the CP violating phase $\delta$ is negligible. In this approximation, the survival probability $P(\nu_\mu \to \nu_\mu)$ for a muon neutrino simplifies to

$$P(\nu_\mu \to \nu_\mu) = 1 - \sin^2(2\Theta_{23}) \cdot \sin^2\left(\frac{1.27 \cdot \Delta m^2_{23}[\mathrm{eV}^2] \cdot L[\mathrm{km}]}{E[\mathrm{GeV}]}\right), \tag{2.10}$$

with $L$ denoting the distance from the source and $E$ the neutrino energy. According to (2.10), the oscillation amplitude depends on the mixing angle $\Theta_{23}$ and the oscillation wavelength on $E$ and $\Delta m^2_{23}$.

A description of current neutrino oscillation experiments, their implications and the best fit values to the five oscillation parameters will be given in Section 2.5.

## 2.3 Neutrino sources

Huge amounts of neutrinos are produced in weak interaction processes and in various sources. Depending on the type of source, the neutrino energies are ranging from a few MeV to extremely high energies of the order of $10^{20}$ eV. Different neutrino sources are discussed in the following, paying special attention to atmospheric neutrinos, which are of relevance for this thesis.

### 2.3.1 Cosmic neutrinos

Cosmic neutrinos naturally appear in the context of primary cosmic rays, which consist of particles from outer space that are permanently impinging on our Earth's atmosphere. These particles are covering a huge energy range and bear precious information about high-energy processes occurring within our galaxy and beyond. The most interesting questions yet to answer are where the particles (especially those with highest energies) are coming from and how they are accelerated.

Cosmic rays consist of charged particles, which are mainly protons but to a small extend also nuclei with higher charge $Z$, as well as electrons. Their energy spectrum has been well measured by many different experiments. It ranges over more than ten orders of magnitude in energy up to extremely high values of the order of $10^{20}$ eV (Fig. 2.2). The flux decreases steeply over thirty orders of magnitude, following a broken power law

$$\frac{dN(E)}{dE} \propto E^{-\gamma}, \tag{2.11}$$

with the energy dependent spectral index $\gamma$, which varies between $\gamma \approx 2.7$ and $\gamma \approx 3.0$. The cause for the distinct features in the spectrum is still subject of discussions, involving different acceleration mechanisms, the composition of the cosmic rays and the transition from galactic to extra-galactic origin.

The information that can be gained from investigating charged cosmic particles is restricted by the fact that they are deflected in interstellar and intergalactic magnetic fields and thus lose their directional information. Only at highest energies the particles

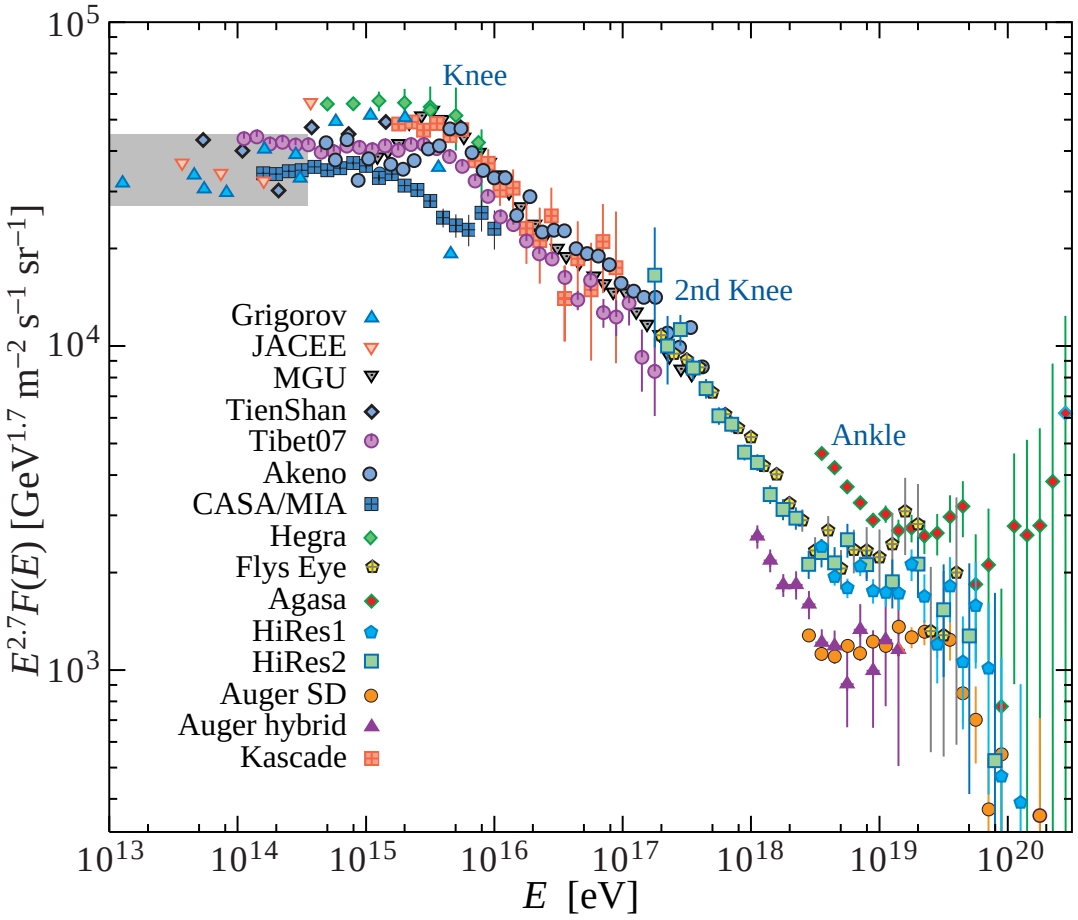



**Figure 2.2:** Cosmic-ray spectrum, multiplied by a factor $E^{2.7}$. See text for further explanation. Figure from [25]. The discrepancy between the different experiments at highest energies ($E \gtrsim 10^{18}\,\mathrm{eV}$) is mainly an issue of the energy calibration of these experiments.

may point back to their sources, because the deflection is small enough to allow for a reasonable directional reconstruction. Furthermore, above an energy of about $6 \cdot 10^{19}\,\mathrm{eV}$ protons are subject to interactions with the cosmic microwave background, which restricts their mean free pathlength to some 10 Mpc [26].

In any source in which charged cosmic particles are accelerated, neutrinos and $\gamma$-photons must be generated in secondary processes. Gamma radiation may, neutrinos will emerge from these sources and can reach the Earth. Investigating them can help to clearly identify hadronic cosmic accelerators.

Cosmic $\gamma$-radiation is being investigated for the last few decades. Especially two imaging air shower Cherenkov telescopes, H.E.S.S. [27] and MAGIC [28], discovered high-energy photon sources. For example, the H.E.S.S. experiment has reported 58

galactic, TeV $\gamma$-photon emitting objects [29] up to now.

Besides the above mentioned hadronic acceleration mechanism, such high-energy photons can also be accelerated in leptonic scenarios via electromagnetic processes. A clear distinction between leptonic and hadronic scenarios from the measured spectra is very challenging and in fact not possible without doubt.

Contrary to this, the detection of cosmic neutrinos would provide a clear indication of the hadronic scenario. Furthermore, due to being electrically neutral and having a tiny interaction cross section (Section 2.4.1), neutrinos can emerge even from the central parts of galactic and extra-galactic objects and travel virtually undisturbedly through the space. They provide valuable information and investigating them will complement our picture of the universe. However, the advantages they have in contrast to other cosmic messengers go along with enormous difficulties in detecting them on Earth. Huge detectors are needed to be able to measure only few of them. Indeed, no significant signal of high-energy cosmic neutrinos has been detected up to now[6]. Good candidates for sources of cosmic high-energy neutrinos are for example active galactic nuclei or gamma-ray bursts.

### 2.3.2 Atmospheric neutrinos

Atmospheric neutrinos are produced uniformly in the Earth's atmosphere by primary cosmic rays interacting with atmospheric particles. They form an irreducible background for the search for cosmic neutrinos, but can also be employed for investigating neutrino oscillations. Studying oscillations of atmospheric neutrinos with the ANTARES neutrino telescope is the topic of this thesis and therefore atmospheric neutrinos will be treated a bit more detailed in this section.

Besides atmospheric neutrinos, muons ($\approx 80\,\%$ of all charged particles) are the main component of the so-called secondary cosmic radiation. Other components, i.e. mesons (pions, kaons, etc.) and baryons (like protons) can also survive down to sea level, however their fraction is small. Electrons and positrons build up a third charged component but as they are absorbed soon after their production, their flux at sea level is small compared to the muon flux (except for very high-energy cosmic particles with $E \geq 10^{16}\,\mathrm{eV}$) [26].

Detecting neutrinos is challenging in general, and so is the detection of atmospheric neutrinos. The measurement of the atmospheric neutrino energy spectrum is still under way, theoretical predictions are continuously probed and models for its calculation are permanently improved.

Two components contribute to the total atmospheric neutrino flux $\Phi_\nu$, the conventional flux $\Phi_\nu^c$ and the prompt flux $\Phi_\nu^p$. The prompt flux of neutrinos results from semi-leptonic decays of particles containing charm quarks. Such particles are increasingly produced in the hadronic component of the secondary cosmic rays with increasing energy of the primary particle. The impact on the neutrino spectrum is very small for

[6] Cosmic low-energy neutrinos ($E = 20-40\,\mathrm{MeV}$) from the supernova explosion SN1987A were observed with the IMB detector [30].

energies below a threshold around $E_\nu^p \approx 10^5\,\mathrm{GeV}$. The energy range above this threshold is not of interest for the investigation of atmospheric neutrino oscillations and therefore the prompt flux will not be further discussed here.

In the energy range below the threshold, where the impact of the prompt flux becomes negligible, neutrinos are predominantly produced in decays of pions and kaons and of the resulting muons:

$$\pi^-/K^- \to \mu^- + \overline{\nu}_\mu \qquad \text{and} \qquad \pi^+/K^+ \to \mu^+ + \nu_\mu, \tag{2.12}$$

$$\mu^- \to e^- + \overline{\nu}_e + \nu_\mu \qquad \text{and} \qquad \mu^+ \to e^+ + \nu_e + \overline{\nu}_\mu. \tag{2.13}$$

The flux of neutrinos resulting from these decays is called the conventional neutrino flux. The expected flux ratio of muon to electron neutrinos $(\nu_\mu + \bar{\nu}_\mu)/(\nu_e + \bar{\nu}_e)$ and the ratio of neutrinos to anti-neutrinos $\nu/\bar{\nu}$ for electron and muon neutrinos as functions of $E_\nu$ are shown in Fig. 2.3. The flavour ratio has a value of about two around 1 GeV and increases up to about seven at 100 GeV, due to muons reaching the ground before decaying. In fact, this calculated ratio is reduced because of the effect of neutrino oscillations, where $\nu_\mu$ oscillate into $\nu_\tau$.

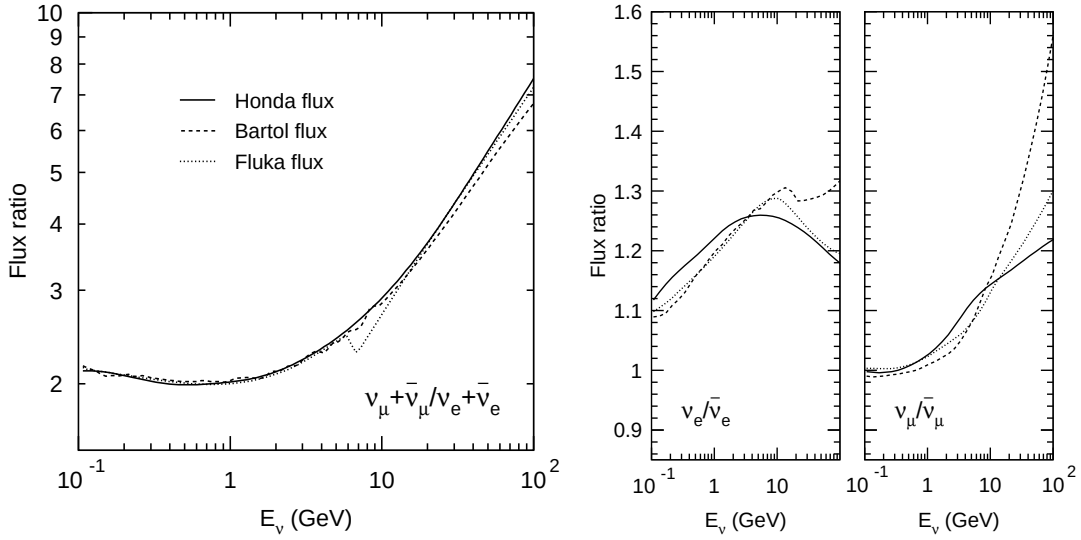



**Figure 2.3:** *Left:* The flux ratio $(\nu_\mu + \bar{\nu}_\mu)/(\nu_e + \bar{\nu}_e)$ averaged over zenith and azimuth versus neutrino energy for different flux models.
*Right:* The flux ratios of $\nu_e/\bar{\nu}_e$ and $\nu_\mu/\bar{\nu}_\mu$ versus neutrino energy for different flux models (same key as left figure). Figures from [31].

If all pions and kaons produced in the atmosphere decay, the energy spectrum of the atmospheric neutrinos theoretically reflects the spectrum of the primary cosmic rays. Indeed, the neutrino flux can be parametrised by a power law ($\frac{dN(E_\nu)}{dE_\nu} \propto E_\nu^{-\gamma}$), and around 10 GeV the spectral index approximately relates to the spectral index of the low-energy primary cosmic rays of about $\gamma \approx 2.7$. At about 100 GeV though, the

spectrum steepens to a value of about $\gamma \approx 3.7$. This relates to an increasing number of hadronic interaction processes of pions and kaons with atmospheric particles. With increasing energy, they are beginning to dominate over the decay processes and therefore less neutrinos are produced in this energy range.

Furthermore, variations in the atmospheric neutrino flux with the zenith angle $\Theta$ (angle between the particle's trajectory and the normal of the Earth's surface) are observed. Given the same track length, vertical particles are crossing more matter than horizontal ones and thus have more opportunities to interact. This results in a vertical flux that is reduced by a factor of about two[7] compared to the horizontal flux in the energy range from about 1 GeV to 1 PeV. The angle-averaged atmospheric muon neutrino and anti-neutrino spectrum is shown in Figure 2.4.

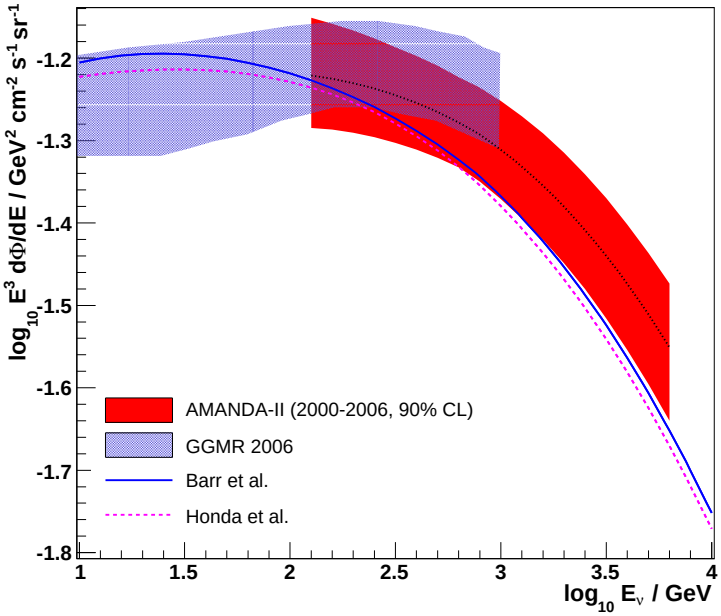



**Figure 2.4:** Angle averaged atmospheric $\nu_\mu + \overline{\nu}_\mu$ flux, multiplied by $E^3$. The red solid band characterises AMANDA-II data, the blue one Super-Kamiokande data. The black dashed line indicates the central best-fit curve. Also shown are predictions by Honda et al. and Barr et al. Picture from [32] (see also for references).

### 2.3.3 Other sources

#### Geological and artificial sources

Neutrinos, besides electrons, are a product of the nuclear $\beta$-decay. Radioactive $\beta$-emitting isotopes occur everywhere in nature, mainly in rocks, hydrogenetic rocks and

[7] The exact value depends on the flux models.

sea water. Mainly $\bar{\nu}_e$ in the MeV range are emitted by $\beta$-decay processes. By far, geological sources contribute most to the amount of terrestrial neutrinos.

Besides natural radioactivity, artificial radioactivity in reactors or particle accelerators is a further terrestrial source of neutrinos. While in reactors only the radioactive $\beta$-decay contributes a nearly pure flux of anti-electron neutrinos in the range of a few MeV, all three neutrino flavours can be produced in particle accelerators up to the higher GeV range.

#### Solar neutrinos

From the standard solar model (SSM) that describes the fusion processes in the sun, we can expect a pure electron neutrino flux up to 14 MeV. The main fusion process in the sun is the pp-cycle which contributes more than 98 % to the total energy release. In each fusion step $\nu_e$ are produced. Contrary to photons, neutrinos are not absorbed by the solar matter and can thus convey information from the inner part of the sun. The energy integrated flux predicted by the SSM is $\Phi_{\text{sol}} \geq 10^{11}\text{cm}^{-2}\text{s}^{-1}$.

## 2.4 Neutrino interactions and detection

Investigating neutrinos in general helps to complete our picture of the fundamental particles and forces in nature. In particular, neutrinos might be a clue to the physics beyond the Standard Model. Many things are already known about neutrinos, but there are still open questions, like the exact mass values, the reason why the neutrino masses are so tiny in relation to other particles, and the phenomena of oscillations and CP violation. Besides the role of neutrinos in the context of the fundamental understanding of matter and interactions, neutrinos also play a role as messenger particles. From detecting cosmic neutrinos, new findings about high-energy phenomena in our universe are expected. Accelerator, reactor, solar and atmospheric neutrinos on the other hand can be employed for the study of neutrino oscillations and CP violation. As neutrinos themselves leave no traces in detectors, they can only be detected by their interaction products. Hence, in this section, fundamental neutrino interactions are described, and how neutrinos can be detected.

The ANTARES neutrino telescope is employing the Cherenkov effect for an optical detection of high-energy neutrinos. With respect to the focus of this thesis, this detection method is stressed and the Cherenkov effect is introduced. An alternative detection technique for neutrinos of highest energies will also be mentioned briefly.

### 2.4.1 Neutrino interactions

The interaction cross section is proportional to the square of the transition matrix element, which in turn is proportional to the inverse of the mass of the force carrier

(i.e. $Z^0$, $W^+$ or $W^-$ bosons) squared:

$$\sigma \propto \mathfrak{M}_{fi}^2 \qquad \text{and} \qquad \mathfrak{M}_{fi} \propto \frac{g^2}{(M^2 - q^2)}, \tag{2.14}$$

with the cross section $\sigma$, the transition matrix element $\mathfrak{M}_{fi}$, the weak charge $g$ (which is of the order of the electrical charge $e$), the momentum transfer $q^2$ and the mass of the force carrier, $M$. Due to the large masses of the bosons, the cross section is small for small values of the momentum transfer. Only if the momentum transfer is very high (at least of the order of the mass of the exchanged boson), the cross section becomes comparable to the cross section of the electromagnetic interaction.

### Charged and neutral currents

The exchange of an electrically neutral $Z^0$ boson is called a neutral current (NC) interaction, while the exchange of an electrically charged $W^\pm$ boson is referred to as charged current (CC) interaction.

The CC interaction can change the flavour of quarks, create a charged lepton/anti-neutrino pair, or convert a charged lepton into a neutrino of the same family or vice versa (the same holds for the anti-particles). In this type of interaction, the identity of all participating particles changes, as well as their electrical charges. Depending on the particles involved, the interactions are categorised as leptonic, semi-leptonic or non-leptonic processes. An example for a leptonic process is the muon decay $\mu^- \rightarrow e^- + \overline{\nu}_e + \nu_\mu$ (accordingly for anti-muons). The pion and kaon decays are examples for semi-leptonic processes: $\pi^-/K^- \rightarrow \mu^- + \overline{\nu}_{mu}$ (accordingly for $\pi^+/K^+$). Only quarks are involved in non-leptonic processes. As they are not of relevance for what follows, they will not be further discussed here.

NC interactions, on the other hand, do not change the flavour or the types of the involved particles, implying that the in- and outgoing particles are identical. An example for a NC reaction is the neutrino scattering from quarks: $\nu_\mu + q \rightarrow \nu_\mu + q$.

### Deep inelastic scattering

When neutrinos are passing through baryonic matter, they can interact with the target nuclei in semi-leptonic processes, either via CC or NC interactions. In the energy range of interest for large-volume neutrino detectors like ANTARES (above 10 GeV), the cross section for such neutrino-nucleon interactions is predominantly deep inelastic scattering[8] [33]. The measure for the inelasticity of such a process is the Bjorken scaling variable $x$, which is defined as:

$$x := \frac{Q^2}{2M(E_\nu - E_l)}, \tag{2.15}$$

[8] Further reaction channels are quasi-elastic and resonant scattering, which will not be discussed here further, but are included in the simulations.

with the mass of the nucleus $M$, the energy of the incident neutrino $E_\nu$ and the outgoing lepton $E_l$ and the negative squared four-momentum transfer between the neutrino and the lepton, $Q^2 = -(p - p')^2$ [34]. In an elastic interaction the Bjorken variable will be $x \equiv 1$, whereas for an inelastic process $0 < x < 1$. The energy difference $E_\nu - E_l$ is converted into a hadronic cascade, originating from the interaction vertex. The cascade mainly consists of pions, kaons, protons, and neutrons. All possible semi-leptonic neutrino-nucleon deep inelastic interaction channels including the hadronic cascades are depicted in Fig. 2.5.

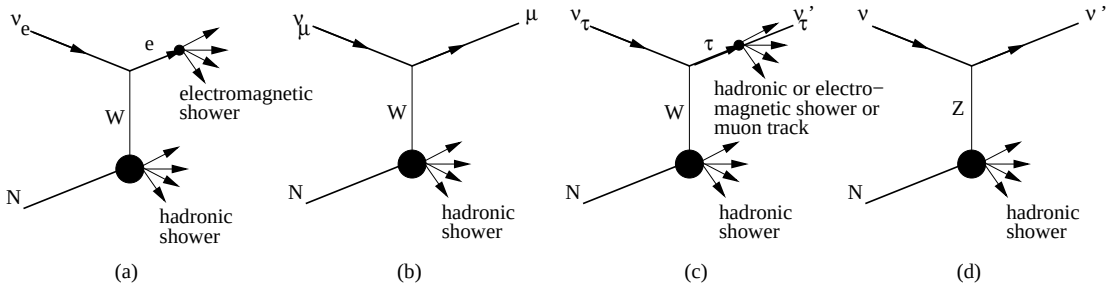



**Figure 2.5:** Neutrino-nucleon deep inelastic interaction channels. Charged current interactions of all three neutrino flavours (a – c) and neutral current interactions (d; identical for all neutrino flavours). See text for details.

In CC processes charged leptons are produced ($e, \mu$ or $\tau$), depending on the flavour of the incoming neutrino. Electrons produced by the interaction of $\nu_e$ (Fig. 2.5a) have only a short mean free path length (in the order of a few cm). Shortly after their generation, they initiate an electromagnetic cascade consisting of $e^+$, $e^-$ and $\gamma$ by iterative energy loss processes (pair production and bremsstrahlung). Muons from $\nu_\mu$ interactions (Fig. 2.5b) have much longer trajectories than electrons, ranging from a few meters around 1 GeV up to several 10 km above 10 TeV (see also Section 3.1). The path length of taus produced in the interaction of $\nu_\tau$ increases with increasing energy, but only beyond 10 TeV it exceeds the length of the hadronic shower (about 10 m). Taus have only a short lifetime and most of their decay modes include the generation of a hadronic cascade as depicted in Fig. 2.5c.

In NC processes the outgoing lepton is a neutrino. The process is identical for all neutrino flavours: as shown in Fig. 2.5d, a $Z^0$ boson is exchanged between the neutrino and the nucleus and, like for the CC processes, a hadronic cascade is generated at the interaction vertex.

As will be discussed in the next subsection and in the following chapter, the CC processes are of main interest for conventional neutrino detectors in the deep-inelastic energy range. Therefore, only the leading-order cross section for the charged current deep inelastic scattering is given here [35]:

$$\frac{d^2\sigma^{cc}}{dxdy} = \frac{2G_F^2 M E_\nu}{\pi} \left( \frac{M_W^2}{Q^2 + M_W^2} \right)^2 [xq(x, Q^2) + x\overline{q}(x, Q^2)(1 - y)^2]. \tag{2.16}$$

$G_F = 1.17 \cdot 10^{-5}\,\mathrm{GeV}^{-2}$ is the Fermi coupling constant, $x$ and $Q^2$ the Bjorken scaling variable and the momentum transfer, respectively. $M$ is the mass of the nucleon, $M_W$ the mass of the W boson and $y = (E_\nu - E_l)/E_\nu$ the difference between the energy of the incident neutrino $E_\nu$ and the outgoing lepton $E_l$ divided by $E_\nu$. $q(x, Q^2)$ and $\overline{q}(x, Q^2)$ are the parton distribution functions for quarks and anti-quarks, respectively. In Fig. 2.6, the total cross sections for neutrino and anti-neutrino nuclear scattering are given as a function of the neutrino energy. It can be seen that below about 5 TeV (the interesting energy range for studying oscillations of atmospheric neutrinos in ANTARES will be around $10 - 100\,\mathrm{GeV}$) the cross section increases linearly with the logarithmic energy.

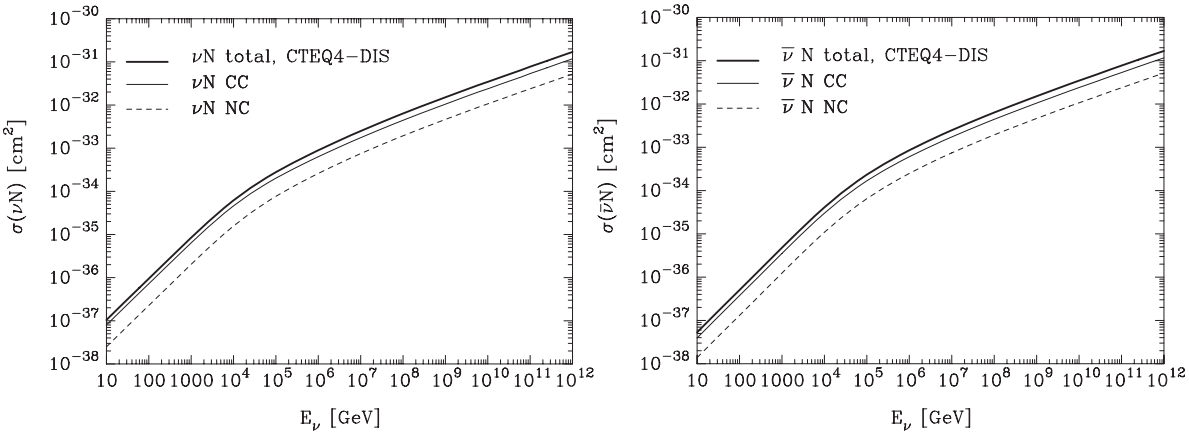



**Figure 2.6:** Cross sections for NC and CC scattering of neutrinos and anti-neutrinos from isoscalar nucleons as a function of the neutrino energy. Figure from [35].

### 2.4.2 Neutrino detection

Due to their tiny cross section, the detection of neutrinos is difficult. Either large neutrino fluxes or large detection volumina are necessary for the successful detection of a significant number of neutrinos. Hence, the choice of the detector type is correlated with the neutrino flux and consequently driven by the required detector size.

Conventional scintillation detectors, for example, can be employed for the detection of reactor and accelerator neutrinos. In such detectors, photons are measured, whose emission is indirectly induced by neutrino interactions[9]. One experiment that is detecting reactor neutrinos with a scintillation detector is, for example, the Double Chooz experiment [36], which is studying neutrino oscillations.

The flux of solar low-energy neutrinos is sufficiently large to measure them also with scintillation detectors (for example the BOREXINO experiment [37]). Alternatively,

[9]The exact mechanism depends on the detection media.

detectors based on radiochemical methods can be used (e.g. the Homestake [18] or the GALLEX [38] experiment). In such detectors, unstable nuclei are generated in neutrino capture reactions, whose radioactive decay can be detected.

When it comes to high-energy, atmospheric or cosmic neutrinos, with fluxes many orders of magnitude below the solar low-energy flux or the flux of reactor and accelerator neutrinos, huge detectors are needed. Typically, such detectors are based on the measurement of Cherenkov light, emitted by charged, secondary high-energy particles from neutrino interactions. The detection media employed are transparent media like water or ice. Natural volumina like the sea or the clear ice beneath the South Pole allow for the construction of large-scale detectors. The Cherenkov effect itself, as well as the optical neutrino detection via Cherenkov light will be discussed in detail below.

Cosmic neutrinos with ultra-high energies ($E_\nu > 10^{18}\,\mathrm{eV}$) may be detected with the Pierre Auger Observatory in western Argentina, which is designed for the detection of extensive air showers, initiated by cosmic (charged) particles with highest energies [39]. Ultra high-energy neutrinos can be identified as highly inclined showers that originate from deep inside the atmosphere, tagged by a significant electromagnetic component [40]. An alternative detection method for cosmic neutrinos with energies above $\approx 1\,\mathrm{PeV}$ uses an acoustical approach and will be introduced below.

### Optical detection - making use of the Cherenkov effect

Detecting high-energy neutrinos is mainly done by making use of the Cherenkov effect. If a charged particle traverses a dielectric medium with a velocity $v = \beta c$ exceeding the velocity of light $c_n = c/n$ in this medium (with the vacuum velocity of light $c$ and the refraction index of the medium $n$), it emits electromagnetic radiation, the so-called Cherenkov light. The molecules in the medium are polarised along the trajectory of the charged particle and act as dipoles. If the velocity of the particle is small ($v < c_n$), the dipoles are oriented isotropically with respect to the position of the particle, and the waves they emit interfere destructively. If $v > c_n$, the symmetry is broken and radiation is emitted under an angle $\Theta_{\mathrm{Ch}}$, which depends on the refraction index of the medium $n$, as well as on the velocity of the particle. $\Theta_{\mathrm{Ch}}$ is given by:

$$\cos\Theta_{\mathrm{Ch}} = \frac{c_n t}{\beta c t} = \frac{1}{\beta n}. \tag{2.17}$$

In the case of charged interaction products of high-energy neutrinos, the velocity almost equals the vacuum velocity of light ($\beta \simeq 1, v \simeq c$). Employing a transparent medium like water or ice as detection medium, high-energy charged particles can therefore be detected by measuring the Cherenkov radiation in photosensor arrays.

This technique is used for example in the ANTARES detector (see Chapter 3), where a volume of about $0.01\,\mathrm{km}^3$ in the Mediterranean Sea is instrumented with photomultipliers. The sea water is employed as detection medium. The IceCube neutrino telescope [41] at the South Pole, uses the clear ice in a depth from about 1 400 to 2 400 m beneath the surface as detection medium, and the Super-Kamiokande experiment [7], designed

for the detection of solar and atmospheric neutrinos, was built up of a large tank of fresh water equipped with photomultipliers.

The intensity of the Cherenkov radiation decreases with increasing wavelength (as $1/\lambda^2$). In the typical efficiency range of a photo-multiplier ($300-600\,\mathrm{nm}$) about 35 000 photons are emitted per meter of track length [42]. An essential condition for the measurement of Cherenkov radiation of individual particles is therefore perfect darkness in the detection volume[10].

While propagating through matter, the Cherenkov photons can be absorbed by or scattered off molecules and particles of the detection medium. Whereas absorption attenuates the intensity of the emitted light, scattering only affects its original angular distribution. By introducing the wavelength-dependent absorption length $\Lambda_{\mathrm{abs}}$ and scattering length $\Lambda_{\mathrm{scat}}$, the effects of photon absorption and scattering can be quantified. Both factors give the length at which on average a fraction of $1/e$ of the originally emitted photons is not absorbed or is unscattered, respectively.

The tracks of the charged particles crossing the detector can eventually be reconstructed from the known positions of the photosensors that are used to measure the light, the arrival times and the amplitudes of the light signals.

### Acoustic detection - an alternative detection mechanism

In natural detection media like sea water or ice, almost arbitrarily large volumina can be instrumented with photosensors to detect even cosmic neutrinos with highest energy and very small fluxes by means of the optical Cherenov technique. The only given restriction is the distance between the photomultipliers, that has to be around the attenuation length (i.e. the combination of absorption and scattering length, see Section 3.3.2) of light in this medium. Consequently, with increasing volume, also the costs for such a detector increase and in real life this usually is the limiting factor. In order to increase the instrumented volume without increasing the costs in the same way, an alternative detection method, basing on the thermo-acoustical model [43] is currently studied within the ANTARES collaboration [44].

The energy deposited by a neutrino interaction with a nucleus leads to a local increase in temperature, and thus to a local expansion of the detection medium[11]. This causes a short acoustic bi-polar signal, which can be detected by sensitive acoustic antennas. As the absorption length of sound in sea water exceeds the absorption length of light by a factor of ten to twenty, the instrumentation density of the detection volume could be significantly reduced.

Within the ANTARES detector an acoustic detection test system is installed, that consists of a three-dimensional array of hydrophones. Except for its small size, the system has all features required for an acoustic neutrino telescope and hence is suited for a feasibility study of a potential future large-scale acoustic neutrino telescope [45].

[10] As is always the case, perfection is almost impossible to obtain. What is meant in principle, is a best possible shielding from sun light or artificial light sources.

[11] The water in the Mediterranean Sea has a temperature of $\sim 14°$ and is not at the anomaly point.

Currently the acoustical background at the ANTARES site is investigated and studies concerning signal classification and source reconstruction are under way.

## 2.5 Overview of experimental neutrino oscillation status

Neutrino oscillations have been firmly established in the last twelve years by a series of different experiments. Such experiments for the determination of neutrino oscillation parameters are in general only sensitive to certain neutrino flavours, depending on the neutrino source and the detection techniques employed. The distance from the source and the energy of the neutrinos determine the flavour transitions that can be measured.

As discussed in Section 2.2, there will be a mixing among all three flavours of neutrinos in general. This three-flavour mixing scenario decouples, given that $\Delta m^2_{12} \ll \Delta m^2_{23}$ and $\Theta_{13} \approx 0$. In this case, the two-generation mixing model is a good approximation and each transition can be described by a two-generation mixing equation.

Typically, the two decoupled two-oscillation scenarios are referred to as "solar" and "atmospheric", because they are accessible via electron neutrinos from the sun and muon neutrinos generated in the Earth's atmosphere, respectively. Even though neutrinos from reactors or accelerators are also used for determining the oscillation parameters, the nomenclature is kept in literature. While the solar oscillation scenario constrains the parameters $\Delta m^2_{12}$, $\Theta_{12}$, and $\Theta_{13}$, the atmospheric scenario constrains the parameters $\Delta m^2_{23}$, $\Theta_{23}$, and $\Theta_{13}$. The only parameter common to both scenarios is $\Theta_{13}$, which may potentially allow for some mutual influence.

### 2.5.1 Neutrino oscillation experiments

Experiments for the investigation of neutrino oscillations employ different sources of neutrinos. In the following, the exemplary mentioned experiments are classified according to their neutrino source.

#### Solar neutrinos

In the late 1960s, the Homestake experiment [18] observed a significant deficit in the solar neutrino flux compared to the predictions of the standard solar model (SSM). Homestake was a radiochemical experiment that detected $\nu_e$ from the sun via the decay of $^{37}$Ar generated in the CC reaction of $\nu_e$ with $^{37}$Cl: $\nu_e + {}^{37}\mathrm{Cl} \rightarrow {}^{37}\mathrm{Ar} + e^-$. The measured deficit could be explained by $\nu_e \rightarrow \nu_\mu$, or $\nu_e \rightarrow \nu_\tau$ oscillations.

Since then, several experiments measured solar neutrinos in different energy bands. The Sudbury Neutrino Observatory (SNO) [46] took data from 1999 to 2006. It was a heavy-water Cherenkov detector designed to detect neutrinos produced by fusion reactions in the Sun. SNO was capable of detecting $\nu_e$ via CC interactions, but could also detect all other neutrino flavours via NC interactions and electron scattering. SNO, and also Super-Kamiokande (see below), allow for a directional and energy reconstruction of

the detected neutrinos. Both detectors have an energy threshold of about $E = 5\,\mathrm{MeV}$[12], and are sensitive to the solar flux contribution of neutrinos from the radioactive $^8$B decay in the sun.

Borexino [48] on the other hand focuses on the measurement of the solar $^7$Be neutrino flux (emerging from the process of $^7\mathrm{Be}^{4+} + e^- \rightarrow {}^7\mathrm{Li}^{3+} + \nu_e$) in the sub-MeV range, but is also capable of measuring $^8$B neutrinos. The detector is taking data since May 2007 and contains an ultrapure liquid scintillator viewed by more than 2 000 photomultipliers. Its energy threshold is at an energy of about 250 keV.

### Atmospheric neutrinos

The Super-Kamiokande experiment [7] is operational since 1996, and the collaboration was the first that published definite evidence for neutrino oscillations from the observation of atmospheric neutrino data [8]. Super-Kamiokande is a water Cherenkov detector and consists of a tank filled with 50 000 tons of ultra-pure water. More than 11 000 photomultipliers measure the light emitted by high-energy charged particles crossing the detection volume. The detector is sensitive to both solar electron neutrinos and atmospheric electron and muon neutrinos.

Super-Kamiokande has measured the flux of atmospheric $\nu_e$ and $\nu_\mu$ as a function of the zenith angle of the neutrinos. It turned out that the $\nu_\mu$ flux from the hemisphere below the detector is smaller than predicted by models without neutrino oscillations, but conversely, the $\nu_\mu$ flux from the hemisphere above the detector, as well as the electron flux from both hemispheres, are consistent with the predictions. This effect of vanishing $\nu_\mu$ from below can be assigned to an oscillation $\nu_\mu \rightarrow \nu_\tau$. Only for those $\nu_\mu$ the path length is long enough to allow for oscillations. On the contrary, the oscillation of atmospheric $\nu_e$ can not be resolved at these path lengths and energies.

### Reactor neutrinos

Employing neutrinos from nuclear reactors is another possibility of investigating the neutrino flavour oscillation phenomenon. The CHOOZ experiment [49, 50], for example, utilised the high intensity and purity of a reactor $\bar{\nu}_e$ flux. It was located at about 1 km from a nuclear reactor near the city of Chooz in France. CHOOZ, that took data in 1997 and 1998, detected the $\bar{\nu}_e$ emerging from the reactor via inverse $\beta$-decay. As for the CHOOZ reactor data, the survival probability of $\bar{\nu}_e$ at an average energy of a few MeV depends only on $\Delta m^2_{23}$ and $\Theta_{13}$, as long as $\Delta m^2_{12} \lesssim 3 \cdot 10^{-4}\,\mathrm{eV}^2$. CHOOZ reported no evidence for neutrino oscillations in the $\bar{\nu}_e$ disappearance mode. It could thus exclude that the observation of a $\nu_\mu/\nu_e$ ratio that is roughly one half of what expected, is due to a transition from $\nu_\mu$ to $\nu_e$. Together with the Super-Kamiokande results it became clear that the effect has to be caused by $\nu_\mu \rightarrow \nu_\tau$ oscillations. From this non-observation, CHOOZ also made an important contribution to the bound on $\Theta_{13}$.

[12] In a recent publication SNO reported an effective electron kinetic energy threshold of $E_{\mathrm{eff}} = 3.5\,\mathrm{MeV}$ [47].

KamLAND [51], a liquid scintillator detector at the underground Kamioka neutrino observatory in Japan, started data taking in 2002. Like CHOOZ, it detects $\bar{\nu}_e$ via inverse $\beta$-decay. The $\bar{\nu}_e$ are produced in nuclear reactors surrounding the detector site. With an energy threshold at about 1.8 MeV and a flux weighted average distance of $\sim 180$ km from the reactors, KamLAND is sensitive to the solar mixing parameters $\Delta m^2_{12}$, $\Theta_{12}$ and $\Theta_{13}$. The spectral information from the experiment allowed for a determination of $\Delta m^2_{12}$ with unprecedented precision. Furthermore, the KamLAND data has helped to constrain the lower bound of the mixing angle $\Theta_{12}$ [52], and has also set upper limits on $\Theta_{13}$ [53]. The detector is currently undergoing a purification upgrade which will enable KamLAND to execute a low energy neutrino program in parallel with the anti-neutrino program.

An upcoming reactor experiment, Double Chooz [36], is intended to improve the results of its predecessor CHOOZ concerning the determination of the $\Theta_{13}$ mixing angle. This will require an increase in statistics, a reduction of the systematic errors below one percent, and a careful control of the backgrounds. For this purpose, two identical detectors are used, one at 400 m distance, the other at 1.05 km distance from the Chooz nuclear cores.

### Accelerator neutrinos

Beam-dumps of proton accelerators are a source of $\nu_\mu$ and $\bar{\nu}_\mu$. Huge amounts of charged pions are generated when high-energy protons hit the beam-dump target and decay mainly to $\mu^+$ and $\nu_\mu$ or $\mu^-$ and $\bar{\nu}_\mu$, respectively.

The K2K neutrino oscillation experiment [54, 55] was located at Kamioka in Japan, 250 km away from the 12 GeV proton synchrotron accelerator KEK. At the KEK site, a neutrino beam was generated and precise measurements of the beam flux were performed. The contamination of the beam with other than $\nu_\mu$ flavours is determined using a set of nearby detectors. By comparing the measurement in Kamioka with the front detector measurement, possible effects due to neutrino oscillations can be investigated. K2K took data from 1999 to 2004. It measured the disappearance of $\nu_\mu$ and was sensitive to the atmospheric oscillation parameters $\Delta m^2_{23}$ and $\Theta_{23}$. As K2K used muon neutrinos from a well-controlled and well-understood beam, this was the first positive measurement of neutrino oscillations in which both the source and the detector were fully under control. The results of K2K are consistent with the oscillation parameters previously measured by the Super-Kamiokande collaboration using atmospheric neutrinos.

T2K is the second generation follow up to the K2K experiment [56]. The J-PARC facility, which is a more powerful proton synchrotron accelerator than KEK used in the K2K experiment, will produce an intense off-axis beam of $\nu_\mu$. Off-axis means that the initial particle beam is directed 2 to 3 degrees away from the detector, which lowers the flux of neutrinos reaching the detector but provides a more desirable neutrino energy spectrum. At the T2K baseline (295 km), maximal neutrino oscillation is expected to occur at energies lower than 1 GeV. The goal of the T2K experiment is to gain a more complete understanding of neutrino oscillation parameters by observing oscillation

from $\nu_\mu$ to $\nu_e$. The probability for this oscillation is controlled by the mixing angle $\Theta_{13}$. T2K could be the first experiment to measure the appearance of $\nu_e$ in a $\nu_\mu$ beam. Furthermore, T2K aims at a precise determination of the atmospheric mixing parameters [57].

The MINOS Collaboration operates another accelerator based neutrino oscillation experiment [58, 59], which started data taking in its final configuration in 2005. MINOS consists of two detectors. The first one is more than 700 km away from the neutrino source, the NuMI ("Neutrinos at Main Injector") beamline at Fermilab. The second detector is located only a few hundred meters away from the accelerator. Neutrino interactions in this detector are used to determine the initial neutrino flux and energy spectrum. Both detectors are steel-scintillator sampling calorimeters made of alternating planes of magnetised steel and plastic scintillators. Muons produced in $\nu_\mu$ interactions are deflected by a magnetic field. It is therefore possible to distinguish between neutrinos and anti-neutrinos, allowing for a search for CPT violation. Like K2K and T2K, MINOS is sensitive to the atmospheric oscillation parameters. In addition to producing precision measurements of $\Delta m^2_{23}$ and $\Theta_{23}$, MINOS also looks for the appearance of $\nu_e$ in the far detector in order to measure or to set a limit on the oscillation probability of $\nu_\mu \rightarrow \nu_e$ and thus on the mixing angle $\Theta_{13}$.

The OPERA neutrino detector [60] is located in the Gran Sasso Laboratory and was completed in 2008. It is exposed to the $\nu_\mu$ beam of the SPS accelerator at CERN. It has been designed to perform the first detection of neutrino oscillations in direct appearance mode through the study of the $\nu_\mu \rightarrow \nu_\tau$ channel. OPERA is a hybrid detector, consisting of two targets, each followed by a muon spectrometer for momentum and charge identification of penetrating particles. The targets are built of 75 000 "bricks" arranged in parallel walls and interleaved with plastic scintillator counters. The bricks are emulsion modules consisting of lead plates interleaved with emulsion films.

The distance between OPERA and CERN is about 730 km and the average neutrino energy is $\sim$ 17 GeV. The number of neutrino interactions expected to occur in the OPERA targets in five years is about 24 000. The number of $\nu_\tau$ charged current interactions is about 120 for $\Delta m^2 = 2.5 \cdot 10^{-3}\,\mathrm{eV}^2$, leading to an observation of about ten $\nu_\tau$ events with less than one background event. The observation of a first $\nu_\tau$ candidate was reported recently [61], with a significance of $2.36\,\sigma$ not being a background fluctuation. This does not allow yet for claiming the observation of $\nu_\mu \rightarrow \nu_\tau$ oscillation and the detection of a few more candidate events will be required to firmly establish $\nu_\mu \rightarrow \nu_\tau$ neutrino oscillations in direct appearance mode.

### 2.5.2 Global fits

Combining the results of several neutrino oscillation experiments, a global analysis of the data can be made and the three-flavour oscillation parameters can be obtained. The current best fit values from such a global analysis, including the data of solar, atmospheric, reactor and accelerator experiments, can be found in Table 2.1. Details on the analysis can be found in [62, 63, 64].

Spectral information from KamLAND data leads to an accurate determination of

**Table 2.1:** Neutrino oscillation parameters obtained from a global three-flavour analysis ($1\,\sigma$ errors). $|\Delta m^2_{23}|$ is given as absolute value as its sign, i.e. the neutrino mass hierarchy, is still unknown.

| | |
|---|---|
| $\sin^2(\Theta_{12}) = 0.318^{+0.019}_{-0.016}$ | $\sin^2(\Theta_{23}) = 0.50^{+0.07}_{-0.06}$ |
| $\Delta m^2_{12} = 7.59^{+0.23}_{-0.18} \cdot 10^{-5}\,\mathrm{eV}^2$ | $\lvert\Delta m^2_{23}\rvert = 2.40^{+0.12}_{-0.11} \cdot 10^{-3}\,\mathrm{eV}^2$ |

the solar mass parameter $\Delta m^2_{12}$. Whereas KamLAND data contribute also to the lower bound on $\Theta_{12}$, the upper bound is constrained by solar data. $|\Delta m^2_{23}|$ is dominated by data from MINOS. Mainly data from Super-Kamiokande determine the atmospheric mixing angle $\Theta_{23}$ and lead to a best fit point at maximal mixing. The mass hierarchy, i.e. the sign of $\Delta m^2_{23}$ remains undetermined by the present data.

The value of the third mixing angle $\Theta_{13}$ is not known at present, but is constrained to be small compared to the other two angles. Like the other oscillation parameters, the limit on $\Theta_{13}$ is obtained from the combination of different data sets; its current value is

$$\sin^2(\Theta_{13}) \leq 0.031 \tag{2.18}$$

at 90 % confidence level [62, 63, 64].

### 2.5.3 The LSND and MiniBoone results

From the LSND experiment, there is a hint for oscillations with a much larger $\Delta m^2$. LSND took data from 1993 to 1998, searching for $\bar{\nu}_\mu \to \bar{\nu}_e$ and $\nu_\mu \to \nu_e$ transitions in an appearance measurement. The neutrinos were generated in the interaction of an intense proton beam at the Los Alamos Neutrino Science Center with different targets. The beam stops were at a distance of 30 to 135 m from the detector and the neutrino energy was about 40 MeV. LSND reported an evidence for neutrino oscillations, based on an event excess in the $\bar{\nu}_\mu \to \bar{\nu}_e$ transition [65]. No clear excess has been observed in the $\nu_\mu \to \nu_e$ transition, but the results are consistent with the anti-neutrino oscillation signal. In conjunction with other available neutrino oscillation limits the LSND data suggest neutrino flavour oscillations with $0.2\,\mathrm{eV}^2 < \Delta m^2 < 10\,\mathrm{eV}^2$.

This result is incompatible with the mass-squared differences required by other experiments within the standard three-flavour framework. Various solutions to the problem are suggested. The standard one is to introduce one or more "sterile" neutrinos (see, for example, [22] and [63], and references therein), which are hypothetical particles that are assumed to interact only gravitationally.

The MiniBooNE experiment [66] is intended to shed more light on this topic and test the evidence of the transitions reported by LSND. The experiment has searched for a $\nu_\mu \to \nu_e$ appearance with a baseline of 540 m and a mean neutrino energy of about 700 MeV and thus a very similar $L/E_\nu$ range as LSND. The results of the experiment were consistent with no oscillations within a two-neutrino appearance-only oscillation

model [67] and were thus incompatible with the LSND results, under the assumption that oscillations of neutrinos and anti-neutrinos are the same.

Later on, in a more direct test of the LSND signal, MiniBooNE performed a search for $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ oscillations. Only recently, in October 2010, MiniBooNE reported an $\bar{\nu}_e$ event excess in this study [68]. The data are consistent with $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ oscillations in the 0.1 to 1.0 eV$^2$ $\Delta m^2$ range and with the evidence for anti-neutrino oscillations from LSND. No clear conclusion can currently be drawn from these results.

# 3 The ANTARES neutrino telescope

The ANTARES collaboration presently counts more than 150 active physicists and technicians from seven European countries, operates a neutrino detector in the Mediterranean deep sea, about 25 km off the French cost near the city of Toulon (Fig. 3.1). The collaboration has been formed in 1996 with the purpose of developing and constructing a deep-sea neutrino telescope. Twelve years later, in May 2008, the construction of the detector was successfully completed [69].

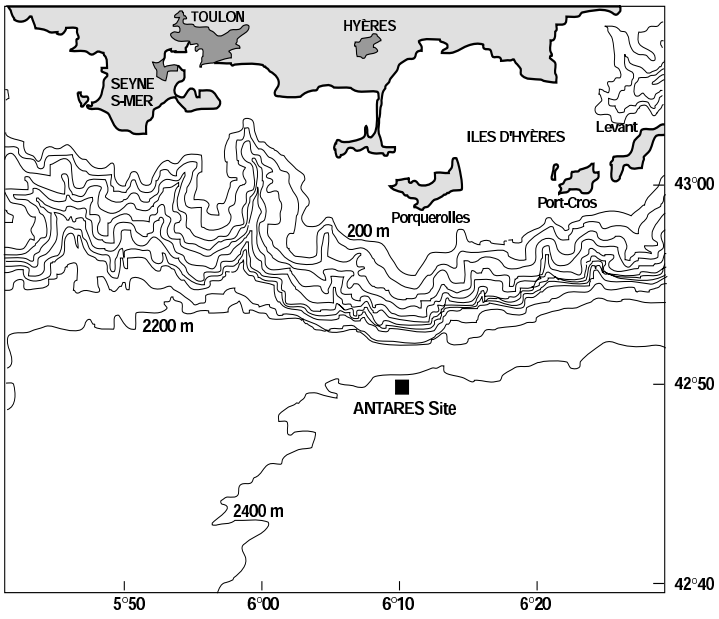



**Figure 3.1:** The location of the ANTARES detector, near the coast of southern France.

The detector itself is a large-scale water Cherenkov detector, designed and optimised for the detection of high-energy extraterrestrial and atmospheric neutrinos. It consists of a three-dimensional array of photosensors, and the sea water serves as detection medium. The detector is installed on the sea bed in a depth of about 2475 m, where it is completely shielded from sunlight. The sea water above the detector additionally serves as shielding against downward-going atmospheric muons.

In the first section of this chapter the detection principle of ANTARES will be explained in detail. The experimental setup, i.e. the layout of the detector and the data

acquisition will be illustrated in Section 3.2. Section 3.3 deals with the ambient conditions at the ANTARES site, like the optical background and the water properties. In Section 3.4, the trigger algorithms and the event building is introduced. The various event types measured by ANTARES are discussed in Section 3.5 and in Section 3.6 the simulation chain for ANTARES is illustrated. Finally, a summary of the software used in the ANTARES collaboration is introduced in Section 3.7, paying particular consideration to the analysis software.

## 3.1 Detection principle

### 3.1.1 Detector description

When high-energy neutrinos interact with nuclei, they generate highly relativistic, charged particles, which in turn emit Cherenkov light while propagating through the sea water. If such an interaction takes place inside or near the instrumented volume of the ANTARES detector, the Cherenkov radiation can be measured. Extremely sensitive photosensors, attached to mooring structures, are the "eyes" of the detector. They are capable of detecting single photons. From the known positions of the photosensors, the measured arrival times of the Cherenkov photons and the amount of light recorded, trajectories of charged particles and, to a certain extent, also the position of the interaction vertex can be reconstructed.

ANTARES is optimised for the detection of muons that are produced in CC interactions of high-energy $\nu_\mu$. Electrons, produced by $\nu_e$, only induce short-range electromagnetic cascades. In the ANTARES perspective, they have a point-like signature, which is difficult to identify and even more difficult to distinguish from the signature of NC interactions.

Muons, on the other hand, have a long trajectory (Fig. 3.2). They lead to a track-like signature in ANTARES, which enables a comparatively easy separation of signals from optical background and a precise directional reconstruction. Furthermore, the detection volume is larger for $\nu_\mu$ compared to $\nu_e$, as the latter can only be measured when the interaction vertex is inside or close to the instrumented volume. The muon track length is determined by the initial muon energy: muons lose energy on their way through matter through various processes (Section 3.1.3) and eventually stop before they decay. Because of their key role, production, propagation, and energy loss of muons in sea water will be discussed in the following subsections.

Even though about one third of the cosmic neutrinos are expected to arrive as $\nu_\tau$, they could only for $E_\nu > 1\,\mathrm{PeV}$ be identified due to their characteristic signature being that of a shower, a track, and a second shower: most decay modes of the tau include the generation of a hadronic or electromagnetic cascade. Because of the short lifetime of taus, their flight distance before decay is short (several ten meters at $E_\nu = 1\,\mathrm{PeV}$), but for taus with $E_\nu > 1\,\mathrm{PeV}$, the track is long enough to distinguish between the shower from the initial interaction of the $\nu_\tau$ and the shower from the decay of the tau. Below this energy, taus would have a point like signature in ANTARES and their detection and

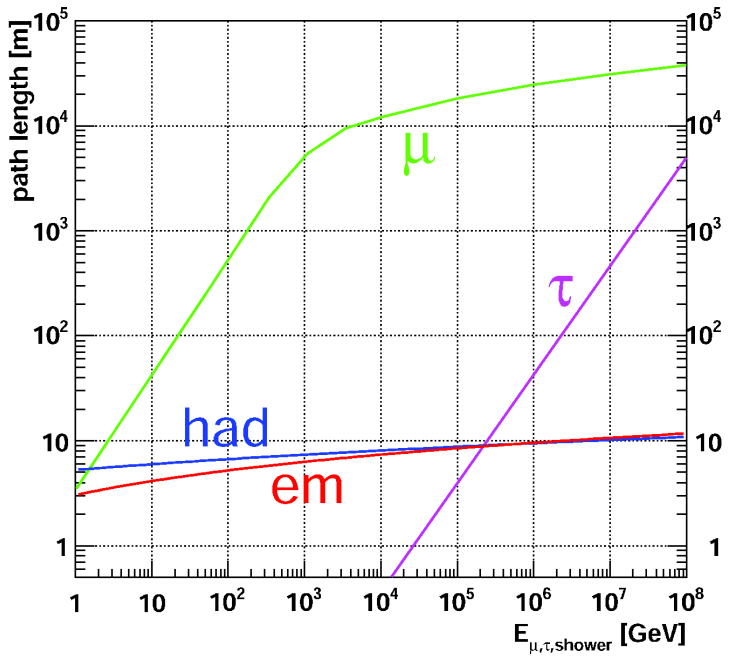



**Figure 3.2:** Range of muons, taus, and hadronic and electromagnetic cascades in sea water, as function of their respective energy.

identification is subject to the same problems as for other events with such signatures. Contrary to muons, energy loss processes play a much smaller role, because the 17 times larger mass of the tau.

### 3.1.2 Muon production

When $\nu_\mu$ interact with nuclei via a CC interaction, a muon is generated. ANTARES is sensitive to muons with an energy above $\sim 10\,\text{GeV}$. As discussed above, the deep inelastic interaction is predominant at these energies. The average scattering angle between $\mu_\nu$ and muon depends on the neutrino energy and is very small for high-energy neutrinos. It can be parametrised by

$$\Delta\Theta_{\text{scat}} = \frac{0.7^\circ}{(E[\text{TeV}])^{0.6}}, \tag{3.1}$$

where $E$ is the energy of the neutrino [70]. For energies of $\sim 500\,\text{GeV}$, the scattering angle between muon neutrino and muon is around 1° and is further decreasing with energy; above 25 TeV it is already smaller than 0.1°. In this energy range, the resolution of the direction measurement of the neutrino is usually limited by the uncertainty on the direction of the reconstructed muon track. For energies in the range of some ten GeV, the deviation between the directions of the $\nu_\mu$ and the muon is a few degrees (see also Section 4.2), and the precision of the direction reconstruction is dominantly limited

by this scattering angle. The fractional energy transfer to the muon ($E_\mu/E_\nu$) depends on the original energy of the neutrino and increases from about 50 % on average at 100 GeV to more than 70 % on average at 1 PeV. This correlation allows for drawing firm conclusions on the neutrinos from measuring the muons.

### 3.1.3 Muon propagation

The muon itself suffers from multiple scattering on its way through rock or sea water, but in general the deviations from the original direction are smaller than the scattering angle between muon neutrino and muon. The muon thus essentially retains the direction of the muon neutrino and therefore points back to the source of the neutrino. This eventually legitimates calling instruments like ANTARES neutrino telescopes and justifies the usage of the term neutrino astronomy.

On their way through matter, muons interact with the medium through various processes. The ionisation (or excitation) of atoms and molecules of the surrounding matter, transferring each time a small amount of energy, is an almost energy independent process. It results in a nearly constant energy loss of about 0.2 – 0.3 GeV per meter water equivalent. Below about 100 GeV the energy loss is dominated by ionisation and, as a consequence, the path length of the muon is linearly correlated with the energy.

Above about 1 TeV, radiative processes start to dominate the total energy loss: in the nuclear electric fields of atoms, muons can radiate photons (bremsstrahlung) or produce electron-positron pairs (pair-production). The energy loss due to these radiative processes increases linearly with energy and is of stochastic nature, implying that the actual energy loss is strongly fluctuating.

In Fig. 3.3, the average differential energy loss of muons propagating through matter is illustrated.

## 3.2 Experimental setup

### 3.2.1 Detector layout

ANTARES consists of 885 photosensors (photomultiplier tubes, PMTs) in total, arranged on twelve so-called lines or strings. From each string, an electro-optical cable runs to a junction box (JB), that again is connected to the shore station in La-Seyne-Sur-Mer via a deep-sea communication cable.

A deadweight at the bottom string socket (BSS) of each line keeps the bottom of the lines down to the sea bed, while buoys at their top ends hold them vertically upright. The strings are free to swing and rotate in the undersea current and have a total length of about 450 m.

One string consists of 25 rigid titanium structures (storeys), which are equidistantly distributed along an electro-mechanical cable (EMC). The lowermost storey is at a height of about 100 m measured from the seabed, and the distance between two neighbouring storeys is about 14.5 m.

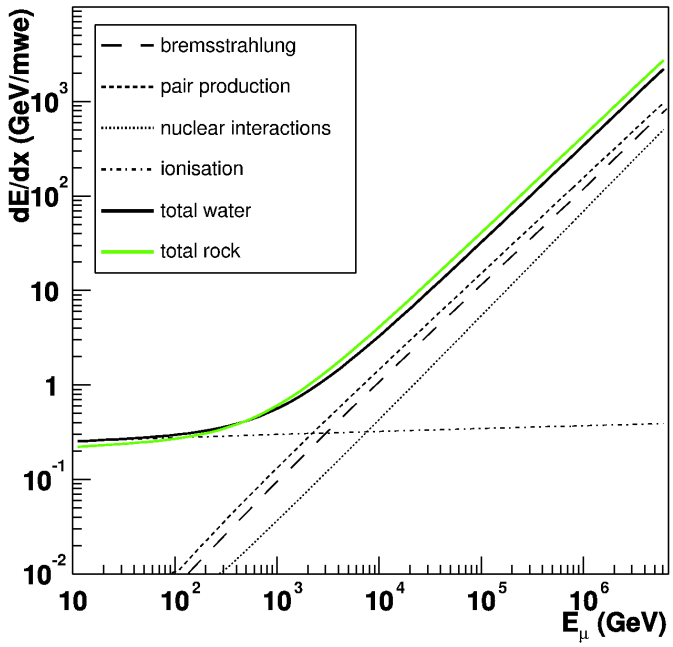



**Figure 3.3:** Differential energy loss of muons propagating through matter in GeV per meter water equivalent (mwe), as a function of the muon energy. The total energy loss for muons propagating through water (black solid line) and through rock (green solid line), as well as the contributions from different interaction processes are indicated (dashed lines).

Three pressure-resistant glass spheres (optical modules, OMs) are fixed to the titanium frames with an angular spacing of 120°. Each of them is housing a 10" photomultiplier tube. The PMTs are orientated downward at an angle of 45° from the horizontal. The OMs are electrically connected to a local control module (LCM) in the central part of the storey. The LCM consists of a titanium cylinder that contains the front-end electronics of the storey. For technical reasons, five neighbouring storeys are grouped together, forming a so-called sector. The second storey of each sector (counted from the bottom) is equipped with a master local control module (MLCM) that gathers the information from the five storeys of the respective sector. A string control module (SCM) in the BSS gathers the data from the MLCMs. The SCMs from all the lines are connected directly to the JB. Line 12 is somewhat exceptional because it consists only of four optical sectors. The five topmost optical storeys are replaced by storeys with acoustical sensors (hydrophones) and other instruments.

The octagonal footprint of the detector with the numbering of the lines can be seen in Fig. 3.4. The spacing between two neighbouring strings is about 60 – 75 m. ANTARES thus covers an area of about 180 m × 180 m on the ground. The layout of the detector is depicted in Fig. 3.5. In both figures, the instrumentation line (IL) is also shown, which contains further acoustical storeys and monitoring instruments.

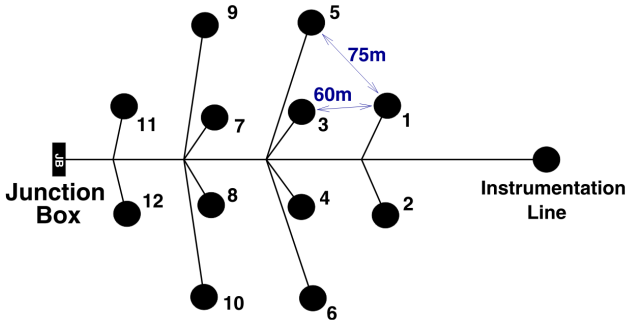



**Figure 3.4:** Footprint of the ANTARES detector.

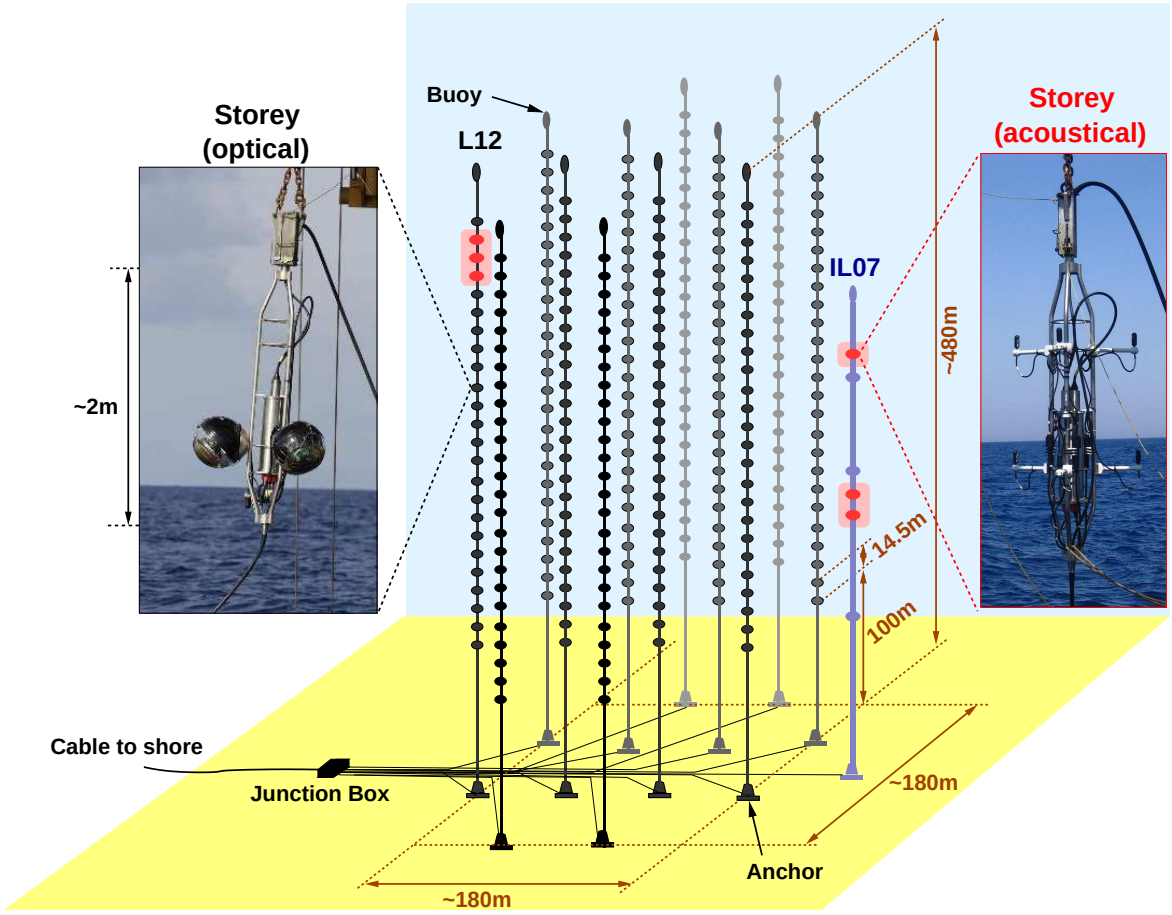



**Figure 3.5:** Schematic view of the layout of the ANTARES detector. Also shown are pictures of a standard storey with optical modules and of an acoustical storey with hydrophones. See text for details.

As the lines are subject to movement by deep sea currents, the positions of all OMs need to be monitored in short time intervals. For this purpose, all lines carry tilt-meters and compasses. Additionally, a system of acoustic transmitters and hydrophones is employed. By measuring the time between emission and reception of acoustical signals, the positions of the hydrophones (which are attached to every fifth storey) can be triangulated. With these positions and the tilt-meter data as input, the shape of each line, including the orientation of each storey, can be reconstructed. This process is generally referred to as alignment.

The downward orientation of the OMs renders them mostly sensitive to particles coming from below. The reason for this design is the flux of atmospheric muons produced in the atmosphere by impinging cosmic particles. It exceeds the flux of muons from atmospheric muon neutrinos by several orders of magnitude, even in the depth of the Mediterranean Sea that already shields the detector (see also Section 3.5). In ANTARES, it is not possible to extract $\nu_\mu$ out of this background, because the signatures of downward-going neutrino-induced muons and of atmospheric muons are identical.

Because of being subject only to the weak interaction, neutrinos are able to cross the whole Earth[13], contrary to electrically charged particles. Upward going muons can therefore only be neutrino induced instead of being of atmospheric origin and hence the identification of neutrino induced muons is straightforward[14].

Another reason for the downward oriented design is that the transparency loss of the OMs due to sedimentation can be reduced. Sediments are covering mostly the top side of the OMs. With downward-facing photomultipliers inside the glass spheres, the loss of transparency is only 1.5 % per year [71].

### 3.2.2 Data acquisition

The analogue PMT signals have to be converted into digital data and be transmitted to the shore station. The data need to be filtered and stored on disk, while information about the measurement settings and the calibration are archived in a database. All these processes are covered by the data acquisition system (DAQ), which will be described in this subsection. For logical reasons, it is divided into an "off"- and "on-shore" part. The first part deals with the hardware components that are located in the deep sea and the related data acquisition processes. The on-shore components and processes are discussed in the second part. In principle, also the trigger algorithms are part of the on-shore DAQ, but they will be discussed in a separate section (Section 3.4).

#### Optical modules, front-end electronics and off-shore DAQ

All components of the OMs are housed inside a glass sphere with a diameter of 43 cm. The essential part of one single module is the 14-stage 10" diameter Hamamatsu photomultiplier R7081-20 [72, 73, 74]. It is optically coupled to the glass sphere by a

[13] Above a certain zenith dependent energy (about 100 TeV at $\Theta = 180°$), the Earth is getting opaque to neutrinos.

[14] Assuming perfect muon track reconstruction.

silicone gel; a $\mu$-metal cage shields it from the terrestrial magnetic field. The average FWHM transit time spread (TTS) of the photomultipliers is 2.6 ns, and the overall detection efficiency (i.e. the combination of the quantum efficiency of the photocathode and the electrostatic collection efficiency) of the light hitting the photocathode surface is around 16 % [75]. The spectrum of this light is the convolution of the Cherenkov light spectrum and of the attenuation due to the gel, the glass sphere and the sea water. The PMTs are sensitive to single photons in a wavelength range between 300 nm and 600 nm. The highest sensitivity is reached around 400 nm, which matches well the wavelength distribution of the photons hitting the PMTs (see also Section 3.3.2).

The angular acceptance of the photomultipliers can be seen in Fig. 3.6 which shows the results of new measurements and simulations [76, 77]. They revealed that the acceptance in the boundary region around $\cos\Theta_c \sim 0$ ($\Theta_c$ is the angle between the PMT axis and the Cherenkov light) is about 20 – 30 % higher than previously thought.

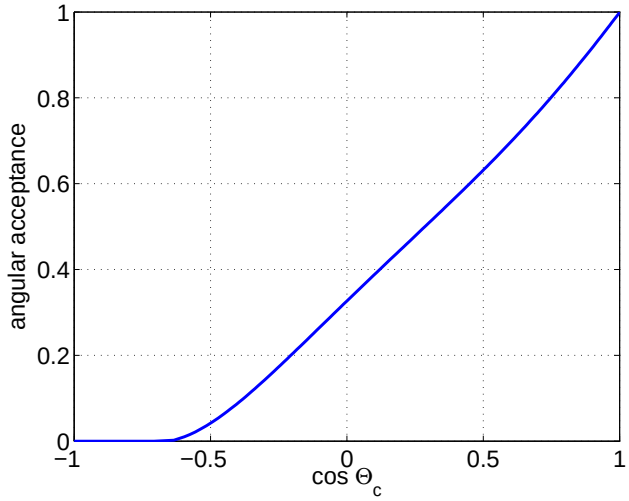



**Figure 3.6:** The angular acceptance of the photomultipliers as a function of the angle $\Theta_c$ between the PMT axis and the Cherenkov light, normalised to one at $\Theta_c = 0$.

The analogue signal of a PMT is recorded and converted into a digital signal by the front-end electronics. Two custom-designed analogue ring sampler chips (ARS) per optical module record the information from the photomultipliers and send the data to a field programmable gate array (FPGA) which buffers and converts the data stream into a dedicated data format [78]. A scheme of the main electronics components of the off-shore DAQ contained in the titanium cylinders of each storey is depicted in Fig. 3.7.

The ARS chips are able to record the data in two different modes, waveform and single photo electron mode. In the waveform mode, which is only used for calibration purposes, the pulse shape of a signal is recorded by sampling the anode signal with a tunable frequency of up to 2 GHz. In order to minimise power consumption and the transmitted data volume, the ARS operate in the single photo electron mode (SPE) by

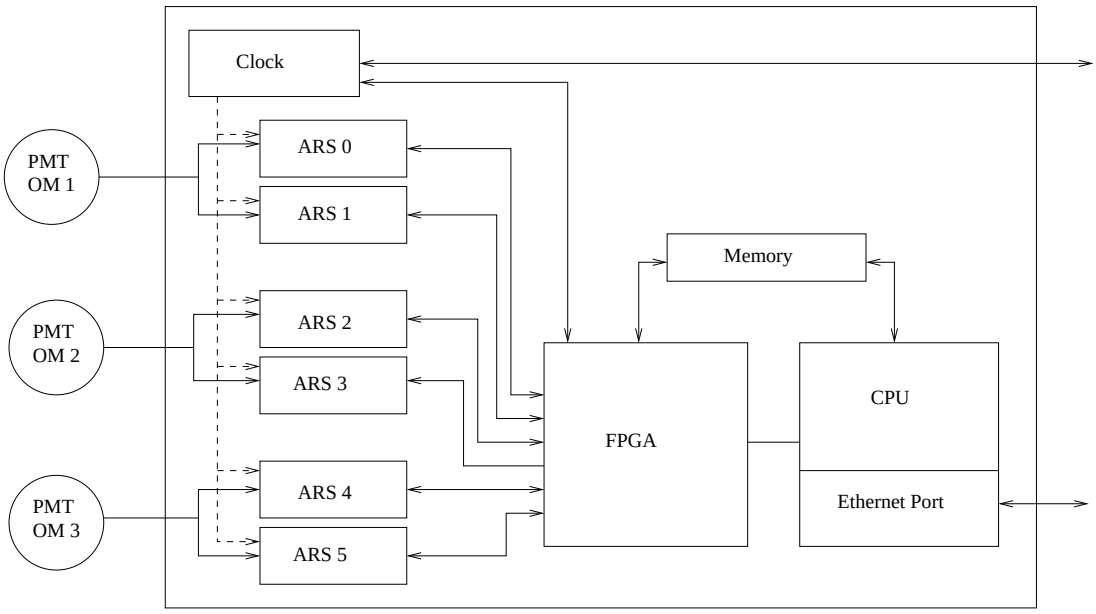



**Figure 3.7:** Scheme of the front-end electronics of the ANTARES optical modules. Each of the three photomultiplier tubes of one ANTARES storey is connected to two ARS chips, which are synchronised by a common clock signal. The FPGA collects the data, which finally is sent to shore by the CPU.

default. As soon as a PMT pulse crosses a tunable threshold, the recording of the data starts and the signal charge is integrated over a certain time window. The threshold is typically set to the equivalent of 0.3 photoelectrons (pe), which is enough to suppress the dark current noise of the PMT. The integration gate is chosen to integrate most of the PMT signal and at the same time limiting the contribution of electronic noise. It is set to 25 ns after the signal crossing the threshold plus 8 ns before. After recording a signal, the ARS has a dead time of about 250 ns. For this reason, a second ARS is connected to each PMT to provide readout functionality during this dead time. The two ARS communicate via a token ring protocol. After the integration gate of the first ARS closes, it takes $10-20$ ns for the second to take over. When the PMT signal crosses the threshold again (at least 40 ns after the first time) it will be recorded by the second ARS.

A timestamp is assigned to every PMT signal at the moment the pulse crosses the threshold. A local clock system in each storey provides the timing information. To provide a common clock signal to all ARS chips, the local systems are synchronised with an on-shore master clock that generates a 20 MHz clock signal, which again is synchronised to the GPS time with an absolute accuracy of 100 ns.

The data from the ARS chips are collected by the FPGA in time frames of 104.858 ms. This large value is chosen in order to minimise the probability of a physical event measured by the detector (which has a typical duration of $1-5\,\mu$s) to cross the boundary of the frame.

Finally, the data is transmitted to shore by a central processing unit (CPU). The CPU is the interface between off-shore data recording and the online data processing system and is connected to a fast Ethernet port (100 Mb/s). For the data transport the TCP/IP protocol is used.

The Ethernet port of each storey is connected with the MLCM of its sector, which has additional functionality compared to the standard LCMs. Within the MLCM, the links from the five storeys of one sector are merged by an Ethernet switch into a single Ethernet link.

The data transport is organised with multiple wavelengths that transmit different data streams along a single fibre. This technique is referred to as dense wavelength division multiplexing (DWDM). The five DWDM channels (one per sector) are optically multiplexed in the string module, located at the bottom of each string.

Each string is connected to the JB via an electro-optical cable. The JB again is connected to the shore station by one single (main) electro-optical cable (MEOC), where a (de-)multiplexer exists for each detector string. The data transfer policy in ANTARES is the "all-data-to-shore" concept, which means that all recorded signals are sent to shore, where different data processing methods are applied to the data.

### On-shore DAQ, data taking and processing

The on-shore part of the DAQ hardware consists of a farm of standard PCs and an Ethernet switch to which all on- and off-shore processors that are part of the DAQ are connected. The processors are forming a large network, where each of the processors is addressable by its unique IP address. This enables communication with all processors in the system and makes the data flow from the ARS to the on-shore processors completely transparent. Also belonging to the on-shore hardware are the master clock system and a DWDM transceiver for each line, which is used to multiplex slow control data streams for initialisation and configuration of the detector and to demultiplex the data streams from the ARS respectively.

Three types of processes have to be organised by the DAQ system: processes for data transfer and communication, processes for the detector operation and processes for data taking and data handling. In total, there are hundreds of processes that all need to be synchronised and, depending on their type and purpose, are running either on off- or on-shore processors.

Details concerning data transfer and communication processes as well as a description of the detector operation processes can be found in [79]. Data processing and filtering of the data are of interest for the subsequent analysis. As the filtering plays an important role, it is described separately in Section 3.4, and only the handling of the data will be discussed in the following.

The raw data from the off-shore processors is sent to shore in time frames. Each ARS sends one frame, and the simultaneously recorded frames of the different ARS chips are all sent to one processor in the on-shore PC farm. The collected bunch of frames now contains the complete data recorded by the ARS chips during this time interval and forms a so-called time slice. The frames of the subsequent time interval are sent

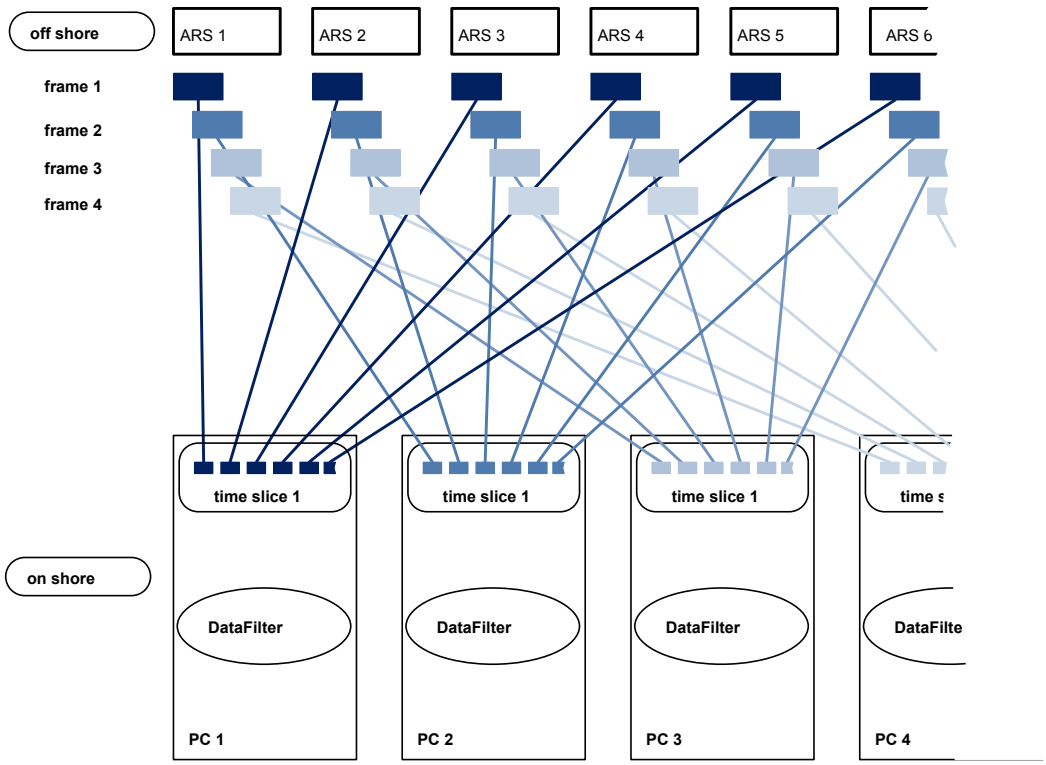



**Figure 3.8:** Scheme of the data processing in ANTARES. All frames belonging to the same time window are sent from the ARS to a single PC, forming a time slice. The data filters running on each PC process the data in the time slice.

to another PC, because only when the processing of one time slice is finished, the PC accepts another one. This processing scheme is illustrated in Fig. 3.8.

The data processing software (i.e. the data filters) implements algorithms (triggers) that are designed to detect physics signals. Various of such triggers are running in parallel, all having certain purposes. The two trigger algorithms that are of relevance for this analysis will be described in Section 3.4.

The physics event candidates selected by the data filter programs are passed to a data writer process, which formats the events and finally writes them to disk in ROOT [80] format. The size of one ROOT file is limited to about 2 GB and one file corresponds to what is called a run in ANTARES. Normally a new run starts when the data limit is reached, but new runs can also be started manually, for example, after changing the detector settings.

## 3.3 Ambient conditions

### 3.3.1 Optical noise

The photomultipliers of ANTARES are sensitive to single photons of the faint Cherenkov radiation emitted by highly relativistic, charged particles. Shielding the detector from surface light is a fundamental condition for measuring these signals. Though no sunlight is able to reach the ANTARES detector in the depth of the Mediterranean Sea, background light signals are produced by different sources and processes.

The intensity of this optical background is specified as the number of measured photons per photomultiplier per second. It is referred to as background rate and fluctuates with time, typically in the range between 50 – 80 kHz, but it has been observed to increase up to more than 100 kHz or even into the MHz range.

The decay of radioactive $^{40}$K, which is naturally contained in salt water, leads to a constant and homogeneous rate of about 30 kHz [81]. The coincidence rate (i.e. the rate of signals on different OMs of the same storey within a time window of 20 ns, see Section 3.4.1) due to $^{40}$K, which is important to consider because it might cause accidental triggers, is about 10 – 15 Hz.

This kind of background is well understood and comparatively easy to simulate. Much more challenging is the contribution of light from living organisms (bioluminescence) to the background rates, because it is fluctuating with time and is not homogeneous. Bioluminescent bacteria and other micro-organisms are more or less uniformly distributed. They produce a single-photoelectron background rate of around 20 – 50 kHz, which is to a certain extent correlated with the sea current. Despite that, light is also emitted by larger multi-cellular deep sea inhabitants, like squids or crustaceans. They are responsible for aperiodic and localised light bursts that can lead to rates of a few MHz in the affected PMTs. Altogether, bioluminescent organisms are responsible for large fluctuations on short (minutes to hours) and long timescales (seasonal) in the measured background rate.

Commonly, the quality of the recorded data with respect to the contamination with optical background is quantified by two values, the baseline rate and the burst fraction. The baseline rate is computed by fitting a Gaussian to the distribution of counting rates from the single PMTs, while the burst fraction is the fraction of time, where the counting rates exceed the baseline rate by 20 %. These two parameters are, among others, used for the assessment of the data quality and are criteria for the selection of runs for the analysis. In Fig. 3.9, the baseline rate and the burst fraction per run are shown, for all physics runs between January 2007 and October 2009. It can be seen, that the baseline rate most of the time fluctuates between about 50 and 80 kHz, but also exhibits peaks up to 300 kHz during the first half of the years 2007 and 2009. The burst fraction mainly ranges between about 5 and 40 % (median value is 18 %), but fluctuates up to 80 %.

According to its definition, the calculation of the burst fraction depends on the baseline rate. In turn, the computing of the baseline rate does not give meaningful results during times of high bioluminescent activity. To improve this somewhat unsatisfactory

situation, an alternative approach was followed for this thesis, see Chapter 5.

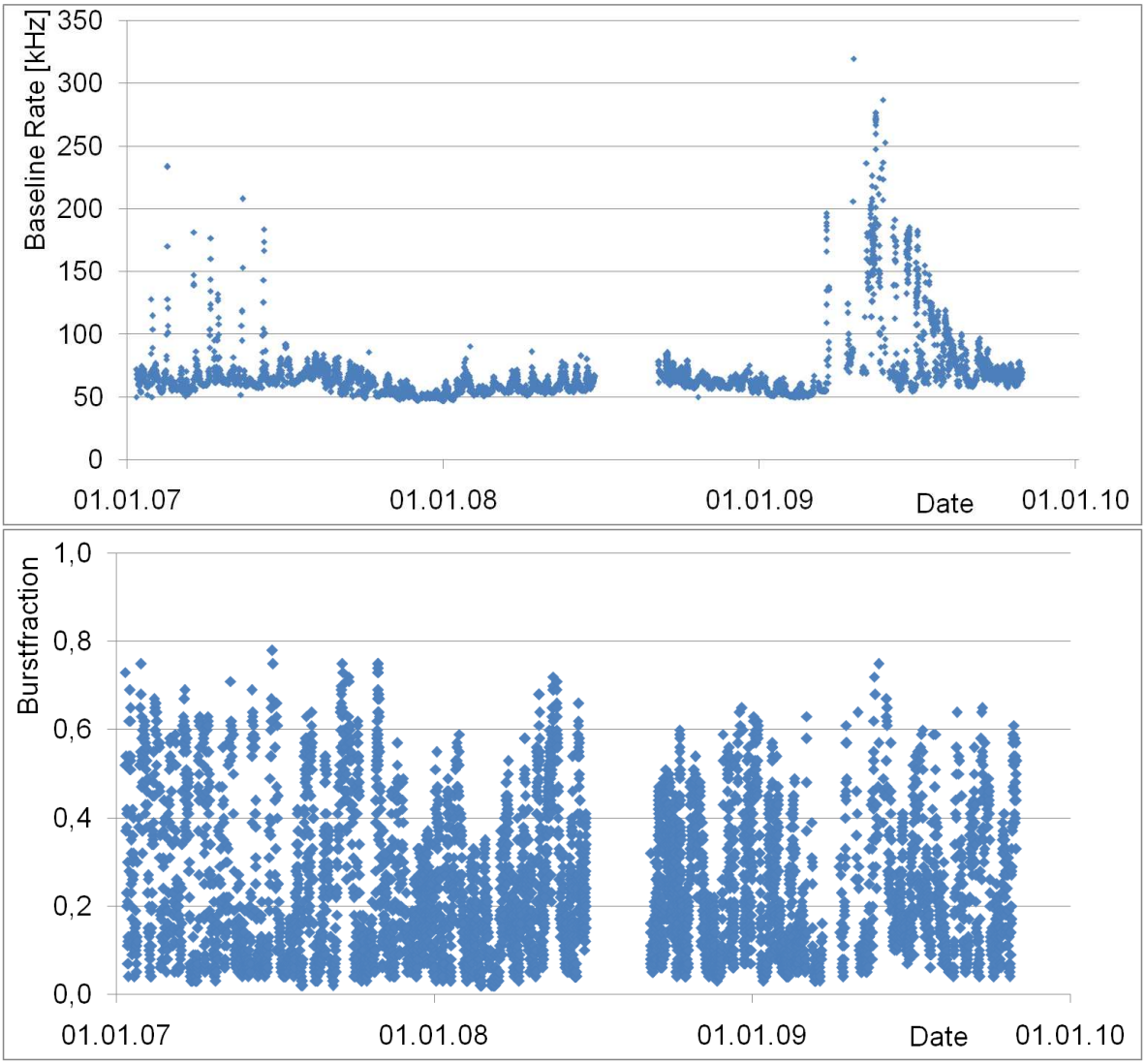



**Figure 3.9:** The baseline rate (above) and the burst fraction (below) for all physics runs between January 2007 and October 2009.

### 3.3.2 Water properties at the ANTARES site

Certain parameters of the sea water have impact on the measurement or are required for the reconstruction of the muon tracks. The detector is affected by deep-sea currents, in which the strings with the photosensors move. The amplitude of this displacement is of the order of several meters and depends on the velocity of the currents, with typical values of $4-10\,\mathrm{cm/s}$ and rare peaks up to more than $30\,\mathrm{cm/s}$.

Absorption and scattering length of light and the refraction index of the sea water, which determines the velocity of light in water, have been studied in detail during several campaigns for the site evaluation [82] and are monitored continuously. Both absorption and scattering length depend on the wavelength of the light. In the wavelength range the PMTs are sensitive (300 – 600 nm), the absorption length has a maximum around 400 – 500 nm, while the scattering length increases with increasing wavelength. Convoluting this with the spectrum of the Cherenkov light, the attenuation of the glass spheres and the optical gel between glass sphere and photomultiplier, results in a peak around 470 nm in the spectral distribution of the photons hitting the photomultiplier.

The attenuation length of light $\Lambda_{\text{att}}$ is defined as

$$1/\Lambda_{\text{att}} = 1/\Lambda_{\text{abs}} + 1/\Lambda_{\text{scatt}}, \tag{3.2}$$

with the absorption length $\Lambda_{\text{abs}}$ and the scattering length $\Lambda_{\text{scatt}}$. As scattered light is not necessarily lost, the scattering length is replaced by an effective scattering length $\Lambda^{\text{eff}}_{\text{scatt}}$ which depends on the average cosine of the scattering angle distribution $\langle\cos\Theta\rangle$:

$$\Lambda^{\text{eff}}_{\text{scatt}} = \frac{\Lambda_{\text{scatt}}}{1 - \langle\cos\theta\rangle}. \tag{3.3}$$

Employing the effective scattering length, the attenuation length becomes an effective attenuation length $\Lambda^{\text{eff}}_{\text{att}}$.

The ANTARES Optical Beacon System has been used to estimate $\Lambda^{\text{eff}}_{\text{att}}$ [82, 83] at the ANTARES site, resulting in a value of about 46 m for a wavelength of 472 nm (corresponding to the wavelength of the LEDs used for the measurements) The values for the absorption and the effective scattering length included in the relevant Monte-Carlo simulations (Section 3.6) are $\Lambda_{\text{abs}} = 55\,\text{m}$ and $\Lambda^{\text{eff}}_{\text{scatt}} \approx 53\,\text{m}$ (both at $\lambda = 472\,\text{nm}$) [84].

The refractive index of sea water connects directly to the velocity of light in water and thus to the Cherenkov radiation angle and the arrival times of the photons at the photosensors. Studies concerning the refractive index of the sea water at the ANTARES site can be found in [85] and [86]. The value of the refractive index for the phase velocity of light used in the simulation and for track reconstruction is $n = 1.35$ (assumed to be constant for the relevant photon spectrum).

## 3.4 Trigger and event building

The measured data consist of light signals, referred to as pulses or hits. From the identity of the ARS and hence the position of the associated OMs, the positions of measured photons are known. The amplitude of a hit is given in the unit of photoelectron equivalents, i.e. the number of photons detected by the PMT. This value is obtained after a charge calibration performed separately for each ARS.

Trigger algorithms are employed to identify Cherenkov photons emerging from high-energy charged particles within the continuous stream of background hits that superpose the Cherenkov signal. If one of the active filters finds that the recorded data are

compatible with a signal, a physics event is created. Only triggered events are stored and the (untriggered) rest of the data is discarded, except for the so-called summary time slices that contain the measured rate of each OM per time slice.

### 3.4.1 Trigger

Various triggers are available for different purposes: there are some universal triggers with different sensitivities, but also triggers for dedicated searches, for example a magnetic monopole trigger. A standard trigger (*3N* trigger) is running by default during all times of regular data taking. Other trigger algorithms are additionally switched on and off as needed. Special trigger algorithms also exist for example for in-situ calibration purposes. The choice of active triggers usually depends on the background rates. When rates are low, more sensitive triggers can be applied, whereas the trigger algorithms need to be strict during periods of high background rates to guarantee that the selected data rate respects the available bandwidth. For this work only the default *3N* trigger and another universal but more sensitive trigger (*2T3* trigger, only active when the background rate is smaller than about 80 kHz) were considered and will be explained in the following.

The time slices contain the calibrated and unfiltered PMT signals, which are referred to as *L0* hits. In a first step, common to all trigger algorithms, a basic hit selection is applied: big hits, having an amplitude larger than a certain value (usually either 3 pe or 10 pe) and coincident hits, that are measured on different OMs of the same storey within a tunable time window (20 ns by default) are tagged as *L1* hits. Background processes like bioluminescence and the decay of $^{40}$K mostly produce uncorrelated single hits, but there is a certain rate of accidental *L1* hits due to background hits that increases with increasing background rate. Consequently, more advanced filter mechanisms are required, but the *L1* selection is forming the basis for all subsequent trigger algorithms.

Within the *L1* hits, the *3N* trigger is looking for causally connected hits. The causality condition is given by

$$|\Delta t| < \frac{1}{c_n} \cdot |\Delta x|, \tag{3.4}$$

where $|\Delta t|$ is the absolute time difference between the hits, $|\Delta x|$ the absolute distance between their positions and $c_n$ the velocity of light in water.

Hits that are causally connected form a cluster. As soon as the size of a cluster exceeds a certain value (five by default), it is accepted for the next step, in which a scan of muon trajectory directions is made. Only if the hits are consistent with a muon from a certain direction, the cluster is accepted [87]. By doing so, the accidental trigger rate can be reduced to less than 0.1 Hz for a background rate of about 100 kHz. The efficiency of this trigger (i.e. the number of triggered neutrino induced muons with hits on at least six floors and two lines, compared to the total number of such events) strongly depends on the energy of the events and ranges from around 5 % at 20 GeV neutrino energy up to about 70 % at 1 PeV [88].

The *2T3* trigger is simpler, but at the same time more sensitive, which leads to a higher accidental trigger rate. It searches for *L1* hits in adjacent or next-to-adjacent

storeys within a time window of 100 ns or 200 ns, respectively. Two of these so called *T3* clusters within a time interval of 2.2 $\mu$s are needed to accept the data as physics event. Two *T3* clusters can share one *L1* hit such that three *L1* hits can be sufficient, if all other conditions are fulfilled.

The *2T3* trigger has an efficiency of around 20 – 30 % between 10 and 100 GeV. For high energies around 1 PeV its efficiency is a few percent higher than that of the *3N* trigger. The accidental trigger rate at 100 kHz background rate is 0.3 Hz, assuming uncorrelated, single photoelectron background. Current investigations indicate, that the optical background also contributes a certain amount of hits with a big amplitude and additional coincident hits. This effect is not yet included in the simulations and the impact on the accidental trigger rate has not been studied.

### 3.4.2 Event building

Finally, the data is stored as physics events which contain a header with some general information and a list of hits. In the hit list all *L0* hits between the first and the last hit of the triggered sequence (i.e. "triggered hits") are stored as well as all *L0* hits within a snapshot window that comprises the 2.2 $\mu$s before the first and 2.2 $\mu$s after the last triggered hit. If within the snapshot window a second hit sequence is triggered (independent of the trigger type), the two events are merged into one physics event.

The size of 2.2 $\mu$s of the snapshot window corresponds to the time a relativistic particle needs for traversing the detector in vertical direction. This ensures that the whole physics signal is contained in a physics event. Altogether, the duration of the data stored in a physics event is $2 \cdot 2.2\,\mu\text{s} + (t_2 - t_1)$, where $t_1$ and $t_2$ are the time of the first and last triggered hit, respectively.

The total data output rate of the complete detector is about 0.3 – 0.5 GB/s, which cannot be stored completely. Consequently, only triggered events or hit sequences (for example minimum bias events that are created with a certain constant frequency without requiring any hit patterns) are stored and the rest of the data is dropped.

## 3.5 Signal and background events

ANTARES is designed for the detection of high-energy cosmic neutrinos. It measures Cherenkov-light signals, and is thus sensitive to all high-energy charged particles that cross the detector or pass near by. Regarding the cosmic neutrino events as signal, all other events are background. Nevertheless, these background events also provide interesting experimental opportunities. Labelling events as "signal" and "background" is a question of definition and is depending on the scientific intention. The present work focuses on low-energy atmospheric neutrinos (10 – 100 GeV) which, in the context of this thesis, are considered as signal, whereas possible high-energy neutrinos are part of the background.

In principle, physics events can be classified in upward-going and downward-going, depending on their zenith angle. Upward going events cross the detector from bottom to

top and their zenith angle is larger than 90°. Events are referred to as downward-going, when they cross the detector from top to bottom and their zenith angle is thus smaller than 90°. Upward going events can only be caused by neutrino induced particles, as neutrinos are the only particles, that can cross the Earth and enter the detection volume from the bottom. Downward going events, on the other hand, can be neutrino-induced particles as well as atmospheric muons from interactions of primary cosmic rays with the Earth's atmosphere. Although the flux of atmospheric muons is reduced by the water above the detector, it exceeds the flux of muons induced by neutrinos by several orders of magnitude.

The bulk of triggered events in ANTARES are therefore muons from three different sources: (downward-going) atmospheric muons, atmospheric muon neutrinos and cosmic muon neutrinos (Fig. 3.10). In Fig. 3.11, the flux of muons from atmospheric muon neutrinos and of atmospheric muons is shown as a function of the zenith angle $\Theta$. As already mentioned above, shower events from NC neutrino interactions or from CC $\nu_\mu$ or $\nu_\tau$ interactions only play a minor role and will not be discussed further.

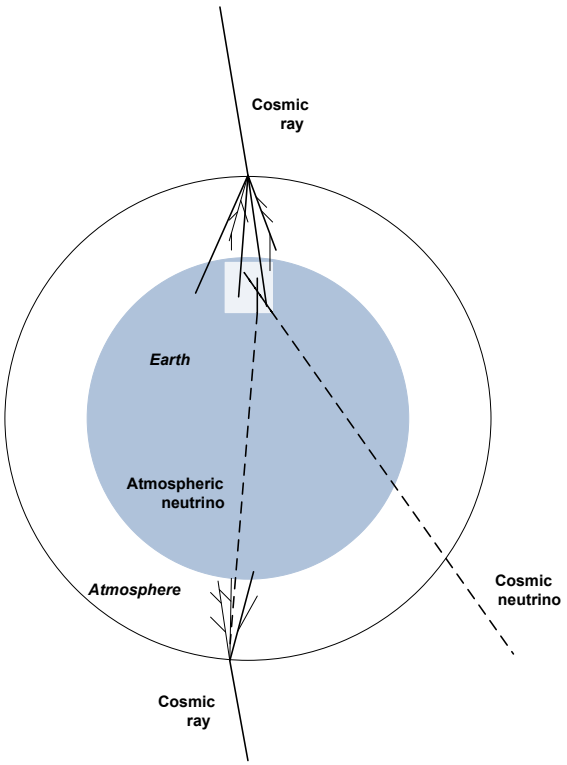



**Figure 3.10:** Scheme of the three muon event sources in ANTARES. Both cosmic and atmospheric (cosmic-ray induced) neutrinos can cross the Earth and produce detectable muons. Cosmic-ray induced muons can reach the detector only from above.

### 3.5.1 Atmospheric muons

Atmospheric muons dominate the event triggers in the ANTARES detector. In ANTARES it is not possible to distinguish atmospheric muons from downward-going neutrino events. Therefore, events coming from above are usually discarded. An exception are analyses that make use of atmospheric muons, like the moon shadow analysis,

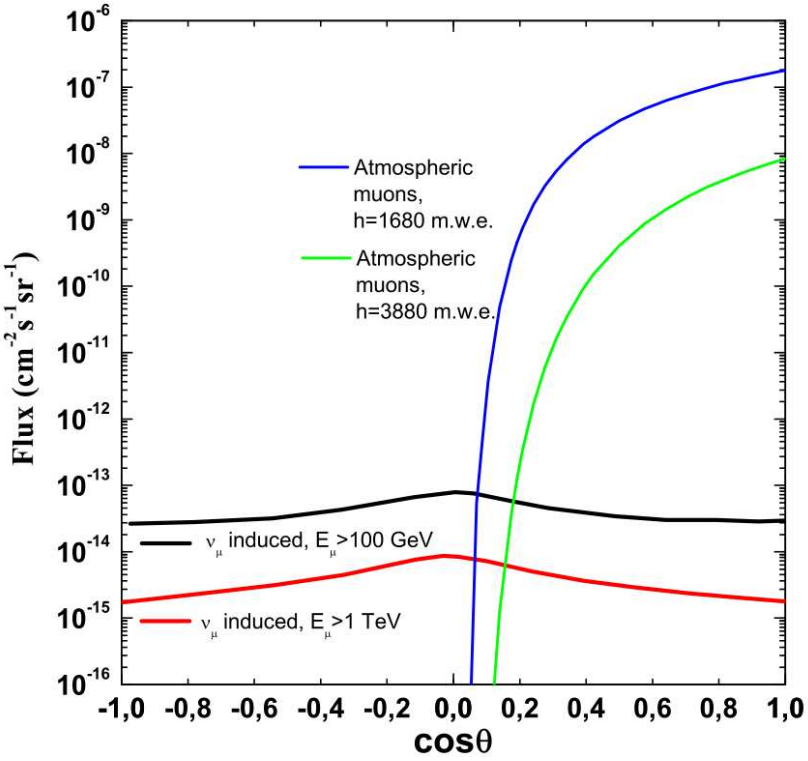



**Figure 3.11:** Fluxes of atmospheric muons and of muons generated by atmospheric neutrinos as a function of the cosine of the zenith angle. Figure from [89] (see also for references).

investigating the shielding effect of the moon on cosmic radiation in order to verify the absolute pointing of the detector [70].

Due to the downward orientation of the OMs, the light emitted by atmospheric muons is dominantly measured near the acceptance limit of the PMTs, and the fraction of scattered light is enhanced. This leads to a larger uncertainty in the reconstruction of the trajectories of atmospheric muons and a certain amount of them will be misreconstructed as upward-going particles. Additionally, atmospheric muons often appear in bundles, which means that several muons cross the detector in parallel. The detected light signals, produced by more than one particle, can fake upward-going neutrino events. As the muon flux is so much larger than the flux of upward-going neutrino-induced muons, even a small percentage of misreconstructed events leads to an unacceptable signal-to-noise ratio. Effective quality cuts are therefore needed to suppress this background.

### 3.5.2 Atmospheric neutrinos

Atmospheric neutrinos are the main neutrino signal (i.e. upward-going particles) in ANTARES. Like atmospheric muons, they also emerge from the interaction of primary

cosmic rays with the atmosphere. Their flux depends on the zenith angle $\Theta = [0°, 180°]$, and is almost symmetric around $\Theta = 90°$ (see Fig. 3.11). However, in ANTARES it is not possible to separate downward-going neutrinos from atmospheric background.

The sensitivity of ANTARES to atmospheric low-energy neutrinos is restricted by the spacing of the OMs and to a certain extent also by the optical background. Events with an energy smaller than 10 GeV are not detectable, because the number of hits measured by the detector (the signature in general) is typically not sufficient for the trigger algorithms. Even if they are triggered, the reconstruction will usually fail.

This thesis therefore deals with the investigation of oscillations of atmospheric neutrinos in the energy range of 10 – 100 GeV (Chapter 6). This is right at the low-energy end of the sensitivity range of ANTARES. The characteristics of such low-energy atmospheric neutrino events in ANTARES will be illustrated in detail in Chapter 4.

### 3.5.3 Cosmic neutrinos

Detecting cosmic neutrinos is what the ANTARES detector is built for. For the measurement of high-energy cosmic neutrinos, atmospheric neutrinos are forming a background, that is difficult to separate. It is not possible to identify individual neutrino events as cosmic neutrinos. Only an excess of neutrino events at a certain energy or from a certain celestial direction over the expected flux from atmospheric neutrinos signals a cosmic origin. No such cosmic high-energy neutrino signal was as yet found by any neutrino telescope.

## 3.6 Simulation

For understanding the detector response to the various processes, detailed Monte-Carlo simulations are necessary. The software chain used for the generation of such simulations consists of several steps and depends on the simulated particles. Resulting from the simulation chain is the detector response to a given input signal (i.e. charged particles) as a set of detected PMT hits, including the arrival time at the respective PMT. Taken into account are the PMT characteristics and photon attenuation due to the surrounding material. The simulation of the read-out electronics is done within an offline analysis framework (Section 3.7) and will be described in detail in Chapter 5, together with the simulation of optical background and unsound OMs.

### 3.6.1 Atmospheric muons

The fast simulation package MUPAGE (v3r4), [90, 91, 92] was used for the simulation of atmospheric muons. Instead of doing a full Monte-Carlo (MC) simulation of atmospheric showers, including the cosmic-ray interactions and the shower propagation in the atmosphere, this MC generator is based on a set of parametric formulae, generating atmospheric muon events directly at the detector.

The parametric formulae are obtained from simulations done with HEMAS [93, 94, 95] and MUSIC [96]. They give the radial distribution of the muons in the shower, as well as the energy spectrum and the flux of the muons in the range of vertical depths from 1.5 to 5 km water equivalent and zenith angles $\Theta < 85°$. HEMAS simulates the interaction of the primary cosmic rays and the propagation of particles in the atmospheric shower. The cosmic ray flux used as input to HEMAS is based on a phenomenological model combining results from various measurements of cosmic rays in the energy range from 10 GeV to 1 EeV. MUSIC propagates the muons from sea-level down to the detector. It takes into account the deflection and energy losses of the muons due to multiple and inelastic scattering, pair production, Bremsstrahlung and knock-on electrons.

Simulations done with MUPAGE do not require the detailed tracking of muons to the detector and are therefore much faster than full simulations. Also no relative event weighting is needed because the events are generated according to their abundance and each set of events corresponds to a certain scalable lifetime.

### 3.6.2 Neutrinos

Neutrino interactions are simulated only within a certain volume around the detector, determined from the maximal neutrino energy and the zenith angle, so that the longest ranging interaction products are able to reach the detector.

The GENHEN (v5r6) software package [97, 98] is used for the simulation of neutrino events, including deep-inelastic scattering using the LEPTO 6.5 package [99] and resonant and quasi-elastic interactions using RSQ [100]. For the parton distributions the functions from CTEQ6D [101, 102] are used and the hadronisation is done with PYTHIA 5.7 and JETSET 7.4 [103].

Each event is generated with a certain weight that needs to be considered when analysing the simulations. This weight includes the probability for a neutrino to reach the interaction point, calculated from the neutrino energy and the column density through the Earth, which is in turn associated with the neutrino direction. It also accounts for the neutrino flux and includes the energy dependent neutrino cross section. If the neutrino interaction is inside the detector region, all originating charged particles are passed to the Cherenkov photon simulation (next subsection), otherwise only the muon is propagated. The energy range covered with the simulations is $10 - 10^8$ GeV.

### 3.6.3 Photon generation

The simulation of Cherenkov light is common for all particles and is restricted to a certain volume around the detector, referred to as "can". The size of the can is determined by the absorption length of light in sea water $\Lambda_{\text{abs}}$ and exceeds the detector volume by $3 \cdot \Lambda_{\text{abs}}$.

Cherenkov light production and the response of the ANTARES detector is simulated by the KM3 (v3r6) package [104], which makes use of a modified version of MUSIC for the propagation of the muons inside the can. The Cherenkov photons are not tracked individually but their flux is calculated from tables, based on the parametrisation of

photon number and arrival time distributions of the emitted Cherenkov light at different distances, positions and orientations with respect to the muon track. The parametrisation are obtained from full simulations with the GEANT3 package [105]. Measured water properties at the ANTARES site are used to simulate absorption, scattering and dispersion of photons.

GEANT3 is also the basis for GEASIM (v4r10) [106] that simulates the propagation and light emission of other charged particles that are, for example, generated in the hadronic shower of neutrino interactions. GEASIM performs a full tracking of the particles simulating all relevant physics processes. The number of Cherenkov photons is calculated taking into account the attenuation of light but without considering the photon scattering. The arrival time of the Cherenkov photons incident on the OMs is calculated analytically for each of the charged particles.

The response of the OMs is simulated in the same way in GEASIM and KM3, calculating the number of photons eventually detected by the PMTs. This number depends strongly on the properties of the OMs including the angular acceptance of the PMT, the wavelength dependent detection efficiency and the absorption of photons in the surrounding material (glass and gel).

## 3.7 ANTARES software

The ANTARES software presently can be divided into three main parts: the DAQ software used for the data acquisition, the simulation software and programs for offline analyses.

The data acquisition software has been described in detail in Sections 3.2.2 and 3.4.

The simulation software, up to the individual Cherenkov photons seen by the OMs, has been discussed in the previous section. The simulated data in ANTARES is centrally generated and provided to the group members, basically at two different simulation levels: the raw simulated data up to OM level, including the PMT simulation as described above, is available in ASCII format. Additionally, simulations are available in ROOT format, already containing simulated optical background and filtered with the data filters (triggers) of the DAQ. For this thesis, an alternative approach was used for the simulation of the real detection conditions like optical background and defective OMs, which will be illustrated in detail in Chapter 5.

Contrary to the DAQ and the simulations, programs for offline analysis are typically designed individually by the group members, depending on their aims and purposes. As in large-scale experiments a lot of physicists are working together, most of them not being professional programmers, a special way of organising the analysis software is needed, to bundle efforts and to make program code easier to implement, to understand and to modify for others. All this is guaranteed by a modular structured software framework named *SeaTray*, that was introduced to ANTARES during 2008 [107, 108]. As most algorithms written in context of this thesis were implemented in this framework, its structure and program logic will be sketched in the following.

The framework consists of modules that operate on the data. So called services are

common tools that can be called by the modules for certain tasks. The data flow itself is organised in four separate, parallel streams: for the detector status, the geometry, the calibration and for physics events. New data in one of the streams, such as a new physics event, or a modification of the geometry due to new alignment data, induces the creation of a new data container, a so called frame. The detector status, the geometry and the calibration change on timescales (minutes to days) much larger than the usual physics event duration, meaning that one frame in general contains exactly the information of one physics event together with its detector status, calibration and geometry. The information in the frame is stored in so called frame objects. The frames are passed from module to module; the modules get the information they need to perform their operations from the frame and finally add new objects to the frame. Modules, frame objects and services have a predefined structure. Users can modify already existing algorithms or easily implement new program code in the prototype structures. The modules usually work independently from each other. Apart from a few default modules like the data I/O modules, they can be arranged and exchanged more or less arbitrarily, depending on the needs of the user. This allows for the flexible combination of different modules without changing the program code.

In the context of this thesis, several contributions to this framework for the general use in ANTARES were made, like the complete low-energy reconstruction program *Posidonia* (Chapter 4). Also all developments concerning data quality, data filtering and the adaption of real detector conditions to simulations (Chapter 5) have been implemented into *SeaTray*, as well as several useful tools like a module for the extraction and storage of Monte-Carlo information.

# 4 Low-energy reconstruction

The energy range relevant for the investigation of oscillations of atmospheric neutrinos with the ANTARES neutrino telescope is around 10 – 100 GeV. This is at the sensitivity limit of the detector at about 10 GeV and is consequently referred to as low-energy range. ANTARES is primarily dedicated to searches for high-energy cosmic neutrinos (starting from about 1 TeV), and so are the available reconstruction algorithms. Low-energy neutrino events need special attention because they typically produce only few hits in the detector, which, like all measured physics signals in ANTARES, are contaminated by optical background hits.

A special reconstruction program named *Posidonia*, optimised for low-energy neutrino events, will be introduced in this section. It was originally developed before the ANTARES detector design was finalised and settled [109, 110, 111]. In the context of this thesis, the program has been recovered and ported to the official ANTARES software framework *SeaTray*. It has been revised, tuned and adjusted to the final ANTARES layout, the hardware settings and the latest Monte Carlo simulations. Within the framework, the program was structured as modularly as possible to enable the easy exchange of individual parts of the reconstruction chain. The reconstruction procedure itself was enhanced and optimised in several ways, for example by applying new hit selections. As the main focus was on the quality and efficiency of the reconstruction, the program is not (yet) optimised concerning computing time. This shortcoming has to be compensated by a careful preselection of the data, which will be discussed in Section 5.4.

The first section of this chapter (4.1) introduces the simulations used for the characterisation of low-energy events and for development and optimisation of the reconstruction algorithm, as well as for the assessment of its performance. The physics of low-energy events and the detector response will be discussed in Section 4.2, then the reconstruction algorithm is illustrated in detail in Section 4.3. The performance of the reconstruction program for simulated low-energy neutrino events is demonstrated in Section 4.4. Also shown is a comparison with one of the two standard reconstruction algorithms in ANTARES, as well as the performance in case of an increased background rate and a degraded detector state. Finally, also the selection of low-energy events with a dedicated containment estimator and the energy reconstruction are assessed.

## 4.1 Simulation of neutrino events and detector conditions

The simulations used in this chapter only contain muon neutrinos reaching the detector from the lower hemisphere (i.e. upward-going, neutrino zenith angle $\Theta_\nu > 90°$). $50 \cdot 10^{10}$ muon neutrinos were simulated in the energy range from $10 - 10^7$ GeV, flat in $\cos\Theta_\nu$ and with an energy spectrum of $E^{-1.4}$. The neutrinos are reweighted to the atmospheric flux (Bartol flux model), assuming one year of data taking.

In the majority of cases only neutrino events with a neutrino energy up to $E_\nu = 200\,\text{GeV}$ or with a muon energy up to $E_\mu = 150\,\text{GeV}$ were used[15]. The number of simulated neutrino events with at least one hit in the detector in the energy range $10\,\text{GeV} < E_\nu < 200\,\text{GeV}$ is 237 690 (the weighted sum is 98 449).

The full simulation procedure up to the level of Cherenkov hits at the OMs has already been described in the last chapter. The electronics simulation is performed within the *SeaTray* framework and will be described in Chapter 5.

Random optical background is added to the simulated signal. Unless otherwise noted, for this analysis a background rate of 60 kHz per OM was added. The detector geometry used is the one of the ideal detector: twelve strings with a total of 885 active OMs (75 OMs on lines 1 – 11, and 60 OMs on line 12), all working and perfectly aligned. This neutrino event sample will later also be employed in the oscillation analysis and will be referred to as "ideal reference sample" (IRS), as it is a measure of what could be obtained under perfect detection conditions and assuming the ability to select a pure neutrino sample.

If triggered events are considered, the events had to pass the ANTARES trigger simulation module, with both the *3N* and the *2T3* triggers active. The module emulates the behaviour of the online data filter of the ANTARES DAQ. The different behaviour of the two trigger algorithms, *3N* and *2T3*, is demonstrated in the following sections. However, it becomes also evident when comparing the fraction of *3N* triggered events to the fraction of *2T3* triggered events: about 99 % of all triggered events are triggered by the *2T3* trigger and only 30 % by the *3N* trigger (events could be selected by both of the triggers and the *3N* trigger is almost a subset of the *2T3* trigger). So generally, the *2T3* trigger is much more efficient concerning low-energy muons. Consequently, the reconstruction quality of the *Posidonia* reconstruction algorithm was evaluated separately for *3N* and *2T3* triggered events (Section 4.4). The weighted sum of triggered low-energy (anti-)neutrinos[16] events within one year of data taking ($E_\nu < 200\,\text{GeV}$,

[15] Although only neutrinos with $E_\nu > 10\,\text{GeV}$ are simulated, the generated muons can have energies $E_\mu < 10\,\text{GeV}$, as only a certain fraction of the neutrino energy is transferred to the muon (Section 4.2).

[16] At this point, the anti-neutrinos were accounted for by simply multiplying the total number of events with a factor of two. In fact, anti-neutrinos have a smaller cross-section than neutrinos within the low-energy range: between 10 and 200 GeV, the ratio of $\nu_\mu/\bar{\nu}_\mu$ is 1 – 2. However, this procedure was considered as acceptable at this point, as only the performance of the reconstruction algorithm was evaluated, which is not sensitive to the charge of the muon. The given absolute event numbers within this chapter are only upper estimates, that are intended to be considered relative to each

60 kHz optical background and ideal detector) is 17 064, corresponding to a number of 31 333 simulated events (both triggers combined).

## 4.2 Low energy events and their signature in ANTARES

As discussed in Section 3.1.2, the angular correlation between muon neutrino and muon decreases with decreasing energy. For muons from low-energy neutrinos with at least five hits from the muon trajectory seen by ANTARES ("muon hits"), the mean scattering angle ranges from about 4° at 10 GeV to 1° – 2° between 50 and 200 GeV (Fig. 4.1). The mean energy transfer from neutrino to muon ranges from 60 % to 80 %.

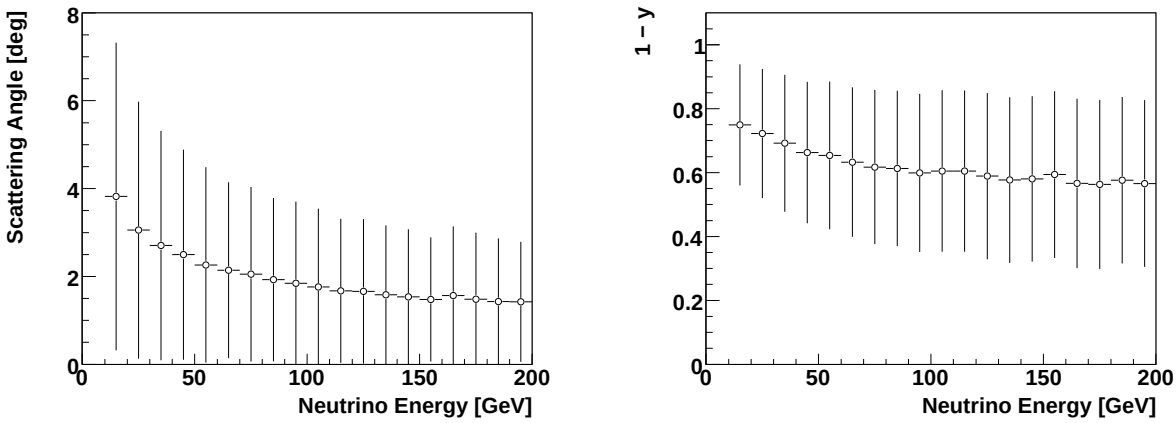



**Figure 4.1:** *Left:* Mean scattering angle between muon neutrino and muon (true simulated angles) as a function of the neutrino energy.
*Right:* Fraction of energy transferred from the muon neutrino to the muon (true simulated energy) as a function of the neutrino energy.
The bars indicate the standard deviations. Only events with at least five detected hits from the muon are considered.

In this low-energy range, muons are minimally ionising particles (MIP). Furthermore, the amount of light produced along the track is much smaller than at higher energies where stochastic energy-loss processes of the muon lead to the emission of additional light. Hence, the energy of the muon is linearly correlated with its track length (Fig. 4.2) and, consequently, the energy of a low-energy muon can be inferred as soon as the range of the muon is reconstructed reliably. The length of the muon trajectory at low energies is of the order of tens to some hundred meters. For a detector with a sensor spacing of 14.5 m (vertical) and 60 – 80 m (horizontal), the reconstruction of low-energy muon tracks is thus very challenging.

other, to evaluate the efficiency loss due to a varying background rate or detector condition.

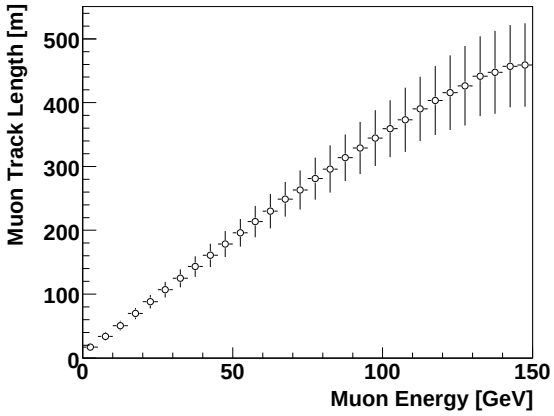



**Figure 4.2:** Mean length of the muon trajectory as a function of the muon energy (true simulated length and energy). The bars indicate the standard deviations. Only events with at least five detected hits from the muon are considered.

Track reconstruction in ANTARES is additionally complicated by light signals due to optical background. The mean number of hits coming from the muon as a function of the energy of the muon is shown in Fig. 4.3. Under favourable environmental conditions, the optical background in the deep sea at the ANTARES site corresponds to a typical rate of 60 kHz per single PMT. For a typical physics event of about 5 $\mu$s length, this results in about 260 single photoelectron hits due to random background noise, which exceeds the mean number of signal hits by far. Consequently, hit filters that extract the signal hits out of the optical background are indispensable. In principle, five hits are sufficient for the track reconstruction, though in the *Posidonia* reconstruction algorithm at least six hits are required for quality reasons.

The non-uniformity of the detector, i.e. the different spacing of the OMs in horizontal and vertical direction, poses complicated problems that have to be addressed when reconstructing low-energy muon tracks. The muon hits can either be distributed along a single line or spread out over different lines, inducing two different topologies of hit patterns. This gives rise to two different classes of low-energy muon events that are consequently referred to as *1D* or single-string events and accordingly *3D* or multi-string events. All analyses based on low-energy muons have to deal with these two distinct classes with their different characteristics and systematics.

From Fig. 4.4 it becomes obvious that low-energy single-string events ($E_\nu < 200\,\mathrm{GeV}$) tend to be "more vertical" and have significantly lower energies than multi-string events. Overall, only about 17 % of all events with at least five muon hits in the detector are single-string events. The exact definition and classification of single- or multi-string events applied for the reconstruction with *Posidonia* will be given within the following section.

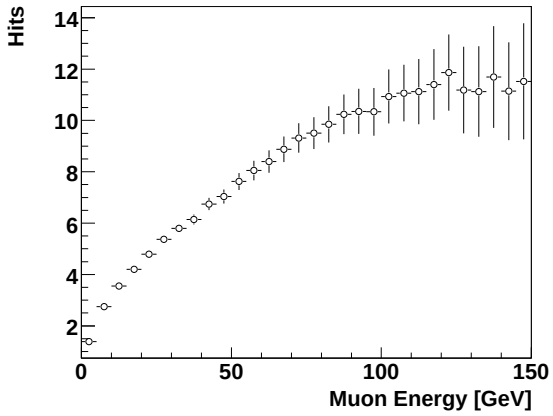



**Figure 4.3:** Mean number of detected hits from the muon trajectory as a function of the (true simulated) muon energy. The bars indicate the error on the mean. Only events with a neutrino energy $10\,\mathrm{GeV} < E_\nu < 200\,\mathrm{GeV}$ are used.

## 4.3 The reconstruction algorithm

The complete reconstruction algorithm, optimised for an analysis in the low-energy regime, basically consists of an independent hit pre-selection and the *Posidonia* reconstruction package, which itself consists of several modules. A schematic view of the *Posidonia* reconstruction chain, including the hit pre-selection is depicted in Fig. 4.5. A detailed description of the reconstruction algorithm will be given in this section.

### 4.3.1 Hit pre-selection

Hit filters are typically assessed by means of their efficiency and their purity. If $s$ ($b$) is the number of signal (background) hits contained in an event, $s_{\mathrm{acc}}$ ($b_{\mathrm{acc}}$) the number of signal (background) hits accepted by the hit filter and $s_{\mathrm{dis}}$ ($b_{\mathrm{dis}}$) the number of signal (background) hits discarded by the filter, the efficiency $R_{\mathrm{eff}}$ of a hit filter is defined as

$$R_{\mathrm{eff}} = \frac{s_{\mathrm{acc}}}{s_{\mathrm{acc}} + s_{\mathrm{dis}}}, \tag{4.1}$$

and the purity $R_{\mathrm{purity}}$ as

$$R_{\mathrm{purity}} = \frac{s_{\mathrm{acc}}}{s_{\mathrm{acc}} + b_{\mathrm{acc}}}. \tag{4.2}$$

Since low-energy events have numbers of signal hits close to the minimum number of hits required to reconstruct an event, the efficiency of a hit selection plays a major role in defining the efficiency of the reconstruction itself. Consequently, before entering the *Posidonia* reconstruction chain, the events are passed to an elaborate hit selection

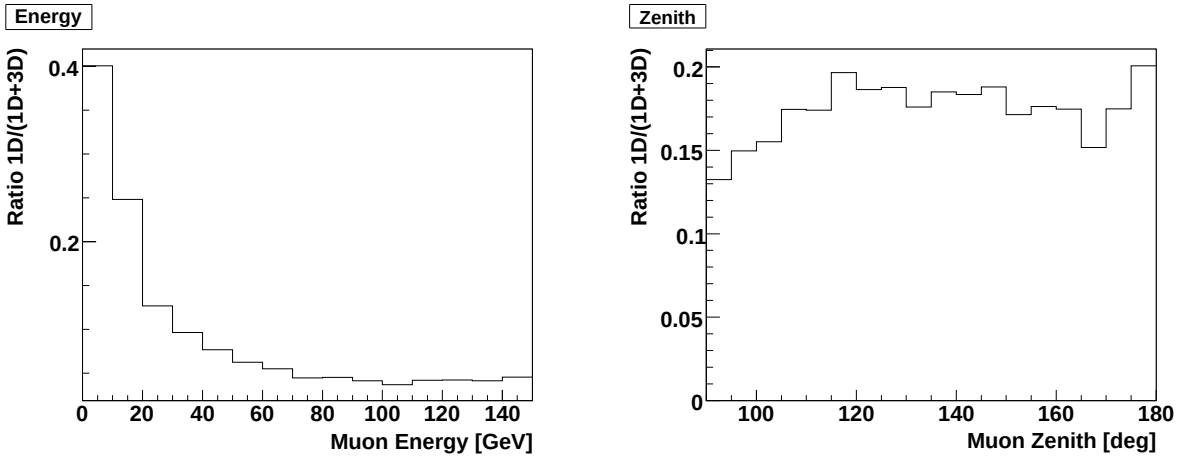



**Figure 4.4:** Ratio of the single-string events to all events as a function of the true muon energy (*left*) and the true zenith of the muon (*right*). Only events with at least five detected hits from the muon were considered. The simulation incorporates neutrinos up to $E_\nu = 200\,\mathrm{GeV}$. No optical noise is included.

algorithm to reject a significant part of the background hits while keeping as many signal hits as possible.

This so-called *HM* hit selection looks for hits close in distance and time and computes the slopes of their connecting lines in the $z$-$t$ diagram (where $z$ is the $z$-coordinate of the hit position $\vec{p}$ and $t$ its time). If the connecting lines of different pairs have related slopes, the pairs are grouped. A detailed description of this hit selection can be found in [112].

The hits remaining after the *HM* hit selection are merged with the so-called *L2* hits, that have a very high purity. In the context of this thesis the name *L2* refers to those hits that are selected by at least one of the two triggers, *3N* and *2T3*, while both triggers are active. In this context the triggers are only used for selecting hits and not as data filters. The trigger algorithms are not only part of the DAQ, but also of the *SeaTray* analysis framework and thus can also be applied offline on real and Monte Carlo data.

In Fig. 4.6 the distributions of efficiency and purity of the *HM*, the *L2* and the combined *HM*/*L2* hit selection are shown. By merging *L2* and *HM* hits, at the expense of a small degradation of the purity, an increase in the efficiency can be obtained.

In Fig. 4.7, the mean efficiency of the merged *HM*/*L2* hit selection compared to the *L1* selection (i.e. the hits selected by the *L1* trigger) is given as a function of the number of hits from the muon (the same simulated data sample was used). Over the whole range, the efficiency could be increased by at least 10 % by applying the *HM*/*L2* hit selection, instead of using the default *L1* selection.

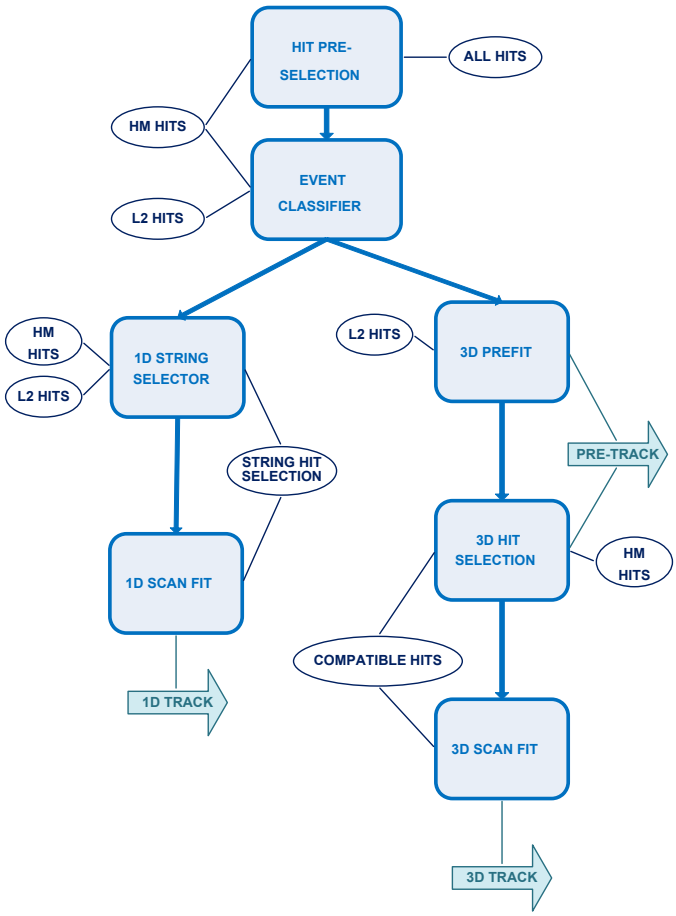



**Figure 4.5:** Scheme of the *Posidonia* reconstruction algorithm as implemented in the *SeaTray* software framework. Besides the modules (blue boxes), the relevant hit collections (dark blue oval shapes) and tracks (arrow shapes) are indicated.

### 4.3.2 *Posidonia:* Event classifier

The first step of the *Posidonia* reconstruction chain is an event classifier, separating the detected events into single-string and multi-string types. According to the classification result, the events are passed to the single-string or multi-string reconstruction module. The classifier is based on the *L2* hits: if there are *L2* hits on more than one string, the event is classified as multi-string event, as single-string event otherwise.

Figure 4.8 shows, in comparison to Fig. 4.4, how the trigger affects the percentage of single-string events. Obviously, both triggers increase the percentage of single-string events. The overall ratio of single- to multi-string events for $E_\nu < 200\,\mathrm{GeV}$ (no optical background) has a mean value of about 25 % for the *3N* trigger and a mean value of about 75 % for the *2T3* trigger (compared to 17 % without trigger, but requiring five muon hits in the detector). Clearly, the *2T3* trigger is more sensitive to single-string

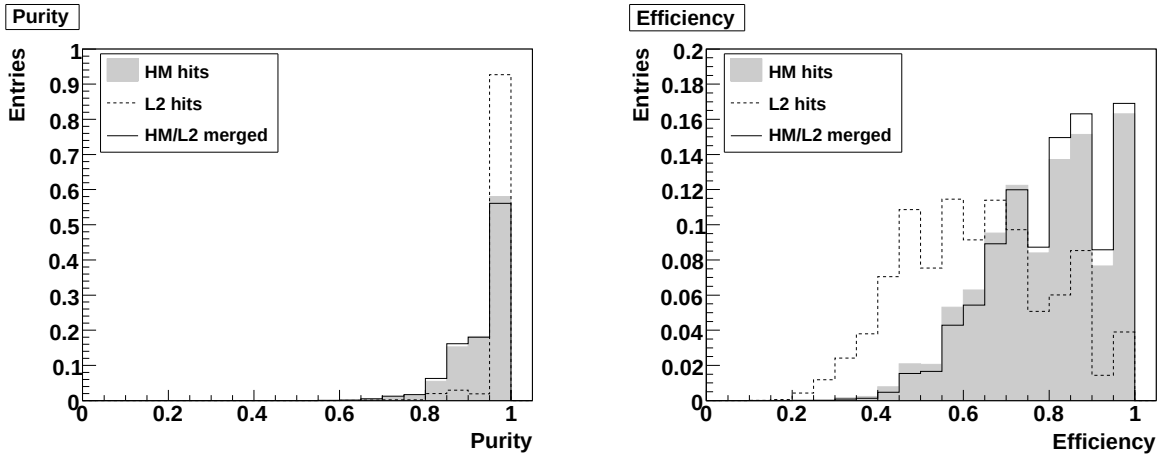



**Figure 4.6:** Distributions of purity (*left*) and efficiency (*right*) of the *HM* hit selection, the *L2* selection and the combination of both. 60 kHz random optical background per OM is added to the simulations, and only triggered events with a muon energy below 150 GeV are considered. The distinct peaks in the distribution of the efficiency are not due to low statistics, but to the small number of signal hits in the low-energy events.

events and thus more sensitive to events below 50 GeV. For both trigger types the percentage of single-string events decreases with increasing muon energy; especially the *3N* trigger enhances the percentage of single-string events above $\Theta = 150°$.

Figures 4.9 and 4.10 show that single-string events tend to be more vertical and have a lower energy than multi-string events. However, even though both the *3N* and the *2T3* triggers preserve the significant energy difference between the single-string and the multi-string events, the angular difference is largely washed out by the *2T3* trigger. It can therefore be concluded that the event characteristics of *3N* and *2T3* triggered events are significantly different, and thus they will be treated separately for the assessment of the reconstruction algorithm.

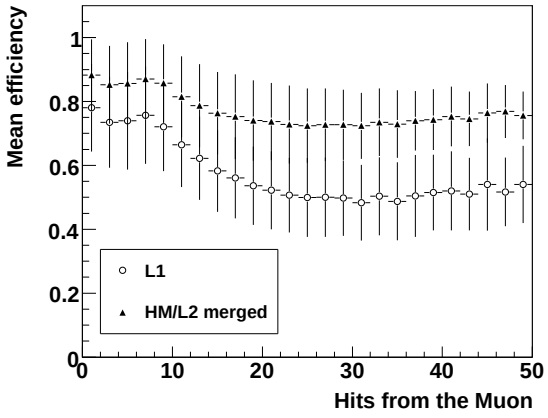



**Figure 4.7:** Mean efficiency of the *L1* hit selection compared to the merged *HM/L2* selection as a function of the muon hits. 60 kHz random optical background per OM is added to the simulations, and only triggered events with a muon energy below 150 GeV are considered.

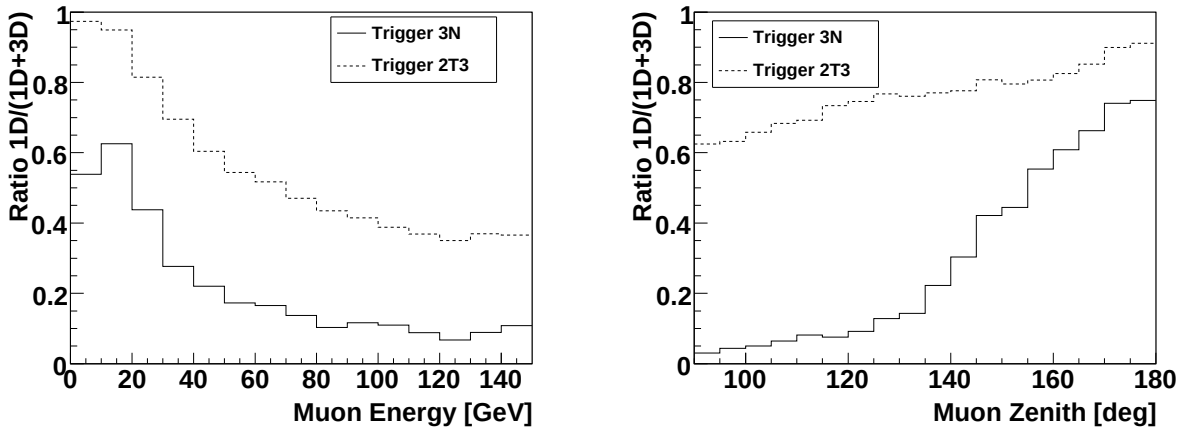



**Figure 4.8:** Ratio of single-string events to all triggered events (*3N* and *2T3* trigger separately) as a function of the true muon energy (*left*) and of the true muon zenith (*right*). No optical noise was added to the simulated signal. Only neutrinos with $E_\nu < 200\,\mathrm{GeV}$ were used.

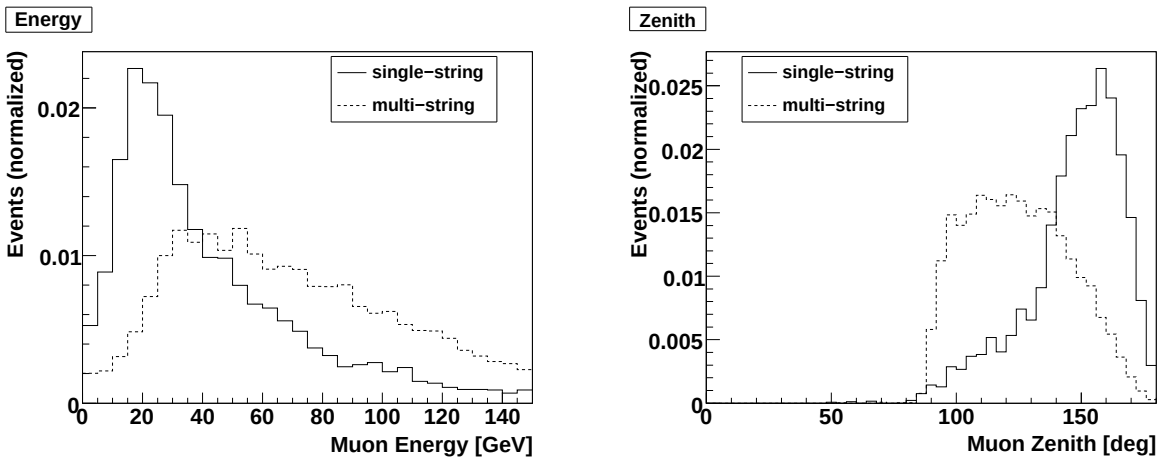



**Figure 4.9:** Distributions of energy and zenith for single- and multi-string events of *3N* triggered events. No optical noise was added to the simulated signal.

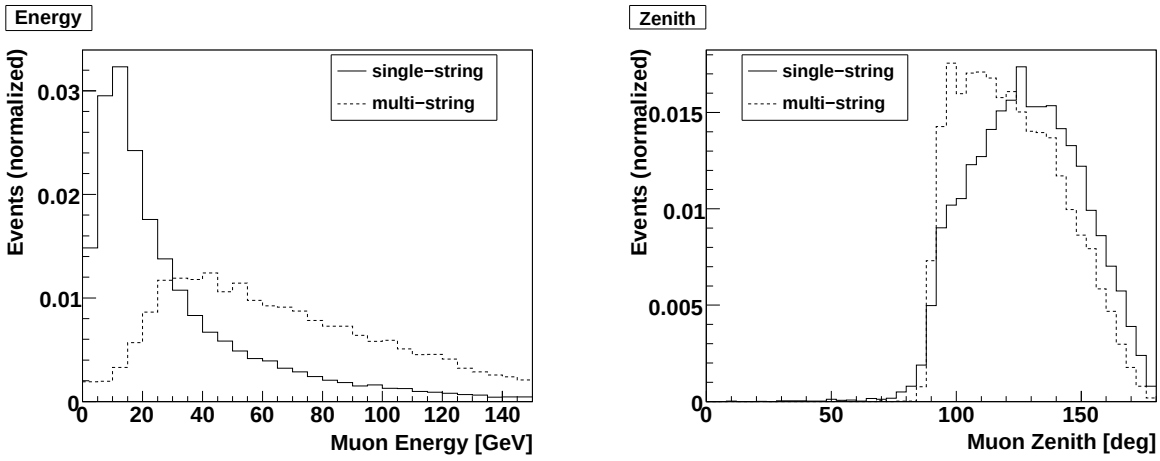



**Figure 4.10:** Distributions of energy and zenith for single- and multi-string events of *2T3* triggered events. No optical noise was added to the simulated signal.

### 4.3.3 *Posidonia:* Track reconstruction

In general, a muon trajectory is fully constrained by five parameters. The direction is given in terms of the zenith angle $\Theta$ and the azimuth angle $\Phi$. By defining a plane perpendicular to the track and through the origin of the coordinate system, the track is finally determined by the time $t_0$ when the muon crosses the plane and the position $\vec{p}$ at $t_0$, which is the point of minimal approach of the track to the origin. Rotating the axes of the coordinate system by $\Theta$ and $\Phi$ (the zenith and azimuth angle of the track), the $z$-coordinate of the plane transforms to $z' = 0$ and $\vec{p}$ can be given in terms of the independent variables $x'$ and $y'$. Therefore, five hits are necessary to completely reconstruct a track, but for numerical and quality reasons this threshold is increased to six. Triggered events have always a sufficient number of hits, albeit sometimes only due to the additional contribution of background hits. In the low-energy regime it happens that after all hit filters an event no longer fulfils this condition and is therefore lost, thus reducing the efficiency of the reconstruction algorithm.

Both the single and the multi-string reconstruction of *Posidonia* are based on a maximum likelihood fit, during which the track parameters are varied until the maximum of the likelihood function is found. The likelihood function,

$$\mathcal{L} = \prod_{i=1}^{n} f_i(\Delta t_i; \mathcal{T}), \tag{4.3}$$

is the product of the probability distribution functions (PDF) $f_i$ for the $n$ hits selected for the reconstruction. For a given track $\mathcal{T}$, $f_i$ can be expressed as a function of the time residuals $\Delta t_i = t_{i,\text{calc.}} - t_{i,\text{meas.}}$, which give the time delay between the calculated arrival time of a Cherenkov photon emerging from the assumed muon track and the measured arrival time. The track $\mathcal{T}$ itself is described by the five track parameters: the time $t_0$, the position $(x', y')$ and the two angles $\Theta$ and $\Phi$. In a first approximation, the PDF decouples into two functions of the track and propagation probabilities, $f_{i1}$, and the acceptance of the photomultiplier, $f_{i2}$, which is a function only of the angle of incidence of the photon onto the photomultiplier:

$$f_i(\Delta t_i; \mathcal{T}) = f_{i1} \cdot f_{i2}. \tag{4.4}$$

The parametrisation of $f_{i1}$ was newly adjusted in the context of implementing the reconstruction algorithm into the *SeaTray* framework. It is given as

$$\begin{aligned} f_{i1} = \frac{1}{a_{\text{norm}}} \cdot \Bigg[ & a_{\text{Gauss}} \cdot \exp\left(\frac{t_{\text{diff}}^2}{2a_\sigma^2}\right) \\ & + a_{\text{late\_factor}}^2 \cdot \frac{(1 + \tanh(t_{\text{diff}}))}{2} \cdot \exp\left(\frac{-t_{\text{diff}}}{a_{\text{late\_exp}}}\right) \cdot \frac{1}{\sqrt{1 + \frac{t_{\text{diff}}^2}{a_\sigma^2}}} \\ & + a_{\text{early\_factor}}^2 \cdot \left(1 - \frac{(1 + \tanh(t_{\text{diff}}))}{2}\right) \cdot \exp\left(\frac{t_{\text{diff}}}{a_{\text{early\_exp}}}\right) \cdot \frac{1}{\sqrt{1 + \frac{t_{\text{diff}}^2}{a_\sigma^2}}} \Bigg], \end{aligned} \tag{4.5}$$

with $t_{\text{diff}} = \Delta t - a_{\text{offset}}$, the time residual $\Delta t = t_{\text{calc.}} - t_{\text{meas.}}$, the normalisation factor $a_{\text{norm}}$, and the seven fit parameters $a_{\text{Gauss}}, a_{\text{offset}}, a_{\sigma}, a_{\text{late\_factor}}, a_{\text{late\_exp}}, a_{\text{early\_factor}}$, and $a_{\text{early\_exp}}$, which are obtained by fitting the distribution of time residuals $\Delta t$.

In Fig. 4.11, the distributions of the time residuals obtained from simulations, and the functions $f_{i1}$ finally implemented in the *Posidonia* fitting algorithm are displayed. For single- and multi-string reconstruction, all the respective hit filters were applied, resulting in two slightly different $f_{i1}$ distributions for the two cases. The values for the parameters and the normalisation factor can be seen in Table 4.1.

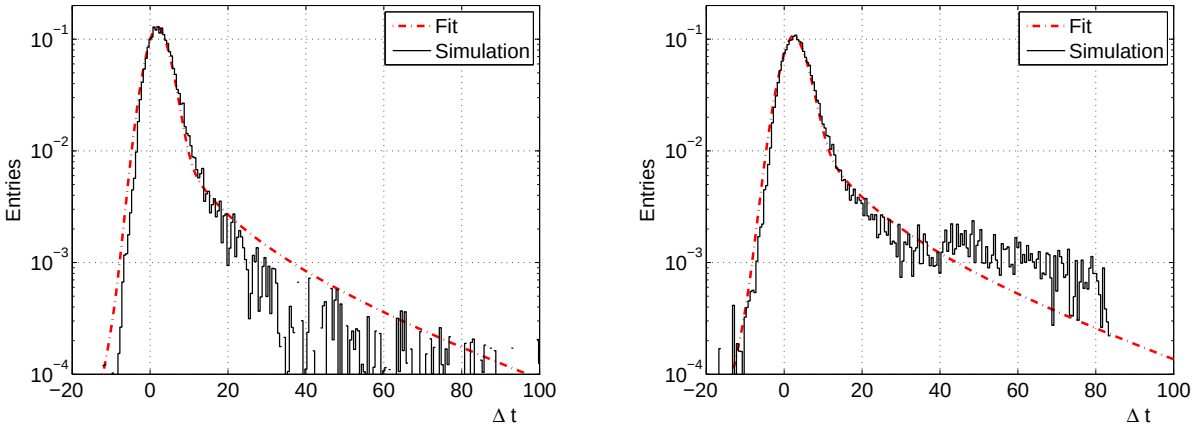



**Figure 4.11:** Distribution of time residuals after all hit filters for single- (*left*) and multi-string events (*right*). The histogram shows the distributions from the simulations, the red line is the resulting parametrisation.

**Table 4.1:** Parametrisation of $f_{i1}$ for single- and multi-string events.

| Parameter | Single-string | Multi-string |
|---|---|---|
| $a_{\text{offset}}$ | 1.9126 | 2.3652 |
| $a_{\text{Gauss}}$ | 0.0969 | 0.0741 |
| $a_{\sigma}$ | 2.9613 | 3.2843 |
| $a_{\text{late\_factor}}$ | -0.1553 | -0.1742 |
| $a_{\text{late\_exp}}$ | 46.7905 | 48.4724 |
| $a_{\text{early\_factor}}$ | -0.2092 | -0.1950 |
| $a_{\text{early\_exp}}$ | 3.0836 | 3.5923 |
| $a_{\text{norm}}$ | 1.0352 | 1.0011 |

For calculating the distribution of the time residuals, only hits from simulated upward-going neutrino events (atmospheric flux) with an energy less than 250 GeV were used, and no optical background was applied. Due to the permanently fluctuating background rates, it is technically almost impossible to implement background in the PDF in a reasonable way. However, because of the good purity of the merged *HM/L2* hit selections,

the neglicence of background in the PDFs can be justified. The deviations between fitted curves and distributions of the time residuals are predominantly in the tails of the distribution and their impact on the reconstruction quality is likely smaller than the neglicence of background. However, the agreement could probably be improved with a different parametrisation.

For practical reasons usually instead of maximising the likelihood $\mathcal{L}$, the negative logarithm of the likelihood is minimised by applying an appropriate minimisation procedure.

The minimiser that is used in *Posidonia* is part of the commercial NAG library [113]. Especially at low energies, where the information is sparse, it requires a reasonable starting point for convergence, since the likelihood function may exhibit several local minima and therefore different starting values for the minimisation procedure may result in different solutions. However, as long as the same hits are used in the fit, those different solutions can be compared using their absolute likelihood value. To account for this behaviour of the minimiser, a five-parameter space was scanned to determine a set of starting points for the minimisation procedure. For this purpose, the zenith $\Theta = [0, \pi]$, the azimuth $\Phi = [0, 2\pi]$ and the horizontal distance from the centre of gravity of the hits $d = [0\,\mathrm{m}, 44\,\mathrm{m}]$ are sampled uniformly in equidistant steps. In total, for each fit 520 starting points are determined. A track is fitted for each such starting point and the fitted track with the smallest log-likelihood value is considered to be the best solution.

This "scan-fit" procedure was already employed in the original *Posidonia* program, but only for the single-string fit. In the context of this thesis, it has also been introduced to the multi-string fit procedure.

### Multi-string reconstruction

In the multi-string reconstruction chain of *Posidonia* the scan-fit procedure described above is preceded by a double-stage hit filter, which consists of a linear prefit and a successive selection of hits compatible with this prefit. The linear prefit is a simple $\chi^2$-fit and is adapted from [114].

It is assumed that all $n$ hits with their positions $\vec{p}_i$ at times $t_i$ are located on the track. Then

$$\vec{p}_i \backsimeq \vec{p} + \vec{v} \cdot t_i, \tag{4.6}$$

where $\vec{p}$ is the position of the track at $t = 0$ and $\vec{v}$ is the velocity vector of the track. The parameters $\vec{p}$ and $\vec{v}$ can be obtained by minimising the quadratic sum

$$\chi^2 = \sum_{i=1}^{n} a_i \frac{(\vec{p}_i - \vec{p} - \vec{v} \cdot t_i)^2}{(\Delta p_i^2 + |\vec{v}|^2 \Delta t_i^2)}. \tag{4.7}$$

To account for the fact that the measured charge is the larger the closer the track passes by the optical module, the $\chi^2$ terms are weighted with the charge $a_i$ of the corresponding hits. As the error $(\Delta p_i^2 + |\vec{v}|^2 \Delta t_i^2)$ is about the same for all hits it can be neglected in this case. The minimisation of the $\chi^2$ is done analytically.

The prefit track only serves as reference track for the following hit selection. Within the preselected hits, all hits are selected with a maximum photon path length from the pre-fitted track of 120 m and time residuals $\Delta t = t_{\text{calc.}} - t_{\text{meas.}}$ in an interval of [-100 ns, 80 ns], where $t_{\text{calc.}}$ is the expected arrival time of the photon assuming the pre-fitted track is the true track (neglecting scattering and attenuation) and $t_{\text{meas.}}$ is the experimentally measured arrival time. In Fig. 4.12 the time residual distributions with respect to the pre-fitted track for signal and background hits can be seen. Their suitability as cut variable becomes evident from the different shapes of the distributions for signal and background hits.

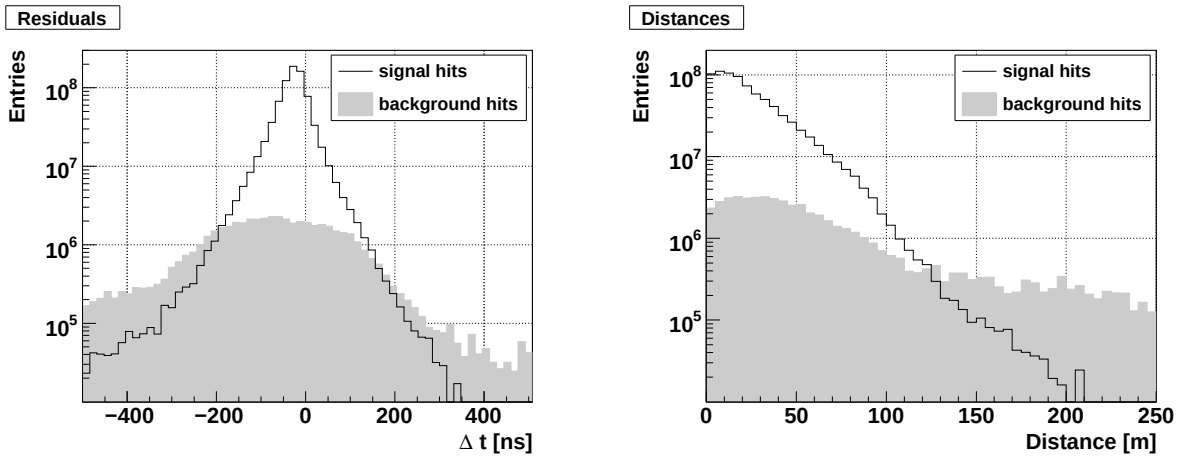



**Figure 4.12:** Distributions of the time residuals $\Delta t$ (*left*) and the photon path length $l$ (*right*) of signal (solid line) and background hits (shaded area) with respect to the pre-fitted track. Without the preceding *HM* hit selection, the time residuals of the background hits would be distributed uniformly over the complete range. The simulation incorporates neutrinos up to $E_\nu = 250\,\text{GeV}$, and includes 60 kHz optical background.

### Single-string reconstruction

The essential difference between the single- and multi-string reconstruction is the hit selection procedure. The main scan fit in the end is basically the same for both reconstruction chains except for the PDF used in the fit.

In order to get a high-quality selection of hits for the single-string fit, the hits have to pass a special single-string hit selection. Like for the multi-string events, the hits have to pass the basic hit selection first. The remaining hits are separated into two groups: coincident hits (according to the coincidence condition for *L1* hits) and the remaining ones, so-called single hits. Those two hit collections are passed to the main causality filter. Each coincident hit is tested, if it is causally connected to all other coincident

hits. This basic causality condition that has to be met is

$$\frac{|\Delta z|}{c_n} + T_{\text{add}} > |\Delta t|, \tag{4.8}$$

with the group velocity of light in sea water $c_n$, the absolute difference of the measured hit times $|\Delta t|$, the absolute difference in distance between the two hits $|\Delta z|$, and an allowance of $T_{\text{add}}$, which is set to 20 ns.

To improve the selection, two additional causality conditions are introduced. They affect hits that are at least four storeys apart and are increasingly restrictive as the number of storeys between the hits is increasing.

The first condition requires a minimum time difference $\Delta t$ for the hits:

$$\frac{\Delta z}{c} - T'_{\text{add}} < \Delta t, \tag{4.9}$$

with the velocity of light in vacuum $c$, which is approximately the velocity of the muon track and $T'_{\text{add}} = 150\,\text{ns}$.

The second additional condition is a slight modification of the basic causality condition. Again the group velocity of light in water is replaced by the velocity of the muon track $c$, and the allowance factor is again $T'_{\text{add}} = 150\,\text{ns}$:

$$\frac{\Delta z}{c} + T'_{\text{add}} < \Delta t, \tag{4.10}$$

This modified condition is more restrictive than the basic causality condition, if the hits are more than four storeys apart. It accounts for the fact, that the Cherenkov light emitted by the muon is attenuated in the sea water ($\Lambda_{\text{att}}^{\text{eff}} \approx 45\,\text{m}$), and the probability of hits being causally connected with the track is small if the calculated photon path length significantly exceeds the attenuation length. Hence, hits on one string, being some storeys apart, can only result from rather vertical tracks (resulting in a constant photon path length over the length of the string) and thus the arrival times of these hits are connected via the velocity of the muon track.

### 4.3.4 *Posidonia:* Containment estimation and energy reconstruction

For the energy reconstruction, different techniques are used, depending on the energy of the muon. The energy loss of muons above 1 TeV is dominated by stochastic radiative processes and elaborate techniques for reconstructing the energy are necessary. Contrary to this, muons within the energy range of interest for an oscillation analysis are minimally ionising particles: their energy loss is almost constant over the complete track length and the energy is linearly correlated with the track length (see also Fig. 4.2). Thus, having a reliable muon track length reconstruction, the energy of the muon can be inferred.

On the other hand, a reasonable and reliable energy estimation is only possible for tracks that are completely contained within the instrumented detector volume. Only

for muons below about 100 GeV this requirement is satisfiable, as the muon track length is about 300 m at this energy.

Selecting such "contained events" from the upward-going neutrino sample therefore automatically implies selecting low-energy events. Again, single- and multi-string events have to be treated separately due to their different topologies. Furthermore, also the definition of the containment volume is different for the two cases.

### Multi-string events

For estimating the containment of an event, the starting point and end point of the track have to be determined. This is done by calculating the point of photon emission with respect to the reconstructed muon trajectory, for every hit of the hit sample used for the fit. Only tracks reconstructed as upward-going are considered at this level, so the point of emission with the smallest $z$-coordinate is assumed to be the starting point of the track and the point with the largest $z$-coordinate is assumed to be the end point.

Then, the expected number of photons is calculated that would have been detected if the track would exceed the so-determined length. Assuming an infinite track length of the muon, for each OM the corresponding point of photon emission along the track is calculated (Fig. 4.13). If this point of emission is beyond the previously calculated track limits, the amount of light (i.e. the number of photons) is estimated, that should have been detected by the given OM, considering its angular acceptance, the distance from the track to the module and the attenuation of light in water. The sum over all OMs finally gives an estimate on the probabilities that the starting or end point, respectively, are located within the instrumented detector volume. The photon distributions for contained and non-contained events, are shown below (Section 4.4.6), where the performance of the reconstruction algorithm, including the containment estimation and energy reconstruction, is discussed. The larger the amount of light that should have been detected, the more likely the track starting or end point is contained. If both the starting and the end point are contained, the event is fully contained in the detector and the track length finally is determined by the difference of the track limits [110].

### Single-string events

Since the single-string tracks tend to be quite vertical, the $z$-positions of the starting and of the end point of the track (which are calculated in the same way as for multi-string events) serve as separation variables. Before concluding on the containment, a cut has to be applied to select only vertical events (e.g. $\Theta_{\text{rec, Posidonia}} > 135°$).

The smaller the absolute $z$-values of the reconstructed starting and end point of the track are, the more likely the track is contained (Fig. 4.14). The $z$-distributions are shown below (Section 4.4.6).

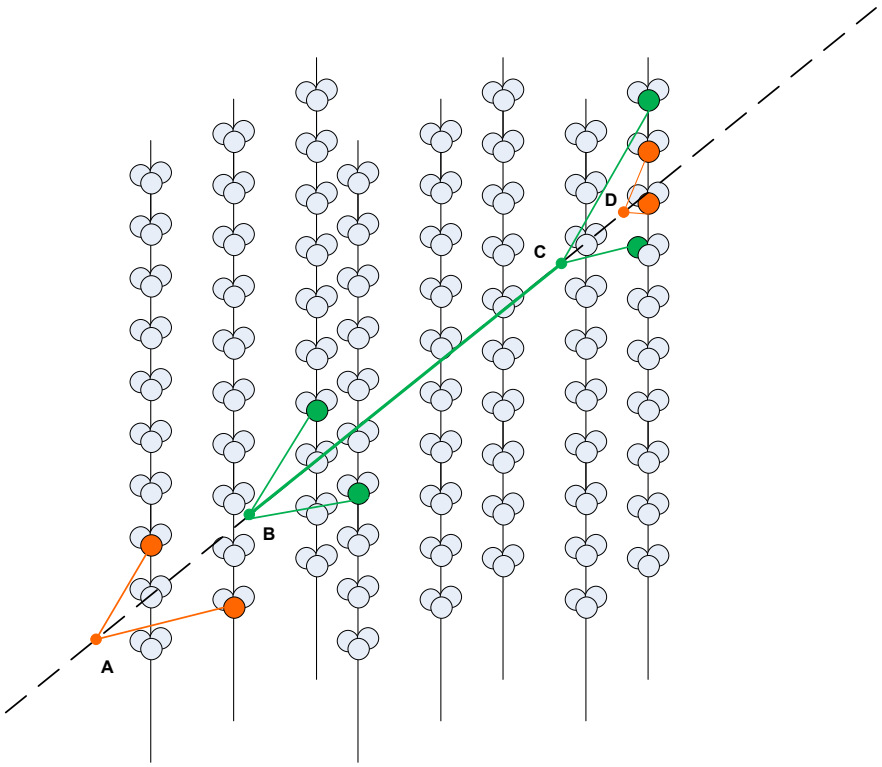



**Figure 4.13:** Scheme of the containment estimator for multi-string events. The starting (B) and stopping point (C) of the track are calculated by using information from OMs with hits (green OMs). If the track exceeds the calculated length, light emitted at points (A) or (D) should have been detected by the marked OMs (orange).

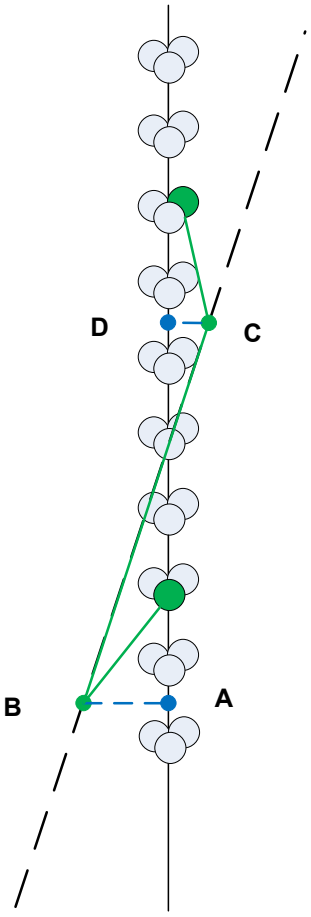



**Figure 4.14:** Scheme of the containment estimator for single-string events. The starting and stopping points (B) and (C), respectively, are calculated in the same way as for multi-string events. The projections of the starting and stopping point on the $z$-axis, (A) and (D), are used to estimate the containment of the track.

## 4.4 Performance of the *Posidonia* reconstruction algorithm

In this section, the performance of the *Posidonia* reconstruction algorithm will be discussed by means of signal events from upward-going low-energy neutrinos. To be able to better judge the performance, it will be compared to the standard ANTARES reconstruction algorithm *BBfit* [115] (*SeaTray* implementation, corresponding to version v3r5). The impact of different background rates and of an imperfect detector will be investigated, as well as the efficiency of the containment estimator and the quality of the energy reconstruction.

### 4.4.1 Multi-string reconstruction quality

In Fig. 4.15 the distributions of the reconstruction error on the zenith, $\Delta\Theta = \Theta_{\mathrm{rec}} - \Theta_{\mathrm{sim}}$, and on the azimuth, $\Delta\Phi = \Phi_{\mathrm{rec}} - \Phi_{\mathrm{sim}}$, are shown. The distributions are shown separately for the two different triggers *3N* and *2T3*. Only few multi-string events were additionally selected by the *2T3*, compared to the *3N* trigger, and thus both trigger samples have a similar behaviour concerning the multi-string fit.

In order to quantify the reconstruction quality, the median reconstruction errors are given. They are calculated from the absolute difference between reconstructed and true angle: 50% of the valid reconstructed events have an absolute reconstruction error (of the respective parameter) smaller than this median value. The median reconstruction errors on the zenith and azimuth are almost the same for both triggers: 0.7° and 1.6° for the *3N*, and 0.8° and 1.7° for the *2T3* trigger, respectively.

Also shown is the error on the zenith of the prefit, which is only used as reference for a hit selection at this point, but will also be used for a preselection cut later on in the context of the analysis. The median values of the absolute errors are 4.7° for the *3N* and 4.5° for the *2T3* trigger.

In Fig. 4.15, also the ratio of well reconstructed multi-string events to all triggered multi-string events is displayed. For low-energy events a reconstruction error of $|\Delta\Theta| < 5°$ defines "well reconstructed". The ratio decreases steeply below 40 GeV but is quite stable around 85 % above 50 GeV. Overall, for both triggers the mean fraction of well reconstructed multi-string events is about 80 %.

### 4.4.2 Single-string reconstruction quality

The percentage of single-string events of all triggered events is 21 % for the *3N* trigger and 73 % for the *2T3* trigger for a background rate of 60 kHz, similar to the values without any optical background. The performance of the single-string reconstruction is expected to be worse than the performance of the multi-string reconstruction due to the lack of three dimensional information and the lower number of signal hits in general. In particular, the azimuth angle can only be determined with bad quality. As it is not of relevance for this analysis anyway, it will be completely neglected in the

following. The distribution of the reconstruction error on the zenith and the ratio of well reconstructed events (defined analogously to the multi-string case) are shown in Fig. 4.16.

For the zenith, a median value of 2.4° for the *3N* events and 6.1° for *2T3* events is obtained. Below 30 GeV, the efficiency decreases less steeply compared to the multi-string reconstruction. Above 50 GeV, the efficiency is only around 70 – 80 % for the *3N*, and around 50 % for the *2T3* trigger. The overall efficiency is 68 % for the *3N* and 44 % for the *2T3* trigger. Noticeable is a significantly larger difference between the *3N* and the *2T3* trigger compared to the multi-string reconstruction, which is due to the increased efficiency of the *2T3* trigger that is basically due to an increase of single-string events. The *2T3* trigger is more sensitive to low-energy events, which contain less hits, and is therefore, on average, less efficient and yields larger reconstruction errors. Additionally, the contamination of the signal with background hits has the more impact the less hits an event has.

### 4.4.3 Overall performance and comparison with a standard ANTARES reconstruction algorithm

In order to compare *Posidonia* with other reconstruction strategies and to assess its overall performance, the single- and multi-string reconstruction will be combined in the following. As there is no overlap in the events, this just corresponds to the sum of the respective distributions. To evaluate the performance of *Posidonia*, it is compared to the performance of the *BBfit* algorithm [115]. As input, exactly the same simulated event sample was used. No quality cuts were applied at this level in either of the studies.

In Fig. 4.17 the distributions of the absolute reconstruction error on the zenith $\Theta$ are shown for *Posidonia* (left) and for *BBfit* (right). With *BBfit*, 92 % of the *3N* and 61 % of the *2T3* triggered events have a valid fit result. For *Posidonia*, these values account for 99 % and 98 % respectively. For *3N* events, the median values of the absolute error on the zenith reconstruction are 1.0° for *Posidonia* and 1.7° for *BBfit*, and for the *2T3* trigger the mean values degrade to 3.9° for *Posidonia* as well as for *BBfit*.

The overall fractions of well reconstructed events with *Posidonia* and *BBfit* differ by 10 to 20 % in the energy range below 50 GeV (Fig. 4.18). The energy-integrated percentage of well reconstructed in all triggered events is 78 % for *3N* and 54 % for *2T3* events for *Posidonia* and 66 % and 33 %, respectively, for *BBfit*.

### 4.4.4 Impact of different background rates

The optical background at the ANTARES site is not constant but changes on time scales of typically days to weeks. 60 kHz is a typical value, but the rates can go up to 150 kHz or even more. At some point, the signal-to-noise ratio becomes too small to extract the signal. The number of fake events increases as well as the number of misreconstructed events, due to a degradation in the performance of the hit filters. For this thesis, a mean background rate (see Section 5.2.1 for details about the calculation) of up to 80 kHz was

considered as acceptable for an analysis. The trigger rate is at the same level as with 60 kHz optical background (a weighted sum of 17 003 events per year with both triggers active). The impact of such an increased rate on the reconstruction is illustrated in Fig. 4.19. As expected, a small decrease in the fraction of well reconstructed events can be observed: the efficiencies are 74 % for *3N* and 53 % for *2T3* events for *Posidonia*, and to 66 % and 31 % respectively for *BBfit*.

This degradation of the reconstruction quality with increasing background rates is mainly due to the imperfect hit selection and has to be taken into account when analysing the data. The decrease in the percentage of single-string events (19 % single-string events with the *3N*, 71 % with the *2T3* trigger) also indicates the selection of additional background hits by the hit filters, resulting in an increase of multi-string events.

### 4.4.5 Impact of missing PMTs

Not only optical background degrades the measurements, but also failure of detector components. Individual OMs, storeys and entire lines can fail to take data for various reasons. To study the impact of real detector conditions, the simulations are masked by taking into account defective OMs or lines absent in the data taking. Possible reasons for the failure of OMs and the way they are treated in data and simulation will be discussed in detail in Chapter 5. For the simulated data this basically means, that modules, storeys or complete lines can be switched off, ignoring all hits theoretically detected by the corresponding modules.

For this study, a detector setup without lines 11 and 12 was chosen. This represents a phase during the installation of the detector, where only the lines 1 to 10 were taking data. Out of the resulting 750 modules, 108 modules were switched off continuously, plus a mean number of 66 randomly distributed modules (different ones from event to event). This results in 23 % bad modules, which is a typical value. The optical background was set to 60 kHz.

Due to the sparser detector and thus less hits, less events are triggered in total and a larger fraction of them is classified as single-string. The number of triggered events in total is reduced by more than a factor two (a weighted sum of 8 401 events per year with both triggers active), and the percentage of single-string events is increased to 24 % for *3N* and 78 % for *2T3*, respectively.

Despite the strongly reduced trigger rate and the shift towards the single-string reconstruction, the overall ratio of well reconstructed events for the *3N* triggered events is the same as for the ideal detector at the same noise level for *Posidonia* and decreases only slightly to 64 % for *BBfit*. The ratio of *2T3* triggered events decreases to 50 % for *Posidonia* and stays at 33 % for *BBfit* (see Fig. 4.20).

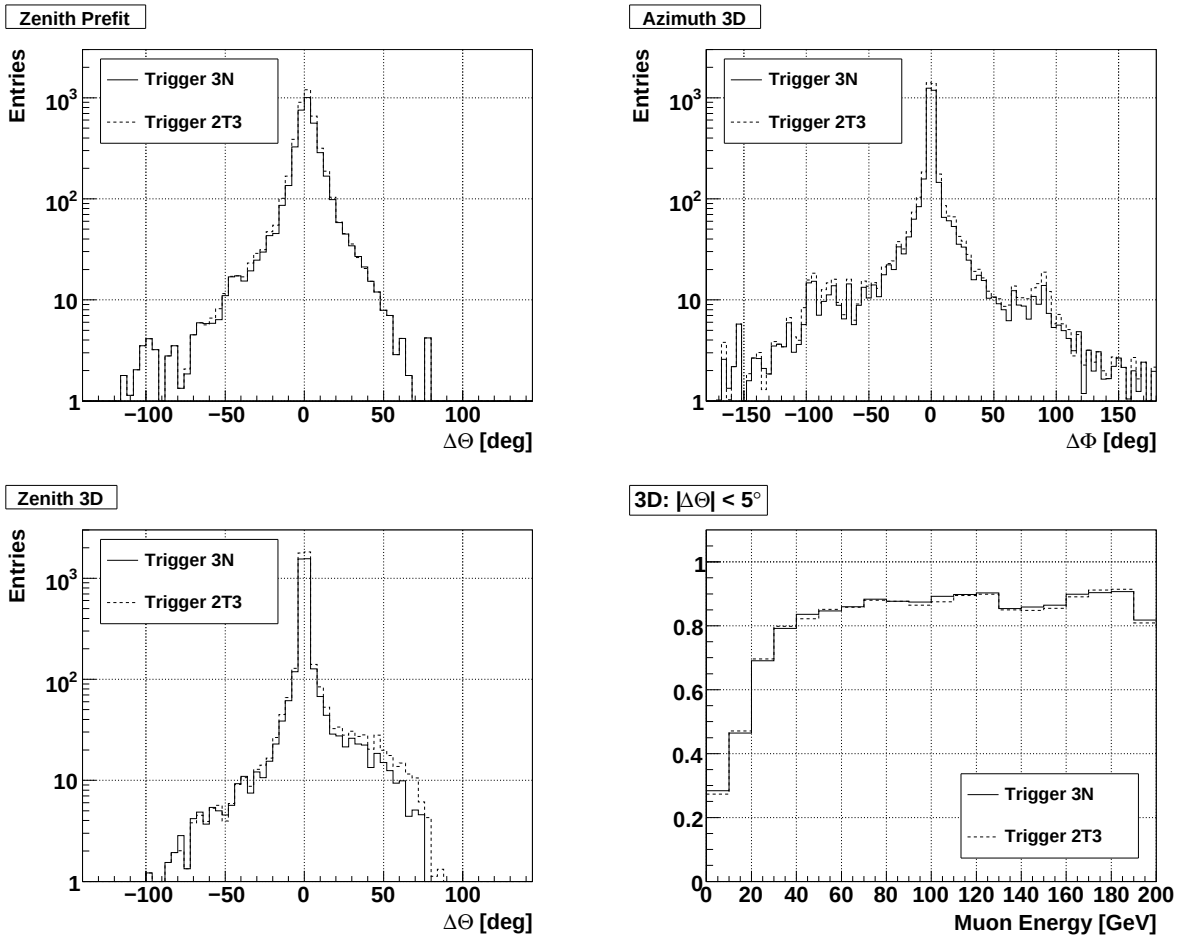



**Figure 4.15:** Distributions of the reconstruction error of the prefit zenith $\Theta_{\text{Prefit}}$ (*upper left*), the multi-string azimuth $\Phi$ (*upper right*), and the multi-string zenith $\Theta$ (*lower left*). All three plots show the distributions of the difference between the reconstructed and the simulated value. The *lower right* plot shows the ratio of well reconstructed events (reconstruction error on $\Theta$ smaller 5°) to all triggered events for the multi-string algorithm as a function of the true muon energy. The background rate is 60 kHz per OM.

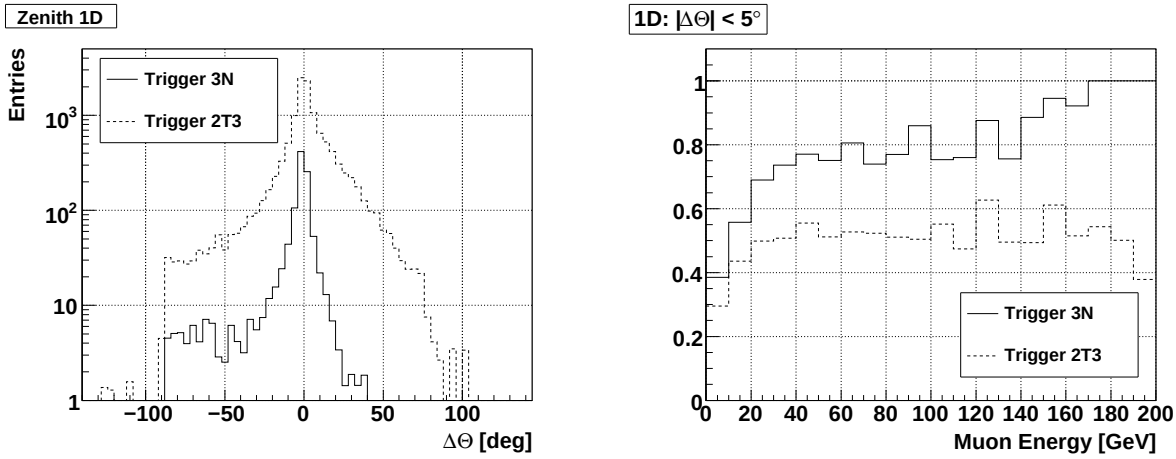



**Figure 4.16:** Distribution of the zenith reconstruction error of the *Posidonia* single-string reconstruction. Ratio of well reconstructed single-string events to all triggered single-string events as a function of the true muon energy (*right*). The background rate is 60 kHz per OM.

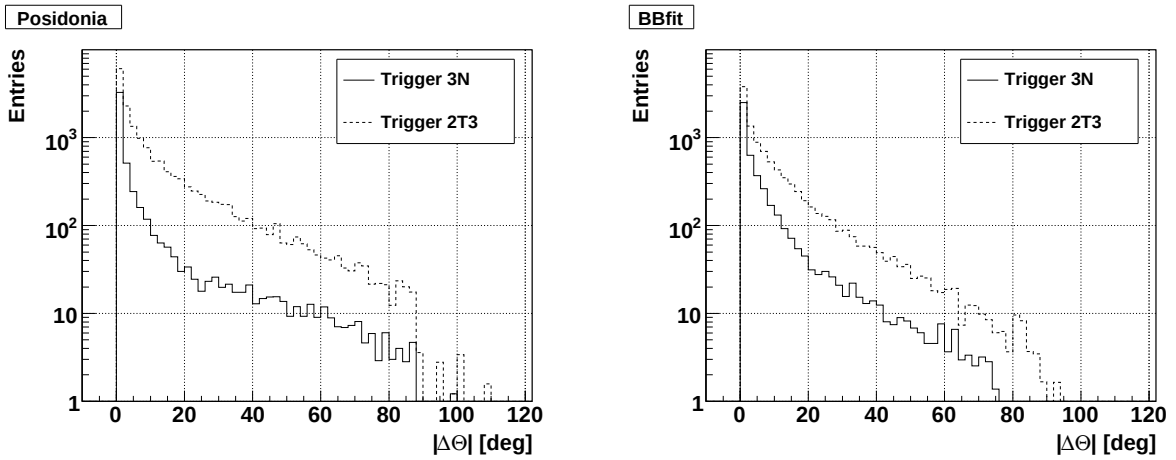



**Figure 4.17:** Distribution of the absolute value of the difference between simulated and reconstructed zenith angle $|\Delta\Theta|$ (in degrees). *Posidonia* reconstruction algorithm (*left*) and standard ANTARES reconstruction algorithm *BBfit* (*right*). The background rate is 60 kHz per OM.

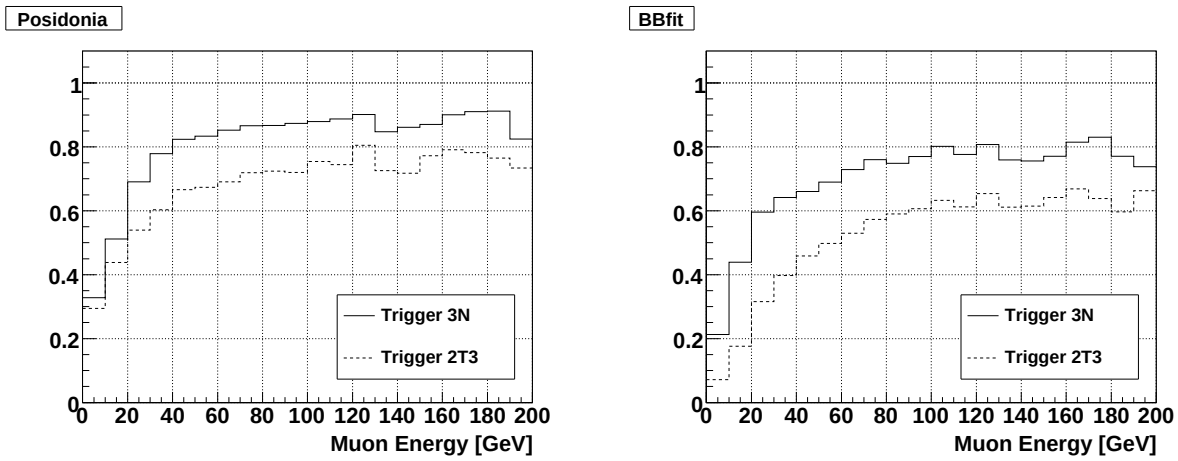



**Figure 4.18:** Ratio of well reconstructed events with the *Posidonia* (*left*) and the *BBfit* reconstruction algorithm (*right*) as a function of the true muon energy. The background rate is 60 kHz per OM.

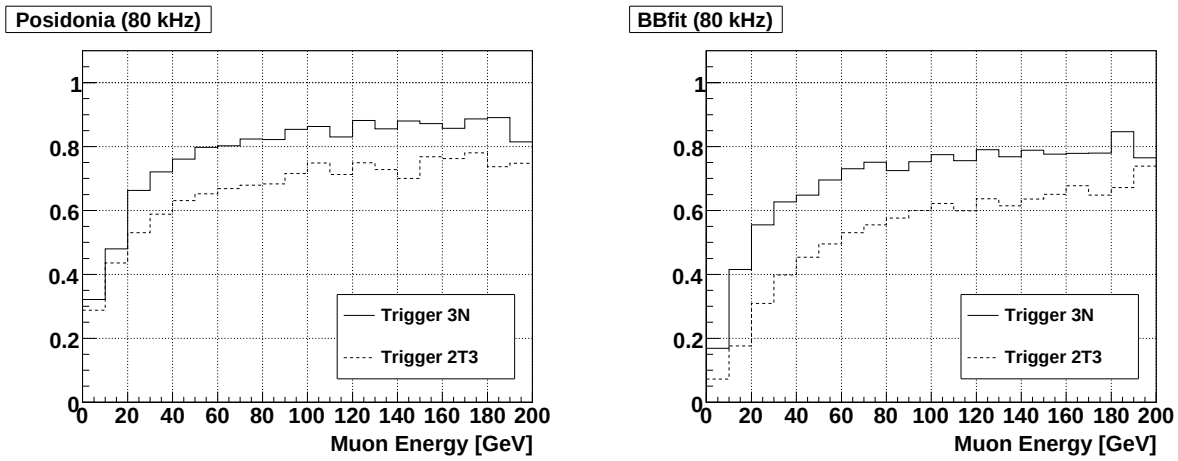



**Figure 4.19:** Like Fig. 4.18, but with a background rate of 80 kHz per OM.

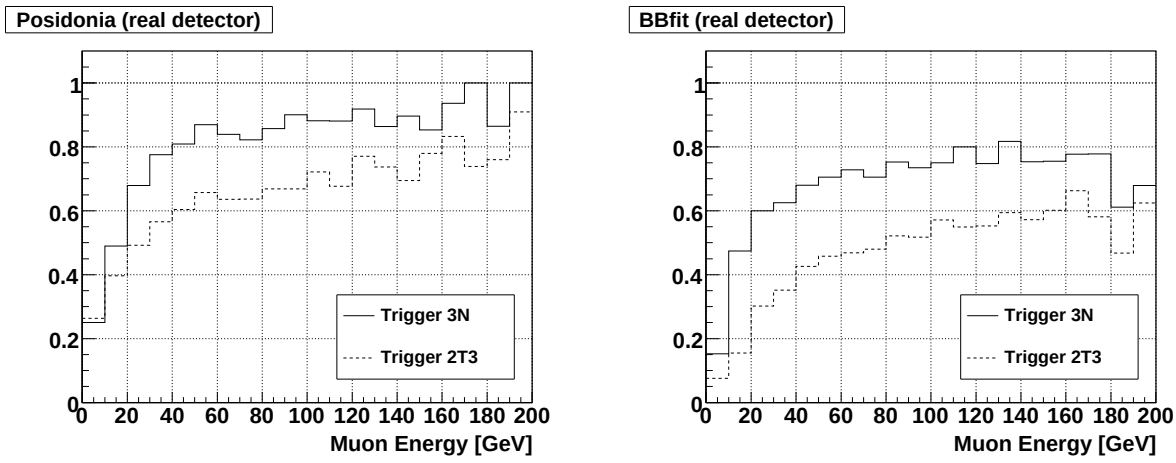



**Figure 4.20:** Ratio of well reconstructed events with the *Posidonia* (*left*) and the *BBfit* reconstruction algorithm (*right*) as a function of the true muon energy. Only 576 OMs are active and the background rate is 60 kHz per OM.

### 4.4.6 Performance of containment estimator and energy reconstruction

The selection of contained events, the calculation of the track length, and the subsequent energy reconstruction can only reasonably be performed for well reconstructed events. Hence, in the following, only events of the ideal reference sample were considered that were reconstructed within 5° of the true muon zenith angle. Additionally, the simulated test sample was extended to muon energies up to $E = 10^7\,\mathrm{GeV}$, in order to better assess the functionality of the containment estimator and the energy reconstruction.

#### Multi-string events

Because of the asymmetric sensitivity of the detector and the seabed below the detector, the reconstruction of the $z$-coordinate of the starting position of the track is biased and a tolerable agreement between simulation and reconstruction is only given up to about $z = -150\,\mathrm{m}$ (origin at detector centre; Fig. 4.21). Therefore, events with an interaction vertex at reconstructed $z$-coordinates of less than $-150\,\mathrm{m}$ are excluded for the containment estimation of multi-string events.

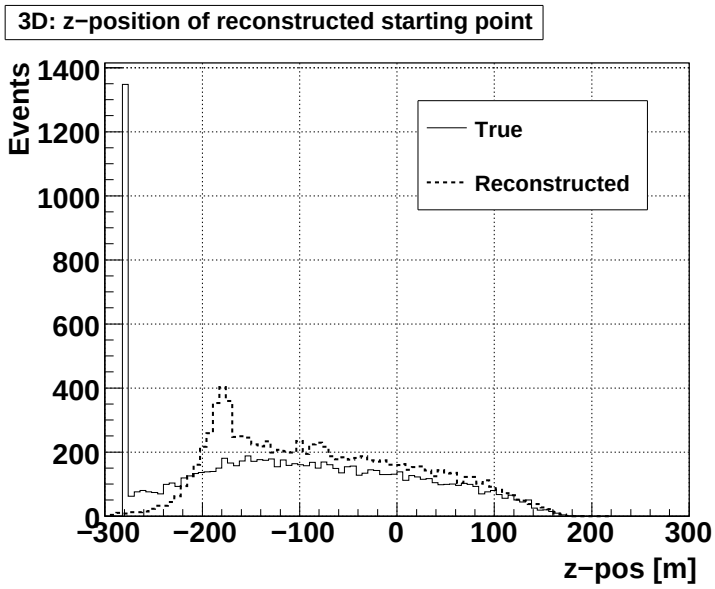



**Figure 4.21:** Distribution of true simulated (dashed line) and reconstructed (dashed line) $z$-coordinates of the starting points of the muon tracks. The sharp peak at $z = -280\,\mathrm{m}$ coincides with the position of the sea floor, and light emission is only simulated above the sea floor. In these cases, the starting point of the muon track is not the interaction vertex.

Starting and stopping points of tracks are called "contained", if they are within the containment volume (CV). The track is called "contained", if both its starting and

stopping points are contained. A rotational symmetry was assumed for the CV of multi-string events, implying a cylindrical shape. The chosen values are a radius of $r_{\mathrm{cv}} = 120\,\mathrm{m}$, and a height $z_{\mathrm{cv}}$ from $-150\,\mathrm{m}$ to $180\,\mathrm{m}$.

The upper row of Fig. 4.22 shows the (normalised) distributions of the calculated number of photons for contained and non-contained starting and stopping points, respectively. The lower row in the figure shows the ratio of contained and non-contained events, that are lost when cutting on the respective number of photons of the containment estimate.

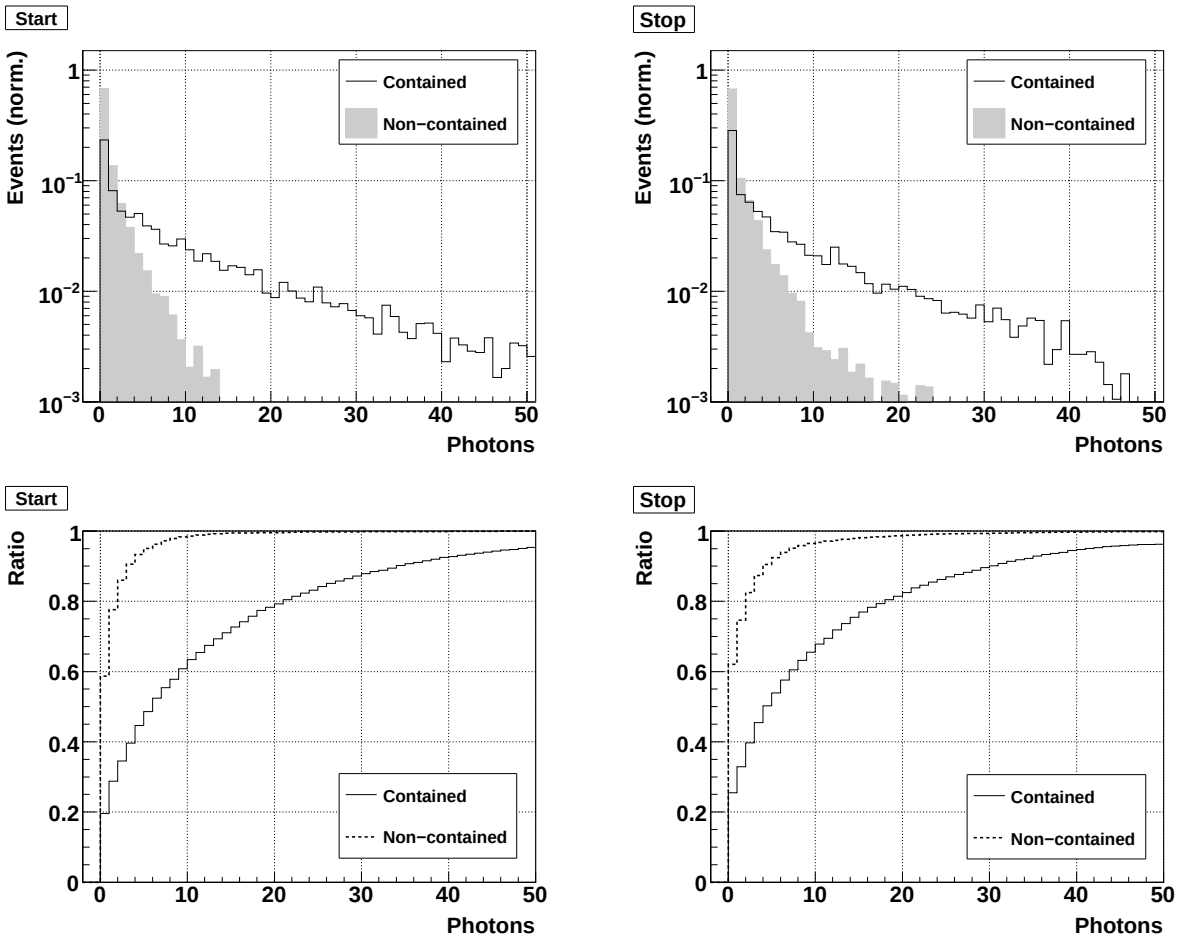



**Figure 4.22:** Distributions of the expected number of photons for contained and non-contained events (*upper row*) and the ratio of suppressed events, when cutting on the specific value (*lower row*). Distributions for the starting point (*left*) and for the stopping point (*right*). Only neutrinos with a zenith reconstruction error of less than 5° were considered.

For the following, a cut value of an expected number of 1.1 photons was chosen, both for the starting and for the stopping point. From originally 6 932 multi-string events, only 1 373 survived this cut. For the starting point, 30 (±1) % of non-contained events

remained after the cut, whereas 76 (±2) % of the contained events were kept. For the stopping point, the values are 31 (±1) % and 71 (±2) %, respectively.

Cutting on the output of the containment estimator is equivalent with selecting low-energy muons, because of the size of the detector, that restricts contained events to track lengths of about 300 m. In Fig. 4.23, the muon energy of the events before and after the cut (both on starting and stopping estimate) can be seen. The mean muon energy of the event sample is reduced from 445 GeV before the cut to 149 GeV after the cut.

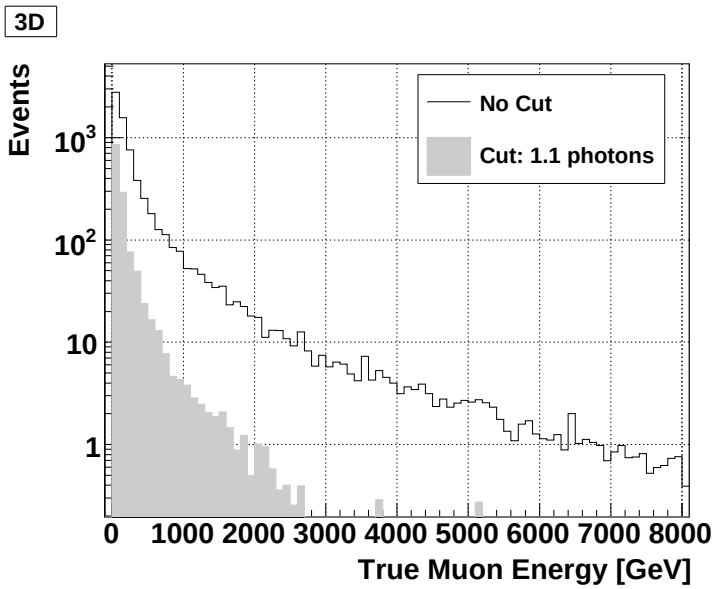



**Figure 4.23:** Distribution of the true muon energy for the event sample before (solid line) and after cutting on the output of the containment estimator (shaded area). Only neutrinos with a zenith reconstruction error of less than 5° were considered.

For the events remaining the cut, the median distance from the reconstructed to the true starting point is 32 m and the median distance from the reconstructed to the true stopping point accounts for 58 m. The muon track length is calculated as the difference between those two values. Figure 4.24, left, shows the mean true track length as a function of the reconstructed track length. The mean reconstructed track length shows a good correlation with the true track length up to about 300 m. In Fig. 4.24, right, the true track length and the reconstructed track length are displayed in a scatter plot.

The reconstructed track length in turn shows a very good linear correlation with the true energy of the muon and the energy of the selected events can be directly inferred. In the upper row of Fig. 4.25, left, the mean true muon energy is shown as a function of the reconstructed muon track length. The green line indicates the result of a straight line fit ($f(x) = a_0 + a_1 \cdot x$) to the distribution, which is then employed

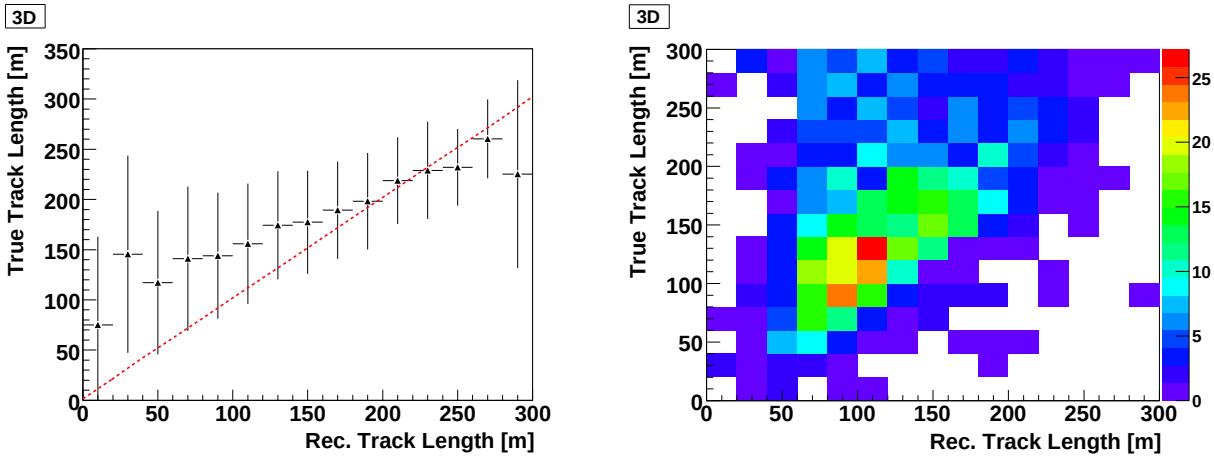



**Figure 4.24:** Mean true muon track length as a function of the reconstructed muon track length (*left*). True muon track length and reconstructed muon track length displayed in a scatter plot (*right*). Only events with a zenith reconstruction error of less than 5° and containment estimation values of more than 1.1 photons for both the starting and stopping point were considered. The bars indicated the standard deviation, and the red dashed line indicates the identity.

for the reconstruction of the energy of the muon. The values obtained from the fit are $a_0 = 27.85\,\text{GeV}$ with an error of 19.80 GeV and $a_1 = 0.34\,\text{GeV/m}$ with an error of $0.12\,\text{GeV/m}$[17]. The mean true muon energy as a function of the reconstructed energy can be seen in the upper row of Fig. 4.25, right. The mean reconstructed energy per energy bin matches almost perfectly the true muon energy. The median reconstruction error on the energy is 35 GeV. In the lower row of Fig. 4.25 the true muon energy and the reconstructed track length (left), as well as the true muon energy and the reconstructed muon energy (right) are displayed in scatter plots.

### Single-string events

A different containment volume needs to be defined for single-string events. Again, a cylindrical shape was chosen such that proximity of the track to the string is assured: the volume has a radius of $r_{\text{cv}} = 60\,\text{m}$ around the string on which the event was detected, and a height $z_{\text{cv}}$ from -150 m to 150 m. Like for the multi-string events, the single-string CV was determined empirically.

As mentioned in the description of the single-string containment estimator and also in the characterisation of the low-energy events, single-string events tend to be "more vertical". Thus, events can be separated into contained and non-contained events by using the $z$-coordinate of the reconstructed starting or stopping position, respectively,

[17] This value compares well to the values reported by the Particle Data Group [20].

as cut parameter. To guarantee the validity of this approach, a cut on the reconstructed zenith angle is set at $\Theta = 135°$.

In the first row of Fig. 4.26, the distributions of the absolute values of the reconstructed $z$-coordinates $|p_z|$ of the starting (left) and stopping point (right) for contained and non-contained events are shown. To enhance features, for this figure containment is required within a radius of $r'_{\mathrm{cv}} = 30\,\mathrm{m}$ around the string and a height $z'_{\mathrm{cv}}$ from $-100\,\mathrm{m}$ to 100 m, whereas non-containment is determined by means of the default single-string CV. The second row of the figure shows the fraction of contained and non-contained events that remain when cutting on the respective $|p_z|$ value of either the starting (left) or the stopping point (right).

The cut applied for the selection of the contained events is $|p_z| < 100\,\mathrm{m}$ both for the starting and the stopping point. 68 % of the events with a contained starting point survive the cut on the starting point, as well as 25 % of the events with non-contained starting point. For the stopping point the values are 68 % and 34 %, respectively. By applying the combined cut on the starting and stopping point, the event sample is reduced from originally 2 834 single-string events to about 982 contained single-string events. As for the multi-string containment estimator, cutting on the single-string containment parameter $|p_z|$ means selecting low-energy muons. This can be seen in Fig. 4.27, where the true muon energy of all single-string events is shown, as well as the true muon energy of contained single-string events. The mean energy of the sample decreased from about 84 GeV before the cut to 44 GeV after the cut.

In Fig. 4.28, right, the mean true muon track length for contained single-string events is plotted as a function of the reconstructed muon track length. It can be seen from the mean values, that a good correlation is given only up to about 170 m reconstructed track length. Therefore, for the single-string energy reconstruction, a further cut is applied on the reconstructed track length at 170 m, which reduces the test sample used in this section by only ten events. The true muon track length and the reconstructed muon track length are displayed in a scatter plot in Fig. 4.28, left.

The correlation between the reconstructed track length and the mean true energy of the muon can be seen in Fig. 4.29, left. The green solid line is the result of a straight line fit, which is employed later on for the reconstruction of the energy. The values obtained from the fit are $a_0 = 4.94\,\mathrm{GeV}$ with an error of 18.6 GeV, and $a_1 = 0.38\,\mathrm{GeV/m}$ with an error of $0.20\,\mathrm{GeV/m}$ . On the right hand side of the same figure, the correlation between the reconstructed muon energy and the true muon energy is shown. The mean reconstructed energy per energy bin matches nicely the true muon energy. The median reconstruction error on the energy is 13 GeV. In the lower row of Fig. 4.29, the true muon energy and the reconstructed track length (left), as well as the true muon energy and the reconstructed muon energy (right) are displayed in scatter plots.

## 4.5 Summary

The percentage of well reconstructed events in the muon energy range below 100 GeV can be increased by up to 20 % by using the *Posidonia* reconstruction algorithm instead

of the ANTARES standard reconstruction program *BBfit*. In the relevant energy range around and below 50 GeV the improvement even exceeds 20 %. This is also the case for increased background rates or a detector with more than 20 % disfunctional OMs. The behaviour of *Posidonia* in the low-energy range is robust against changing environmental conditions, which eventually helps to reduce systematics. Because of the correlation between track length and energy, the selection of contained events is equivalent to selecting low-energy tracks. For contained tracks, a reliable energy reconstruction is possible. Both for single- and multi-string events, a good correlation between reconstructed and true track length and thus between reconstructed track length and energy is obtained, up to track lengths of 170 m and 300 m for single- and multi-string tracks, respectively.

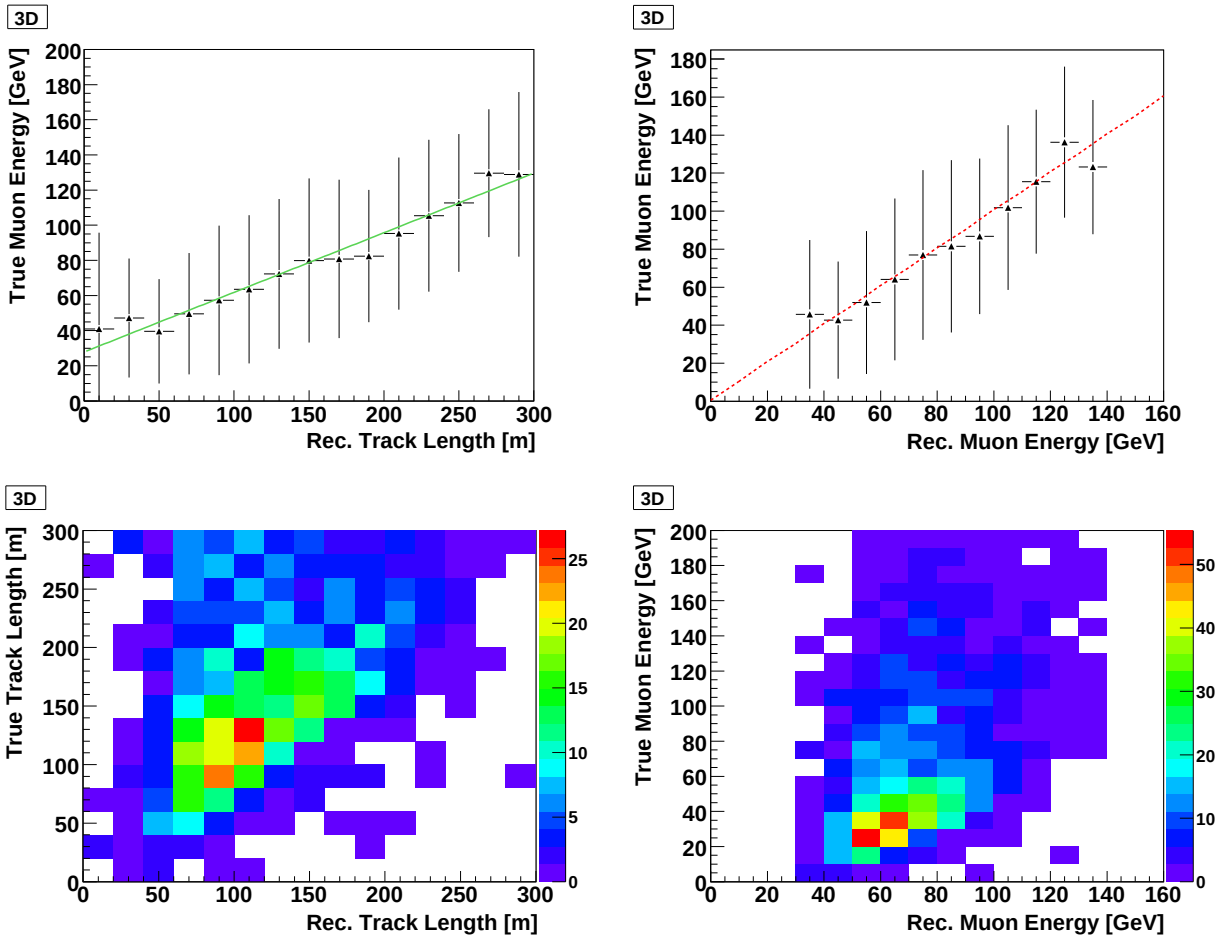



**Figure 4.25:** *Upper left:* Mean true muon energy as a function of the reconstructed muon track length. The green line indicates the result of a fit to the distribution.
*Upper right:* Mean true muon energy as a function of the reconstructed muon energy. The red dashed line indicates the identity.
Both figures: the bars indicate the standard deviation.
*Lower row:* True muon energy and reconstructed muon track length (*left*), as well as true muon energy and reconstructed muon energy (*right*) displayed in scatter plots. Only neutrino events with a zenith reconstruction error of less than 5° and containment estimation values of more than 1.1 photons for both the starting and stopping point were considered.

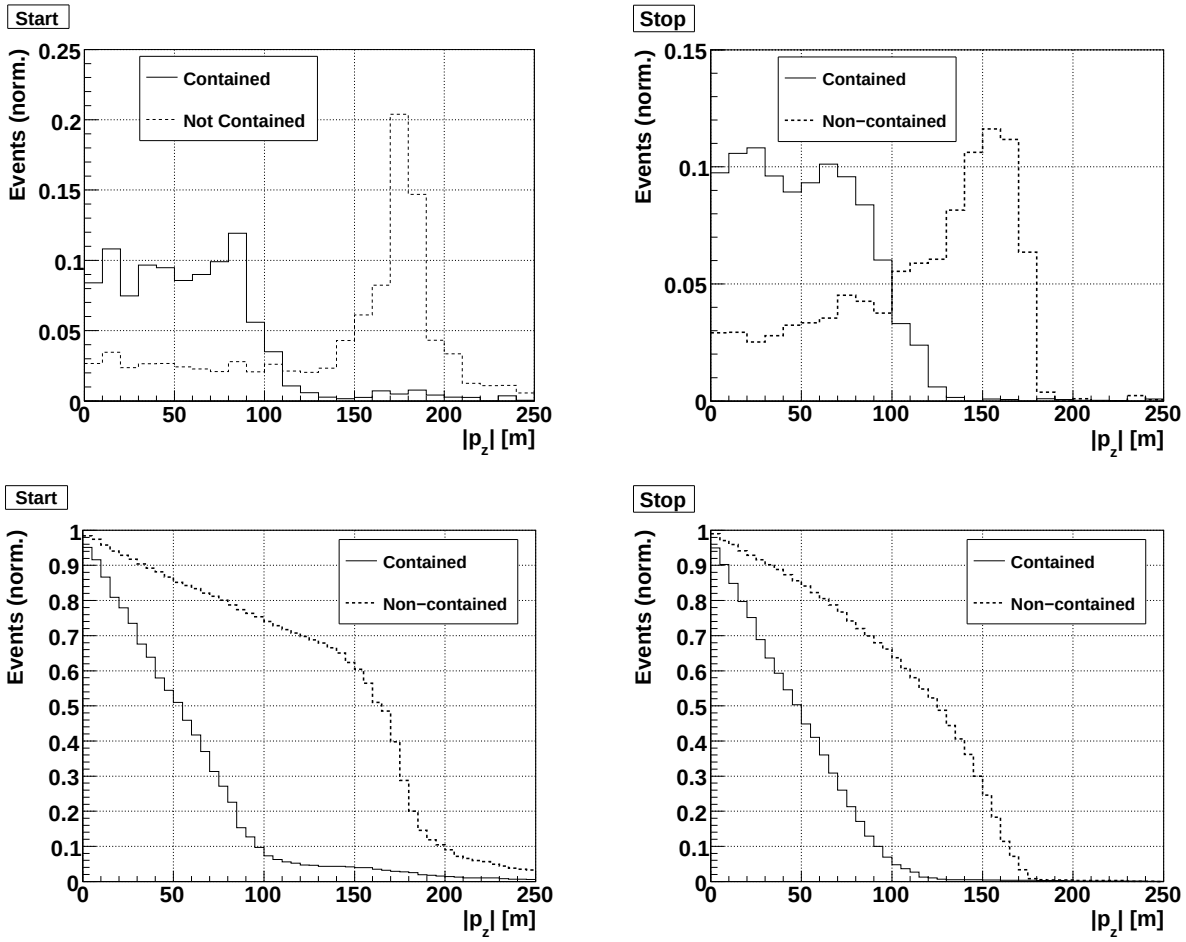



**Figure 4.26:** Distributions of the absolute $z$-coordinate of the reconstructed starting (*left*) and stopping point (*right*) of the track for contained and non-contained single-string events (*upper row*) and the fraction of events lost when cutting at $|p_z|$ (*lower row*).

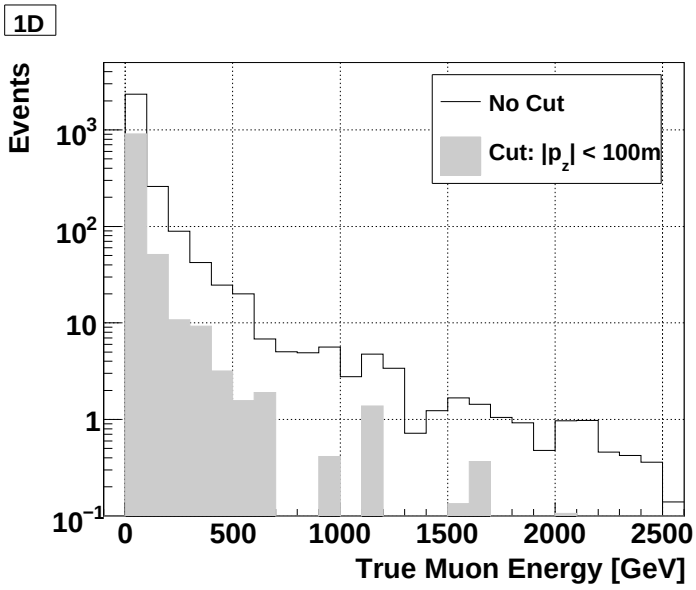



**Figure 4.27:** Distribution of the true muon energy of the events before cutting on the containment parameter $|p_z|$ (solid line) and muon energy of the remaining events after the cut on $|p_z| < 100\,\mathrm{m}$ both for starting and stopping point (shaded area). Only neutrinos with a zenith reconstruction error on the muon of less than 5° and a reconstructed muon zenith of less than 135° were considered.

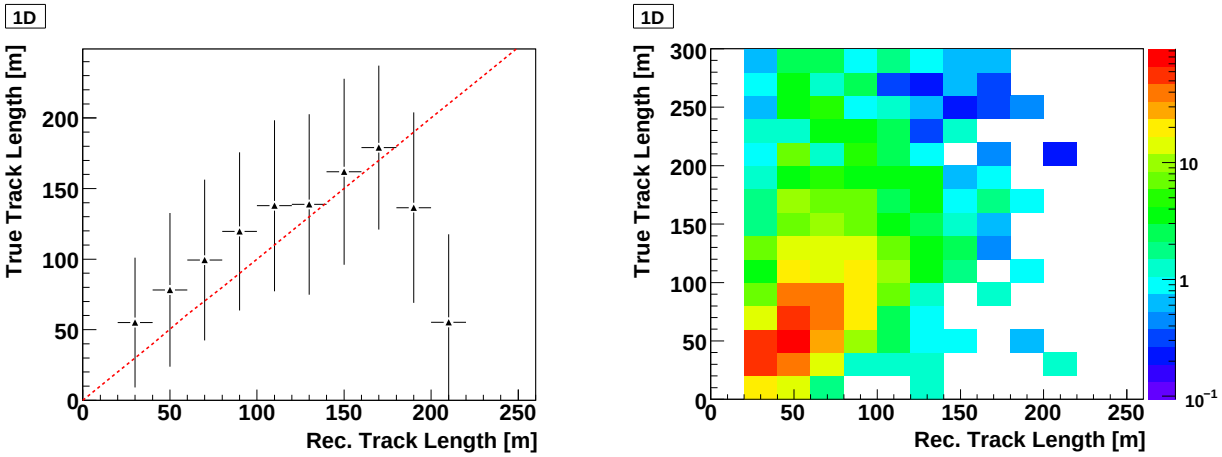



**Figure 4.28:** Mean true muon track length as a function of the reconstructed muon track length (*left*). True muon track length and reconstructed muon track length displayed in a scatter plot (*right*, event numbers given on a log scale). Only neutrinos with a zenith reconstruction error on the muon of less than 5°, a reconstructed muon zenith of less than 135°, and $|p_z| < 100\,\mathrm{m}$ for both the starting and the stopping point were considered. The bars indicate the standard deviation and the red dashed line indicates the identity.

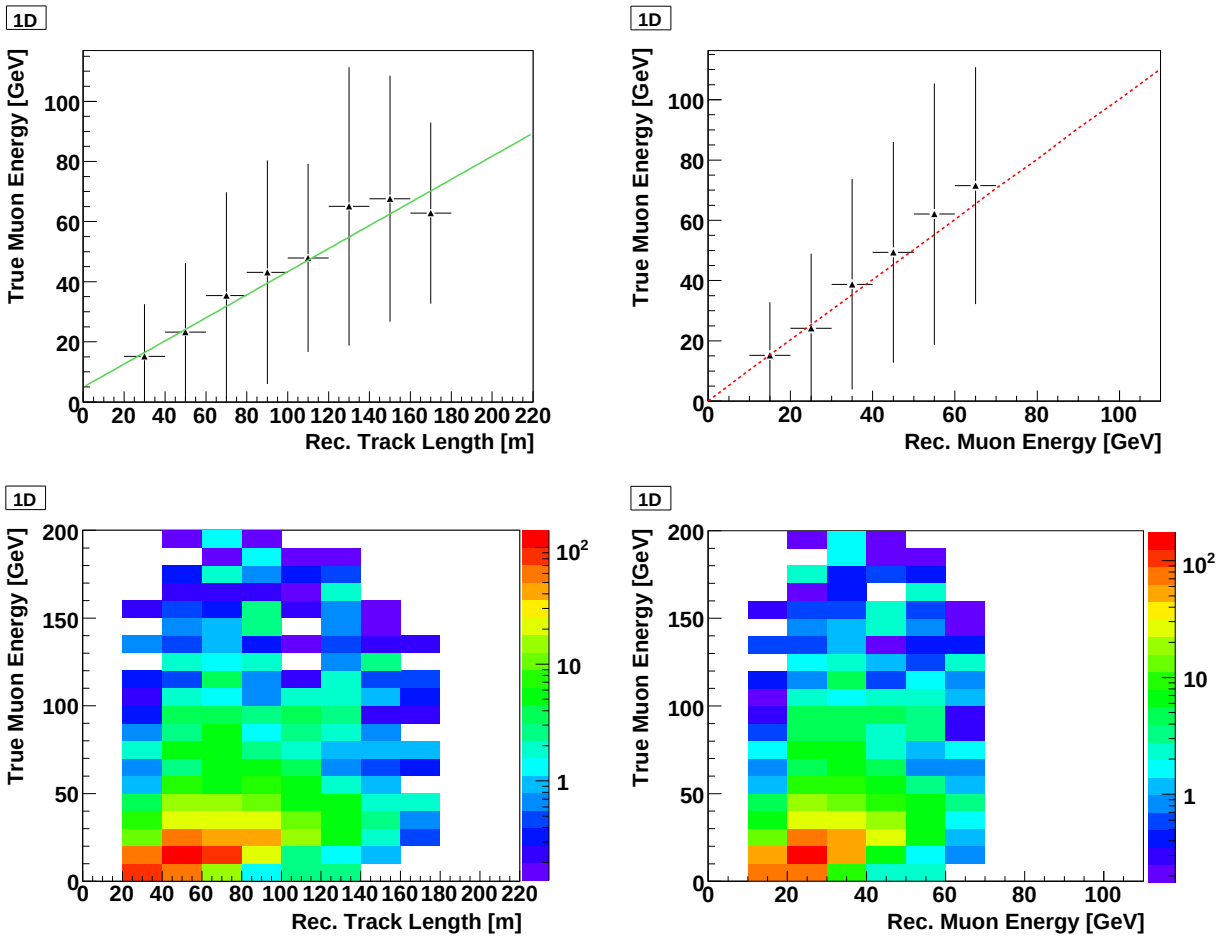



**Figure 4.29:** *Upper left:* Mean true muon energy as a function of the reconstructed muon track length. The green line indicates the result of a fit to the distribution.
*Upper right:* Mean true muon energy as function of the reconstructed muon energy. The red dashed line indicates the identity.
Both figures: the bars indicate the standard deviation.
*Lower row:* True muon energy and reconstructed muon track length (*left*), as well as true muon energy and reconstructed muon energy (*right*) displayed in scatter plots. The event numbers are given on a log scale. Only neutrino events with a zenith reconstruction error of less than 5° and a reconstructed muon zenith of less than 135°, and $|p_z| < 100\,\mathrm{m}$ for both the starting and the stopping point were considered.

# 5 Processing of data and simulations

No measurement process is ideal and measured signals typically suffer from background from different sources. Additionally, every measurement process can be subject to various kinds of problems, possibly having impact on the recorded data. It is therefore mandatory to inspect and evaluate the quality of recorded data. Only data that are matching specific standards and conditions may be used for an analysis. Furthermore, remaining imperfections have to be included in the simulations, to reproduce the real detection process as well as possible and to guarantee good agreement between simulations and data. Finally, after the track reconstruction process, quality cuts have to be applied, in order to select well reconstructed events and to obtain a high-quality sample of reconstructed events for the subsequent analysis.

In this chapter, data quality assessment and data selection for the following oscillation analysis are discussed. In Section 5.1, processes with impact on the data quality will be illustrated. Several of such processes that are affecting the data, like miscalibration or the contamination due to bioluminescence are not yet completely understood.

In principle, simulated data, without and with optical background included, is provided by an ANTARES working group to the whole collaboration. At the time the processing of data and simulation for this analysis should start, there were certain shortcomings in the consistency of data and simulation. To overcome these problems, with special focus to the low-energy regime, an alternative approach was made to assess the quality of ANTARES data and to simulate optical background and different kinds of detector problems.

Data assessment and selection, and the data processing chain are described in Section 5.2; the simulation of the detector electronics and real data taking conditions are illustrated in Section 5.3. The common reconstruction procedure of data and simulations is described in Section 5.4. In order to further suppress muonic background and badly reconstructed muon neutrinos after event reconstruction, quality cuts are applied using a neural net. To finally obtain a sample of low-energy neutrino events and to be able to do a reliable energy reconstruction, only events where the muon track is contained in the detector will be considered. Both procedures are subject of Section 5.5.

## 5.1 Processes with impact on data quality

Constant processes like homogeneous, random background and also some kinds of failure like permanently defective optical modules, can typically be considered in the simulations very easily. More challenging are processes that occur rarely, have fluctuating intensity or are not yet well understood. Three main factors, the alignment, optical background and failures of OMs and LCMs (including miscalibration of OMs) are known of having impact on the data and will be discussed in the following subsection.

### 5.1.1 Alignment

The strings of the ANTARES detector are free to move in the deep-sea currents. The exact positions of the OMs are therefore unknown and need to be continuously monitored. For this purpose, compasses, tilt-meters and an acoustic positioning system are employed. Data from these instruments is recorded every few minutes (currently two, formerly six minutes). The combined information allows for the reconstruction of the shapes of the ANTARES strings, and the particular detector geometry can be obtained. This calculation is currently done offline. The knowledge of the position of the optical modules, i.e. the coordinates of the hits recorded by the modules, is essential for the track reconstruction and the errors on the reconstructed trajectories are increasing with the error on the positions of the optical modules.

The alignment procedure can fail if no data from the hydrophones and/or the compasses and tiltmeters are available. Reasons for missing data are for example problems with the slow control, which is controlling the readout of various monitoring instruments, or problems during sending and writing the compasses and tiltmeter and/or hydrophone data to the data base. Also, when data from a certain number of tiltmeters from an individual string are missing, the shape reconstruction of the respective strings is not possible.

Currently, if an aligned geometry is missing, either for the complete detector or for individual strings, the so called nominal geometry is employed, which corresponds to a perfectly aligned detector with the line shapes being straight lines and all OMs oriented vertically downward. Periods of missing alignment can affect complete runs, but they also occur within runs, so that for parts of a run only the nominal geometry is available. When the nominal geometry is employed, the errors on the line shape increase with increasing displacement of the lines, and in general they are expected to be larger than with an aligned geometry. Furthermore, when switching from an aligned to nominal geometry or vice versa, a discontinuity occurs, with an impact that is difficult to estimate. Therefore, one should avoid using the nominal geometry for an analysis.

### 5.1.2 Optical noise

Optical background mainly consists of contributions from decays of $^{40}$K and of bioluminescence. The $^{40}$K decay is homogeneous at a constant level of about 30 kHz per

optical module. Bioluminescence on the other hand is emitted by organisms and is not constant in time. Even more, macroorganisms can emit large amounts of light, so called bursts, causing a local increase in the optical background rate. This kind of background is not yet well investigated and understood and still bears a lot of potential for marine biology research. Nevertheless, it has to be treated for an acceptable agreement between data and simulation.

Investigations show that contributions from bioluminescence can be separated into a rather constant and homogeneous part that, together with the $^{40}$K decays, contributes a rather stable rate (referred to as baseline), and the above mentioned bursts. The baseline typically ranges between 55 and 70 kHz and mainly consists of single photoelectron pulses. During bursts, rates up to several MHz can be observed and the number of coincidences (hits on different OMs at the same storey within some ns) and big hits (amplitudes of more than one photoelectron) is expected to increase. Beyond a certain rate (roughly at about 250 kHz), due to saturation of the data processing system, affected optical modules only send empty data frames (which contain no recorded signals but only some header information) to the on-shore processors, leading effectively to a vanishing rate of these OMs.

### 5.1.3 OM and LCM/SCM failure

The optical modules of ANTARES are custom designed, complex, and extremely sensitive measuring instruments, performing their task in a very hostile environment. The same holds for the local and string control modules (LCM/SCM) that consist of titanium cylinders, housing the electronics for the individual storeys and strings. Naturally, the housings, as well as the electronics and the sensors, can malfunction in several ways, in particular as ANTARES is the first fully operational optical deep-sea neutrino telescope[18] and a pilot project for future, much larger telescopes.

Penetration of water into glass spheres of OMs or electronics containers had happened rarely up to now. More frequent are problems with the electronics or the sensors, such that affected modules eventually might have to be turned off permanently. From time to time the connection to individual LCMs or SCMs is lost and no data are recorded or transmitted for a certain time period from a whole storey or even a complete sector. Typically, the LCMs/SCMs can be "revived" in-situ (when data taking pauses), but it also happens that complete sectors are permanently lost, at least unless the string is recovered from the sea bed and repaired. This sums up to a certain number of OMs that are not taking part in the data-taking process and can be considered as "off". Some modules are affected permanently or for a long period of time and are never taking data during that time. Such behaviour is straightforward to consider in the simulations, but when it comes to detecting units that are only temporarily affected, the task is more complex.

OMs can exhibit further problems that renders them unusable for analysis, like

[18] The DUMAND project [116] that existed between 1975 and 1995 aimed at constructing the first deep-ocean high-energy neutrino detector, but was cancelled due to technical problems.

recording high (above 250 kHz) or unusually low rates (below 40 kHz), or a permanent, recurring, or transient abnormal fraction of hits with large amplitude. Permanent problems might be due to a bad charge calibration, whereas locally clustered, transient high rates are probably caused by local bioluminescence bursts.

Low or high rates, as well as sending empty frames, are typically transient states of an OM, that randomly affect a certain percentage of OMs in a mainly unpredictable and highly stochastic way. Their impact is difficult to estimate and can disturb the measurement. Badly calibrated modules producing fake large hits can lead to fake triggers and also bias reconstruction algorithms that are sensitive to hits with large amplitudes.

## 5.2 Data processing

For a neutrino oscillation analysis at the edge of detector sensitivity, a data high-quality data set is necessary. Within the context of this thesis, methods to assess the quality of the recorded data and to select data have been developed.

The information obtained from data quality assessment is also employed to mask the raw signal simulations (resulting from the simulation chain described in Section 3.6; these simulations are conducted on the basis of a perfect 12-line detector and without optical background). This procedure will be described in detail in Section 5.3.

### 5.2.1 Data quality assessment and data selection

For this analysis, all data recorded between February 2007 and end of October 2009 were taken into account. Only runs have been chosen for further processing that have a duration of more than 15 minutes and contain more than 1 000 events. In addition, they had to be marked with a PHYSICS tag, which implies data taking for physics purposes (instead of calibration for example). Runs with the label SCAN in the run or trigger setup were excluded, because such runs are employed for the tuning of the high voltage of the PMTs and might show unexpected behaviour. Furthermore, by means of parameters already available from the data base, the remaining runs were tested for technical problems, for example for problems with the data filters.

After this rough and rather standard selection on a run-by-run basis, the data selection procedure has been applied that has been devised within context of this thesis. The method is based on an event-by-event inspection. Instead of deselecting complete runs due to temporary problems only the affected parts of the data are dismissed. For each event the availability of alignment data is checked and two quality assessment parameters, the so-called bad channel ratio and the mean rate, are calculated. An event is selected only if both the aligned geometry is available and none of the two quality assessment parameters is exceeding a certain threshold.

### Missing alignment data

The settings of the offline analysis framework *SeaTray*[19] were such, that if a single two (six) minute period of alignment data was missing, the values of the preceding period were used. If the data of the preceding period were also missing, the data from the subsequent period were used. This is commonly considered to be an acceptable solution, as the deep-sea currents typically do not change significantly on such time scales. No further extrapolation was done if more than three subsequent periods of alignment data are missing, resulting in periods without available alignment. For these periods, the nominal geometry was employed.

If only individual strings are unaligned due to missing data, the whole detector was assumed to be not aligned and also the nominal geometry was used. As the strings are typically moving in parallel, the shape of strings with missing alignment data can theoretically be inferred from the neighboring aligned strings. This modification and improvement of the alignment procedure was implemented only towards the end of this work and could therefore not be used for the analysis.

As already mentioned in Section 5.1.1, it has to be avoided to switch between nominal and aligned geometry within the analysis procedure. Principally, two solutions are possible to treat this problem. The first one is to consequently use the nominal geometry, which might be reasonable when reconstructing with a robust algorithm. The second solution is to use the aligned geometry and skip all data without alignment information. As the *Posidonia* reconstruction algorithm is expected to be very sensitive, only data with full alignment information are used in the final analysis. Whereas accepting only runs with complete coverage of alignment data would significantly reduce the available amount of data, working on an event-by-event basis allows for excluding shorter periods without alignment information.

### Bad quality of the data

Two quality assessment parameters, the bad channel ratio and the mean rate are calculated for each event and is based on the information of the particular frame that contains the event. Consequently, two events within the same frame have the same bad channel ratio and mean rate.

The bad channel ratio is given by the number of OMs that are in a failure state divided by the number of all theoretically active modules. An OM can be either in one of four predefined failure states or in the state OK, when recording data regularly. The different failure states are either OFF, when the module is not sending any data, HIGH or LOW, when it is measuring high (more than 250 kHz) or low (less than 40 kHz) rates, or EMPTY, when it is sendig empty frames. Another important state was lately included, hinting at bad calibration. It could not be considered for this data production any more, but should be in future analyses.

During times of high bioluminescent activity like local bursts, affected modules are typically either in state HIGH or EMPTY and should not be considered for track recon-

[19]The alignment version used for the analysis in this thesis was V:0.992.

struction. Especially OMs with very high rates or big hits due to such bioluminescence bursts can bias the reconstruction result. Also, events that are triggered only by such OMs should be discarded. A reasonable and feasible solution to reduce the impact of bioluminescent activity is to ignore data from OMs being in any failure state (see Section 5.3). For data selection purposes events are discarded, where the bad channel ratio exceeded a certain limit, which was chosen to be 30 % for this analysis.

The mean rate is the mean counting rate of all OMs being in OK state in the particular frame. This rate will later be used for masking the simulations with optical background. Only events with mean rates below 80 kHz were accepted for the analysis data set.

#### Calculating the effective lifetime

Selection on an event-by-event basis immediately generates the demand for a tool to calculate the effective lifetime: dismissing individual events is equivalent to turning the detector off during these periods. This procedure is legitimate, as long as the decision about dismissing events is based on parameters that are not correlated with the physics signal but depend only on background processes. This is the case for the alignment data, as well as for the mean rate and the bad channel ratio.

The calculation of the effective lifetime (of a run for example) is done by adding up the extended time intervals of each accepted event, which are half of the period between the preceding and the following event.

Indeed, when an event has to be discarded, it is unknown how much of the time intervals between the current event and the preceding and following event, respectively, is affected and so this approach is only an approximation. Nevertheless, it is acceptable because processes that lead to the dismission of events typically affect several minutes (e.g. missing alignment), implying that uncertainties in the range of a few ms to s, which is the typical duration between two triggered events, are negligible. Furthermore, the uncertainties are expected to cancel to a large extend.

### 5.2.2 Organising the data

ANTARES is a "living" detector with varying detection conditions, concerning optical background, size of the detector, number of OMs in failure states, and trigger settings. The optical background shows, besides fluctuations due to bursts, variations on larger time scales (days to weeks) that have to be considered in the simulation. The size of the detector is changing with the number of active lines. It was increased during the construction of the detector and repeatedly changed afterwards due to failures of complete lines, which had to be recovered and repaired before they were redeployed and resumed data taking. Additionally, temporary failures of OMs and LCMs/SCMs cause varying bad channel ratios. Another issue are changing trigger settings: the threshold for *L1* big hits is either 3 or 10 photoelectrons, depending on the settings chosen, and thus affects the trigger efficiency.

To adequately incorporate all these factors in the analysis, the data set remaining after the selection procedure described above has been grouped into different setups

of similar conditions. For each setup, simulations of (anti-)neutrino and background events (atmospheric muons) have been adjusted and processed separately (compare Section 5.3). Altogether, 23 setups of rather homogeneous conditions have been defined (see Table 5.1). The effective lifetime of these setups is ranging from less than two days to almost forty days. In total, the data corresponds to an effective data taking time of about 250 days. As can be seen from the table, the assignment of the runs to the setups is not strictly chronological, implying that several succeeding runs can belong to different setups, depending for example on their background rate.

**Table 5.1:** List of the 23 data setups. Simulations are adjusted separately to each of the setups, incorporating the respective detection conditions like background rate and trigger settings.

| Setup | First Date | Last Date | Missing Lines | Mean Rate [kHz] | Bad Channel Ratio | Life-time [d] | Trigger 3N | Trigger 3T | Threshold of big *L1* hits [pe] |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 27.03.07 | 29.05.07 | 6 - 12 | 66.8 | 0.11 | 21.05 | x | - | 10 |
| 2 | 29.05.07 | 03.11.07 | 6 - 12 | 64.0 | 0.12 | 39.13 | x | - | 3 |
| 3 | 10.09.07 | 03.12.07 | 6 - 12 | 59.4 | 0.19 | 30.5 | x | - | 3 |
| 4 | 03.06.07 | 17.08.07 | 6 - 12 | 72.7 | 0.12 | 25.95 | x | - | 3 |
| 5 | 05.06.07 | 27.11.07 | 6 - 12 | 74.0 | 0.20 | 13.87 | x | - | 3 |
| 6 | 05.04.07 | 28.05.07 | 6 - 12 | 74.4 | 0.17 | 4.69 | x | - | 10 |
| 7 | 03.01.08 | 25.02.08 | 11, 12 | 56.9 | 0.18 | 9.94 | x | - | 3 |
| 8 | 22.12.07 | 27.02.08 | 11, 12 | 56.4 | 0.23 | 13.29 | x | - | 3 |
| 9 | 22.12.07 | 27.02.08 | 11, 12 | 67.3 | 0.24 | 9.49 | x | - | 3 |
| 10 | 27.02.08 | 02.03.08 | 11, 12 | 67.3 | 0.24 | 3.08 | x | x | 3 |
| 11 | 03.03.08 | 22.04.08 | 4, 11, 12 | 66.6 | 0.23 | 13.79 | x | x | 3 |
| 12 | 04.03.08 | 25.04.08 | 4, 11, 12 | 59.1 | 0.19 | 18.23 | x | x | 3 |
| 13 | 30.05.08 | 06.10.08 | - | 60.2 | 0.16 | 5.23 | x | x | 10 |
| 14 | 31.05.08 | 05.10.08 | - | 69.2 | 0.17 | 5.51 | x | x | 10 |
| 15 | 01.06.08 | 07.10.08 | - | 68.4 | 0.22 | 5.41 | x | x | 10 |
| 16 | 11.10.08 | 30.12.08 | - | 62.1 | 0.20 | 2.58 | x | x | 3 |
| 17 | 07.10.08 | 27.12.08 | - | 70.4 | 0.21 | 3.65 | x | x | 3 |
| 18 | 04.06.08 | 05.10.08 | - | 74.4 | 0.24 | 4.61 | x | x | 10 |
| 19 | 02.06.08 | 18.09.08 | - | 75.6 | 0.16 | 3.41 | x | x | 10 |
| 20 | 10.01.09 | 28.01.09 | 10 | 56.8 | 0.23 | 7.33 | x | x | 3 |
| 21 | 22.01.09 | 28.01.09 | 10 | 66.5 | 0.26 | 1.98 | x | x | 3 |
| 22 | 11.06.09 | 30.06.09 | 10, 12 | 65.5 | 0.26 | 5.56 | x | x | 3 |
| 23 | 18.08.09 | 20.08.09 | 9, 10, 12 | 63.7 | 0.17 | 1.84 | x | x | 3 |

### 5.2.3 Data processing chain

The first steps in the data processing chain are default procedures and are accomplished by means of so-called tasks provided within the *SeaTray* framework. They take care of reading event by event from the specified ANTARES run files (available in .root format) by implementing the required services and calling the corresponding modules. Additional information is gathered from the data base, like the valid calibration version and the requested geometry. If no aligned geometry is available, the nominal geometry is employed. In this case, all OMs of the affected events are flagged "unaligned" to store this information for later use. Furthermore, a filter is installed to discard so-called minimum bias events. They are not suitable for track reconstruction in general, because they contain only optical background[20]. Finally, the raw hits of the events are calibrated according to the actual calibration version.

After these standard tasks, the newly developed event-by-event selection of the data follows. In a first step, all hits measured by OMs in any of the failure states are discarded. Then the events are re-triggered with the simulated trigger. This is necessary because OMs being in state HIGH, for example, could have caused fake triggers. Furthermore, the offline calibration is taken into account for the re-triggering, which is recommended to be used instead of the preliminary online calibration considered for the online triggering of data.

The events are then passed to two filters. The first one discard events where no aligned geometry is available, the second one discards events with the mean rate or the bad channel ratio exceeding the limit of 80 kHz and 0.3, respectively. Additionally, all events triggered within the first 120 s at the beginning of each run are skipped to get rid of start-up problems due to the transition phase of the data taking system. Finally, linked to this event filtering procedure, the effective lifetime is calculated.

The remaining events are stored in the *SeaTray* specific .i3 data format and are passed to the reconstruction chain (Section 5.4). A scheme of the data processing chain can be seen in Fig. 5.1. The steering script to perform this data processing in *SeaTray* can be found in Appendix B.

## 5.3 Detector simulation and simulation processing chain

The general simulation procedure of physics events in the ANTARES detector (i.e. charged particles, that are traversing the detector) up to the level of the photocathode of the PMT has been described in Section 3.6. It already includes the angular acceptance and the detection efficiency of the PMT. Still, the simulation of the dynodes of the PMT and the ARS, that are responsible for the integration and recording of the PMT signals, as well as optical background and failures of OMs have yet to be

[20]The minimum bias events are triggered in certain time intervals, independent of any hit patterns. They are intended for example for background investigations.

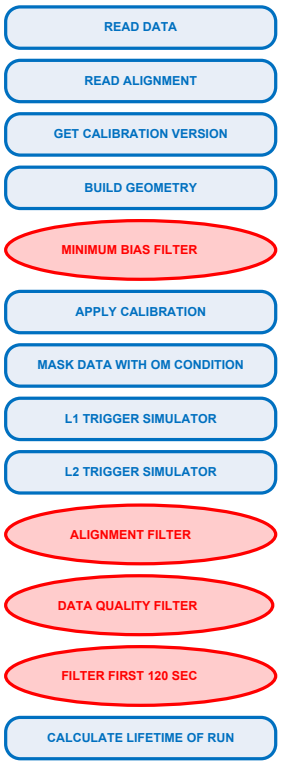



**Figure 5.1:** Scheme of the data processing chain of the present thesis. The data has been processed within the *SeaTray* software framework.

included. This simulation of the "real" detector is done within the *SeaTray* framework, which allows for an individual adjusting of a simulated event on the fly.

For the present study, simulated event samples were adjusted separately to each of the 23 data setups. All information necessary, like the mean background rate and information about the condition of the OMs, is gathered from the respective data setup the simulated events are being adjusted to. The simulated event sample incorporates atmospheric muons simulated with MUPAGE, as well as upward-going muon neutrino and muon anti-neutrino events (atmospheric flux, Bartol model) with an energy between $10 - 10^7$GeV. Not included are neither NC events, nor electron or tau (anti-)neutrinos, as their contribution to the number of triggered events in ANTARES can be safely neglected.

Random optical background hits are added to the raw 12-line simulations with a certain fixed rate for each OM. Currently, only single photoelectron hits with one dedicated rate for all OMs can be included in *SeaTray* - it is not (yet) possible to include *L1* hits, i.e. coincident hits or hits with large amplitude. After including the background hits, the simulation of the photomultiplier dynodes and the read-out electronics follows. Contrary to the simulation of the background, the electronics simulation is a default procedure in ANTARES.

The photocathode signals are not simulated as waveforms integrated over a certain time as done in reality. Instead, the process is simplified: the signal as well as the background hits are stored as single photoelectron hits, only containing position and timing information. The integration time and the two ARS chips in the OMs are

simulated such that when further hits appear on the same module within the integration gate of 33 ns, they are added up to the first one, resulting in a hit with an amplitude corresponding to the number of recorded hits and the time stamp of the first hit. After the integration gate of the first ARS has closed and after an additional dead time of about 8 ns, the second ARS takes over. Hits within the 8 ns dead time are discarded. The amplitude of the integrated signal is finally smeared by a Gaussian, whose width depends on the square of the integrated amplitude.

The effective differences between simulation and reality are subtle: in reality, the integration gate of 25 ns starts as soon as the signal has crossed the threshold. In order to account for the rising edge of the signal below the threshold, the integrated charge within the 8 ns time window before the threshold crossing is added. Furthermore, the integration of the second ARS is only triggered, when the signal has dropped beneath the threshold and crossed it again. In summary, these two factors lead to an overestimated charge in the simulations, though the effect is expected to be small in comparison to other effects like miscalibration.

Succeeding the dynode and ARS simulation is the simulation of OM failure states, that has been developed within the context of this thesis. The implementation is such, that dedicated OMs can be tagged as OFF, plus a certain number of randomly selected OMs changing from event to event. This accounts for the behaviour of the detector: some OMs never take part in the data taking process and a certain number of OMs are in (usually) transient states like HIGH or EMPTY. Consequently, those hits of the simulated events, that are to be detected by OMs tagged as OFF, are discarded.

Finally, the event building process with the snapshot window around the triggered hit sequence is emulated. At this level, the events are also stored in .i3 format and should mimic the data that have undergone the procedure described above, to a high degree. A scheme of the simulation processing chain can be seen in Fig. 5.2. The steering script to perform this simulation processing in *SeaTray* can be found in Appendix B.

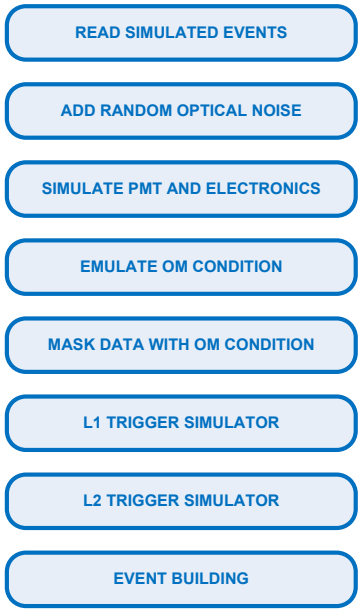



**Figure 5.2:** Scheme of the simulation processing chain of the present thesis. The simulations have been processed within the *SeaTray* software framework.

Yet, discrepancies between data and simulations still exist, though they are typically decreasing with more stringent hit selections. As examples, low-level distributions of four arbitrarily selected data runs from four different setups (Table 5.2) and the respective simulations are shown. The distributions of the number of hits per event and the distributions of the integrated charge of the events are shown in Figs. 5.3 – 5.6 for *L0* hits, *L1* hits and *L2* hits. For better comparability, the distributions are normalised relative to each other. The distributions of the integrated charge of the events exhibit larger differences than the distributions of the number of hits. Reasons for the discrepancies in the charge distributions are manifold: besides the inaccuracy in the ARS simulation discussed above, also a bad charge calibration can become evident at this level. Furthermore, hits with amplitudes larger than one photoelectron (pe) or coincident hits can also occur in background processes. This is currently not not simulated in *SeaTray*.

**Table 5.2:** Runs for the low-level data-simulation comparison.

| Run | Setup | Date | Missing Lines | Mean Rate [kHz] | Bad Ch. Ratio | Threshold of big *L1* hits [pe] |
|---|---|---|---|---|---|---|
| 30208 | 3 | 06.11.07 | 6 - 12 | 61 | 0.18 | 3 |
| 33617 | 11 | 18.04.08 | 4,11,12 | 65 | 0.2 | 3 |
| 34711 | 13 | 09.06.08 | - | 61 | 0.16 | 10 |
| 38084 | 16 | 23.12.08 | - | 62 | 0.21 | 3 |

That a larger percentage of big and/or coincident hits might indeed be to blame (independently from what causes them) becomes obvious from the distributions of the event duration with respect to *L1* and *L2* hits, i.e. the time interval between first and last *L1* and *L2* hit, respectively (Figs. 5.7 – 5.10). Especially the *L1* time interval is, on average, larger in data than in simulation. The first peak around 300 – 400 ns is due to the muon event, whereas the second peak, starting at about 1 200 ns, is due to *L1* hits within the background. As some studies indicate too few coincident hits in the simulation [117] and others too few big hits [118], the problem appears to be complex and further investigation is necessary in the future.

As already mentioned and obvious from Figs. 5.3 – 5.6, a general improvement in the agreement between data and simulation with the more stringent hit selections can be observed. This also becomes evident from the distributions of the number of lines with *L1* and *L2* hits, respectively. Generally, an acceptable agreement on the *L2* hit level can be observed providing confidence in this hit selection. The *L2* hits are consequently used as a basis for the *Posidonia* event classifier and the linear prefit.

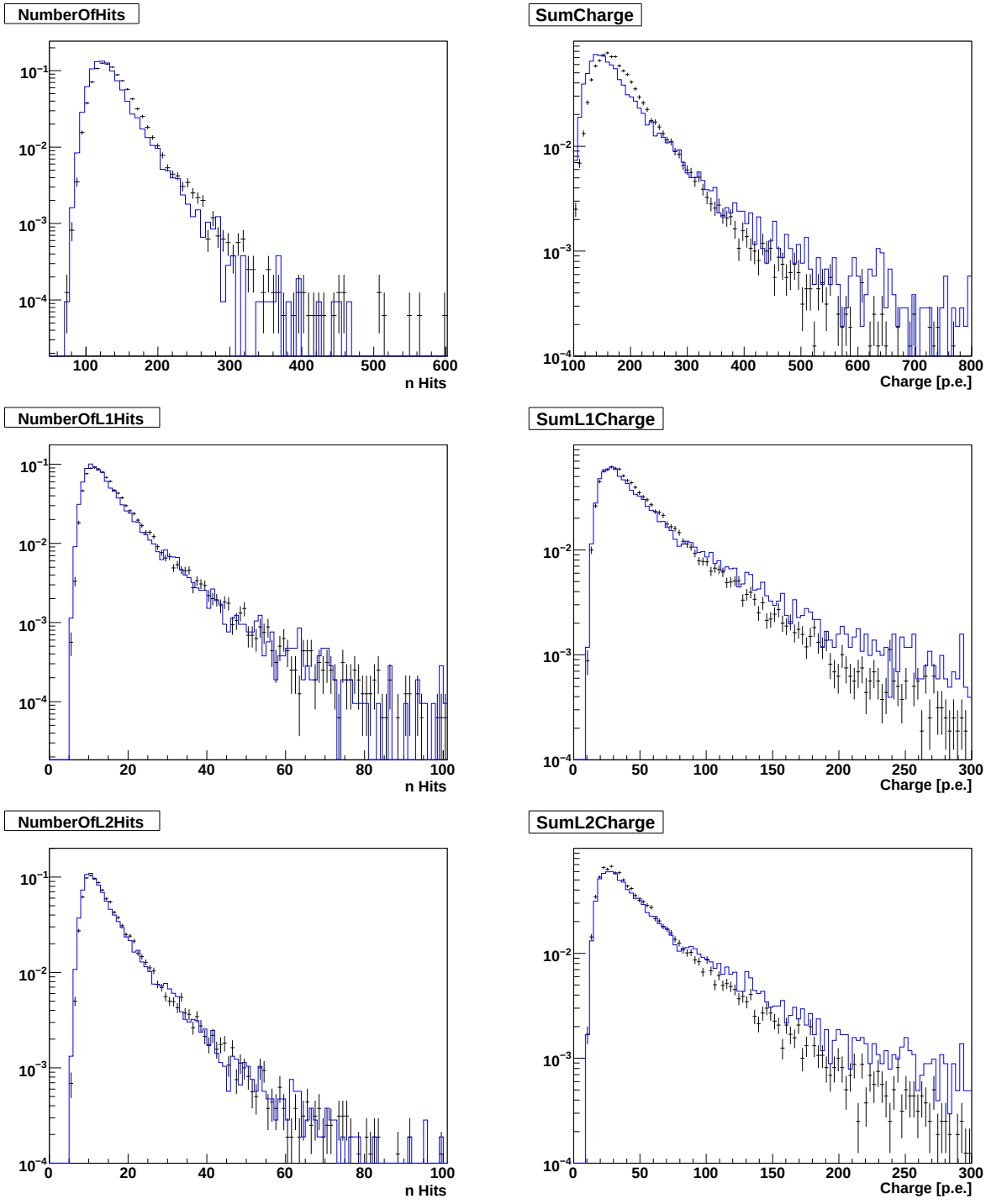



**Figure 5.3:** Low-level distributions of ANTARES run **30208** (black) and the corresponding simulation (blue line). The distributions are normalised relative to each other.
*Left:* Distribution of the number of hits per event (from top to bottom: *L0*, *L1*, *L2*).
*Right:* Distribution of the integrated charge (i.e. sum of the hit amplitudes) of the events in photoelectrons. Same order as left column.

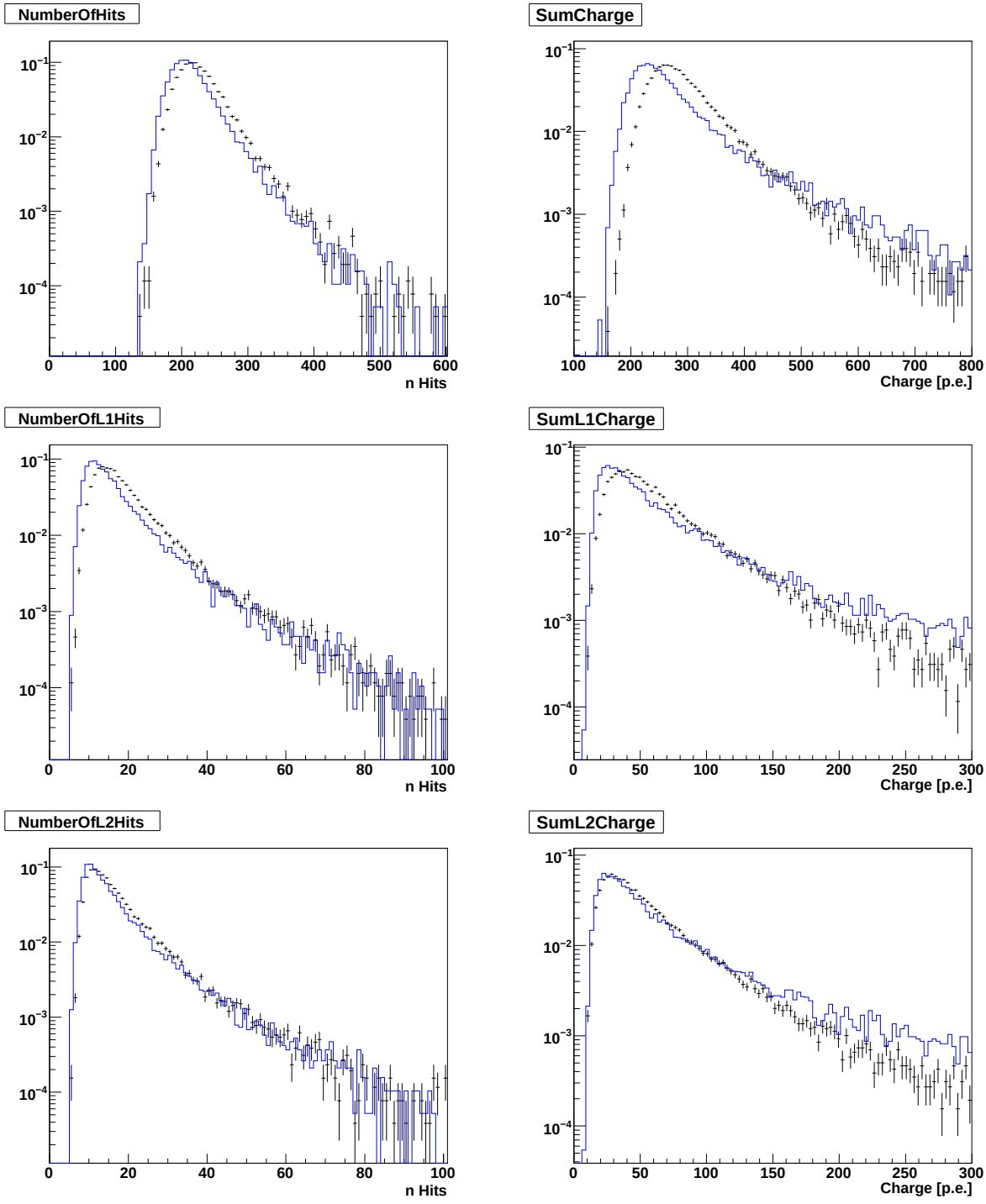



**Figure 5.4:** Low-level distributions of ANTARES run **33617** (black) and the corresponding simulation (blue line). The distributions are normalised relative to each other.
*Left:* Distribution of the number of hits per event (from top to bottom: *L0*, *L1*, *L2*).
*Right:* Distribution of the integrated charge (i.e. sum of the hit amplitudes) of the events in photoelectrons. Same order as left column.

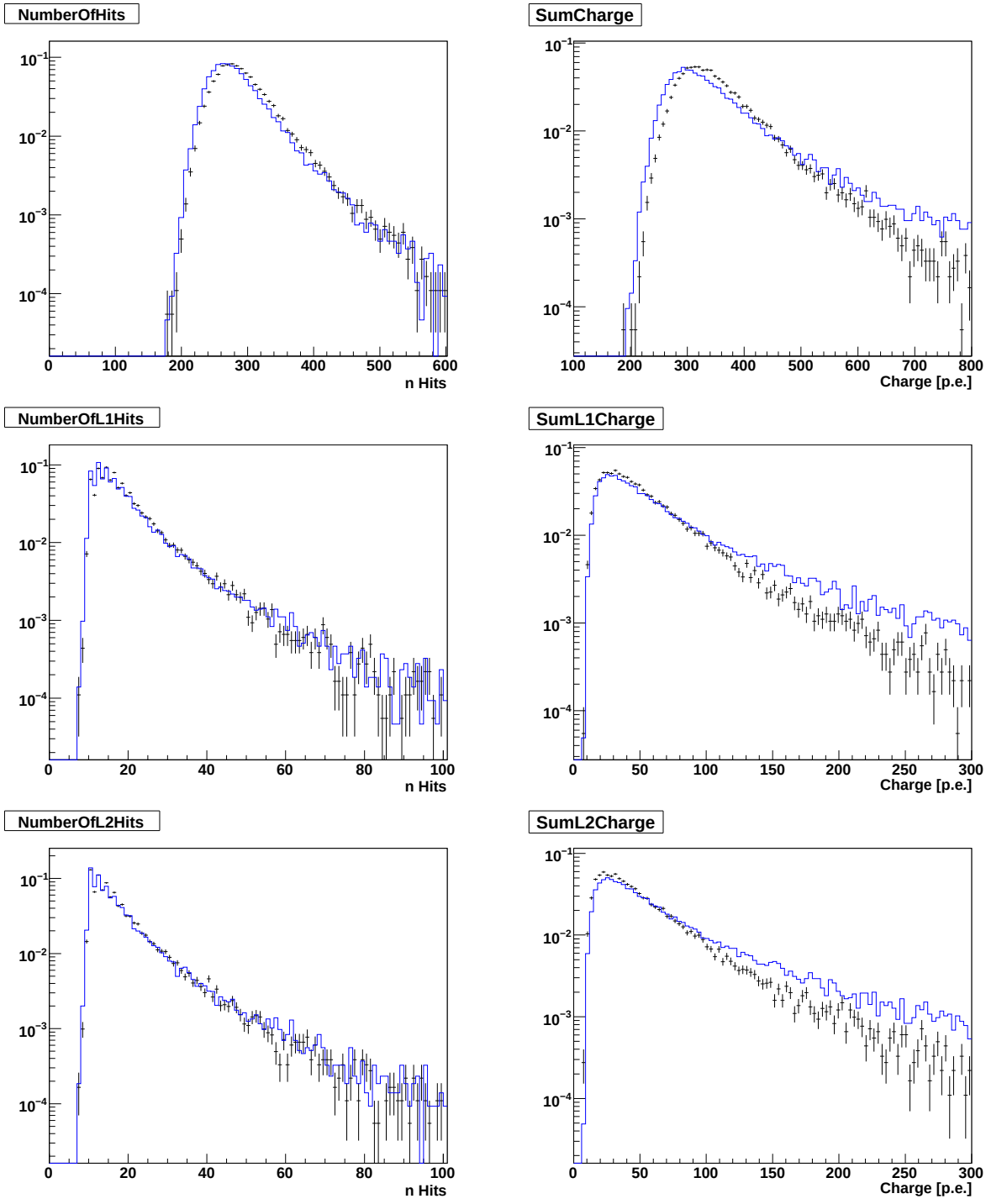



**Figure 5.5:** Low-level distributions of ANTARES run **34711** (black) and the corresponding simulation (blue line). The distributions are normalised relative to each other.
*Left:* Distribution of the number of hits per event (from top to bottom: *L0*, *L1*, *L2*).
*Right:* Distribution of the integrated charge (i.e. sum of the hit amplitudes) of the events in photoelectrons. Same order as left column.

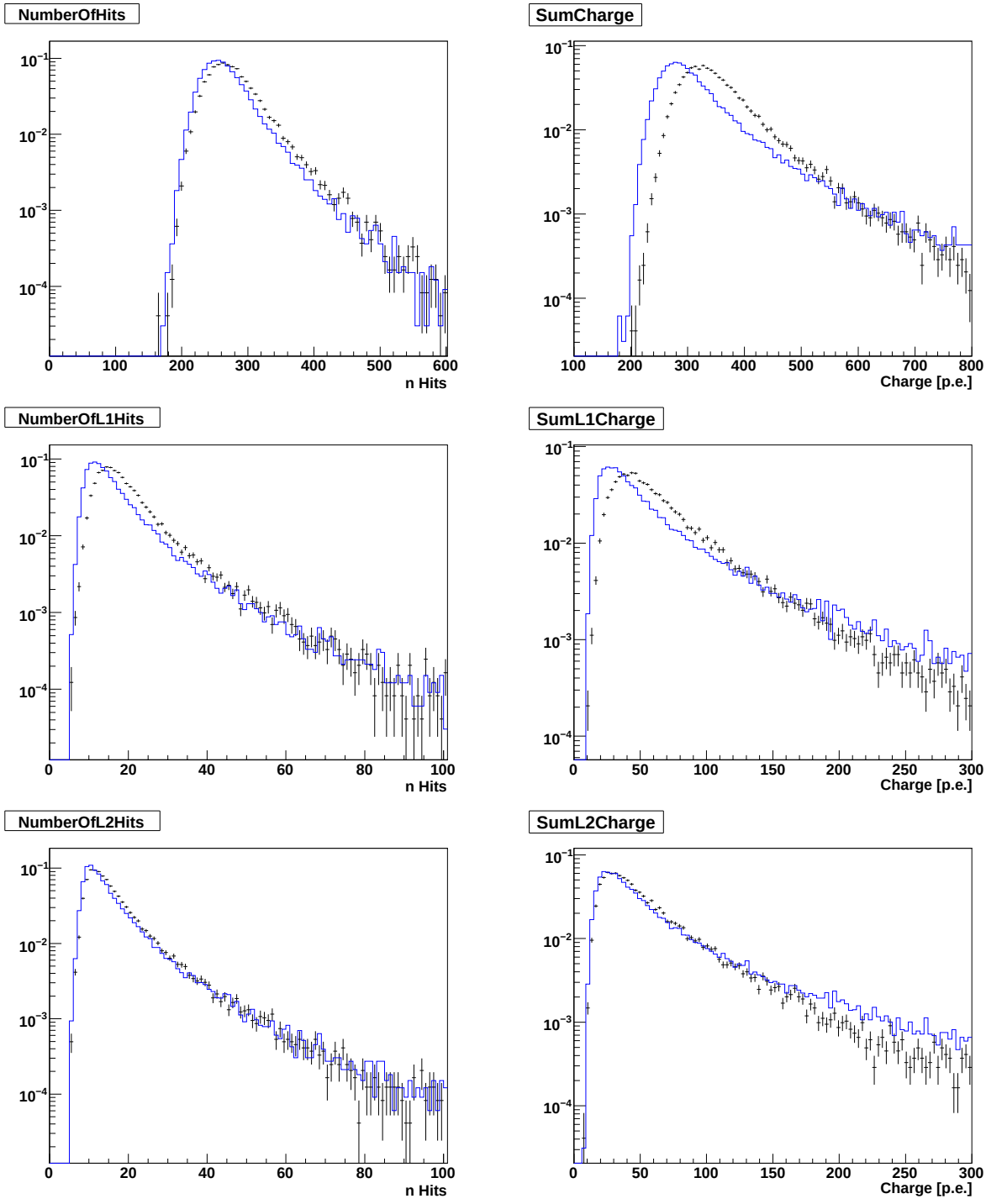



**Figure 5.6:** Low-level distributions of ANTARES run **38084** (black) and the corresponding simulation (blue line). The distributions are normalised relative to each other.
*Left:* Distribution of the number of hits per event (from top to bottom: *L0*, *L1*, *L2*).
*Right:* Distribution of the integrated charge (i.e. sum of the hit amplitudes) of the events in photoelectrons. Same order as left column.

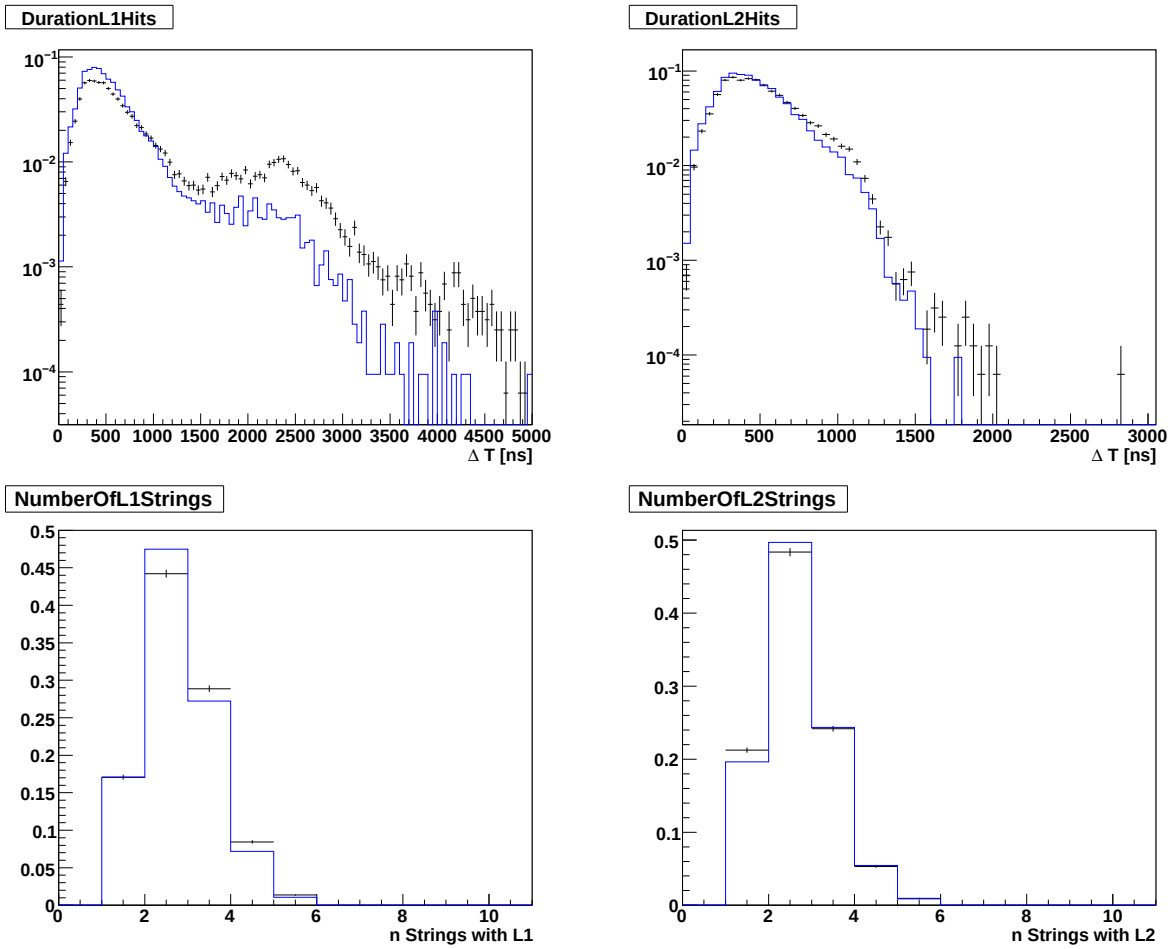



**Figure 5.7:** Low-level distributions of ANTARES run **30208** (black) and the corresponding simulation (blue line). The distributions are normalised relative to each other.
*Upper row:* distribution of the time difference between first and last *L1* hit (*left*) and *L2* hit (*right*), respectively.
*Lower row:* distribution of the number of lines with *L1* hits (*left*) and *L2* hits (*right*).

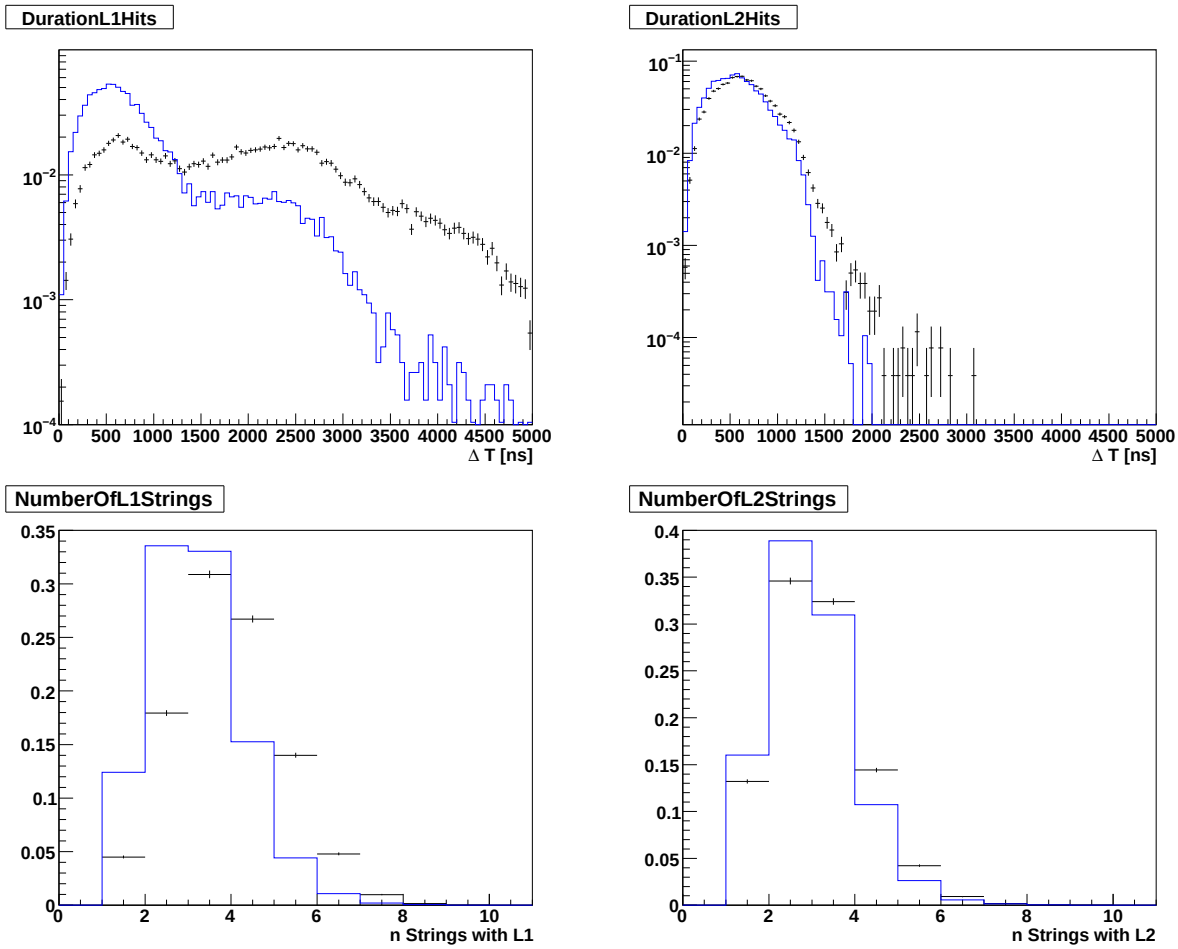



**Figure 5.8:** Low-level distributions of ANTARES run **33617** (black) and the corresponding simulation (blue line). The distributions are normalised relative to each other.
*Upper row:* distribution of the time difference between first and last *L1* hit (*left*) and *L2* hit (*right*), respectively.
*Lower row:* distribution of the number of lines with *L1* hits (*left*) and *L2* hits (*right*).

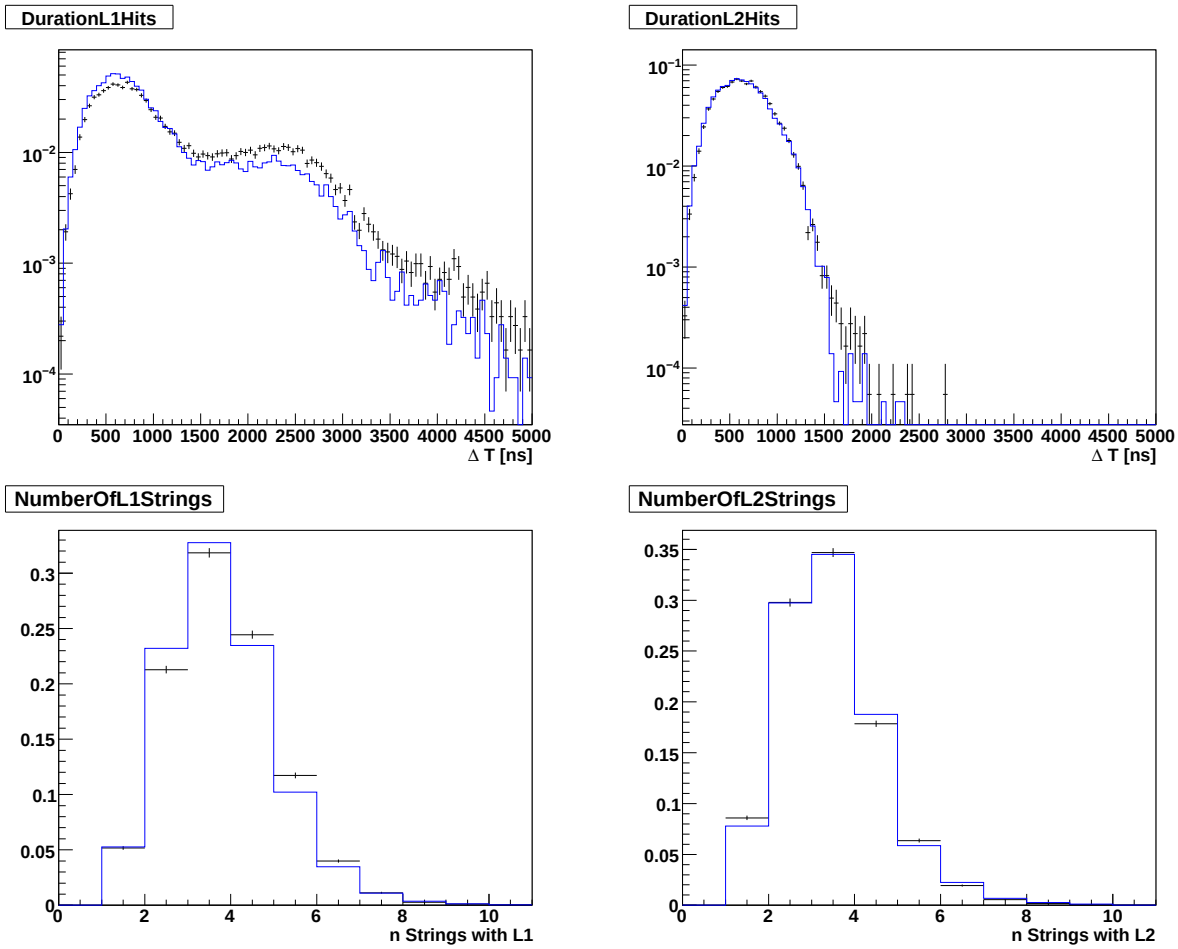



**Figure 5.9:** Low-level distributions of ANTARES run **34711** (black) and the corresponding simulation (blue line). The distributions are normalised relative to each other.
*Upper row:* distribution of the time difference between first and last *L1* hit (*left*) and *L2* hit (*right*), respectively.
*Lower row:* distribution of the number of lines with *L1* hits (*left*) and *L2* hits (*right*).

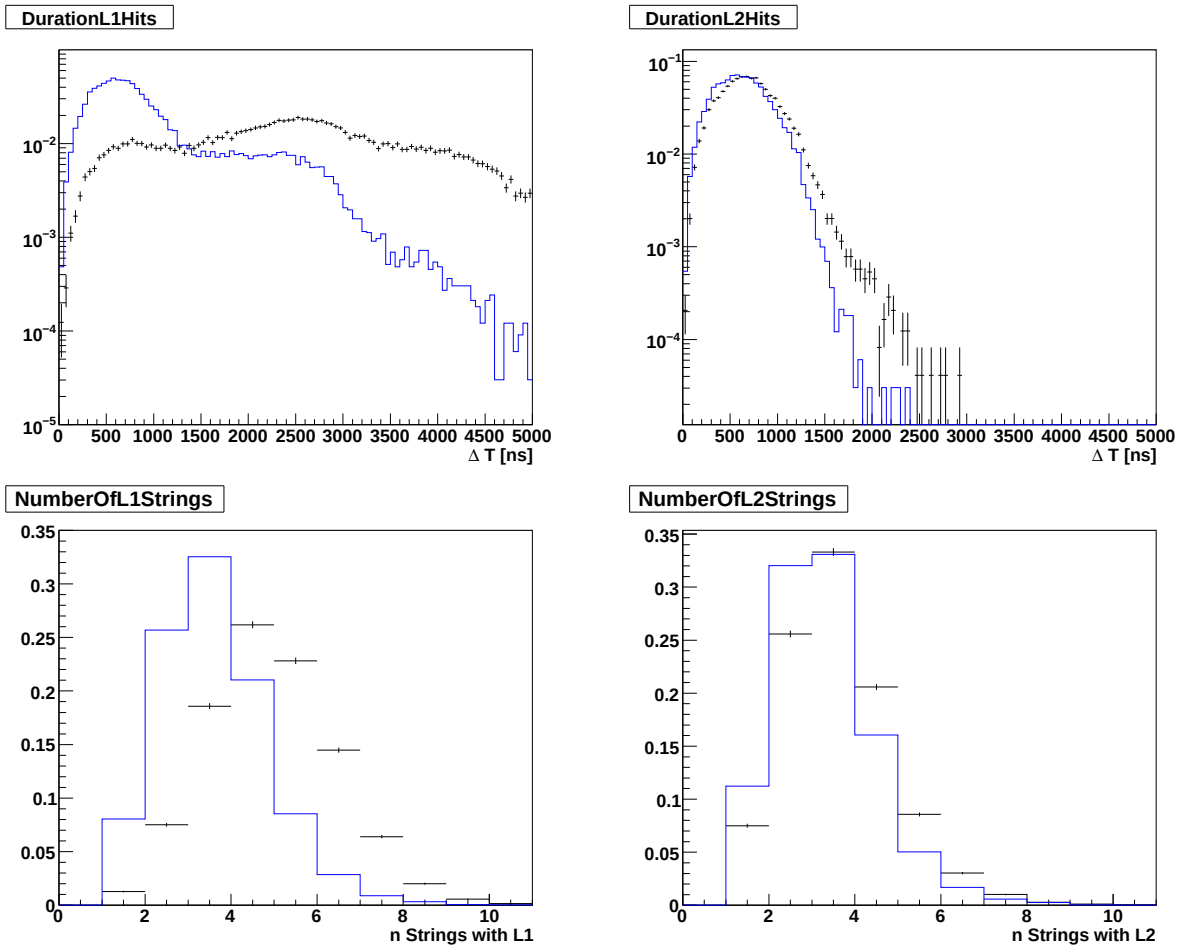



**Figure 5.10:** Low-level distributions of ANTARES run **38084** (black) and the corresponding simulation (blue line). The distributions are normalised relative to each other.
*Upper row:* distribution of the time difference between first and last *L1* hit (*left*) and *L2* hit (*right*), respectively.
*Lower row:* distribution of the number of lines with *L1* hits (*left*) and *L2* hits (*right*).

The agreement between data and simulation of the distributions of the *HM* hits ranges somewhere between the agreement of the *L1* and *L2* selection. This is understandable, as, conversely to the *L1* selection, the *HM* selection bases on causality connections, and the *L2* hits are a subset of the *HM* hits. The distributions of the *HM* hits of the four example runs can be see in Figs. 5.11 – 5.14.

A compromise had to be found between having a sufficient number of hits left after the basic hit selection and, at the same time, having a good agreement between data and simulations. The *HM* selection was considered as acceptable to serve as a basis for the *Posidonia* reconstruction algorithm, which moreover contains further dedicated single- and multi-string hit selections. The efficiency of only the *L2* hit selection would have been too low to serve as initial hit sample, i.e. there would have been too few hits left after the *L2* hit selection, especially for events with lowest energy $E < 50\,\mathrm{GeV}$.

None of the deviations between data and simulation is yet perfectly understood. They are not a specific problem of the described data selection procedure nor the data and simulation processing chain. Attempts made in the ANTARES collaboration to emulate the detection conditions using another software program and following a different approach, showed similar results. The extend of the disagreement varies with the runs and no underlying pattern could be identified. Besides discrepancies in the shape of the individual distributions, disagreements in the total number of triggered events can be observed. The reasons are still unknown and need to be carefully investigated in the future. Trying to understand the differences is a major task that has already started to be tackled and improvements in the agreement can be expected within the next months. Lastly, efforts within the collaboration of considering hits with large amplitudes in the optical background, resulted in an overall improvement of the data-simulation agreement. For this analysis, this innovation could not be included anymore.

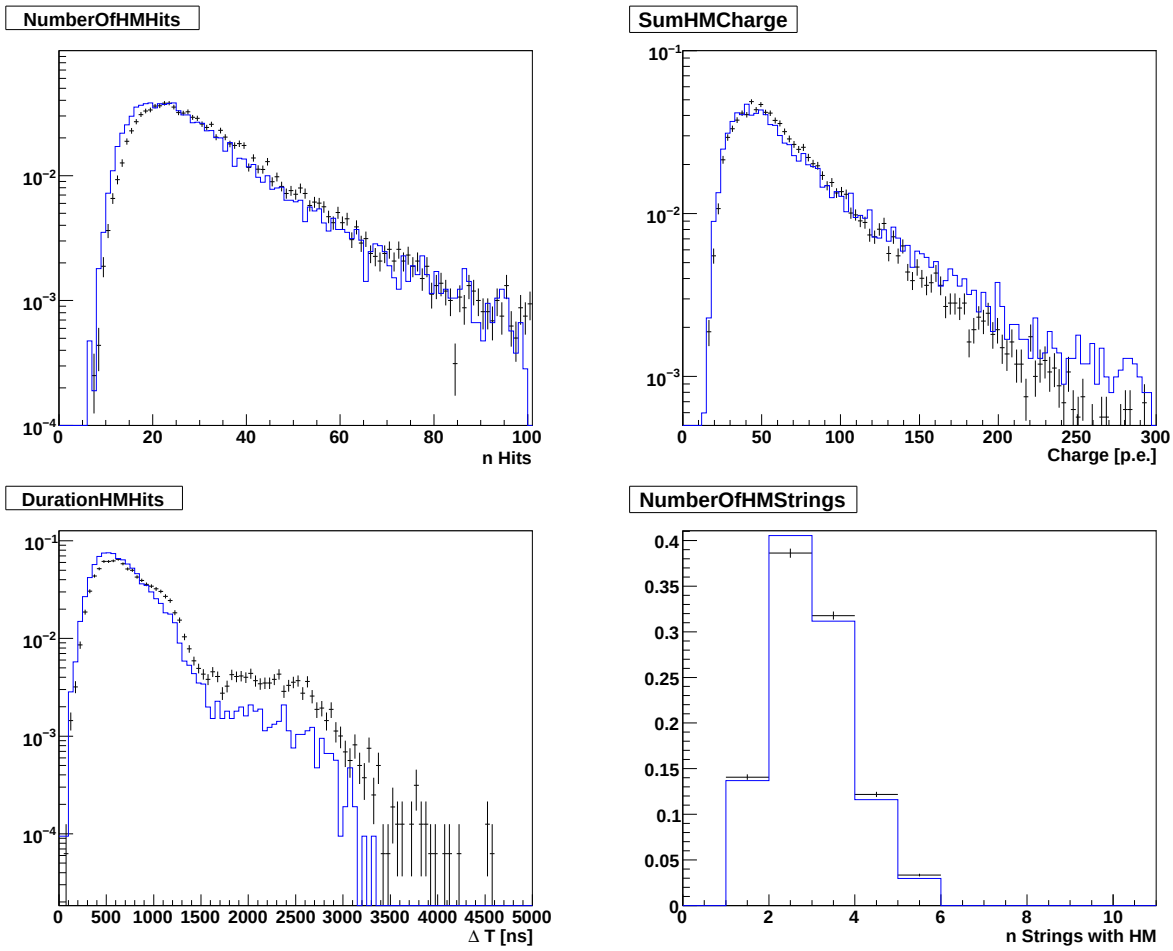



**Figure 5.11:** Low-level distributions of the *HM* hits of ANTARES run **30208** (black) and the corresponding simulation (blue line). The distributions are normalised relative to each other.
*Upper row*: distribution of the number of *HM* hits (*left*) and distribution of the integrated charge of the *HM* hits (*right*).
*Lower row*: distribution of the time difference between the first and the last *HM* hit of an event (*left*) and the number of lines with *HM* hits (*right*).

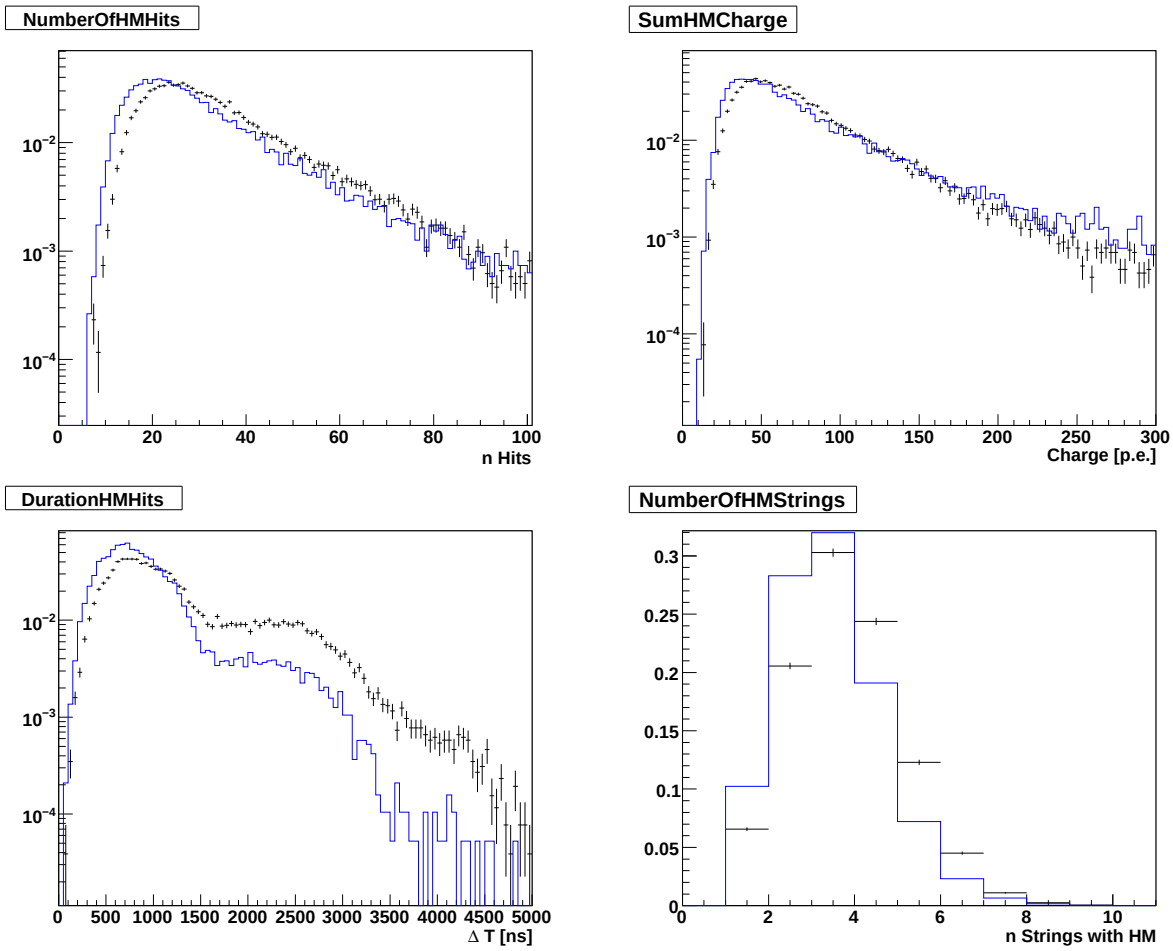



**Figure 5.12:** Low-level distributions of the *HM* hits of ANTARES run **33617** (black) and the corresponding simulation (blue line). The distributions are normalised relative to each other.
*Upper row*: distribution of the number of *HM* hits (*left*) and distribution of the integrated charge of the *HM* hits (*right*).
*Lower row*: distribution of the time difference between the first and the last *HM* hit of an event (*left*) and the number of lines with *HM* hits (*right*).

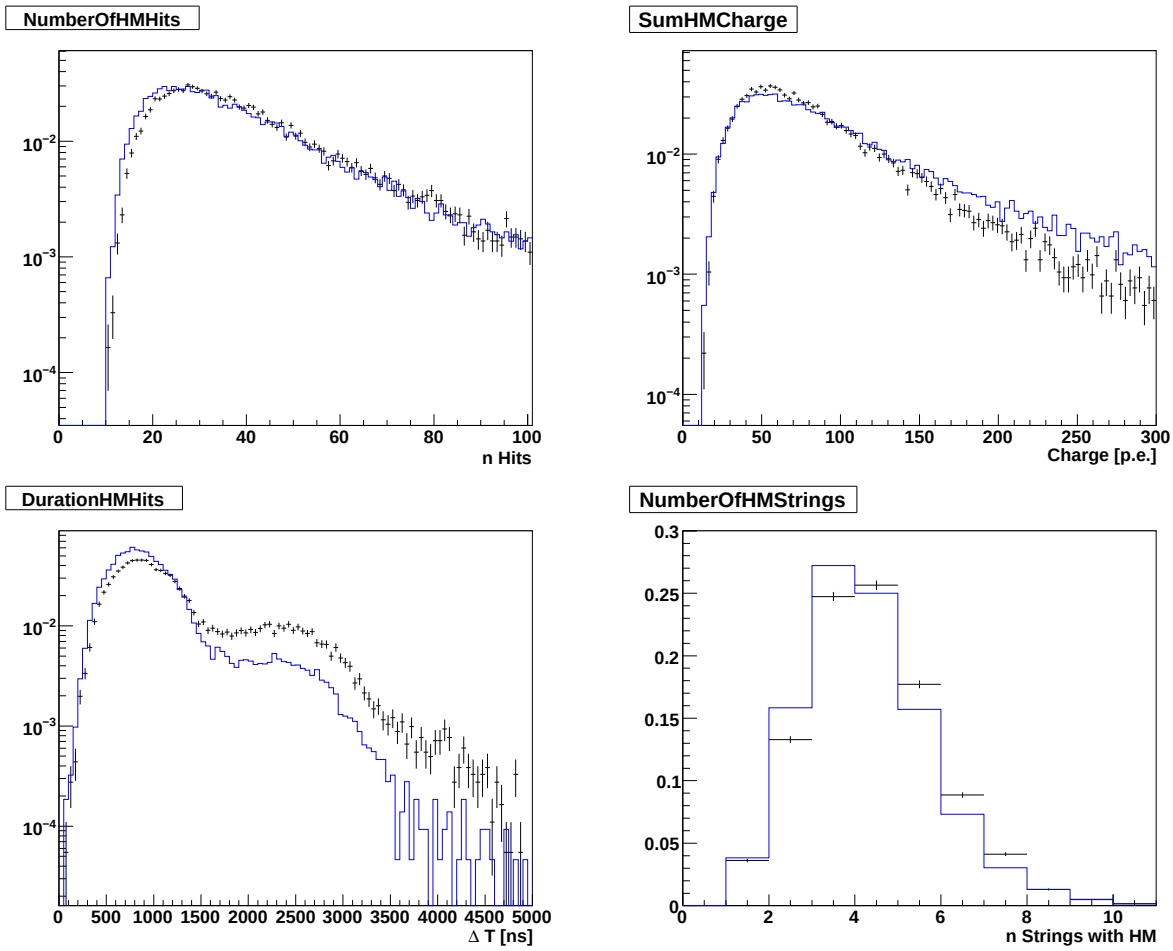



**Figure 5.13:** Low-level distributions of the *HM* hits of ANTARES run **34711** (black) and the corresponding simulation (blue line). The distributions are normalised relative to each other.
*Upper row*: distribution of the number of *HM* hits (*left*) and distribution of the integrated charge of the *HM* hits (*right*).
*Lower row*: distribution of the time difference between the first and the last *HM* hit of an event (*left*) and the number of lines with *HM* hits (*right*).

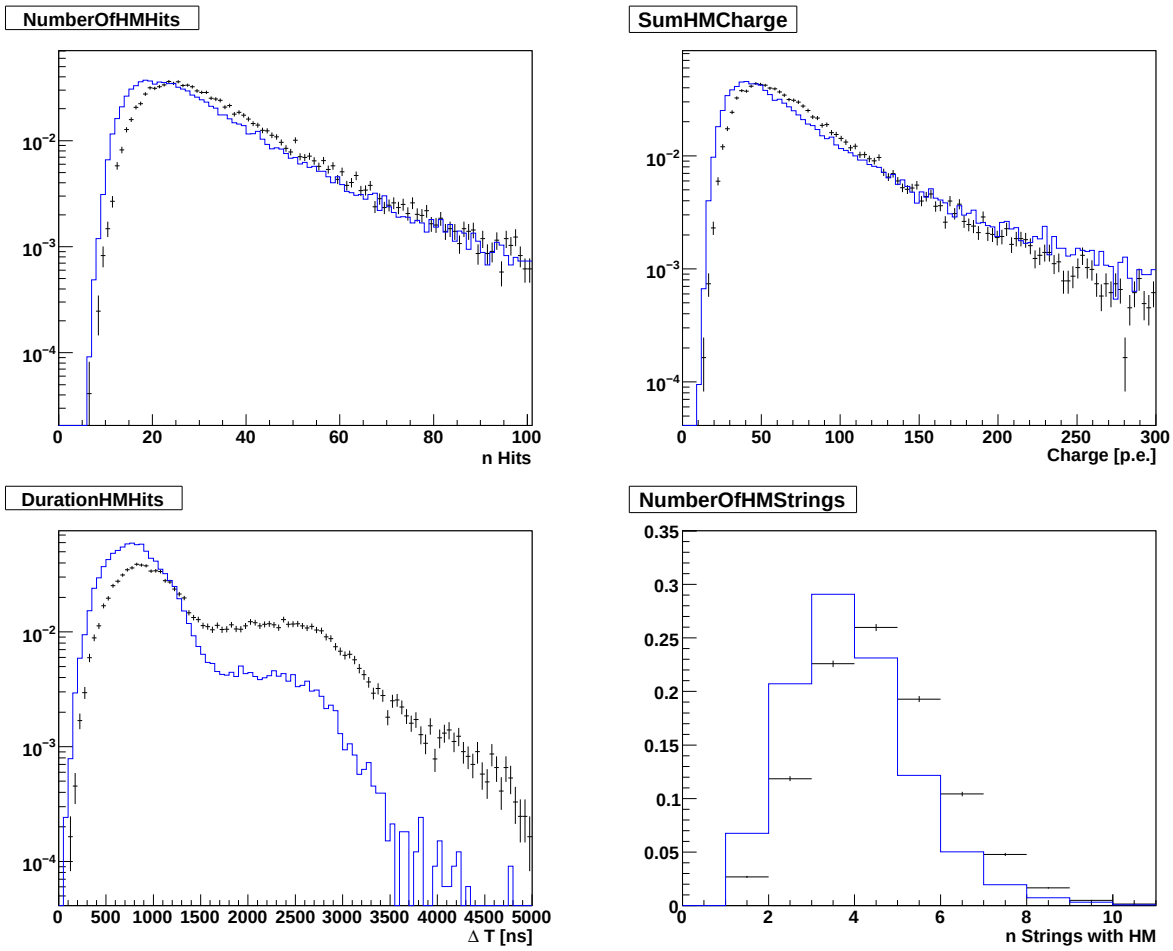



**Figure 5.14:** Low-level distributions of the *HM* hits of ANTARES run **38084** (black) and the corresponding simulation (blue line). The distributions are normalised relative to each other.
*Upper row*: distribution of the number of *HM* hits (*left*) and distribution of the integrated charge of the *HM* hits (*right*).
*Lower row*: distribution of the time difference between the first and the last *HM* hit of an event (*left*) and the number of lines with *HM* hits (*right*).

## 5.4 Analysis reconstruction chain

After data and simulations have undergone their respective processing chain, they are passed to a common reconstruction chain. This sequence of *SeaTray* modules does not only include the mere reconstruction algorithm but also a basic filter for the rejection of atmospheric muon events. Their number typically exceeds the number of upward-going neutrino events by several orders of magnitude (see also Fig. 3.11). The *Posidonia* reconstruction algorithm is not optimised concerning computing time and it would therefore be impossible to reconstruct all recorded ANTARES events within an acceptable amount of time, not even by employing a powerful computing cluster. To avoid the time consuming reconstruction of clear atmospheric downward-going muon signatures with *Posidonia*, a basic cut was applied on the data. For this purpose, the *BBfit* reconstruction algorithm was employed, which proved to be quite robust.

The distributions of the *BBfit*-reconstructed zenith angles of the events of the four example runs from the previous subsection are shown in Fig. 5.15. For better comparability, the distributions are normalised relative to each other. The agreement of the distributions is highly acceptable, especially when considering the fact that the distributions contain (almost) exclusively atmospheric muons Such atmospheric muon events have a large fraction of Cherenkov photons that are multiple scattered or that hit the acceptance areas of the OMs in regions where the uncertainties are large.

For further processing, all events with a valid *BBfit*-reconstructed zenith angle of less than 115° are discarded. Only events without valid *BBfit* result, or with a zenith angle of more than 115° (i.e. upward-going) are passed to the *Posidonia* algorithm. This procedure reduces the amount of data by 60 – 80 %, depending on the optical background level and active triggers. On the other hand, the number of upward-going low-energy neutrinos with a true (simulated) neutrino incident angle of more than $\Theta_\nu = 115°$ were only reduced by about 4 %.

Later on, cuts on the final *Posidonia*-reconstructed zenith angle are applied at $\Theta_{\text{Pos.}} = 115°$ for multi-string and $\Theta_{\text{Pos.}} = 135°$ for single-string events, to suppress misreconstructed atmospheric muons and to select vertical single-string events for the containment selection. The cut on $\Theta_{\text{BBfit}} = 115°$ was chosen with respect to these subsequent cuts and, on the other hand, with respect to the need for efficient data reduction.

The *BBfit* reconstruction algorithm is followed by the *HM* hit selection, the *Posidonia* event classifier and, for the multi-string events, the linear prefit. To further reduce the amount of data, a cut at $\Theta_{\text{Prefit}} = 115°$ is imposed on the zenith angle of the linear prefit for events without valid *BBfit* result. The reconstruction efficiency of the linear prefit is higher than the efficiency of *BBfit*, but the angular resolution is slightly worse.

The remaining events are passed to the full *Posidonia* reconstruction procedure for single- and multi-string events, respectively. A scheme of the analysis reconstruction chain is displayed in Fig. 5.16. The steering script to process this reconstruction chain in *SeaTray* can be found in Appendix B.

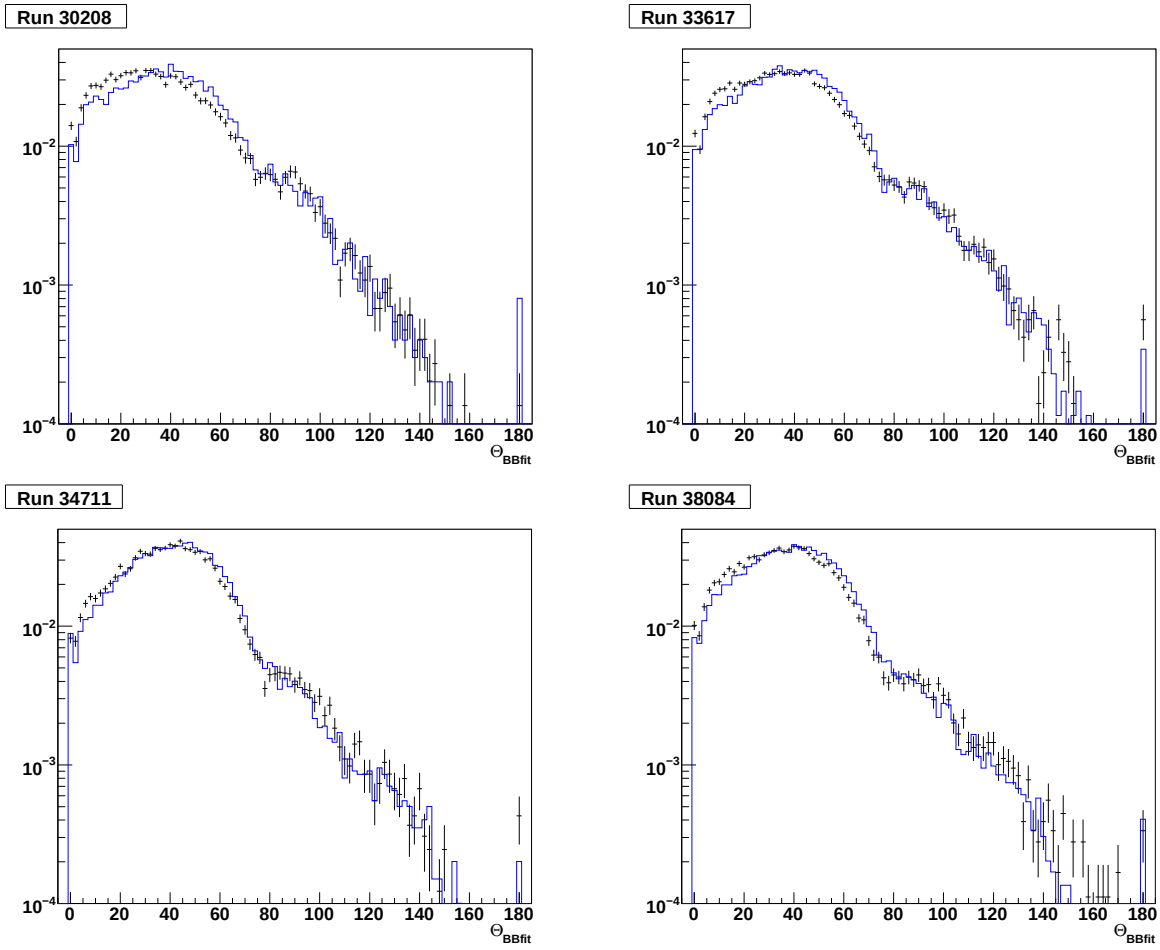



**Figure 5.15:** Zenith angle of four ANTARES runs, reconstructed with the *BBfit* algorithm. Data (black) and simulation (blue line) are normalised relative to each other. Events with $\Theta_{\mathrm{BBFit}} < 115°$ are discarded for further processing.

## 5.5 Event selection

In Fig. 5.17, the reconstructed zenith of the full simulation sample (FSS), i.e. the combination of all simulations of the 23 setup samples is shown. Except for the preselection cuts described in the last section, no further cuts are applied. Also indicated in the figure are the well reconstructed upward-going neutrino events, which directly demonstrates the need for effective quality cuts: the overwhelmingly large background of misreconstructed atmospheric muons has to be suppressed to detect the tiny signal within. This is typically done by cutting on parameters that are indicating the fit quality. This cutting procedure can be enhanced by making use of multivariate analysis methods, like an artificial neural network (ANN). The parameters that are used as input to the ANN are illustrated within the first subsection, as well as the functioning of the

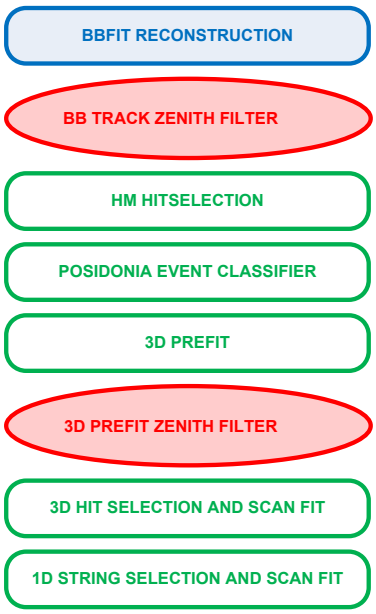



**Figure 5.16:** Scheme of the reconstruction chain used for the analysis.

ANN. The efficiency of the quality cuts are presented, as well as relevant distributions after application of the quality cuts.

Further event selection is necessary to finally obtain a set of low-energy neutrinos for an oscillation analysis, and to be able to reconstruct the energy reliably. For this purpose, after application of the quality cuts the containment estimation method described in Chapter 4 is employed. The performance of the containment estimation on the simulated events is evaluated in the second subsection, and the relevant distributions are shown.

### 5.5.1 Selection of well reconstructed events

#### Principle and performance

The separation of well reconstructed from badly reconstructed events is done by means of an artificial neural network. For this purpose the "MLP" network, one of the classifiers embedded in TMVA (Toolkit for Multivariate data Analysis with ROOT [119]), was employed. TMVA is a ROOT-based software package containing various classifiers. They can be used to combine several unsatisfactorily separating distributions of parameters to a single cut parameter according to a mathematical function mapping: N:= $p_i \longmapsto \mathrm{N}(p_i) = a$, with the input parameters $p_i$ and the output parameter $a$. All classifiers need to be trained by providing them a representative sample of (simulated) data together with the information which events are "signal" ($\widehat{=}$ well reconstructed events, here: $\Delta\Theta < 5°$) and which are "background" ($\widehat{=}$ badly reconstructed events, here: $\Delta\Theta > 10°$).

In the remaining event sample (after the cuts on the *Posidonia*-reconstructed zenith and on the output of the ANN), the contamination with badly reconstructed events (almost exclusively atmospheric muons) has to be minimised, while maximising at the

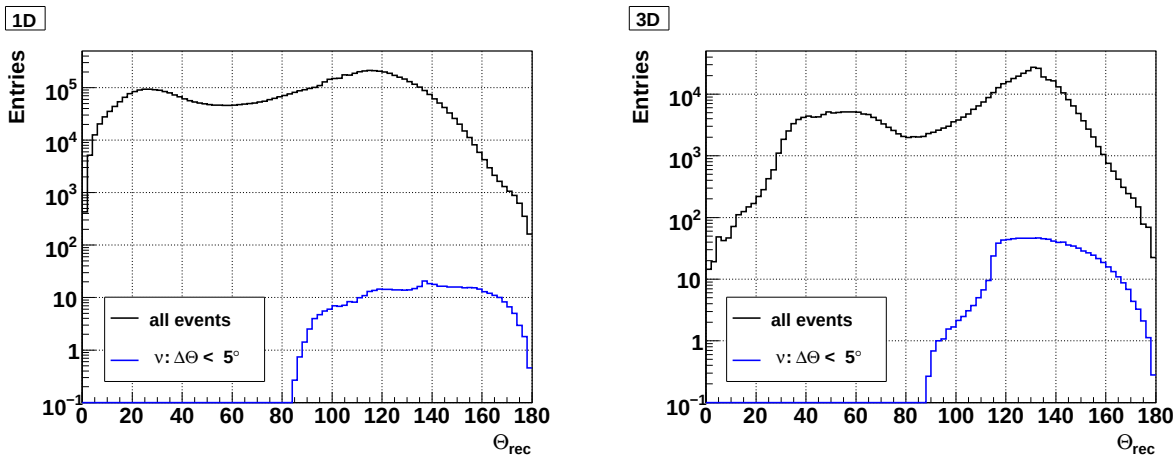



**Figure 5.17:** Reconstructed zenith of the full simulation sample (FSS) after the complete reconstruction chain, without any quality cuts (only the preselection cuts described in Section 5.4 are applied). All reconstructed events (black) and upward-going neutrino events (blue) with a reconstruction error on the zenith angle of less than 5° (i.e. "well" reconstructed events). *Left:* single-string reconstruction. *Right:* multi-string reconstruction.

same time the number of well reconstructed events. Intermediate events with an error on the reconstructed zenith angle between five and ten degrees are acceptable to be kept in the remaining event sample, but they are not of importance for the choice of the cut value. Hence, for the training of the ANN, the definition of signal and background was chosen such that two distinct samples of well and badly reconstructed events are used.

Single- and multi-string events are treated separately with two ANNs trained individually. This is necessary because the input distributions for the two event classes exhibit different shapes and mean values. For the multi-string reconstruction, six parameters were ascertained as "quality" parameters, suitable as input parameters to the ANN:

- Three fit error estimates from the *Posidonia* final fitting routine calculated from the singular values of the Jacobian matrix for the zenith, the azimuth and the time, given as decadic logarithm.

- The number of unscattered Cherenkov photons emitted by the muon track, referred to as "number of direct hits". This value is calculated by using the fitted track as reference and select only hits with time residuals between $-5$ and $+10$ ns.

- The minimal negative log-likelihood $-\ln\mathcal{L}$ of each event divided by the number of degrees of freedom, i.e. the number of hits used for the fit. This fraction is also given as decadic logarithm $\log(-\ln\mathcal{L}/\mathrm{ndof})$.

- The absolute difference between the final zenith of the fit with *Posidonia* and the

linear prefit $\Delta\Theta = |\Theta_{\mathrm{FinalFit}} - \Theta_{\mathrm{PreFit}}|$. The prefit, which is only used for hit selection purposes, is quite robust though not very precise. If the final fit solution differs significantly from the solution of the prefit, this is indicating very likely a bad quality of the final fit.

The distributions of these parameters can be seen in Fig. 5.18. They are shown separately for signal events (upward-going neutrinos, reconstructed within $\Delta\Theta < 5°$) and for two types of background events: upward-going neutrinos reconstructed with $\Delta\Theta > 10°$ and misreconstructed atmospheric muons with a reconstructed zenith of more than 115°. The peaks of signal and background distributions of the input parameters are well distinguishable from each other. Employing the powerful ANN optimises the cut on combinations of these input parameters.

No prefit is available for single-string events, making this input parameter obsolete. Therefore, only the first five parameters described above serve as input variables to the ANN. The error estimate on the azimuth is also employed as quality parameter, even if the azimuth itself is not of relevance for the single-string fit. The individual parameter distributions are slightly degraded compared to the multi-string fit in terms of separation power. The corresponding distributions can be seen in Fig. 5.19. Obviously, direct and consecutive cuts on the parameters will not give satisfying results and the usage of multivariate data analysis methods like the ANN are mandatory.

Attempts were made to obtain a further quality parameter by performing the single-string fit twice. While the first fit was applied as before, hits from other strings were included in the second fit procedure, which are close to the trajectory obtained from the first fit. The result was not satisfying and did not enhance the quality selection procedure.

The MLP neural network is said to perform stable even with linearly and non-linearly correlated input variables [120]. Hence, no bias has to be expected by usage of linearly correlated parameters and all quality parameters described above can be employed as input to the respective (single- or multi-string) network. The output value of the ANN finally ranges between $a = -1$ and $a = 1$. The smaller the value, the more likely the event is a background event.

The Monte-Carlo simulation sample used for training and testing the network incorporates upward-going neutrino events and atmospheric muons with a 60 kHz optical background rate. The detector geometry employed was the perfectly aligned 12-line detector with all OMs active. For low-energy neutrinos up to 200 GeV, the sample is identical to the ideal reference sample (IRS, see Sec. 4.1). This training and testing sample will be referred to in the following as "global test sample" (GTS).

Before training, the preselection cut on *BBfit* at $\Theta_{\mathrm{BBfit}} = 115°$ (see previous subsection) was applied, corresponding to the procedure in the final processing scheme. Additionally, only events having a *Posidonia*-reconstructed zenith of more than $\Theta_{\mathrm{Pos.}} = 115°$ were accepted for the training[21].

[21] For both single- and multi-string events. The more restrictive cut on $\Theta_{\mathrm{Pos.}} = 135°$ for the single-string events is applied afterwards.

After training the network, its performance is evaluated by means of a test sample, which is identical to the sample used for training itself but contains different events. In Fig. 5.20, the results of the ANN training are shown: the distributions of the ANN output for signal and background are depicted. Efficiency, purity and background rejection are given as a function of the cut value. Finally, the so called ROC curve (receiver operating characteristic) enables the evaluation of the multivariate method. The larger the area beneath the curve, the better the method performs. As can be seen, the separation for the multi-string fit with the ANN is almost perfect, while for the single-string fit it is worse but still satisfying.

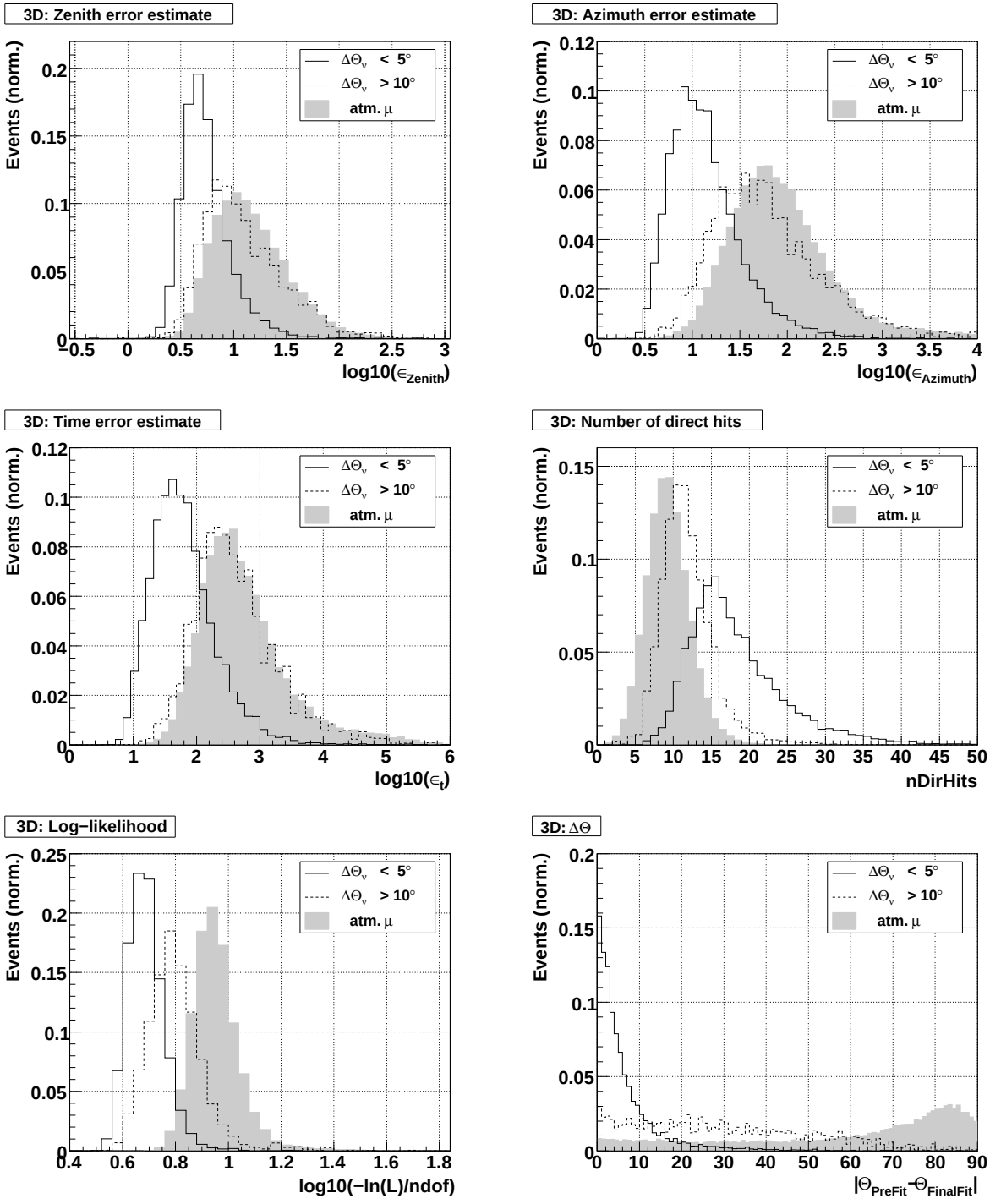



**Figure 5.18:** Normalised distributions of the ANN input parameters for multi-string events. The solid lines show the distributions for well reconstructed upward-going neutrino events, the dashed lines for badly reconstructed upward-going neutrino events, the shaded areas for misreconstructed atmospheric muons. Only events with a reconstructed zenith angle of $\Theta_{\text{Pos.}} > 115°$ are taken into account.

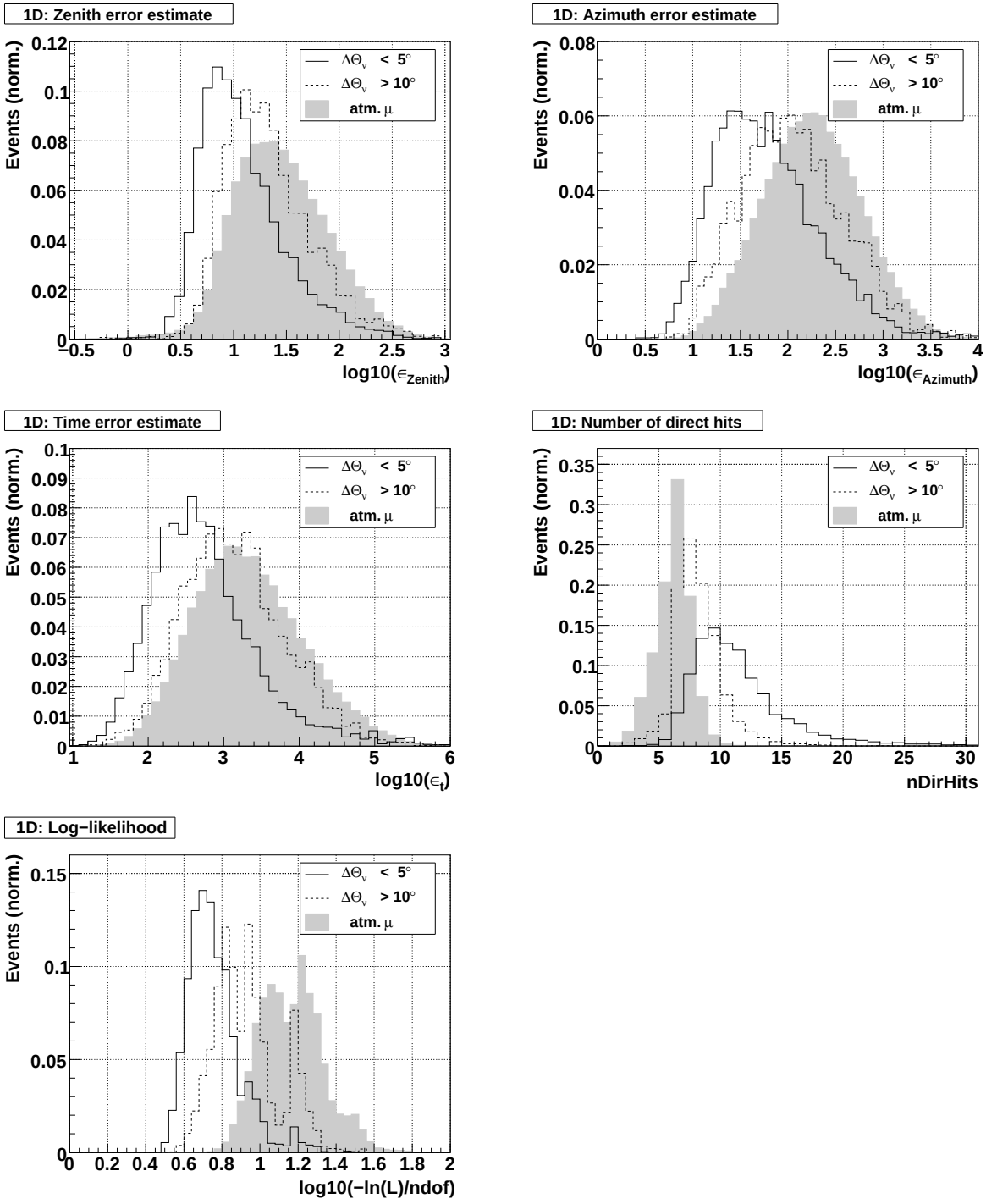



**Figure 5.19:** Normalised distributions of the ANN input parameters for single-string events. The solid lines show the distributions for well reconstructed upward-going neutrino events, the dashed lines for badly reconstructed upward-going neutrino events, the shaded areas for misreconstructed atmospheric muons. Only events with a reconstructed zenith angle of $\Theta_{\text{Pos.}} > 135°$ are taken into account.

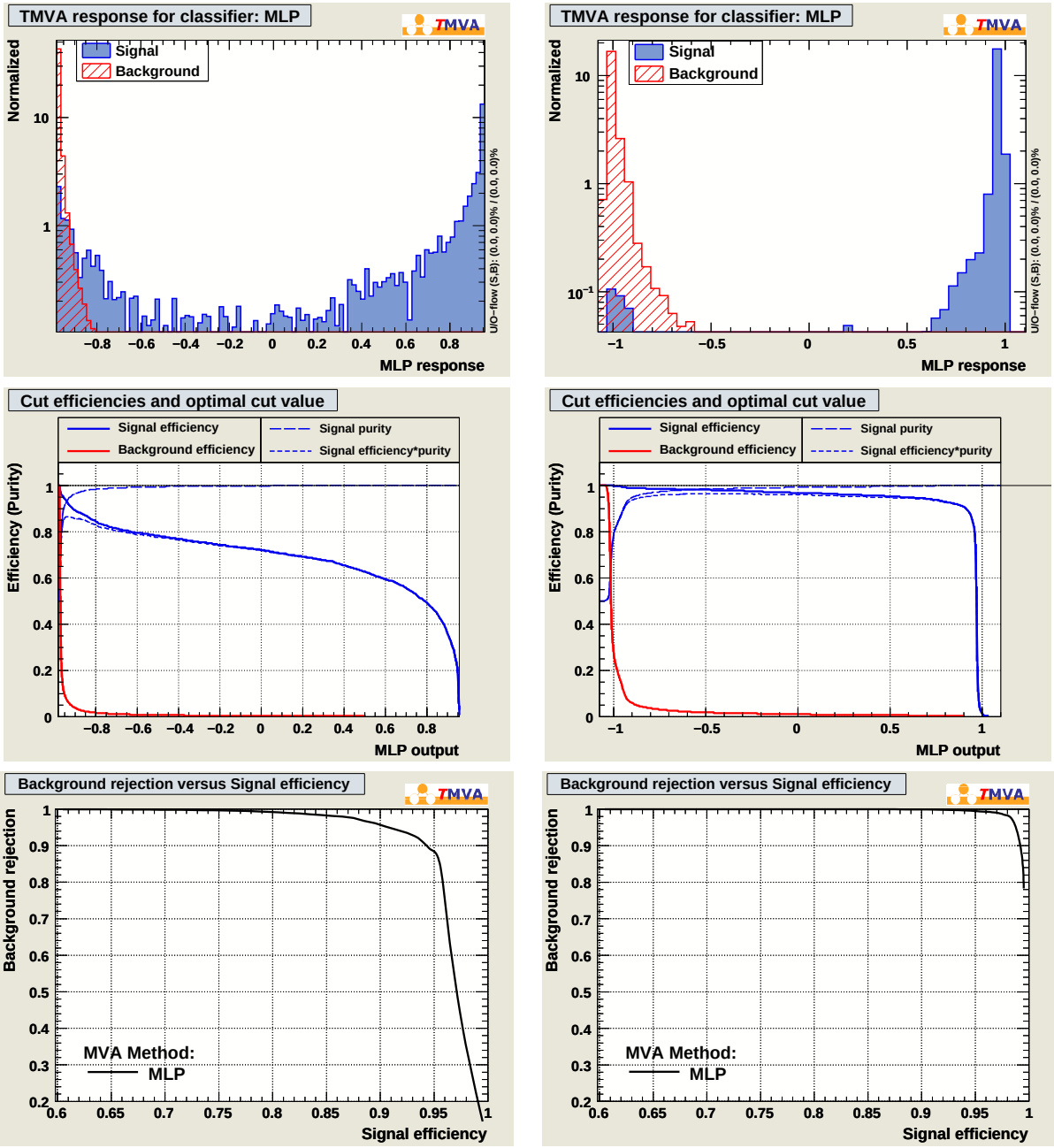



**Figure 5.20:** Performance of TMVA on the test sample. *Left:* Single-string reconstruction. *Right:* multi-string reconstruction.
*First row:* Distribution of the ANN output parameter for signal and background.
*Second row:* Purity and efficiency as a function of the cut value.
*Third row:* Signal efficiency as a function of the background efficiency.

### Application to full simulation setup

Neither the shape of the detector, nor the background rate (ranging from 60 – 80 kHz) show significant impact on the distributions of the quality parameters serving as input to the ANN. This nice agreement allows for a direct application of the network, trained with the GTS, to all different simulated setup samples. In Fig. 5.21, the output distributions of the ANN are shown for the GTS and the FSS. The distributions of the GTS and the FSS both for well and for badly reconstructed events agree well. The individual parameter distributions of the GTS used for training in comparison to the FSS can be found in Appendix A.

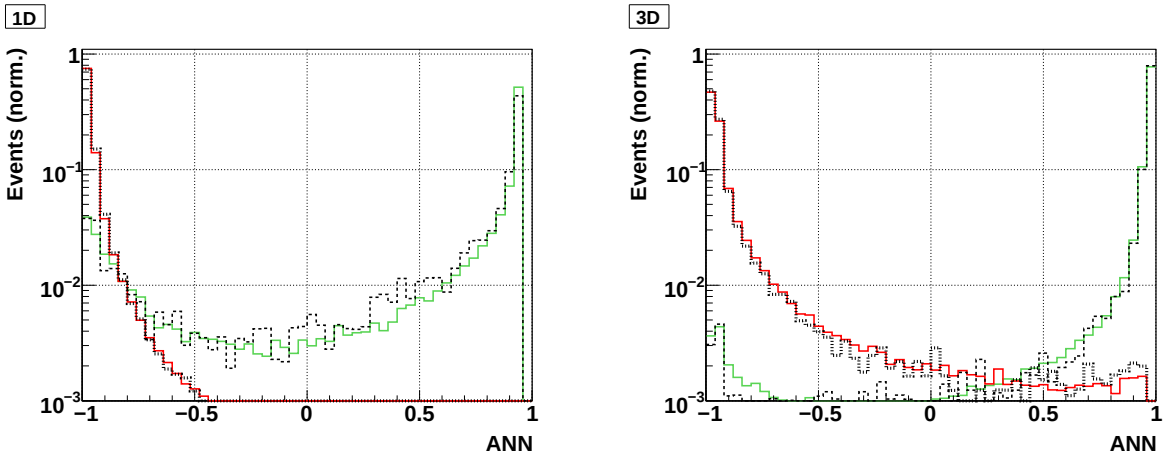



**Figure 5.21:** Neural network output distributions of well and badly reconstructed events of the global test sample (GTS) used for training (black dashed and black dotted lines, respectively), and of the full simulation sample (FSS, green and red lines, respectively).

In order to effectively suppress misreconstructed muons and to achieve an acceptable signal-to-background ratio, restrictive cuts have to be imposed on the reconstructed zenith angle and at large output values $a \in [-1; 1]$ of the ANN. The chosen cut values on the *Posidonia*-reconstructed zenith angle are $\Theta_{\text{Pos.}} = 115°$ for multi-string events and $\Theta_{\text{Pos.}} = 135°$ for single-string event. The stronger cut on single-string events is due to the required verticality for the subsequent containment selection. The reason for not accepting multi-string events up to a zenith angle of $\Theta = 90°$ is the typically worse reconstruction quality of rather horizontal events.

In Fig. 5.22, the true muon zenith angles of the FSS and the ANN output are displayed in a scatter plot. Obviously, the ANN nicely separates upward-going neutrino events from downward-going atmospheric muons, except for upward-going single-string events with a true muon zenith below about 135°, which will be discarded anyway. The cut values on the ANN output $a$ were determined by making use of these distributions and were chosen to be at $a = 0.93$ for single-string events and $a = 0.95$ for multi-string

events.

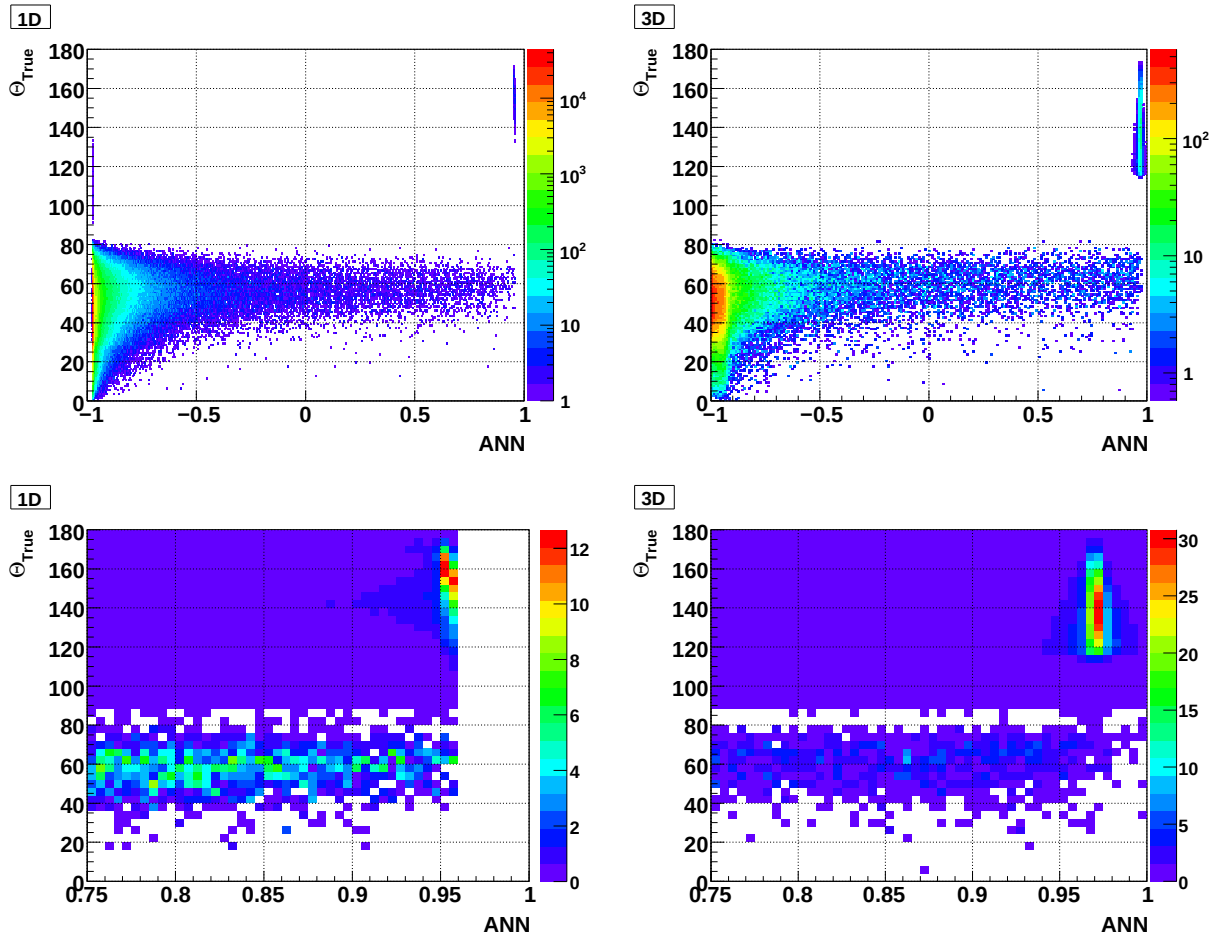



**Figure 5.22:** True (simulated) muon zenith angle $\Theta_{\text{True}}$ and the output of the neural network displayed in a scatter plot. The plots in the *upper row* cover the whole range and are shown on log scale, the plots in the *lower row* are a zoom of the region of interest, shown on normal scale. *Left:* single-string reconstruction. *Right:* multi-string reconstruction.

After cutting on the respective *Posidonia*-reconstructed zenith angle of 135° and 115° for the single- and multi-string events, as well as on the output of the neural network, the percentage of well reconstructed neutrino events out of all remaining events accounts for 65 % for the single-string events and for 87 % for multi-string events. The contamination of the remaining single-string event sample with atmospheric muons is about 21 % and the contamination of the multi-string event sample is only 6 %. The missing percentage of events (14 % and 7 %, respectively) are neutrino events that are reconstructed with a zenith error of more than 5°.

29 % of all well reconstructed single-string neutrinos remain after these two cuts, as well as 79 % of the well reconstructed multi-string events. The percentage of well

reconstructed single-string neutrino events out of all valid reconstructed single-string neutrino events before the cut accounted for 48 %, and for multi-string events, respectively, the percentage accounted for 82 %. After the cut, these values improve to 82 % for the single-string events and 93 % for the multi-string events. Both in the single- and the multi-string case, more than 99.9 % of the background events (i.e. atmospheric muons and badly reconstructed neutrinos) are rejected by applying the cuts.

In Fig. 5.23, the distribution of the zenith angle of the FSS after application of the cuts are shown. In addition to the distribution of the events remaining after the cut, also the distribution of the zenith angle of the well reconstructed neutrino events before and after the cut are indicated.

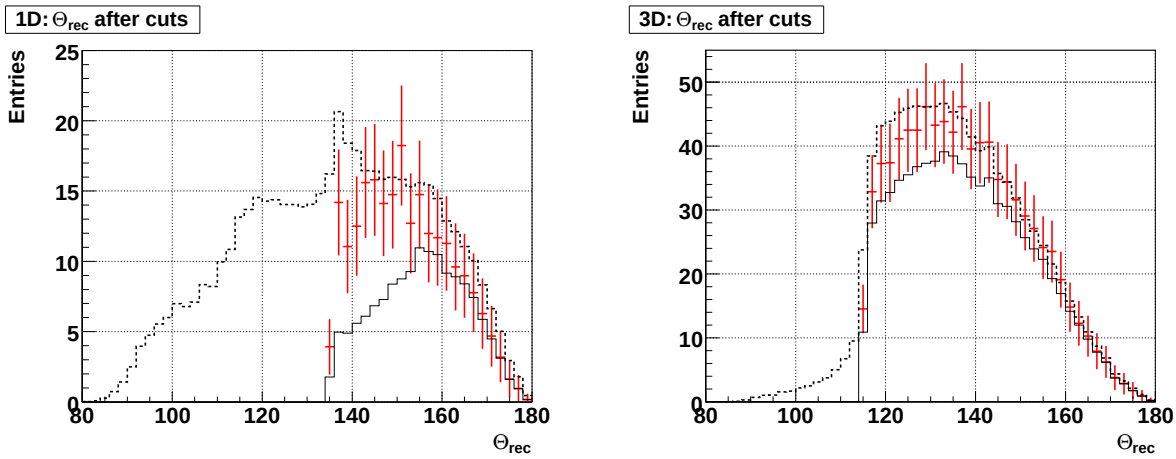



**Figure 5.23:** Distributions of the reconstructed zenith angles of the FSS after the quality cuts (red). *Left:* single-string reconstruction. *Right:* multi-string reconstruction. Also indicated are the distributions of well reconstructed neutrino events before the cuts (black dashed line) and after the cuts (black solid line).

### 5.5.2 Selection of low-energy neutrino events

Having now rejected almost all misreconstructed events, the energy of the remaining events has to be calculated. For this purpose, the procedure described in Sections 4.3.4 and 4.4.6 is applied. The containment of single- and multi-string events is estimated and their track length is calculated, from which the energy finally is inferred.

The cuts have been determined and evaluated by means of the ideal reference sample and they are now transferred one-to-one to the FSS. In detail, single-string events are discarded if the reconstructed starting or stopping point of the trajectory exceeds a vertical distance of 100 m from the detector centre in upward or downward direction. Furthermore, events are rejected if their reconstructed track length exceeds 170 m. Multi-string events are discarded, if the reconstructed starting point has a $z$-value of less

than −150 m (with respect to the detector centre) and if either the containment estimate of the starting or the stopping point is smaller than the equivalent of 1.1 photon.

In Fig. 5.24, the distributions of the reconstructed zenith angles of the FSS after the containment selection are shown. In addition to the distributions of the events remaining after cuts and containment selection, also the distributions of well reconstructed neutrino events before and after cuts and containment selection are indicated. Obviously, the remaining sample has a high purity. In detail, the remaining single-string sample consists of 61 % well reconstructed neutrino events, the multi-string sample of 86 %. The contamination with atmospheric muons accounts for 19 % and 4 % for the single-string and multi-string event sample, respectively. As the cuts had to be chosen very strict, the signal efficiency of both the cuts and the containment selection combined is only 7 % and 17 % for single- and multi-string events, respectively.

On the other hand, the strong selection leads to a good zenith and energy reconstruction quality of the remaining event sample. In Fig. 5.25, the absolute reconstruction error on the zenith and on the energy for both the single- and the multi-string reconstruction can be seen. The median reconstruction error on the zenith angle accounts for 3.0° for single-string events and for 0.6° for multi-string events. The median reconstruction error on the energy is 25 GeV for single- and 42 GeV for multi-string events, which is only a factor of 2 and 1.2, respectively, larger than the median error on the IRS (Sec. 4.4.6).

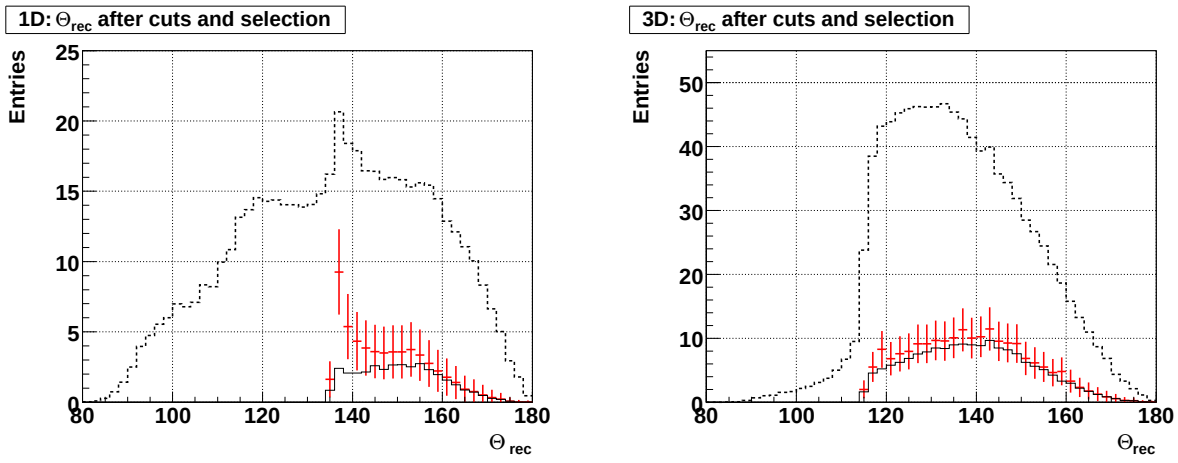



**Figure 5.24:** Distributions of the reconstructed zenith angles of the FSS after cuts and containment selection (red). *Left:* single-string reconstruction. *Right:* multi-string reconstruction. Also shown are the distributions of well reconstructed neutrino events before the application of quality cuts and containment selection (black dashed line) and afterwards (black solid line).

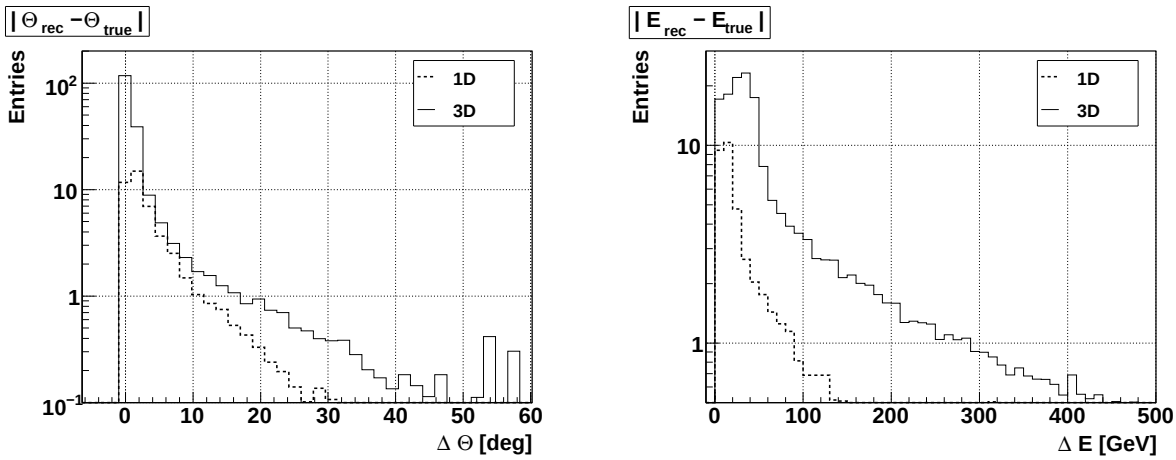



**Figure 5.25:** Distributions of the reconstruction errors on the zenith (*left*) and on the energy (*right*) of the final event sample, remaining after quality cuts and containment selection.

### 5.5.3 Application on the data

Having now determined the event selection procedure and all cut values, the selection can be applied to the data. For this purpose, the ANN output of the data is compared to the ANN output of FSS (Fig. 5.26). It can be seen that discrepancies between the distributions exist, which are larger for multi-string events than for single-string events. On a log-scale, the discrepancies increase with increasing ANN output values. An increase in the event numbers of both data and simulations around $a = 0.95$ (single-string events) and $a = 0.96$ (multi-string events) can be seen. As these ANN output values are typical values of well reconstructed neutrinos (see also Fig. 5.22), it can be assumed that the cut values applied on the data do also mainly select neutrino events.

Trying to understand the differences in the ANN output distributions can only be done by studying the input parameter distributions, which are shown in Figs. 5.27 and 5.28. Already included in these distributions are the cuts on $\Theta_{\text{rec}} = 115°$ and $\Theta_{\text{rec}} = 135°$ for multi- and single-string events, respectively. While all of the distributions of the single-string reconstruction show an acceptable agreement between data and simulation, clear discrepancies are apparent in the multi-string distributions of the number of direct hits, the negative log-likelihood divided by the number of degrees of freedom and by the difference between the zenith angles obtained from the prefit and the final fit.

The reasons that cause these differences are (yet) unknown. It has already been mentioned that investigations of the deviations between data and simulation (in low-level distributions, as well as in more advanced distributions like the parameters presented here) are still under way. Lately, it has been observed, for example, that there are different time offsets between hits detected at different lines [121]. This has an impact only on multi-string events and can explain the observed discrepancies in the corresponding

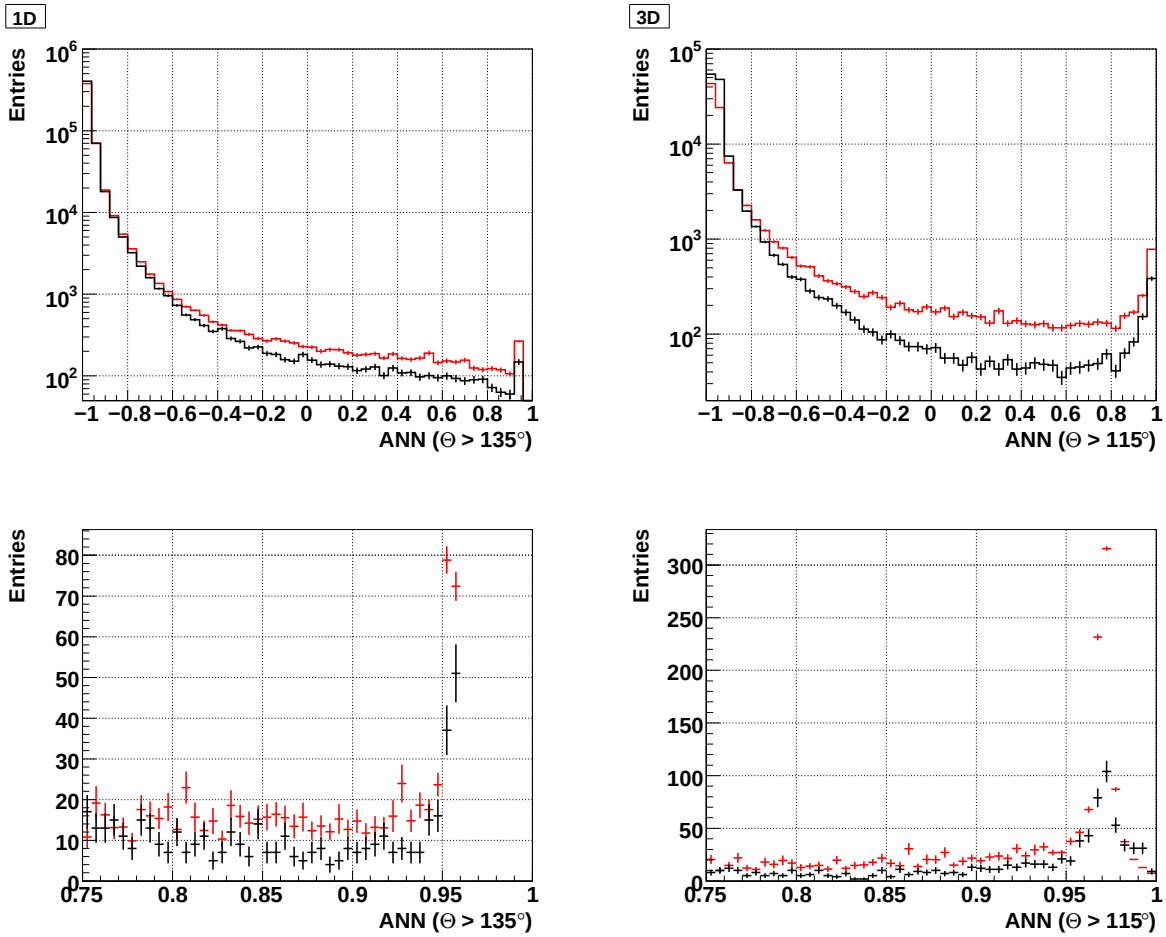



**Figure 5.26:** *First row:* distribution of the ANN output of data (black) and of the full simulation sample (red) for single-string events (*left*) and for multi-string events (*right*).
*Second row:* Zoom of the interesting region of the ANN output distributions.

distributions. As the single-string events are not affected, this plausibly confirms the observation that smaller discrepancies between data and simulations are apparent in the single-string than in the multi-string distributions. This and other effects that might emerge with further investigations, have to be studied carefully in the near future.

Other problems already mentioned, like hits with amplitudes of more than one photo-electron within the background hits, are already under investigation. Recent progress in this field could not be considered for this analysis any more. In summary it can be said, that the collaboration is currently spending a lot of effort in studying, understanding and correcting the discrepancies between data and simulation and for future analysis much better agreement can be expected. Nevertheless, this disagreement can not be investigated and solved in the present analysis. The analysis is therefore continued,

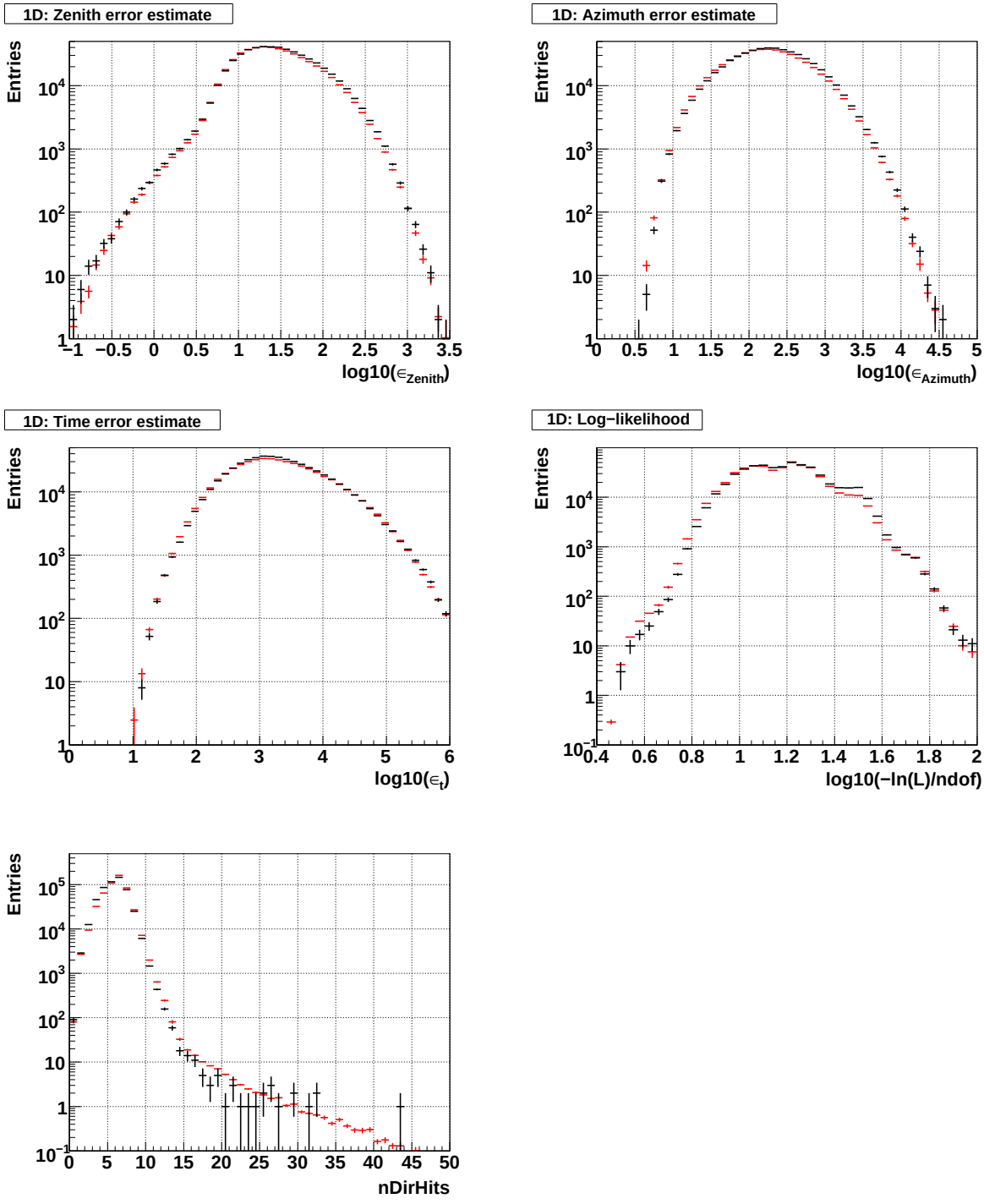



**Figure 5.27:** Single-string: distributions of the neutral network input parameters of the data (black) and the full simulation sample (red).

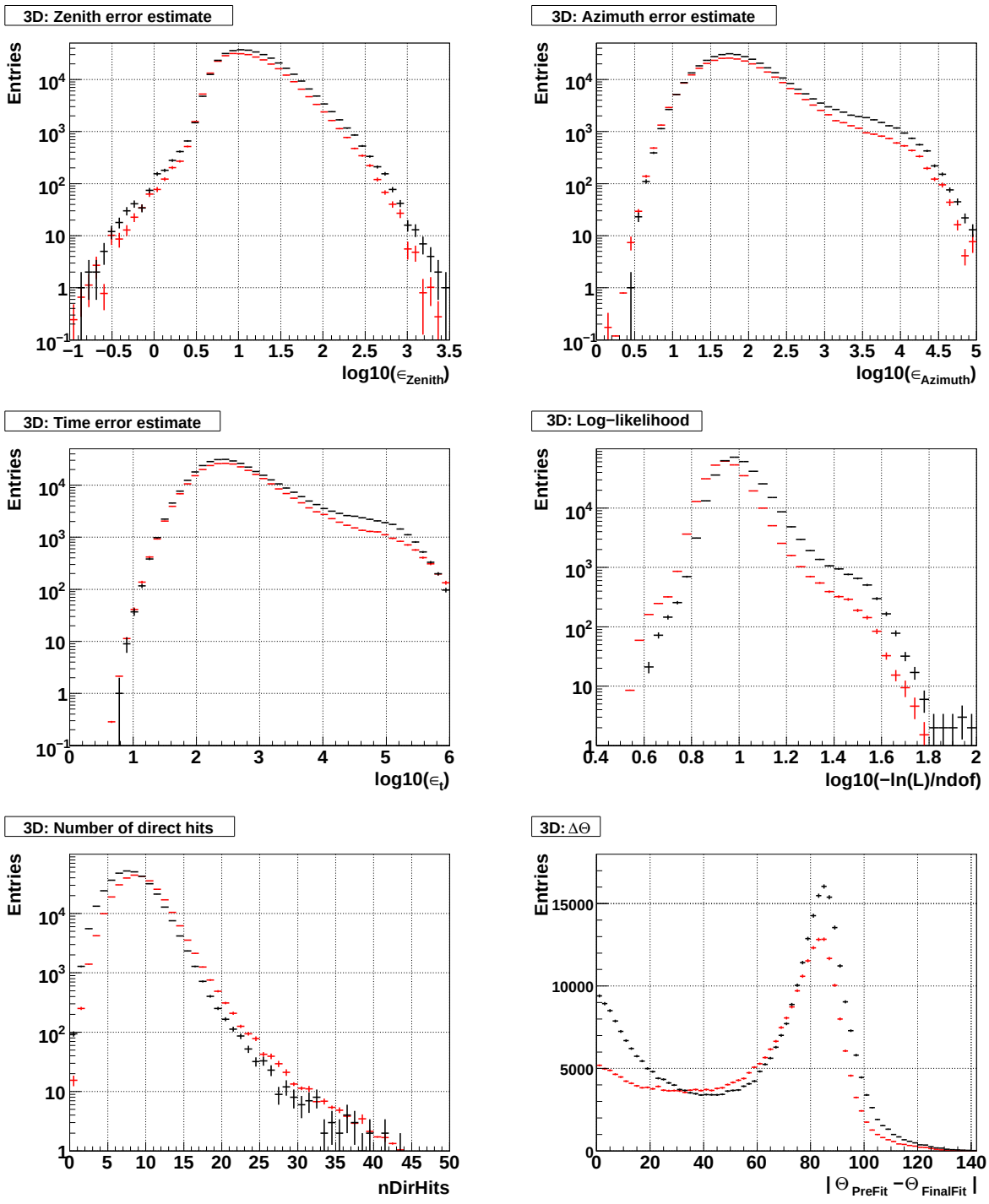



**Figure 5.28:** Multi-string: distributions of the neutral network input parameters of the data (black) and the full simulation sample (red).

ignoring the problems at this level. But in any case they have to be kept in mind when evaluating the results of the analysis.

The distributions of the reconstructed zenith angles of data and simulations after applying the quality cuts on the reconstructed zenith and the output of the ANN are shown in Fig. 5.29. As expected from the ANN output distributions, the number of events in the data is smaller than the number of events in the simulation. Though the statistics is low, it does not seem to be merely a normalisation problem. It shall be pointed out that the simulated neutrino sample is already weighted with the oscillation probability based on the oscillation parameters published by the MINOS collaboration (Chapter 6). Without considering oscillatory effects in the simulation, the discrepancies would be larger.

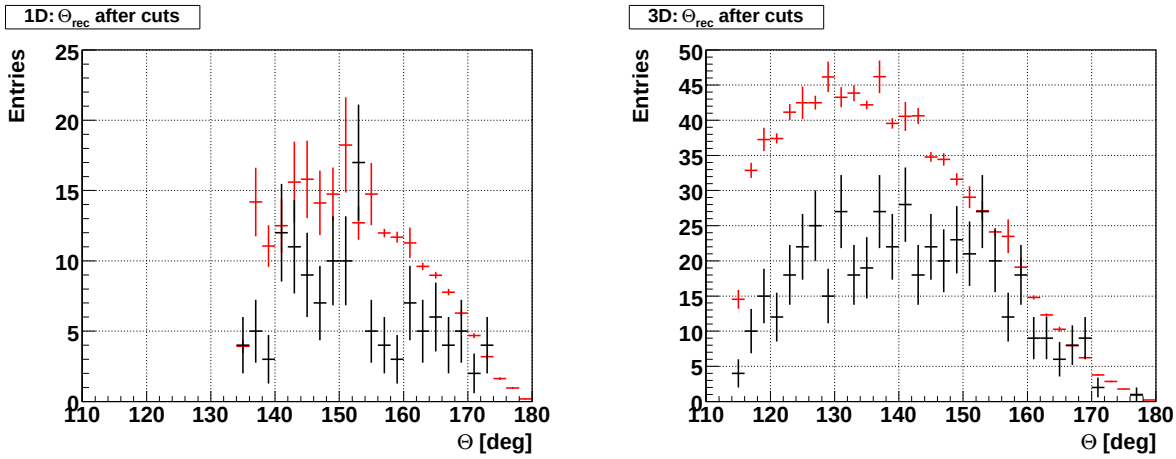



**Figure 5.29:** Distributions of the reconstructed zenith angles of data (black) and simulations (red) after application of the quality cuts on the reconstructed zenith angle and the output of the neural network. *Left:* single-string reconstruction. *Right:* multi-string reconstruction.

Applying the containment selection leads to the final event sample which will be used for the subsequent analysis. In Fig. 5.30, the distribution of the reconstructed zenith angles as well as the distributions of the reconstructed energy of data and simulations after the application of the containment selection are shown. In total numbers, the data is reduced from 133 single-string events after the cuts to 29 after the containment selection, whereas the simulated single-string events are reduced from 226 to 57 events. The multi-string data events are reduced from 487 to 72 events and the simulated multi-string events from 875 to 195.

Combining single- and multi-string events, there are 101 data events remaining in total, compared to 252 simulated events, which corresponds to a factor of about 2.5. The combined distributions of reconstructed zenith and energy can be seen in Fig. 5.31. The distributions of the simulations without oscillation hypothesis are also indicated. In Fig. 5.32 finally, the $E/\cos\Theta'$-distribution, which is the basis of the subsequent

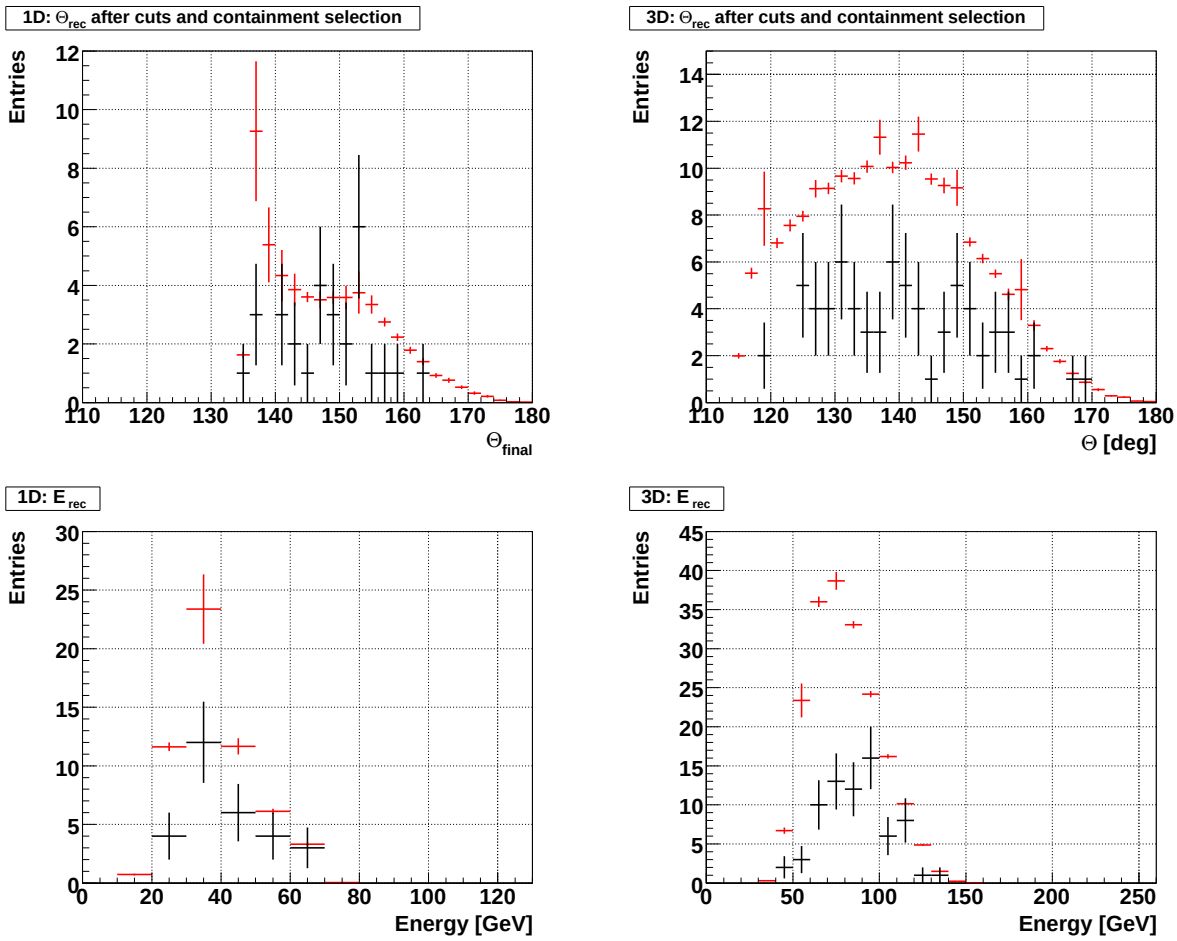



**Figure 5.30:** Distributions of the reconstructed zenith angle (*upper row*) and of the reconstructed energy (*lower row*) of the data (black) and the simulations (red) after application of the containment selection.
*Left:* single-string reconstruction. *Right:* multi-string reconstruction.

oscillation analysis, is depicted (with $\Theta' := \pi - \Theta$, and the zenith angle $\Theta$). The distribution of the simulation is shown with and without oscillation hypothesis.

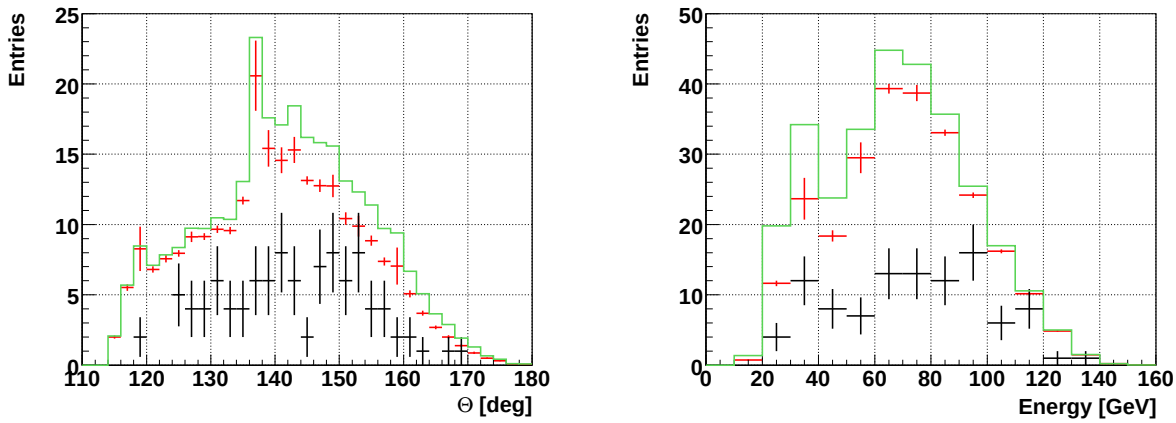



**Figure 5.31:** Combination of single- and multi-string events remaining after all cuts and selection. Distributions of the reconstructed zenith (*left*) and of the reconstructed energy (*right*). In addition to the data (black) and the simulations based on the MINOS oscillation hypothesis (red line), the simulations are also shown without oscillation hypothesis (green line).

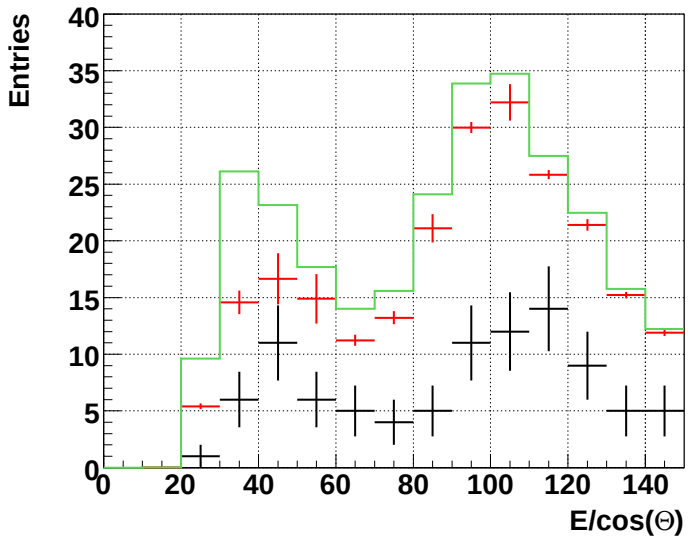



**Figure 5.32:** The resulting $E/\cos\Theta'$ distribution. In addition to the data (black) and the simulations based on the MINOS oscillation hypothesis (red line), the simulations are also shown without oscillation hypothesis (green line).

# 6 Studying oscillations of atmospheric neutrinos

The phenomenon of neutrino oscillations was first proposed in the late 1950s [3, 4], and the theory has been developed during subsequent years [5, 6]. In 1968, a deficit in the solar neutrino flux was observed [17, 18], which at that time could not be assigned unambiguously to the oscillation of neutrino flavours. Definite experimental evidence for neutrino oscillations has been reported only thirty years later in 1998 by the Super-Kamiokande collaboration [8]. Since then, various experiments could confirm this effect and measure the relevant (solar and/or atmospheric) oscillation parameters with increasing precision.

Neutrino oscillations offer a great opportunity for neutrino telescopes such as ANTARES: oscillations have not been studied yet in this energy range and the ANTARES data could therefore help to complement the picture. Such an analysis is ambitious, as it operates at the edge of the ANTARES sensitivity. The sparse instrumentation of the detector, the optical background, as well as misreconstructed atmospheric muons render a reliable zenith and energy reconstruction of low-energy tracks very challenging. Investigating the data with respect to a signature of oscillations is therefore also valuable for understanding the detector in the low-energy regime.

In this chapter, the neutrino oscillation analysis of the ANTARES data is presented. In Section 6.1 it is introduced, how neutrino oscillations are probed with ANTARES. The detailed analysis method is illustrated in Section 6.2 and its performance is evaluated in Section 6.3. Finally, the analysis of the ANTARES data set and a discussion of the results are subject of Section 6.4.

## 6.1 Probing neutrino oscillations with ANTARES

The ANTARES detector is sensitive to muons generated by upward-going atmospheric neutrinos with energies down to about 10 GeV. The survival probability of atmospheric neutrinos, given fixed oscillation parameters $\Delta m^2_{23}$ and $\sin^2(2\Theta_{23})$, depends only on the energy $E_\nu$ of the neutrino and the distance $L_\nu$ travelled by the neutrino (see Equation (2.10), Section 2.2). For upward-going atmospheric neutrinos generated in the Earth's atmosphere and detected with ANTARES, the distance $L_\nu$ can be inferred from the mean diameter of the Earth $L_0 = 12\,740\,\mathrm{km}$[22] and the zenith angle of the neutrino

[22]The deviations of the diameter due to the non-spherical shape of the Earth, which are less than $\pm 30\,\mathrm{m}$, can safely be neglected.

$\Theta_\nu$ (Fig. 6.1):

$$L_\nu = -\cos\Theta_\nu \cdot L_0 = \cos(\pi - \Theta_\nu) \cdot L_0. \tag{6.1}$$

Making use of $L_\nu$, Equation (2.10) reads as

$$P(\nu_\mu \to \nu_\mu) = 1 - \sin^2(2\Theta_{23}) \cdot \sin^2\left(16.2 \cdot 10^3 \frac{\Delta m^2_{23}[\mathrm{eV}^2]}{x}\right), \tag{6.2}$$

with $x = E_\nu[\mathrm{GeV}]/\cos(\pi - \Theta_\nu)$, where $E_\nu$ is the energy of the incident muon neutrino in units of GeV, and $\Theta_\nu$ its zenith angle (in the following: $\Theta'_\nu := \pi - \Theta_\nu$). Incorporating the results of the MINOS experiment published in 2008 [122], $\Delta m^2_{23} = (2.43 \pm 0.13) \times 10^{-3}\,\mathrm{eV}^2$ and $\sin^2(2\Theta_{23}) > 0.95$ both at 68 % C.L., with the best fit values $\Delta m^2_{23} = 2.43 \times 10^{-3}\,\mathrm{eV}^2$ and $\sin^2(2\Theta_{23}) = 1$, the minimum in the survival probability of atmospheric muon neutrinos with largest $x$ is expected to occur around $x \approx 25\,\mathrm{GeV}$, the next one around $x \approx 8.3\,\mathrm{GeV}$ (Fig. 6.2).

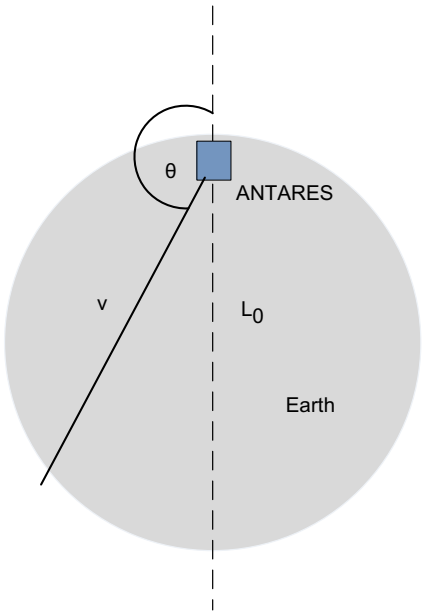



**Figure 6.1:** The distance of flight of the neutrino can be inferred from the Earth's diameter $L_0$ and the zenith angle of the neutrino $\Theta$.

In Fig. 6.3, the survival probability of atmospheric muon neutrinos is depicted (blue line). Also shown are the distributions of the muons generated by the neutrinos (red line), both with true muon energy $E_{\mu,\mathrm{true}}$ and true muon zenith $\Theta_{\mu,\mathrm{true}}$, which could theoretically be obtained given a perfect track and energy reconstruction of the detected particle, and with reconstructed muon energy $E_{\mu,\mathrm{rec}}$ and zenith $\Theta_{\mu,\mathrm{rec}}$ (black points). Statistics is reduced when employing the reconstructed values, because of the necessary quality cuts: only events reconstructed within 5° of the true zenith angle are considered, and for the energy reconstruction the cuts specified in Section 4.4.6 are applied.

It can be seen that the shape of the muon distributions follow the distribution of the parent neutrino, though the minimum is shifted and washed out due to the kinematics of the neutrino interaction (see also Section 3.1.2). The distribution of the reconstructed muon is additionally smeared due to the errors on the reconstructed zenith and energy.

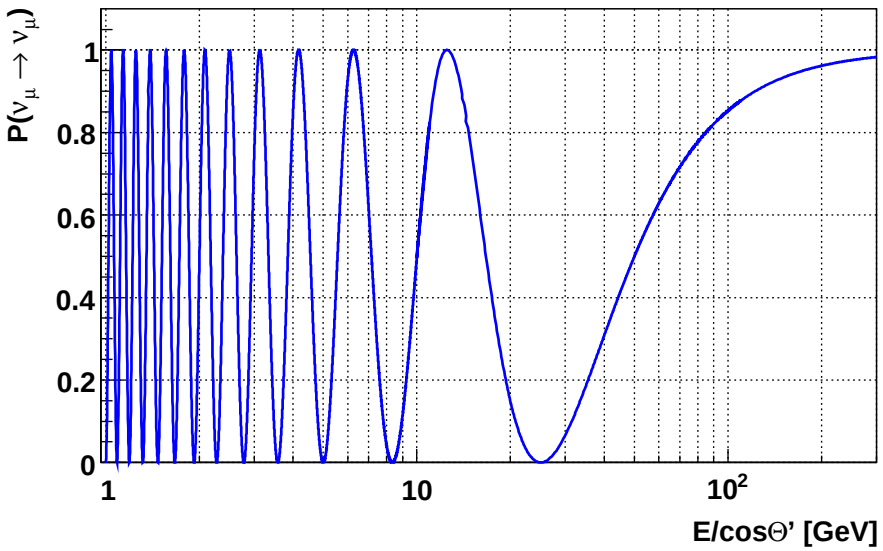



**Figure 6.2:** The survival probability of muon neutrinos as a function of $E_\nu/\cos\Theta'_\nu$, based on the MINOS oscillation scenario ( $\Delta m^2_{23} = 2.43 \times 10^{-3}\,\mathrm{eV}^2$ and $\sin^2(2\Theta_{23}) = 1$).

Yet, provided enough statistics, and given an effective zenith and energy reconstruction algorithm such as *Posidonia*, detection of the atmospheric neutrino oscillation signature in the ANTARES data seems feasible.

The distributions shown in Fig. 6.3 are based on the extended ideal reference sample described in Section 4.4.6, where only neutrinos with energies of more than 10 GeV are simulated. The impact of neutrinos with energies below 10 GeV on the distribution of reconstructed muons can be safely neglected, because of the strongly decreasing effective area (see below) for both triggered and selected events within this energy range: the effective area of triggered events drops by a factor of about 10 between 50 GeV and 20 GeV, and the effective areas of both selected and contained events drop by a factor of about 14 (see below, Fig. 6.4).

Statistics is one of the crucial points in performing an oscillation analysis with the ANTARES data. Muons with energies below 100 GeV are at the lowermost end of the sensitivity range (compare Section 4.2), which is mainly determined by the sparse instrumentation of the detector and the optical background. The neutrino effective area is given by

$$A_{\mathrm{eff}}(E_\nu, \Theta_\nu) = V_{\mathrm{eff}}(E_\nu, \Theta_\nu) \cdot \sigma(E_\nu) \cdot \rho N_{\mathrm{A}} \cdot P_{\mathrm{Earth}}(E_\nu, \Theta_\nu), \tag{6.3}$$

with the neutrino interaction cross-section $\sigma(E_\nu)$, the transmission probability of neutrinos through the Earth $P_{\mathrm{Earth}}(E_\nu, \Theta_\nu)$ and the target nucleon density $\rho N_{\mathrm{A}}$. The effective area is assumed to be independent of the azimuth angle of the neutrino $\phi_\nu$. This is a

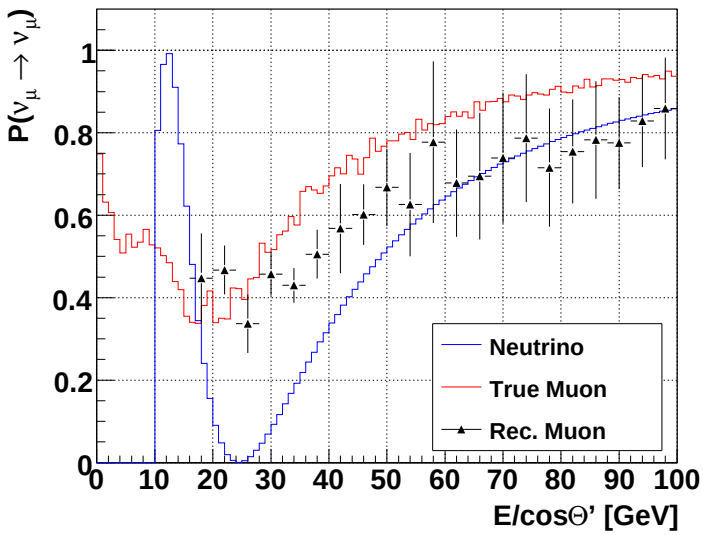



**Figure 6.3:** Survival probability of neutrinos (blue line) and their respective muons as function of $E/\cos\Theta'$ (red line). Employing the reconstructed muon zenith and muon energy (black) instead of the simulated true muon values reduces statistics due to necessary quality cuts.

good estimate for the full 12 line ANTARES detector, but is only a first approximation when lines are missing, like during the 5 line period. The effective volume

$$V_{\text{eff}} = \frac{N_{\text{sel}}(E_\nu, \Theta_\nu)}{N_{\text{gen}}(E_\nu, \Theta_\nu)} \cdot V_{\text{gen}}, \tag{6.4}$$

is obtained by multiplying the generation volume $V_{\text{gen}}$ by the ratio of selected events $N_{\text{sel}}$ (e.g. triggered or reconstructed events; events after quality cuts, etc.) to the total number of generated events $N_{\text{gen}}$. The expected event rate for a given flux model $\Phi_\nu = \mathrm{d}N_\nu/\mathrm{d}E_\nu$ can then be calculated by integrating over the neutrino energy $E_\nu$ and space angle $\Omega_\nu$:

$$\dot{N}_\nu(E_\nu) = \int A_{\text{eff}}(E_\nu, \Theta_\nu) \cdot \Phi(E_\nu, \Theta_\nu, \phi_\nu) dE_\nu d\Omega_\nu. \tag{6.5}$$

The neutrino effective area for triggered low-energy neutrinos in ANTARES (both *3N* and *2T3* triggers active) are shown in Fig. 6.4 (solid line). Additionally indicated are areas for selected neutrinos remaining after application of the quality cuts (Section 5.5.1; dashed line), as well as for the neutrinos after applying the containment selection (Section 5.5.2; shaded area). In the energy range of $10-250\,\text{GeV}$, the area drops towards smaller neutrino energies by more than two orders of magnitude, reflecting the fading sensitivity. As the oscillation signature itself is a reduced number of events with respect to the expected number without oscillation, special care is necessary in neutrino disappearance studies such as the present analysis.

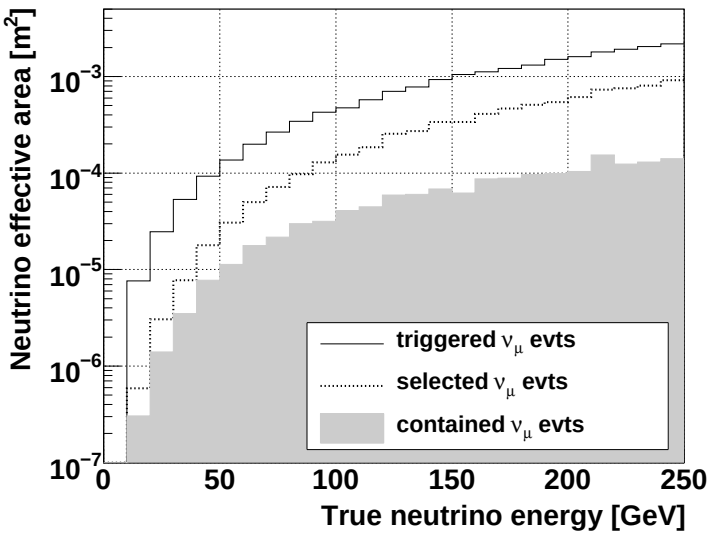



**Figure 6.4:** Neutrino effective area of triggered events (solid line), of events after application of the quality cuts (dashed line) and of events after the containment selection (shaded area).

From the neutrino effective areas, only a small low-energy event statistic can be expected from ANTARES, due to the diminishing sensitivity and the size of the detector (Fig. 6.5, solid line. Both *3N* and *2T3* trigger active). Applying the quality cuts reduces the number of events by a factor of four in the energy range from 10 – 150 GeV (dashed line), applying additionally the containment selection (shaded area) reduces them by a factor of twenty compared to the triggered number of events. Eventually about 1 500 (low-energy) neutrino events can be expected after one full year of ANTARES data taking with both *3N* and *2T3* trigger active and after the application of quality cuts and containment selection[23] (in the simulation, the full 12-line detector with all OMs active, and a constant optical background rate of 60 kHz is assumed).

## 6.2 Description of the analysis method

Assuming the mixing parameters published by MINOS, the first minimum in the atmospheric muon neutrino survival probability is expected to become apparent in the range of 10 – 60 GeV in the ANTARES $E/\cos\Theta'$ spectrum (compare Fig. 6.3). The present neutrino oscillation analysis is designed to detect the minimum in the spectrum and to determine the mixing parameters using the ANTARES data. For this purpose, the

[23] This roughly estimated number is slightly overestimated, as the interaction cross sections of neutrinos and anti-neutrinos have been assumed to be identical, which in fact they are not (compare also Section 4.1).

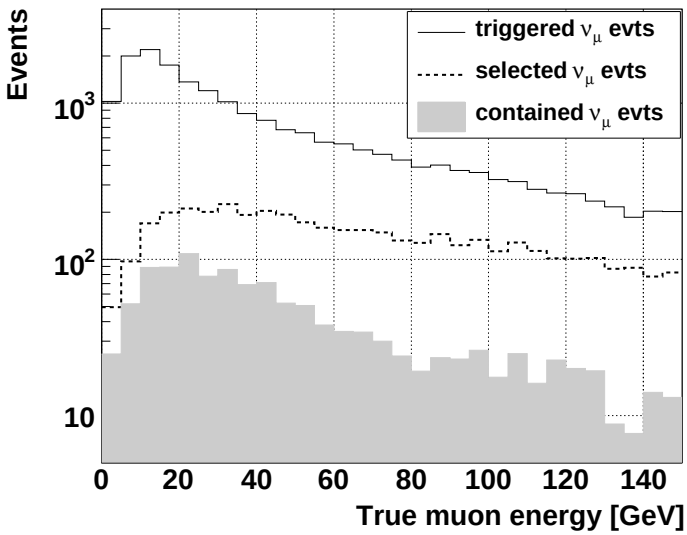



**Figure 6.5:** Expected number of atmospheric muon neutrino events as a function of the true muon energy, after trigger (solid line), after quality cuts (dashed line) and after containment selection (shaded area). The simulated event sample is the extended ideal reference sample (Section 4.4.6).

$E/\cos\Theta'$ distribution of the data is compared to different simulated oscillation scenarios, i.e. assuming different values of $\sin^2(2\Theta)$ and $\Delta m^2$. For the computation of the confidence interval, a frequentist approach for the analysis of small signals [123] is employed.

A basic, simulated $E/\cos\Theta'$ distribution is obtained from the full Monte-Carlo software chain (Section 3.6 and Section 5.3), after reconstruction and event selection. It incorporates the ANTARES detector acceptance and efficiencies, and corresponds to a spectrum in absence of neutrino oscillations. Different oscillation scenarios $T$, each corresponding to a certain oscillation parameter pair $\sin^2(2\Theta)$ and $\Delta m^2$, are generated from this spectrum by weighting the individual events with the muon neutrino survival probability determined by the respective oscillation parameters.

The performance of the analysis is not tested with the measured data, but with toy-data sampled from a full Monte-Carlo simulation incorporating a certain oscillation hypothesis (parent scenario). For this study, the mixing parameters published by MINOS are taken as granted and hence the MINOS oscillation scenario is used as parent scenario. The sampling is done by randomly generating a given number of events, distributed according to the histogram bin contents of the parent spectrum. For the final analysis eventually, the toy-data will be replaced with the measured ANTARES data.

The comparison between a (toy-)data $E/\cos\Theta'$ distribution and different oscillation scenarios is done by means of a $\chi^2$ statistics, that is constructed directly from the likelihood function [124]. Given (toy-)data in the form of a histogram, with $k$ bins labelled

by the index $i$ running from 1 to $k$ and being induced by a certain oscillation scenario $N \equiv (n_i|\sin^2(2\Theta_{\text{true}}), \Delta m^2_{\text{true}})$. Let
$n_i$ = the number of events in the $i$-th bin, and
$\boldsymbol{n} = (n_1, n_2, ..., n_k)$.
The task is to fit to the (toy-)data a theoretical distribution depending on $j$ parameters: in the present case $j = 2$ for the two oscillation parameters. This is done by comparing the (toy-)data to different model histograms, each based on a particular oscillation scenario
$T \equiv (\mu_i|\sin^2(2\Theta), \Delta m^2)$, with
$\mu_i$ = the number of events predicted by the model to be in the $i$-th bin, and
$\boldsymbol{\mu} = (\mu_1, \mu_2, ..., \mu_k)$.
The (toy-)data is described best by the theoretical histogram basing on the oscillation parameter pair $T_{\text{best}}$, that minimises the $\chi^2$ statistics.

Detecting atmospheric neutrinos with ANTARES is a Poisson process and accordingly, the entries of the (toy-)data histogram are Poisson-distributed. For such a Poisson-distributed histogram $\boldsymbol{n}$, the likelihood function is given as

$$L(\boldsymbol{n};\boldsymbol{\mu}) = \prod_{i=1}^{k} \exp(-\mu_i) \cdot \frac{\mu_i^{n_i}}{n_i!}, \tag{6.6}$$

which is the product of the probabilities of all $k$ bins, to have $n_i$ events in the $i$-th bin, given an expectation of $\mu_i$ events. According to [124], this likelihood function can be converted into the form of a general $\chi^2$ statistic by making use of the theorem on the likelihood ratio test for goodness-of-fit. This will be sketched in the following.

Let $\boldsymbol{m}$ be the true (unknown) values of $\boldsymbol{n}$ that one would get if there were no errors. The likelihood ratio $\lambda$ is then defined by

$$\lambda = \frac{L(\boldsymbol{n};\boldsymbol{\mu})}{L(\boldsymbol{n};\boldsymbol{m})}. \tag{6.7}$$

According to the likelihood ratio theorem the "likelihood"-$\chi^2$, defined by

$$\chi^2_\lambda = -2\ln\lambda = -2\ln L(\boldsymbol{n};\boldsymbol{\mu}) + 2\ln L(\boldsymbol{n};\boldsymbol{m}), \tag{6.8}$$

asymptotically for $k \to \infty$ obeys a $\chi^2$ distribution. As the second term is independent of $\boldsymbol{\mu}$, minimisation of $\chi^2_\lambda$ is entirely equivalent to maximisation of the likelihood function $L(\boldsymbol{n};\boldsymbol{\mu})$. Replacing the unknown true values $\boldsymbol{m}$ by their bin-by-bin maximum likelihood estimation $\boldsymbol{n}$ (i.e. the mean of the Poisson distribution of each bin) leads to the likelihood ratio

$$\lambda = \frac{L(\boldsymbol{n};\boldsymbol{\mu})}{L(\boldsymbol{n};\boldsymbol{n})}. \tag{6.9}$$

With equation (6.6) this leads to

$$\begin{aligned}\chi^2_\lambda = -2\ln\lambda &= -2\left(\ln L(\boldsymbol{n},\boldsymbol{\mu}) - \ln L(\boldsymbol{n},\boldsymbol{n})\right)\\ &= -2\left[\sum_{i=1}^{k}\left(-\mu_i + \ln\frac{\mu_i^{n_i}}{n_i!}\right) - \sum_{i=1}^{k}\left(-n_i + \ln\frac{n_i^{n_i}}{n_i!}\right)\right]\\ &= 2\sum_{i=1}^{k}\left(\mu_i - n_i + n_i\ln\frac{n_i}{\mu_i}\right). \end{aligned} \tag{6.10}$$

The confidence region is determined by a technique proposed by Feldman and Cousins (FC) [123]. For each point on the $\sin^2(2\Theta) - \Delta m^2$ plane (i.e. for each oscillation scenario $T$), a large number of simulated experiments $n_{\mathrm{Exp}}$ with ANTARES acceptance and resolution are generated by sampling from the respective simulated oscillation scenario $T = T_{\mathrm{s}}$. For each experiment, $\chi^2_{\mathrm{s}}$ of the currently tested oscillation scenario $T_{\mathrm{s}}$ is calculated, and $T_{\mathrm{best}} \equiv (\mu_{\mathrm{best},i}|\sin^2(2\Theta_{\mathrm{best}}), \Delta m^2_{\mathrm{best}})$ with the corresponding $\chi^2_{\mathrm{best}}$ is determined, which is the smallest $\chi^2$ for physically allowed values of $\sin^2(2\Theta)$ and $\Delta m^2$. Then, the difference in $\chi^2$ between $T_s$ and $T_{\mathrm{best}}$ is calculated.

$$\Delta\chi^2_\lambda = \Delta\chi^2 = \chi^2_{\mathrm{s}} - \chi^2_{\mathrm{best}} = 2\sum_{i=1}^{k}\left[\mu_i - \mu_{\mathrm{best},i} + n_i\ln\left(\frac{\mu_{\mathrm{best},i}}{\mu_i}\right)\right], \tag{6.11}$$

For the construction of the confidence region, a dedicated ordering principle is suggested by FC, basing on the difference in $\chi^2$: the $\Delta\chi^2$ of all simulated experiments for the currently tested point on the $\sin^2(2\Theta) - \Delta m^2$ plane are arranged in increasing order and a $\Delta\chi^2_{\mathrm{c}}$ is determined for each point on the plane such, that for $\alpha\%$ of the experiments $\Delta\chi^2 < \Delta\chi^2_{\mathrm{c}}$. The confidence region at $\alpha\%$ confidence level (C.L.) is then given by all points in the parameter plane for which

$$\Delta\chi^2_{\text{(toy-)data, s}} < \Delta\chi^2_{\mathrm{c}}, \tag{6.12}$$

with

$$\Delta\chi^2_{\text{(toy-)data, s}} = \chi^2_{\text{(toy-)data, s}} - \chi^2_{\text{(toy-)data, best}}, \tag{6.13}$$

where $\chi^2_{\text{(toy-)data, best}}$ corresponds to the oscillation scenario $T_{\text{(toy-)data, best}}$ that minimises $\chi^2$ for the (toy-)data, and $\chi^2_{\text{(toy-)data},s}$ is the $\chi^2$ between (toy-)data and the currently tested point $T_{\mathrm{s}}$ on the $\sin^2(2\Theta) - \Delta m^2$ plane.

The advantage of constructing the confidence region according to the frequentist approach of FC is that a correct coverage is given, except for a slight conservatism resulting from the discreteness of the problem (i.e. the measured number of events $n$ distributed over $i$ bins). Furthermore, the underlying ordering principle disregards statistical fluctuations with no information on the parameters. Finally, FC acceptance regions are never empty for any choice of the confidence level, because every point in the data space belongs at least to the FC acceptance region of the parameter $T_{\mathrm{best}}$ that maximises its probability [123, 125].

## 6.3 Performance of the analysis

The two oscillation parameters $\sin^2(2\Theta)$ and $\Delta m^2$ have different impact on the $E/\cos\Theta'$ distribution. Whereas $\sin^2(2\Theta)$ affects the amplitude of the oscillation, i.e. the depth of the minimum in the survival probability of muon neutrinos, $\Delta m^2$ determines the position of the minimum. With decreasing $\sin^2(2\Theta)$ values, the total number of events in the $E/\cos\Theta'$ distribution increases because of the increasing survival probability of the muon neutrinos. A maximum is reached at $\sin^2(2\Theta) = 0$, where the survival probability is exactly one, i.e. the absence of oscillations. On the other hand, the minimum in the distribution is shifted towards larger $E/\cos\Theta'$ values with increasing values of $\Delta m^2$. This implies for the ANTARES spectrum that a larger number of events is affected by muon flavour suppression (compare with the neutrino effective areas of ANTARES in Section 6.1), leading to a decrease in the integrated event number within the relevant $E/\cos\Theta'$ range from about 10 to 150 GeV.

As becomes obvious from Fig. 6.6 (left), variation of $\sin^2(2\Theta)$ at a given $\Delta m^2$ value (here: $\Delta m^2 = 2.4 \cdot 10^{-3}\,\mathrm{eV}^2$) affects mainly the region where the maximum in the $E/\cos\Theta'$ spectrum of the muon events appears (here: around 30 GeV), whereas an increase of $\Delta m^2$ at given $\sin^2(2\Theta)$ (Fig. 6.6, right; maximal mixing at $\sin^2(2\Theta) = 1$ is assumed) leads to an overall reduction of the event numbers in the spectrum between 10 and 150 GeV. This effect due to the kinematics of the neutrino interaction and the uncertainty on the reconstructed muon track (i.e. $E_{\mu,\mathrm{rec}}$ and $\Theta_{\mu,\mathrm{rec}}$) has already been mentioned in context of Fig. 6.3.

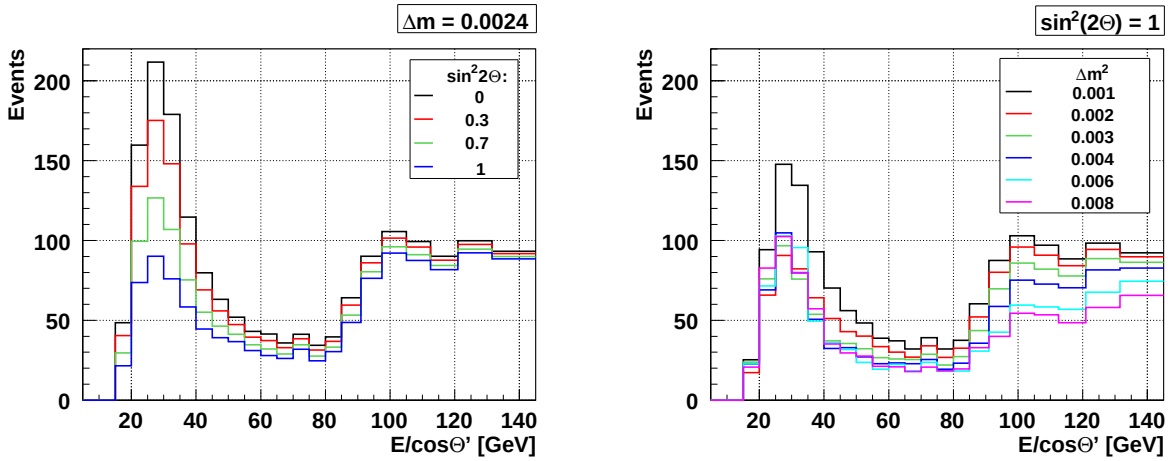



**Figure 6.6:** Impact of different $\sin^2(2\Theta)$ and $\Delta m^2$ values on the reconstructed $E/\cos\Theta'$ distribution. *Left:* Distributions for $\Delta m^2 = 2.4 \cdot 10^{-3}\,\mathrm{eV}^2$ and different values of $\sin^2(2\Theta)$. *Right:* Distributions for $\sin^2(2\Theta) = 1$ (maximal mixing) and different values of $\Delta m^2$.

The shape of the $E/\cos\Theta'$ distribution, as well as the total number of events, are affected by both neutrino oscillation parameters. However, because of uncertainties in

the absolute normalisation of the neutrino flux of about $10-20\%$, it is reasonable to use only the shape of the $E/\cos\Theta'$ distribution for the estimation of the ANTARES sensitivity to neutrino oscillations. At the expense of reducing sensitivity of the $\chi^2$ minimisation, the normalisation is fitted as free parameter by normalising the full Monte-Carlo simulations relative to the (toy-)data set. Additionally, also an absolute normalisation of the Monte-Carlo simulations is tested. In this case, the fixed number of toy-data events that are to be sampled from the particular oscillation scenario is smeared with a Poisson distribution. By comparing the results obtained from relative and absolute normalisation of the simulations, the loss in sensitivity due to a free normalisation parameter can be assessed.

### 6.3.1 Performance of the $\chi^2$ minimisation

In a first step the robustness of the $\chi^2$ minimisation procedure is investigated, which indicates the sensitiveness to statistical fluctuations in the (toy-)data. Three different types of Monte-Carlo samples are used for this purpose:

- The extended ideal reference sample (IRS, Section 4.4.6), for which only events that are reconstructed within 5° error on the zenith angle are used. The containment selection cuts are applied according to Section 4.4.6.
- The global test sample (GTS, Section 5.5.1). The impact of the event selection via neural network, as well as the impact of the contamination of the event sample with atmospheric muons will become obvious in comparison to the ideal reference sample.
- The full simulation sample (FSS) that additionally considers the data taking conditions according to the 23 setups used for the analysis (Section 5.5).

Toy-data is sampled from the MINOS scenario as described in the previous section. For the test, the sampling and the subsequent comparison to different oscillation scenarios is repeated 1 000 times for each of the three simulation test samples. Only physically allowed oscillation parameters are used to find the best scenario $T_{\text{best}}$: $\Delta m^2$ is scanned in steps of $1 \cdot 10^{-4}\,\text{eV}^2$ from $0\,\text{eV}^2$ to $0.01\,\text{eV}^2$, and $\sin^2(2\Theta)$ is sampled in steps of 0.01 from 0 to 1. Both absolute and relative normalisation are tested. Furthermore, in order to estimate the impact of different event statistics, the relative normalisation is tested with small toy-data statistics containing 150 neutrino events and large toy-data statistics, containing 1 000 neutrino events. For each toy-data sample, the best oscillation scenario is determined by $\chi^2$ minimisation, as described in detail in Section 6.2.

In Fig. 6.7, the muon neutrino survival probability as function of $E/\cos\Theta'$ for the three different Monte-Carlo samples is displayed. All three samples exhibit a clear oscillation signature, even though the distinctness of the minimum in the survival probability is reduced when passing on from IRS to GTS and FSS. This effect is due to the event selection via neural network, which is less efficient than allowing for a reconstruction error on the zenith angle of only 5°. Furthermore, the contamination with misreconstructed

atmospheric muons, as well as a decreasing reconstruction quality (i.e. increasing errors on the reconstructed values) when incorporating experimental detection conditions in the simulations contribute to this effect. Because of the reduced spectral distortion, a sensitivity loss in the determination of the confidence region has to be expected in the described order of the Monte-Carlo samples: IRS $\rightarrow$ GTS $\rightarrow$ FSS. It shall be noted, that no events with a reconstructed muon energy below 15 GeV are left in any of the three samples after all selection cuts. Furthermore, the large values of more than 0.8 in the survival probability in the 15 to 20 GeV bin of the GTS and FSS sample compared to the rather small value of about 0.45 in the IRS simulation sample are not caused by an increasing survival probability of the muon neutrino below the minimum at 25 GeV, but have to be ascribed to smaller statistics in that bin (compare with Fig. 6.3, for the distribution of the true muon values $E_{\mu,\mathrm{true}}/\cos\Theta'_{\mu,\mathrm{true}}$).

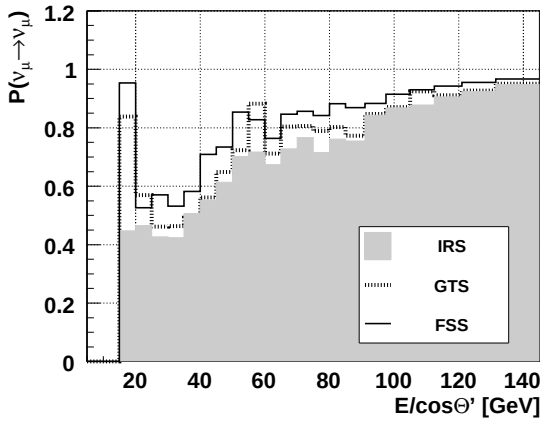



**Figure 6.7:** Muon neutrino survival probability as a function of $E/\cos\Theta'$ (with reconstructed muon energy $E$ and muon zenith angle $\Theta' = \pi - \Theta_\mu$), for three different Monte-Carlo samples. See text for details.

In Fig. 6.8 the distributions of the best oscillation scenarios $T_{\mathrm{best}}$ can be seen for the ideal reference and the global test sample. Additionally, also the mean values $\langle\Delta m^2\rangle$ and $\langle\sin^2(2\Theta)\rangle$ of the oscillation parameters, as well as the respective standard deviation $\sigma$ are given. As expected, the spread in $\Delta m^2$ increases when passing from IRS to GTS, due to the washed out minimum in the $E/\cos\Theta'$ spectrum. Conversely, the spread in $\sin^2(2\Theta)$ is more or less constant in both samples when using toy-data sets of the same size. This is due to the fact that this parameter is mainly sensitive to the event statistics in the individual bins. When using the small toy-data samples, the standard deviation of the $\sin^2(2\Theta)$ parameter approximately doubles. Furthermore, the spread increases also for the $\Delta m^2$ parameter, when reducing the event statistics of the toy-data sample.

The most stable fit results are obtained for both IRS and GTS, when using absolute normalisation (see Fig. 6.9). In this case, the mean statistics of the toy-data sets

are determined by the number of events in the respective simulated MINOS oscillation scenario (1 176 events in the IRS sample, 1 030 events in the GTS sample) and they are smeared with a Poisson distribution to obtain the number of events to sample.

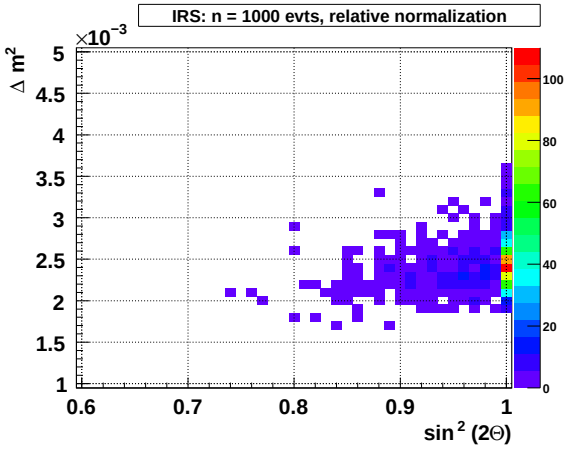
As discussed above, a reduction in the event numbers due to smaller $\sin^2(2\Theta)$ can be compensated to some extend by larger $\Delta m^2$ values. The $E/\cos\Theta'$ spectra exhibit similar shapes, and this consequently leads to wrong results in the minimisation process of the $\chi^2$. When using simulations with absolute normalisation, the dispersion of the best fit results in the $\sin^2(2\Theta) - \Delta m^2$ space indicates the relation between the total number of events in the observed energy regime and the oscillation parameters, although the effect is not very significant due to the smearing of the number of sampled events and the large event statistics of more than 1 000 events in both cases. However it can be concluded that the distribution of the best fit results using simulations with absolute normalisation is more physically motivated than the distribution of the best fit results using simulations with relative normalisation.

The results using the FSS are shown in Fig. 6.10. The stability of the fit is significantly reduced compared to IRS and GTS, due to the less distinct oscillation signature in the spectrum. Whereas with large sampling statistics (1 000 events), the distribution of the best fit results has a broad peak around the true input mixing parameters, this is no longer the case for the small sampling statistic with only 150 events: the peak is shifted towards smaller $\Delta m^2$ values. Using absolute normalisation, the distribution exhibits several peaks at different $\Delta m^2$ values around the true input value. Again, the correlation between smaller $\sin^2(2\Theta)$ and larger $\Delta m^2$ values is apparent, even though not very explicit.

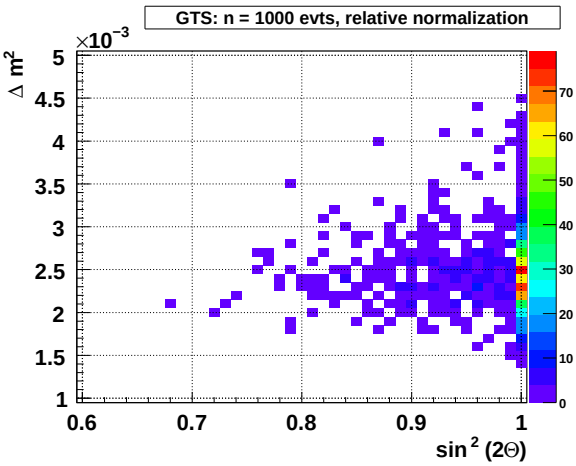
### 6.3.2 Confidence region

Confidence regions are computed to report uncertainties on results of experiments. The method to construct confidence regions proposed by Feldman and Cousins (FC) [123] has been described above. Its main advantage is that it gives correct coverage, except for a slight conservatism due to the discreteness in the Poisson case. Besides, it avoids unphysical confidence intervals. In the following, tests are performed using the FSS and constructing confidence intervals at $\alpha = 90\%$ C.L. for the three different kinds of toy-data samples: fitting the normalisation of the simulated oscillation scenarios as free parameter (a mean of 1 000 and 150 events, respectively, in the toy-data) and using absolute normalisation (a mean of 228 events in the toy-data). The number of degrees of freedom in the $\chi^2$ minimisation procedure is 20 in all cases using the FSS, corresponding to the 20 bins of the toy-data $E/\cos\Theta'$ histograms considered for the $\chi^2$ minimisation.

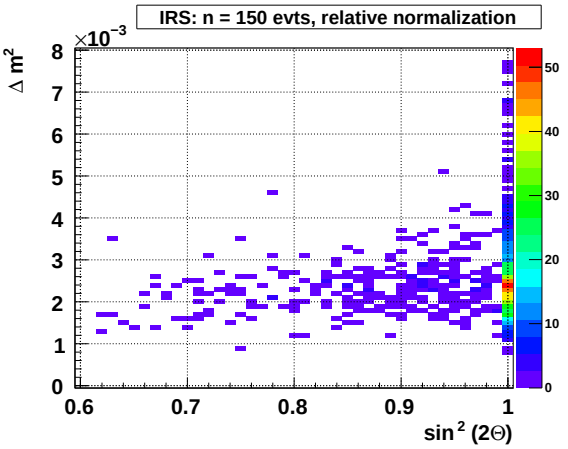
In addition to the FC confidence region, also approximate confidence regions are given, which are constructed from the difference in $\chi^2$ between the parameters $T_{\text{best}}$ with smallest $\chi^2$ within the scanned parameter space, and the $\chi^2$ of a particular point in the parameter space $T_{\text{s}}$. As done in the last section, the toy-data is obtained by sampling the respective number of events from the full Monte-Carlo simulation based on the MINOS oscillation scenario. In both the absolute and relative normalisation



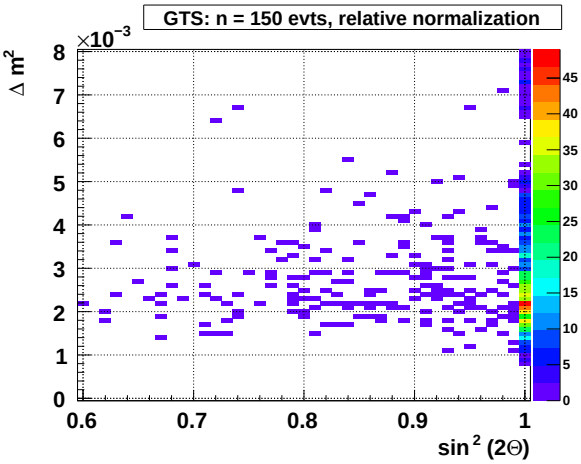

$\langle \Delta m^2 \rangle = 2.41 \cdot 10^{-3}\,\mathrm{eV}^2$, $\sigma = 0.26 \cdot 10^{-3}\,\mathrm{eV}^2$
$\langle \sin^2(2\Theta) \rangle = 0.97$, $\sigma = 0.04$




$\langle \Delta m^2 \rangle = 2.48 \cdot 10^{-3}\,\mathrm{eV}^2$, $\sigma = 0.44 \cdot 10^{-3}\,\mathrm{eV}^2$
$\langle \sin^2(2\Theta) \rangle = 0.97$, $\sigma = 0.05$




$\langle \Delta m^2 \rangle = 2.54 \cdot 10^{-3}\,\mathrm{eV}^2$, $\sigma = 0.90 \cdot 10^{-3}\,\mathrm{eV}^2$
$\langle \sin^2(2\Theta) \rangle = 0.96$, $\sigma = 0.08$

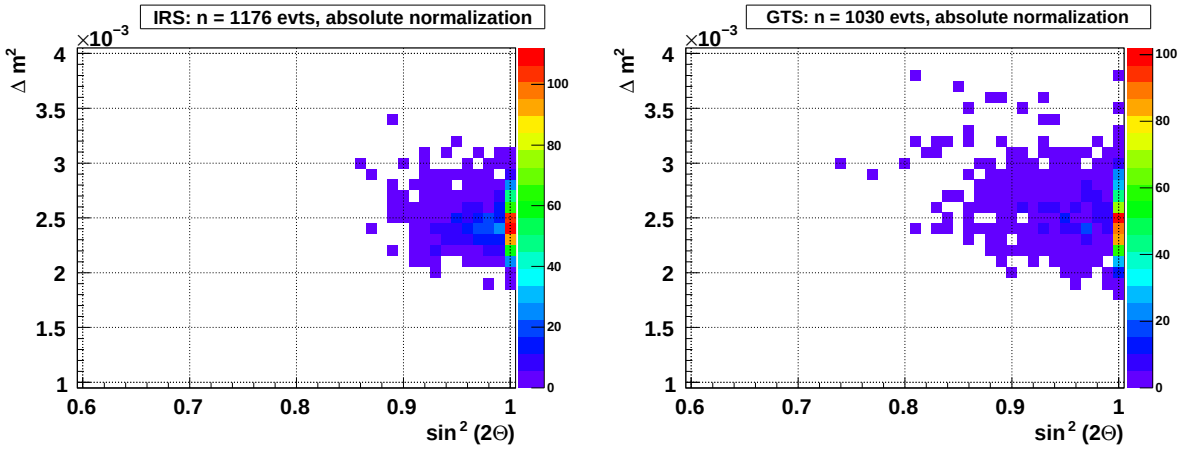



$\langle \Delta m^2 \rangle = 2.80 \cdot 10^{-3}\,\mathrm{eV}^2$, $\sigma = 1.27 \cdot 10^{-3}\,\mathrm{eV}^2$
$\langle \sin^2(2\Theta) \rangle = 0.97$, $\sigma = 0.08$

**Figure 6.8:** Distribution of $T_{\mathrm{best}}$ of 1 000 toy-data samples using two different types of simulation: *Left:* IRS. *Right:* GTS. All simulations are normalised relative to the respective toy-data. Two different toy-data statistics are tested: a large statistics (*upper row*, 1 000 neutrino events) and a small statistic (*lower row*, 150 neutrino events). See text for details.

case, the number of events is smeared with a Poisson distribution.

For the computation of the confidence intervals, only a physically allowed parameter space is scanned. Contrary to the study on the best fit results described above, $\Delta m^2$ is now computed and given on log scale instead of linear scale: $\Delta m^2$ is scanned from $10^{-5}\,\mathrm{eV}^2$ to $1\,\mathrm{eV}^2$ in steps of $10^{0.1}\,\mathrm{eV}^2$. The $\sin^2(2\Theta)$ parameter is scanned in steps of 0.02 from 0 to 1.




$\langle \Delta m^2 \rangle = 2.46 \cdot 10^{-3}\,\mathrm{eV}^2$, $\sigma = 0.20 \cdot 10^{-3}\,\mathrm{eV}^2$
$\langle \sin^2(2\Theta) \rangle = 0.98$, $\sigma = 0.03$

$\langle \Delta m^2 \rangle = 2.53 \cdot 10^{-3}\,\mathrm{eV}^2$, $\sigma = 0.30 \cdot 10^{-3}\,\mathrm{eV}^2$
$\langle \sin^2(2\Theta) \rangle = 0.97$, $\sigma = 0.04$

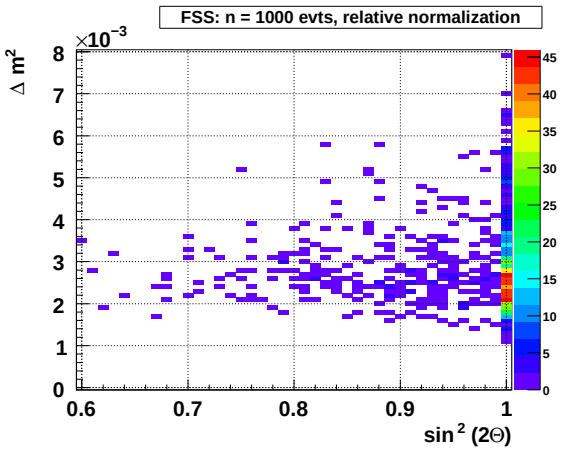
**Figure 6.9:** Distribution of $T_{\mathrm{best}}$ of 1 000 toy-data samples using two different different types of simulation: *Left:* IRS. *Right:* GTS. The simulations are normalised to an absolute value.

For the construction of the confidence region, $n_{\mathrm{Exp}} = 100$ experiments are generated for each point in the parameter space. The number of events to sample for one of these experiments is determined from the particular oscillation scenario $T_{\mathrm{s}}$ at the given point in the parameter space. When using simulations with relative normalisation, this number of events is rescaled with a factor given by the mean number of events sampled for the toy-data (here: either 1 000 or 150 events), divided by the total number of events contained in the FSS histogram based on the MINOS mixing parameters, from which the toy-data is sampled. Eventually, the number of events to be sampled for one experiment is also smeared with a Poisson distribution.

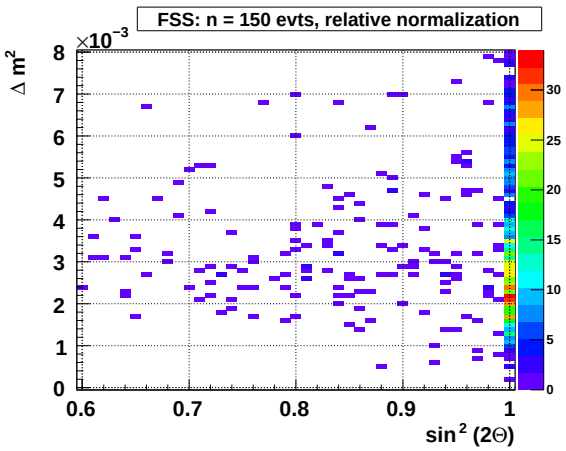
Typically, confidence regions are given as smoothed areas, since the entries in the confidence plane are not necessarily simply connected. This pathology arises from the fact that the measured parameter, i.e. the number of events $n$, is discrete. The frazzling of the confidence region increases with decreasing event statistic. The smoothed confidence interval stretches over the outermost entries in the $\Delta m^2 - \sin^2(2\Theta)$-plane. In the present study, the confidence regions are not smoothed in order to see all effects including decreasing event statistics.

Figure 6.11 shows confidence regions for exemplary toy-data samples with large (left) and small (right) statistic, calculated with relative normalisation of the simulated oscillation scenarios. The presented example of a large toy-data sample has a best $\chi^2 = 16.1$. The best fit values within the scanned physical region are $\Delta m^2 = 10^{-2.7}\,\mathrm{eV}^2$ and $\sin^2(2\Theta) = 1$. The best fit values of the presented small toy-data example inside the scanned physical region are $\Delta m^2 = 10^{-2.5}\,\mathrm{eV}^2$ and $\sin^2(2\Theta) = 1$ with $\chi^2 = 19.9$.

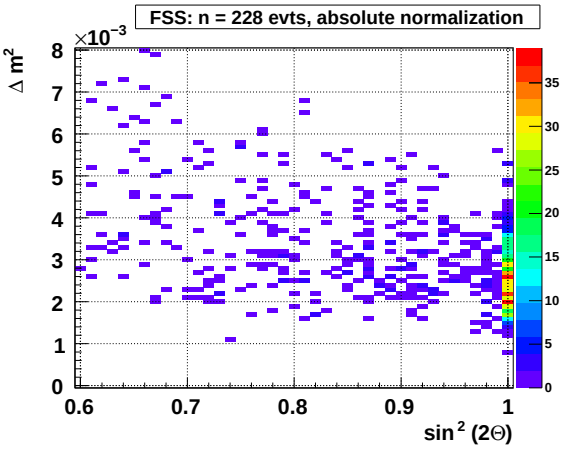
Besides the 90 % C.L. interval constructed according to the FC approach (first row),




$\langle \Delta m^2 \rangle = 2.83 \cdot 10^{-3}\,\mathrm{eV}^2$, $\sigma = 0.93 \cdot 10^{-3}\,\mathrm{eV}^2$
$\langle \sin^2(2\Theta) \rangle = 0.96$, $\sigma = 0.07$




$\langle \Delta m^2 \rangle = 3.2 \cdot 10^{-3}\,\mathrm{eV}^2$, $\sigma = 1.56 \cdot 10^{-3}\,\mathrm{eV}^2$
$\langle \sin^2(2\Theta) \rangle = 0.97$, $\sigma = 0.08$




$\langle \Delta m^2 \rangle = 2.91 \cdot 10^{-3}\,\mathrm{eV}^2$, $\sigma = 1.04 \cdot 10^{-3}\,\mathrm{eV}^2$
$\langle \sin^2(2\Theta) \rangle = 0.94$, $\sigma = 0.11$

**Figure 6.10:** Distribution of the best fit results of the full simulation sample using relative normalisation with two different event statistics for the toy-data spectrum (*left*: 1 000 events, *right:* 150 events) and using absolute normalisation (228 events with Poisson smearing, corresponding to the $\sim$250 days of lifetime, compare Chapter 5).

also the $\chi^2$ contours in the $\Delta m^2 - \sin^2(2\Theta)$-plane are depicted (second row). Additionally, the contours of the 68 % ($\Delta\chi^2 = 2.30$), 90 % ($\Delta\chi^2 = 4.61$) and 95 % C.L. ($\Delta\chi^2 = 5.99$) intervals basing on the minimum inside the physical region (2 degrees of freedom) are shown (third row)[24]. They are obtained from the difference in $\chi^2$: $\Delta\chi^2 = \chi^2_\mathrm{s} - \chi^2_\mathrm{best}$, with the actual given oscillation scenario $\chi^2_\mathrm{s}$ and the scenario with

[24]The $\Delta\chi^2$ values are obtained from tables using a standard statistics books [126].

the maximal likelihood $\chi^2_{\text{best}}$ (within the physical region). $\Delta\chi^2$ between the best fit and the no-oscillation scenario is 27.7 for the large sample and 7.5 for the small sample. The $\Delta\chi^2$ between best fit and the parameter point closest to the MINOS input mixing parameters ($\Delta m^2 = 10^{-2.6}$, $\sin^2(2\Theta) = 1$) is 0.6 for the large and 0.1 for the small sample.

As obvious from the figure, the event statistic has significant impact on the topology of the $\chi^2$ distribution and consequently on the size of the confidence region, independently of how it is constructed. Whereas the true input mixing parameters can be obtained with good precision having a large event statistic of 1 000 events in the toy-data sample, the 90 % confidence level region is larger for the small toy-data sample containing only 150 events. Extrapolating the available ANTARES data used in this analysis, 1 000 events could be measured within 6 – 7 years of effective lifetime. Basing the estimation on more optimistic assumptions i.e. for example all detector lines taking data constantly, an improved event selection or energy reconstruction procedure, etc., this statistic could probably also be obtained within 1 or 2 years from now on. In the case of the large event statistic, the approximate 90 % confidence region is a good approximation to the confidence region constructed with FC. This is no longer the case for the small event sample, where the approximate region is larger, indicating an overcoverage.

In Fig. 6.12, two further examples of small toy-data samples with 150 sampled events are shown, having different parameter sets $T_{\text{best}}$. The minimum $\chi^2$ of the sample to the left is at $\Delta m^2 = 10^{-2.1}\,\text{eV}^2$ and $\sin^2(2\Theta) = 1$ with $\chi^2 = 22.6$, and the best parameters of the sample to the right are $\Delta m^2 = 10^{-2.5}\,\text{eV}^2$ and $\sin^2(2\Theta) = 1$ with $\chi^2 = 15.3$. Compared to the first toy-data sample with small statistic shown in Fig. 6.11, the confidence intervals are enlarged and very different in shape. For the sample to the left the approximate acceptance regions are the regions left of the respective contours. For the sample to the right, the respective approximate confidence regions are to the right of the 68 % and the 90 % contour, but it is to the left of the 95 % contour. The difference in $\chi^2$ between the best fit result and the MINOS scenario are 4.7 for the toy-data sample to the left, and 0.1 for the sample to the right. The difference between best fit and the no-oscillation scenario are 2.1 and 5.1, respectively.

As obvious from Fig. 6.12 and also from Fig. 6.11 (second row), the $\chi^2$ distribution of the small toy-data sets have a flat topology. This is due to low event statistics and to the consequently less significant oscillation spectrum. The effect of this indistinct $\chi^2$ distribution became already apparent in the distribution of the best fit result, where are larger spread, i.e. a less stable $\chi^2$ minimisation, could be observed (Section 6.3.1). Consequently, as obvious from the figures, also size and shape of the confidence region become increasingly variable with decreasing event numbers. It shall be pointed out, that the fluctuations are due to different shapes of the toy-data sets. The smaller the mean of the events to be sampled, the larger the fluctuations in the event numbers in the toy-data set are, and the less significant the shape of the spectrum is. Conversely, the $\Delta\chi^2_{\text{c}}$ distribution, even though computed separately for each toy-data sample, is more or less constant.

With large toy-data samples containing 1 000 events, the shape of the sampled toy-

data distribution and consequently the best fit result (i.e. the parameter set which gives the small es $\chi^2$) is less variable, resulting in a more constant shape and size of the confidence region.

The increased sensitivity due to using absolute normalisation partially compensates for the low toy-data event statistics, and the $\chi^2$ distribution is very distinct. But nevertheless, due to the unavoidable variations in the shape of the $E/\cos\Theta'$ spectrum of different toy-data samples, fluctuations of the best fit scenario $T_{\text{best}}$ and accordingly in size and shape of the confidence region become apparent. Two typical results for two different toy-data samples using simulations normalised to an absolute value can be seen in Fig. 6.13.

The best fit parameters of the sample to the left are $\Delta m^2 = 10^{-2.5}\,\text{eV}^2$ and $\sin^2(2\Theta) = 1$ with $\chi^2 = 11.3$. $\Delta\chi^2$ between best fit and the scenario with no oscillations is 18.3. Between best fit and MINOS scenario, that again served as parent scenario for the toy-data sample, it is 0.8. The sample to the right has the best fit value at $\Delta m^2 = 10^{-2.4}\,\text{eV}^2$ and $\sin^2(2\Theta) = 0.62$ with $\chi^2 = 14.2$. $\Delta\chi^2$ between best fit and the scenario with no oscillations is 7.7, and between best fit and the MINOS scenario it is 3.1. The fact that smaller $\sin^2(2\Theta)$ values are compensated by larger $\Delta m^2$ values, as already mentioned above, becomes apparent again from the topology of the $\chi^2$ distribution of both samples and from the shape of the confidence regions, especially of the toy-data sample to the right.

Summing up it can be said that the results that can be expected from the analysis of the ANTARES data will be difficult to interpret due to the low event statistic, which is with 91 events in the considered $E/\cos\Theta'$ spectrum range even smaller than the toy-data statistics used for testing the analysis (see next section). From the results of the small toy-data samples with a mean of 150 sampled events and relative normalisation it becomes apparent that the confidence regions can assume significantly different shapes, depending on the particular shape of the toy-data sample. Using simulations with absolute normalisation, the specific shape of the $\chi^2$ distribution and thus also of the confidence regions is rather constant, but the size of the confidence region is variable.

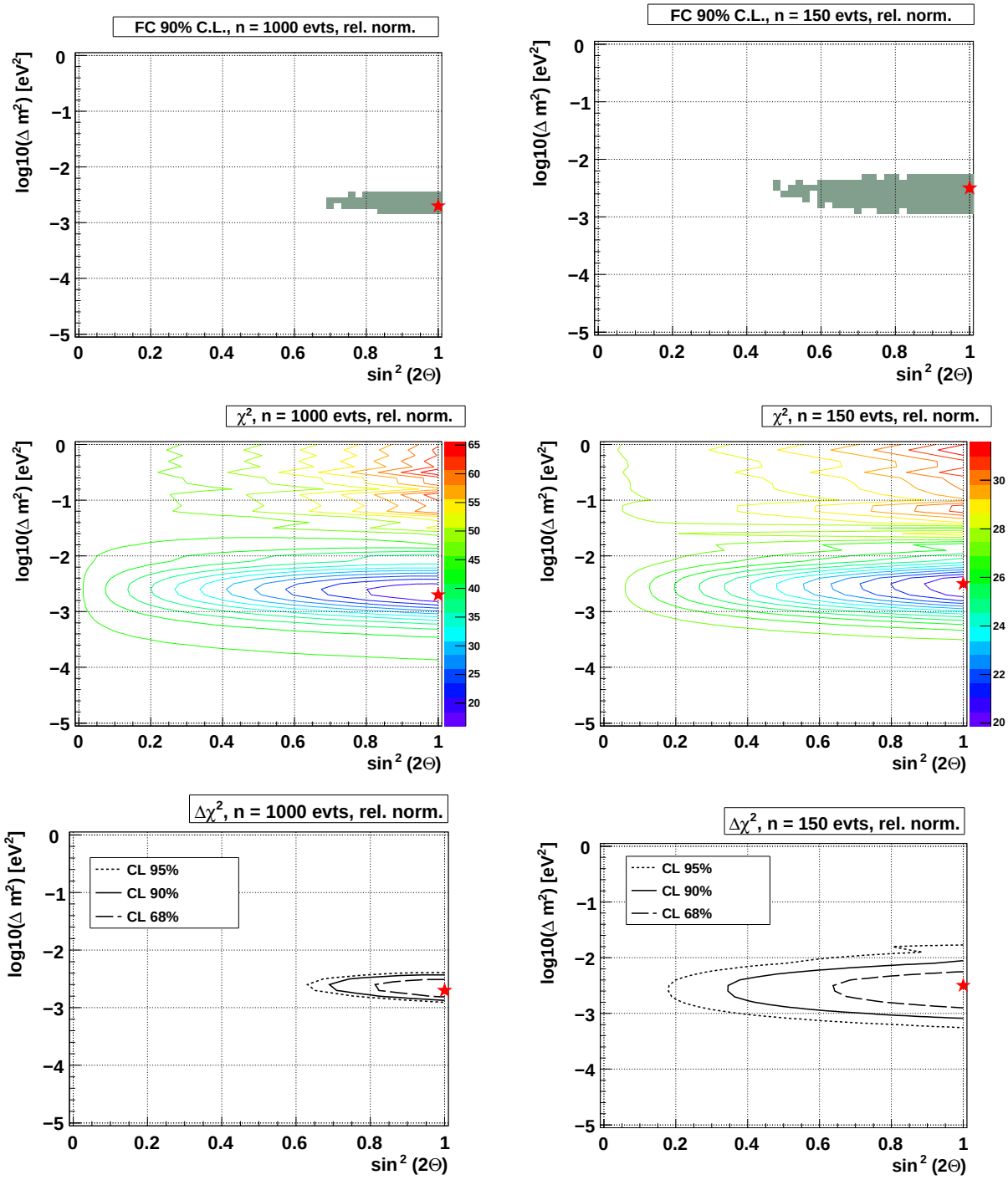



**Figure 6.11:** Confidence region at $\alpha = 90\,\%$, computed according to the Feldman-Cousins approach (*first row*), for simulations with relative normalisation and 1 000 sampled events (*left*) and 150 sampled events (*right*). *Second row:* distribution of the $\chi^2$. *Third row:* approximate 68 %, 90 % and 95 % confidence regions. The best fit to the toy-data samples $T_{\text{best}}$ is indicated by a red star.

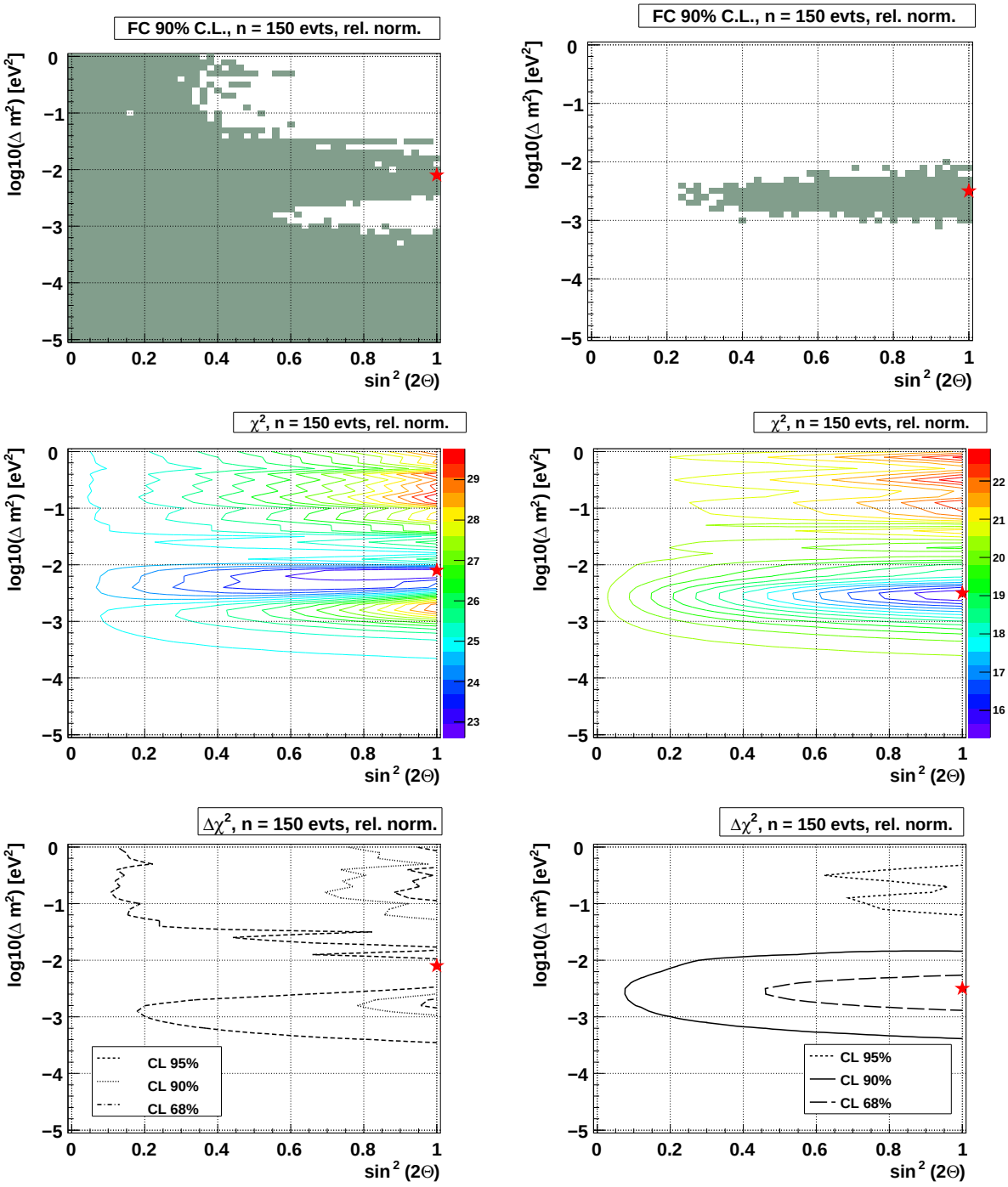



**Figure 6.12:** Confidence region at $\alpha = 90\%$, according to the Feldman-Cousins construction (*first row*), the $\chi^2$ distributions (*second row*) and the approximate confidence regions constructed from the difference in $\chi^2$ (*third row*) for two different toy-data samples with small statistic (150 events). The best fit to the toy-data samples $T_{\text{best}}$ is indicated by a red star. See text for details.

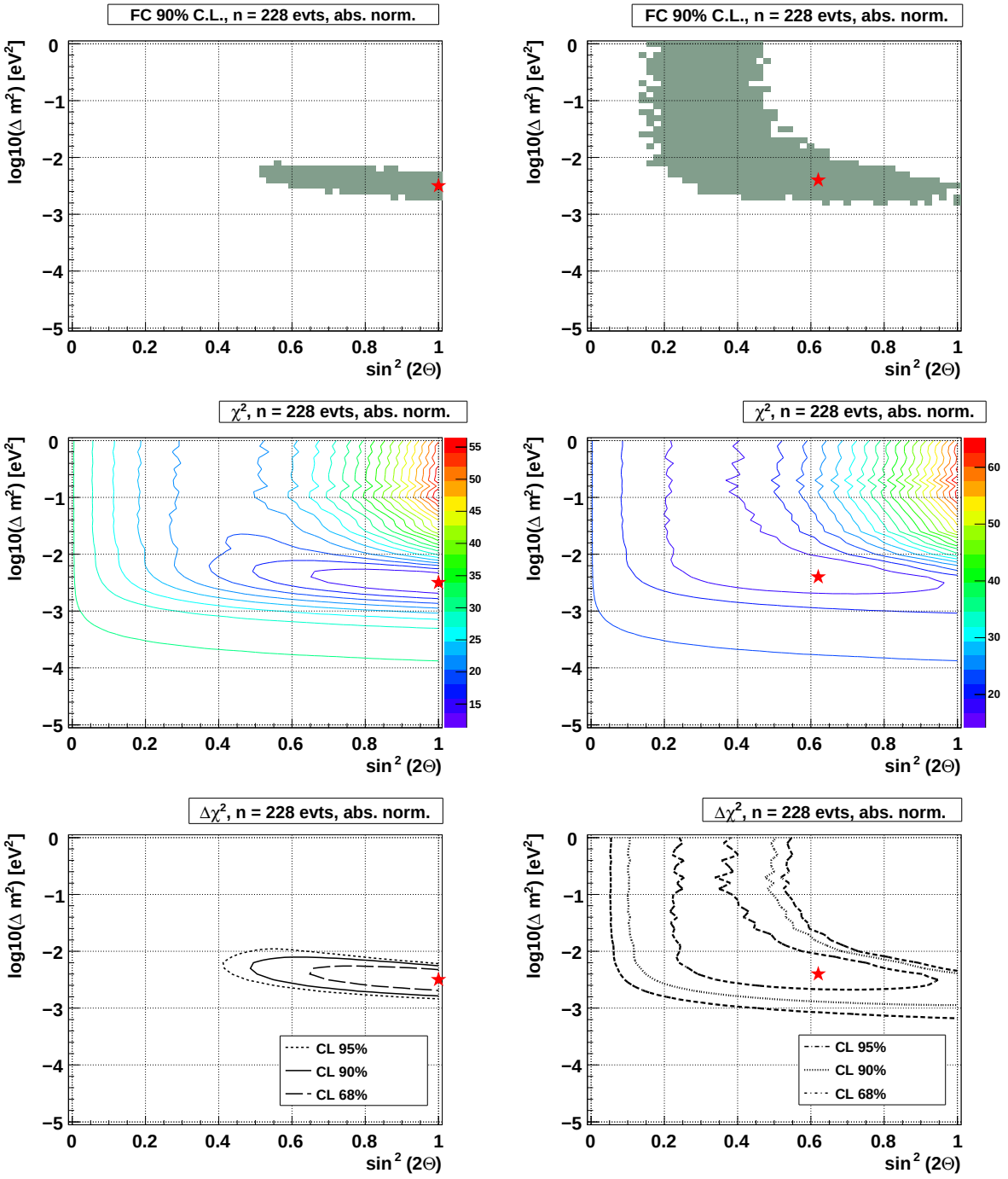



**Figure 6.13:** Confidence regions at $\alpha = 90\%$ C.L., according to the Feldman-Cousins construction (*first row*), the $\chi^2$ distributions (*second row*) and the approximate confidence regions constructed from the difference in $\chi^2$ (*third row*) for two different toy-data samples. Simulations are normalised to an absolute value, and both toy-data samples have a mean number of 228 events, corresponding to the expectation with the MINOS oscillation scenario. The best fit to the toy-data samples $T_{\text{best}}$ is indicated by a red star. See text for details.

## 6.4 Analysis of the ANTARES data

The $E/\cos\Theta'$ spectrum of the FSS and the data is depicted in Fig. 6.14, using the particular histogram binning that is also employed for the $\chi^2$ minimisation. The simulated spectrum based on the MINOS mixing parameters is displayed, as well as the simulated spectrum based on the no-oscillation scenario. They are shown with an absolute normalisation (left), and with normalisation relative to the data (right). In the energy range considered for the $\chi^2$ minimisation ($5\,\text{GeV} < E/\cos\Theta' < 145\,\text{GeV}$), the data sample contains 91 neutrino candidate events, in contrast to the simulations with absolute normalisation, that contain a mean of 228 events assuming the MINOS scenario and a mean of 272 events assuming the scenario without oscillations.

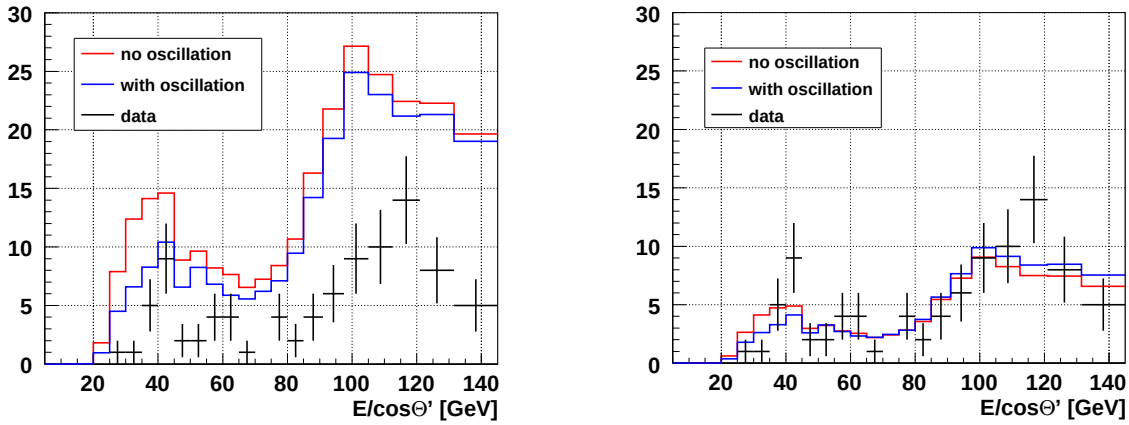



**Figure 6.14:** The $E/\cos\Theta'$ spectrum of the data (black points) and the simulations without oscillation hypothesis (red line) and assuming oscillation with the MINOS parameter set (blue line). *Left*: absolute normalisation of the simulations. *Right:* relative normalisation of the simulations.

Despite the significant discrepancies between data and simulation that already became evident in Chapter 5, an oscillation analysis is performed. However, the apparent problems in the agreement between data and simulations, as well as the small event statistic, have to be kept in mind, when interpreting the results.

In Fig. 6.15, the constructed FC confidence regions (first row), as well as the respective $\chi^2$ distributions (second row) are shown, using relative (left) and absolute normalisation (right). The best fit for the relative normalised simulations is at $\Delta m^2 = 10^{-2.8}\,\text{eV}^2$ and $\sin^2(2\Theta) = 1$ with $\chi^2 = 20.3$ for 20 degrees of freedom. The $\Delta\chi^2$ between best fit and no oscillation scenario is 2.3, and between best fit and MINOS scenario 0.9. Even though the best fit parameter set is close to the expected value, the result is not significant. This becomes clearly obvious from the large acceptance region that covers almost the complete considered parameter space.

For the simulations with absolute normalisation, the best fit parameters are $\Delta m^2 =$

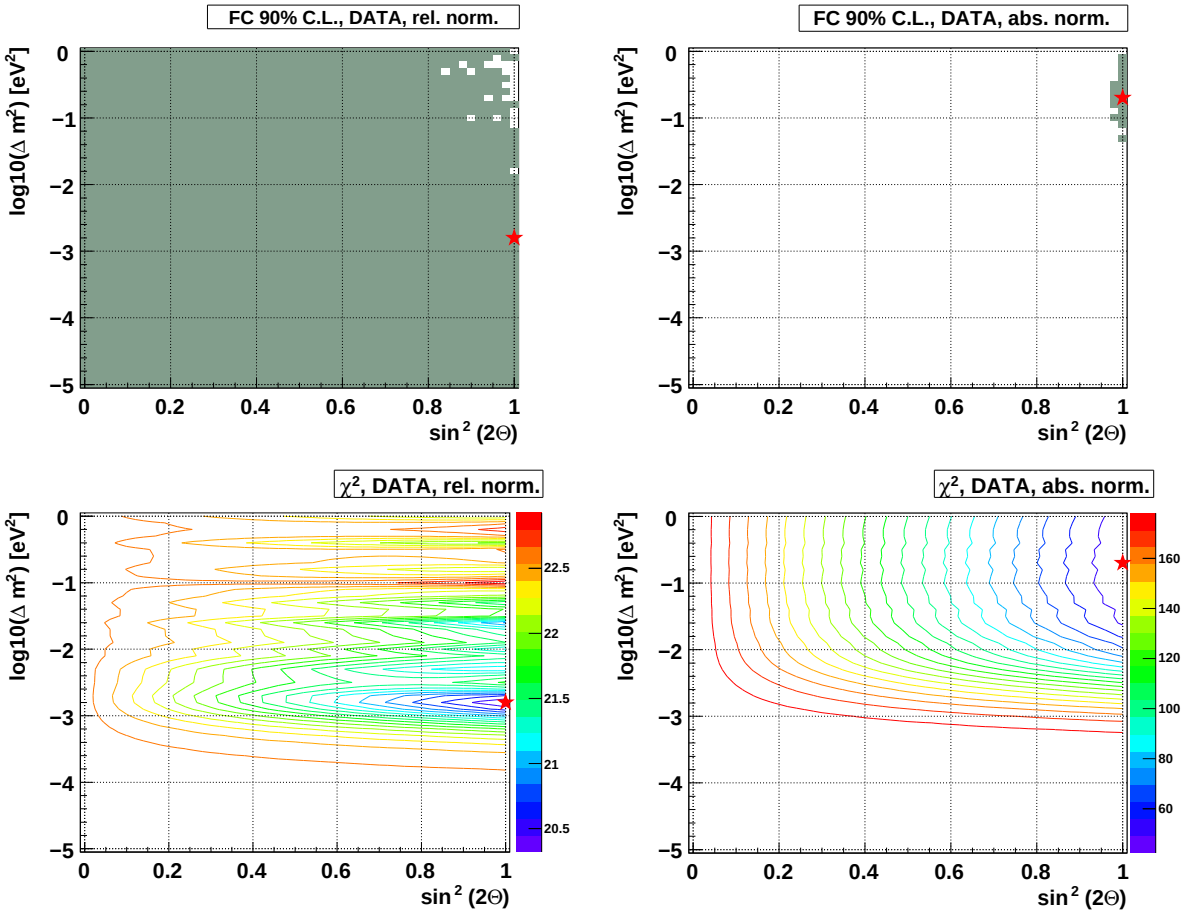



**Figure 6.15:** Confidence region at $\alpha = 90\%$ C.L., according to the Feldman-Cousins construction (*first row*) and the $\chi^2$ distributions (*second row*) for the measured data and simulations with relative normalisation (*left*) and with absolute normalisation (*right*). The best fit to the data is indicated by a red star. See text for details.

$10^{-0.7}\,\mathrm{eV}^2$ and $\sin^2(2\Theta) = 1$, close to the boundary of the scanned $\Delta m^2$ parameter space. The $\chi^2 = 42.4$, indicating a considerably bad fit. The $\Delta\chi^2$ between best fit and no oscillation scenario and best fit and MINOS scenario are 135.6 and 80.0, respectively.

From the available ANTARES data, no conclusion can yet be drawn on the true neutrino mixing parameters. Even though the scenario with the MINOS mixing parameters is more likely in both cases (using relative and using absolute normalisation), the confidence regions are too large to draw meaningful conclusions. Nonetheless, testing the analysis with Monte-Carlo simulations proved that a good estimation of the oscillation parameters is possible with the given resolution of the zenith and energy reconstruction of the muon, as soon as a large enough event statistic is available. It became apparent

from the test, that with a number of 1000 events for example, a good estimation of the oscillation parameters is feasible. This statistics can be expected within two years of full data taking, provided a complete 12-line detector and reasonable data taking conditions (optical background rates below 100 kHz). Furthermore, by improving the event selection and especially the selection of low-energy events, the number of available events can also be increased.

Moreover, the result of the analysis does not allow for inferring about the understanding of the detector in the low-energy range, mainly because of the small event statistics. It has been clearly revealed with the selection of low-energy events (Chapter 5) for the oscillation analysis, that further studies are necessary to eventually obtain a good understanding of the detector.

No systematic error estimation is yet included in the computation of the confidence region, as the large discrepancies between data and simulations have to be treated first. As long as the agreement is unsatisfactory, a systematic error estimation is of minor use. Nevertheless, the main sources of systematic errors shall be discussed in short.

The first point to mention is the uncertainty in the atmospheric neutrino flux, which is roughly about 20 % in the energy range of interest. As the shape of the atmospheric neutrino spectrum is much better understood than the absolute normalisation, the $E/\cos\Theta'$ spectrum is expected to be dominantly affected in terms of statistics. Fitting the normalisation as a free parameter neutralises the impact of variations in the overall flux normalisation, which will consequently only be reflected in variations of the necessary period of data taking time. The same holds for uncertainties in the absolute detector acceptance. An under- or overestimation of the absolute detector acceptance would cause a larger or smaller event statistic as expected, but will have no impact on the shape of the measured $E/\cos\Theta'$ spectrum.

Probably the main systematics is the detector energy acceptance. For low-energy neutrinos up to $E_\nu = 150\,\mathrm{GeV}$, as required for the oscillation analysis, the detection efficiency is steeply increasing with energy. Thus, the calibration of the detector energy acceptance is one of the main challenges for future low-energy analyses. The energy acceptance shape can be affected by the detector environment like water transparency and optical background level.

The alignment of the detector strings and the timing resolution both impact the zenith resolution of the track reconstruction. A poorer resolution both of the alignment and the timing will consequently lead to a degradation of the oscillation pattern. Systematic timing shifts between the individual strings, as detected in the ANTARES data only recently, will also deteriorate the zenith reconstruction. All three effects are problematic insofar, as they might have an energy dependent impact.

Potential contamination of the detected signal with misreconstructed downward-going atmospheric muons is particularly dangerous. Underestimation of such contamination, probably accompanied by an overestimation of the absolute detection efficiency could result in an $E/\cos\Theta'$ spectrum that mimics the effect of neutrino oscillations or blurs it completely.

All possible systematic effects have to be studied carefully in the future. Their impact

on the result of the analysis can be significant, especially at the required sensitivity level.

# 7 Conclusion and outlook

The ANTARES detector is designed to detect high-energy cosmic neutrinos that are probably produced in astrophysical objects such as active galactic nuclei. It is installed on the seabed of the Mediterranean Sea, about 24 km off the French coast near the city of Toulon. ANTARES consists of a sparse three-dimensional photosensor array, covering a volume of about $0.03\,\mathrm{km}^3$, and the sea water itself is employed as detection medium. Cherenkov light, emitted by high-energy charged particles traversing the instrumented volume will be detected by the photosensors. Such particles are for example generated in charged current weak interactions of neutrinos with nuclei. ANTARES is optimized for the detection of muon neutrinos, as the long trajectory of muons generated in charged current high-energy muon neutrino interactions with nuclei allows for a precise directional reconstruction of a few tenth of a degree. The length of the muon track depends on the energy of the muon and ranges from a few ten meters at 10 GeV muon energy up to some kilometers at 1 PeV. The coarse instrumentation of the detector, as well as the impact of optical background caused mainly by Potassium-40 decays and bioluminescence, determine the energy threshold of ANTARES, which is at about 10 to 20 GeV.

The expected cosmic neutrino signal is embedded in a background of atmospheric neutrinos, generated by charged cosmic particles impinging the Earth's atmosphere. Atmospheric neutrinos are the main neutrino signal in ANTARES. Being an irreducible background for the detection of cosmic neutrinos, atmospheric neutrinos themselves offer an opportunity for different types of analyses. One possibility is to study the quantum mechanical phenomenon of neutrino oscillations, which has been done in the context of this thesis. Neutrinos are generated and detected as pure flavour eigenstates but their time evolution is described in mass eigenstates. As mass and flavour eigenstates of neutrinos are not identical and a pure flavour eigenstate is a composite of the three mass eigenstates, this leads to interference effects and hence to a change in the flavour composition. Due to the values determined in various experiments so far, the general three flavour mixing scenario can be simplified to a two flavor mixing approximation. The survival probability for a muon neutrino then depends only on the mixing angle $\Theta_{23}$, the mass difference squared $\Delta m^2_{23}$, as well as upon $E_\nu/\cos\Theta'_\nu$, with the neutrino energy $E_\nu$ and the angle $\Theta'_\nu = \pi - \Theta_\nu$ with the neutrino zenith angle $\Theta_\nu$. Using dedicated reconstruction algorithms, such as the one developed and optimized within the scope of this thesis, the zenith angle and the energy of a muon detected with ANTARES can be computed. Eventually, the known kinematics of neutrino scattering allows to infere the direction of the original neutrino to sufficient, albeit energy dependent, precision. According to the two flavor mixing scenario approximation and given values for $\Theta_{23}$ and $\Delta m^2_{23}$, the flavor oscillations will become evident in a deficit in the

measured $E_{\mu,\mathrm{rec}}/\cos\Theta'_{\mu,\mathrm{rec}}$ distribution.

Assuming for $\Theta_{23}$ and $\Delta m^2_{23}$ values that have been published in 2008 by the MINOS experiment, the only minimum in the muon neutrino survival probability within the ANTARES sensitivity range is expected around $E_\nu/\cos\Theta'_\nu = 25\,\mathrm{GeV}$. Due to the kinematics and the errors on the reconstruction, this results in a broad deficit of events in the ANTARES data in the reconstructed $E_{\mu,\mathrm{rec}}/\cos\Theta'_{\mu,\mathrm{rec}}$ spectrum between 10 and 50 GeV. As this energy range is at the sensitivity threshold of ANTARES, an analysis aiming at the determination of the so-called atmospheric oscillation parameters $\Theta_{23}$ and $\Delta m^2_{23}$ is very challenging.

This thesis is devoted to an analysis of the ANTARES data with respect to a neutrio oscillation signature. For this purpose, a dedicated low-energy reconstruction alogrithm has been implemented into the official ANTARES software framework. The algorithm had originally been developed before the ANTARES detector design was finalized. In the context of this thesis, the algorithm has been revised and adjusted to the final detector layout and hardware settings, as well as to the latest simulations. Furthermore, it has been tuned and improved in several ways.

Besides an efficient low-energy reconstruction algorithm, also a proper data selection is important for a neutrino oscillation analysis. Building upon an existing set of simple quality criteria, a data selection concept has been devised that is based on the environmental and detector conditions at the respective time of the data taking. Furthermore, the detector simulations have been refined such that the time varying detector status is taken into account.

After identifying muon tracks in the selected data and the subsequent reconstruction of their zenith angle with the above mentioned low-energy reconstruction algorithm, effective quality cuts are necessary to suppress misreconstructed atmospheric muon events, which exceed the atmospheric neutrino signal by three to four orders of magnitudes. To this end, a neural network has been employed that effectively combines the information of several input parameters into one single cut parameter. The energy reconstruction is applied to all events remaining after the cut on the output of the neural network. It consists at first of a selection of so-called contained events (meaning that their starting and stopping point is within the instrumented detector volume), allowing for a reliable track length reconstruction and at the same time, reducing the mean energy of the remaining event sample. The energy of the muons can then be inferred from the calculated track length.

The resulting $E_{\mu,\mathrm{rec}}/\cos\Theta_{\mu,\mathrm{rec}}$ spectrum is investigated with respect to neutrino oscillations. It is compared to simulated spectra basing on different oscillation scenarios (i.e. different pairs of oscillation parameters $\Theta_{23}$ and $\Delta m^2_{23}$; the mixing angle is given here in the form $\sin^2(2\Theta_{23})$). Minimizing a $\chi^2$ function for data and simulations determines the most likely oscillation scenario. For the construction of a confidence region, a frequentist approach to the analysis of small signals is employed. In order to be able to evaluate the result obtained from the analysis of the data, the analysis and the construction of the confidence region was tested beforehand with toy-data samples, that were generated assuming the mixing parameters published by MINOS.

The toy-data study shows that with an assumed event statistics of 150 neutrino events

in the spectrum, the $\chi^2$ distribution does not allow for a clear discrimination between different scenarios. Consequently, for different samples of toy-data basing on the same MINOS oscillation scenario, confidence regions very different in size and shape are obtained. Using a toy-data sample with a statistics of 1 000 neutrino events instead, the $\chi^2$ distribution is much more distinct, leading to confidence regions that are smaller and more robust in size and shape than obtained with the small toy-data sample.

Applying the analysis to the data, the best fit parameters are $\Delta m^2_{23} = 10^{-2.8}$ and $\sin^2(2\Theta_{23}) = 1$, which is close to the values measured by MINOS, which are $\Delta m^2_{23} = 10^{-2.6}$ and $\sin^2(2\Theta_{23}) = 1$. The number of neutrino candidate events in the relevant energy range considered for the analysis is 91, which is even less than the statistic of the small toy-data sample. The result can therefore not be expected to be significant. Consequently, the confidence region covers the complete tested parameter plane. Using an absolute instead of a relative normalization of the tested $E_{\mu,\mathrm{rec}}/\cos\Theta'_{\mu,\mathrm{rec}}$ distribution between data and simulation, the minimum yields a best fit result at $\Delta m^2_{23} = 10^{-0.7}$ and $\sin^2(2\Theta_{23}) = 1$, completely off the expected $\Delta m^2_{23}$ value.

This is due to the fact that the agreement between data and simulations is not very satisfying. This problem becomes not only evident in the final $E_{\mu,\mathrm{rec}}/\cos\Theta'_{\mu,\mathrm{rec}}$ spectrum, but also at a more basic level. This disagreement is not a specific problem of the presented work but at present still a general deficit of the ANTARES detector simulations. It is currently investigated and progress has been made recently, but could not be considered for this thesis anymore.

However, the study with the toy-data proved that having a good event statistics, a proper determination of the oscillation parameters with an acceptably small confidence region can be expected. A statistics of 1 000 neutrino events for example can be obtained within six to seven years extrapolating from the currently available data. Basing the estimation on a bit more optimistic assumptions, like a fully operational detector, low optical background rates and for example a further increase in the reconstruction and event selection efficiency, such a neutrino statistic might be obtained in one to two years from now on. The analysis itself can then easily be repeated and, given the results for atmospheric oscillation parameters from other experiments, might give a valuable clue to the understanding of the ANTARES detector in the low-energy range.

# A Distribution of quality parameters

Neither the shape of the detector, nor the background rate (ranging from 60 – 80 kHz) are showing significant impact on the distributions of the quality parameters serving as input to the neutral network. In this Appendix, the distributions of the training sample (global test sample, GTS, basing on ideal detector: 12 strings, 60 kHz optical background, all modules active) are shown in comparison to the full simulation sample (FSS), which is the combination of all 23 setup simulations with their different numbers of active lines, background rates and bad channel ratios. It can be seen, that all distributions agree acceptably. Consequently, the network does not have to be trained separately for each of the different setups. Instead, direct application of the neural network trained with the GTS to all different setup simulations is possible. In Figs. A.1 and A.2 the distributions of the single- and multi-string parameters are shown, respectively.

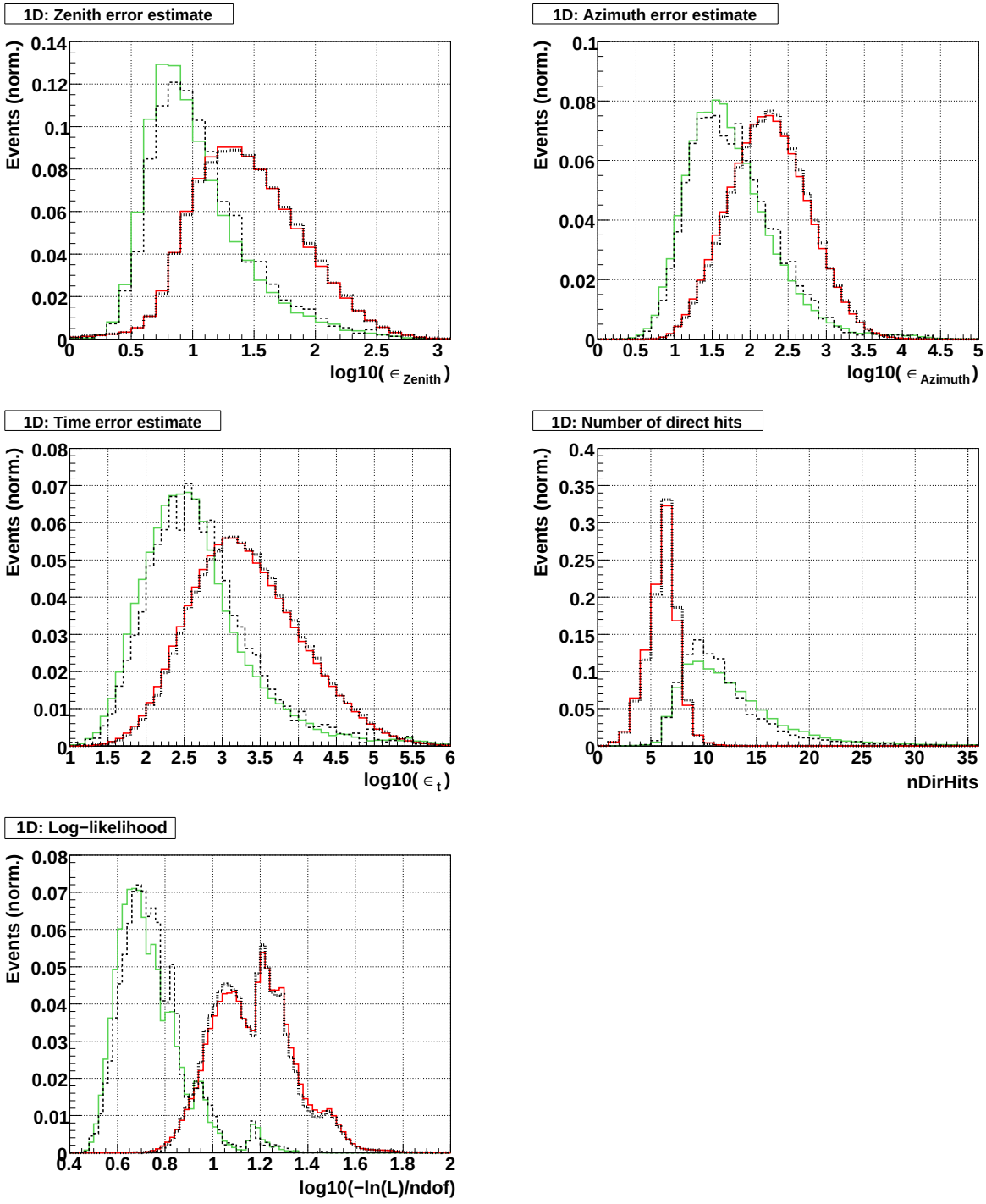



**Figure A.1:** Neural network input distributions of well ($\Delta\Theta < 5°$) and badly reconstructed single-string events ($\Delta\Theta > 10°$) of the global test sample (black dashed and black dotted lines, respectively), used for training the network, and the full simulation sample (green and red lines, respectively).

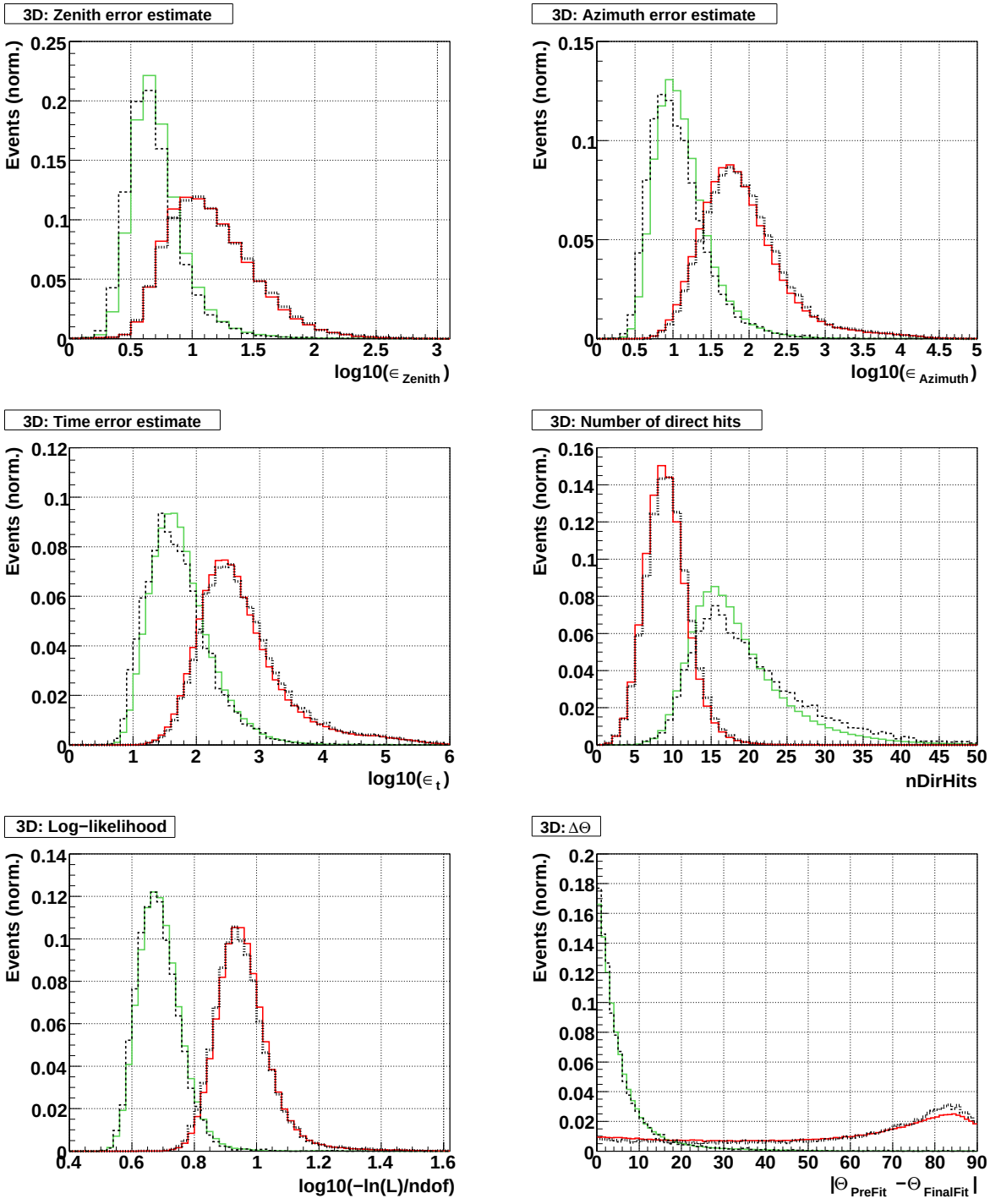



**Figure A.2:** Neural network input distributions of well ($\Delta\Theta < 5°$) and badly reconstructed multi-string events ($\Delta\Theta > 10°$) of the global test sample (black dashed and black dotted lines, respectively), used for training the network, and the full simulation sample (green and red lines, respectively).

# B Python steering scripts for processing the data

The Python steering script to perform the data processing with *SeaTray*:

```
#xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
# Script start
#xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

#!/usr/bin/env python

from I3Tray import *
from os.path import expandvars

import os
import sys
import tasks.antares

tray = I3Tray()

load("libdataclasses")
load("libphys-services")
load("libdataio")
load("libantares-reader")
load("libhit-selector")
load("libantares-tools")
load("libinterfaces")
load("libposidonia")
load("libicepick")
load("libantares-qa")


#========================================================================
# Set parameters
#========================================================================
storeydist = 14.5
coincidenceWindow = 20.*I3Units.ns
evtfile = <ANTARES data file>
outfile = <path> + <filenmame> + ".root"
outi3file = <path> + <filenmame> + ".i3.gz"
outtxtfile = <path> + <filenmame> +  ".txt"

#========================================================================
# Access to database
#========================================================================
# load basic functionality for Antares data
tasks.antares.exp.addReadAntExpData(tray, evtfile=evtfile,
           fallbackToTriggerGeometry=True)

#========================================================================
# Filter Minimum Bias
#========================================================================
# filter minimum bias events from Antares raw data
tasks.antares.exp.addFilterMinimumBiasEvents(tray)

#========================================================================
# Calibration
#========================================================================
# ANATARES raw data calibration
tasks.antares.exp.addCalibration(tray)


#========================================================================
```

```
# Mask Pulses
#========================================================================
tray.AddModule("AntMaskOMCondition", "MaskPulses")(
   ("InputRecoPulses", "CalibratedPulses"),
   ("OutputRecoPulses", "PureCalibratedPulses"),
   ("OMConditionMap", "OMCondition")
   )

#========================================================================
# Calculate the mean rate and defective channel ratio
#========================================================================
tray.AddModule("AntCalcEventInfo", "CalcEventParams")(
   ("InputOMCondition","OMCondition"),
   ("EventHeader","I3EventHeader")
   )


#========================================================================
# The L1 trigger
#========================================================================
# this is the L1 trigger (4 modules)
# find big hits
tray.AddModule("I3HitSelectorModule<I3LowHighCutOff>",
          "HighThreshSelection")(
   ("INmap", "PureCalibratedPulses"),
   ("OUTmap", "PureCalibratedPulses_high_thresh"),
   ("MinNPE", bigHitAmp),
   ("MaxNPE", 1.e7)
   )

# find coincidences within the remaining hits
tray.AddModule("I3HitSelectorModule<I3LocalCoincidences>",
           "LocalCoincidenceSelection")(
   ("INmap", "PureCalibratedPulses_high_thresh_deselected"),
   ("OUTmap", "PureCalibratedPulses_coincidences"),
   ("Coincidence", coincidenceWindow),
   ("OnlyOneHitPerCoincidence", True),
   ("NoFrameNamingPostfix", True),
   ("WriteSelectedHits", True),
   ("WriteDeselectedHits", False)
   )

# build a combined list of big hits and floor coincidences
tray.AddModule("I3HitMerger","MergeSelections")(
   ("INmap1", "PureCalibratedPulses_high_thresh_selected"),
   ("INmap2", "PureCalibratedPulses_coincidences"),
   ("OUTmap", "PureCalibratedPulses_with_redundancies")
   )

# clean up this list by removing redundant coincidences
tray.AddModule("I3HitSelectorModule<I3RemoveRedundantCoincidences>",
          "RemoveRedundantCoincidences")(
   ("INmap", "PureCalibratedPulses_with_redundancies"),
   ("OUTmap", "PureCalibratedPulses_L1"),
   ("NoFrameNamingPostfix", True),
   ("WriteSelectedHits", True),
   ("WriteDeselectedHits", False),
   ("Coincidence", coincidenceWindow)
   )

#========================================================================
# The L2 trigger
#========================================================================
tray.AddModule("I3AntTriggerSimulator","TriggerSimulator")(
   ("InputRecoPulses", "PureCalibratedPulses_L1"),
   ("OutputRecoPulses", "PureCalibratedPulses_L2"),
   ("OutputTriggeredBool", "Triggered"),
   ("OutputTriggeredNum", "NumTriggers"),
   ("TriggerHierarchyName", "SimulatedTriggers"),
   ("WriteOutUntriggeredEvents", True),
   ("UseFloorPositions", True),
   ("DoTrigger3D", False),
   ("DoTrigger3N", True),
   ("DoTrigger3S", False),
   ("DoTrigger3T", True),
   ("DoTrigger1D", False)
   )
```

```
#=========================================================================
# Rebuild L1 Hits
#=========================================================================
tray.AddModule("I3HitSelectorModule<I3LocalCoincidencesWithHits>",
          "RebuildTriggeredHitsL1")(
   ("INmap", "PureCalibratedPulses"),
   ("ReferenceHitSeriesMap", "PureCalibratedPulses_L1"),
   ("OUTmap", "FullL1List"),
   ("CoincidenceGate", coincidenceWindow),
   ("NoFrameNamingPostfix", True),
   ("WriteSelectedHits", True),
   ("WriteDeselectedHits", False)
   )

#=========================================================================
# Rebuild L2 Hits
#=========================================================================
tray.AddModule("I3HitSelectorModule<I3LocalCoincidencesWithHits>",
          "RebuildTriggeredHitsL2")(
   ("INmap", "PureCalibratedPulses"),
   ("ReferenceHitSeriesMap", "PureCalibratedPulses_L2"),
   ("OUTmap", "FullL2List"),
   ("CoincidenceGate", coincidenceWindow),
   ("NoFrameNamingPostfix", True),
   ("WriteSelectedHits", True),
   ("WriteDeselectedHits", False)
   )


#=========================================================================
# Pick the events with alignment
#=========================================================================
# the icepick to select events
tray.AddModule("I3IcePickModule<I3AlignmentFilter>", "FilterAlignment")(
   ("DiscardEvents",False),
   ("OMConditionMap","OMCondition"),
   ("DecisionName","PassAlignment"),
   )

#=========================================================================
# Pick the events with good quality
#=========================================================================
# the icepick to select events
tray.AddModule("I3IcePickModule<I3DataQualityFilter>",
          "FilterDataQuality")(
   ("DiscardEvents",False),
   ("NoiseMeanRateName", "EventMeanRate"),
   ("NoiseLimit", 80.0),
   ("DefectiveChannelRatioName","DefectiveChannelRatio"),
   ("ChannelLimit",0.3),
   ("DecisionName","PassDataquality"),
   )


#=========================================================================
# Calculate lifetime and discard events that did not pass the filters
#=========================================================================
tray.AddModule("AntCalcLifetime", "CalcLifetime")(
   ("EventHeader","I3EventHeader"),
   ("TriggerHierarchy", "SimulatedTriggers"),
   ("InputRecoPulses", "PureCalibratedPulses"),
   ("DataqualityBool", "PassDataquality"),
   ("AlignmentBool", "PassAlignment"),
   ("DiscardEvents", True),
   ("SkipFirstNSeconds", 120),
   ("OutfileNameTxt", outtxtfile),
   )

#=========================================================================
# Write a .i3 file
#=========================================================================
tray.AddModule("I3Writer","Writer")(
   ("filename", outi3file)
   )

#=========================================================================
# Finish script
```

```
#========================================================================
tray.AddModule("TrashCan","TheEnd")
tray.Execute()
tray.Finish()

#xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
# Script end
#xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

The Python steering script to perform the processing of the simulations with *SeaTray*:

```
#xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
# Script start
#xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
#!/usr/bin/env python

from I3Tray import *
from os.path import expandvars

import os
import sys
import tasks.antares
import string

tray = I3Tray()

load("libdataclasses")
load("libphys-services")
load("libdataio")
load("libantares-reader")
load("libhit-selector")
load("libantares-tools")
load("libinterfaces")
load("libicepick")


#========================================================================
# Set parameters
#========================================================================
infile = <MC input file>
outfile = <path> + <filenmame> + ".root"
outi3file = <path> + <filenmame> + ".i3.gz"

# set according to data file
noise = 60.
bigHitAmp = 10.

geofile = expandvars("/pi1/common/antsoft/DETECTOR/r12_c00_s01.det")
background_rate = noise*1000
storeydist = 14.5
seed = <seed>

# these are the values used by the ANTARES tool "TriggerEfficiency"
coincidenceWindow = 20.*I3Units.ns
ARSThresholdAmp = 0.3
ARSIntegrationTime = 40.*I3Units.ns
switchTime = 7.*I3Units.ns
ARSDeadTime = 250.*I3Units.ns

# these are the values used by the ANTARES tool "TriggerEfficiency"
# determine number of events, generated for this input file
numberOfGeneratedEvents = 0
mupage=False
if string.find(infile, "mupage") is not -1:
   irradiationTime = 312* 1.*I3Units.s
   numberOfGeneratedEvents = 208.*30.*24.*3600.
   mupage=True
elif string.find(infile, "anumu") is not -1:
   irradiationTime = 1.*I3Units.s
   numberOfGeneratedEvents = 13*1E10
elif string.find(infile, "numu") is not -1 and
          string.find(infile, "anumu") is -1:
   irradiationTime = 1.*I3Units.s
   numberOfGeneratedEvents = 30*1E10


#========================================================================
# Install the services
#========================================================================
# fill the geometry stream with data from the ANTARES detector file
tray.AddService("I3AntTextFileGeometryServiceFactory","geometry")(
   ("OMAngularParametrization", "Spring09"),
   ("AntaresGeoFile",geofile)
   )

# read the text file
tray.AddService("I3AntTxtReaderServiceFactory","anttxtreader")(
```

```
   ("Filename",infile),
   ("Year",2007),
   ("DAQTime",173356670000000001),
   ("MCSeriesName","EvtMCHitSeries"),
   ("RawSeriesName","EvtRawHitSeries")
   )

# add empty streams where necessary
tray.AddService("I3EmptyStreamsFactory","empty_streams")(
   ("InstallGeometry",False),
   ("InstallEvent",False),
   ("InstallStatus",True),
   ("InstallCalibration",True),
   )

# install random number service
tray.AddService("I3SPRNGRandomServiceFactory","random")(
   ("Seed",seed),
   ("NStreams",seed+1),
   ("StreamNum",1)
   )

tray.AddModule("I3Muxer","muxer")


#=====================================================================
# Add the noise
#=====================================================================
tray.AddModule("I3NoiseHitsAdder","AddNoise")(
   ("WhiteNoiseRatePerPMT",noise*I3Units.kilohertz),
   ("NoiseOffsetTime",2500.*I3Units.ns),
   ("RemoveOldNoise",True),
   ("InputMCHits","EvtMCHitSeries"),
   ("OutputMCHits","EvtMCHitSeries_WithNewNoise")
   )

#=====================================================================
# The PMT simulator
#=====================================================================
tray.AddModule("I3AntPMSimulator","simpm")(
   ("InputMCHits","EvtMCHitSeries_WithNewNoise"),
   ("OutputRecoPulses","RecoPulseSeriesAfterARS"),
   ("OutputHitRelationMap","RecoPulseToMCHitRelation"),
   ("NumARS",2),
   ("IntegrationTimeARS",ARSIntegrationTime),
   ("SwitchTime", switchTime),
   ("DeadTimeARS",ARSDeadTime),
   ("AmplitudeThresholdARS",ARSThresholdAmp),
   ("DoTrigEffTTS", True),
   ("DoTrigEffGainRandomization", True)
   )


#=====================================================================
# Detector emulation
#=====================================================================
# set according to data file
tray.AddModule("AntEmulateOMCondition", "DetectorEmulation")(
   ("OMsLine1", [9,40,48,65]),
   ("OMsLine2", [6,10,19,28,64]),
   ("OMsLine3", [1,31,41,50,54,57,62,66,73]),
   ("OMsLine4", [11,15]),
   ("OMsLine5", [1,2,3,15,34,41,46,48,55,66,70,71,73,75]),
   ("OMsLine6", [12,13,14,15,555]),
   ("OMsLine7", [7,31,33,39,46,47,49,56,58,59,63,64,74]),
   ("OMsLine8", [27,49]),
   ("OMsLine9", [15,222,41,444]),
   ("OMsLine10", [4,5,6,13,14,21,35,59,67,68]),
   ("OMsLine11", [1,2,3,4,7,12,15,39,67,70,71]),
   ("OMsLine12", [5,17,28,33,34,35,36]),
   ("OutputOMConditionMap", "OMCondition"),
   ("NumberOfDeadOMs", 15)
   )

#=====================================================================
# Mask the pulses
#=====================================================================
tray.AddModule("AntMaskOMCondition", "MaskPulses")(
```

```
   ("InputRecoPulses", "RecoPulseSeriesAfterARS"),
   ("OutputRecoPulses", "FinalPulses"),
   ("OMConditionMap", "OMCondition")
   )


#========================================================================
# The L1 trigger
#========================================================================
tray.AddModule("I3HitSelectorModule<I3LowHighCutOff>",
          "HighThreshSelection")(
   ("INmap", "FinalPulses"),
   ("OUTmap", "FinalPulses_high_thresh"),
   ("MinNPE", bigHitAmp),
   ("MaxNPE", 1.e7)
   )

tray.AddModule("I3HitSelectorModule<I3LocalCoincidences>",
           "LocalCoincidenceSelection")(
   ("INmap", "FinalPulses_high_thresh_deselected"),
   ("OUTmap", "FinalPulses_coincidences"),
   ("Coincidence", coincidenceWindow),
   ("OnlyOneHitPerCoincidence", True),
   ("NoFrameNamingPostfix", True),
   ("WriteSelectedHits", True),
   ("WriteDeselectedHits", False)
   )

tray.AddModule("I3HitMerger","MergeSelections")(
   ("INmap1", "FinalPulses_high_thresh_selected"),
   ("INmap2", "FinalPulses_coincidences"),
   ("OUTmap", "FinalPulses_with_redundancies")
   )

tray.AddModule("I3HitSelectorModule<I3RemoveRedundantCoincidences>",
           "RemoveRedundantCoincidences")(
   ("INmap", "FinalPulses_with_redundancies"),
   ("OUTmap", "FinalPulses_L1"),
   ("NoFrameNamingPostfix", True),
   ("WriteSelectedHits", True),
   ("WriteDeselectedHits", False),
   ("Coincidence", coincidenceWindow)
   )

#========================================================================
#The L2 trigger
#========================================================================
tray.AddModule("I3AntTriggerSimulator","TriggerSimulator")(
   ("InputRecoPulses", "FinalPulses_L1"),
   ("OutputRecoPulses", "FinalPulses_L2"),
   ("OutputTriggeredBool", "Triggered"),
   ("OutputTriggeredNum", "NumTriggers"),
   ("TriggerHierarchyName", "SimulatedTriggers"),
   ("WriteOutUntriggeredEvents", False),
   ("UseFloorPositions", True),
   ("DoTrigger3D", False),
   ("DoTrigger3N", True),
   ("DoTrigger3S", False),
   ("DoTrigger3T", True),
   ("DoTrigger1D", False)
   )

#========================================================================
# Event building
#========================================================================
tray.AddModule("AntSnapshotMaker", "SnapshotMaker")(
   ("InputAllRecoPulses", "FinalPulses"),
   ("InputTriggeredPulses", "FinalPulses_L2"),
   ("OutputSnapshotPulses", "EventPulses"),
   )

#========================================================================
# Rebuild L1 hits
#========================================================================
tray.AddModule("I3HitSelectorModule<I3LocalCoincidencesWithHits>",
          "RebuildTriggeredHitsL1")(
   ("INmap", "EventPulses"),
   ("ReferenceHitSeriesMap", "FinalPulses_L1"),
```

```
   ("OUTmap", "FullL1List"),
   ("CoincidenceGate", 20.*I3Units.ns),
   ("NoFrameNamingPostfix", True),
   ("WriteSelectedHits", True),
   ("WriteDeselectedHits", False)
   )

#========================================================================
# Rebuild L2 hits
#========================================================================
tray.AddModule("I3HitSelectorModule<I3LocalCoincidencesWithHits>",
          "RebuildTriggeredHitsL2")(
   ("INmap", "EventPulses"),
   ("ReferenceHitSeriesMap", "FinalPulses_L2"),
   ("OUTmap", "FullL2List"),
   ("CoincidenceGate", coincidenceWindow),
   ("NoFrameNamingPostfix", True),
   ("WriteSelectedHits", True),
   ("WriteDeselectedHits", False)
   )


#========================================================================
# Write a .i3 file
#========================================================================
tray.AddModule("I3Writer","Writer")(
   ("filename", outi3file)
   )

#========================================================================
# Finish script
#========================================================================
tray.AddModule("TrashCan","TheEnd")
tray.Execute()
tray.Finish()

#xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
# Script end
#xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

The Python steering script to perform the full reconstruction processing with *SeaTray*, as done within context of this thesis. The reconstruction procedure is the same both for data and simulations.

```
#xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
# Script start
#xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
#!/usr/bin/env python

from I3Tray import *
from os.path import expandvars

import os
import sys
import string

tray = I3Tray()

load("libdataclasses")
load("libphys-services")
load("libantares-reader")
load("libhit-selector")
load("libantares-tools")
load("libicepick")
load("libposidonia")

#========================================================================
# Set parameters
#========================================================================
infile = <input file> + ".i3.gz"
outi3file = <path> + <filenmame> + ".i3.gz"

storeydist = 14.5

# When processing data, replace 'EventPulses' by 'PureCalibratedPulses'

#========================================================================
# File reader
#========================================================================
tray.AddService("I3ReaderServiceFactory", "Reader")(
   ("Filename", infile ),
   )

tray.AddModule("I3Muxer","muxer")

#========================================================================
# BBfit track reconstruction
#========================================================================
tray.AddModule("I3BBFitReco", "BBReco")(
   ("InputPulseSeriesMap","EventPulses"),
   ("InputTriggeredPulseSeriesMap","FinalPulses_FullL1List"),
   ("FitResultNameTrack","BBTrack"),
   ("FitResultNameBrightPoint","BBBright"),
   ("FitParamsName","BBInfo"),
   ("AmplitudeCut",2.5)
   )

#========================================================================
# BBfit track zenith filter
#========================================================================
tray.AddModule("OscillationZenithFilter","ZenithFilterBBTrack")(
   ("BBTrack","BBTrack"),
   ("MinZenith", 115.),
   )


#========================================================================
#HM Hitselection
#========================================================================
tray.AddModule("I3HitSelectorModule<I3ClusterHits>", "XClusterHits")(
   ("MaxTimeDifference",150.0*I3Units.ns),
   ("MaxDistance",3*storeydist*I3Units.m),
   ("MinHits",3),
   ("MaxRMSInc",4.0*I3Units.degree),
   ("OriginHitOnly",True),
   ("INmap", "EventPulses"),
```

```
   ("OUTmap", "XClusterHits"),
   ("WriteDeselectedHits", False),
   ("NoFrameNamingPostfix", True),
   )
tray.AddModule("I3HitSelectorModule<I3ClusterHits>", "YClusterHits")(
   ("MaxTimeDifference",50.0*I3Units.ns),
   ("MaxDistance",2*storeydist*I3Units.m),
   ("MinHits",2),
   ("OriginHitOnly",True),
   ("INmap", "EventPulses"),
   ("OUTmap", "YClusterHits"),
   ("WriteDeselectedHits", False),
   ("NoFrameNamingPostfix", True),
   )
tray.AddModule("I3HitSelectorModule<I3ClusterHits>", "ZClusterHits")(
   ("MaxTimeDifference",200.0*I3Units.ns),
   ("MaxDistance",5*storeydist*I3Units.m),
   ("MinHits",4),
   ("OriginHitOnly",True),
   ("MaxRMSInc",4.0*I3Units.degree),
   ("INmap", "EventPulses"),
   ("OUTmap", "ZClusterHits"),
   ("WriteDeselectedHits", False),
   ("NoFrameNamingPostfix", True),
   )

tray.AddModule("I3HitSelectorModule<I3LowHighCutOff>", "XLargeHits")(
   ("INmap", "XClusterHits"),
   ("MaxNPE", 1000000),
   ("MinNPE", 1.5),
   ("OUTmap", "XLargeHits"),
   ("WriteDeselectedHits", False),
   ("NoFrameNamingPostfix", True),
   )
tray.AddModule("I3HitSelectorModule<I3LowHighCutOff>", "YLargeHits")(
   ("INmap", "YClusterHits"),
   ("MaxNPE", 1000000),
   ("MinNPE", 2.0),
   ("OUTmap", "YLargeHits"),
   ("WriteDeselectedHits", False),
   ("NoFrameNamingPostfix", True),
   )
tray.AddModule("I3HitSelectorModule<I3LowHighCutOff>", "ZLargeHits")(
   ("INmap", "ZClusterHits"),
   ("MaxNPE", 1000000),
   ("MinNPE", 2.0),
   ("OUTmap", "ZLargeHits"),
   ("WriteDeselectedHits", False),
   ("NoFrameNamingPostfix", True),
   )
tray.AddModule("I3HitSelectorModule<I3IdenticalHits>", "XYClusterHits")(
   ("ReferenceHitSeriesMap","YClusterHits"),
   ("INmap", "XClusterHits"),
   ("OUTmap", "XYClusterHits"),
   ("WriteDeselectedHits", False),
   ("NoFrameNamingPostfix", True),
   )
tray.AddModule("I3HitSelectorModule<I3IdenticalHits>", "YZClusterHits")(
   ("ReferenceHitSeriesMap","YClusterHits"),
   ("INmap", "ZClusterHits"),
   ("OUTmap", "YZClusterHits"),
   ("WriteDeselectedHits", False),
   ("NoFrameNamingPostfix", True),
   )
tray.AddModule("I3HitSelectorModule<I3IdenticalHits>", "XZClusterHits")(
   ("ReferenceHitSeriesMap","XClusterHits"),
   ("INmap", "ZClusterHits"),
   ("OUTmap", "XZClusterHits"),
   ("WriteDeselectedHits", False),
   ("NoFrameNamingPostfix", True),
   )
tray.AddModule("I3HitMerger", "Meta_One")(
   ("INmap1", "XYClusterHits"),
   ("INmap2", "YZClusterHits"),
   ("OUTmap", "Meta_One"),
   )
tray.AddModule("I3HitMerger", "Meta_Two")(
   ("INmap1", "Meta_One"),
```

```
   ("INmap2", "XZClusterHits"),
   ("OUTmap", "Meta_Two"),
   )
tray.AddModule("I3HitMerger", "Meta_Three")(
   ("INmap1", "Meta_Two"),
   ("INmap2", "XLargeHits"),
   ("OUTmap", "Meta_Three"),
   )
tray.AddModule("I3HitMerger", "Meta_Four")(
   ("INmap1", "Meta_Three"),
   ("INmap2", "YLargeHits"),
   ("OUTmap", "Meta_Four"),
   )
tray.AddModule("I3HitMerger", "Meta_Five")(
   ("INmap1", "Meta_Four"),
   ("INmap2", "ZLargeHits"),
   ("OUTmap", "ClusterHits"),
   )

#=========================================================================
# Merge with L2 hits
#=========================================================================
tray.AddModule("I3HitMerger","MergeSelections")(
   ("INmap1", "ClusterHits"),
   ("INmap2", "FullL2List"),
   ("OUTmap", "HMHits")
   )


#=========================================================================
# Posidonia event classifier
#=========================================================================
tray.AddModule("PosidoniaEventClassifier","Classifier")(
   ("EventHeader","I3EventHeader"),
   ("InputAllRecoPulses","HMHits"),
   ("InputCoincidentPulses","FullL2Pulses"),
   ("NStrings", 13),
   ("NStringsWithCoincs", 2),
   ("LogicalOperator", "and"),
   ("OutputInfoName","EventInfo"),
   ("OutputName","ClassificationResult")
   )

#=========================================================================
# 3D Prefit
#=========================================================================
tray.AddModule("PosidoniaLinearPrefit","Prefit")(
   ("InputRecoPulses","FullL2Pulses"),
   ("EventClassificationResult", "ClassificationResult"),
   ("OutputTrack","3DPrefit"),
   )

#=========================================================================
# Posidonia zenith filter
#=========================================================================
tray.AddModule("OscillationZenithFilter","ZenithFilterPrefit")(
   ("BBTrack","BBTrack"),
   ("ParticleKey","3DPrefit"),
   ("PosidoniaClassificationResult","ClassificationResult"),
   ("MinZenith", 115.),
   )

#=========================================================================
# Posidonia 3D hit selection
#=========================================================================
tray.AddModule("I3HitSelectorModule<I3SmallTimeResidualsAndDistances>",
         "SmallTimeResidualSelection")(
   ("INmap", "HMHits"),
   ("OUTmap", "HitsCompatibleWithPrefit"),
   ("WriteDeselectedHits", False),
   ("NoFrameNamingPostfix", True),
   ("ReferenceParticleName", "3DPrefit"),
   ("TimeResidualsInUnitsOfRMS", False),
   ("MinDt",  -100.*I3Units.ns),
   ("MaxDt",    80.*I3Units.ns),
   ("MaxDist", 120.*I3Units.m),
   ("AlwaysKeepAmpThreshold", 999999.)
   )
```

```
#========================================================================
# Posidonia 3D scan fit
#========================================================================
tray.AddModule("PosidoniaMainFit", "3DScanFit")(
   ("EventClassificationResult","ClassificationResult"),
   ("InputRecoPulses","HitsCompatibleWithPrefit"),
   ("OutputFinalFit", "3DFinalfit"),
   ("OutputFitInfo", "3DFitinfo"),
   ("OutputAmplitudes", "3DAmplitudes"),
   ("OutputResiduals", "3DResiduals"),
   ("Do1DScanFit", "Do1DScanFit"),
   )

#========================================================================
# Posidonia: information about 3D track, containment estimation
#========================================================================
tray.AddModule("PosidoniaCalcTrackInfo", "CalcTrackInfo3D")(
   ("InputFitCollection","HitsCompatibleWithPrefit"),
   ("InputContainmentCollection","HMHits"),
   ("InputTrack","3DFinalfit"),
   ("OutputTrackInfo","3DTrackinfo"),
   ("InputEventType", "3D"),
   )

#========================================================================
# Posidonia 1D string selector
#========================================================================
tray.AddModule("Posidonia1DStringSelector","StringSelector")(
   ("EventClassificationResult", "ClassificationResult"),
   ("InputCoincidentPulses", "FullL2Pulses"),
   ("InputSinglePulses", "HMHits"),
   ("OutputHitCollectionName","1DCollection"),
   ("OutputMainStringNumber","Mainstring"),
   ("Do1DScanFit", "Do1DScanFit")
   )

#========================================================================
# Posidonia 1D scan fit
#========================================================================
tray.AddModule("Posidonia1DScan","1DScanFit")(
   ("InputPosidoniaHitCollection","1DCollection"),
   ("InputMainString","Mainstring"),
   ("OutputFinalFit", "1DFinalfit"),
   ("OutputFitInfo", "1DFitinfo"),
   ("OutputAmplitudes", "1DAmplitudes"),
   ("OutputResiduals", "1DResiduals"),
   )

#========================================================================
#1D CalcTrackInfo
#========================================================================
tray.AddModule("PosidoniaCalcTrackInfo", "CalcTrackInfo3D")(
   ("InputFitCollection","1DCollection"),
   ("InputContainmentCollection","HMHits"),
   ("InputTrack","1DFinalfit"),
   ("OutputTrackInfo","1DTrackinfo"),
   ("InputEventType", "1D"),
   )

#========================================================================
# Write a .i3 file
#========================================================================
tray.AddModule("I3Writer","Writer")(
   ("filename", outi3file)
   )

#========================================================================
# Finish script
#========================================================================
tray.AddModule("TrashCan","TheEnd")
tray.Execute()
tray.Finish()

#xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
# Script end
#xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

# Zusammenfassung

Im Jahre 1911 entdeckte Victor Hess die kosmische Strahlung (Nobel Preis 1936), und öffnete damit ein völlig neues Fenster zu unserem Universum. Kosmische Strahlen sind geladene Botenteilchen aus dem All, die permanent auf unsere Erdatmosphäre auftreffen. Sie tragen Information, welche helfen kann, noch offene astrophysikalische Fragen zu beantworten.

Kosmische Strahlung wird hauptsächlich auf der Erde über Sekundärstrahlung (auch atmosphärische Strahlung genannt) gemessen, die durch die Wechselwirkung der kosmischen Teilchen mit Gasmolekülen der Erdatmosphäre entsteht. Dabei werden neue Teilchen erzeugt, die wiederum wechselwirken und weitere Teilchen erzeugen. Eine hadronische oder elektromagnetische Kaskade (Schauer) bildet sich aus. Über den Nachweis dieser Sekundärstrahlung kann auf das Primärteilchen rückgeschlossen werden.

Das Spektrum der primären kosmischen Strahlung reicht über viele Größenordnungen bis jenseits von $E = 10^{20}\,\mathrm{eV}$. Dies übersteigt die Energie, die in den aktuell leistungsfähigsten Teilchenbeschleunigern erzeugt werden kann um mehrere Größenordnungen. Obwohl über die kosmische Strahlung bereits viel bekannt ist, gibt es immer noch ungeklärte fundamentale Fragen wie beispielsweise Fragen nach ihrer Herkunft und ihren Beschleunigungsprozessen. Da geladene Teilchen in galaktischen und intergalaktischen magnetischen Feldern abgelenkt werden, treffen sie isotrop auf die Erde und lassen keine Rückschlüsse über ihre Herkunft zu. Nur bei höchsten Energien ist die Ablenkung klein genug, so dass die Teilchen auf ihren Ursprungsort zurückdeuten.

Neben geladenen Teilchen zählen auch hochenergetische Photonen ($E > 100\,\mathrm{keV}$) und Neutrinos zur kosmischen Strahlung[1]. Während $\gamma$-Strahlen durch verschiedene Mechanismen in leptonischen (z.B. Synchrotronstrahlung) oder hadronischen Prozessen (z.B. Pion-Zerfall) erzeugt werden können, werden Neutrinos nur in hadronischen Szenarien erzeugt. Der Nachweis von hochenergetischen Neutrinos würde daher direkt die Existenz relativistischer Hadronen beweisen. Neutrinos propagieren im Wesentlichen ungestört durch das Universum, da sie in magnetischen Feldern nicht abgelenkt werden und sie einen sehr kleinen Wechselwirkungsquerschnitt haben. Allerdings sind diese Vorteile auch gleichzeitig der entscheidende Nachteil beim Nachweis von Neutrinos: aufgrund des kleines Wirkungsquerschnittes sind sie schwer nachzuweisen, und große Detektoren sind nötig, um zumindest einige wenige kosmische Neutrinos zu messen.

Das ANTARES Neutrino Teleskop wurde gebaut um hochenergetische kosmische Neutrinos zu detektieren. Der Detektor ist etwa 24 km vor der französischen Küste in der Nähe der Stadt Toulon am Grund des Mittelmeeres verankert. ANTARES nutzt das

[1] Einige Autoren verwenden den Ausdruck "kosmische Strahlung" nur für die geladene Komponente der Strahlung.

Meereswasser als Detektionsmedium und besteht aus einer dreidimensionalen Anordnung von Photomultipliern, die ein Volumen von etwa $0.03\,\mathrm{km}^3$ instrumentieren. In der schwachen charged-current Wechselwirkung von hochenergetischen Neutrinos mit Nukleonen werden relativistische, geladene Teilchen erzeugt. Beim Durchgang durch ein transparentes Medium emittieren diese Cherenkovlicht, welches von den Photomultipliern detektiert wird.

ANTARES ist auf den Nachweis von Myon-Neutrinos optimiert. In charged-current Wechselwirkungen mit Nukleonen erzeugen diese Myonen, welche aufgrund ihrer lange Spurlänge eine präzise Spurrekonstruktion bis auf einige zehntel Grad ermöglichen. Die Länge der Myonspur ist von der Energie des Myons abhängig und reicht von einigen zehn Metern bei 10 GeV bis zu einigen Kilometern bei 1 PeV. Durch die Instrumentierung des Detektors und den im Detektor vorhandenen optischen Untergrund (hauptsächlich Zerfälle von $^{40}$K und Bioluminescenz) liegt die Nachweisschwelle von Myonen mit dem ANTARES Detektor bei einer Energie von etwa 10 bis 20 GeV.

Das häufigste Neutrinosignal in ANTARES sind atmosphärische Neutrinos, welche Teil der sekundären kosmischen Strahlung sind. Ein potentielles kosmisches Neutrinosignal ist in diesen irreduziblen Untergrund eingebettet, und kann nur durch einen Überschuss an Ereignissen aus einer bestimmten Himmelsrichtung oder bei sehr hohen Energien identifiziert werden. Atmosphärische Neutrinos selbst ermöglichen allerdings verschiedenen Arten von Studien. Das Phänomen der Neutrino-Oszillationen beispielsweise, welches Thema der vorliegenden Arbeit ist, kann anhand atmosphärischer Neutrinos untersucht werden.

Neutrinos werden als reine Flavour-Eigenzustände erzeugt und detektiert, wohingegen ihre Zeitentwicklung in Masse-Eigenzuständen beschrieben wird. Flavour- und Masse-Eigenzustände von Neutrinos sind nicht identisch. Ein reiner Flavour-Eigenzustand ist aus den drei Masse-Eigenzuständen zusammengesetzt, was bei einer zeitlichen Entwicklung zu Interferenzeffekten, und somit zu einer Änderung der Flavour-Zusammensetzung führt. Aufgrund der Werte, die in verschiedenen Experimenten bisher bestimmt wurden, kann das allgemeine Drei-Flavour Mischungsszenario durch zwei Zwei-Flavour Szenarien angenähert werden. Die Überlebenswahrscheinlichkeit für ein atmosphärisches Myon-Neutrino hängt danach nur vom Mischungswinkel $\Theta_{23}$ (oft auch in der Form $\sin^2(2\Theta_{23})$ angegeben), von der Differenz der Massenquadrate $\Delta m^2_{23}$, sowie von der Neutrinoenergie $E_\nu$ ab. Berücksichtigt man Neutrinos mit unterschiedlichen Zenitwinkeln, so ergibt sich eine Abhängigkeit von $E_\nu/\cos\Theta'_\nu$, mit dem Winkel $\Theta'_\nu = \pi - \Theta_\nu$ und dem Zenitwinkel des Neutrinos $\Theta_\nu$.

Der Zenitwinkel und die Energie eines von ANTARES nachgewiesenen Myons können mit einem geeigneten Rekonstruktionsalgorithmus, wie er im Rahmen der Arbeit entwickelt und optimiert wurde, berechnet werden. Die Kinematik der Neutrinowechselwirkung lässt es schließlich zu, (mit energieabhängiger Genauigkeit) auf die Richtung und die Energie des ursprünglichen Neutrinos rückzuschließen. Entsprechend dem Zwei-Flavour Mischungsszenario und gegebenen Werten für $\Theta_{23}$ und $\Delta m^2_{23}$, manifestiert sich eine Flavour-Oszillation in einem Defizit in der $E_{\mu,\mathrm{rec}}/\cos\Theta'_{\mu,\mathrm{rec}}$ Verteilung.

Unter der Annahme von Werten für $\Theta_{23}$ und $\Delta m^2_{23}$, die im Jahr 2008 von der MINOS Kollaboration veröffentlicht wurden [122], liegt ein Minimum der Myon-Neutrino Über-

lebenswahrscheinlichkeit innerhalb der ANTARES Sensitivität bei etwa $E_\nu / \cos\Theta'_\nu = 25\,\mathrm{GeV}$. Aufgrund der Kinematik und der Fehler durch die Rekonstruktion führt dies zu einem breiten Defizit an Neutrino-Ereignissen in den ANTARES Daten im rekonstruierten $E_{\mu,\mathrm{rec}} / \cos\Theta'_{\mu,\mathrm{rec}}$ Spektrum zwischen 10 und 50 GeV. Da dieser Energiebereich an der Sensitivitätsschwelle von ANTARES liegt, ist eine Analyse, die auf die Bestimmung der sogenannten atmosphärischen Oszillationsparameter $\Theta_{23}$ und $\Delta m^2_{23}$ abzielt eine große Herausforderung.

Im Rahmen der vorliegenden Arbeit wurden die Daten des ANTARES Neutrino Teleskops auf eine Signatur von Neutrino-Oszillationen hin untersucht. Zu diesem Zweck wurde ein dedizierter Niederenergie-Rekonstruktionsalgorithmus in das offizielle ANTARES Software-Framework implementiert. Der Algorithmus war ursprünglich zu einer Zeit entwickelt worden, in der das endgültige Design von ANTARES noch nicht festgelegt war [109, 110, 111]. Im Rahmen der Arbeit wurde der Algorithmus überarbeitet und an das finale Detektorlayout und die aktuellen Simulationen angepasst. Außerdem wurde der Algorithmus auf unterschiedliche Weise verbessert, beispielsweise durch das Einfügen einer neuen Wahrscheinlichkeitsdichtefunktion (PDF), oder durch das Verwenden neuer Hitfilter.

Für eine Neutrino-Oszillationsanalyse mit ANTARES ist neben einem effizienten Niederenergie-Rekonstruktionsalgorithmus auch eine gute Datenauswahl wichtig. Im Rahmen der Arbeit wurde ein Konzept zur Selektion der ANTARES-Daten erarbeitet, welches die Umgebungsbedingungen und den Zustand des Detektors zur Zeit der Datennahme ermittelt und prüft. Weiterhin wurden die Detektorsimulationen so verändert, dass der sich zeitlich ändernde Detektorzustand berücksichtigt wird.

In den selektierten Daten wurden Myonspuren identifiziert, und der oben genannte Rekonstruktionsalgorithmus wurde angewendet um die Zenitwinkel der Myonspuren zu berechnen. Da fehlrekonstruierte atmosphärische Myonen aus atmosphärischen Schauern das atmosphärische Neutrinosignal um drei bis vier Größenordnungen übertreffen, sind effektive Qualitätsschnitte notwendig, um diese zu unterdrücken. Zu diesem Zweck wurde ein neuronales Netz verwendet, welches die Information verschiedener Eingangsparameter effektiv zu einem einzelnen Schnittparameter kombiniert. Nur für Ereignisse, die durch den Schnitt auf den Ausgangsparameter des neuronalen Netzes selektiert wurden, wurden eine Energierekonstruktion durchgeführt.

Die Energierekonstruktion besteht in einem ersten Schritt aus der Auswahl von so genannten "contained" Ereignissen, deren Start- und Endpunkt im Detektor enthalten sind. Durch diese Auswahl wird die mittlere Energie der selektierten Myon-Ereignisse reduziert, und eine verlässliche Spurlängenrekonstruktion ist möglich. Die Energie der einzelnen "contained" Ereignisse kann schließlich über die berechnete Spurlänge ermittelt werden.

Das so erhaltene $E_{\mu,\mathrm{rec}} / \cos\Theta'_{\mu,\mathrm{rec}}$ Spektrum wurde mit simulierten, auf unterschiedlichen Oszillations-Szenarien (d.h. unterschiedlichen Werten für die Oszillationsparameter $\Theta_{23}$ und $\Delta m^2_{23}$) beruhenden Spektren verglichen. Über die Minimierung einer $\chi^2$ Statistik [124] wurde das wahrscheinlichste Oszillationsszenario bestimmt. Für die Konstruktion einer 90 % Konfidenzregion wurde ein frequentistischer Ansatz zur Analyse kleiner Signale verwendet [123].

Um das Ergebnis der Datenanalyse einordnen und bewerten zu können, wurden die Analyse und die Konstruktion der Konfidenzregion vor Anwendung auf die Daten mit Toy-Daten getestet. Diese Toy-Daten basieren auf einem Oszillations-Szenario gemäß den von MINOS veröffentlichten Werten [122] ($\Delta m^2_{23} = 10^{-2.6}\,\mathrm{eV}^2$ und $\sin^2(2\Theta_{23}) = 1$). Die Toy-Daten Studie zeigte, dass bei einer kleinen Ereignis-Statistik von 150 Neutrinos die $\chi^2$ Verteilung keine klare Unterscheidung zwischen verschiedenen Szenarien zulässt. Dementsprechend variieren die Konfidenzregionen unterschiedlicher Toy-Datensätze (jeweils auf dem MINOS Szenario beruhend und mit gleicher Statistik von 150 Ereignissen) stark in Größe und Form. Unter Verwendung von (ebenfalls auf dem MINOS Szenario beruhenden) Toy-Datensätzen mit einer großen Ereignis-Statistik von 1 000 Neutrinos dagegen ist die $\chi^2$ Verteilung wesentlich ausgeprägter und führt dementsprechend zu kleineren Konfidenzregionen. Diese sind stabiler in Form und Größe als die Konfidenzregionen der kleinen Toy-Datensätze mit 150 Neutrino-Ereignissen.

Die Anwendung der Analyse auf die Daten ergab als besten Fit die Werte $\Delta m^2_{23} = 10^{-2.8}$ und $\sin^2(2\Theta_{23}) = 1$, was nahe an den MINOS Werten liegt. Die Daten enthalten im für die Analyse berücksichtigten Energiebereich 91 Ereignisse. Dies liegt noch unter der Statistik der kleinen Toy-Datensätze mit 150 Ereignissen. Aus der Untersuchung mit den Toy-Daten lässt sich der Schluss ziehen, dass das Ergebnis nicht signifikant ist. Dies spiegelt sich auch in der 90 % Konfidenzregion der Daten wieder, welche die komplette getestete Parameterebene bedeckt.

Durch Verwendung einer absoluten anstelle einer relativen Normierung der simulierten $E_{\mu,\mathrm{rec}}/\cos\Theta'_{\mu,\mathrm{rec}}$ Verteilungen ergab sich der beste Fit zu $\Delta m^2_{23} = 10^{-0.7}$ und $\sin^2(2\Theta_{23}) = 1$. Dies ist weit abseits des erwarteten $\Delta m^2_{23}$ Wertes, der durch andere Oszillationsexperimente, wie beispielsweise MINOS, bereits gut bestimmt ist.

Die Ursache dafür ist vermutlich, dass die Übereinstimmung zwischen Daten und Simulationen nicht zufriedenstellend ist. Dieses Problem wird nicht nur im finalen $E_{\mu,\mathrm{rec}}/\cos\Theta'_{\mu,\mathrm{rec}}$ Spektrum deutlich, sondern bereits in elementareren Verteilungen, wie der Anzahl von Hits pro Ereignis oder der gemessenen Lichtmenge pro Ereignis. Diese Diskrepanz ist kein spezifisches Problem der hier präsentierten Arbeit, sondern noch ein generelles Problem der ANTARES Detektorsimulation. Der ANTARES Detektor befindet sich nach der Fertigstellung im Sommer 2008 noch in seiner Commissioning Phase, und es wird aktuell innerhalb der Kollaboration daran gearbeitet, die Übereinstimmung zwischen Daten und Simulation zu verbessern. Es wurden auf diesem Gebiet bereits große Fortschritte erzielt, die jedoch in der vorliegenden Arbeit nicht mehr berücksichtigt werden konnten.

Die Studie mit den Toy-Daten hat gezeigt, dass eine gute Bestimmung der Oszillationsparameter mit einer kleinen Konfidenzregion möglich ist, wenn die Ereignisstatistik ausreichend groß ist. Eine Anzahl von 1 000 Neutrino Ereignissen, mit der in der Studie gute Ergebnisse erzielt werden konnten, kann beispielsweise innerhalb von sechs bis sieben Jahren Datennahme erhalten werden (unter Extrapolation der vorliegenden Daten). Gründet man die Abschätzung auf optimistische Annahmen, wie beispielsweise ein vollständig funktionierenden Detektor, niedrige optische Untergrundraten und möglicherweise einen weiteren Anstieg der Effizienz von Rekonstruktion und Ereignisselektion, so kann solch eine Ereignis-Statistik bereits innerhalb von ein oder zwei

Jahren erzielt werden. Die Analyse kann zum gegebenen Zeitpunkt leicht wiederholt werden. Unter Einbeziehung der Ergebnisse anderer Experimente zur Bestimmung der Oszillationsparameter könnten dann auch wertvolle Schlüsse über das Verständnis des ANTARES Detektors im Niederenergiebereich gezogen werden.

# Bibliography


[1] URL *http://icecube.wisc.edu/*

[2] URL *http://www.km3net.org/home.php*

[3] B. Pontecorvo, Journal of Experimental and Theoretical Physics 6, 429, 1957

[4] B. Pontecorvo, Journal of Experimental and Theoretical Physics 7, 172, 1958

[5] Z. Maki, M. Nakagawa, S. Sakata, *Remarks on the Unified Model of Elementary Particles*, Progress of Theoretical Physics, 28, , pp. 870 – 880, 1962

[6] B. Pontecorvo, Journal of Experimental and Theoretical Physics 26, 984, 1968

[7] URL *http://www-sk.icrr.u-tokyo.ac.jp/sk/index-e.html*

[8] The Super-Kamiokande Collaboration: Y. Fukuda et al., *Evidence of oscillation of atmospheric neutrinos*, Physical Review Letters, 81, pp. 1562 – 1567, 1998

[9] D. Griffiths, *Einführung in die Elementarteilchenphysik*, Akademie Verlag, 1996, ISBN 3-05-501627-0

[10] C. L. Cowan, F. Reines, *The Neutrino*, Nature, 178, pp. 446 – 449, 1956

[11] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, R. P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Physical Review, 105, pp. 1413 – 1415, 1957

[12] M. Goldhaber, L. Grodzins, A. W. Sunyar, *Helicity of neutrinos*, Physical Review, 109, pp. 1015 – 1017, 1958

[13] The OPAL Collaboration: M. Z. Akrawy et al., *Measurement of the $Z^0$ mass and width with the opal detector at LEP*, Physics Letters B, 231, 4, pp. 530 – 538, 1989

[14] G. Danby, J. M. Gaillard, K. A. Goulianos, L. M. Lederman, N. B. Mistry, M. Schwartz, J. Steinberger, *Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos*, Physical Review Letters, 9, pp. 36– 44, 1962

[15] URL *http://physicsworld.com/cws/article/news/2807*

[16] The DONUT Collaboration: K. Kodama et al., *Observation of tau neutrino interactions*, Physics Letters B, 504, 3, pp. 218 – 224, 2001

[17] R. Davis, Jr., D. S. Harmer, K. C. Hoffman, *Search for neutrinos from the sun*, Physical Review Letters, 20, 21, pp. 1205 – 1209, 1968

[18] R. Davis, *A Review of the Homestake Solar Neutrino Experiment*, Progress in Particle and Nuclear Physics, pp. 13 – 32, 1994

[19] C. Kraus et al., *Latest results of the Mainz Neutrino Mass Experiment*, European Physical Journal C, 33, pp. 805 – 807, 2004

[20] K. Nakamura et al. (Particle Data Group), Journal of Physics G, 37, 075021, 2010, (URL: http://pdg.lbl.gov)

[21] H. Nunokawa, S. Parke, J. W. F. Valle, *CP violation and neutrino oscillations*, Progress in Particle and Nuclear Physics, 60, pp. 338 – 402, 2008

[22] R. N. Mohapatra et al., *Theory of neutrinos: a white paper*, Reports on Progress in Physics, 70, pp. 1757 – 1867, 2007

[23] J. M. Conrad, *Neutrino Experiments*, arXiv: 0708.2446v1 [hep-ex], 2007

[24] N. Schmitz, *Neutrinophysik*, Teubner Studienbücher Physik, 1997, ISBN 3-519-03236-8

[25] http://pdg.lbl.gov C. Amsler et al. (Particle Data Group), Physics Letters B, 667, 1, 2008

[26] C. Grupen, *Astroparticle physics*, Springer-Verlag Berlin, 2005, ISBN 3-540-25312-2

[27] URL *http://www.mpi-hd.mpg.de/hfm/HESS/*

[28] URL *http://wwwmagic.mppmu.mpg.de/index.en.html*

[29] URL *http://www.mpi-hd.mpg.de/hfm/HESS/pages/home/sources/*, 16.11.2010

[30] R. M. Bionta et al., *Observation of a Neutrino Burst in Coincidence with Supernova 1987A in the Large Magellanic Cloud*, Physical Review Letters, 58, 14, pp 1494 – 1496, 1987

[31] The Super-Kamiokande Collaboration: Y. Ashie et al., *Measurement of atmospheric neutrino oscillation parameters by Super-Kamiokande I*, Physical Review D, 71, 11, p. 112005, 2005

[32] The IceCube Collaboration: R. Abbasi et al., *Determination of the Atmospheric Neutrino Flux and Searches for New Physics with AMANDA-II*, arXiv: 0902.0675v2 [astro-ph.HE], 2009

[33] P. Lipari, *Neutrino oscillation studies and the neutrino cross section*, Nuclear Physics - Proceedings Supplements, 112, pp. 274 – 287, 2002

[34] B. Povh, K. Rith, C. Scholz, F. Zetsche, *Teilchen und Kerne*, Springer-Verlag Berlin, 2004, ISBN 3-540-68075-6

[35] R. Gandhi, C. Quigg, M. H. Reno, I. Sarcevic, *Ultrahigh-energy neutrino interactions*, Astroparticle Physics, 5, pp. 81 – 110, 1996

[36] URL *http://doublechooz.in2p3.fr/*

[37] The Borexino Collaboration: G. Alimonti et al., *Science and Technology of Borexino: A Real Time Detector for Low Energy Solar Neutrinos*, Astroparticle Physics, 16, pp. 205 – 234, 2002

[38] T. A. Kirsten, *Gallex solar neutrino results*, Progress in Particle and Nuclear Physics, 40, pp. 85 – 99, 1998

[39] URL *http://www.auger.org/*

[40] H. J. Blümer, D. Gora, M. Roth, A. Tamburro, *Neutrino Detection with the Surface Array of the Pierre Auger Observatory*, arXiv: 0810.1860v2 [astro-ph], 2008

[41] A. Karle for the IceCube Collaboration, *IceCube - the next generation neutrino telescope at the South Pole*, arXiv: 0209556v1 [astro-ph], 2002

[42] R. Bruijn, *The Antares Neutrino Telescope: Performance Studies and Analysis of First Data*, Ph.D. thesis, University of Amsterdam, 2008

[43] G. A. Askaryan, B. A. Dolgoshein, A. H. Kalinovsky, N. V. Mokhov, *Acoustic Detection of High Energy Particle Showers in Water*, Nuclear Instruments and Methods, 164, pp. 267 – 278, 1979

[44] The Antares Collaboration: J. A. Aguilar et al., *AMADEUS - The Acoustic Neutrino Detection Test System of the ANTARES Deep-Sea Neutrino Telescope*, Nuclear Instruments and Methods in Physics Research A, 626 – 627, pp. 128 – 143, 2011

[45] R. Lahmann for the Antares Collaboration, *Status and First Results of the Acoustic Detection Test System AMADEUS*, arXiv: 0901.0321v1 [astro-ph.IM], 2010

[46] URL *http://www.sno.phy.queensu.ca/*

[47] The SNO Collaboration: B. Aharmim et al., *Low-energy-threshold analysis of the Phase I and Phase II data sets of the Sudbury Neutrino Observatory*, Physical Review C, 81, 055504, 2010

[48] URL *http://borex.lngs.infn.it/*

[49] URL *http://www.pi.infn.it/chooz/*

[50] M. Apollonio et al., *Limits on Neutrino Oscillations from the CHOOZ Experiment*, Physics Letters B, 466, pp. 415 – 430, 1999

[51] URL *http://kamland.lbl.gov/*

[52] The KamLAND Collaboration: S. Abe et al., *Precision Measurement of Neutrino Oscillation Parameters with KamLAND*, Physical Review Letters, 100, 221803, 2008

[53] The KamLAND Collaboration: A. Gando et al., *Constraints on* $\Theta_{13}$ *from A Three-Flavor Oscillation Analysis of Reactor Antineutrinos at KamLAND*, arXiv:1009.4771v2 [hep-ex], 2010

[54] URL *http://neutrino.kek.jp/*

[55] The K2K Collaboration: M. H. Ahn et al., *Measurement of Neutrino Oscillation by the K2K Experiment*, Physical Review D, 74, 072003, 2006

[56] URL *http://jnusrv01.kek.jp*

[57] Y. Oyama, *Results from K2K and status of T2K*, Nuclear Science and Safety in Europe 2006, pp. 113 – 124, 2006

[58] URL *http://www-numi.fnal.gov/*

[59] The MINOS Collaboration: D. G. Michael et al., *Observation of Muon Neutrino Disappearance with the MINOS Detectors in the NuMI Neutrino Beam*, Physical Review Letters, 97, 191801, 2006

[60] URL http://operaweb.lngs.infn.it/

[61] The OPERA collaboration: N. Agafonova et al., *Observation of a first* $\nu_\tau$ *candidate in the* OPERA *experiment in the CNGS beam*, Physics Letters B, 691, pp. 138 – 145, 2010

[62] T. Schwetz, M. Tórtola, J. W. F. Valle, *Three-flavour neutrino oscillation update*, New Journal of Physics, 10, 113011, 2008

[63] T. Schwetz, *Neutrino oscillations: Present status and outlook*, Pramana - Journal of Physics, 72, 1, pp. 119 – 129, 2009

[64] M. Mezzetto, T. Schwetz, $\Theta_{13}$*: phenomenology, present status and prospect*, Journal of Physics G: Nuclear and Particle Physics, 37, 103001, 2010

[65] The LSND Collaboration: A. Aguilar et al., *Evidence for neutrino oscillations from the observation of* $\bar{\nu}_e$ *appearance in a* $\bar{\nu}_\mu$ *beam*, Physical Review D, 64, 112007, 2001

[66] URL *http://www-boone.fnal.gov/*

[67] The MiniBooNE Collaboration: A. A. Aguilar-Arevalo et al., *Search for Electron Neutrino Appearance at the $\Delta m^2 \sim 1\,eV^2$ Scale*, Physical Review Letters, 98, 231801, 2007

[68] The MiniBooNE Collaboration: A. A. Aguilar-Arevalo et al., *Event Excess in the MiniBooNE Search for $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ Oscillations*, Physical Review Letters, 105, 181801, 2010

[69] The Antares Collaboration: J. A. Aguilar et al., *ANTARES, the $1^{st}$ operational Neutrino Telescope in the Mediterranean Sea*, in preparation

[70] F. Fehr, *Systematic studies, calibration, and software development for event reconstruction and data analysis using the ANTARES deep-sea neutrino telescope*, Ph.D. thesis, University Erlangen-Nuremberg, 2009

[71] The Antares Collaboration: P. Amram et al., *Sedimentation and fouling of optical surfaces at the ANTARES site*, Astroparticle Physics, 19, pp. 253 – 267, 2003

[72] S. Basa, *Photomultipliers activity inside the ANTARES project*, Nuclear Instruments and Methods in Physics Research A, 442, pp. 91 – 98, 2000

[73] The Antares Collaboration: P. Amram et al., *The ANTARES optical module*, Nuclear Instruments and Methods in Physics Research A, 484, pp. 369 – 383, 2002

[74] The Antares Collaboration: J. A. Aguilar et al., *Study of large hemispherical photomultiplier tubes for the ANTARES neutrino telescope*, Nuclear Instruments and Methods in Physics Research A, 555, pp. 132 – 141, 2005

[75] G. D. Hallewell, *Construction and operation of the ANTARES underwater neutrino telescope*, Nuclear Instruments and Methods in Physics Research A, 595, pp. 51 – 53, 2008

[76] M. Anghinolfi, H. Costantini, K. Fratini, D. Piombo, M. Taiuti, *New measurement of the angular acceptance of the Antares Optical Module*, Antares Internal Note, ANTARES-OPMO-2008-001, 2008

[77] M. Anghinolfi, H. Costantini, V. Kulikovsky, M. Taiuti, *GEANT4 simulation of the* Antares *optical module*, Antares Internal Note, ANTARES-OPMO-2009-003, 2009

[78] The Antares Collaboration: J. A. Aguilar et al., *Performance of the front-end electronics of the ANTARES neutrino telescope*, Nuclear Instruments and Methods in Physics Research A, 622, 1, pp. 59 – 73, 2010

[79] The Antares Collaboration: J. A. Aguilar et al., *The data acquisition system for the ANTARES neutrino telescope*, Nuclear Instruments and Methods in Physics Research A, 570, pp. 107 – 116, 2007

[80] URL *http://root.cern.ch/*

[81] J. Brunner, *Upgrade on K40 simulation*, ANTARES Internal Note, ANTARES-PHYS-2006-005, 2006

[82] The ANTARES Collaboration: J. A. Aguilar et al., *Transmission of light in deep sea water at the site of the ANTARES neutrino telescope*, Astroparticle Physcis, 23, pp. 131 – 155, 2005

[83] J. A. Aguilar, *Measurements of the attenuation length in ANTARES with the Optical Beacon system*, ANTARES Internal Note, ANTARES-SITE-2007-001, 2007

[84] A. Margiotta, private communication, May 2010

[85] J. Brunner, *The refraction index at the Antares site*, ANTARES Internal Note, ANTARES-SITE-2000-001, 2000

[86] S. Mangano, *Measurement of the group refractive index in the Antares site with the Optical Beacon system*, *ANTARES-PHYS-2010-007*, 2010

[87] M. de Jong, *The* ANTARES *Trigger Software*, ANTARES Internal Note, ANTARES-SOFT-2005-005, 2005

[88] S. Escoffier, *Performance of the T3 Triggers on MC Data*, ANTARES Internal Note, ANTARES-SOFT-2008-009, 2008

[89] T. Chiarusi, M. Spurio, *High-energy astrophysics with neutrino telescopes*, The European Physical Journal C - Particles and Fields, 65, 3 – 4, pp. 649 – 701, 2010

[90] Y. Becherini et al., *A parameterisation of single and multiple muons in deep water or ice*, Astroparticle Physics, 25, pp. 1 – 13, 2006

[91] G. Carminati et al., *Atmospheric MUons from PArametric formulas: A fast GEnerator for neutrino telescopes (MUPAGE)*, ANTARES Internal Note, ANTARES-PHYS-2006-003, 2006.

[92] G. Carminati et al., *MUPAGE User Guide*, ANTARES Internal Note, ANTARES-SOFT-2007-004, 2007

[93] C. Forti et al., *Simulation of atmospheric cascades and deep underground muons*, Physical Review D, 42, pp. 3668 – 3689, 1990

[94] E. Scapparone, *HEMAS: a Monte Carlo code for hadronic, electromagnetic and TeV muon components in air shower*, arXiv: 9902043v1 [physics.data-an], 1999

[95] G. Battistoni, M. Carboni, C. Forti, J. Ranft, *Release of a new version (V.07-1) of the HEMAS-DPM Monte Carlo: Description and user manual*, INFN-AE-99-07

[96] P. Antonilo, C. Ghetti, E. V. Korolkova, V. A. Kudryavtsev, G. Sartorelli, *A three-dimensional code for muon propagation through the rock: Music*, Astroparticle Physics, 7, pp. 357 – 368, 1997

[97] D. Bailey, *Genhen v5r1: Software Documentation*, ANTARES Internal Note, ANTARES-SOFT-2002-004, 2002

[98] A. Labbate, T. Montaruli, I. A. Sokalski, *GENHEN v6: ANTARES neutrino generator extension to all neutrino flavors and inclusion of propagation through the Earth*, ANTARES Internal Note, ANTARES-SOFT-2004-010, 2004

[99] G. Ingelman, A. Edin, J. Rathsman, *Lepto 6.5 - a monte carlo generator for deep inelastic lepton-nucleon scattering*, Computer Physics Communications, 101, pp. 108 – 134, 1997

[100] G. Barr, *The separation of signals and background in a nucleon decay experiment*, Ph.D. Thesis, Keble College, University of Oxford, 1987

[101] H. L. Lai et al., *Global QCD analysis and the CTEQ parton distributions*, Physical Review D, 51, p. 4763, 1995

[102] J. Pumplin, D. R. Stump, J. Huston, H.-L. Lai, P. Nadolsky, W.-K. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, Journal of High Energy Physics, 07:012, 2002

[103] T. Sjöstrand, *High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4*, Computer Physics Commuications, 82, p. 74, 1994

[104] S. Navas, L. Thompson, *KM3 User Guide and Reference Manual*, ANTARES Internal Note, ANTARES-SOFT-1999-011, 1999

[105] Application Software Group, *GEANT - Detector Description and Simulation Tool*, CERN Program Library Long Writeup, vol. W5013, 1993

[106] J. Brunner, *Geasim*, http://antares.in2p3.fr/internal/software/geasim.html, 2000

[107] T. Eberl, C. Kopper, *The SeaTray software framework*, ANTARES Internal Note, ANTARES-SOFT-2009-013, 2009

[108] C. Kopper, *A software framework for KM3NeT*, Nuclear Instruments and Methods in Physics Research A, 602, pp. 107 – 110, 2009

[109] C. Cârloganu, J. Carr, *On the Atmospheric Neutrino Oscillation Study in ANTARES*, ANTARES Internal Note, ANTARES-PHYS-1999-014, 1999

[110] C. Cârloganu, *Caractérisation des performances à basse énergie du futur télescope sous-marine à neutrinos ANTARES et leur application à l'étude des oscillations des neutrinos atmosphériques*, Ph.D. thesis, University of Marseille, 1999

[111] C. Cârloganu *Sensitivity of the ANTARES detector to atmospheric neutrino oscillations from the analysis of upgoing, contained muon neutrino events*, ANTARES Internal Note, ANTARES-PHYS-2001-009, 2001

[112] H. Motz, Ph.D. thesis, University Erlangen-Nuremberg, in preparation

[113] URL *http://www.nag.co.uk/*

[114] V. J. Stenger, *Track Fitting for the Dumand Octogon*, HDC-1-90, 1990

[115] J. Brunner, *The BBfit Reconstruction algorithm*, ANTARES Internal Note, ANTARES-SOFT-2009-003

[116] URL *http://www.phys.hawaii.edu/dmnd/dumand.html*

[117] F. Schöck: *Towards oscillation analysis: basic data-MC comparisons and first look with Posidonia*, Talk at the ANTARES collaboration meeting, CERN, 2010

[118] A. Heijboer: *Point source search: progress on data/mc agreement*, Talk at the ANTARES collaboration meeting, Gandia, 2009

[119] URL *http://tmva.sourceforge.net*

[120] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, *Toolkit for multivariate Data Analysis with* ROOT - *Users Guide*, arXiv:physics/0703039, 2009

[121] A. Heijboer: *High resolution point source search* , Talk at the ANTARES collaboration meeting, Clermont-Ferrand, 2010

[122] The MINOS Collaboration: P. Adamson et al., *Measurement of Neutrino Oscillations with the MINOS Detectors in the NuMI Beam*, Physical Review Letters, 101, 131802, 2008

[123] G. F. Feldman, R. D. Cousins, *A Unified Approach to the Classical Statistical Analysis of Small Signals*, Physical Review D, 57, pp. 3873 – 3889, 1998

[124] S. Baker, D. Cousins, *Clarification of the use of chi-square and likelihood functions in fits to histograms*, Nuclear Instruments and Methods in Physics Research, 221, pp. 437 – 442, 1984

[125] P. Creminelli, G. Signorelli, A. Strumia, *Frequentist analyses of solar neutrino data*, arXiv: hep-ph/0102234v6, 2005

[126] R. J. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*, John Wiley & Sons, 1989, ISBN 0-471-92295-1

# Danksagung

Viele Menschen haben zum Gelingen dieser Doktorarbeit beigetragen. An erster Stelle möchte ich mich bei Herrn Prof. Dr. Uli Katz für die Möglichkeit zur Promotion, die gute Unterstützung und die vielen hilfreichen Kommentare bedanken. Vielen Dank auch an Herrn Prof. Dr. Christian Stegmann für die Erstellung des Zweitgutachtens.

Ein besonders großes Dankeschön geht an Dr. Alexander Kappes für die unzähligen, aber stets hilfreichen Diskussionen, für viele aufmunternde Worte und für die große Unterstützung. Dr. Claudio Kopper möchte ich für seine enorme Hilfsbereitschaft, von der ich besonders profitiert habe, herzlich danken. Ich habe sehr viel von Claudio gelernt, und vieles wäre ohne seine Hilfe nicht möglich gewesen, oder hätte deutlich länger gedauert.

Bedanken möchte ich mich bei der gesamten ANTARES Kollaboration, und besonders bei der Erlanger ANTARES Gruppe, für die tolle Arbeitsatmosphäre und gute Zusammenarbeit. Persönlich erwähnen möchte ich Holger Motz für die freundliche Bereitstellung seiner Hit Selektion, Frau Prof. Dr. Gisela Anton und Dr. Felix Fehr. Besonders danken möchte ich Dr. Thomas Eberl für das Korrekturlesen der Arbeit, für das Erstellen diverser Computer-Skripte, sowie für sein erstaunliches Gedächtnis, ohne das oftmals umfangreiche Recherchen nötig gewesen wären. Ebenso möchte ich mich bei Dr. Kay Graf für tausende von Dingen bedanken, die ich gar nicht alle aufzählen kann. Kay steht stets als Ansprechpartner für Probleme aller Art zur Verfügung, insbesondere für organisatorische oder technische Fragen, und seine Geduld ist nahezu unerschöpflich.

Die Prozessierung der Daten und Simulationen wäre ohne das Computercluster des ECAP nicht denkbar gewesen. Ich möchte mich daher ganz herzlich bei allen Administratoren bedanken, die sehr viel Zeit investieren, damit andere gut arbeiten können. Besonders erwähnen möchte ich hierbei Max Neff, Ralf Auer, Klaus Geyer, Bernhard Glück, sowie nochmals Dr. Kay Graf und Dr. Claudio Kopper.

Diese Arbeit würde nicht in der jetzt vorliegenden Form existieren, ohne die Zusammenarbeit mit Dr. Cristina Cârloganu vom LPC in Clermont-Ferrand. Ich habe sehr viel von Cristina gelernt, nicht nur in physikalischer Hinsicht. Die Zusammenarbeit war wirklich sehr bereichernd und hat viel Spaß gemacht.

Mein Büro werde ich sicherlich sehr vermissen, was auch an meiner netten Bürokollegin Stefanie Wagner liegt, mit der man hervorragend physikalische, aber auch computertechnische und insbesondere menschliche und zwischenmenschliche Probleme besprechen kann.

Am meisten habe ich aber meinem Mann, Dr. Fabian Schöck, zu danken, der mich stets unterstützt und mir immer wieder Kraft gegeben hat, diesen nicht ganz einfachen Weg zu bestreiten.