Printed by Presses Universitaires de Bruxelles, December 2014
Cover design in collaboration with Yiannis Papadopoulos

Development of a sub-glacial array of radio antennas for the detection of the flux of GZK neutrinos

Thesis submitted for the degree of Doctor of Science by

Thomas Meures

Prepared under the supervision of

Prof. Dr. Kael Hanson

At the Interuniversity Institute for High Energies
(ULB - VUB)

Jury members:

Prof. Dr. Stijn Buitink
Prof. Dr. Kael Hanson
Dr. Gilles De Lentdecker
Prof. Dr. Olaf Scholten
Prof. Dr. Petr Tiniakov

Université Libre de Bruxelles, October 2014

Abstract

GZK neutrino are interesting messenger particles since, if detected, they can transmit us exclusive information about ultra-high energy processes in the Universe. These particles, which hold energies above 10^{16} eV, interact very rarely. Therefore, detectors of several gigatons of matter are needed to discover them. The ARA detector is planned and currently being constructed at the South Pole. It is designed to use the Askaryan effect, the emission of radio waves from neutrino induced cascades in the South Pole ice, to detect neutrino interactions. With antennas distributed in 37 stations in the ice, such interactions can be observed in a volume of several hundred cubic kilometers. Currently 2 ARA stations have been deployed in the ice and are taking data since the beginning of the year 2013.

The first part of this thesis summarizes the current theories concerning the GZK mechanism and the Askaryan effect to explain the interest in GZK neutrinos and in the used detection method.

In the second part the ARA detector is described and calibrations of different detector parts are presented. In this work, the digitization chips have been calibrated concerning their timing precision and signal amplitude. In this way a timing precision of 100 ps between antennas could be achieved. Furthermore, the geometry of the antenna clusters is determined by fits based on external signals to allow for a proper radio vertex reconstruction.

In the third part of the thesis the development of methods to distinguish radio signals from thermal noise are presented. Moreover, a reconstruction method, developed to determine the position of radio sources, is described. With only two stations operational a discovery of GZK neutrinos is not expected and in fact no signal candidate has been found in the analysis of the data. A neutrino flux limit is calculated. This limit is not competitive yet with the current best limits, but very promising for the full ARA detector. The work shows that after completion this detector is expected to be capable of a neutrino discovery.

Résumé

Les neutrinos GZK sont des messagers intéressants étant donné qu'ils peuvent nous apporter des informations exclusives sur des processus d'ultra-haute énergie se déroulant dans l'Univers. Ces particules, qui ont des énergies au-dessus de 10^{16} eV, interagissent très rarement. Dès lors, des détecteurs de plusieurs gigatonnes de matière sont nécessaires pour découvrir ces neutrinos GZK. Le détecteur ARA a été conçu de manière à utiliser l'effet Askaryan, l'émission d'ondes radio par des cascades électromagnétiques produites par les neutrinos dans la glace du Pôle Sud, afin de révéler les interactions des neutrinos. Grâce à des antennes distribuées en 37 stations dans la glace, les interactions des neutrinos peuvent être observées dans un volume de plusieurs kilomètres cubes. Actuellement, 2 stations d'ARA ont été déployées dans la glace et enregistrent des données depuis le début de l'année 2013.

La première partie de cette thèse décrit les théories actuelles concernant le mécanisme GZK et l'effet Askaryan afin d'expliquer l'intérêt de ces neutrinos et la façon de les détecter.

Dans la deuxième partie, une description du détecteur ARA et de quelques-unes de ses calibrations est présentée. Dans le cadre de ce travail, les puces de digitalisation ont été calibrées du point de vue de la précision temporelle et de l'amplitude des signaux. De cette manière, une précision en temps de 100 ps a été obtenue. De plus, la géométrie des groupes d'antennes a été déterminée dans des fits des signaux externes, pour permettre des reconstructions précises de l'origine des ondes radio.

La troisième partie de cette thèse présente une description du développement des méthodes utilisées afin de distinguer les signaux radio du bruit de fond thermique. Un algorithme de reconstruction, développé pour déterminer la position de l'émission des ondes radio enregistrées, y est également présenté. En utilisant ces méthodes et algorithmes, les données de la première année d'opération du détecteur ont été analysées. Avec seulement deux stations opérationnelles, on ne s'attend pas à la découverte de neutrinos GZK et, en effet, aucun candidat neutrino n'a été trouvé dans les données analysées. Néanmoins, ce travail démontre qu'une recherche de ces neutrinos est possible avec ARA. Une limite sur le flux des neutrinos GZK a été calculée. Bien que non compétitive avec les meilleures limites, la limite obtenue promet beaucoup quant à la sensibilité du détecteur complet d'ARA.

Contents

II Detector hardware and calibration efforts 33

Appendices

1

Introduction

The investigations, conducted in the frame of this thesis, are dedicated to the development and verification of techniques for the detection of ultra-high energy neutrinos, using the coherent radio emission from neutrino-induced cascades. Such neutrinos are particularly interesting since they can provide us with exclusive information about ultra-high energy processes in the Universe. The described research is embedded in the field of high energy astroparticle physics which will be shortly introduced in the first part of this chapter. Furthermore, neutrinos, their properties and their advantages as cosmic messenger particles are discussed. In the last section an overview about the content of this thesis is presented.

1.1 Ultra-high energy astroparticle physics

The era of astroparticle physics started with the discovery of cosmic rays in 1912 by Victor Hess [1]. He measured an increase of radiation with rising altitude. This increase could only be explained by radiation coming from above and being absorbed in the atmosphere. Nowadays we know that this extraterrestrial radiation is composed of electromagnetic waves in different frequency regimes and of particles. The highest energy cosmic radiation is generally divided into three parts: gamma rays, neutrinos and cosmic rays, where “cosmic rays” refers to all charged particles approaching us from outside our atmosphere. These particles are produced with different energies in interactions within various cosmic phenomena and therefore carry information about a large variety of processes from different stages of the Universe. This thesis focuses on neutrinos that can be produced alongside with cosmic rays or in the interaction of cosmic rays with other particles or electromagnetic radiation.

Since the first observation, many experiments have been conducted to investigate cosmic rays. Widely used measurement techniques were electrometers (Hess), bubble and cloud chambers (Anderson [2]) and later photo-multiplier and fluorescence techniques, which are still employed in modern cosmic ray detectors. There are three main detector types: ground based detectors, detectors mounted on balloons or detectors mounted on satellites. The ground based experiments usually look for cosmic ray induced air showers, while the balloon and satellite based detectors measure the primary cosmic ray particle.

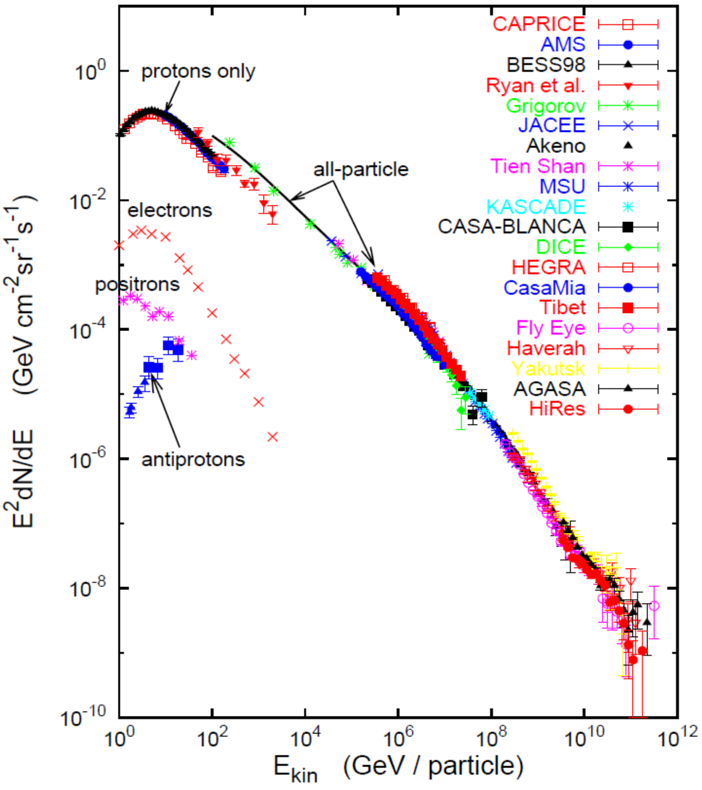


Figure 1.1: *The cosmic ray flux spectrum of various particle components versus energy as a summary of measurements until 2005 [3].*

The photo-multiplier and fluorescence techniques allow to cost efficiently build ground based detectors with a large field of view. This makes them especially interesting for the detection of highest energy cosmic rays. Their flux dN/dE drops steeply with rising energy E, on average as $dN/dE = E^{-3}$, as shown in Figure 1.1. Therefore, ultra-high energy cosmic ray interactions are very rare and large detectors are needed to collect statistically significant event samples. Experiments based on the photo-multiplier and fluorescence techniques are for example the two first large scale air-shower arrays: the Akeno Giant Air Shower Array (AGASA) [4] and the High Resolution Fly's Eye Observatory (HiRes) [5]. Besides these widely used methods, the detection of radio emission from cosmic-ray induced air showers has developed recently from a research and development phase into a competitive measurement technique. A prominent representative detector using such technology is the LOFAR detector in the Netherlands [6]. Figure 1.1 shows the cosmic ray spectrum as measured before 2006 by the most important cosmic ray experiments until then [3]. These experiments represent all three detector types: ground based, balloon based and satellite based.

In the meanwhile, the shown measurements have been refined, amongst others by the Pierre Auger Observatory [7], Telescope Array [8] and IceTop [9]. The spectrum is well measured and the average energy dependence is well understood. However, large scale features in the cosmic ray spectrum are not completely explained yet. These structures are kinks in the spectrum, normally referred to as knee at around 10^{15} eV, second knee at around 10^{17} eV and ankle at around 10^{19} eV. These features are suspected to be transition regions for different acceleration mechanisms and the confinement of cosmic rays in astronomical objects. However, these hypotheses could not be confirmed by now. Interesting at the highest energies is the ankle, which may indicate the transition region from dominantly galactic cosmic rays to extra-galactic cosmic rays. Our galaxy is regarded as too small to accelerate particles to energies above 10^{19} eV which is why particles at those energies are expected to originate from extra-galactic objects. The cosmic ray spectrum ends with a cutoff which is currently expected to be caused by the so called GZK mechanism, named after its explorers Greisen, Zatsepin and Kuzmin. This mechanism will be discussed in detail in Chapter 2.
The study of ultra-high energy cosmic rays is particularly interesting since they are produced in processes with collision energies unreachable by man-made experiments. Investigations can thus reveal new phenomena in the field of particle physics. In addition to that, the research on cosmic rays allows a deep understanding of high energy processes in the universe. The most prominent source candidates for ultra-high energy cosmic rays are Supernovae, Gamma-ray Bursts and Active Galactic Nuclei. A possible correlation of cosmic rays to such sources is limited due to deflections of the charged particles by magnetic fields. In this respect an alternative cosmic messenger, the neutrino, is interesting since its path through the universe is widely undisturbed, as will be illustrated in the next section.

1.2 The neutrino - a perfect cosmic messenger?

The neutrino is an elementary particle and part of the Standard Model of particle physics which comprises all known fundamental particles and the interactions between them. The neutrino does not hold an electric charge, its mass is very low and it only interacts weakly. Its interaction cross section with matter is extremely small [10, 11]. The neutrino was postulated by Wolfgang Pauli in 1930 to explain the energy spectrum of β-decays [12]. In spite of abundant existence, the neutrino was only discovered in 1956 by Frederick Reines [13, 14] due to its low interaction rate.

There are three types of neutrinos, the muon neutrino ν_μ, the electron neutrino ν_e and the tau neutrino ν_τ. They are the lightest leptons with a yet undetermined mass. Only upper limits could be placed for the sum of all three neutrino masses by investigations of the cosmic microwave background. This limit is currently at 0.44 eV [15, 16]. Experiments like KATRIN are dedicated to measure the neutrino mass with a sensitivity of 0.2 eV in the β-decay of tritium [17]. In the Standard Model of particle physics, the neutrino was assumed to hold no mass. Only the discovery of neutrino oscillations with Super Kamiokande in 1998 excluded the possibility of a massless particle [18]. The neutrino masses are a mixture of three mass eigenstates and oscillations between them are possible. The mixing matrix for such oscillations is understood to a high level and matrix elements are measured with a high precision by different experiments (Super Kamiokande [19], Daya Bay [20], Double Chooz [21], RENO [22]). However, not all mysteries about neutrinos have been resolved yet. To explain anomalies in oscillations of reactor neutrinos [23], the existence of right handed sterile neutrinos [24] is currently being investigated. Further unknown parameters are the absolute values of the mass eigenstates and the question if a neutrino is its own anti-particle [25].

The above mentioned characteristics of neutrinos render them a very interesting messenger particle for cosmic phenomena. Cosmic ray particles all have an electric charge. They are deflected by magnetic fields and can be absorbed in matter. Also gamma rays are eventually blocked by opaque objects. The neutrino instead, due to its low mass, the low interaction cross section and the nonexistent charge, can pass cosmological distances nearly unaffected by magnetic fields and surrounding objects. Neutrinos can thus be used to observe very distant sources with a relatively high angular precision. However, the detection of neutrinos is very challenging due to their low interaction cross section. Most neutrinos we see at Earth are generated by cosmic rays that hit our atmosphere. Up to an energy of 100 TeV, they dominate the observed neutrino flux. Cosmic neutrinos can only be discovered as deviations from this flux either in number or in energy. From the expected flux and the measured interaction cross sections it was deduced that a gigaton detector is needed to realize neutrino astronomy [26].
A few selected objects could already be investigated with smaller scale detectors:

- The first discovered extraterrestrial neutrinos came from the sun and were observed by the Homestake Chlorine Detector [27]. The detected particles allowed for a good improvement in the understanding of particle physics processes in the sun.

- In 1987, neutrinos with energies in the MeV range, originating from the supernova explosion SN1987A, were observed, leading to an improved understanding of astrophysical processes [28]. This was a milestone in neutrino astrophysics since the first neutrinos from outside our solar system had been observed. The neutrinos were measured by the Kamioka Observatory [29], the Irvine-Michigan-Brookhaven detector (IMB) [30] and the Baksan Neutrino Observatory [31]. Using the difference in arrival time of neutrinos and visible light from this supernova, an upper limit of 10 eV could be placed on the mass of ν_e.

The currently most sensitive detector for cosmic neutrinos with energies between 10^{10} eV and 10^{18} eV is the IceCube Observatory [32]. Recently, IceCube measured tens of neutrinos at much higher energies than those from SN1987A [33]. Due to their energy distribution they are inconsistent with an atmospheric neutrino flux. Events have been found with energies in the upper TeV and lower PeV range, which are the highest energy neutrinos ever observed. With such a discovery, IceCube leads the transition from neutrino searches to neutrino astronomy. However, one should note that the found neutrinos are limited to an energy range of up to roughly 2 PeV.
A very interesting part of the incoming neutrino spectrum, the GZK neutrinos, has not been measured yet. Such highly energetic neutrinos are expected to be produced in the GZK mechanism, which will be explained in Chapter 2. In this mechanism ultra-high energy cosmic rays interact with the cosmic microwave background radiation and produce a neutrino flux, expected to peak at particle energies around 10^{18} eV. The cross section for this interaction has a resonance at cosmic ray energies of roughly 10^{19} eV which could explain the cutoff seen in the cosmic ray spectrum. The resulting neutrino flux can give information about the acceleration mechanism, the source distribution and the composition of ultra-high energy cosmic rays. However, the neutrino flux at Earth is expected to be very low and detectors even bigger than the IceCube detector are needed to efficiently measure GZK neutrinos.
This thesis is dedicated to a neutrino search with the **A**skaryan **R**adio **A**rray (ARA), a neutrino detector at South Pole. As we will see, this detector is specialized for the discovery of neutrinos with energies above 10 PeV, the region of GZK neutrinos.

1.3 Thesis overview

This thesis is divided into three parts:

- a summary of the theory which the performed research is based on,
- the description of the instrument used and the calibration of different components,
- the presentation of algorithms for the analysis of recorded data concerning neutrino candidates.

A short overview of these parts is given in the following.

I Theory

The following two chapters of this thesis are dedicated to explain the theoretical background of the performed research. In Chapter 2 the **GZK mechanism** will be explained. This includes the presentation of resulting neutrino fluxes, the dependencies on cosmic

ray properties and a discussion of the potential information which can be deduced from a measured neutrino flux.
Chapter 3 introduces the **Askaryan effect**, which allows us to cost-efficiently detect ultra-high energy neutrinos. Particle cascades, induced in dense media by interactions of ultra-high energy neutrinos, are expected to emit radio wave pulses. These pulses can travel distances on a kilometer scale in matter like ice or rock salt before being absorbed. A detector with widely spread radio antennas can thus be build to detect these pulses and cover a detection area of several hundred square kilometers at low costs.

II Detector hardware and calibration efforts

Once completed, the **ARA detector**, described in Chapter 4, is planned to become a detector that is sufficiently large to detect GZK neutrinos. It will consist of antenna clusters inside the ice, spaced by 2 km. Each of these clusters can operate as an autonomous neutrino detector, covering a volume of a few cubic kilometers. For the neutrino search performed in this thesis, ten months of data from the first two detector stations have been used.
For the data analysis a proper calibration of the instrument is very important. Parts of the calibration have been performed in the framework of this thesis. In Chapter 5 the **calibration of the ARA digitizer chip** is explained and results are shown. This calibration allows for proper timing correlations between incoming signals, which are crucial for radio vertex reconstruction and thus neutrino detection. Furthermore, a **calibration of the station geometries and cable delays** has been performed, which is described in Chapter 6. The knowledge of the exact station geometry and possible delays between different antennas is again important for event reconstruction.

III Data analysis

In the last part of this thesis, the analysis of the recorded data is described. This analysis is based on **neutrino signal simulations** explained in Chapter 7. These simulations are used to produce a neutrino sample to train algorithms and cuts on. For background noise only a small sample is produced from simulations, used for first checks of the algorithms. The primarily used noise sample to determine cut values and background estimations is a 10% subset of the recorded data. This is not expected to contain neutrino events due to their small interaction probability. Two main **analysis algorithms** are developed in the frame of this thesis, explained in Chapter 8. One is a **thermal noise filter**, checking recorded events for their consistency with an incoming radio wavefront. The second algorithm is a **matrix based reconstruction**, used to determine the incoming direction of impulsive signals and to distinguish neutrino signals from other radio sources in that way. The usage of these methods to discriminate neutrinos from the background noise is explained in Chapter 9.

The results of this analysis are presented and neutrino limits obtained are shown in Chapter 10. As mentioned above, the here performed analysis is a **blind analysis**, trained on a 10% subset of the full data. Only when cuts are finalized and a proper background estimation is calculated, the analysis is applied to the complete set of recorded data.

The thesis is concluded with a short **review** of the presented analysis and the description of **possible improvements** in Chapter 11. In addition to that the investigation of a **new detection technique** is discussed, utilizing RADAR reflections off of neutrino-induced cascades. This method could provide the possibility of building even more efficient detectors for the detection of GZK neutrinos.

Part I

Theory

2

The GZK neutrino flux

The main research topic of this thesis are the so-called GZK neutrinos. Shortly after the discovery of the **C**osmic **M**icrowave **B**ackground (CMB) an interaction between this omnipresent photon radiation and **U**ltra-**H**igh **E**nergy **C**osmic **R**ays (UHECRs) was predicted in the mid sixties by Greisen [34], Zatsepin and Kuzmin [35], named the GZK mechanism. In this interaction, pions are generated with a resonance in the cross section for cosmic ray energies slightly above the production threshold. Due to the resonance, the mean free path of cosmic rays with sufficient energy is reduced to some tens of Mpc. Since no source candidates of such high energy cosmic rays have been observed within this distance, a cutoff in the cosmic rays spectrum is expected. Furthermore, a guaranteed flux of neutrinos was predicted by Berezinsky and Zatsepin in 1968 [36] to result from the GZK mechanism. This flux is a decay product from the generated pions and is estimated to be very small. Interactions of neutrinos from this flux are expected to happen at a rate of less than once per year in one gigaton of target material. Therefore, extremely large detector volumes are needed to investigate the GZK neutrino flux.

The GZK interaction depends on different factors. Most prominent among these factors are the mass composition of UHECRs, the distribution of sources throughout the Universe and their acceleration properties. Information about all three astrophysical features can be obtained by investigating the neutrino flux. Moreover, due to the predicted cutoff in the cosmic-ray spectrum, only a very small number of cosmic rays with energies above the cutoff are expected to reach the Earth. Thus, they can only be studied via GZK neutrinos.

The present chapter is a summary of various theoretical discussions. Neutrino production via the GZK mechanism is explained and the dependencies of the flux on different prerequisites are discussed. In addition, the informative potential of a neutrino flux measurement about cosmic rays in this energy region is shortly presented.

2.1 Neutrino production in the GZK mechanism

In 1965 the Cosmic Microwave Background radiation was observed by Penzias and Wilson [37] as an unknown black body radiation of 3.5 K. This result was first interpreted by Dicke et al. as a hint for the Big Bang to be the origin of our Universe [38]. Nowadays, very precise measurements of the CMB exist which confirm its general isotropy in intensity and Kelvin scale temperature as well as its black-body radiation spectrum [39]. It was realized as a necessary consequence of this omnipresent, isotropic radiation that cosmic ray particles (in the following represented by protons) would interact with it and produce pions through the so-called GZK mechanism [34, 35]. Such pions decay rapidly on their way through the Universe into final states that include neutrinos. The dominant reactions in the GZK mechanism are sketched in Equations 2.1 and 2.2.

$$\begin{aligned} p+\gamma_{CMB} \rightarrow \quad & \pi^{+}+n && (2.1)\\ \pi^{+} \rightarrow \quad & \mu^{+}+\nu_{\mu} \\ \mu^{+} \rightarrow \quad & e^{+}+\nu_{e}+\bar{\nu}_{\mu} \\ p+\gamma_{CMB} \rightarrow \quad & \pi^{0}+p && (2.2)\\ \pi^{0} \rightarrow \quad & \gamma+\gamma \end{aligned}$$

This generation of pions is also denoted as photopion production.
Both, charged and neutral pions are produced in the interactions in Equations 2.1 and 2.2, but neutrinos are only produced in the decay of charged pions. Neutral pions instead decay mostly into gamma rays. Neglecting subsequent decays and interactions of produced neutrons, roughly 1/3 of the proton energy is transferred to neutrinos in photopion production. This ratio increases with rising proton energies due to an increasing cross section for $\pi^{\pm}$ compared to π^{0} [40], which is illustrated in Figure 2.1. The lower energy threshold E_{th} for the GZK mechanism can be calculated via the center of mass energy to

$$E_{th} \approx \frac{6.8 \cdot 10^{16}\ \mathrm{eV}^{2}}{E_{\gamma}}. \tag{2.3}$$

In this equation, E_{γ} is the energy of the interacting photon from the CMB. The threshold energy is inversely proportional to the CMB energy spectrum, which shifts upwards with a growing redshift of the interaction [40]. For example, assuming a CMB peak energy of 1 meV, the threshold cosmic ray energy is $6.8 \cdot 10^{19}$ eV. Slightly above this threshold, the cross section for pion production rises to a peak value of roughly 500 μb at the $\Delta^{+}(1232)$ resonance (Figure 2.1). From the energy dependence of the shown cross section the shape of the resulting neutrino flux spectrum can be deduced (Figure 2.2). A two peak spectrum is expected. The muon neutrino flux will peak at a neutrino energy corresponding to the resonance energy of the interaction. The electron neutrino spectrum will have two peaks, one corresponding to the resonance and one additional low energy component from the decaying neutron in Equation 2.1. In general, the neutrino flux $F(E_{\nu})$ can be

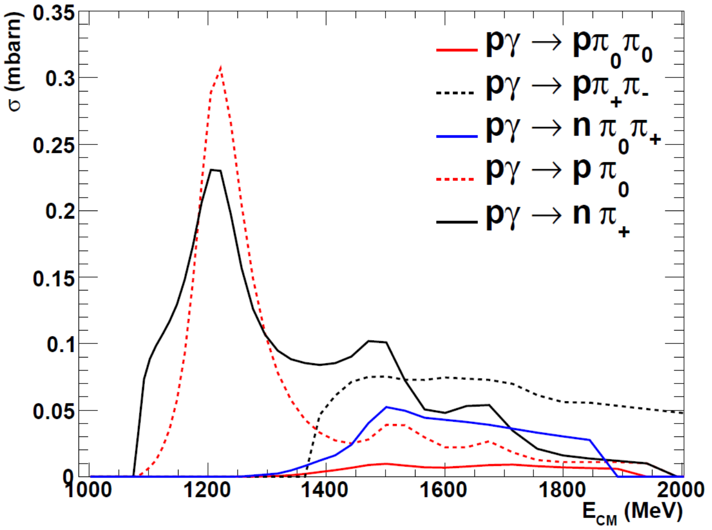


Figure 2.1: *The cross section for pion production in the interaction of protons and gamma rays. Interesting for the GZK mechanism is the* Δ^+ *resonance, where the cross section peaks to around* 500 µb *in total [41].*

calculated by integrating a source function L and the neutrino yield Y over the proton energy E_p and the redshift z [40]:

$$F(E_\nu) = \frac{c}{4\pi E_\nu} \cdot \int\int L(z, E_p^s) Y(E_p^s, E_\nu, z) \frac{dE_p^s}{E_p^s} dz. \tag{2.4}$$

The source function $L(z,\ E_p^s)$ describes the density and energy spectrum of the UHECRs. It depends on the production energy at the sources, their redshift distribution throughout the Universe and the energy loss of the cosmic ray particle during propagation. Furthermore, it scales inversely with the adiabatic expansion of the Universe. The neutrino yield function $Y(E_p^s,\ E_\nu,\ z)$ can be derived from Monte Carlo simulations with the given prerequisites of pion production cross section and the energy spectrum of the CMB. Apart from its influence on the CMB spectrum, the redshift of the interaction becomes important in the yield function if it is very small, comparable to the interaction length of the UHECRs. In this case, not all energy is transferred in photopion production and a lower number of neutrinos than expected is produced. This effect vanishes at distances above roughly 200 Mpc, where the number of produced neutrinos saturates [40].

An example for a flux calculated from Equation 2.4 under certain assumptions for L and Y is displayed in Figure 2.2. This flux can be taken as a possible example flux. Models used, especially the model for the source function, have big uncertainties, which make absolute flux predictions with the current knowledge impossible. Still it can be noted that the flux expectation is very low. Given the low cross section for neutrino interactions, different flux models result in less than one interaction per year per cubic kilometer of dense matter. It can be easily deduced that a detector exposure of several hundred km^{3}yr is needed to investigate the flux properties.

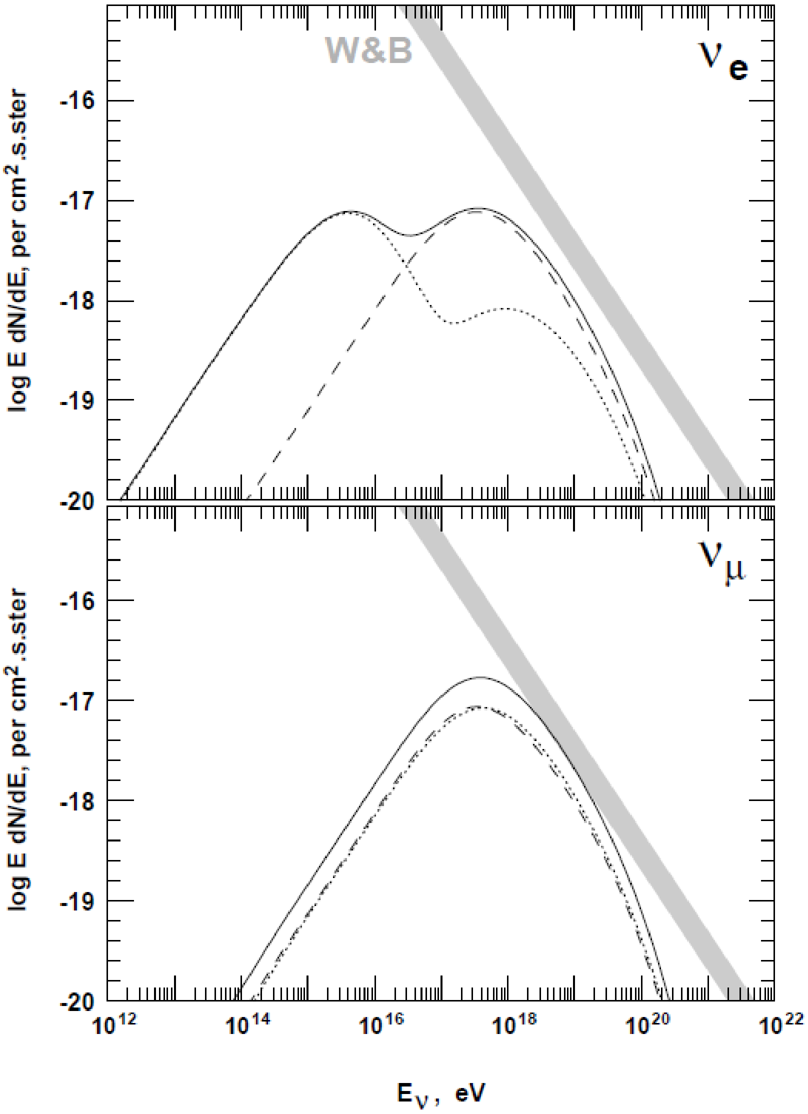


Figure 2.2: *An example of a neutrino flux generated by ultra-high energy protons, plotted versus energy. The upper panel shows the flux for* ν_e*, the lower panel for* ν_μ*. Neutrinos and antineutrinos are denoted by the dashed and dotted lines respectively. The solid line indicates the total flux from neutrinos of one flavor [40].*

As explained, the cross section for photopion production has a resonance at $\Delta^+(1232)$. The resulting mean free path of cosmic rays with energies above approximately 10^{19} eV drops drastically to less than 10 Mpc. Until now, very few source candidates with high enough acceleration power to produce such energetic cosmic rays have been observed within this range. Thus, nearly no cosmic rays with energies above the resonance are expected to arrive at Earth. This expected cutoff has indeed been measured by the High Resolution Fly's Eye Observatory (HiRes [5], now included in the Telescope Array Observatory) and confirmed by the Pierre Auger Observatory [42], as shown in Figure 2.3. Still, these measurements do not provide evidence for a strong appearance of the

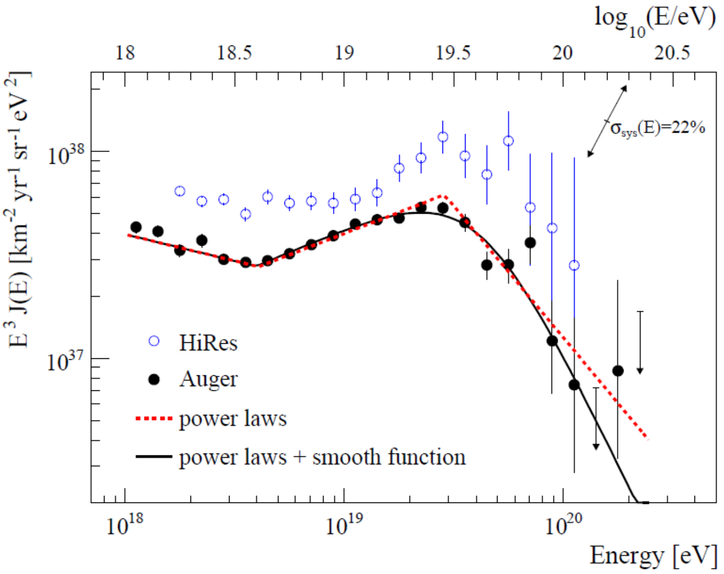


Figure 2.3: *The GZK cutoff as measured by the Pierre Auger Observatory and the HiRes Observatory [42].*

GZK process and there can be other reasons for the cutoff, i.e. source exhaustion and the cosmic ray composition. These conditions and their influence on the GZK mechanism will be explained in the following section.

2.2 Dependencies and implications

The dependencies of the GZK neutrino flux which have been explained so far are the cross section of the interaction and the CMB spectrum, which are measured very precisely. In Equation 2.4, these dependencies determine the neutrino yield function Y, which is thus very well understood. The main uncertainties for the neutrino production are instead part of the source function L, determining the energy and density of UHECR at a certain point in the Universe. The most important parts of the source model are:

- the cosmic ray composition,
- the emission spectrum of cosmic ray sources,
- the source evolution throughout the Universe.

These conditions and their influence on the GZK mechanism are strongly connected. Therefore, in the following only a qualitative explanation is attempted.

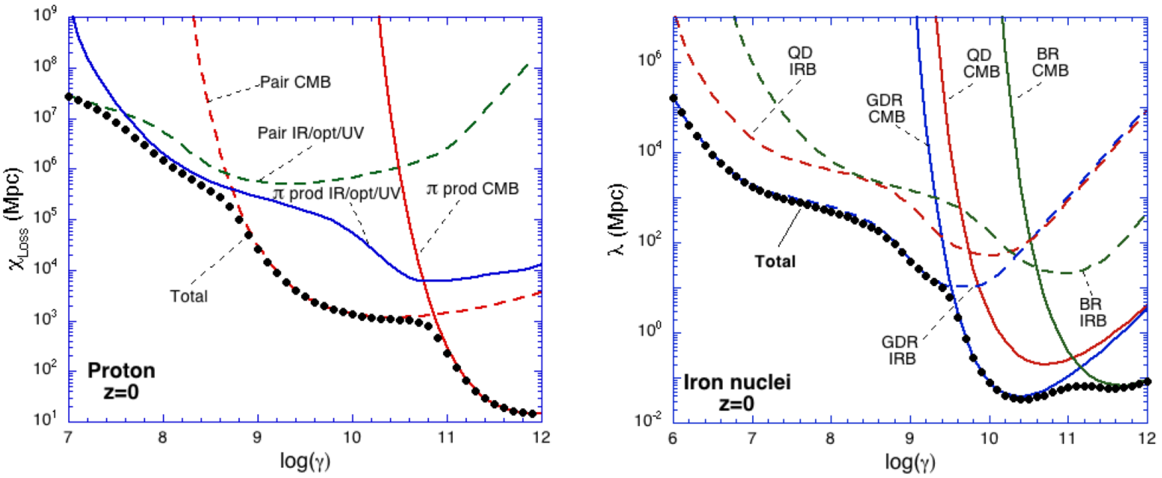
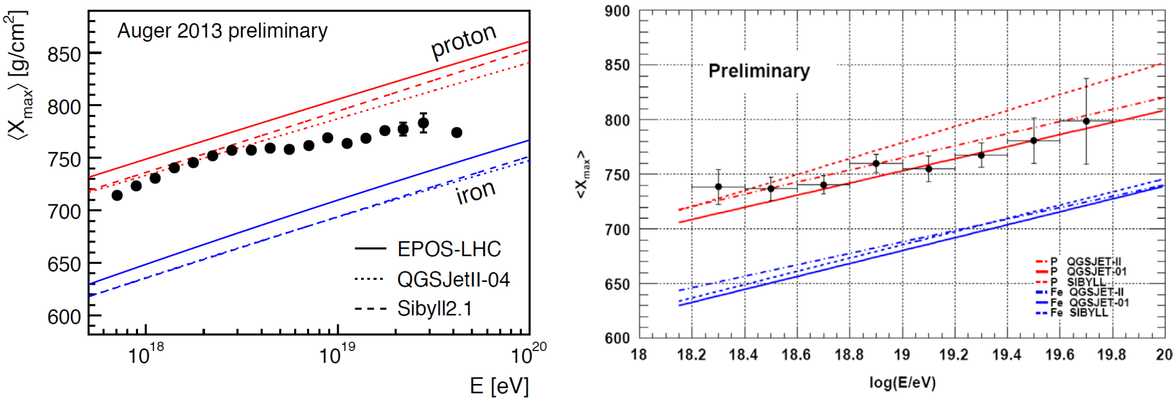


Figure 2.4: *The attenuation length for different energy loss mechanisms for protons (left) and iron nuclei (right). For the proton, e^+e^- pair production on the CMB is dominant up to the threshold energy for the GZK mechanism. For iron, the photo disintegration processes are dominantly the Giant Dipole Resonance (GDR) and the Quasi-Deuteron process. Photopion production is denoted as Baryonic Resonance (BR) [43].*

2.2.1 The composition of ultra-high energy cosmic rays

The mass composition of the cosmic rays with energies above the photopion production threshold has a strong influence on the resulting neutrino flux due to the varying energy loss mechanisms and the limited production energy. If we assume cosmic rays at these energies to be dominated by protons, non-continuous energy losses other than the GZK mechanism can be neglected as shown in Figure 2.4. The energy loss mechanisms for heavy nuclei instead differ strongly from those for protons. The dominant interactions with radiation are here photo disintegration processes, reducing the particle energy by emission of nucleons and alpha particles. In these processes neutrinos are only rarely produced through subsequent decays. The amount of energy per nucleon can be assumed constant in photo disintegration. Therefore efficient photopion production can only happen at energies higher than $A \cdot E_{th}$, with A being the atomic number of the cosmic ray. Assuming a certain energy spectrum of UHECR, it becomes apparent that pure protons are highly favorable over heavier nuclei for the production of GZK neutrinos, since the threshold energy for photopion production is much lower.

The mass composition of UHECR is currently measured by two experiments, the Pierre Auger Observatory and Telescope Array, which disagree in their results [7, 44]. These two experiments measure the mass composition in multiple ways, for example via the average atmospheric depth of the maximum of cosmic-ray air showers in our atmosphere. From simulations, showers from heavy primary nuclei are expected to develop less deeply in the atmosphere than proton showers. The shower depth measurements of the two experiments are displayed in Figure 2.5. While the Auger data tends to a heavy composition, Telescope Array measurements predict a proton dominance in UHECRs. This disagreement illustrates the difficulties of measuring such a composition and consequently making predictions for the GZK neutrino flux. On the other hand a well understood neutrino flux measurement could put strong constraints on the mass composition.



(a) **(b)**

Figure 2.5: *The atmospheric depth of the shower maximum of air showers produced by cosmic rays with energies above* 10^{18} eV *measured by* ***(a)*** *the Pierre Auger Observatory [7] and* ***(b)*** *the Telescope Array Experiment [44]. The expectation for the depth of iron primary particles and protons is indicated in each plot by the red (proton) and blue (iron) lines. Differently dashed lines are produced with different shower simulations (different hadronic models) as indicated in the legends.*

2.2.2 The cosmic ray emission spectrum

The acceleration and emission of UHECRs is one more important parameter in the GZK neutrino production. So far, no acceleration mechanism has been proven to produce cosmic ray energies significantly above 10^{16} eV [45]. Under the assumption of dominantly heavy UHECRs, this can be extended to around 10^{18} eV [46], which is still below the threshold for photopion production. The most prominent model for high energy acceleration is the acceleration through shocks, also referred to as first order Fermi acceleration, in objects like Supernova Remnants, Active Galactic Nuclei and Gamma-ray Bursts. In this model, particles are accelerated during statistical crossings of electromagnetic shock fronts, gaining a small amount of energy with each crossing. The resulting spectrum from this shock acceleration follows a power law as $dN/dE = E^{-2}$ [45]. Particles can only be accelerated in this way, as long as they are contained in the object. Once they reach a critical energy, they escape and are not further accelerated. This critical energy is limited by the Hillas condition [47]

$$B \cdot L > \frac{2E}{Z\beta}. \tag{2.5}$$

In this equation B denotes the magnetic field in the object and L its size. The parameter E stands for the energy of the particle, Z for its charge and $\beta = v/c$ denotes its speed relative to the speed of light. In current observations no objects have been found which allow acceleration to sufficient energies for the GZK mechanism.
Other, more exotic Top-Down models consider decays of super-massive particles into Standard-Model particles, which in principle could provide enough energy. In such decays ultra-high energy photons would be produced which have not been observed so far. Experiments like Auger strongly constrain their flux in the latest measurements and rule out most of the Top-Down models [48].

The existence of UHECRs on the other hand has been proven, as visible in Figure 2.3. Cosmic rays are at least accelerated to energies of 10^{20} eV but their abundance at these energies throughout the Universe is mostly unknown. It is possible to extrapolate the primary cosmic ray spectrum from the one observed on Earth when accounting for all known energy loss mechanisms. Due to large uncertainties on these losses, for example due to the unknown mass composition, the extrapolation will not be able to provide a very precise result. If cosmic rays are not accelerated to energies sufficiently high above the threshold energy for photopion production, this interaction does not occur very often and the neutrino flux measured at Earth will be relatively low. An indication for no acceleration beyond the Δ resonance would be the absence of the pronounced peak in the GZK neutrino spectrum. Consequently, measurements of the neutrino flux in the GZK energy range will provide information about the existence of cosmic rays with energies above 10^{20} eV and about their energy spectrum.

2.2.3 The evolution of cosmic ray sources

The distribution or evolution of cosmic ray sources in the Universe is important due to continuous energy loss of cosmic ray particles during propagation. It has been shown in [40] that neutrino production from protons rises up to a propagation distance of about 200 Mpc. Thus, if produced closer than this distance, less neutrinos are produced by emitted protons. In addition, the redshift of the interaction is important due to a changing CMB spectrum. At higher redshifts, the mean photon energy is higher and the threshold energy for photopion production consequently lower. The redshift also influences the neutrino flux shape through adiabatic energy loss of neutrinos propagating in an expanding Universe, which does however not affect their abundance.
In summary more distant sources appear to be favorable for GZK neutrino production. For the distribution of sources normally two possibilities are considered: isotropic distribution or a distribution coupled to the star formation rate.
A possible star formation rate, shown in Figure 2.6, is discussed in [49], deduced from measurements of ultraviolet and far infrared radiation. Compared to an isotropic assumption, the star formation rate is higher at large redshifts. Hence, an assumption of isotropically distributed sources would result in a conservative limit for the neutrino flux.

The coupling to the measured star formation rate though seems to be a more realistic model to follow.

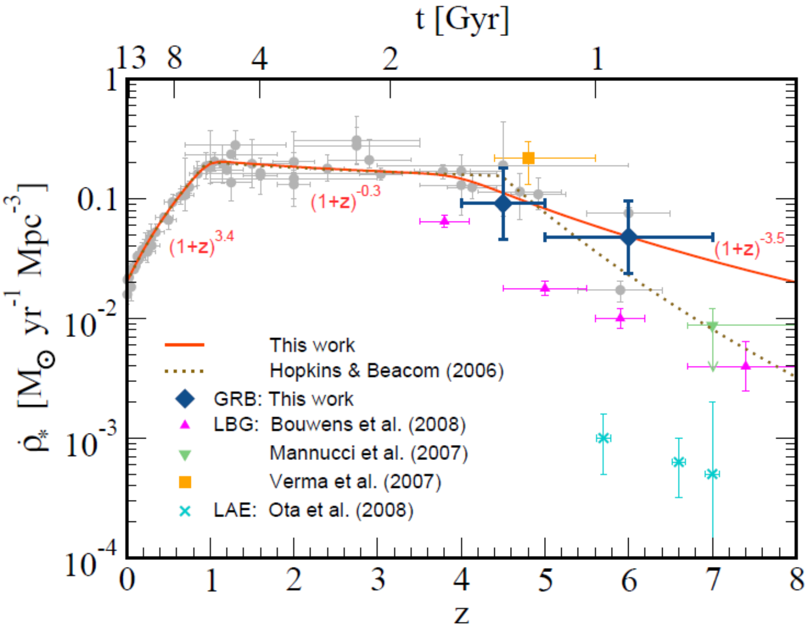
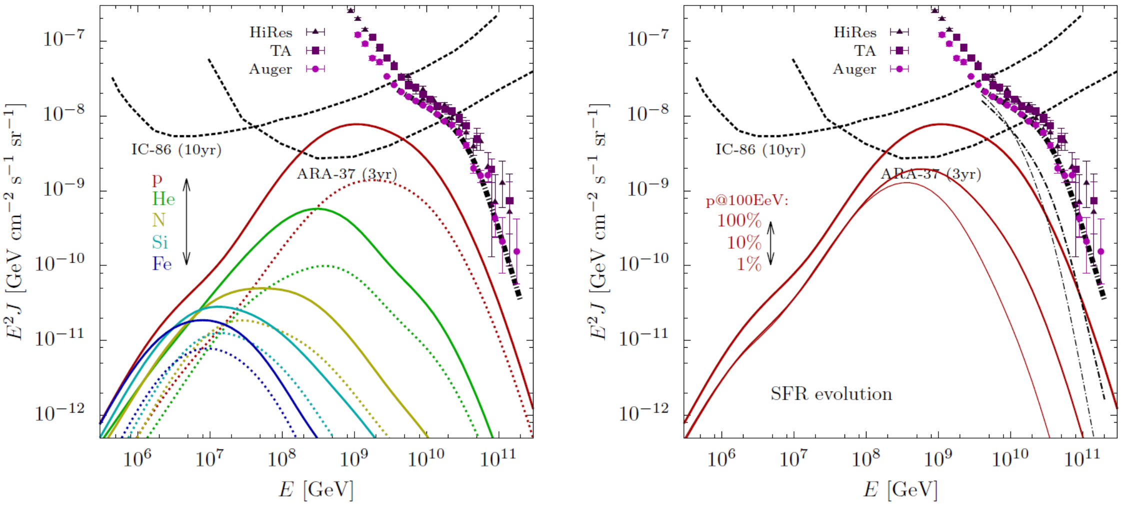


Figure 2.6: *The star formation rate from measurements of UV radiation (light grey dots) including Gamma-ray Burst measurements with fit models shown as solid and dashed lines [49].*

2.2.4 Neutrino flux expectations under varying assumptions

In Figure 2.7 neutrino flux expectations are shown which are calculated under various conditions as described in [50]. The mass composition is investigated, on the one hand for cosmic rays of a single type but with varying mass (Figure 2.7a), on the other hand for differently weighted combinations of protons with other cosmic rays (Figure 2.7b). In Figure 2.7a two cases are plotted for the source evolution: the conservative isotropic assumption and the more realistic source evolution coupled to the star formation rate, shown in Figure 2.6. The cosmic ray injection spectrum is extrapolated from latest flux measurements at Earth, accounting for redshift effects and other energy losses like photo disintegration and Bethe-Heitler pair production. For the observed cosmic ray spectrum the data from the Pierre Auger Observatory are chosen, which are systematically lower than other measurements, as can be seen from Figure 2.3. In this way, the cosmic ray spectrum is assumed conservatively.

The two plots in Figure 2.7 show clearly the strong influence of changing cosmic ray masses and source evolution on the neutrino flux. As expected the highest flux is produced from pure protons under the assumption of a non-isotropic source evolution. It



(a) **(b)**

Figure 2.7: *The expectation of neutrino fluxes under various conditions.* ***(a)*** *The colored lines show neutrino fluxes under the assumption of different nuclei, dominant in UHECRs, for isotropic source distribution (dashed line) and source evolution following the star formation rate (solid line). The black dashed lines show the expected sensitivities of the IceCube detector and the full ARA detector. The data points display cosmic ray flux measurements from the Pierre Auger Observatory, Telescope Array and HiRes. This data is used to calculate the source spectrum for the cosmic rays.* ***(b)*** *The red lines indicate neutrino fluxes under the assumption of different amounts of proton contribution to UHECRs [50].*

should be noted that a mixed composition, including a small amount of protons, provides still a more efficient neutrino production than pure relatively light nuclei like helium. The shape of the spectrum and the peak position changes strongly for different cases, which supports the statement that a measurement of the GZK neutrino flux will provide valuable information about the discussed preconditions.

An interesting fact to mention in this GZK summary is the possibility to measure the GZK mechanism in photons instead of neutrinos. Gamma rays, produced during the GZK mechanism in the pion decay, accumulate in two energy regions after traveling through the universe. One region is above 10^{19} eV, close to the production energy. This is populated by photons which did not lose enough energy to enter the region of efficient production of e^+e^- pairs on the CMB or infrared background radiation. If the photon energy drops low enough to around 10^{18} eV, they will lose their energy rapidly in pair production and accumulate instead at energies right below the pair production threshold in the GeV range.
The Fermi-LAT experiment has set an upper limit on the GZK neutrino flux by measuring and limiting gamma ray fluxes between 100 MeV and 100 GeV [51]. The limit is

calculated and shown in detail in [52]. In the highest energy region, around 10^{19} eV the Pierre Auger Observatory is currently conducting a gamma ray search and has provided limits for this region [48]. Since direct neutrino detection experiments currently put more stringent limits on the GZK flux than these two gamma ray measurements they will not be further discussed. Still they should be mentioned as an additional tool for the investigation of the GZK mechanism from a completely different detection method, which can become competitive in the coming years.

3

The Askaryan effect in dense media

As we have seen in the last chapter, GZK neutrinos only arrive very rarely at Earth, i.e. they interact roughly once per 1 $\mathrm{km}^3/\mathrm{yr}$ in dense matter like ice and water. Large detector volumes are necessary to efficiently catch such neutrinos. In this chapter the emission of coherent radio waves from an excess negative charge in cascades, induced by ultra-high energy neutrinos, as first discussed by Askaryan in the sixties ([53, 54]) will be presented. The attenuation length for so-called Askaryan emission can be on the order of one kilometer in ice or similar radio-transparent materials, which renders feasible the construction of large neutrino detectors such as the Askaryan Radio Array (ARA).
In this chapter, we first discuss electromagnetic cascades and the origin of the negative charge excess. Then the emission from such imbalanced charge is explained and finally verification measurements of this emission are presented.

3.1 Emission of radio waves from neutrino-induced cascades

3.1.1 Negative charge excess in electromagnetic cascades

An electromagnetic cascade can be initiated by the interaction of an ultra-high energy neutrino in ice. Depending on the interaction type it can be generated alone, containing nearly exclusively electrons, positrons and photons, or with a hadronic component. For the explanation of the emission of radio waves in this chapter we will focus on the electromagnetic part of such cascades. This part mainly evolves by the emission of Bremsstrahlung from charged particles in the atomic fields and the pair production of

electrons and positrons. These processes are dominating the longitudinal shower development and, due to their nature, charges of both sign are produced in equal numbers. Still, by closer examination of the energy loss mechanisms of electrons, positrons and photons, an excess of electrons of some ten percent, building up during the development of the cascade, can be understood.
After the prediction and first calculations by Askaryan, a very detailed description of the energy loss mechanisms can be found in [55]. There, four main mechanisms are mentioned, which are responsible for the shower development, namely:

- emission of Bremsstrahlung by positrons and electrons in the fields of the atoms,
- pair production by photons,
- elastic Rutherford scattering of electrons and positrons on the atoms,
- interactions with atoms, namely Compton scattering of photons and at lower energies annihilation of positrons and other ionization processes.

As explained above, Bremsstrahlung and pair production are the main processes to determine the shower development. One characteristic value for these mechanisms is the interaction length after which a photon is emitted or a pair is produced. It is similar for both processes and depends on the medium in which the particle cascade evolves. In approximation the interaction length grows logarithmically with the particle energy. It depends on the cross section for such processes. This cross section is particularly large compared to other energy losses, described in the following, which makes Bremsstrahlung and pair production dominant at the energies considered. As noted above, there is no charge asymmetry expected so far since electrons and positrons are always produced in pairs and lose their energy equally. It should be further mentioned that in principle the cross section for Bremsstrahlung grows (logarithmically) with energy and the cross section for pair production converges to a constant value at the highest energies. This is only the case as long as the screening effect of a collective atomic potential does not play a role in the energy loss. This screening appears when the interaction length exceeds the atomic distance in the medium. Single atomic fields gather in a global potential and give more freedom to the electron (positron). Thus, above a certain energy, the cross section for Bremsstrahlung and pair production is strongly reduced. This effect is called the LPM effect, after its explorers Landau, Pomeranchuk and Migdal [56, 57]. The threshold energy for this effect has been calculated in [58] to be of roughly 2 PeV in media like water and ice. Consequently, we expect a strong elongation of the shower above this energy. In the following first energies will be regarded at which the LPM effect does not play an important role. Its influence will be discussed in a separate section.

Another energy loss channel appears through elastic scattering on the atoms which is the main cause for the lateral expansion of the shower. The first generated shower particles are highly boosted along the incident direction of the neutrino. Only after undergoing several elastic deflections is their path changed. This has been explored by Molière, who calculated a characteristic deviation angle and a radius for different materials [59]. To summarize, in this energy loss no particles are added to the shower but it is slightly broadened in the lateral dimensions.

The last mentioned energy loss mechanism is the one disturbing the balance in charge. Bremsstrahlung photons can drag atomic electrons into the shower by Compton scattering. Due to the high number of photons in the shower, this is the dominant effect to build up a net negative charge. At still lower energies, further contributions come from the annihilation of shower positrons with atomic electrons and ionization of the atoms by positrons and electrons. Either a positron is taken out of the shower, or additional electrons are dragged along.

3.1.2 Coherent emission of radio waves from a charge excess in electromagnetic cascades

To understand the emission of **R**adio **F**requency (RF) radiation from the full electromagnetic cascade, one first has to understand this emission for a single particle moving through matter with relativistic speed, exceeding the speed of light in the medium. This emission has to be summed for all particles in the shower to describe the overall produced electromagnetic radiation. In the following, the calculations from [55] are summarized as an example of how the radiation from electromagnetic showers can be determined.
To start, one can use the Lienard-Wiechert retarded potential produced by a moving and accelerated particle to calculate the emitted electric and magnetic field. Given the relations

$$\vec{E} = -\nabla\phi - \frac{\delta\vec{A}}{\delta t} \qquad \text{and} \tag{3.1}$$

$$\vec{B} = \nabla \times \vec{A}, \tag{3.2}$$

with the scalar potential ϕ and vector potential $\vec{A}$, one can calculate the electric and magnetic field from

$$\vec{E}(t,\vec{x}) = \frac{e\mu_r}{4\pi\epsilon_0}\left[\frac{\vec{u}-n\vec{\beta}}{\gamma^2|1-n\vec{u}\vec{\beta}|^3R^2}\right]_{ret} + \frac{e\mu_r}{4\pi\epsilon_0 c}\left[\frac{\vec{u}\times((\vec{u}-n\vec{\beta})\times\dot{\vec{\beta}})}{|1-n\vec{u}\vec{\beta}|^3R}\right]_{ret}, \tag{3.3}$$

$$\vec{H}(t,\vec{x}) = \sqrt{\frac{\epsilon_r\epsilon_0}{\mu_r\mu_0}}\vec{u} \times \vec{E}(t,\vec{x}). \tag{3.4}$$

The used parameters are:

- e the charge of the particle,
- μ_r, μ_0 and ϵ_r, ϵ_0 the permeability and the permittivity of the medium and vacuum respectively,

- $\vec{\beta} = \vec{v}/c$ the particle velocity vector per speed of light,
- γ the Lorentz factor,
- $\vec{u}$ a unit vector pointing from the particle position to the observation point,
- R the distance between the observer and the emitting particle at time of emission.

All numbers in the brackets denoted as "ret" are to be taken at the emission time, also referred to as retarded time t_r.
The radiation, captured at a distant observation point, is expected to come from the second term. It scales with $1/R$ whereas the first term scales with $1/R^2$ and is thus dominant in the near field. Furthermore, only the second term is proportional to the acceleration $\dot{\vec{\beta}}$ and is thus the source for electromagnetic radiation.
In further steps in [55] the Fourier transform of such electric field is introduced as

$$\vec{E}(\omega, \vec{x}) = 2 \int e^{i\omega t} \vec{E}(t, \vec{x}) dt. \tag{3.5}$$

Using the substitution [60]

$$\frac{d}{dt}\left(\frac{\vec{\beta}_T}{1 - n\vec{u}\vec{\beta}}\right) = \left[\frac{\vec{u} \times ((\vec{u} - n\vec{\beta}) \times \dot{\vec{\beta}})}{|1 - n\vec{u}\vec{\beta}|^3}\right], \tag{3.6}$$

this results at large distances R to

$$\vec{E}(\omega, \vec{x}) = \frac{e\mu_r}{2\pi\epsilon_0 c^2} \cdot i\omega \cdot \int dt_r c\vec{\beta}_T \cdot \frac{e^{i\omega(t_r + R(t_r)/c)}}{R(t_r)}, \tag{3.7}$$

with ω the emission frequency and $\vec{\beta}_T$ the average speed component, perpendicular to the line of sight. Note that the time Δt in which acceleration and deceleration of the particle happen is assumed to be much smaller than the period of the regarded waves which implies $\omega\Delta t \ll 1$. We have further assumed R to be very large, in particular much larger than the particle track length over which the emission takes place. This assumption implies that R stays approximately constant for all points along the track. Given this, the Fourier transform can be calculated to

$$\vec{E}(\omega, \vec{x}) = \frac{e\mu_r}{2\pi\epsilon_0 c^2} \cdot i\omega \frac{e^{ikR}}{R} e^{i(\omega - \vec{k}\vec{v})t_1} \vec{v}_T \left[\frac{e^{i(\omega - \vec{k}\vec{v})\delta t} - 1}{i(\omega - \vec{k}\vec{v})}\right]. \tag{3.8}$$

In Equation 3.8 t_1 denotes the start time of the particle track, δt the time from appearance to absorption of the particle, i.e. the duration of radiation, and $\vec{v}_T$ the component of the particle speed, perpendicular to the line of sight. The wave vector is given by $\vec{k} = \omega n/c \cdot \vec{u}$, with $\vec{u}$ the unity vector in direction of observation. It can be calculated that for $\lim(\omega - \vec{k}\vec{v}) \to 0$, the case of observation under the so-called Cherenkov angle Θ_C, the term in the brackets converges to δt. Under this angle, finally one can conclude that the emission field strength scales proportional to the length of the particle track

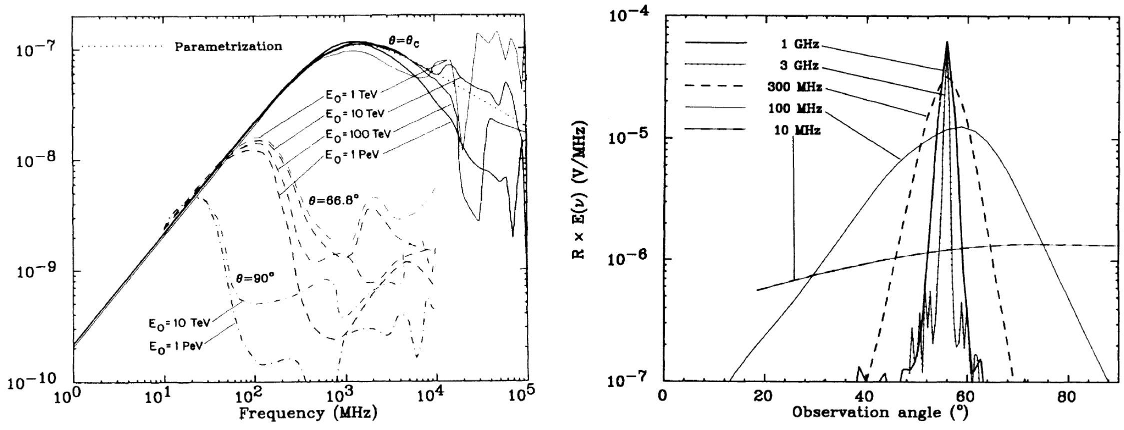
$|\vec{v}_T| \cdot \delta t$ and to the emission frequency ω.
This result for one single particle can be used together with shower simulations to compute the full emission spectrum of an electromagnetic cascade.
In the combined emission of many particles there are three frequency regimes to be regarded.

- First, there is the regime of very short wavelength, thus high frequencies, in which the emission of the shower particles is completely incoherent. This is the case for optical light, also referred to as Cherenkov light. All particles take part in the emission and the field strength scales with the square root of the sum of all tracks. Measurements show that its reach in various transparent media is not very high, i.e. less than 100 m [61]. This is not unexpected due to the high frequency and the radiation is thus not suitable to be used in large GZK neutrino detectors because a wide sensor spacing, which requires long signal reach, is not possible.

- The other extreme is the regime of very low frequencies, below 10 MHz at which the emission of all particles becomes coherent. Then, only the negative excess charge contributes to the emission, but due to the coherence the emitted field strength scales linearly with the sum of all excess tracks. However, the emission is not expected to be very strong due to its proportionality with frequency. In practice, this radiation is moreover uninteresting because it requires physically large antennas which can be difficult to handle in some environments.

- The third regime lies between the two. It is the transition region between full incoherence and full coherence. In this frequency range, the coherence depends strongly on the observation angle. Maximum coherence can be reached when observing at the Cherenkov angle. It should be noted that in this case the peak frequency only depends on the lateral cascade dimension, which is independent from the primary energy [55]. The regarded frequencies are in the radio range, which is most interesting for neutrino detection because the attenuation of the emission is low in dense media and average signals can be detected well up to distance scales in excess of 1 km.

Figure 3.1a illustrates the above explained behavior very well. In this plot the electric field emitted by the shower is shown as function of the observation frequency, normalized by its primary energy. At low frequencies, hence the case of fully coherent emission, the field scales linearly with primary energy and frequency and is only mildly dependent on the observation angle. The frequency range of the partially coherent regime expands from around 100 MHz to several GHz. It can be seen that the signal depends strongly on the observation angle and peaks at Θ_C. The fully incoherent regime then develops in the high GHz range. It can be noted that the angular signal distribution is mostly energy independent in the shown range. Figure 3.1b shows more intuitively that the angular spread around the Cherenkov peak depends strongly on the observation frequency. In



(a) (b)

Figure 3.1: *(a) The electric field normalized by the primary energy versus observation frequency, displayed for different energies and observation angles, as denoted in the plot. (b) The electric field for a primary energy of* 1 PeV *versus observation angle around* Θ_C*; data are plotted for different observation frequencies, as denoted in the plot [55].*

addition, this angular spread depends on the lateral cascade dimensions. The emission pattern is comparable to the Fraunhofer diffraction pattern of an elongated radiation source with its maximum at the Cherenkov angle. Due to the mild energy dependence of the cross section of Bremsstrahlung and pair production in this energy range, the angular emission pattern should be nearly unaffected by the primary energy, which is visible in Figure 3.1a. In the following section we will see that for energies above those in the plot the diffraction pattern is actually changed.

3.1.3 The influence of the LPM effect

In [55] only showers from particles with primary energies of maximal 10 PeV and mainly in the TeV range are regarded. At those energies, the earlier mentioned LPM effect does not play an important role for the shower development. This effect is treated in more detail in [62] where showers with primary energies in the PeV range are discussed. As already explained, the LPM effect reduces the cross sections for Bremsstrahlung and pair production above a certain energy, which leads to an elongation of the shower. However, the main contribution to Cherenkov radiation has been found to come from low energy particles in the shower whose tracks are not affected by the elongation [55]. Furthermore, the total number of particles in a shower is unaffected by the LPM effect and simulations show that there is no change in the lateral shower dimensions.

There is thus no influence on the general signal strength of coherent Cherenkov emission expected from the LPM effect and the only difference is the elongated particle distribution along the shower axis. Hence, the Fraunhofer diffraction pattern will be changed.

In [62] the angular emission spectrum around the Cherenkov angle Θ_C is very well approximated by a Gaussian distribution

$$E(\omega, R, \Theta) = E(\omega, R, \Theta_C) \cdot e^{-\ln 2 \left(\frac{\Theta - \Theta_C}{\Delta\Theta}\right)^2}, \tag{3.9}$$

with

$$E(\omega, R, \Theta_C) = \frac{1.1 \cdot 10^{-7}}{R} \cdot \frac{E_0}{1\ \text{TeV}} \cdot \frac{\nu}{\nu_0} \cdot \frac{1}{1 + 0.4 \left(\frac{\nu}{\nu_0}\right)^2} \frac{\text{V}}{\text{MHz}}. \tag{3.10}$$

The dependencies are the observation frequency ν, the distance R and the observation angle Θ. The LPM effect enters in the determination of $\Delta\Theta$ as

$$\Delta\Theta = \begin{cases} 2.7° \cdot \frac{\nu}{\nu_0} \cdot E_0^{-0.03}, & \text{if } E_0 < 1\ \text{PeV} \\ \\ 2.7° \cdot \frac{\nu}{\nu_0} \left(\frac{E_{LPM}}{0.14 E_0 + E_{LPM}}\right)^{0.3}, & \text{otherwise.} \end{cases} \tag{3.11}$$

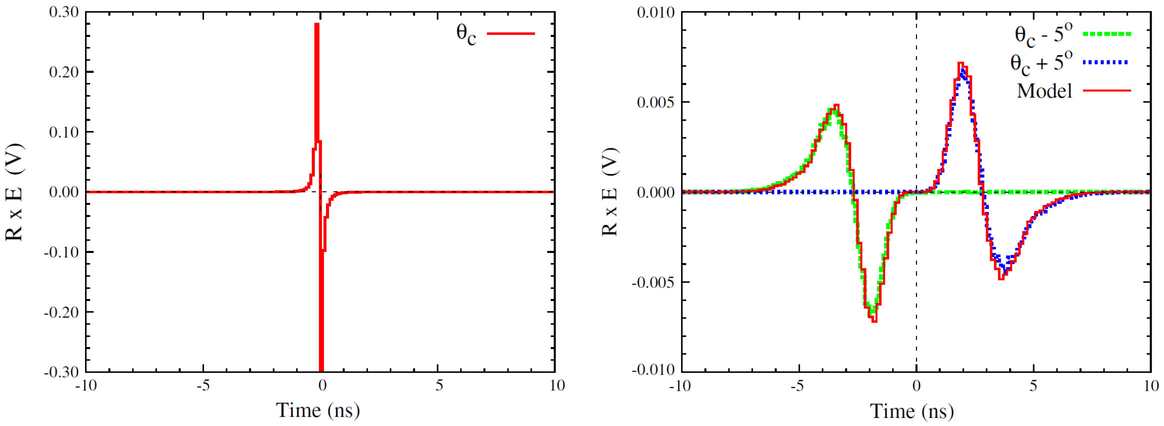
It can be seen that the influence of the primary energy only becomes important for energies above the threshold energy for the LPM effect, denoted as E_{LPM}.

3.1.4 Askaryan emission in the time domain

The Askaryan emission, which is described above in the frequency domain, has also been modeled in time domain as explained in [63]. This is especially interesting for the simulation of the Askaryan effect. Assuming a particle track with only acceleration in the beginning and deceleration in the end and a constant speed in between, one can obtain the electric field of a single moving particle in time domain

$$\begin{aligned} E(t, \Theta) = -\frac{1}{R} \frac{e\mu_r}{4\pi\epsilon_0 c^2} \bar{v}_T \Bigg[& \frac{\delta(t - \frac{nR}{c} - (1 - n\beta cos\Theta) t_1)}{1 - n\beta cos\Theta} \\ & - \frac{\delta(t - \frac{nR}{c} - (1 - n\beta cos\Theta) t_2)}{1 - n\beta cos\Theta} \Bigg]. \end{aligned} \tag{3.12}$$

The times t_1 and t_2 are here the start and end time of the track. The δ-functions stand for the instantaneous acceleration and deceleration of the particle. One can note that these functions overlap for the Cherenkov angle $\cos\Theta_C = 1/n\beta$. Their spread becomes larger with rising difference between observation and Cherenkov angle. Hence, the resulting pulse will be expected as a bipolar signal, very sharp at the Cherenkov angle and more spread at angles differing from this. Pulses calculated in a full shower simulation are displayed for the Cherenkov angle in Figure 3.2a and off the Cherenkov angle in Figure 3.2b, showing the above described behavior.



(a) **(b)**

Figure 3.2: *The Askaryan signal of a shower with* 1 PeV *primary energy, observed under* ***(a)*** *the Cherenkov angle and* ***(b)*** 5° *above and below that angle [63].*

3.1.5 Simulation codes

Most of the above discussed results have been produced through simulations rather than analytical calculations. There are in fact several codes for the estimation of the Askaryan effect in dense media: GEANT based simulations [64] or the ZHaireS simulation [65]. It should shortly be mentioned that the Askaryan effect is as well observed in cosmic ray induced air showers. The difference is the density of the medium which changes the calculations accordingly. In addition, the influence of the Earth's magnetic field becomes important for air showers [66] as opposed to the relatively small showers in dense media. Apart from that, the principle features stay the same. Also in this field a number of simulations have been developed, based on different approaches, such as the MGMR [67] and EVA [68] code, SELFAS2 [69] and the CoREAS simulation [70]. Also the the ZHaireS code can be used to model cosmic ray induced air showers. These simulation codes will not be discussed in further details since they are not very relevant for the topic of this thesis. However, some of the examples can be adapted to dense media without large difficulties an can thus be interesting for future modeling.

3.2 Verifications of the Askaryan effect

After being postulated and modeled in different ways, the Askaryan emission has been measured in the meanwhile in different media at accelerator facilities. These measurements have been performed with a particle beam directed into a sand [71], rock salt [72] and ice target [73] at SLAC. The target volumes, reaching several tons, have been instrumented with antennas distributed around the targets. For sand and salt, a photon beam was used to avoid transition radiation caused by electrons passing the boundaries between different media. For the measurement with ice, a dense electron beam was used.

The energy of a primary particle is simulated by different photon or electron intensities. Below E_{LPM}, the primary energy only influences the density of a particle shower. It has been shown that mostly low energy particles contribute to the Askaryan signal. Therefore, the change of the density of low energy particles in a shower can be used to emulate different primary energies with regards to the Askaryan emission.

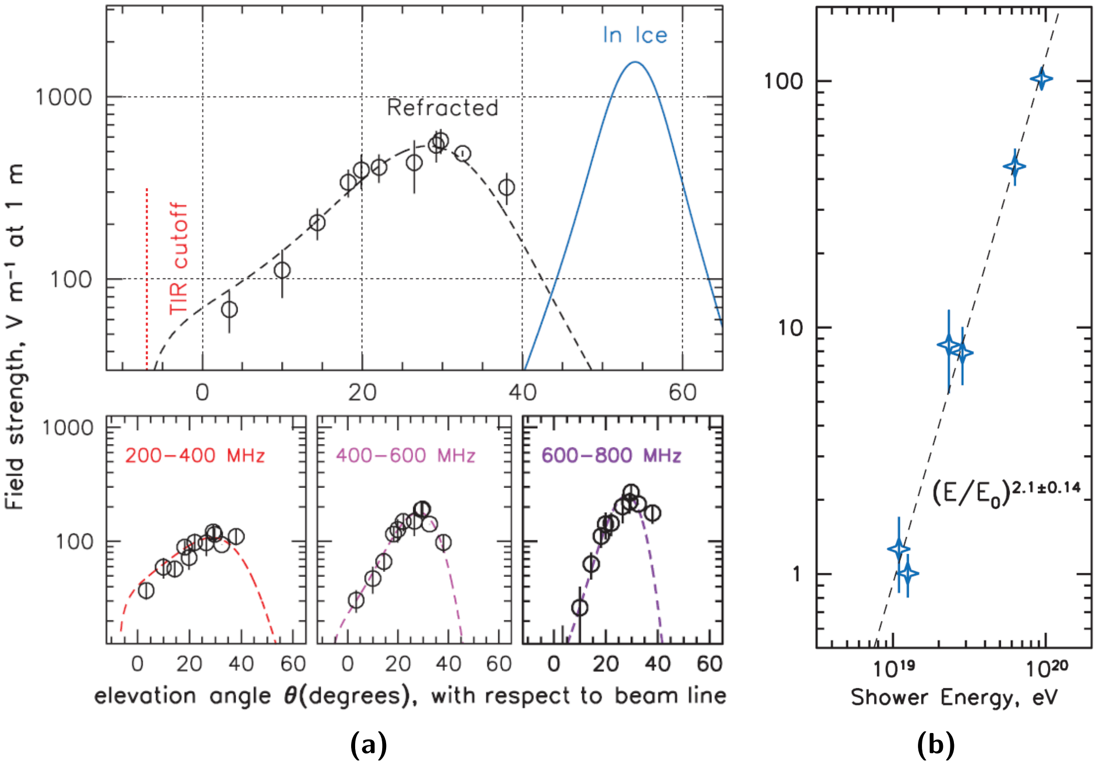


Figure 3.3: *(a) The diffraction pattern of Askaryan emission at different frequencies measured with an ice target at SLAC. The receiving antennas were mounted outside the target. The dashed line shows the expected angular spectrum after refraction.* ***(b)*** *The pulse power versus emulated shower energy, showing a nearly quadratic dependency [73].*

In the above mentioned experiments various features of the Askaryan effect have been verified which clearly exclude other radiation types to be their source. These features are the coherence, the quadratic energy dependence of emission power (Figure 3.3b), the linear polarization of the emission and the diffraction pattern around the Cherenkov angle (Figure 3.3a). The receiving antennas used in the measurements of the diffraction patterns were mounted outside the ice target, in air. Therefore, a correction factor has to be applied to account for refraction at the ice-air interface. The corrected distribution is denoted by the dashed lines in Figure 3.3a. It can be seen that this expectation is well reproduced by the data.

It should be noted that in addition to the measurements described above Askaryan emission has also been observed from cosmic ray induced air showers [74].

Part II

Detector hardware and calibration efforts

4

The Askaryan Radio Array (ARA)

The Askaryan Radio Array (ARA) is currently being constructed as a detector for GZK neutrinos based on the Askaryan emission from neutrino-induced cascades, as described in the previous chapter. The detector is deployed in the ice at the South Pole next to the IceCube detector. In the frame of this thesis, data recorded by the already constructed parts of the detector are analyzed to search for neutrino signatures.
In the first part of this chapter, the design of the ARA detector is described, including the explanation of its geometry, antennas used and electronics. The second part of this chapter presents the setup and performance of the first installed detector stations, which delivered the data that is analyzed in this thesis. Furthermore, a comparison of ARA to other neutrino observatories is discussed.

4.1 The general detector setup

The ARA detector is planned to consist of widely spread antenna clusters that can each operate as autonomous detectors of the Askaryan emission from neutrino-induced cascades. Such clusters are currently deployed in the South Pole ice, next to the IceCube detector [32], as shown in Figure 4.1. As explained in chapter 2 and 3, GZK neutrinos interact very rarely and need large detector volumes of several cubic kilometers to be efficiently detected. The Askaryan emission from neutrino-induced cascades allows to build such large detectors due to the long attenuation length of radio waves in media like ice and salt, which permits a wide sensor spacing. The South Pole ice sheet with its thickness of nearly 3 km and a basically unlimited area provides, as one of the few places on Earth, a large and continuous radio transparent medium, perfectly suited for neutrino detection. Moreover, the noise conditions are very good, due to its remoteness

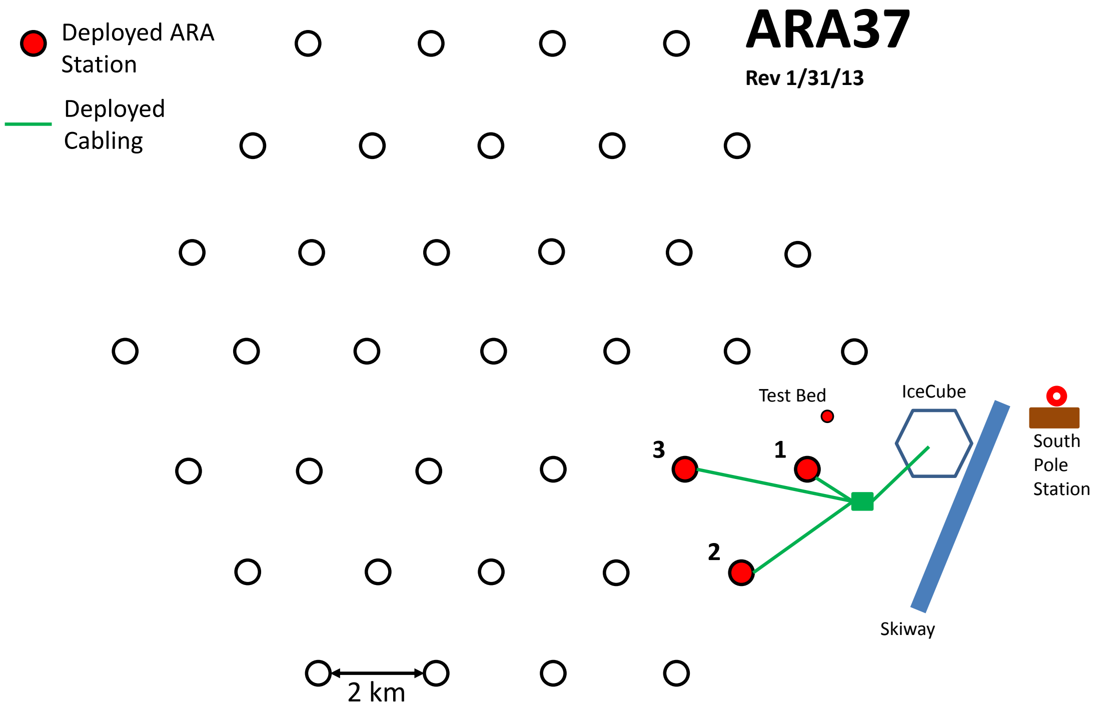


Figure 4.1: *The ARA station map at South Pole as planned, next to the IceCube detector. Open circles indicate future antenna clusters, red filled circles indicate already deployed clusters until* 2014*.*

from human activities, and an existing science infrastructure, the Amundsen-Scott station, renders the construction of a detector like ARA feasible. The baseline design of the full detector foresees 37 antenna clusters, spaced by 2 km at completion, and will be referred to as ARA37 in the following. The stations will cover an area of more than 100 km^2. A detailed description of the initial detector design as it is summarized in this chapter is given in [75].

A single antenna cluster, in the following referred to as an ARA station, consists of 20 in-ice antennas and four surface antennas. Of the in-ice antennas, 16 form the measurement core of the station, mounted on four strings in a close-to-cubical structure (Figure 4.2). The sides of this cube amount to 20 m each and the full system is installed at an average depth of 180 m. For a clear polarization measurement of incoming signals, two of the four in-ice antennas on a string are horizontally polarized **(Hpol)** and two vertically polarized **(Vpol)**, mounted in alternating order. All 16 antennas are designed to be sensitive in a range between 150 and 850 MHz. Outside the antenna cluster, at a distance of about 40 m from its core, four calibration antennas are deployed with the ability to send radio pulses of both polarizations. Two such pulsers, one of each po-

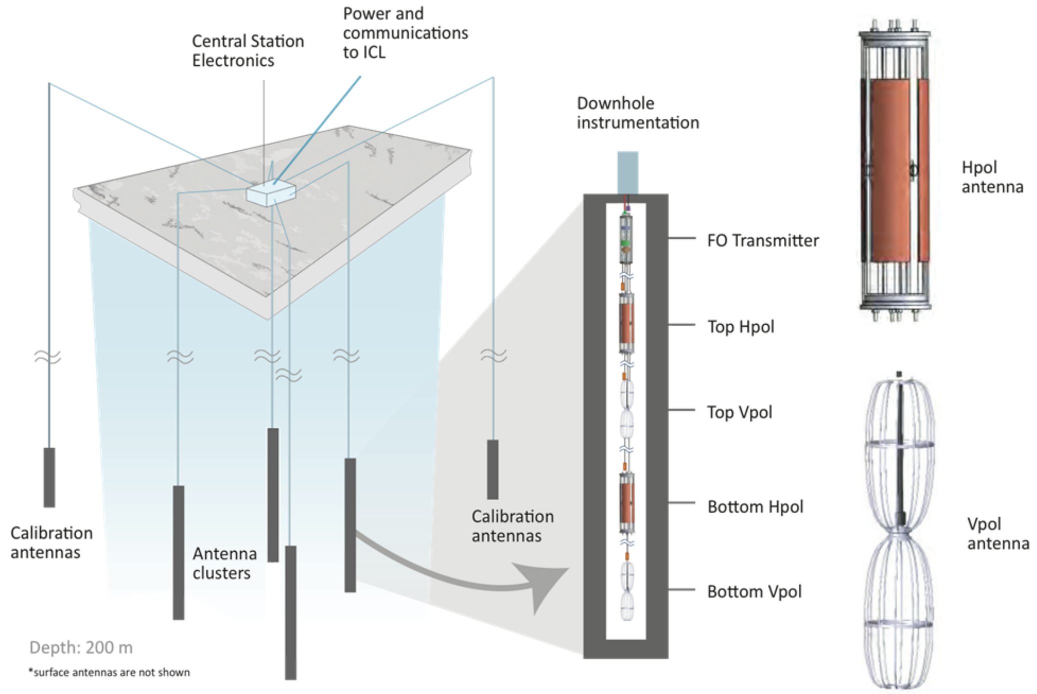
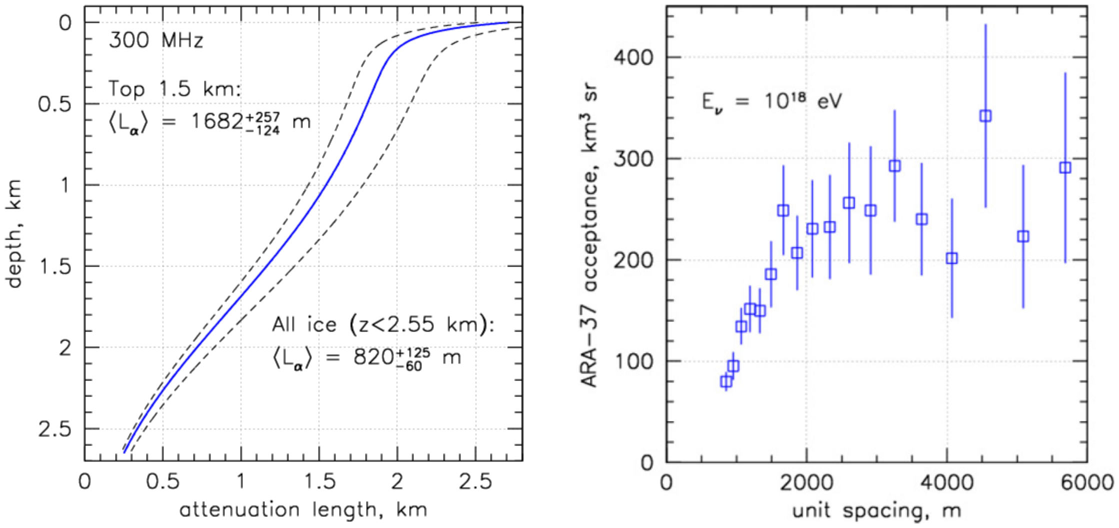


Figure 4.2: *Schematic view of an ARA station, with embedded string details. Surface antennas are not shown in this figure.*

larization, are mounted on a common string. These antennas are used for calibration, mainly of timing and geometry, as well as for the determination of station performance. Their signal is modeled to be similar to the expected Askaryan signals. Hence, it can be used to tune analysis parameters as well as vertex reconstructions. The surface antennas operate at a lower frequency than the in-ice antennas, namely in a range between 30 and 300 MHz. Their main purpose is to investigate air shower and cosmic radio signals. They can further be used as a veto for the in-ice station.

Every station can serve as an autonomous neutrino detector. Therefore, their spacing is chosen to maximize the effective volume of the full ARA array while not counting on coincident events. Such coincidences are in any case expected to be rare due to the comparably tight Cherenkov cones of Askaryan emission. The effective volume of the detector depends strongly on the attenuation length, which has been measured at South Pole by the ARA prototype station, referred to as the ARA TestBed. Its setup and performance is described in detail in [75] and first constraints on a neutrino flux from this detector are presented in [76]. The main purpose of the ARA TestBed was to investigate measurement conditions for future ARA stations as well as to test detector hardware. In terms of measurement conditions, the main noise sources and the attenuation of radio waves in South Pole ice were investigated. The main ARA noise has been found to be thermal in nature, which determines the trigger thresholds. The attenuation length has been measured by using pulsers deployed deep in the ice with the last IceCube strings. With such pulsers the attenuation length could be determined to roughly 820 m on average throughout the whole ice for 300 MHz signals (Figure 4.3a).



(a) **(b)**

Figure 4.3: ***(a)*** *The depth profile of the attenuation length for radio waves propagating through South Pole ice, with an all-ice average of* 820 m *[77].* ***(b)*** *The detector acceptance of ARA37 for different station spacings, determined from simulations [75].*

Considering this result, simulations show that the effective volume of the full detector saturates around a station-spacing of 2 km (Figure 4.3b), which has been chosen for the final design. A more detailed depth profile of the attenuation length can be found in [75] and [78].

The station depth is critical due to a changing index of refraction in the ice. The ice sheet at South Pole is about 3000 m thick. It was formed over thousands of years by accumulating snow on the surface. With time, overlaid snow layers get compacted by pressure from layers above and eventually turn into solid ice. Thus, the density of the ice changes strongly in the top 200 m, called the firn, whereas it becomes approximately constant in the solid ice region below. This and a changing temperature cause the change in index of refraction throughout the ice depth. Assuming such a changing index of refraction, Snell's law predicts a down-bending of the path of radio waves in the vertical direction. Different measurements, models and implications for the changing index of refraction are described in [79]. Starting at low values, the index of refraction is expected to converge to the average value of 1.78 below an ice depth of roughly 200 m. There are different models describing the existing measurements. In the model which is

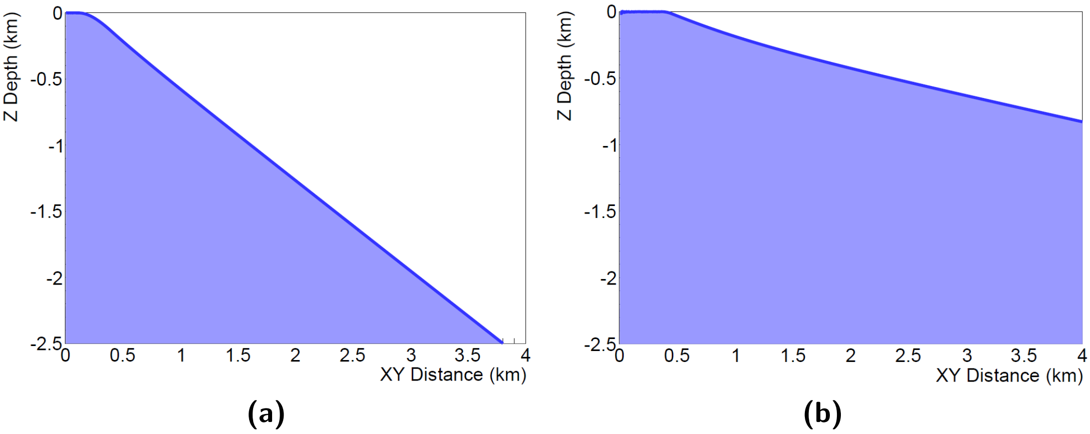


Figure 4.4: *Region of ray-tracing solutions for a receiver antenna at a depth of* 25 m *(**left**) and* 200 m *(**right**). Only a signal emitted in the blue shaded area can reach the antenna [76].*

used in ARA for simulations and reconstructions, the refractive index n as function of the depth z in meters is calculated as

$$n(z) = 1.78 - 0.43 \cdot e^{0.016 \cdot z}, \tag{4.1}$$

with $z = 0$ at the ice surface and $z < 0$ below.

Given such a refraction profile, radio waves cannot reach the detector from everywhere in the ice due to the ray-bending effect. The closer an emission point lies to the surface, the closer a receiver at a fixed depth has to be to receive the signal. This shadowing effect is illustrated in Figure 4.4. It becomes apparent that there are areas in the ice, which are nominally inside the sensitivity reach of a receiver, but which are inaccessible due to ray-bending. Figure 4.4 shows the area from which a signal can reach a receiver at depth of 25 m and 200 m shaded in blue. A deeper deployment appears to be favorable to maximize the detector volume, which has been confirmed by simulations, as shown in Table 4.1.

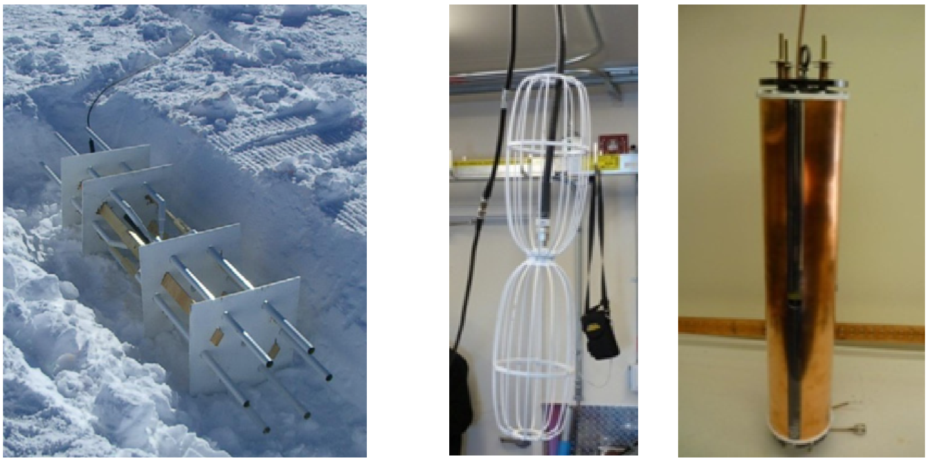
The final choice of 180 m average antenna depth allows us to avoid the region of strongest change and to keep drilling costs affordable.

Table 4.1: *The simulated effective volume of one single antenna cluster with the antennas deployed at different average depths, for a neutrino energy of* $10^{18.5}$ eV*. Note that the parameters for this simulations are different from those of the final ARA stations [77].*

Station depth / m	effective Volume / km^{3}sr
100	3.8 ± 0.2
200	4.6 ± 0.2
300	5.2 ± 0.2

4.1.1 ARA antennas

There are three different types of antennas used in the ARA stations: low-frequency surface antennas, vertically polarized in-ice antennas (**Vpol**) and horizontally polarized in-ice antennas (**Hpol**).



(a) (b)

Figure 4.5: ***(a)*** *A surface antenna installed on top of the ARA stations.* ***(b)*** *The two antenna types used for the in-ice array of the ARA stations:* ***left:*** *the so-called birdcage antenna (Vpol),* ***right:*** *the quad-slot antenna (Hpol) [75].*

Surface antennas are installed on the ice, in a regular triangle of roughly 100 m at each side, with one antenna in the center. A picture of a surface antenna as installed in ARA is shown in Figure 4.5a. The antenna's sensitivity is optimized between 30 and 300 MHz, which is best for cosmic ray induced air showers. Since the surface array is not used in the analysis of this thesis, it will not be discussed in detail.

The in-ice antennas are used to detect neutrino-induced cascades in the ice. As explained in Chapter 3, the coherent radio emission from such cascades is expected to be strongest at frequencies between a few hundred MHz and several GHz. The in-ice antennas are designed to be sensitive in a range between 150 and 850 MHz, which is in general the lower end of the Askaryan emission spectrum. While the peak emission of the expected cascades lies at around 2 GHz, these antennas will detect lower frequency emission, which is of inferior strength but more widely distributed in angle as has been explained in Chapter 3. Such wider angular distribution makes lower frequencies more favorable for a detection with small antenna clusters.

Given the constraint that the antenna holes can be drilled into the ice with a maximum diameter of 15 cm, the design of horizontally polarized antennas is particularly

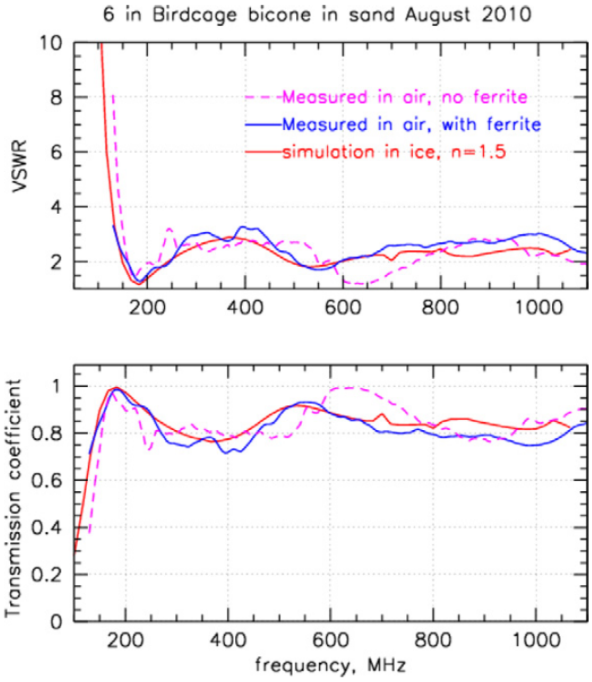
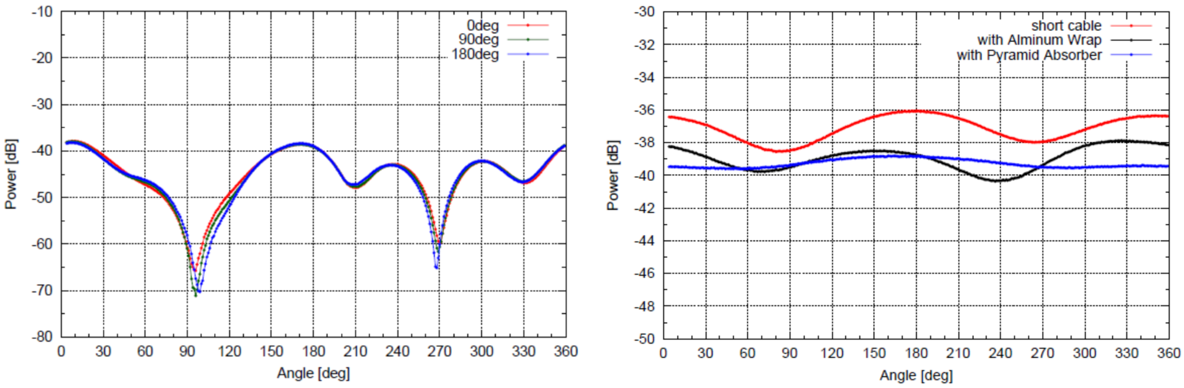


Figure 4.6: *The Voltage Standing Wave Ratio* ***(top)*** *and corresponding transmission coefficient* ***(bottom)*** *of an ARA Vpol antenna. Measurements have been taken with and without ferrite material inside the antenna for testing purposes. The finally used ARA Vpol antenna does not contain ferrite material [75].*



(a) **(b)**

Figure 4.7: ***(a)*** *The zenith gain pattern of the ARA Vpol antenna in air for a frequency of* 500 MHz. *The antenna is rotated completely to check its symmetry. The points of* 0°, 180° *and* 360° *indicate a vertical orientation to the radio source. The colors indicate measurements under different azimuth angles [80].* ***(b)*** *The directional gain pattern of a typical ARA Vpol antenna at a frequency of* 500 MHz*, showing a close to isotropic azimuthal behavior [81].*

challenging. The vertically polarized antennas have a simple wire cage dipole design and their length can be easily adapted to the wavelength of interest (Figure 4.5b). The transmission coefficient and directional gain patterns for a Vpol antenna are shown in Figure 4.6 and 4.7. The transmission coefficient, a measure for the antenna sensitivity, is above 80% over a broad frequency range. Furthermore, the azimuthal gain pattern is very close to an isotropic acceptance. The horizontally polarized antennas instead need special designs to achieve sensitivity in the MHz range. Slotted cylinders show the best results under the given requirements. With the electromagnetic field resonating in the slot between two conducting sheets, the antenna shows a clean polarization and its sensitivity can be tuned to the given requirements, strongly independent from the shape of the antenna body. In the ARA station, copper cylinders cut into four segments by regular slots are mounted around sticks of special magnetic materials. Such ferrite materials additionally push the lower bound of sensitivity to smaller frequencies while the four slots provide a near-isotropic azimuthal gain pattern (Figure 4.8b). However, the Hpol sensitivity, especially at low frequencies, is worse compared to the Vpol antennas, as it is visible when comparing Figure 4.6 and 4.8a. In these two plots, the measure used to represent the efficiency of an antenna is the **V**oltage **S**tanding **W**ave **R**atio **(VSWR)**. Its relation to the reflected power P_r and the transmitted power P_t on an antenna is:

$$\frac{P_r}{P_t} = \frac{(VSWR-1)^2}{(VSWR+1)^2}.$$

Hence, the closer the VSWR value is to 1 the better is the sensitivity of an antenna. This is additionally illustrated in Figure 4.6, which shows the transmission coefficient of a Vpol antenna along with its VSWR.

4.1.2 Signal chain

The in-ice signal chain is displayed in Figure 4.9a. Due to strong South Pole communication transmitters, operating at frequencies of 150 and 450 MHz, the signal of the in-ice antennas has to be refined by notch filters[1], before being fed into the **L**ow **N**oise **A**mplifiers (LNA[2]). If unfiltered, such communication signals would saturate the highly sensitive LNAs and thus blind the detector for the cascade emission, if not damage the electronics. Notch filters are able to sharply exclude signals of narrow frequency bandwidth. In addition to that the uninteresting range of frequencies below 150 MHz and above 850 MHz are excluded to reduce noise contributions to a minimum. The LNAs then provide a gain of roughly 80 dB and the full chain of notch filter and LNA only adds 80 K of system noise, which is well below the thermal noise of the antennas.
From the LNAs the antenna signals are transmitted to the surface by Optical ZONU links[3]. In such links, the incoming electric signal is translated into a modulated laser

[1] model LARK MCN814-038.

[2] model MITEQ AFS3-00200120-10-1P-4-L.

[3] model OZ450-LNA.

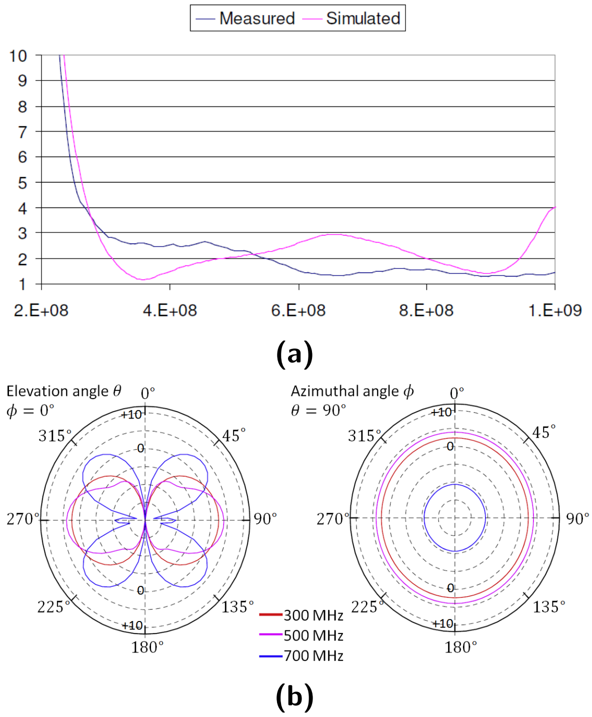
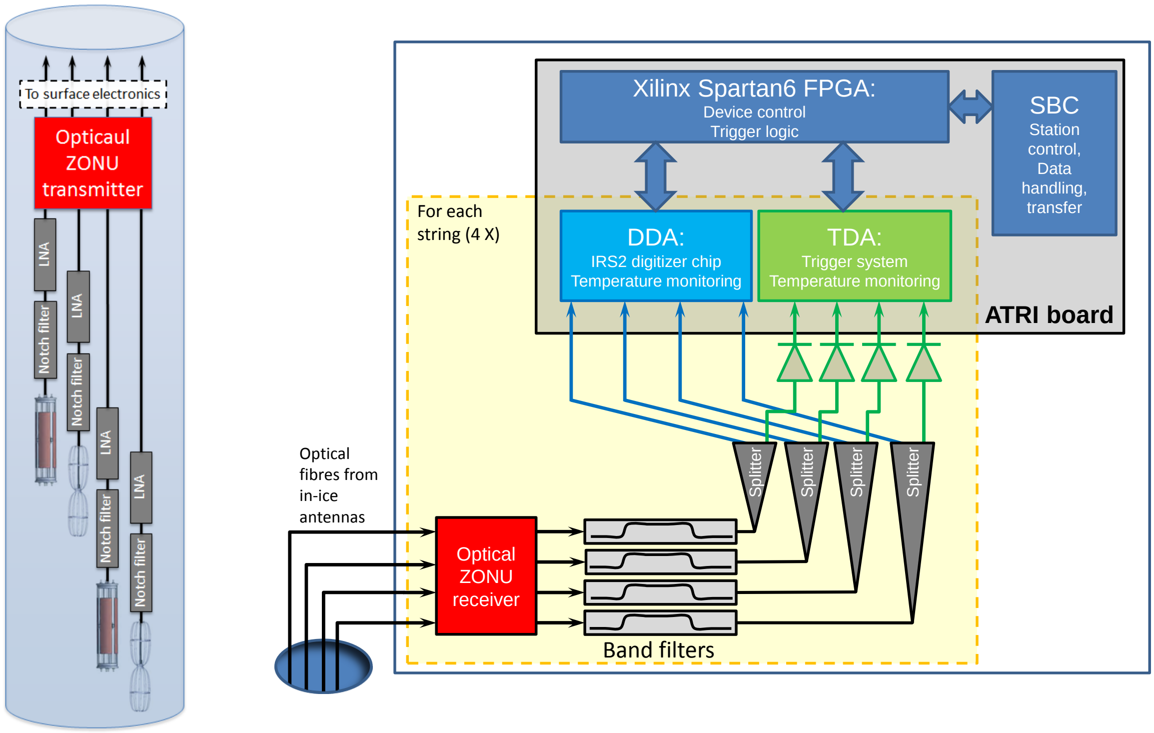


Figure 4.8: ***(a)*** *The measured VSWR versus frequency in Hz of a quad-slot antenna.* ***(b)*** *Its simulated directional gain in dB. In difference to Figure 4.7a angles of* 90° *and* 270° *indicate a vertical antenna orientation towards the source in the zenith response [75].*

beam and transferred through optical fiber cables to the surface, where an inverse conversion to electric signals takes place. Such modules guarantee a transmission without loss and free of possible electromagnetic noise pickup on the way to the surface.

The Data AcQuisition system (DAQ) of the ARA detectors collects all signals from the in-ice antennas at the surface. They are first cleaned through bandpass filters, then split and equally fed into the ARA trigger system and the digitizer chip, the **I**ce **R**ay **S**ampler (IRS2), as shown in Figure 4.9b. For a remote experiment like ARA, a high frequency digitizer running at a low power consumption is a crucial device. With the IRS2 chip, a sampling speed of up to 4 GHz can be reached at a power consumption of less than 20 mW per channel. Moreover, it can sample 8 channels at the same time and has a storage buffer of 32768 samples per channel, which allows for a digitization latency of several micro-seconds. This storage buffer is randomly accessible for the chip internal digitizer. Therefore, the IRS2 chip can in principle operate without dead time, since data can be stored in other free blocks, while a few are locked for digitization. The possibility of precise waveform capture is very important for a proper vertex reconstruction in ARA. The arrival time difference of a signal at two antennas has to be determined with a precision of down to 100 ps, given the size of an ARA station. This can be achieved through cross correlation of precisely sampled waveforms as produced



(a) (b)

Figure 4.9: *(a) A schematic view of the key components in the ARA in-ice readout chain and (b) the DAQ system. Parts in the yellow shaded area are displayed for one string only, for visibility reasons. In reality they appear four times in each DAQ.*

by the IRS2. A more detailed description of the functionality and performance of this chip will be given in Chapter 5.

The IRS2 is mounted on the **D**igitizing **D**aughterboard for **A**RA, short DDA. On this board, incoming data is once more cleaned through bandpass filters, limiting the incoming signal to a range of 150 to 850 MHz. Such a limitation is important to avoid aliasing by high frequency noise at the given sampling speed, normally set to 3.2 GS/s for ARA. In addition, the DDA contains various **D**igital to **A**nalog **C**onverters (DACs) and control signals to configure the sampling and digitization inside the IRS2 chip.

For the trigger system the incoming signals from all antennas are each passed through an integrating tunnel diode[4], which sums the signal power over a few nanoseconds. In that way, the ratio of expected radio signals to thermal noise can be enhanced; additionally short impulsive signal pulses are broadened to be more easily detectable by later stage electronics. This broadening is crucial, since the electronics run typically on a clock frequency of 100 MHz. Hence, short GHz pulses cannot be detected. The diode outputs are fed into the core trigger system implemented on the **T**rigger **D**aughterboard

---

[4] Advanced Control ACTP-1629NC3-39.

for **A**RA (TDA). After the signal is processed through bandpass filters and amplifiers on the TDA, logic signals are discriminated from the diode output by a dynamically set threshold. Whenever the incoming signal is above threshold the logic output is set high ('1'), which is counted as an antenna trigger. The trigger rate depends on the relation between the threshold and the noise RMS of the incoming signal. Therefore, the threshold is adapted by a DAC to changing noise conditions, requiring a constant trigger rate on each input line (Figure 4.10). The further processing of the obtained logic signals is then done in a **F**ield **P**rogrammable **G**ate **A**rray (FPGA[5]) which allows the implementation of complex trigger algorithms based on hit patterns from all incoming antenna signals. The FPGA is the main core of the DAQ system, as it controls and manages all devices of the TDA and DDA and their interaction, via implemented firmware.

As a first trigger a multiplicity condition is set up in the ARA firmware, waiting for at least 3 of 8 Vpol or 3 of 8 Hpol channels to cross the threshold within a time window of 110 ns. The event rate which can currently be accommodated by the transfer system to the data storage is at most 30 Hz. Thus, for this multiplicity trigger, the initial thresholds have to be set high enough to reduce the event rate to less than 10 Hz to protect the transfer from saturation. Under a thermal noise assumption, the single channel rate R which results in a final trigger rate r for such multiplicity trigger can be calculated as:

$$R = \left(\frac{r}{2 \cdot N \cdot \binom{M}{N} \cdot T^{N-1}} \right)^{1/N} \approx 10 \text{ kHz}, \tag{4.2}$$

$$\text{with} \qquad N = 3, M = 8, r = 4 \text{ Hz}, T = 110 \text{ ns}.$$

This leads to the necessity of a very high threshold, put to around 6 standard-deviations above the thermal noise floor. Alternatives for the currently used trigger logic, such as the algorithm described in Section 8.3.2, and their functionality at lower thresholds are currently under investigation.

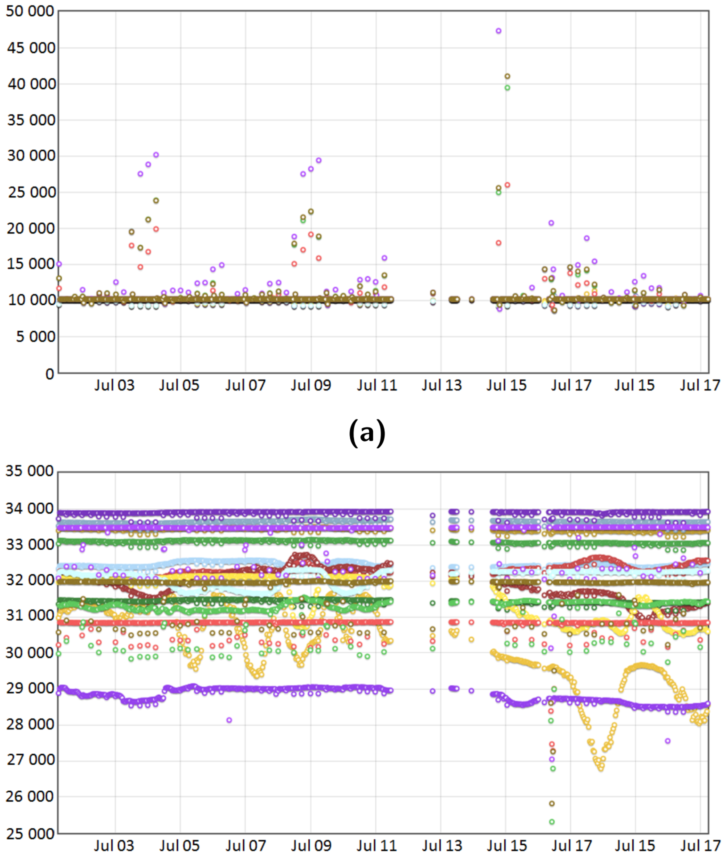
Once a trigger occurs, an **event** is recorded. This means, digitization of data on all channels from the corresponding time frame is asserted in the IRS2 chip. The digitized data are then stored on the internal memory of the station, before being transferred via fiber links to the IceCube Laboratory, where it is stored on disk arrays. For a typical radio-triggered event 400 ns of data are recorded for each antenna channel. This can vary depending on the trigger occurrences.

Data transfer, generally all networking and the FPGA configuration is managed by a **S**ingle **B**oard **C**omputer (SBC[6]), running a lightweight Linux version. Through the SBC, the station can be maintained, reconfigured and monitored by scientists in the northern hemisphere via internet connection. All described parts, the TDA, DDA, FPGA and the

---

[5] model Xilinx Spartan6.

[6] model Kontron, Com Express Mini



(a)

(b)

Figure 4.10: ***(a)*** *The trigger rates in Hertz on single antenna channels for all in-ice antennas of ARA03, recorded over a period of* 20 *days. The rates are relatively flat and close to the nominal value of* 10 kHz*. Outliers are limited to the beginning of a new run, when the threshold still has to be tuned to the correct value.* ***(b)*** *The corresponding antenna thresholds, continuously adapted to keep the trigger rates constant under changing noise conditions.*

SBC are gathered on the **A**RA **T**riggering and **R**eadout **I**nterface, the ATRI-board. On this board all necessary interfaces between the devices, their power supply and various monitoring routines are established.
South Pole is a remote place with limited data transfer capacities. Therefore, only a small fraction of the collected data, approximately 0.5%, can be transferred north to be directly available for analysis. So far the largest amount of the data can only be hand-carried out during the summer season, when the South Pole Station is accessible. Recently, filtering algorithms have been put into place to transfer the most interesting signal candidates alongside with a small randomly picked subset of events. This will allow for a faster data analysis in the future.

4.2 Status and performance of the currently installed ARA stations

In the summer season 2011 - 2012 a first station, ARA01, was deployed in South Pole ice at a distance of two kilometers from the edge of the IceCube neutrino observatory (Figure 4.1). Due to difficulties in drilling, the in-ice antennas of this station could only be installed at a depth of 100 m instead of 180 m. However, the intrinsic station geometry and electronics follow the general baseline design. After various failures and repairs of the used electronics, this station started taking data in early 2014.

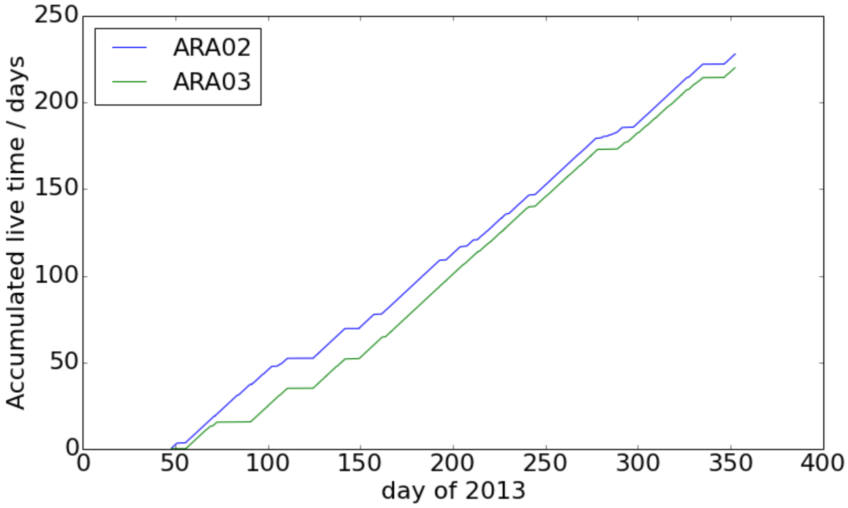


Figure 4.11: *The live time accumulated by ARA02 and ARA03 during the year of 2013.*

After an upgrade of the drilling system, two more stations, ARA02 and ARA03, were deployed at the full depth of 200 m during the season 2012 - 2013. Apart from one single misbehaving channel (Figure 4.15, 4.16), these two stations are successfully taking data since station commissioning in February 2013. Channel 15 (D4BH) on ARA02 shows a very high and unstable noise RMS. The suspicion is that the LNA of this channel got damaged during deployment. Unfortunately, this channel does not produce useful data and will be excluded from further analysis. Figure 4.11 shows the accumulated live time of the two stations during the year 2013. Apart from data taking gaps due to station failures, maintenance or power outages at South Pole, a certain amount of operational dead time has to be subtracted from the time of data taking to get to the actual live time. This dead time is mainly caused by the digitizer. Due to certain features in the DAQ system, the data sampling is stopped during digitization of a triggering event. Such digitization takes a few hundred microseconds which, at an event rate of less than 10 Hz, amounts to less than 10 ms dead time per second. Thus, while the station is running, it accepts incoming events more than 99% of the time. The small fraction of digitization dead time is monitored in the ARA stations and can be accounted for in the live time calculation. The final accumulated live time during the year 2013 amounts to 228.0 days for ARA02 and to 220.1 days for ARA03.

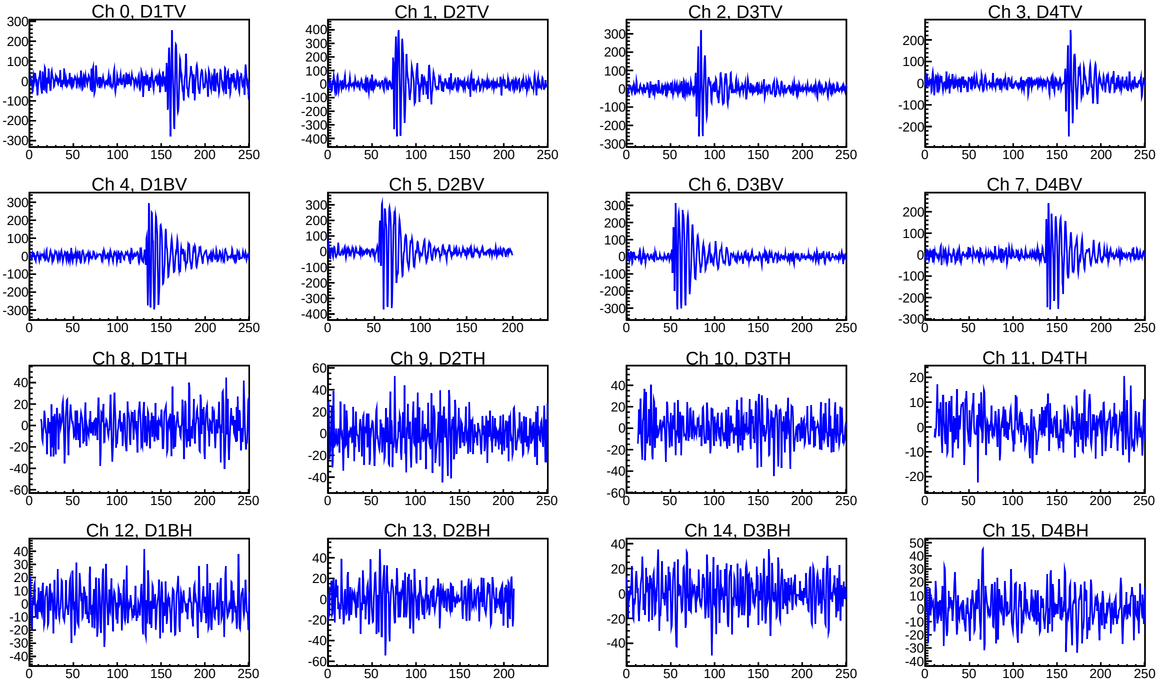


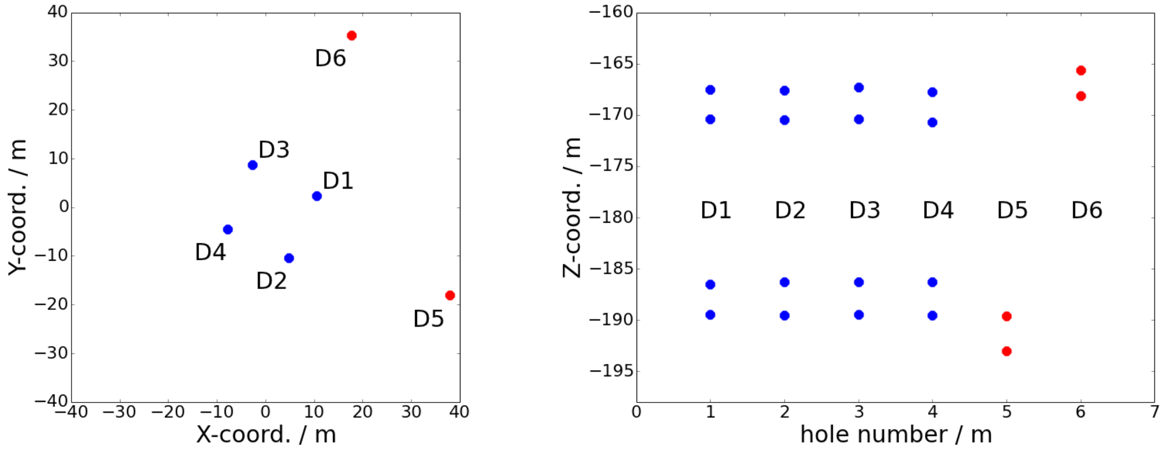
Figure 4.12: *The antenna response on a signal from Vpol calibration pulser D5VP on ARA03 in mV versus ns. Only Vpol antennas react to this signal.*

As the RMS distributions in Figure 4.15 and 4.16 show, noise conditions appear to be very stable during the year, apart from a few episodic noise incidents. Station functionality can be tested with the various calibration pulsers. In Figure 4.12 a signal received from one of the calibration pulsers is displayed, demonstrating the functionality of the in-ice Vpol antennas on station ARA03. The plot shows the clean polarization of the Vpol emitter and Hpol receivers, which do not show any reaction to a rather strong vertically polarized signal.

The antenna positions in station-centric coordinates are shown in Tables 4.2 and 4.3, as well as in Figure 4.13 and 4.14.

Table 4.2: *The coordinates of all ARA02 antennas in the station-centric coordinate system (Section 6.1).*

	Polarization	Antenna name	X [m]	Y [m]	Z [m]
Pulser 5	horizontal	D5HP	37.87	-18.05	-189.64
	vertical	D5VP	37.87	-18.05	-192.93
Pulser 6	horizontal	D6HP	17.75	35.34	-164.65
	vertical	D6VP	17.75	35.34	-167.94
Hole 1	vertical	D1TV	10.59	2.34	-170.41
		D1BV	10.59	2.34	-189.50
	horizontal	D1TH	10.59	2.34	-167.49
		D1BH	10.59	2.34	-186.55
Hole 2	vertical	D2TV	4.82	-10.38	-170.49
		D2BV	4.82	-10.38	-189.54
	horizontal	D2TH	4.82	-10.38	-167.57
		D2BH	4.82	-10.38	-186.26
Hole 3	vertical	D3TV	-2.68	8.68	-170.39
		D3BV	-2.68	8.68	-189.44
	horizontal	D3TH	-2.68	8.68	-167.27
		D3BH	-2.68	8.68	-186.32
Hole 4	vertical	D4TV	-7.72	-4.47	-170.67
		D4BV	-7.72	-4.47	-189.56
	horizontal	D4TH	-7.72	-4.47	-167.72
		D4BH	-7.72	-4.47	-186.28



(a) **(b)**

Figure 4.13: ***(a)*** *The ARA02 geometry in* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *coordinates.* ***(b)*** *The depth of ARA02 antennas in the ice, as noted in Table 4.2. Blue dots represent the receiver antennas. Red dots represent the calibration pulsers.*

Table 4.3: *The coordinates of all ARA03 antennas in the station-centric coordinate system (Section 6.1).*

	Polarization	Antenna name	X [m]	Y [m]	Z [m]
Pulser 5	horizontal	D5HP	38.01	-16.40	-194.90
	vertical	D5VP	38.01	-16.40	-198.18
Pulser 6	horizontal	D6HP	17.94	36.39	-184.73
	vertical	D6VP	17.94	36.39	-188.01
Hole 1	vertical	D1TV	4.41	-9.39	-173.39
		D1BV	4.41	-9.39	-192.45
	horizontal	D1TH	4.41	-9.39	-170.27
		D1BH	4.41	-9.39	-189.20
Hole 2	vertical	D2TV	10.69	3.51	-173.97
		D2BV	10.69	3.51	-192.70
	horizontal	D2TH	10.69	3.51	-171.02
		D2BH	10.69	3.51	-189.74
Hole 3	vertical	D3TV	-2.01	9.41	-174.11
		D3BV	-2.01	9.41	-192.67
	horizontal	D3TH	-2.01	9.41	-170.63
		D3BH	-2.01	9.41	-189.71
Hole 4	vertical	D4TV	-8.10	-3.71	-173.55
		D4BV	-8.10	-3.71	-192.60
	horizontal	D4TH	-8.10	-3.71	-170.59
		D4BH	-8.10	-3.71	-189.48

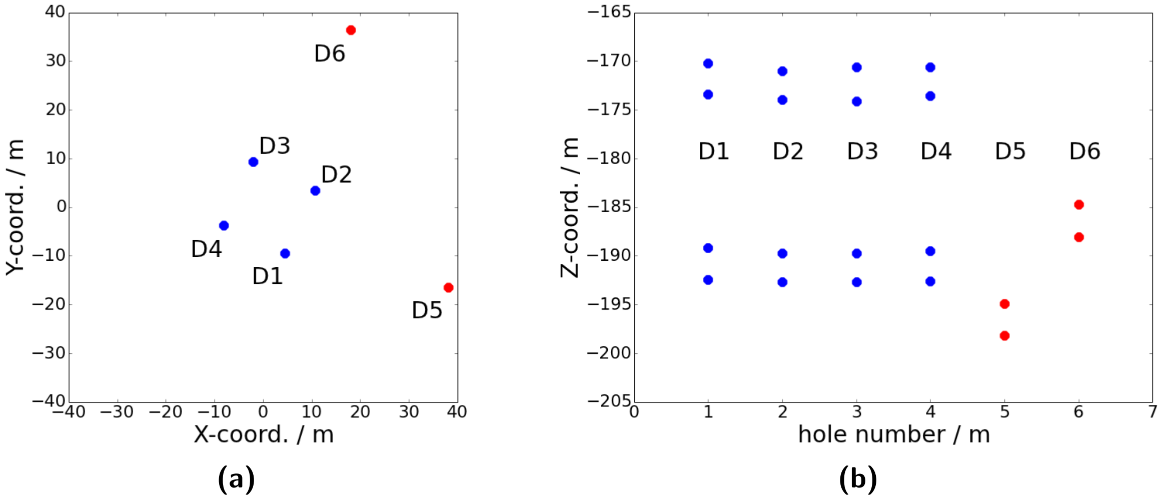


Figure 4.14: *(a) The ARA03 geometry in* X *and* Y *coordinates. (b) The depth of ARA03 antennas in the ice, as noted in table 4.3. Blue dots represent the receiver antennas. Red dots represent the calibration pulsers.*

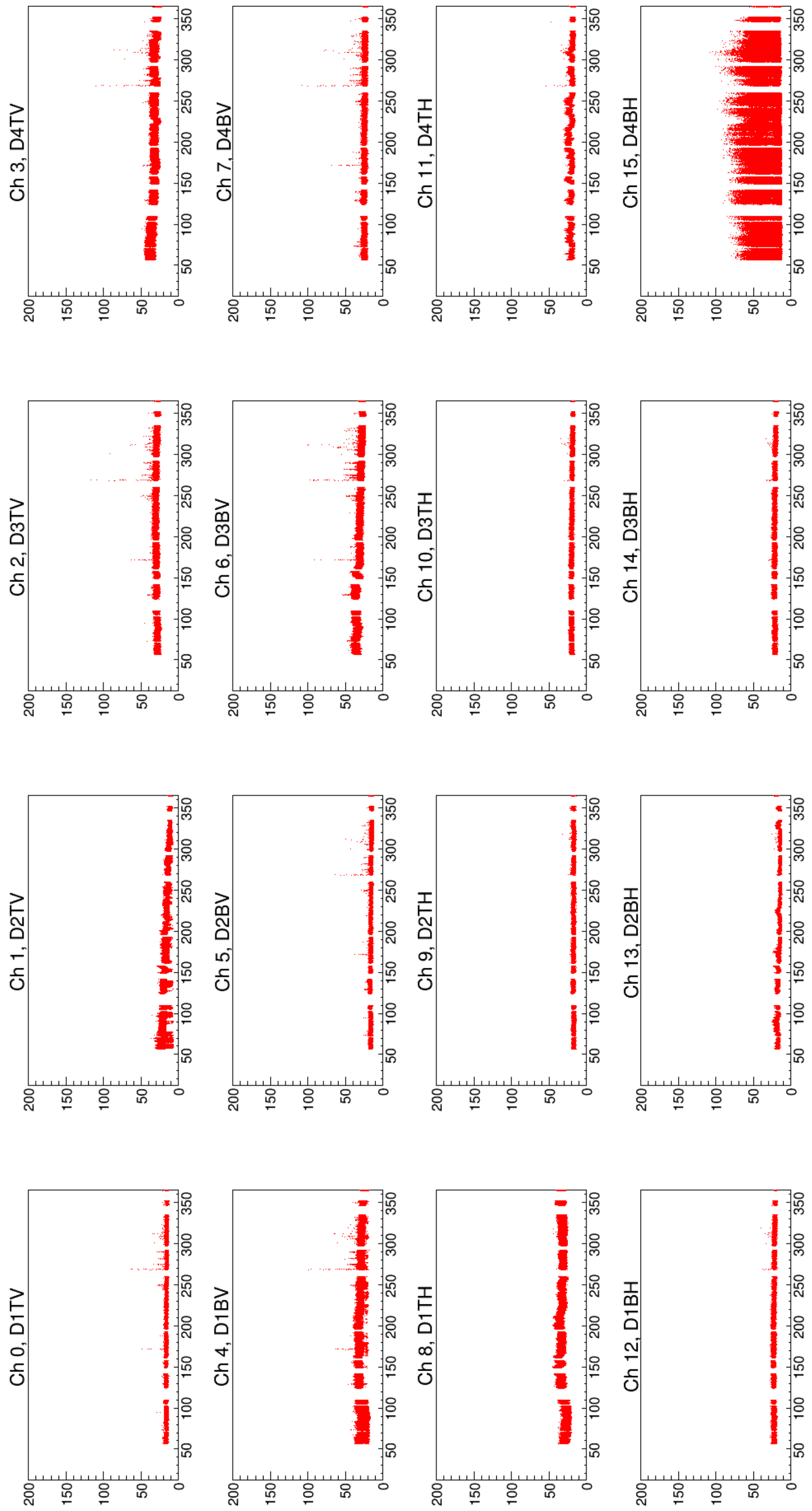


Figure 4.15: *The ARA02 RMS per event in mV versus the day of 2013 for all 16 in-ice channels, showing the very noisy D4BV channel.*

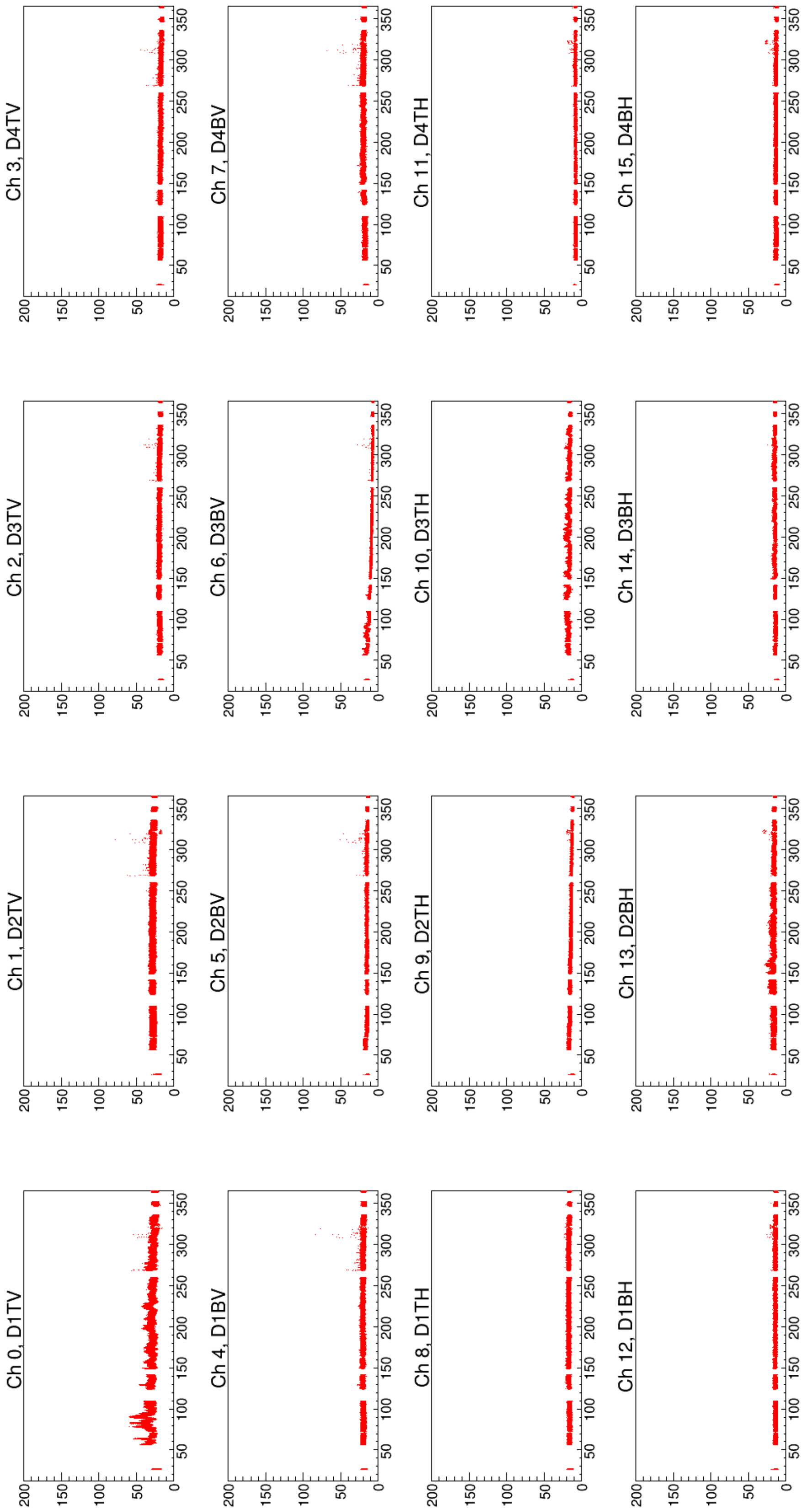


Figure 4.16: *The ARA03 RMS per event in mV versus the day of 2013 for all 16 in-ice channels.*

4.3 The ARA detector in comparison to other neutrino detectors

Apart from the ARA detector there is a number of already built and planned experiments which are sensitive to neutrinos with energies above 10^{16} eV. To be mentioned here are operating detectors like the IceCube Observatory, the ANITA balloon experiment and the Pierre Auger Observatory [32, 82, 7]. They are currently the most sensitive detectors to a GZK neutrino flux (Figure 4.17).

The IceCube detector is an in-ice sensor array, detecting neutrinos through the Cherenkov light emission from secondary particles in the ice. It consists of 5160 optical modules, deployed on 86 strings in the ice below 1500 m. With an instrumented volume of about 1 km^3 it is currently the most sensitive detector for neutrinos from 10^{10} eV to roughly 10^{19} eV. In fact, IceCube has been able to claim the first observation of astrophysical high-energy neutrinos [33]. However, no detected event deposited an energy significantly higher than 2 PeV in the detector, which does not fit expectations from GZK neutrinos. The limits shown in Figure 4.17 have been derived in [83].

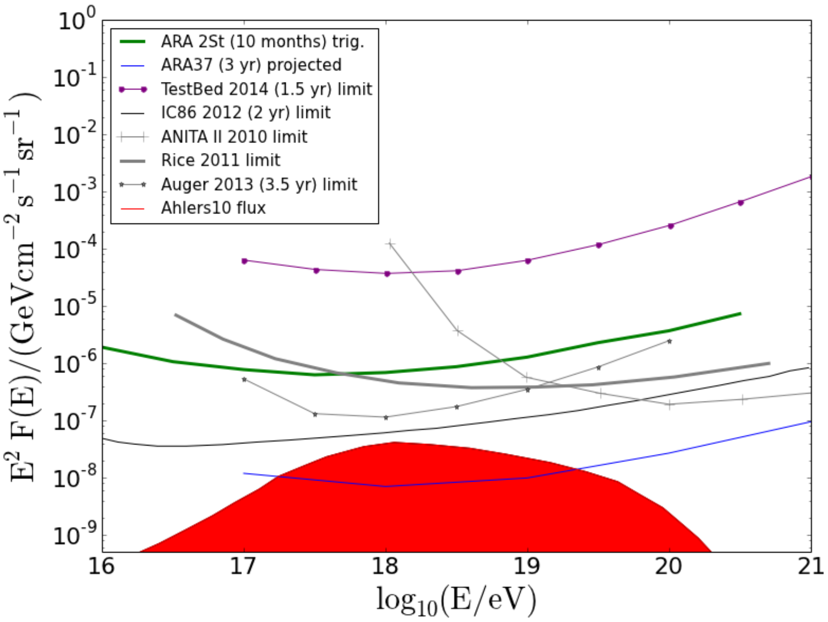


Figure 4.17: *The expected sensitivity and recent limits from current and future GZK neutrino detectors compared to a flux estimation from [52].*

The ANITA experiment consists of a dense array of radio antennas, mounted on the payload of a helium balloon [82]. The antennas operate in a similar frequency range as the ARA antennas and ANITA is also looking for coherent radio Cherenkov emission from neutrino-induced cascades. ANITA has performed two flights of roughly 30 days each at an altitude of around 35 km. The goal is to detect radio signals that leave the ice at the surface and travel nearly unattenuated through air to the balloon payload. By this means, a volume on the order of 10^6 km^3 can be surveyed. However, the solid angle of observation is very limited. ANITA is currently the most sensitive neutrino detector for energies above $10^{19.5}$ eV.

The Pierre Auger Observatory is primarily built as an air shower array next to the Argentinian Andes [84]. It consists of an array of 1660 water Cherenkov detectors, the surface detectors, spaced by 1.5 km and distributed over an area of roughly 3000 km^2. Each water tank is equipped with three photo-multipliers to detect Cherenkov emission from particles produced in air showers. In addition, 27 fluorescence detectors are placed at various locations around the surface detectors, overlooking the full array. Recently, radio antennas have been added to investigate a third detection method. The highest sensitivity of the Pierre Auger Observatory to GZK neutrinos is provided through the so-called Earth-skimming effect. Tau neutrinos, approaching from the horizon, are expected to produce τ particles in the mountains surrounding the detector. These particles can decay above the detector and induce a highly inclined air shower [85]. In principle there are no tau neutrinos produced in the GZK mechanism. Due to neutrino oscillations we expect however all three flavors to arrive in equal parts at Earth. In Figure 4.17 the limit of the Pierre Auger Observatory on GZK neutrinos from [85] is shown.

Another radio detector of GZK neutrinos is the RICE array, the direct predecessor of the ARA detector [86]. It is a small scale array of radio antennas deployed above the AMANDA detector at South Pole. Its purpose was, amongst others, the investigation of measurement conditions and testing of equipment. During its operation, the array collected more than 10 years of data, providing the neutrino limit shown in Figure 4.17.

In addition to the above mentioned, several neutrino detectors are planned or currently under construction, like the ARA detector. Another radio detector in dense media is the ARIANNA experiment. This array is currently being deployed in the Ross ice shelf at the antarctic coast. The detector principle is the same as for ARA. Differences are the antenna depth and the surrounding conditions [87]. The antennas in ARIANNA are mounted close to the surface instead of in the deep ice. The Ross ice shelf is only a few hundred meters thick, floating on seawater. Furthermore, there is no thick firn layer as for example at the South Pole, which allows for a shallow antenna deployment. The idea is to achieve an enhancement of the effective volume by the reflections of neutrino signals off of the top and bottom ice boundaries. The expected sensitivity is comparable to the one of the ARA detector.

There are several more approaches and ideas to detect neutrinos in the GZK range and above. There is acoustic neutrino detection in ice and water [88, 89] and radio detection of neutrino interactions in the moon [90]. Both are mostly suited for very high energies above 10^{20} eV, which only overlaps with the high end of the ARA sensitivity spectrum and the GZK neutrino flux.

In Figure 4.17, the sensitivity of two ARA stations after ten months of operation and the sensitivity of the full ARA37 detector are shown at trigger level. The sensitivies are obtained from simulations as described in Chapter 7. One can note that the two ARA stations are already close to be competitive with the current best detectors. The full ARA37 detector will improve the currently best sensitivity to GZK neutrinos significantly.

5

The calibration of the Ice Ray Sampler (IRS2) chip

In this chapter, the calibration of the IRS2 chip [91], the data sampling and digitization chip in ARA, is described. This custom ASIC is recording the signal, received by the ARA antennas, at a sampling speed of several GHz. Such a precise waveform sampling is crucial for the determination of signal timing on a sub-nanosecond level, which is important for vertex reconstruction. Data recorded by this chip need to be calibrated in timing and amplitude. As we will see in the following, a precise sample timing has the strongest impact on the parameters used in the analysis. Voltage calibration plays a secondary role, but still adds improvement to the data. This will become more important once ARA has recorded neutrino signals, which need to be analyzed precisely in signal amplitude, frequency content and polarization.

In the following, first the working principle of the IRS2 chip is described with focus on the parts to be calibrated. After that the calibration methods are explained and finally results and quality checks on this calibration are discussed.

5.1 The IRS2 chip

The IRS2 is the second generation of the IRS chip, a custom ASIC for high speed sampling and digitization [92]. It is designed for radio astroparticle physics and other experiments which require high numbers of channels with deep analog buffers and waveforms sampled at GHz rates. Such experiments are often located in remote places like the South Pole, on balloons or satellites, where low power consumption is an important feature of all used components. The IRS2 chip has the possibility to sample up to 8 channels simultaneously at a speed of up to 4 GS/s and with a power consumption of less than 20 mW per channel. Moreover, it provides an analog storage buffer of 32 kS per channel, allowing for a digitization latency of several microseconds. Digitization is performed with the Wilkinson method, which is relatively slow but precise, physically compact and of low power consumption [93, 94].

A chip such as the IRS2 can be realized by using compact **C**omplementary **M**etal **O**xide **S**emiconductor (CMOS) technology. By combining n-channel and p-channel **F**ield-**E**ffect **T**ransistors (FETs), power consumption can be limited to the instance of a logical switch. Therefore, the general power consumption is low compared to other continuously power drawing technologies. It increases however proportional to the frequency of operations which determines the speed of the particular device. In the IRS2 chip this problem is avoided by using special techniques like **S**witched **C**apacitor **A**rrays (SCA) and delay lines. Such setups, as will be explained in the following, allow to run single components at a fraction of the required sampling speed. In that way the power consumption of the chip is reduced significantly. In addition to that, the deep sampling buffer of 32 kS per channel, which can be randomly accessed in blocks of 64 samples, gives time for an advanced pre-selection of data. This allows the digitization to be run at low speed without inducing significant dead time.

The used technology has the above described benefits of high sampling rates at low power consumption but renders the calibration of the sampling device more complicated than for ordinary digitizing systems.

The calibration discussed in this chapter is focused on two main points: the precision of the sample timing and the conversion of the **A**nalog to **D**igital **C**onverter (ADC) output to a corresponding input voltage. This chapter will not give a detailed description of the digitizer chip but focus on the main functionality, which is important for a proper interpretation of the collected data.

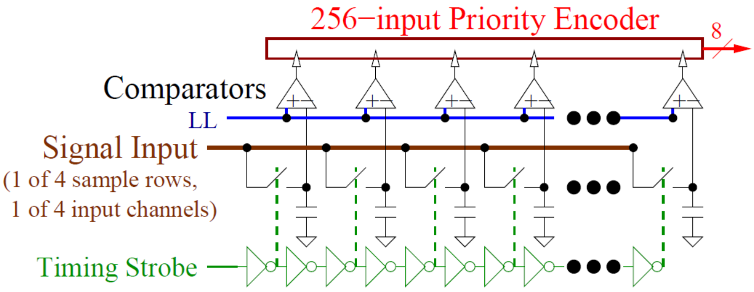


Figure 5.1: *A* ***S****witched* ***C****apacitor* ***A****rray (SCA) with delay line. The signal comes in on the brown line and is at closed switch stored on the respective capacitor below. The delay line is drawn in green, fed with the timing strobe, which is transferred to each switch through the delay elements, indicated as side-pointing triangles. The connection to the switches is indicated by the vertical green dashed line [94].*

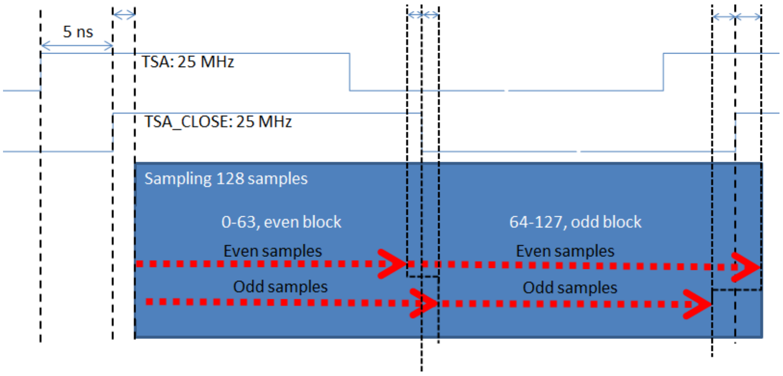


Figure 5.2: *The schematic sampling process in the IRS2 chip; a sampling round of 128 samples is initialized by the rising TSA close signal. Red lines indicate the delay lines with even and odd samples. Possible shifts in time during the sampling are marked with the red arrows and dashed vertical lines.*

5.1.1 Sample timing

The data sampling in the IRS2 is performed by an SCA, as shown in Figure 5.1. In the IRS2 chip this array consists of 128 capacitors, each connected to the input line through a switch. A closed switch means the given capacitor runs in tracking mode, carrying the same voltage as currently applied to the input line. Once the switch is opened the capacitor holds its voltage, which in this state can be transferred to the analog storage array. This will be referred to as storage mode. The switches are controlled by the timing strobe (TSA-CLOSE signal in Figure 5.2). This is a slow clock signal which is fed through a delay line to reach the switches in sequence with a time difference of a few hundred picoseconds between two switches. Its frequency is set up such that all 128 capacitors are reached within one period. The used delay line consists of a row of current starved inverters each delaying the timing strobe signal by a few hundred picoseconds. To achieve different sampling times, the switches of the SCA are connected to the delay line separated by two delay elements each. Once the timing strobe passes through a number of delays and reaches a switch, this switch is opened and the capacitor changes from tracking mode into storage mode. This process is repeated for each capacitor with the frequency of the timing strobe.

In the IRS2 chip the 128 sampling switches are connected alternately to two delay lines, run in parallel and shifted by half the delay between two switches. The time of delay between switches can be adjusted via the input V_{adj} down to a minimum delay of 0.5 ns. The minimum shift between the two delay lines is accordingly 0.25 ns.

In the ARA detector, the timing strobe runs at a frequency of 25 MHz, the delay lines are shifted by 0.31 ns between each other and their delay elements are tuned to 0.625 ns. In this way, a sampling speed of 3.2 GS/s is achieved, when combining even and odd samples. The SCA is set up as a ring buffer. Once the last capacitor has been locked to storage mode, sampling starts again on the first. In the meanwhile, data is transferred to the storage array in a ping pong fashion. Once half of the SCA is sampled, this part is transferred while sampling proceeds in the second half of the capacitors.

It becomes apparent that not many elements in this setup are operating at a frequency close to the achieved sampling frequency, which allows for the low power consumption of the chip as explained in Section 5.1.

There are different things to be calibrated in the sampling system of the IRS2. Due to process variation during the chip production, there can be systematic offsets of several 100 ps between the delay elements on one line [95]. These need to be determined and corrected for each of the 128 samples, since they will disturb the timing of the sampled waveform and add high frequency noise components. Both render the waveforms less precise for further analysis. In addition, the average speed, tunable by V_{adj}, has to be measured and corrected towards the nominal value for the two delay lines.

5.1.2 Sample storage and digitization

Once samples are locked on the sampling capacitors they are transferred to the capacitors of the analog storage array. The transfer is performed in a ping pong fashion. From the 128 sampling capacitors, voltages are transferred in sample blocks of 64, while the other half of the capacitors are still available for sampling. The settling time of one block, the time that 64 samples need to be properly stored, amounts to 16 ns. With such a settling time and by using the ping pong transfer method, dead time can be completely avoided for a sampling speed of up to 4 GS/s. The capacitors of the sampling array can be switched back into tracking mode before being reached by the next rising edge of the timing strobe.

Of the above mentioned sample blocks, 512 can be stored in the analog buffer of the IRS2 chip. Thus, the buffer can hold in total 32768 samples, which allows for temporary buffering of up to 10 µs at a sampling speed of 3.2 GS/s. Once stored, an external trigger decides if a given sample block is locked and digitized or instead will be overwritten in one of the following sampling cycles. The storage and digitization process is guided by an FPGA, which selects the address of a block for storage as well as it selects blocks for digitization. In principle this selection can be performed freely within the block array and blocks queuing for digitization are temporarily excluded from the sampling buffer. FPGA logic can thus be implemented to completely avoid digitization dead time up to a trigger rate of 1 kHz. This feature is novel and not entirely functional yet since digitization and storage can not properly be performed at the same time on the IRS2. However, problems have been identified and future chip generations should allow for a simultaneous execution of the two processes. In ARA the blocks are currently accessed in sequence and for the used trigger rate of a few Hertz a dead time of roughly 0.1% has to be accepted.

After a sample block is positively tagged through an external trigger, the samples are digitized with a classical Wilkinson method [93]. A comparator voltage is ramped through an extensive voltage range and time is counted until this comparator matches the sample voltage. The count result is saved as ADC counts. The counting speed is determined by the Wilkinson clock, which in the IRS2 chip is running at a speed of 740 MHz [91]. Both, the ramping speed as well as the clock speed can be adjusted via an external voltage V_{dly}. In the ARA stations V_{dly} is automatically adjusted to compensate for temperature changes and to keep the conversion gain constant. In the current settings, a range of about 1.5 V input voltage is encoded with a 12 bit counter, providing a precision of about 0.4 mV. It should be noted that from this method we expect a linear relation between ADC counts and the input voltage, which is important for the calibration.

On the chip there are different ways to corrupt the incoming voltages during sampling, storing, transfer and digitization. Parts of these processes are individual for each of the 32768 storage cells, due to their physical location on the chip. Moreover, every storage cell is unique due to small differences during the production of the chip [95]. Hence, they have to be treated individually for each single sample. These differences during various stages of sample processing are very complex to model individually and the recorded calibration data does not provide enough information to address a full analysis of such processes. Therefore, in the here described calibration the focus lies on providing an ADC to voltage conversion gain for the full sampling chain of each single storage cell.

5.2 The calibration

The calibration of the IRS2 chip for ARA is designed in a way to calibrate the timing precision and ADC to voltage conversion for the full sampling and digitization chain and not for single processing components inside the chip. Such more detailed calibration would be very difficult due to limited documentation and data. It is also not crucial for the given application. The parameters that will be analyzed in the following are:

- the timing correction and precision of the SCA,
- the ADC to voltage conversion factors for incoming RF signals for each storage element,
- the frequency response of the different channels,
- the influence of temperature on the sample timing.

5.2.1 The data set

In the stations ARA02 and ARA03, 8 IRS2 chips are used. The calibration data for all chips was taken in freezers or atmospheric chambers in the laboratory to emulate conditions of operation in South Pole ice. The chip is installed on the Digital Daughter board for ARA (DDA). Input sine waveforms of different frequency and amplitude were generated with an external waveform generator and fed to the DDA. On the DDA a bias voltage was applied to the RF inputs, raising the zero level to 1.25 V, since the IRS2 chip can only digitize positive voltages [96]. The bias is removed in the data via a pedestal subtraction, with average pedestals for each sample, measured once per day during operation. The incoming RF channels on the DDA are fed through a low- and a high-pass filter consecutively, limiting the signal bandwidth from roughly 150 MHz to 850 MHz.

The above explained voltages V_{adj} and V_{dly} are set in a way, to allow for a proper sampling speed and ADC-to-voltage conversion gain inside the IRS2 chip. Earlier checks revealed a rising corruption of sample timing and voltage with channel number, which are not yet fully understood but expected to be caused by the chip architecture. For this reason the RF channels used are typically channel 0 to 3. On two chips per station additionally channel 4 and 5 are used for the surface antennas, which operate at lower frequency. The data of a given frequency and amplitude is taken in runs of a few thousand waveforms.

5.2.2 Timing calibration

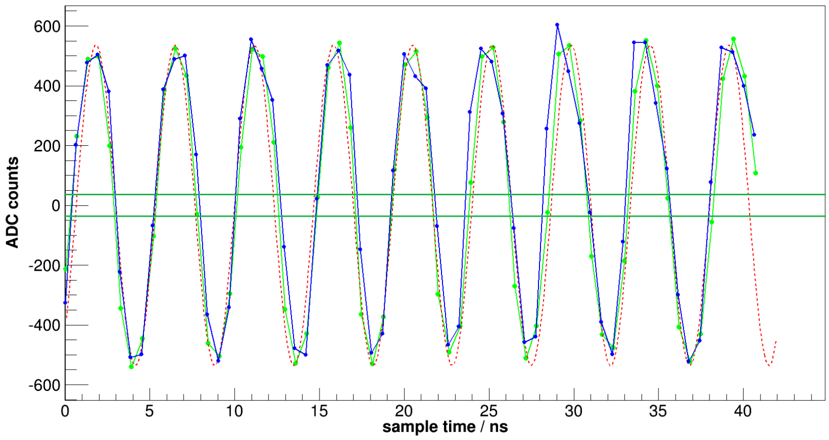


Figure 5.3: *The application of timing calibration on an example waveform with odd samples (**green**), even samples (**blue**) and the fit to the odd samples (**red**). From all samples between the horizontal green lines a timing correction towards the fit is calculated.*

The timing calibration is applied to correct for an imprecision in the V_{adj} voltage and to calculate a systematic correction for each individual delay element in the SCA. The goal of this calibration is to achieve an individual sampling error of $O(100\text{ ps})$.

The calibration is performed by fitting sine waves to blocks of 40 ns of the recorded data and correct for the difference in multiple iterations. The data blocks are fit separately for even and odd samples, which are running on two separated delay lines. The first iterations are focused on the calibration of the average sampling speed. Each waveform i in the data set is fit by

$$f_i(t) = A_i \cdot \sin(2\pi k_i t - \phi_i), \tag{5.1}$$

with individual amplitude A_i, frequency k_i and phase ϕ_i. Since the waveforms are pedestal corrected and consequently expected to be symmetric around zero, the fit does not need a vertical offset. An example of a calibration waveform with fit can be seen in Figure 5.3. From the average of the frequency parameter k_i a correction factor for the sampling speed is calculated as

$$k_s = \frac{f_{in}}{< k >}, \tag{5.2}$$

where f_{in} is the set input frequency of the calibration data and $< k >= 1/N \sum k_i$ the average sine wave frequency of the fit. From the fits above also the shift between the two delay lines is calculated as

$$c_s = \frac{|\phi_{i,even} - \phi_{i,odd} - \phi_0|}{\phi_0}. \tag{5.3}$$

The parameter ϕ_0 is the nominal delay of 0.31 ns. Samples are counted from 0 and even samples are assumed to start at a fixed time t_B, the starting time of a sampling round. The correction factors k_s and c_s are applied to the sampling for the next iteration.

After correcting the overall sampling speed in a few iterations, the single sample jitter for each of the 128 delay elements is recorded as the timing difference between the fit at the sample voltage and the sample timing itself. To minimize the influence of voltage errors, only samples with an absolute ADC count of less than 30 are taken into account. At those points the derivative of the wave is very high and timing errors have a maximal influence on the sample position relative to the sine fit: the influence of incorrect voltages is minimal. The uncertainty on the pedestal values can be estimated as ±5 ADC counts. Considering a sine wave frequency of 214 MHz and an amplitude of 300 ADC counts for the calibration waveform, this leads to an error of about 12 ps which is negligible.

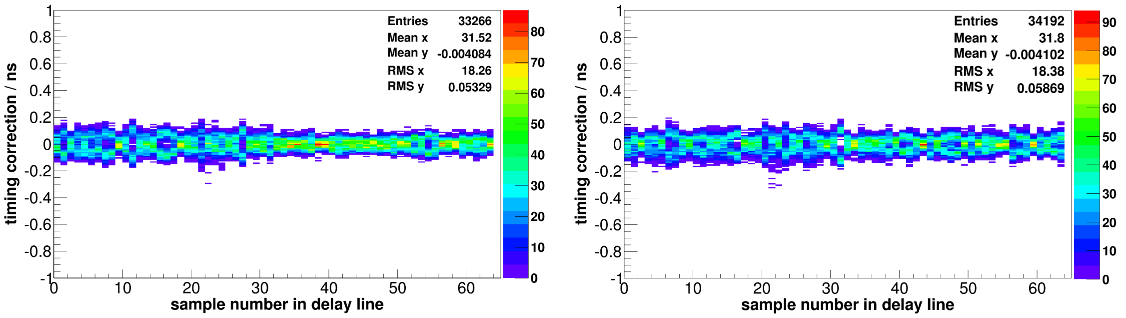
The calibration constant c_j for each sample j at nominal time t_j and recorded with an ADC count U_j is finally calculated as

$$c_j = \frac{|f^{-1}(U_j) - t_j + d_0|}{d_0}. \tag{5.4}$$

The parameter d_0 denotes the nominal delay time of 1/1.6 ns. The calibrated true time for each sample is then calculated from

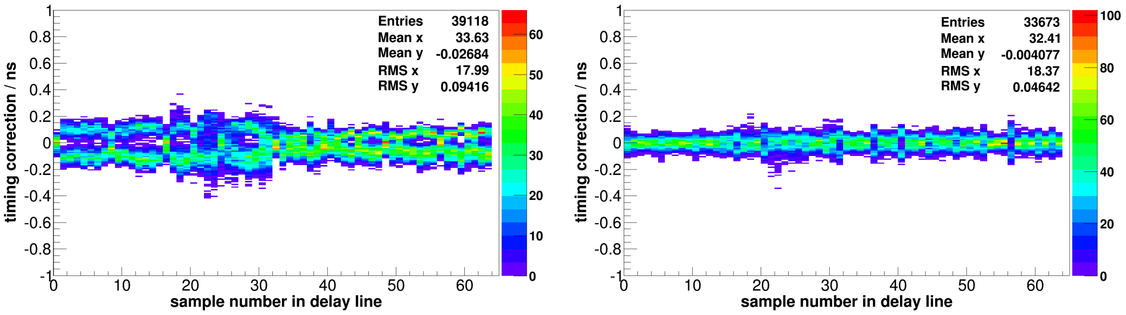
$$t_{cal,j} = \begin{cases} t_B + k_s * d_0 \sum c_j, & \text{even sample j} \\ \\ t_B + k_s * d_0 \sum c_j + c_s * \phi_0, & \text{odd sample j.} \end{cases} \tag{5.5}$$

The offset time t_B for each sampling round is to be taken in steps of 40 ns. This calibration process is repeated several times and the calibrated times are used as input for the next iteration of calibration constants. The difference $f^{-1}(U_j) - t_j$ after 6 iterations for the first channel of one IRS2 is plotted in Figure 5.4. In this case a good convergence has been reached. However, the distributions for some samples in the plot show an unexpected spread, which appears to be even more pronounced in examination of the time difference for higher channel numbers, shown in Figure 5.5.



(a) Even samples **(b)** Odd samples

Figure 5.4: *Collected timing correction for samples of the two delay lines after 6 iterations of individual sample correction, for channel 0 on DDA2 in station ARA03.*



(a) Even samples **(b)** Odd samples

Figure 5.5: *Collected timing correction for samples of the two delay lines after 6 iterations of individual sample correction, for channel 2 on DDA2 in station ARA03.*

The observed spread is connected to the linearity of the samples, which is observed to degrade with rising channel number on a chip. In general, for the first four channels reasonable calibration values can be obtained while the calibration of higher channel numbers looks much more difficult. Such behavior can be reproduced in a similar way on all chips.

The quality degradation is not unexpected given the novel architecture of the IRS2 but has not been fully understood yet. It was found that the affected samples react differently on the rising and falling edge of the waveform (Figure 5.6). Investigations to resolve this behavior didn't lead to a clear result. Several possible sample corrupting processes have been investigated as summarized in the following:

- **A sample offset**, caused by incorrect pedestal treatment can be excluded as a source of error. Waveforms without signal have been checked and the pedestal correction works well within a range of ±5 ADC counts, which is negligible as an error source.

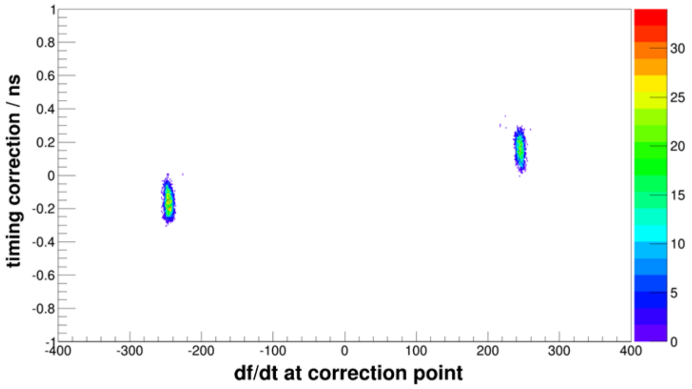


Figure 5.6: *A histogram of the timing correction versus the derivative at the correction point, within ±30 ADC counts, for all even samples of one channel. A clear difference between the rising and the falling edge can be observed. Samples on the rising edge need a positive correction, i.e. they come early compared to the fit, samples on the falling edge appear to be delayed.*

- **Inter channel crosstalk** has been tested by injecting sine waves to single channels, as opposed to all channels at the same time. The neighbors of the connected channel are then examined for signal leakage. Such leakage was measured to be limited to about 5% of the signal in the neighboring channel. Hence, this effect can as well be excluded to have an influence on the timing correction.

- **Inter sample crosstalk** is very difficult to measure since there are many factors playing a role at the same time. In the given time frame no investigation could be lead, to proof or disproof the existence of such crosstalk. Still, if only looking at a first order effect, crosstalk between two samples would be expected to be symmetric in both directions. Therefore, it can not explain the systematic shift between the rising and the falling edge.

- A so called **memory effect** describes a remnant charge on a given sample capacitor after one sample round of 128 samples. This charge will then add to the new charge coming in with a new signal input on that capacitor. In such a way, a sample is influenced by the value sampled 40 ns ago. This effect is not very likely to take place, since capacitors are normally running in tracking mode for some amount of time during which remnant charges should be completely erased. The memory effect could be excluded as the main corrupting influence on the timing calibration with a calibration using a 200 MHz sine input. At that frequency the sampled voltage on a certain capacitor is always the same after 40 ns. Only an error on the input frequency and on the earlier described timing strobe can disturb this equality. However, in the range of their possible error the influence is very small. In the 200 MHz calibration the corruption of samples has been found similarly pronounced as at other frequencies. Hence, a significant influence of the memory effect can be excluded.

As explained, the above mentioned effects can be excluded as main error sources on the sample calibration in the IRS2 chip. The cause of the systematic sample corruption remains unknown and needs more detailed investigations about the chip, which can not be conducted in the frame of this work, due to limited material available. Measurements, to understand the observed features could be fine voltage and frequency scans with high statistics, allowing for detailed fits. Furthermore, narrow pulses should be applied in a regular manner. Their deformation after digitization could reveal inter sample crosstalk. To understand external influences on the chip behavior, calibrations could be performed with different bias voltages.

In any case, the observed data corruption does not cause systematic errors bigger than 150 ps which, as we will see in the following, are acceptable for the given goal of reconstruction precision.

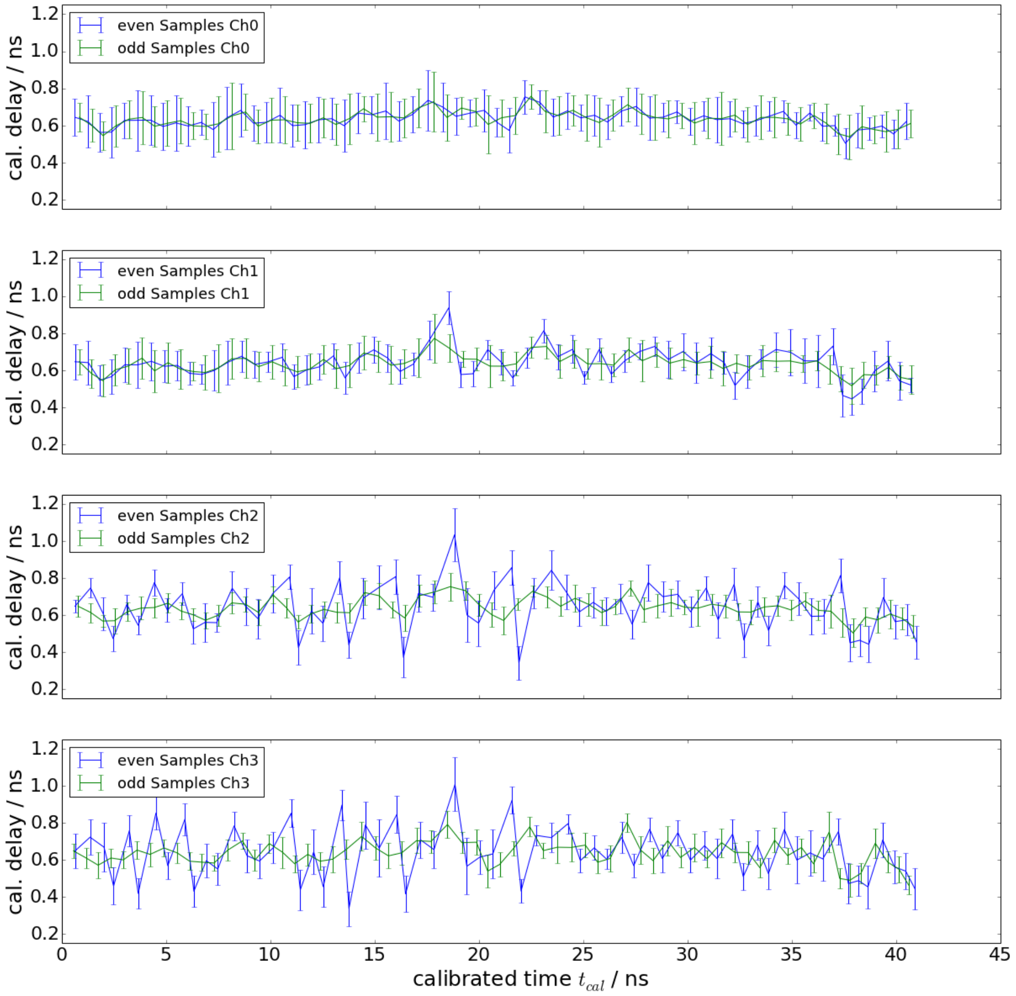


Figure 5.7: *Timing calibration for the first four channels of DDA0 in station ARA03. On the Y axis, the corrected data is plotted with its error. On the X axis the calibrated time is plotted relative to* t_B*. One can see the decrease of regularity and preciseness for channel 2 and channel 3.*

The calibration of the first four channels of a given chip are shown in Figure 5.7 and can be taken as examples for all data channels in the ARA stations.
One can see that the first two channels show a fairly stable width of the used delay elements with a solid relation between even and odd samples. Higher channels 2 and 3 have a more rough behavior. Especially the even samples seem to differ strongly amongst themselves according to the calibration. This seems to be a virtual difference, caused by the behavior described above, which we do not entirely understand. For those channels, the even samples will be excluded in further analysis. This means, the sampling frequency for those channels will be reduced to about 1.6 GS/s, which is acceptable because input signals are limited by low-pass filters to a maximum frequency of 850 MHz. Aliasing will thus be a small effect.

5.2.3 Voltage calibration

The voltage calibration is executed on timing calibrated waveforms. While the timing calibration can be performed rather uninfluenced by the correct calibration of the ADC-to-voltage conversion gain, the voltage calibration, as it is performed, strongly depends on the timing calibration.

A common method for the voltage calibration of a digitizer chip is to inject a large number of input sine waves and to histogram the ADC counts for each sample to be calibrated. The distribution of such points will peak at the maxima of the sine wave, since its derivative is minimal at these points. The resulting peak positions can be used as the corresponding ADC count to the input peak to peak voltage. The advantage of such a method is to be independent of a prior timing calibration. A disadvantage on the other hand is the high statistics which are needed to produce properly peaked histograms with a reasonable error on the peaks. Especially for the IRS2 chip this method is very time consuming, since 32768 individual samples need to be calibrated in that way for each channel. The statistical method becomes even more difficult, when the ADC-to-voltage conversion is not perfectly linear and the peaks are washed out by non linear contributions, as it is observed in the case of the IRS2. The available calibration data do indeed not contain enough statistics to properly apply the above described method and the chip has to be calibrated in a different way for the reasons given above. An alternative technique, tested in the scope of this work, is a correlation-based method which allows for a simultaneous calibration of timing and conversion gain [97]. In this method timing and voltage can in principle be calibrated independently and the needed statistics are limited. However, this method was found to be not applicable due to strong non-linearities in the conversion gain.

The currently applied voltage calibration is based on timing calibrated sine waves of different amplitudes. These waves are fit with a similar function as used for the timing calibration

$$f_i(t) = A_{in} \cdot \sin(2\pi k_i t - \phi_i). \tag{5.6}$$

In this case the amplitude A_{in} is fixed to the given input amplitude, whereas frequency and phase shift are left free. Whenever the relation

$$\left| \frac{df}{dt}(x_i) \right| < 0.45 \cdot A_{in} \tag{5.7}$$

at the time x_i of a sample is true, the ADC count of the sample and the amplitude of the fit at that point are recorded. By this, the error caused by an incorrect timing calibration is minimized, while a data set of reasonable size can be taken for each of the 32768 samples on each channel. The data is separated in bins of 30 mV in fit voltage and averaged (Figure 5.8). It becomes evident that the resulting data is not symmetric around the used zero level, which as explained above is actually at a bias of 1.25 V. In principle there is no reason to expect the chip to show a non symmetric behavior around the bias voltage, as long as saturation can be excluded. Hence, this feature is most likely to be induced by the input chain on a DDA board. In the case of ARA the calibration is used to correct for this asymmetry. In addition to the asymmetry it becomes evident that the ADC-to-voltage conversion is not linear and needs some higher order addition to describe the data properly. This was expected from the features observed during timing calibration.

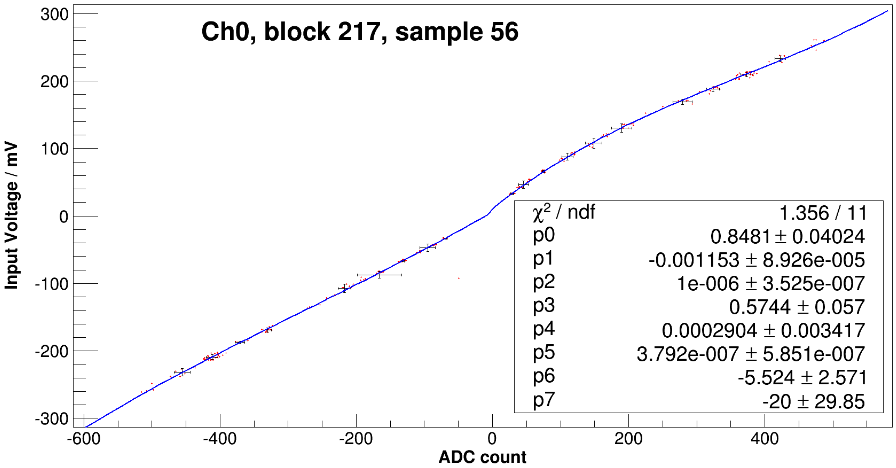


Figure 5.8: *The data collected to calculate the ADC-to-voltage conversion for one storage sample. The* ***red*** *dots are the total collected data,* ***black*** *points show the average over a range of* 30 mV *and the* ***blue*** *line represents the broken 3rd order polynomial fit to the average. Parameters p6 and p7 hold the respective shift in y direction. Parameters p0 to p5 hold the higher order constants of a 3rd order polynomial for the positive part (p0, p1, p2) and the negative part (p3, p4, p5).*

To account for the asymmetry and the non linearity, the collected data will be fit separately for negative and positive values with a 3rd order polynomial. For each sample the

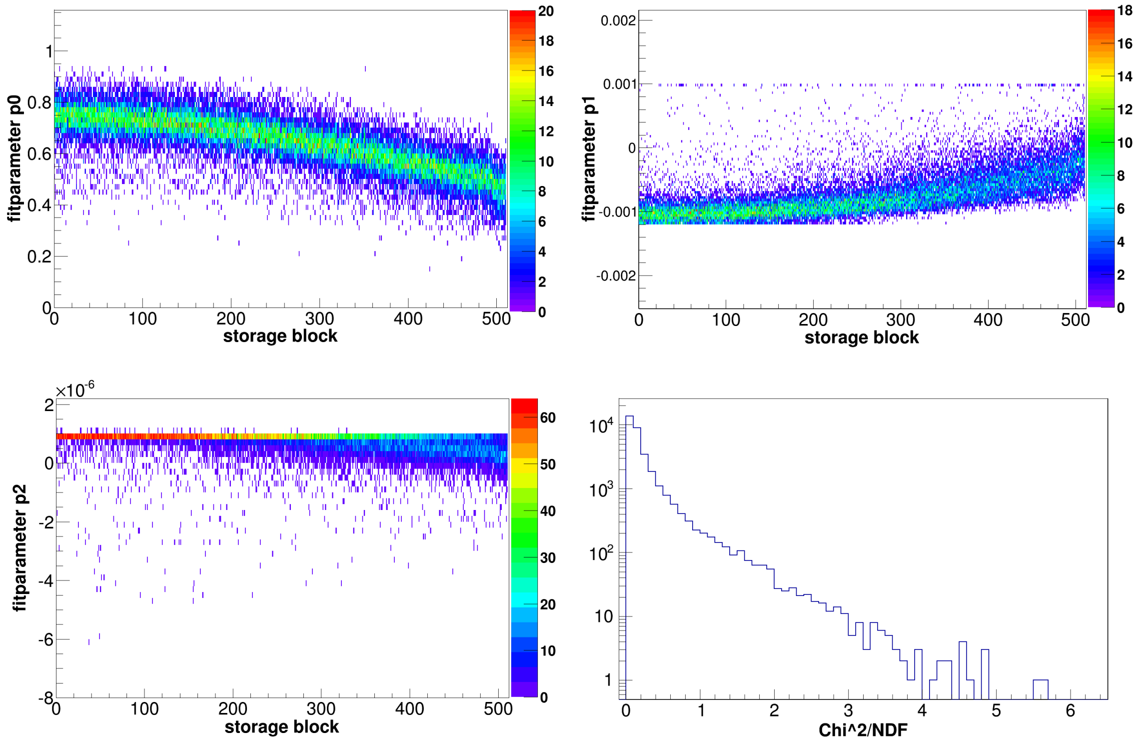
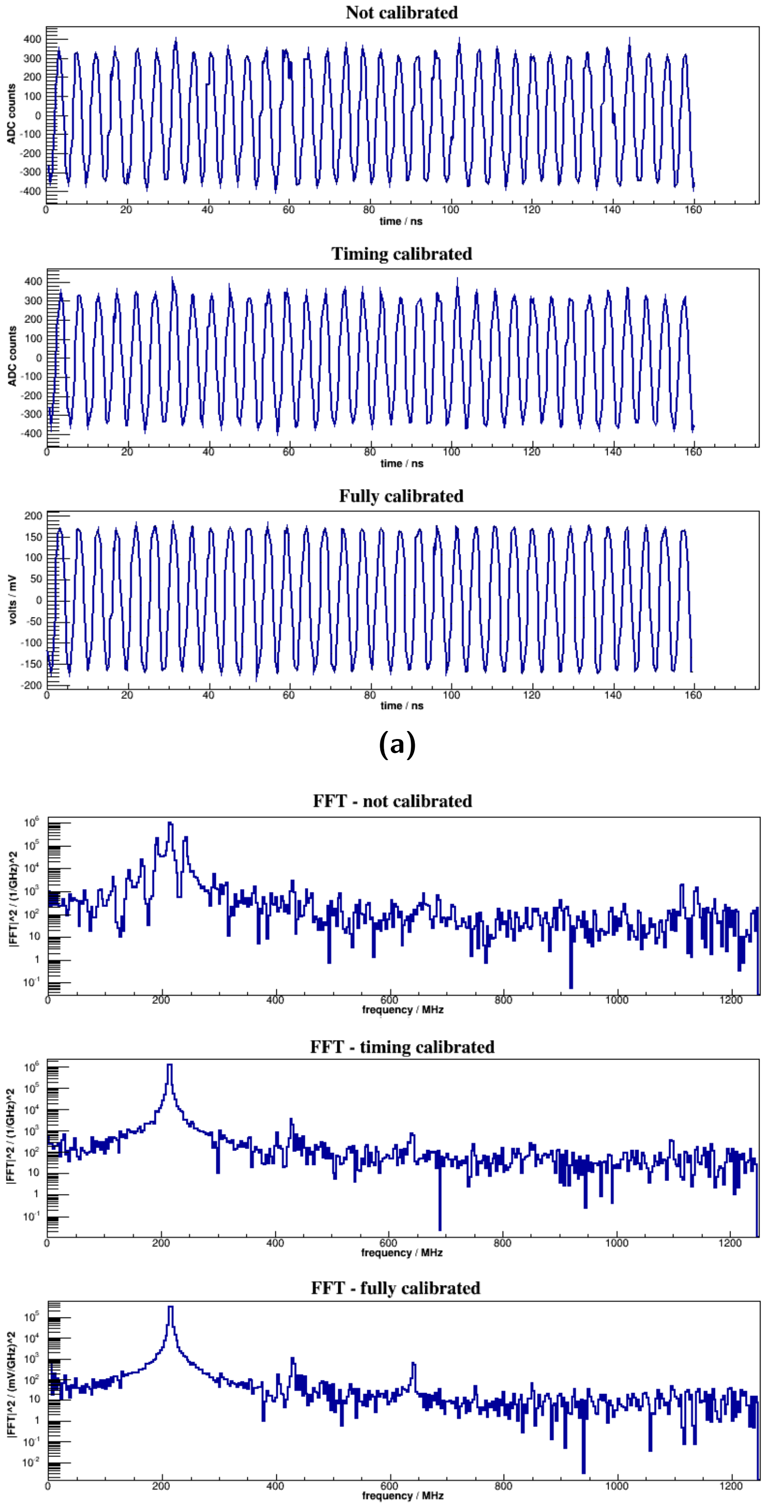


Figure 5.9: *The fit results for all* 32768 *samples on channel 2 in ARA03, plotted as a function of the block number of each sample. The shown parameters are the 1st (p0), 2nd (p1) and 3rd (p2) order parameters of the positive part of the fit. Second and third order parameters are limited, to enhance the fit stability.*

fit is performed up to 20 times with randomly changed seeds, to obtain a $\chi^2/NDF < 2$. Such a fit can be seen in Figure 5.8. The fit results for the positive part of all storage samples in one channel are shown in Figure 5.9 versus block number. The clear dependence on the block number corresponds to the physical position of a certain storage block between the digitization unit and the SCA and is most likely a transfer effect. The plots nicely show that transferred sample charges are changed differently, depending on the path they have to cross inside the chip. The parameters for the negative part of the fit behave similarly. Stringent limits are applied to high order parameters, to enhance the robustness of the fit and to avoid non physical values. The limits are asymmetric to account for possible saturation effects at the highest and lowest voltages respectively. The χ^2 values shown in Figure 5.9 are an indicator for the fit quality and are not to be interpreted statistically because of the fit repetition. It has been observed that samples at the same position in a block and in neighbouring blocks normally show a very similar behavior in voltage conversion gain. Therefore, calibration for samples with a χ^2 value bigger than 1 will be discarded and the conversion factor of the same sample number in a neighboring block, provided it has a better χ^2 value, will be used. In this way, a relatively stable voltage calibration can be achieved for all 40 channels in the two ARA stations.



(a)

(b)

Figure 5.10: *The calibration steps presented on channel 0 on DDA2, in* ***(a)*** *waveforms and* ***(b)*** *the corresponding FFTs. A clear improvement can be observed for the timing calibration. Contribution of voltage calibration is not very pronounce, but visible in the waveforms.*

The improvement of data quality during different steps of the calibration can be seen in the example waveforms and FFTs in Figure 5.10. Especially the FFT spectrum approximates the spectrum of a perfect sine wave better with every step. The visible first and second order harmonics are likely to be caused by the remaining errors on non-linear components in the sample timing and voltage.

A slight offset is caused by the first block of data in the waveform, i.e. the first 20 ns. Due to a feature inside the ARA DAQ, the first block of an event can be corrupted, which is why in later analysis it will be cut from the waveform.

5.2.4 The frequency response of the IRS2 chip

To determine the frequency response of the IRS2 chip, waveforms of different frequency are produced and fed into the system to compare the spectral response. A power spectrum is produced with an FFT as shown in Figure 5.11. The square root of the peak maximum is proportional to the voltage response amplitude at the given frequency. To minimize influences from noise the sum of the five bin contents around the maximum are taken as the actual response factor A_i, where i denotes the frequency. A comparison factor for different frequencies can be calculated in the following way:

$$R_i = \sum_{j=i-2}^{i+2} \frac{\sqrt{A_j}}{V_{i,in}}. \tag{5.8}$$

The unit of R_i is GHz^{-1}.

As a reference, the measurement at 214 MHz is taken since in that frequency region no signal loss is expected [98]. The final response G_i is calculated for each frequency as

$$G_i = 20 \cdot lg(\frac{R_i}{R_{214}}). \tag{5.9}$$

The frequency response for one example chip is shown in Figure 5.12. Unfortunately there are uncertainties about the input voltage due to missing information in the calibration data. The shown frequency response is thus not to be taken as a valid result. This section is used to present a possible approach to calculate the frequency response.

5.2.5 The temperature dependence of the sample timing

The width of the delay elements in the SCA of the IRS2 can be systematically affected by temperature changes [95]. This dependency can be measured using a reference clock which is continuously running on the highest channel of each chip. While the input clock is a temperature stabilized oscillator, the recorded waveform can differ depending on the sample timing. Such a clock waveform can be seen in Figure 5.13.

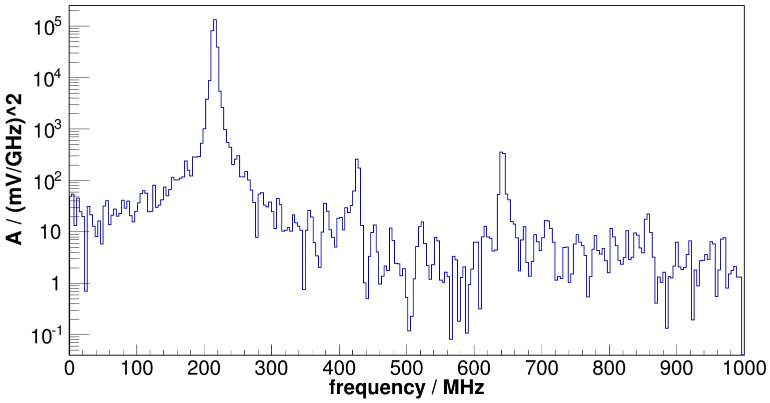


Figure 5.11: *Example spectrum for the determination of the frequency response. The response power is the sum of 5 bins around the main peak.*

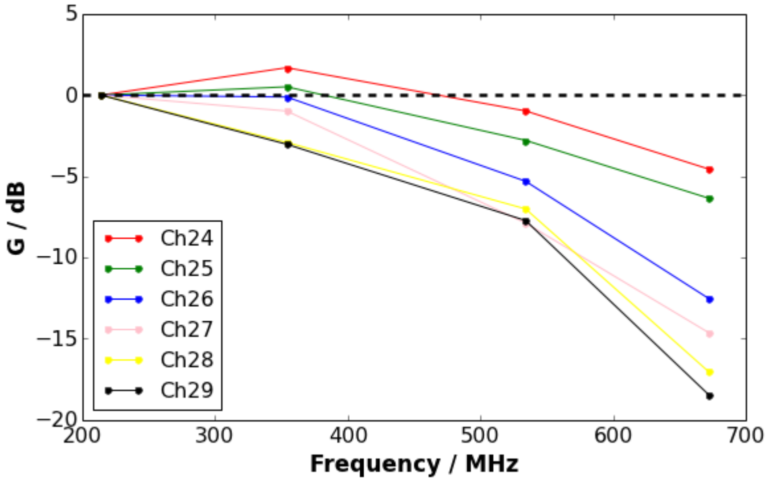


Figure 5.12: *The frequency response G calculated for one IRS2 chip in the ARA02 electronics. The response at* 214 MHz *is used as the reference value.*

From the given waveform the mean is calculated and by interpolation the crossing time of such mean value from different edges in the waveform is calculated. In Figure 5.14 histograms of the nominal crossing time difference divided by the crossing time difference between two rising edges are shown. The mean value $corr_T$ of these histograms can be used as a correction factor for the sampling speed as:

$$sp_{real} = \frac{sp_{ideal}}{corr_T}. \tag{5.10}$$

The edges are chosen to be as far apart as possible within the window of one sampling round of 40 ns, to minimize the influence of errors on the correction factor.

The above described procedure is performed for selected data runs, taken at different temperatures to investigate the behavior of the factor $corr_T$. From the results, shown in Figure 5.15, no clear temperature dependence can be deduced. The observed change of the sample timing on the IRS2 chip is less than 1% over a range of more than 20 °C.

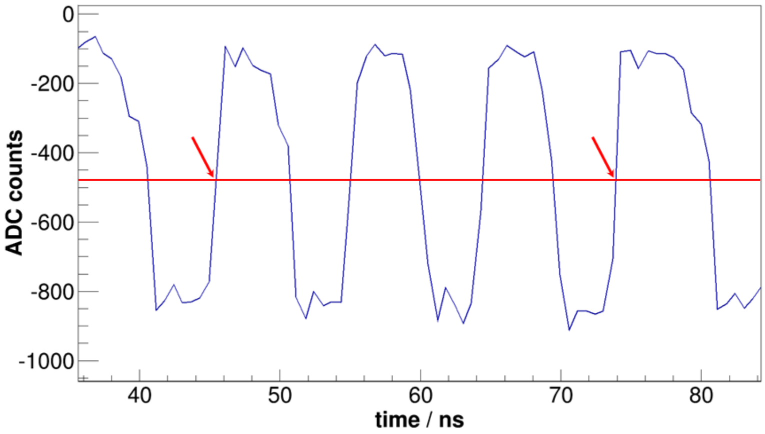


Figure 5.13: *The waveform of the* 100 MHz *reference clock as sampled on channel 7 on each IRS2. The red line indicates the mean value and the arrows indicate the crossings which will be used for sampling speed determination.*

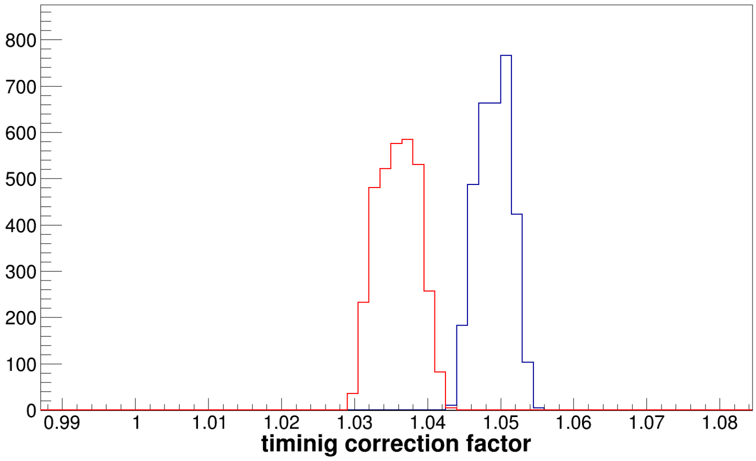
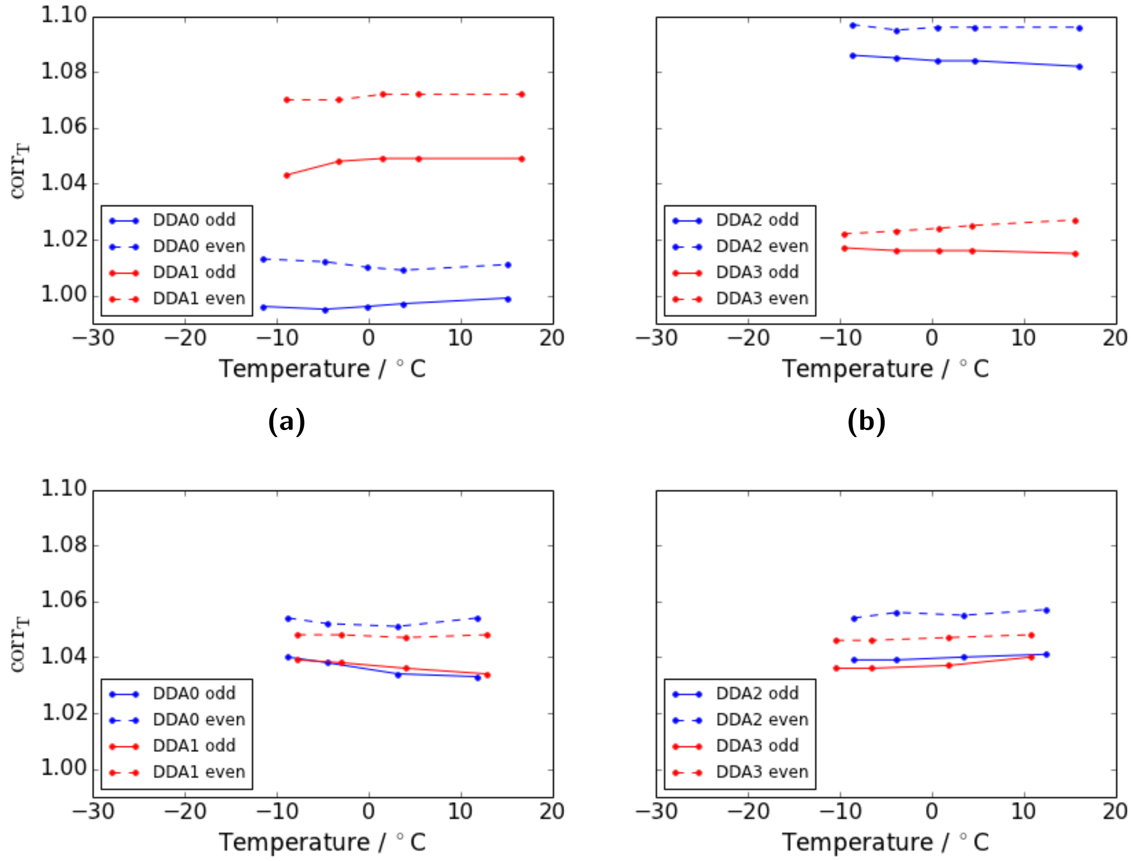


Figure 5.14: *The nominal time difference of* 30 ns *divided by the time differences between two rising edges as indicated in Figure 5.13. Values bigger than* 1 *indicate that the time difference was below nominal. Thus, the real sampling speed was lower than ideally expected.*

This is negligible compared to the errors of the timing calibration. The above described procedure is used in the same way to calibrate changes in the V_{adj} value relative to the settings during the calibration runs.



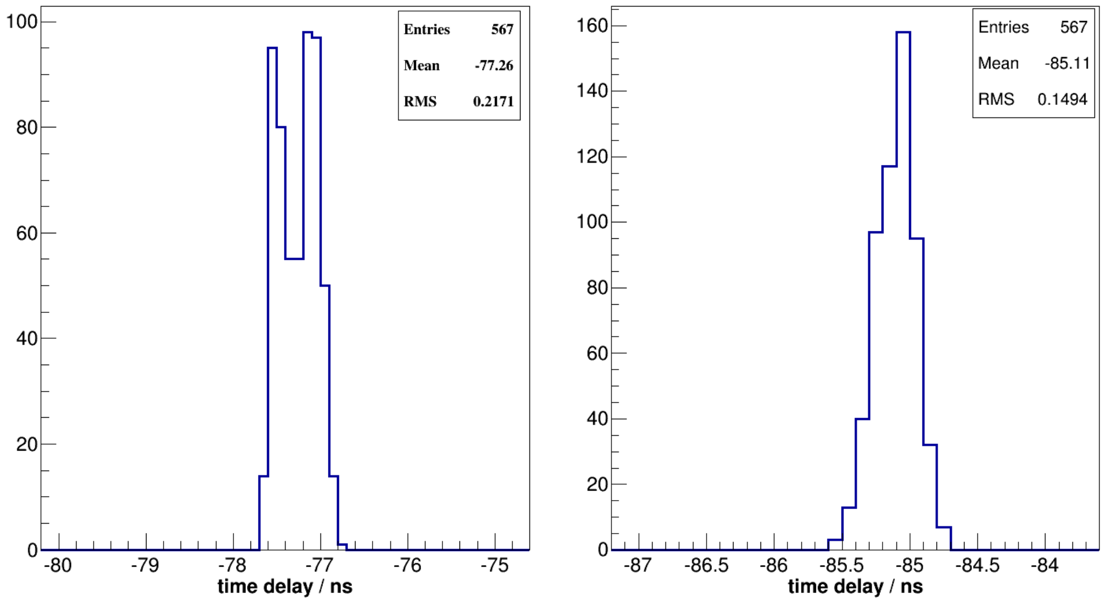
(a) **(b)**

(c) **(d)**

Figure 5.15: *The temperature correction factor corr$_T$ for all DDA boards, holding the IRS2 chips, in* **(a,b)** *ARA02 and* **(c,d)** *ARA03.*

5.2.6 Calibration checks

The timing and voltage calibration is important to achieve a good reconstruction quality and to obtain correct FFT spectra. The FFT spectra were checked during the calibration process, as displayed in Figure 5.10. The influence of the calibration on the reconstruction will be presented in this section as a quality check of the calibration. For this check, recorded signals sent by the calibration pulsers from a distance of about 40 m to the station core are used. The maximum correlation time between two waveforms is calculated as it will be used in the later described reconstruction. Details of this timing determination are given in Section 8.4.1. The precision of the correlation time depends on the timing calibration as well as on the voltage calibration of the two correlated waveforms. Hence, it can be used to determine the quality of the data. Distributions of the correlation time between two channels are shown in Figure 5.16 without V_{adj} correction and in Figure 5.17 with V_{adj} correction. It can be noted that the V_{adj} correction and the timing calibration provide a strong improvement to the standard ARA calibration. The voltage calibration instead appears to be not very important in this check.



(a) **(b)**

Figure 5.16: *The maximum correlation time for calibration pulser signals on two channels of station ARA02, for* **(a)** *the previous standard ARA calibration,* **(b)** *the here described calibration. The correction for changes in the sampling speed settings via V_{adj} is not included.*

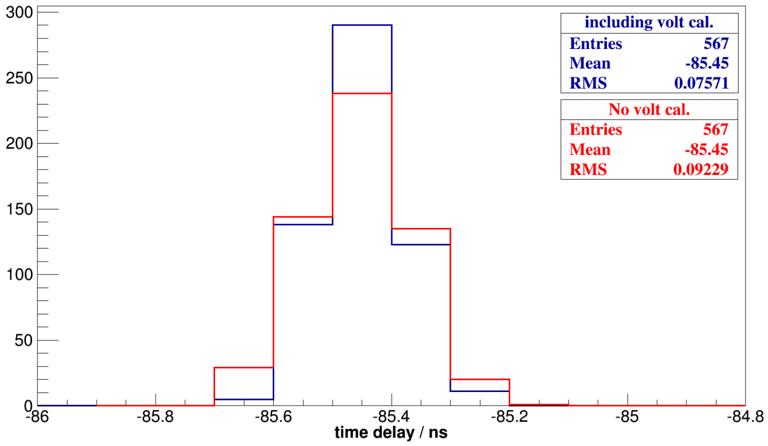


Figure 5.17: *The maximum correlation time for calibration pulser signals on two channels of station ARA02 including all corrections determined during the calibration. For the red distribution, no voltage calibration has been applied.*

The correlation is checked systematically for different calibration pulsers and various channel combinations. The two important quality parameters to measure are here the stability of such a correlation and the actual timing precision. The stability is determined by counting the fraction of correlations which can clearly be separated from the main peak in timing distribution. Normally the stability drops for either very weak waveforms with a low signal to noise ratio or for very strong waveforms which saturate the readout electronics. A more detailed examination of the stability of the timing correlation will be given in Section 8.4.1. The precision of good correlations is defined as the RMS of the main peak. The results for the two calibrated stations are shown in Table 5.1 and Table 5.2. One can see that the timing precision for the main peak is distributed in a range between 48 ps and 158 ps for the used data sample. This is a good base for proper reconstruction precision, as will become apparent in the following chapters. Systematic errors on the timing in the full signal chain, including the IRS2 chip, and their reduction will be discussed in Chapter 6.

Table 5.1: *Correlation results for ARA02.*

Ch #1	Name	Ch #2	Name	Failure in %	Main peak RMS / ns
0	D1BV	8	D2BV	0	0.048
0	D1BV	16	D3BV	0	0.083
0	D1BV	24	D4BV	0	0.056
8	D2BV	16	D3BV	0	0.064
8	D2BV	24	D4BV	0	0.059
16	D3BV	24	D4BV	0	0.065
1	D1TV	9	D2TV	0	0.063
1	D1TV	17	D3TV	0	0.141
1	D1TV	25	D4TV	0	0.056
9	D2TV	17	D3TV	0	0.139
9	D2TV	25	D4TV	0	0.073
17	D3TV	25	D4TV	0	0.150
2	D1BH	10	D2BH	0	0.069
2	D1BH	18	D3BH	0	0.152
2	D1BH	26	D4BH	1.1	0.089
10	D2BH	18	D3BH	0	0.158
10	D2BH	26	D4BH	5.3	0.107
18	D3BH	26	D4BH	1.9	0.144
3	D1TH	11	D2TH	0	0.071
3	D1TH	19	D3TH	0	0.122
3	D1TH	27	D4TH	0	0.075
11	D2TH	19	D3TH	0	0.110
11	D2TH	27	D4TH	0	0.086
19	D3TH	27	D4TH	0	0.122

Table 5.2: *Correlation results for ARA03.*

Ch #1	Name	Ch #2	Name	Failure in %	Main peak RMS / ns
0	D1BV	8	D2BV	2.9	0.118
0	D1BV	16	D3BV	0	0.058
0	D1BV	24	D4BV	0	0.053
8	D2BV	16	D3BV	10.1	0.115
8	D2BV	24	D4BV	11.0	0.099
16	D3BV	24	D4BV	0	0.056
1	D1TV	9	D2TV	0	0.049
1	D1TV	17	D3TV	0	0.055
1	D1TV	25	D4TV	0	0.061
9	D2TV	17	D3TV	0	0.052
9	D2TV	25	D4TV	0	0.081
17	D3TV	25	D4TV	0	0.065
2	D1BH	10	D2BH	2.2	0.103
2	D1BH	18	D3BH	3.0	0.094
2	D1BH	26	D4BH	15.6	0.110
10	D2BH	18	D3BH	0	0.098
10	D2BH	26	D4BH	0.9	0.100
18	D3BH	26	D4BH	2.0	0.088
3	D1TH	11	D2TH	0	0.073
3	D1TH	19	D3TH	0.2	0.106
3	D1TH	27	D4TH	0	0.072
11	D2TH	19	D3TH	0	0.125
11	D2TH	27	D4TH	0	0.108
19	D3TH	27	D4TH	0.4	0.150

6

The calibration of the ARA station geometry

To allow for a proper vertex reconstruction, which is important for neutrino identification and later astrophysical research, we need to know the relative timing between antennas and their geometrical positions very precisely. Given the small side length of roughly 20 m of the cubical ARA stations we need a relative timing precision of about 2 ns or a knowledge of the antenna position within 30 cm, to achieve an angular resolution of roughly 1°. To allow for a good stability of used reconstruction algorithms, an even better knowledge of the station geometry is required. Furthermore, it can be deduced from calculations shown in Section 3.1.3 that we need to know the distance to the vertex of a neutrino induced cascade to calculate the energy of the primary particle. This distance can be derived using the curvature of the incoming wavefront. Especially for far vertices this curvature is very faint. On the baseline of 20 m, timing precision has to be on the order of 100 ps to obtain a reasonable radial precision. From Section 5.2.6 we know that this precision can in principle be achieved with the used sampling devices, if systematic errors on the antenna positions and cable delays are small.

In this chapter, a first attempt to calibrate the geometry and cable delays between antenna strings of the ARA stations is described. First, the used coordinate system and the expected uncertainties about the station setup are discussed. Then, a method for calibration is explained and results are presented including cross checks and an outlook on reconstruction precision.

6.1 ARA coordinates

For the understanding of plots with positions and direction, a short explanation of the definitions used should be presented. These definitions are based on the description of

ARA coordinates in [99].
All positions in this thesis are given in Cartesian (x, y, z)-coordinates relative to the ARA station position at the surface, or in polar (ϕ, θ, r)-coordinates relative to the center of the regarded ARA station. The used directions are shown in Figure 6.1. The station center is defined to be at a depth of $z = -180$ m. The z-axis points upwards, perpendicular to the earth surface at the location of the station. The ice at South Pole is drifting away from the mountain to the coast by several meters each year. The x-axis has been chosen tangential to the surface and in direction of the approximated ice flow at the station. The angle θ denotes the zenith angle of an event, relative to the positive z-axis, where 0° corresponds to positions vertically above the station and 90° to the horizon. The azimuth angle is taken relative to the positive x-axis. In most cases events are regarded separately for the two analyzed stations and the above described coordinates are used. For the analysis of coincident events between the stations, the ARA global coordinate system is used with its center at the nominal center for the full ARA37 detector. It should be noted that the coordinate systems of single stations are slightly rotated against each other and the global ARA coordinates. This needs to be accounted for in coordinate transformations.

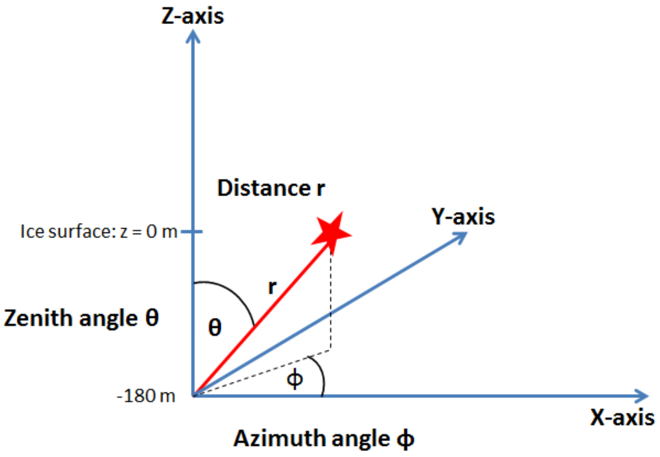


Figure 6.1: *The used coordinates in this analysis. All values are relative to the ARA station center.*

6.2 Uncertainties on the ARA station geometry and cable delays

The measurement antennas of one station are hanging on four strings lowered into 200 m deep holes in the ice. The holes are drilled such that they are vertical within some decimeters on the full length of 200 m. The antennas are hanging above each other at the bottom 20 m of a string. Therefore, the difference in XY-position between antennas from the same string can be assumed to be less than 5 cm. Also their relative depth is well known due to the fixed cable length in between. Accounting for stretching due to temperature effects, an error of 10 cm can be assumed. The global XY-position of a string is determined via a GPS survey. The precision of such survey is estimated to be 20 cm. The global depth of the antennas in the hole is very difficult to monitor precisely from the ice surface. This parameter is assumed to be correct within 3 m.

The cable delays between antennas on one string have been measured in the laboratory. Temperature changes are generally very small which is why we can assume the relative cable delay between antennas on one string to be precise down to 50 ps. In this estimate we included all electronic devices in the signal chain and assume them to behave equal in timing. This might be slightly incorrect but there is no data to determine their delay and we have no hint to believe that it would differ strongly. Measurements of the relative delay between two strings in a station are not as precise. The time delays have only been determined for parts of the signal chain. Long cables, many connectors and electronic devices can induce relatively large errors. In total we will assume an error of 3 ns on the relative string delays.

6.3 The fit of the geometry parameters

6.3.1 The general approach

To calibrate the geometry and cable delays of the strings of an ARA station we have timing measurements available from the four in-ice calibration antennas. These are located at a distance of about 40 m from the center of each station at a depth similar to the measurement antennas (Table 4.2 and 4.3). They are deployed on two different strings with similar geometrical precision as the measurement antennas. From the relative arrival times of signal from the four different sources we can construct a system of equations. If the rank is sufficient, its solution can provide us with the correct positions and cable delays for all antennas involved.

In practice one can calculate the arrival time t_i of a signal at an antenna a_i, emitted at time t_k by a source s_k from

$$c^2(t_k - t_i)^2 = (x_k - x_i)^2 + (y_k - y_i)^2 + (z_k - z_i)^2, \tag{6.1}$$

where c is the speed of light in ice at the station depth and x, y and z are the coordinates of the source and antenna respectively. Since we only measure relative arrival time differences between two antennas, we choose one antenna a_{ref} as a reference with time $t = 0$ and express all other times relative to this. If we now subtract two equations for two antennas, we get

$$\begin{aligned} -c^2(dt^2_{k,i,ref} - dt^2_{k,j,ref}) &= x_k \cdot 2(x_i - x_j) + y_k \cdot 2(y_i - y_j) \\ &+ z_k \cdot 2(z_i - z_j) - t_{k,ref} \cdot 2c^2(dt_{k,i,ref} - dt_{k,j,ref}) - r_i^2 + r_j^2. \end{aligned} \tag{6.2}$$

The variable $r_i = \sqrt{x_i^2 + y_i^2 + z_i^2}$ denotes the distance of the antenna i to the center of the coordinate system and $dt_{k,i,j}$ denotes the time difference measured between the antennas i and j for a signal from source k. We can replace the location differences in x, y and z as $x_{ij} = x_i - x_j$. Equation 6.2 can be reordered and results to

$$\begin{aligned} 0 = c^2(dt_{k,i,ref}^2 - dt_{k,j,ref}^2) + x_k \cdot 2x_{ij} + y_k \cdot 2y_{ij} \\ +z_k \cdot 2z_{ij} - t_{k,ref} \cdot 2c^2 dt_{k,i,j} - r_i^2 + r_j^2. \end{aligned} \tag{6.3}$$

This equation can be set up for every antenna pair in the station. Since the signals we get from the 4 sources are either vertically or horizontally polarized, we will have 8 antennas receiving signal from each source. This gives us in principle 112 equations, of which however only 28 are independent. This is the upper limit of free parameters we can determine with the given setup. Still, all 112 equations can be useful to reduce the error of the following fit and to render it more robust. To perform the fit with these equations, one can use available minimizer algorithms to find the minimum χ^2, defined as

$$\begin{aligned} \chi^2 = \sum \big[c^2(dt_{k,i,ref}^2 - dt_{k,j,ref}^2) + x_k \cdot 2x_{ij} + y_k \cdot 2y_{ij} \\ +z_k \cdot 2z_{ij} - t_{k,ref} \cdot 2c^2 dt_{k,i,j} - r_i^2 + r_j^2 \big] . \end{aligned} \tag{6.4}$$

Equation 6.1 is only valid when assuming a constant speed of light between the signal source and the receiving antenna. As explained in Section 4.1, measurements show that this assumption is not valid for South Pole ice. The speed of light changes with depth and radio ray paths are curved due to the ray-tracing effect. For the in-ice calibration pulsers, differences of up to a nanosecond compared to the assumption of a constant speed can be expected in relative timing between two antennas at different depth. Such a difference is small but can be significant for the goal of achieving a precision of some 100 ps. To minimize systematic uncertainties caused by the changing speed of light, we calculate a correction factor for the measured timing.
As a model for the dependence of the speed of light on the depth z we use

$$c(z) = \frac{c_0}{n(z)} = \frac{c_0}{1.78 - 0.43 \cdot e^{0.016 \cdot z}}, \tag{6.5}$$

where c_0 denotes the speed of light in vacuum. For a source k and a receiver i at different depth, one can approximate the ray-traced propagation time of the signal fairly well by integrating the time on a straight path between them. The result is

$$t_{app,k,i} = \frac{1.78 * (z_i - z_k) - \frac{0.43}{0.016} * (e^{0.016 z_i} - e^{0.016 z_k})}{c_0 * \frac{z_i - z_k}{D}}, \tag{6.6}$$

with D the distance between signal source and antenna. We define the geometrical time difference under the assumption of constant speed of light to

$$t_{geo,k,i} = \frac{D}{c(-180\,\mathrm{m})}. \tag{6.7}$$

This is the time to be used in Equation 6.4. The measured time differences are assumed to be ray-traced times and thus correspond to $t_{app,k,i}$. The difference between geometrical time and ray-traced time can then be calculated to $t_{corr,k,i} = t_{app,k,i} - t_{geo,k,i}$ and the measured times can be corrected to

$$t_{k,i} = t_{meas,ki,i} - t_{corr,k,i}. \tag{6.8}$$

The time difference, to be used in the fit is then $t_{k,i}$. The values $t_{app,k,i}$ and $t_{geo,k,i}$ are calculated for the nominal geometry of the fit station. The fit is executed several hundred times with continuously corrected timings, based on the resulting geometry. After the first iteration, the refinement of the timing correction with the new geometry is very small and is buried in the errors of the timing measurement. It thus acts more like a slightly randomized seed in further iterations. However, its application on the original geometry has been found to add non negligible improvements to the fit results.

6.3.2 Initial assumptions

The number of free parameters we have to fit is in general 4 per involved antenna, which means 80 in total. Since our system of equations does not provide us with enough constraints to determine all these free parameters we have to make a few assumptions based on the error estimates from Section 6.2. These assumtions are:

- Strings are perfectly vertical and the antennas on one string all have the same x and y position.
- The relative depth of antennas on the same string is perfectly known due to the stiffness of the used cables.
- The relative cable delay between antennas on a string is perfectly known while differences in the electronics are assumed to be negligible.

The above described assumptions leave 4 parameters per string to be fit for the measurement strings. For the calibration pulsers we are not interested in the cable delay and thus have only 3 free parameters. In total this gives us 22 free parameters for the full system. To give the station a fixed position in space and a fixed angular orientation we need to define a reference string of the measurement array and a reference pulser string for each station. The coordinates of the reference measurement string are fixed in the calibration fit. To constrain the station in azimuthal rotation, the azimuth angle for the selected reference pulser is fixed, which reduces its free parameters to the distance and the depth relative to the center of the coordinate system. Finally, this leaves us with 17 unknowns, which can be fully determined with a system of 28 equations.
The errors we assume on these free parameters are summarized in Table 6.1.

Table 6.1: *The generally assumed errors and fit limits for the free parameters of the ARA stations. "Ref pulser" denotes the pulser which has been chosen as reference.*

Parameter	Assumed error	Lower limit	Upper limit	unit
string x	0.2	-2	2	m
string y	0.2	-2	2	m
string z	3	-10	10	m
string delay	3	-15	15	ns
pulser2 x	0.2	-3	3	m
pulser2 y	0.2	-3	3	m
pulser2 z	3	-10	10	m
ref pulser z	3	-10	10	m
ref pulser distance	0.4	-40	40	m

The fit limits from Table 6.1 are slightly adapted for the two stations to enhance the stability of the fit. The reference string and pulser are determined by testing all combinations and selecting the best result in the reconstruction cross check, which is described in the following section.

6.3.3 Results and cross checks

The fit distributions of all iterations are shown in Figure 6.2 and 6.3. The final results for the two stations are displayed at the end of this chapter in Table 6.3 and 6.4. It becomes visible from the mean errors that the calibration in the given geometrical configuration works best for the depth correction. It should be noted that from station ARA02 two antennas, D3BV and D3BH, had to be excluded due to a systematic timing offset which is not understandable under the above described assumptions. The suspicion is that there are unknown systematic offsets in the cable delays or parts of the electronics on those antenna channels. The error source could not be isolated with the available data. More detailed measurements and investigations are needed to understand this offset. For the current calibration and further analysis, the two channels are excluded. This has a strong influence on the precision of the geometrical calibration of string D3, which is well visible in Figure 6.2.
One can deduce from the tables that the corrections fit very well the assumed errors on the given parameters, noted in Table 6.1. This gives us confidence that the assumptions are taken in a reasonable manner.

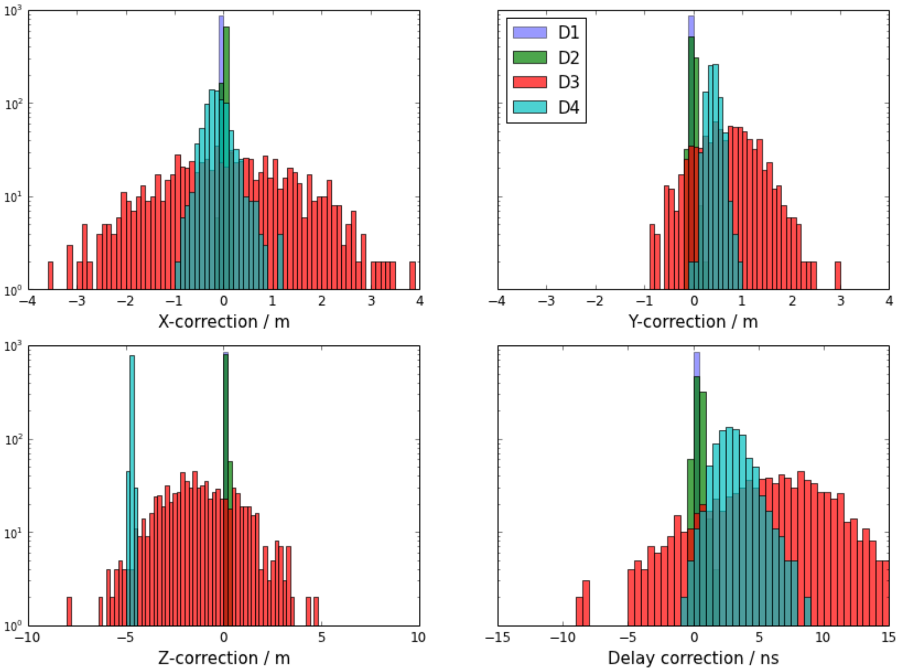


Figure 6.2: *The geometry fit results for the four measurement strings on ARA02. The four panels show the correction in X, Y, Z and cable delay for each string, as denoted on the X-axes. The fixed reference string is D1.*

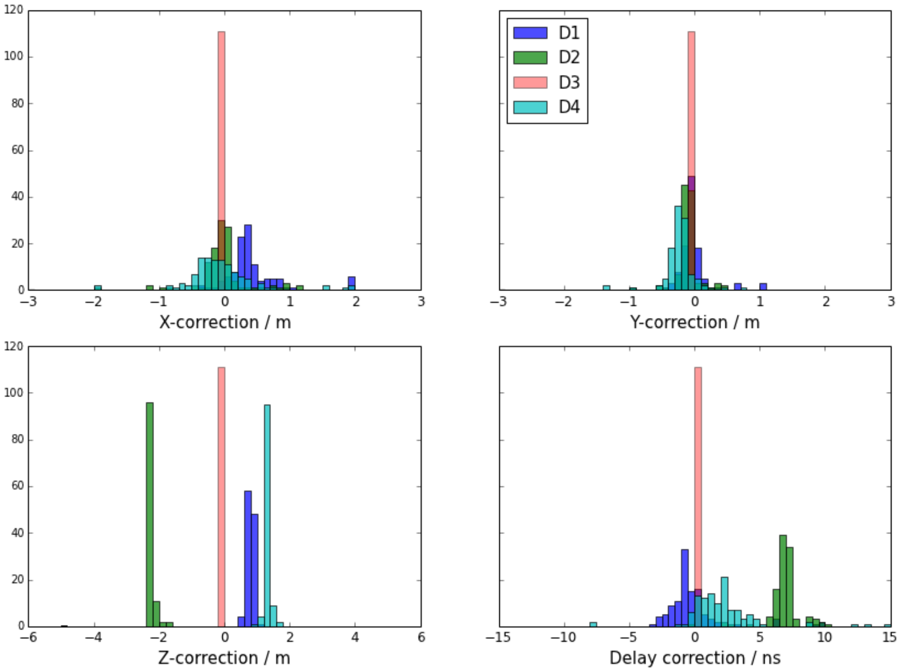
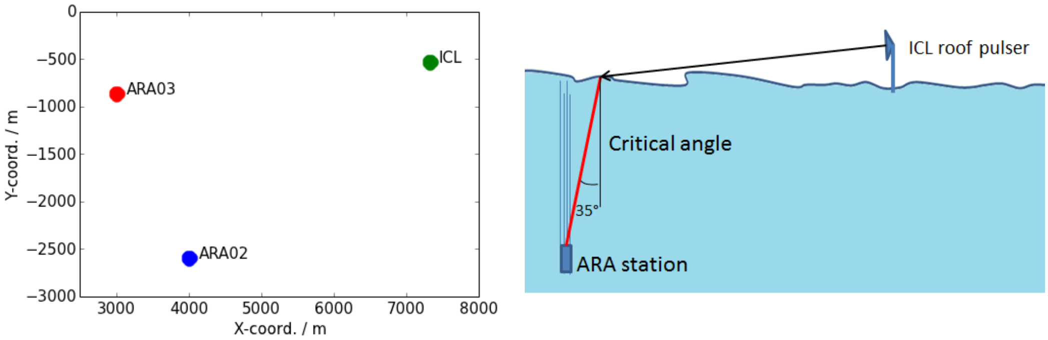


Figure 6.3: *The geometry fit results for the four measurement strings on ARA03. The four panels show the correction in X, Y, Z and cable delay for each string, as denoted on the X-axes. The fixed reference string is D3.*

As explained above, the calibration of geometry and cable delays is most important for the reconstruction of radio vertices around the stations. Therefore, a vertex reconstruction is used as a cross check of the calibration. The used reconstruction method is described in Section 8.4. The cross check we perform is a comparison of the reconstruction of different sources before and after calibration. The regarded parameters are the following:

- **The reconstructed mean position** gives us information about a systematic offset in position and rotation of the station. Such offset is in principle not problematic for neutrino identification. For a signal source with known position, we can use the offsets to correct the station position.
- **The RMS on the reconstruction** of many events allows us to determine the stability of the reconstruction. Since in different cases different channel combinations are used, a small RMS will show us that all channel combinations reconstruct to the same vertex. This implies that the geometry fit worked out well and the station is very self-consistent.
- **The residual of the reconstruction** indicates the discrepancy between the measured timings and the actually reconstructed vertex. For perfect geometry and timing, this residual should approach zero. When calculated for many events, it thus gives us a measure of systematic errors in the intrinsic station geometry. One should note that a small error will be induced by the fact that the reconstruction uses a constant speed of light rather than the more correct ray-traced times.



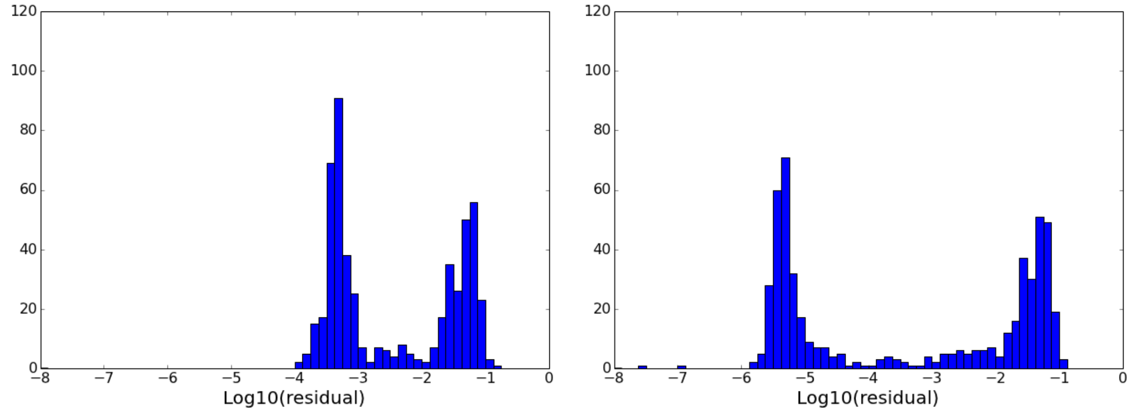
(a) **(b)**

Figure 6.4: *(a) The relative positions of ARA02, ARA03 and the IceCube Laboratory (ICL) in global ARA coordinates. The pulser, mounted on the roof of the ICL is located at a distance of roughly* 4000 m *from both stations. (b) Incoming radio wave fronts from the roof pulser appear refracted into the ice under the critical angle of* 34.7° *at the antennas of the ARA stations.*

Apart from the in-ice pulsers, we have one additional signal source available, which is mounted on the roof of the **I**ce**C**ube **L**aboratory (ICL) at the surface of the ice, in the following denoted as the roof pulser. The position of this source compared to the two stations is shown in Figure 6.4a. This pulser is not used in the geometry fit and can thus be a source for an independent cross check. When deriving the incoming angle of the signal, one has to take into account the air to ice interface and calculate refracted angles using Snell's law. Since the pulser is very far away and mounted close to the ice surface, the signal at the antennas of the ARA stations will be received under the critical angle θ_c (Figure 6.4b). In spite of the changing index of refraction, this can be simply calculated as

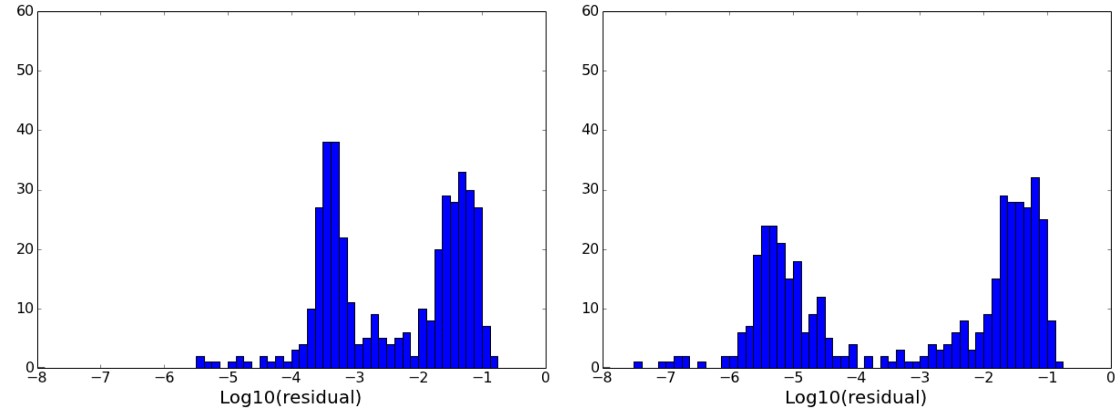
$$\theta_c = \arcsin\left(\frac{1}{n(-180\ \mathrm{m})}\right) = 34.7^\circ, \tag{6.9}$$

with $n(-180\ \mathrm{m})$ the index of refraction in the ice at the average station depth of 180 m.



(a) **(b)**

ARA02

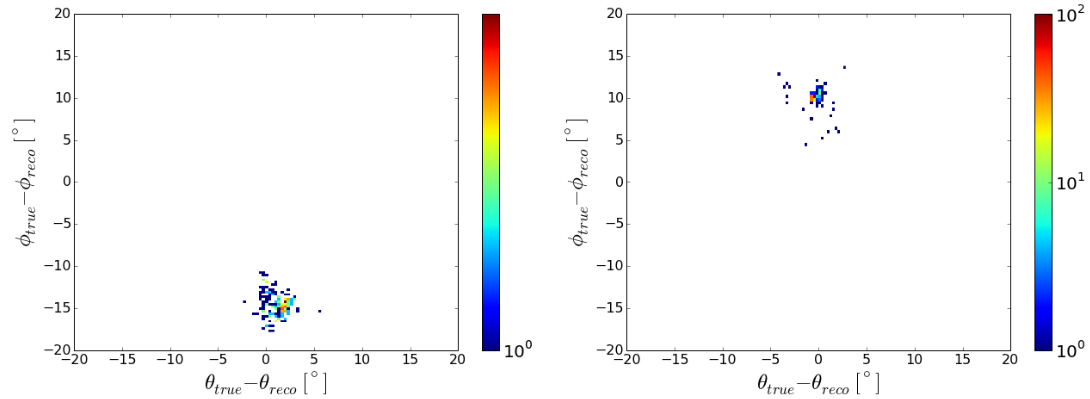


(c) **(d)**

ARA03

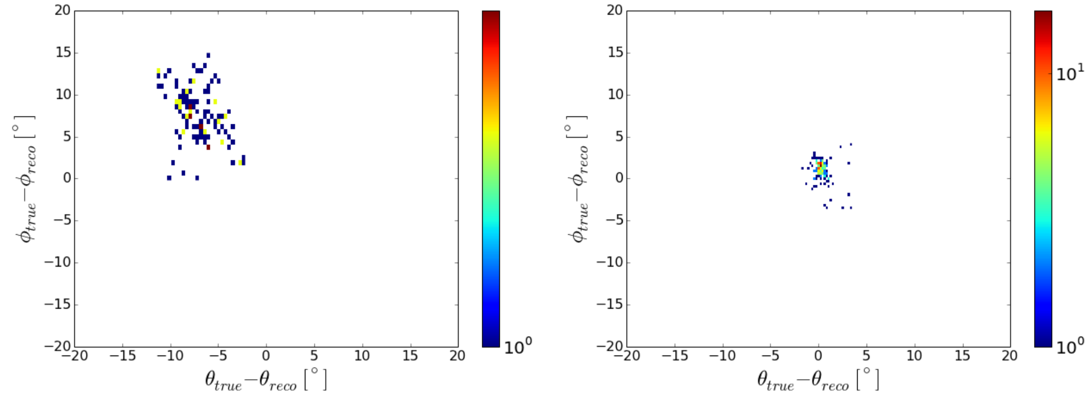
Figure 6.5: *The residuals for reconstructions of the roof pulser with ARA02 and ARA03 from dedicated data runs* ***(a,c)*** *before and* ***(b,d)*** *after station calibration. The relatively high residuals are most likely connected to noise events which have been wrongly selected as impulsive events in this case.*

As an example, the reconstructions of the roof pulser with ARA02 and ARA03 are plotted in Figures 6.5 and 6.6. In these plots a comparison of the reconstruction before and after geometry calibration is shown. In all three parameters, Zenith (θ), Azimuth (ϕ) and the residual, a clear improvement is visible. The parameters, identified as quality indicators for the geometrical calibration are summarized in Table 6.2 for the roof pulser reconstruction from both stations. It is visible that the reconstruction improves not only in RMS and residual, but also in the systematic offset of the mean value. For the azimuth angle this is not of great important since an offset just implies a systematic rotation of the station. The reconstruction of the correct zenith angle to within less than 0.5° shows that the assumption of perfectly vertical strings and a well know delay and depth relation between the antennas on one string is a valid approach.



(a) **(b)**

ARA02



(c) **(d)**

ARA03

Figure 6.6: *The reconstruction of the zenith angle θ and azimuth angle ϕ for the roof pulser with ARA02 and ARA03 from a dedicated data run* ***(a,c)*** *before and* ***(b,d)*** *after station calibration. The data are plotted in comparison to the true position in ϕ and the critical angle of 34.7° in θ. Reconstruction quality cuts are chosen to be loose in this case. Before correction 161 (ARA02) and 119 (ARA03) events pass these cuts, after correction 267 (ARA02) and 193 (ARA03) events pass.*

Table 6.2: *The roof pulser reconstruction results for the two ARA stations, before and after geometrical calibration.*

Station	Parameter	Nominal value	Reconstruction mean		Reconstruction RMS	
			no cal.	cal.	no cal.	cal.
ARA02	Zenith θ	34.7°	35.8°	35.1°	1.1°	0.6°
	Azimuth ϕ	265.2°	279.9°	255.2°	1.3°	0.8°
	residual	–	$4 \cdot 10^{-4}$	$2 \cdot 10^{-5}$	–	–
ARA03	Zenith θ	34.7°	41.8°	34.4°	2.0°	0.6°
	Azimuth ϕ	237.7°	245.3°	238.7°	3.1°	1.2°
	residual	–	$3 \cdot 10^{-4}$	$2 \cdot 10^{-5}$	–	–

6.3.4 Errors on the geometrical calibration

The main contributions to the final systematic error in this geometrical calibration are the error on the average correlation timing used and an error in the assumptions taken on fixed geometrical parameters. The error on the correlation timing can be caused by saturation, noise and an error on the timing calibration of the digitizer chip. The differences between strings will be taken care of by the cable delay fits and should not cause further problems. Errors between timing or position of antennas on the same string have an influence on how well the station will be intrinsically consistent. As explained, this consistency will be reflected in the residual and, assuming different channel selections for the reconstruction of different events, in the RMS of a reconstruction. In conclusion, the systematic error on the geometrical calibration is thus expressed in the reconstruction of radio sources around the station. Such a source has been shown in Figures 6.6 and 6.5 and more will be presented in the following chapters. Furthermore, systematic errors on the antennas have an influence on the final rotation of the station. As has been mentioned, this is not problematic and can be accounted for, if needed.

Table 6.3: *The geometry and cable delay corrections as fit results for station ARA02. The reference string in this fit is string D1 and the reference pulser is D6.*

Parameter	string	nominal value	correction	correction mean error	unit
x	D1	10.59	0	0	m
	D2	4.82	0.013	0.001	m
	D3	-2.68	0.1	0.05	m
	D4	-7.72	-0.1	0.01	m
y	D1	2.34	0	0	m
	D2	-10.38	-0.01	0.002	m
	D3	8.68	0.7	0.02	m
	D4	-4.47	-0.1	0.003	m
z	D1	-189.50	0	0	m
	D2	-189.54	0.1	0.001	m
	D3	-189.44	-1.2	0.07	m
	D4	-189.56	-4.7	0.002	m
cable delay	D1	0	0	0	ns
	D2	0	0.4	0.01	ns
	D3	0	6.1	0.2	ns
	D4	0	3.1	0.05	ns
x	D5 pulser	37.87	-0.016	-	m
y	D5 pulser	-18.05	0.28	-	m
z	D5 pulser	-192.93	-4.6	-	m
z	D6 pulser	-167.94	-5.35	-	m
distance	D6 pulser	39.55	0.31	-	m

Table 6.4: *The geometry and cable delay corrections as fit results for station ARA03. The reference string in this fit is string D3 and the reference pulser is D6.*

Parameter	string	nominal value	correction	correction mean error	unit
x	D1	4.41	0.4	0.04	m
	D2	10.69	-0.05	0.04	m
	D3	-2.01	0	0	m
	D4	-8.10	-0.05	0.05	m
y	D1	-9.39	-0.02	0.02	m
	D2	3.51	-0.1	0.02	m
	D3	9.41	0	0	m
	D4	-3.71	-0.2	0.02	m
z	D1	-192.45	0.8	0.008	m
	D2	-192.70	-2.2	0.01	m
	D3	-192.67	0	0	m
	D4	-192.60	1.3	0.009	m
cable delay	D1	0	-0.3	0.2	ns
	D2	0	6.8	0.1	ns
	D3	0	0	0	ns
	D4	0	2.1	0.3	ns
x	D5 pulser	38.01	0.23	-	m
y	D5 pulser	-16.40	0.18	-	m
z	D5 pulser	-198.18	4.9	-	m
z	D6 pulser	-188.01	-3.95	-	m
distance	D6 pulser	40.57	0.36	-	m

Part III

Data analysis

7

Simulations

The data analysis developed in this work is partly based on neutrino and noise simulations which are needed to test algorithms and to estimate a final analysis efficiency. From this efficiency the sensitivity of the detector can be calculated and an expected number of detected GZK neutrinos for a given flux can be predicted. The systematic errors on these values depend strongly on the precision of the simulation. In the present chapter first the official ARA simulation is described, which is used in this analysis with small modifications. Furthermore, the methods to calculate effective volumes, effective areas, the detector sensitivity and the expected number of neutrinos in ARA are presented.

7.1 The ARA simulation

The neutrino and noise simulations used in the frame of this thesis are produced with the official ARA simulation code **AraSim**. This simulation tool has been developed within the collaboration over several years. Detailed descriptions of the status of AraSim can be found in [100] and [76]. The main techniques used to simulate neutrino signals on the ARA stations are summarized in the following points:

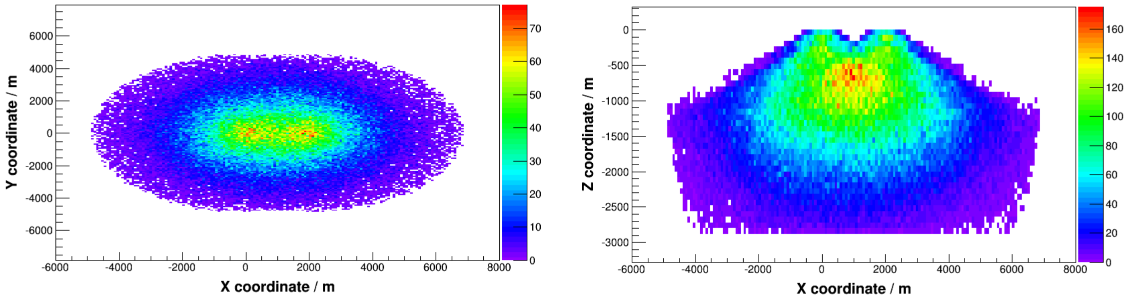
1. Neutrino interactions with an isotropically distributed incoming angle are forced at random places within a defined cylinder around the station. Each such generated event gets a weight ω_i according to the probability of the interaction, which depends on the location and incoming direction of the neutrino. The weight is calculated from the interaction length of neutrinos in the matter passed by the particle to reach the generated interaction point, as derived in [101]. The resulting weight is used for the calculation of effective volumes and sensitivities (Section 7.2).

2. For each generated vertex the interaction type is simulated and the energy transfer to the hadromic or purely electromagnetic cascade is modelled according to [101]. For the energy range of ARA most signal is expected from the electromagnetic component of hadronic cascades. These cascades only hold roughly 20% of the neutrino energy whereas the transfer to purely electromagnetic cascades is much higher [11]. However, they are favorable in terms of the Askaryan signal for ARA since they are nearly unaffected by the LPM effect, which strongly limits the visibility of purely electromagnetic showers [65].
   The Askaryan emission for the given shower is modeled in the frequency domain based on [62]. The Askaryan signal is propagated through the ice to the detector, accounting for attenuation and ray-tracing effects. All possible solutions for ray paths between the signal source and an antenna are taken into account.

3. The antenna responses are modeled separately for Hpol and Vpol antennas using a NEC2 simulation. The signal is propagated through the antenna and the full signal chain. The frequency response of the signal chain is assumed to be the same for each receiver channel and is modeled according to a laboratory measurement. Noise is generated as flat thermal noise at the temperature of the ice around the antennas. It is propagated through the measurement chain alongside with the signal. The resulting sum of signal and noise is transformed into a time domain waveform.

4. To simulate the detector trigger, a model of the diode output is used and convolved with the incoming time domain waveform. The trigger logic is applied as in the ARA stations (Chapter 4.1.2) and the power threshold on the diode output is adapted to achieve the same final trigger rate as the real detector stations.

As detector geometries and antenna configurations, one can currently choose between the geometry of the TestBed (Section 4) and an idealized cubical ARA station geometry.



(a) **(b)**

Figure 7.1: *The vertex distribution of triggering events in the ice around ARA02 (at [2000, 0, -180]) and ARA03 (at [0, 0, -180]).* ***(a)*** *The XY-distribution,* ***(b)*** *the distribution of depth Z versus X.*

For the analysis described in this thesis the simulation has been slightly adapted, to fit better to the real detector setup. Neutrino signals have been simulated for a two station detector with the relative positions and geometrical shape of ARA02 and ARA03. This geometrical shape includes the corrections measured in Chapter 6. To account for saturation effects in the timing correlation for reconstructions, a random perturbation is added to the timing of saturated events according to the model described in Section 8.4.1. We note that this disturbance is an individual adaptation for the analysis of this work, and it does not necessarily hold for other timing algorithms.

The neutrino energy has been generated in 20 discrete energy bins between 10^{16} and $10^{20.75}$ eV, spaced equally by 0.25 on a logarithmic scale. For each bin 10000 events have been generated in a cylinder with adapted radius. The radii are chosen to exceed the furthest possible triggering vertex, to guarantee the capture of all possible events. The vertices of events that triggered at least one of the stations are shown in Figure 7.1. The zenith distribution of all triggering events is shown in Figure 7.2. This distribution shows that the stations have their highest sensitivity for zenith angles greater than 100°, i.e. from slightly below the horizon. Such sensitivity behavior is caused by the shadowing effect from ray-tracing, as explained in Section 4.1. In the depth distribution of Figure 7.1 one can clearly see that regions above a certain incoming angle are not reachable due to this effect. We should remind ourselves that the triggering events are forced interactions in the ice. The weighting factor which accounts for the interaction cross section is not applied in the distributions shown in Figure 7.1 and 7.2.

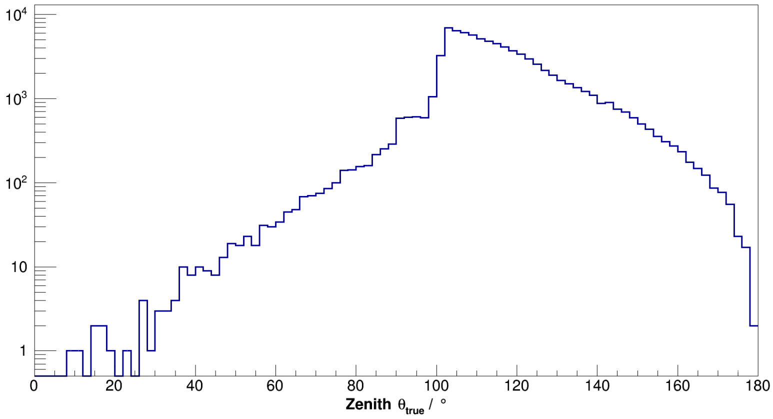


Figure 7.2: *The zenith angle θ_{true} of triggered neutrino vertices relative to the station centers.*

The above described simulation is used in the development of cut algorithms, vertex reconstruction and the calculations of detector sensitivity, unless specified otherwise. In some cases it is adapted to look at neutrino distributions for each station separately to perform a direct comparison to the recorded data. Apart from the signal simulation, a small pure noise sample is generated, which is modeled according to the individual noise frequency spectrum of the ARA03 antenna channels. Details on this noise modeling are presented in Appendix A. Time constraints did not allow to include this model in the signal simulation yet. The noise sample will be used for testing during the development of cut and reconstruction algorithms, described in Chapter 8. Due to the high trigger threshold used in the ARA stations (Section 4.1.2), the generation of triggering thermal noise events requires the generation of a large number of events and thus enormous computing resources. Therefore, it is difficult to generate sufficient noise in a realistic manner within reasonable time to determine cuts and their efficiency on background. The investigation of the cut efficiency is thus performed on a subset of the data recorded in the two stations, which is assumed to contain only background events.

7.2 Effective volume, effective area and neutrino limits

For the parameters, estimated in the following, we generally distinguish between an estimation at trigger level, derived from all triggering events, and an estimation at analysis level, derived from events passing the analysis cuts after background rejection. In this chapter the estimations at trigger level will be presented. Calculations at analysis level are performed in an analogous way and are presented in Section 10.1.

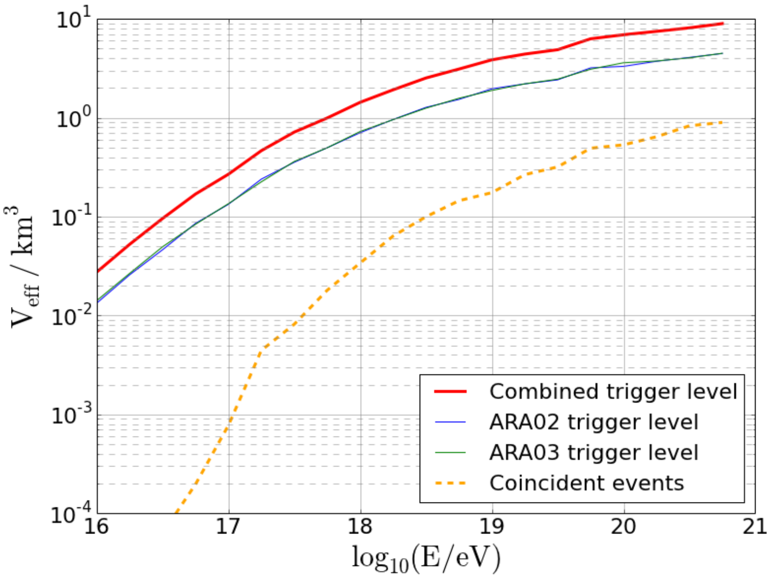


Figure 7.3: *The effective volume at trigger level of the two station simulation combined and separately and for each station, plotted versus neutrino energy. The dashed line shows the contribution of coincident events to the effective volume.*

From the fraction of triggered events and the number of generated events we can derive an effective volume for the currently available two station detector, which allows us to estimate the effective area and its sensitivity. The effective volume V_{eff} is calculated to be

$$V_{\text{eff}} = V_{\text{gen}} \cdot \frac{1}{N_{\text{gen}}} \cdot \sum_i \omega_i, \tag{7.1}$$

with N_{gen} being the number of generated events in a volume V_{gen} and ω_i being the weight of an event i, described in the previous section. The weights are summed over all triggering events. The effective volume of the two ARA stations at trigger level as a function of energy is shown in Figure 7.3. The two stations show a similar effective volume, as is expected. In a two station setup, there is a chance that one single event will trigger both stations. The effective volume of such coincident triggers is represented by the dashed orange line in Figure 7.3. These events amount to about 5% of the total effective volume, which changes with energy. They have to be treated separately. In case of a coincident trigger in both stations only one of them is counted for the effective volume.

In ARA the neutrino effective area A_{eff} of the detector can be calculated from the effective volume V_{eff} to

$$A_{\text{eff}} = \frac{V_{\text{eff}}}{L_{\text{int}}}, \tag{7.2}$$

with the interaction length L_{int}. This length depends on the cross section σ for a neutrino interaction in the regarded matter and the number density of the medium,

$$L_{int} = \frac{1}{\sigma \cdot N_A \cdot \rho}. \tag{7.3}$$

The energy dependent cross section σ is taken from [101]. N_A denotes the Avogadro constant and $\rho = 0.92\ \text{g/cm}^3$ is the density of the ice. A_{eff} as a function of energy for ARA02 and ARA03 is shown in Figure 7.4. This effective area can be used to directly compare neutrino detectors. From the effective area and the live time we derive the neutrino sensitivity. Since ARA02 and ARA03 accumulated different live times during operation (Figure 4.11), the detector sensitivity has to be calculated separately for both stations. The detector sensitivity to an isotropic neutrino flux is given by

$$S(E) = \frac{1}{4\pi \cdot (A_{\text{eff},2} \cdot T_2 + A_{\text{eff},3} \cdot T_3)}, \tag{7.4}$$

with the detector live time T_2 and T_3 for the two stations. The derived sensitivity is typically used to compare the ARA detectors to other neutrino telescopes and neutrino fluxes. It is plotted in Figure 4.17, multiplied by the neutrino energy. In addition, one can calculate the expected number of neutrinos N_ν to

$$N_\nu = 4\pi \cdot \left(T_2 \cdot \int dE A_{\text{eff},2} \frac{dN}{dE} + T_3 \cdot \int dE A_{\text{eff},3} \frac{dN}{dE} \right). \tag{7.5}$$

In this equation, dN/dE denotes the differential neutrino flux, predicted in different ways from various models. Assuming the moderately conservative neutrino flux estimation from [52] the number of neutrinos, expected to trigger the two ARA stations within

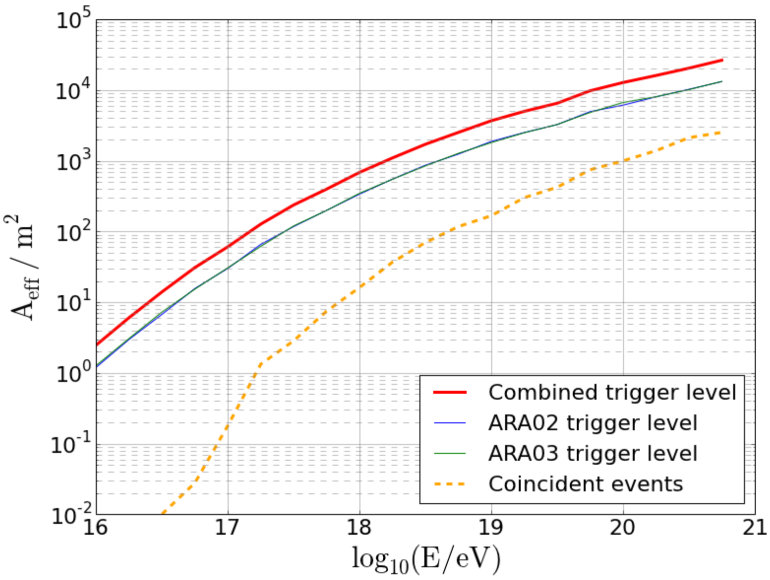


Figure 7.4: *The effective area at trigger level of the two ARA stations combined, separated and for only coincident events, plotted versus neutrino energy.*

roughly 220 days of live time can be calculated to 0.18. This number is most interesting for a discovery instrument like ARA.

8

Algorithms for the ARA data analysis

8.1 The signal to noise ratio

In this thesis, a data analysis will be presented which is based on simulated data and the analysis of neutrino-like events produced by the in-ice calibration pulsers. For comparison between the recorded data and simulation, a useful parameter is the signal to noise ratio of a waveform or an event, defined as

$$SNR = \frac{S}{N}. \tag{8.1}$$

The parameters S and N denote the strength of signal and noise in the waveform. In our case this is defined via an energy envelope E_j, at sample j, calculated through a sliding integration of the waveform amplitude U_i of a sample i as

$$E_j = \sqrt{\frac{1}{n} \sum_{i=j}^{j+25\text{ ns}} U_i^2}. \tag{8.2}$$

The parameter $n = 50$ denotes here the number of samples which are integrated. All waveforms are interpolated to a sampling time difference of 0.5 ns for the SNR value. The signal is defined as $S = E_{j,max}$ of the given waveform. As noise value N the envelope of a data interval far away from the signal time is taken.
The SNR for an event is defined as the average of the 3 highest waveform $SNRs$ to reduce the influence of noise fluctuations. This parameter is useful for comparison between simulation in signal and noise in terms of the influence of the signal shape and given detector features on the analysis parameters. It does not depend directly on the frequency content of a waveform and the event SNR does not depend on the time window, trigger

settings and difference in signal gain between different antenna channels. Therefore the efficiency and robustness of the analysis can well be verified using the SNR.
One should note that the SNR defined here will be biased to values greater than one, due to the selection of the maximum value as signal. Pure noise distributions show an SNR of 1.9 ± 0.2. It should thus only be used as a comparator if greater than 2.

8.2 Thermal noise rejection

Most of the recorded background data in the ARA stations is thermal noise, which triggers at a rate of 5 Hz. Neutrino-like impulsive signals, which are generally anthropogenic noise, appear rarely within this background. In a first analysis step, impulsive events need thus to be identified within the noise background by relatively simple algorithms. These algorithms should not need large computing resources to manage the data analysis within a reasonable amount of time. In further steps, the relatively low number of identified impulsive events can instead be analyzed with highly sophisticated algorithms. In this section, a simple method to distinguish incoming radio signals from thermal noise is presented. In addition to that, the possibility to implement this algorithm as a trigger and the possible enhancement in sensitivity are discussed. The discussed algorithm and results on its efficiency are either updates to or partly published in [102].

8.3 The time sequence algorithm

This algorithm can be regarded as a boosted hit count. It is performed on each event in three steps:

- Calculation of an energy envelope for each waveform to enhance signal from noise fluctuations,
- construction of a hit time pattern via threshold discrimination,
- measurement of the similarity between the hit pattern and an incoming flat wavefront.

The energy envelope E is calculated as described in Section 8.1, with the difference that the integration time is chosen to be 5 ns instead of 25 ns. The purpose of this envelope is to reduce uncorrelated noise fluctuations compared to the signal, as can be seen in the example waveform displayed in Figure 8.1.

For the envelope E the average value μ_E is calculated, which corresponds to the RMS of the waveform. To avoid bias in case of the presence of strong signal, the waveform is divided into four segments and the average of the two segments with the lowest values

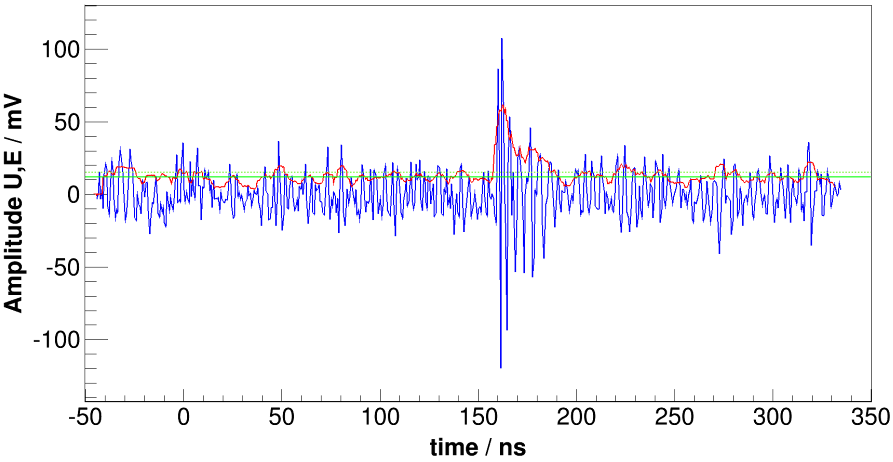


Figure 8.1: *An example waveform (blue) with energy envelope (red) and the average μ_E (green solid line) and RMS σ_E (green dashed line). Note the enhancement of the signal envelope compared to noise fluctuations relative to μ_E.*

is used. In addition to μ_E, the RMS of the envelope σ_E itself is computed as

$$\sigma_E = \sqrt{\frac{1}{N} \cdot \sum (E_j - \mu_E)^2}. \tag{8.3}$$

A signal bias is avoided in the same way as for μ_E. With these two parameters a discrimination threshold th_E is defined to select hits from the waveform envelope. A data point j is selected as a hit if E_j is above the threshold value

$$th_E = \mu_E + 4 \cdot \sigma_E. \tag{8.4}$$

This threshold is calculated for each antenna channel in each event separately and will thus dynamically adapt to changing noise and other environmental conditions. The selected hits are recorded with a precision of 5 ns. Multiple hits within these 5 ns are not counted.

Examples of the hit patterns generated using the method described above are shown in Figure 8.2 and 8.3 for a simulated signal event and for simulated noise. In the next step of the algorithm, the patterns are analyzed concerning their correlation to an incoming flat wavefront.

The ARA stations have a close to cubical geometry. If a flat radio wavefront hits a station, antennas with similar spatial relation between each other will show a similar hit time difference. In the applied algorithm, antenna pairs are formed from selected hit combinations between antennas of the same polarization and their hit time difference is calculated. This time difference is divided by the distance to diminish the influence of deviations from a cubical structure of the ARA stations. The hit pairs are separated into five categories, as indicated in Figure 8.4:

- One vertical group: pairs of antennas on the same string.

- Two lateral groups: pairs of antennas at the same depth on different strings in the two lateral dimensions.
- Two diagonal groups: pairs of antennas at the same depth on different strings, connected diagonally.

All pairs of one group have similar spatial relations between the partner antennas. One should note at this point that the algorithm works in principle in the same way if the station strings are mounted in a parallelogram instead of a square in the xy-plane. For the hit pairs found, the normalized time difference

$$dt_P = \frac{t_1 - t_2}{distance} \cdot 15 \tag{8.5}$$

is filled into a separate histogram for each group. The scaling factor 15 (18 for vertical pairs) is applied to allow for an integer binning of the histograms, which keeps the

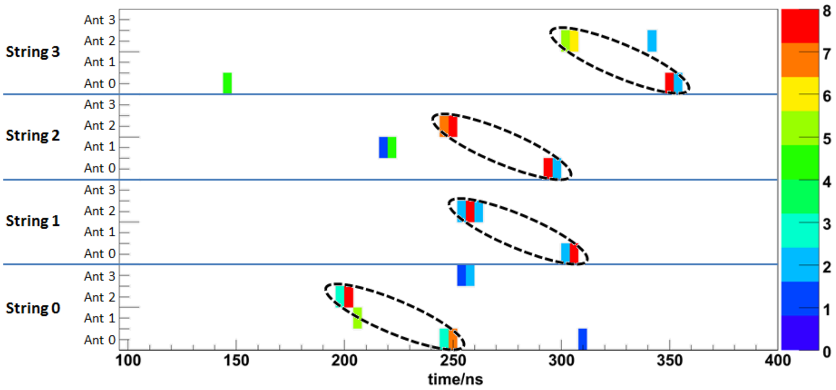


Figure 8.2: *An example signal hit pattern containing hits at times that the envelope waveform E exceeded the threshold th_E. The hits originating from the incoming signal wavefront are indicated with the black dashed ellipses. The number of hits per bin, indicated by the color scale, is ignored and the pattern contains* 25 *hits for further analysis.*

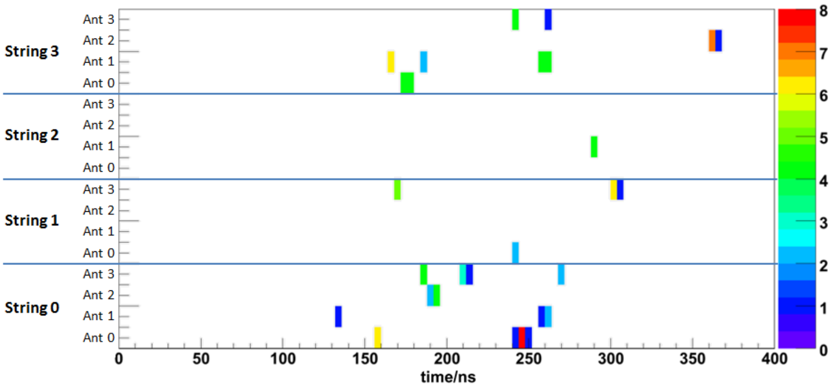


Figure 8.3: *An example noise hit pattern generated in the same way as the signal hit pattern in Figure 8.2. The number of hits considered for further analysis is* 28. *Note that waveforms only start at roughly* 120 ns.

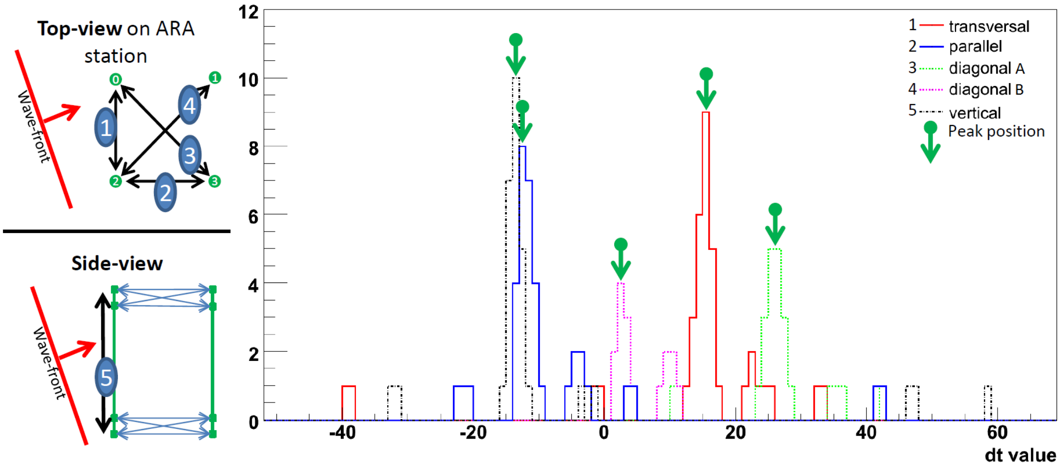


Figure 8.4: *The five histograms from different hit pair groups for the event shown in Figure 8.2. The spatial orientation of the pairs, corresponding to the histograms is indicated on the left. The quality parameter of this event is* 1.6*.*

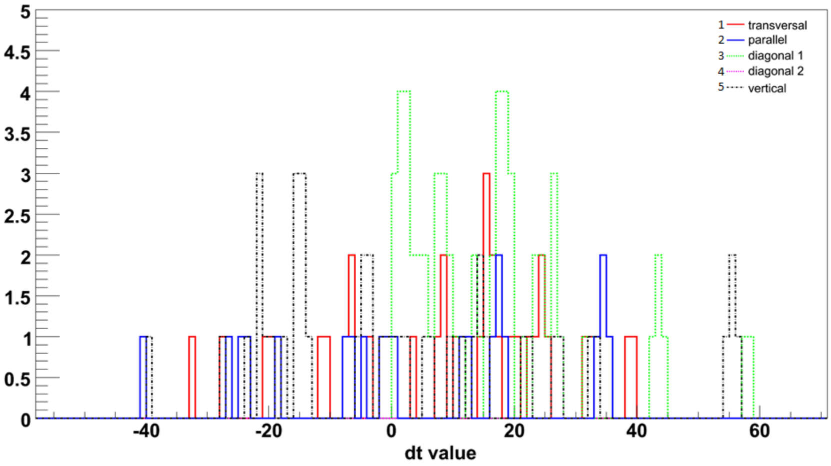


Figure 8.5: *The five histograms from different hit pair groups for the event shown in Figure 8.3. The spatial orientation of the histograms is indicated in Figure 8.4. The quality parameter of this event is* 0.5*.*

algorithm numerically simple. This is especially interesting for the potential usage as a trigger algorithm (Section 8.3.2). The histograms for a signal and a noise event are shown in Figure 8.4 and 8.5 respectively.

The expectation is that signal histograms show a concentration of hit pairs into peaks whereas noise histograms should contain random distributions. This behavior is well visible in the examples in Figure 8.4 and 8.5. The height of the peaks is used to calculate a quality parameter QP for the analyzed event:

$$QP = \frac{1}{T} \cdot \sum_{i=0}^{5} \sum_{j=max-1}^{max+1} H_{i,j}. \tag{8.6}$$

For each histogram i, the count $H_{i,j}$ is summed at the maximum bin and the two neighbors. This value is summed for all five pair groups. The result is normalized by a constant T proportional to the duration of the recorded waveforms, to compensate for statistically higher hit counts due to longer waveforms. The inclusion of neighboring bins has been found to significantly improve the efficiency of the algorithm as it compensates for the non perfectly cubical geometry and the deviation from a flat wavefront.
The two events compared in Figure 8.2 and 8.3 clearly show the signal boost compared to noise. Even with a smaller number of direct hits (signal 25, noise 28), the quality parameter for the signal event is significantly higher than for the noise event (signal 1.6, noise 0.5).

The quality parameter for a set of simulated neutrino events and simulated thermal noise (Chapter 7) is shown in Figure 8.6. One can see that the signal can be separated very efficiently from noise by a cut on QP. Figure 8.6 visualizes as well the character of QP, which is boosted compared to a simple hit count for the envelope waveform. Total hit count values are normalized by a constant to achieve a noise reduction to 1% at the same value as for QP. The distributions are not expected to match, neither for signal nor for noise, because of the different ways of calculation. However the reduction to 1% can be used as a benchmark for comparison. A clear enhancement of signal to noise can be seen for QP compared to a simple hit count.

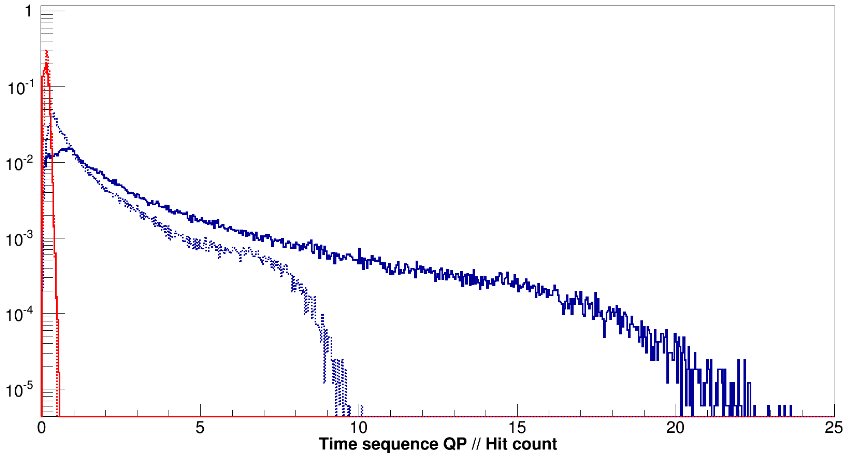
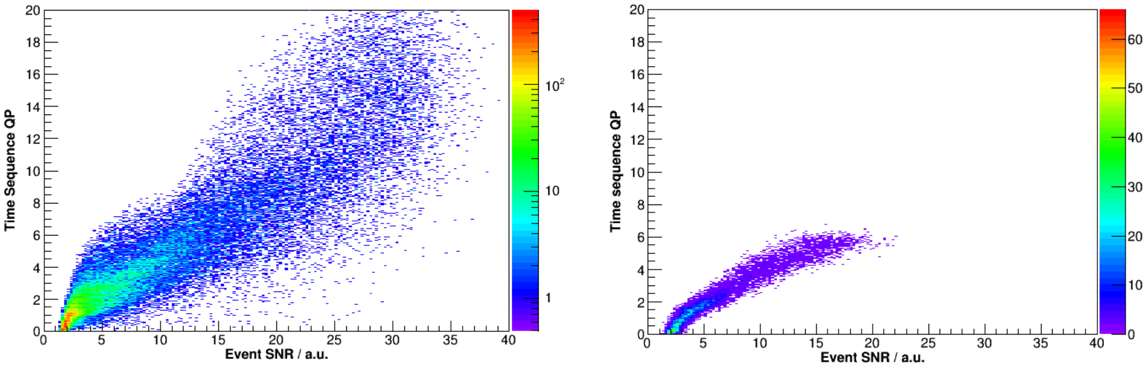


Figure 8.6: *The quality parameter* QP *(solid line) and the hit count (dashed line) for simulated neutrinos* ***(blue)*** *and thermal noise events* ***(red)*** *as described in Chapter 7. The hit count values are scaled such that the cumulative noise distributions reach* 99% *at the same value. Bins are normalized by the total event count of each distribution.*

One should note that simulations are done for the original ARA station geometry. Even with the deviations from a cubical structure the algorithm can be applied efficiently. The reason for this is the fact that the largest changes from a cubical structure are in the z-direction for both ARA stations (Section 6.3.3). This does not have a strong influence on QP for wavefronts coming in from relatively close to the horizon which are the majority (Chapter 7). Also, the assumption of a flat waveform appears to be a fair approach to obtain an efficient signal discrimination. One can calculate the distance at which a wavefront does not appear as flat anymore with a binning precision of 5 ns of the hit patterns and an antenna pair distance of 20 m. This distance is 234 m. Events coming from vertices nearer to the station will deviate from the flat wavefront assumption and are expected to show a slightly reduced quality parameter. This can have an influence on the calibration pulsers which are located at distances of around 40 m. Neutrino interactions are unlikely to appear within this range.

8.3.1 Efficiency comparison of simulation and data

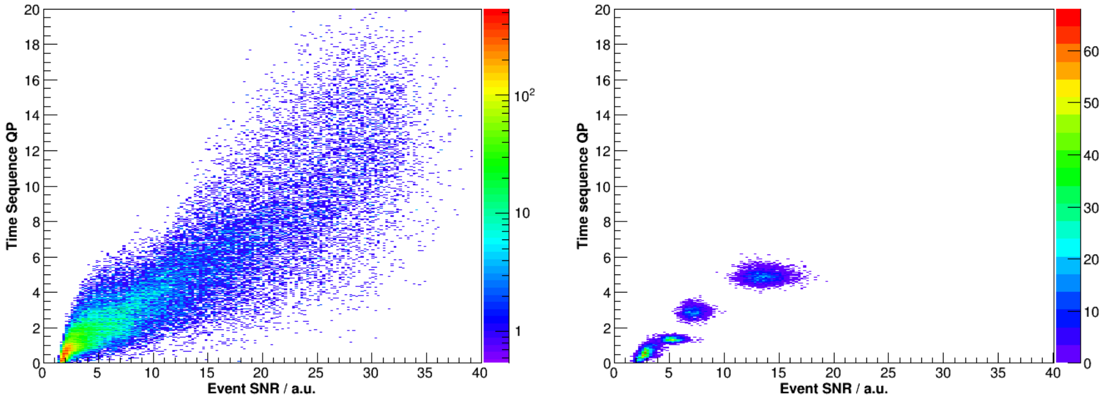
To compare the quality parameter for simulated events and data, amplitude scans of the in-ice calibration pulsers are used and compared to the simulated neutrinos. This comparison is of course limited by the fact of having only 4 close-by calibration sources available. These cannot cover all possible distances and angular positions as they are simulated for the neutrinos. Such comparison is shown for ARA03 in Figure 8.7 in plots of QP versus the SNR (Section 8.1). One can see that the ARA03 pulsers fit the distribution of simulated neutrinos very well in QP. For ARA02 the stepwise attenuation could not be performed as detailed as for ARA03. Still, data with different pulsers at different levels of attenuation from operation throughout the year could have been gathered as shown in Figure 8.8. The distributions fit nicely into the simulated data.



(a) (b)

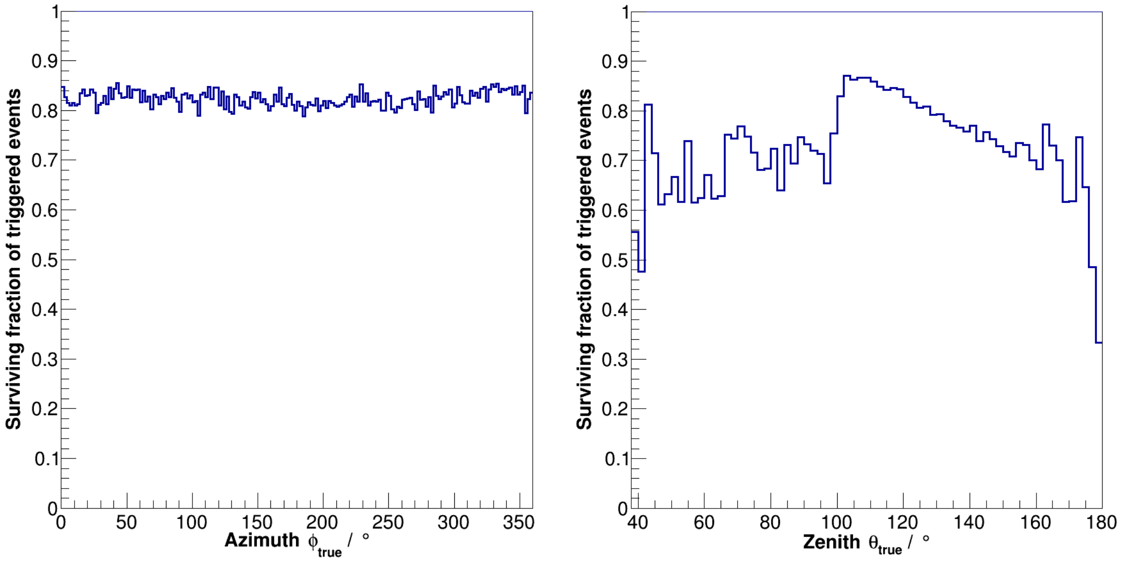
Figure 8.7: *The quality parameter QP versus SNR of simulated neutrinos for ARA03* ***(a)*** *and stepwise attenuated calibration pulsers recorded by ARA03* ***(b)****.*

The angular acceptance after a cut on the quality parameter is shown in Figure 8.9. The azimuthal efficiency appears to be isotropic, while the zenith acceptance favors events originating from close to the horizon. The enhanced efficiency for events coming from slightly below the horizon can be explained by the ray-tracing effect in the ice. Due to this effect, the wavefront from events in the ice always comes in under an elevated zenith



(a) **(b)**

Figure 8.8: ***(a)*** *The quality parameter* QP *versus* SNR *of simulated neutrinos for* $ARA02$ *and* ***(b)*** *in-ice calibration pulsers at different levels of attenuation, recorded by* $ARA02$*.*



(a) **(b)**

Figure 8.9: *The angular acceptance after a cut on the quality parameter* QP *at* 0.6 *for simulated neutrinos. The histograms show the fraction of triggered events, passing this cut.* ***(a)*** *Azimuthal acceptance,* ***(b)*** *zenith acceptance (limited to statistically relevant regions). Note that the true zenith angle indicates the location of the emission vertex relative to the ARA station.*

angle, compared to the vertex position. A horizontally approaching wavefront belongs thus to a slightly negative vertex position.

It has been shown in this section that the quality parameter QP is a very efficient discriminator between radio signal sources and thermal noise. In a comparison between data and simulation versus SNR, systematic differences appear to be small for ARA03 and for ARA02. It should be mentioned that such comparison is limited to the four point sources in the ice. The discrimination efficiency is isotropic in azimuth, but depends on the incoming zenith angle. This does not cause a great neutrino signal loss, since best efficiency is achieved for the zenith angle which is highest populated by triggering neutrinos.

8.3.2 Filter and trigger studies

As mentioned in the previous section, the time sequence algorithm is relatively simple but very efficient. Its limited operations could be implemented in an FPGA and the algorithm could thus be used as an event trigger. The ARA trigger system is based on an integrating tunnel diode which produces waveforms similar to the envelope E from Equation 8.2, with an integration time of a few nanoseconds. The trigger logic, implemented in an FPGA, processes these waveforms as described in Section 4.1.2. It is looking for coincident hits within a time window of 110 ns on at least 3 antennas of the same polarization to trigger an event. This algorithm has a limited efficiency in discriminating signal from noise. To keep the event rate low enough for a proper operation of the readout electronics, a high power threshold has to be used to identify hits. The time sequence algorithm is in principle simple enough to be implemented in the FPGA instead. The tunnel diode in the trigger system produces waveform outputs similar to the energy envelope (Section 4.1.2), used in the here described algorithm. This output can thus be used as input envelope waveforms to generate the hit patterns. A small study has been performed to check the possible gain from such an implementation on simulated neutrino events.

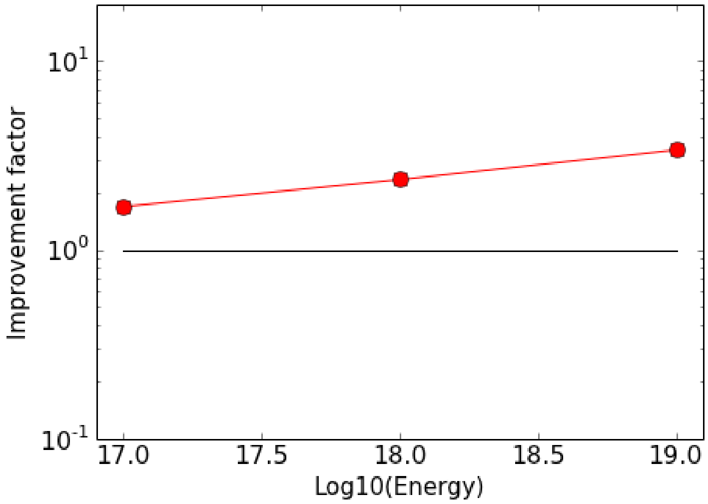


Figure 8.10: *The efficiency improvement of the time sequence algorithm over the current ARA trigger versus energy, for simulated neutrino events recorded by an ideal cubical station.*

The investigation has been pursued on a separately simulated neutrino and noise sample, using an idealized cubical station geometry. The neutrino signal is simulated for three discrete primary energies of 10^{17}, 10^{18} and 10^{19} eV. The current ARA trigger and the time sequence algorithm have been used to trigger events. The diode threshold is adjusted for each trigger algorithm, to keep the final noise event rate constant at 10 Hz, as it is required by the readout electronics. On the used sample, the time sequence algorithm allows for a threshold reduction by a factor of 1.7, compared to the standard trigger. The possible improvement factor in effective volume is shown in Figure 8.10. This small investigation shows very encouraging results. However, the result needs to be reproduced under more realistic conditions and the trigger has to be tested on real hardware to really understand the final gain in sensitivity.

8.4 The vertex reconstruction algorithm

The primary goal of event reconstruction in the scope of this thesis is the distinction of signals generated by neutrino interactions from those produced by other impulsive radio sources. Some of these sources are the in-ice calibration pulsers installed with the ARA stations. Others are buildings, airplanes and human activities at the surface of the ice. An angular reconstruction allows us to correlate events with a given emitter location and to exclude them from our signal sample by angular cuts. This will lead to an efficiency loss on neutrino signals, which depends on the precision to which we can reconstruct events. A radial reconstruction of events is very difficult with the given station geometry. This can thus not be used for distinction. However, the resulting signal loss is small and a radial reconstruction therefore not crucial. It only becomes important once a signal candidate has been found to determine its primary energy. The expectation for this to happen is small with the current detector setup.

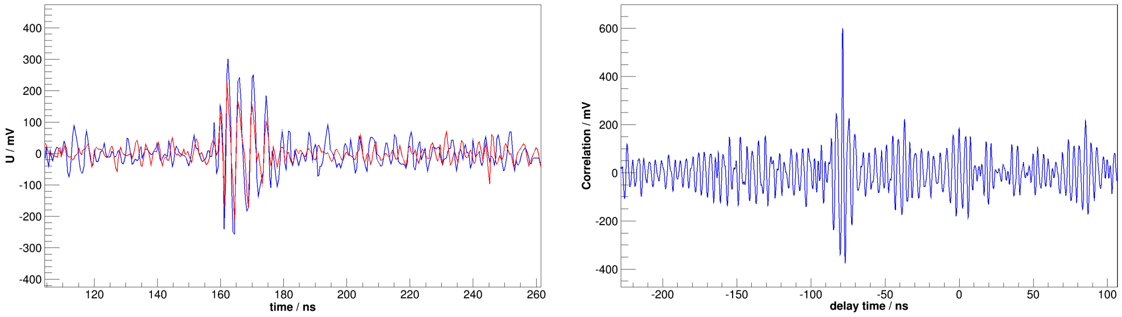
In this chapter a matrix based reconstruction method for radio sources is presented with details about timing determination, the mathematical algorithm and the final performance on measured and simulated data.

8.4.1 Timing determination

The simplest way to perform a radio vertex reconstruction in ARA is to measure the arrival time differences of signal at different antennas and use their relative positions to trace the signal back to the vertex location. The arrival time difference can be obtained with different methods. Their efficiency depends generally on the received signal shape. For a radio signal source in the ice we expect the waveforms reaching different ARA antennas, to be of very similar shape. Effects like dispersion and reflections at the surface normally affect the signal for all antennas similarly due to their vicinity within the cluster. Such effects can thus be expected to cause only small differences. Also the change in antenna gain and spectral response of the signal chain electronics is rather small and mostly induces overall differences in amplitude rather than changing the shape of the waveform. For these reasons, a good approach to determine the arrival time difference on two antennas is to use the cross correlation between them. The cross correlation g between two functions f_1 and f_2 can be calculated via the Fourier transform $\mathcal{F}$ as

$$g = f_1 \star f_2 = \mathcal{F}^{-1}\left((\mathcal{F}(f_1))^* \cdot \mathcal{F}(f_2)\right). \tag{8.7}$$

The correlation function g holds the correlation at a time shift dt between the two antennas. The maximum of g arises at the time dt when the two waveforms overlap perfectly. Figure 8.11 shows two signal waveforms, shifted against each other by the maximum correlation time, and their cross correlation function.

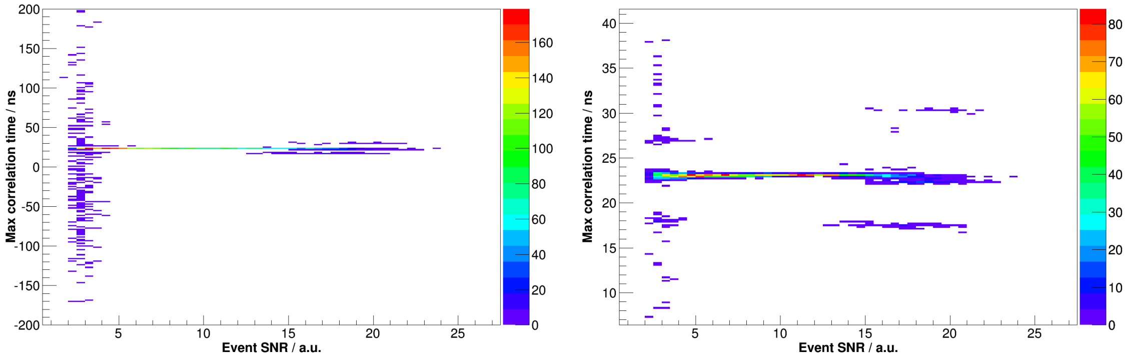


(a) **(b)**

Figure 8.11: *An example correlation of signals on two channels in ARA02.* ***(a)*** *The two waveforms shifted against each other by the maximum correlation time* $dt = -78.4$ ns. ***(b)*** *The correlation function for the two waveforms before shifting. The maximum of this graph is used to determine the delay between two signals.*

The precision of timing determination by cross correlation depends on the signal to noise ratio since noise randomly disturbs the correlation. As shown in Figure 8.12, the timing correlation for very weak waveforms with an SNR below 4 is basically random but improves strongly with rising SNR. On the other hand, the timing correlation gets distorted for extremely strong waveforms. This happens due to saturation effects in the readout chain and mostly in the digitizer chip. Waveforms can only be handled properly up to amplitudes of roughly 600 mV. Beyond that the digitizer starts to flip samples in sign, and the voltage calibration becomes imprecise.

The result can be seen in Figure 8.12 for high SNR values. The maximum correlation time can jump in steps of a few nanoseconds. The step size of these jumps is connected to the oscillation feature, visible in the correlation graph in Figure 8.11. This oscillation depends on the frequency content and shape of the compared waveforms. The observed saturation effect is difficult to model since it depends on many factors. Moreover, there are only limited waveform sources available to study the distortion from saturation. In Figure 8.13 a projection on the time axis of the saturated SNR range is shown for the



(a) **(b)**

Figure 8.12: *The correlation of two channels for calibration pulsers of varying signal strength indicated by the SNR of the event on the X-axis.* ***(a)*** *The full distribution,* ***(b)*** *a zoom into the region of the main peak. SNR regions of proper correlation can be clearly distinguished from regions of strong noise influence and disturbance due to saturation effects.*

strongest distortion found in calibration pulser events. The disturbance can be modeled using a combination of Gaussian distributions $G(\mu, \sigma)$ as

$$f(dt_s) = \sum_i \left[G(\mu = b_i, \sigma = 0.2) \cdot G(\mu = 0, \sigma = 3.1)\right](dt_s), \tag{8.8}$$

with the mean value μ and the width σ for the Gaussian distributions. The parameter b_i denotes the position of the side peaks. The function $f(dt_s)$ can be used as a conservative probability distribution to model a random disturbance for saturated waveforms in simulation. The positions of the side peaks are chosen as $b_i = \pm i \cdot 6$.

There are some aspects that should be noted about the correlation amplitude. The correlation graph is examined without normalization. This leads to a bias to small correlation times for noise and weak signals since the amplitude grows with the overlap of the waveforms. The proper normalization by the waveform overlap has been tested and found to reduce stability for the reconstruction performed here. Therefore, it has not been applied.

The finally found maximum amplitude A_{corr} of the correlation graph will be of further use in the reconstruction algorithm. For this purpose, it is normalized by the sum over the voltages U_1 and U_2 of the two waveforms as

$$A^2 = \frac{A_{\text{corr}}^2}{\sqrt{\sum U_1^2 \cdot \sum U_2^2}}. \tag{8.9}$$

This normalization is applied to avoid high correlation values for pure noise waveforms from channels with a high noise RMS. The value A^2 is used as the correlation amplitude for channel selection in the reconstruction. This will be described in Section 8.4.3.

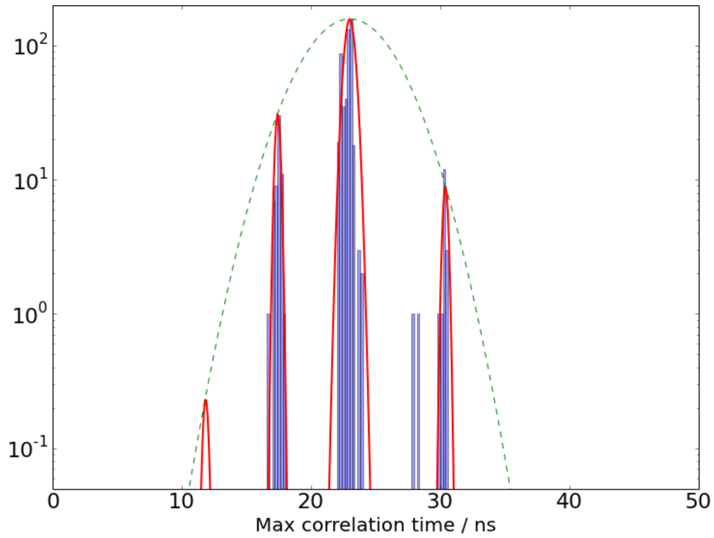


Figure 8.13: *The distribution of the maximum correlation time in the region of saturation. The red line shows the deduced model for the measured effect. The model is a base Gaussian distribution indicated by the green dashed line, multiplied by local Gaussian distributions at the side peaks.*

8.4.2 Matrix based vertex reconstruction

The reconstruction algorithm used in this analysis is similar to Bancroft's method for GPS position reconstruction [103]. As with the geometrical station fitting, described in Section 6.3.1, one starts from the relation

$$c^2(t_v - t_i)^2 = (x_v - x_i)^2 + (y_v - y_i)^2 + (z_v - z_i)^2. \tag{8.10}$$

The emission vertex coordinates are denoted x_v, y_v, z_v and emission time by t_v. The coordinates of the receiving antenna are represented by x_i, y_i, z_i and its arrival time by t_i. The speed of light in the propagation medium is assumed to be $c = 0.3/1.755$ m/ns, which is the speed of light in ice at a depth of 180 m, calculated from the index of refraction model in Equation 4.1. Since in ARA we only measure relative time differences between the antennas available, we subtract Equation 8.10 for two antennas i and j. After reordering this becomes

$$\begin{aligned} \boldsymbol{x_v} \cdot 2x_{ij} + \boldsymbol{y_v} \cdot 2y_{ij} + \boldsymbol{z_v} \cdot 2z_{ij} - \boldsymbol{t_{v,ref}} \cdot 2c^2 dt_{ij} \\ = r_i^2 - r_j^2 - c^2(dt_{i,ref}^2 - dt_{j,ref}^2). \end{aligned} \tag{8.11}$$

The variables with annotation a_{ij} represent the difference $a_i - a_j$. The emission time relative to a chosen reference antenna is given by $t_{v,ref}$ and the time difference between two antennas is denoted as dt_{ij} or $dt_{i,ref}$ respectively. The parameter $r_i = \sqrt{x_i^2 + y_i^2 + z_i^2}$ denotes the distance of an antenna i to the station center. It is apparent that Equation 8.11 is linear in the vertex coordinates and emission time. When setting up a system of equations with all possible antenna pairs, the linearity allows us to solve the equations for the vertex coordinates analytically. One can note that at least 4 independent antenna pairs a_{ij} and thus 5 antennas are needed, including the reference antenna, for an unambiguous solution from this method. In principle a position can be reconstructed with only 4 receivers. In our method we sacrifice one additional antenna for the linearity of the system. However, due to the close to cubical structure of the ARA detector stations, it is advised to use at least 5 antennas to avoid failures from small timing errors. The final system of equations can be written as a matrix equation

$$\mathbf{A}\vec{v} = \vec{b}, \tag{8.12}$$

with the matrix $\mathbf{A}$ of rank $N \times 4$, the vertex vector $\vec{v} = (x_v, y_v, z_v, t_{v,ref})$ and the right-hand side of Equation 8.11 in the vector $\vec{b}$. The rank N corresponds to the number of antenna pairs used.

An appropriate way to solve Equation 8.12 is to apply matrix decomposition methods. For this, algorithms from the C++ library Eigen are used [104]. Various algorithms have been tested and the LU decomposition method with full pivoting has been found as a fast algorithm with the most stable solutions. In this method, the matrix $\mathbf{A}$ is decomposed into a lower triangular matrix $\mathbf{L}$ and an upper triangular matrix $\mathbf{U}$. The system of equations can then be solved in two steps in a trivial manner. The full pivoting is in principle a permutation of rows and columns, used to make the algorithm fast and robust [105]. The matrix $\mathbf{A}$ is decomposed into

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{L}\mathbf{U}\mathbf{Q}^{-1}. \tag{8.13}$$

The matrices $\mathbf{P}$ and $\mathbf{Q}$ are the permutation matrices. This method in principle allows us to perform several thousand reconstructions per second.

The above described algorithm is expected to be very stable for reconstructions with widely separated receiving antennas. In the ARA stations, antennas are very close to each other and they are used to determine vertex positions far away in the ice. Reconstructions performed in such a setup are strongly affected by timing errors and their stability can be reduced. To counteract these errors, the method is modified by changing the emission time t_v from a reconstructed parameter into a fixed negative value, which is scanned for the best reconstruction of an event. This allows us to avoid reconstructions to unreasonably small radii or to positive emission times, which are always connected to an incorrect position. The requirement of a minimum negative emission time, which corresponds to a minimum vertex distance, renders the angular reconstruction very robust. For the reconstructions in the t_v scan a residual is defined as

$$res = \left| \frac{\vec{b}}{|\vec{b}|} - \frac{\mathbf{A}\cdot\vec{v}}{|\mathbf{A}\cdot\vec{v}|} \right|^2 \cdot \frac{1}{N_{chp}}. \tag{8.14}$$

This residual uses the two sides of the equation normalized. Without normalization, the residual depends strongly on the distance to the event, and smaller distances will

be highly favored due to timing errors. The normalization compensates the distance dependence to a certain extent and the angular reconstruction results to be much more stable. The scan is performed in 200 timing steps between −150 ns and −22900 ns, which corresponds to a distance of roughly 4000 m. This decreases the speed of an event reconstruction dramatically since it has to be performed 200 times.
The found best residual of a reconstructed event is the main quality indicator for its position determination. It can separate good from bad reconstructions very well, as visible in Figure 8.14. The shown azimuthal reconstruction becomes very tight with decreasing residual.

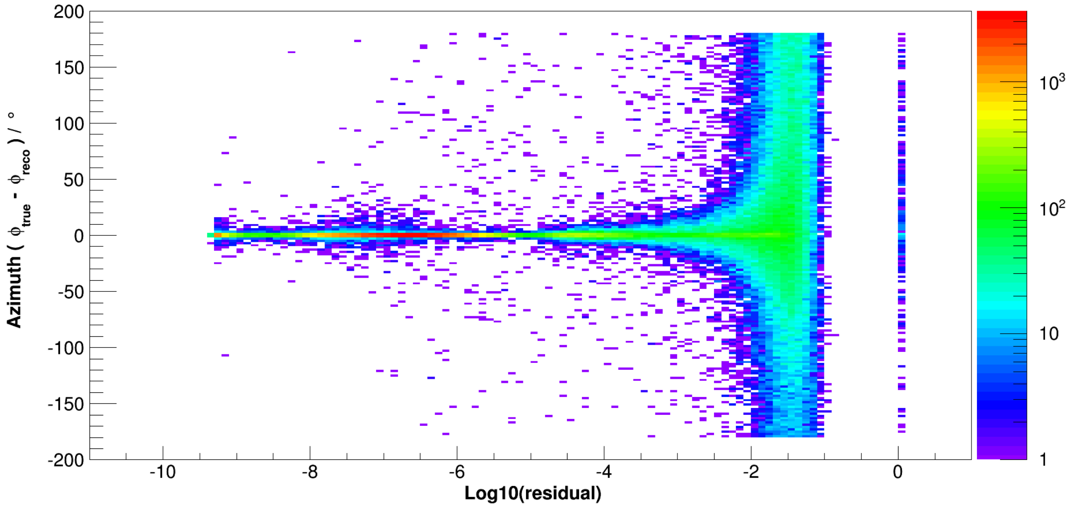


Figure 8.14: *The azimuthal reconstruction of simulated events versus the residual of the reconstruction. The residual is set to* 1 *if the number of available channel pairs is not sufficient for a reconstruction.*

In addition, the residual can also be used to distinguish signal events from thermal noise. If the correlation amplitudes in a noise event are sufficient to pass the threshold for reconstruction, the times will be random and the residual thus much higher than for a timing pattern connected to an incoming wavefront. The residual for simulated signal and simulated thermal noise is shown in Figure 8.15. One can see that a very effective separation can be achieved by applying a residual cut. Although this is not the primary purpose of the reconstruction algorithm, it can help to render final cuts on thermal noise more efficient.

A closer investigation should be dedicated to the features in the residual distribution of the signal. Good reconstructions appear in the main peak at logarithmic residuals between −6 and −8. Very weak signals, coming for example from very distant interactions of low energy, form the noise like distribution of high residuals. A further feature is a small bump between logarithmic residuals of −4 and −5. It is difficult to extract the reason for this peak from the simulated data. Investigations have shown that well shaped

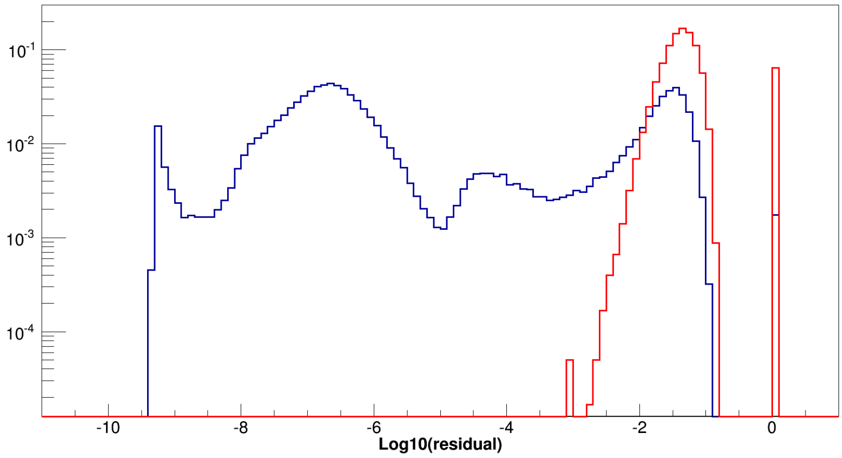


Figure 8.15: *The residual for simulated neutrino events* ***(blue)*** *and thermal noise events* ***(red)*** *as described in Section 7. The histograms are normalized by the total event count. Events with insufficient selected channel pairs have a residual of* 1.

signal events with good signal correlation times but one single noise timing in the reconstruction, peak in that region. Further research is needed to confirm this explanation. The rise at lowest residuals is an accumulation due to a limit set in the reconstruction algorithm.

8.4.3 Channel pair selection

For the reconstruction it is crucial to avoid noise influences and to select channel pairs with robust correlation timings. This is especially challenging for faint events with low SNR. For the presented reconstruction, channel selection is performed in two steps:

- In the first step the antenna pairs are ranked by their normalized correlation amplitude A^2 (Equation 8.9). Two signal waveforms are expected to show a stronger correlation than correlations with noise waveforms. A high threshold is set to distinguish good correlations from bad correlations. This threshold is lowered stepwise, until a sufficient number of channel pairs is found or until a lower threshold limit is reached. In principle 7 pairs are sufficient to guarantee the usage of 5 antennas and thus 4 independent pairs. Since there is a realistic chance of selecting noise correlations amongst the channel pairs, this is not sufficient and reconstructions are relatively unstable. A selection of 10 channels has been found to enhance reconstruction stability and efficiency significantly.

- The second step of channel pair selection is performed with the reconstruction itself. Due to the limited efficiency of the amplitude ranking there is a chance of having selected pure noise waveforms for the reconstruction. Such noise influences will, apart from disturbing the reconstruction precision, strongly increase the residual since their timing does not fit to a reconstructed point. In a second stage reconstruction all pairs including a given antenna are excluded and the residual is checked for the reduced set. The channel with the highest improvement, if any, is excluded from the reconstruction of the event. The goal of this exclusion is to clean the reconstructed event of possible noise influences. The second stage channel selection is only performed if more than 10 pairs have been selected in the first place. Furthermore, when excluding a channel, the reduction of pairs is limited to the point of keeping at least 10.

As mentioned in Section 4.2 and 6.3.3, there are a few channels which are never used in reconstructions in station ARA02: channel $D3BV$, $D3BH$ and $D4BH$. They are avoided due to not understood systematic timing offsets and noise problems. The selection possibility for this station, especially for horizontally polarized events is thus limited and the reconstruction is expected to be of lower precision and robustness.

8.4.4 Reconstruction quality parameters

When performing an event reconstruction, apart from the residual, a few more indicators for the reconstruction quality are used to mark badly reconstructed events. A few of them have already been mentioned during the description of the algorithm. The complete overview with a small explanation is listed below.

- **The number of channel pairs** found with a sufficiently high correlation amplitude A^2, has to be at least 10. With a lower number, the chance of misreconstructions and low residuals for thermal noise events rises. This would render a later distinction between signal and noise more difficult.

- The scanned **emission time** has to be less than its initial value. The timing scan is performed, starting at -150 ns. Events with a high noise contamination in the correlation times often reconstruct to the smallest possible radius. The emission time is thus required to be smaller than -155 ns.

- Related to the minimum negative emission time, a **minimum radius** of 26 m is required. This radial requirement is independent from the emission time requirement since it is related to the center of the coordinate system as opposed to the reference antenna.

- Consistency checks between timing and location are used as well to check the quality of a reconstruction. One such consistency check is the measurement of the **discrepancy between the emission time and the distance** of the reconstructed vertex to the reference antenna. A parameter R_{Err} is calculated as

$$R_{Err} = \frac{|t_{v,ref} \cdot c - R_{v,ref}|}{R_{v,ref}}, \tag{8.15}$$

with the emission time $t_{v,ref}$ relative to the reference time and the distance $R_{v,ref}$ between the reconstructed vertex and the reference antenna. A relation between the angular reconstruction quality and R_{Err} is plotted in Figure 8.16, after applying all other quality cuts. One can see that the reconstruction quality rises strongly with falling R_{Err}. A requirement of less than 0.15 for this quality parameter appears to ensure a high precision.

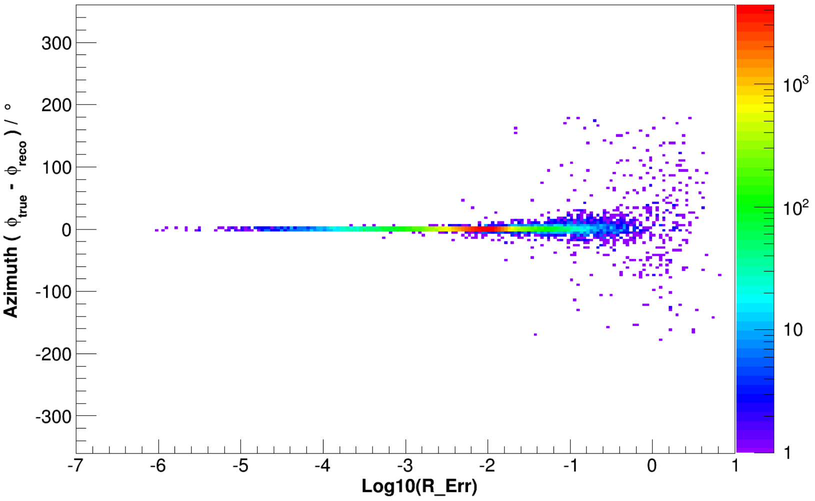


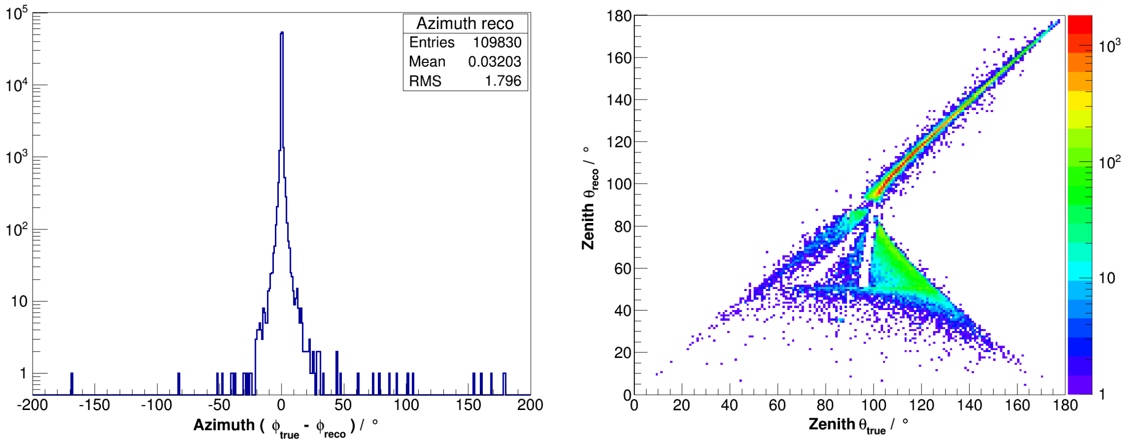
Figure 8.16: *The azimuth reconstruction versus the logarithmic radial discrepancy* R_{Err} *for simulated neutrino events in the ice.*

- A further consistency check is related to the **zenith angle**. After reconstruction, the correlation time of the vertical antenna pair with the highest A^2 is used to calculate an expected zenith angle for the event. This is easily possible since the antennas are assumed to have a perfectly vertical spatial relation. If the reconstructed zenith angle does not agree with the expected angle the chance of a bad reconstruction is high. The limit for well reconstructed events is set at a difference of 20°.

If an event reconstruction fails one of the above mentioned criteria, it is excluded from the event sample for further analysis.

8.4.5 Performance on simulation and recorded data

The reconstruction has been tested on simulated data, as well as on known radio signal sources at South Pole. The results for simulations are shown in Figure 8.17 after all quality cuts and a residual cut at 10^{-4}. Of the complete simulated data sample, which includes the features of ARA02 and ARA03, on average 66% are marked as properly reconstructed events. In azimuth an angular precision of 1.8° can be reached in that way. For a small number of events the determination of the azimuth angle fails. The suspicion is that they are related to rarely appearing pathological channel pair combinations which are not handled properly in the algorithm. However, this number is very small and negligible for the purpose of a neutrino search. The zenith reconstruction shows interesting features which are connected to ray-tracing and surface reflections. The reconstruction is performed assuming the speed of light to be equal to the average speed of light at the station depth and constant throughout the ice. This does not affect the azimuthal reconstruction, but can disturb the reconstructed zenith angle severely. In Figure 8.17 events which occurred at high zenith angles, thus below the station, are reconstructed to low zenith angles above the station. The radio rays of these events have been found to undergo surface reflections or strong ray-tracing effects, which cannot be accounted for in the present reconstruction. About 30% of all events are reconstructed upwards and have to be treated carefully in further analysis. Still, events reconstructed as coming from the deep ice (zenith $> 90°$) show a solid relation between true and reconstructed zenith. A slight bow is visible in the distribution which again is caused by ignoring the changing index of refraction in the ice.

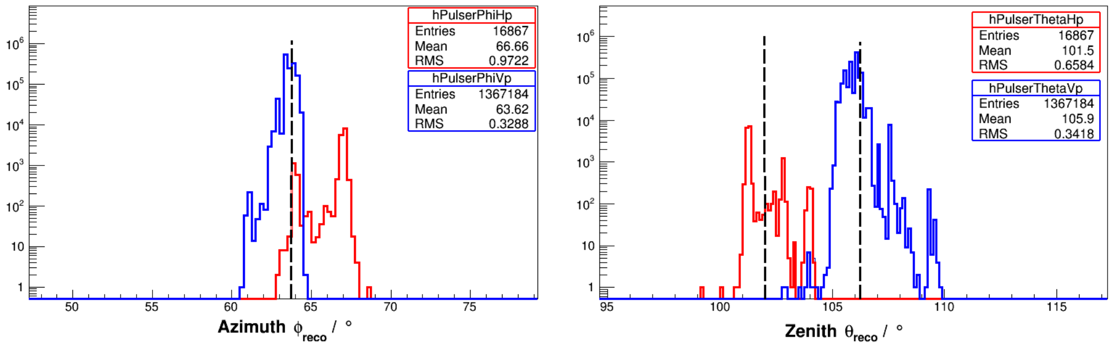


(a) **(b)**

Figure 8.17: *The angular reconstruction of simulated neutrino events hitting the two ARA stations as described in Chapter 7.* ***(a)*** *The difference between the true and the reconstructed azimuth angle.* ***(b)*** *The reconstructed zenith angle versus the true zenith angle of the event.*

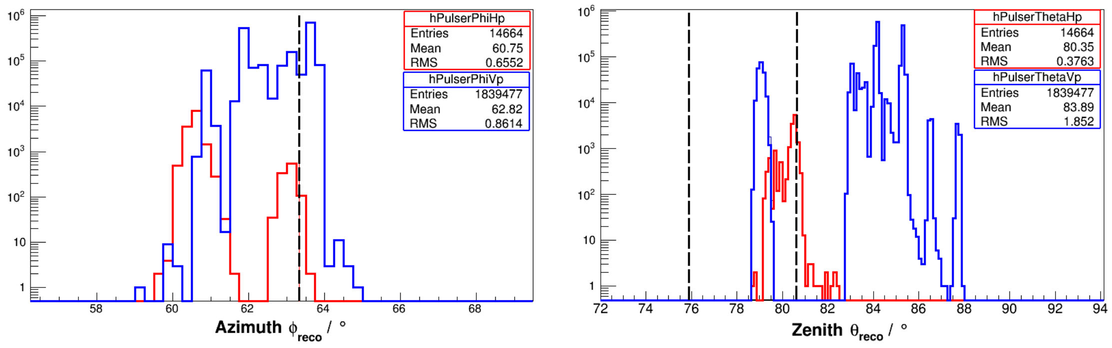
As discussed, the two distributions in Figure 8.17 reveal the limitations of the current algorithm. In further analysis, we have to exclude events reconstructed to the surface of the ice since they are most likely anthropogenic noise. Due to the misreconstruction in zenith, we will lose sensitivity when cutting on small zenith angles. However, this loss is not dramatic and still allows for a high analysis efficiency.

The radial reconstruction appears to not be very efficient and is not presented here. This will need improvement for the investigation of candidate neutrino events but is not of great importance for the distinguishing of signal events from noise.



(a) **(b)**

Figure 8.18: *The angular reconstruction of the two ARA03 D6 pulsers from* 10% *of the data collected in* 2013. *Hpol reconstruction is shown in red, Vpol in blue. The black dashed lines indicate the nominal pulser position.*



(a) **(b)**

Figure 8.19: *The angular reconstruction of the two ARA02 D6 pulsers from* 10% *of the data collected in* 2013. *Hpol reconstruction is shown in red, Vpol in blue. The black dashed lines indicate the nominal pulser position.*

To test the reconstruction on real data, we have five known radio sources available in the vicinity of the ARA stations. Those are the four in-ice pulsers and the roof pulser mounted on the ICL. Results for the roof pulser reconstruction have been presented in Section 6.3.3 under loose quality conditions. The in-ice pulsers are triggered every second during operation. Figure 8.18 shows the reconstruction of the ARA03 pulsers in hole D6 from 10% of data throughout the year 2013. Settings of polarization and amplitude have been changed during that time.

One can note that the distributions are not very smooth. A reason for that is the fact that the pulser was operated partly at maximum emission power, which saturated the station and disturbed the timing correlation (Section 8.4.1). The second reason is the still imperfect station geometry calibration. Antennas are still slightly off from their nominal position, which is why reconstructions including different antennas lead to different results. This is visualized in Figure 8.20 in the zenith distribution plotted versus

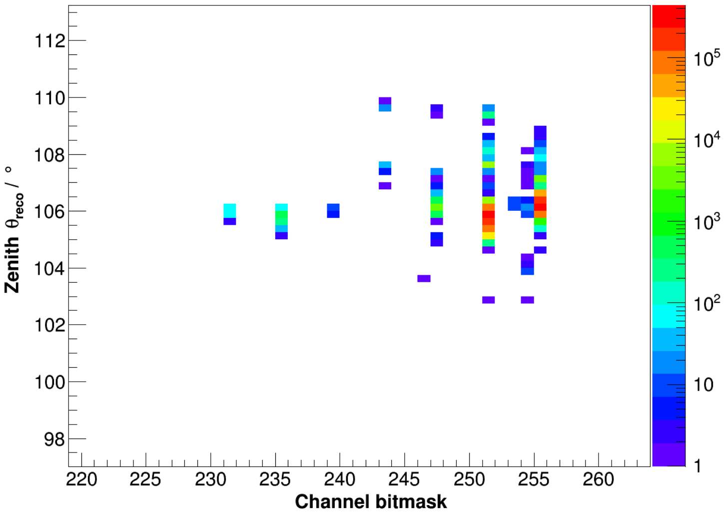


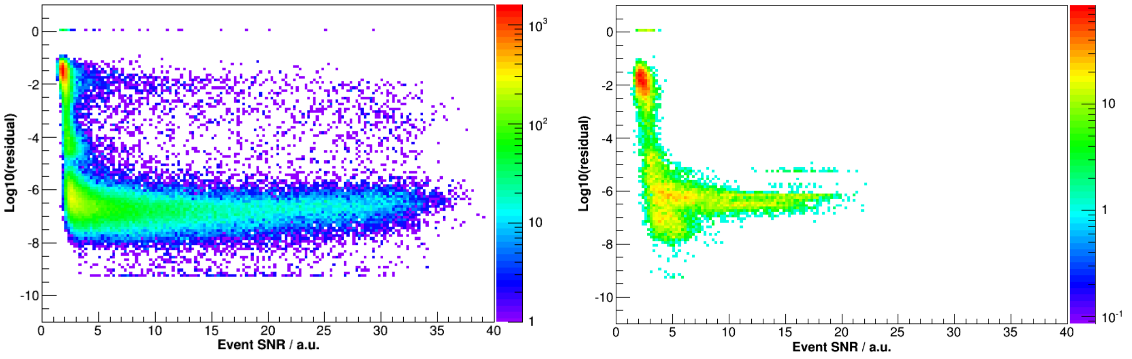
Figure 8.20: *The reconstructed zenith angle of Vpol events recorded in* 2013 *with ARA03 versus the bitmask, which represents the combination of used channels in a reconstruction.*

the bitmask of used channels for the Vpol events. Each used antenna channel is represented by one bit of a 8bit binary number, according to its channel number. The bitmask is the decimal number corresponding to a given combination of used channels. Clear differences in zenith reconstruction can be observed for different channel combinations. Still, the general reconstruction is very stable and the RMS is always less than 1° for both polarizations.

For ARA02 a similar behavior can be observed for the angular reconstruction (Figure 8.19). However, the zenith reconstruction shows a systematic offset, which could not be completely compensated in the geometry calibration of the station. ARA02 reconstructions are in general expected to be weaker, due to three missing channels and a less convincing geometrical calibration (Section 6.3.3).

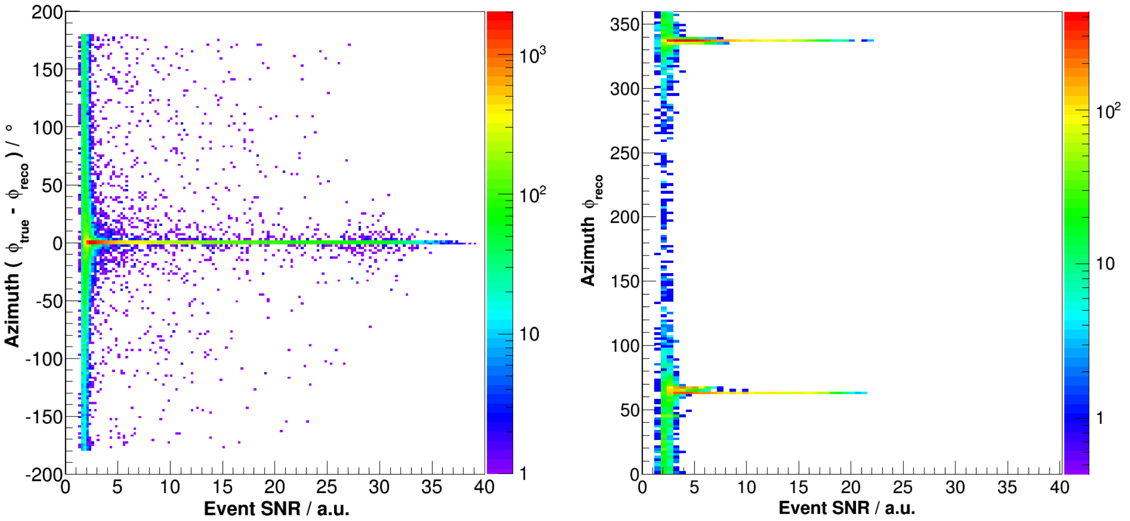
It should be noted that the reconstructions of the in-ice calibration pulsers are very robust, performing millions of reconstructions without failure. An additional investigation of the robustness will be shown in Section 9.3 for the development of angular cuts.

A cross check between recorded data and the simulation can be performed as for the time sequence quality parameter QP by looking at stepwise attenuated signal from different pulsers versus event SNR. Such a comparison for ARA03 can be seen for the residual in Figure 8.21 and for the azimuthal reconstruction in Figure 8.22. The data include all four in-ice calibration pulsers around the station. Both parameters show a good agreement between simulation and real data. This gives us confidence that the simulated events are well modeled as far as the reconstruction algorithm is concerned and systematic errors on the efficiency and precision of the reconstruction are small. Furthermore, in the simulations the station geometries and timings are assumed to be perfectly known. The good agreement in Figure 8.21 and 8.22 show that systematic errors on the station



(a) **(b)**

Figure 8.21: *The reconstruction residual versus SNR for **(a)** simulated neutrinos in ARA03 and **(b)** stepwise attenuated calibration pulsers D5 and D6 recorded by ARA03.*

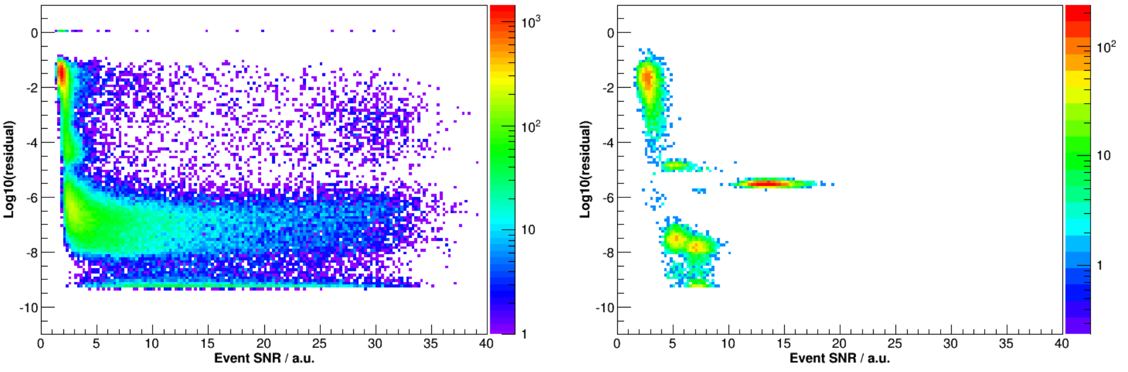


(a) **(b)**

Figure 8.22: *The reconstructed azimuth angle versus SNR compared to the true value for **(a)** simulated neutrinos and **(b)** stepwise attenuated calibration pulsers D5 and D6 recorded by ARA03. No quality criteria are applied.*

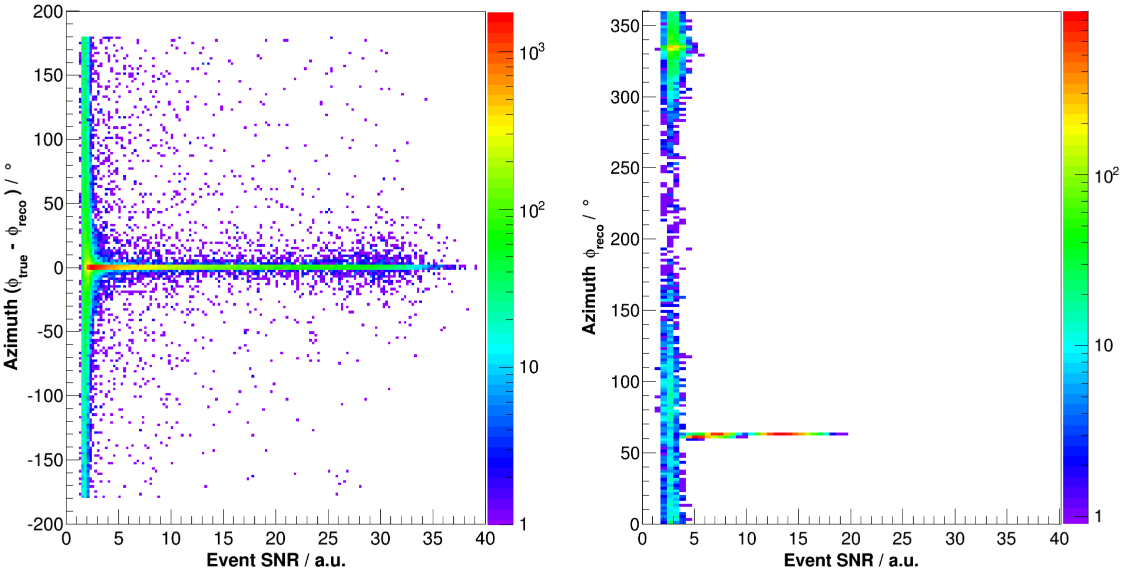
geometry only have a small influence on the reconstruction quality. The final systematic errors on the analysis efficiency will be presented in Section 10.1.

The SNR comparison for ARA02 is plotted in Figure 8.23 (residual) and Figure 8.24 (azimuth reconstruction). As mentioned in Section 8.3.1, for this station there are less detailed data available. In Figure 8.23 differences in the behavior of the residual can be seen for events with $SNR > 10$. This is most likely caused by the saturation effects which are observed in these events (Section 8.4.1). The same pulser with weaker signal ($5 < SNR < 10$) has a much lower residual in the reconstruction. In the azimuthal distribution in Figure 8.24 one can see that signal from the D5 pulsers (at azimuth of



(a) **(b)**

Figure 8.23: *The reconstruction residual versus SNR for* ***(a)*** *simulated neutrinos in ARA02 and* ***(b)*** *calibration pulsers under different attenuation, recorded by ARA02.*



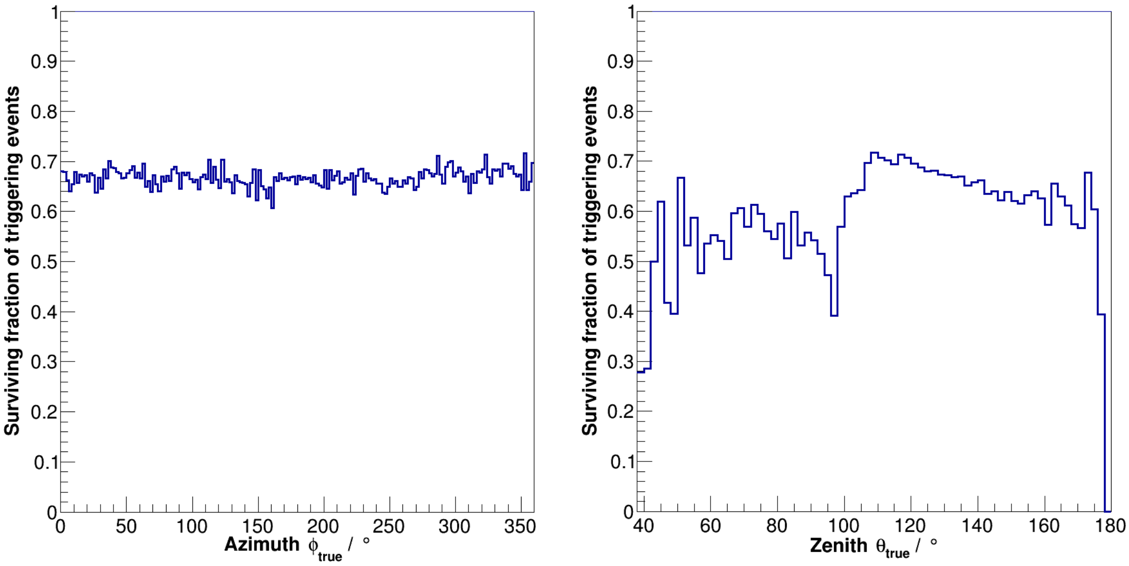
(a) **(b)**

Figure 8.24: *The reconstructed azimuth angle versus SNR compared to the true value for* ***(a)*** *simulated neutrinos and* ***(b)*** *calibration pulsers recorded by ARA02. No reconstruction quality criteria are applied.*

335°) is very weak. In fact these pulsers have hardware problems and are not very useful for the presented comparison.

In conclusion one should note that the possibilities to test the reconstruction on recorded data are limited since we only have a few point sources available.

Figure 8.25 shows the angular dependence of the reconstruction efficiency for simulated data. As for the time sequence quality parameter (Figure 8.9), the azimuthal efficiency is isotropic while the zenith reconstruction shows a maximum right below the horizon.



(a) **(b)**

Figure 8.25: *The angular acceptance after a cut on the residual at* 10^{-4} *and reconstruction quality cuts for simulated neutrinos. The histograms show the fraction of triggered events passing this cut.* ***(a)*** *Azimuthal acceptance,* ***(b)*** *zenith acceptance (limited to statistically relevant regions).*

9

Signal discrimination in the ARA02-ARA03 data

This chapter discusses how the previously described algorithms are applied on the recorded data to separate signal candidates from the background noise. As explained in the last chapter, the algorithms need to reject thermal noise and impulsive background events which originate from the in-ice calibration pulsers and anthropogenic noise sources above the surface of the ice. Apart from neutrinos, there is no source inside the ice expected to produce strong enough radio signals to be visible by the ARA stations. The neutrino interactions themselves are very rare (less than one per cubic kilometer per year). Therefore, the goal is to reject the background to 100% to allow for a neutrino discovery.

The pursued strategy to reject the background noise is in principle performed in two steps. Firstly we ensure with the time sequence algorithm (Section 8.3), some special event rejection and the reconstruction (Section 8.4) to only collect impulsive and properly reconstructed events. Secondly, of the remaining events, those which correlate to known background sources are excluded by angular cuts on their reconstruction.

The analysis is performed as a blind analysis, developed on a 10% subset of the data and, when finalized, applied to the full recorded event sample. In the following, the data subset is denoted as burn sample. Due to the low probability of a neutrino being included in this sample, the data are treated as pure noise. In this way, an estimation of background events passing the cuts can be extrapolated for the full set. The advantage of this method is that the analyzer is prevented from inducing a bias in the event selection, when developing cuts directly on the final data sample. A difficulty in the case of ARA is the fact that the experiment is still in an initial phase and background sources are not

very well explored yet. There is a chance for unexpected background events or detector features to appear only in the full data sample. These might need a special investigation after running the finalized analysis on the complete set of recorded data.
In the following, first, a cut for events with sparking strings is described. Then, the application of the time sequence algorithm in combination with the reconstruction is discussed and finally the angular cuts are explained and a background estimation for the full data is presented.

9.1 String spark rejection

As mentioned above the ARA detector is still in its initial phase and this analysis is the first performed on data of the two ARA stations. We need thus to examine the detector for special features to be able to exclude them from the signal sample.
Such a feature has been found in events like the one shown in Figure 9.1. The signal on string D1 is very strong compared to all other strings. This suggests a source close to this string or in the electronics of the string itself. The non existence of signal on channel 8 (D1TH) shows that not all channels are affected, which renders an external source very unlikely. The reason for such events can probably be found in the electronics of the string.

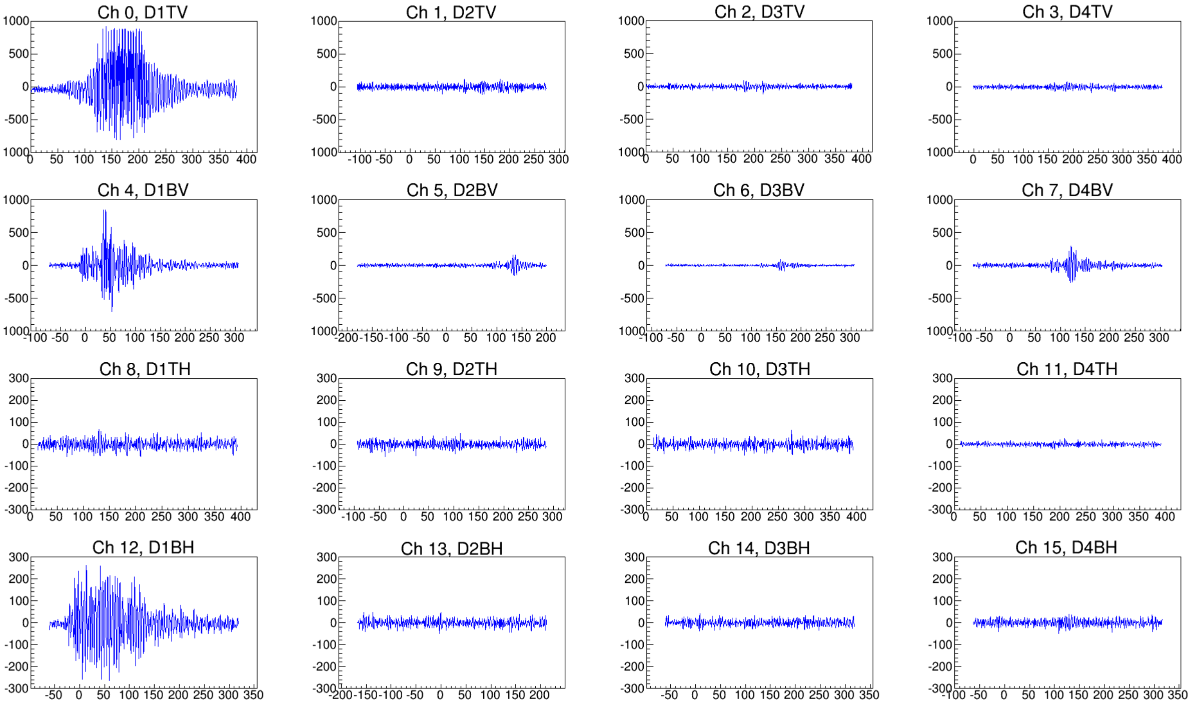
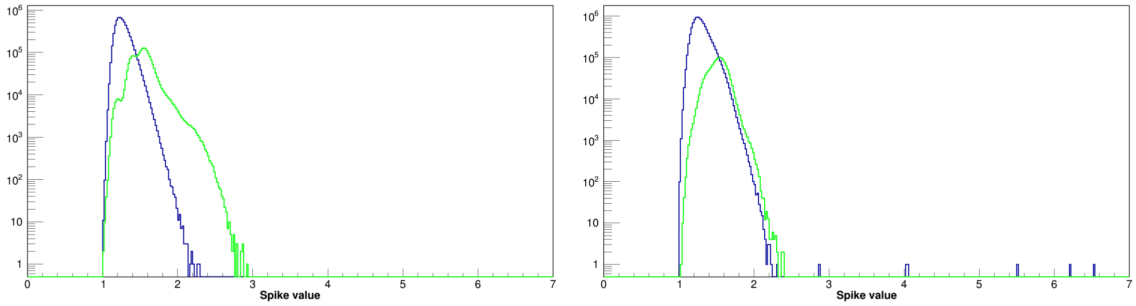


Figure 9.1: *A special event found in the burn sample of ARA03. The waveforms are plotted as mV versus ns Scales of Vpol antennas and Hpol antennas are set equal to visualize the features of this event.*

The event presented in Figure 9.1 is a background event but strongly impulsive. Therefore, the chance for this event to be reconstructed and to end up in the signal sample rises significantly. A special cut is designed to identify the feature of such sparking strings and tag the events as background. In this cut, the peak voltage on each channel is calculated and averaged for each string to M_s. The value of each string is then compared to the average peak of the three other strings and a ratio R_s is calculated to

$$R_s = \frac{M_s}{\frac{1}{3} \cdot \sum_{j \neq s} M_j}. \tag{9.1}$$

The maximum R_s is selected as a spark parameter for the event. Figure 9.2 shows the distributions of R_s for the two ARA stations. In ARA03 five similar events to the one in Figure 9.1 have been found with spark values greater than 2.8. In ARA02 no such event could be identified. The final cut will be performed at 2.6.



(a) **(b)**

Figure 9.2: *The spark value for recorded RF triggers (blue) and calibration pulsers (green) for* ***(a)*** *ARA02 and* ***(b)*** *ARA03. The events similar to Figure 9.1 show up with values above* 2.8.

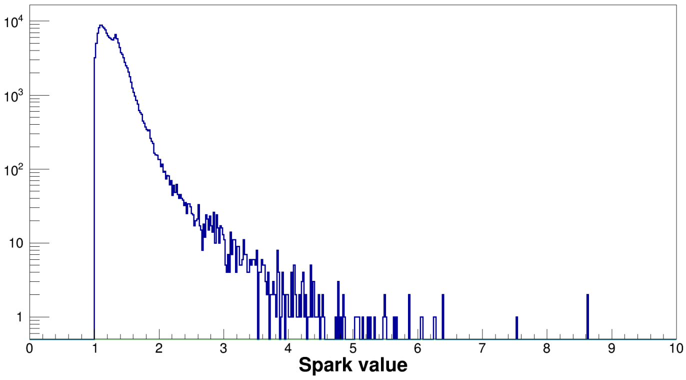


Figure 9.3: *The spark value for simulated neutrinos.*

The string spark rejection affects a small number of calibration pulser events, which contain however mostly saturated waveforms and thus of low quality. Furthermore, about 0.4% of the triggering neutrinos show a spark value above 2.6 which is a well acceptable signal loss for this cut (Figure 9.3). The readout window in the simulation is shorter than the one of the ARA stations. Therefore, we are slightly penalized by signals dropping out of this window. The observed signal loss of 0.4% can consequently be considered as a conservative value.

9.2 Cut adjustments for background rejection

The two main cut parameters in this analysis are the time sequence parameter QP (Section 8.3) and the residual of the reconstruction (Section 8.4). The time sequence quality parameter QP cut ensures to reject thermal noise, while a cut on the residual ensures to select only properly reconstructed events. Figure 9.4 and 9.5 show 2-dimensional distributions of the two parameters for simulated neutrinos and data in the two ARA stations. The plots for the two ARA stations are divided in RF-triggered events and calibration pulser events. From the RF events, the known regions of impulsive background sources are excluded by angular cuts on the reconstructed position. The calibration pulser events are recorded over the full year from various pulsers at different levels of attenuation.

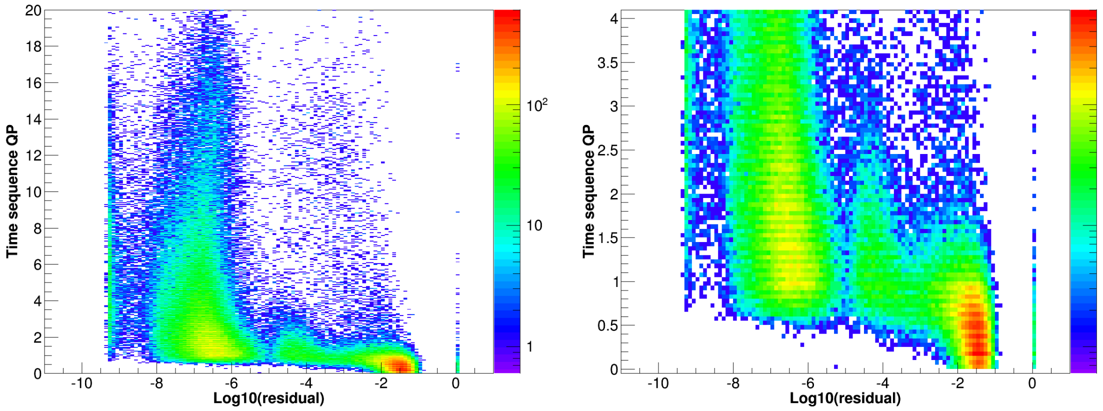
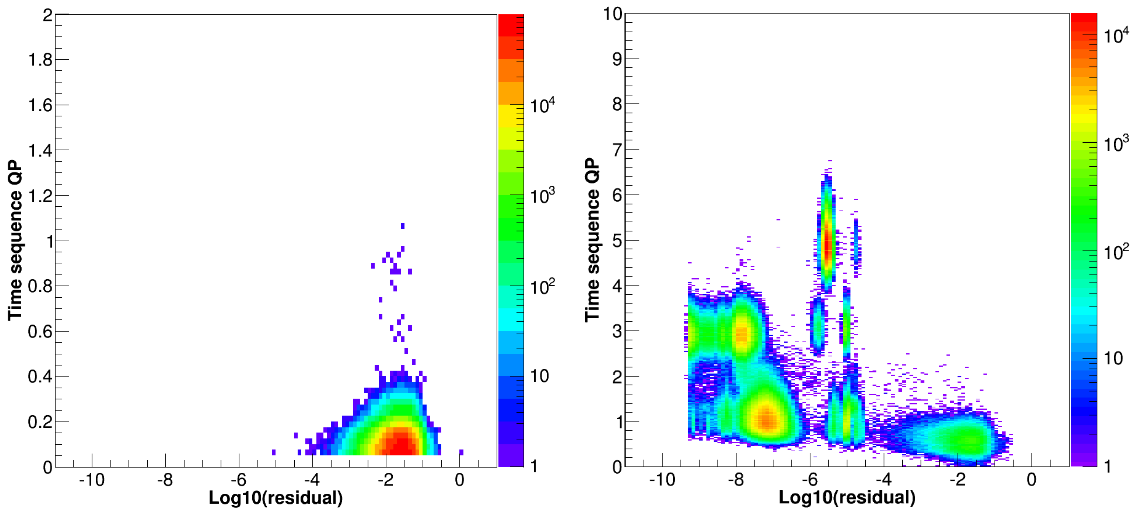
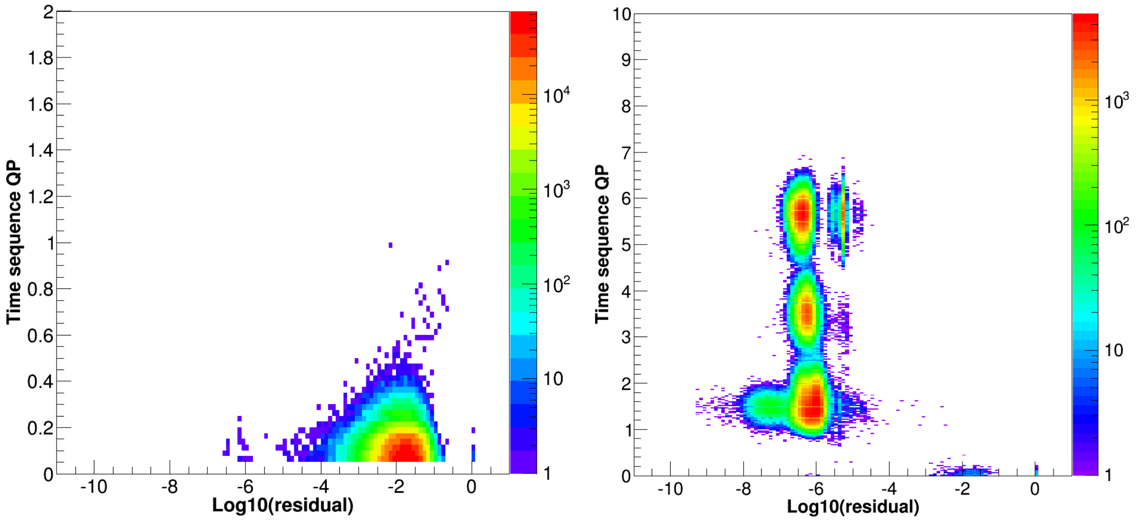


Figure 9.4: ***Left panel:*** *The distribution of QP versus the logarithmic residual of the reconstruction for simulated events.* ***Right panel:*** *A zoom into lower QP regions. The color indicates the event count for each bin.*

The observed behavior is similar in simulation and data. For small QP, the two parameters show a visible level of correlation. Once a certain QP level is reached, the correlation vanishes and the residual remains basically constant with further rising QP. In the case of the calibration pulsers, this can depend on the position and type of the pulser. In ARA02 some pulsers are not functioning properly and their signal output is very low. This is the reason for the relatively high accumulation of events with low QP and large residual. Deviations in the residual between the behavior of simulated events (Figure 9.4) and calibration events (Figure 9.5) can be explained by saturation for high QP



ARA02



ARA03

Figure 9.5: *The distribution of QP versus the logarithmic residual of the reconstruction for ARA02* ***(top)*** *and ARA03* ***(bottom)****. The left panels show the recorded RF-triggers in 2013. The right panels show the events tagged as calibration pulser events from different positions and after differing attenuation. The color indicates the event count for each bin.*

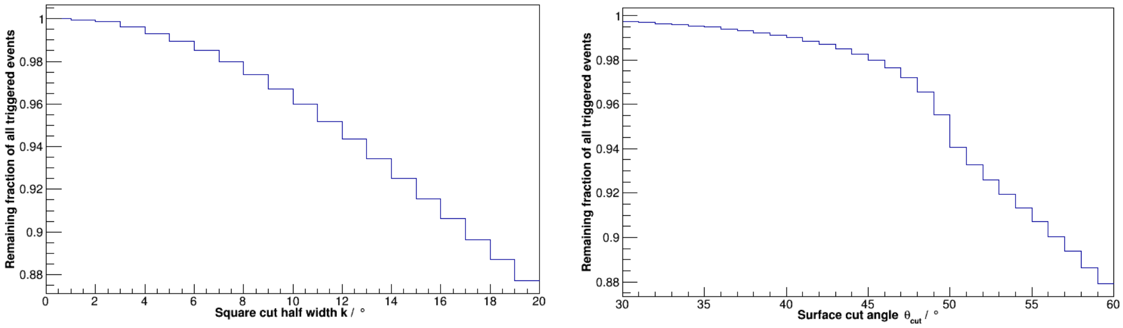
values and by uncertainties on the geometry in the two ARA stations. These differences are strongest in ARA02 but still not dramatic for the cut efficiency.

We can divide the distributions of Figure 9.4 and 9.5 in three regions:

- Top left: well reconstructed impulsive events (high QP, low residual)
- Top right: Impulsive events with failing reconstruction (high QP, high residual)
- Bottom: Thermal noise with random reconstruction and residual (low QP)

The first region holds the interesting events. The other two regions need to be excluded by cuts on the two parameters.

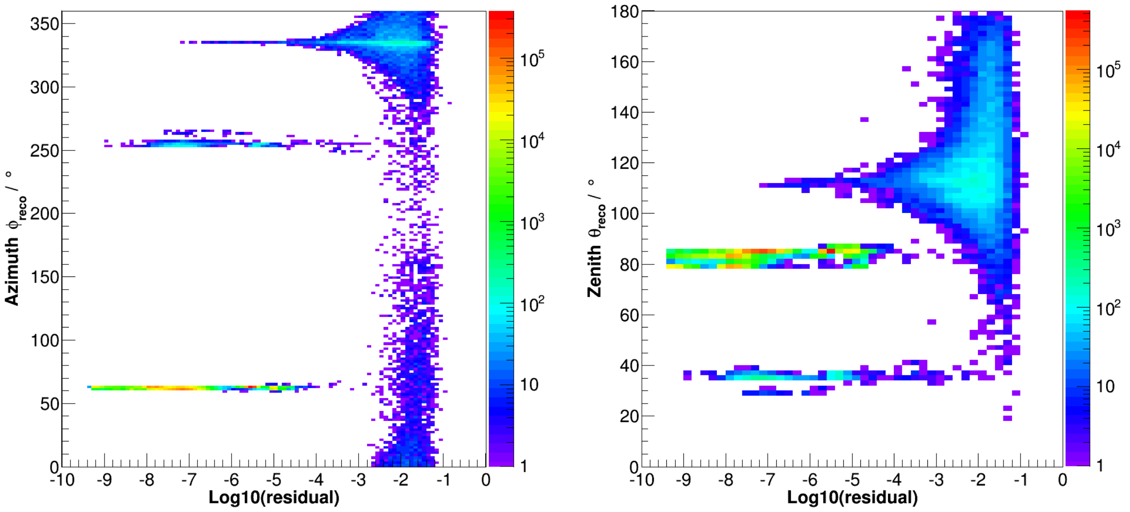
It has been shown that the two cut parameters used are very efficient in separating thermal noise from impulsive signal (Section 8.3) and to guarantee a proper reconstruction (Section 8.4). Under the assumption that the burn sample does not contain neutrino signals, we have to tune the cuts in such a way that all passing events can be correlated to known background sources and excluded by reasonably small angular cuts.
The efficiency loss on neutrino events due to angular cuts is shown in Figure 9.6. The in-ice calibration pulsers will be excluded by rectangular boxes around the zenith and azimuth angle of the pulser position. Figure 9.6a shows the efficiency loss on simulated neutrino events by a square box cut with side lengths of an angle of $2 \cdot k$ around the pulser positions. Even though for ARA03 the zenith angle of all pulsers falls directly into the region of highest neutrino acceptance, the loss is limited and of less than 5% for a square width of $2 \cdot k = 20°$. The surface events are received under the critical angle of 35° or below for signal emitted above the ice. Radio waves, produced directly on top of the ice can be received under an angle of up to roughly 50°. Such events could be generated by vehicles or other human activity and are thus not expected to appear during the austral winter. The efficiency loss for a zenith cut at 50° is about 7% (Figure 9.6b). We can conclude that the efficiency loss due to angular cuts is small if exclusion regions have roughly the above mentioned sizes.



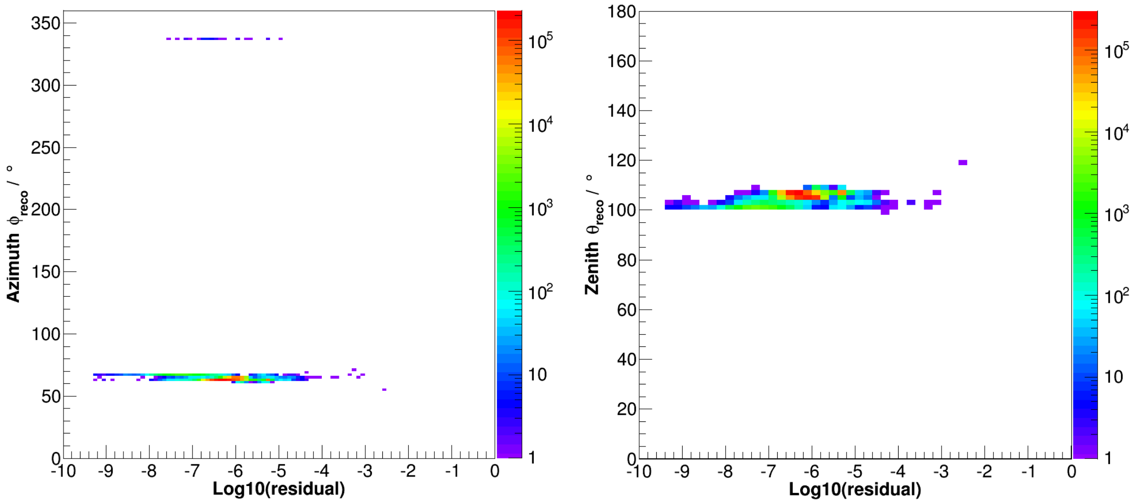
(a) **(b)**

Figure 9.6: ***(a)*** *The fraction of events surviving a square cut of solid angle* $(2k)^2$ *around the calibration pulser directions.* ***(b)*** *The fraction of events passing a surface cut, excluding all events with a reconstructed zenith angle smaller than* θ_{cut}.

The final values for the two main cut parameters have been determined in an iterative process. Cut requirements of $QP > 0.6$ and $residual < 10^{-4}$ have been found to allow for total background rejection by tight angular cuts in both stations, while keeping a high neutrino efficiency. In this optimization, the residual determines the reconstruction precision and plays thus the biggest role for the size of angular exclusion regions. Figure 9.7 illustrates how the angular precision develops with the residual of the recon-



ARA02



ARA03

Figure 9.7: *The reconstructed incoming angles (**Left:** azimuth, **Right:** zenith) for all calibration events on ARA02 and the in-ice calibration events on ARA03, plotted versus the residual. Reconstruction quality cuts and a QP cut at* 0.6 *have been applied.*

struction for different calibration sources in the two ARA stations after a cut on QP. The calibration sources of ARA03 are both working very well and reconstruct generally good. The D5 pulser in ARA02 only produces very weak signals, which are however useful to visualize the precision dependence of the residual cut in more detail. A cut at $residual < 10^{-4}$ allows for angular exclusion regions similar to those mentioned earlier. The final determination of the angular cuts is described in Section 9.3. They are tuned to reject all possible background to an expectation of 0.01 event in the full data sample of each station. A summary of the final cut values is shown in Table 9.1 and 9.2.

It should be mentioned that the cuts have not been optimized in a strictly systematic procedure. Due to the complex connection between all involved parameters and a missing realistic simulation of the full data such optimization was not possible within the scope of this work. The potential improvement, to be gained over the found cut efficiency, can in any case expected to be limited.

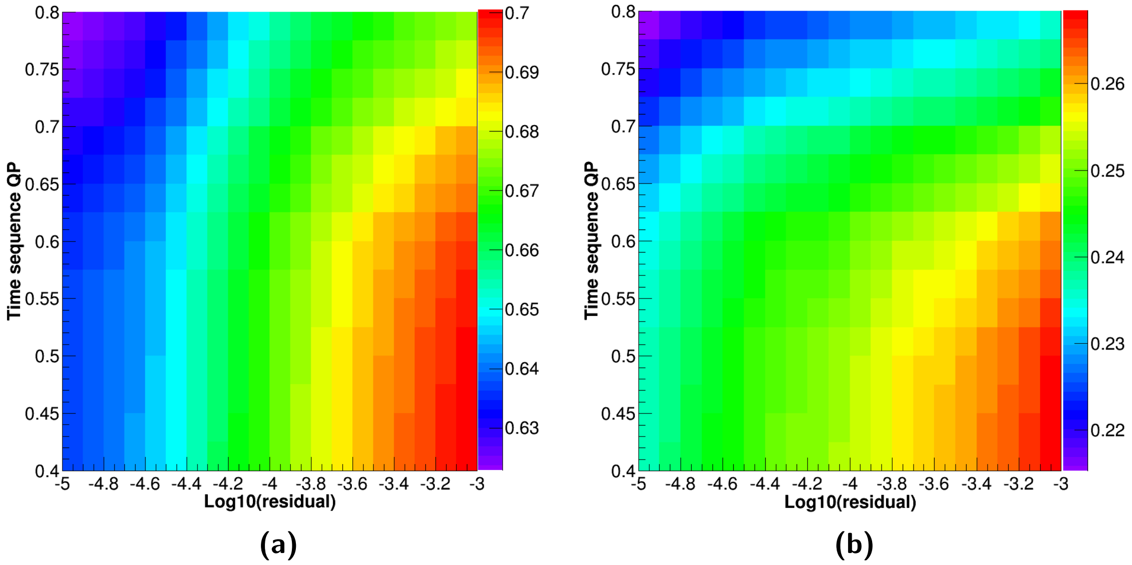


Figure 9.8: *(a) The fraction of events surviving a combination of* QP *and residual cut, shown on a color scale. The indicated efficiency in a bin corresponds to a* QP *and residual cut on the top left border of the bin. (a) All generated neutrinos. (b) Neutrinos with energies below* 10^{17} eV.

The cut efficiencies of various combinations of QP and the residual are displayed in Figure 9.8. Efficiency distributions, zoomed into realistic parameter regions, are plotted for the full simulated neutrino sample and for neutrinos with energies below 10^{17} eV. The low energy events are in particular interesting since they are most strongly affected by the cut levels. The efficiency of the full parameter space in Figure 9.8a is fairly high and only shows a mild change over a wide region of combinations. For low energies the efficiency is lower, but again insensitive to a parameter change within the shown range. Consequently, drastic changes of the cut parameters would be required to achieve a significant improvement in neutrino efficiency. Finally one should note that under the selected cut requirements, of the calibration pulsers shown in Figure 9.5, 95% are kept for ARA02 and more than 99% are kept for ARA03.

9.3 The background estimation

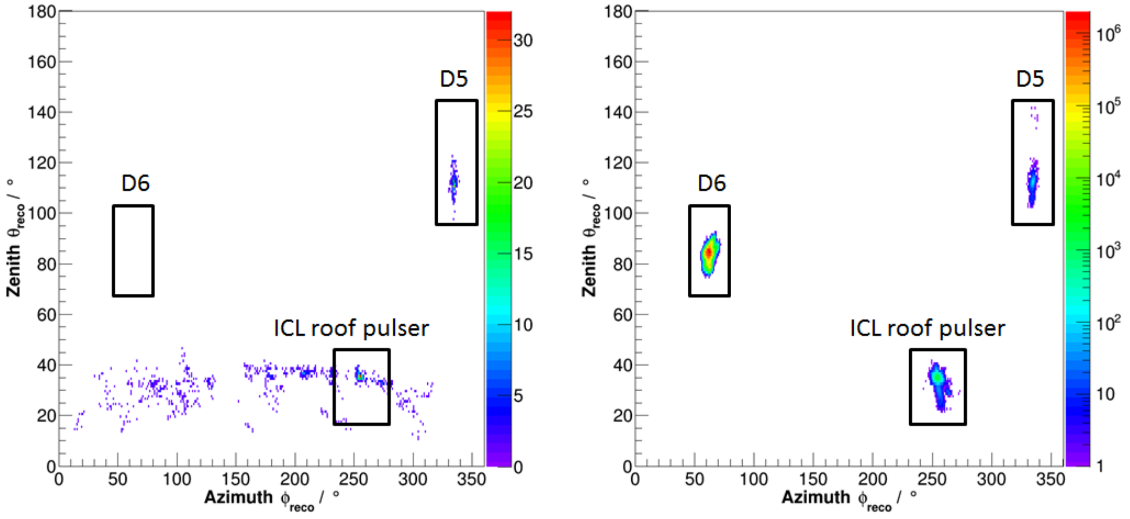
The estimation of the number of expected background events passing the applied analysis cuts is difficult for the two ARA stations. Since simulations don't provide enough events, the background expectation has to be extrapolated from the 10% burn sample, under the assumption that all potential noise sources are represented. Calibration pulsers need to be included in the set of sources since the pulser tagging in the DAQ is not absolutely fail-safe. There is thus a chance that calibration events are marked as normal RF triggers and added to the analyzed data sample.
The number of background events is estimated for each cut parameter separately, after applying all other cuts. We calculate the background for:

- **angular cuts** after applying the QP cut and the residual cut,
- **the residual cut** after applying the QP cut and the angular cuts,
- **the QP cut** after applying the angular cuts and the residual cut.

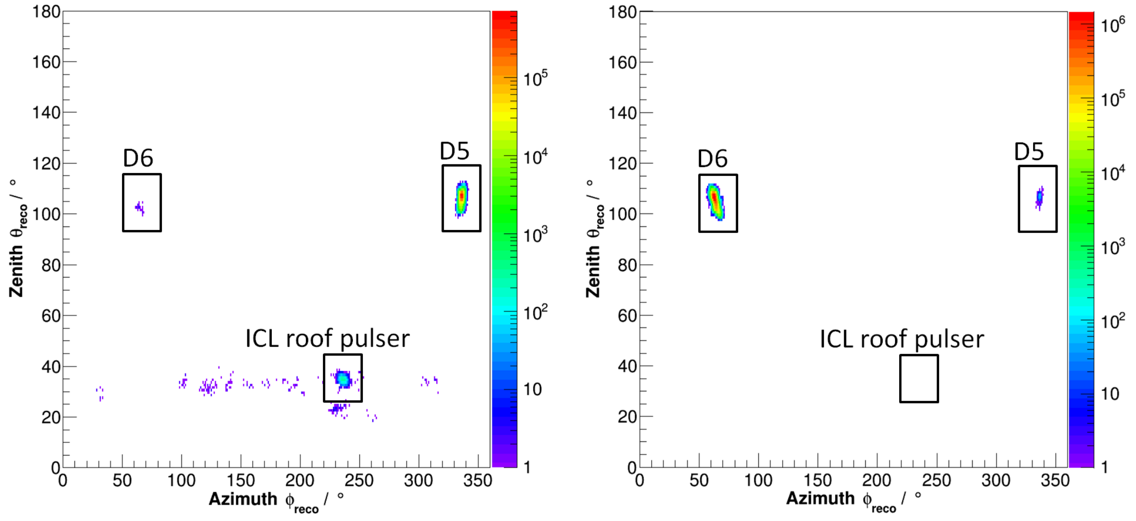
Furthermore, the relation between QP and the residual on thermal noise is investigated and an upper limit on the background contribution is placed.

The background estimation for **angular cuts** depends on the the reconstruction precision and stability, which are determined by uncertainties on the station geometry and on the signal timing. It can be assumed that all systematic uncertainties, i.e. the station geometry and saturation effects are represented in the burn sample. To estimate the behavior on a factor 10 bigger data sample, we add a statistical distortion to the timing pattern of every event with $QP > 0.6$. The distortion is added to every timing correlation as a random number, generated from a Gaussian distribution with $\sigma = 2.0$ ns. Considering the timing precision measured in Section 5.2.6 and Section 8.4.1 this is an over estimation of more than a factor 10. The motivation for a such chosen perturbation is to obtain a conservative estimate on the angular cuts and at the same time to test the robustness of the reconstruction against timing disturbances. Each event is distorted and reconstructed ten times and angular cuts are applied to the resulting data sample.

Sky-maps generated by this approach can be seen in Figure 9.9. One should note that even with big timing distortions, the reconstruction appears to be very robust. The known calibration sources are marked in these maps with black rectangles. One can clearly see that a significant amount of RF events are correlated to the positions of the calibration pulsers and indeed wrongly tagged. The surface events ($\theta_{reco} < 50°$) in the left panels are connected to human activity on top of the ice. There appears to be less activity recorded by ARA03 which can be explained by the larger distance to the station facilities compared to ARA02. On the basis of such sky-maps angular cuts and their background expectation are determined by fitting the tail of a cumulative event



ARA02



ARA03

Figure 9.9: *The angular distribution of all events passing the analysis cuts for ARA02* ***(top)*** *and ARA03* ***(bottom)****. Each event is 10 times disturbed in timing and reconstructed. The events correlated to calibration sources are marked with black frames.* ***Left Panel:*** *Events tagged as RF triggers.* ***Right panel:*** *Events tagged as calibration pulser events.*

distribution for a given angle with a Gaussian function (Figure 9.10):

$$f(x) = Constant \cdot e^{\frac{(x-Mean)^2}{2Sigma^2}} . \tag{9.2}$$

The expected number of passing background events is given by the fit value at the cut position. Figure 9.10 shows the cumulative distribution of events in azimuth for the D5VP pulser in ARA03 with fit. The cut is adjusted to 328° and $f(328°) = 9 \cdot 10^{-10} \pm 4 \cdot 10^{-10}$ events are expected to contribute to the background sample. This sounds like a very conservative cut definition. It is however adapted to the fact that the tail of the distribution is very steep. A summary of all cuts with their background estimation for the two ARA stations is presented in Table 9.1 and 9.2 at the end of this section. One outstanding cut is here the upper zenith cut on D5 pulsers in ARA02. In Figure 9.9 it can be seen that a number of calibration pulser events reconstructed to an unreasonably high zenith angle. To account for such instabilities the cut on this particular ARA02 pulser has to be widened.

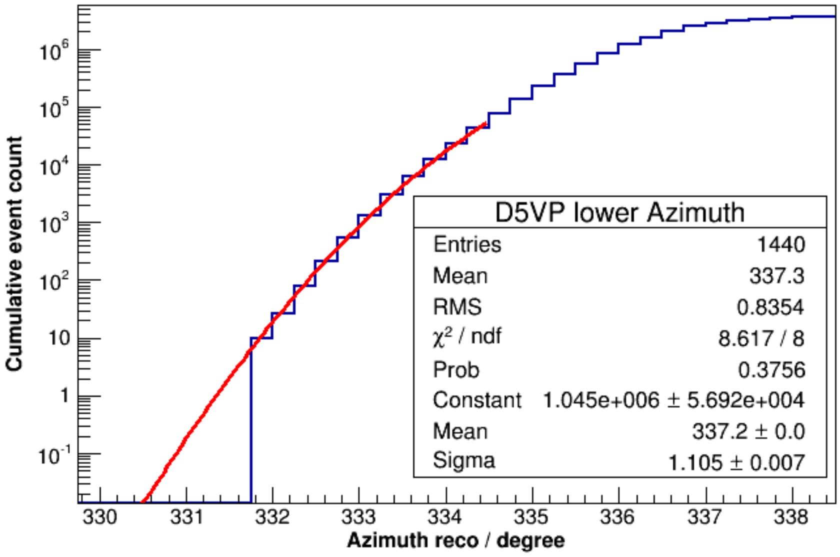


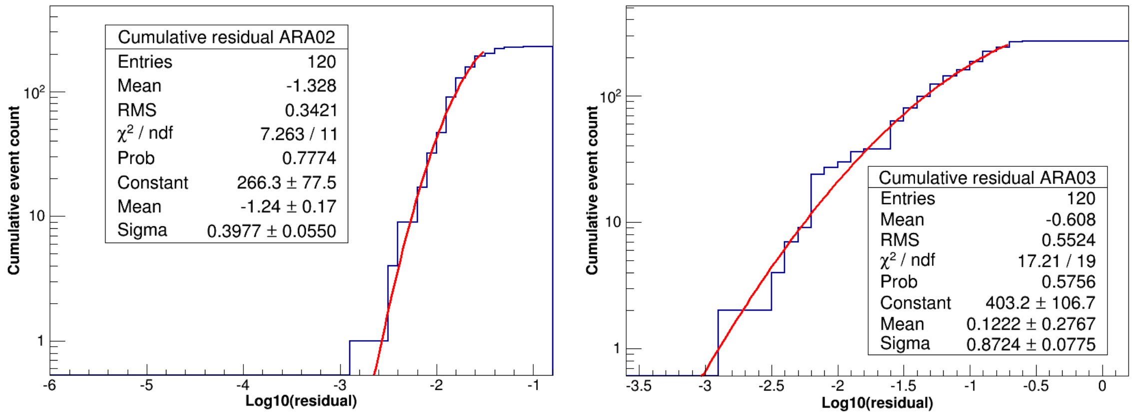
Figure 9.10: *The cumulative distribution of D5VP events in azimuth for ARA03 with Gaussian fit (red line). The resulting fit parameters are used to determine the cut value and the background expectation.*

After the determination of the final angular cuts, we can estimate the number of noise events outside the excluded regions which pass the **residual cut** of 10^{-4}. The calculation is performed in the same way as for the angular cuts. The cumulative residual distributions for ARA02 and ARA03 from the 10 times scrambled events are shown in Figure 9.11 with the appropriate fits. Results on the expected number of background events are shown in Table 9.1 and 9.2.

For the **time sequence quality parameter** QP a background estimation as described above is not possible. The generation of a 10 times larger event sample with realistically simulated disturbances concerning QP is much more complicated than for the reconstruction. For the background estimation we can only use the available burn sample and assume that the full data set will behave accordingly. The cumulative distributions of QP for the two ARA stations, after angular and residual cuts, are shown in Figure 9.12. Their tail can be fitted by an exponential distribution

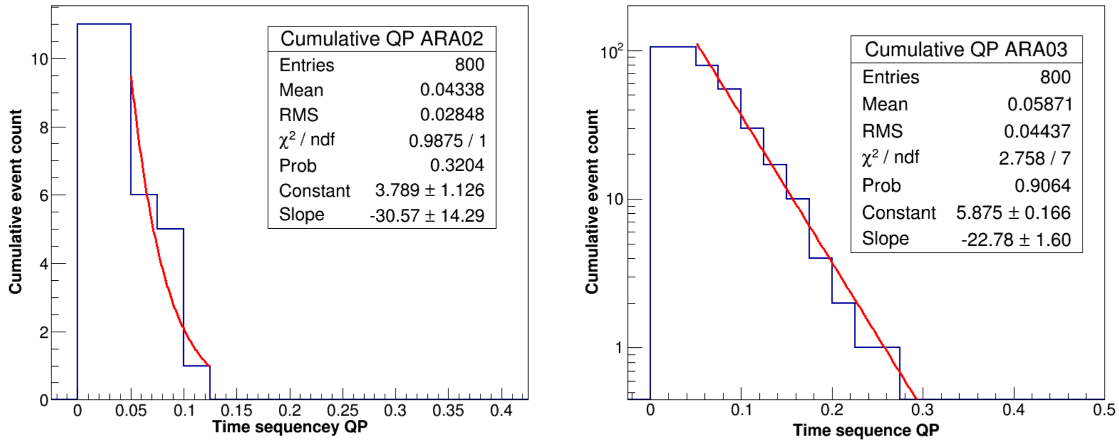
$$f(x) = e^{Constant+Slope \cdot x} \tag{9.3}$$

and the background expectation for the full data sample is calculated as $B_{QP} = f(0.6) \cdot 10$. The results are listed in Table 9.1 and 9.2.



(a) **(b)**

Figure 9.11: *The background estimation from the cumulative distribution of the residual for events reconstructed outside the angular cut regions for* ***(a)*** *ARA02 and* ***(b)*** *ARA03.*



(a) **(b)**

Figure 9.12: *The background estimation from the cumulative distribution of QP for events reconstructed outside the angular cut regions for* ***(a)*** *ARA02 and* ***(b)*** *ARA03.*

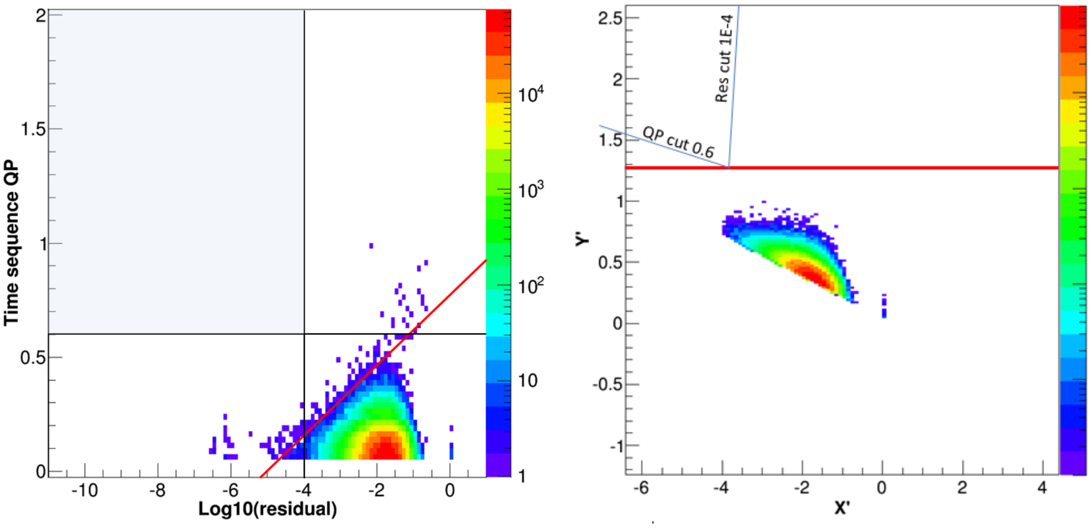
So far, the background has been estimated separately for the different analysis parameters after application of all other cuts. Events, not regarded in the above estimation are thermal noise events with $QP < 0.6$ and $residual > 10^{-4}$ (bottom right of the plots in Figure 9.5). Their behavior concerning the cut combination is difficult to predict. However the border of the two-dimensional distribution towards the cut parameters can be described by an inclined line (Figure 9.13a). The function for this line can be determined to

$$y(x) = y_0 + 0.17 \cdot x, \tag{9.4}$$

with x the residual value and y the QP value. The inclination of 0.17 is the same for both ARA stations. An upper limit on the background from thermal events can be placed by calculating the expected number of events passing a cut inclined by 0.17 through the point $(-4, 0.6)$. This cut is shown as an example for ARA03 in Figure 9.13b. For the estimation, we apply a rotational transformation to the coordinates as follows:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \tag{9.5}$$

The rotation angle $\alpha = \arctan 0.17$ is calculated from the inclination of the line.

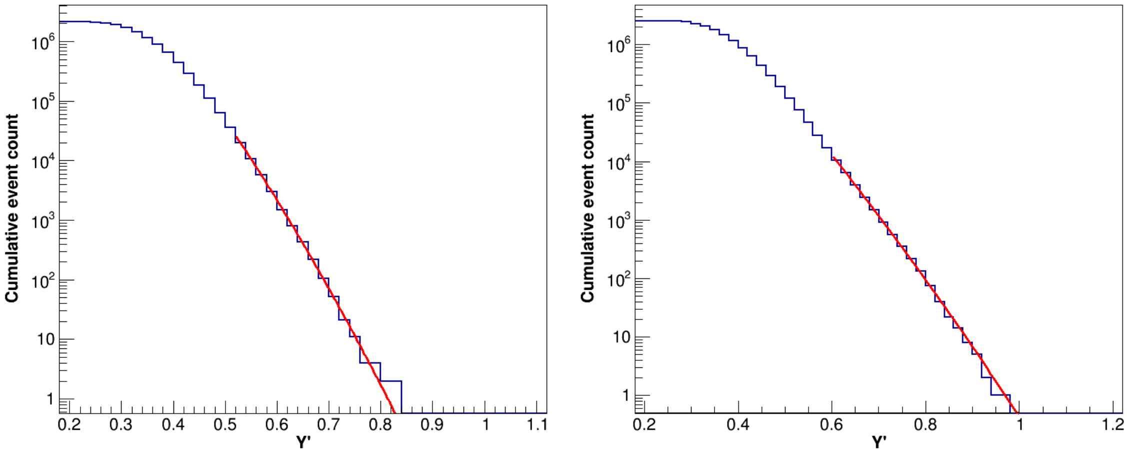


(a) **(b)**

Figure 9.13: ***(a)*** *The distribution of QP versus logarithmic residual for ARA03 (Figure 9.5) with data cuts indicated as black lines. The region of passing events is shaded in light blue. The line fit to the slope of the background distribution is drawn in red.* ***(b)*** *The same distribution after the coordinate transformation from Equation 9.5 with sketched QP and residual cut lines.*

The transformed distribution can then be projected onto the Y'-axis and a fit to the cumulative distribution can be performed as described earlier. The projected cumulative distributions for both ARA stations with fit are shown in Figure 9.14. The numbers obtained for the background are listed as "Thermal" in Table 9.1 and 9.2. As mentioned above, this is an upper limit to the background expectation from thermal noise. Since the contribution is negligible in the total background expectation for both stations, the exact number is not of importance.

In summary, the total background estimation for the full data sample results to 0.009 ± 0.010 for ARA02 and 0.011 ± 0.015 for ARA03.



(a) **(b)**

Figure 9.14: *The distribution from Figure 9.13b b, projected onto the Y'-axis and plotted cumulative.* ***(a)*** *For ARA02 and* ***(b)*** *for ARA03. The cut corresponding to the red line in Figure 9.13b is at* 1.26.

Table 9.1: *The background estimation for ARA02 from all angular cuts, the QP cut, residual cut and the upper limit on the thermal noise background. The errors are calculated via standard Gaussian error propagation.*

Source	cut parameter	value	Background estimate
D6VP	Lower θ	72°	$2 \cdot 10^{-10} \pm 3 \cdot 10^{-10}$
	Upper θ	95°	$6 \cdot 10^{-08} \pm 2 \cdot 10^{-08}$
	Lower ϕ	53°	$2 \cdot 10^{-05} \pm 9 \cdot 10^{-06}$
	Upper ϕ	74°	0.0007 ± 0.0005
D6HP	Lower θ	72°	0.0008 ± 0.001
	Upper θ	95°	$1 \cdot 10^{-09} \pm 4 \cdot 10^{-09}$
	Lower ϕ	53°	$3 \cdot 10^{-07} \pm 2 \cdot 10^{-07}$
	Upper ϕ	74°	0.0008 ± 0.0008
D5VP	Lower θ	88°	0.0009 ± 0.002
	Upper θ	150°	0.001 ± 0.007
	Lower ϕ	326°	0.0003 ± 0.0007
	Upper ϕ	342°	0.0007 ± 0.0008
D5HP	–	–	–
Surface cut	Upper θ	55°	0.003 ± 0.007
Residual	–	$< 10^{-4}$	$9 \cdot 10^{-09} \pm 7 \cdot 10^{-08}$
QP	–	> 0.6	$5 \cdot 10^{-06} \pm 4 \cdot 10^{-05}$
Thermal	–	> 1.26	$< 1 \cdot 10^{-08} \pm 1 \cdot 10^{-08}$
TOTAL			0.009 ± 0.010

Table 9.2: *The background estimation for ARA03 from all angular cuts, the QP cut, residual cut and the upper limit on the thermal noise background. The errors are calculated via standard Gaussian error propagation.*

Source	cut parameter	value	Background estimate
D6VP	Lower θ	93°	$2 \cdot 10^{-18} \pm 4 \cdot 10^{-18}$
	Upper θ	115°	$1 \cdot 10^{-09} \pm 2 \cdot 10^{-09}$
	Lower ϕ	54°	$9 \cdot 10^{-06} \pm 1 \cdot 10^{-05}$
	Upper ϕ	73°	$3 \cdot 10^{-05} \pm 3 \cdot 10^{-05}$
D6HP	Lower θ	93°	$9 \cdot 10^{-15} \pm 6 \cdot 10^{-15}$
	Upper θ	115°	$1 \cdot 10^{-30} \pm 2 \cdot 10^{-29}$
	Lower ϕ	54°	$1 \cdot 10^{-15} \pm 1 \cdot 10^{-15}$
	Upper ϕ	73°	0.0006 ± 0.0003
D5VP	Lower θ	93°	$2 \cdot 10^{-34} \pm 4 \cdot 10^{-34}$
	Upper θ	116°	$2 \cdot 10^{-10} \pm 9 \cdot 10^{-10}$
	Lower ϕ	328°	$9 \cdot 10^{-10} \pm 4 \cdot 10^{-10}$
	Upper ϕ	345°	$6 \cdot 10^{-08} \pm 6.0 \cdot 10^{-10}$
D5HP	–	–	2 events within D5VP cuts
Surface cut	Upper θ	50°	$7 \cdot 10^{-09} \pm 1 \cdot 10^{-08}$
Residual	–	$< 10^{-4}$	0.006 ± 0.014
QP	–	> 0.6	0.004 ± 0.004
Thermal	–	> 1.26	$< 0.0005 \pm 0.0004$
TOTAL			0.011 ± 0.015

10

Data analysis results

In Chapter 8 and 9 the development of cuts for the analysis of data from the stations ARA02 and ARA03 has been described. The cuts have been adjusted on a 10% subset of the data recorded in the year 2013. The current chapter describes the result after applying the developed analysis cuts on the full recorded data set. The events, surviving the QP-cut, the residual cut and all minor quality requirements are shown on sky maps in Figure 10.1 and 10.2. The angular cuts are sketched as black boxes around the rejected area and as a line for the surface cut. One can see that no event passes the final angular selection. There are thus no neutrino candidates found in this data analysis.

In the following section the number of neutrino events which we expect from the presented analysis is calculated and a neutrino flux limit is derived. Furthermore, systematic uncertainties on the neutrino sensitivity are discussed. In the last section of this chapter, the analysis of surviving events is described, which can be examined for interesting features and cross checks on the reconstruction quality.

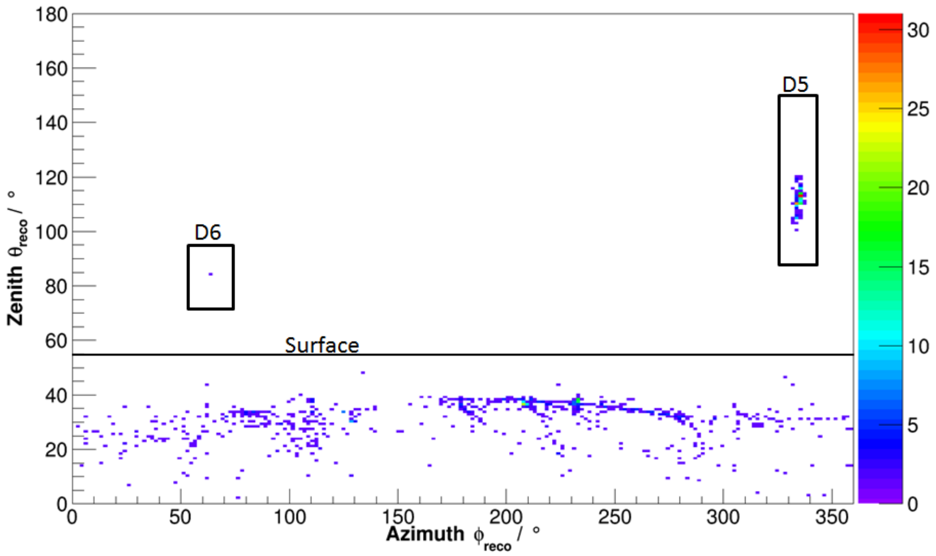


Figure 10.1: *The sky map of all events passing the analysis cuts (QP, residual) in ARA02. The angular cuts are drawn as black rectangles for the in-ice pulsers and as a line for the surface cut. Events with* $\theta < 55°$ *are excluded.*

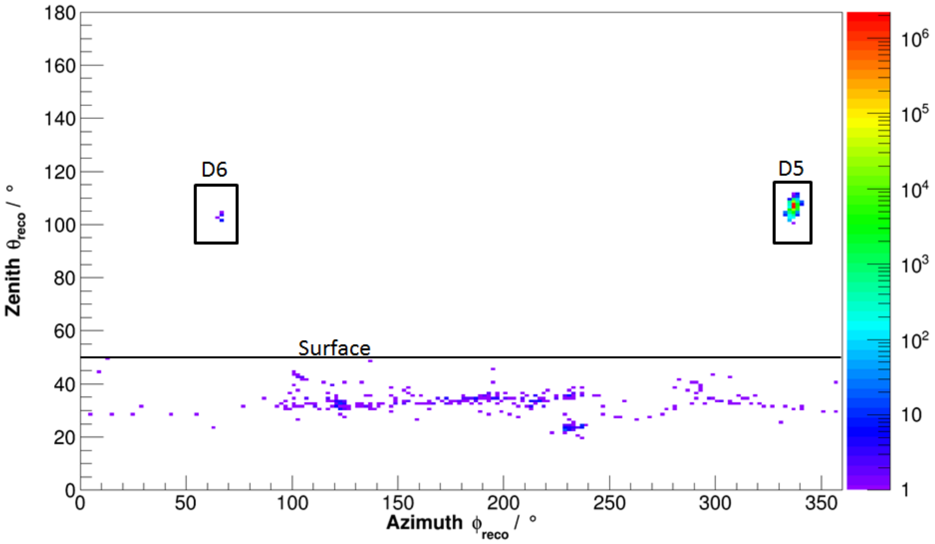


Figure 10.2: *The sky map of all events passing the analysis cuts (QP, residual) in ARA03. The angular cuts are drawn as black rectangles for the in-ice pulsers and as a line for the surface cut. Events with* $\theta < 50°$ *are excluded.*

10.1 Signal estimation and neutrino flux limits

The expected number of neutrino events passing the analysis cuts is calculated using the simulation described in Chapter 7. In this chapter the effective volume and effective area have been calculated for all triggering events. Such an estimation can be done accordingly for events passing the analysis cuts. The resulting V_{eff} and A_{eff} are shown in Figure 10.3 and Figure 10.4 respectively, in comparison to the trigger level result. It becomes visible in both figures that the analysis efficiency is low (roughly 10%) for neutrino energies of 10^{16} eV, but increases with rising energy. At 10^{18} eV, the point of strongest neutrino flux expectations, the analysis efficiency goes up to roughly 60% which is a promising result.

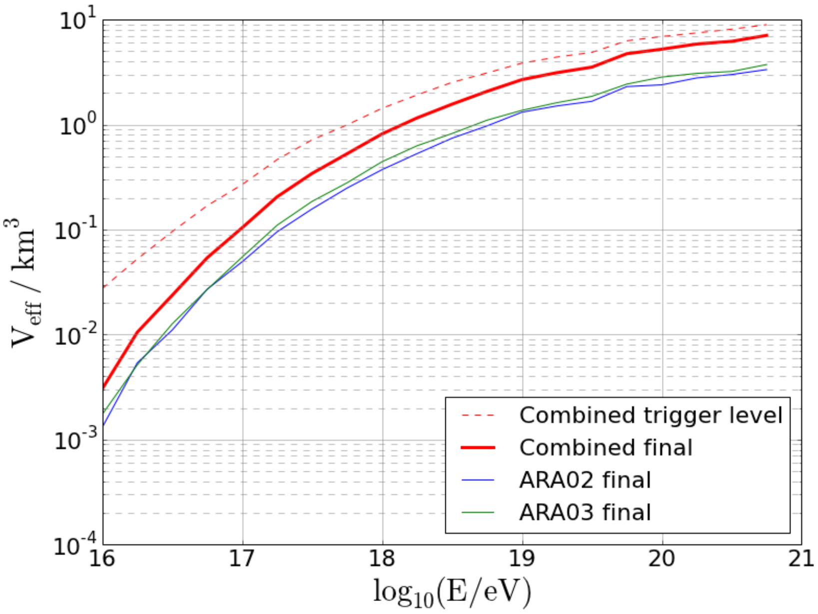


Figure 10.3: *The effective volume of the two ARA stations at trigger level and after applying all analysis cuts, plotted versus neutrino energy.*

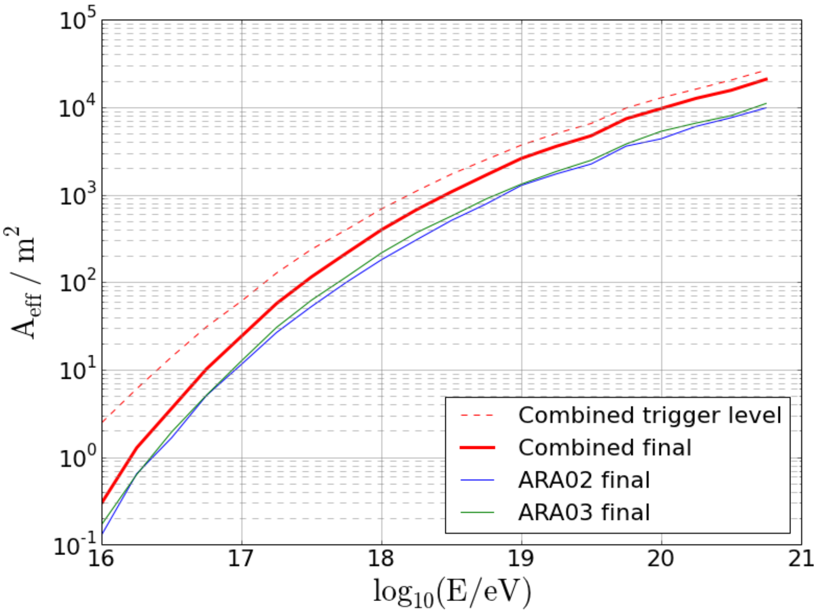


Figure 10.4: *The effective area of the two ARA stations at trigger level and after applying all analysis cuts, plotted versus neutrino energy.*

From the effective areas and the live time of each detector (Section 4.2), the number of expected neutrinos can be calculated using Equation 7.5. Note that both live times are corrected by a factor of 0.9, since the 10% burn sample, used for the development of analysis cuts, has to be excluded from the full data set. The input flux we assume for the calculation is the moderately conservative model developed in [52].

The resulting number of expected neutrinos for the used model and for the presented data analysis is

$$N_\nu = 0.103 \pm 0.002(stat). \tag{10.1}$$

This is consistent with the result of no found signal candidate. The chance to discover a neutrino will grow with a higher live time and more detector stations deployed in the future.

With the obtained results we can calculate a Poisson 90% confidence limit on the differential neutrino flux according to [106] for zero background and zero signal expectation:

$$E^2 F_{up}(E) = E^2 S(E) \cdot \frac{2.44}{dE} = ES(E) \cdot \frac{2.44}{0.5 \cdot ln(10)}, \tag{10.2}$$

with $S(E)$ the energy dependent detector sensitivity calculated in Equation 7.4. For the binning of this limit, half decade energy bins on a logarithmic scale have been chosen. The factor $1/(0.5 \cdot ln(10))$ is applied to account for the relation

$$dE = E \cdot ln(10) dLog_{10}(E), \tag{10.3}$$

under the assumption of constant values for $E^2F(E)$ within one bin. The resulting limit from the two station data analysis is shown in Figure 10.5 with systematic errors (Section 10.2). This limit does not add constraints to the predicted neutrino flux and is not in the range of the currently best limit, set by IceCube. However this is expected to improve with more ARA stations deployed, better analysis techniques and a longer time of operation. The projected sensitivity after three years of operation for the full ARA37 array is shown as a blue line in Figure 10.5. With such a sensitivity, according to the current flux predictions a discovery of GZK neutrinos can be expected.

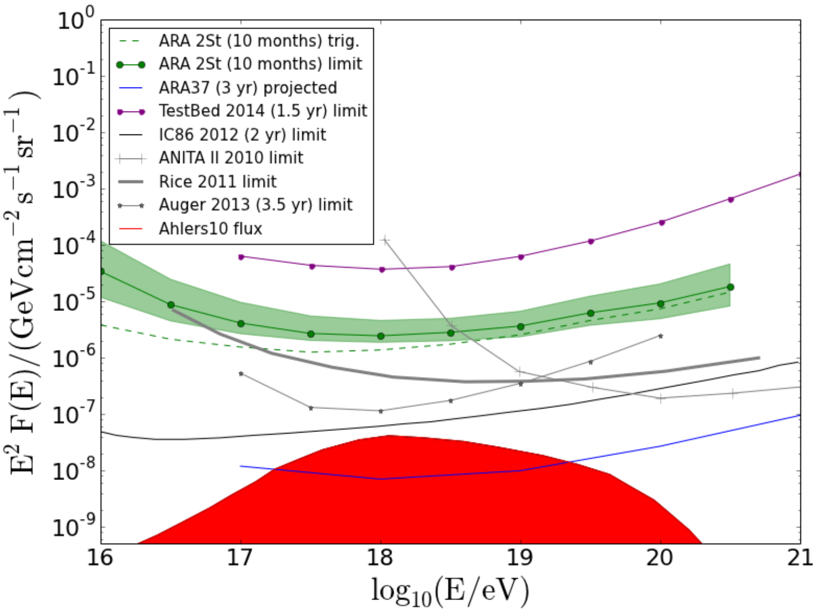


Figure 10.5: *The energy dependent neutrino limit from the 10 months data analysis of the two ARA stations with systematic errors (green shaded band, derived in Section 10.2) compared to current limits and future sensitivities (Section 4.3).*

10.2 Systematic errors

There are various sources of systematic errors in this analysis. They are connected to our understanding of the detector and to the limited knowledge about interaction cross sections, the Askaryan radiation model and ice properties. This section summarizes how influences of the various error sources on the effective area of the detector are estimated. Further details on this estimation can be found in Appendix B.

- The error on the interaction cross section of neutrinos in matter is derived in [101]. To estimate the influence on the here presented results, simulated data is produced using an upper and lower limit value of the cross section. The difference to the average result is used as a systematic error.

- The error on the Askaryan emission model could not be estimated to a satisfactory level within the frame of this thesis. Detailed understanding of the cascade development and the emission under various conditions is needed to properly simulate uncertainties and calculate systematic errors for ARA. This uncertainty thus needs to be determined in future investigations.

- The Askaryan signal is propagated through the ice using a parametrization of the attenuation length measured with deep ice pulsers and the ARA Testbed. An older parametrization is based on measurements of waves reflected off of the bedrock under the ice [78]. As described above, a systematic error is derived from the difference of the two models found in the simulation. In this case, the model used in the presented analysis is found to be more optimistic in the comparison. However, it is based on more detailed measurements and therefore assumed to be more realistic.

- The antenna response and signal chain has been modeled from simulations and lab measurements. A comparison between a response measurement in air and the used simulation shows an average difference of roughly 30% in the antenna frequency response. Since the presented analysis is purely based on time domain signals, apart from the cross correlation algorithm, the expected error is calculated by changing the overall signal response amplitude. The gain of the signal chain has been measured to a precision of 20%. The full response level is thus changed by 36% to account for the antenna and signal chain.
- The systematic errors in station geometry and signal shape in terms of analysis efficiency can be estimated by the SNR comparisons for calibration pulser waveforms presented in Section 8.3.1 and 8.4.5. For each energy bin the analysis efficiency for simulated data and for calibration pulsers is compared as function of the event SNR. The average difference is calculated for an $SNR > 2$ using the number of events in each SNR bin as weights (Appendix B). The efficiency error due to angular cuts is assumed to be negligible since it has been shown that the reconstruction is very robust against timing and geometrical errors (Section 8.4.5 and 9.3).

Table 10.1: *The systematic error on signal expectation from different sources.*

Source	positive error	negative error
Cross section	0.029(28%)	0.022(21%)
Attenuation	–	0.010(9%)
Signal Chain	0.022(22%)	0.036(35%)
Analysis efficiency	0.018(17%)	0.018(17%)
TOTAL	0.042(41%)	0.049(47%)

The uncertainties on the number of neutrinos, obtained from the above described estimations are shown separately in Table 10.1. The total error is calculated from the combination of all uncertainties to the total error on the effective area. One can note that the dominant errors are due to the uncertainties on theoretical calculations of the cross section and due to the assumed error on the signal chain. The error due to the cross section is caused by the direct dependency of the effective area on the cross section through the interaction length (Appendix B). An improvement in the theoretical estimations can reduce this error significantly. The error on the signal chain can be improved for future analysis with detailed calibration measurements of the antennas and

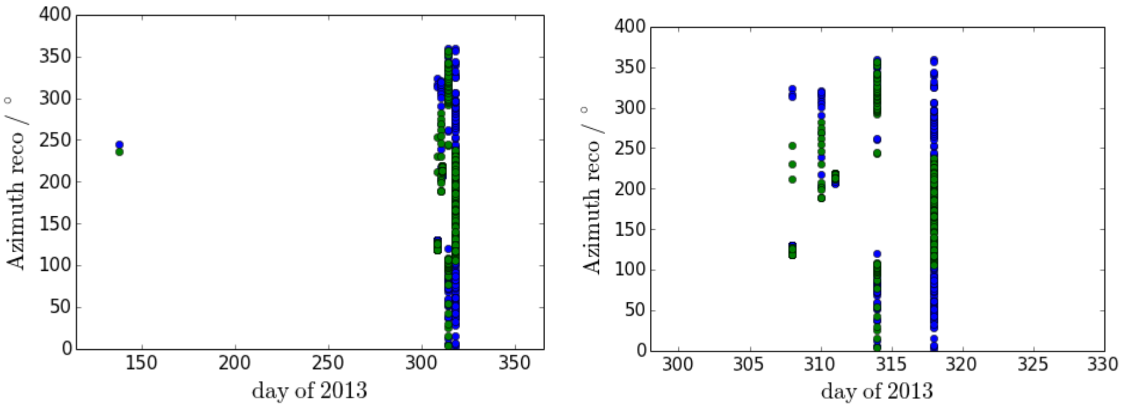
other electronic components, which falls beyond the scope of the presented work.
One can see in Figure 10.5 that the systematic errors in the energy range below 10^{17} eV and above 10^{20} eV rise strongly compared to central energy regions. This does not affect the result of the expected number of neutrinos significantly since such energy regions only contribute by less than 1% to the total value. The final number of neutrino events in this analysis, expected from the flux calculated in [52], can be quoted as

$$N_\nu = 0.10 \pm 0.002\ (stat) + 0.04 - 0.05\ (sys). \tag{10.4}$$

10.3 Coincident event analysis

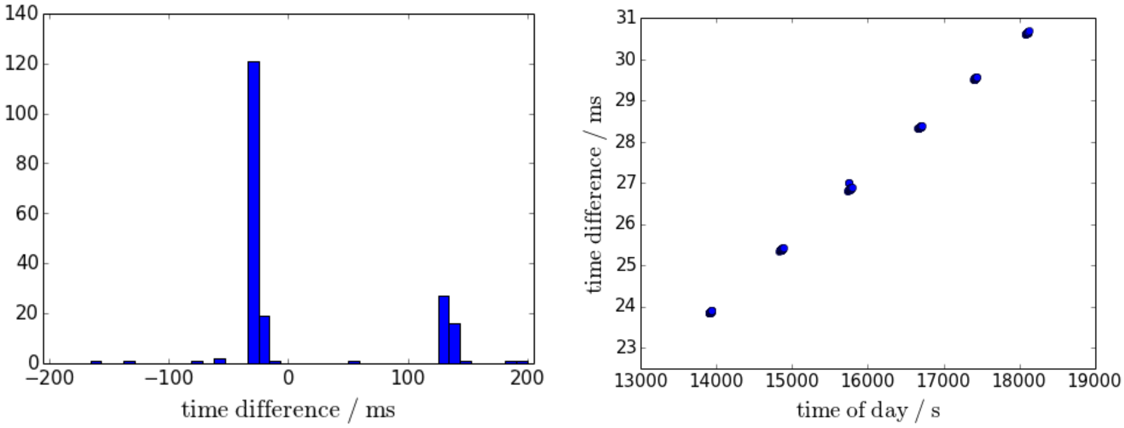
In the full data sample only events coming from the surface and from the in-ice pulsers have survived the analysis cuts. In this section we investigate on surface events which are coincident between the two ARA stations. Using such coincidences we can cross check the reconstruction results and selected station performance parameters. The research, presented in this section, has been performed in collaboration with Prof. David Seckel (University of Delaware).
Coincident events are determined using the trigger time in the two stations. This time is provided on a second base from GPS timing and on a base of 10 ns by station-internal 100 MHz clocks. The used clocks are high performance rubidium clocks, nominally disciplined by the GPS pulse per second. When disciplined, these clocks are expected to be synchronized on a sub-nanosecond level. Events are tagged as coincident if they trigger both stations within a time window of 7 μs, which corresponds to the station separation by 2 km. Unfortunately it appears that the two station clocks were not perfectly synchronized during the year 2013. They show a systematic drift against each other which needs to be determined and accounted for in the trigger times. Therefore, we check events appearing in the same GPS second with the goal to identify coincidences and to perform a clock drift correction. The azimuth reconstruction for such events is shown in Figure 10.6 versus the day of 2013. One event has been found during the Austral winter (day 138), which reconstructs in the direction of the IceCube Laboratory (ICL) and the South Pole for both stations. The clock drift can not be determined with such an isolated event, which is why we can not claim a coincidence.
During the Austral summer instead, the period of human activity around the station facilities, many events appeared within a short time frame. Figure 10.7a displays a histogram of trigger time differences for events occurring in the same GPS second on day 318. When zooming into the main peak at roughly −28 ms, we discover a systematic relation between ARA02 and ARA03 (Figure 10.7b). As expected, a regular drift is visible in the time difference of the events. This drift can be determined to 1.6 μs/s. The systematic behavior gives us confidence that the shown events are indeed coincident events on the two stations.

One can note that the events appear in groups, separated by about 15 minutes and



(a) **(b)**

Figure 10.6: *The azimuth reconstruction of events triggering both stations (blue: ARA02, green: ARA03) in the same second versus the day of 2013.* ***(a)*** *for the full year,* ***(b)*** *for the times of human activity during the austral summer.*



(a) **(b)**

Figure 10.7: ***(a)*** *A histogram of the time difference between events on ARA02 and ARA03 on day 318.* ***(b)*** *The time differences of the main peak at around* −28 ms *versus the second of day 318.*

spread over a time of 1 hour. We select one particularly interesting group and perform a parallax reconstruction of the events using the reconstructed azimuth angles from the two stations (Figure 10.8a). These azimuth angles are now corrected by the systematic offset found on the ICL roof pulser position (Table 6.2). The found events form a smooth track which crosses ARA02. We can thus conclude that we are seeing regular radio emissions from a moving object. There is only one event reconstructed far off track. Following the general curve, this event would be expected to reconstruct exactly above ARA02. A small error on the azimuth reconstruction of ARA03 causes the point to be significantly of track. When adding a systematic correction of 1.5° to the reconstructed angle from ARA03 all events line up nicely (Figure 10.8b). From the difference in timing

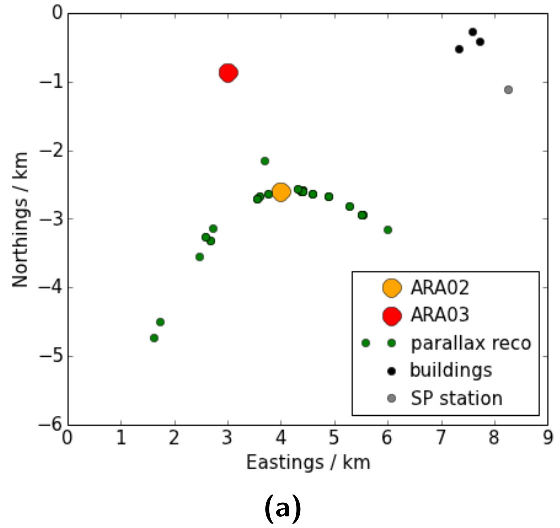
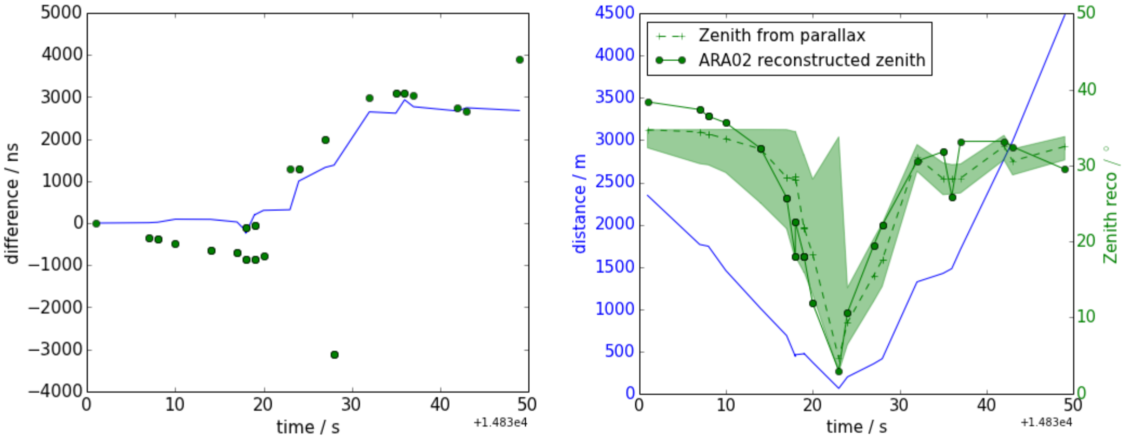


Figure 10.8: *A top view of the two ARA stations with the parallax reconstruction of coincident events (green dots) happening on day 318 at 04:00 am* ***(a)*** *for the normal reconstruction,* ***(b)*** *after adding a systematic azimuth offset of* 1.5° *to the azimuth from ARA03.*

and position of two points one can calculate the speed of the emission source to roughly 400 km/h on average. This corresponds to the speed of an airplane crossing the South Pole station. The track itself shows by its smoothness that the reconstruction algorithm we are using gives us reasonable results.

An additional cross check of the reconstruction algorithm is based on the trigger time difference between the two stations after clock drift correction (Figure 10.9a). In comparison to this, we plot the time difference calculated from the parallax position and height, which we derive using the zenith angle reconstruction from ARA03. The ice-air interface (Section 6.3.3) has been taken into account for this calculation. One can see in Figure 10.9a that the general trend is reproduced nicely. Such a cross check is particularly interesting since the trigger time is technically unrelated to the reconstruction.
In Figure 10.9b the reconstructed zenith angle from ARA02 is compared to the zenith estimation from the parallax position. The height above the ice is again calculated using ARA03. A good agreement can be observed within the shown error band and deviations are usually smaller than 5°. The shown error is derived from the statistical errors on the reconstruction in ARA02 and ARA03 presented in Table 6.2 for the ICL roof pulser. One should note that Figure 10.9b shows the reconstruction of ARA02 which is generally of lower quality than the one from ARA03.

A further interesting feature of the vertical event above ARA02 is the expected total horizontal polarization. The main frequency content of the observed events is recorded



(a) **(b)**

Figure 10.9: ***(a)*** *The trigger time difference between ARA02 and ARA03 after clock drift correction compared to the expected time difference from parallax reconstruction.* ***(b)*** *The zenith angle reconstructed by ARA02 (green solid) and from the parallax reconstruction (green dashed), assuming a height of* 500 m*. The green shaded area indicates the error, resulting from azimuth reconstruction errors of the two stations. The blue line indicates the horizontal distance of the vertex to ARA02 from parallax reconstruction.*

below 200 MHz in the ARA stations. Due to the low sensitivity of the Hpol antennas in this frequency range, most signal is expected to appear in the Vpol antennas independent of the actual signal polarization. This is observed in a moderately inclined event shown in Figure 10.10. The highly inclined event reconstructed to occur directly above ARA02, as expected, shows a strong enhancement of the Hpol component compared to the Vpol signals (Figure 10.11).

From the presented cross checks we can conclude that reconstructions are possible with the ARA stations and show consistent results between the two stations and in a cross check with a technically unrelated parameter.
When examining other event tracks a similar behavior can be observed. The reconstructed tracks are relatively smooth and it can be concluded that the used cut algorithms ensure very well the selection of non-thermal and well reconstructable events in the recorded data. The current ARA stations are thus well capable of the identification and localization of radio sources like neutrino-induced cascades.

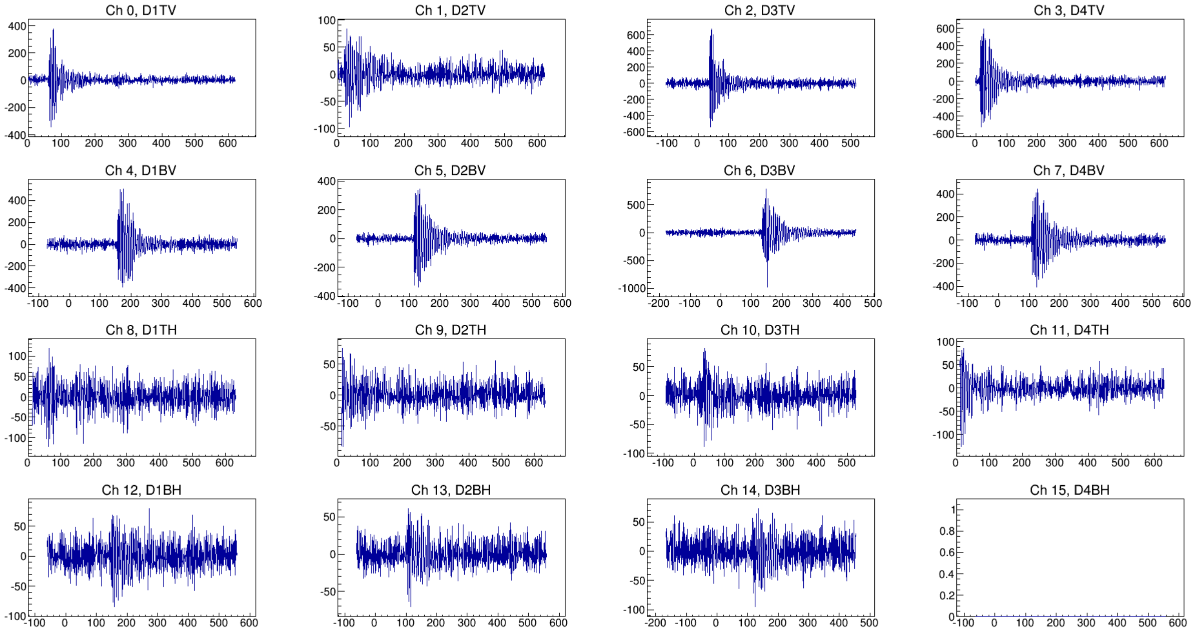


Figure 10.10: *A example event from the beginning of the track as recorded by ARA02. The event is reconstructed to* $(5722\ \mathrm{m}, -2983\ \mathrm{m}, 212\ \mathrm{m})$ *in Figure B.2.*

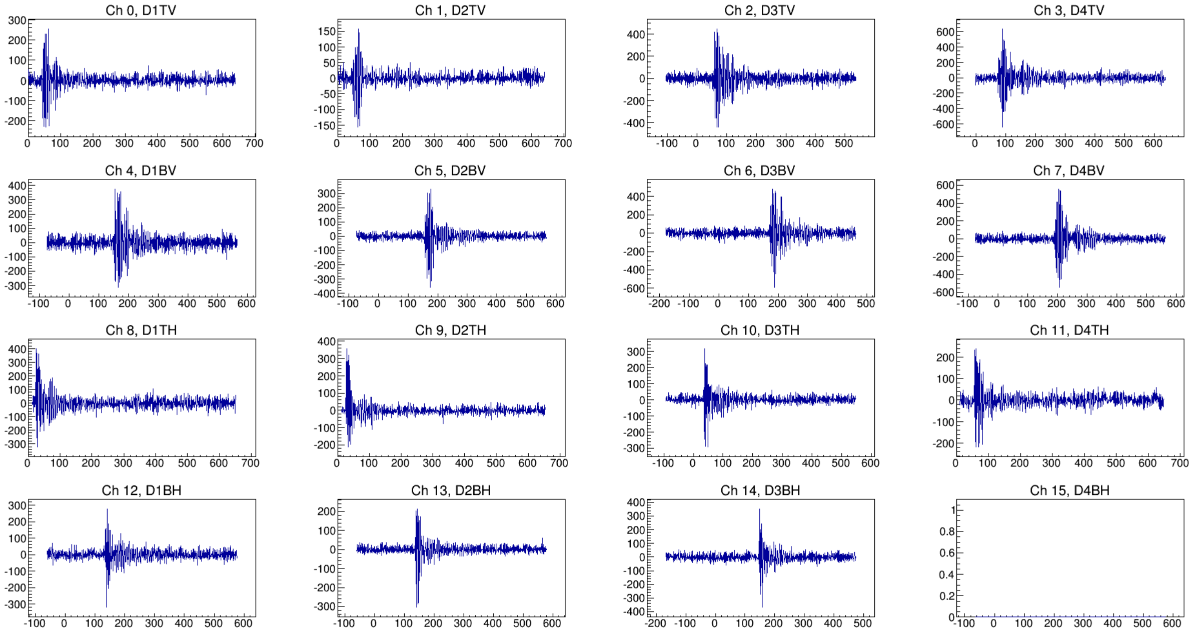


Figure 10.11: *The event occurring vertically above ARA02. The waveforms are plotted in mV versus ns. The event is reconstructed to* $(3975\ \mathrm{m}, -2541\ \mathrm{m}, 322\ \mathrm{m})$ *in Figure B.2.*

11

Summary and outlook

The aim of this work was to prepare and optimize the currently deployed parts of the **A**skaryan **R**adio **A**rray (ARA) for GZK neutrino detection. Various components of the hardware have been tested, programmed and calibrated and a data analysis has been performed to show that a neutrino search is possible and that a good sensitivity can be expected for the full ARA37 detector.
In the first part of this chapter the work presented in this thesis is summarized and possible improvements to the detector performance and the analysis techniques are discussed. In the second part investigations towards a new detection method using RADAR reflections off of neutrino-induced cascades are described.

11.1 Summary and discussion

In the first part of the thesis, theoretical discussions of the GZK mechanism are summarized to explain the interest in the detection of GZK neutrinos in the context of high energy astroparticle physics. This interest is based on the potential information about ultra-high energy cosmic rays which the neutrinos can deliver. Furhtermore, the emission of coherent radio waves from neutrino-induced cascades is described, which allows to measure GZK neutrinos cost-efficiently with radio detectors like ARA.
Subsequently, we describe calibrations that have been performed in order to analyze the recorded data properly. The used digitizer chip in ARA, the IRS2, has been calibrated in timing and voltage conversion gain. Cross checks show that after this calibration signal timing can be determined with a precision of roughly 100 ps. Moreover, the station geometry and cable delays have been calibrated. This calibration is proven to add a significant improvement to radio vertex reconstructions.

The data analysis described in this thesis is based on data recorded in 2013 within 10 months of operation by the first two ARA stations. In the scope of this analysis, a thermal noise filter and a reconstruction algorithm have been developed to perform a search for neutrino induced cascades in South Pole ice. The noise filter provides an efficient separation of impulsive radio signals from thermal background noise. With the developed reconstruction algorithm, incoming angles for radio wavefronts can be determined with a precision of a few degrees. Tests have been performed which proof the robustness of the algorithm against strong perturbations in timing.
The developed methods have been used to extract neutrino induced radio signals from thermal and anthropogenic background. They have been applied in a way to cut the expected background in the data set to 0.009 ± 0.010 events in ARA02 and 0.011 ± 0.015 events in ARA03. From simulations, the efficiency of the analysis on triggering neutrinos and an effective area have been calculated to estimate the neutrino sensitivity and the expected number of signal events. No neutrino candidate has been found within this analysis which is consistent with the expectation of 0.10 neutrinos from a moderately conservative neutrino flux estimation [52]. A limit on the differential neutrino flux has been calculated which does not add a constraint to the current best limit but is a very promising result considering the current size of the detector, the building costs and the relatively short time of operation (Figure 10.5).

There are a few aspects which have been improved about the ARA detector during the past year and which will positively affect future analyses:

- The remoteness of the detector stations and limited satellite connectivity rendered a regular monitoring of the station operation difficult. Down times of several days, due to problems like power cuts and DAQ crashes, caused a loss of roughly 25% operation time. Since December 2013 new monitoring techniques have been implemented which allow us to react on a time scale of a few hours to debug failures in the detector. As a result, the live time increased significantly.
- From experience gained in the detector understanding, changes in the DAQ software and detector configuration have been applied to optimize the data acquisition and to prevent problems, as encountered in the performed analysis. To name one example: the triggering and data acquisition have been adapted to discrepancies found in the cable delays. Such delays caused an efficiency loss of roughly 7% in the analysis of data from the year 2013.

In addition to problems in the detector operation, weaknesses in the data analysis should be pointed out which cost efficiency and can probably be improved significantly in the future:

- The channel selection in the reconstruction algorithm is performed in a relatively simple way. The result is a relatively high chance of noise selection which causes bad reconstructions and a rise of the reconstruction residual above the cut value, which reduces the efficiency. The usage of more sophisticated algorithms for the channel pair selection can improve this point for the future.

- Furthermore, the reconstruction algorithm can be improved in its handling of reflected signals and ray-tracing effects. The sensitivity gain in this case will be mainly due to the applied surface cut. Between 5% and 10% of the events are currently lost because of misreconstructions to low zenith angles (Figure 8.17 and 9.6).
- Including ray-tracing effects in the reconstruction will further allow for a proper distance determination which is not yet implemented. Apart from adding a slight improvement to the analysis efficiency, the distance information is important for the energy resolution on future neutrino discoveries.
- In the geometrical calibration of station ARA02, for two of the channels no sensible positions and cable delays could be determined due to not explainable timing offsets. The exclusion of these channels leads to uncertainties on the reconstruction and consequently to wider angular cuts. An improvement of the geometrical calibration can allow to tighten these cuts and to increase the efficiency.
- The response of the antennas and the signal chain have not been properly calibrated yet due to constraints in man power. Therefore, a detailed Fourier analysis of the recorded signals has been avoided in the present analysis. In addition, significant systematic errors have to be taken into account due to the missing calibration. A proper understanding of the signal chain can improve these systematic errors and allow for frequency domain based analysis techniques to improve the efficiency.

As explained, the developed analysis methods can be further improved and replaced by better techniques with a better understanding of the ARA stations in the coming years. The sensitivity for the final ARA37 detector at trigger level and in the current configuration has been shown in Figure 10.5. This sensitivity cuts deeply into the GZK neutrino flux prediction which renders a discovery of neutrinos with ARA37 very likely.

Currently, investigations are being performed to improve the detector configuration. As a result of the presented analysis, the reconstruction of the distance of radio vertices has been found to be difficult with the antenna spacing used in ARA02 and ARA03. Such a reconstruction is however important for the determination of the neutrino energy for signal events. A wider antenna spacing is currently tested in simulations since it is expected to improve the distance reconstruction significantly.
Moreover, the geometry of the calibration setup can still be optimized. As shown in Chapter 6, the antenna positions are determined well in relative depth but there is room for improvement in the calibration of cable delays and the X-Y position of the strings. The installation of additional calibration pulsers can add significant constraints for the

mentioned parameters and improve our knowledge about the station geometry. A good position for such pulsers could be on two of the four measurement strings in the center between the receiving antennas. In this way, significantly higher costs for additional hole drilling could be prevented. Furthermore, the pulser positions would be locked to the string positions and the distance between strings could be determined with a very good precision.

Apart from changes in the geometry, also the data acquisition could be optimized with an improved trigger system. As presented in Section 8.3.2, simulations show that the time sequence algorithm would allow for a lowering of the current trigger threshold by a factor of 1.7, if used as a trigger. Additional algorithms are being tested and methods to test new triggers on real data are in development, in order to improve the ARA trigger systems and enhance the sensitivity of the detector.

This enhancement could also be achieved by the implementation of a PCI-express bus for the data transfer between the ATRI board and the SBC, to replace the currently used and much slower USB connection (Section 4.1.2). A PCI-express bus could increase the acceptable event rate by approximately a factor of 50, which would allow for the usage of a lower trigger threshold and consequently a higher sensitivity for ARA.

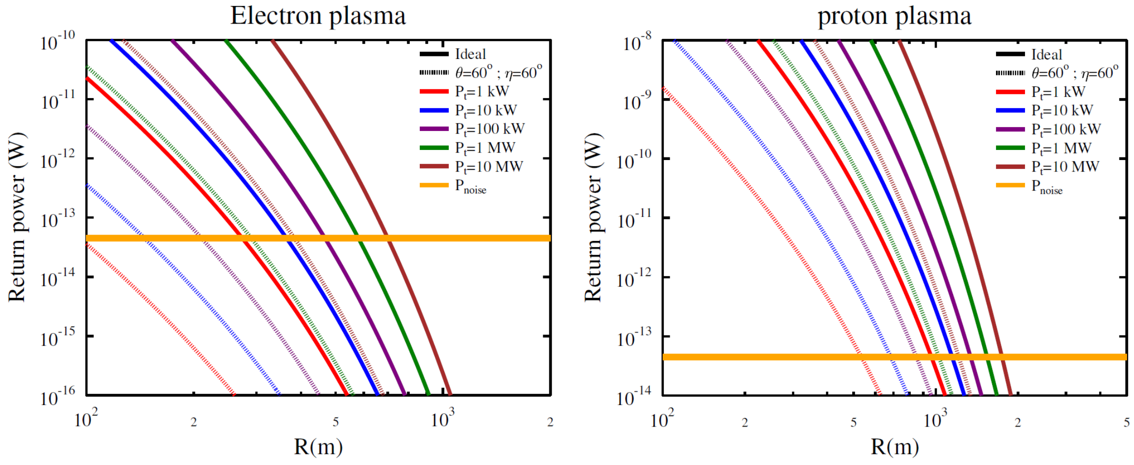
11.2 Outlook

In spite of the discovery potential of ARA37, Figure 10.5 shows that in an energy range between 10^{15} eV and 10^{17} eV none of the existing and planned detectors provides a satisfying sensitivity for the study of GZK neutrinos. This sensitivity gap could be closed by utilizing a new detection technique, proposed in [107]. In this paper the detection of neutrino-induced cascades via RADAR signals, bounced off of the induced plasmas in the cascade, is discussed. Such a method is currently tested for air-shower detection at the Telescope Array site [108]. For ice, small scale measurements have been performed so far to investigate the plasma properties [109].

Conducting plasmas are formed in neutrino induced cascades in ice mainly via ionization of the lattice molecules. This leaves us with a relatively short living electron plasma and a more stable plasma with properties equal to free protons. Electromagnetic waves can efficiently be reflected off of a plasma if the frequency does not exceed a characteristic plasma frequency ν_P. On the other hand, the reflection is only efficient if the period of the incoming wave does not exceed the lifetime of the plasma and the wavelength does not exceed its geometrical dimensions. The characteristic plasma frequency increases with the charge density which again is proportional to the energy of the particle that induced the plasma. Hence, we can deduce a lower energy limit for RADAR detection from the lifetime and geometry of the produced plasmas.

In [107] ionized electrons and proton-like particles from neutrino induced cascades have been found to allow for a RADAR reflection at reasonable frequency. The lifetime of the electrons was measured in [110, 111] to between 0.5 ns and 10 ns. For the proton-like plasma a lifetime of more than 10 ns was found [112, 113]. The geometrical shape of the respective plasma is determined in [107] based on the parametrization of electromagnetic cascades. The lower limit on a RADAR frequency for the electron plasma, under the conservative assumption of a 1 ns lifetime, has been found to be 1 GHz. This corresponds to a lower energy limit of 4 PeV. For the proton-like plasma the lower frequency limit is determined to be 50 MHz, setting an energy limit of 20 PeV. The RADAR detection technique could thus close the current sensitivity gap for GZK neutrinos.

An advantage of RADAR reflections compared to other detection methods is the control over the input signal. As long as the plasma frequency is not exceeded, the return power depends mainly on the distance and the input power and can thus be very high. In [107] the return power has been calculated for the electron and the proton-like plasma for a RADAR signal propagating through South Pole ice. The result is shown in Figure 11.1.



(a) **(b)**

Figure 11.1: *The RADAR return power versus the distance of emitter and receiver to a **(a)** electron and **(b)** proton-like plasma. Different colors indicate different levels of input power. The estimation of thermal noise power on the antennas is indicated by the horizontal yellow line. The RADAR frequencies assumed are **(a)*** 1 GHz *for the electron plasma and **(b)*** 50 MHz *for the proton plasma [107].*

This figure nicely displays the range of possible RADAR detections, defined by the crossing point of the return signal power and the thermal noise power. The presented curves have been calculated for a single antenna. The signal to noise ratio is expected to rise with the square root of the number of antennas. The presented ranges in Figure 11.1 look very promising to allow for a large sensor spacing and consequently for a very cost efficient detector setup. Especially for the proton-like plasma the range is large since a far lower RADAR frequency can be used, which is favorable in the wave propagation in ice.

As mentioned above, the electron lifetime has been assumed conservatively in the given limits. Currently, experiments are planned for a detailed investigation of the plasma properties to confirm and refine the results of previous measurements and the above mentioned theoretical calculation. If the lifetime and geometry of such plasmas is found to be in a usable range, RADAR detection of neutrino induced cascades could become an interesting alternative for building GZK neutrino detectors in ice.

Appendices

A

Noise modeling for AraSim

In this appendix the modeling of thermal noise according to the frequency response of ARA03 is described. This modeling has been performed with information from [114] and under close guidance of Eugene Hong and Dr. Carl Pfendner (both Ohio State University).

The main noise component in ARA is Gaussian distributed thermal noise originating from the ice surrounding the ARA antennas. When Gaussian noise is transformed into the frequency domain by a Fourier transformation, the distribution of the absolute magnitude of events in one frequency bin can be modeled by a Rayleigh distribution:

$$f(x) = \frac{x}{\sigma^2} e^{-\frac{1}{2}\frac{x^2}{\sigma^2}}. \tag{A.1}$$

In Figure A.1 this distribution is shown for noise events on ARA03. The σ-value of the Rayleigh distribution depends on the frequency response of the recording antenna and signal chain. We can thus use fits of Rayleigh functions to the recorded noise distribution in each frequency bin to model the noise on the ARA stations.

For each frequency bin we generate a random complex number with arbitrary phase and a magnitude following the Rayleigh function obtained by the fit. In the ARA simulation the Box-Muller algorithm is used to generate this complex number [115]. The obtained frequency spectrum is transformed into the time domain by an inverse Fourier transform. The result is a noise waveform which on average follows the fit frequency spectrum of the used data. In this way, a realistic estimation of thermal noise can be obtained. The average spectrum of noise events, simulated with the method described above, compared to the average spectrum of original noise data from station ARA03, is shown in Figure

A.2. In general we find a good agreement. However, for one channel with very low noise amplitude the fits did not converge properly and the response is overestimated in the simulation. This is a small effect and can be neglected in the general result.

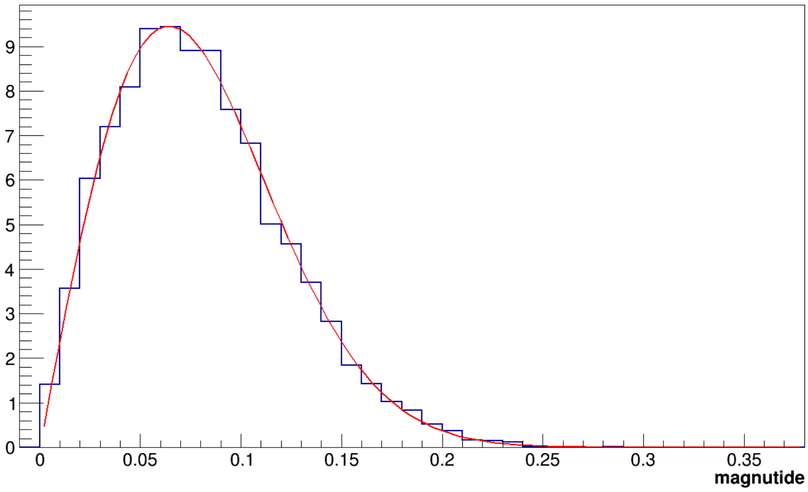


Figure A.1: *The magnitude distribution in mV of a single frequency bin for noise events in ARA03 (blue), normalized by the number of data points in the waveform and fit by a Rayleigh function (red).*

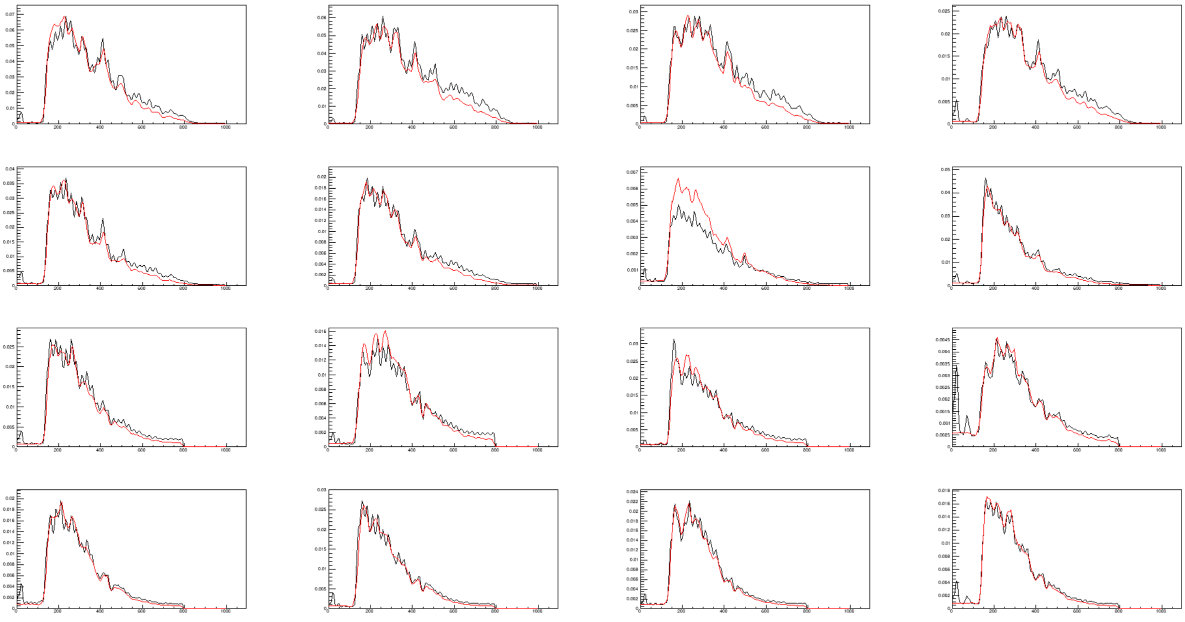


Figure A.2: *The average frequency spectrum for simulated noise events (red) and noise events in ARA03 (black) for all 16 in-ice antennas. The y-axis shows the magnitude in mV normalized by the number of data points in the waveform, the x-axis shows the frequency in MHz.*

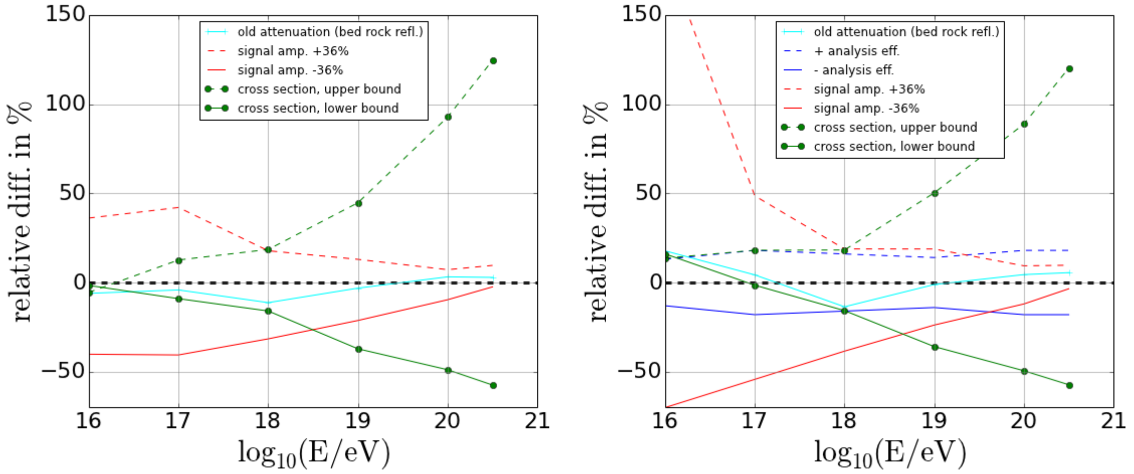
B

Details on the systematic error estimation

In Section 10.2 we summarized the derivation of systematic errors on the analysis presented in this thesis. In the following we explain a few more details on the energy dependence and on selected error estimations.

Figure B.1 shows the relative systematic errors on the effective area as function of energy at trigger level (a) and after application of the analysis cuts (b). One finds that uncertainties on the attenuation length and the analysis efficiency only have a mild influence on the effective area and do not show a strong energy dependence. The biggest contributions to the error originate from the uncertainties on the cross section and on the signal chain. Their dominance shifts with energy: at low energies the error on the signal chain is the strongest while at higher energies the cross section error becomes dominant. In the following, the origin of these influences is described in more details.
The strong influence of the signal chain at low energies is due to the fact that the signal amplitudes of most events are close to the trigger threshold. At higher energies the signal abundantly exceeds the threshold and a degradation or elevation does not have a strong influence on the trigger and analysis efficiency.

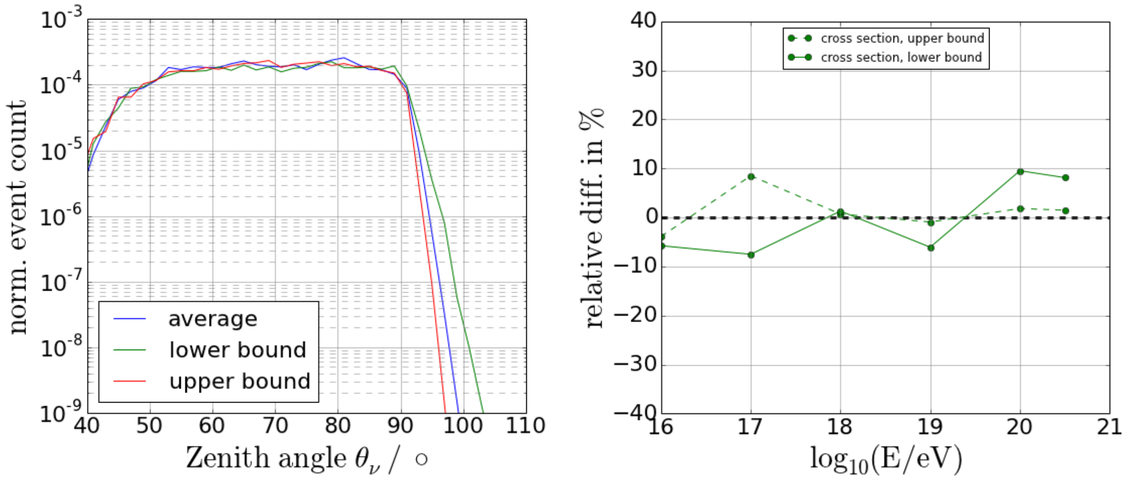
The error on the cross section arises from the fact that the neutrino interaction cross section at the energies of GZK neutrinos is unmeasured but has to be extrapolated using theoretical calculations. Uncertainties go up to 100% at 10^{20} eV [101]. This error affects mostly the neutrinos reaching the detector through the Earth. Their contribution to the effective volume is however very small (Figure B.2a) and consequently a mild influence of the cross section on the effective volume is observed (Figure B.2b). For this reason, the error on the effective area is proportional to the uncertainty of the cross section (see



(a) **(b)**

Figure B.1: *The relative changes due to systematic uncertainties on given input parameters plotted versus energy:* ***(a)*** *at trigger level,* ***(b)*** *after analysis cuts.*

Equation 7.2 and 7.3) and grows with rising energy.



(a) **(b)**

Figure B.2: ***(a)*** *The acceptance of the two ARA stations versus zenith angle of the incoming neutrino at* 10^{19} eV *at differently assumed interaction cross sections, according to the values and uncertainty bounds derived in [101].* ***(b)*** *The relative difference of the effective volume of the two ARA stations for cross sections assumed at the upper and lower uncertainty bound, compared to the average value.*

The error on the analysis efficiency is calculated using the SNR comparison between simulated data and calibration pulsers, as shown in Figure 8.7, 8.8, 8.21 and 8.23. The same distributions are produced for each energy bin separately and analysis cuts are applied. For SNR bins of $width = 1$ the cut efficiency on calibration pulser events and simulated events is compared and the difference is used as the systematic error for that bin. The analysis efficiency for simulated neutrinos with a primary energy of 10^{18} eV compared to calibration pulser events on ARA03 is shown in Figure B.3. The bias on the SNR value for pure noise has been calculated to 1.9 ± 0.2. Therefore, only events with $SNR > 2$ are used for the error estimation. The final error for a given neutrino energy is derived as the mean value of the differences in efficiency $\Delta_{eff,i}$, weighted by the number of triggering events N_i in a SNR bin i:

$$\sigma_{sys} = \frac{\sum (|\Delta_{eff,i}| \cdot N_i)}{\sum N_i}. \tag{B.1}$$

One should note that most of the simulated events have a low SNR which is why these bins get the strongest weight. To give one example the derived relative systematic error for 10^{18} eV is 16%.

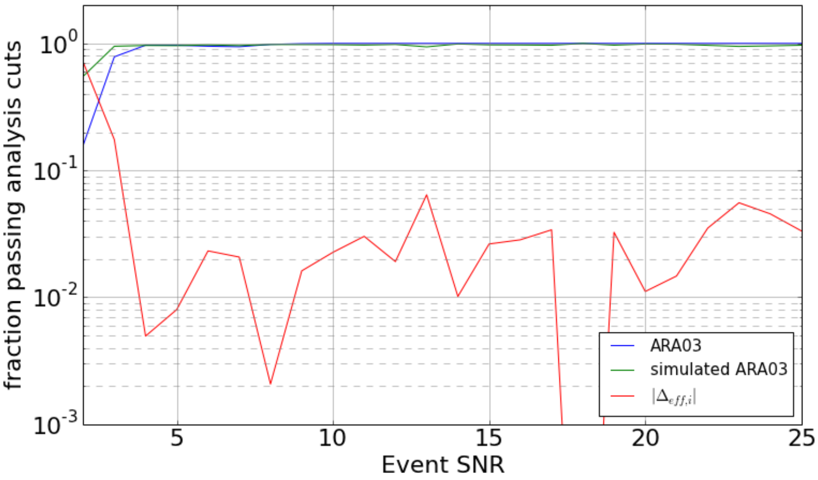


Figure B.3: *The analysis efficiency for simulated neutrino events (green) with a primary energy of* 10^{18} eV *and calibration pulser events on ARA03 (blue). The difference in efficiency is plotted in red.*

List of Figures

List of Tables

Bibliography

[1] V. F. Hess, *Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten*, *Physikalische Zeitschrift* **XIII** (1912) 1084–1091.

[2] C. D. Anderson, *The Positive Electron*, *Phys. Rev.* **43** (1933) 491–494.

[3] A. M. Hillas, *Cosmic Rays: Recent Progress and some Current Questions*, *ArXiv Astrophysics e-prints* (2006) [`astro-ph/0607109`].

[4] **AGASA** Collaboration, N. Chiba et al., *Akeno Giant Air Shower Array (AGASA) covering 100 km2 area* , *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **311** (1992), no. 1–2 338 – 349.

[5] R. Abbasi et al., *Measurement of the flux of ultra high energy cosmic rays by the stereo technique*, *Astroparticle Physics* **32** (2009), no. 1 53 – 60.

[6] **LOFAR** Collaboration, P. Schellart et al., *Detecting cosmic rays with the LOFAR radio telescope*, *Astronomy and Astrophysics* **560** (2013) A98.

[7] **Pierre Auger** Collaboration, A. Letessier-Selvon et al., *Highlights from the Pierre Auger Observatory*, *Brazilian Journal of Physics* (2014) [`arXiv:1310.4620`].

[8] **Telescope Array** Collaboration, T. Abu-Zayyad et al., *The surface detector array of the Telescope Array experiment* , *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **689** (2012), no. 0 87 – 97.

[9] **IceCube** Collaboration, M. G. Aartsen et al., *Measurement of the cosmic ray energy spectrum with IceTop-73*, *Phys. Rev. D* **88** (2013) 042004.

[10] J. Beringer et. al. (PDG), *Neutrino Cross Section Measurements*, *Physical Review D* **86** (2012).

[11] R. Gandhi, C. Quigg, M. H. Reno, and I. Sarcevic, *Ultrahigh-energy neutrino interactions* , *Astroparticle Physics* **5** (1996), no. 2 81 – 110.

[12] W. Pauli, *Dear radioactive ladies and gentlemen*, *Phys. Today* **31N9** (1978) 27.

[13] F. Reines, *The neutrino: from poltergeist to particle*, *Rev. Mod. Phys.* **68** (1996) 317–327.

[14] C. Cowan, F. Reines, F. Harrison, H. Kruse, and A. McGuire, *Detection of the free neutrino: A Confirmation*, *Science* **124** (1956) 103–104.

[15] V. M. Hannen, *Direct neutrino mass determination: Status and prospects*, *Journal of Physics: Conference Series* **375** (2012), no. 4 042004.

[16] S. Hannestad, A. Mirizzi, G. G. Raffelt, and Y. Y. Wong, *Neutrino and axion hot dark matter bounds after WMAP-7*, *Journal of Cosmology and Astroparticle Physics* **2010** (2010), no. 08 001.

[17] **KATRIN** Collaboration, J. Wolf et al., *The KATRIN neutrino mass experiment*, vol. 623, pp. 442 – 444, 2010. 1st International Conference on Technology and Instrumentation in Particle Physics.

[18] **Super-Kamiokande** Collaboration, Y. Fukuda et al., *Evidence for Oscillation of Atmospheric Neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562–1567.

[19] **Super-Kamiokande** Collaboration, Y. Fukuda et al., *The Super-Kamiokande detector*, *Nucl. Instrum. Meth.* **A501** (2003) 418–462.

[20] **Daya Bay** Collaboration, F. P. An et al., *Observation of Electron-Antineutrino Disappearance at Daya Bay*, *Phys. Rev. Lett.* **108** (2012) 171803.

[21] **Double Chooz** Collaboration, Y. Abe et al., *Indication of Reactor $\bar{\nu}_e$ Disappearance in the Double Chooz Experiment*, *Phys. Rev. Lett.* **108** (2012) 131801.

[22] **RENO** Collaboration, J. K. Ahn et al., *Observation of Reactor Electron Antineutrinos Disappearance in the RENO Experiment*, *Phys. Rev. Lett.* **108** (2012) 191802.

[23] G. Mention et al., *Reactor antineutrino anomaly*, *Phys. Rev. D* **83** (2011) 073006.

[24] K. Nakamura et. al. (PDG), *Review of Particle Physics*, *Journal of Physics G: Nuclear and Particle Physics* **37** (2010), no. 7A 075021.

[25] E. Majorana, *Teoria simmetrica dell'elettrone e del positrone*, *Il Nuovo Cimento* **14** (1937), no. 4 171–184.

[26] M. Markov, *On high energy neutrino physics*, in *Proceedings of the 1960 Annual International Conference on HEP at Rochester*, pp. 578–581, 1960.

[27] B. T. Cleveland et al., *Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector*, *The Astrophysical Journal* **496** (1998), no. 1 505.

[28] W. Arnett, J. N. Bahcall, R. Kirshner, and S. Woosley, *SUPERNOVA SN1987A*, *Ann. Rev. Astron. Astrophys.* **27** (1989) 629–700.

[29] K. Hirata et al., *Observation of a neutrino burst from the supernova SN1987A*, *Phys. Rev. Lett.* **58** (1987) 1490–1493.

[30] R. M. Bionta et al., *Observation of a neutrino burst in coincidence with supernova 1987A in the Large Magellanic Cloud*, *Phys. Rev. Lett.* **58** (1987) 1494–1496.

[31] E. Alexeyev, L. Alexeyeva, I. Krivoscheina, and V. Volchenko, *Detection of the neutrino signal from sn1987a using the inr baksan underground scintillation telescope*, in *Neutrino Physics* (H. Klapdor and B. Povh, eds.), pp. 288–298. Springer Berlin Heidelberg, 1988.

[32] **IceCube** Collaboration, R. Abbasi et al., *The IceCube data acquisition system: Signal capture, digitization, and timestamping* , *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **601** (2009), no. 3 294 – 316.

[33] **IceCube** Collaboration, M. Aartsen et al., *Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data*, *Phys.Rev.Lett.* **113** (2014) 101101, [`arXiv:1405.5303`].

[34] K. Greisen, *End to the Cosmic-Ray Spectrum?*, *Phys. Rev. Lett.* **16** (1966) 748–750.

[35] G. T. Zatsepin and V. A. Kuzmin, *Upper limit of the spectrum of cosmic rays*, *JETP Lett.* **4** (1966) 78–80.

[36] V. Beresinsky and G. Zatsepin, *Cosmic rays at ultra high energies (neutrino?)* , *Physics Letters B* **28** (1969), no. 6 423 – 424.

[37] A. A. Penzias and R. W. Wilson, *A Measurement of Excess Antenna Temperature at 4080 Mc/s*, *Astrophysical Journal* **142** (1965) 419.

[38] R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson, *Cosmic Black-body Radiation*, *Astrophysical Journal* **142** (1965) 414.

[39] G. F. Hinshaw et al., *Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Results*, *apJS.* (2012).

[40] R. Engel, D. Seckel, and T. Stanev, *Neutrinos from propagation of ultrahigh energy protons*, *Phys. Rev. D* **64** (2001) 093010.

[41] K. K. Andersen and S. R. Klein, *High energy cosmic-ray interactions with particles from the Sun*, *Phys. Rev. D* **83** (2011) 103519.

[42] **Pierre Auger** Collaboration, J. Abraham et al., *Measurement of the energy spectrum of cosmic rays above 1018 eV using the Pierre Auger Observatory*, *Physics Letters B* **685** (2010), no. 4–5 239 – 246.

[43] D. Allard et al., *Cosmogenic neutrinos from the propagation of ultrahigh energy nuclei*, *Journal of Cosmology and Astroparticle Physics* **2006** (2006), no. 09 005.

[44] **Telescope Array** Collaboration, P. Tinyakov et al., *Latest results from the telescope array* , *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **742** (2014), no. 0 29 – 34.

[45] A. Bell, *Cosmic ray acceleration* , *Astroparticle Physics* **43** (2013) 56 – 70.

[46] A. Hillas, *Where do 1019 eV cosmic rays come from?* , *Nuclear Physics B - Proceedings Supplements* **136** (2004), no. 0 139 – 146. CRIS 2004 Proceedings of the Cosmic Ray International Seminars: GZK and Surroundings.

[47] A. M. Hillas, *The Origin of Ultra-High-Energy Cosmic Rays*, *Annual Review of Astronomy and Astrophysics* **22** (1984), no. 1 425–444.

[48] J. Abraham et al., *Upper limit on the cosmic-ray photon fraction at EeV energies from the Pierre Auger Observatory*, *Astroparticle Physics* **31** (2009), no. 6 399 – 406.

[49] H. Yüksel, M. D. Kistler, J. F. Beacom, and A. M. Hopkins, *Revealing the High-Redshift Star Formation Rate with Gamma-Ray Bursts*, *The Astrophysical Journal Letters* **683** (2008), no. 1 L5.

[50] M. Ahlers and F. Halzen, *Minimal cosmogenic neutrinos*, *Phys. Rev. D* **86** (2012) 083010.

[51] **Fermi LAT** Collaboration, A. A. Abdo et al., *Spectrum of the Isotropic Diffuse Gamma-Ray Emission Derived from First-Year Fermi Large Area Telescope Data*, *Phys. Rev. Lett.* **104** (2010) 101101.

[52] M. Ahlers et al., *GZK neutrinos after the Fermi-LAT diffuse photon flux measurement*, *Astroparticle Physics* **34** (2010), no. 2 106 – 115.

[53] G. A. Askaryan, *Excess negative charge of an electron-photon shower and its coherent radio emission*, *JETP* **41** (1962) 616 – 618.

[54] G. A. Askaryan, *Coherent radio emission from cosmic showers in air and in dense media*, *JETP* **48** (1965) 988 – 990.

[55] E. Zas, F. Halzen, and T. Stanev, *Electromagnetic pulses from high-energy showers: Implications for neutrino detection*, *Phys. Rev. D* **45** (1992) 362–376.

[56] L. D. Landau and I. I. Pomeranchuk, *The limits of applicability of the theory of Bremsstrahlung by electrons and of the creation of pairs at large energies*, *Dokl. Akad. Nauk SSSR* **92** (1953) 535.

[57] A. B. Migdal, *Bremsstrahlung and Pair Production in Condensed Media at High Energies*, *Phys. Rev.* **103** (1956) 1811–1820.

[58] T. Stanev et al., *Development of ultrahigh-energy electromagnetic cascades in water and lead including the Landau-Pomeranchuk-Migdal effect*, *Phys. Rev. D* **25** (1982) 1291–1304.

[59] H. A. Bethe, *Molière's Theory of Multiple Scattering*, *Phys. Rev.* **89** (1953) 1256–1266.

[60] K. D. de Vries. private communication, 2014.

[61] J. Lundberg et al., *Light tracking through ice and water—Scattering and absorption in heterogeneous media with Photonics* , *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **581** (2007), no. 3 619 – 631.

[62] J. Alvarez-Muñiz and E. Zas, *Cherenkov radio pulses from EeV neutrino interactions: the LPM effect* , *Physics Letters B* **411** (1997), no. 1–2 218 – 224.

[63] J. Alvarez-Muñiz, A. Romero-Wolf, and E. Zas, *Čerenkov radio pulses from electromagnetic showers in the time domain*, *Phys. Rev. D* **81** (2010) 123009.

[64] S. Hussain and D. W. McKay, *Comparative study of radio pulses from simulated hadron-, electron-, and neutrino-initiated showers in ice in the GeV-PeV range*, *Phys. Rev. D* **70** (2004) 103003.

[65] J. Alvarez-Muñiz, W. R. J. Carvalho, M. Tueros, and E. Zas, *Coherent Cherenkov radio pulses from hadronic showers up to EeV energies*, *Astroparticle Physics* **35** (2012), no. 6 287 – 299.

[66] K. D. de Vries, O. Scholten, and K. Werner, *Macroscopic geo-magnetic radiation model; polarization effects and finite volume calculations* , *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **662, Supplement 1** (2012), no. 0 S175 – S178.

[67] O. Scholten, K. Werner, and F. Rusydi, *A macroscopic description of coherent geo-magnetic radiation from cosmic-ray air showers* , *Astroparticle Physics* **29** (2008), no. 2 94 – 103.

[68] K. Werner, K. D. de Vries, and O. Scholten, *A realistic treatment of geomagnetic Cherenkov radiation from cosmic ray air showers* , *Astroparticle Physics* **37** (2012), no. 0 5 – 16.

[69] V. Marin and B. Revenu, *Simulation of radio emission from cosmic ray air shower with SELFAS2*, *Astroparticle Physics* **35** (2012), no. 11 733 – 741.

[70] T. Huege, M. Ludwig, and C. James, *Simulating radio emission from air showers with CoREAS*, `arXiv:1301.2132`.

[71] D. Saltzberg et al., *Observation of the Askaryan Effect: Coherent Microwave Cherenkov Emission from Charge Asymmetry in High-Energy Particle Cascades*, *Phys. Rev. Lett.* **86** (2001) 2802–2805.

[72] P. W. Gorham et al., *Accelerator measurements of the Askaryan effect in rock salt: A roadmap toward teraton underground neutrino detectors*, *Phys. Rev. D* **72** (2005) 023002.

[73] **ANITA** Collaboration, P. W. Gorham et al., *Observations of the Askaryan Effect in Ice*, *Phys. Rev. Lett.* **99** (2007) 171101.

[74] **Pierre Auger** Collaboration, A. Aab et al., *Probing the radio emission from air showers with polarization measurements*, *Phys. Rev. D* **89** (2014) 052002.

[75] **ARA** Collaboration, P. Allison et al., *Design and initial performance of the Askaryan Radio Array prototype EeV neutrino detector at the South Pole* , *Astroparticle Physics* **35** (2012), no. 7 457 – 477.

[76] **ARA** Collaboration, P. Allison et al., *First Constraints on the Ultra-High Energy Neutrino Flux from a Prototype station of the Askaryan Radio Array*, 2014.

[77] A. Connolly. private communication, 2011.

[78] S. Barwick, D. Besson, P. Gorham, and D. Saltzberg, *South Polar in situ radio-frequency ice attenuation*, *Journal of Glaciology* **51** (2005), no. 173 231–238.

[79] I. Kravchenko, D. Besson, and J. Meyers, *In situ index-of-refraction measurements of the South Polar firn with the RICE detector*, *Journal of Glaciology* **50** (2004), no. 171 522–532.

[80] R. Gaïor. private communication, 2014.

[81] T. Kuwabara. private communication, 2014.

[82] P. W. Gorham et al., *Observational constraints on the ultrahigh energy cosmic neutrino flux from the second flight of the ANITA experiment*, *Phys. Rev. D* **82** (2010) 022004.

[83] **IceCube** Collaboration, M. Aartsen et al., *Probing the origin of cosmic-rays with extremely high energy neutrinos using the IceCube Observatory*, *Phys.Rev.* **D88** (2013) 112008, [`arXiv:1310.5477`].

[84] J. Abraham et al., *Properties and performance of the prototype instrument for the Pierre Auger Observatory*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **523** (2004), no. 1–2 50 – 95.

[85] **Pierre Auger** Collaboration, P. Abreu et al., *Ultrahigh Energy Neutrinos at the Pierre Auger Observatory*, *Adv. High Energy Phys.* **2013** (2013) 708680, [`arXiv:1304.1630`].

[86] I. Kravchenko et al., *Updated results from the RICE experiment and future prospects for ultra-high energy neutrino detection at the south pole*, *Phys. Rev. D* **85** (2012) 062004.

[87] S. W. Barwick for the ARIANNA Collaboration, *Performance of the ARIANNA Prototype Array*, in *Proceedings of 33rd International Cosmic Ray Conference*, 2013.

[88] Y. Abdou et al., *Design and performance of the South Pole Acoustic Test Setup*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **683** (2012), no. 0 78 – 90.

[89] J. Aguilar et al., *AMADEUS—The acoustic neutrino detection test system of the ANTARES deep-sea neutrino telescope* , *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **626–627** (2011), no. 0 128 – 143.

[90] O. Scholten et al., *Improved Flux Limits for Neutrinos with Energies above* 10^{22} eV *from Observations with the Westerbork Synthesis Radio Telescope*, *Phys. Rev. Lett.* **103** (2009) 191301.

[91] G. S. Varner, *IRS2 - 8-channel, GSPS Transient Waveform Recorder with GigaHertz bandwidth and Fast, Selective Window Readout*, *Internal ARA document* (2011).

[92] G. S. Varner et al., “Specifications for the IceRay Sampler (IRS) ASIC.” `http://www.phys.hawaii.edu/~varner/IRS_spec_v01.pdf`, 2009.

[93] W. R. Leo, *Techniques for Nuclear and Particle Physics Experiments: A How-To Approach.* Springer-Verlag, Berlin, 1994.

[94] L. L. Ruckman and G. S. Varner, *The PRO1 ASIC for fast Wilkinson encoding*, *Journal of Instrumentation* **3** (2008), no. 12 P12003.

[95] G. Varner et al., *The large analog bandwidth recorder and digitizer with ordered readout (LABRADOR) ASIC*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **583** (2007), no. 2–3 447 – 460.

[96] P. Allison, *ATRI revision C information*, *Internal ARA document* (2011).

[97] K. Nishimura and A. Romero-Wolf, *A Correlation-based Timing Calibration and Diagnostic Technique for Fast Digitizing ASICs*, *Physics Procedia* **37** (2012), no. 0 1707 – 1714.

[98] G. Varner, *IRS2 design review*, *Internal ARA document* (2010).

[99] A. Connolly, *Defining ARA Coordinate Systems*, *Internal ARA document* (2013).

[100] E. S. Hong, A. Connolly, and C. G. Pfendner for the ARA Collaboration, *Trigger and data filtering approaches in the askaryan radio array*, *Proceedings of the 33rd International Cosmic Ray Conference, Rio de Janeiro* (2013).

[101] A. Connolly, R. S. 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* **83** (2011) 113009.

[102] T. Meures for the ARA Collaboration, *Trigger and data filtering approaches in the Askaryan Radio Array*, *Proceedings of the 33rd International Cosmic Ray Conference, Rio de Janeiro* (2013).

[103] S. Bancroft, *An Algebraic Solution of the GPS Equations*, *Aerospace and Electronic Systems, IEEE Transactions on* **AES-21** (1985), no. 1 56–59.

[104] “Eigen3 documentation.” `http://eigen.tuxfamily.org`, 2014.

[105] W. H. Press, *Numerical recipes 3rd edition: The art of scientific computing.* Cambridge University Press, New York, 2007.

[106] G. J. Feldman and R. D. Cousins, *Unified approach to the classical statistical analysis of small signals*, *Phys. Rev. D* **57** (1998) 3873–3889.

[107] K. D. de Vries, K. Hanson, and T. Meures, *On the feasibility of RADAR detection of high-energy neutrino-induced showers in ice*, *Astroparticle Physics* **60** (2015), no. 0 25 – 31.

[108] M. A. B. Othman et al., *Air Shower Detection by Bistatic Radar*, *AIP Conference Proceedings* **1367** (2011), no. 1.

[109] M. Chiba et al., *Measurement of a phase of a radio wave reflected from rock salt and ice irradiated by an electron beam for detection of ultra-high-energy neutrinos*, *AIP Conference Proceedings* **1535** (2013), no. 1.

[110] J. Verberne et al., *Excess electrons in ice*, *Nature* **272** (1978) 343 – 344.

[111] M. P. De Haas, M. Kunst, J. M. Warman, and J. B. Verberne, *Nanosecond time-resolved conductivity studies of pulse-ionized ice. 1. The mobility and trapping of conduction-band electrons in water and deuterium oxide ice*, *The Journal of Physical Chemistry* **87** (1983), no. 21 4089–4092.

[112] M. Kunst and J. Warman, *Proton mobility in ice*, *Nature* **288** (1980) 465 – 467.

[113] M. Kunst, J. M. Warman, M. P. De Haas, and J. B. Verberne, *Nanosecond time-resolved conductivity studies of pulse-ionized ice. 3. The electron as a probe for defects in doped ice*, *The Journal of Physical Chemistry* **87** (1983), no. 21 4096–4098.

[114] M. J. Mottram, *A Search for Ultra-high Energy Neutrinos and Cosmic-Rays with ANITA-2*. PhD dissertation, University College London, 2012.

[115] G. E. P. Box and M. E. Muller, *A Note on the Generation of Random Normal Deviates*, *The Annals of Mathematical Statistics* **29** (1958), no. 2 610–611.

Acknowledgements

There are many people that I am very grateful to and which contributed to the completion of this work. Probably I will forget some names which should be in the following list. Please forgive me if I forgot somebody in the moment that I am writing this.

Many thanks go to my supervisor Kael Hanson, for always being available for discussions and questions. Thanks for creating a relaxed and familiar atmosphere. In your guidance I found this a nice mix of input and freedom to develop my own ideas. Furthermore, thank you for giving me the opportunity to take responsibilities within the ARA project and to spend two very exciting month at the South Pole and at many other places.

I would like to thank the members of my Jury: Stijn Buitink, Gilles De Lentdecker, Olaf Scholten and Petr Tiniakov. Thank you to agree to read my work, to spend the time to evaluate it and to supply corrections and suggestions for improvement. Special thanks go to Petr Tiniakov and Gilles De Lentdecker as members of my Committee d'accompagnement, who had to listen to my annual reports and supply many signatures.

Thanks to the members of the IceCube group in Brussels, which I have encountered. I got a warm welcome from everybody and always found you a very enjoyable company. During coffee breaks, pizza lunches and other social activities one could forget all work related (or unrelated) problems. Furthermore, the Belgian colleagues were always available to help out the foreigners amongst us with administrative issues. I would especially like to thank the current members, namely: David Heereman, Jan Lüneman, Gwenhaël De Wasseige, Martin Casier, Lionel Brayeur, Geraldina Golup, Elisa Pinat, Jan Kunnen, Giuliano Maggi, Aongus Ó Murchadha, Krijn De Vries, Catherine De Clercq, Nick Van Eijndhoven and Kael Hanson. Thanks for the support when finishing my thesis, for discussions about cosmic rays and the Askaryan effect, and for all the proof reading and rehearsal talks.

I would like to thank my office mates Rachel, David, Aongus, Mark and Alberto to cope with hours of my ARA phone calls, for medical and life advices, and for always creating a nice atmosphere in the office.

Thanks to the DAQ team of the lab which gave me a good start into my working life and which I always enjoyed working with.

I would like to thank Audrey Terrier, Marleen Goeman, Danielle Peymans and Fatimé Pero for their help with the university administration. I think Marleen and Audrey are unsurpassed in making the IIHE a nice place to work at. In fact I want to thank all members of the IIHE for the good working environment which I encountered in the lab. This was not possible without the friendly lead from the directors of the lab: Catherine De Clercq, Daniel Bertrand, Pierre Marage, Jorgen D'Hondt and Laurent Favart. In particular I would like to thank the former director Pierre Marage, who would always listen to the trouble of a PhD student and help where he could.

Many thanks go to Dave, Dennis, Jim, Jef, Darrel, Terry, James, Gary and Andrew for allowing the German PhD student on the drill team and giving me the best time.

I would like to thank the ARA Collaboration for their support and scientific input during my PhD. This support went from hosting me for test measurements, providing simulations and continuously poking into my analysis until it was solid work.

Danke David und Bettina für Rauchergespräche und für die Möglichkeit alles Gute und Schlechte irgendwo loszuwerden. Danke für schöne Erinnerungen an abendliche Gespräche in „Le Phare“, oder am Strand von Buzios.

Thanks to Yiannis Papadopoulos for the friendly help with the cover design.

Thanks to all my friends for all the small things, which are not measurable, but which in the end keep you going.

Ich möchte meinen Eltern und Geschwistern danken: dafür, dass ihr immer an mich geglaubt habt und mich konstant unterstützt habt, in dem was ich tue. Das hat mich schlussendlich bis hierher gebracht. Ich habe viel Glück mit meiner Familie.

Danke Kathleen, dass du mich durch die stressige Zeiten begleitet hast und nie den Mut verloren hast. Danke vor allem dafür, dass du mich immer daran erinnerst, dass die Doktorarbeit nicht das wichtigste im Leben ist.