ECAP-2013-023

Studies on the Selection of Neutrino-like Signals for the Acoustic Detection Test Device AMADEUS

Studien zur Selektion von neutrino-artigen Signalen für das Testsystem zur akustischen Detektion AMADEUS

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

vorgelegt von
Max Neff

aus Karlsruhe

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

Tag der mündlichen Prüfung: 14. Mai 2013
Vorsitzender der Promotionskommission: Prof. Dr. Johannes Barth
Erstberichterstatter: Prof. Dr. Ulrich F. Katz
Zweitberichterstatter: Prof. Dr. Sean Danaher

Contents

Chapter 1

Introduction

Astroparticle physics is an emerging field that combines astronomy and particle physics to extend the knowledge about the universe to the highest energies. The field unites three major branches: cosmic ray, gamma ray, and neutrino astronomy. The detection of *ultra-high-energy cosmic rays* (UHECRs), charged particles exceeding an energy of $5 \cdot 10^{19}$ eV, and the determination of their composition is starting to give an insight into the acceleration processes of hadrons in extragalactic sources. For charged particles, depending on their energy, deflection in intergalactic magnetic fields smears out the reconstructed position of sources. TeV gamma rays, which can be produced by hadron and lepton accelerating sources, only give a picture of the surface of the source, as the optical thickness of the acceleration region causes the absorption of these photons before they can escape. The observation of neutrinos with energies ranging from GeV to EeV can open another unique window to the understanding of particle accelerating sources. These neutrinos travel nearly unobstructed through the universe. Their small cross section allows them to escape dense objects and, as neutral and light particles, they are not deflected on their way to earth. Neutrinos are thus valuable messengers providing information about the acceleration region of their sources and the underlying mechanisms, as neutrinos would prove the acceleration of hadrons. Combined observations can complete the picture of the fundamental processes driving the non-thermal universe.
Neutrino telescopes currently in operation are based on the detection of Cherenkov light that is produced in a transparent medium by charged secondary particles of a neutrino interaction. Cherenkov light is emitted as a charged particle travels at a speed exceeding the phase speed of light in a given medium. This light can be detected by *photomultiplier tubes* (PMTs), sensors that are sensitive to single photons. Given the frequency spectrum of Cherenkov light, the frequency-dependent attenuation of light in water or ice yields a typical attenuation length of the order of 50 m. This length scale is one of the determining factors of the spacing of photosensors in a neutrino telescope. A huge volume of transparent medium has to be instrumented to detect neutrino interactions from the expected low flux of neutrinos from outside the solar system at a significant rate. Furthermore, the detector must be shielded against background signals from cosmic rays, which is attained by placing it in sufficient depth. Consequently, neutrino telescopes are only feasible in natural reservoirs of water or ice. ANTARES[1] is such

[1] *Astronomy with a Neutrino Telescope and Abyss Environmental Research* (ANTARES)

a neutrino telescope among others, most notable IceCube[2] and the future detector KM3NeT[3] based on the detection of Cherenkov light. For the detection of cosmogenic neutrinos in the *ultra-high-energy* (UHE) regime with energies exceeding 10^{17} eV, the size of Cherenkov detectors currently in operation or planned for the future presumably becomes insufficient to detect a significant number of events. A complementary approach is the use of acoustic pressure pulses produced by a particle cascade that evolves when a neutrino interacts in a medium like water or ice. According to the thermo-acoustic model developed by Askariyan [1], the fast deposition of energy by the cascade in the medium and the ensuing local heating process leads to a characteristic *bipolar pressure pulse* (BIP). Caused by the coherent emission over the cylindrical geometry of the cascade, the pressure wave propagates through the medium in a disk-like shape perpendicular to the main axis of the cascade. Given the expected low flux of neutrinos with energies in excess of 10^{17} eV, a potential acoustic neutrino telescope must have large dimensions of presumably $\gtrsim 100\,\mathrm{km}^3$. Acoustic detection has the advantage over detection of Cherenkov light that the attenuation length is of the order of 1 km for the peak spectral density of around 11 kHz [2]. Hence, the spatial configuration of an acoustic detector is not limited by the attenuation length.
To explore the feasibility of an acoustic neutrino detector, the AMADEUS[4] project was devised as part of the ANTARES neutrino telescope. Other notable currently active projects in this field are: the NEMO-OνDE[5] acoustic system also located in the Mediterranean Sea, the *South Pole Acoustic Test Setup* (SPATS) [3] as part of IceCube, and the prototype device for acoustic neutrino detection in Lake Baikal [4]. The AMADEUS systems consists of an array of a total of 36 acoustic sensors arranged in six clusters of sensors distributed over the ANTARES detector on different length scales. The main goal of the AMADEUS project is to measure the ambient noise level at the ANTARES detector site on a timescale of several years. The results of this measurement will allow an estimation of a lower limit of the signal-to-noise ratio for a detectable signal. Furthermore, the rate and correlation length of BIP-like background events, which imitate the acoustic signature of a UHE neutrino, and the corresponding sources are of interest. These studies are crucial to assess the feasibility of an acoustic neutrino telescope in the Mediterranean Sea.
The main objective of this work is to research possible ways to classify acoustic signals produced by a neutrino interaction in the presence of acoustic background. This requires a suitable simulation chain to reproduce the acoustic signature of a neutrino interaction, the transient background, and their propagation to the sensors within the detector. An ambient noise model for the Mediterranean Sea and a simulation of the data acquisition hardware used in the AMADEUS detector are implemented. A pre-selection scheme for neutrino-like signals based on matched filtering is employed, as it is used for on-line filtering of the AMADEUS data. A detailed description of the developed simulation chain will be given in Chap. 3. In Chap. 4, the analysis chain consisting of reconstruction techniques for the arrival time, the incident direction of the pressure wave, and the source position of an acoustic signal will be described. In addition, the signal classification strategy will be presented. It is based on machine learning algorithms, which were trained and tested with features extracted from simulated data.

[2] *IceCube South Pole Neutrino Observatory* `http://icecube.wisc.edu`

[3] *A multi-km*^{3} *sized Neutrino Telescope* `http://www.km3net.org`

[4] *ANTARES Modules for Acoustic Detection under the Sea* (AMADEUS)

[5] *Ocean noise Detection Experiment* (OνDE), a project at the site of the *Neutrino Mediterranean Observatory* (NEMO). `http://nemoweb.lns.infn.it/`

Furthermore, studies of an effective volume and a transient-free, limit-setting potential of the AMADEUS detector were performed with the simulation and analysis chain. Together with an analysis of transient signals in the experimental data taken with AMADEUS, this will be shown in Chap. 5. Finally, Chap. 6 concludes with a summary.
In the next chapter, the sources of UHE neutrinos and the principles of their detection will be discussed; the setup of the ANTARES and AMADEUS detectors will be described in more detail.

Chapter 2

Neutrinos and their Detection

Contents

The basic properties and interaction channels of neutrinos and their impact on neutrino detection are presented in this chapter. An overview of the possible sources of UHE neutrinos and the assumed underlying mechanisms to accelerate them to energies in the UHE regime are given. Furthermore, the principles of optical and acoustic detection of neutrinos are explained, as well as their implementation in the ANTARES optical Cherenkov neutrino telescope and the subsystem AMADEUS for acoustic detection.

2.1 Neutrino properties and interactions

Neutrinos are light neutral leptons with spin 1/2 (cf. Tab. 2.1). Experimentally proven are three neutrino flavours: electron-neutrino, muon-neutrino, and tau-neutrino. Neutrinos are created and detected with a well-defined flavour. As neutrinos have mass and the flavour eigenstates are not the same as their mass eigenstates, they are able to oscillate among the three flavours, while propagating through space. For three-flavour oscillations, the unitary transformation

$m < 2.2\,eV$	$m < 0.17\,MeV$	$m < 15.5\,MeV$
${}^{0}_{1/2}\nu_e$	${}^{0}_{1/2}\,\nu_\mu$	${}^{0}_{1/2}\,\nu_\tau$
electron-neutrino	muon-neutrino	tau-neutrino
$m = 0.511\,MeV$	$m = 105.7\,MeV$	$m = 1.777\,GeV$
${}^{-1}_{1/2}\,e$	${}^{-1}_{1/2}\,\mu$	${}^{-1}_{1/2}\,\tau$
electron	muon	tau

Table 2.1: Overview of the six flavours of leptons forming three generations and their basic properties: upper mass limit, charge, and spin [5].

relating the flavour and mass eigenstates can be written as:

$$|\nu_\alpha\rangle = \sum_i U_{\alpha i}\,|\nu_i\rangle\ , \tag{2.1}$$

where $|\nu_\alpha\rangle$ is a neutrino flavour eigenstate for a specific flavour α, $|\nu_i\rangle$ is a neutrino mass eigenstate with $i \in \{1,2,3\}$, and $U_{\alpha i}$ is an unitary transform matrix, known as the *Pontecorvo-Maki-Nakagawa-Sakata matrix* (PMNS matrix). This is a 3×3 matrix, if the three Standard Model neutrinos are considered. The propagation of the mass eigenstates $|\nu_i\rangle$ can be described by plane wave solutions of the form (in the following natural units are used):

$$|\nu_i(t)\rangle = e^{-i(E_i t - \mathbf{p}_i\mathbf{x})}|\nu_i(0)\rangle\,, \tag{2.2}$$

where E_i is the energy related to the mass eigenstate i, t is the time since the propagation started at $t_0 = 0$, $\mathbf{p}_i$ is the momentum vector, and $\mathbf{x}$ is the position vector of the particle relative to the source. In the ultra-relativistic limit, $E = |\mathbf{p}_i| = p_i \gg m_i$, the energy of the mass eigenstate can be approximated as:

$$E_i \approx E + \frac{m_i^2}{2E}\,. \tag{2.3}$$

Using $L = t$, where L is the distance traveled, the wave-function translates into:

$$|\nu_i(L)\rangle = e^{-i m_i^2 L/2E}|\nu_i(0)\rangle\,. \tag{2.4}$$

The time evolution of the mass eigenstates depends on the neutrino mass, and thus results in different propagation speeds. As the mass eigenstates are combinations of flavour eigenstates, this difference in speed causes interference between the corresponding flavour components of each mass eigenstate. A neutrino created with a given flavour changes its flavour during propagation caused by constructive interference. The probability that a neutrino originally of flavour α will later be observed as having flavour β is:

$$P_{\nu_\alpha\to\nu_\beta} = |\langle\nu_\beta(0)|\nu_\alpha(t)\rangle|^2 = \left|\sum_i U^*_{\alpha i}U_{\beta i}e^{-i m_i^2 L/2E}\right|^2 \tag{2.5}$$

$$\begin{aligned} P_{\nu_\alpha\to\nu_\beta} = \delta_{\alpha\beta} \quad &- \quad 4\sum_{i>j}\Re(U^*_{\alpha i}U_{\beta i}U_{\alpha j}U^*_{\beta j})\sin^2(\tfrac{\Delta m^2_{ij}L}{4E}) \\ &+ \quad 2\sum_{i>j}\Im(U^*_{\alpha i}U_{\beta i}U_{\alpha j}U^*_{\beta j})\sin(\tfrac{\Delta m^2_{ij}L}{2E})\,, \end{aligned} \tag{2.6}$$

where $\Delta m_{ij}^2 \equiv m_i^2 - m_j^2$. An oscillation length L_{ij} can be defined as the absolute value of:

$$L_{ij} \equiv \frac{4\pi E}{|\Delta m_{ij}^2|} \,. \tag{2.7}$$

The sources of neutrinos are assumed to be distributed over cosmological distances that are far greater than L_{ij}. The neutrinos from these sources are generated over a wide range of energies. So the sine term in the probability function (Eq. 2.6 real part) is averaged out:

$$\frac{1}{L_{ij}} \int_0^{L_{ij}} \sin^2(\frac{2\pi x}{2L_{ij}}) dx = \frac{1}{2} \,. \tag{2.8}$$

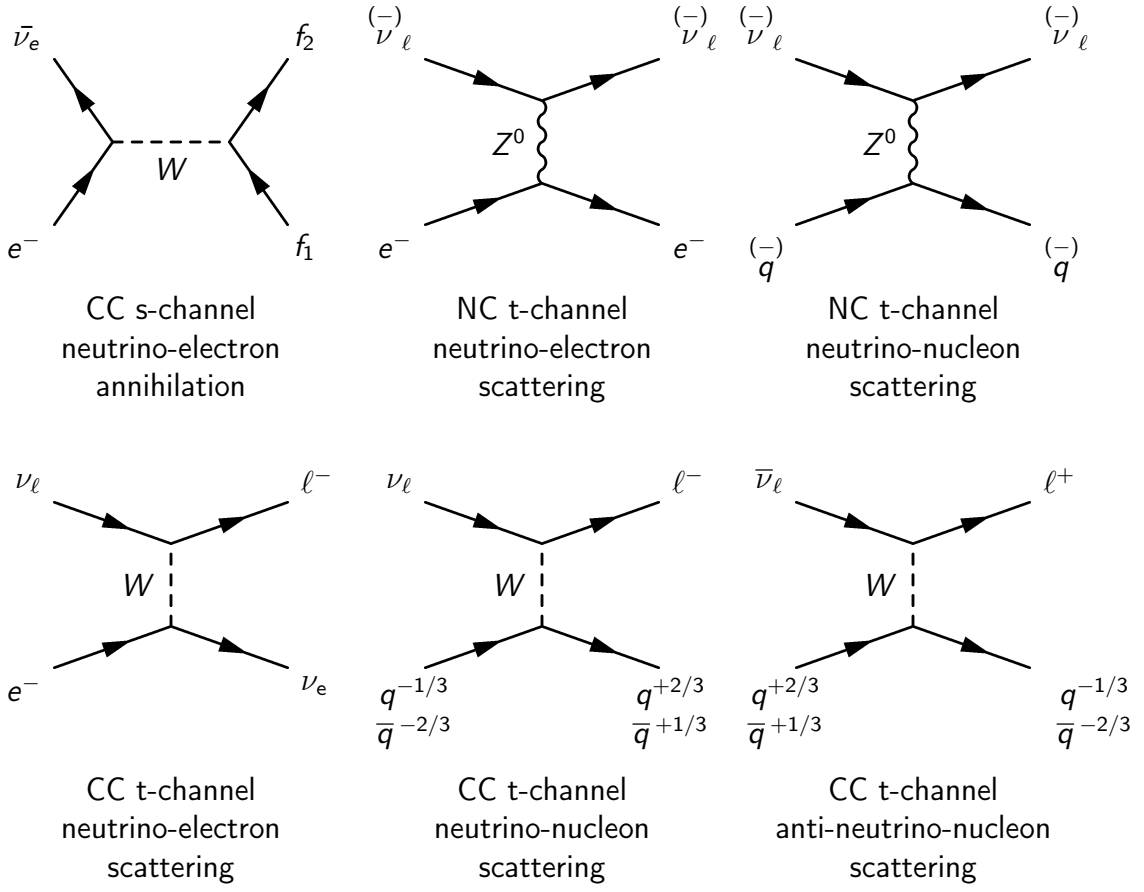
The relative numbers of neutrinos with a specific flavour $\alpha = e, \mu, \tau$ arriving on earth, $\tilde{N}_{\nu_\alpha}$, are calculated from the corresponding probabilities as:

$$\tilde{N}_{\nu_\alpha} = \sum_\beta P_{\nu_\beta \to \nu_\alpha} \,. \tag{2.9}$$

As neutrinos from cosmological sources are primarily produced via pion decay (cf. Sec. 2.2), the relative numbers of generated neutrino flavours are $\nu_e : \nu_\mu : \nu_\tau = 1 : 2 : 0$. At earth, these numbers are changed due to oscillation into $1 : 1 : 1$ and result in equal probabilities for measuring one of the three neutrino flavours [6].
The basic properties of neutrinos and neutrino oscillation having been discussed, the possible interactions of neutrinos within the Standard Model are presented. Neutrinos only interact via the weak force. So the cross section for neutrino interactions are low, allowing them not only to escape from dense sources, but also making them difficult to detect. The Feynman diagrams of the possible interactions [7] are given in Fig. 2.1. A neutrino or anti-neutrino of any flavour can interact with a shell electron or a nucleon of the target medium (e.g. water) by the exchange of the corresponding weak force vector boson (t-channel). For the *charged current* (CC) interaction, this is the corresponding $W^\pm$ vector boson and, for the *neutral current* (NC) interaction, a Z^0 is exchanged. The annihilation (s-channel) of an electron anti-neutrino with a shell electron is also possible. The contribution of this interaction channel can be generally neglected, except for the resonant formation of the intermediate vector boson W^- at an energy of $E_\nu = 6.3 \cdot 10^{15}\,\text{eV}$, where the cross section exceeds that of the t-channel cross sections by an order of magnitude. For the acoustic detection, this energy is below the detection threshold of about $E_\nu = 10^{18}\,\text{eV}$ (cf. Sec. 2.3.2). For ultra-high energies, the effect of the resonance is not significant.
The double differential cross section of the CC/NC interactions of neutrinos or anti-neutrinos with nucleons [7–9] can be calculated as:

$$\frac{\mathrm{d}^2\sigma_i(\overset{(-)}{\nu} N)}{\mathrm{d}x\mathrm{d}Q^2} = \frac{G_F^2 M_{V_i}^4}{2\pi x(Q^2 + M_{V_i}^2)^2}[Y_+ F_2^{\overset{(-)}{\nu} N,i}(x, Q^2) - y^2 F_L^{\overset{(-)}{\nu} N,i}(x, Q^2) \overset{(-)}{+} Y_- x F_3^{\overset{(-)}{\nu} N,i}(x, Q^2)], \tag{2.10}$$

where i represents the current: neutral or charged. Anti-neutrino interactions are indicated by (-), which also applies for the sign $\overset{(-)}{+}$. M_V is the mass of the corresponding vector boson, which is exchanged, and Q^2 is the four-momentum transfer. The Fermi constant is defined



CC s-channel neutrino-electron annihilation

NC t-channel neutrino-electron scattering

NC t-channel neutrino-nucleon scattering

CC t-channel neutrino-electron scattering

CC t-channel neutrino-nucleon scattering

CC t-channel anti-neutrino-nucleon scattering

Figure 2.1: The Born level Feynman diagrams for reactions of neutrinos (ν_ℓ) and anti-neutrinos ($\bar{\nu}_\ell$) with matter. Time axis is pointing to the right. ℓ denotes any charged lepton, $q(\bar{q})$ any quark (anti-quark), where in relevant cases the charge is given as superscript. For the CC neutrino annihilation diagram the outgoing fermion f_1 can be any of $\bar{\nu}_\ell, q^{+2/3}, \bar{q}^{+1/3}$ and the other fermion f_2 accordingly $\ell^-, q^{-1/3}, \bar{q}^{-2/3}$.

as: $G_F = \frac{\sqrt{2}}{8}\frac{g^2}{M_W^2}(\hbar c)^3$, where g is the weak force coupling constant and M_W is the mass of the W-boson. F_2, F_3, and F_L are the structure function in the *deep inelastic scattering* (DIS) regime. $Y_\pm$ is defined as $Y_\pm \equiv 1 \pm (1-y)^2$ and $x = \frac{Q^2}{2M(E-E')}$ is the fraction of the nucleon's momentum carried by the struck quark. Finally, $y = \frac{E-E'}{E}$ is the fraction of the lepton's energy loss in the nucleon rest frame, E and E' are the lepton's energy in the initial and the final state, respectively.
The predictions of neutrino cross sections in the UHE regime rely on extrapolation of experimental data of the structure functions over several orders of magnitude, thus, introducing sizeable uncertainties. In the context of UHE neutrino detection, these predictions are used for both the estimate of the opaqueness of the earth and the rate of detectable interactions for a given neutrino flux model and detector.
As the cross section increases with energy, the earth gets opaque for neutrinos in the UHE regime. The mean free path can be defined as:

$$\lambda = \frac{1}{n\sigma_{tot}(E_\nu)}, \tag{2.11}$$

where n is the number of target particles per unit volume and σ_{tot} is the total cross section of the relevant interactions. Furthermore, n can be expressed using the depth-dependent function of the density ρ of the earth: $n = \rho(r)/m_n$, where m_n is the mass of the neutron and r is the distance to the centre of the earth, leading to a depth-dependent mean free path $\lambda(r)$. The earth is approximated as a sphere, and the *Preliminary Reference Earth Model* (PREM) [10] is used to model the earth's density profile. Typically, the detector would be placed under water, so the earth is assumed to be covered with 2.5 km of water. The probability of a neutrino reaching an interaction vertex in 2 km water depth is a function of the incident angle ϕ and energy of the neutrino E_ν is defined as:

$$p(\phi, E_\nu) = e^{-\int 1/\lambda(r,E_\nu)ds} = e^{-d_{WE}(\phi)/\lambda_{water}(E_\nu)}\,. \tag{2.12}$$

The integral in the equation above is performed over the flight path of the neutrino. The distance that the neutrino travels through matter of varying densities can be expressed as the equivalent distance d_{WE} in water. Fig. 2.2 shows the resulting probability distribution for a neutrino reaching the detector as a function of the incident angle. For the upper hemisphere, this probability is about 1 for neutrinos with an energy of $10^6 - 10^{12}$ GeV. The probability decreases rapidly for incident angles below the horizon as the neutrino has to penetrate the earth. For acoustic detection, neutrino energies greater than 10^9 GeV are of significance, thus the upper hemisphere with a margin of 10° below the horizon is sufficient as solid angle accessible by a detector. The probability of a neutrino reaching the detector with an incident angle below 10° is less than 10^{-3} at 10^9 GeV.

2.2 Ultra-high-energy neutrino sources and production mechanisms

The sources of UHE neutrinos and their underlying production mechanisms are discussed in this section. Neutrinos can be produced be pp and $p\gamma$ interactions of UHE protons with nuclei

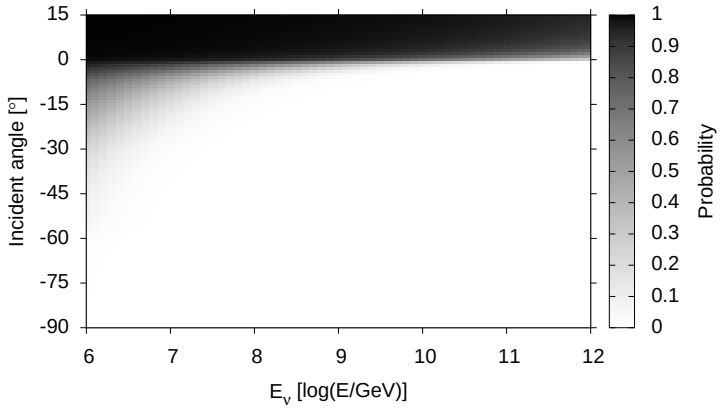


Figure 2.2: The probability distribution of a neutrino reaching a point in 2 km water depth as a function of the incident angle and the neutrino energy. An angle of 0° denotes the horizon, negative angles indicate that the neutrino is coming from below.

or photons in the ambient plasma:

$$p + p/\gamma \rightarrow N\pi + X\,, \tag{2.13}$$

where $N \geq 1$ is the number of pions produced, and X represents other final state particles. Also, neutrinos can be generated by the decay of exotic heavy particles and by dark matter particle annihilation. So based on their production mechanisms, the possible sources can be divided into two types: hadron accelerating sources and so-called top-down sources [11]. Most of the processes[1] have in common that the neutrinos are produced in the chain of pion decay:

$$\begin{aligned}
\pi^+ &\rightarrow \mu^+ + \nu_\mu \\
&\hookrightarrow \mathrm{e}^+ + \nu_e + \bar{\nu}_\mu \\
\pi^- &\rightarrow \mu^- + \bar{\nu}_\mu \\
&\hookrightarrow \mathrm{e}^- + \bar{\nu}_e + \nu_\mu \\
\pi^0 &\rightarrow \gamma + \gamma\,.
\end{aligned}$$

This mechanism leads to the composition of flavours produced in cosmological sources of 1 : 2 : 0 as mentioned before. The π^0 decay is a source of high-energy gamma rays. As TeV gamma ray sources can have hadronic or leptonic production mechanisms, the detection of UHE neutrinos would prove the hadronic production of photons in these sources.
The currently accepted theory for the acceleration of charged particles in astrophysical sources is that of Fermi acceleration [12–14]. In this stochastic process, a charged particle gains a constant fraction of energy, while elastically scattered by magnetic turbulences or irregularities moving with a characteristic velocity **u**. The particle is confined in the region of the inhomogeneous magnetic fields due to deflection until it possesses sufficient energy to escape. The maximum energy achieved in this process depends on the magnetic field and the size of the acceleration region [15]:

$$E_{max} = \beta_s ZBR\,, \tag{2.14}$$

[1] Exotic heavy particles can have decay channels that directly produce neutrinos.

where E_{max} is the maximum energy in units of 10^{18} eV, $\beta_s = u/c$ is the shock velocity, Z is the charge of the accelerated particle in units of the electron charge, B is the magnetic field in the acceleration region in μG, and R is the radius of this region in kpc.
A *gamma-ray burst* (GRB) [16, 17] is a short and rather directed emission of high-energy radiation. It releases $\approx 10^{54} - 10^{56}$ GeV on time-sales ranging from ≈ 30 ms to several minutes. This makes them to the most (electromagnetically) luminous objects in the universe. There are two main types of processes that produce GRBs: the core-collapse of a massive star resulting in a black hole or two compact massive objects, like two neutron stars, merge into a black hole (cf. Fig. 2.3). The two processes can be distinguished by the duration of the GRB, the first is of long-duration ($\gtrsim 2$ s) and the second one is of short duration ($\lesssim 2$ s). The high material inflow rates accompanying these events result in the formation of relativistic jets. It is assumed that velocity variations in the outflow of the jets produce internal relativistic shock fronts capable of accelerating electrons. These electrons lose energy via synchrotron radiation and produce the gamma ray burst observed. However, the gamma rays could also be produced by π^0-decay, which would indicate the acceleration of hadrons. The so-called afterglow of a GRB is a wide-spectrum emission following the burst. It is produced by the creation of a highly relativistic external shock front as the jet encounters the interstellar medium. The afterglow emission seems to originate from the same mechanism as the burst itself. Fermi acceleration processes in this turbulent source may also allow for an efficient acceleration of protons to ultra-high energies in excess of 10^{20} eV. The detection of neutrinos from GRBs would prove this hypothesis.
The next source discussed is among the most powerful known cosmological sources: the *active galatic nucleus* (AGN) [18, 19]. Their luminosity outperforms that of a typical galaxy by several orders of magnitude, while the emission comes only from the compact region in the centre of the host galaxy. A super-massive black hole forms an accretion disk, fuelled by the host galaxy, and thus ejects mass in jets perpendicular to the accretion disk (cf. Fig. 2.4). Fermi acceleration seems to describe the mechanism that can efficiently accelerate protons in the multiple shock fronts occurring in the AGN jets, especially in their inner parts (sub-parsec scale). Neutrinos are again produced via pion decay by the collision of accelerated protons with photons from the accretion disc and with photons produced in the jet. This model is compatible to the spectra observed in various wavelengths.
After discussing the production of UHE neutrinos at proton accelerating sources, the mechanisms that generate UHE neutrinos at the propagation through space is described. The prediction of the *Greisen-Zatsepin-Kuzmin* (GZK) cut-off [20, 21] in the UHECR spectrum caused by the interaction of these protons with the *cosmic microwave background* (CMB) implies the existence of a UHE neutrino flux. The protons can interact with the CMB resulting in a Δ-resonance, which further decays into a charged pion and a neutron:

$$p + \gamma_{CMB} \rightarrow \Delta^+(1232) \rightarrow n + \pi^+ \,. \tag{2.15}$$

The energy of CMB photons are distributed according to a Plank distribution for a black body radiation with an average temperature of 2.7 K. The energy required to produce the Δ-resonance, and thus for the pion production, is $E_{th}^{\Delta} \approx 6 \cdot 10^{19}$ eV. This production limits the distance from which protons with energies greater than E_{th}^{Δ} can reach the earth and results in the so-called GZK horizon of the order of 100 Mpc. The GZK neutrino flux seems to be "guaranteed", as production mechanisms for protons exceeding this energy threshold are known,

like that mentioned above. The observations made by the AUGER collaboration [22, 23] are compatible with the existence of the GZK cut-off, but uncertainties remain in the interpretation of the experimental results. Also the composition of UHECRs, see e.g. [24] and references therein, has a strong effect on the expected neutrino flux. If a substantial fraction of the cosmic ray primaries are heavy nuclei rather than protons, they would preferentially lose energy through photo-disintegration. So the corresponding neutrino flux may be substantially depleted. This makes GZK neutrinos one of the prime candidates for UHE neutrino astronomy.
Apart from hadron accelerating sources or the GZK cut-off, there are other possible scenarios to produce UHE neutrinos. One theory is the Z-Burst model [25,26]], in which it is assumed that it exists a *cosmic neutrino background* (CνB) comparable to the CMB originating from relic neutrinos, which decoupled from the remaining matter about 1 s after the Big Bang. This CνB is not yet experimentally proven to exist. This is complicated to obtain as these neutrinos have energies of only a few meV [27]. However, if it exists, a flux of UHE neutrinos from distant sources could interact via resonance with the CνB producing a Z-boson. On average, this decays into about one baryon-anti-baryon pair and pions, which further decay into neutrinos and photons. The primary UHE neutrino would require an energy of $E_\nu^{primary} \gtrsim M_Z/(2m_\nu^{relict})$. This energy is sensitive to the mass of the relic neutrinos m_ν^{relict}. If the momentum of the secondary particles produced by the Z-burst points in the direction of the earth and occurs within the GZK distance, then some of the photons and nucleons from that burst may initiate an air shower in the atmosphere of the earth exceeding the GZK energy limit. So the Z-burst model circumvents the energy limitation introduced by the GZK cutoff, but the derived neutrino fluxes, if normalised to the measured cosmic ray spectra, are excluded by experiments, see e.g. [28].
Another theory is the so-called top-down scenario [29], involving the decay or annihilation of exotic superheavy particle, often denoted as X-particles, with a typical mass of the order of the grand unified theory's energy scale $M_{GUT} \simeq 10^{24}$ eV. Candidates of these X-particles are: topological defects caused by symmetry-breaking in the early universe, which naturally occurs in the time-varying gravitational field of the expanding universe at the post-inflationary stage, or super-massive dark matter.

2.3 HE/UHE Neutrino detection principle

2.3.1 Optical detection principle

The optical detection of neutrinos via Cherenkov radiation, as used in neutrino telescopes, will be introduced. The typical energy range of neutrinos, in which neutrino telescopes operate, is $10-10^7$ GeV. As depicted in the sketch (Fig. 2.5), neutrinos interact with matter surrounding the detector producing secondary particles such as electrons, taus, and muons. The latter are favoured for the optical detection, because of their large mean free path and lifetime in this energy range. As the earth acts as a shield against all particles except neutrinos, the detection of upward-going muons as a signature of muon neutrino interactions can be used. The detection medium needs to be transparent and of sufficiently large homogenous volume to detect the expected low flux of cosmological neutrinos at a significant rate. So usually a natural reservoir of water or ice is used, where the muon emits Cherenkov light. Basically, Cherenkov light is emitted when a charged particle travels at a speed greater than the speed

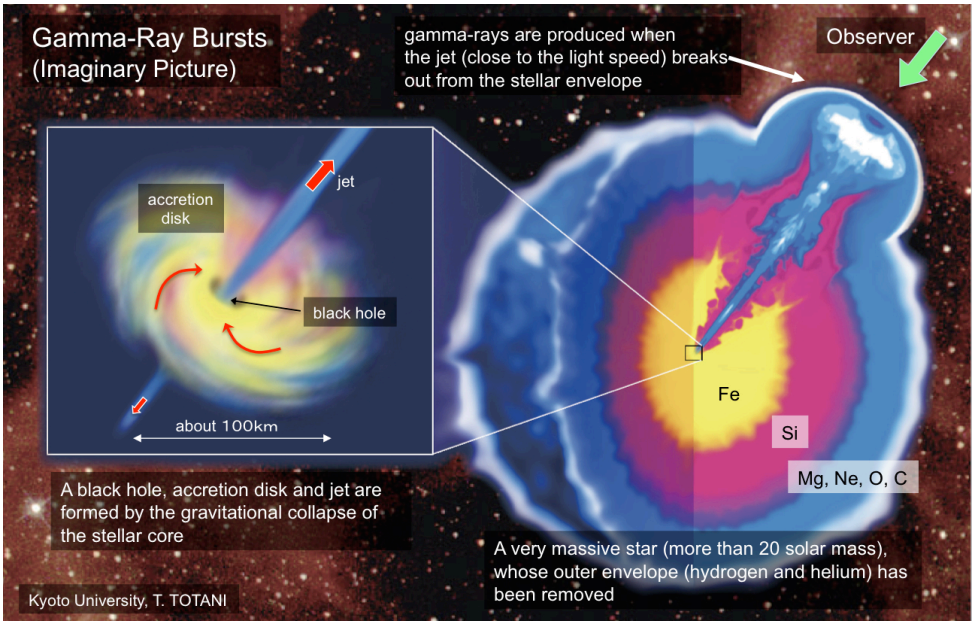


Figure 2.3: A Sketch showing one proposed mechanism resulting in a gamma ray burst: the core collapse of a massive star. Adapted from [30].

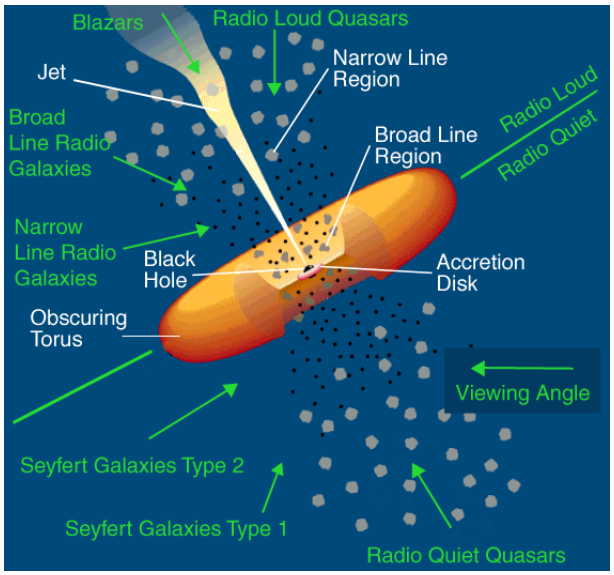


Figure 2.4: A sketch of the current paradigm for AGNs (not to scale). The central black hole is surrounded by a luminous accretion disk. Broad emission lines are produced in clouds (dark grey spots) orbiting above the disk and perhaps by the disk itself. A thick torus obscures the broad-line region from transverse lines of sight; some continuum and broad-line emission can be scattered into those lines of sight by hot electrons (black dots) that pervade the region. A hot corona above the accretion disk may produce the hard X-ray continuum. Narrow lines are produced in clouds (grey spots) much further from the central source. Adapted from [19].

of light in a given medium. The emission has a cone-like geometry that is characterised by the Cherenkov angle θ_C:

$$\cos(\theta_C) = \frac{1}{n\beta}, \tag{2.16}$$

where n is the refraction index and $\beta = v/c$ the velocity of the muon represented as the fraction of the speed of light in vacuum. For sea water, θ_C is typically about 42°. The direction of the muon trajectory corresponds to that of the neutrino within a few degrees at 10 GeV due to relativistic forward boost. The track of flight through the detector is reconstructed from the known position of the sensors and the arrival time of the light registered by the sensor. To detect the Cherenkov light, PMTs are used with optimised peak quantum-efficiency for blue light. Although the Cherenkov emission spectrum is continuous and peaked in the ultraviolet region, the attenuation of light in water is minimal for blue light.

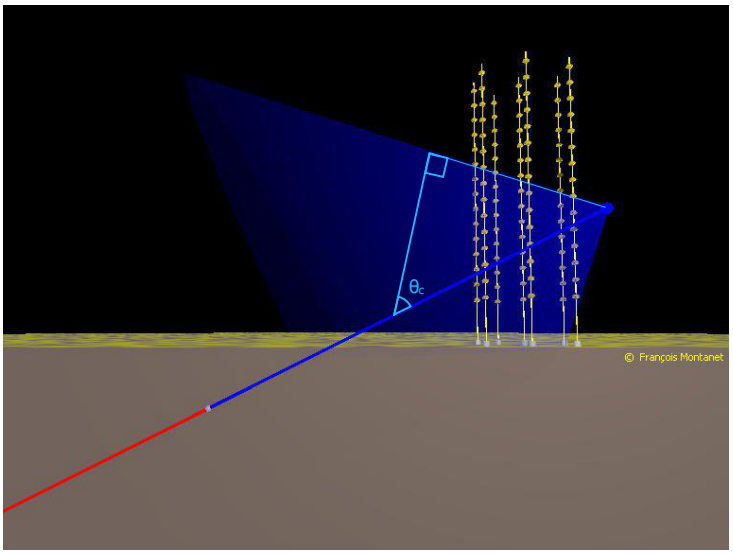


Figure 2.5: Sketch of a neutrino (red line) interacting in the vicinity of the ANTARES detector producing a muon (blue line) and the characteristic Cherenkov cone emitted by the muon. Adapted from [31].

2.3.2 Acoustic detection principle

As already mentioned, the acoustic particle detection is based on the thermo-acoustic model developed by Askariyan [1]. This model describes the generation of an acoustic pulse due to the local heating of the medium by the energy deposition of a particle cascade. Such a particle cascade evolves in the surrounding medium when a UHE neutrino interacts. The fast deposition of energy by the cascade in the medium and the ensuing local heating process leads to the characteristic *bipolar pressure pulse* (BIP). Due to the cylindrical geometry of the energy deposition volume, the coherent superposition of primary wavelets leads to a propagation through the medium in a disk-like shape perpendicular to the main axis of the cascade. This shape is often referred to as "pancake". The coherence length $L_c \approx \Delta f / c_s$ of the broadband pulses need to exceed the radial dimension of the heated volume to achieve coherence. The

thermo-acoustic model [1, 32] can be derived from the Euler Equation for a compressible fluid with the three coordinates $i = 1, 2, 3$:

$$\frac{\partial(\rho v_i)}{\partial t} = -\sum_{j=1}^{3} \frac{\partial \Pi_{ij}}{\partial x_j} \,, \tag{2.17}$$

with mass density ρ, velocity vector $\mathbf{v}$ of the fluid and momentum-density tensor including the pressure p:

$$\Pi_{ij} = p\delta_{ij} - \rho v_i v_j \,, \tag{2.18}$$

Taking the partial derivative $\partial/\partial x_i$ of component i of Eq. 2.17 and using the mass continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) = 0 \,, \tag{2.19}$$

a non-linear wave equation can be derived:

$$\frac{\partial^2 \rho}{\partial t^2} = \sum_{i,j=1}^{3} \frac{\partial^2 \Pi_{ij}}{\partial x_i \partial x_j} \,. \tag{2.20}$$

A separation in two spatial regions is used as an approach to solve Eg. 2.20: One region called I for "interaction", where the energy is deposited by the particle interactions with the medium and thus the wave excited. The hydrodynamic ("acoustic") region called A is the second one, where the acoustic wave propagates through the medium. This spatial split is applied to the momentum density tensor and is rewritten as:

$$\Pi_{ij} = \Pi_{ij}^{A} + \Pi_{ij}^{I}. \tag{2.21}$$

The change of density of the medium $d\rho$ can also be split into these two regions. $d\rho$ is composed of a dynamic part, resulting from the change of the pressure, and a part caused by heat transfer to or from the system. Under the assumption that the mole number N is constant, the entropy S and pressure p characterise the state and the total derivative of ρ can be separated in the following form:

$$\mathrm{d}\rho = \left.\frac{\partial \rho}{\partial p}\right|_{S} \mathrm{d}p + \left.\frac{\partial \rho}{\partial S}\right|_{p} \mathrm{d}S \tag{2.22}$$

In thermodynamics, the entropy S in a reversible system is related to the heat Q and the temperature T. C_p is the heat capacity at constant pressure and c_p is the specific heat capacity at constant pressure for a given mass m; the bulk volume expansion coefficient α for a Volume V and the adiabatic sound velocity are defined as:

$$\mathrm{d}S = \frac{\delta Q}{T} \,;\; C_p \equiv m\, c_p = \left. T \frac{\partial S}{\partial T}\right|_{p} \,;\; \alpha = \frac{1}{V}\frac{\partial V}{\partial T} \,;\; c_s = \sqrt{\frac{\partial p}{\partial \rho}}. \tag{2.23}$$

For a reversible transformation, the integral $\int \delta Q/T$ is independent of the path (see e.g. [33]). Hence, for a reversible isothermal expansion, the heat deposition can be expressed using the

exact differential dQ. Obviously, the process discussed is not isothermal, as the temperature change is eventually responsible for the pressure wave. However, the temperature change resulting from the interaction of a neutrino compared to the absolute temperature of the surrounding medium is negligible. So the change of density $d\rho$ as expressed in Eq. 2.22 can be reformulated in terms of thermodynamical equations as:

$$\mathrm{d}\rho = \frac{1}{c_s^2}\mathrm{d}p - \frac{\alpha}{c_p}\frac{\mathrm{d}Q}{V} \tag{2.24}$$

$$\Rightarrow \mathrm{d}p = \mathrm{d}p_A + \mathrm{d}p_I = c_s^2\mathrm{d}\rho + c_s^2\frac{\alpha}{c_p}\frac{\mathrm{d}Q}{V}\,, \tag{2.25}$$

which splits the pressure change into an acoustic part $\mathrm{d}p_A = c_s^2\mathrm{d}\rho$ and an interaction part $\mathrm{d}p_I = c_s^2\alpha/c_p\mathrm{d}\epsilon$, where $\epsilon \equiv Q/V$ is the energy deposition density. The quantities c_s, α, and c_p are assumed to be constant when performing integrations. Eq. 2.21 combined with Eq. 2.18 and the result of Eq. 2.24 can be rewritten as:

$$\Pi_{ij} = p_A\delta_{ij} + \Pi_{ij}^I = c_s^2\rho\delta_{ij} + \Pi_{ij}^I, \tag{2.26}$$

with the term $-\rho v_i v_j$ being neglected. The latter seems justified as v_i can be assumed to be small compared to c_s and constant or only slowly varying with x_i. As pressure changes will only be measured in the acoustic region, in the following $p \equiv p_A$ will be used to formulate the wave equation. Using Eq. 2.26, the wave equation (Eq. 2.20) translates for pressure differences $p' = p - p_0$ (p_0 being the static pressure) to:

$$\nabla^2 p' - \frac{1}{c_s^2}\frac{\partial^2 p'}{\partial t^2} = \frac{\partial^2\Pi_{ij}^I}{\partial x_i \partial x_j}\,.$$

The general solution for the wave in the acoustic region can be written as [34]:

$$\begin{aligned} p'(\mathbf{r},t) &= \frac{1}{4\pi}\int_V \mathrm{d}V' \frac{1}{|\mathbf{r}-\mathbf{r}'|}\frac{\partial^2\Pi_{ij}^I(\mathbf{r}',t')}{\partial x_i' \partial x_j'} \\ &\approx \frac{1}{4\pi c_s^2}\int_V \mathrm{d}V' \frac{\delta_{ij}\cdot\ddot{\Pi}_{ij}^I(\mathbf{r}',t')}{|\mathbf{r}-\mathbf{r}'|}\,, \end{aligned}$$

where for the last conversion, partial integration and the total derivative $\frac{\mathrm{d}}{\mathrm{d}x_i} = \frac{\partial}{\partial x_i} + \frac{1}{c_s}\frac{\partial}{\partial t}$ have been used repeatedly. Terms that are suppressed more than linear in $|\mathbf{r}-\mathbf{r}'|$ were neglected. In the affected terms, the expressions have to be evaluated at the retarded time t' which is given by $t' = t - |\mathbf{r}-\mathbf{r}'|/c_s$.
For the momentum density tensor in the interaction region, an isotropic energy deposition without momentum transfer to the medium through the term $-\rho v_i v_j$ in Eq. 2.18 is assumed:

$$\Pi_{ij}^I = p_I\delta_{ij} = c_s^2\frac{\alpha}{C_p}\epsilon\delta_{ij}\,.$$

Under these assumptions, a thermo-acoustic pressure wave that is generated by heating up the medium is described by:

$$p'(\mathbf{r},t) = \frac{1}{4\pi}\frac{\alpha}{c_p}\int_V \frac{\mathrm{d}V'}{|\mathbf{r}-\mathbf{r}'|}\,\frac{\partial^2}{\partial t^2}\,\epsilon(\mathbf{r}',t')\,. \tag{2.27}$$

The pressure field resulting from a particle interaction in a medium is thus determined by the spatial and temporal distribution of the energy deposition density ϵ. The amplitude of the resulting acoustic wave is governed by the thermodynamic properties c_s, C_p and α, which depend primarily on the temperature of the medium.
The thermo-acoustic model was proved by experiments, e.g., with proton and laser beams [35]. Simulations [36–38] showed that the frequency range of the generated sound waves spans from 1 to 100 kHz with a maximum spectral density around 11 kHz for a distance from the source of about 1 km. The calculated bipolar signal has a peak-to-peak amplitude of approximately 170 mPa per 10^{20} eV of cascade energy. The pulse shape and the distinctive geometry of the propagation are the main features to sought. The AMADEUS detector uses an array of clusters of sensors to find the typical signature of a neutrino-induced pressure wave. The sensors and read-out electronics are tuned to record these pulses. The design of the ANTARES and AMADEUS detector and their components are described in the following sections.

2.4 The ANTARES neutrino telescope

The ANTARES neutrino telescope [39] is located in the Mediterranean Sea, 40 km off the coast of Toulon, southern France in a water depth of about 2500 m. It was completed in May 2008 and has operating since this time. The detector is designed to detect the Cherenkov light emitted from charged secondary particles of a neutrino interaction. It consists of 12 vertical structures (*detection lines* [2]) to detect the Cherenkov light. Each line is fixed to the seabed by an anchor and is held vertically by a buoy on the top. Fig. 2.6 shows a sketch of the complete ANTARES detector including the acoustic setup described in Sec. 2.5. The 12 detection lines are placed on the seabed in an octagonal shape with four lines in a quadratical layout in the centre. The distance between two neighbouring lines ranges from 60 to 80 m covering a total area of $\approx 180 \times 180\,\mathrm{m}^2$ on the seabed. Each detection line has a total height of 480 m and comprises 25 so-called storeys arrayed at distances of 14.5 m from each other, starting 100 m above the seabed. An *Optical Module* (OM) [40] houses one 10" photomultiplier tube inside a pressure resistant glass sphere with outer diameter of 43 cm as photo-sensitive sensor for the detection of Cherenkov light. Each standard storey consists of a support frame that holds three 45°-downward-looking OMs and a titanium container with the required electronics (*Local Control Module* (LCM)) for the purposes of data acquisition, control, and monitoring. An extra line — the *Instrumentation Line* (IL07) — is equipped with sensors to monitor environmental parameters. The detector is connected to the on-shore control room via an electro-optical deep sea cable providing electrical power and data transmission.
As the detection lines are anchored at the bottom of the lines, they can swing and rotate due to the force applied by the undersea currents. To determine the positions of the storey with a precision of about 20 cm — as required to achieve the specified pointing precision of reconstructed muon tracks in the sub-degree range — the detector is equipped with an acoustic positioning system [41]. The system consists of acoustic transceivers at the anchor of each line. Along each detection line, five positioning hydrophones receive the signals of the emitters. By performing multiple time delay measurements and using these to trilaterate the individual hydrophones, the line shapes can be reconstructed relative to the positions of the

[2] The expressions "line" and later in the text "storey" are used according to the ANTARES terminology.

emitters. Currently, the sequence of signal emissions required for the positioning is emitted every two minutes. The signal emitted is a sine wave of particular frequency in the range of 40 – 60 kHz and the duration of the emission is 2 ms. For the positioning, the signal of the corresponding line is identified by the know emission time and the frequency.

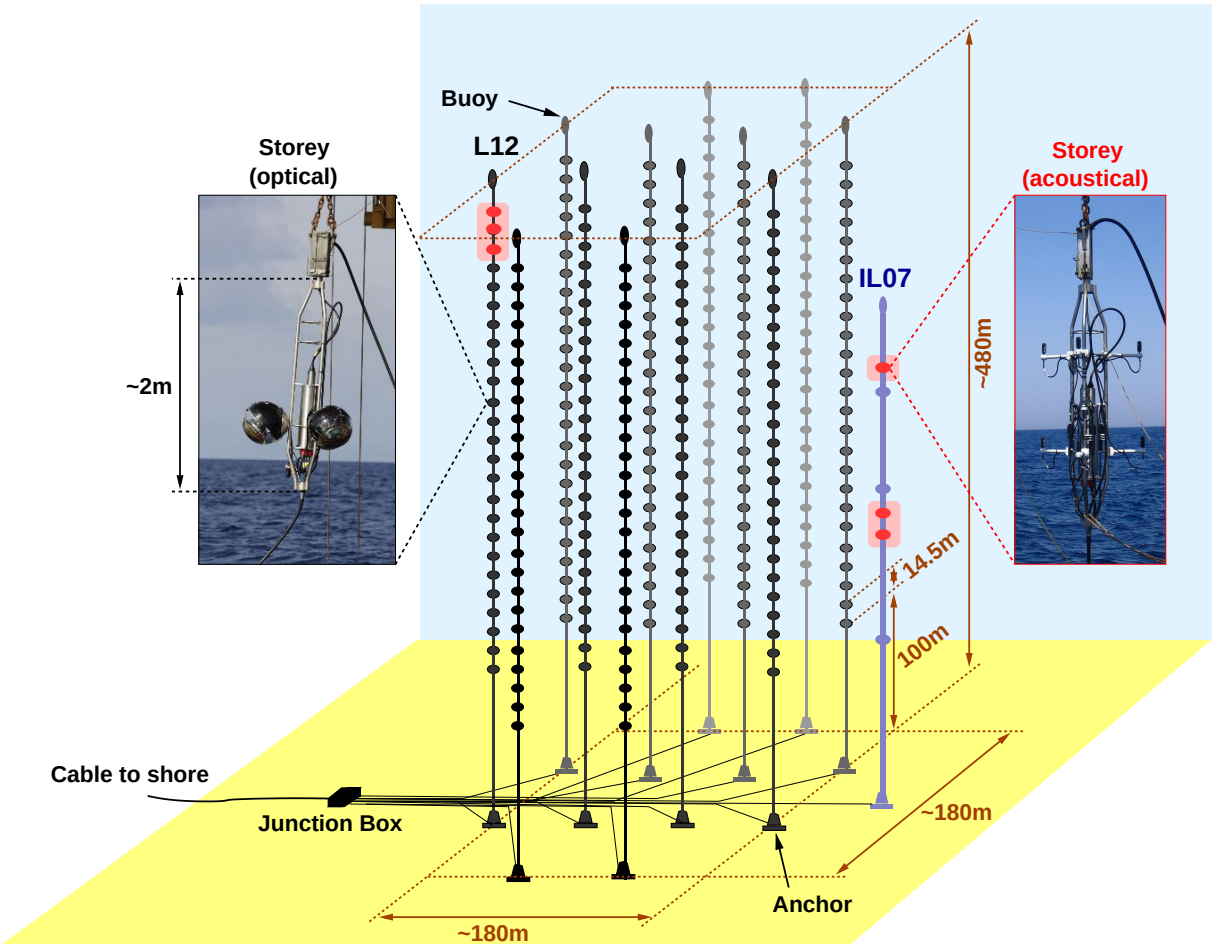


Figure 2.6: Sketch of the ANTARES detector in its design configuration with the so-called acoustic storeys on Line 12 and on the Instrumentation Line (IL07) highlighted. For further description see the text.

2.5 The AMADEUS detector

In contrast to neutrino telescopes currently in operation, acoustic neutrino detection uses the effect that UHE neutrinos interacting in a homogenous medium cause a detectable acoustic pulse according to the thermo-acoustic model [1]. The AMADEUS detector [42] was designed to prove this principle and to conduct a feasibility study in a realistic setup as it can be used for a future large-scale detector. The detector is integrated into the ANTARES neutrino telescope. As a subsystem, the AMADEUS detector shares the infrastructure with the ANTARES detector. The main goal of the AMADEUS project is to study the feasibility of an acoustic neutrino telescope. This includes the following studies as part of the research:

- The development and testing of hardware components, including sensors and read-out electronics suitable for future detection systems.
- Measurement of the ambient noise level at the ANTARES site on a timescale of several years. These results will allow for an estimate of the lower threshold of the signal-to-noise ratio of a detectable signal.
- The rate and correlation length of transient background signals that can mimic the acoustic signature of a neutrino interaction and their corresponding sources are researched.
- The development of a simulation and analysis chain with a modular architecture that is highly flexible and makes it easy to adapt to different detector geometries, environmental conditions or hardware used.

The last two points are discussed in this work.
In this section, the experimental setup of the AMADEUS detector with its components are described. As the geometry of the detector and its hardware components (sensors and read-out electronics) are integrated in the simulation, here an overview of the technical aspects is given, while the detailed description of the implementation in the simulation is given in the corresponding Sec. 3.3.

2.5.1 Experimental setup

The components of the acoustic detection setup are designed to perform measurements in an arrangement of acoustic sensors allowing for coincidence studies at different length scales. In total, six storeys of the ANTARES detector are equipped with six sensors (cf. Sec. 2.5.2) each representing clusters of sensors (*acoustic storeys*): three acoustic storeys are located at the top of Line 12 and on the IL07, respectively (cf. Fig. 2.6). These lines are separated by a horizontal distance of 240 m. The vertical distance between the acoustic storeys on Line 12 is the standard distance of 15 m and the lowest storey is 390 m above the sea floor. The IL07 comprises a total of six non-standard storeys with distances of 80 m between Storey 1 and 2 as well as Storey 4 and 5. The Storeys 2, 3, and 6 are acoustic storeys, which therefore have vertical interspaces of 15 m and 110 m, respectively. So, the storeys are 180 m, 195 m, and 305 m above the sea floor. The maximum distance of two acoustic storey is 340 m [42]. This setup realises three different distance scales: sensors within a storey have spacings of about 1 m, adjacent storeys of about 15 m, and the distances are from about 100 m to 340 m for storeys on the IL07 and between storeys on different lines. The setup of acoustic storeys within the ANTARES neutrino telescope and the different length scales are shown in Fig. 2.7.
As for the ANTARES detector, the relative positions of the acoustic storeys within the detector have to be continuously monitored. The emitter signals of the ANTARES acoustic positioning system, as described in Sec. 2.4, are used. The time difference between the known emission time and reception of the signal can be calculated. The positions of the AMADEUS sensors can be reconstructed using the signals from multiple emitters and their known positions at the anchors of the lines. The positioning accuracy for each hydrophone shows statistical uncertainties of a few millimetres for the hydrophones. The final measurement is expected to be dominated by systematic uncertainties due to the actual physical size of the receiving piezoelectric sensor, their relative positions within the storey, and the speed of sound in sea water. For the *Acoustic*

Modules (AMs) (cf. Sec. 2.5.2), the position reconstruction is less precise and statistical and systematic uncertainties are expected to be of the same order of magnitude [42]. In the next section, the acoustic sensors used in the AMADEUS detector and the configurations, in which the sensors are mounted, will described.

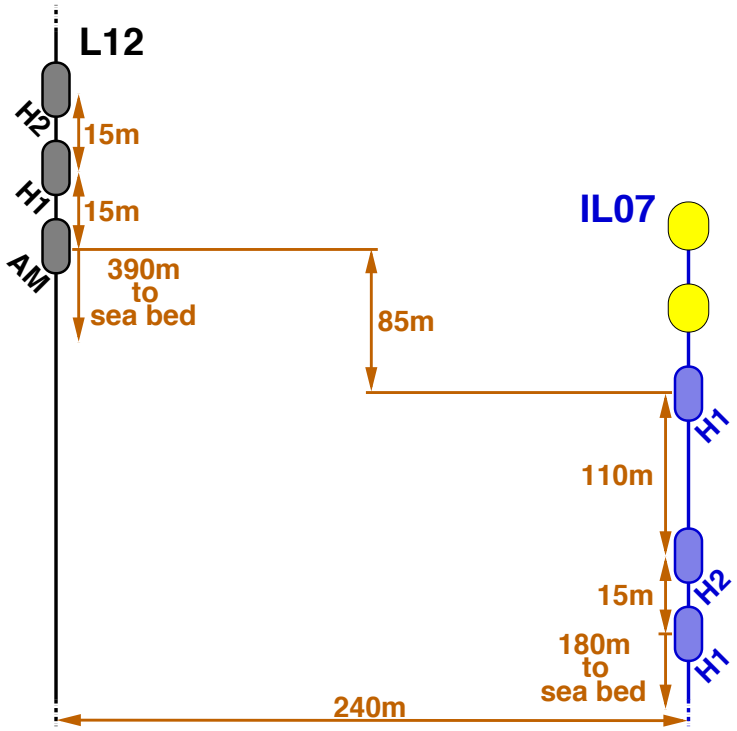


Figure 2.7: A sketch of the different length scales realised in the acoustic setup. H1 indicates storeys equipped with custom-designed hydrophones, H2 these with commercial ones and AM refers to Acoustic Modules (see text).

2.5.2 Acoustic sensors

Two types of acoustic sensors are used in the AMADEUS setup [42]: hydrophones and *Acoustic Modules* (AMs) (cf. Fig. 2.8). The three acoustic storeys on the IL07 are holding hydrophones, on Line 12 one of the acoustic storeys is equipped with AMs, and one has downward pointing hydrophones in contrast to the other storeys holding upward pointing ones. Both types of sensors are based on piezoelectric ceramics that convert pressure waves into voltage signals, which are then amplified for read out [43]. To protect the sensor in the deep sea environment, the ceramics and amplifiers are moulded in polymer plastics in the case of the hydrophones. For the AMs, they are glued to the inside of glass spheres identical to those used for the ANTARES OM (cf. Fig. 2.9). For future applications, this design allows for the integration of acoustic sensors and optical sensors, together, into the same pressure housing, thereby avoiding the need for additional mechanical structures [44]. The hydrophones used in the AMADEUS detector are either from a commercial supplier (*HighTech Inc. hydrophones* (HTIs)[3]) or developed and produced at the *Erlangen Centre for Astroparticle Physics* (ECAP), dubbed *custom-designed*

[3] `http://www.hightechincusa.com/`

hydrophones (LTIs). As the frequency range of the expected acoustic signal from a neutrino interaction spans from 1 to 100 kHz with a peak spectral density around 11 kHz, the sensors are tuned to be sensitive (around -145 dB re. 1V/µPa) [45, 46] in this range and provide a low inherent noise level. The sensitivity of one of HTI and LTI hydrophones is shown in Fig. 2.10 together with a scheme of the a LTI hydrophone.

(a) Acoustic storey with hydrophones pointing up (standard)

(b) Acoustic storey with hydrophones pointing down

(c) Acoustic storey with Acoustic Modules

Figure 2.8: Photographs of the three different storeys of the AMADEUS system [42] during their deployment: (a) A standard acoustic storey, equipped with hydrophones pointing up; (b) the central acoustic storey on Line 12 with the hydrophones pointing down; (c) the lowermost acoustic storey on Line 12 equipped with Acoustic Modules.

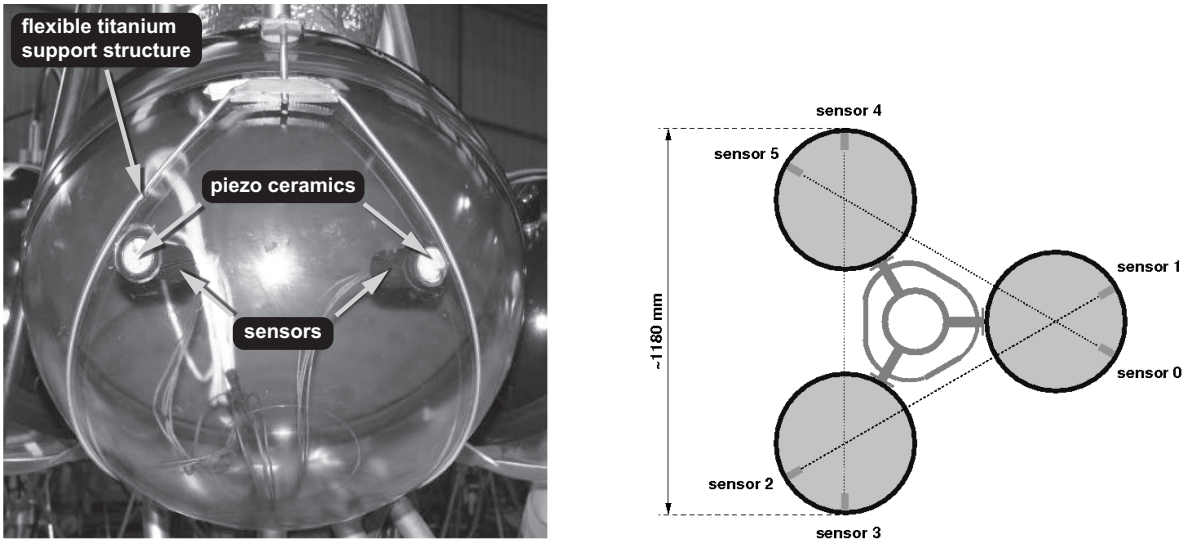
2.5.3 Acoustic data acquisition

The data acquisition of the AMADEUS detector [42] consists of one part off-shore and another on-shore. Off-shore, the digitisation and the data transfer are handled within the LCM on each storey by custom-designed digitisation boards (AcouADC-board[4]) and the standard ANTARES *data acquisition* (DAQ)[5]. Figure 2.11(a) shows a fully equipped LCM of an acoustic storey. From the left to the right, the following boards are installed: a *Compass board*, which measures the tilt and the orientation of the storey; three AcouADC-boards[6] (cf. Fig. 2.11(b)); a DAQ-board that transmits the data to shore; and a *Clock board* that provides the timing signals

[4] *acoustic analog-to-digital converter board* (AcouADC)

[5] According to the ANTARES "all data to shore" strategy [47], a completely unfiltered data stream is transmitted to shore.

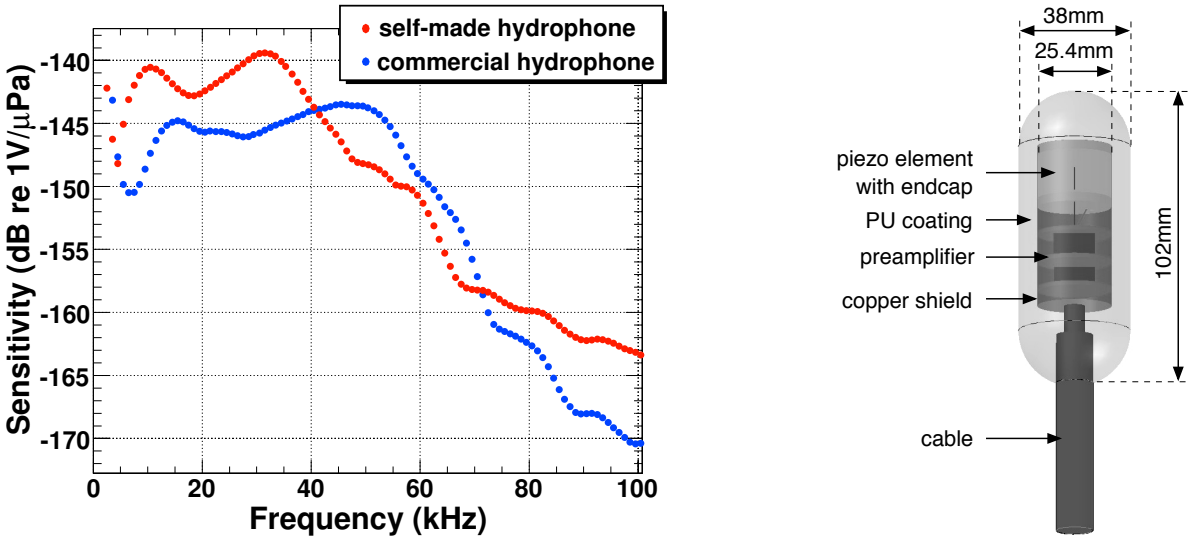
[6] Each board is designed to process the differential signals from two acoustic sensors, which results in a total of three such boards per storey.



(a) Acoustic Module (AM)

(b) Horizontal cross section

Figure 2.9: (a) Photograph of an Acoustic Module (AM) before deployment; (b) horizontal cross section of the acoustic storey holding Acoustic Modules in the plane of the sensors. The dotted lines are collinear with the longitudinal axis of the sensors and indicate the arrangement of the sensor within the storey. The lines intersect at angles of 60° at the centres of the glass spheres. Adapted from [42].

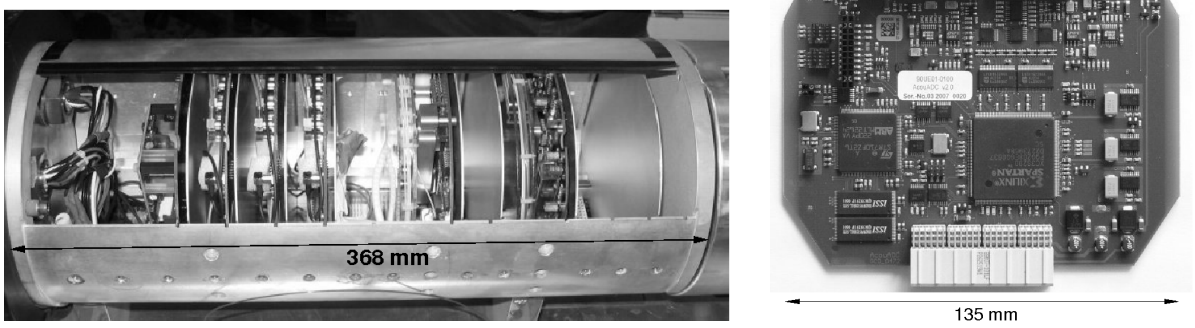


(a) Hydrophone sensitivity

(b) LTI scheme

Figure 2.10: In (a) The sensitivity in the horizontal direction, which is the result of a calibration measurement [45, 46], for one of the commercial and one of self-made hydrophones is shown. In (b), a scheme of a LTI hydrophone is given. Adapted from [42]

to correlate measurements performed in different storeys. The ANTARES neutrino telescope provides the capability of a 50 nanosecond timing resolution[7] and a simultaneous data transmission of up to 30 MBits/s per storey, which is suited for the acquisition of acoustic data. On-shore, a dedicated computer-cluster is used to process and store the acoustic data arriving from the storeys and to control the off-shore DAQ [48]. On this cluster different data filtering schemes are implemented. Transient signals are selected by a variable threshold, which is self-adjusting to the changing conditions of the deep sea. For BIPs a pattern recognition based on cross-correlating the output of the sensors with a pre-defined bipolar pulse is used. To study the ambient noise in the deep sea a certain amount of unfiltered data is stored as well. In Sec. 3.4, the on-line filter system is described in more detail, as well as its implementation in the simulation.



(a) Local Control Module (LCM) (b) AcouADC-board

Figure 2.11: (a) A LCM during assembly, equipped with three AcouADC-boards, before insertion into the titanium container. The sockets for external connection to the sensors (not visible in this picture) are attached to the lid of the container on the left-hand side of the photograph. From left to right, the following boards are installed: a Compass board; three AcouADC-boards; a DAQ-board; a Clock board. (b) Photograph of an AcouADC-board. The four connectors for the two differential input signals are located at the top, the analogue signal processing electronics is covered by metal shields. Adapted from [42]

.

AcouADC-board

The AcouADC-board is shown in Fig. 2.11(b) and consists of an analogue and a digital part [49]. The analogue part amplifies the voltage signals from the acoustic sensors by adjustable factors between 1 and 562 and applies a band frequency filter to the resulting signal. The digital part digitises and further processes the acoustic data. The digitisation is performed by an *analog-to-digital converter* (ADC) with a 16-bit resolution and a sampling rate of 500 thousand samples per second (kSPS). To have a flexible setup, a *field-programmable gate array* (FPGA) as

[7] In fact, the ANTARES detector is capable of providing sub-nanosecond precision for the synchronisation of the optical data recorded by the PMTs. This precision, however, is not required for the acoustic data.

data processor and a *micro-controller* (μC) are employed. Within the FPGA, which reads the digitised acoustic data from the ADC, the data can be further processed, e.g., being down-sampled to reduce data traffic, and formatted for transfer to the DAQ-board, which handles the transmission to the on-shore control room. As a standard setting, a downsampling of a factor of two (250 kSPS) is used, which is sufficient to study the background noise up to 100 kHz without exceeding the bandwidth limit for a continuous and simultaneous transmission of the data of all sensors to shore. The μC can be controlled from on-shore and is used to adjust settings of the analogue part and the data processing. Furthermore, the μC can be used to update the firmware of the FPGA in situ.

System response

In the deep sea, transient signals exhibit a huge variety of signal shapes, and the background noise covers a broad frequency range. Therefore, a precise understanding of the data-acquisition hardware concerning its frequency behaviour and the resulting signal distortion on the path from the sensors to the digitisation is essential. By calibrating the complete data taking chain — deducing the transfer function of the system — it is possible to reconstruct the real acoustic signal from the recorded one with high precision within the sensitive frequency range of the setup. Fig. 2.12 shows the measured filter characteristics of the analogue part of an AcouADC-board for a gain of 10 in amplitude (20 dB) and different settings of downsampling. The parameterised transfer function is also shown. This filter suppresses frequencies below about 4 kHz and above $\approx$ 130 kHz. The high-pass part cuts into the trailing edge of the low frequency noise of the deep sea acoustic background [50] and thus protects the system from saturation. The low-pass part efficiently suppresses frequencies above the Nyquist frequency of 250 kHz for the digitisation frequency of 500 kHz to eliminate aliasing effects. The inherent noise of the AcouADC-board is about -150 dB re. 1 V^2/Hz and the dynamic range achieved is from about 1 mPa to 10 Pa (including the adjustable gain) in RMS over the frequency range from 1 to 100 kHz. This allows studies of the acoustic background in the deep sea and of BIPs from noise sources imitating neutrino signatures under essentially all prevailing conditions [50].

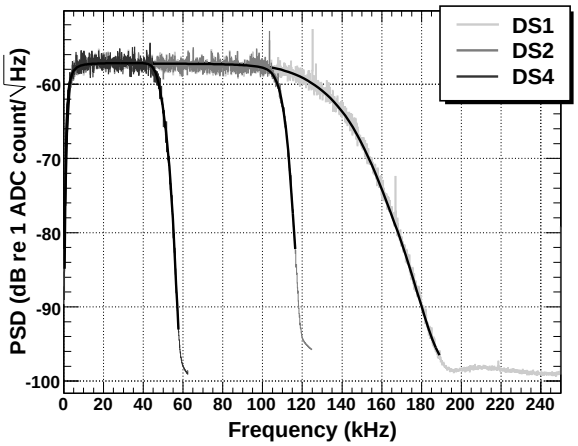


Figure 2.12: Measured filter characteristics of an AcouADC-board in the frequency range from 1 to 250 kHz for a downsampling of a factor of one (DS1), two (DS2) and four (DS4). The solid black lines correspond to the calculated transfer functions [49].

Chapter 3

Simulation Chain

Contents

In this chapter, the simulation chain that was developed is presented. The simulation covers all aspects required for acoustic neutrino detection — from the generation of the UHE neutrino-induced *bipolar pressure pulse* (BIP) and its propagation to the sensors within the detector to the acoustic environment of the deep sea. This includes transient and ambient noise models for the Mediterranean Sea, as derived from AMADEUS data. The detector geometry[1] and hardware components used in the simulation chain are equivalent to those of the AMADEUS detector. The simulation chain is designed within the SeaTray/IceTray software

[1] The known position and orientation of the sensors within the detector are used as fixed parameters. A simulation of the movement of the lines due to sea current is not yet implemented.

framework [51–53]. Its modular architecture is highly flexible and easy to adapt to different environmental conditions, data acquisition hardware, or detector geometries. The simulation chain consists of a number of modules (cf. Fig 3.1), which build on each other to create a simulated event that corresponds to the output of the detector. The tasks of the modules are the following:

1. The parameters of the incident neutrino — its energy, flight direction and the position of the interaction vertex — are chosen. The distributions that the parameters follow and their ranges are user-defined within the limits of the simulation.

2. The energy deposition of the UHE neutrino-induced hadronic shower in water is reproduced at the position of the interaction vertex.

3. The formation and propagation of the acoustic signal generated by a UHE neutrino interaction is calculated.

4. The reproduction of the deep-sea environment at the AMADEUS site, the ambient and transient noise conditions, is performed.

5. The DAQ-hardware — the system response and inherent noise of the acoustic sensors and corresponding read-out electronics — is simulated.

6. The simulation of the on-line filter system used in the AMADEUS detector for data reduction and waveform pre-selection is applied.

In the following sections, the design of the simulation chain and the capabilities of the different modules are described in detail.

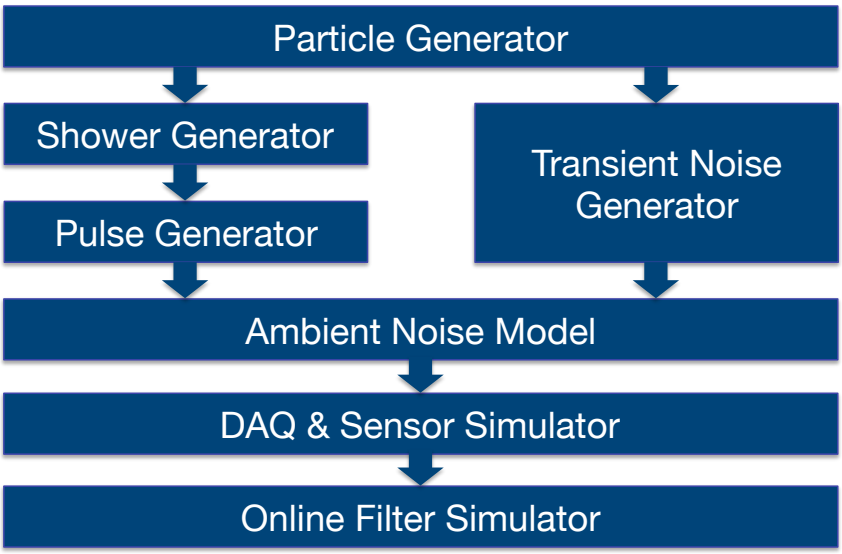


Figure 3.1: Overview of the different modules of the simulation chain.

3.1 Neutrino-induced acoustic pulse simulation

The first major step of the simulation chain is the calculation of the acoustic pulse from a hadronic cascade produced by a UHE neutrino interaction in water. Following the thermo-acoustic model [1] (cf. Sec. 2.3.2), an acoustic pulse is generated by the energy deposition of a particle shower in the surrounding medium. The fast local heating caused by this energy deposition leads to a BIP. Due to the coherent emission from the cylindrical geometry of the shower the wave propagates through the medium in a disk-like shape perpendicular to the main axis of the cascade. Firstly, the simulation of the energy deposition by the cascade is described and subsequently the calculation of the resulting acoustic pulse.

3.1.1 Calculation of the total shower energy

The total amount of the shower energy E_{sh} that is deposited by a neutrino-induced cascade is calculated from the neutrino energy E_ν using the energy transfer represented by the kinematic variable y:

$$E_{sh} = y \cdot E_\nu \tag{3.1}$$

The distribution of the energy transfer y depends on the neutrino cross section, which is a function of the energy, the flavour, and interaction channel, as described in Sec. 2.1. A parametrisation of the y-distribution derived in the work of Connolly et. al. [8] is used in this simulation to determine the total energy deposited by the cascade (cf. Fig. 3.2). It is assumed that the ratio of neutrinos and anti-neutrinos from cosmological sources of any flavour is 1 : 1. The proportion of cross sections from NC or CC interactions is about $\sigma_{NC}/\sigma_{CC} = 0.415$, according to values of the cross sections taken from [54]. This is used to determine the current of the neutrino interaction, as this is needed for the calculation of y from the parametrisation. The energy range of the incoming neutrinos is set from 10^9 GeV, where a neutrino-induced interaction starts to produce an acoustically detectable signal, to 10^{12} GeV, to which the parametrisation is well-defined. A uniform distribution of the neutrino energy is used to maintain a sufficient rate at high energies. For shower energies smaller than 10^6 GeV, the energy is set to 10^6 GeV to avoid exceeding the valid energy range of the total energy deposited by the cascade described in the next section.

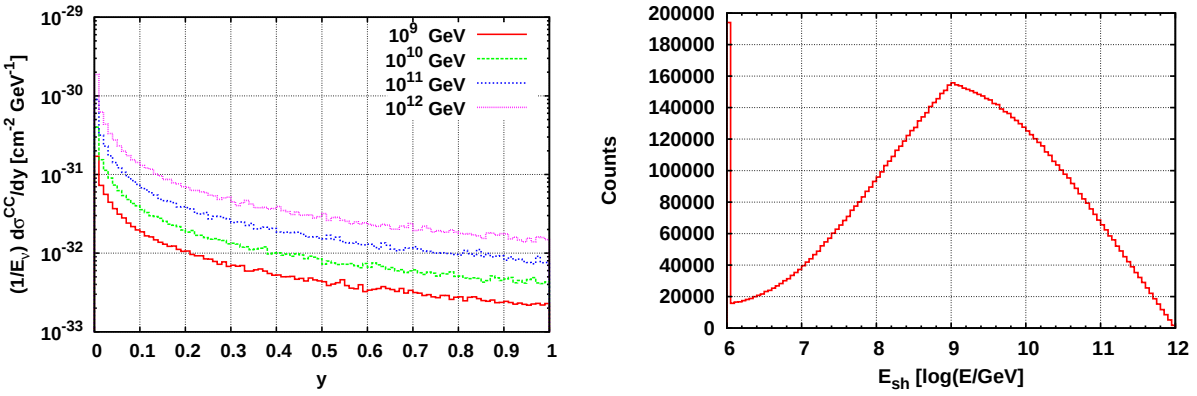
3.1.2 Shower generation

As starting point, the work done by the ACoRNE[2] Collaboration was used for the shower simulation [36,37][3]. The basic idea of that work was to use a modified version of the CORSIKA[4] simulation [55] in order to reproduce the hadronic showers in water and the resulting energy deposition. The distribution of this energy deposition was then parameterised, in order to reduce the required time for the simulation to run. The energy deposition distribution used in the simulation chain presented here is based on their parametrisation. This incorporates the result of 1500 CORSIKA showers in water for each of the half magnitude steps in the energy range between 10^5 GeV and 10^{12} GeV. Utilising the cylindrical symmetry of the cascade, each of

[2] *Acoustic Cosmic Ray Neutrino Experiment* (ACoRNE)

[3] Matlab source code available at `http://www.hep.shef.ac.uk/research/acorne/shower.php`

[4] *Cosmic Ray Simulations for Kascade* (CORSIKA)



(a) y-distribution

(b) Total shower energy

Figure 3.2: (a) The y-distribution of neutrino-nucleon charged current interaction for neutrino energies ranging from 10^9 GeV to 10^{12} GeV. (b) The simulated distribution of the total shower energy using the y-distributions for the neutrino energy ranging from 10^9 GeV to 10^{12} GeV. For shower energies smaller than 10^6 GeV, the energy is set to 10^6 GeV to avoid exceeding the valid energy range of the total energy deposited by the cascade.

the deposited energy distributions are represented by a matrix with 100 rows corresponding to the longitudinal dimension and 20 columns corresponding to the radial dimension. The actual parametrisation of the shower is performed by using the *singular value decomposition* (SVD) method. The main properties of the SVD [56] will be introduced and an overview on how the parametrisation was performed is given. The SVD of a $m \times n$ matrix $\mathbf{M}$ can be decomposed to the following form:

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^* \tag{3.2}$$

where $\mathbf{U}$ is a $m \times m$ unitary matrix, $\mathbf{\Sigma}$ is a $m \times n$ rectangular diagonal matrix with nonnegative real numbers on the diagonal, and $\mathbf{V}^*$ (the conjugate transpose of $\mathbf{V}$) is a $n \times n$ unitary matrix. The diagonal entries σ of $\mathbf{\Sigma}$ are the singular values of $\mathbf{M}$, which are commonly listed in descending order. The m columns of $\mathbf{U}$ and the n columns of $\mathbf{V}$ are called the left and right singular vectors of $\mathbf{M}$, respectively. The Eckart-Young theorem [57] asserts that the truncated singular value decomposition, obtained by discarding all but the first k largest singular values σ and their corresponding left and right singular vectors, is the best rank-k approximation in the sense of least squares to the original matrix.
For the derivation of the parametrisation, the SVD was performed two times, once for the radial and once for the longitudinal distribution of the deposited energy as a function of the total shower energy. In each decomposition step, the main components of the energy deposition can be separated from the fluctuations between simulated showers by only taking the largest singular values into account. As the value of the singular value σ can be interpreted as the importance for the approximation, the number of singular values to choose for the approximation can be estimated by defining a parameter α as:

$$\alpha = \frac{\sum_{i=1}^{i=k} \sigma_i}{\sum_{i=1}^{i=n} \sigma_i} \tag{3.3}$$

This procedure [36,37] leads to a linear parametrisation of the deposited energy by the shower for each of the energy steps. In comparison with the original shower data, the parametrisation is able to recreate the energy deposition within an average accuracy of 5 % in profile and magnitude. The mean and standard deviation of the linear parameters within the parametrisation for each of the 15 half decades of the energy range were used to reproduce the statistics of the showers. To overcome the limitation of the energy stepping, interpolation between steps were used to model intermediate values. As the corresponding interpolated singular vectors are not necessarily orthogonal, which would normally be the case for the SVD, the resulting correlations between the singular vectors have to be reintroduced by multiplying them with the matrix square root of the correlation matrix.
The longitudinal and radial energy density distribution of a 10^{11} GeV shower is shown in Fig. 3.3(a). In order to overcome limitations in the original binning of radial dimension r and the longitudinal dimension z, the energy distribution was linear interpolated in the two dimensions. From this energy distribution, a three dimensional representation — so-called *Monte Carlo* (MC) shower — is produced with a point density proportional to the energy density distribution. Typically 10^6 MC points are used to create the MC shower, this is a trade-off between runtime and accuracy. The resulting MC shower is shown in Figs. 3.3(b) – (d) as projection of three-dimensional MC point distribution of the MC shower to a plane defined by the axis of the coordinate system. The size of the bins is 0.001 m for x and y, and 0.005 m for z, respectively. The MC shower is about 20 m long and has a Molière radius[5] of about 10 cm. After the energy deposition within the cascade has been simulated, the MC point shower is translated to the position of the interaction vertex set in the volume around the detector and aligned according to the flight direction of the neutrino. For the energy range under consideration, the flight direction of the UHE neutrino and direction of the main axis of the shower can be assumed to be equal. Thereafter, the acoustic pulse and its propagation to the sensors within the detector are calculated.

3.1.3 Acoustic pulse calculation

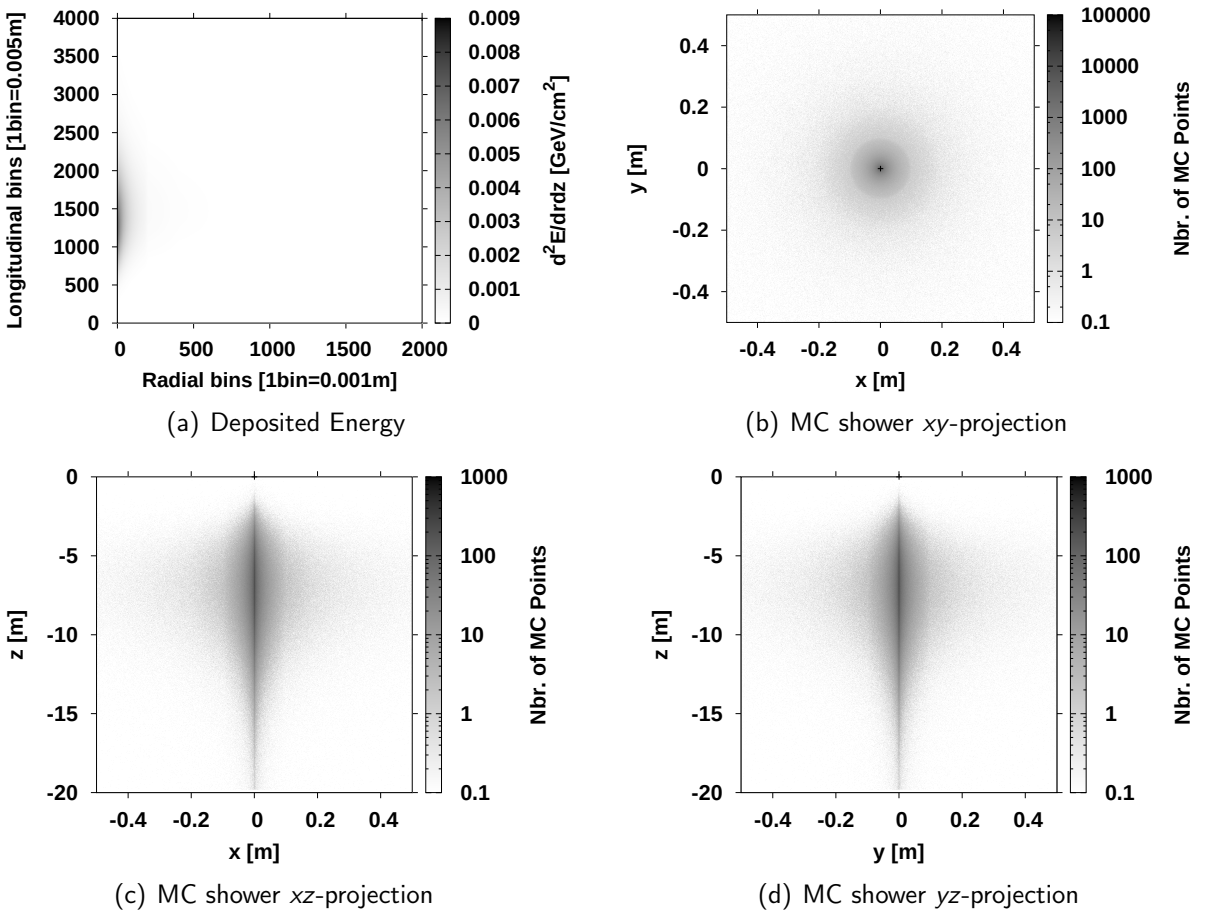
As mentioned in the previous section, the hadronic shower deposits energy in the medium, in this case in water, resulting in a local heating. With respect to hydrodynamical time scales, the energy deposition can be assumed instantaneous at time t_0, and the dissipation of the energy is slow in comparison. The energy deposition density $\epsilon(\mathbf{r}, t)$ can be factorized into a spatial and temporal part using the Heaviside function:

$$\epsilon(\mathbf{r}, t) = \tilde{\epsilon}(\mathbf{r})\Theta(t - t_0) \;\Rightarrow\; \frac{\partial}{\partial t}\epsilon(\mathbf{r}, t) = \tilde{\epsilon}(\mathbf{r})\delta(t - t_0)\,. \tag{3.4}$$

Assuming a total energy deposition E in a cylindrical volume containing the shower, the spatial part $\tilde{\epsilon}(\mathbf{r})$ can be expressed for the longitudinal and radial positions z and r in the shower as:

$$E\tilde{\epsilon}(\mathbf{r}) = \frac{1}{2\pi r}\frac{\mathrm{d}^2 E}{\mathrm{d}r\mathrm{d}z}\,. \tag{3.5}$$

[5] By definition, this is the radius of a cylinder containing on average 90 % of the shower's energy deposition. As the energy deposition is proportional to the point density in this case, the 90 %-quantile of the distance distribution of the points from the showers main axis is used.



(a) Deposited Energy

(b) MC shower *xy*-projection

(c) MC shower *xz*-projection

(d) MC shower *yz*-projection

Figure 3.3: Energy density distribution and generated MC shower for a total shower energy is 10^{11} GeV. (a) The longitudinal and radial distribution of the deposited energy of the shower in water. (b) – (d) show the projection of three-dimensional MC point distribution of the MC shower to a plane defined by the axis of the coordinate system. Bin-sizes for x and y are 0.001 m and 0.005 m for z, respectively.

Inserting Eq. 3.4 and Eq. 3.5 into the expression for the pressure p, following [32], as derived in Sec. 2.3.2:

$$p(\mathbf{r}, t) = \frac{\alpha}{4\pi c_p} \int \frac{\mathrm{d}^3 r'}{|\mathbf{r} - \mathbf{r}'|} \frac{\partial^2}{\partial t^2} \epsilon(\mathbf{r}', t - \frac{|\mathbf{r} - \mathbf{r}'|}{c_s}), \tag{3.6}$$

where α is the thermal expansion coefficient, c_p is the specific heat capacity at constant pressure and c_s the speed of sound in the medium. Eq. 3.6 can be further reduced to:

$$p(\mathbf{r}, t) = \frac{E\alpha}{4\pi c_p} \int \frac{\mathrm{d}^3 r'}{R} \tilde{\epsilon}(\mathbf{r}') \frac{\mathrm{d}}{\mathrm{d}t} \delta(t - \frac{R}{c_s}), \tag{3.7}$$

where $R = |\mathbf{r} - \mathbf{r}'|$ is the distance between the shower maximum and the sensor. A velocity potential can be defined, following [36, 37]:

$$E_{xyz}(t) = \frac{E\alpha}{4\pi c_p} \int \frac{\mathrm{d}^3 r'}{R} \tilde{\epsilon}(\mathbf{r}') \delta(t - \frac{R}{c_s}). \tag{3.8}$$

The velocity potential E_{xyz} is evaluated at the position of the sensor. For the Mediterranean Sea, the water temperature can be assumed constant and $\alpha = 2.0 \cdot 10^{-4}$ /K at 287 K, $c_p = 3.8 \cdot 10^3$ J/kg/K, and the speed of sound $c_s = 1542$ m/s at the detector centre are used. As described before, the distribution of points within the MC shower is proportional to the energy density distribution of the shower. So the pressure at a sensor in a distance R from the shower can be numerically calculated. The signal propagation time from each point within the MC shower to a sensor in the detector is calculated and entered into a histogram with a bin-width according to the sampling rate (here 1 µs) and a size sufficient to hold the distribution (here 2^{15} bins). After normalising each bin with the number of points in the shower, scaling it with the constant term (see Eq. 3.8), and dividing by the mean distance to the shower, this results in the velocity potential $E_{xyz}(t)$ as shown in Fig. 3.4(a) for a total shower energy of 10^{11} GeV and a distance of 1 km. The Fourier Transform of Eq. 3.7 including Eq. 3.8 can be written as:

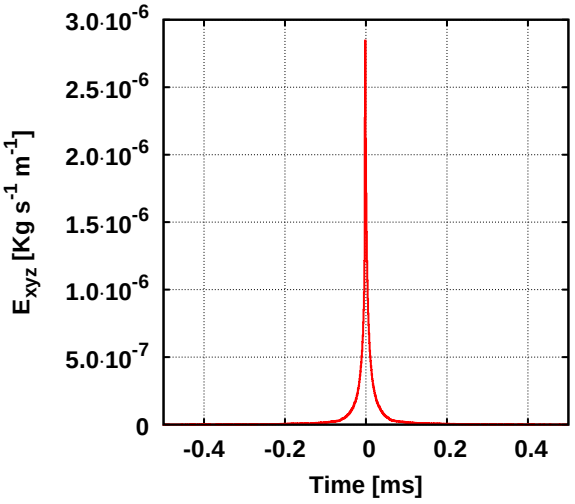
$$p(\omega) = \int \frac{\mathrm{d}}{\mathrm{d}t} E_{xyz}(t) \mathrm{e}^{-i\omega t} \mathrm{d}t = i\omega \int E_{xyz}(t) \mathrm{e}^{-i\omega t} \mathrm{d}t = i\omega E_{xyz}(\omega)\,, \tag{3.9}$$

taking into account the basic property of the *Fourier transformation* (FT) that the derivative in the time domain is the same as multiplying by $i\omega$ in the frequency domain. $E_{xyz}(\omega)$ is derived from the histogram $E_{xyz}(t)$ using the *fast Fourier transformation* (FFT). Also in the frequency domain, the frequency-dependent sound attenuation in sea water can be accounted for by multiplying it with the pressure signal. The sound attenuation model used in this simulation chain is discussed below. The resulting acoustic signal as a function of time for a total shower energy of 10^{11} GeV and a distance of 1 km is shown in Fig. 3.4(b). The bipolar pulse shape can be clearly recognised. The corresponding relative *power spectral density* (PSD) of the acoustic signal as function of the frequency is presented in Fig. 3.4(c). The spectrum is broad, as 90 % of the spectral power is contained within the frequency range from 1 to 100 kHz with a maximum spectral density around 11 kHz.

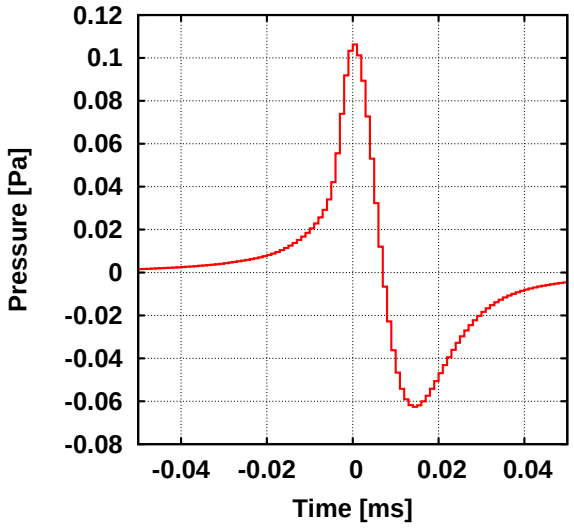
Attenuation

The attenuation of sound is described by an exponential function of the sound intensity I:

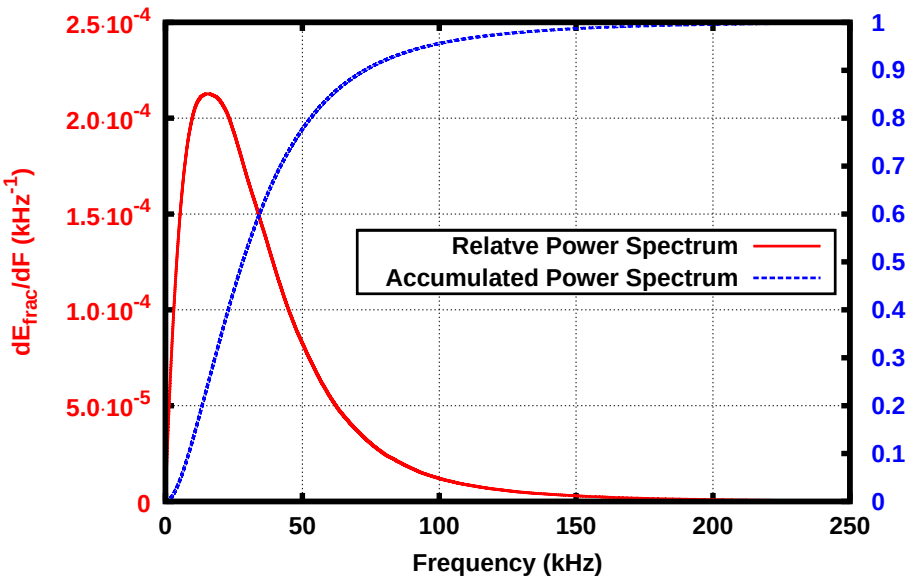
$$I = I(r_0) \exp(-\eta r / r_0)\,, \tag{3.10}$$



(a) Velocity potential



(b) Resulting bipolar pulse



(c) Relative PSD

Figure 3.4: The velocity potential $E_{xyz}(t)$, the resulting acoustic signal and its PSD calculated for a MC shower of a total energy of 10^{11} GeV and a distance of 1 km. (a) The velocity potential $E_{xyz}(t)$ as a function of the flight time, which is normalised to the mean flight time. (b) The calculated acoustic signal at the sensor as a function of the flight time and (c) the relative PSD of the signal as function of the frequency.

where η is the attenuation coefficient, r is the distance from the emitter to the receiver, and r_0 is a reference distance, which is typically 1 m. Expressed as intensity level L, this translates into a linear equation:

$$L(r) = -\alpha r \text{ with} \tag{3.11}$$

$$\alpha = 10 \log(e)\eta\,, \tag{3.12}$$

where α is the logarithmic attenuation coefficient (in dB/m or dB/km). Following the approach of Ainslie and McColm [58], the frequency-dependent attenuation coefficient $\alpha(f)$ for sea water can be represented by the contribution of the processes involved:

$$\alpha(f)(\text{in dB/km}) = \sum_{j=1,2,3} \alpha_j = \sum_{j=1,2} A_j \left(\frac{f_j f^2}{f_j^2 + f^2} + \mathrm{i} \frac{f_j^2 f}{f_j^2 + f^2} \right) + A_3 f^2\,. \tag{3.13}$$

The first term describes absorption by chemical relaxation processes. In this model, boric acid (α_1) and magnesium sulphate (α_2) contribute to the attenuation. Both can be represented as complex high-pass filters with specific cut-off frequencies f_j; thus they introduce a phase-shift, as described by Liebermann [59]. The second term introduces viscous absorption caused by particle motion (α_3). A_j are constants depending on environment conditions like pressure, temperature, salinity and acidity. The absorption coefficients α_j can be parametrized as:

$$\alpha_1 = 0.106 \exp\left(\frac{pH - 8}{0.56}\right) \left(\frac{f_1 f^2}{f_1^2 + f^2} + \mathrm{i} \frac{f_1^2 f}{f_1^2 + f^2} \right) \text{, with } f_1 = 0.78 \sqrt{\frac{S}{35}} \exp\left(\frac{T}{26}\right) \text{ kHz} \tag{3.14}$$

$$\alpha_2 = 0.52 \left(1 + \frac{T}{43}\right) \frac{S}{35} \exp\left(-\frac{D}{6}\right) \left(\frac{f_2 f^2}{f_2^2 + f^2} + \mathrm{i} \frac{f_2^2 f}{f_2^2 + f^2} \right) \text{, with } f_2 = 42 \exp\left(\frac{T}{17}\right) \text{ kHz} \tag{3.15}$$

$$\alpha_3 = 4.9 \cdot 10^{-4} \exp\left[-\left(\frac{T}{27} + \frac{D}{17} \right) \right] f^2 \tag{3.16}$$

T is the water temperature in °C, D is the depth in km, pH the acidity and S the salinity in ppt. For the Mediterranean Sea at the depth of the centre of the AMADEUS detector, $D = 2.15\,\text{km}$, $T = 13.2\,°\text{C}$, $S = 38.5\,\text{ppt}$, and $pH = 8.1$ are used.

Characteristics of the acoustic signal

Measures have to be defined to characterise the acoustic signal of a neutrino-induced interaction, such as the one calculated from the simulation of the deposited energy by a cascade. A peak can be defined as the point, where a curve has an extremum. So the peak pressure of an acoustic signal can be defined as the maximal pressure p reached by the time-domain waveform relative to the equilibrium. Also the peak-to-peak amplitude can be formulated as: $p_{\text{peak-to-peak}} = |\max p| + |\min p|$ and the asymmetry of the waveform as: $\frac{|\max p| - |\min p|}{|\max p| + |\min p|}$. The duration of the BIP is defined as the time between the occurrence of the positive and the negative peak of the waveform.
As mentioned before, the pulse propagates through the water in a disk-like shape perpendicular to the cascade's main axis — the so-called pancake. The emission angle of this propagation

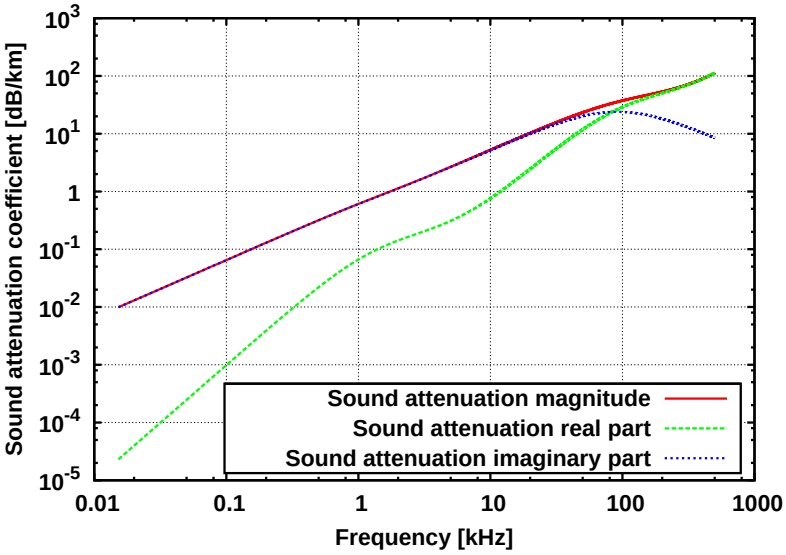
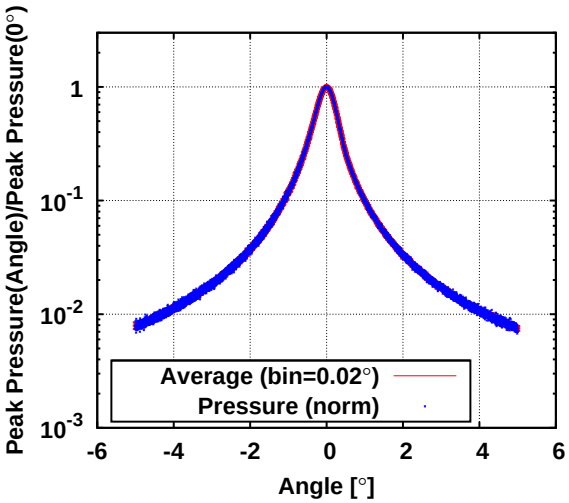


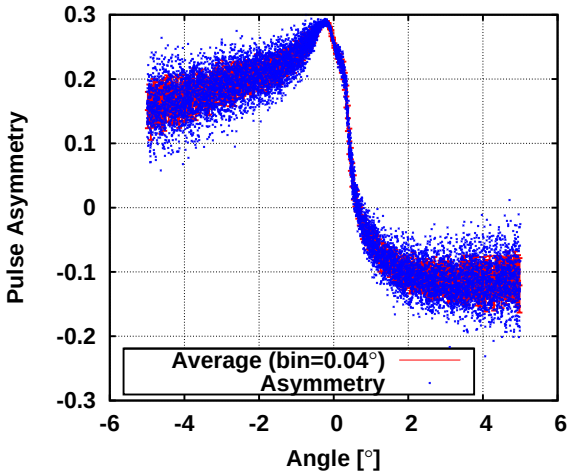
Figure 3.5: The attenuation coefficient *alpha* given in dB/km as a function of the frequency in kHz.

pattern can be defined as the angle between the plane through the shower maximum perpendicular to the main axis of the shower and the receiver. A positive emission angle indicates that the receiver is below this plane. The dependency on the relative peak pressure from the emission angle is shown in Fig. 3.6(a). The opening angle of the pancake can be defined as the *full width at half maximum* (FWHM) of the distribution, which is about 0.6°. Due to the geometric spread and attenuation in sea water, as discussed above, the peak pressure is a function of the distance r from the cascade to the receiving sensor. It is expected that in the near-field the amplitude has a $1/\sqrt{r}$-dependency, in the far-field a $1/r$-dependency, and in the attenuation dominated region a $1/r^2$-dependency [60]. This dependency is shown in Fig. 3.6(c), until about 100 m the near-field dependency is dominant, between 100 – 2000 m the transition between near- and far-field occurs, and for greater distances the attenuation becomes dominant. Fig. 3.6(b) shows the asymmetry of the pulse in relation to the emission angle defined above. For positive emission angles, the pulse initially becomes more symmetric while moving out of the median plane and then the asymmetry becomes negative at larger angles. For negative angles, the asymmetry of the pulse decreases but not to the extent as for positive ones. The shower maximum is not centred along the main axis of shower, it is located at about 1/4 of the total length of the shower. This results in the difference between positive and negative angles in the angular distribution of the asymmetry of the pulse. The asymmetry as a function of the distance r is presented in Fig. 3.6(d). In the near-field region the determination of the asymmetry is problematic as small variations in the position of the shower maximum have a sizable impact on this. For larger distances, the asymmetry of the pulse decreases without becoming completely symmetric, which is caused by the complex nature of the attenuation.
The PSDs as function of the distance, angle, and total shower energy is shown in Fig. 3.7; the dependence on the PSD from the distance, as shown in Figs. 3.7(a) – (b), is the following: While in the near-field the PSD is broad and flat, the main spectral component evolves in the range of 5 – 50 kHz peaking around 10 – 15 kHz as the distance grows into the tran-

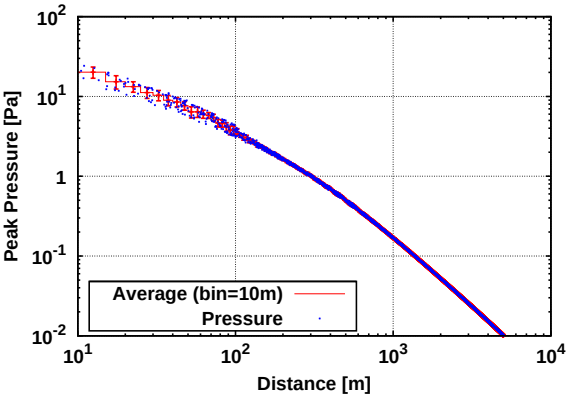
sition region between near- and far-field. The effect of the angle relative to the receiver is that the peak spectral density is shifted below 15 kHz and the power contained in this peak increases for larger angles. (cf. Figs. 3.7(c) – (d)). The total energy of shower has minor effect on the PSD. The PSD is shifted to higher frequencies by about 1 kHz per decade of energy (cf. Figs. 3.7(e) – (f)).



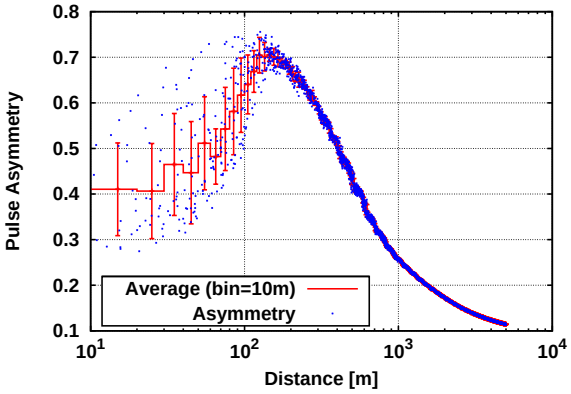
(a) Peak pressure angular distribution



(b) Asymmetry angular distribution



(c) Peak pressure over distance distribution

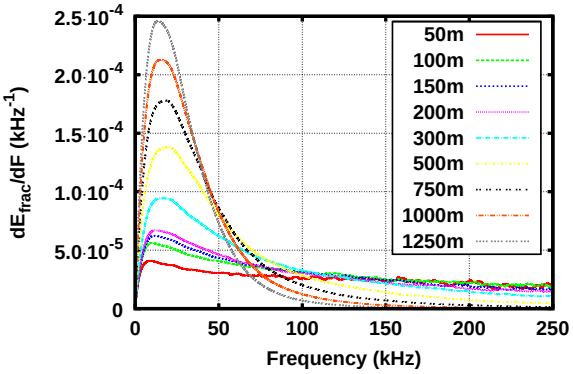


(d) Asymmetry over distance distribution

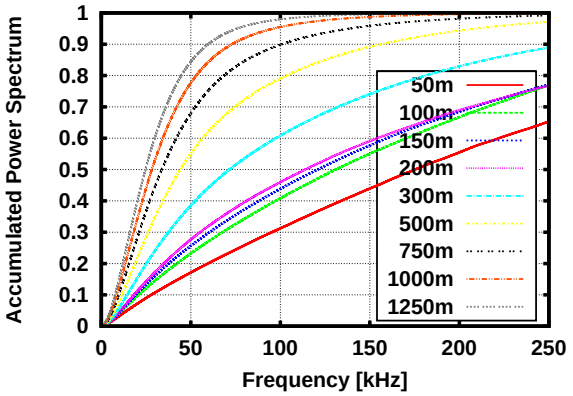
Figure 3.6: For all four figures, the total shower energy is 10^{11} GeV. In (a), the peak pressure for a distance of 1000 m is shown as function of the emission angle defined in the text and, in (c), as function of the distance from the cascade to the receiver at an emission angle of 0°. In (b) and (d) the asymmetry is given as function of the emission angle and distance, accordingly.

3.2 Transient and ambient noise simulation

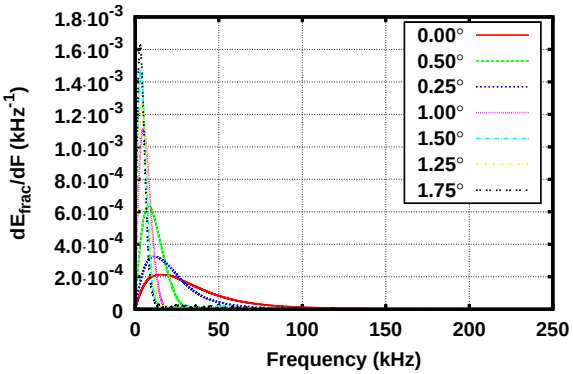
The background for acoustic neutrino detection in the deep sea consists of two different types of noise: transient and ambient noise. Transient noise signals have short durations and amplitudes that exceed the ambient noise level ($S/N > 1$). Transients can mimic bipolar pulses



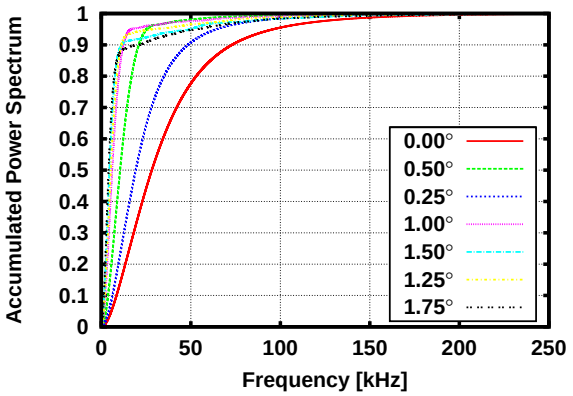
(a) Rel. PSD for different distances



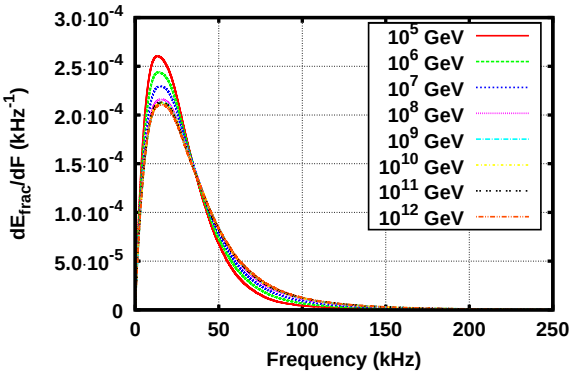
(b) Rel. PSD for different distances (acc.)



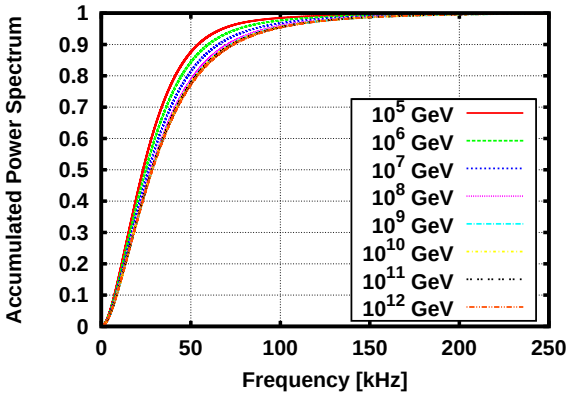
(c) Rel. PSD for different angles



(d) Rel. PSD for different angles (acc.)



(e) Rel. PSD for different energies



(f) Rel. PSD for different energies (acc.)

Figure 3.7: The PSDs of a neutrino-induced BIP is presented as relative and accumulative function of the frequency. The dependency on the PSD for different distances ((a)−(b)), angles to the shower ((c)−(d)) and energies of the shower ((e)−(f)) is shown. Described in more detail in the text.

from UHE neutrino-induced showers. In the simulation, four types of transient signals based on observations with the AMADEUS detector are implemented: bipolar and multipolar pulses, sinusoidal signals, and signals with *Brown noise* frequency characteristics. Sources of these four types can either be marine mammals or anthropogenic sources, such as shipping traffic. The ambient noise is mainly caused by agitation of the surface of the sea [50], e.g., by wind, breaking waves, spray, and cavitations. Thus it is correlated to the weather conditions, mainly to the wind speed. The model used for the simulation of the ambient noise is based on the so-called Knudsen spectra [61], which are adapted to the deep sea by applying attenuation effects.

3.2.1 Transient noise simulation

The simulation chain is capable of generating typical transient signals present at the ANTARES site. Four different typical types of transient signal are implemented: bipolar and multi-polar pulses, which can be caused by shipping traffic or marine mammals; sinusoidal signals such as reflected signals from the ANTARES acoustic positioning system (cf. Sec. 2.4); and signals that have Brown noise characteristics and do not belong to one of the previous categories. For example, the time-domain waveform and the corresponding PSD of the four signal types are shown in Fig. 3.8 – 3.11. The initial pressure amplitude p_0 of all four signals types, which is set at the source position of the pressure pulse, follows a definable frequency distribution. The other parameters, namely the duration and the frequency, of the signal are selected from a uniform distribution.
For the transient BIP, the following parametrisation is used:

$$p = -p_0 \frac{t - t_0}{\tau} \exp\left(\frac{-(t - t_0)^2}{2\tau^2}\right), \tag{3.17}$$

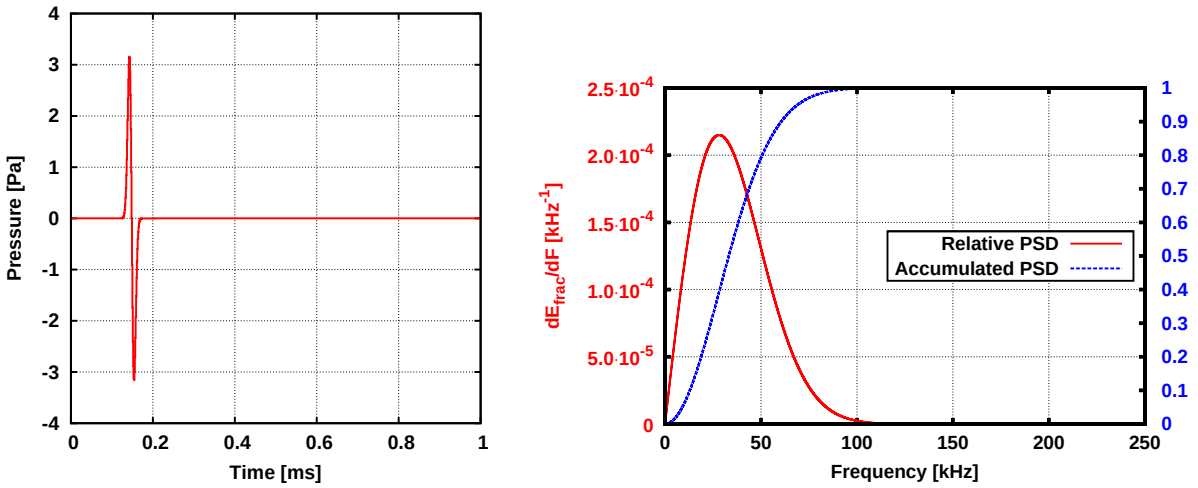
where t_0 is the time at the centre (zero-crossing) of BIP and $\tau \in [5\mu s, 50\mu s]$ is the duration of the signal. In Fig. 3.8, the resulting waveform of a symmetric BIP is shown. The broad PSD is comparable to the one of a neutrino-induced BIP.
For the multipolar signal a windowed sine function is used:

$$p = \frac{1}{2}\left[1 - \cos\left(2\pi \frac{|t - t_0|}{\tau}\right)\right] \sin\left(\omega(t - t_0)\right), \tag{3.18}$$

where the first term is a window function defining the length of the multipolar signal by choosing τ from 40 – 4096 µs and $|t - t_0| < \tau/2$ is required for the start time t_0. The frequency ω of the sine wave ranges from 5 – 50 kHz. The waveform of this signal is given in Fig. 3.9. The corresponding PSD of this signal is dominated by the main frequency of the sine term.
The sinusoidal signals have a duration between 1 µs and 4 µs, and their frequencies range from 40 – 60 kHz (cf. Fig. 3.10). This is the same frequency range, in which the emitters of the ANTARES positioning system work.
In Fig. 3.11, the last signal type is shown that correspond to a signal of random shape with a Brown noise frequency characteristic[6]. The duration of the signal ranges from 40 – 4096 µs. As

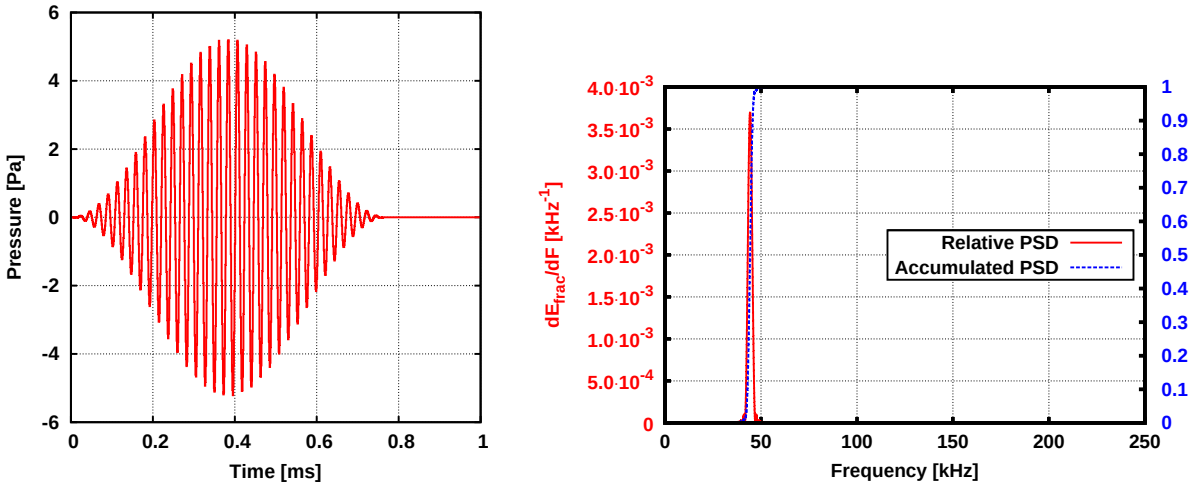
[6]Brown noise is defined by spectral density that is proportional to $\frac{1}{f^2}$. Brown noise can be produced by integrating white noise. That is, whereas white noise can be produced by randomly choosing each sample independently, Brown noise can be produced by adding a random offset to each sample to obtain the next one.

described for the simulation of neutrino-induced pressure pulses, the source of the transient signal is set to a random position within a given volume around the detector and propagated to the sensors. The signal types can be generated separately or all together following a defined frequency distributions. For most applications a uniform distribution is chosen.



(a) Waveform (b) PSD

Figure 3.8: Transient signal of the BIP type shown in the time domain (a) and the corresponding PSD (b).

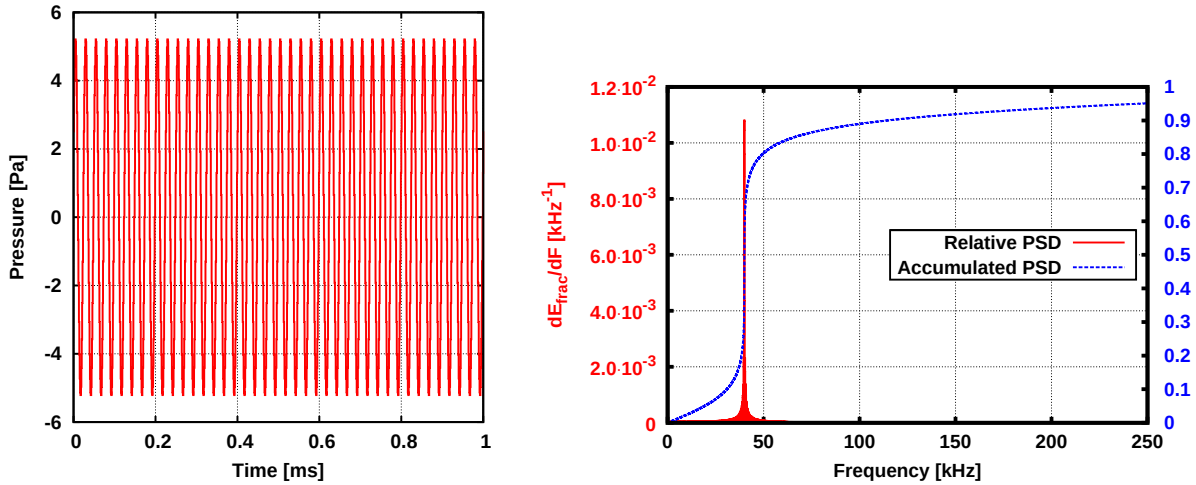


(a) Waveform (b) PSD

Figure 3.9: Transient signal of the multipolar type shown in the time domain (a) and the corresponding PSD (b).

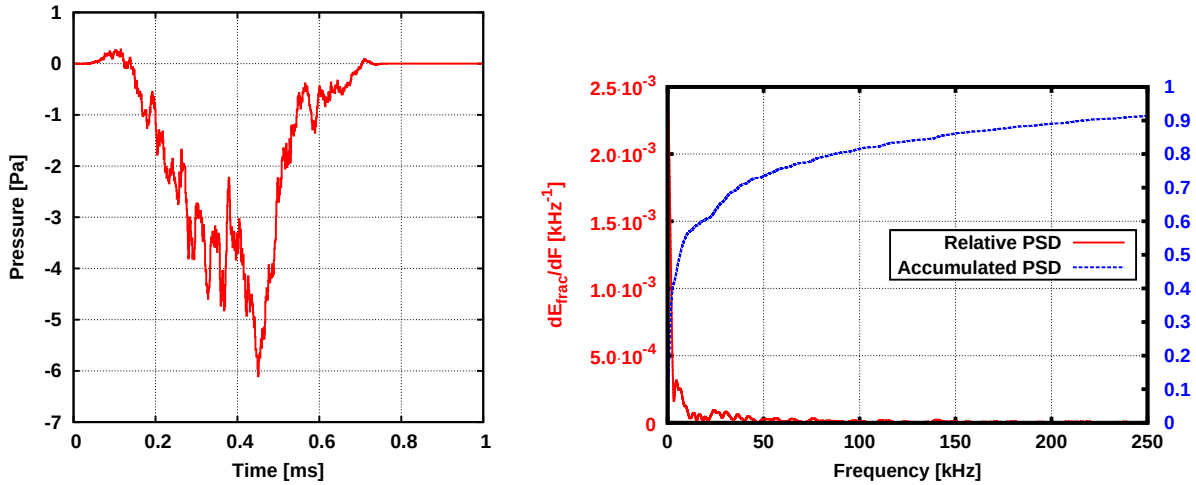
3.2.2 Ambient noise simulation

Following the work of Knudsen [50, 61], the ambient noise spectrum can be parameterised in three independent frequency regimes, where the noise characterised by its source: distant



(a) Waveform (b) PSD

Figure 3.10: Transient signal of the type representing the echoes of the positioning system shown in the time domain (a) and the corresponding PSD (b).



(a) Waveform (b) PSD

Figure 3.11: Transient signal of the type representing signal with Brown noise features shown in the time domain (a) and the corresponding PSD (b).

anthropogenic noise, sea surface noise, and Brownian motion of water molecules, is dominant. Under the assumption that these sources are incoherent, the final power spectrum is calculated by adding the individual spectra. In the following, the parametrisation of these spectra and their implementation in the simulation are discussed.

Distant anthropogenic sources

In the range between 10 – 1000 Hz, the ambient noise is dominated by distant shipping and industrial noise. The spectral density of this type of sources is proportional to $\frac{1}{f^\alpha}$ with $\alpha \in [1, 2]$. Due to the great distance to the sources, higher frequencies are absorbed. Nearby shipping traffic would contribute to the transient noise.
Noise level $\left(\text{dB re } 1\,\mu\text{Pa}^2/\text{Hz}\right)$:

$$NL_{ship}(f) = \begin{cases} NL_{100} & \text{for} \quad f \leq 100\,\text{Hz} \\ NL_{100} - 20\log(\frac{f}{100\,\text{Hz}}) & \text{for} \quad f > 100\,\text{Hz} \end{cases} \tag{3.19}$$

The scaling factor NL_{100} is between 60 – 90 dB depending on the density of distance anthropogenic sources. A uniform frequency distribution is assumed, as the spatial and temporal distribution of these sources is unknown.

Sea surface noise

The noise from the surface of the sea itself is caused by wind, breaking waves, spray, and cavitation. This noise type has its main influence between 1 – 100 kHz and can be parametrized as the following:

$$NL_{surface}(f) = \begin{cases} NL_{1k} & \text{for} \quad f \leq 1\,\text{kHz} \\ NL_{1k} - 17\log(\frac{f}{1\,\text{kHz}}) & \text{for} \quad f > 1\,\text{kHz} \end{cases} \tag{3.20}$$

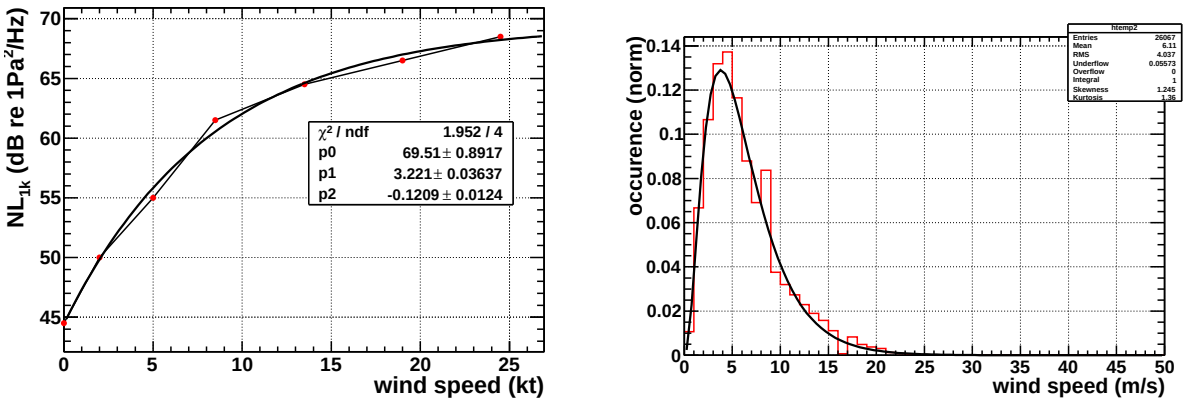
Discrete values of NL_{1k} and their dependency on the wind speed are provided in Tab. 3.1. The wind speed measured at the surface of the sea and the noise level are correlated. This was verified by measurements of the ambient noise level using the AMADEUS detector [62]. The noise level was derived from the minimum bias data-set (cf. Sec. 3.4.1), which consists of 10 second long data samples roughly every hour for each sensor. Using the values in Tab. 3.1, a continuous dependency on $NL_{1k}(c_w)$ (cf. Fig. 3.12(a)), from the wind speed c_w in knots can be derived using the following fit:

$$\begin{aligned} NL_{1k}(c_w) &= p_0 - 1\,\text{dB re } 1\,\text{Pa}^2/\text{Hz} \cdot \exp(p_1 + p_2 c_w) \text{ for } c_w \text{ in kt, where} \\ p_0 &= 69.5 \pm 0.9 \text{ in dB re } 1\,\text{Pa}^2/\text{Hz} \\ p_1 &= 3.22 \pm 0.04 \\ p_2 &= -0.121 \pm 0.012 \text{ in } 1/\text{kt} \end{aligned} \tag{3.21}$$

The distribution of wind speeds as measured by the Cap Cepet weather station, which is the nearest weather station to the detector, are given in Fig. 3.12(b). The distribution includes hourly measurements between 2007 – 2009 and was parameterised using an adapted Maxwell

Beaufort Force	Sea State	Windspeed (knots)	NL_{1k} (dB)	Description
0	0	0	44.5	Calm
1	0.5	1 – 3	50.0	Light Air
2	1	4 – 6	55	Light Breeze
3	2	7 – 10	61.5	Gentle Breeze
4	3	11 – 16	64.5	Moderate Breeze
5	4	17 – 21	66.5	Fresh Breeze
6	5 – 6	22 – 27	68.5 – 70	Strong Breeze
7	7	28 – 33		Near Gale
8	8	34 – 40		Gale
9	9	41 – 47		Strong Gale
10	9	48 – 55		Storm
11	9	56 – 63		Violent Storm
12	9	$\geq$ 64		Hurricane

Table 3.1: Connections between the windspeed in knots ($1\,kt \equiv 0.5144 m/s$), the sea state, and NL_{1k} [50], as described in the text. Adapted from [49]



(a) $NL_{1k}(c_w)$ (b) Wind speed distribution

Figure 3.12: (a) The scaling factor NL_{1k} of the noise from the surface of the sea as a function the wind speed c_w including the values from in Tab. 3.1 and the fit as described in the Eq. 3.21. (b) The frequency distribution of wind speeds (in m/s) as measured at Cap Cepet weather station (hourly measurements in 2007 – 2009) and the fit by an adapted Maxwell distribution (cf. Eq. 3.21).

distribution:

$$\begin{aligned} P(c_w) &= p_0 \cdot c_w^{p_1} \cdot \exp(p_2 \cdot c_w^{p_3}) \text{ for } c_w \text{ in m/s, where} \\ p_0 &= 0.144 \pm 0.641\,\text{dB re } 1\,\text{Pa}^2/\text{Hz} \\ p_1 &= 2.59 \pm 4.24 \\ p_2 &= -1.35 \pm 3.52 \\ p_3 &= 0.728 \pm 0.61 \end{aligned} \tag{3.22}$$

Thermal noise

The thermal noise is caused by the Brownian motion of the water molecules and mainly contributes in the region above 100 kHz. The following parametrisation is used:

$$NL_{therm}(f) = -75 + 20 \cdot \log(f) \tag{3.23}$$

Combined ambient noise spectrum

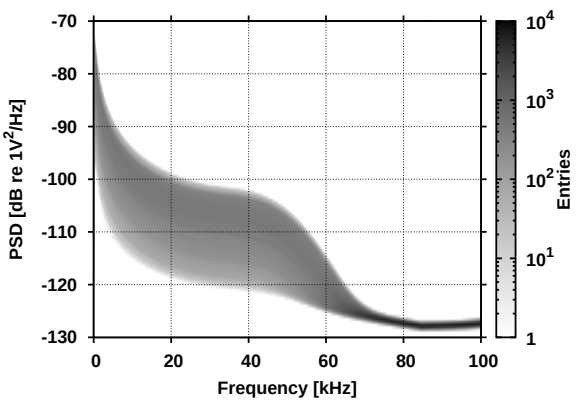
The ambient noise sources are assumed to be incoherent, so the summation of the noise levels is valid:

$$NL_{tot} = NL_{ship} \oplus NL_{surf} \oplus Nl_{therm} = 10 \cdot \log\left(10^{NL_{ship}/10} + 10^{NL_{surf}/10} + 10^{NL_{therm}/10}\right)$$

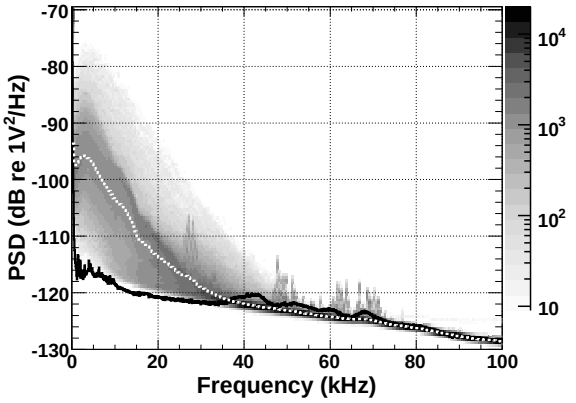
The Knudsen spectra [50] were measured in shallow water, and were thus adapted to the deep sea by applying attenuation effects. The source of the ambient noise is mainly the surface region above the detector. This is used as a starting point and the attenuation described in Sec. 3.1.3 is applied. The PSD of the ambient noise as simulated and measured are shown in Fig. 3.13; the scatter plots present the PSD of the ambient noise for different levels of the wind speed and shipping traffic. The comparison between the simulated and measured spectrum reveals that the model overestimates the noise level for frequencies below 10 kHz and between 40 – 60 kHz. The first region appears affected by the wind speed dependency on NL_{1k} and its parametrisation. The second region results from an underestimation of the surface area above the detector contributing to the surface noise. These effects could not be researched further as part of this work. The mean noise level $\langle \sigma_{noise} \rangle$ calculated from the simulation output is 27 ± 5 mPa for the frequency range from 1 – 50 kHz and for 95 % of time the noise level is smaller than $2\langle \sigma_{noise} \rangle$, and is in good agreement with the values for the noise measurements taken from [62]. The measured mean noise is level $\langle \sigma_{noise} \rangle = 25^{+7}_{-5}$ mPa and also for 95 % of time the noise level is smaller than $2\langle \sigma_{noise} \rangle$. This is reproduced by the simulation as shown in Fig. 3.14, where the resulting distribution of the simulated relative noise level and the measured one is shown, and are in good agreement.

3.3 DAQ-hardware simulation

The simulation of the DAQ-hardware comprises of two parts: the simulation of the acoustic sensors and of the read-out electronics. The DAQ-hardware includes acoustic sensors using the piezoelectric effect (hydrophones) to convert the pressure signal into a voltage signal, and read-out electronics (AcouADC-board) to amplify, filter, and digitise the voltage signal (cf. Sec. 2.5.2 and Sec. 2.5.3). The inherent noise and the system transfer function, for both the

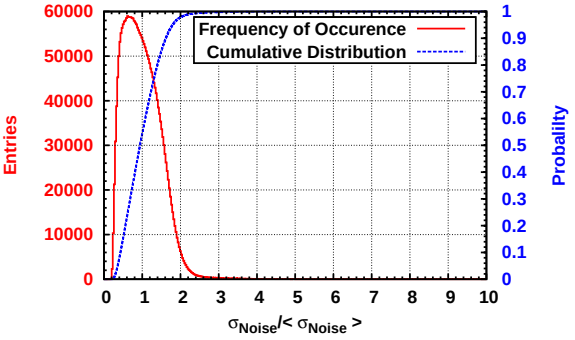


(a) Simulated noise PSD



(b) Measured noise PSD

Figure 3.13: (a) The PSD of the ambient noise as produced by the model described in the text. This scatter plot shows the PSD for the different levels of the weather conditions and shipping activities. (b) The PSD of the ambient noise measured with one HTI sensor on the topmost storey of the IL. The calibrated input voltage of the AcouADC-board is used. Shown in shades of grey is the occurrence rate in arbitrary units, where dark colours indicate higher rates. Shown as a white dotted line is the median value of the in-situ PSD and as a black solid line the noise level recorded in the laboratory prior to deployment [42].



(a) Simulated occurrence distribution

(b) Measured occurrence distribution

Figure 3.14: Frequency of occurrence distribution of the ambient noise level as function of the ratio between the noise level σ_{noise} of the sample and the mean noise level $\langle\sigma_{noise}\rangle$. The left scale represents the number of occurrences and the right scale the probability that a noise level occurs. (a) The result of the simulation based on an ambient noise model as described in the text. (b) the frequency of occurrence distribution for the ambient noise as measured by the AMADEUS detector [62] in the range 1 – 50 kHz, relative to the mean ambient noise recorded over the complete period of about two years that was used for the analysis (left scale, filled histogram). Also shown is the cumulative distribution, normalised to the total number of entries of the distribution (right scale).

sensors and the electronics, were measured in the laboratory. From this data, a parametrisation of the power spectral density of the inherent noise and the system response function of the hardware was derived. Sensors normally show a directional dependency on their sensitivity, therefore signal and ambient noise have to be treated separately. In this case, the incident direction of the noise is the surface of the sea above the detector. Signal and ambient noise are then superimposed. The inherent noise of the sensor is added. To simulate the effect of the read-out electronics, the resulting waveform is convoluted with the system transfer function and the inherent noise is added.

3.3.1 Sensor simulation

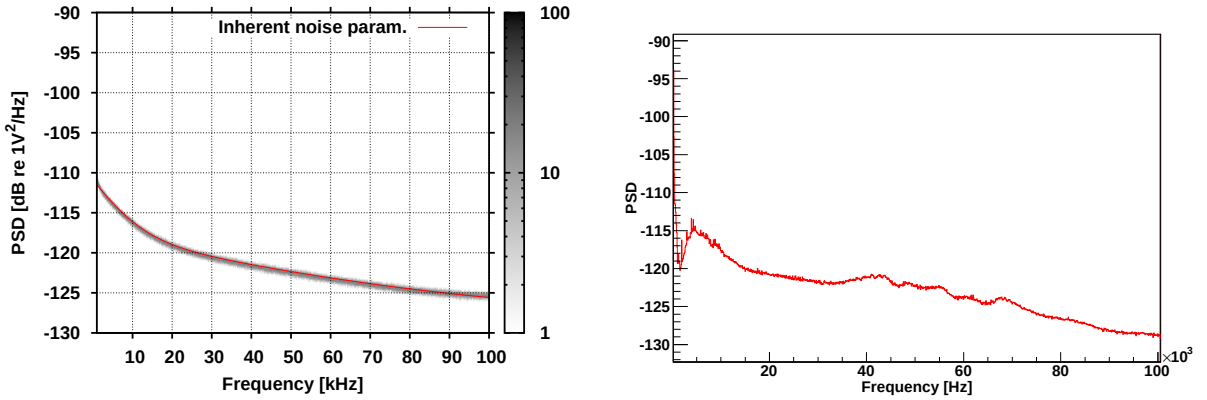
Different sensor types are incorporated into the AMADEUS detector, hydrophones from different manufactures and the AMs, where the acoustic sensor is included in a glass sphere (cf. Sec. 2.5.2). For the simulation, only the characteristics of one type of sensor is currently used as model for the 36 sensors included in the AMADEUS detector. As characteristics of the HTI hydrophones are comparable and stable over time, one HTI hydrophone included in the detector (HTI 24 on Storey 6 of the IL07) was chosen to represent the acoustic sensors in the simulation. The parametrisation of the inherent noise spectrum $N^{HTI}_{inherent}$ for HTI 24 is stated in Eq. 3.24 and is shown in Fig. 3.15 (b).

$$\begin{aligned} NL^{HTI}_{inherent}(f) &= 10 \cdot \log(10^{\frac{p_0+p_1 f}{10}} + 10^{\frac{p_2+p_3 f}{10}}) \\ p_0 &= -1.18 \cdot 10^2 \pm 3.42 \\ p_1 &= -1.1 \cdot 10^{-4} \pm 4.97 \cdot 10^{-5} \\ p_2 &= -1.12 \cdot 10^2 \pm 5.31 \\ p_3 &= -7.7 \cdot 10^{-4} \pm 9.22 \cdot 10^{-5} \end{aligned} \tag{3.24}$$

The system transfer function of the hydrophone is directional, and it shows a dominant dependency on the zenith angle, therefore signal and ambient noise have to be treated separately. The incident direction of the noise is assumed to be the surface of the sea above the detector. From measurements in the laboratory, a parametrisation of the system response function was derived in steps of 15° for the zenith angle [45, 46]. The sensitivity of the hydrophone used is shown in Fig. 3.16 as the function of the frequency and the zenith angle. The dependency on the azimuth angle is reasonably flat, and thus negligible. For the parametrisation of the system transfer function a 8th order polynomial fit to the measured magnitude $A(f)$ and phase $\phi(f)$ of the hydrophones response was used:

$$\begin{aligned} H(f) &= A(f) \exp\left(\mathrm{i}\phi(f)\right) \text{ with} \\ A(f) &= \begin{cases} \sum_{i=0}^{8} a_i f^i & \text{for } f \geq 10\,\text{kHz} \\ \text{const} & \text{for } f < 10\,\text{kHz} \end{cases} \\ \phi(f) &= \sum_{i=0}^{8} b_i f^i \end{aligned} \tag{3.25}$$

The calibration setup was not optimised for frequencies below 10 kHz, thus the magnitude is assumed constant below that frequency. However, this is acceptable, as the frequency range significant for neutrino detection is above 10 kHz.
By multiplying the pressure signal with the system transfer function $H(f)$ in the frequency domain and performing the inverse Fourier transform, the resulting voltage signal is used for the simulation of the read-out electronics, which is described below.



(a) Simulated inherent noise

(b) Measured inherent noise

Figure 3.15: (b) shows the PSD of the measured inherent noise and the (a) the parametrisation of the inherent noise of the hydrophone (HTI 24) and the AcouADC-board.

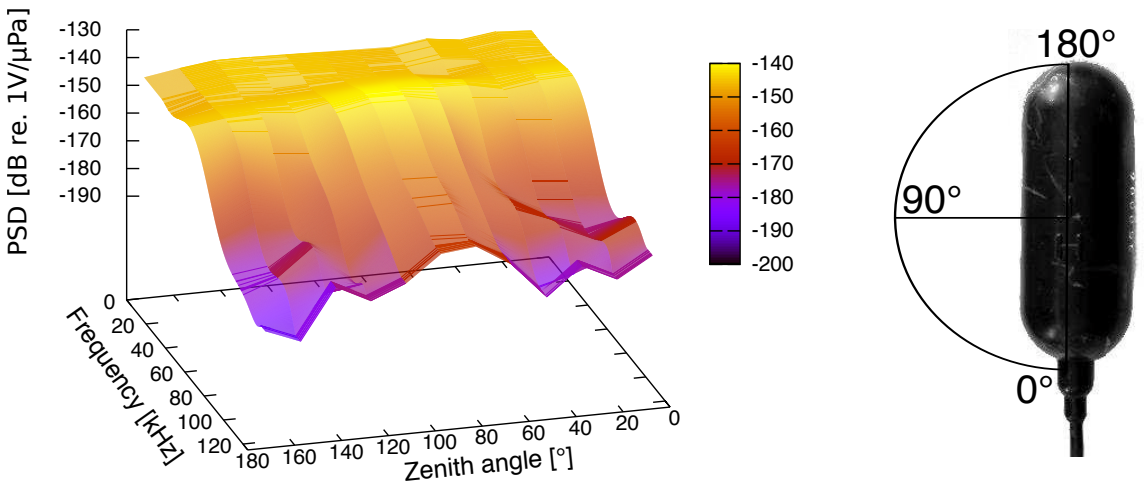


Figure 3.16: The directional sensitivity of the hydrophone used (HTI 24) is shown as the function of the frequency and the zenith angle. A zenith angle of 90° defines the horizontal plane through the centre of the hydrophone as depicted.

3.3.2 AcouADC simulation

The AcouADC-board filters, samples, and digitises the incoming voltage signal from the sensors. As for the sensors, the inherent noise and the system transfer function was measured in the laboratory and a parametric model was derived [49]. A simple noise model is assumed for inherent noise spectrum of the AcouADC-board: quadratic addition of a low-pass filter and a constant background. The parametrisation is given in Eq. 3.26 and shown in Fig. 3.15.

$$\begin{aligned} NL_{inherent}^{AcouADC}(f) &= 10 \cdot \log\left(\left(\frac{p_0}{\sqrt{1+(f/p_2)^{2p_3}}}\right)^2 + p_2^2\right) \\ p_0 &= -1.18 \cdot 10^2 \pm 3.42 \\ p_1 &= -1.1 \cdot 10^{-4} \pm 4.97 \cdot 10^{-5} \\ p_2 &= -1.12 \cdot 10^2 \pm 5.31 \\ p_3 &= -7.7 \cdot 10^{-4} \pm 9.22 \cdot 10^{-5} \end{aligned} \tag{3.26}$$

The system transfer function of the AcouADC-board is a function of the gain, downsampling, and range settings of the ADC. A parametrisation was derived from the calibration measurements using the complex-valued filter transfer function $H(f)$, equivalent to Eq. 3.25. The characteristics of the AcouADC-board are composed of simple models for filters and labelled by i, j and k. For $A(f)$ the following formulation is used:

$$A(f) = G_{tot} \cdot \prod_i A_{high,i}(f) \cdot \prod_j A_{low1,j}(f) \cdot \prod_k A_{low2,k}(f)\,, \text{ with} \tag{3.27}$$

$$A_{high,i}(f) = \left(\frac{f/f_{high,i}^0}{\sqrt{1+(f/f_{high,i}^0)^2}}\right)^{n_{high,i}}\,, \tag{3.28}$$

$$A_{low1,j}(f) = \left(\frac{1}{\sqrt{1+(f/f_{low1,j}^0)^2}}\right)^{n_{low1,j}}\,, \text{ and} \tag{3.29}$$

$$A_{low2,k}(f) = \frac{1}{\sqrt{1+(f/f_{low2,k}^0)^{2n_{low2,k}}}}\,. \tag{3.30}$$

And for $\phi(f)$:

$$\phi(f) = \phi_{tot} + \sum_i \phi_{high,i}(f) + \sum_j \phi_{low1,j}(f) + \sum_k \phi_{low2,k}(f) \text{ with} \tag{3.31}$$

$$\phi_{high,i}(f) = n_{high,i} \cdot \left(\frac{\pi}{2} - \arctan\left(f/f_{high,i}^0\right)\right)\,, \tag{3.32}$$

$$\phi_{low1,j}(f) = -n_{low1,j} \cdot \arctan\left(f/f_{low1,j}^0\right)\,, \text{ and} \tag{3.33}$$

$$\phi_{low2,k}(f) = d_{low2}^G \cdot 2\pi f\,. \tag{3.34}$$

The functions subscripted with *high* and *low*1 model the high- and low-pass filters[7] integrated on the AcouADC card, respectively. For those filters the corner frequencies f^0 and the order n were calculated from the values of the actual composition of the filter elements. The functions

[7] Passive or active RC/LRC filters with resistors, capacitors, and/or inductors.

subscripted with *low*2 describe generalised low-pass filters and are used to model a root-raised cosine (*RRC*) low-pass filter. This steep 10th order root-raised cosine filter suppresses high frequency components to eliminate aliasing effects of the digitisation. This filter exhibits an almost linear phase response with frequency, which was approximated by the group delay (d^G_{low2}). The total phase offset ϕ_{tot} was found to be -0.02 ± 0.15 rad, compatible with no additional total phase delay.
Additional low-passes — also with the generalised form A_{low2} — were used to model the finite impulse response (FIR) filters realised in the FPGA [63], which down-samples the digitised data. Those digital FIR filters use the same filter core for all boards and have a frequency-independent group delay. The total delay of the FIR filters depends on the downsampling (DS) setting: for DS1 no delay is induced, for DS2 and DS4 the delay corresponds to 64 samples of the raw data, i.e. 128 µs. The parameters of the transfer function used in the simulation are listed in Tab. 3.2. The parametrisation of the amplitude and phase response of the AcouADC-board used in the simulation are given in Fig. 3.17 as a function of the frequency.
The system transfer function and the inherent noise are applied to the data coming from the sensor simulation. Finally, the voltage signal is mapped according to the 16-bit resolution of the ADC with an input range of -2 to 2 V as used in the AMADEUS detector.

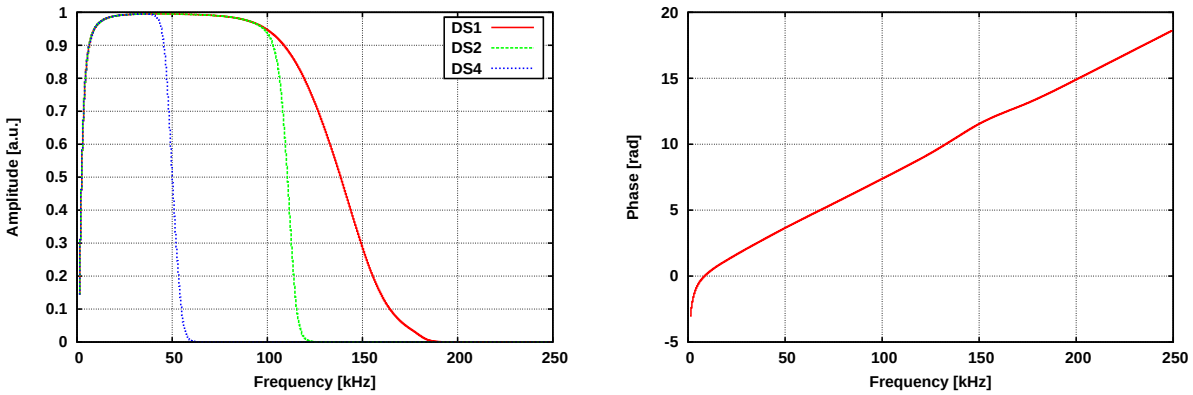
Subscript	Pass Type	Corner Frequency f^0 (kHz)	Order n
high,1	RC, high	1.129	1
high,2	RC, high	1.881	2
high,3	RC, high	0.03183	1
low1,1	LRC, low	1022	3
low2,1	RRC, low	128.0	5.000
low2,2	RRC, low	150.0	10.00
low2,3	RRC, low	178.1	18.45
low2,4	FIR, DS2	109.1	22.00
low2,5	FIR, DS2	112.8	38.44
low2,4	FIR, DS4	47.96	13.76
low2,5	FIR, DS4	54.03	24.38

$G_{tot} \in \{1, 10, 100\}$ $\quad \phi_{tot} = -0.05$ rad $\quad d^G_{l2} = 12.39\,\mu s$

Table 3.2: Parameters of the filter transfer function described in the text.

3.4 On-line filter simulation

The final step in the simulation chain is the simulation of the on-line filter. The now digitised signals from the simulation are subjected to a filter system equivalent to the one used in the experiment [48, 64], where it is used to identify pre-defined signal types and thus reduce the amount of data stored for off-line analysis, like signal classification and position reconstruction. The filter setup consists of three filters: A minimum bias filter; an amplitude threshold for transient signals, which is self-adjusting to the changing ambient noise conditions; and a pulse shape recognition filter for bipolar signals, which is based on a cross correlation technique. The



(a) Amplitude response (b) Phase response

Figure 3.17: The system response of the filters integrated in the AcouADC-board is shown as function of the frequency; (a) The amplitude response is given for the downsampling 1, 2 and 4. (b) The corresponding phase response is given. For more information see text.

filters are applied to each sensor within the detector. If one of the last two filters is activated by a signal, a coincidence test is performed to check whether or not the waveform recorded by the sensors corresponds to same signal source. A predefined set of sensors is required to record the signal in a given time window. So the signal has to pass the filter and the coincidence test to form an event (cf. Fig. 3.18). The filters and the coincidence test will now be described in more detail.

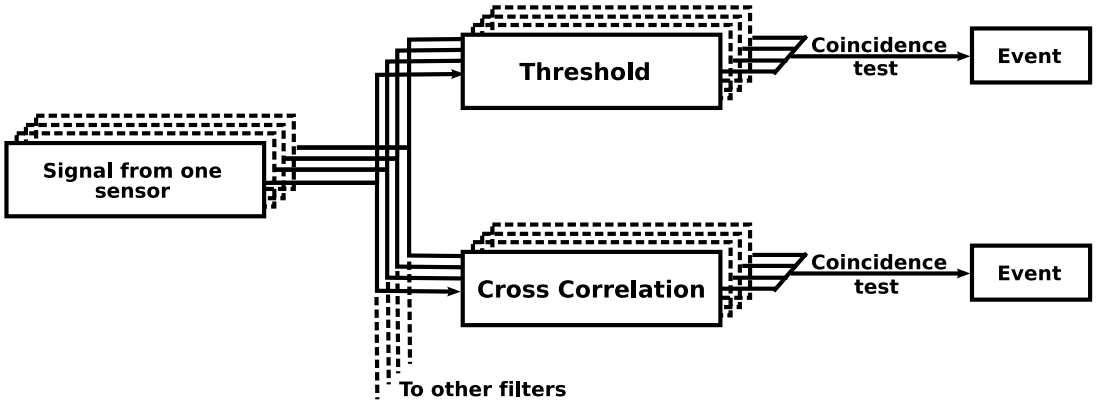


Figure 3.18: Overview of the main concept of the AMADUES online filter system.

3.4.1 Minimum bias filter

To obtain data with a minimum bias, a recording of the complete incoming data stream is performed at a regular interval. In the current operation of the AMADEUS detector, a sample of 10 s is acquired every hour. This accumulates to about four Gigabyte of minimum bias data per day for 36 sensors. This data is mainly used for ambient noise studies of the deep sea environment at the ANTARES site. Although it is used in the experimental setup and delivers

valuable information about the ambient noise, it was not necessary to implement this filter in the simulation.

3.4.2 Threshold filter

By definition, transient signals have a short time duration and an amplitude that exceeds a signal-to-noise ratio of one, thus a threshold is suitable to select these kind of signals. Furthermore, a threshold does not bias the selection with the requirement of specific frequency information of the signal. The ambient noise level, as described in Sec. 3.2.2, changes on different time scales and is unpredictable. Therefore, a variable threshold is needed, which is adaptive to the current noise conditions. This is implemented by using the *root mean square* (RMS) of the data sample multiplied by a predefined scaling factor to set the threshold. To prevent statistically appearing outliers triggering the selection, a configurable number of samples has to exceed the threshold. If these criteria are satisfied, a window around the signal is selected for storing, hence the complete data of the sensor will not be stored, only the parts containing the selected signals. The runtime complexity of this filter scales linearly with the number of data points N contained in the data stream ($\mathcal{O}(N)$). The simplicity and the limited usages of resource are the main advantages of this algorithm.
One known source of transient signals are the transducers of the ANTARES acoustic positioning system (cf. Sec. 2.4), which are located at the anchors of each detection line. As all lines are only anchored on the bottom side and moving with the sea currents, a precise positioning of the storeys (cf. Sec. 2.5.1) is crucial to reconstruct the position of unknown sources and thus the signal must be recorded. As the emitted sinusoidal signals have high amplitude and narrow bandwidth, they are easy to identify in the time domain and the power density spectrum. To guarantee a sufficient data reduction rate, only the incipient part of the waveform, the arrival time, and the peak frequency are stored for this type of signal.

3.4.3 Pulse shape recognition filter

The aim of this filter is to find any neutrino-like signal in the data and to characterise the similarity of these signals as compared to a given bipolar signal sample. For this the cross correlation is used, which is defined for two discrete real samples g and h of length N as the following:

$$(g \star h)[j] := \sum_{k=-N/2}^{N/2} g[j]h[j+k]. \tag{3.35}$$

In the resulting function a peak will arise at a lag k, where the similarity between the two samples is most pronounced. According to the cross-correlation theorem, it holds:

$$\mathcal{F}\{(g \star h)(\tau)\} := \mathcal{F}\{g(\tau)\}^* \mathcal{F}\{h(\tau)\}, \tag{3.36}$$

where $\mathcal{F}\{g(\tau)\}^*$ denotes the complex conjugate of the Fourier transformed function $g(\tau)$. An efficient numerical calculation is possible using a Fast Fourier Transformation algorithm, which has a runtime complexity of $\mathcal{O}(Nlog(N))$. The filter uses the result of cross correlation between the incoming data stream from the sensors and a predefined bipolar signal to determine the similarity. As a reference signal, a bipolar pulse is used according to the one that is produced by

a 10^{20} eV shower at a distance of 300 m perpendicular to the shower axis. If the data contains a signal similar to a bipolar pulse, a peak arises at the lag τ (k, in the discrete case) in the resulting function of the cross correlation. A threshold is applied to this function, which is comparable to the one described above (cf. Sec. 3.4.2). The peak with the highest amplitude above the threshold marks the centre of the sample of 640 samples equivalent to 2560 µs that is selected for storage.

3.4.4 Coincidence

After a filter was activated by a signal within the data of the sensors, a coincidence test is performed. Three types of tests were implemented. On the first level, the number of coinciding signals for the six sensors of one storey is checked. The time window for the coincidence within a storey is of about one millisecond, which corresponds to the longest travel time of an acoustic wave through the storey. Then a coincidence criterium for the number of storeys (cf. Fig. 3.19) can be set. Finally, a test between the storeys of the two lines can be performed. For the coincidence test between the storeys within the line or between the lines, the window size is limited to the length of the time-slice of recorded data that is sent to shore, which is 104.864 ms for technical reasons, corresponding to about 150 m. As the maximal separation of acoustic storeys in the AMADEUS detector is about 350 m, the recorded waveform can be distributed between two time-slices depending on the incident angle of the signal. In this case, it is not possible to bring them into coincidence. The number of coinciding sensors or storeys is configurable. The current setup requires that at least four sensors in one storey and at least two storeys have responded to one of the filters described above.

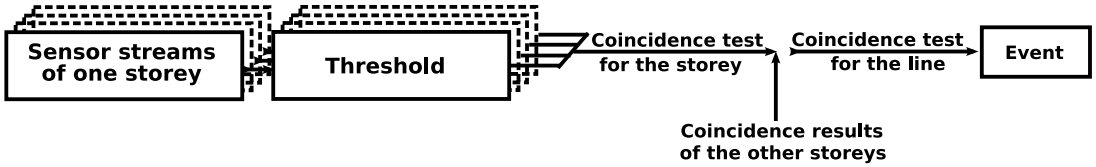


Figure 3.19: Overview of the coincidence test used in the AMADEUS on-line filter system.

3.4.5 Filter efficiency

The efficiency and purity of the filter system were analysed with simulated data according the simulation chain as described. For this analysis, bipolar pulses from point sources were produced with an initial pressure amplitude at the source of $50 - 10^5$ Pa 1 m as a uniform distribution in a volume of $10 \times 10 \times 2.5\,\mathrm{km}^3$. The signal-to-noise ratio (S/N) is defined as:

$$S/N = \frac{A}{4\sigma_{noise}}\,, \tag{3.37}$$

where A is the maximal amplitude of the BIP and σ_{noise} is the standard deviation of the noise sample in the time domain[8]. The efficiency and purity of a filter are defined as:

$$\text{efficiency} = \frac{N_{\text{filtered}}}{N_{\text{generated}}} \tag{3.38}$$

$$\text{purity} = \frac{N_{\text{true}}}{N_{\text{filtered}}}, \tag{3.39}$$

where $N_{\text{generated}}$ is the number of simulated signals, N_{filtered} is the number of signals selected by the filter, and N_{true} is the number of correctly selected signals. The correctness of a selected event is validated by comparing the expected arrival time at a given sensor with the arrival time of the selected signal. This difference needs to be in a time window of 128 µs. Fig. 3.20 shows the resulting efficiency and purity of the AMADEUS on-line filter as used in the experiment. The scaling factor, as described in Sec. 3.4.3, used for the threshold of the pulse shape recognition filter is set to 4.5. The coincidence requirements as mentioned above are used. The efficiency reaches 100 % for a S/N greater than 1.5 and the purity is at a S/N of 1 at 100 %. This high purity is mainly caused by the strong requirements of the coincidence test.

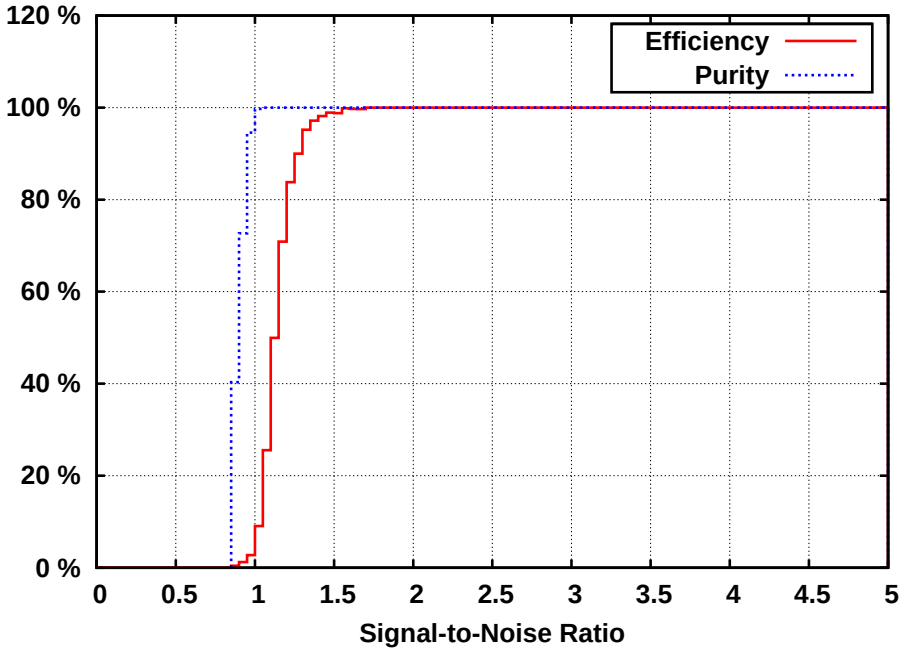


Figure 3.20: Simulation result for the filter efficiency (red, solid) and purity (blue, dashed) for bipolar pulses as function of the signal-to-noise ratio.

[8]This definition was chosen so that the maximal value taken from a noise sample and used for A has an S/N of about one.

Chapter 4

Analysis Chain

Contents

In this chapter, the analysis chain is described in detail. It is comprised of several reconstruction techniques: The determination of the arrival time of a signal, reconstruction of its incident direction and acoustic source position; and signal classification strategies to identify whether the signal is bipolar or not (cf. Fig. 4.1). The simulation chain was used to study and characterise the signal classification and reconstruction algorithms and their performance.

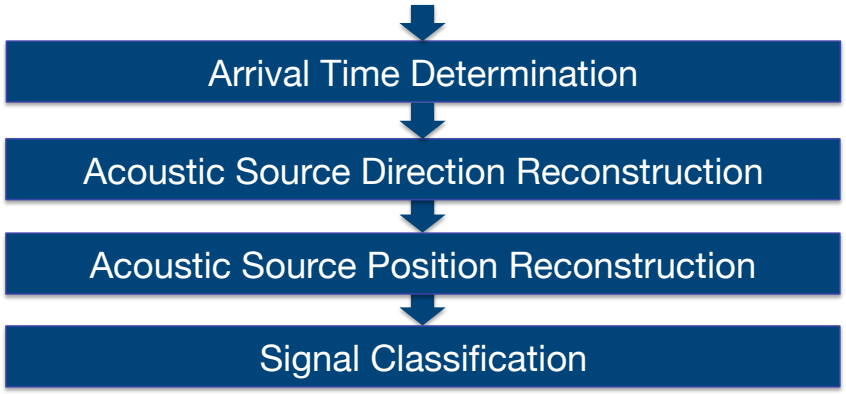


Figure 4.1: Overview of the different modules of the analysis chain.

4.1 Determination of the time of arrival

The determination of the time of arrival of a signal is crucial for the direction and position reconstruction of an acoustic source. Two types of signals are of prime interest for the experiment: BIPs, because they are the acoustic signature of a UHE neutrino interaction, and the signals emitted by the ANTARES acoustic positioning system. The characteristics of these two types of signals are too different to cover them with one algorithm. In a first step, the waveform sample that was selected by the on-line filter system is up-sampled based on an outlier-robust spline interpolation [65] to a sampling rate of 1 µs. The different types of sensors have different sensitivities, so, in order to handle the signals form these sensors in a uniform way, the samples are normalised to their maximum amplitude.
For the BIPs, the arrival time is determined by a cross-correlation with a pre-defined bipolar pulse. According to the properties of the cross-correlation, the time difference between the pre-defined pulse and the pulse contained in the sample can be obtained (cf. Sec. 3.4.3).
For the signals of the ANTARES acoustic positioning system, the envelope of the sample is formed, and a threshold on the envelop's amplitude is applied to determine the arrival time. The envelope is created by rectifying the waveform and smoothing it by a low-pass filter with a cut-off frequency of 50 kHz. These two procedures achieve a precision of about 1 µs for the corresponding signal types (cf. Fig. 4.2).
The pre-selection made by the on-line filter is not ideal. Thus, it is possible to select a signal formed in the ambient noise, a reflected signal, or an extraneous signal that may have originated from a source other than the actual source of interest. To minimise the effect of these coincidences, a causality check is performed for the arrival times within a local sensor cluster (acoustic storey, cf. Sec. 2.5.1). This tests whether the time differences determined for the sensors are smaller than or equal to the maximum travel time required for a sound wave to pass through the storey. Outliers are removed from the dataset that is used for further analysis.
For high-amplitude signals, such as the emissions of the acoustic positioning system, a waveform distortion — so-called clipping — can occur when the amplifiers involved in the DAQ electronics are overdriven and attempt to deliver an output voltage beyond their maximum capability. This can make it difficult to determine the arrival time consistently. For example, this can occur for the emissions of the positioning system.

4.2 Acoustic source direction reconstruction

The direction reconstruction [66, 67] of the incoming signal is required for the subsequent reconstruction of the acoustic source position, as described below. In the AMADEUS detector, local clusters of sensors, the so-called acoustic storeys (cf. Sec. 2.5.1), are used with sensor spacings of the order of 1 m. The direction reconstruction is performed for each storey within the AMADEUS detector, and is based on a least square fit. The difference of the arrival time between two of the sensors within an acoustic storey is compared to the expected time difference for an assumed incident direction:

$$\min \left\{ \sum_{i=1}^{N} (\Delta t_{measured_i} - \Delta t_{expected_i}(\theta, \phi))^2 \right\}, \tag{4.1}$$

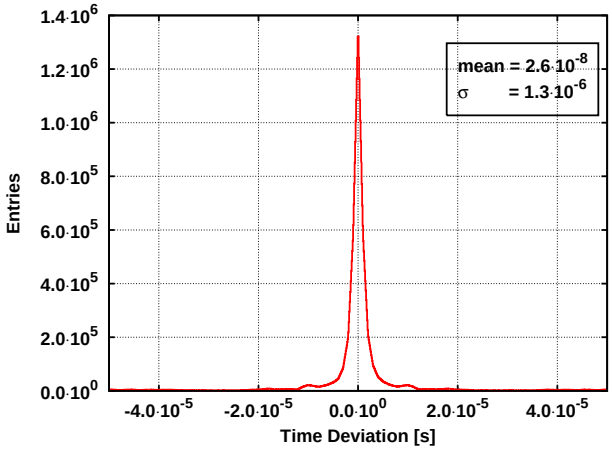
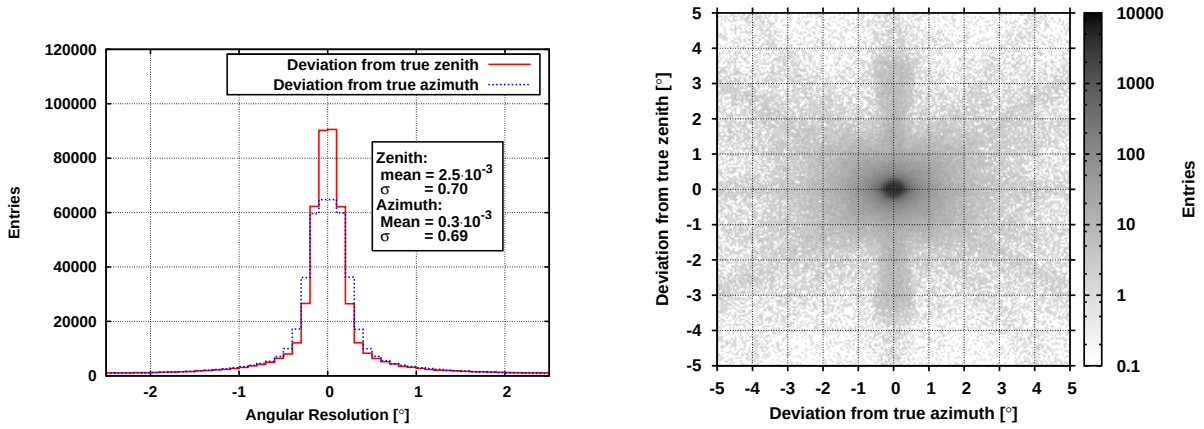


Figure 4.2: Difference between time of arrival (ToA) as reconstructed and simulated: The distribution of arrival times has a mean of about 0.01 µs and a sigma of about 1 µs.

where $i \in 1..N$ is i-th pair of sensors out of N combinations (without repetition) within a cluster of sensors, Δt is the difference of the arrival time, and θ and ϕ are the zenith and azimuth angle, respectively. The expected difference of the arrival time is calculated in-place from the known sensor positions in the cluster for a given detector geometry. The current implementation of this method uses a simplex minimisation algorithm [68], which normally converges after a few tens of iterations. The Pearson correlation coefficient r is used as an estimate of goodness of the least squares fit. A cut on r can be set to reduce the number of mis-reconstructed directions, which are due to possible outliers in the determination of the arrival time. In practice, $r > 0.999$ is required.

For the determination of the angular resolution of the direction reconstruction, transient acoustic sources were generated in a cube that measured of $5 \times 5 \times 2.5\,\mathrm{km}^3$ around the centre of the AMADEUS detector using the simulation chain. As shown in Fig. 4.3(a), the angular resolution obtained with this method is centred around zero, and the standard deviation of the distribution is about 0.7° for both zenith and azimuth angle, respectively. The resolution of the reconstruction is related to the uncertainty of the determination of the arrival time of about 1 µs. The tails in the distribution of the zenith and azimuth angles (cf. Fig. 4.3(b)) are due to the regular build acoustic storey, as it introduces symmetry effects. These effects cannot be excluded by an r-cut, so further consistency checks are required for the position reconstruction of an acoustic source, as described in the next section.

In order to evaluate and validated the performance of the direction reconstruction on experimental data, the emissions of the acoustic positioning system of the ANTARES neutrino telescope were used. At the bottom of each detection line in the ANTARES detector, where is an acoustic transducer that emits a defined signal with a characteristic frequency every 2 min (cf. Sec. 2.4). The positions and emission times of these emitters are known. Line 6 and Line 9 were not operational, and Line 10 was only partly operational at this time. The emitters on Line 4 and Line 12 were malfunctioning, and thus they were only partly operational due to communication problems with the corresponding hardware that controls the emissions. The distribution of the reconstructed directions of these events is shown in Fig. 4.4 for the three storeys equipped with HTI hydrophones (Storey 2, Storey 6, and Storey 22). The HTI hydrophones are suited best for this study, because there are only small differences in their



(a) Angular resolution

(b) 2d angular resolution

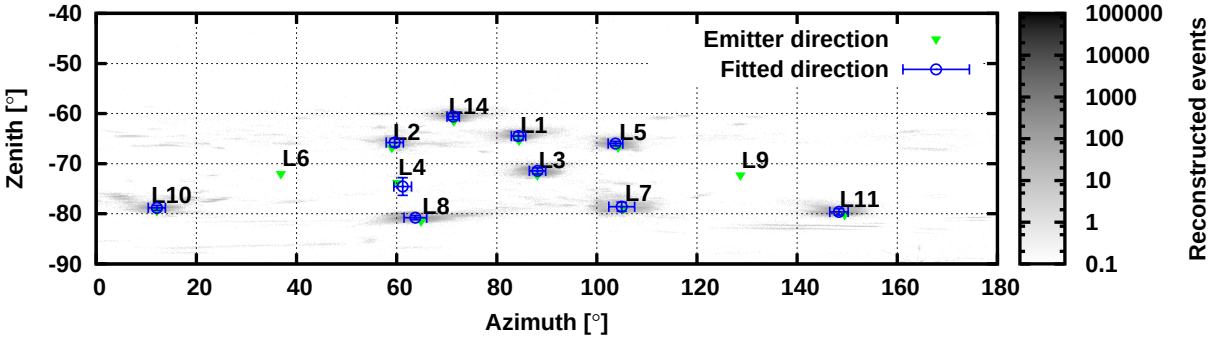
Figure 4.3: (a) Angular resolution of the direction reconstruction algorithm shown for the zenith and azimuth angles. The mean of the distribution is around zero, and the sigma is about 0.7° for both the zenith and azimuth angles. (b) The angular distribution is shown as 2D-histogram. The ray like pattern results from the symmetry effects introduced by the regular build acoustic storeys.

characteristics. For each of the storeys and each of the acoustic emitters, a two-dimensional, Gaussian distribution was fitted to the corresponding direction distribution. Lines 1 and 2 have a low number of reconstructed directions for Storeys 2 and 6 on the IL07. As these lines are the closest lines to the IL07, the probability that waveform clipping can occur increases, and this results in an insufficient determination of the arrival time. The resulting mean reconstructed direction and its uncertainty are given in Fig. 4.4 as blue markers and the known position of the emitters are given as green triangles. The resolution of the direction reconstruction is calculated as the mean uncertainty of the azimuth and zenith angles, which are $1.6 \pm 0.2°$ and $0.6 \pm 0.1°$, respectively. The uncertainty of the zenith angles is in agreement with the simulation result of 0.7°. For the azimuth angle, the uncertainty is dominated by the uncertainty of the positioning of the storey within the detector. The positioning is conducted every 2 min, but rotational movements of the storey around the z-axis of the order of 1° can occur on smaller time scales, which were measured with the Compass board (cf. Sec. 2.5.3) included in each storey [69]. This results in a higher-than-expected uncertainty of the reconstruction of the azimuth angle.

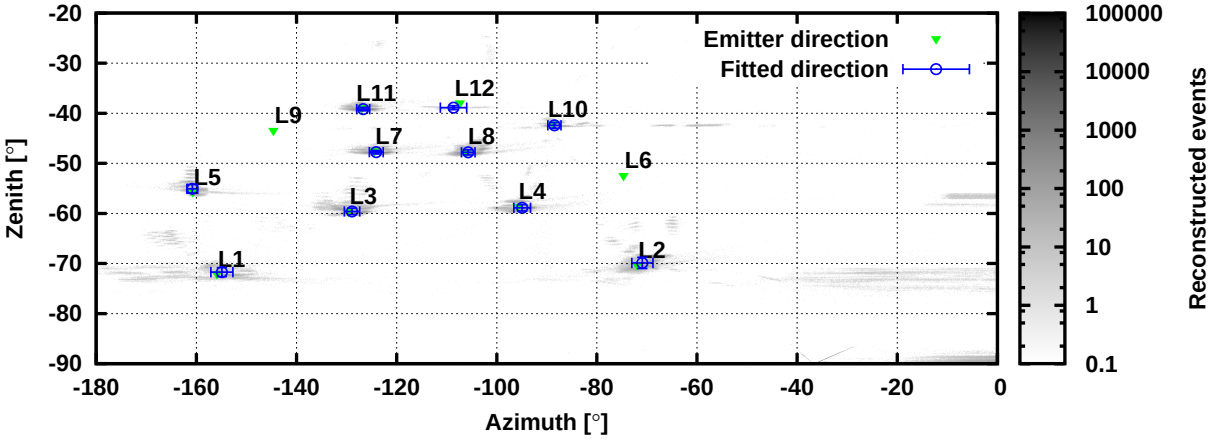
4.3 Acoustic source position reconstruction

The position reconstruction of the acoustic source uses the previously determined direction of the incoming signal [66,67]. If the directions were reconstructed for at least two of the acoustic storeys, a simplified ray tracing technique is performed. A search is conducted to identify the intersection point or the nearest approach of the rays, starting at the storeys and pointing in the reconstructed direction. The algorithm employed[1] minimises the distance L between the

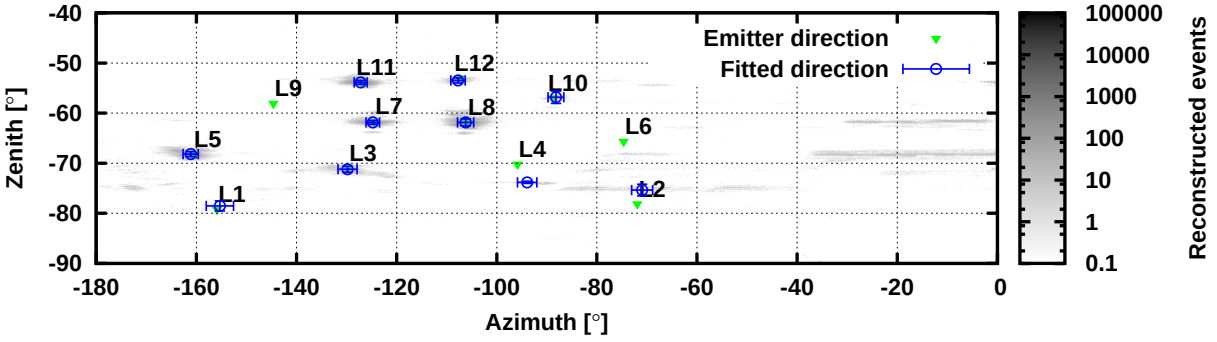
[1] The implementation uses a simplex minimisation algorithm [68].



(a) Line 12 Storey 22



(b) IL Storey 2



(c) IL Storey 6

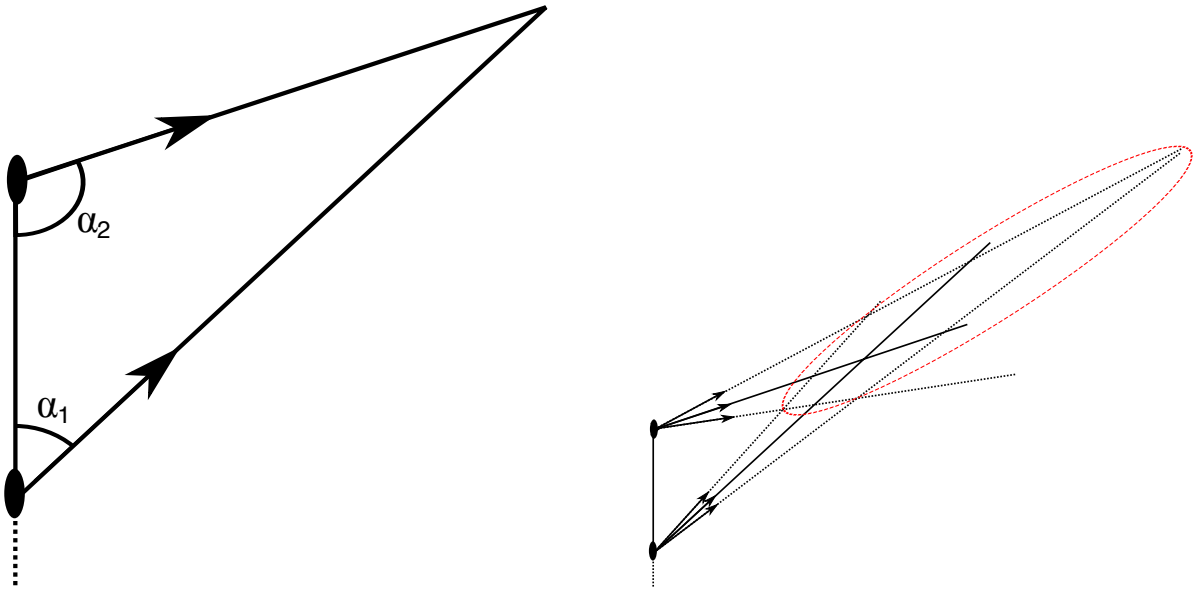
Figure 4.4: The distribution of direction reconstructed events of the ANTARES positioning system emitters is shown as recorded for the three storeys 22, 2, and 6. Also given is the direction, under which the emitters are located (green) and the fitted direction (blue).

rays and a proposed source position $\mathbf{x}$:

$$\begin{array}{ll} \min & \left\{\sum_{i=1}^{N} L_i(\mathbf{x})\right\} \\ \min & \left\{\sum_{i=1}^{N} |\mathbf{x}-\mathbf{r}_i|^2\right\} \\ \min & \left\{\sum_{i=1}^{N} |\mathbf{x}-(\mathbf{r}_0+\mathbf{e_p}d)_i|^2\right\}\,, \end{array} \tag{4.2}$$

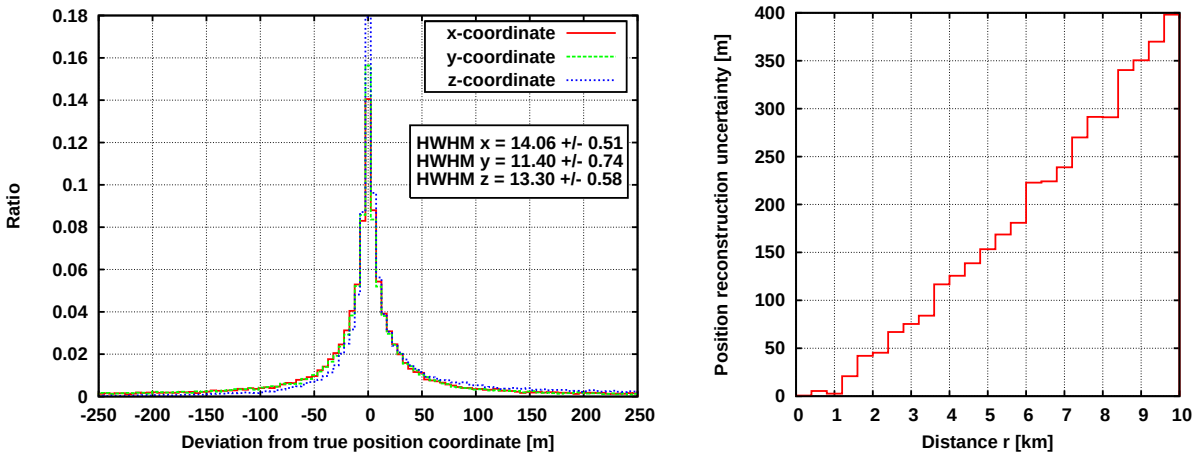
where $i \in 1..N$ is the i-th ray, $\mathbf{r} = \mathbf{r}_0 + \mathbf{e_p}d$ represents the current position of the ray, where r_0 is the position of the cluster, $\mathbf{e_p}$ is the unit vector of the starting direction, and d is the distance between the sensor cluster and the proposed source position $\mathbf{x}$. In this approach, straight rays are used, not including the depth-dependent speed of sound profile expected in the Mediterranean Sea (cf. App A).
As for the arrival time and direction reconstruction, a test of the input parameters, in this case against mis-reconstructed directions, is implemented. This cross-check is based on the requirement that the rays from two different storeys must have at least a nearest approach to each other so that it is closer than the distance between these storeys. For a possible intersection point, any two rays must fulfil the following requirements: First, the sum of the angles α_1 and α_2 that are defined as the angle between the direction vector and the connection line between the storeys, respectively, have to be less than 180° (cf. Fig. 4.5(a)). Second, the angle β is defined as the angle between one of the direction vectors and the plane spanned by the other direction vector and the connection line between the storeys. This angle β should be zero, a tolerance of $\pm 2°$ between the ray and the plane is allowed for.
In Fig. 4.6(a), the distributions of the deviation of the x, y, and z coordinates of the reconstructed position from the true position of the vertex as set by the simulation are shown. The acoustic sources for this study were simulated in a cube of the size of $5 \times 5 \times 2.5\,\mathrm{km}^3$ around the detector centre. The *half width at half maximum* (HWHM) of the distributions is better than 15 m for each coordinate. The position reconstruction depends on the precise knowledge of the position of the acoustic storey and the reconstructed direction. In the simulation, the positions of the storeys are fixed. The uncertainty of the direction reconstruction, which applies for each of the storeys, affects the ray tracing used for position reconstruction, as sketched in Fig. 4.5(b). The influence of the direction reconstruction's uncertainty increases with distance, as the rays become almost parallel. This results in a distance-dependency of the mean uncertainty of the position reconstruction, which increases with distance, as shown in Fig. 4.6(b).
As for the direction reconstruction, the performance of position reconstruction was tested with the emitters of the acoustic positioning system. The mean position and the uncertainty were derived by fits of Gaussian distributions to the corresponding distributions of the reconstructed x, y, and z coordinates for each of the emitters. The results are given in Fig. 4.7. As mentioned in the previous section, the emitters of Line 6, 4, 9, and 12 were either not operational or only partly so at this time. The mean deviation of the mean reconstructed position and the true position of the emitters are:

$$\begin{aligned} \langle x \rangle &= 0.8 \pm 0.4\,m \\ \langle y \rangle &= 1.0 \pm 0.4\,m \\ \langle z \rangle &= 1.1 \pm 0.1\,m\,. \end{aligned}$$



(a) Geometric cross-check (b) Effect of the direction uncertainty

Figure 4.5: (a) A Sketch of the geometric cross check to determine whether the rays starting from the storeys (black dots) are able to intersect (not to scale). (b) A Sketch of the effect of the direction reconstruction's uncertainty on the ray starting tracing technique (not to scale). For an acoustic source with fixed position, which is constantly emitting, the reconstructed positions would lie in this oval shape pointing in the direction of the source. The further away the source is, the greater the extent to which the shape is stretched out.



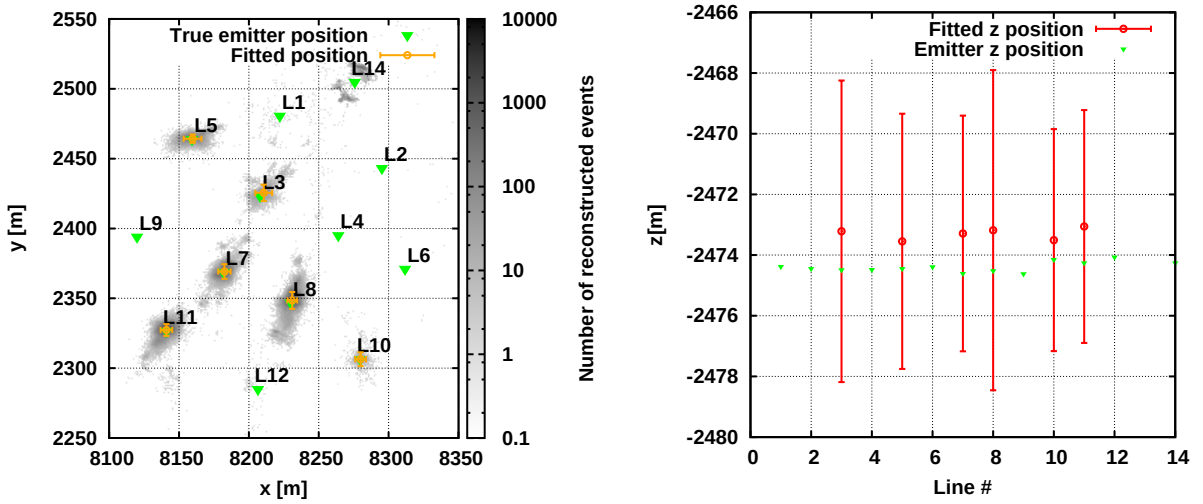
(a) Position resolution (b) Uncertainty dependency

Figure 4.6: (a) Distributions of the deviation of the x, y, and z coordinates of the reconstructed position from the true position of the vertex. The acoustic sources were generated in a cube with the dimensions of $5 \times 5 \times 2.5\,\mathrm{km}^3$ around the detector centre. (b) Dependency on the uncertainty of the position reconstruction with the distance r between source and detector.

The resolution of the position reconstruction derived for these emissions are:

$$\begin{aligned} \langle \sigma_x \rangle &= 4.7 \pm 0.5\, m \\ \langle \sigma_y \rangle &= 4.9 \pm 0.5\, m \\ \langle \sigma_z \rangle &= 4.3 \pm 0.4\, m\,. \end{aligned}$$

The reconstructed position of the emitters is in agreement with their true position and the resolution is better than 5 m for each coordinate. As noted above, the uncertainties of the position reconstruction depend on the distance to the acoustic source. So this result is only valid in the range of a few 100 m and it is comparable with the resolution derived from the simulation, which is, for x, y, and z, 3.3 ± 0.1 m, 3.0 ± 0.1 m, and 4.2 ± 0.1 m, respectively. As mentioned in the previous section, the movement of the lines with the sea current and the resulting uncertainties have not yet been included in the simulation. Thus, the resolution derived from the simulation is generally smaller than that from the emissions of the positioning system.



(a) Position reconstruction x,y-coordinate

(b) Position reconstruction z-coordinate

Figure 4.7: (a) Distribution of the x and y coordinate of the reconstructed positions. Also the fitted position and the uncertainty are given (orange marker) and the true position of the emitters also are given (green markers). (b) The fitted z position of the emitters (red markers) and the their uncertainty are shown and the true z-coordinate of the emitters is given.

4.4 Signal classification

In the deep sea, the background of transient signals is very diverse. Approaches such as matched filtering, which was used for the on-line filter, are insufficient to distinguish between neutrino-like signals and other transient signals with similar waveforms that form the acoustic background for neutrino detection in the deep-sea acoustic environment. A classification strategy [70] based on machine learning algorithms was analysed with the goal of determining a robust and effective way to perform this task. This study incorporates the complete simulation

chain, as described in previous chapter, feature extraction techniques, and the signal classification based on machine learning algorithms. For a well-trained model, a recognition error on the level of a few percent is achieved for strong classifiers, such as Random Forest and Boosting Trees, using the extracted features of the signal as input and utilising dense clusters of sensors instead of single sensors.

4.4.1 Feature extraction

As a first step in the feature extraction process, the simulation of the AMADEUS on-line filter system (cf. Sec. 3.4), which is equivalent to the one installed in the experiment, is used to pre-select signals in order to maintain a sufficient signal-to-noise ratio. In particular, the pulse shape recognition filter, which is based on the matched filtering technique, is used to pre-select bipolar pulses. A bipolar pulse, equivalent to that expected to be produced by a 10^{20} eV shower at a distance of 300 m perpendicular to the shower axis (cf. Sec. 3.1.3), was used as the reference signal for the matched filter. In the next step, the characteristics — the features — of the pre-selected signals are extracted. The resulting feature vector contains the time and frequency domain characteristics of the signal as well as the results of a matched filter bank, which was tuned for neutrino-like signals. This bank consists of six reference signals that correspond to angles of $0^\circ - 6^\circ$ in steps of one degree with respect to the plane perpendicular to the shower axis of a 10^{20} eV shower at a distance of 300 m. This plane intersects the shower axis at the position of the shower maximum. In the time domain, the following features are extracted from the waveform:

- The number of positive and negative peaks above a threshold given by the RMS of the sample multiplied by a constant factor of 4.5 is extracted.
- The total duration of the signal is defined as the time above this threshold.
- The peak-to-peak amplitude between the largest positive peak and the largest neighbouring negative peak and the corresponding asymmetry and duration, which are used as feature.

In the frequency domain, the main frequency component and its spectral power are extracted from the PSD of the sample. Furthermore, the number of spectral peaks is used as a feature. From the results of the matched filter bank, the best match that with the largest peak is taken into account. From this matched filter output the number of peaks and the amplitude, width, and the integral of the largest peak are stored in the feature vector.
As an independent feature vector, the waveform selected by the on-line filter itself can be subjected to the classification algorithms. The standard waveform sample has a duration of 640 samples equivalent to 2560 µs, which is extensive as a feature vector. Thus, only 128 samples around the largest peak were used.

4.4.2 Classification algorithms

The classification strategy stems from machine learning algorithms that have been trained and tested with data from the simulation [70]. Either the extracted feature vector or the waveform selected by the on-line filter itself can be used as input; As output, either binary class labels

(bipolar or not) or multiple class labels (one for each signal type in the simulated data) can be predicted. It should be noted that some of all the algorithms that were used are restricted to binary class labels as output[2]. So, all algorithm are operated with this output mode. The same training and testing datasets are used as input to the different classification algorithms under consideration. The simulation of transient signals, as described in Sec. 3.2.1, was used to create these datasets, including the four types of signals: bipolar pulse, multipolar pulse, sine wave, and strong Brown noise, which were equally distributed. The predictions of the signal type made by the algorithms for the individual sensors are combined to create a new feature vector, which is used as input in order to train and test classifiers for the acoustic storeys. Thus, a prediction of the signal type can be made for each of the acoustic storeys in the detector. In the final step, the resulting predictions for the storeys are combined to create a response of the detector by majority vote. The data taken with the detector is pre-selected by the on-line filter system to maintain a sufficient data reduction and signal-to-noise ratio. This also implies that not every sensor within a storey or every storey within the detector will have recorded data for a given signal. Thus, it is necessary to deal with missing data in the input of the classification algorithms for the storeys and the detector on the whole. All of the algorithms that were studied, can handle missing data within the feature vector. The following algorithms [71] have been researched for individual sensors and clusters of sensors[3]:

Naïve Bayes: This simple classification model is based on applying the Bayes' theorem and assuming that the features are conditionally independent of one another for each class [72]. For a given feature vector $\mathbf{f}$ containing N features, the independent class c is selected using probabilities acquired from the training data. Using the Bayes' theorem, the conditional probability for class c given a feature vector $\mathbf{f}$ can be written as:

$$p(c|\mathbf{f}) = \frac{p(c)\ p(\mathbf{f}|c)}{p(\mathbf{f})}. \tag{4.3}$$

Since the feature vector $\mathbf{f}$ is known, the probability $p(\mathbf{f})$ is a constant scaling factor s. So the numerator is equivalent to a joint probability model, which can be reformulated using the chain rule for repeated applications of the definition of conditional probability:

$$p(c, \mathbf{f}) = \frac{1}{s} p(c) \prod_{n=1}^{N} p(f_n|c, f_{n-1}). \tag{4.4}$$

Under the "naïve" assumption that a feature $f_i \in \mathbf{f}$ is conditionally independent from all other features $f_j \forall j \neq i$, it follows that:

$$p(f_i|c, f_j) = p(f_i|c)\,,\ \forall i \neq j. \tag{4.5}$$

Using this, Eq.4.4 can be rewritten as:

$$p(c, \mathbf{f}) = \frac{1}{s} p(c) \prod_{n=1}^{N} p(f_n|c). \tag{4.6}$$

[2]Although multiple class labels are not included as output in this study, initial tests showed that this output mode is less effective than the binary mode by more than factor of two. Also the main concern is on classifying bipolar pluses, which renders the correct identification of the other signal types less important

[3]The descriptions of the algorithms given are summaries of the basic concepts, please refer to the references given for more details.

So, a classifier can be constructed from this result by introducing a decision rule that selects the most probable hypothesis.

Decision Tree: This classification model stems from a tree-like structured set of rules [73]. Starting at the root, the tree splits up at each node based on the input feature that has not already been used as an ancestor node and provides the optimal decision rule at this node. The two criteria that are most commonly used to determine the relevance of a feature, and thus to determine this optimal rule, are the information gain and the so-called Gini impurity. The information gain, as defined in information theory, measures the difference in information entropy caused by the split. The Gini impurity indexes the change of the purity of the outcome that arise, when a split is relayed on one feature in contrast to another [74]. Typically, the best split is based on one feature, but, in order to handle missing data, additional features can be used to split — often referred to as surrogate splits. While training a Decision Tree, it is probable that the number of splits inflates and results in a sizeable complex tree. Thus, the tree is evaluated from the leaves, which are representing the class labels, to the root. Redundant paths, which are less likely to improve the performance, are removed. The path from the root of the tree to one of the leaves is used as a classifier and defines one decision rule.

Random Forest: A Random Forest [75] is a collection of Decision Trees. The classification works as follows: The Random Forest takes the input feature vector, makes a prediction with every tree in the forest, and outputs the class label that received the majority of the votes. The trees in the forest are trained with randomly sampled sets, including replacements from the original training data. These subsets have the same size as the original dataset. Statistically, about one-third of the feature vectors in the training data are left out of these sets. This so-called *out-of-bag* (oob) data are used to estimate the training error in-place. The trees use a fixed number m of input variables for splitting at a node. The input variables are chosen at random from all features $M \gg m$. The Decision Trees in the ensemble are trained until they are fully developed and they are not pruned. Pruning is not required because the "strong law of large numbers" shows that Random Forests always converge, meaning that over-fitting does not occur.

Boosting Trees: Boosting Trees combine the performance of many so-called weak classifiers to produce a powerful classification scheme [76]. A weak classifier is only required to be better than a random decision. However, many of them, when smartly combined, result in a strong classifier. Decision trees are used as weak classifiers in this boosting scheme. In an iterative process, the Decision Trees are trained with the feature vectors, while maintaining weights corresponding to each of the feature vectors. Initially, all weights are set to have the same value, but, during an iteration, the weights of incorrectly classified vectors are increased so that the classifier is forced to focus on the feature vectors in the training set that are complicated to predict. The training of the trees is "boosted". The final decision is made by a majority vote of the weighted outcome of the Decision Trees. This weight is calculated from the training error of the corresponding Decision Tree. In contrast to a Random Forest, the Decision Trees are not necessarily full-grown trees, since, typically, they consist of only a few nodes.

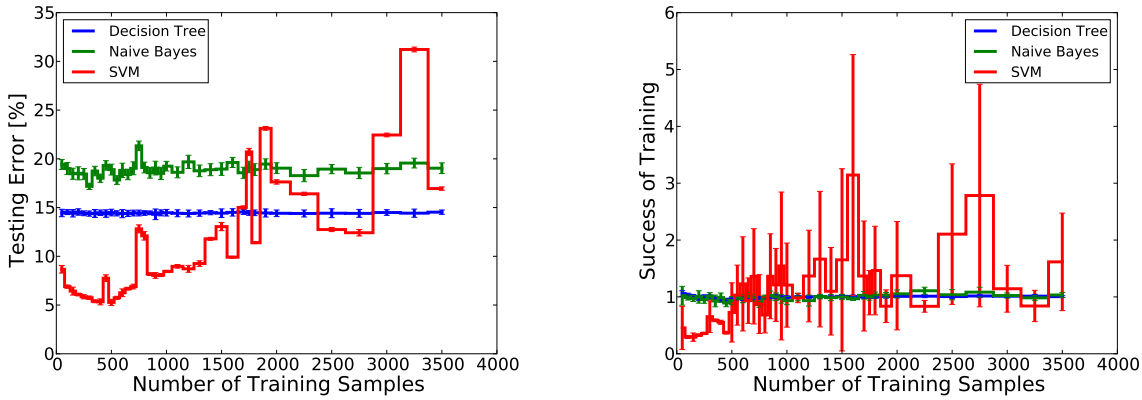
Support Vector Machine: A SVM [77,78] maps the feature vectors into higher-dimensional feature space using a kernel function, which defines the inner product in this feature space. A hyper-plane is searched so that the margin between this hyper-plane and the nearest feature vectors from both of the two labels of a binary class is maximal. The feature vectors that have the smallest margin to the hyper-plane are called support vectors. The positions of the other feature vectors in the feature space does not affect the hyper-plane and thus the decision rule. For the prediction of a feature vector, the margin of the hyper-plane that separates the class labels is calculated, and the sign of the margin determines the corresponding class label.

Over-fitting occurs when a classification model becomes sensitive to random noise instead of the underlying relationship. In general, over-fitting is found for models that are excessively complex, e.g., an unreasonable number of features relative to the number of observations is used or the observations used as training set have excessively specific characteristics such that a generalisation is problematic to accomplish. A model that has been over-fit has poor predictive performance, because it can exaggerate minor fluctuations in the data. Most classifiers have incorporated methods to reduce the possibility of over-fitting. Cross-validation is one of these techniques. The training dataset is partitioned into k complementary subsets. One of the k subsets is kept for validation of the model, while the other $k-1$ subsets are used for the training. This procedure is repeated k-times, where each subset is used once for validation. The results can be averaged and used to estimate the training error. This method takes all observations for both training and validation into account. Although most of different algorithms under consideration can calculate the training error in-place, cross-validation is also used to maintain a consistent way of calculating the training error.

4.4.3 Predictive performance

In this section, the performance results of the classification strategy are described. Two indicators are used to measure the performance of the classification: the *testing error*, which is the error of the prediction with respect to the simulation truth including mis-predicted BIPs and false alarms, and the *success of training*, which is the ratio testing error to training error and the ratio indicates whether the model is under-trained (<1) or over-trained (>1). As an overall result, binary class labels as output are more than twice as effective as multiple class labels. The binary class labels are the standard output of the subsequent results that are presented. Weak classifiers, such as Naïve Bias and Decision Trees, have high testing error above 14 % and they are neither more robust against changing ambient noise conditions nor significantly faster than other classifiers (cf. Fig. 4.8(a)). Although the SVM is a strong classifier, its significant numerical complexity and lack of robustness disqualify it for further applications (cf. Fig. 4.8(b)). Thus the most favourable classifiers are Random Forest and Boosting Trees. In addition, the usage of clusters shows a substantial improvement over individual sensors. Random Forest and Boosting Trees are robust and stable, and they produce well-trained models. The elapsed time for processing one event is less than a second. For the individual sensors and the extracted features as input, a testing error of about 5 % for the Boosting Trees and for the Random Forest of about 10 % is achieved, which is further improved by combining the sensors into clusters, similar to an acoustic storey in the AMADEUS detector. The testing errors are well below 2 % for successfully trained models (cf. Fig. 4.9(a) and Fig. 4.9(b)). Using the extracted

waveform as input yields similar results, the Random Forest achieves a testing error of about 6 % and the Boosting Trees of about 12 %. These errors are also improved, when combining the individual sensors to clusters, resulting in a testing error below 4 % (cf. Fig. 4.10(a) and Fig. 4.10(b)).
An ensemble of the classifiers is used to determine the signal type of an event detected by the AMADEUS system to account for the varying and very diverse transient noise conditions at the ANTARES site. This ensemble of classifiers consists of Random Forest and Boosting Trees classifiers to make predictions for the signals recorded by individual sensors and subsequently to make a prediction for the corresponding acoustic storey. From the latter, the prediction of an event's signal type for the complete detector is derived by majority vote. As the training and testing of the classifiers was performed using the simulation of transient signals, the performance of the resulting ensemble of classifiers was validated by the simulation of the neutrino-induced acoustic signals (cf. Sec. 3.1). The ensemble achieved a testing error of 1.3 ± 0.3 % for the prediction of bipolar pulses on the detector level.

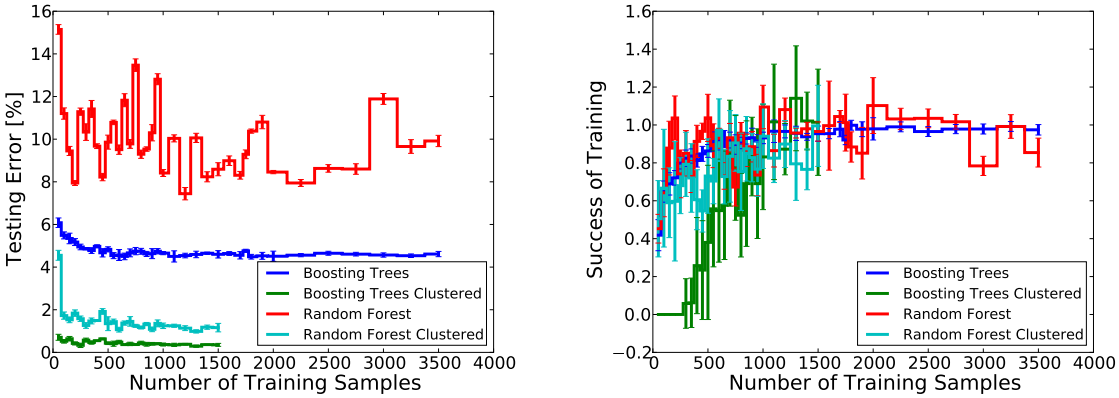


(a) Testing error

(b) Success of training

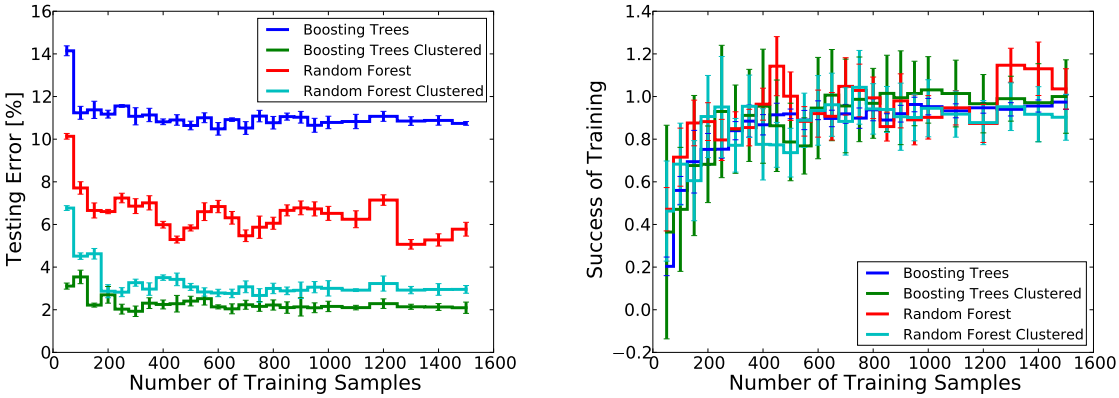
Figure 4.8: (a) The testing error is shown as a function of the training samples for Decision Tree, Naïve Bias and SVM classifiers. (b) The success of the training is shown as a function of the training samples for Decision Tree, Naïve Bias and SVM classifiers. A value of one indicates that the model is well-trained. The extracted feature vector is used as input and the binary class labels are used as output for the individual sensors.

The results show that the use of machine learning algorithms is a robust, effective and efficient classification strategy. The classifiers perform well under different levels of ambient noise and are able to distinguish between bipolar signals — one of the characteristics of the acoustic signature of a neutrino interacting — and other signals, especially to differentiate them from short, multi-polar signals. In Sec. 5.1, the classification strategy is used to perform an analysis of the temporal and spatial distribution of the background of the bipolar signals. For further applications, the strategy must be extended to classify neutrino-like events with all their features, in particular their disk-like spatial propagation. For the AMADEUS detector, the determination of this propagation pattern is not feasible due to the geometry of the detector.



(a) Testing error (b) Success of training

Figure 4.9: (a) The testing error is shown as a function of the training samples for Random Forest and Boosting Trees classifiers. (b) The success of the training is shown as a function of the training samples for Random Forest and Boosting Trees classifiers. A value of one indicates that the model is well-trained. The extracted feature vector is used as input and the binary class labels are used as output for the individual sensors and the clusters of sensors (indicated by "clustered").



(a) Testing error (b) Success of training

Figure 4.10: (a) The testing error is shown as a function of the training samples for Random Forest and Boosting Trees classifiers. (b) The success of the training is shown as a function of the training samples for Random Forest and Boosting Trees classifiers. A value of one indicates that the model is well-trained. The extracted waveform of the signal is used as input and the binary class labels are used as output for the individual sensors and the clusters of sensors (indicated by "clustered").

4.4.4 Density-based spatial clustering

This is an approach to classify signals by their spatial distribution rather than their waveform characteristics. Under the assumption that density-clustered events can be associated with sources of transient background signals, such as ships and marine mammals, these events can be identified and rejected. The *density-based spatial clustering of applications with noise* (DBSCAN) algorithm [79] is used to recognise the density-clustered events within the data. The basic idea is that a point p is direct density-reachable from a point q if the distance d between them is less than a given distance ϵ and that the point q is connected to more than a given number of other points N_{min}. So a minimal number of points is required to form a cluster. The DBSCAN algorithm starts with an arbitrary starting point that has not been visited yet. This point's ϵ-neighbourhood is retrieved, and if it contains a number of points greater than N_{min}, a cluster is started. Otherwise, the point is labeled as noise. This point might later be found in a sufficiently sized ϵ-neighbourhood of a different point and hence be made part of a cluster. The advantages of this algorithm are that only two parameters are needed: The maximal distance between points ϵ and the minimal number of points N_{min} for a cluster to form. The algorithm can discover clusters of arbitrary numbers and shapes. For the analysis of data taken with the AMADEUS detector, this method can be used to identify sound-emitting objects. It is assumed that the sources of the transient background, such as ships and marine mammals, are located at or near the surface of the sea. Since the position reconstruction has considerable uncertainties, the incident direction retrieved from the direction reconstruction is used instead of locating the position of the acoustic source at the surface of the sea. This is done by calculating the intersection point of a straight line pointing into the direction reconstructed for this acoustic storey and the surface of the sea. The DBSCAN algorithm performs an analysis of the clusters of points projected onto the surface of the sea in this manner. The minimal size of a cluster N_{min} is set to three, and the distance $d < \epsilon = 100\,\mathrm{m}$ is calculated as shown:

$$d = \sqrt{\Delta x^2 + \Delta y^2 + (c\Delta t)^2}, \tag{4.7}$$

where Δx, Δy are the differences of corresponding coordinates in the projection onto the surface of the sea and t is the time between events. The time is taken into account for the distance calculation to avoid clusters that contain events that are not necessarily from the same source because of time between them. It is assumed that the sources are constantly emitting and $c = 10\,\mathrm{m/s}$ is the assumed average speed of the object, e.g, a vessel or a sea mammal.

To determine moving, sound-emitting objects from the results of the clustering, straight tracks are assumed and searched for this specific shape of clusters. This search is performed by fitting a straight line to the points within a cluster and evaluating the Pearson correlation coefficient r as an estimate of the goodness of the straight-line model for the given set of data points. So, it is possible to separate moving acoustic sources at or near the surface of the sea from other spatially-clustered events. In Fig. 4.11, the surface projection of signals that have been identified as moving, sound-emitting objects is given. The cluster analysis can be used as a veto for the identification of neutrino-like signals, because the probability of neutrino interactions to form spatial clusters is negligible.

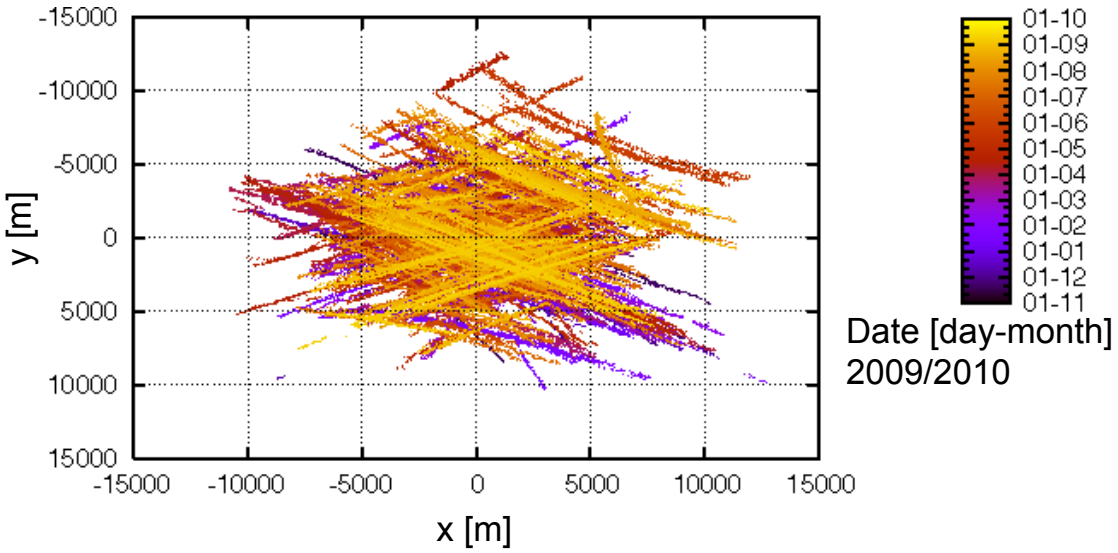


Figure 4.11: The surface projection of signals that had been identified as moving emitters by the DBSCAN algorithm as described in the text. The x and y are the coordinates relative to the detector centre at the surface of the sea. The colour encodes the date. The measurement period was 11/2009 - 09/2010.

Chapter 5

Analysis and Results

Contents

The capabilities of the simulation and analysis chains were discussed in the previous chapters. It was shown that the simulation reproduces neutrino-induced, acoustic signals as well as transient signals, the data acquisition of the AMADEUS detector, and the ambient background noise in the Mediterranean Sea. The analysis chain was presented, which includes techniques for the reconstruction of incident direction of a recorded signal, the reconstruction of the position of acoustic sources, and a classification strategy for neutrino-like bipolar pulses. In this chapter, the calculation of an effective volume and a transient-free, limit-setting potential of the AMADEUS detector are presented. Furthermore, the results of an analysis of the transient background for about 10 month's worth of data are discussed.

5.1 Transient background at the AMADEUS site

In the following, the results of researching the background of transient signals at the AMADEUS detector site in the Mediterranean Sea are presented. As BIP-like waveforms can mimic the acoustic signature of a neutrino and can be produced by various sources, the spatial and temporal distribution of these events is of special interest. Depending on the size and geometry of a future detector, it will be challenging to verify the disk-like shape in which a sound wave propagates when it is produced by a neutrino interaction. In consequence, the background of BIP-like signals is at least partially irreducible and thus a limiting factor for the feasibility of acoustic neutrino detection. For a future, large-scale detector, simulations based on the results

of this background study need to be implemented in order to quantify and optimise the effects of the geometry of such a detector.
This study includes data from 1583 selected runs from 11/2009 to 09/2010, which are about 3754 hours of measurement time. During this period, all six storeys of the AMADEUS detector were basically operational. The following requirements were used to determine the runs that could be included in this analysis:

- The time synchronisation of the detector was working correctly, meaning that the starting time of the run was synchronised between the acoustic storeys (cf. Sec. 2.5.3) and that the time when the run started was recorded. This is required for the positioning of the acoustic storeys.
- The position calibration of the individual acoustic storeys within the AMADEUS detector was available (cf. Sec. 2.5.1).
- The run was not marked as a special run, e.g. for testing hardware or software configurations.
- No failure of the hardware or software involved in the data taking has occurred during the run. The data acquisition of the AMADEUS detector is robust, but occasionally interruptions occur occasionally, resulting in possible losses of data.
- Only the data pre-selected by the on-line filter used for selecting BIP-like signals was used (cf. Sec. 3.4). The data selected by the threshold and minimum bias filters were not considered.

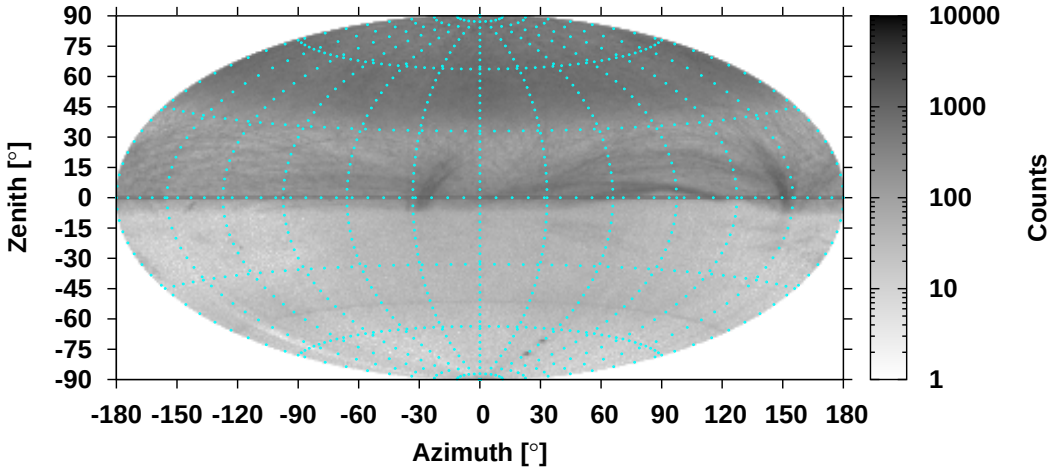
5.1.1 Directional distribution

The directional distribution of transient signals selected by the on-line filter of the AMADEUS detector is analysed, as is the effect of the classification of BIP-like signals. The direction reconstruction is performed for each storey, as described in Sec. 4.2. A cut on the Pearson correlation coefficient selecting reconstructed directions with $r > 0.999$ is applied. This is used to reduce the influence of mis-reconstructed directions, and 31 % of the data for which the direction was reconstructed passed this quality cut. The resulting angular distributions for each storey is shown in Fig. 5.1. The distributions vary between the different storeys due to the different types of sensors used. The storeys equipped with HTI hydrophones have comparable results (cf. Fig. 5.1). The angular distributions of the Storey 3 and Storey 23 have a slightly different behaviour, due to the higher sensitivity of the LTI hydrophones and, consequentially, higher probability that signal clipping distorts the determination of the arrival times. Furthermore, the hydrophones of Storey 22 are mounted in a different way, i.e., pointing downward while the others were pointing upward (cf. Sec. 2.5.1). Storey 21, holding the AMs, have poorer performance due to the alignment of the sensors. They are arranged in the horizontal plane, so their lack of expansion along the z-axis leads to a higher probability of mis-reconstructions due to symmetry effects, especially for the zenith angle (cf. Fig. 5.7, at the end of this chapter).
The main contribution of the transient signals has an arrival direction that points to the upper hemisphere to the extent of about 7° below the horizon (cf. App. A). Taking refraction due to the depth-dependent speed of sound into account, this is in agreement with the assumption

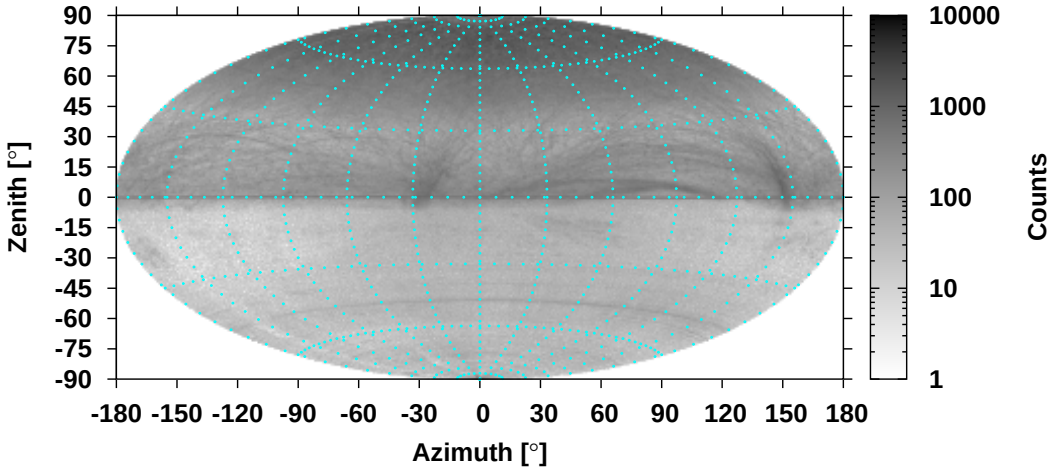
that transient signals are predominantly coming from sources at or near the surface of the sea. In the upper hemisphere, curved structures can be identified, which can be associated with sound-emitting objects moving at or near the surface of the sea, such as ships or marine mammals. The lines converge in spots located at the horizon, where the objects are leaving the surface area accessible by the AMADEUS detector. The direction in which the spots are pointing is correlated to harbours along the coastline of France (Marseille and Toulon) and Corsica.
In Fig. 5.2 the directional distribution of BIP-like signals is shown for Storeys 2, 3, and 22, which are equipped with HTI hydrophones (cf. Fig. 5.8 for Storeys 6, 21 and 23). These signals were selected by the classification strategy described in Sec. 4.4. The number of signals is reduced by more than a factor of 30 with respect to all of the transient signals reconstructed. For the resulting sample, mainly the area above the detector contributes to the distribution of reconstructed directions. The density of the reconstructed events remains high. As the curved lines, which indicate the presence of sound-emitting, moving objects such as ships, are still visible, these objects contribute, as expected, to the rate of transient bipolar pulses that can mimic the acoustic signature of a neutrino interaction. In the following, the spatial distribution of position-reconstructed events is analysed.

5.1.2 Spatial distribution

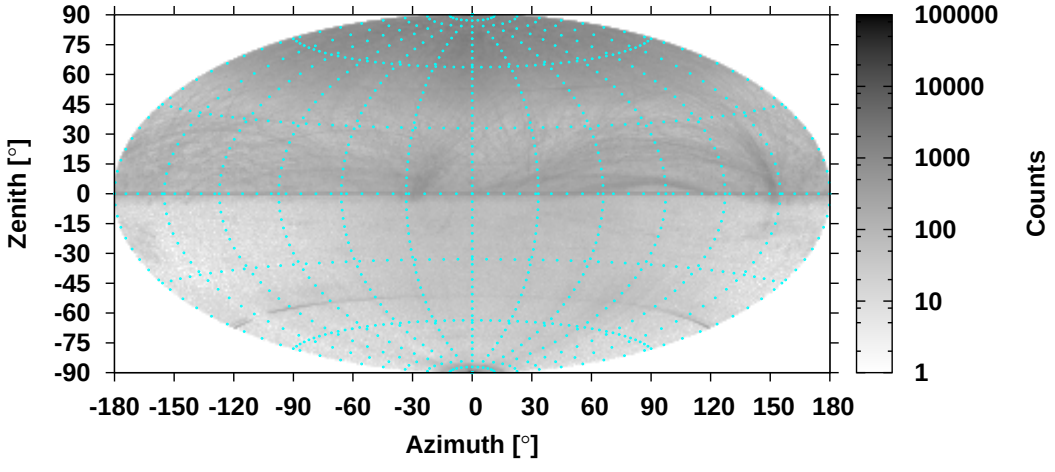
The spatial distribution of transient signals as reconstructed by the acoustic source position reconstruction algorithm is presented. The density of the transient background, especially of BIP-like signals, can be used to identify regions in which the probability of undisturbed detection operations is high. These regions can define a fiducial volume for neutrino searches. For this analysis, the position reconstruction described in Sec. 4.3 was used; For 17 % of all events selected by the on-line filter, the acoustic source position has been successfully reconstructed. In Fig. 5.3, the density of transient signals as a function of the depth z and the distance r from the detector centre is given. In Fig. 5.3(a), the overall density of the transient signals, as reconstructed for the signals selected by the AMADEUS on-line filter, is shown. The signal density is high over the displayed volume of $7.85 \cdot 10^2\,\mathrm{km}^3$. The mean density is about $15 \cdot 10^3\,\mathrm{km}^{-3}\,\mathrm{year}^{-1}$ for the volume given in the figure. Due to the challenges of the position reconstruction described in Sec. 4.3, the reconstructed positions at distances greater than about 1 km from the detector have sizeable uncertainties. Briefly said, the small, two-dimensional detector, in conjunction with the fact that there are small inaccuracies in the angular reconstruction, lead to sizeable uncertainties of the position reconstruction for sources at great distances. If the classification strategy is applied, the density of signals is significantly reduced. Mainly, the volume around the detector contributes, as shown in Fig. 5.3(b). The cluster analysis based on the DBSCAN algorithm, as described in Sec. 4.4.4, can be applied to the data, and the results are shown in Fig. 5.3(c). If the position of the source is assumed at the surface of the sea, a significant part of the event density can be associated with sound-emitting objects that form spatial clusters. The result of this clustering technique can be used as a veto and in combination with the classification. The spatial density of BIP-like events that are not spatially correlated, is shown in Fig. 5.3(d). The rate of these BIP-like signals in the same fiducial volume, as above, is about $100\,\mathrm{km}^{-3}\,\mathrm{year}^{-1}$. For further applications, this rate can be further reduced by taking into account the disk-like propagation pattern of a BIP originating



(a) Line 12 Storey 22

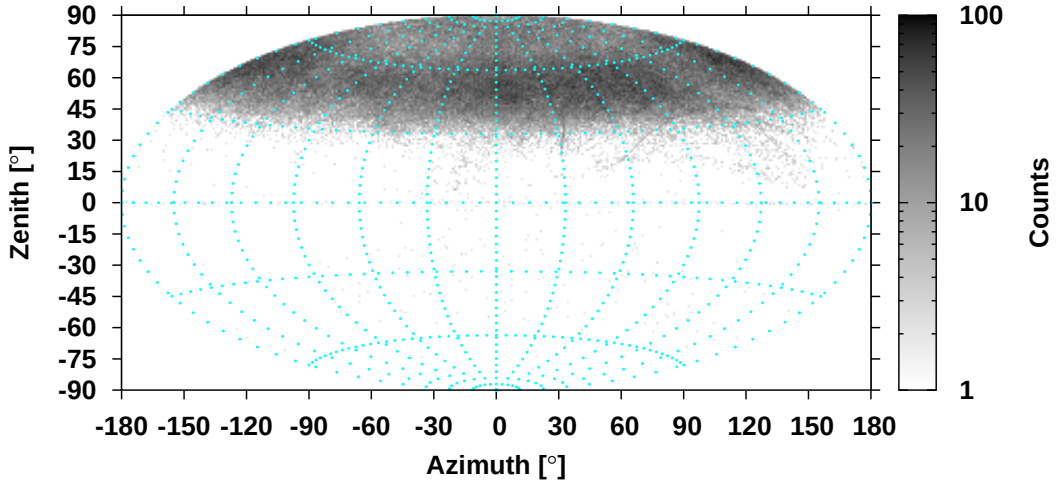


(b) IL Storey 6

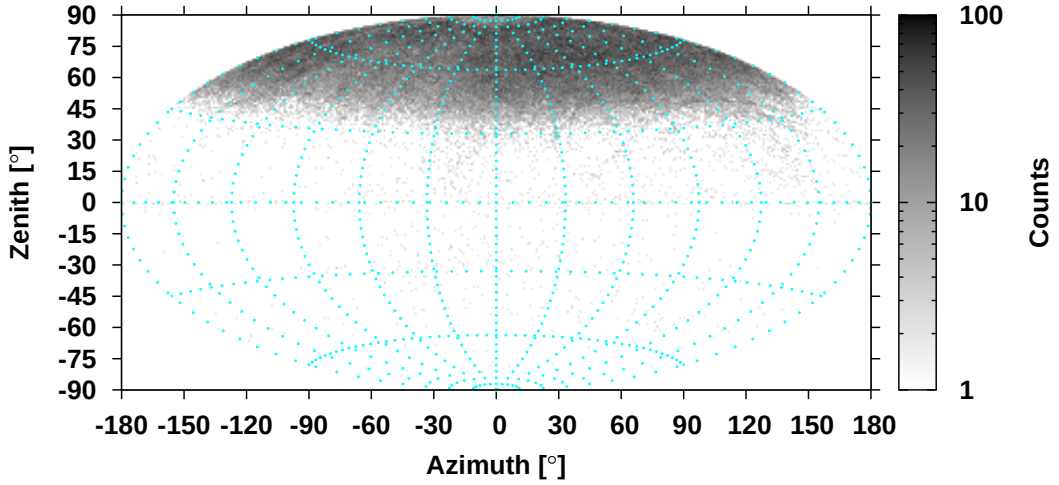


(c) IL Storey 2

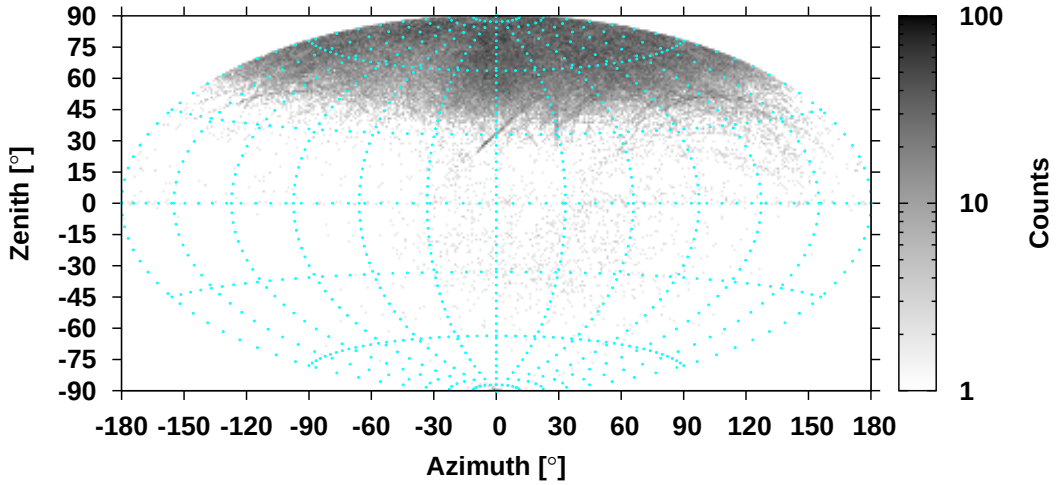
Figure 5.1: Directional distribution of transient signals for Storeys 2, 6, and 22 of the AMADEUS detector: These three storeys are equipped with HTI hydrophones.



(a) Line 12 Storey 22



(b) IL Storey 6



(c) IL Storey 2

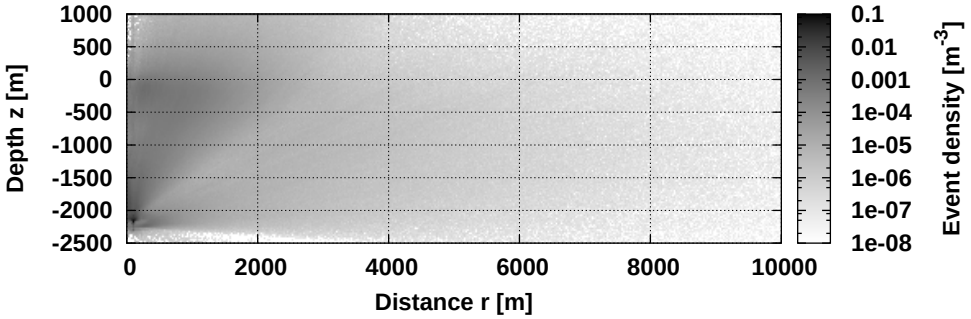
Figure 5.2: Directional distribution of transient BIPs as selected by the classification strategy for Storeys 2, 6, and 22 of the AMADEUS detector: These three storeys are equipped with HTI hydrophones.

from a UHE neutrino interaction in water. For the AMADEUS detector, the determination of this propagation pattern is barely feasible due to the geometry of the detector. For a future, large-scale detector with an optimised configuration, which would result in a better precision of the position reconstruction, the volume near the surface of the sea should be excluded in the search for neutrinos.

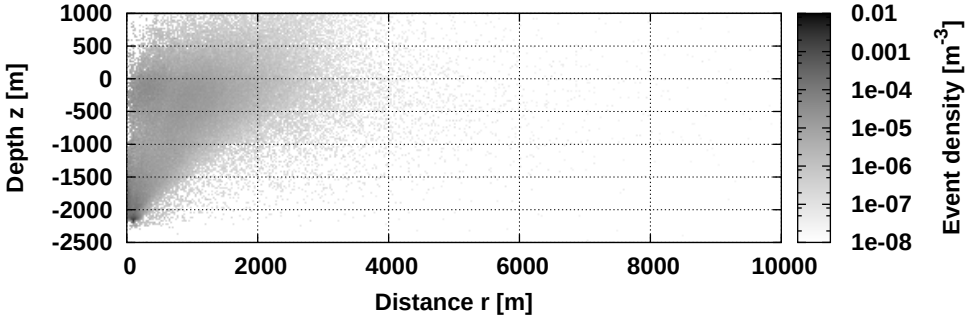
5.1.3 Temporal distribution

As was the case for the spatial distribution, the temporal distribution of signals can be used to define times that are favourable for the detection of neutrinos. The temporal distribution of transient signals as a function of the time of the day is shown in Fig. 5.1.3 for the complete dataset and three subsets. The mean rate of the position-reconstructed signals (events) for each bin of one hour was calculated. The distribution for the complete dataset used in this analysis is shown in red. The distribution shows that the mean rate of events at night is as much as five times greater than during the day. The overall mean rate is about 0.3 Hz. The dotted blue curve represents the temporal distribution of events that the spatial clustering algorithm (cf. Sec. 4.4.4) identified as part of a spatial cluster, when the source is assumed to be at the surface of the sea. This distribution is almost identical to the distribution of the complete data. The correlation coefficient between the two distributions is 99.8 %. So, most of the sources of the recorded signals originate at or near the surface of the sea, as mentioned above. As part of the clustering method, a search for sound-emitting objects moving along straight tracks was performed. The temporal occurrence of events identified as such objects (11 % of all clustered events) is shown as the dotted pink curve. This distribution also is correlated (correlation coefficient $< 80\,\%$) with the temporal structure of all events. The main contribution of the events from these moving sound emitters is around four o'clock in the morning and 11 o'clock in the evening. Those periods approximately coincide with the times where ferries (from Marseille, France to Ajaccio, Corsica and back) would be near the ANTARES detector. The possibility of high bio-acoustic activity at night time was also researched, however, so far no positive indications were found. The resulting distribution after classification for bipolar pulses (dashed green line) has an overall mean rate of 0.01 Hz. The distribution is a highly correlated subset of all reconstructed events, which indicates that the sources of transient signals, most probably shipping traffic and marine mammals, can emit various waveform types including bipolar pulses. Ships can produce BIPs by cavitation or by using sonar systems. Sea mammals, like dolphins, use bipolar pulses for echo sounding and other types of signals for communication between individuals. As excluding the times of high activity would reduce the measurement time significantly, an exclusion of the sources of the transient signals based on their position is favoured. Using the spatial clustering technique as a veto and the classification of BIP-like signals, the mean rate of the remaining events is about 0.002 Hz.

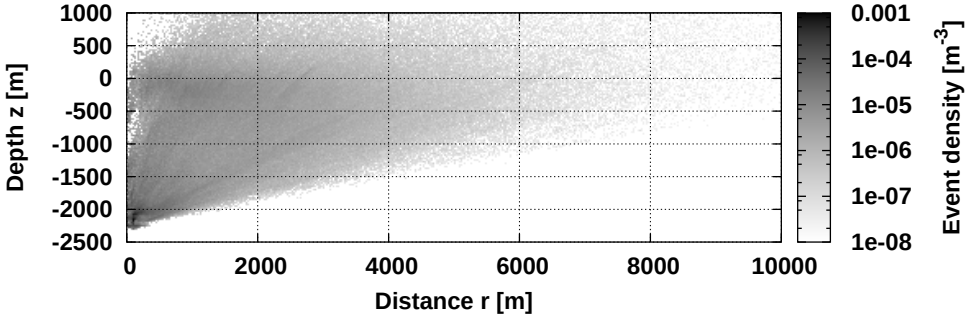
To summarise, the rate of BIP-like signals, which can mimic neutrino signatures, is around 100 events per year per km^3 at the AMADEUS detector site. The main contribution is from objects that can be located at or near the surface of the sea and they seem to emit various types of waveforms for a given period of time. So, it would be ideal to exclude that region of the surface of the sea from the fiducial volume. For the AMADEUS detector, this is hardly feasible because the position reconstruction is severely hampered by the size and geometry of



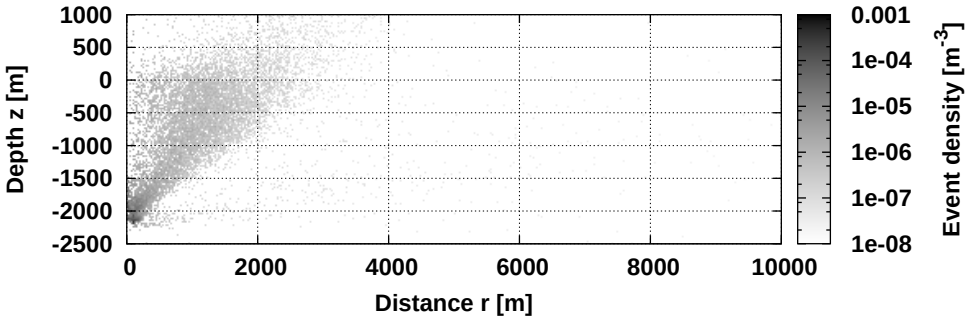
(a) All events



(b) After classification



(c) Result of the cluster analysis



(d) After classification and cluster analysis

Figure 5.3: Spatial density of transient signals as a function of the depth z and the distance r. In (a) the density of the transient signals as filtered by the AMADEUS online-filter is shown. The density of BIPs-like signals as selected by the classification strategy is given in (b). In (c), the spatial density of the events that have been identified by the cluster analysis is shown. In (d), the resulting density distribution after classification and cluster analysis (see text).

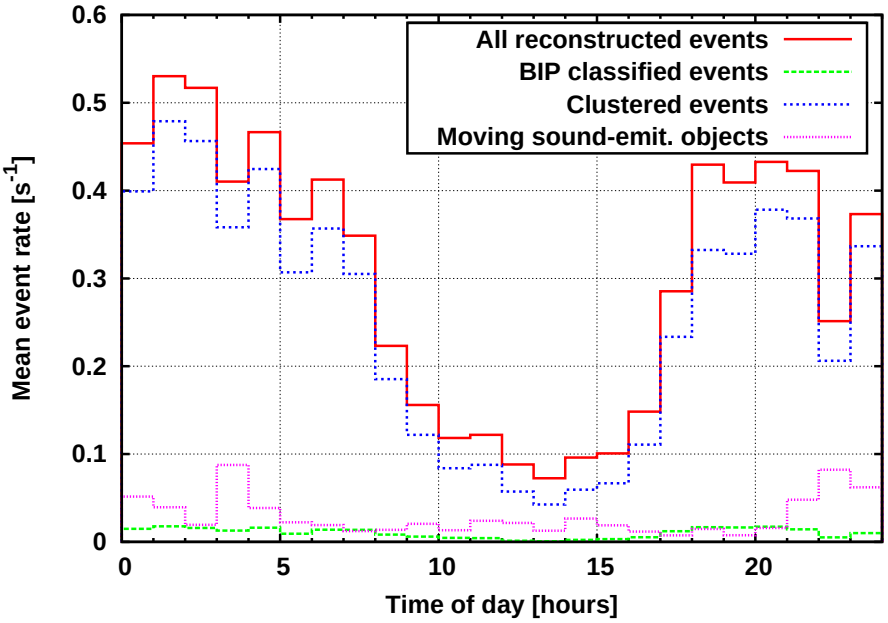


Figure 5.4: Temporal distribution of the reconstructed events over the time of the day for all events (red, solid), for events classified as bipolar (green, dashed), for events that spatially cluster, when projected to the surface of the sea (blue, dotted), and for clustered events identified as moving sound-emitting objects (pink, small dotted).

the AMADEUS detector.

5.2 Effective volume of the AMADEUS detector

An effective volume of the AMADEUS detector is presented and discussed in this section. The simulation and analysis chain, as described before, was used to simulate the data needed for this study. Following the approach derived in [80], an effective volume V_{eff} can be defined as:

$$V_{\text{eff}}(E_\nu) = \frac{\sum_{N_{\text{gen}}} \delta_{\text{selected}} p(E_\nu, \mathbf{r}, \mathbf{e_p})}{N_{\text{gen}}} V_{\text{gen}}, \tag{5.1}$$

where N_{gen} is the number of generated neutrino interactions in a volume V_{gen} around the detector centre, and $p(E_\nu, \mathbf{r}, \mathbf{e_p})$ is the probability that the neutrino can reach the interaction vertex set in the simulation. The term $\delta_{\text{selected}} \in \{0, 1\}$ expresses that the probability is only taken into account for the effective volume V_{eff}, if the pressure pulse corresponding to the neutrino interaction was selected by the on-line filter within a time window of 128 µs around the expected arrival time. So the effective volume depends on the cuts chosen to select the neutrino events and is specific to the analysis presented in this work. The probability that the neutrino reaches the vertex is defined as:

$$p(E_\nu, \mathbf{r}, \mathbf{e_p}) = \mathrm{e}^{-d_{\text{WE}}(\mathbf{r}, \mathbf{e_p})/\lambda_{\text{water}}(E_\nu)}, \tag{5.2}$$

where $\mathbf{r}$ is the position of the interaction vertex, and $\mathbf{e_p}$ is the unit vector of the direction of the flight trajectory. The mean free path $\lambda_{\text{water}}(E_\nu)$ of the neutrino, as described in Sec. 2.1, is

inversely proportional to the neutrino's total cross section. The total cross section as a function of the energy E_ν was parameterised using values from [9]. The distance d_{WE} is the water equivalent of the distance traveled through matter of varying densities encountered along the flight path. For the determination of the density distribution over the flight path, the *Preliminary Reference Earth Model* (PREM) [10] was used to model the earth's density profile. In addition, it is assumed that the earth is covered by water with a depth of 2.5 km and that the detector is placed on the sea floor.
For the calculation of the effective volume, $N_{\mathrm{gen}} = 10^7$ neutrinos with energies uniformly distributed between 10^9 GeV and 10^{12} GeV were simulated. The simulation is limited to 10^{12} GeV (cf. Sec. 3.1). A uniform energy distribution was chosen to ensure a sufficient number of events over the entire energy range under consideration. The interaction vertexes of these neutrinos were chosen in a cylindrical volume of 1200 km^3 around the AMADEUS detector. The headings of the flight paths were ranging from $0^\circ - 360^\circ$ in the azimuth and from $0^\circ - 100^\circ$ in the zenith[1], meaning that the solid angle was extended from the upper hemisphere to 10° below the horizon. Neutrinos entering the generation volume from below the horizon will traverse an increasing amount of matter. Below 10°, the probability is practically zero that a neutrino in the energy range under consideration would reach the interaction vertex. To determine the random coincidences formed by the ambient noise a separate set of simulated data was created that did not contain any signals. The corresponding model used to simulate the ambient noise is described in Sec. 3.2.2. The effective volume has been calculated for three different "Stages" that describe increasingly realistic conditions and reconstruction requirements:

Stage 1: The ambient noise is assumed to be minimal, matching the lowest exspected noise level (sea state 0) at all times, and the coincidence requirement for the on-line filter simulation is that at least two sensors on one storey must respond.

Stage 2: The complete ambient noise model (cf. 3.2.2) and the standard on-line filter used for the AMADEUS detector (cf. 3.4) are used, which requires at least four sensors on two storeys each.

Stage 3: In addition to Stage 2, the reconstruction of the acoustic source position is required to have a maximum deviation of less than 100 m from the simulated position.

The results of this study are shown in Fig. 5.5 for the three Stages. For Stage 1, the effective volume obtained is consistent with the random coincidences up to a neutrino energy of $1.8 \cdot 10^{10}$ GeV and for 10^{12} GeV the effective volume is above 2 km^3. The requirements of this Stage 1 are minimal and so this can be seen as an idealised detection threshold of the AMADEUS detector. For Stage 2, in which the complete ambient noise model and the standard settings for the on-line filter of the AMADEUS detector are used, the effective volume exceeds the background of random coincidences with statistical significance at $1.8 \cdot 10^{11}$ GeV, and it reaches a size of about 0.1 km^3 at an energy of 10^{12} GeV. The effective volume for Stage 3 shows that the determination of the position of the interaction vertex has sizeable uncertainties for a small detector that essentially has a two-dimensional configuration. The effective volume is below 0.01 km^3 at the highest energy in the given range, and the detection threshold is $3.2 \cdot 10^{11}$ GeV.

[1] A zenith angle of 0° corresponds to a neutrino coming from straight above.

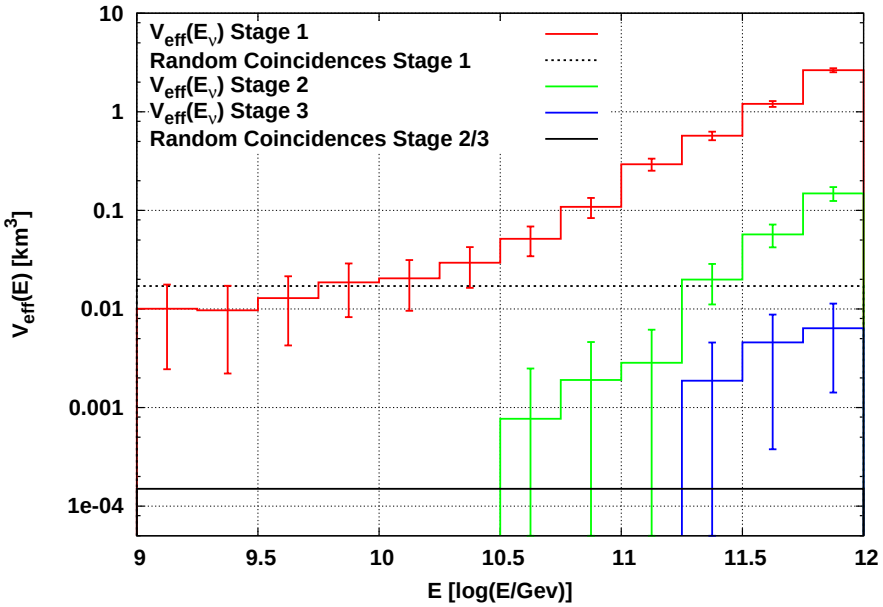


Figure 5.5: The effective volume of the AMADEUS detector as a function of the logarithmic neutrino energy for the three different stages (1: red, 2: green, and 3: blue) as described in the text. Also shown are the random coincidence rates for the Stage 1 (black solid line) and the Stage 2 and 3 (black dashed line).

5.3 Transient-free, limit-setting potential of the AMADEUS detector

The calculation of the transient-free, limit-setting potential of the AMADEUS detector to a flux of UHE neutrinos is described in this section. To estimate the potential of the acoustic detection method for UHE neutrinos, a flux limit for the AMADEUS detector is calculated assuming that the background of transient signals can be completely suppressed. However, this is not possible for the AMADEUS detector without further improvements of the background reduction technique, which is beyond the scope of this work. This estimate is used mainly to determine how the lower energy threshold of a neutrino flux measurement is affected by the ambient background. Following the approach derived in [81], an model-independent flux limit can be calculated as:

$$\Phi_{90\,\%\,\mathrm{CL}} = \frac{N_{90\,\%\,\mathrm{CL}}}{\Omega T E_\nu [V_{\mathrm{eff}}(E_\nu)/\lambda_{\mathrm{water}}(E_\nu)]}, \tag{5.3}$$

where $N_{90\,\%\,\mathrm{CL}} = 2.44$ is the 90% *confidence level* (CL) [82] from an average upper limit that would be obtained by an ensemble of experiments with the expected background of zero events and no true signal observed, Ω is the solid angle, T is the assumed integrated measurement time, V_{eff} is the effective volume and λ_{water} is the mean free path of the neutrino as described in the previous section.

The effective Volume V_{eff} for the three so-called Stages defined in the previous section were used for the transient-free limit estimate. The assumed integrated measurement time used is one year; the energy range is from 10^9 GeV to 10^{12} GeV and the solid angle is the same as that used for the calculation of the effective volume. In Fig. 5.6, the estimate of the AMADEUS

transient-free, limit-setting potential is shown together with the theoretical cosmogenic neutrino flux predictions [83]. The calculated transient-free limit estimates of the AMADEUS detector are characterised mainly by the effective volumes and energy thresholds derived for the three Stages. In particular the effective volume reflects the detector's small physical size and basically two-dimensional configuration, which is also the reason why the requirement for the precision of the position reconstruction used for Stage 3 is difficult to accomplish for the AMADEUS detector. Nevertheless, the transient-free, limit-setting potential is promising and encourages further studies for large-scale detectors. A larger physical size with an optimised configuration of the sensors — taking advantage of the large attenuation length — would increase the effective volume significantly and presumably would lower the energy threshold.
In Sec. 5.1, the rate of BIP-like signals was derived as about 100 events per year per km^3. A significant reduction of this background rate is possible by excluding the region of the sea's surface, where most of the transient signals originate, furthermore, by taking into account the disk-like propagation pattern of a BIP originating from a UHE neutrino interaction. This is hardly feasible for the AMADEUS detector in the current setup, but it would be possible in a large-scale three-dimensional detector.

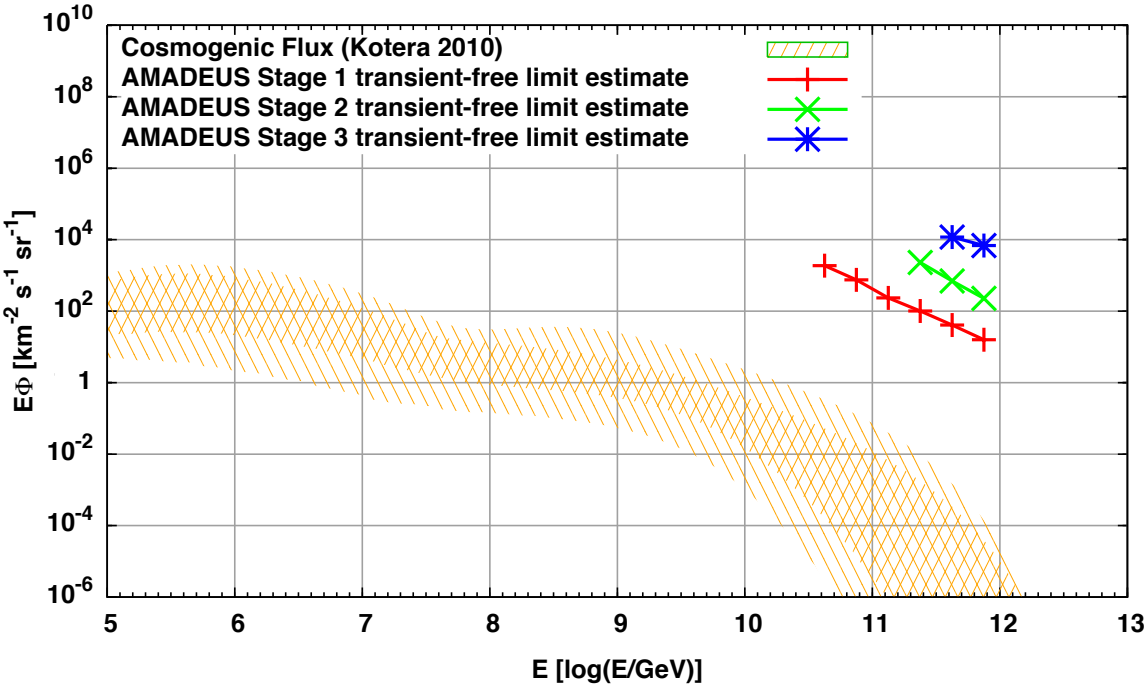
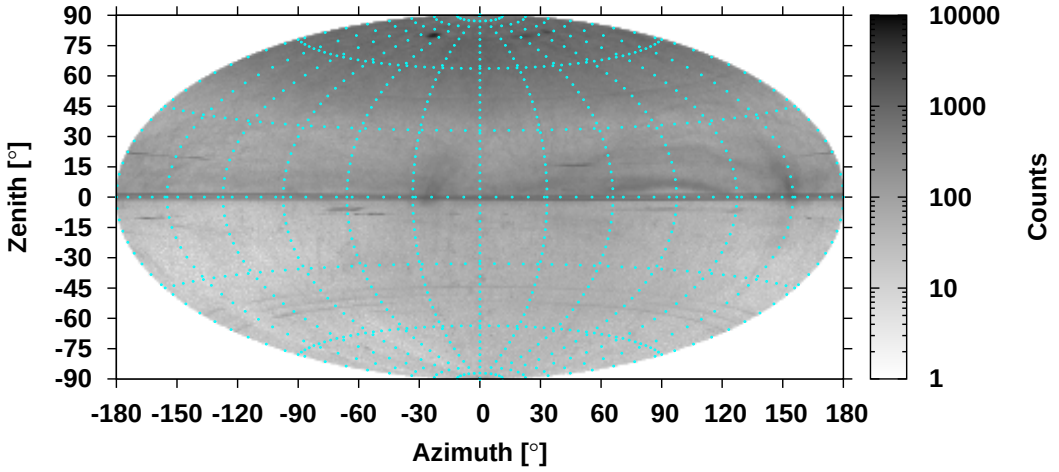
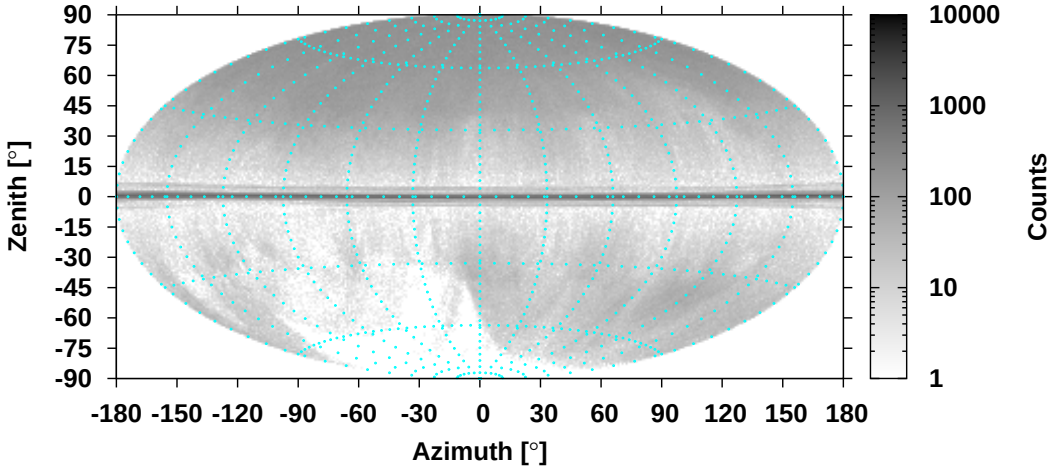


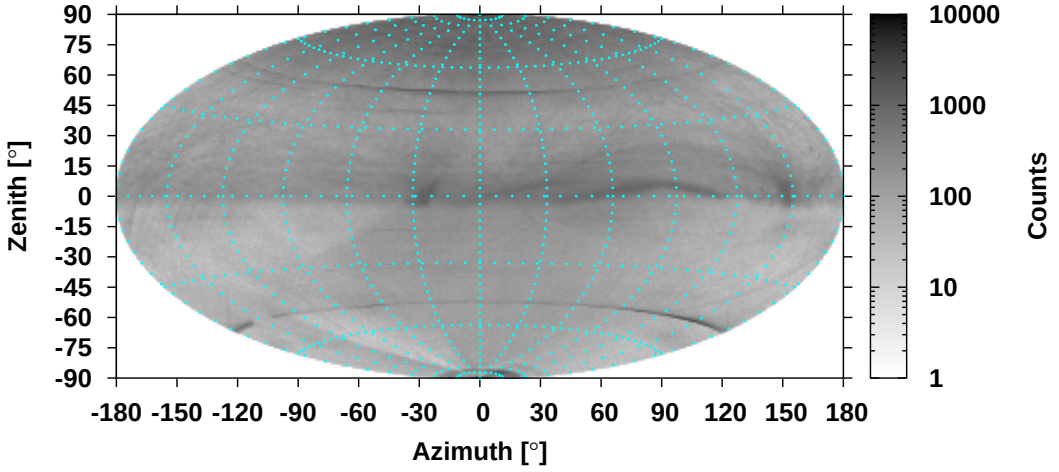
Figure 5.6: The transient-free, limit-setting potential of the AMADEUS detector is shown together with predictions of the cosmogenic neutrino flux [83] (orange pattern).



(a) Line 12 Storey 23

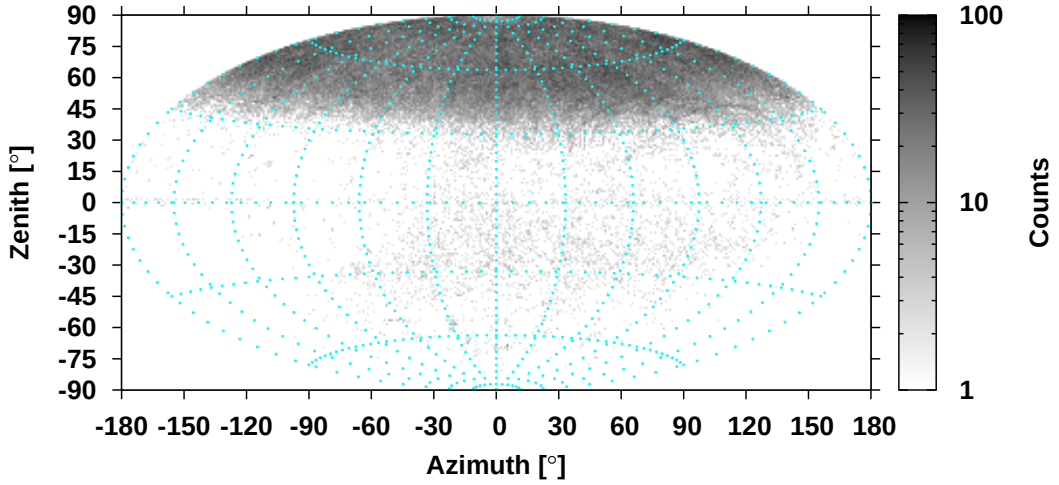


(b) Line 12 Storey 21

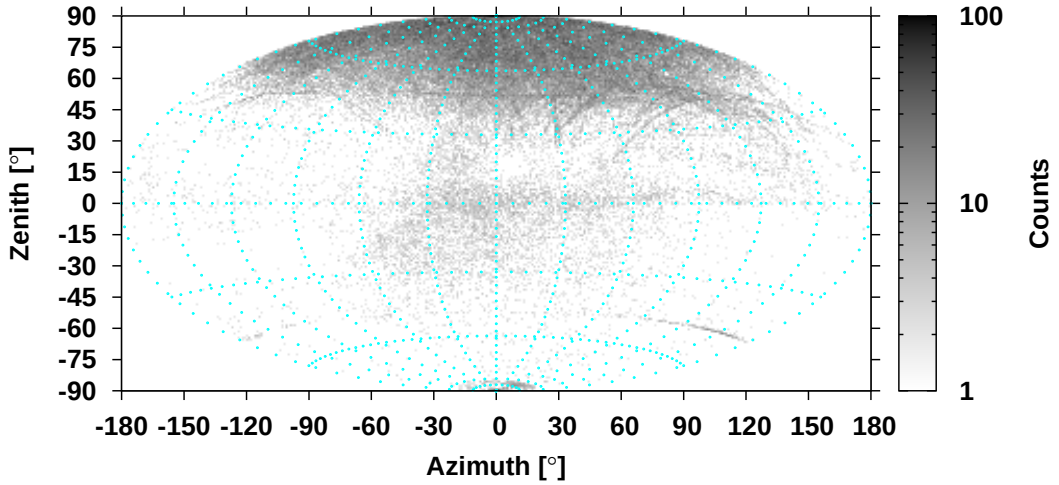


(c) IL Storey 3

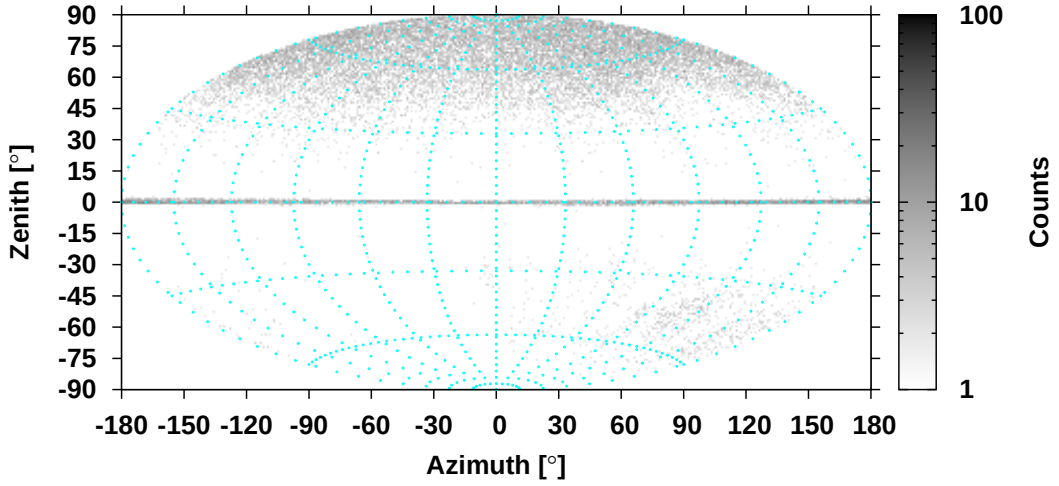
Figure 5.7: Directional distribution of transient signals for Storeys 3, 21, and 23 of the AMADEUS detector.



(a) Line 12 Storey 23



(b) IL Storey 3



(c) Line 12 Storey 21

Figure 5.8: Directional distribution of transient BIPs as selected by the classification strategy for Storeys 3, 21, and 23.

Chapter 6

Summary

The main objective of this work was to study possible methods for classifying neutrino-induced acoustic signals in the presence of acoustic background in the deep sea. This required the development of a suitable simulation and analysis chain. This chain was designed within the SeaTray/IceTray software framework. Its modular architecture is highly flexible and making it easy to adapt to different environmental conditions, different data acquisition hardware, and different detector geometries. In this framework, analysis tools can be implemented efficiently or extended in a consistent way for further application. The simulation chain is capable of reproducing all of the aspects required for acoustic neutrino detection — the generation of the acoustic signal of an *ultra-high-energy* (UHE) neutrino that induces a particle cascade, the acoustic deep-sea environment with its variable and diverse noise conditions, and different detector geometries and hardware components. In the current stage of development, the detector geometry, data acquisition hardware, and background models used in the simulation chain are equivalent to those of the AMADEUS[1] detector. The AMADEUS system consists of an array of 36 acoustic sensors arranged in six clusters distributed over the ANTARES[2] neutrino telescope on different length scales. The ANTARES detector, of which AMADEUS is a subsystem, is located in the Mediterranean Sea, 40 km off the coast of Toulon, southern France in a depth of about 2500 m.

The simulation chain consists of modules that build on each other to create a simulated event that corresponds to the output of the detector. The first major step of the simulation chain is the calculation of the acoustic pulse from a hadronic cascade produced by a UHE neutrino interaction in the water. Following the thermo-acoustic model, an acoustic pulse is produced by the energy deposition of a hadronic shower originating from a UHE neutrino interaction. The local heating caused by fast deposition of the energy leads to a characteristic *bipolar pressure pulse* (BIP). Due to coherent emission over the cylindrical geometry of the hadronic shower, the resulting pressure pulse propagates through the medium in a disk-like shape perpendicular to the main axis of the cascade. In the simulation, a parametrisation of the energy deposited by the cascade was used to calculate the resulting acoustic signal, including complex attenuation over the path from the origin of the shower to the position of the acoustic sensor.

The background for acoustic neutrino detection in the deep sea consists of two different types

[1] *ANTARES Modules for Acoustic Detection under the Sea* (AMADEUS)

[2] *Astronomy with a Neutrino Telescope and Abyss Environmental Research* (ANTARES)

of acoustic noise: transient and ambient noise. Transient noise signals have short durations and amplitudes that exceed the ambient noise level. These signals can mimic bipolar pulses from neutrino interactions. In the simulation, four types of transient signals were implemented based on observations made with the AMADEUS detector: bipolar and multipolar pulses, sinusoidal signals, and signals with Brown noise frequency characteristics. These four types of signals can be generated by anthropogenic sources, such as shipping traffic, and by marine mammals. The ambient noise is caused mainly by the agitation of the seaâĂŹs surface, e.g., by wind, breaking waves, spray, and cavitations. Thus, it is correlated to the weather conditions, mainly wind speed. The model used for the simulation of ambient noise was based on the so-called Knudsen spectra, which were adapted to the deep sea enviornment by applying attenuation effects. The resulting noise model was in general agreement with the measured noise spectra and the noise levels that occurred.
The simulation of the data acquisition hardware comprises the simulation of the acoustic sensor's response and the read-out electronic's response to an incoming signal. This includes sensors using the piezoelectric effect to convert the pressure signal into a voltage signal and read-out electronics to amplify, filter, and digitise this signal. The inherent noise and the system transfer function, for both the sensors and the electronics, were measured in the laboratory. From these data, parametrisations of the inherent noise spectrum and the system response function for each part were derived and used to simulate the effect of the data acquisition hardware. The sensors show a directional dependency on the sensitivity, therefore signal and ambient noise must be treated separately. For the ambient noise, the sources defining the incident direction is the surface of the sea area above the detector.
Finally, an implementation of the AMADEUS on-line filter system was applied, which, for real data, was used to reduce the amount of data that had to be stored and to select signals for further off-line analysis. The filter of main importance for this work was based on a matched filtering technique, which uses a defined bipolar pulse as a reference to select signals with bipolar shape. In addition, a coincidence test between the sensors was performed. In the current setup, at least four sensors in one cluster of sensors and at least two clusters must have responded to the filter.
Furthermore, the analysis chain uses strategies for the reconstruction of the acoustic source position and the classification of bipolar pulses, which are the acoustic signature of a UHE neutrino interaction. The precise reconstruction of the arrival time of the signal is crucial for the direction and position reconstruction of the acoustic source. For bipolar pulses, the arrival time is determined by performing up-sampling of the filtered waveform sample and cross-correlation with a pre-defined bipolar pulse. This procedure achieves a precision of about 1 µs. Due to the narrow opening angle of the acoustic emission of a neutrino interaction, local clusters of sensors would be preferred in the design of a potential future acoustic neutrino detector. Such clusters, consisting of several sensors arranged with interspaces of a few meters, also have advantages for the coincidence test used by the on-line filter and for the reconstruction of acoustic source positions. The direction reconstruction is based on a least squares fit of the measured arrival times compared to the expected arrival times at a given sensor in the cluster for an assumed incident direction. The angular resolution reached with this algorithm is centred around zero and the standard deviation of the distribution is about 0.7° for both the zenith and azimuth angles. The application of the algorithm on simulated acoustic sources was validated successfully by using experimental data of the emissions of the acoustic posi-

tioning system of the ANTARES detector. Each detection line in the ANTARES detector has an acoustic transducer at the anchor of the line, which emits every 2 min.
The position reconstruction of the acoustic source was obtained using a ray tracing technique. If the incident direction was reconstructed for at least two of the sensor clusters, the intersection point of the rays, starting from the storeys and pointing into the reconstructed direction, was searched for. In the implementation of this algorithm, the point was used at which the sum of the squared distances to the rays was minimal. From the simulation of acoustic sources in a cube of the size of $5 \times 5 \times 2.5\,\mathrm{km}^3$ around the centre of the detector, a resolution better than 15 m for each coordinate was derived. The uncertainty of the direction reconstruction results in an increasing uncertainty of the position reconstruction with the distance to the acoustic source. As for the direction reconstruction, the position reconstruction was verified with the emissions of the acoustic positioning system. The reconstructed positions are in agreement with the true positions of the emitters. The resolution of the position reconstruction for these emitters is about 5 m.
A classification strategy based on machine learning algorithms was deployed as a robust, effective, and efficient way to determine bipolar signals in the diverse transient background. The classifiers included in the strategy were trained and tested with data from the complete simulation chain. Two classifiers — the Random Forest and the Boosting Trees algorithms — have achieved the best results for predicting the data of individual sensors and for the combined result of clusters of sensors. For individual sensors, the classification error is of the order of 10 % for a well-trained model. The combined results of the individual sensors in a cluster were used as new input for training. This method obtains a classification error of less than 2 %.
On that basis, an analysis of the transient background as measured at the AMADEUS site was performed. The spatial and temporal distributions of events for which the position had been reconstructed were researched. The rate of BIP-like signals, which can mimic neutrino signatures, is around 100 events per year per km^3 in the vicinity of the AMADEUS detector. The objects that contribute to these signals are located mainly at or near the seaâĂŹs surface. These objects, most probably ships and marine mammals, can emit various waveform types including bipolar pulses. So, it would be ideal to exclude the region of the surface of the sea from the fiducial volume. For the AMADEUS detector, which has a particular geometry and is rather small, this is barely feasible due to the uncertainty of the position reconstruction.
The calculation of an effective volume and a transient-free, limit-setting potential of the AMADEUS detector to a UHE neutrino flux was performed based on the simulation and analysis chain that was developed. The expression "transient-free, limit-setting potential" was introduced to define a limit that could be established with the AMADEUS detector, if it is assumed that the remaining transient background can be suppressed completely. It is used mainly to determine how the lower energy threshold of a neutrino flux measurement is affected by the ambient background. The effective volume of the AMADEUS detector was calculated from 10^7 simulated neutrinos in a volume of $1200\,\mathrm{km}^3$ with energies distributed uniformly between 10^9 GeV, at which a neutrino-induced interaction begins to produce an acoustically detectable signal, and 10^{12} GeV, the energy up to which the underlying parametrisations of the simulation are specified. The solid angle, from which the incident direction of the neutrinos was chosen, was extended from the upper hemisphere to 10° below the horizon. The probability of a UHE neutrino to reach the AMADEUS detector after propagating through an angle-dependent amount of soil and water was taken into account. Below that angle and in the energy range

under consideration, the probability of a neutrino to reach the interaction vertex is practically zero. The assumed integrated measurement time was one year. The ambient noise model for the Mediterranean Sea, which is included in the simulation, was used for this study, but transient sources were not considered. The effective volume and the transient-free, limit-setting potential were evaluated for different stages describing the increasingly realistic conditions of the ambient noise level and of the on-line filter settings.
The detection threshold of the AMADEUS detector for idealised conditions is at about $1.8 \cdot 10^{10}$ GeV and the effective volume exceeds $2\,\mathrm{km}^3$ at 10^{12} GeV. This indicates the potential of acoustic detection of UHE neutrinos. The transient-free, limit-setting potential derived for the AMADEUS detector is determined mainly by the detector's effective volume, which, for the current setup of the detector, is less than $0.1\,\mathrm{km}^3$ at 10^{12} GeV, and the threshold energy of $1.8 \cdot 10^{11}$ GeV.
For a small detector, such as AMADEUS, the transient-free, limit-setting potential is promising and encourages further studies for larger-scale detectors. These studies can be based on the simulation and analysis chain developed for this work. Furthermore, the measurement of the rate of BIP-like signals from the observations made with the AMADEUS detector is an important step forward in the effort to reach a conclusion on the feasibility of the acoustic neutrino detection in the Mediterranean Sea.

Kapitel 7

Zusammenfassung

Hauptziel der vorgestellten Arbeit war es zu untersuchen, wie man das akustische Signal, das von einer Neutrino-Interaktion in Wasser herrührt, von transienten Untergrundsignalen, die in der Tiefsee vorkommen, unterscheiden kann. Hierzu wurden Algorithmen, die auf maschinellem Lernen beruhen, untersucht. Um diese Algorithmen zu analysieren, wurde eine geeignete Simulations- und Analysekette entworfen und implementiert. Dies geschah im Rahmen des SeaTray/IceTray Software Frameworks, dessen modularer Aufbau eine einfach handhabbare und flexible Umgebung darstellt, um sie an unterschiedliche zu simulierende Detektorgeometrien, Datennahme-Elektronik und Umgebungsbedingungen anzupassen. Des Weiteren lässt sich auch die Analyse-Software leicht in dieses Framework integrieren und erweitern.
Die erstellte Simulationskette ist in der Lage, alle Prozesse, die für die akustische Teilchendetektion notwendig sind, nachzubilden. Dies umfasst die Erzeugung der akustischen Signatur einer ultrahoch-energetischen Neutrino-Wechselwirkung in Wasser, den variablen und vielfältigen akustischen Untergrund in der Tiefsee und einen entsprechenden Detektor und dessen Datennahme. In der momentanen Ausbaustufe der Simulation wurde die Detektorgeometrie und Datennahme des AMADEUS[1] Detektors nachempfunden und ein Modell für den akustischen Untergrund im Mittelmeer verwendet. Der AMADEUS Detektor besteht aus insgesamt 36 akustischen Sensoren, die in jeweils 6 Clustern von Sensoren über das ANTARES Neutrino-Teleskop verteilt sind. Dabei wurden unterschiedliche Abstände zwischen den einzelnen Clustern realisiert. Das ANTARES[2] Neutrino-Teleskop und das dazugehörige Subsystem AMADEUS sind im Mittelmeer, 40 km vor der französischen Küste nahe Toulon, in einer Tiefe von 2500 m installiert. Die Simulationskette besteht aus einzelnen Modulen, die aufeinander aufbauen, um ein Ereignis zu generieren, das dem des AMADEUS Detektors entsprechen würde.
Der erste Schritt in der Simulation ist die Berechnung des akustischen Signals, das durch eine ultrahoch-energetische Neutrino-Interaktion in Wasser ausgelöst wird. Entsprechend dem thermo-akustischen Modell entsteht ein akustischer Puls, wenn durch eine hadronische Kaskade, die durch eine Neutrino-Wechselwirkung erzeugt wurde, Energie im sie umgebenden Medium deponiert wird. Diese schnell ablaufende Energiedeposition und die darauf folgende lokale Erwärmung des Mediums führt zu einem charakteristischen bipolaren Druckpuls, der sich aufgrund der kohärenten Emission über die zylindrische Geometrie der Teilchenkaskade

[1] *ANTARES Modules for Acoustic Detection under the Sea* (AMADEUS)
[2] *Astronomy with a Neutrino Telescope and Abyss Environmental Research* (ANTARES)

hinweg scheibenartig ausbreitet. In der Simulation wird eine Parametrisierung dieser Energiedeposition einer Teilchenkaskade verwendet. Daraus wird das akustische Signal errechnet, das an den entsprechenden Sensoren innerhalb des Detektors ankommt, einschließlich der komplexen Abschwächung, der das Signal auf seinem Weg zum Sensor unterworfen ist.
Der akustische Untergrund in der Tiefsee besteht aus zwei Anteilen: dem Umgebungsrauschen und dem transienten Rauschen. Transiente Rauschsignale sind von kurzer Dauer und besitzen eine Amplitude, die das Niveau des Umgebungsrauschens übersteigt. Transiente Signale können unter anderem bipolar sein und damit die akustische Signatur einer Neutrino-Wechselwirkung imitieren. Vier verschiedene Formen von transienten Signalen, die auf Beobachtungen mit dem AMADEUS Detektor beruhen, sind in der Simulation enthalten: bipolare und multipolare Pulse, sinusförmige Signale und Signale, die die Charakteristik von braunem Rauschen haben und sich in keine der anderen Gruppen einordnen lassen. Diese Signale können von Meerestieren oder anthroposophischen Quellen wie Schiffen stammen. Das Umgebungsrauschen rührt hauptsächlich von der Meeresoberfläche her, verursacht durch Wind, Wellen, Gischt und eingeschlossene Luftbläschen, deren Vorkommen von den Wetterbedingungen und im Besonderen von der Windgeschwindigkeit abhängen. In der Simulation wird das Umgebungsrauschen über ein Modell erzeugt, das auf den Knudsen-Spektren beruht. Diese wurden durch die Hinzunahme von Abschwächungseffekten an die Tiefsee angepasst. Das resultierende Modell ist grundsätzlich in Übereinstimmung mit den gemessenen Rauschspektren und der Häufigkeitsverteilung der Rauschniveaus.
Die Simulation der Datennahme-Elektronik besteht zum einen aus der Simulation der akustischen Sensoren, die den piezo-elektrischen Effekt ausnutzen, und zum anderen aus der Simulation der Elektronik, die für die Verstärkung, Filterung und Abtastung der von den Sensoren erzeugten Spannungssignale zuständig ist. Sowohl für die Sensoren als auch für die Elektronik waren Labormessungen der Systemtransferfunktion und des inhärenten Rauschens vorhanden und diese wurden in Form einer Parametrisierung in die Simulation integriert.
Letzter Teil der Simulation ist die Implementation des im AMADEUS Detektor verwendeten Filtersystems, das zur Vorselektierung von Signalen zur späteren Analyse und zur Datenreduktion dient. Für die Selektion von bipolaren Signalen wird eine Matched-Filter-Technik verwendet und ein Koinzidenztest durchgeführt. Diese Funktionalität wurde in die Simulation integriert.
Die Analysekette enthält Methoden, um den akustischen Quellort eines Signals zu rekonstruieren und eine Strategie zur Klassifizierung von bipolaren Signalen. Für die Rekonstruktion des Quellorts ist eine präzise Bestimmung der Ankunftszeit eines Signals am Sensor von großer Wichtigkeit. Mit den verwendeten Algorithmen ist eine Bestimmung der Ankunftszeit mit einer Genauigkeit von 1 µs möglich. Aus früheren Studien ist bekannt, dass lokale Cluster von Sensoren wegen des engen Öffungswinkels der akustischen Emission einer UHE-Neutrino-Wechselwirkung vorteilhaft sind. Auch erleichtert dies den Koinzidenztest während der Vorselektierung der Signale durch das Filtersystem. Dieses Design ist in AMADEUS umgesetzt. Dadurch wird es ermöglicht, für jeden lokalen Cluster die Richtung eines ankommenden Signals zu rekonstruieren. Die Richtungsrekonstruktion basiert auf einem Least-Square-Fit, bei dem die gemessenen Ankunftszeiten an den Sensoren in einem Cluster mit den erwarteten für eine vorgegebene Richtung verglichen und minimiert werden. Die aus der Simulation bestimmte Auflösung der Richtungsrekonstruktion ist 0.7° für den Azimut- und Zenitwinkel. Dies konnte auch aus den Experimentdaten für die Emissionen des akustischen Positionierungssystems, das in ANTARES integriert ist, nachvollzogen werden.

Der Quellort wird aus der vorhergegangenen Richtungsrekonstruktion über ein Strahlverfolgungsverfahren bestimmt. Wenn für mindestens zwei der lokalen Cluster die Richtung rekonstruiert wurde, kann der Ort bestimmt werden, in dem man von den Clustern ausgehend Strahlen in diese Richtung aussendet und den Schnittpunkt der Strahlen bestimmt. In der Implementation dieses Algorithmus wurde der Punkt gesucht, für den die Summe der Abstandsquadrate zwischen den Strahlen minimal ist. Diese Methode wurde ebenfalls mit den Daten des akustischen Positionierungssystems validiert und für diese Emissionen konnte eine Auflösung der Positionsbestimmung von ungefähr 5 m bestimmt werden. Allerdings zeigte die Simulation, dass die Genauigkeit der Positionsbestimmung des akustischen Quellorts mit der Entfernung abnimmt, wie es auch qualitativ aus geometrischen Überlegungen zu erwarten ist. Für simulierte Ereignisse in einem Würfel von $5 \times 5 \times 2.5\,\mathrm{km}^3$ wurde eine Auflösung von besser als 15 m für jede Positionskoordinate erreicht.
Eine Strategie zur robusten, effektiven und effizienten Klassifizierung von bipolaren Signalen, die auf Algorithmen des maschinellen Lernens basieren, wurde entwickelt. Diese Algorithmen wurden mithilfe der Simulation trainiert und getestet. Die Algorithmen Random Forest und Boosting Trees haben sich als am besten geeignet herausgestellt, um die Daten von einzelnen Sensoren und von denen der lokalen Cluster zu klassifizieren. Es konnte für einzelne Sensoren ein Klassifizierungsfehler von unter 10 % erreicht werden unter Verwendung eines austrainierten Modells. Wenn die Vorhersagen für die einzelnen Sensoren in einem lokalen Cluster kombiniert wurden, war es möglich, einen Klassifizierungsfehler von unter 2 % zu erreichen.
Aufbauend auf der Analysekette wurde der am AMADEUS Detektor gemessene Untergrund an transienten Signalen untersucht. Hierzu wurden die zeitliche und örtliche Verteilung der ortsrekonstruierten Signale analysiert. Die Rate an bipolaren Signalen, die die akustische Signatur einer Neutrino-Interaktion nachahmen können, beträgt ungefähr 100 Ereignisse pro Jahr und km^3 in der Umgebung des AMADEUS Detektors. Hauptsächlich verantwortlich für diese Signale scheinen Objekte zu sein, die sich an oder nahe der Meeresoberfläche befinden. Diese Objekte, höchstwahrscheinlich Schiffe und Meerestiere, können verschiedene Signalformen aussenden, unter anderem bipolare Pulse, in kurzer zeitlicher Abfolge. Davon ausgehend wäre es sinnvoll, die Region nahe und an der Meeresoberfläche von der Suche nach Neutrinos auszuschließen. Dies ist für den AMADEUS Detektor jedoch kaum möglich, da es dafür an der Genauigkeit der Ortsrekonstruktion mangelt.
Mithilfe der Simulations- und Analysekette wurden ein effektives Volumen und ein "transient-free, limit-setting potential" für den AMADEUS Detektor errechnet. Der Ausdruck "transient-free, limit-setting potential" wurde eingeführt, um ein Limit zu definieren, dass vom AMADEUS Detektor aufgestellt werden könnte, wenn man annimmt, dass der noch verbleibende Untergrund an transienten Signalen komplett unterdrückt werden kann. In der Hauptsache kann es dazu verwendet werden, um zu untersuchen, wie sich das ambiente Untergrundrauschen auf die Messung des Neutrinoflusses auswirkt. Für die Berechnung des effektiven Volumens wurden 10^7 Neutrinos simuliert, deren Energie gleichverteilt gewählt wurde zwischen 10^9 GeV, ab dieser Energie wird ein akustisch detektierbares Signal von einer Neutrino-Interaktion erzeugt, und 10^{12} GeV, bis zu dieser Energie sind die in der Simulation verwendeten Parametrisierungen spezifiziert. Die Interaktionsverticies der Neutrinos wurden in einem Volumen von $1200\,\mathrm{km}^3$ festgelegt. Der Raumwinkel, aus dem die Flugrichtung der Neutrinos ausgewählt wurde, setzte sich aus der oberen Hemisphäre zuzüglich 10° unterhalb des Horizonts zusammen. Darüber hinaus ist die Wahrscheinlichkeit, dass ein Neutrino im betrachteten Energiebereich den Vertex

erreicht, nachdem es eine vom Einfallswinkel abhängige Menge an Erde und Wasser durchdrungen hat, praktisch gleich null. Die angenommene integrierte Messzeit war ein Jahr. Der ambiente Untergrund, der in der Simulation enthalten ist, wurde für diese Studie verwendet, transiente Signale wurden nicht mit eingeschlossen. Das effektive Volumen und das "transient-free, limit-setting potential" wurden für drei verschiedene sogenannte Stufen untersucht, die immer realistischer werdende Bedingungen für das ambiente Rauschen und für die Einstellungen des Online-Filters beschreiben. Die Energieschwelle des AMADEUS Detektors liegt unter idealisierten Bedingungen bei $1.8 \cdot 10^{10}$ GeV und das effektive Volumen überschreitet ein Volumen von $2\,\mathrm{km}^2$ bei 10^{12} GeV. Dies zeigt bereits das Potenzial der akustischen Detektion von UHE-Neutrinos. Das "transient-free, limit-setting potential" ist hauptsächlich durch das effektive Volumen des Detektors, das für die derzeitige Detektorkonfiguration $0.1\,\mathrm{km}^3$ bei 10^{12} GeV ist, und durch die hohe Energieschwelle von $1.8 \cdot 10^{11}$ GeV bestimmt. Das errechnete "transient-free, limit-setting potential" des AMADEUS Systems ist vielversprechend und sollte Studien für einen zukünftigen großvolumigen Detektor anregen, die auf der entwickelten Simulations- und Analysekette beruhen könnten. Des Weiteren ist mit der Bestimmung der Rate an bipolaren Untergrundsignalen ein weiterer großer Schritt in Richtung einer Entscheidung über die Machbarkeit der akustischen Neutrino-Detektion im Mittelmeer getan.

Appendix A

Acoustic Ray Tracing in Sea Water

The approaches to reconstruct the position of the acoustic source, as described on Sec. 4.3, is based on the assumption that the speed of sound in sea water does not depend on the depth of the water, so the path trough the medium is linear. The method of ray tracing is introduced here to overcome this simplification. The speed of sound in water depends on the pressure, temperature, and salinity. The speed of sound profile $c_s(z)$for a detector site can be determined by either direct measurement or implicitly by measuring the quantities mentioned above. The so-called UNESCO[1] algorithm [84–86] is the standard method for calculating the depth-dependent speed of sound from pressure, water temperature, and salinity data. The derivation of the ray tracing equation is based on Fermat's principle of least time and the calculus of variations following [87–89]. Fermat's principle can be expressed as a variation on the travel time T for a path between two fixed points:

$$\delta T = \delta \int \mathrm{d}t = 0. \tag{A.1}$$

In order to find the path that minimises the travel time, the differential time $\mathrm{d}t$ is substituted with the differential path length $\mathrm{d}t/\mathrm{d}s = 1/c_s$:

$$\delta T = \delta \int \frac{1}{c_s}\mathrm{d}s = 0. \tag{A.2}$$

The path length $\mathrm{d}s$ in Cartesian coordinates x_i is $\mathrm{d}s = \sqrt{\mathrm{d}x_i \mathrm{d}x_i}$. The Einstein summation is assumed and we will explore the effect of variation on $\mathrm{d}s$:

$$\delta \mathrm{d}s = \delta(\mathrm{d}x_i \mathrm{d}x_i)^{\frac{1}{2}} = \frac{1}{\mathrm{d}s}\delta(\mathrm{d}x_i \mathrm{d}x_i) = \frac{1}{\mathrm{d}s}\mathrm{d}x_i \delta \mathrm{d}x_i = \frac{\mathrm{d}x_i}{\mathrm{d}s}\delta \mathrm{d}x_i = \frac{\mathrm{d}x_i}{\mathrm{d}s}\frac{\delta \mathrm{d}x_i}{\mathrm{d}s}\mathrm{d}s = \dot{x}_i \delta \dot{x}_i \mathrm{d}s \tag{A.3}$$

[1] *United Nations Educational, Scientific, and Cultural Organisation* (UNESCO)

The variation on the travel time can be rewritten as:

$$\delta T = \delta \int \frac{1}{c_s} \mathrm{d}s \tag{A.4}$$

$$= \int \delta(\frac{1}{c_s} \mathrm{d}s) \tag{A.5}$$

$$= \int (\delta \frac{1}{c_s} \mathrm{d}s + \frac{1}{c_s} \delta \mathrm{d}s) \tag{A.6}$$

$$= \int (\delta \frac{1}{c_s} + \frac{1}{c_s} \dot{x}_i \delta \dot{x}_i) \mathrm{d}s \tag{A.7}$$

$$= \int (\delta x_i \partial_i \frac{1}{c_s} + \frac{1}{c_s} \dot{x}_i \delta \dot{x}_i) \mathrm{d}s \tag{A.8}$$

The term $\delta(x_i \partial_i \frac{1}{c_s})$ represents the change in the travel time caused by the variation on the speed of sound over another path and $\dot{x}_i \delta \dot{x}_i$ induce a change due to the variation on the path length.

$$\delta T = \int (\delta(x_i \partial_i \frac{1}{c_s}) + \dot{x}_i \delta \dot{x}_i) \mathrm{d}s \tag{A.9}$$

$$= \int \delta x_i \partial_i \frac{1}{c_s} \mathrm{d}s + \int \dot{x}_i \delta \dot{x}_i \mathrm{d}s \tag{A.10}$$

$$= \int \delta x_i \partial_i \frac{1}{c_s} \mathrm{d}s + [\frac{1}{c_s} \dot{x}_i \delta x_i]_a^b - \int \delta x_i \frac{d}{\mathrm{d}s} (\frac{1}{c_s} \dot{x}_i) \mathrm{d}s \tag{A.11}$$

$$= \int \delta x_i (\partial_i \frac{1}{c_s} - \frac{d}{\mathrm{d}s} (\frac{1}{c_s} \dot{x}_i)) \mathrm{d}s \tag{A.12}$$

As Eq. A.2 must be true for all variations δx_i, the term:

$$\partial_i \frac{1}{c_s} - \frac{d}{\mathrm{d}s} (\frac{1}{c_s} \dot{x}_i) = 0 \tag{A.13}$$

must vanish everywhere along the path. So we can write the ray equation as:

$$\ddot{x}_i = c_s \partial_i \frac{1}{c_s} - \dot{x}_i (\dot{x}_j c_s \partial_j \frac{1}{c_s}). \tag{A.14}$$

Under the assumption that the profile of speed of sound profile depends only on the depth, the ray propagates in the plane represented by the vertical z-axis and the horizontal r-axis. The angle θ is defined between the ray and the horizontal. In this case, $\dot{\mathbf{x}}$ can be defined as: $\dot{x} = \mathbf{e}_r \cos\theta + \mathbf{e}_z \sin\theta$ and Eq. A.14 can be rewritten as:

$$\frac{1}{c_s}(\mathbf{e}_z \cos\theta - \mathbf{e}_r \sin\theta)\dot{\theta} = \mathbf{e}_r \partial_r \frac{1}{c_s} + \mathbf{e}_z \partial_z \frac{1}{c_s} - (\mathbf{e}_r \cos\theta + \mathbf{e}_z \sin\theta)(\cos\theta \partial_r \frac{1}{c_s} + \sin\theta \partial_z \frac{1}{c_s}). \tag{A.15}$$

Now, an expression for $\dot{\theta}$ can be written as:

$$\frac{1}{c_s}\dot{\theta} = \partial_z \frac{1}{c_s} \cos\theta - \partial_r \frac{1}{c_s} \sin\theta \tag{A.16}$$

The derivative $\dot{\theta}$ can be reformulated as: $\dot{\theta} = \cos\theta d\theta/dr$. As mentioned earlier, the profile of the speed of sound profile is assumed to be only z-dependent, thus the second term can be neglected. This leads to the first ray equation:

$$\frac{\mathrm{d}\theta}{\mathrm{d}r} = -\frac{\partial_z c_s}{c_s} \tag{A.17}$$

The second ray equation can be derived from the definition of θ as:

$$\frac{\mathrm{d}z}{\mathrm{d}r} = \tan\theta \tag{A.18}$$

As the travel time of the sound wave is of interest, a third equation can be formulated as:

$$\frac{\mathrm{d}t}{\mathrm{d}r} = \frac{\sec\theta}{c_s} \tag{A.19}$$

The ray equations can be integrated using an *ordinary differential equation* (ODE) solver from initial conditions: the depth z_0 and the starting angle θ_0. The profile of the speed of sound profile at the ANTARES site (cf. Fig. A.1(a)) was measured during a sea campaign [90]. The measurement included the velocity of the sound from the surface of the sea to the sea floor in 2475 m in steps of 1 m. The measured values were interpolated using a spline interpolation and the derivatives were calculated numerically.

This ray tracing algorithm was used to determine the maximal distance accessible by the AMADEUS detector to a source located at the surface of the sea and the corresponding incident angle. This distance is about 32 km for the upmost acoustic storey, which is 410 m above the sea floor, and the incident angle is about 5.5° below the horizon. The surface of the sea area accessible by the AMADEUS detector is about 32 km^2. In Fig. A.1(b), the paths that result from the profile of the speed of sound for different starting angles is shown. It is assumed that the profile of the speed of sound and the the depth of the water do not vary in the sea surrounding the acoustic receiver. This is, however, not completely valid for the AMADEUS site, since the depth of the water decreases in the direction of the coast after about 10 km to about 200 m.

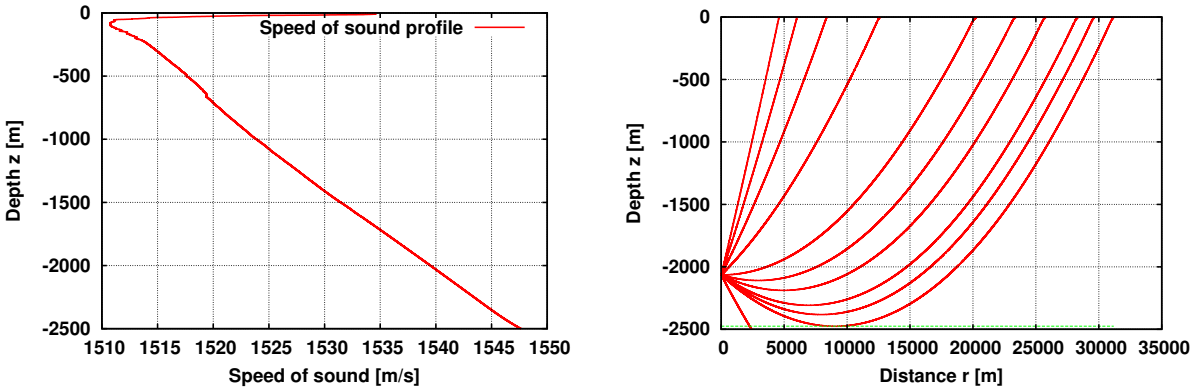
A simple ray tracing technique, assuming straight rays, was introduced in Sec. 4.3 as position reconstruction method. This was extended with the ray equations derived above in order to incorporate the effect of the speed of sound profile. The ray tracing technique was tested with the emissions of the acoustic positioning system integrated in the ANTARES detector as for the original position reconstruction method. The mean position and the uncertainty was derived by fits of Gaussian distributions to the corresponding distributions of the reconstructed x, y, and z coordinates for each of the emitters. The results are given in Fig. A.2. As mentioned in Sec. 4.3, the emitters of Line 6, 4, 9, and 12 were not operational or only partially operational at this time. The resolution of the ray tracing derived for these emissions is:

$$\begin{aligned} \langle\sigma_x\rangle &= 5.0 \pm 0.6\, m \\ \langle\sigma_y\rangle &= 5.4 \pm 0.7\, m \\ \langle\sigma_z\rangle &= 4.5 \pm 1.1\, m\,. \end{aligned}$$

The mean deviation of mean reconstructed position and the true position of the emitters is:

$$\begin{aligned} \langle x\rangle &= 1.0 \pm 0.5\, m \\ \langle y\rangle &= 1.8 \pm 0.4\, m \\ \langle z\rangle &= 23.9 \pm 0.5\, m\,. \end{aligned}$$

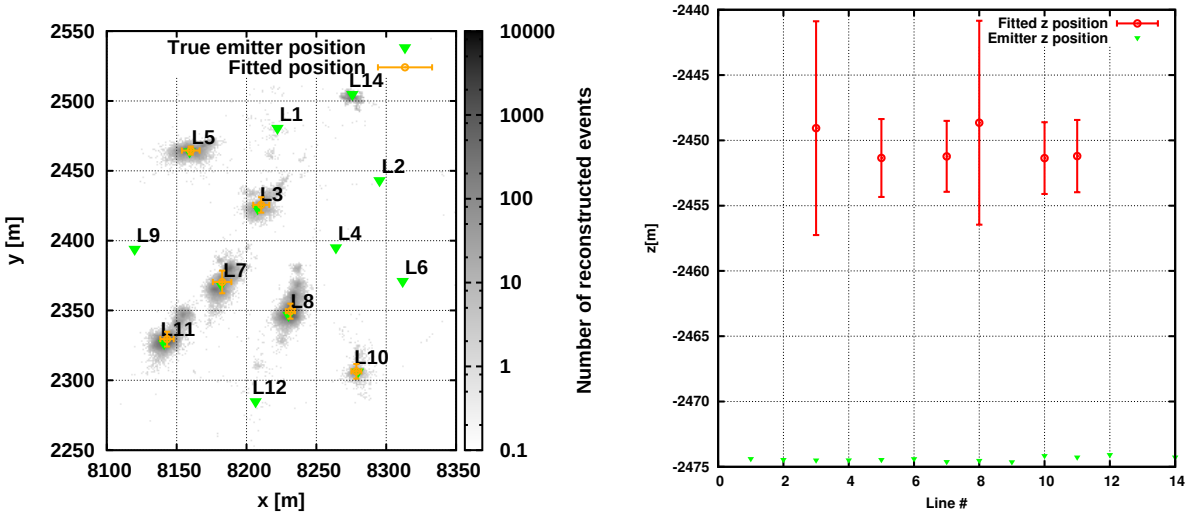
The resolution if this ray tracing technique is better than 6 m for each coordinate, which is in agreement with the result of about 5 m provided by the straight-ray method. The reconstructed position of the emitters is in agreement with their true position only for the x and y coordinates. For z, a significant offset of about 24 m between the reconstructed coordinate and the true coordinate is seen. Although the effect of profile of the speed of sound should be less than a few metres for the distances of a few 100 m between the emitters of the positioning system and acoustic storeys in the AMADEUS detector. This was tested with the simulation in which it was assumed that the velocity of sound did not change. No positive indication of the origin of the offset had been determined at the time of writing. Also, the integration of the effect of the profile of the velocity of sound into the simulation would require further work that is beyond the scope of this thesis.



(a) Profile of the speed of sound at ANTARES site

(b) Example of ray tracing

Figure A.1: (a) The profile of the speed of sound as a function of the depth, which was derived from measurements at the ANTARES site. (b) An example of the ray tracing technique with rays starting at a depth of 2065 m (upmost acoustic storey) for different starting angles. The sea floor is at 2475 m (dashed line).



(a) Ray tracing x,y-coordinate

(b) Ray tracing z-coordinate

Figure A.2: (a) The distribution of the x and y coordinates of the reconstructed positions. Also, the fitted position and the uncertainty are given (orange markers) and the true position of the emitters (green markers). (b) The fitted z positions of the emitters (red markers) and their uncertainty are shown, and the true z-coordinate of the emitters is given.

Bibliography

[1] G. Askariyan, B. Dolgoshein, A. Kalinovsky, and N. Mokhov, "Acoustic detection of high energy particle showers in water," *Nucl. Instrum. Meth.*, vol. 164, no. 2, pp. 267–278, 1979.

[2] F. H. Fisher and V. P. Simmons, "Sound absorption in sea water," *J. Acoust. Soc. Am.*, vol. 62, no. 3, pp. 558–564, 1977.

[3] S. Boeser, *et al.*, "Feasibility of acoustic neutrino detection in ice: Design and performance of the South Pole Acoustic Test Setup (SPATS)," *arXiv preprint*, 2008, arXiv:0807.4676.

[4] K. Antipin *et al.*, "A prototype device for acoustic neutrino detection in Lake Baikal," in *Proceedings of the 30th International Cosmic Ray Conference (ICRC2007), Merida, Mexico, July 3–11*, 2007, arXiv:0710.3113.

[5] J. Beringer *et al.*, "The Review of Particle Physics," *Phys. Rev. D*, vol. 86, no. 010001, 2012. Particle Data Group.

[6] H. Athar, M. Jeżabek, and O. Yasuda, "Effect of neutrino mixing on high-energy cosmic neutrino flux," vol. D 62, p. 103007, 2000, arXiv:hep-ph/0005104v2.

[7] R. Gandhi, C. Quigg, M. Hall Reno, and I. Sarcevic, "Ultrahigh-energy neutrino interactions," *Astropart. Phys.*, vol. 5, no. 2, pp. 81–110, 1996, arXiv:hep-ph/9512364v1.

[8] A. Connolly, R. Thorne, and D. Waters, "Calculation of high energy neutrino-nucleon cross sections and uncertainties using the Martin-Stirling-Thorne-Watt parton distribution functions and implications for future experiments," *Phys. Rev. D*, vol. 83, no. 11, p. 113009, 2011, arXiv:1102.0691v2.

[9] A. Cooper-Sarkar, P. Mertsch, and S. Sarkar, "The high energy neutrino cross-section in the Standard Model and its uncertainty," *J. High Energ. Phys.*, vol. 2011, no. 8, pp. 1–20, 2011, arXiv:1106.3723v2.

[10] A. M. Dziewonski and D. L. Anderson, "Preliminary reference Earth model," *Phys. Earth Planet. Inter.*, vol. 25, no. 4, pp. 297 – 356, 1981.

[11] V. Berezinsky, "High Energy Neutrino Astronomy," 02 2011, arXiv:1102.3591v1.

[12] E. Fermi, "Galactic Magnetic Fields and the Origin of Cosmic Radiation.," *Astrophys. J.*, vol. 119, p. 1, Jan. 1954.

[13] E. Fermi, "On the Origin of the Cosmic Radiation," *Physical Review*, vol. 75, no. 8, pp. 1169–1174, 1949.

[14] F. Rieger, V. Bosch-Ramon, and P. Duffy, "Fermi acceleration in astrophysical jets," *Astrophys. Space. Sci.*, vol. 309, no. 1-4, pp. 119–125, 2007, arXiv:astro-ph/0610141.

[15] A. M. Hillas, "The Origin of Ultra-High-Energy Cosmic Rays," *Annu. Rev. Astron. Astrophys.*, vol. 22, pp. 425–444, 1984.

[16] N. Gehrels and P. Mészáros, "Gamma-Ray Bursts," *Science*, vol. 337, no. 6097, pp. 932–936, 2012.

[17] S. Woosley and J. Bloom, "The Supernova Gamma-Ray Burst Connection," *Annu. Rev. Astron. Astrophys.*, vol. 44, pp. 507–556, 2006, arXiv:astro-ph/0609142.

[18] S. Bianchi, R. Maiolino, and G. Risaliti, "AGN Obscuration and the Unified Model," *Adv. Astron.*, vol. 2012, p. 17, 2012. Article ID 782030.

[19] P. Padovani, "Gamma-Ray Emitting AGN and Unified Schemes," in *Very High Energy Phenomena in the Universe; Moriond Workshop* (Y. Giraud-Heraud and J. Tran Thanh van, eds.), p. 7, 1997, arXiv:astro-ph/9704002.

[20] K. Greisen, "End to the Cosmic-Ray Spectrum?," *Phys. Rev. Lett.*, vol. 16, pp. 748–750, Apr 1966.

[21] G.T. Zatsepin and V.A. Kuz'min, "Upper limit of the spectrum of cosmic rays," *J. Exp. Theor. Phys.*, vol. 4, p. 78, 1966. Russian original: Pis'ma Zh. Éksp. Teor. Fiz. **4** (1966) 114.

[22] J. Abraham, *et al.*, "Measurement of the energy spectrum of cosmic rays above 10^{18} eV using the Pierre Auger Observatory," *Phys. Lett. B*, vol. 685, no. 4–5, pp. 239–246, 2010.

[23] P. Sokolsky *et al.*, "Final Results from the High Resolution Fly's Eye (HiRes) Experiment," *arXiv preprint*, 2010, arXiv:1010.2690.

[24] P. Abreu, *et al.*, "The Pierre Auger Observatory II: Studies of Cosmic Ray Composition and Hadronic Interaction models," *arXiv preprint*, 2011, arXiv:1107.4804.

[25] T. Weiler, "Resonant Absorption of Cosmic-Ray Neutrinos by the Relic-Neutrino Background," *Phys. Rev. Lett.*, vol. 49, pp. 234–237, Jul 1982, arXiv:astro-ph/9803095.

[26] T. J. Weiler, "Cosmic-ray neutrino annihilation on relic neutrinos revisited: a mechanism for generating air showers above the Greisen-Zatsepin-Kuzmin cutoff," *Astro. Phys.*, vol. 11, no. 3, pp. 303 – 316, 1999, arXiv:hep-ph/9710431.

[27] R. E. Lopez, S. Dodelson, A. Heckler, and M. S. Turner, "Precision Detection of the Cosmic Neutrino Background," *Phys. Rev. Lett.*, vol. 82, pp. 3952–3955, May 1999.

[28] G. P. W. *et al.*, "Observational Constraints on the Ultra-high Energy Cosmic Neutrino Flux from the Second Flight of the ANITA Experiment," vol. D 82, p. 022004, 2010, arXiv:1003.2961v3. Erratum: arXiv:1011.5004v1.

[29] D. V. Semikoz and G. Sigl, "Ultra-high energy neutrino fluxes: new constraints and implications," *J. Cosmol. Astropart. Phys.*, vol. 2004, no. 04, p. 003, 2004, arXiv:hep-ph/0309328.

[30] T. Totani, "GRB Sketch." Homepage, 1 2009. http://www.kusastro.kyoto-u.ac.jp/~totani/en/grb_image/index.html.

[31] "Optical Cherenkow Detection Principle Sketch." ANTARES Homepage, 2 2011. http://antares.in2p3.fr/Gallery/3D/index.html.

[32] J. G. Learned, "Acoustic radiation by charged atomic particles in liquids: An analysis," *Phys. Rev. D*, vol. 19, pp. 3293–3307, Jun 1979.

[33] K. Huang, *Statistical Mechanics*. John Wiley & Sons, 2nd ed., 1987. ISBN 0-471-81518-7.

[34] L.D. Landau and E.M. Lifschitz, *Course of Theoretical Physics, Vol.6: Fluid Mechanics*. Pergamon Press, 2nd ed., 1987.

[35] K. Graf, *et al.*, "Testing Thermo-acoustic Sound Generation in Water with Proton and Laser Beams," *Int. J. Mod. Phys. A21S*, vol. 1, pp. 127–131, 2006, arXiv:astro-ph/0509450.

[36] S. Bevan, *et al.*, "Simulation of Ultra High Energy Neutrino Interactions in Ice and Water," *Astropart. Phys.*, vol. 28, no. 3, pp. 366–379, 2007, arXiv:0704.1025.

[37] S. Bevan, *et al.*, "Study of the acoustic signature of UHE neutrino interactions in water and ice," *Nucl. Instrum. Meth. A*, vol. 607, no. 2, pp. 398–411, 2009, arXiv:0903.0949.

[38] T. Sloan, "Simulation of Cosmic Ray neutrinos Interactions in Water," *J. Phys. Conf. Ser.*, vol. 81, no. 1, p. 012001, 2007, arXiv:astro-ph/0609617.

[39] M. Ageron *et al.*, "ANTARES: the first undersea neutrino telescope," *Nucl. Instrum. Meth. A*, vol. 656, no. 1, pp. 11–38, 2011, arXiv:1104.1607.

[40] P. Amram, *et al.*, "The ANTARES optical module," *Nucl. Instrum. Meth. A*, vol. 484, no. 1, pp. 369–383, 2002, arXiv:astro-ph/0112172v1.

[41] S. Adrián-Martínez, *et al.*, "The positioning system of the ANTARES Neutrino Telescope," *J. Inst.*, vol. 7, no. 08, p. T08002, 2012, arXiv:1202.3894.

[42] J. Aguilar *et al.*, "AMADEUS - The Acoustic Neutrino Detection Test System of the ANTARES Deep-Sea Neutrino Telescope," *Nucl. Instrum. Meth. A*, vol. 626–627, pp. 128–143, 09 201, arXiv:1009.4179.

[43] K. Graf, *et al.*, "Integration of Acoustic Neutrino Detection Methods into ANTARES," in *Journal of Physics: Conference Series*, vol. 81, p. 012012, 2007, arXiv:astro-ph/0703442.

[44] A. Enzenhöfer, *et al.*, "Development of combined Opto-Acoustical sensor Modules," in *Proceedings of the ARENA 2010*, vol. 662, pp. S203–S205, Elsevier, 2012, arXiv:1104.3061.

[45] C. L. Naumann, *et al.*, "Development of Acoustic Sensors for the ANTARES Experiment," *Int. J. Mod. Phys. A21S*, vol. 1, pp. 92–96, 2006, arXiv:astro-ph/0511243.

[46] C.L. Naumann, *Development of Sensors for the Acoustic Detection of Ultra High Energy Neutrinos in the Deep Sea*. PhD thesis, Univ. Erlangen-Nürnberg, FAU-PI4-DISS-07-002, 2007.
obtainable from http://www.antares.physik.uni-erlangen.de/publications.

[47] J. Aguilar, *et al.*, "The data acquisition system for the ANTARES neutrino telescope," *Nucl. Instrum. Meth. A*, vol. 570, no. 1, pp. 107–116, 2007, arXiv:astro-ph/0610029v1.

[48] M. Neff, "Studie zur akustischen Teilchendetektion im Rahmen des ANTARES-Experiments: Entwicklung und Integration von Datennahmesoftware," diploma thesis, Univ. Erlangen-Nürnberg, FAU-PI1-DIPL-07-003, 2007.
obtainable from http://www.antares.physik.uni-erlangen.de/publications.

[49] K. Graf, *Experimental Studies within ANTARES towards Acoustic Detection of Ultra-High Energy Neutrinos in the Deep Sea*. PhD thesis, Univ. Erlangen-Nürnberg, FAU-PI1-DISS-08-001, 2008.
obtainable from http://www.antares.physik.uni-erlangen.de/publications.

[50] R. Urick, *Ambient noise in the sea*. Peninsula Publishing, 1986.

[51] C. Kopper, "A software framework for KM3NeT," in *Proceedings of the 3rd International Workshop VLVnT*, vol. 602, pp. 107–110, 2009.

[52] A. Würstlein, "Development of a Software Framework for the ANTARES Acoustic Data and Simulations within the Framework," diploma thesis, Univ. Erlangen-Nürnberg, no. ECAP-2010-047, 2009.
obtainable from http://www.antares.physik.uni-erlangen.de/publications.

[53] M. Neff *et al.*, "Simulation Chain for Acoustic Ultra-high Energy Neutrino Detectors." acepted for publication in Nuclear Inst. and Methods in Physics Research, A as part of the proceedings for the 4rd International Workshop VLVnT 2011, 2012.

[54] R. Gandhi, C. Quigg, M. H. Reno, and I. Sarcevic, "Neutrino interactions at ultrahigh energies," *Phys. Rev. D*, vol. 58, p. 093009, Sep 1998, arXiv:hep-ph/9807264.

[55] D. Heck, *et al.*, *CORSIKA: a Monte Carlo code to simulate extensive air showers.*, vol. 6019. FZKA, 1998.

[56] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 3rd ed., Sept. 2007.

[57] C. Eckart and G. Young, "The approximation of one matrix by another low rank," *Psychometrika*, vol. 1, pp. 211-218, 1936.

[58] M. A. Ainslie and J. G. McColm, "A simplified formula for viscous and chemical absorption in sea water," *J. Acoust. Soc. Am.*, vol. 103, pp. 1671–1672, Mar. 1998.

[59] L. N. Liebermann, "The Origin of Sound Absorption in Water and in Sea Water," *J. Acoust. Soc. Am.*, vol. 20, no. 6, pp. 868–873, 1948.

[60] V. Niess and V. Bertin, "Underwater acoustic detection of ultra high energy neutrinos," *Astroparticle Physics*, vol. 26, no. 4, pp. 243–256, 2006, arXiv:astro-ph/0511617.

[61] V. Knudsen, R. Alford, and J. Emling, "Underwater ambient noise," *J. Mar. Res*, vol. 7, no. 3, pp. 410–429, 1948.

[62] R. Lahmann, "Status and recent results of the acoustic neutrino detection test system AMADEUS," in *Proceedings of the ARENA 2010*, vol. 662, Supplement 1, pp. S216 – S221, 2012, arXiv:1104.3041.

[63] Xilinx, Inc., *XILINX LogiCORE FIR Compiler v1.0*, 2006. DS534, available at http://www.xilinx.com.

[64] M. Neff, *et al.*, "AMADEUS on-line trigger and filtering methods," in *ARENA 2008* (F. Ameli *et al.*, eds.), vol. 604, pp. S185 – S188, 2009.

[65] H. Akima, "A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures," *J. ACM*, vol. 17, pp. 589–602, Oct. 1970.

[66] C. Richardt, *et al.*, "Reconstruction methods for acoustic particle detection in the deep sea using clusters of hydrophones," *Astro. Phys.*, vol. 31, no. 1, pp. 19–23, 2009, arXiv:0906.1718.

[67] C. Richardt, *Acoustic particle detection – Direction and source location reconstruction techniques*. PhD thesis, Univ. Erlangen-Nürnberg, ECAP-2010-046, 2010. obtainable from http://www.antares.physik.uni-erlangen.de/publications.

[68] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization," *Comput. J.*, vol. 7, no. 4, pp. 308–313, 1965.

[69] V. Bertin, "Compass-board readout with reduced time interval." private communication.

[70] M. Neff, *et al.*, "Signal classification for acoustic neutrino detection," in *4th International workshop on Acoustic and Radio EeV Neutrino detection Activities* (W. Barletta *et al.*, eds.), vol. 662, Supplement 1, pp. S242 – S245, 2012, arXiv:1104.3248.

[71] OpenCV, "Online documention of the Open Source Computer Vision library." http://opencv.willowgarage.com/, August 2011.

[72] D. Lewis, "Naive (Bayes) at forty: The independence assumption in information retrieval," in *Machine Learning: ECML-98* (C. Nédellec and C. Rouveirol, eds.), vol. 1398 of *Lecture Notes in Computer Science*, pp. 4–15, Springer Berlin Heidelberg, 1998.

[73] L. Breiman, *Classification and regression trees*. Belmont, Calif: Wadsworth International Group, 1984.

[74] L. E. Raileanu and K. Stoffel, "Theoretical Comparison between the Gini Index and Information Gain Criteria," *Ann. Math. Artif. Intel.*, vol. 41, pp. 77–93, 2000.

[75] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[76] Y. Freund and R. Schapire, "A desicion-theoretic generalization of on-line learning and an application to boosting," in *Computational Learning Theory* (P. Vitányi, ed.), vol. 904 of *Lecture Notes in Computer Science*, pp. 23–37, Springer Berlin Heidelberg, 1995.

[77] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[78] C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discov.*, vol. 2, no. 2, pp. 121–167, 1998.

[79] M. Ester, H. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the 2nd International Conference on Knowledge Discovery and Data mining*, vol. 1996, pp. 226–231, AAAI Press, 1996.

[80] R. Lahmann, "Ultra-High-Energy Neutrinos and Their Acoustic Detection in the Sea," diploma thesis, Univ. Erlangen-Nürnberg, ECAP, 2012.
obtainable from `http://www.antares.physik.uni-erlangen.de/publications`.

[81] N. G. Lehtinen, P. W. Gorham, A. R. Jacobson, and R. A. Roussel-Dupré, "FORTE satellite constraints on ultrahigh energy cosmic particle fluxes," *Phys. Rev. D*, vol. 69, p. 013008, Jan 2004, `arXiv:astro-ph/0309656`.

[82] G. J. Feldman and R. D. Cousins, "Unified approach to the classical statistical analysis of small signals," *Phys. Rev. D*, vol. 57, pp. 3873–3889, Apr 1998, `arXiv:physics/9711021`.

[83] K. Kotera, D. Allard, and A. Olinto, "Cosmogenic Neutrinos: parameter space and detectabilty from PeV to ZeV," *J. Cosmol. Astropart. Phys.*, vol. 2010, no. 10, p. 013, 2010, `arXiv:1009.1382`.

[84] C.-T. Chen and F. J. Millero, "Speed of sound in seawater at high pressures," *J. Acoust. Soc. Am.*, vol. 62, no. 5, pp. 1129–1135, 1977.

[85] V. A. D. Grosso, "New equation for the speed of sound in natural waters (with comparisons to other equations)," *J. Acoust. Soc. Am.*, vol. 56, no. 4, pp. 1084–1091, 1974.

[86] C. C. Leroy, S. P. Robinson, and M. J. Goldsmith, "A new equation for the accurate calculation of sound speed in all oceans," *J. Acoust. Soc. Am.*, vol. 124, no. 5, pp. 2774–2782, 2008.

[87] J. Bowlin, J. Spiesberger, T. Duda, and L. Freitag, *Ocean Acoustical Ray-Tracing Software RAY.* Defense Technical Information Center, 1992.

[88] L. Brekhovskikh and Y. Lysanov, *Fundamentals of ocean acoustics*. Springer, 2003.

[89] T. Duda, "Derivation of ray equations in moving media in angle/depth form." aviable at `http://www.whoi.edu/science/AOPE/cofdl/tim/Ray_eqs.pdf`, 7 2005.

[90] D. Lefèvre, "Data from an INSU sea campaign at the ANTARES site March 2010." private communication, 11 2010.

List of Abbreviations and Acronyms

ANTARES Astronomy with a Neutrino Telescope and Abyss Environmental Research . . . 89

AMADEUS ANTARES Modules for Acoustic Detection under the Sea 89

SPATS South Pole Acoustic Test Setup 6

ACoRNE Acoustic Cosmic Ray Neutrino Experiment 33

CORSIKA Cosmic Ray Simulations for Kascade 33

ECAP Erlangen Centre for Astroparticle Physics 24

UHE ultra-high-energy 89

BIP bipolar pressure pulse 89

OM Optical Module 21

AM Acoustic Module 23

LCM Local Control Module 21

IL07 Instrumentation Line deployed 2007, redeployed 2011 21

MC Monte Carlo 35

SVD singular value decomposition 34

FT Fourier transformation 37

FFT fast Fourier transformation 37

DAQ data acquisition 25

HTI HighTech Inc. hydrophone . . . 24

LTI custom-designed hydrophone . . . 24

AcouADC acoustic analog-to-digital converter board . . . 115

ADC analog-to-digital converter . . . 27

FPGA field-programmable gate array . . . 27

μC micro-controller . . . 28

CC charged current . . . 11

NC neutral current . . . 11

DIS deep inelastic scattering . . . 13

CL confidence level . . . 84

PSD power spectral density . . . 37

PREM Preliminary Reference Earth Model . . . 13

UNESCO United Nations Educational, Scientific, and Cultural Organisation Intergovernmental Oceanographic Commission . . . 97

DBSCAN density-based spatial clustering of applications with noise . . . 73

PMNS matrix Pontecorvo-Maki-Nakagawa-Sakata matrix . . . 10

GRB gamma-ray burst . . . 15

AGN active galatic nucleus . . . 15

CMB cosmic microwave background . . . 15

CνB cosmic neutrino background . . . 16

UHECR ultra-high-energy cosmic ray . . . 5

GZK Greisen-Zatsepin-Kuzmin . . . 15

PMT photomultiplier tube . . . 5

HWHM half width at half maximum 64

FWHM full width at half maximum 40

RMS root mean square 55

oob out-of-bag 69

ODE ordinary differential equation 99

List of Figures

List of Tables