Search for Ultra-high Energy Neutrinos with the Askaryan Radio Array

By

Ming-Yuan Lu

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Physics)

at the
UNIVERSITY OF WISCONSIN–MADISON
2020

Date of final oral examination: March 27th 2020

The dissertation was reviewed by the following members of the Final Oral Committee:

Albrecht Karle, Professor, Physics
Keith Bechtol, Assistant Professor, Physics
Kael Hanson, Professor, Physics
Sebastian Heinz, Professor, Astronomy

ABSTRACT

Neutrinos act as unique messengers that bear exclusive information to help us understand the ultra-high energy universe of cosmic rays. The Askaryan Radio Array (ARA) is a neutrino telescope array under phased deployment near the South Pole aiming to discover and determine the ultra-high energy neutrino flux via detection of the Askaryan signal from neutrino-induced charged-particle showers in the Antarctic ice. This novel detection channel makes ARA the most cost-effective neutrino observatory in probing the neutrino flux from 100 PeV to 10 EeV.

This thesis focuses on the data analysis of two ARA stations from 2013 to 2016 with a goal of searching for a diffuse neutrino flux. Event selection criteria reject low-quality data and thermal background. An interferometric technique taking into account the curved paths traveled by electromagnetic radiation in inhomogeneous ice is used for event reconstruction and signal-background differentiation. No neutrino candidate is observed in the end. As a result, the analysis places a 90 percent CL upper limit on the diffuse neutrino flux, which represents the best-reported limit from radio neutrino detectors at the energy of 1 EeV.

The prospects of future analyses utilizing four times the presented data, as well as the design implication for upcoming radio neutrino detectors, are discussed towards the end.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Chapter 1

Introduction

Scientists discovered cosmic rays over a century ago. In his pioneering works in 1911 and 1913 (Nobel Prize in Physics 1936), Victor Francis Hess measured the level of radiation along increasing altitudes. And while the radiation decreases initially, over roughly 1km an enhance radiation was observed [1]. Hess concluded that the source of the said radiation was extraterrestrial, and this marks the discovery of cosmic rays.

In the first half of the 20th century, cosmic rays served the discovery channels for a number of elementary particles including positrons [2], muons [3], pions [4] [5], etc. In modern-day physics, "cosmic rays" usually refer specifically to charged nuclei, while other types of cosmic particles include electrons/positrons, gamma rays, and neutrinos. We can also make the distinction between "primary" and "secondary" cosmic rays, in which "primary" refers to particles accelerated at the source, and "secondary" refers to daughter particles from interactions of primary particles and interstellar gas. These secondaries carry information relevant to the sources and acceleration mechanism of their parent cosmic rays. We should mention that, these leptons and photons could be products of processes such as dark matter annihilation or exotic particle decay, and therefore are interesting to be probed on their own. The field of particle astrophysics seeks to detect, characterize the existing cosmic particles, and understand their origins as well as the production mechanisms. Furthermore, our understanding of astronomy and particle physics can be tested and broadened through studies in particle astrophysics.

In this thesis, we will specifically focus on UHE neutrinos and their role as the indicator to hadronic processes involving UHE cosmic rays (UHECR).

1.1 Cosmic Rays

The questions surrounding cosmic rays that physicists have yet to fully answer include: the origin of cosmic rays, the acceleration mechanism of cosmic rays, their source distribution, the composition of cosmic ray particles, their interaction with ambient matter or radiation, and how they propagate from sources to

our planet. Furthermore, cosmic rays carry information concerning particle physics at energies otherwise inaccessible on Earth, as well as the workings of astrophysical objects that produces these particles. Large-scale detectors such as the Telescope Array (TA) and the Pierre Auger Observatory (PAO) are constructed to measure characteristics such as spectrum, anisotropy, and composition of cosmic rays, with the ultimate goal of answering the above questions.

1.1.1 Spectrum

The measured cosmic rays spectrum is shown in Fig. 1.1. Below $10^{15}eV$, the spectrum follows a power law of spectral index of ~ -2.7. This is believed to be a component of galactic origin, in which supernova remnants are likely to be the particle accelerator [6]. The power-law shape of the spectrum can be explained by the non-thermal process of Fermi acceleration [7]. A transition region can be observed from 1PeV to 1EeV, with several spectral feature in this region. Most notably are the "knee" at around $10^{15} \sim 10^{16}$ eV, the "2nd knee" near 10^{17}eV, and the ankle at several time 10^{18}eV, in the "ultra-high-energy cosmic ray" (UHECR) regime. These features likely correspond to a transition from galactic to extragalactic cosmic rays, and effects of source injection, composition, propagation in and magnetic confinement of the Galaxy should be taken into account. Physicists are actively working on reducing the systematic uncertainties in these measurements in the hope for a coherent explanation.

Fig. 1.2 shows the highest-energy end of the spectrum. Spectral hardening can be observed around the "ankle". A possible explanation is the superseding of a higher-energy component, such as extragalactic cosmic rays, over a less energetic population. Above $\sim 5 \times 10^{19}$ eV, the flux level drops sharply, consistent with the prediction from the Greisen-Zatspin-Kuzmin (GZK) [9, 10] effect, in which UHECRs engage in photo-pion production with cosmic microwave background (CMB) photons. This GZK suppression is reported from the HiRes experiment [11], TA [12], and PAO [13]. In section 1.3.2, we will further explore the GZK process its byproduct.

1.1.2 Anisotropy

Cosmic ray flux is nearly isotropic across most energies, with observed anisotropy at the scale of $\sim 10^{-3}$ around a few TeV [14] [15] [16]. Such anisotropy is considered to be related to the direction of the local Galactic magnetic field, the direction of sources, and the motion of the solar system.

1.1.3 Composition

The compositions of cosmic rays is correlated with source composition, particle lifetime, and propagation models. Primary cosmic rays consist mostly of light nuclei - $\sim$ 74% are protons and $\sim$ 70% of the rest are helium nuclei. In the UHE regime, where this thesis is concerned with, PAO measures an inclination

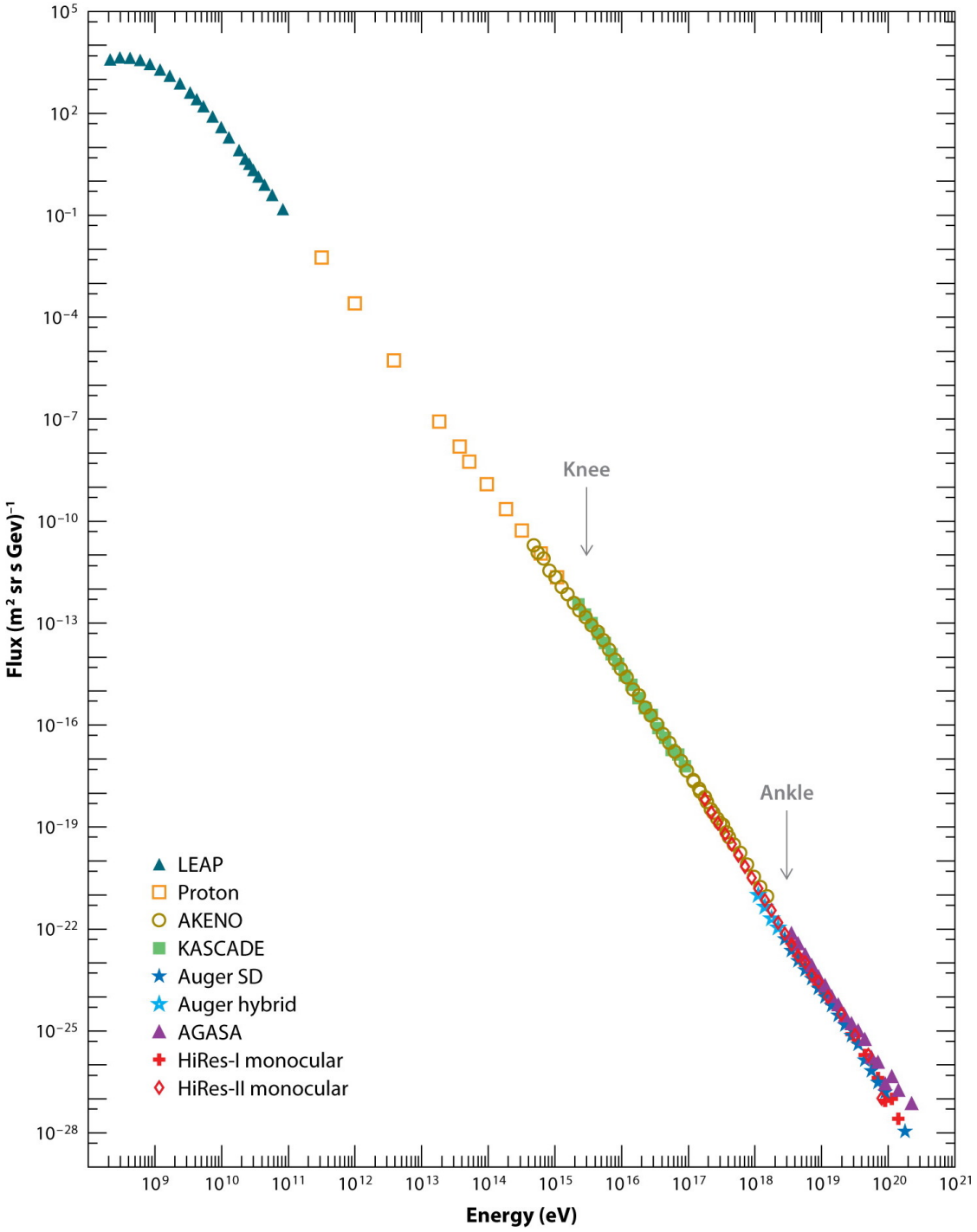


Figure 1.1: Measurements of the cosmic ray spectrum as compiled from several experiments across a wide range of energies[8].

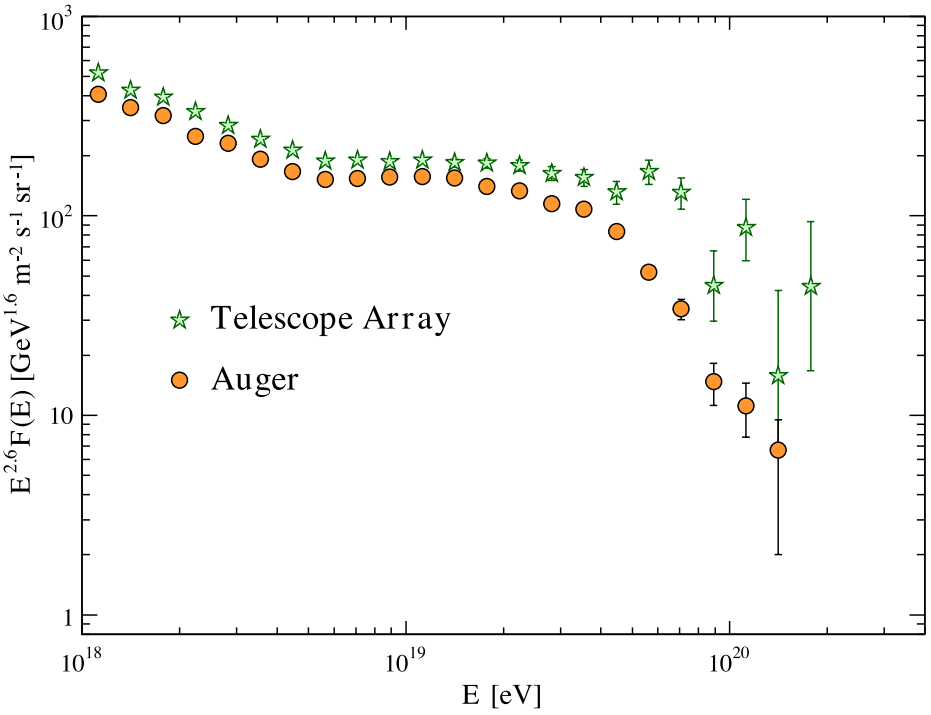


Figure 1.2: Measurements of the UHECR spectrum from PAO [13] and TA [12], displaying the described "ankle" feature and the GZK cutoff. [8].

towards heavier mass composition beginning near 2×10^{18} eV, as shown in Fig. 1.3 [17], while TA reports mass composition consistent with predominantly light nuclei up to $10^{19.1}$ eV (Fig. 1.4 [18]). While the measurements imply distinct pictures, one must note that the two experiments employ different analysis methods , and the results cannot be directly compared. The two colaborations have carried out joint analyses, where TA simulated and analyzed data based on an energy-dependent mixture of light and heavy elements produced by PAO. The results indicated compatibility between the two experiments from $10^{18.2}$ to 10^{19} eV [19]. Further studies are underway to understand and compare the two datasets.

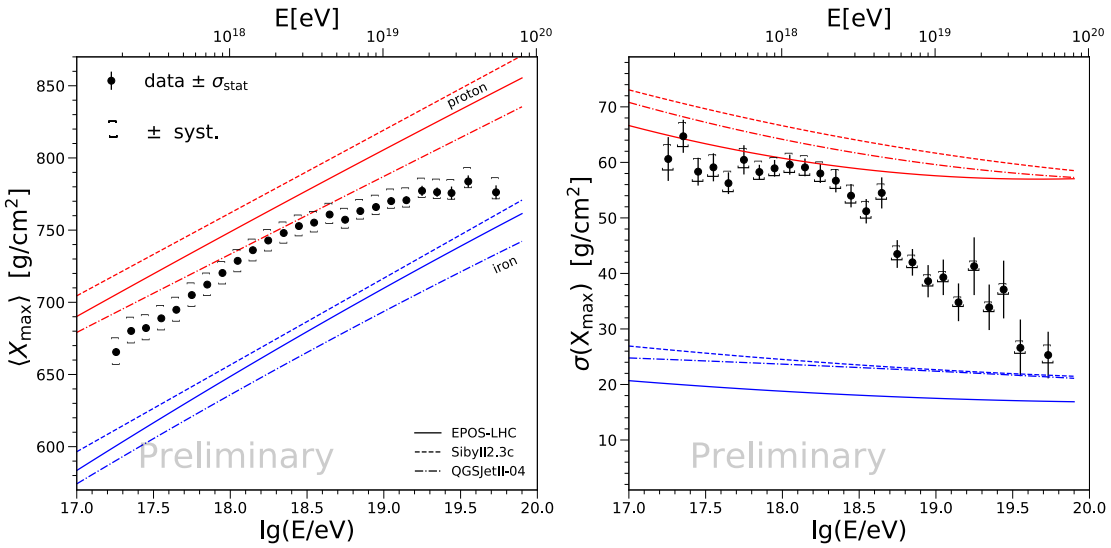
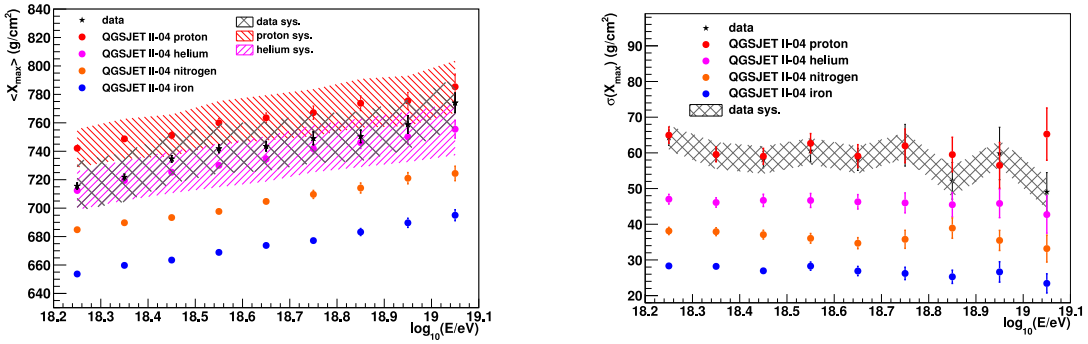


Figure 1.3: Mean ($< X_{max} >$, left) and dispersion ($\sigma(X_{max}$, right) of the depth of air shower maximum measured by PAO, compared to the predictions from proton or iron nuclei using various hadronic models. [17]



(a) TA 10 year hybrid $\langle X_{\max} \rangle$

(b) TA 10 year hybrid $\sigma(X_{\max})$

Figure 1.4: Mean ($< X_{max} >$, left) and dispersion ($\sigma(X_{max}$, right) of the depth of air shower maximum measured by TA, compared to the predictions from proton, helium, nitrogen, and iron nuclei using the QGSJET II-04 hadronic model [18].

1.2 The Neutrino-Cosmic Ray Connection

Directly probing cosmic rays is not the only approach towards cosmic ray physics. Acting as "smoking guns", secondary particles such as photons or neutrinos are also powerful messengers that carry information regarding the accelerators of cosmic rays. Neutrinos are nearly-massless, charge-neutral elementary particles

in the lepton family that virtually only interact through the weak force. It was first postulated by Wolfgang Pauli as a hypothetical particle to ensure energy conservation in the β decay process [20]. In 1956, it was finally detected by Reines and Cowan [21, 22]. Their experiment utilizes the reaction

$$\bar{\nu_e} + p \rightarrow e^+ + n \tag{1.1}$$

to capture neutrinos from a reactor beam. The coincidence of the γ ray from electron-positron pair annihilation and the other γ ray emission from neutron capture forms a unique signature of this antineutrino interaction.

In the Standard Model (SM) of elementary particles, neutrinos are theorized as massless particles having three flavors eigenstates. These three flavors are electron neutrinos nu_e, muon neutrinos ν_μ, and tau neutrinos ν_τ. However, the phenomenon called "neutrino oscillation" [23], where the flavor of a neutrino does not stay unchanged during its propagation, implies that neutrinos have small but non-zero masses. Indeed, cosmologists have placed upper limits on the sum of the masses of the three neutrino mass eigenstates [24], and experiments that seek to measure the model-independent absolute neutrino masses are also underway [25, 26]. The magnitude of this flavor mixing is determined by three "mixing angles" θ_{12}, θ_{23}, θ_{13} and one CP-violating phase δ_{CP} in the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [27].

Having laid out the basic properties of neutrinos, we can clearly see that neutrinos are a unique kind of messenger for astrophysical events. Cosmic rays below $\sim 10^{18}$eV are deflected by magnetic fields during propagation, and UHECRs above 10^{18}eV, while the deflection is small enough due to high rigidity, are soon met with the GZK cutoff that causes the horizon to be confined within $\sim$ 10's of Mpc from the acceleration sites. On the other hand, γ rays pair produce above $\sim$ TeV with cosmic microwave background (CMB) or infrared (IR) radiation, again limiting the horizon of high energy γ rays to several tens of Mpc only. Neutrinos are chargeless, and only interact weakly. Thus, it is an ideal messenger for pointing back to cosmic ray sources that are otherwise opaque to comic rays or γ rays. In the next section, we give an overview of the field arising from observing the universe through neutrinos, namely neutrino astronomy.

1.3 Neutrino Astronomy

The first observation of extraterrestrial neutrinos was made in the Homestake experiment in 1968, where solar neutrinos were observed [28]. This marked the start of the "neutrino astronomy" era, as opposed to the well-established "photon" astronomy. In this experiment, electron neutrinos from solar fusion

$$4p \rightarrow \alpha + 2e^+ + 2\nu_e \tag{1.2}$$

were captured using the inverse β process with

$$\nu_e + {}^{37}\mathrm{Cl} \rightarrow {}^{37}\mathrm{Ar} + e^-. \tag{1.3}$$

The experiment placed a tank containing 100,000 gallons of perchloroethylene underground to provision a chlorine-rich target that is shielded from cosmic rays at the surface. The above neutrino capture process has a reaction threshold of 0.814 MeV, and the argon produced was collected and measured periodically to count the number of neutrino captures. The experiment saw a neutrino flux at a level of only one-third of the solar physics prediction. This is known as the "solar neutrino problem". In hindsight we understand this as a result of neutrino flavor oscillation.

Neutrinos can also be used to assemble a holistic picture of astronomical phenomena. For example, the Type II supernova SN 1987A, was observed not only in visible light, X-ray, and radio frequencies, but also with neutrinos. Two to three hours before the visible light of the supernova explosion reached Earth, three separate neutrino detectors (Kamiokande II [29], IMB [30], Baksan [31]) observed an excess of neutrinos from SN 1987A. The early arrival of neutrinos can be explained by the fact that neutrino emission occurs simultaneously as the core collapses, but optical emission does no occur until the shockwave reaches the stellar surface. In Fig. 1.5, the energy and timing of the burst of events are shown. These observations are consistent with the core-collapse supernova theory [32], and combined with optical observations greatly improved our understanding of the physical processing underlining supernova explosions.

So far we have only discussed astrophysical neutrinos in the MeV range, where they are mostly produced by nuclear reactions. In the high energy regime ($\geq$ GeV), neutrinos are produced through hadronic interactions between protons/ions with matter or radiation. The processes we usually consider are:

$$\begin{aligned} p + \gamma &\rightarrow p + \pi^0 \\ &\rightarrow n + \pi^+ \end{aligned} \tag{1.4}$$

$$\begin{aligned} p + p &\rightarrow p + p + \pi^0 \\ &\rightarrow p + n + \pi^+ \end{aligned} \tag{1.5}$$

$$\pi^0 \rightarrow 2\gamma \tag{1.6}$$

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \rightarrow e^+ + \nu_e + \bar{\nu}_\mu + \nu_\mu \tag{1.7}$$

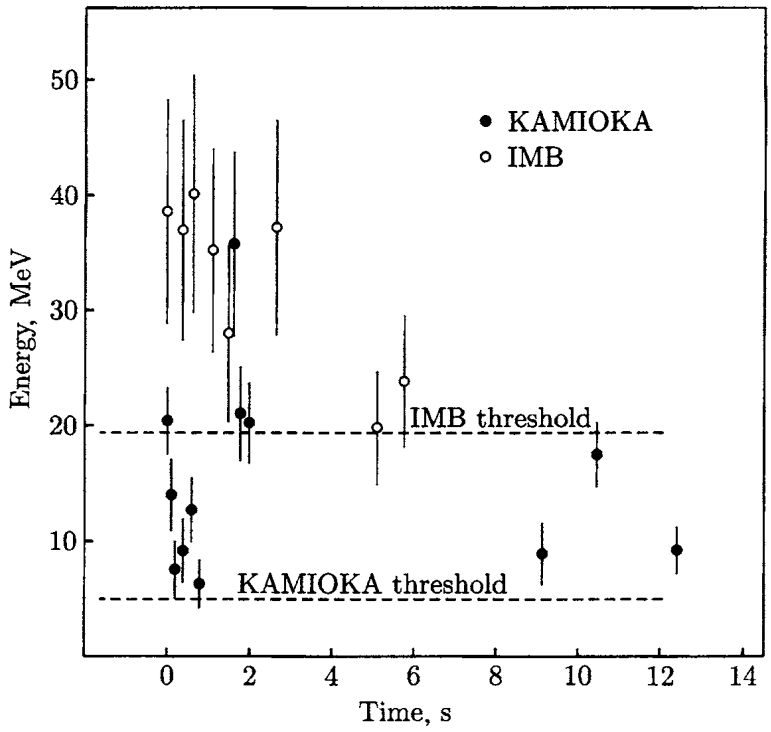


Figure 1.5: Burst of neutrino events from SN 1987A measured by Kamionkande II and IMB water Cherenkov detectors. Kamionkande II had an effective energy threshold of $\sim$ 6MeV, seeing 12 events in two time groups. IMB had an effective energy threshold of roughly 20MeV, seeing 8 events. [33]

Therefore, neutrinos represents unambiguous signs for hadronic accelerators. The spectrum shape and normalization of the neutrino flux then provide insights into the spectrum, composition, cutoff energy, source evolution, and acceleration mechanism for cosmic rays. Such pp and $p\gamma$ interactions also entail the neutrino flavor ratio $\nu_e : \nu_\mu : \nu_\tau$ be $1 : 2 : 0$ at the source. But due to neutrino oscillation, at Earth we can expect a uniform ratio of $1 : 1 : 1$ [34, 35]. Other processes can also contribute to the neutrino flux. For example, neutrons as secondaries in sources that are optically thin to neutron interactions decay into $\bar{\nu_e}$, though with energies $\geq 10^2$ times less than those from pion decays. Still other processes include kaon decays, multi-pion production, and higher resonances. For a complete treatment, see [36, 37]. Numerous astrophysical objects are identified as candidate neutrino sources. These include blazar flares, active galactic nuclei (AGNs), gamma ray bursts (GRBs), galaxy clusters, starburst galaxies (SBGs), tidal disruption events (TDEs), white dwarf (WD) mergers, neutron star (NS) mergers, blackhole (BH) mergers, and the aforementioned supernovae, to name a few [38, 39].

Modern neutrino detectors such as IceCube [40] are capable of pointing back to neutrino sources with approximately one degree resolution for ν_μ-induced track-like events above $\sim$ TeV. This makes associating astrophysical neutrinos with sources possible. The IceCube Neutrino Observatory [41] is a cubic-kilometer glacial ice Cherenkov high-energy neutrino detector built at the South Pole. It consists of 86 cabled "strings", each instrumented with 60 Digital Optical Modules (DOMs) from the depth of 1450 to 2450 meters below the ice surface. Each DOM contains a photomultiplier tube (PMT) that collects Cherenkov radiation from particle interactions in ice. IceCube reconstructs the energy, direction, and flavor of each detected event with the radiation detected from the DOMs. Roughly speaking, we can classify the events into two categories based on the event topology - tracks and cascades. Tracks pertain to the topology generated by muons zipping through IceCube, leaving a near-linear trace of radiation along the direction of momentum. The muons could be from ν_μ charged-current (CC) interactions in the glacial ice, or from cosmic ray muons (secondary muons from cosmic ray and atmospheric molecule collisions). Cascades refer to the quasi-spherical event topology as a result of hadronic showers from neutral-current (NC) interactions between all three flavors of neutrinos and the ice, or hadronic + electromagnetic showers from ν_e or ν_{tau} CC interactions. A rarer "double-cascade" topology from CC interactions of ν_{tau} arises where the CC interactions generates the first shower, and the τ lepton decay generates the second. This represents a unique detectable signature for τ neutrinos.

1.3.1 Astrophysical Neutrinos

An astrophysical neutrino flux was established by IceCube in 2013 [42]. In IceCube, using an all-sky high energy starting event (HESE) sample including both cascades and tracks. This was later confirmed in a complementary analysis using through-going ν_μ-induced muon tracks from the northern hemisphere [43]. Fig. 1.6 illustrates the HESE events and Fig. 1.7 shows the flux measurements from the through-going track analysis. According to the uniform flavor ratio assumption we can use a simple scaling factor of three to convert between the single-flavor and the all-flavor normalization. IceCube has since gathered and analyzed more data, the the current best-fit flux is consistent with a single unbroken power-law model. Assuming the power law fit takes the form:

$$\frac{d\Phi_{\nu+\bar{\nu}}}{dE} = \Phi_{\text{astro}} \cdot (\frac{E_\nu}{100\text{TeV}})^{-\gamma} \cdot 10^{-18}[\text{GeV}^{-1}\text{cm}^{-2}\text{s}^{-1}\text{sr}^{-1}] \tag{1.8}$$

, the best-fit power-law spectral index γ from 7.5 years of HESE is $2.89^{+0.20}_{-0.19}$, and the best-fit all-flavor flux normalization Φ_{astro} is $6.45^{+1.46}_{-0.46}$. For the latest through-going track analysis, almost 10 years of data was used. The best-fit single power law $\gamma = 2.28^{+0.08}_{-0.09}$ and the best-fit single flavor flux normalization $\Phi_{\text{astro},\mu} = 1.44^{+0.25}_{-0.24}$ [44]. Fig. 1.8 shows these power law parameters and their confidence regions from these works respectively.

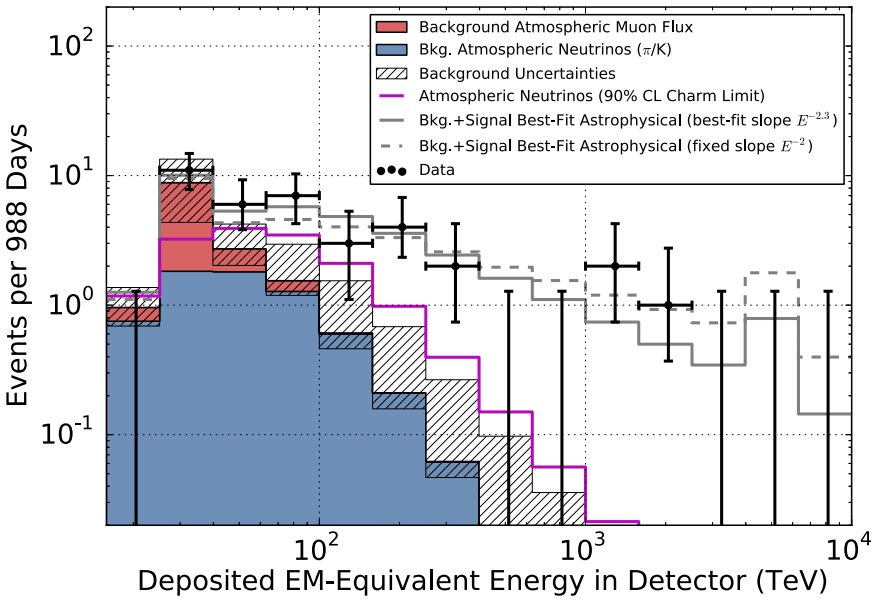


Figure 1.6: Observed events versus background predictions from three years of IceCube HESE sample in [42]. Background uncertainties represent the sum of uncertainties from all backgrounds. An astrophysical component stands out from these backgrounds, and rejects a purely atmospheric explanation at 5.7σ. A total of 37 neutrino candidate events were observed.

To date, no evidence of a steady-state astrophysics point source using the IceCube high energy track events is found [45]. The astrophysical flux remains consistent with being diffuse. One way to piece together a holistic, multi-wavelength and multi-messenger picture of high energy neutrino sources is to coordinate telescope observations triggered on realtime alerts. To achieve this, not only do neutrino telescopes like IceCube need to report precise measurement of the time, energy, and direction of neutrino events, but also telescopes across a wide electromagnetic spectrum should be ready to follow up the neutrino alert in a timely manner. IceCube has set up a realtime alert system [46] that issues alerts through AMON (Astrophysical Multimessenger Observatory Network) [47], which leverages the GRB Coordinates Network — GCN [48], since April 2016. One of the most remarkable discoveries was the high energy neutrino event IceCuve-170922A on 22 September 2017, with an energy of $\sim$ 290 TeV [49]. As shown in Fig. 1.9, this event coincides spatially with a known γ ray source TXS 0506+056, a BL Lacertae (BL Lac) type blazar in flaring state, to within 0.1°. A blazar is a subcategory of AGNs where the AGN's jet points close to our line of sight. A multi-wavelength campaign soon followed from radio to γ ray, contributing to characterizing the variability and energetics of the blazar. The precise reconstructed direction and multi-wavelength follow-ups allow the chance correlation of IceCube-170922A and the blazar to be rejected at a 3σ level. Eventually,

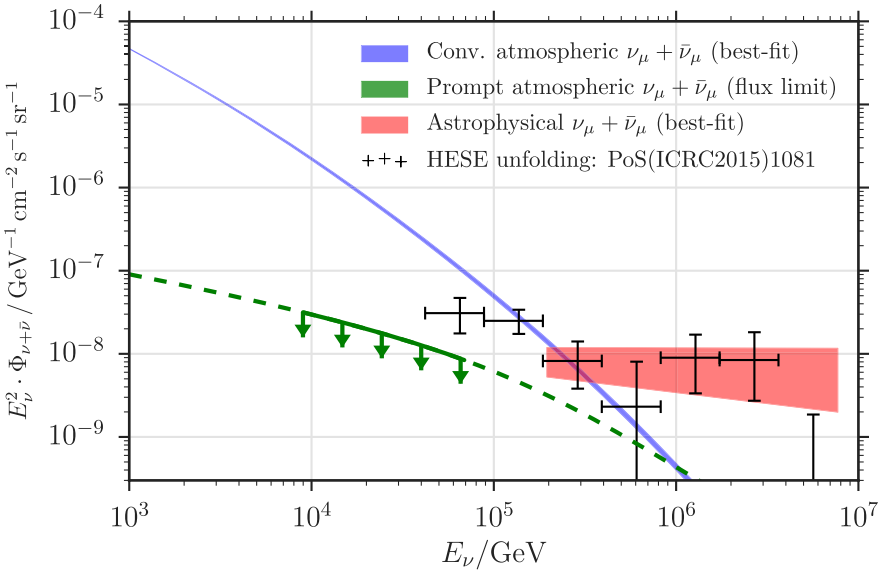


Figure 1.7: The best-fit unbroken power law spectrum of ν_μ track events with six years of IceCube northern hemisphere data (red) in [43]. The width of the astrophysical neutrino fit corresponds to the one-sigma variation of the two model parameters (Φ_{astro} and γ) from the best-fit values, assuming no correlation. The HESE unfolding data is from an updated work to [42]. A purely atmospheric origin of the observed neutrino candidates is rejected at 5.6σ.

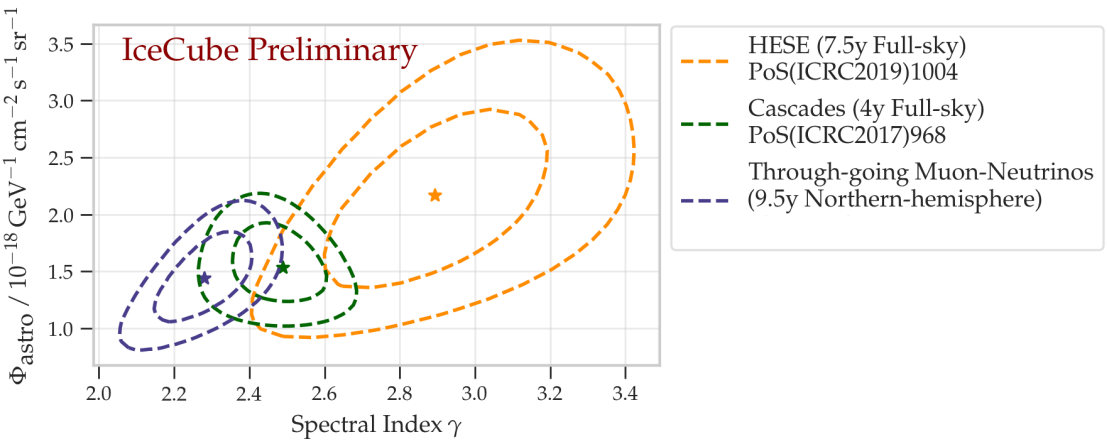


Figure 1.8: Latest single power law fit results for three datasets from IceCube — through-going tracks (blue), starting tracks + cascades (orange, normalization divided by three for single-flavor comparisons), contained cascades (green). The contours represent 68% and 99% uncertainties.

this observation suggests blazars can contribute to a fraction of high energy neutrinos. To establish an unambiguous correlation between high energy neutrinos and blazars, more observation is needed.

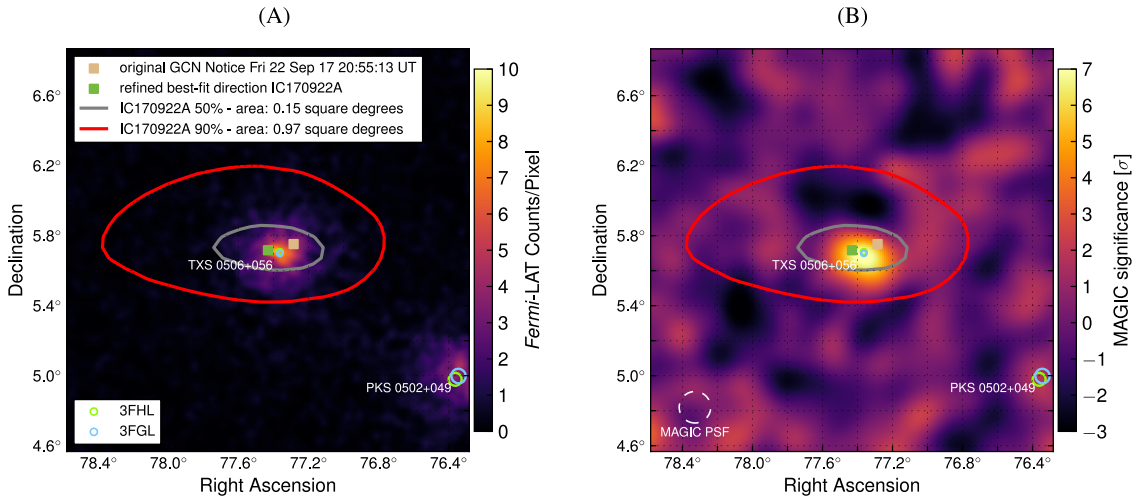


Figure 1.9: Spatial correlation observations with IceCube-170922A from *Fermi*-LAT (A) and MAGIC (B) [49]. Equatorial coordinates of IceCube-170922A is overlaid with 9.5-year photon counts above 1 GeV in the *Fermi*-LAT map, and with γ ray signal significance above 90 GeV in the MAGIC map. The tan box is the neutrino direction from the initial alert and the green box indicates the best-fit position from the follow-up reconstruction. The gray and red contours represent the 50% and 90% neutrino containment regions accounting for both systematic and statistical uncertainties. Green circles represent γ ray sources from the Third Catalog of Hard *Fermi*-LAT Sources (3FHL) [50] and blue circles represent sources from the *Fermi*-LAT Third Source Catalog (3FGL) [51].The circle size indicates the 95% C.L. uncertainty in the source position.

1.3.2 Cosmogenic (GZK) Neutrinos

As noted earlier, the cosmic ray spectrum sees a suppression near the so-called GZK cutoff at $\sim 5\times10^{19}$eV. In principle this could be due to interactions with the CMB photons, or could be a result of maximal energy being reached in cosmic ray accelerators. The detection of a cosmogenic/GZK neutrino flux could discern the two scenarios. The cosmogenic flux arises given that there is a UHECR flux above the cutoff energy. Such UHECRs would undergo the GZK process, essentially a photohadronic interaction with CMB photons, to produce secondary neutrinos. Specifically, the GZK process is:

$$
\begin{aligned}
p + \gamma_{\mathrm{CMB}} &\rightarrow \Delta^* \rightarrow n + \pi^+ \\
n &\rightarrow p + e^- + \bar{\nu_e} \\
\pi^+ &\rightarrow \mu^+ + \nu_\mu \\
\mu^+ &\rightarrow e^+ + \bar{\nu_\mu} + \nu_e
\end{aligned}
\tag{1.9}
$$

As opposed to (1.4), the neutron β decay is explicitly considered here because the pγ_{CMB} photohadronic interaction takes place after the cosmic ray nuclei has escaped the source environment, meaning the daughter neutron will decay instead of interact with ambient matter. Since UHECRs above 10^{20}eV are observed, and CMB photons permeate all space, the cosmogenic flux can be considered as a "guaranteed" flux, the existence of which was first pointed out by Berezinsky and Zatsepin [52]. The threshold energy for the GZK mechanism can be derived from considering the Lorentz invariance of the total four-momentum vectors in the $p + \gamma \rightarrow n + \pi$ process:

$$
(\mathbf{p}_p + \mathbf{p}_\gamma)^2 = (\mathbf{p}_n + \mathbf{p}_\pi)^2 \tag{1.10}
$$

At the production threshold, the neutron and pion are produced at rest with zero momentum. Hence,

$$
\mathbf{p}_p^2 + \mathbf{p}_\gamma^2 + 2\mathbf{p}_p\mathbf{p}_\gamma = -(M_n + M_\pi)^2 c^2 \tag{1.11}
$$

, where M_n and M_π are the rest mass of neutrons and pions respectively. Invoking the on-shell energy-momentum formula $E^2 = (pc)^2 + (Mc^2)^2$, the above equation becomes

$$
-M_p^2 c^2 + 2\mathbf{p}_p\mathbf{p}_\gamma = -(M_n + M_\pi)^2 c^2. \tag{1.12}
$$

To maximize the collision energy, we consider the case where the proton and the photon collide head-on. In the relativistic limit, the 4-momentum of the proton can be written as $\mathbf{p}_p = (E_p/c, E_p/c, 0, 0)$, and the 4-momentum of the photon is $\mathbf{p}_\gamma = (E_\gamma/c, -E_\gamma/c, 0, 0)$. Therefore, we have

$$
-M_p^2 c^2 - 4\frac{E_p E_\gamma}{c^2} = -(M_n + M_\pi)^2 c^2. \tag{1.13}
$$

Solving for E_p, we end up with

$$
E_p = \frac{-(M_p c^2)^2 + (M_n c^2 + M_\pi c^2)^2}{4E_\gamma}. \tag{1.14}
$$

The CMB photon can be estimated using $E_\gamma = k_B T$, where $T = 3$ Kelvin. Plugging in the following numbers:

$$\begin{aligned} M_p &= 938.27\,\mathrm{MeV/c^2} \\ M_n &= 939.57\,\mathrm{MeV/c^2} \\ M_{\pi^\pm} &= 139.57\,\mathrm{MeV/c^2} \\ E_\gamma &= 2.63 \times 10^{-10}\,\mathrm{MeV} \end{aligned} \tag{1.15}$$

, we arrive at the proton threshold energy $E_p = 3 \times 10^{20}$ eV. If we properly integrate over the CMB Planck spectrum and all collision angles, the threshold becomes 5×10^{19} eV. The mean free path of the interaction can be computed as $\lambda = 1/(\sigma_{p\pi} N_{\mathrm{CMB}})$, where $\sigma_{p\pi} = 250mb$ and we take $N_{\mathrm{CMB}} = 5 \times 10^8 m^{-3}$. This gives $\lambda \approx 3$ Mpc. The fractional energy loss of UHE protons in photopion process is $\dfrac{\Delta E}{E} \approx \dfrac{M_\pi}{M_p} \approx 0.1$. Therefore, the spherical horizon to which UHE protons can travel from their origins is approximately 30 Mpc. Equivalently, this means the highest energy cosmic ray protons observed likely originated within tens of megaparsecs from our galaxy. In terms of neutrino flux, roughly 60% of the final flux is generated within the first 50 Mpc [53]. Note that if UHECR are heavier nuclei like iron, the GZK suppression observed would not be due to photopion production. Instead, there will still be an observable suppression due to photonuclear interactions. In this scenario, the absorption of the photon excites nucleons and disintegrates the nuclei through the giant dipole resonance [54]. The photo-disintegration threshold scales with the nucleon number A as $E \geq 10^{18} \times A$ eV. This implies that even for heavy nuclei the cutoff will still be present around 10^{19} to 10^{20} eV [55]. Nevertheless, in this scenario the main energy loss mechanism for cosmic rays will be photo-disintegration and therefore the cosmogenic flux expected from UHE protons will be absent.

Here we lay out a basis for the derivation of the cosmogenic spectrum from propagation of cosmic rays over cosmic distances assuming a pure-proton composition and a uniform isotropic distribution of sources. Integrating over proton energy at source E_p and redshift z, the local neutrino flux of flavor i can be written as [53]:

$$F_i(E_{\nu_i}) = \frac{c}{4\pi E_{\nu_i}} \int\int L(z, E_p) Y(E_p, E_{\nu_i}, z) \frac{dE_p}{E_p} dz. \tag{1.16}$$

$L(z, E_p)$ represents the density of UHECR sources as a function of redshift. Specifically, it is composed of the cosmological source evolution $H(z)$, cosmological expansion $\eta(z)$ (carrying dependencies on cosmological constants matter density Ω_M and vacuum density Ω_Λ), and the normalized version of source spectrum $L_0(E_p) = E^{-\alpha} \times e^{-E/E_{\mathrm{cutoff}}}$ as:

$$L(z, , E_p) = H(z)\eta(z)L_0(E_p). \tag{1.17}$$

The neutrino yield function

$$Y(E_p, E_\nu, z) = E_{\nu_i} dN_{\nu_i}/dN_p dE_{\nu_i} \tag{1.18}$$

can be evaluated using $Y(E_p, E_\nu, z) = Y((1+z)E_p, (1+z)^2 E_\nu, 0)$, where the $(1+z)$ scaling accounts for the redshift at production. As is apparent from this formulation, the final neutrino spectrum depends on the cosmic ray source evolution, source spectrum, and cosmological parameters adopted. Additionally, the composition of cosmic rays obviously plays a vital role in determining the energy span and normalization of the flux. Fig. 1.10 illustrates the resulting cosmogenic neutrino flux for a range of varied parameters, along with some of the projected sensitivities from neutrino detectors. The low flux level indicates that to collect sufficient statistics a detector of order ~ 100's of square kilometer is required. Measurement of the flux can then be used to constrain the flux model parameters above. In the next section, we describe the fundamental requirements to such a large-scale detector, and introduce radio detection in ice.

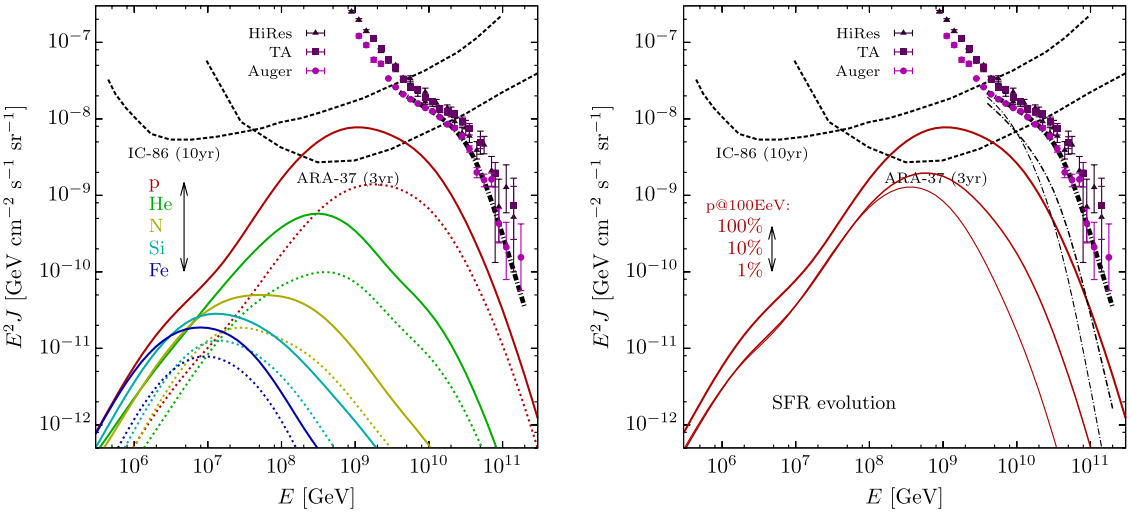


Figure 1.10: Left: cosmogenic flux assuming different compositions for UHECR above 4×10^{18} eV. Solid lines indicates a source evolution rate following the star formation rate, and dashed lines mean no source evolution. Right: (from top to bottom) the red curves indicate the cosmogenic spectrum assuming different fractions of contributing protons [56]. The figure also shows projected sensitivities for leading experiments such as IceCube and ARA.

1.4 Ultra-high Energy Neutrino Detectors

A UHE neutrino detector typically relies on the detection of neutrino-nucleon deep inelastic scatterings (DIS). DIS induce particle showers bearing a fraction of the primary incident particle, depending on the interaction channel (NC or CC) and the inelasticity y. For NC interactions of all three neutrino flavors or CC interactions of ν_μ or ν_τ, hadronic showers are induced. For CC interactions with ν_e, the final shower consist of both the hadronic and the electromagnetic showers. Regardless of the shower type, the shower particles and their electromagnetic emissions are the detectable signature of the primary neutrino. Due to the low neutrino flux ($\sim$<1 km^{-2}yr^{-1}), a large target mass for DIS conversion is usually required . This implies a vast natural formation such as the atmosphere, the sea, mountain ranges, or glacial ice. A detector covering a large area as such therefore must consider its scalability and cost required for extensive deployment. The location should also be as background-free as possible. Here we discuss two detection channels — Cherenkov radiation and Askaryan effect, with an emphasis on the latter.

Water Cherenkov detectors can search for, in addition to cosmic rays, UHE neutrino-induced air showers and/or tau-induced air showers, where the tau lepton comes from tau neutrino conversions in nearby mountains or the Earth's crust at near horizontal angles. This latter case is known as "earth-skimming". Restricting candidate events to highly-inclined young showers, the Pierre Auger Observatory has placed stringent upper limits on the single-flavor UHE neutrino flux [57]. Cherenkov detectors in glacial ice such as IceCube [58] can be used to search for UHE neutrinos as well. After analyzing data nine years of data in its "extremely-high-energy" diffuse neutrino flux search, IceCube placed the most stringent limit below $\sim 2 \times 10^{18}$eV in the UHE regime.

The next section describes the Askaryan effect and presents a simple derivation. We will then discuss the detection of this signal in various natural settings, with a focus on glacial ice where ARA is deployed.

1.4.1 The Askaryan Effect

Named after its conceiver G.A. Askaryan [59, 60], the Askaryan effect predicts a $\sim$ 20% negative-over-positive charge asymmetry being developed in charged particle showers induced by high energy DIS in dense media. The electromagnetic emission from charged particles in the net negatively charged shower adds up coherently for wavelengths much larger than the lateral shower size, typically $O(10)$ cm, giving rise to a unique signal with enhanced low-frequency radio-wavelength power centered around the Cherenkov cone with respect to the shower axis in the 0.1 to 1 GHz frequency range. In the time domain, the Askaryan signal is a bipolar broadband pulse of width $\sim O(100\,\text{ps})$. In terms of polarization, the signal is highly polarized in the plane formed by the line-of-sight and the shower axis. The charge asymmetry is due to

Compton scattering, atomic electrons knocked out by shower electrons/positrons, or pair annihilation with electrons in the media where the particle shower develops.

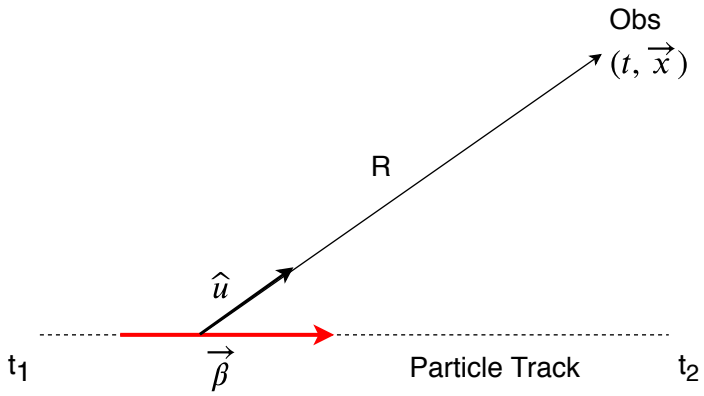


Figure 1.11: Schematic diagram of the particle track being considered in the endpoint formulation used to evaluate the Askaryan radiation fields.

To derive the Askaryan radiation fields in the far field, we follow the endpoint formulation [61]. In this formulation, we consider a relativistic particle accelerated along a track from time t_1 to t_2 with instantaneous velocity of β. The goal is to derive the electric and magnetic fields for an observer at time t and distance R in the direction relative to the particle specified by the unit vector $\hat{\mathbf{u}}$. Fig. 1.11 illustrates this geometry. Using Maxwell's equations, the Liénard-Wiechert potentials can be obtained for the relativistic case. This gives the electric field at $(t, \mathbf{x})$:

$$\mathbf{E}(t,\mathbf{x}) = \frac{e\mu_r}{4\pi\epsilon_0}\left[\frac{\hat{\mathbf{u}} - n\boldsymbol{\beta}}{\gamma^2|1-n\hat{\mathbf{u}}\boldsymbol{\beta}|^3R^2}\right]_{\text{ret}} + \frac{e\mu_r}{4\pi\epsilon_0}\left[\frac{\hat{\mathbf{u}}\times((\hat{\mathbf{u}}-n\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}})}{|1-n\hat{\mathbf{u}}\boldsymbol{\beta}|^3R}\right]_{\text{ret}} \tag{1.19}$$

, where n is the index of refraction of the medium, μ_r is the relative permeability μ/μ_0, and 'ret' indicates the evaluation at retarded time $t^{'} = t - nR/c$. The term on the left represents the static field that does not propagate to the far field. Therefore we will only consider the radiation term on the right. At any $(t, \mathbf{x})$, the magnetic field can be calculated as

$$\mathbf{H}(t,\mathbf{x}) = \sqrt{\frac{\epsilon_r\epsilon_0}{\mu_r\mu_0}}\hat{\mathbf{u}}\times\mathbf{E}(t,\mathbf{x}). \tag{1.20}$$

Next, working in the frequency domain, the Fourier transformed radiation term becomes

$$\mathbf{E}(\omega,\mathbf{x}) = \frac{e\mu_r}{2\pi\epsilon_0c^2}i\omega\int dt_r c\boldsymbol{\beta}_\perp\frac{e^{i\omega(t_r+R(t_r)/c)}}{R(t_r)}. \tag{1.21}$$

$\boldsymbol{\beta}_\perp$ is the particle velocity projected to the plane perpendicular to the observing direction. Integrating over the endpoints, we obtain

$$\mathbf{E}(\omega, \mathbf{x}) = \frac{e\mu_r}{2\pi\epsilon_0 c^2} i\omega \frac{e^{ikR}}{R} e^{i(\omega - \mathbf{kv})t_1} \mathbf{v}_\perp \left[\frac{e^{i(\omega - \mathbf{kv})\delta t} - 1}{i(\omega - \mathbf{kv})} \right] \tag{1.22}$$

If we consider the conditions for the Askaryan effect, namely that ω is small and the observer is near the Cherenkov angle $\theta_c = cos^{-1}(1/\beta n)$, then the above becomes

$$R \cdot \mathbf{E}(\omega, \mathbf{x}) = \frac{e\mu_r i\omega}{2\pi\epsilon_0 c^2} \mathbf{v}_\perp \delta t e^{i(\omega t_1 - \mathbf{kr}_1)} e^{ikR} \tag{1.23}$$

Under this last condition, the total electric field at the observer is simply the coherent sum of contributions of all tracks $\mathbf{v}_\perp \delta t$ from the excess shower charge. Note that although the Askaryan signal is centered around the Cherenkov angle, it is of a 'bremsstrahlung' nature from the acceleration of the particles along the endpoints, a mechanism distinct from the well-known Vavilov-Cherenkov radiation.

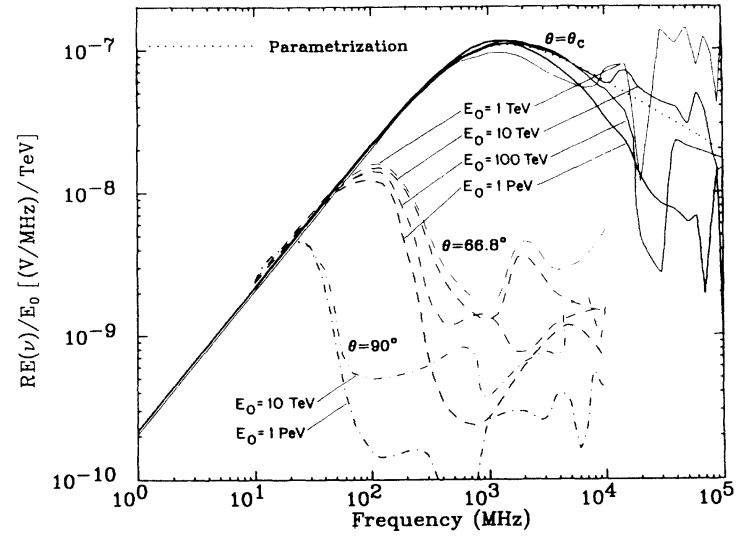
Monte Carlo simulations can follow this prescription and simulate the Askaryan radiation by taking into account all particle track lengths. The pioneering work is referred to as the ZHS simulation demonstrated in [61]. Fig. 1.12 shows key results from such simulations — the frequency and angular spectrum of the signal. Notice that the field strength scales roughly linearly with the frequency until the coherence can no longer be maintained even at θ_c around 1 GHz. This justifies the frequency window of 150 to 850 MHz chosen by ARA. Furthermore, the angular pattern suggests that a full-band frequency content analysis can reveal the off-cone angle in the event of a detection. The in-ice frequency spectrum for a shower with energy E_0 at the Cherenkov cone below 1 GHz can be parametrized as

$$R|\mathbf{E}(\omega, R, \theta_c)| = 1.1 \times 10^{-7} \frac{E_0}{1\text{TeV}} \frac{\nu}{50\text{MHz}} \frac{1}{1 + 0.4(\nu/500\text{MHz})^2} \text{V / MHz.} \tag{1.24}$$

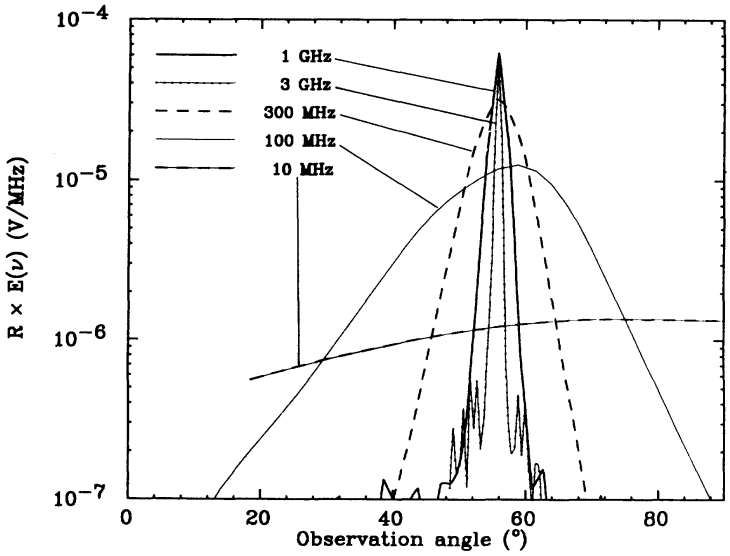
The angular distribution can then be parametrized as

$$E(\omega, R, \theta) = E(\omega, R, \theta_c)\ exp\left[-\frac{1}{2} \left(\frac{\theta - \theta_c}{2.4^\circ \times 500\text{MHz}/\nu} \right)^2 \right] \tag{1.25}$$

The energy in the Askaryan signal therefore scales with E_0^2, where E_0 scales linearly with the charge excess which also is linearly proportional to the incident particle energy. The fact that the Askaryan emission energy is $\propto E^2_{\text{primary}}$ is different from the case of the Vavilov-Cherenkov radiation, in which the emission energy is $\propto E_{\text{primary}}$.



(a)



(b)

Figure 1.12: (a): The energy-normalized Askaryan electric field strength from charge excess showers from 1 TeV up to 1 PeV with different amounts of deviation from the Cherenkov angle in ice as simulated by the ZHS code. At the Cherenkov angle the coherence of the emission is kept to frequencies up to 2 GHz, and is well-characterized by the parametrization presented in the text. (b): the distribution of the Askaryan electric field as a function of the observation angle (polar angle of the observation relative to the shower axis) for a one PeV shower.

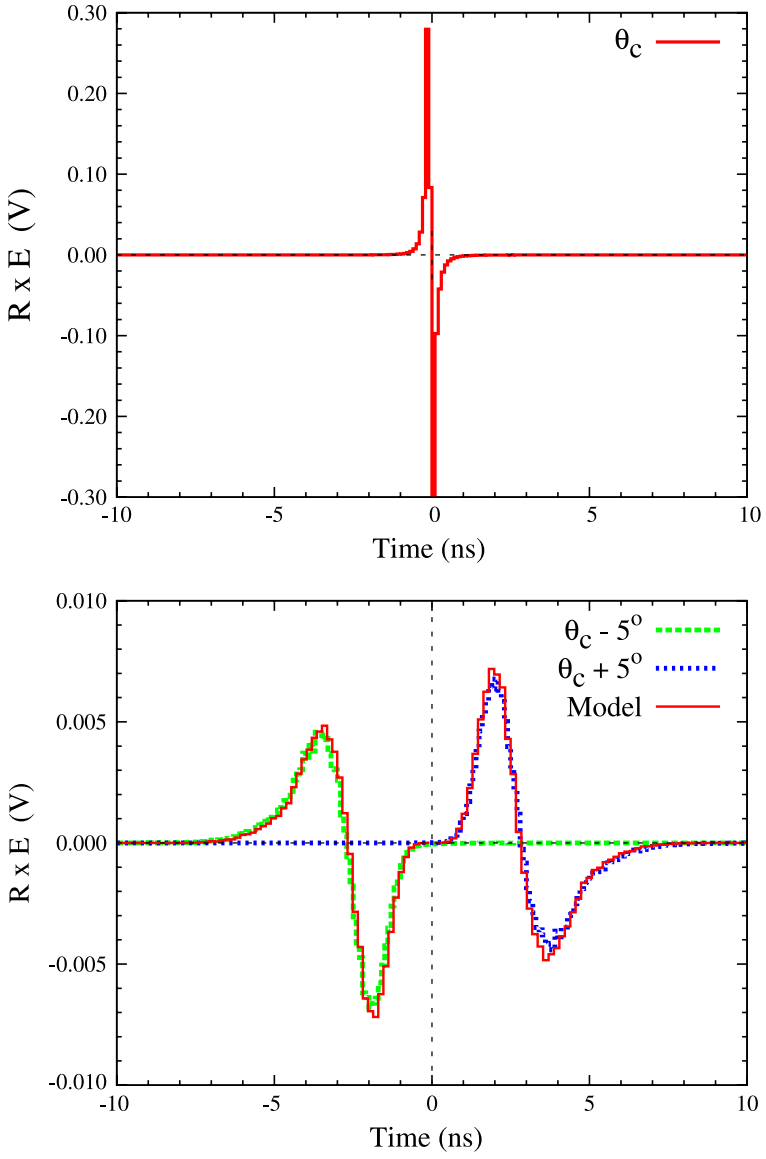


Figure 1.13: Time domain Askaryan signal waveform simulated at a shower energy of one PeV with the ZHS code. Top: electric field versus time for an observer at the Cherenkov angle. Bottom: observer at ±5° from the Cherenkov angle. The model (red histogram) represents a model parametrization of the signal presented in [62].

One can also use full Monte Carlo simulation to compute the in-ice Askaryan signal in the time domain. Such a signal from a one PeV shower is illustrated in Fig. 1.13. However, semi-analytical approaches can be taken for a quicker computation. In ARA, the simulation estimates the field strength following the treatment of [63], where the idea is to only use Monte Carlo to generate a longitudinal charge excess profile that is later convoluted with a parametrized *form factor* that accounts for the shower lateral span. However, to expedite the process even more in ARA the Monte Carlo shower generation process is replaced by functional form shower profiling. This will be discussed further in section 3.2. Finally a vector potential in the far field at all angles can be obtained as

$$\mathbf{A}(\theta,t) = \frac{\mu}{4\pi R} sin\theta \hat{\mathbf{p}} \int_{-\infty}^{\infty} dz' Q(z') F_p \left(t - \frac{nR}{c} - z' \left[\frac{1}{v} - \frac{ncos\theta}{c} \right] \right). \tag{1.26}$$

In the above expression, $\hat{\mathbf{p}}$ is the unit vector in the polarization direction $-\hat{\mathbf{u}} \times (\hat{\mathbf{u}} \times \hat{\boldsymbol{\beta}})$. $Q(z')$ is the longitudinal charge excess profile from simulation. The form factor is given by

$$F_p(t - \frac{nR}{c}) = \frac{4\pi R A(\theta_c, t)}{\mu sin\theta_c \int dz' Q(z')} \tag{1.27}$$

where, for electromagnetic showers,

$$RA(\theta_c,t) = -4.5 \times 10^{-14} E[\text{Vs/TeV}] \times \begin{cases} t \geq 0,\ exp(-|t/\text{ns}|/0.057) + (1 + 2.87|t/\text{ns}|)^{-3} \\ t < 0,\ exp(-|t/\text{ns}|/0.030) + (1 + 3.05|t/\text{ns}|)^{-3.5} \end{cases}$$

, and for hadronic showers [64]

$$RA(\theta_c,t) = -3.2 \times 10^{-14} E[\text{Vs/TeV}] \times \begin{cases} t \geq 0,\ exp(-|t/\text{ns}|/0.043) + (1 + 2.92|t/\text{ns}|)^{-3.21} \\ t < 0,\ exp(-|t/\text{ns}|/0.065) + (1 + 3.0|t/\text{ns}|)^{-2.65}. \end{cases}$$

Then the electric field is simply $\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}$.

The Askaryan signal and its properties have been confirmed in silica sand [65] and rock salt [66]. The in-ice confirmation in 2006 at the SLAC National Accelerator Laboratory was performed by the ANtarctic Impulsive Transient Antenna (ANITA) collaboration [67], and properties including the energy dependence, frequency spectrum, and angular dependence were checked to be in good agreement with theoretical predictions. At the time of writing no Askaryan signal from naturally-occurring particles in ice has been confirmed. However, coherence emission due to the same charge excess mechanism in air showers has been observed in the atmosphere by large radio air shower arrays [68].

1.4.2 Radio Neutrino Detectors

To detect the Askaryan signal from UHE neutrinos, experiments can take advantage of natural target masses that are radio-transparent. Here we focus on the use of glacial ice as the target mass, and finish the section with a brief description of detectors searching for the Askaryan pulses from the Moon and large water bodies.

1.4.2.1 Detectors in Glacial Ice

The glacial ice is an ideal target for radio detectors due to its transparency to radio frequency (RF) transmission. Sites such as Antarctica and Greenland possess clean glacial ice where the depth at locations

can be up to four kilometers and nearly three kilometers respectively, and are nearly background-free due to their remoteness. The ANITA [69] experiment uses a scientific balloon to carry a payload consisting of broadband antennas and hover over the Antarctic continent at an altitude of ∼37km to cover approximately 1 million square kilometers of area. Currently ANITA-IV places the best UHE neutrino flux upper limit in the range $\sim 4\times10^{19}$ to 10^{23}eV [70]. To lower the energy threshold, instead of mounting antennas to a balloon payload where the amount of power from Askaryan emission inevitably undergo inverse-square law scaling before reaching the payload, the antennas can be embedded in-ice or deployed on the ice surface, essentially bring the detector closer to neutrinos. Such an approach also avoids the seasonal availability issue from balloon flights and grants virtually year-round access to the detector for maintenance needs. One experiment in this latter category is the Antarctic Ross Ice-Shelf Antenna Neutrino Array (ARIANNA) [71]. It deploys antenna clusters on the surface of the Ross Ice-Shelf in Antarctica to utilize not only the ice but also the sea water underneath the ice shelf for potential reflected signal.

The South Pole has a history of neutrino telescopes from the AMANDA experiement [72] to IceCube, providing a basis of infrastructure for constructing new detectors. Furthermore, the temperature and index of refraction of the South Pole ice have been well-characterized. ARA is deployed near the South Pole and embed clusters of antenna underneath the ice surface to take advantage of the long radio attenuation length in the 3-km deep ice sheet and the established infrastructure. The electrical properties of glacial ice will be explored in the next section.

1.4.2.2 Electrical Properties of Glacial Ice

The electrical properties of a dielectric including ice can be characterized by the permittivity $\epsilon = \epsilon_0(\epsilon' - i\epsilon'')$, where ϵ_0 is the permittivity in free space, and $\epsilon_r \equiv \epsilon' - i\epsilon''$ is the complex relative permittivity, or dielectric constant. The index of refraction is then:

$$n = \sqrt{\mathrm{Re}(\epsilon_r)} = \sqrt{\epsilon'} \tag{1.28}$$

In the RF regime, aside from geometrical spreading, the energy loss of electromagnetic waves can be mainly attributed to absorption during propagation. The absorption arises due to the alternating electric field polarizing the medium molecules that absorbs the energy through rotation or resonance that leads to the scattering of energy. This absorption loss can be quantified using the attenuation coefficient α [73]

$$\alpha[\text{dB/m}] = 8.686\frac{\omega}{2c}\sqrt{\epsilon'}tan\delta \tag{1.29}$$

$$\text{Loss tangent } tan\delta \equiv \epsilon''/\epsilon' \tag{1.30}$$

$$L_\alpha = 20/(\alpha ln 10) \propto \frac{1}{\nu}\frac{\epsilon'}{\epsilon''} \tag{1.31}$$

, where L_α is the attenuation length, the distance over which the electric field strength would drop by a factor of e when traveled.

The attenuation length as a function of in-ice depth or frequency is usually most relevant to our application. These dependencies are driven by two main factors — the ice temperature T_{ice} and the frequency. Since $\epsilon' \approx 3.19$ in the whole RF band above 100 MHz and depends only weakly on temperature (2% decrease from 0°C to -75°C, see [73]), the dependencies almost entirely come from the explicit frequency factor $1/\nu$ and $\epsilon''(\nu, T_{\text{ice}})$. Hence, to extract L_α as a function of depth for any frequency, one approach is to first conduct in-situ measurements of the average attenuation length $< L_\alpha >$ in the range of depths of interest. Next, one can assume that L_α follows an ice temperature-versus-depth profile such as Fig. 1.14 to work out the L_α depth profile. Following this procedure, ARA used an in-ice pulser and a TestBed station that are 3.16 km apart to produce the result shown in Fig. 1.15. At 300 MHz, $< L_\alpha >= 1660\ m$ above the top 1500 m, and for the full depth range $< L_\alpha >= 820\ m$. This radio transparency, especially in shallow ice, speaks to the uniqueness of the South Pole as a detector site, ensuring that the needed $O(100km^2)$ detector area can be covered efficiently with $O(10)$ relatively simple radio detector stations. Notice that the attenuation length increases for shallower ice, consistent with previous studies that ϵ'' is a strong function of T_{ice} and decreases as T_{ice} drops [74]. The ice temperature monotonically decreases from bottom to top, and thus the L_α shape. The ARA simulation package `AraSim` takes the depth-dependency into account for realistic effective area estimation.

In terms of the frequency scaling of L_α, one can in principle derive from measured values of ϵ'' versus frequency. However, data in the 100's of MHz range are insufficient and inconclusive [76]. As such, direct in-situ measurements of L_α were carried out at several Antarctic locations including the South Pole. Fig. 1.16 demonstrates the result from vertical bottom-bounce experiments at the South Pole [77]. In general the attenuation becomes more prominent as the frequency increases, but not to the point where L_α simply scales with $1/\nu$ and ϵ'' shows no frequency dependence. Instead L_α exhibits near-linear scaling with frequency. Since this dependency is not a strong function, `AraSim` uses only the attenuation length at 300 MHz from the same TestBed measurement in the previous paragraph for simplicity when simulating the propagation of Askaryan signals. This also avoids the need for an assumption on the unknown bedrock reflection coefficient which bottom-bounce experiments require.

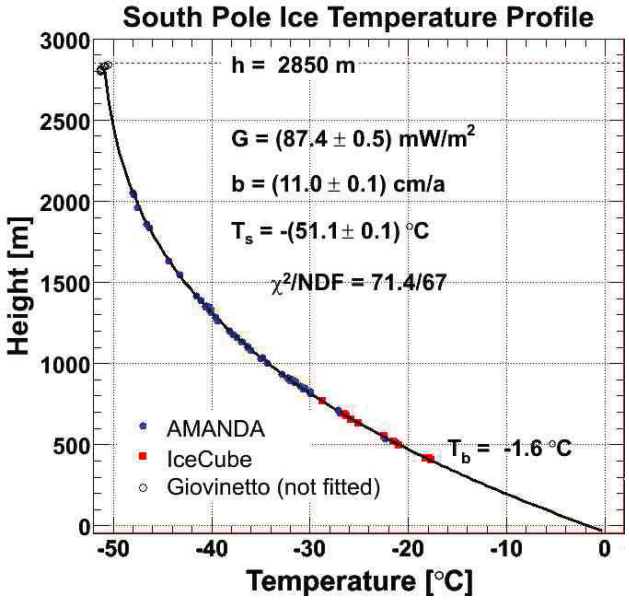


Figure 1.14: Measurements of the glacial ice temperature profile from deep AMANDA and IceCube boreholes [75].

Lastly, the depth dependence of the index of refraction n should be considered. As explained earlier, the index of refraction has negligible dependence on either frequency or ice temperature. However, the index of refraction follows closely the ice density $\rho(z)$. The ice sheet is formed by the compacting of snowfall. Therefore, the ice density increases rapidly in the first $\sim$ 150 m and asymptotically approaches the deep ice value, as shown in Fig. 1.17. The measured index of refraction as a function of depth in the firn and beyond at the South Pole is shown in Fig. 1.18. The firn is defined as the depth above which air bubbles in ice can still exchange air with the surface, and is around 100 m. Notice that although a strong correlation can be drawn from the density profile and the index of refraction profile, field measurements of $n(z)$ need not depend on prior knowledge of $\rho(z)$. For example, in [79], $n(z)$ is determined by directly measuring the propagation time of electromagnetic signals through ice at different depths. In this thesis, an exponential model is used to fit the $n(z)$ data from [79], and the fitted model is

$$n(z) = 1.78 - 0.43e^{0.0132z} \tag{1.32}$$

, where $z = 0$ for the ice surface and $z < 0$ for the direction in-ice.

Due to the depth gradient in the index of refraction, electromagnetic waves propagating through depth differences do not travel in a straight path but instead undergo bending effects according to Snell's law. This ray-bending feature of the South Polar ice is illustrated in Fig. 1.19. One apparent outcome of this

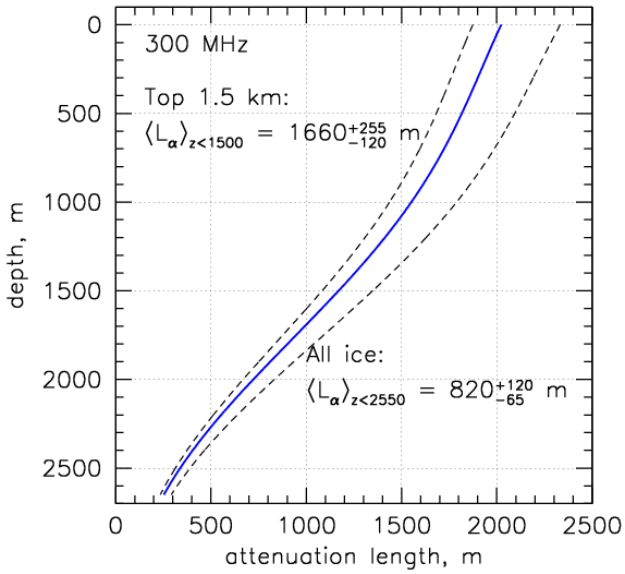


Figure 1.15: Average attenuation length $< L_\alpha >$ as measured by the ARA TestBed coupled with the controlled signal from a preinstalled deep (2450 m) broadband pulser from IceCube string 1 (pulser later referred to as "IC1D"). $< L_\alpha >$ is calculated from the Friis equation, and the depth profile is extracted assuming that the dependence is dominated by the $T_{\text{ice}}(z)$ profile. The measurement assumes a frequency of 300 MHz. Shallow ice is also demonstrated to have a much longer attenuation length.

feature is the need to consider such bending when reconstructing received signals. `RaySolver` codes have been developed to semi-analytically calculate the bending and obtain ray "solutions" that represent viable paths connecting one point in ice to another. `RaySolver` will be described in more detail in section 3.3. In section 4.2.1, a fast arrival time computation software application is introduced to account for the distorted signal timing arising from ray-bending.

Another critical consequence of ray-bending is the introduction of "shadow regions". For a given receiving target in a given South Polar $n(z)$ model, no physical solutions for propagation exists that connects the target to an emission source in its shadow region. The depth dependence of the shadow region is illustrated in Fig. 1.20. This has important implications to the experiment design since shadow regions make a certain volume of ice inaccessible as target mass and in turn reduce the detector sensitivity. The shadowing effect is the strongest in the firn region, and this drives the choice of the depth of ARA stations to go below 150 m, finally deploying at around 200 m. The spatial boundary separating the shadow region and the accessible region is referred to as the "shadow boundary".

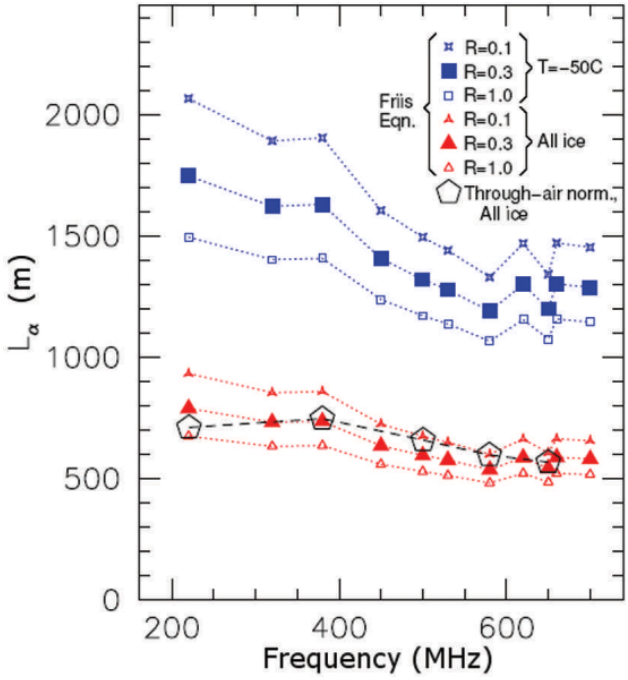


Figure 1.16: Attenuation length $< L_\alpha >$ versus radio frequency from [77]. Lower data (red) represent averaged L_α over the full ice depth with different bedrock reflectivity assumptions. Upper data (blue) show the deconvolved estimate of L_α at $T_{\text{ice}} = -50°C$, using the $\epsilon''(\nu)$ parametrization from [74] and the temperature profile from [78]

Having introduced the idea of shadow regions, it should be noted that for any propagation endpoints in ice where both the source and the target are accessible, namely outside of the shadow region, we consider two ray solutions that connect the two points. The first solution travels the minimal distance consistent with Snell's law to connect the two points, and is called the "direct ray". The second solution undergoes more bending or could reflect off the ice surface. It is referred to as the "refracted/reflected ray". Again, these are exemplified in Fig. 1.19. If the source is at the shadow boundary, the two solutions coalesce and become degenerate. Inclusion of the refracted ray not only adds to the detectable phase space, but also could provide additional information that can be used in interaction vertex reconstruction. In section 4.4.2 this will prove to be valuable.

1.4.2.3 Detectors in Other Media

Other experiments attempt to observe the UHE neutrino flux in different target media. The lunar Askaryan effect is a pursued method. In this method, ground-based radio telescopes or satellites search for

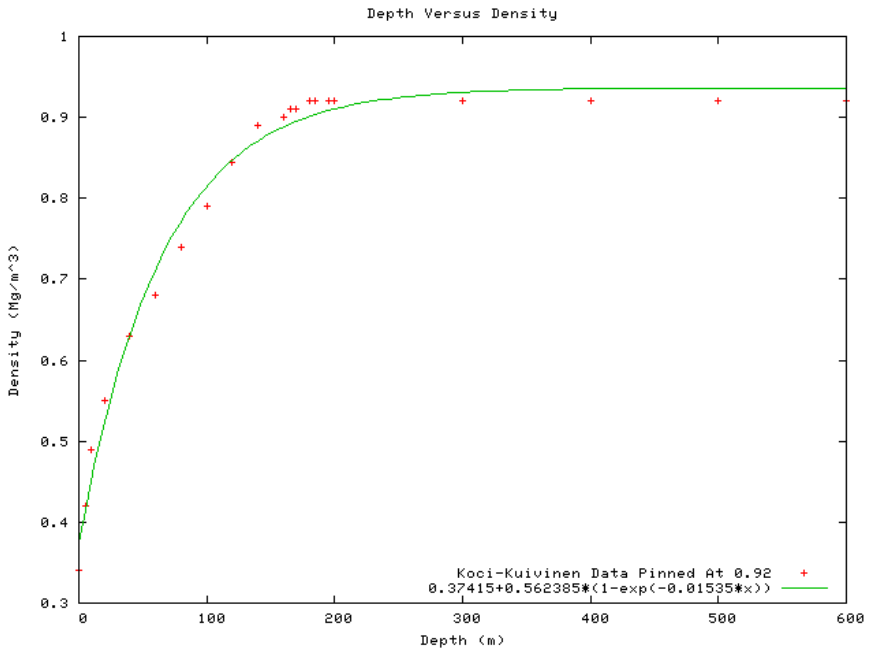


Figure 1.17: Density as a function of in-ice depth measured in [80] from ice cores, compiled and fitted by [81]. The red crosses mark the data points. The green line represents a model fit to the data.

RF Askaryan pulses arising from neutrino interactions in the Moon's outer layer. The techqniue is most sensitive to the spectrum above $\sim 10^{20}$eV or higher [83, 84, 85, 86, 87].

In water, the thermal-acoustic effect [88] manifests as quasi-instantaneous, microscopic heating of a small volume of water due to the deposited energy from the neutrino-induced shower. The heating, of order $\sim \mu K$ causes the volume of water to expand and give rise to an acoustic pulse. The acoustic signal peaks at 5 10kHz, and has a bipolar time-domain shape of amplitude several mPa. Hydrophone arrays have been deployed in sea water and fresh water bodies, and such detectors are also most sensitive in the region above 10^{20} eV or higher [89].

The two kinds of UHE neutrino detectors introduced here are not capable of extending the observsation of the IceCube astrophysical flux and probing the transition of the diffuse neutrino flux from astrophysical to cosmogenic origins due to their high energy threshold. On the other hand, ground-based detectors such as ARA is in the niche for supplementing the IceCube flux and placing meaningful upper limits or making discoveries in the 10 PeV to 100 EeV range. In the next chapter, the ARA detector which produced the data that are analyzed in this thesis will be characterized in detail.

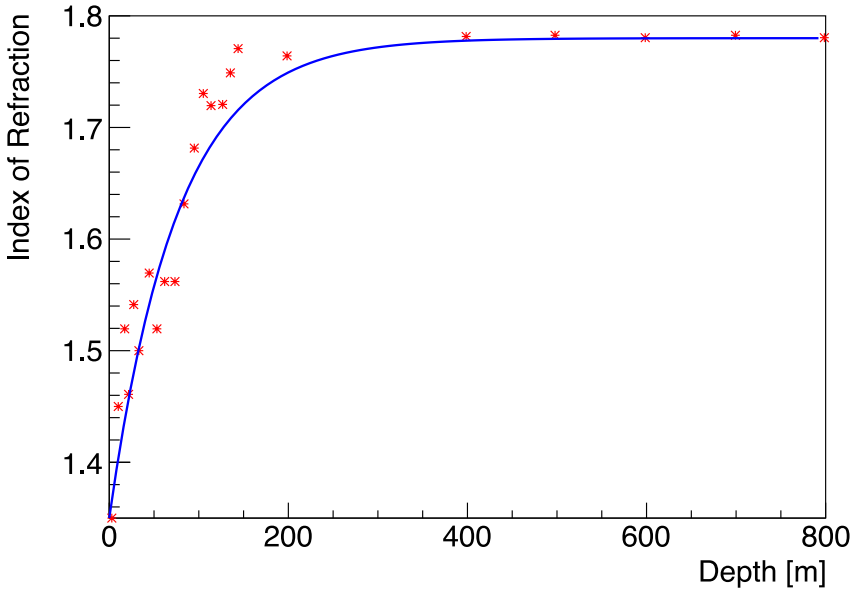


Figure 1.18: Index of refraction measurement at the South Pole made by the Radio Ice Cherenkov Experiment (RICE) by probing signal from radio transmitters lowered to various depths. In this previoud work [64], the data points (red) are digitized from [79], and fitted with an exponential model (green) that produces Eqn. 1.32.

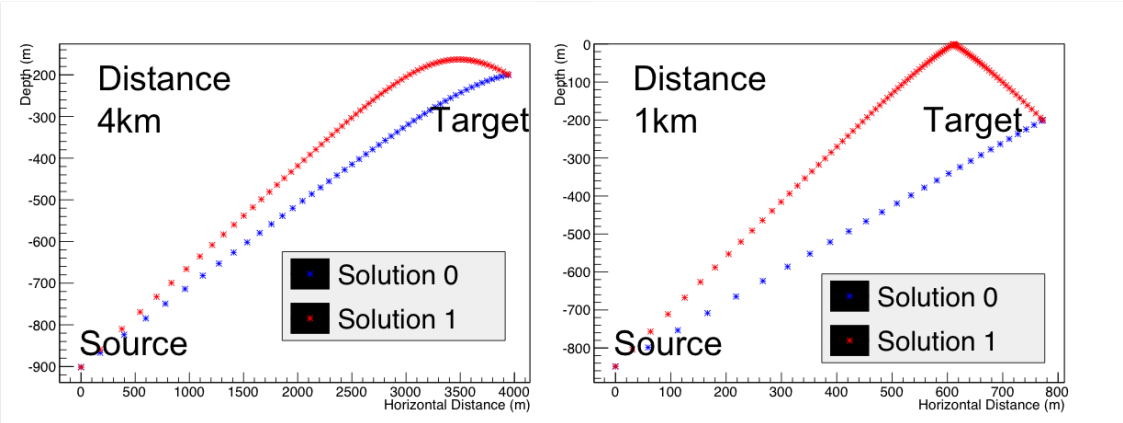


Figure 1.19: Ray paths connecting sources to a target antenna at a depth of 200m. Left: the horizontal displacement between the source and the target is 4 km. The blue path (Solution 0) marks the direct solution, while the red path (Solution 1) marks the second, in this case refracted, solution. Notice how the refracted solution is received at the target as an down-going signal. Right: the horizontal distance between the source and the target becomes 1 km and the vertical distance remains close to that of the left. Here Solution 1 reflects off the surface (depth = 0). These figures assume the $n(z)$ model in Eqn. 1.32

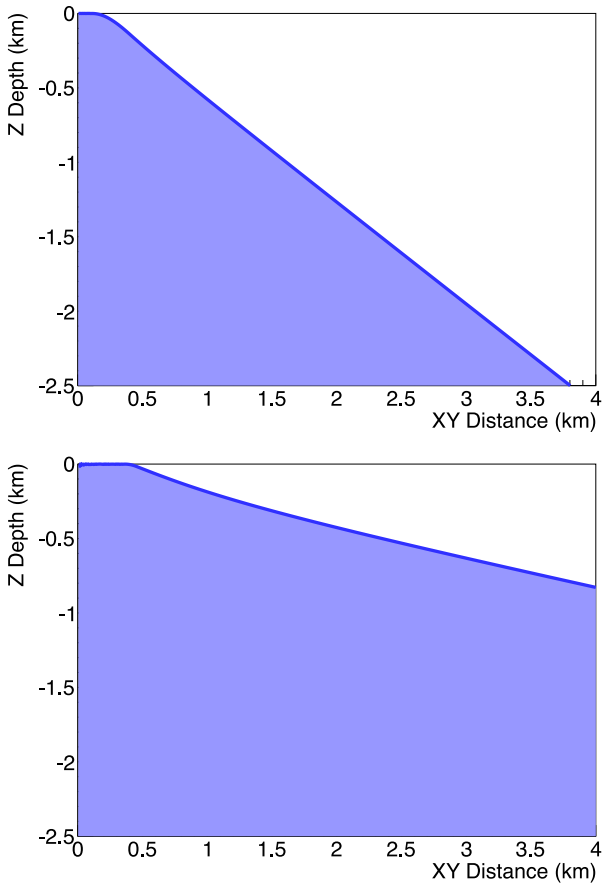


Figure 1.20: Blue: accessible region for ray solutions to the target antenna in consideration. White: shadow region. Top: a target antenna at a depth of 25 m. Bottom: antenna at 200 m depth [82]. The effective volume of an ARA station increases by a factor of three going from an antenna depth of 25 m to 200 m for $E_\nu = 10^{18} eV$, and a factor of four at $E_\nu = 10^{19} eV$.

Chapter 2

Askaryan Radio Array

The Askaryan Radio Array (ARA) is the detector with which we search for UHE neutrinos in this work. In this chapter, an overview of the ARA detector design and critical hardware components will be presented. ARA is designed and built by participating institutions in the ARA collaboration with a combination of custom and off-the-shelf hardware controlled by self-developed firmware and software. The array consists of standalone detector units called "stations". Each station operates as an independent neutrino detector equipped with antenna clusters, a data acquisition (DAQ) system and its own communication system. While at the time of writing there are five stations in the array with varied depths and features, the data used in analysis here are produced solely from stations number two (A2) and number three (A3). These two stations are built with the same design, and indeed perform similarly in the field. As a a result this chapter will focus on them. We first describe the overall array, and proceed to describe the individual ARA station.

2.1 Array Design

The initial design of ARA planned to deployed 37 in-ice stations in a hexagonal grid arrangement with inter-station spacing of 2km (Fig. 2.1). This would cover approximately $100km^2$ of surface area and view more than 200 Gt of radio-transparent glacial ice. As a result of this large detector mass, the array would be sufficiently sensitive for cosmogenic flux measurements. Each station is cable-linked to the IceCube Lab (ICL), which is about halfway amid the edge of ARA and the Amundsen-Scott station, for power and communication. The 2 km inter-station spacing is determined by simulation at $E_\nu = 10^{18}eV$ taking into account attenuation and the effect of the shadow region, which shows marginal overlap of visible volume between two stations at this distance, thus maximizing the effective volume. The simulation result is shown in Fig. 2.2 [90].

The construction of the array has taken a phased approach, where currently five stations are deployed and a new design — the Radio Neutrino Observatory (RNO) will take over in the future [91]. Fig. 2.3 shows the as-built layout of the five-station array.

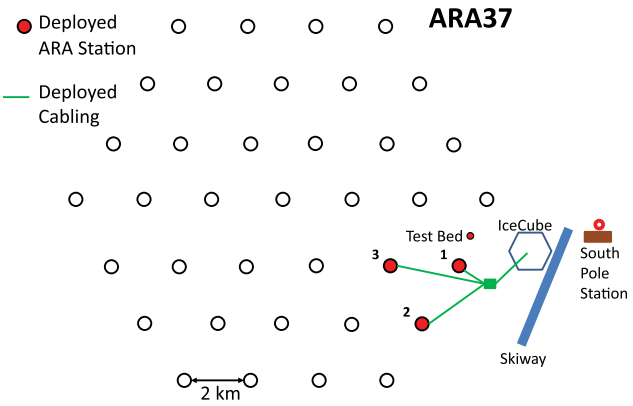


Figure 2.1: The planned ARA detector with 37 standalone detector stations. The map shows the detector's relative position to the South Pole Station, and the ICL. Cables (green) indicate power and data transfer lines. The green box indicates the location of "Wind Turbine 3" (WT3) where a junction box is situated that serves as a midway point for the cables. It is clear from this layout more than $O(100)km^2$ of ice will be covered by the detector.

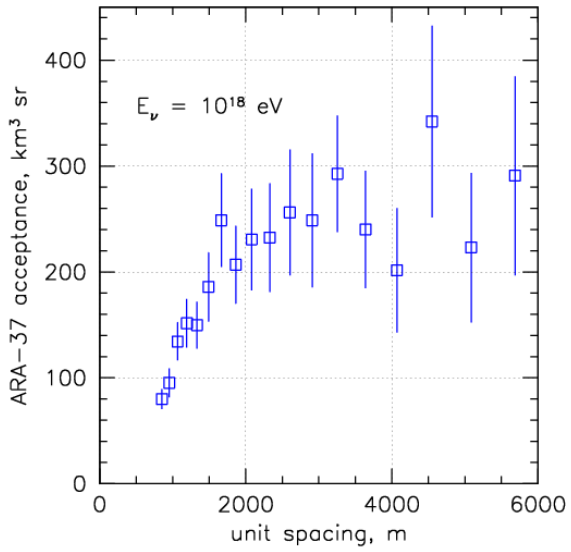


Figure 2.2: Simulated volumetric acceptance of ARA-37 at a neutrino energy of 10^18 eV as a function of the spacing between individual ARA stations. The data plateauing beyond 2000 meters indicates no further sensitivity gain for expansion beyond that distance.

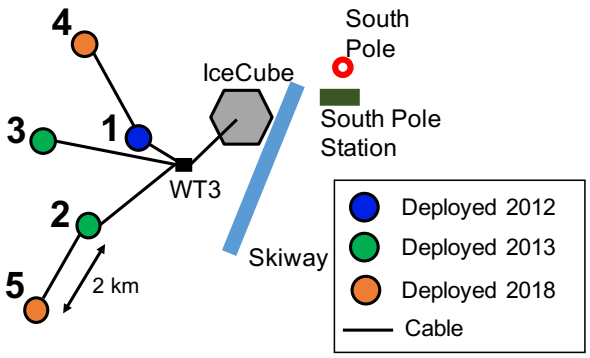


Figure 2.3: The as-built map of the five deployed ARA stations.

2.2 The ARA Station

Each ARA station is designed to be an standalone neutrino detector with isolated trigger, readout, and communication capabilities. The design also allows data gathered from a station to be analyzed independently of other stations to measure the signal arrival time, frequency content, and polarity. Each station is equipped with sixteen down-hole measurement RF antennas installed on four deep strings. Eight of these measurement antennas are designed to only pick up vertically polarized RF signals, and the other eight respond only to horizontally polarized signals. They are referred to as "Vpols" and "Hpols" respectively. The need to measure both polarizations is manifested in the reconstruction of the neutrino direction in which the location of the detector on the Cherenkov cone must be known. On each measurement string, two Hpol-Vpol pairs are installed. The station dimension has gone through several design changes. In stations one to three (A1 to A3), the vertical Hpol-Vpol distance in a pair is two meters, and the separation between two pairs is roughly 15 m, making the vertical dimension of the station essentially $\sim$ 20 m. The bottom Vpols are at an depth of approximately 80 m for A1, and 200 m for A2 and A3. On the surface, the strings are evenly spaced on a circle with a diameter of 20m. For A4 and A5, the vertical separation of the antenna cluster are increased to 30 m, with the bottom depth unchanged. The diameter of the circle on which the strings are spaced is also extended to 40 m. This scaled-up geometry has been shown by simulation to enhance the sensitivity of the stations, for example, by 25% for 10^{18} eV neutrinos with basic reconstruction qualities considered. The longer baselines are provide better reconstruction resolution.

The RF signal chain connecting the antenna from each string performs amplification, RF-over-fiber (RFoF) conversion, and bandpass operation to deliver the electric signal to the central DAQ box at the ice surface. The full setup is illustrated in Fig. 2.4. The physical dimension of the station serves as baselines for

reconstructing the signal arrival direction based on arrival time differences. The four-string setup ensures sufficient baselines can be formed for reconstruction.

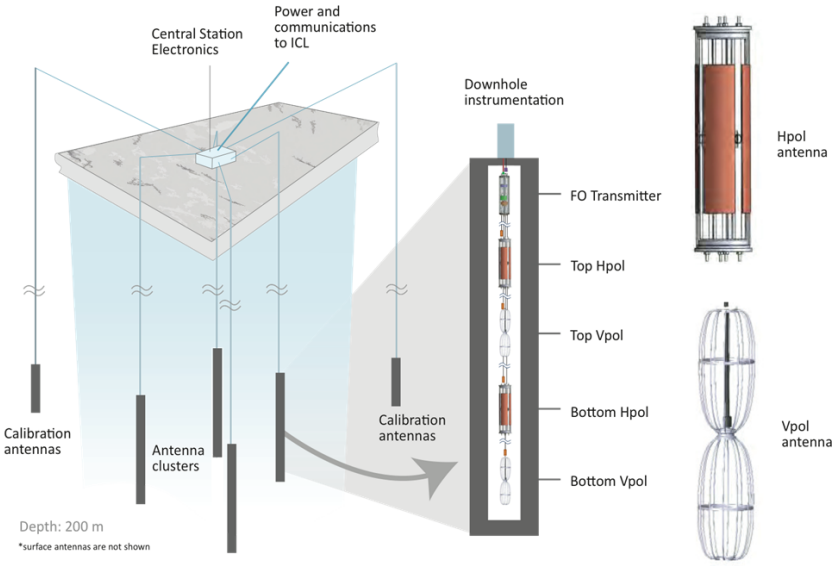


Figure 2.4: The components of an ARA station. The figure illustrates the surface DAQ box, the two calibration strings, and the four measurement strings down to 200 meters. The zoomed-in inset shows the two down-hole Hpol-Vpol pairs installed on each string, separated by 14 meters. On the right, the two antenna types are visualized. It should be mentioned that the ARA station was designed with a surface system of four surface antennas that are sensitive to extended air showers (EAS), but the surface system is currently not involved in neutrino searches and thus neglected here.

To calibrate the various station properties and monitor station health, active broadband pulser units are installed on two additional strings $\sim$ 40m away from the station center. Each such calibration string is equipped with one Hpol antenna and one Vpol antenna serving as transmitters. The pulsing mode and the exact transmitter used can be selected with DAQ control scripts. By default the calibration pulser fires through one Vpol with 1 pulser-per-second (PPS). Using the signal arrival time, the station geometry can be calibrated beyond deployment logs, and the trigger logic can also be checked. Furthermore, the attenuation level of the pulser unit can be adjusted to trace out the in-situ signal efficiency as a function of signal-to-noise (SNR) parameters. In this study, calibration pulser events are often used as a proxy for triggers from bright neutrino events.

The choice of the station depth of 200 m is such that the effective volume is maximized while keeping the drilling cost affordable. As described in section 1.4.2.2, to avoid shadow regions limiting the detector volume, the depth of the detector should ideally go beyond the firn layer. ARA uses the ARA Hot Water Drill (ARAHWD) for drilling and the holes have a diameter of 15 cm. Each 200 m hole takes roughly one day to complete.

The following sections highlight the antenna design and the signal chain including a complete diagram of the DAQ box design. They are followed by a depiction of the operation status of the two stations.

2.2.1 Antenna

The measurement antennas are designed to be broadband in the 150 to 850 MHz range and azimuthally symmetric, all while still fit in the 15cm-diameter holes. These general requirements are placed for the purpose of detecting the broadband Askaryan signal from neutrino interactions that could take place in any direction in the ice. An extra challenge faced is the existence of the vertical cable-passthrough, which is not canonical in tradition antenna designs. Eventually, the Vpol antenna is designed to be a relatively simple wired-cage bicone antenna with a dipole antenna pattern. The realized gain as a function of frequency and the angular gain pattern are shown in Fig. 2.5. The realized gain is the gain taking into account the losses from the antenna and the impedance mismatch at the feed. For the Hpol antenna, a quad-slotted-cylinder design was chosen with ferrite-loading to improve its sensitivity at lower frequency. The working principle of the Hpol antenna is such that the varying magnetic field from the incident electromagnetic wave induces current in the conducting sheets forming the cylinder, giving rise to a horizontally polarized electric signal. Again, the frequency-dependent realized gain and the angular gain pattern of an Hpol is shown in Fig. 2.6. The Vpol and the Hpol antennas can be visualized in Fig. 2.4.

2.2.2 Signal Chain

The signal chain is summarized in Fig. 2.7, wherein the full logic of the DAQ box is also presented. The registered signals on down-hole antennas are bandpass filtered between 150MHz to 850MHz, and notched filtered at 450MHz to eliminate out-of-band and narrow-band South Pole communications frequencies. Signals are then amplified by the Low Noise Amplifier (LNA) by > 75 dB, converted to optical signals to prevent loss and interference, and cabled to the surface where the DAQ box sits. The noise added by the down-hole signal chain is ≈ 132 K, and the ambient noise is ~ 248 K [94]. At the DAQ box, the signals are converted back to electric analog, bandpassed gain, and are equally split into two paths - trigger and digitization.

Implemented by an on-board field programmable gate array (FPGA model Xilinx Spartan6), the trigger logic removes delays arising from different cable lengths, sand searches for three coincident power excursions over a dynamic threshold among eight electric channels of the same polarization in a preset time frame

(*trigger window*). The power is obtained by convoluting the waveform with the few-nanosecond response of a tunnel diode (Advanced Control ACTP-1629NC3-39). The threshold is servo-ed to keep an RF trigger rate of approximately 5Hz from thermal backgrounds, which dominates ARA trigger. Note that no mixed-polarization trigger is checked by default. The trigger operation is handled by the Trigger DAughterboard (TDA). Stemming from the multiplicity trigger condition, the station RF trigger rate R can be related to the single channel trigger rate r by

$$R = 2 \cdot N \cdot C_M^N r^M \delta t^{M-1} \tag{2.1}$$

with $N = 8$ being the total number of antennas of a polarization, $M = 3$ being the multiplicity condition, and δt being the trigger window. The trigger window took two values — 110 ns and 170 ns during different times in the data taking period in this analysis, as will be summarized in section 5.3. For a station trigger rate of 5 Hz this implies a single channel trigger rate of either 10.7 kHz or 8.0 kHz. A threshold scan of the single channel trigger rate versus diode-integrated power threshold in root-mean-square units is shown in Fig. 2.8 using simulation. With the above single channel trigger rates the thresholds can be determined to be above $\gtrsim 6\sigma_{diode}$. These threshold values are critical inputs to our simulation to be described in chapter 3. The maximal sustainable station trigger rate is $\sim$ 30 Hz, as capped by the data transfer system of the DAQ. In additional to RF triggers, the station forced a software trigger at 1 Hz as a means to gathering sub-threshold thermal background data, or "min-bias" data. Overall, the station triggers at roughly 7 Hz as a combination of the RF trigger (5 Hz), the software trigger (1 Hz), and the calibration pulser trigger (1 Hz).

Once triggered, digitization of buffered voltage values with the Ice Ray Sampler (IRS2) ASIC on all channels is carried out at 3.2GS/s. Such a collection of digitized waveforms from trigger constitutes an ARA event. Mounted on the Digitizing DAughterboard (DDA), the IRS2 chip can sample up to eight channels simultaneously, and allows random access to its buffer area of 2^{15} samples per channel, which amounts to $\approx 10\mu s$ of allowed latency. However, as it is used, the buffer-writing is halted during event digitization where analog-to-digital conversion takes place. This results in a dead time of roughly a few hundred microsecond for each trigger. Overall at a trigger rate of *leq* 10 Hz this sums to only $\sim$ one millisecond dead time per second. The digitization is carried out through analog-to-digital converters (ADCs), where voltage values registered on the Switch Capacitor Arrays (SCA) on the IRS2 chip is read out as digital ADC counts using a comparator.

The control of the FPGA, networking, and data transfer in the DAQ is handled by an in-box Single Board Computer (SBC, Kontron Com Express Mini). Recorded events are stored locally in memory before being piped through fiber cables to ICL, where they are persisted on disk in compressed formats. Annually

an ARA station generate approximately ten terabytes of raw data. The data processing levels for analyzers will be introduced in section 5.2. The SBC serves as the interfacing device to experimenters that update firmware, supply DAQ control scripts, or configure data transfer. As shown in Fig. 2.7, the ARA Trigger and Readout Interface (ATRI) board assembles the SBC, FPGA, DDA, and TDA, while providing power supplies and additional monitoring peripherals.

2.2.3 Operation Status

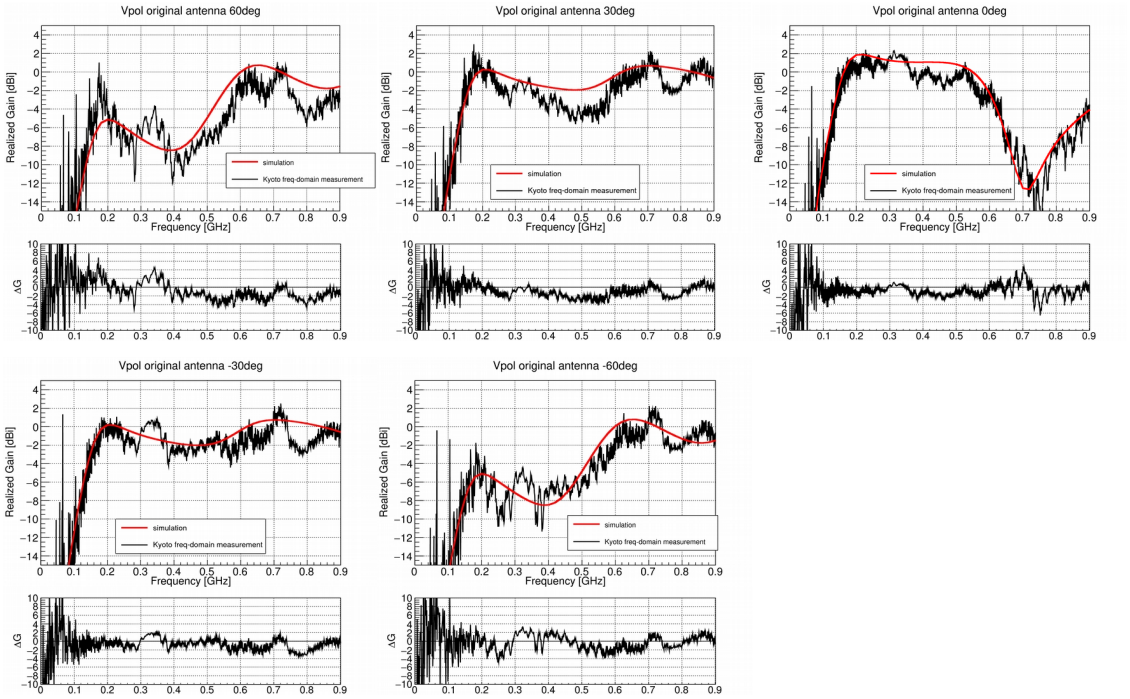
As stated earlier, currently five ARA stations are in operational modes. These five stations were deployed from 2012 to 2018. Each station operates independently of the other with the trigger logic and the readout mechanism as described in the previous section with the exception of station 5 (A5), where an advanced phased triggering is instrumented in addition to the multiplicity trigger. It will be discussion further in chapter 6. Hence, each station is monitored in a standalone fashion mainly using the *AWARE* web monitoring tool built by the ARA collaboration. Fig. 2.9 is an example of *AWARE* used in keeping track of the A3 station trigger rate and the per-channel voltage root-mean-square values. These are quantities often queried to inspect interference or fluctuation in the radio background. Such monitoring has proved to be immensely valuable. Benefits include identifying anomalies in operation in a more timely manner for logging and fixing, and configuration changes can be verified with a short turnaround, to name a few.

The data taking is bounded by runs. The data transfer is done in whole runs as well. A regular run spans six hours, while special calibration runs are usually much shorter (below one hour). The calibration runs in ARA may include special runs where external calibration sources are operated, periods where the local calibration pulser is not running with default settings, or when the DAQ configuration is being tested with certain changes. They are described and logged on the internal ARA Wiki webpages [1]. A more detailed description and reconstruction of a subset of the external calibration sources will be presented in section 4.4. The total operational live time for stations A2 and A3 from 2013 to 2016, which are the years this thesis focuses on, minus known calibration runs are 1141 days and 1021 days respectively. The fractional live time in this period can be visualized in Fig. 2.10.

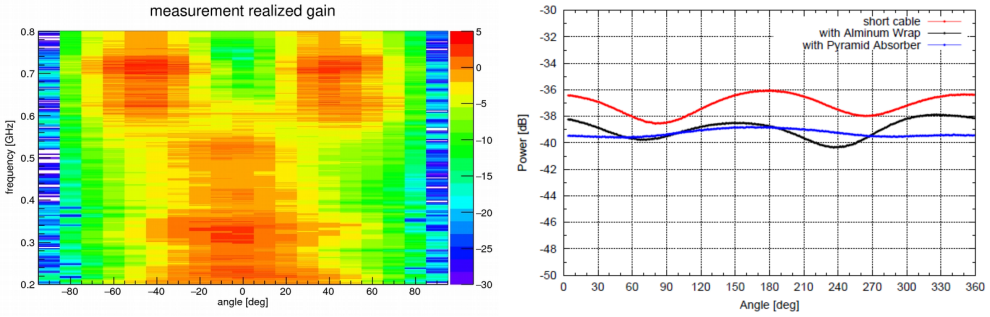
Finally, Fig. 2.11 and Fig. 2.12 exemplify waveforms captured from ARA. Each figure consists of 16 panels representing the eight Vpol channels and the eight Hpol channels. In Fig. 2.11, an event from the calibration pulser D6BV in A2 is shown. The Vpol antenna response to the pulse is clearly visible, and no cross-polarization signal is observed. In Fig. 2.12, a simulated event corresponding to a 10^{18} eV neutrino is shown. The Askaryan signal can be observed in both polarizations in this event, and using the time differences we can reconstruct the incoming angle of the wavefront. Waveforms as such are the basis of event

[1] `http://ara.icecube.wisc.edu/wiki/index.php/Run_Log_2013`, available years are 2013 to 2019 at the time of writing.

reconstruction and signal/noise distinction. In the next chapter, `AraSim`, the simulation software package that is used to generate such simulated waveforms is introduced.

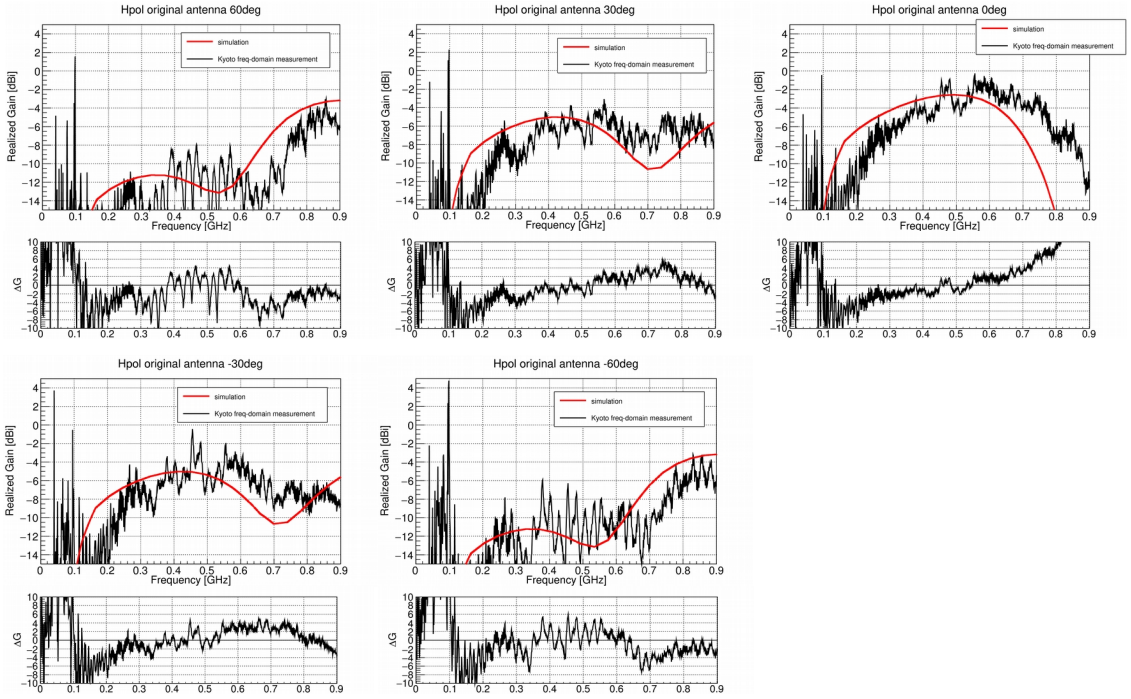


(a)

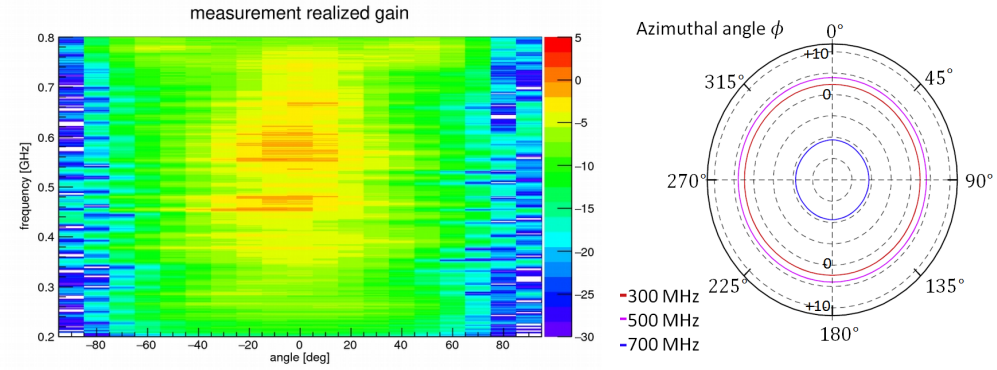


(b)

Figure 2.5: (a): anechoic chamber measurement of the realized gain for the cabled Vpol antenna (black), compared to XFdtd simulation (red) at selected elevation angles. An elevation angle of 0° is the direction perpendicular to the antenna's long side. A cable connects to the antenna feed and goes through the "top" side (+90°). The difference between measurement and simulation is generally ± 4 dB in the frequency range of interest. (b) left: measured realized gain versus frequency and elevation angle. The frequency interval is 1 MHz, and the angular interval is 10°. The angular asymmetry near 0.5 GHz can be attributed to cable effects. (b) right: measured realized gain pattern at 500 MHz and 0° elevation angle versus the antenna azimuth, showing a nearly invariant response. The legend describes how the attached cable is shielded. All measurements in-air. (a) and (b, left) extracted from [92]. (b, right) taken from [93].



(a)



(b)

Figure 2.6: (a): Hpol counterparts of Fig. 2.5a. The difference between measurement and simulation is generally ± 10 dB, larger compared to the Vpol. The high-frequency response is especially poorly understood and is being investigated. The Hpol antenna is generally less sensitive than Vpols, as shown by the absolute scale of the gain curves. (b) left: Hpol counterpart of Fig 2.5b (left). The apparent angular asymmetry near 0.5 GHz can be attributed to cable effects. (b) right: *simulated gain pattern* at selected frequencies at an elevation angle of 0° as a function of the antenna azimuth showing a uniform response.
All measurements in-air. (a) and (b, left) extracted from [92]. (b, right) taken from [90].

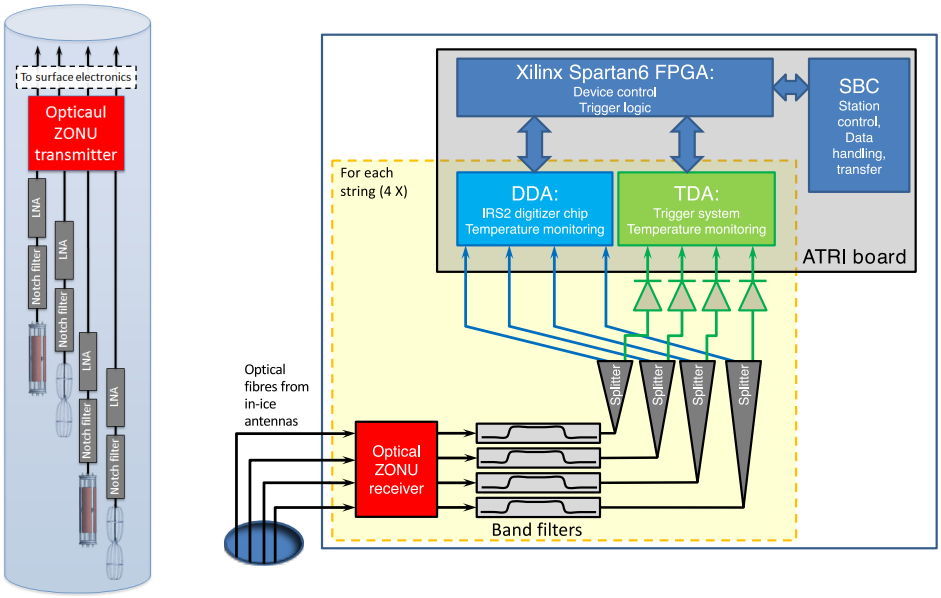


Figure 2.7: Left: down-hole signal chain showing key components to each measurement antenna. Right: the surface DAQ box showing the signal chain and the control logic. The section within the yellow frame is common for all four strings [94].

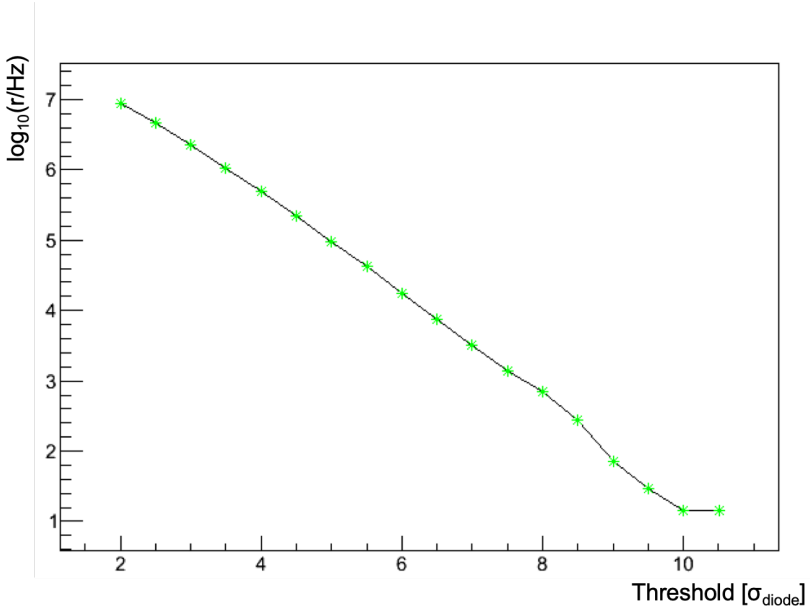
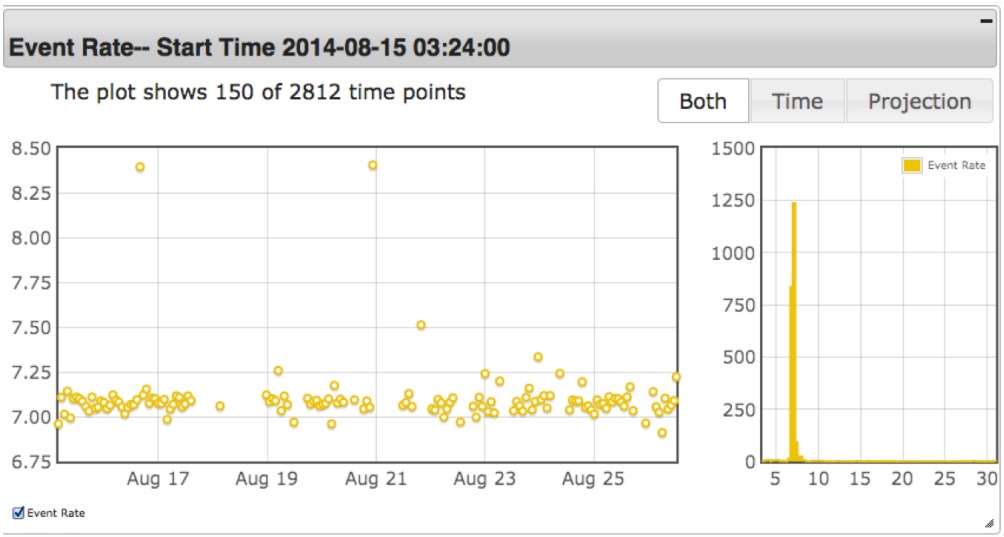
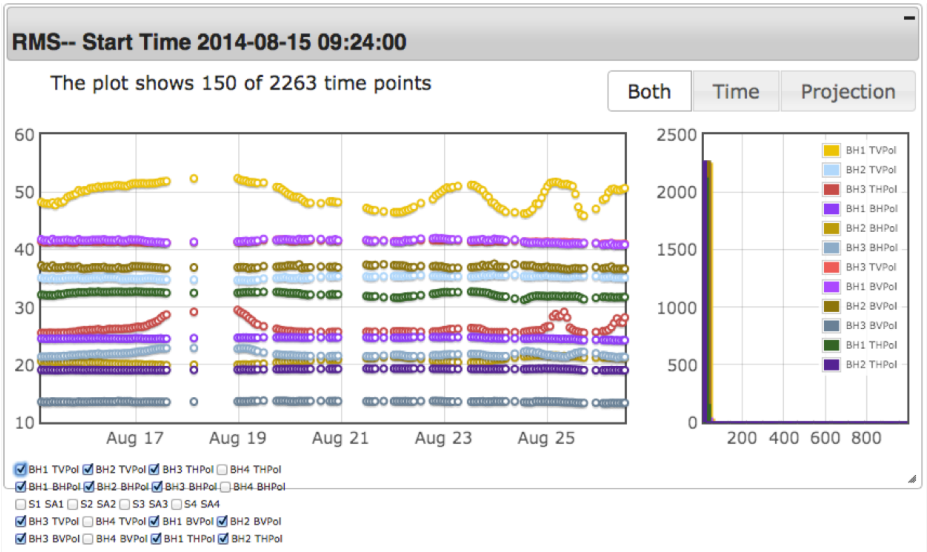


Figure 2.8: The single channel trigger rate r at selected trigger thresholds in unit of the root-mean-square of the diode-convoluted noise simulation.



(a)



(b)

Figure 2.9: (a): The station trigger rate for A3 in a 12-day period in 2014. The total rate is the sum of the RF trigger rate (5 Hz), the calibration pulser trigger rate (1 Hz), and the software trigger rate (1 Hz). The gap is a period of station down time. Data points that deviate significantly from the 7 Hz baseline are an artifact of the rate calculation in the software when dealing with run starts. (b): The root-mean-square voltage of each channel for the same time period. Note that most channels are stable, and the variation in the unstable channels is counter-balanced by per-channel threshold adjustments in the servo mechanism to maintain an overall stable station rate.

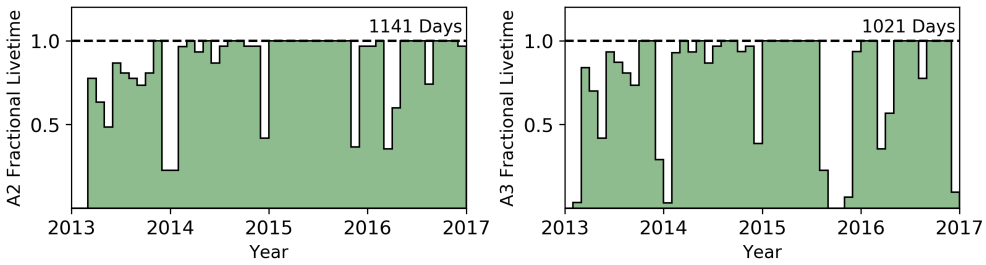


Figure 2.10: Left: fractional live time for A2 from 2013 to 2016, excluding known calibration runs. Right: fractional live time for A3 from 2013 to 2016, excluding known calibration runs. [95]

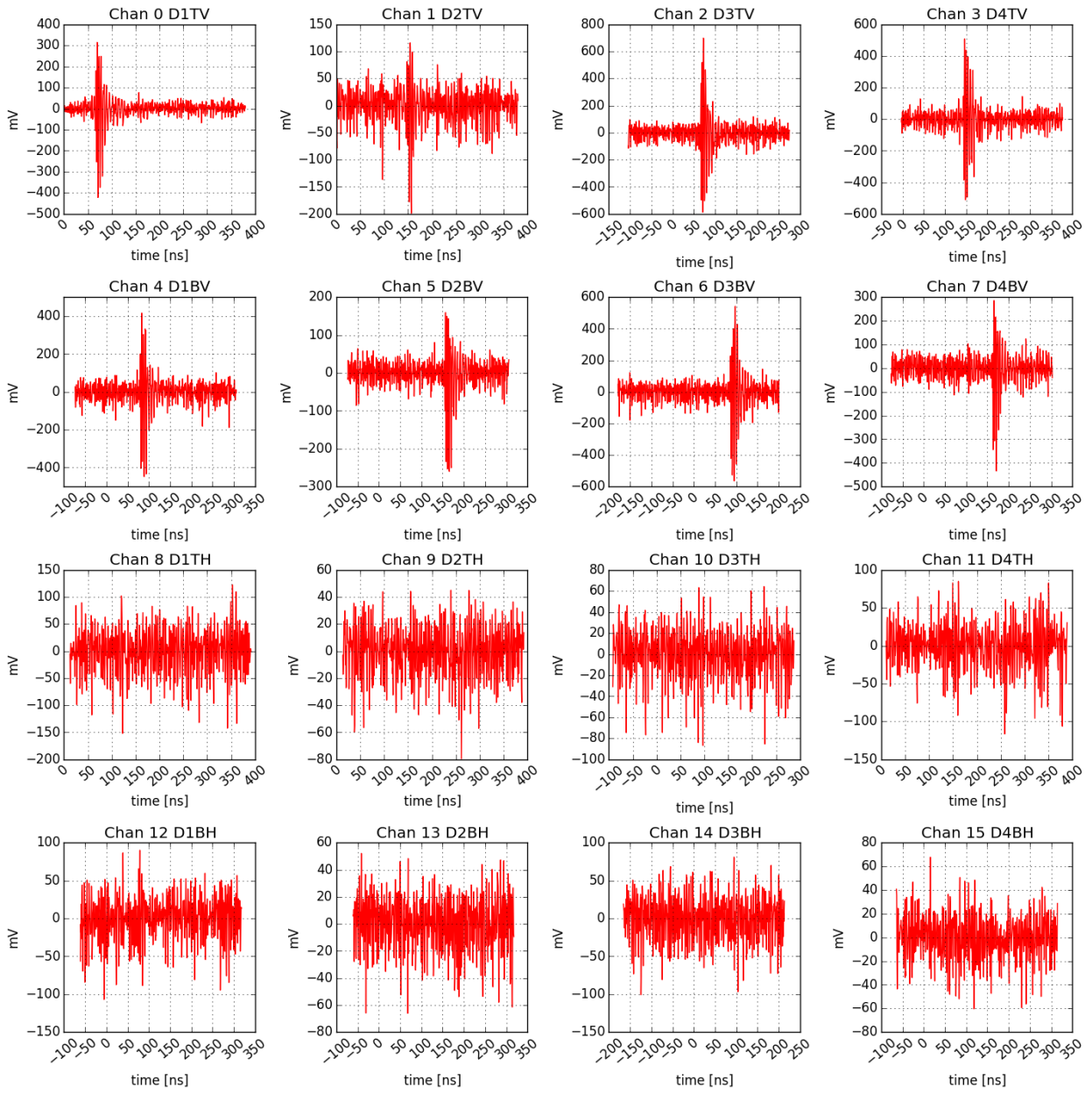


Figure 2.11: A calibration pulser event from A2 (run 2084 event 5). The top two rows are the Vpol channels, and the bottom two rows are the Hpol channels.

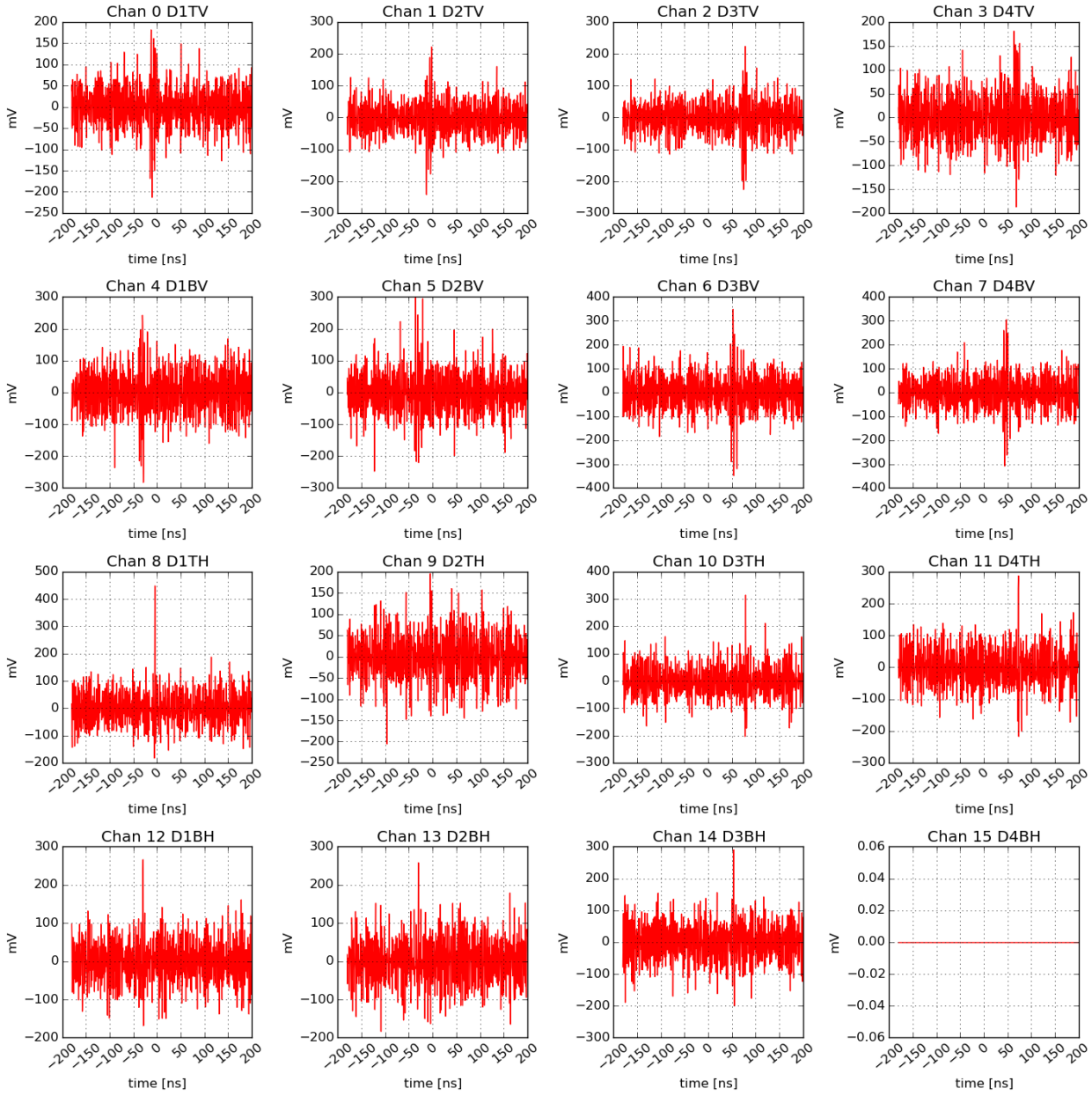


Figure 2.12: Simulated neutrino event with a neutrino energy of 10^{18} eV with A2. In this event antennas of both polarization observe the Askaryan signal. Notice that D4BH (chan 15) is zero-ed out to mimic the actual procedure of analyzing collected A2 data due to a broken channel.

Chapter 3

Simulation

In order to assess the sensitivity of the detector, a software simulation package `AraSim` [1] [82] is developed by the ARA collaboration. `AraSim` represents an end-to-end simulation tool that deals with Monte Carlo signal generation, propagation, detector response and data readout for a sub-glacial radio detector like ARA. These stages will be described in this chapter. In this thesis `AraSim` is used as the sole source of neutrino event simulation. The feature development, debugging, and code maintenance responsibilities of `AraSim` is shared among ARA members including myself.

One key step in using `AraSim` is the use of a setup script, where various simulation parameters can be specified by the user. Once the script is submitted with the `AraSim` run, the simulation is configured according to the parameters therein. Therefore, setup scripts used in this work will be shown at the end of the chapter and in the appendix for reproducibility. The `AraSim` SVN revision number used in producing simulation for estimating neutrino efficiency in this work is 746. The project has since been moved to GitHub for version control (see footnote).

3.1 Detector Setup

In `AraSim`, the user can specify the geometric formation of antennas by selecting from a set of pre-defined detector configurations with the setup parameter. The most frequent modes are the ideal station and the "as-built" station modes. In the ideal station mode, the dimension of the station follows the nominal values of the design, and in the "as-built" modes, actual depths and string spacings from the deployment log plus correction subsequent calibrations [93] are used instead. The latter information in stored in SQLite database that is common to `AraSim` and the ARA analysis software AraRoot [2] [96]. The sensitivity variation purely due to such dimensional differences is negligible. For a developer, it is also possible to implement a customized geometry.

[1] `https://github.com/ara-software/AraSim`

[2] `http://www.hep.ucl.ac.uk/uhen/ara/araroot/branches/3.13/`

3.2 Signal Generation

Once the antenna location is set, the software proceeds to randomly generate neutrino interaction vertices in a defined ice volume around the detector. There are two common ways to define the generation volume - spherical and cylindrical. In the spherical approach, the vertices are randomly generated in a spherical volume centering the detector center with a uniform probability density function (PDF). The radius of the sphere is a parameter in the setup script. Only vertices that are generated in-ice are kept for next steps. This approach fulfills the goal of Monte Carlo sampling the detection phase space while also provides a straightforward way to compute the effective area A_{eff}, an equivalent detector area that is sensitive at 100% efficiency to the neutrino flux. The cylindrical mode generates vertices in a cylindrical ice volume with a user-specified radius and a depth of three km, the depth of the ice sheet at the South Pole. The detector is on the central axis of the cylinder at a depth configured by the setup script. In this mode the ice volume is also uniformly randomly sampled. This setup allows the derivation of the effective volume V_{eff}, which stands for the detector volume assuming the detection efficiency is 100% to a neutrino flux. This quantify can be later converted to A_{eff} to calculate the neutrino event rate. We will cover these calculations in section 3.5. The efficiency as estimated by the two modes are mutually consistent. Past ARA analyses and publications have used the cylindrical mode. Consequently, I adopt the same mode in this work for ease of comparison. The radius of the cylinder should be chosen so that all triggerable neutrino can be included. The values of the radius at each energy are empirically obtained and listed in Table. 3.1:

Table 3.1: `AraSim` volume radius (POSNU_RADIUS) for each simulated neutrino energy bin

$log_{10}(E_\nu/eV)$	16	16.5	17	17.5	18	18.5	19	19.5	20	20.5	21
Radius [km]	3	3.5	4	4.75	5.5	6.25	7	8	9	10.5	12

Once a vertex location is randomly chosen in the generation volume, the simulation assigns particle properties including primary energy, direction, flavor, neutrino/antineutrino, and type of interaction (NC/CC). The energy can be specified by the user to be drawn from a spectrum. In this work we generate simulation at fixed energies from 10^{16} eV to 10^{21} eV in half-decade steps, essentially drawing from a $F = dN/dE = E^{-1}$ spectrum. The outcome can be reweighted to simulate other spectra if desired. The direction of a neutrino is randomly assign in the 4π phase space, following the assumption for a diffuse cosmogenic neutrino flux. The flavor is randomly assigned based on a $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$ assumption. The $\nu/\bar{\nu}$ assignment is based on [97], where $\nu_e : \bar{\nu_e} = 0.78 : 0.22$, and $\nu_{\mu/\tau} : \bar{\nu}_{\mu/\tau} = 0.61 : 0.39$. The NC/CC type is determined based on

the prescription from [98] [3], where in the relevant energy range we have $\sigma_{NC} : \sigma_{CC} = 0.3134746 : 0.6865254$. Finally, the inelasticity y is chosen according to the prescription in [99].

An additional quantity arising from the Monte Carlo forced-interaction nature of the simulation is the event weight. Given the cross section of the interaction, the vertex location and direction of the neutrino, the weight calculates the survival probability that the neutrino reaches the vertex through Earth. Considering a beam of neutrinos directed at a target medium with initial intensity I_0, the intensity at depth l is expressed as

$$I(l) = I_0 exp(-n\sigma l). \tag{3.1}$$

Here, n is the number density of the medium, σ is the total cross section for ν-nucleon interaction. It is convenient to define the interaction length L_{int} as $L_{int} = 1/(n\sigma)$. It follows that the survival probability of a neutrino at an assigned vertex location in `AraSim` is the compound of the survival probabilities of the neutrino through all layers of Earth through which it traveled. The weight is then

$$w = \prod_{\text{layer}} e^{-\frac{l_{\text{layer}}}{L_{\text{int,layer}}}} \,|^{\text{vert}} \tag{3.2}$$

, where l_{layer}, $L_{\text{int,layer}}$ are evaluated for each Earth layer (crust, mantle, core) and the ice thickness traversed by the neutrino up to the vertex.

With the properties of the vertex determined, to generate the Askaryan signal, `AraSim` follows the formulation described in section 1.4.1. The shower profile to be used is, for hadronic showers, the Gaisser-Hillas function [100] with in-ice parameters. The total number of particles at shower depth X is

$$N_{\text{total}}(X) = S_0 \frac{E_0}{E_c} \cdot \frac{X_{\max} - \lambda}{X_{\max}} e^{X_{\max}/\lambda - 1} \cdot (\frac{X}{X_{\max} - \lambda})^{X_{\max}/\lambda} \cdot e^{-X/\lambda} \tag{3.3}$$

, where X has unit of $g \cdot cm^{-2}$, $S_0 = 0.1842$, E_0 is the shower energy in GeV, $E_c = 0.17006$ GeV, X_{max} is the depth of the shower maximum, $\lambda = 113.03$ g/cm^2 [64]. The electromagnetic shower profile is parametrized by the Greisen function [101] as

$$N_{\text{total}}(X) = \frac{0.31}{\sqrt{ln(E_0/E_c)}} \cdot e^{X - 1.5X ln(3X/(X + 2ln(E_0/E_c)))} \tag{3.4}$$

, in which the critical energy $E_c = 0.073$ GeV in ice. Noting that the radiation depends on the charge excess profile, `AraSim` uses a constant charge excess factor of 25% to convert the shower profiles from both

[3] With additional correction from CTEQ6-DIS parton distribution functions

the hadronic and the electromagnetic shower. The charge excess profile (equivalent to $Q(z)$ in section 1.4.1) is then

$$Q(X) = q \cdot 0.25 \cdot (N_{\text{total, had}} + N_{\text{total, EM}}) \tag{3.5}$$

, where q is the charge for an electron, and the allocated shower energies for the hadronic and the electromagnetic components are given by the inelasticity. In addition to the primary shower, `AraSim` also considers secondary showers from μ/τ after ν_μ or ν_τ CC interactions. The energy of the resulting electromagnetic or hadronic showers are calculated with interaction probabilities obtained from the Muon Monte Carlo (MMC) [102]. Between the primary and the secondary shower, `AraSim` generates the Askaryan signal according to the shower that carries the most energy. The simulation then proceeds to treat propagation effects to emulated the radiation incident upon a receiving antenna. This is discussed in the following section.

3.3 Signal Propagation

`AraSim` allows users to select either a bulk-ice constant index of refraction ice mode, or an exponential $n(z)$ profile (1.32), with the latter being the default. As described in section 1.4.2.2, the depth-dependent index of refraction of the polar ice gives rise to ray-bending effects. To find the first solution (direct ray) connecting two endpoints in ice, `AraSim` employs a semi-analytical algorithm. The goal of the algorithm is to find the launch angle of the ray path at the source that would reach the target while being compliant with Snell's law. Consider the two endpoints forming an $x - z$ plane, according to Snell's law:

$$n(z_0)sin(\theta_0) = n(z)sin(\theta) \tag{3.6}$$

, where 0 denotes quantities at the source, the θ denotes the launch angle of the ray measured from vertical . Given an exponential $n(z)$ model of the form $n(z) = A + Be^{Cz}$, we can solve for θ as

$$\theta = arcsin(sin\theta_0 \frac{A + Be^{Cz_0}}{A + Be^{Cz}}) \tag{3.7}$$

Using the definition of θ, we have $dx/dz = tan\theta$. Substituting and rearranging we get

$$x - x_0 = \int_{z_0}^{z} tan(arcsin(sin\theta_0 \frac{A + Be^{Cz_0}}{A + Be^{Cz'}}))dz'. \tag{3.8}$$

If we define $\sigma_0 \equiv sin(\theta_0)$, $n_0 \equiv n(z_0)$, $n \equiv n(z)$, we can arrive at

$$\frac{C\sqrt{A^2-\sigma_0^2 n_0^2}}{\sigma_0 n_0}(x_0-x) = ln(\frac{\sqrt{n^2-\sigma_0^2 n_0^2}+\sqrt{A^2-\sigma_0^2 n_0^2}}{n-A}+\frac{A}{\sqrt{A^2-\sigma_0^2 n_0^2}}) - ln(\frac{\sqrt{n_0^2-\sigma_0^2 n_0^2}+\sqrt{A^2-\sigma_0^2 n_0^2}}{n_0-A}+\frac{A}{\sqrt{A^2-\sigma_0^2 n_0^2}}). \tag{3.9}$$

As a result, given the location of the endpoints and $n(z)$, σ_0 can be numerically solved for from this analytical expression. `RaySolver` then proceeds to trace the path step-by-step. This approach is used to efficiently find the first solution. In the case where a numerical convergence cannot be found, the code resorts to a "trial-and-error" ray-casting approach. To find the second solution, again the "trial-and-error" method is taken. One a path is found, the signal strength is attenuated over the path length with the $1/distance$ factor and the attenuation according to section 1.4.2.2. The `RaySolver` also gives the arrival time, polarization, and the view angle of the signal at each antenna. These quantities are then folded with the `AraSim` antenna response to produce input signal to the simulated signal chain. The next section describes this process.

3.4 Detector Response

Given the electric field incident on an antenna, we can use the antenna effective height to convert it to the voltage measured. Let us first define the effective power collection area of the antenna as

$$A_{\text{eff,ant}} = \frac{Gc^2}{4\pi f^2}. \tag{3.10}$$

Here G is the gain, and f is the frequency in consideration. Then, the effective height is defined as

$$\mathbf{h}_{\text{eff}} = s\sqrt{\frac{R_r A_{\text{eff,ant}}}{\eta}}\widehat{\mathbf{h}} \tag{3.11}$$

, where R_r is the radiation resistance of the antenna, $\widehat{\mathbf{h}}$ stands for the unit vector along the antenna axis, and $\eta = \sqrt{\mu/\epsilon}$ is the impedance of the medium. When the impedance is matched, we have

$$V = \frac{\mathbf{E}\cdot\mathbf{h}_{\text{eff}}}{2}. \tag{3.12}$$

The gain patterns used by `AraSim` in this work are NEC2 [103] simulations with the environment parameters changed to mimic in-ice conditions for bottom Vpols (cable on one side), top Vpols (cabled through both sides), and Hpols, similar to shown Fig. 2.5 and Fig. 2.6. The electric signal is then passed through the response of the filters and the amplifiers, with the amplitude response taken from lab measurements, and the frequency response taken from a similar filter model from the Qusc Studio [104]. At this stage, the

signal is superimposed with a long ($\sim 6\mu$s) noise waveform at offsets corresponding to the arrival time from `RaySolver`. The time-domain noise waveform is generated from the Fourier transform of a Johnson-Nyquist spectrum with the mean noise power

$$< P >= k_B T \Delta f \tag{3.13}$$

, where k_B is the Boltzmann constant, T is the noise temperature, and Δf is the bandwidth. Clearly, the spectral density is then $k_B T$ [W/Hz]. To obtain the noise temperature, there are two sources of noise that need to be considered. The first the the noise from the environment. This includes the black body radiation of the ice, its reflection off the ice surface, and the noise from the sky including the atmospheric and the galactic background. We estimate this ambient noise as been by an antenna via folding the antenna gain pattern with these radiation sources, integrating out to the horizon as set by the attenuation length. This gives

$$T_{\text{ant}} = T_{\text{ice}} + T_{\text{refl}} + T_{\text{sky}}. \tag{3.14}$$

While the black body radiation, the galactic and atmospheric radiation are all frequency-dependent, in `AraSim` they are only estimated at 300 MHz. Eventually for a Vpol T_{ant} is evaluated to be 247 K, and for an Hpol it is 249 K. Since the ambient temperature represents incident noise power on an antenna, it needs to be couple with the transmission coefficient to account for the loss in the antenna feed. The transmission coefficients are simulated as in Fig. 3.1. At 300 MHz the transmission coefficients for the three groups shown are ≈ 0.8.

The second source of noise is the internal noise in the signal chain, to which the largest contributor is the LNA. The noise factor F is used to represent this internal noise power:

$$F = \frac{SNR_{in}}{SNR_{out}} = \frac{\frac{S_{in}}{N_{in}}}{\frac{S_{out}}{N_{out}}} = 1 + \frac{N_{\text{sc}}}{N_{\text{ant}}} = 1 + \frac{T_{\text{sc}}}{T_0}. \tag{3.15}$$

In the above equation, N_{sc} stands for the noise power in the signal chain, and N_{ant} denotes the ambient noise. T_{sc} and T_0 stand for the equivalent noise temperature of the signal chain and of the environment respectively. Given lab measurements of the channel-by-channel noise factor, we assume a linear scaling with temperature to account for the cold ice at the antenna depth and set $T_0 = 220$ K. Finally, this gives us

$$T_{\text{sc}} = 220K \cdot (F - 1). \tag{3.16}$$

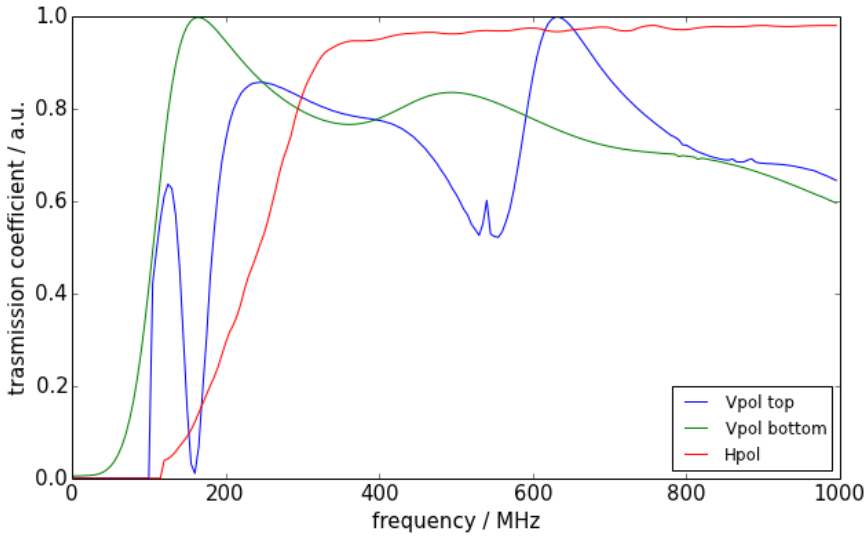


Figure 3.1: The simulated transmission coefficients for three antenna groups top Vpols, bottom Vpols, and Hpols. The Vpols are simulated with XFdtd, and the Hpols are simulated with NEC2. The features of the top Vpol can be attributed to cable effects.

At 300 MHz, F takes the value of ≈ 1.6. And the total noise temperature is thus

$$T_{\text{tot}} = t \cdot T_{\text{ant}} + T_{\text{sc}} \tag{3.17}$$

, where t represents the transmission coefficient. For a more detailed derivation, see [94]. Note that for both A2 and A3 the noise factors in the simulation assume values solely from A2 measurements, due to the A2 dataset being more comprehensive and the fact that the two stations share the same design.

With the total noise temperature obtained, `AraSim` uses Eq. 3.13 to generate a frequency-domain noise power spectrum. Random phases from $0 \sim 2\pi$ are assigned to each frequency bin before the time domain waveform is produced through Fourier transform.

At this point, `AraSim` convolves the pure noise waveform with a model for the tunnel diode response to determine the power root-mean-square σ_{diode} to be used in trigger checks. The diode response can be visualized in Fig. 3.2. On the other hand, the signal waveform, with noise now added, is convoluted with the same diode response. The trigger check function then walks through the convoluted signal to check for signal channel triggers, defined as power excursions over a user-defined multiple of σ_{diode}. The station trigger condition is then the same multiplicity trigger as described in section 2.2.2. Once triggered, the code stops

checking for further excursion, and output waveforms in the same format of a real station. `AraSim` also writes out information regarding the primary and the interaction with the waveforms.

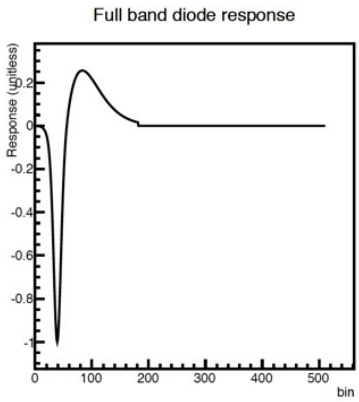


Figure 3.2: The simulated tunnel diode response used in `AraSim`. Each bin on the horizontal axis is 0.5 ns.

3.5 Sensitivity

As `AraSim` is designed and implemented, it provides all information necessary to estimate the sensitivity of ARA. Given a sample of triggered events, the detection efficiency at the trigger level could be expressed as

$$\epsilon_{\text{trig}} = \frac{\sum_{i}^{\text{trig}} w_i}{N_{\text{sim}}} \tag{3.18}$$

, where the numerator is the sum of weights for all triggered events, and N_{sim} denotes the total number of vertices thrown in the simulated volume, triggered or not. The effective volume is then

$$V_{\text{eff}} = V_{\text{ice}} * \epsilon_{\text{trig}}. \tag{3.19}$$

Note that sometimes instead of V_{eff}, which has a volumetric unit, $V\Omega_{\text{eff}}$ is used as a metric to gauge detector sensitivity. This latter expression carries an extra solid angle unit. For example, the unit of $V\Omega_{\text{eff}}$ would be $km^3 \cdot sr$. Since `AraSim` randomly assign the direction of each primary neutrino, the obtain V_{eff} is readily average over the 4π solid angle phase space. Thus, the $V\Omega_{\text{eff}}$ is simply $4\pi \cdot V_{\text{eff}}$.

A slightly more useful expression is the effective area A_{eff}, which goes directly into the calculation of event rates given a flux model. We can convert the effective volume to the effective area by

$$A_{\text{eff}} = \frac{V_{\text{eff}}}{L_{\text{int}}}. \tag{3.20}$$

Similarly, another convention exists to consider $A\Omega_{\text{eff}} = 4\pi \cdot A_{\text{eff}}$, with an extra unit of steradian. The trigger level A_{eff} as computed with `AraSim` samples with half-decade energy bins over an E^{-1} spectrum for both A2 and A3 are plotted in Fig. 3.3. Above 10^{17} eV, the effective area of A2 and A3 are within $\sim 10\%$ of each other. For each station, we can further define the "exposure" as $4\pi \cdot A_{\text{eff}} \cdot T$, where T is the live time of the station. The combined exposure from two stations is simply $4\pi(T_2 A_{\text{eff},2} + T_3 A_{\text{eff},3})$. And the sensitivity is the inverse of the exposure. To calculate the event count given a differential spectrum dN/dE for the two station, we can use the following expression:

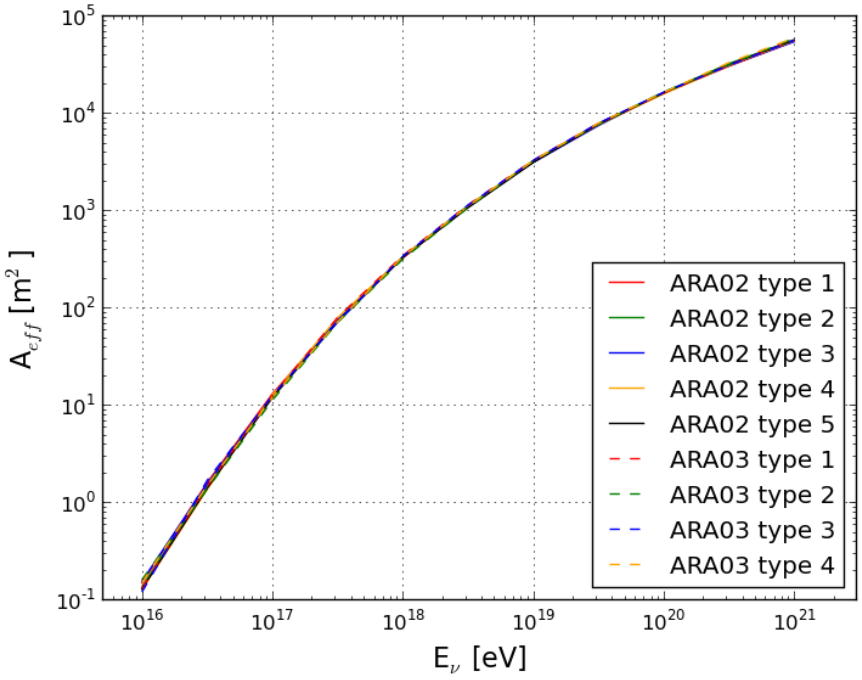


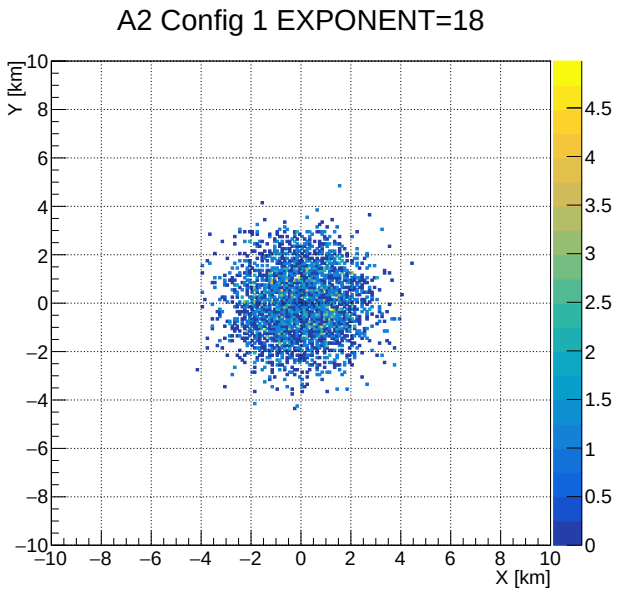
Figure 3.3: Trigger effective areas as simulated by `AraSim` for A2 and A3. For each station, depending on the DAQ setting the live time is divided into multiple configurations (here referred to as "types"). We simulated these configurational variations to produce realistic estimates of the sensitivity. As the figure shows, the discrepancy among the effective areas are negligible at the trigger level. In chapter 5, the configurations will be described.

$$N = \int 4\pi \frac{dN}{dE} A_{\text{eff},2} T_2 dE + \int 4\pi \frac{dN}{dE} A_{\text{eff},3} T_3 dE. \tag{3.21}$$

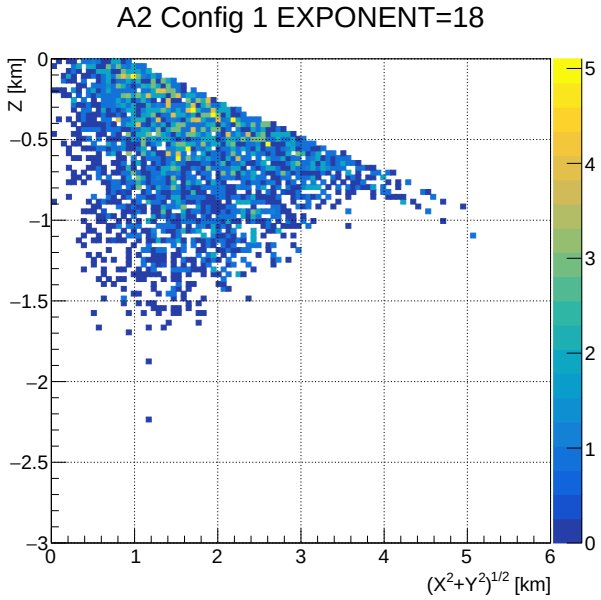
So far we have covered the evaluation of the trigger level sensitivity. At the analysis level, the same quantities can also be calculated by multiplying the trigger efficiency ϵ_{trig} by the analysis efficiency ϵ_{ana}, which can be obtained after analysis level event selection and cuts are defined. We will revisit this topic in section 5.7.

Fig. 3.4 illustrates the vertex distribution of triggered events with $E_\nu = 10^{18}$ eV around A2. The azimuthally symmetric response is manifested by the ϕ-independent distribution of vertices on the X-Y plane. In the R-Z distribution (right), the shadow region is clearly visible, and the distribution weighs more heavily toward the shadow boundary, as vertices in the ice below the station are less likely to cause a trigger due to the side-viewing nature of the antenna gain pattern, and that the interaction weights are suppressed due to Earth absorption in that region. In Fig. 3.5, the arrival direction of triggered events are shown. In this figure the convention of θ_{arr} is such that a vertically down-going event as seen at the South Pole has a θ_{arr} of 0°, and a vertically up-going event has a θ_{arr} of 180°. The sky coverage of an ARA station ranges from $\sim 45^\circ$ above the horizon to $\sim 5^\circ$ below the horizon, in which most events consist of mildly down-going to horizontal neutrinos.

Finally, it should be mentioned for reproducibility, the `AraSim` setup scripts used to generate signal simulation in this work are attached as appendix A.



(a)



(b)

Figure 3.4: Vertex locations for triggered events in A2 assuming a primary neutrino energy of 10^{18} eV. The station center in these figures is (X, Y, Z) = (0, 0, -177) [m]. (a): distribution on the X-Y plane, showing a uniform trigger efficiency in the azimuth. (b): distribution on the R-Z plane, where R= $\sqrt{X^2+Y^2}$. At this energy the effective volume is attenuation-limited, as no physical boundary is met (e.g. bedrock) before distant vertices cease to trigger due to attenuation.

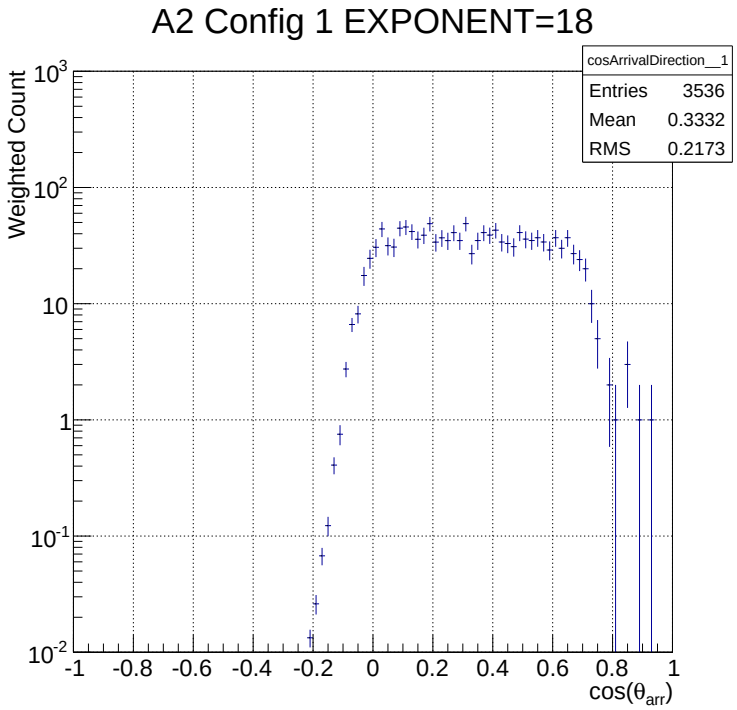


Figure 3.5: The cosine of the arrival direction of triggered events in A2 using a sample of 10^{18} eV neutrinos. The convention of the angle is described in the text. We observe that the angular coverage is roughly 50° in the Southern Hemisphere for an ARA station.

Chapter 4

Data Analysis Techniques

The collected ARA data consist of waveform data and metadata indicating the conditions under which the waveforms are recorded, as defined by the ARA analysis software **AraRoot** [1]. Given the waveforms of an event, it is most critical to determine if the event is of a thermal nature, from a background source, or can be considered a potential neutrino event (neutrino candidate). In the last case, we should be able to reconstruct properties of the neutrino and the interaction vertex. To achieve this, a reconstruction algorithm is needed. To reconstruct the neutrino arrival direction and its energy or flavor, multiple arrival time and frequency spectrum measurements on the Cherenkov cone are needed. This is being studied actively. Instead, in this thesis we focus on reconstructing the arrival direction of the signal with respect to the station and the direction of the vertex. Such information is sufficient to reject backgrounds and identify neutrino candidates.

Generally speaking, radio detector reconstruction algorithms can be classified into three categories — (a) interferometry-based, (b) hit-time based, and (c) template-matching. In (b), each waveform is scanned to obtain a hit-time representing the location of the signal in the waveform. Combining measured hit-times from multiple channels, a fit could be performed against predicted hit-times from simulation (e.g. `RaySolver`) where a minimizer determines the best-fit vertex direction. This is the approach adopted in the previously-published 10-month A2+A3 diffuse analysis [94]. In (c), templates of simulated neutrino signal are put together. Each data event is then compared to the templates, after which an event may be elevated to become a neutrino candidate if the similarity is high. In this work, we adopt approach (a), where the full waveform is used to cross-correlate with other waveforms to create a full interferometric map of the sky as seen by the station. The brightest spot on this sky map is taken as the reconstructed vertex. The following sections will detail this algorithm. Further event selection cuts can be placed using information from the reconstruction. For example, we can devise reconstruction quality cuts and geometric cuts that exclude events from above the ice surface or known anthropogenic sources. In comparison, (a) is advantageous to (b) in that it avoid the complexity and limits of extracting the hit-time from a waveform. However, the trade-off is that (a) is more susceptible to fluctuations in waveforms, including from Continuous Wave (CW)

[1] `http://www.hep.ucl.ac.uk/uhen/ara/araroot/branches/3.13/`

interference, that may lead to extra features in the sky map which could overpower the "true" vertex. The advantage of (c) is that a lower energy threshold can be reached from template-matching. The trade-off is that the generation of templates needs to be sophisticated enough, and that any non-neutrino physics background, for example EAS, will likely be missed out. Both (b) and (c) are currently being developed for future analyses.

With a physical dimension of 20 m, to reconstruct the direction of the vertex to 1° resolution, the timing precision should be $\sim$ 1 ns. In order to reconstruct the distance to the vertex through resolving the wavefront curvature as seen by a station, a timing precision of $\sim$ 100 ps is needed. While a precision of 100 ps is achievable with ARA, the determining of the vertex distance beyond $\sim$ 500 meters is still currently unattainable due to systematic errors on the antenna locations and cable delays. We introduce another potential route towards the reconstruction of the vertex distance using double pulses from the two ray paths in section 4.4.2.

4.1 Station-centric Coordinates

The reconstruction works in the station-centric coordinate system, where the origin of the system lies in the station center with a depth of 180 meters below the ice surface. The x-axis of this system is aligned with the ice flow direction at the South Pole. The z-axis points upward towards the zenith. The reconstructed angles that appear in the following texts will be measured with respect to these axes. The relationship of this system to other coordinate systems in use at the South Pole is detailed in appendix C. The position of other notable objects near ARA is summarized in [105]. The antenna locations of A2 and A3 in their respective local coordinates are visualized in Fig. 4.1 and Fig. 4.2.

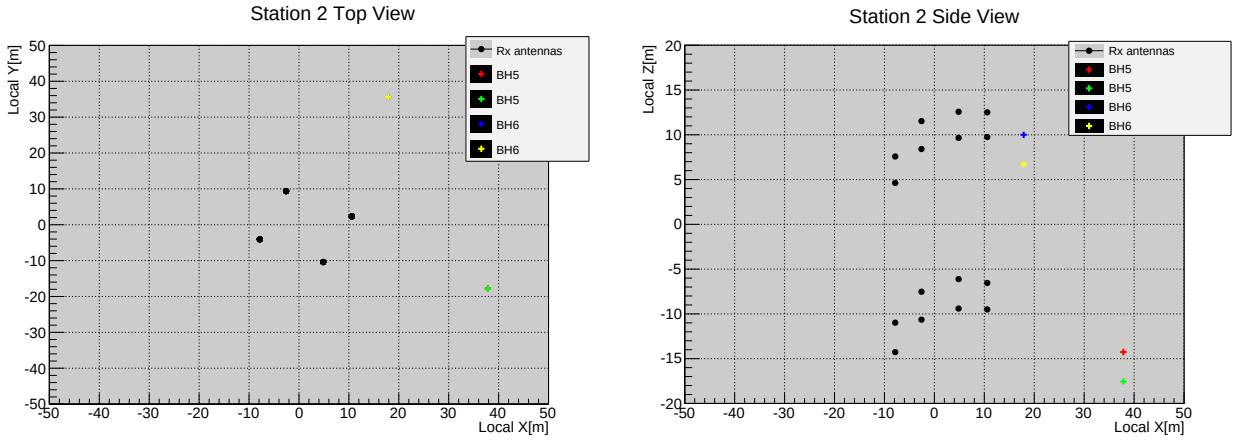
4.2 Interferometry

In an interferometric reconstruction, we assume a set of putative vertex locations. For each location, the signal arrival times (delays) to antennas are calculated. Taking any pair of antennas (a baseline), we compute the cross correlation value according to the difference in delays. The cross correlation of two continuous real functions $f(t)$ and $g(t)$ is computed as (F denotes Fourier transform):

$$(f \otimes g)(\tau) = \int_{-\infty}^{\infty} f(t)g(t+\tau)dt \tag{4.1}$$

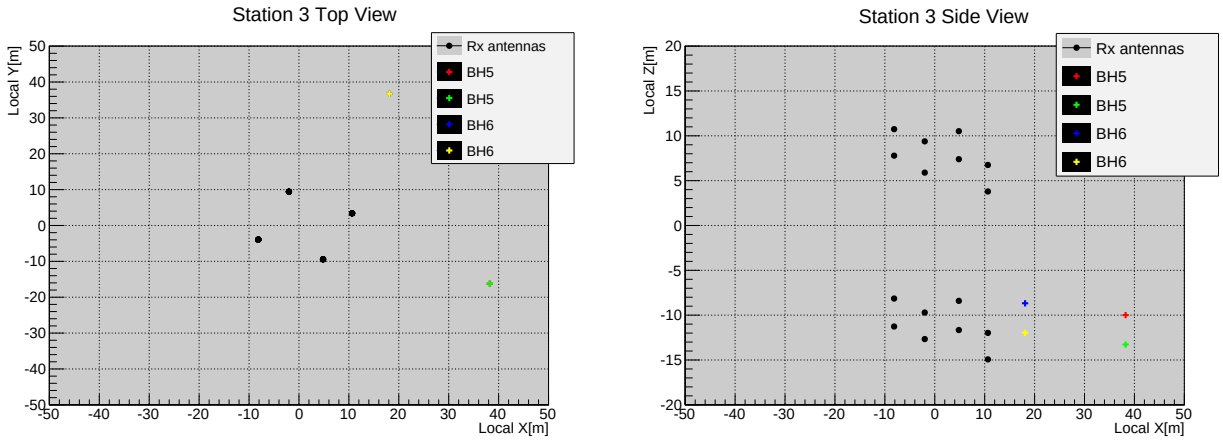
$$F\{f \otimes g\} = F\bar{\{}f\} \cdot F\{g\} \tag{4.2}$$

f and g represent waveforms. τ denotes the delay difference, and the actual operation of the cross correlation is performed in Fourier-transformed domain. This is later transformed back to time domain



(a) (b)

Figure 4.1: A2 antenna locations in the A2 station-centric coordinate system. (a): top view. "BH" stands for Borehole. BH5 and BH6 are calibration pulser strins. (b): side view. In each calibration pulser string, the top point is the location of the Hpol calibration antenna, and the bottom point is the Vpol calibration antenna.



(a) (b)

Figure 4.2: A3 antenna locations in the A3 station-centric coordinate system. (a): top view. "BH" stands for Borehole. BH5 and BH6 are calibration pulser strins. (b): side view. In each calibration pulser string, the top point is the location of the Hpol calibration antenna, and the bottom point is the Vpol calibration antenna.

where $(f \otimes g)(\tau)$ can be read off efficiently. The cross correlation function of waveforms from two channels can be visualized in Fig. 4.3.

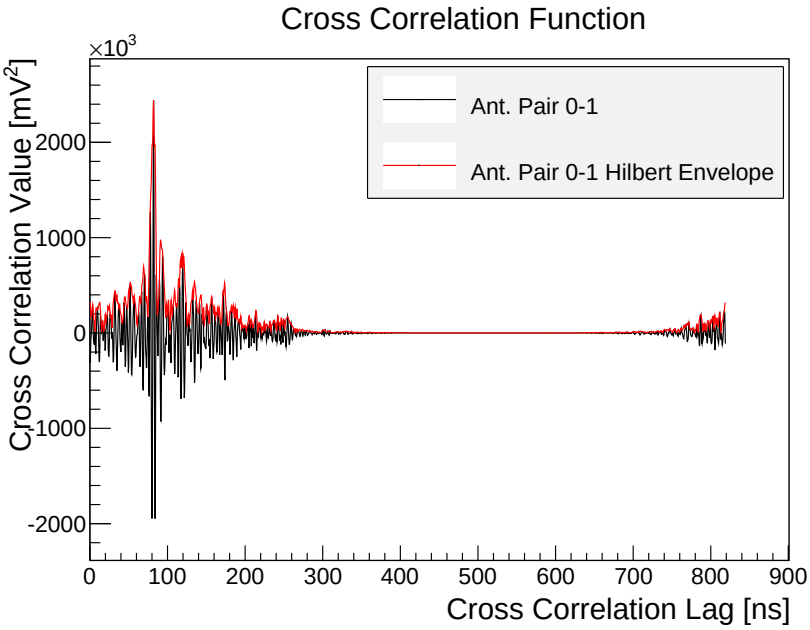


Figure 4.3: Cross correlation function (black) and its Hilbert envelope (red) of channels 0 (D1TV) and 1 (D2TV) from an calibration pulser event in A3. The cross correlation is performed in a wrap-around manner. A lag t' of channel 0 from channel 1 with $t' > T$ where T is the waveform length indicates that channel 0 is shifted to *lead* channel 1 by $2T - t'$.

A important step here after obtaining the cross correlation function is that we apply Hilbert envelope smoothing to the correlated function. The Hilbert envelope of a continuous function f can be calculated as

$$H(t) = \sqrt{f^2(t) + h^2(f(t))} \tag{4.3}$$

, where $h(f)$ denotes the Hilbert transform, which we calculate in the frequency domain by

$$F\{h(f)\} = e^{-\frac{\pi}{2}i} \cdot F\{f\}. \tag{4.4}$$

This has been empirically shown to produce more robust and consistent reconstruction results on calibration sources. We attribute this to our lack of understanding of the high frequency response of the detector. A Hilbert envelope serves to filter out higher frequency oscillations, thus avoiding unwanted cross correlation features from arising. A comparison of the reconstruction of a calibration pulser event from A3 is presented in Fig. 4.4. It is clear that before applying the Hilbert envelope, the reconstruction exhibits highly oscillatory features on the sky map that prevents a consistent localization of the peak coherence from event to event.

In contrast, after enveloping the "hotpsot" becomes more robust with an angular uncertainty of only $\sim 1^\circ$ event-to-event, as will be see in section 4.4.1.

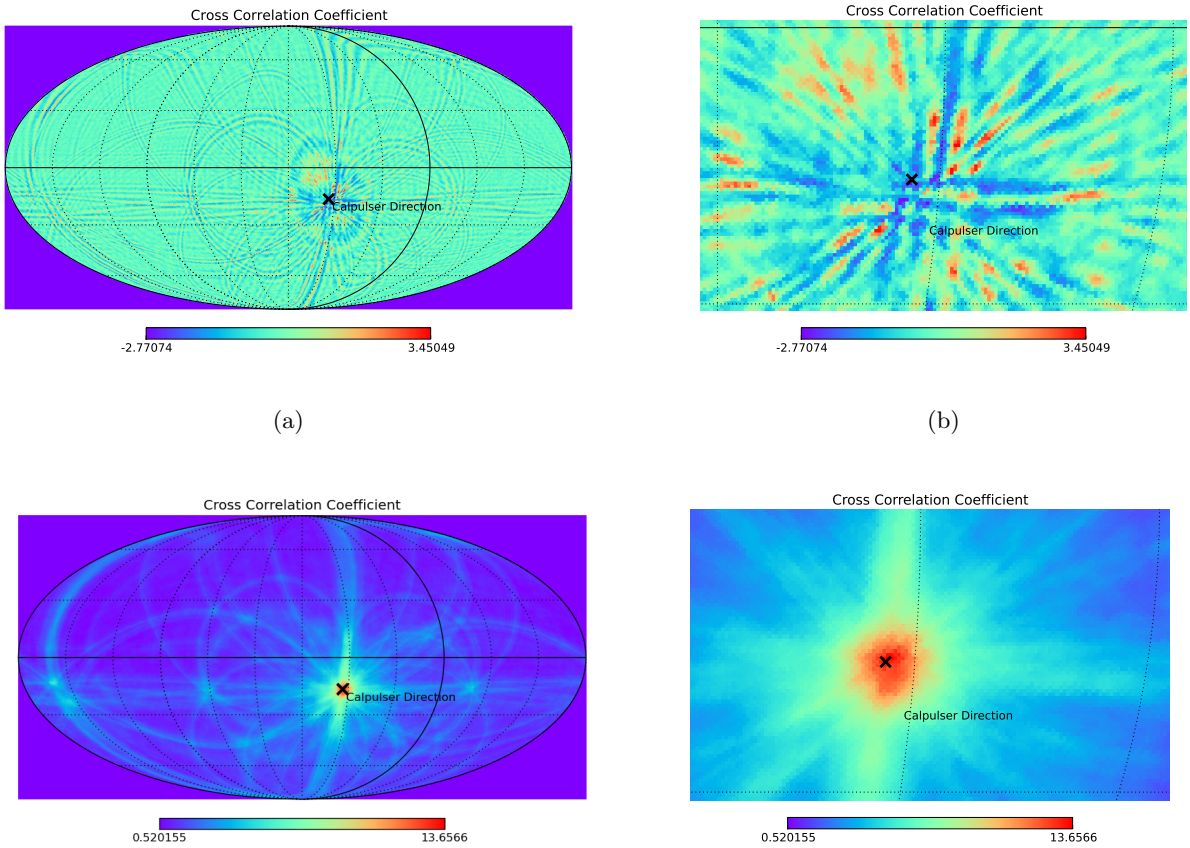


Figure 4.4: Calibration pulser event from A3 reconstructed at a distance of 40m from station center. (a): reconstruction sky map without Hilbert enveloping. The cross marks the nominal direction of the calibration pulser in use. The color scale represents the coherence value. (b): zoomed-in view of a $60^\circ \times 30^\circ$ box around the calibration pulser in (a). The oscillatory features obscure the resolving of the signal direction. (c): the same sky map when the cross correlation functions are Hilbert enveloped. (d): zoomed-in view of (c).

At this point, given a putative source direction $\hat{r}$ and its delays, we then define the coherence value as the sum of the enveloped cross correlation values from all same-polarization baselines after normalizing for channel-to-channel power level differences and number of baselines used. This is expressed as

$$C(\hat{r}) = \frac{2}{N_A(N_A-1)} \sum_{i=1}^{N_A} \sum_{j>i} \frac{H\left[(v_i \otimes v_j)(\hat{r})\right]}{\sqrt{\int_0^T v_i^2 dt \int_0^T v_j^2 dt}} \tag{4.5}$$

, in which $v_{i/j}$ is the voltage waveform of antennas i and j, N_A is the number of antennas involved in interferometry, and the denominator serves to normalize the cross correlation and remove effects from unequal noise levels in each channel. Physically, the coherence value can be understood as a metric of how well the signals in an event align with each other once shifted according to the delays. This is assuming the signals observed across antennas are similar, which is usually valid for signal sources far away. This also implies that only channels of the same polarization should be cross correlated with each other due to the common antenna response.

We define the putative vertex locations as the pixel centers of a collection of *HEALPix* [106] sky maps centered around our station at different radial distances. All sky maps shown in this contribution have 196,608 pixels (mean pixel spacing 0.4581°) unless otherwise specified. For neutrino reconstruction, 50 skymaps at radial distances from 40m to 5000m in equi-distance steps are generated. This is illustrated in Fig. 4.5. Delays are computed with the `Radiospline` [2] software package, which will be detailed in the following section. Note that we consider only vertex positions within the shadow boundary, and delays from only the direct rays are included. The use of Radiospline readily accounts for the ray-bending effect in this reconstruction technique. The reconstruction software is open-sourced and can be found at `https://github.com/mingyuanlu/ARA_reconstruction` with full prerequisite specification, installation instructions, and usage example. The software makes use of the `OpenCL` heterogeneous computing platform to perform cross-platform parallel computation that expedites the interferometry. Similar to `AraSim`, the reconstruction code reads a setup script that specified parameters that configures the runtime behavior, including the use of CPUs or GPUs, the number of pixels in a sky map, the number of sky maps, et cetera.

4.2.1 Radiospline

The clock time taken when calculating a ray path given a pair of endpoints with `RaySolver` is on average 0.21 ms on a typical CPU. Considering the aforementioned reconstruction setup, simply generating all needed delays comes with a heavy overhead of tens of minutes. `Radiospline` [107] is a C++ software tool that speeds the ray generation even further from `RaySolver`. It does so by interpolating a precomputed table of rays with B-splines [108] to enable fast and arbitrary ray property lookup. B-splines are essentially piecewise polynomials that can be linearly combined to perform curve-fitting. This brings the computation time of a ray to $\sim 0.4\ \mu s$. `Radiospline` is capable of computing rays with one endpoint in air and the other in ice. It now can provide lookup for the the arrival time of the first ray and the second ray, as well as

[2] `https://github.com/WIPACrepo/radiospline`

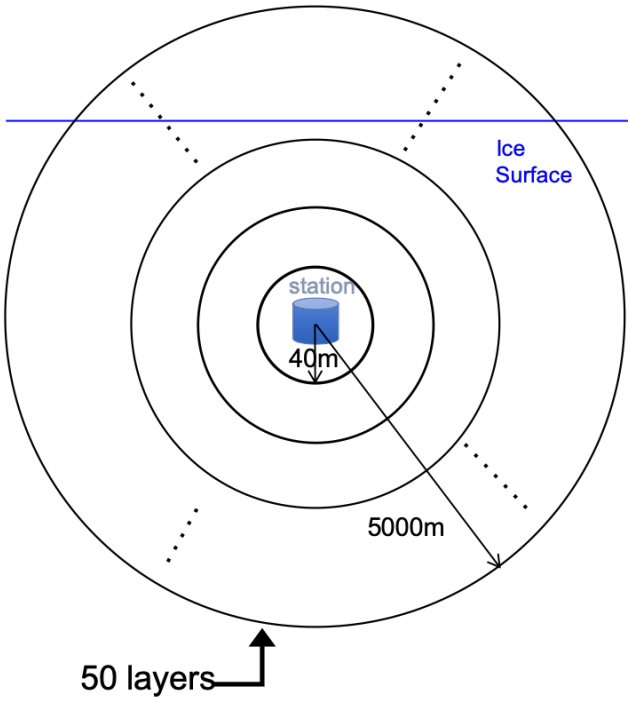


Figure 4.5: Diagram of the multiple sky map setup where each sky map represents a HEALPix map at a fixed distance from the station center. The distance ranges from 40 m to 5000 m in 50 steps, referred to as "layers" in the figure.

the arrival angle of the first ray. In the following we use the construction of the spline table which fits the arrival time of the first ray to illustrate the process, which can be generalized for other ray properties.

To generate an arrival time table for a set of grid points, we can reduce the dimensionality of the table to three variables: the source depth z_s, the target depth z_t, and the horizontal distance in between r. The arrival time for each grid point is precomputed with `RaySolver`. The table is created for both in-ice and in-air sources. The range of z_t is -20 to -250 meters, the range of z_s is +100 to -2850 meters, and r goes from 0 to 5000 meters. Next, we use B-splines to interpolate the tabulated $t_{arr}(z_s, z_t, r)$ values. The main difficulty in the process is the discontinuity of the ray at the shadow boundary, over which no physical arrival time can be calculated, as shown in Fig. 4.6. This is circumvented by artificially removing the discontinuity with a fitted extension of the arrival time across the shadow boundary, and creating an addition table that tabulates the shadow boundary $z_{shadow}(z_s, z_t, r)$, over which the spline-fit is performed. This way, the arrival time fit can avoid artifacts due to discontinuities at the shadow boundary, and a check against the shadow boundary table prevents the return of an artificial delay from the shadow region. The full ray lookup procedure can be summarized in Fig. 4.7. Given a pair of endpoints, the code first consults the shadow boundary table to

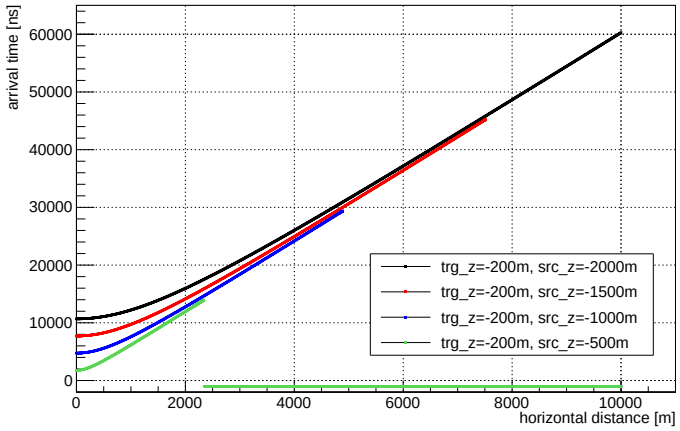


Figure 4.6: Arrival time to a 200m-deep taregt from various sources depths and distances as computed by `RaySolver`. `Radiopline` interpolation is performed over these data. The arrival time beyond the shadow region is manually set to -1000 ns here. The artifacts in the fit arising from such breaks and the solution is described in text.

determine if a ray solution exists. If so, depending on whether the one of the endpoints is in air or not, the in-air table or the in-ice table is used to evaluate the interpolated arrival time. We estimate the accuracy of the spline fit as compared to `RaySolver` to be 0.1 ns in air and 0.2 ns in ice. This accuracy is sufficient for the purpose of reconstruction.

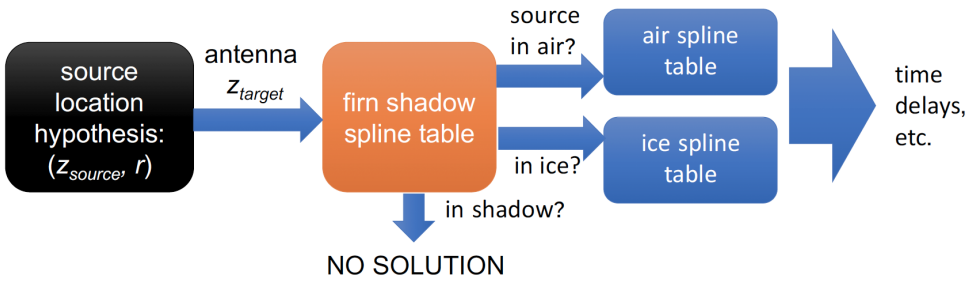


Figure 4.7: Given the location of the target and the source, `Radiospline` looks up the ray property according to this flow chart.

4.3 Neutrino Reconstruction

Neutrino simulation generated with `AraSim` can be reconstructed with the interferometric framework described so far. Fig. 4.8 is the reconstruction sky map for an 10^{18} eV neutrino, where the hot spot marks the reconstructed direction. The vertex of this event is 2581 meters away from the station center, and the reconstructed direction deviates from the true vertex direction by $< 0.3°$.

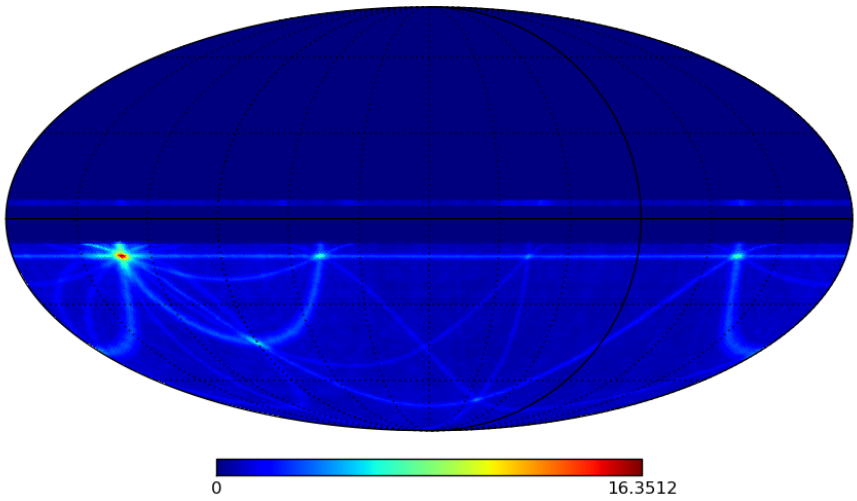


Figure 4.8: The reconstruction sky map containing the pixel of maximal coherence value using the default `Radiospline` reconstruction. The hotspot in red is formed by the intersection of "rings" on the sky map by each baseline, thus representing the direction where the corresponding `Radiospline` delays best agree with the signal arrival times from waveforms. There are another bright spots visible on the sky map but only correspond to the intersection of a subset of all available baselines. Notice the dark blue region represents pixels where no ray solution is found, either because they represent points above the ice that are not supported by `Radiospline`, or because they are in the shadow region for that sky map (layer) distance. The thicker black lines mark zeros for the zenith and the azimuth.

Fig. 4.9 demonstrates our ability to reconstruct vertex direction to a zenith resolution of 1.30° and a azimuth resolution of 0.31°. With the multi-layer sky map approach described here, we are not able to resolve the vertex distance. This, combined with ray-bending, limits us from attaining sub-degree zenith resolution. Fig. 4.10 shows the dependency of the angular precision on the coherence value, which exhibits the expected trend that more coherent signals can be reconstructed with better precision. Notice that in the figures presented in this section the coherence normalization is different from Eq. 4.5 by a constant factor. The scaling does no affect the behavior.

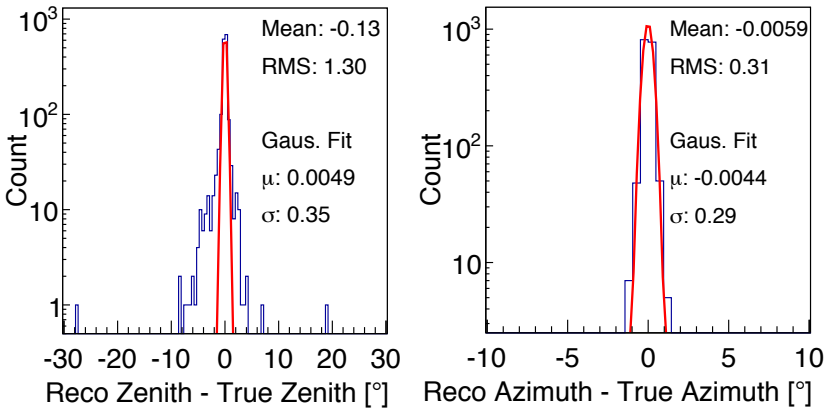
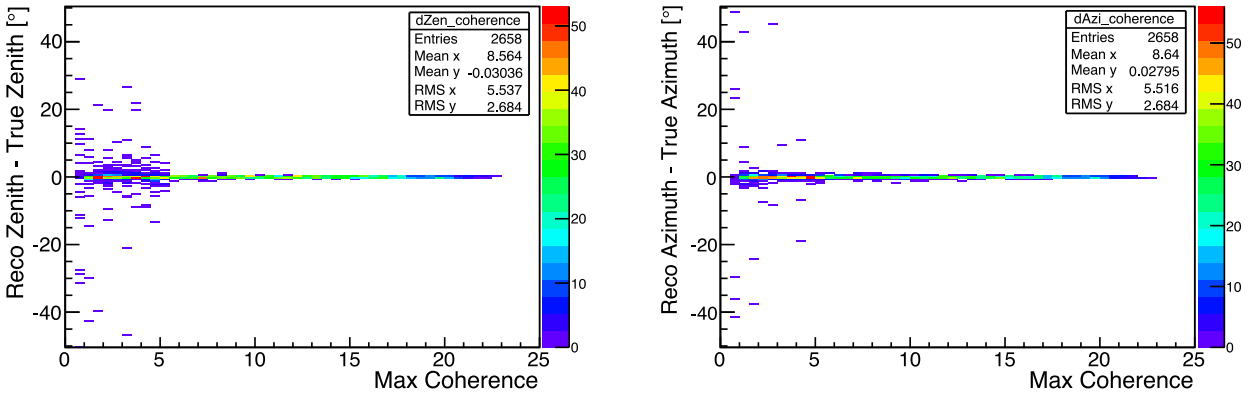


Figure 4.9: Angular accuracy and precision of the `Radiopline` reconstruction tested with 10^{18} eV neutrino sample from `AraSim`. The red curves represent Gaussian fits to the histograms. A simple reconstruction quality cut is applied here to remove poorly reconstructed events with low coherence values, as seen in Fig. 4.10.

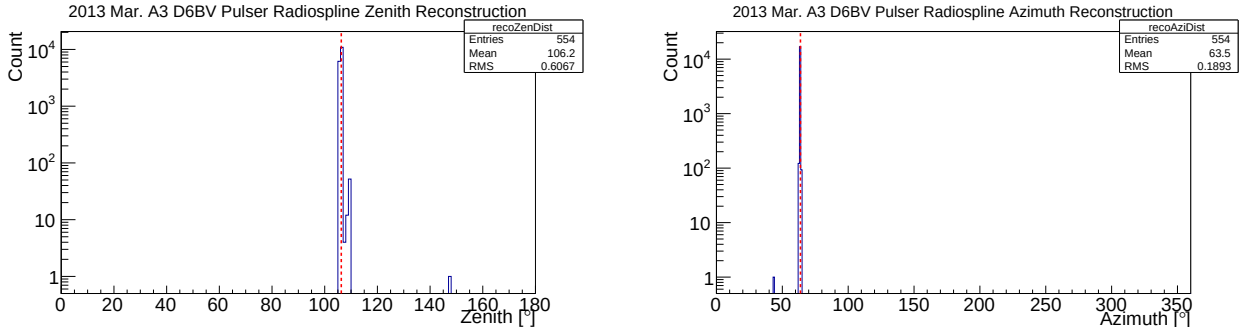


(a) (b)

Figure 4.10: Deviation of the reconstructed vertex angle from simulation-true values. (a): zenith (b): azimuth. Above a coherence value of ~ 5 the reconstructed angles agree with true values to within 1°. Events with low coherence values typically have small or no signal contents due to noise triggering or the event being distant/off-cone. See text for comment on coherence value nomalization.

4.4 Calibration Source Reconstruction

In the UHE regime, virtually no physics background is present in ice for ARA. For example, geomagnetic and Askaryan radiation from EAS form down-going backgrounds that are typically reject by a surface



(a) (b)

Figure 4.11: Comparison of the reconstructed calibration pulser (A3 D6BV) zenith and azimuth with the nominal values taken from calibrated station geometry files. The red dashed lines mark the nominal values. Data are drawn from random 1% of all calibration pulsers events from A3 in the month of March in 2013. The distance of the layer used is 40 meters.

geometric cut. Penetrating high-energy muon bundles capable of generating dense cascades in ice due to catastrophic energy loss is $\sim 10^4$ times smaller than the UHE neutrino flux at 10^{18} eV. Therefore, ARA relies on multiple calibration sources in-ice and near-surface to calibrate its sensitivity and to ensure operational integrity. In this section, the reconstruction of these calibration sources and their implication on the diffuse search will be described.

4.4.1 Calibration Pulser

A critical test to the validity of the reconstruction algorithm is to reconstruct the local calibration pulsers. These pulser events are characterized by strong signal contents (SNR ≥ 5), well-defined timing patterns, and an almost-constant presence at a fixed rate. First, a separate study uses the arrival times of the calibration pulser signal to calibrate the location of the antennas and the cable delays of each measurement string [93]. My contribution to this stage is the further calibration of A2 in channels that were left out as outliers in the initial effort, specifically D3BV and D6BV. This work is detailed in appendix B. With the station geometry calibrated to the best of our knowledge, it is also crucial that the calibration pulsers can be localized after being reconstructed for the purpose of background rejection. As can be seen in Fig. 4.4c, the interferometric algorithm succeeds in pinpointing the pulser direction. In Fig. 4.11, A3 D6BV calibration pulser events from a month in 2013 are reconstructed at a distance of 40 meters using all eight Vpols. We summarize the reconstruction results as compared to nominal values in Tab. 4.1, which demonstrates that the reconstruction satisfies the 1° angular precision requirement, and the accuracy is within the angular error. Given that the

40-m sky map proves to be fitting for identifying the calibration pulser in the local sky, in the diffuse search we use the same setup to geometrically reject potential backgrounds from the direction of such pulsers.

Table 4.1: Comparison of the mean of the distribution from Fig. 4.11 to the nominal values of A3 D6BV zenith and azimuthal angles.

	Nominal [°]	Reco Mean [°]	Diff. [°]
Zenith	106.3	106.2	-0.1
Azimuth	63.8	63.5	-0.3

4.4.2 Deep Pulser

Due to the kilometer-scale RF attenuation length, the detectable volume of ARA is largely deep ice, especially at depths of 1000-2000 m. It is therefore important to understand the sensitivity of the detector for emission sources in those distance. Anticipating this, during the last year of IceCube construction three pulsers were deployed on IceCube strings 1 (1400 m, 2450 m) and 22 (1400 m). They are referred to as the deep pulsers IC1S, IC1D, IC22S respectively. I carried out a study gathering all known information regarding the position of these deep pulsers, and the report is attached as appendix C. This provides the nominal values of the orientation of the deep pulsers to be used in comparison with reconstruction outcomes, and establishes the horizontal distances of the deep pulsers to be roughly 3400 m from A2, and 4000 m from A3. The pulsers were operated in 2015 as part of the calibration of A2 and A3, and again in 2017. Although the 2017 run is outside of the diffuse search period of this thesis, due to its valuable data we study it here. Due to difficulties in identifying events from IC1D with timing consistently, we focus on reconstructing events from IC1S and IC22S here. Fig 4.12 shows the ARA and IceCube layout, with two rays propagating from IC1S to A2. The presence of two rays from source to receiver is generic for our geometry. These rays are calculated according to our default $n(z)$ model Eq. 1.32.

Fig. 4.13 shows a deep pulser event from A2-IC1S. The pulses are emitted from a Vpol antenna and the Vpol response is dominant; however, there is a 10% Hpol signal also received. Moreover, the Hpol signals arrive before the Vpol signals. This is evidence for birefringence, in which the index of refraction of ice differs depending on polarization. This is explored further in [110]. Considering just Vpol signals from the same layer, it is evident that the first set of pulses, which correspond to the first ray, is a up-going signal consistent with the expectation from the relative geometry. Most antennas show two pulses. The separation between the two pulses is commensurate with the predicted difference in arrival times for pulses traveling along the first and second rays from IC1S to A2. The quality of the pulses is visible from the leading edge. The first ray travels just through deep ice. Apparently there is little scattering and the leading edge is quite sharp. However, the pulses identified as the second rays have a precursor which may be due to scattering in

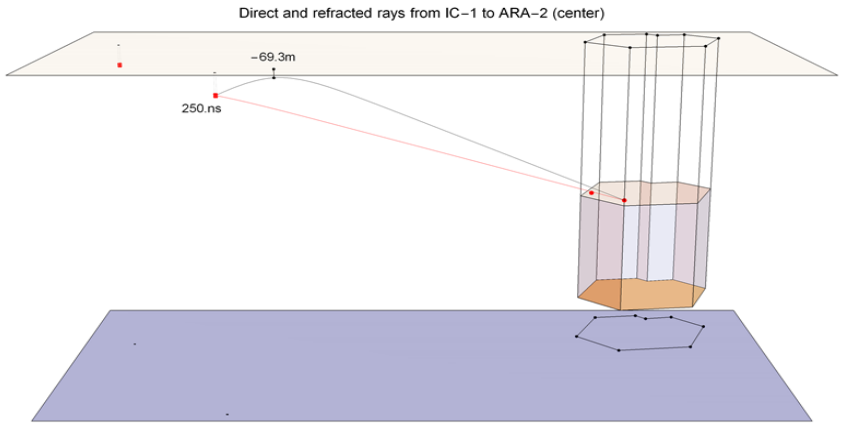


Figure 4.12: View of A2, A3 and IceCube [109]. The two deep pulsers (IC1S, IC22S) are located at a depth of 1400 m, at a horizontal distance of 3400 m from A2. When a pulser fires the radiation travels to a station via two paths. Both IC1S and IC22S can be seen by A2 and A3. Here only the rays from IC1S to A2 are shown. The minimum depth of the second ray and the difference in arrival times are shown on the figure.

the firn layer near the top of the pulse trajectory. Similar features can be observed for A3. Fig. 4.14 gives an example event from A3-IC22S.

In Fig. 4.15, the result of reconstructing IC22S using A3 is shown using first rays in the Vpols. The angular error of $\sim 1^\circ$ indicates a successful reconstruction that validates the station geometry, the index of refraction model, and the paradigm of geometric optics operating on multi-km scales. To utilize both pulses from the first ray and the second ray, pulse "snippets" of 200 ns are cropped from the full wavefrom for each Vpol channel, and interferometry involving the delays from both rays is run with these snippets. The snippets are defined in this exercise by visually identify rising edges for the signals, and capture 50 ns pre-edge plus 150 ns post-edge data. We note that this optimized procedure obviously is not applicable generally, but is tractable for this small-event sample. Future analyses wishing to perform similar reconstruction for a general dataset may need algorithms to identify the double pulses in each waveform ahead of running interferometry.

Fig. 4.16 and Fig. 4.17 demonstrate the reconstruction results from the two stations and the two deep pulsers we consider here using both rays in each case, supporting the conclusion that the ARA station is indeed sensitive to RF sources in the deep ice and is capable of source reconstruction. The accuracy of the reconstruction is summarized in Tab. 4.2.

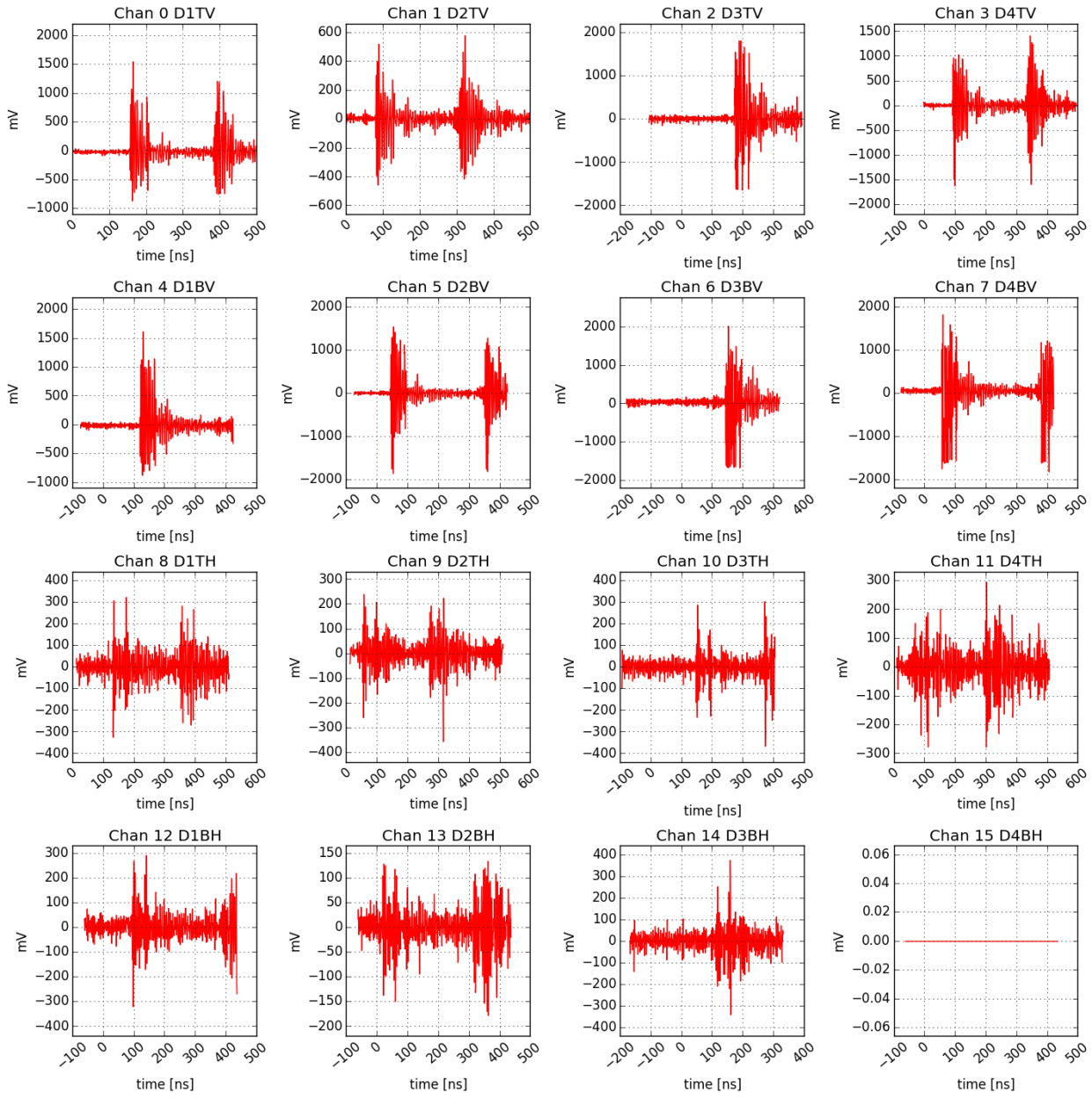


Figure 4.13: A2 run 8573 event 403, a deep pulser event from IC1S in early 2017. The second pulse is absent in D3TV and D3BV due to extra cable delays in D3, leading to late signal being left out from the readout window.

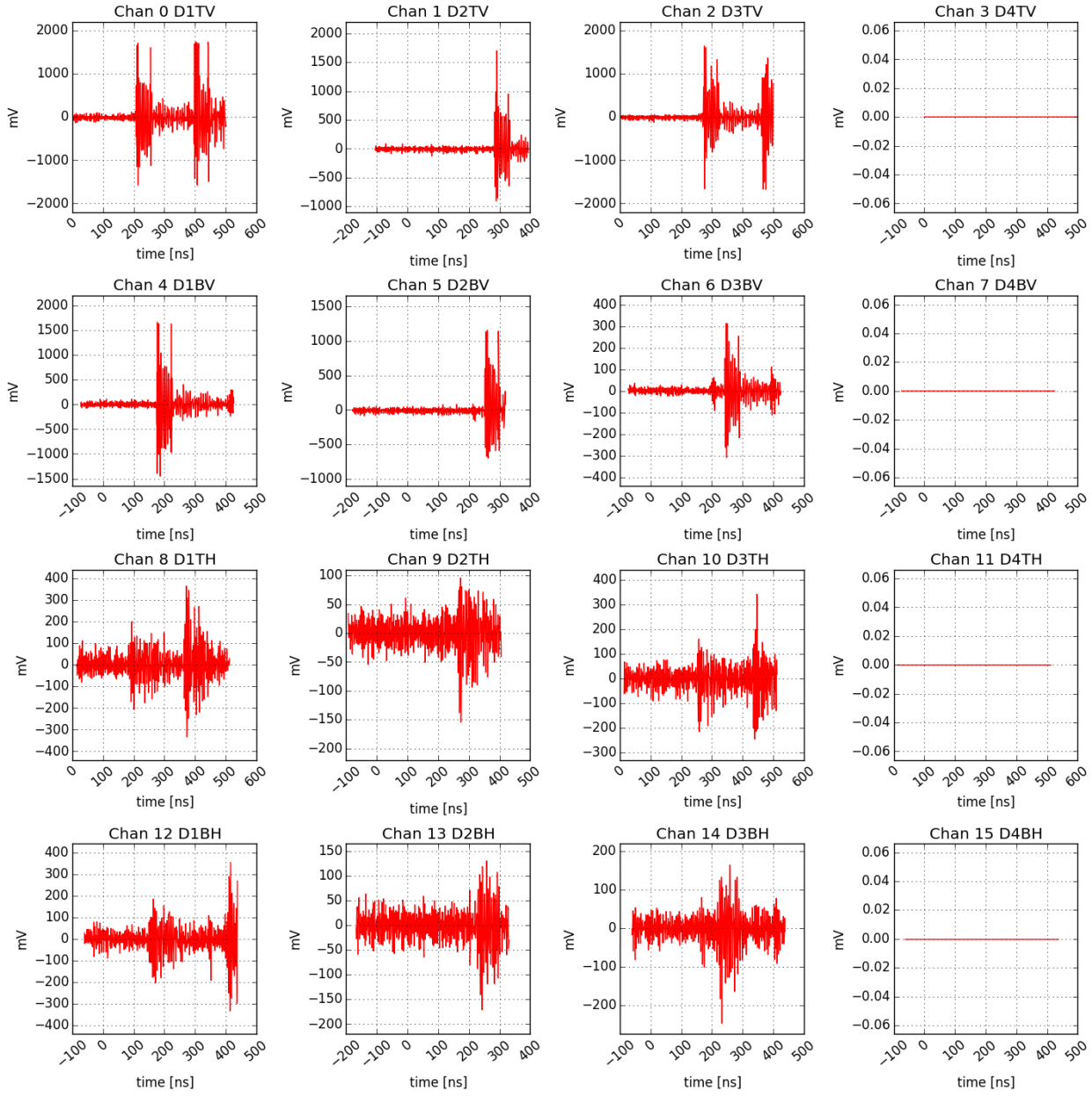


Figure 4.14: A3 run 3811 event 12472, a deep pulser event from IC22S in early 2015. The second pulse is absent in D2TV and D2BV due to extra cable delays in D2. D4 in A3 produces corrupted waveforms, and is zero-ed out.

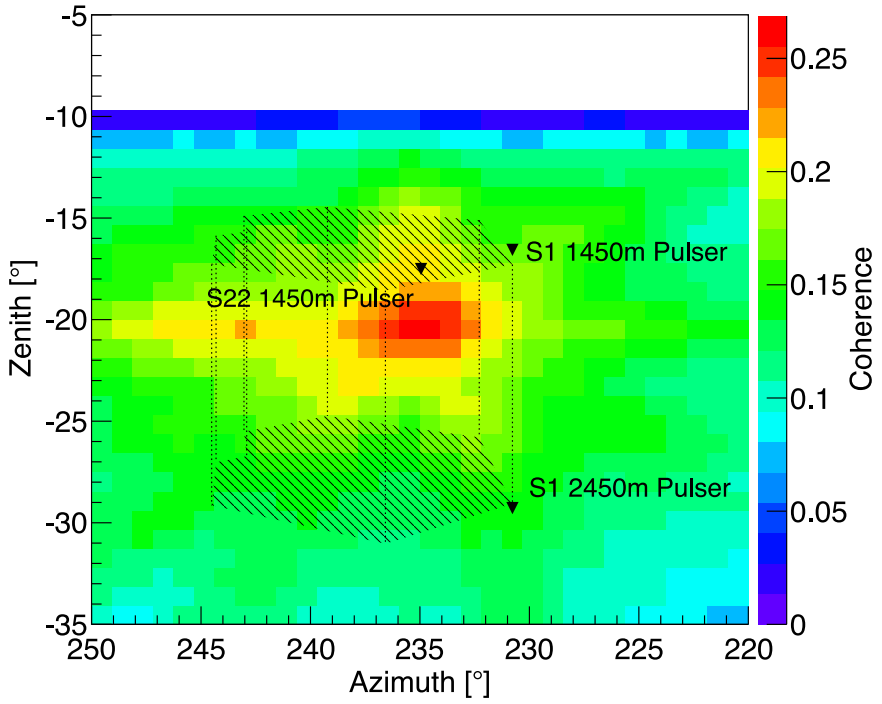


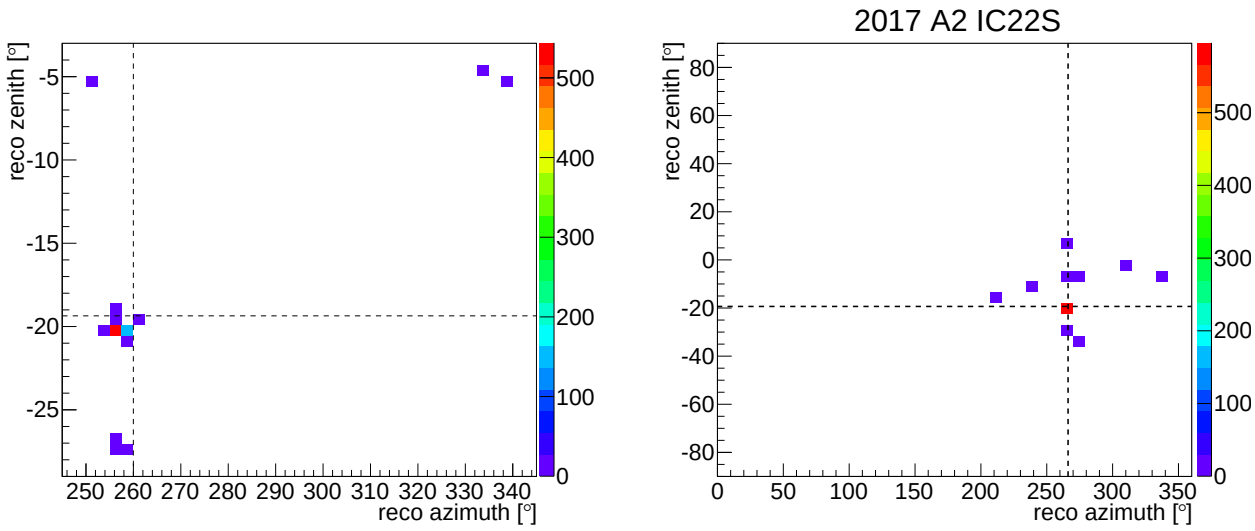
Figure 4.15: Reconstruction of the "sky" position of IC22S as seen from A3 (run 3811 event 12472). This reconstruction uses just the Vpol direct rays. The shaded hexagonal cylinder represents IceCube projected on the local sky of A3.

Table 4.2: Comparison of the nominal deep pulser locations and the double-pulse reconstructed results. The last column is the percentage differences of the reconstructed range r from the nominal range

	Nominal			Reconstructed			
	θ [°]	ϕ [°]	r [m]	θ [°]	ϕ [°]	r [m]	Δr/r [%]
A2 IC1S	-19.36	259.99	3666	-20.09	257.6	4215	+15
A2 IC22S	-19.65	266.17	3609	-20.74	264.4	4896	+36
A3 IC1S	-16.57	230.77	4269	-18.21	231.7	4711	+10
A3 IC22S	-17.51	234.95	4040	-18.84	235.0	4298	+6

4.4.2.1 Deep pulser range reconstruction

Double pulse events as seen from the deep pulsers provide unique access to the range/distance reconstruction. As mentioned earlier, with the current station size and the systematic uncertainties, resolving



(a) (b)

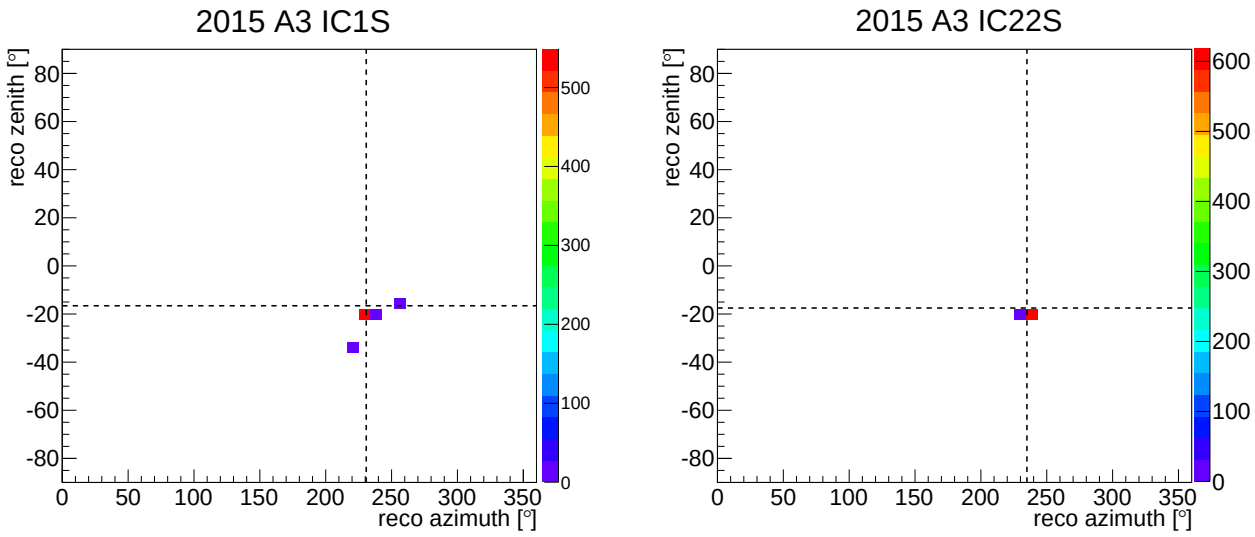
Figure 4.16: Results of the angular reconstruction of deep pulsers (a): IC1S (b): IC22S using Vpols from A2 run 8573. Delays from both rays are computed with `Radiospline` and the snippets are cross correlated accordingly. The black dashed lines mark the nominal direction of each deep pulser.

the distance of the vertex in deep ice is unattainable. The refracted/reflected ray, however, provides additional information that improves the range reconstruction ability. These rays can be viewed as received from a "imaginary" detector symmetrically above the ice surface, effectively augmenting the vertical detection baseline to twice the station depth, or roughly 400 meters.

Using the aformentioned waveform snippets, we are able to reconstruct the range of the deep pulsers to approximately 15%, as tabulated by Tab. 4.2. The cumulative results over many events are shown in Fig. 4.18. Deviations in elevation reconstruction of order one degree typically translate into range errors of tens of percent. The capability to reconstruct the range of a distant in-ice source has important implications as it can be translated to an estimate of the shower energy in the case of neutrino detection. This work is the first in the field to demonstrate the viability of achieving such reconstruction using double-pulse events with data. Future detector designs are actively taking the double-pulse sensitivity into account to further exploit this class of particularly physics-rich events.

4.4.3 ICL Rooftop Pulser

In addition to reconstructing in-ice sources, ARA also needs to effectively reject backgrounds from the ice surface and above. RF interference propagating at the ice surface could reach the detector through downward refraction at the total internal reflection (TIR) angle, and following the Snell's law the signal is incident upon



(a) (b)

Figure 4.17: Results of the angular reconstruction of deep pulsers (a): IC1S (b): IC22S using Vpols from A3 run 3811. Delays from both rays are computed with `Radiospline` and the snippets are cross correlated accordingly. The black dashed lines mark the nominal direction of each deep pulser.

the station at the critical angle of $sin^{-1}(\frac{1}{n_{ice}(-180m)}) = 34.7°$. To test the ability to reconstruct a signal emitted at the ice surface, we use a radio pulser mounted on the rooftop of ICL as a calibration source. This pulser consists of a dual-polarization quad-ridge horn antenna from the ANITA experiment and a pulser unit that can be configured to pulse at a fixed rate in either polarization. The pulser is sufficiently far away from A2 and A3 ($\sim$ 4 km) with a small elevation off ground ($\sim$10 m) that the incident angle on the ice surface can be considered close to 90°. The geometry of the rooftop pulser relative to ARA and the signal path is illustrated in Fig. 4.19. In Fig. 4.20 and Fig. 4.21, the waveforms of rooftop pulser events from A2 and A3 are presented. During the run, the pulser is transmitting in the horizontal polarization, and so in the interferometry we use only the Hpols to cross correlate.

When reconstructing sources from above the ice surface, we take a modified approach to the interferometry. Specifically, instead of obtaining delays from the in-air table of `Radiospline`, we reconstruct with delays generated according to pixel locations from a single sky map at a distance of 5 km assuming a constant index of refraction (bulk-ice) of 1.76. This latter choice is referred to as the *quasi-planewave* mode, since the wavefront from a distant source (5 km) is virtually planar as perceived by the station. The use of the quasi-planewave reconstruction is motivated by the inability to reconstruct consistently the zenith and azimuth the the ICL rooftop pulser signal using the default multi-layer `Radiospline` framework. We

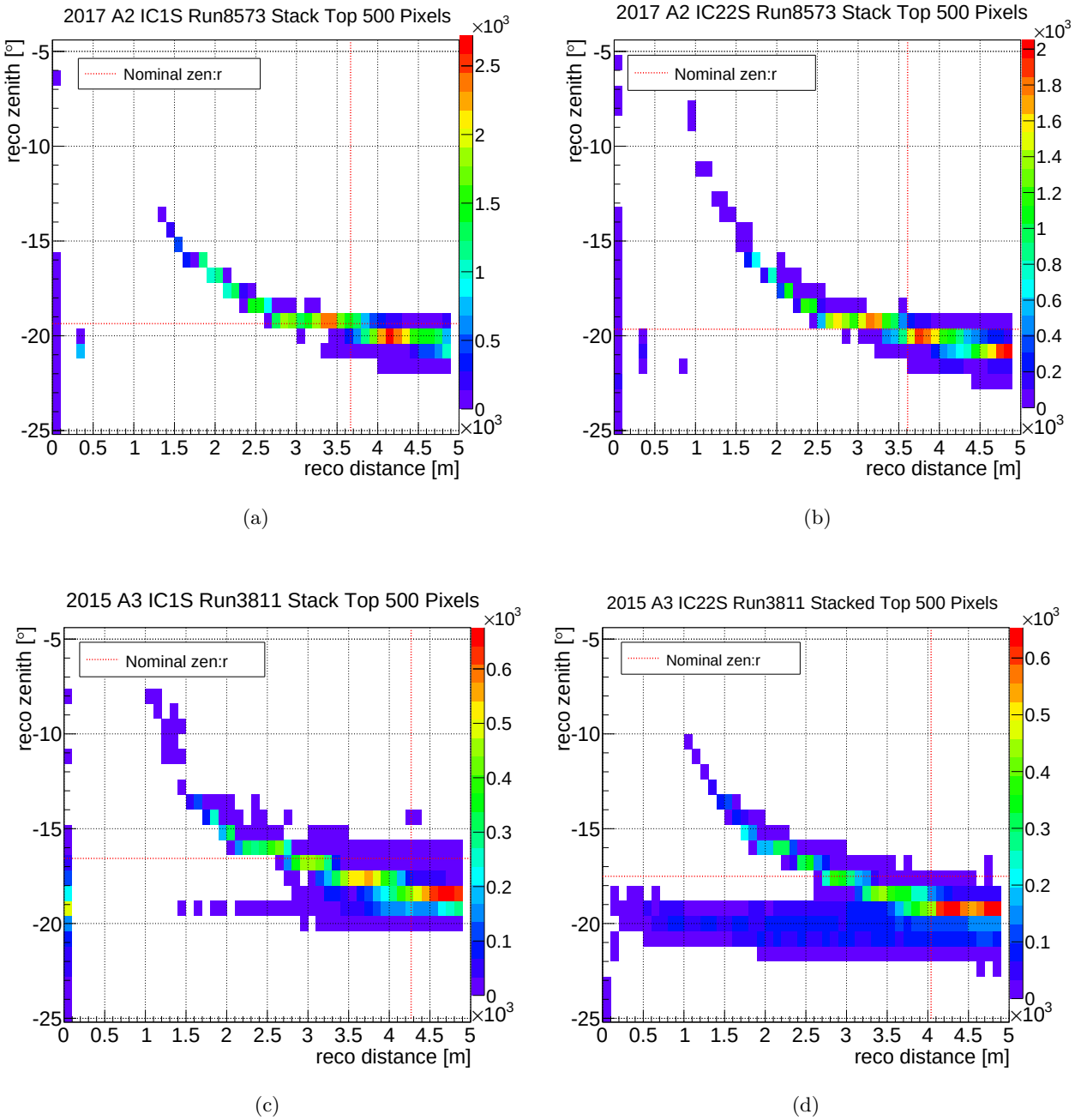
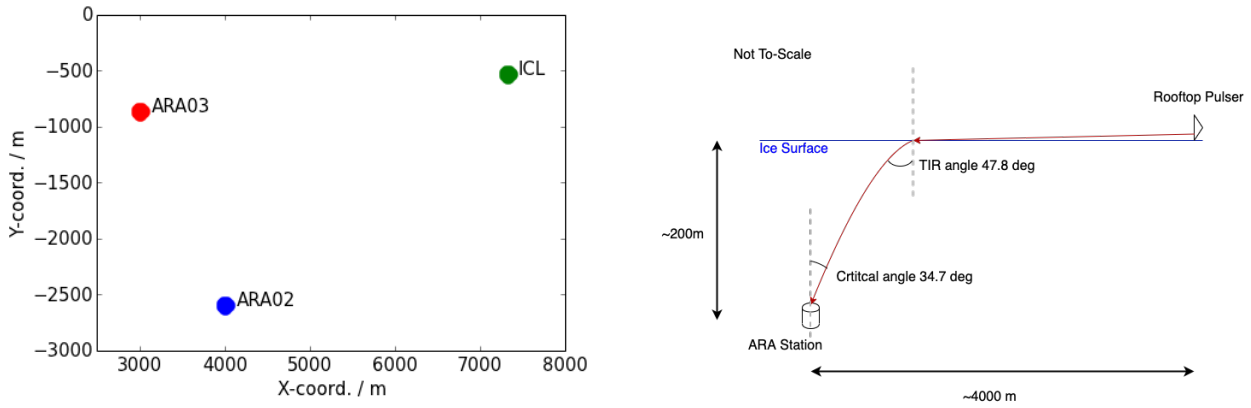


Figure 4.18: 500 pixels with top coherence values in each deep pulser event reconstruction in the deep pulser run, stacked across events in the zenith and range space. This illustrates the uncertainty of the reconstruction and the correlation of the zenith and the range. The red dahsed lines mark the nominal values in each axis.

attribute this inability to the multiple entry points of the signal path from the rooftop pulser to each in-ice



(a) (b)

Figure 4.19: Diagram of the ICL rooftop pulser geometry and signal path. (a): top view of the positions of the ICL and A2/A3 [93]. (b): illustration of the signal path and relevant angles from the rooftop pulser to an ARA station.

antennas. The default mode seeks a maximal phase-match signal point source, which does not exist for the case of the surface-propagating signal such as from the rooftop pulser. The quasi-planewave approach is empirically shown to consistently reconstruct the direction of the rooftop pulser signal with both A2 and A3. Fig. 4.22 and Fig. 4.23 illustrate the reconstruction results. Tab. 4.3 summarized the accuracy and the precision of the reconstruction as compared to nominal values. Aside from the $\sim 10^\circ$ azimuth offset in A2, which can be explained by a systematic rotation of the station-centric coordinate system with respect to the global system, the reconstruction agrees with the "known" values to $\leq 1^\circ$. The systematic rotation issue is unimportant for the station-centric neutrino search we perform here.

Table 4.3: Comparison of the nominal rooftop pulser direction and the quasi-planewave reconstructed results.

	Nominal		Reconstructed			
	θ [°]	ϕ [°]	θ [°]	ϕ [°]	θ RMS [°]	ϕ RMS [°]
A2	55.3	265.2	55.48	257.1	0.27	0.92
A3	55.3	237.7	55.59	237.1	0.34	0.78

Considering the success of the quasi-planewave approach in reconstructing surface events, we incorporate it as a standard method in software. At the analysis stage, every event is reconstructed with the default multi-layer `Radiospline` approach as well as the quasi-planewave approach. Surface background rejection

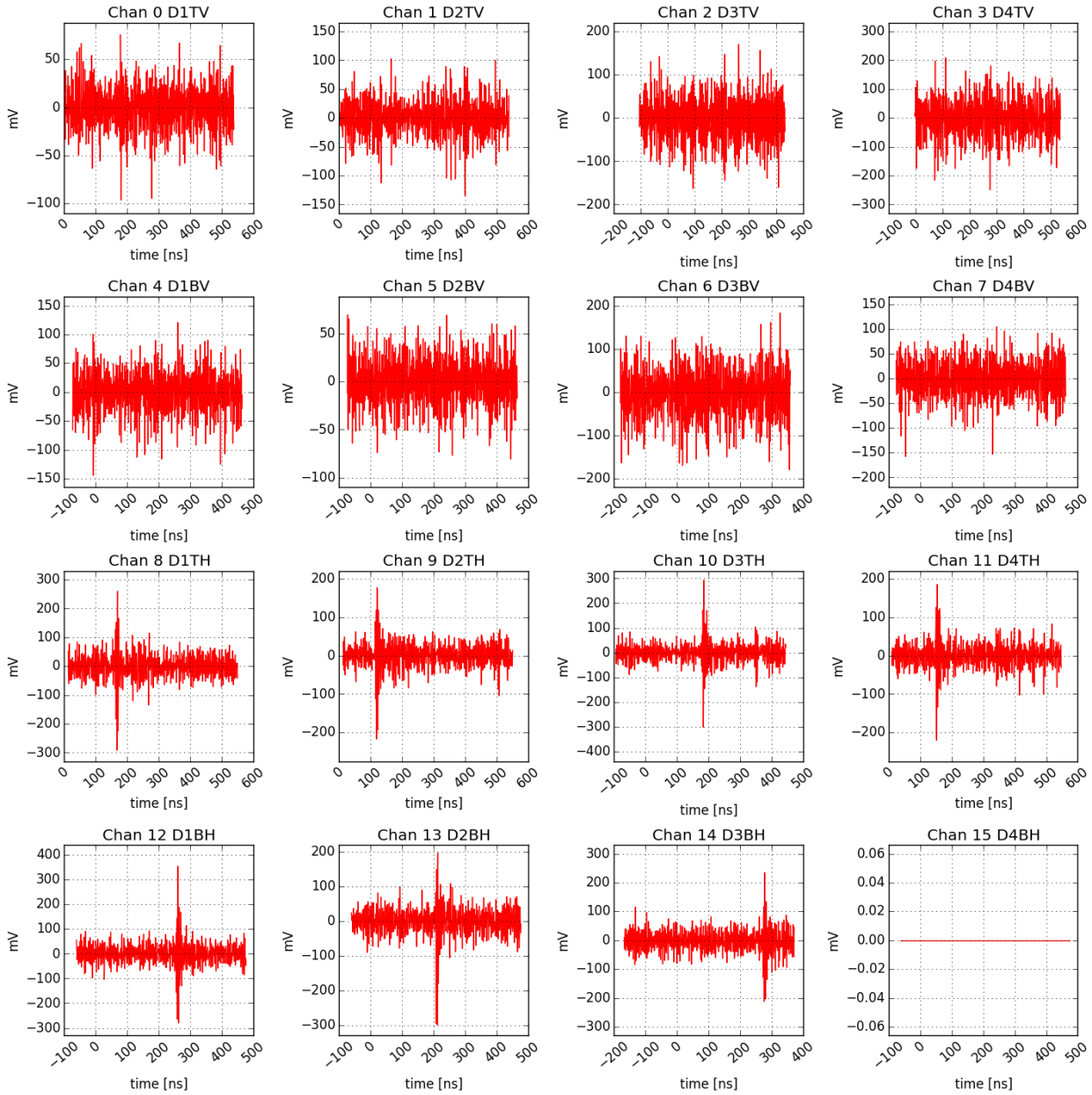


Figure 4.20: A2 run 1122 event 103, an Hpol rooftop pulser event in early 2013.

is done by cutting events where the quasi-planewave reconstructed arrival angle is above a surface cut angle that is derived from data. This is described in section 5.6.1.4.

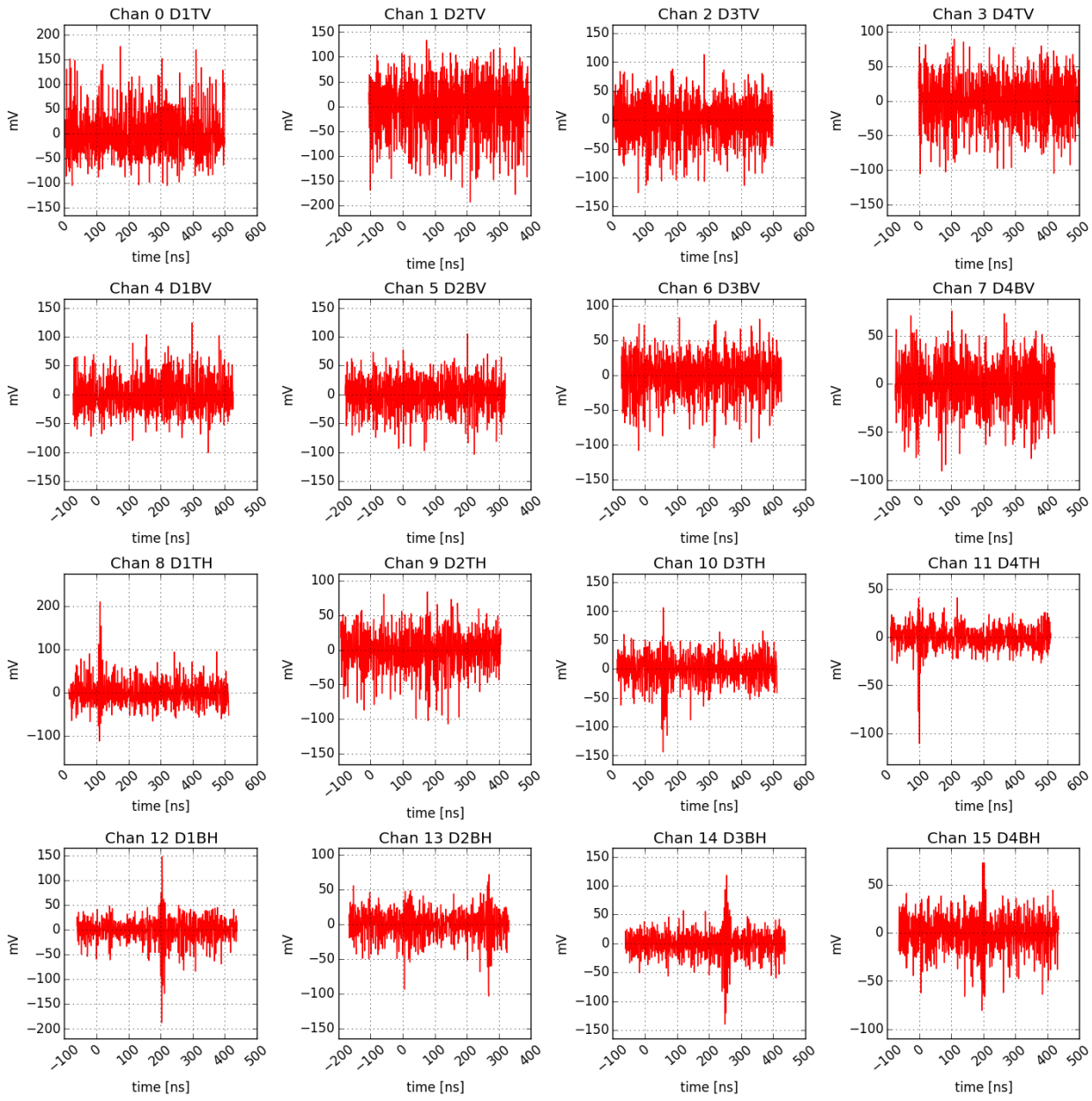
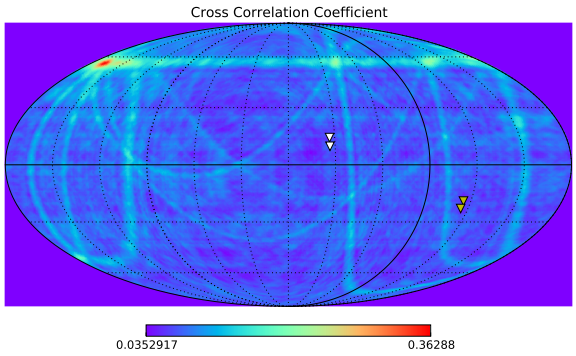
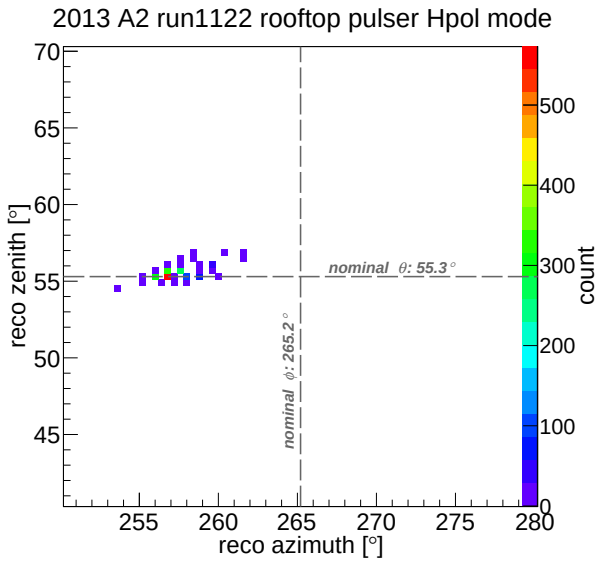


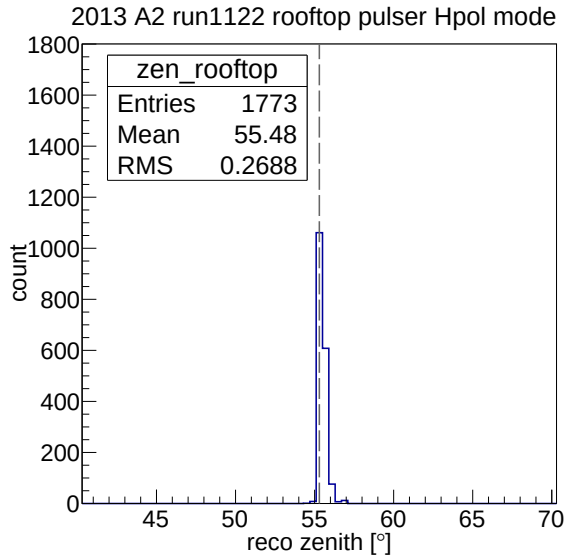
Figure 4.21: A3 run 64 event 4203, an Hpol rooftop pulser event in early 2013. At this very early run D4 of A3 still produced useful data and so is included here.



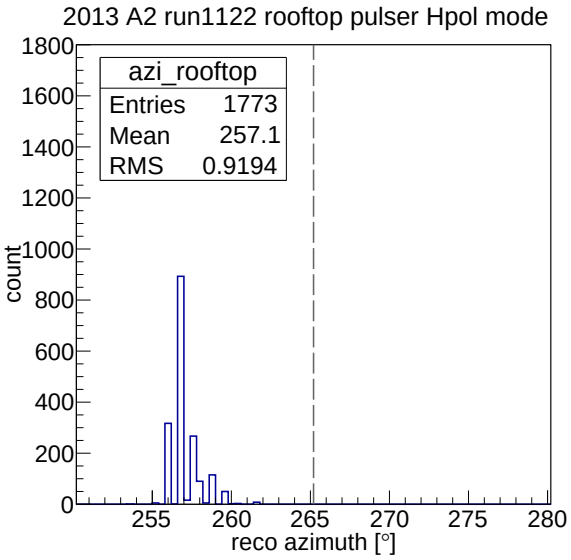
(a)



(b)

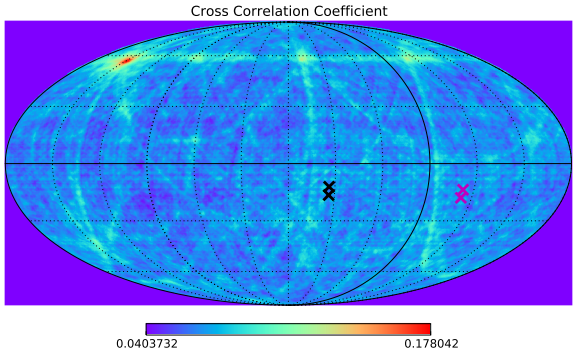


(c)

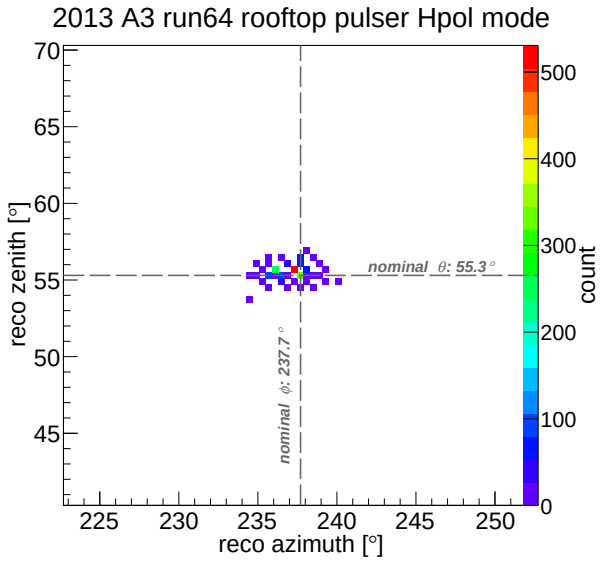


(d)

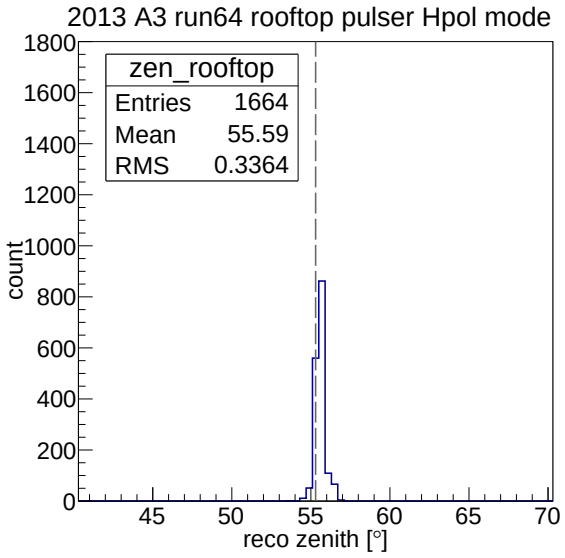
Figure 4.22: Hpol reconstruction results of the ICL rooftop pulser with A2 run 1122. Quasi-planewave mode is used. (a): event 103 reconstruction sky map. The triangles mark the directions of the local calibration pulsers, which are not relevant here. (b): comparison of the reconstructed and nominal direction of the rooftop pulser. The systematic bias in the azimuth indicates a rotational correction needs to be applied to the station-centric coordinate system relative to the South Pole global coordinate system. (c),(d): histograms of the reconstructed zenith and azimuth. The grey dashed lines show the nominal values. Again, the precision of the reconstruction is typically $\leq 1°$.



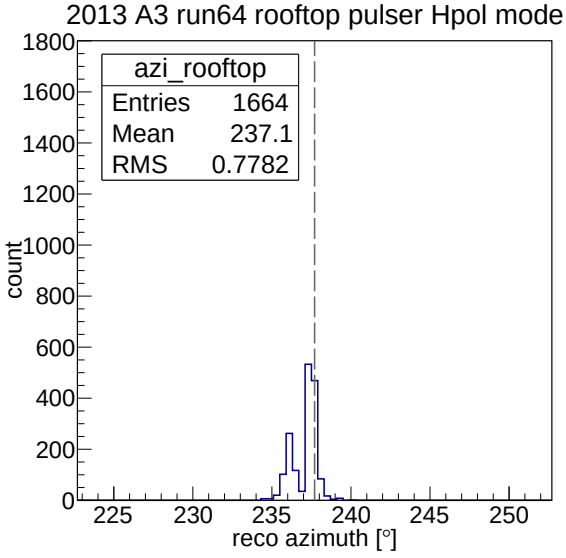
(a)



(b)



(c)



(d)

Figure 4.23: Hpol reconstruction results of the ICL rooftop pulser with A3 run 64. Quasi-planewave mode is used. (a): event 4203 reconstruction sky map. The crosses mark the directions of the local calibration pulsers, which are not relevant here. (b): comparison of the reconstructed and nominal direction of the rooftop pulser. (c),(d): histograms of the reconstructed zenith and azimuth. The grey dashed lines show the nominal values. Again, the precision of the reconstruction is typically $\leq 1°$.

Chapter 5

Diffuse Neutrino Search

In this chapter, we are going to describe the data and methods used in the diffuse neutrino search. Such an analysis can be carried out after we have established a reconstruction method and have calibrated the detector to the best of our knowledge — both of which are done and described in the previous chapter. The chapter is structured to introduce the blind analysis procedure first, and proceeds to describe the data and event selection criteria. Finally a two-stage analysis approach is presented, leading to the final background estimates and signal sensitivity.

5.1 Blind Analysis

We follow the procedure of a blind analysis in this search. A random set of % of all collected data is selected as the unblinded dataset, while the remaining 90% of the data is kept blinded. The randomness is at the event level, where each event has a 10% chance of ending up in the unblinded set. In principle, we use the unblinded set to identify backgrounds, and develop cuts to distinguish neutrino signals from backgrounds. Once we understand the types and frequency of backgrounds in our data, and defined a set of cuts to reject them, we proceed to apply the same set of cuts to the 90% blinded dataset, an act referred to as "unblinding". Ideally the event selection and cuts are "frozen" at this stage, and discovered events in the 90% set after unblinding can be considered neutrino candidates. The purpose of a blind analysis is to avoid potential biases from the experimenters, which could be a result of cut-tuning for significance after having already examined all obtained data. The underlying assumption for freezing the cuts at unblinding is that the backgrounds present in the full data can be fully characterized by the 10% sample. In practice, this is not always the case, and further study may be needed to determine whether the events in the full sample after the cuts have been applied are indeed neutrino candidates. The proper procedure depends on the findings of unblinding.

5.2 Data Processing

Data from 2013 to 2016 for A2 and A3 are used in this search. The raw data from disks are uploaded to the WIPAC [1] data warehouse. The data processing team in ARA then converts raw, or "L0", data to the "L1" dataset that can be consumed by analyzers. The conversion from L0 to L1 data involves first decompressing raw data. Consistency checks follow to make sure the data structures are not corrupted. The processor then proceeds to produce `ROOT` [2] files containing event `TTrees`.

The processing of relevant data is done by Christian Bora and Andrew Shultz. The processed L1 blinded and unblinded data are located in different directory paths in the data warehouse to avoid confusion. Notice that the blinded data contains the 10% unblinded data. In other words, the 10% unblinded sample is duplicated to a separate "ublinded" directory. For 2013, the unblinded and blinded sample can be found at (in the WIPAC server):

`/data/wipac/ARA/2013/filtered/burnSample1in10/`

`/data/wipac/ARA/2013/filtered/full2013Data/`

In and after 2014, the unblinded and blinded data can be found at:

`/data/wipac/ARA/$YEAR/unblinded/L1`

`/data/wipac/ARA/$YEAR/blinded/L1/`

Currently the collaboration shares common L0 and L1 data. In the future, we propose a five-level hierarchy (L0 - L4) of data processing that is summarized in Fig. 5.1.

L0 is the raw data, either from shipped ARA data disks or transferred via satellite to the data warehouse as filtered data. Data at this level is in compressed format. L0 to L1 processing involves decompressing and consistency checks. L1 is the uncalibrated data in `ROOT` files. L1 to L2 processing involves several operations. We subtract the pedestal values from each sample and convert the ADC counts to voltages (mV). Additional calibration on this ADC-to-Voltage conversion is also applied according to the works of Thomas Meures [93]. We also apply cable delays correction and common mode correction here. L2 hence represents calibrated data. L2 to L3 processing is where we apply physics filters. For a diffuse search, a thermal filter can be implemented that filters 99% of thermal events. One can implement adequate filters for other types of physics. For example, a GRB correlation study. Filters run in parallel. L3 therefore is filtered data sets as result of different filters. L3 to L4 processing involves applying reconstruction to reject backgrounds and search for neutrino candidates. Different sets of L3 data can be reconstructed with the same algorithm, or each set can receive its dedicated reconstruction. After reconstruction, a set of event summary can be generated. This could be the event trigger settings, the root-mean-square of the waveforms,

[1] Wisconsin IceCube Particle Astrophysics Center

[2] `https://root.cern.ch/`

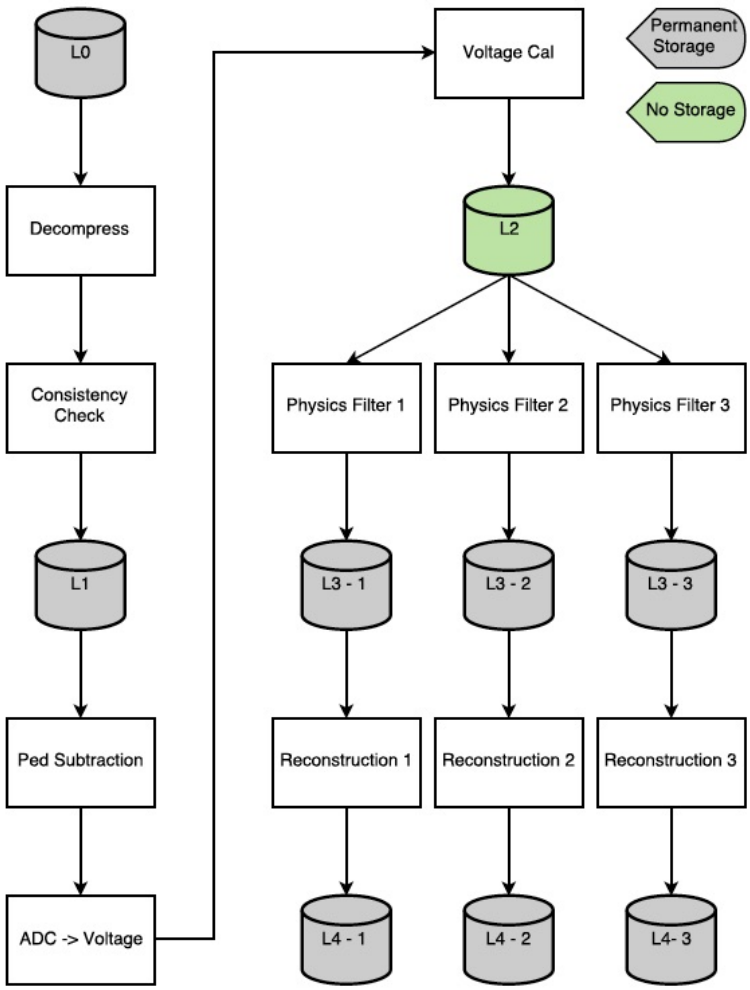


Figure 5.1: Multi-stage data processing plan for ARA. Multiple filters and reconstructions can run in parallel. The coloring scheme denotes the storage plan. Currently stages up to L1 are devised.

the reconstructed direction, the residual or χ^2 values as a result of the reconstruction, et cetera. L4 thus stands for the event summary for each L3 set. The staged processing can in principle be automated once the filters and reconstruction codes are in place. The collaboration is currently exploring options in this direction.

5.3 DAQ Configurations

During the four years of data-taking considered in this thesis, the DAQ configuration went through several changes. We consider five changes that directly affect the trigger and analysis efficiency. The first is the L1 trigger masking, where particular channels can be masked from participating in trigger-forming. A2 D4BH has been observed to be an unstable channel with large RMS variation and unphysical waveform features. Therefore, it is masked in 2014. The second DAQ change considered is the length of the readout window. This directly control the duration of the recorded waveforms. For both station, this is expanded from 400 ns to 520 ns in mid-2014. The third DAQ change is the trigger window, from 110 ns to 170 ns. This is motivated by simulation studies showing the 170 ns trigger window covers all possible timing patterns from any incident angle. This change is also implemented in 2014. The fourth change is whether the trigger delay correction is turned on in trigger-forming of the FPGA or not. As mentioned before, the cable length from the antenna to the DAQ varies among channels, meaning that the data stream from each channel is delayed differently. These delays are tabulated in Tab. 5.1 and Tab. 5.2.

Table 5.1: Cable delay for each antenna. This applies to all strings in both stations.

Antenna	Delay [ns]
BV	81.4
BH	73.2
TV	15.4
TH	7.2

Table 5.2: Extra cable delays for each string in A2 and A3 due to physically longer cables. Both stations have an extra 100 ns delay to one of its strings. The total delay of a channel is the sum of the value from Tab. 5.1 and this table.

String	A2 Delay [ns]	A3 Delay [ns]
D1	0	0
D2	0	100
D3	100	0
D4	0	0

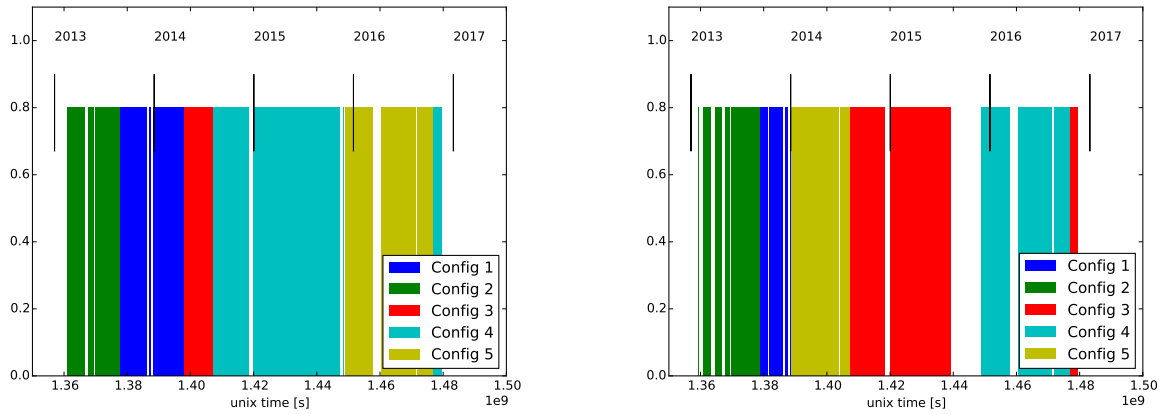
To compensate for this when forming multiplicity triggers, the firmware delays earlier signal so that data stream from all channels are aligned. This correction has been unexpectedly turned off twice for both stations. The last change is the analysis-level channel masking. This is devised specifically for A3 D4, where after run 1946 (Jan. 4 2014) all channels from string D4 stop producing normal waveforms due to digitizer issues. Hence, all D4 channels are masked out in analysis. We summarize the combination of the four changes in Tab. 5.3 and Tab. 5.4.

Table 5.3: DAQ configurations for A2 from 2013 to 2016 and respective live times.

A2 Config	Trig. masking	Readout win.	Trig. win.	Trig. delay corr.	Ana. masking	Live time
1	none	400 ns	110 ns	on	none	179 days
2	none	400 ns	110 ns	off	none	143 days
3	D4BH	400 ns	110 ns	on	none	95 days
4	D4BH	520 ns	170 ns	on	none	439 days
5	D4BH	520 ns	170 ns	off	none	287 days

Table 5.4: DAQ configurations for A3 from 2013 to 2016 and respective live times.

A3 Config	Trig. masking	Readout win.	Trig. win.	Trig. delay corr.	Ana. masking	Live time
1	none	400 ns	110 ns	on	none	80 days
2	none	400 ns	110 ns	off	none	148 days
3	none	520 ns	170 ns	on	none	346 days
4	none	520 ns	170 ns	off	none	260 days
5	none	400 ns	110 ns	on	D4	191 days



(a) (b)

Figure 5.2: Time span of each configuration in (a): A2 and (b): A3. Gaps represent down times. The vertical axis is arbitrary.

Each combination is referred to as the "DAQ configuration", or simply "configuration". In total there are five configurations for A2 and five configurations for A3. Fig. 5.2 shows the span of the configurations. In Fig. 3.3, the trigger level effective areas of the configurations are compared. Notice that A3 configuration 5 is an analysis level choice, and at the trigger level it is identical to configuration 1. At the trigger level, the variation in the effective area is negligible, suggesting that no significant efficiency loss occurred due to sub-optimal configuration parameters. To be conservative, in this analysis we generate simulation files specific to each configuration, with `AraSim` modified to mimic the DAQ differences. For each configuration, the event selection criteria and cuts are defined independent of other configurations, and the sensitivity of the station in each configuration is estimated with the configuration-specific simulation. When calculating the total exposure, the exposures from different configurations can be simply summed.

5.4 Event Preparation

Before the events can be used in analysis, several steps must be taken to prepare the waveforms. For data events, these steps are:

1. Convert ADC counts to voltage and apply pedestal values. This takes the ADC count values in the raw ARA data structure `rawAtriStationEvent` and convert them back to voltage values stored in `UsefulAtriStationEvent` data structure. Plus, "pedestal values" are applied to each sample to account for the residual electric potential in each capacitor in the Switched Capacitor Array when no external potential is applied. The pedestal values are extracted from dedicated pedestal runs that are

launched every five normal runs. In this analysis for each normal run we use the closest pedestal run as the source of pedestal values.

2. Apply cable delays. Just as in the trigger check, the waveforms are shifted according to their respective delays to restore the timing as received by the antenna.

3. Invert waveforms. For A3, the sign of voltage values in channels 0, 4, and 8 are inverted in the signal chain. They are corrected here.

4. Remove first block of samples. The waveforms are read out in IRS2 "blocks" of capacitor values. Each block contains 64 samples. It has been empirically shown that the first block of samples in every channel in A2 and A3 exhibits a systematic negative shift as compared to later blocks. Therefore it is removed in our analysis.

5. Remove corrupted channels. Certain channels within an event are known to produced unusable data due to brokem signal chains. The waveforms of these channels are zero-ed out. They are

   - A2 channel 15 (D4BH)
   - A3 channel 3, 7, 11, 15 (D4) on and after early 2014 run 1946

6. Zero-meaning individual waveforms. This step makes sure the mean of each waveform is zero by shifting waveforms vertically. The shift is calculated on a waveform-by-waveform basis. This prevents the DC component in each waveform from interfering the interferometry.

7. Sample interpolation. The time intervals between digitized samples are generally not uniform even after calibration, producing out-of-band spectral artifacts. This step interpolates Vpol waveform with a constant sample interval of 0.4 ns, and Hpol waveforms with 0.625 ns.

8. Zero-padding. We pad waveforms in configurations where the trigger delay is on to 2048 samples. Otherwise, the padding is 3072 samples. This facilitates the computation of the Fast Fourier Transform (FFT). The location of the recorded signal in the zero-padded waveform depends on its relative timing to the earliest signal.

For simulated events, only the first-block removal, the zero-meaning, and the zero-padding are needed. But in addition, to simulate high-amplitude saturation caused by the digitizer above $\sim$ 600 mV in real data, the simulated waveform values above 600 mV are flattened out at that threshold. This is motivated by the fact that beyond this threshold, the digitizer flips the sign of the samples, and the voltage calibration is no longer valid.

Once the events are prepared, they can be checked against event selection criteria to be discussed next.

5.5 Event Selection

In this section, the event selection criteria are described. Events that pass these criteria are entered into the interferometric reconstruction. The event selection can be categorized intro two levels: run level and event level. Each level addresses data that are inadequate for the diffuse neutrino search at different scales. In the run level selection, runs that are dedicated calibration runs, or have known anomalies across all events, are excluded as a whole. Note that certain runs are added to the list of exclusion only after being inspected with the unblinded sample. They will be described when we introduce analysis cuts. In the event level selection, specific events exhibiting readout issues are discarded. In addition, only regular RF triggers are selected. Finally, an event filter is devised to reduce the amount of data for reconstruction using a simple criterion. We described these selections in more detail next.

5.5.1 Run Level Selection

At the run level, known calibration runs and known anomalous runs are precluded from being analyzed. For A2, they are

- Calibration pulser attenuation sweep. Run 3139 - 3162, 3164 - 3187, run 3289 - 3312, run 7625 - 7686
- L2 trigger masking issue. Run 3464 - 3504
- D1 glitch runs. Run 3, 11, 59, 60, 7100
- Surface pulsing. Run 2884 - 2918, 2938 - 2939, 4872 - 4873, 4876
- ICL rooftop pulsing. Run 3120, 3242, 6527
- Trigger window scan. Run 3578 - 3598
- Calibration pulser noise mode test. Run 4820 - 4825, 4850 - 4854, 4879 - 4936, 5210 - 5277
- A2D6 pulser lift. Run 6513
- Deep pulser run. Run 4785, 4787, 4795 - 4800

For A3, they are

- Miscellaneous early test. Run 22 - 62
- ICL rooftop pulsing. Run 63 - 70, 333 - 341, 2235, 2328, 6041

- Calibration pulser attenuation sweep. Run 346 - 473, 2251 - 2274, 2376 - 2399, 7126 - 7253
- RF event tagging issue + excessive surface noise. Run 479 - 508
- Deep pulser run. Run 3810 - 3811, 3820 - 3822
- Calibration pulser noise mode test. Run 3844 - 3860, 3881 - 3891, 3916 - 3918, 3920 - 3975, 4009 - 4073
- Surface pulsing. Run 3977 - 3978
- Temporary anomalous data-taking. Run 4914 - 4960
- RF event tagging issue. Run 7125
- Excessive software trigger ($\geq 20\%$). 74 runs as shown in link [3]
- Runs where no 10% unblinded data is found. Run 546, 1386, 3429, 3461, 6002-6004, 6006

In both A2 and A3, excluding the above runs results in only $\sim 3\%$ loss in the detector live time.

5.5.2 Event Level Selection

Using `AraRoot` flags, only RF triggers are selected for reconstruction. Software triggers and calibration pulser triggers are identified by their respective flags and skipped. Furthermore, the event level selection deals with various kinds of digitization errors, and for each kind the unblinded sample is used to develop a custom algorithm with which the erroneous event can be rejected. The "glitch" modes common to both stations are

1. First-N event corruption. After run 0003 (Nov. 25 2015) for A2, the first four events in each run are always unusable due to waveforms exhibiting anomalous multi-peak features, presumably due to digitization issues. In A3, after run 6002 (Nov. 29 2015) the same pathology can be found in the first five events of each run. To reject this class of events, we simply skip the first four (five) events in the runs for A2 (A3) after said time. The live time loss is negligible.

2. Mistagged software trigger. These are software trigger events that are not properly tagged by `AraRoot` as software triggers. To tag a software trigger or a calibration pulser event, the software checks the trigger fraction-of-a-second (`rawAtriStationEvent::timestamp`) which, for a periodic 1 Hz trigger, should assume similar values across the run. If this value falls within a predefined range, the

[3] `http://ara.icecube.wisc.edu/wiki/images/1/18/A23_analysis_A3_software_dominated_run_list.txt`

event is tagged. However, due to clock drift issues, occasionally a software trigger or a calibration pulser event could be mistagged. To detect mistagged software triggers, we take advantage of the fact that software triggered events are configured to have shorter waveforms (at most 15 blocks) compared to regular RF data (at least 20 blocks). The number of samples is checked to make sure only events with long enough waveforms are accepted. This does not apply to mistagged calibration pulser events, though. In section 5.6.1.3, a reconstruction-based method is used to rejects those events.

3. Cut-wave events. A small set of events demonstrate a glitch in which the read-out waveform consists of only less than five samples. Again, the number of samples in each waveform is checked to detect such events.

The glitch modes specific to A2 are

1. Repetitive sample time. This refers to events where at least one waveform contains consecutive digitized samples whose timings are the same when read out from raw data. This is checked simply comparing the time interval between consecutive samples. This error is extremely rare — only two such events are found in the burnsample.

2. "Offset-block" events. A2 waveforms are observed to sometimes contain blocks of samples that are apparently vertically offset from the baseline values. One example is given in Fig. 5.3. A data-motivated algorithm is implemented to reject such events. First, a rolling mean over 64 samples across each waveform is calculated. The maximal value of the rolling mean is compared with a threshold (Vpol: $\pm$ 20 mV, Hpol: $\pm$ 12 mV) to determine if and when the channel has an offset block. It is further noticed that offset blocks often appear in channels of the same string at around the same position. Hence, we define a offset-block string as having offset blocks in both Vpols and at least one Hpol, and their offset position is within a predefine range of 10 ns. An event is vetoed if more than one offset block string is present.

3. Block-gap events. By default IRS2 stores sampled values in consecutive blocks. However, A2 events have been observed to be comprised of non-consecutive blocks with a gap in between block numbers, resulting in waveforms that are difficult to interpret. We make sure only events with consecutive block numbers are accepted.

After unblinding, we would find out that with the 100% sample no events with repetitive sample time or block gaps are found. This suggests the software process of extracting the 10% unblinded sample from the 100% set could be a root cause for such error modes.

The glitch mode specific to A3 is

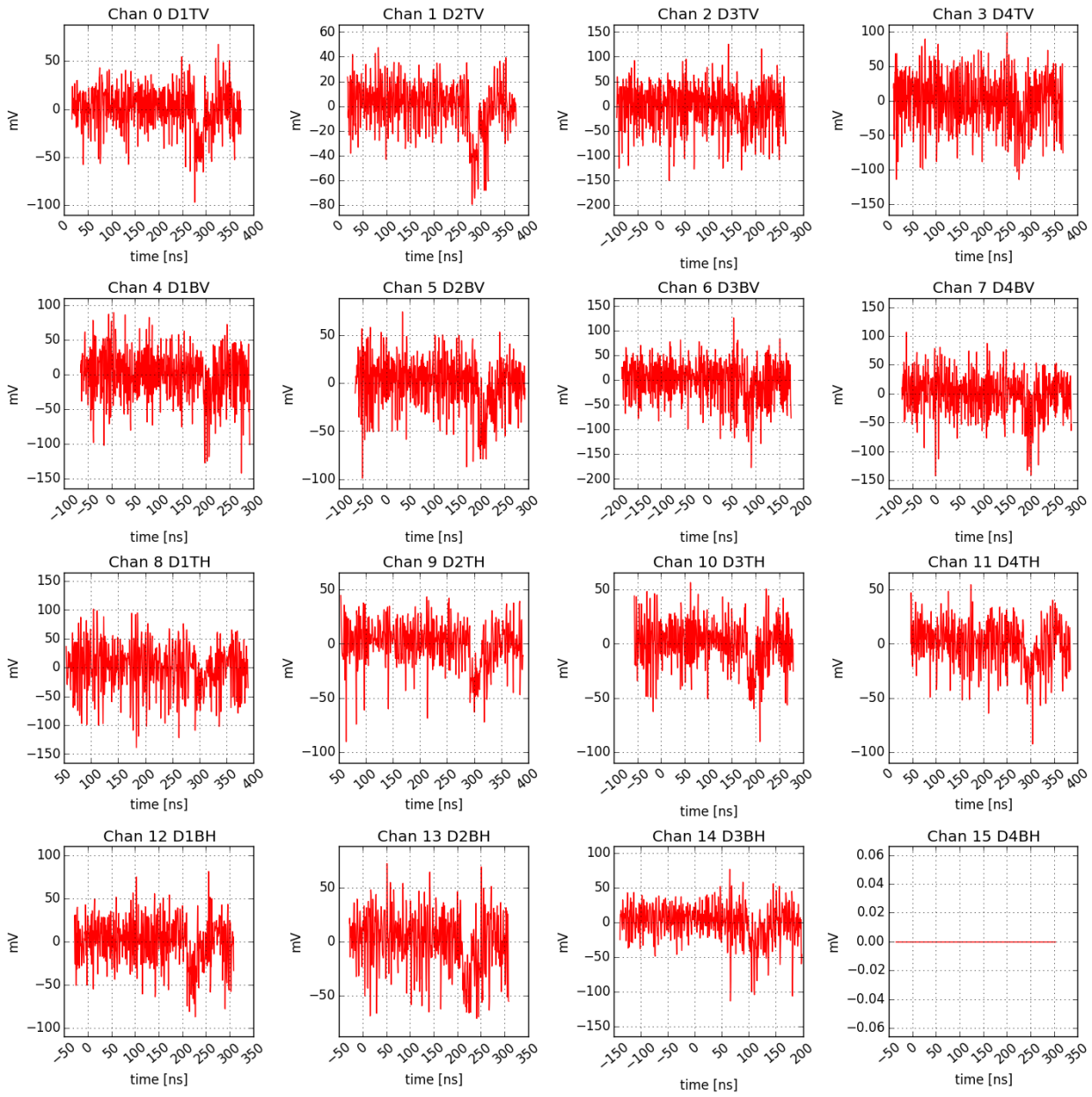


Figure 5.3: Example of an offset-block event in A2. Run 2823 event 95294. The offset block can be observed to occur in the same position for channels on the same string.

1. "Cliff" events. In these events the baseline of the waveforms shifts noticeably in an event, causing a discontinuity at the block boundary. See Fig. 5.4 for an example. To rejects such events, for each waveforms the medians of the voltage values in the first block and in the last block are compared. If all four channels in any single string demonstrates median differences larger than threshold values, the event is considered erroneous and rejected. The threshold values for strings 1, 2, 3 are 100 mV, 45 mV, and 100 mV. String 4 is zero-ed out in the event preparation stage.

Overall these selections have negligible impact on the signal efficiency, as we will see in section 5.7. RF events that pass these selections are entered in the event filter described next.

5.5.3 Event Filter

Since the interferometric reconstruction is the most complex and time-consuming part in the analysis pipeline, and we expect most ($\geq 99\%$) triggers to be upward thermal fluctuation, we try to further reduce the amount of reconstruction needed by cutting thermal events. A simple event n-channel event filter is placed for this purpose. The goal of the filter is to reject $\sim 99\%$ of thermal noise without using sophisticated algorithms. We filter events based on the event's signal-to-noise ratio (SNR). For each waveform, we generate a power envelope E as

$$E_j = \sqrt{\frac{1}{n}\sum_{i=j}^{j+25\text{ns}} V_i^2} \tag{5.1}$$

, where E_j is the envelope value at sample j, n is the number of samples in the 25-ns window, and V_i is the waveform voltage value at sample i. The SNR is then defined as

$$\text{SNR} \equiv E_{j,\max}/\sigma_{E_{j,\text{noise}}}. \tag{5.2}$$

In the above equation, $E_{j,\max}$ is the maximum of E_j, and $\sigma_{E_{j,\text{noise}}}$ is the RMS value of the half of the envelope that does not contain the maximum. Given the per-waveform SNR, the "event SNR" is taken to be the third largest SNR of all Vpols. This brings us to an important choice in this analysis - only events with sufficient Vpol signals are filtered and reconstructed. This choice is motivated by the fact that Vpol antennas are more sensitive and better understood than Hpols, especially at lower frequencies. Simulation studies have shown that the majority ($\sim 70\%$) of simulated signal events contain Vpol triggers.

With this event SNR definition, the 1% RF pass rate threshold for each configuration is obtain from the SNR distribution of the 10% unblinded sample. Fig. 5.5 shows the fractional cumulative distribution for A2 configurations, and Fig. 5.6 shows its A3 counterpart. In Tab. 5.5, the SNR thresholds are compiled for

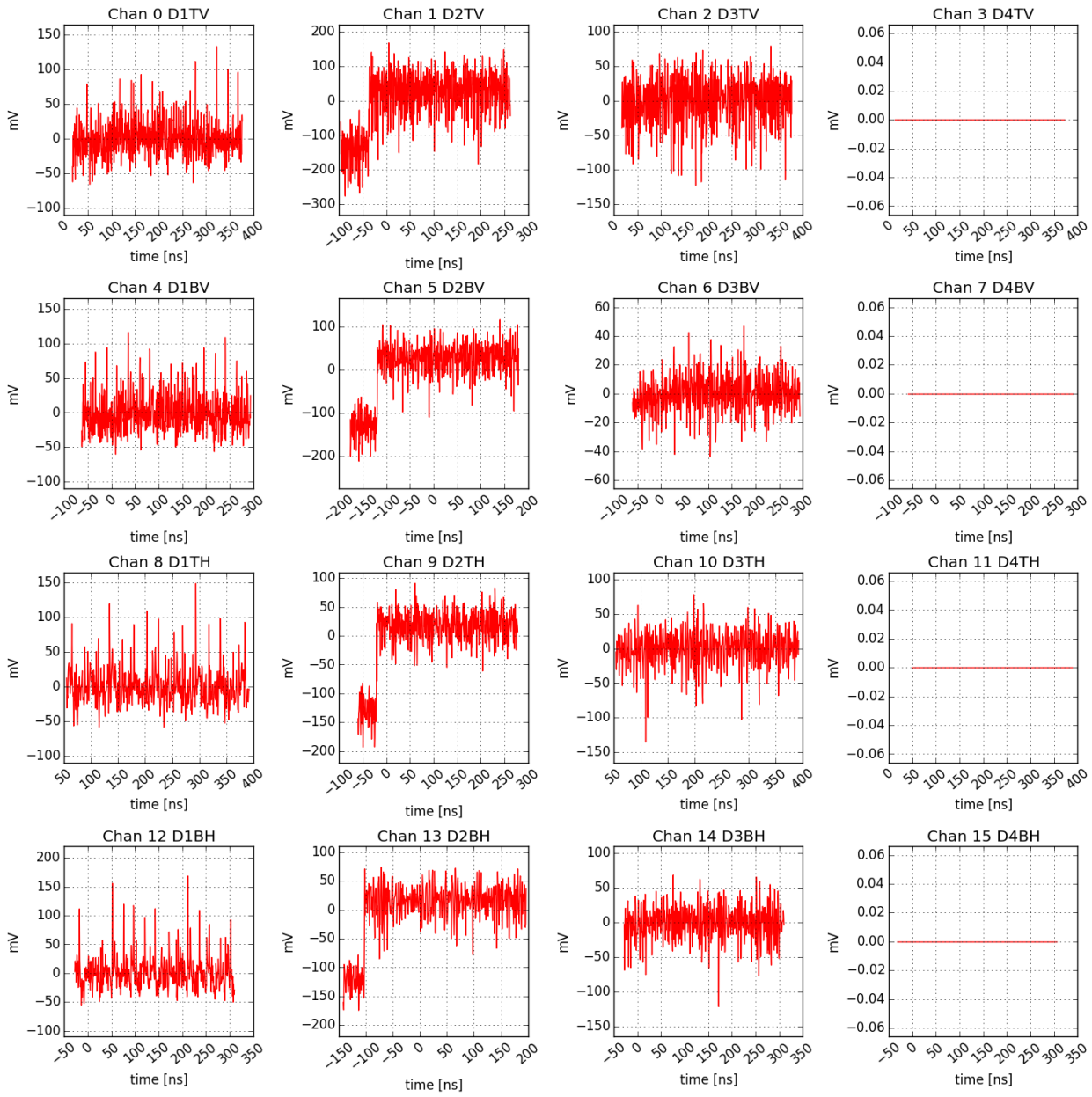


Figure 5.4: Example of a cliff event in A3. Run 1795 event 52890. The cliff feature can be observed in all channels in D2.

all configurations in both stations. The impact of the threshold on the signal efficiency will be discussed in section 5.7. At 10^{18} eV, the filter is approximately 70% efficient.

Table 5.5: N-channel filter event SNR thresholds for each DAQ configuration in A2 and A3, each giving a 1% pass rate of RF triggers.

Config	A2	A3
1	5.3221	4.9701
2	5.7219	5.5527
3	5.2975	4.9541
4	5.6626	4.9645
5	5.6150	4.9701

5.6 Data Analysis

All surviving events after the event filter are reconstructed with the `Radiospline` interferometric reconstruction prescribed in section 4.2. The result of the reconstruction, specifically the coherence value and the reconstructed signal direction, are combined with other waveform features to form multi-variate cuts. The cuts are developed with the 10% unblinded sample with one exception, and will be described in this section. Given the very low expected signal in data ($O(0.1)$ events per year), the goal of the analysis is to reduce the background level to one order of magnitude less than that of the signal, specifically ~ 0.01 background events per 228 days for A2 and 220 days for A3 as reached in the previous ARA analysis [94]. This is consistent with past ARA analyses [94]. The cuts will be tuned to meet this requirement, while optimizing for signal efficiency at 10^{18} eV when applicable. The signal efficiency impact of the cuts will be shown in section 5.7.

The analysis follows the standard procedure of a blind analysis. Using the unblinded sample, we seek to understand the various sorts of backgrounds that the data contain. Given the low expected signal rate, we assume the 10% sample to consist of only background, and the cuts are developed with the goal of completely rejecting the 10% sample as neutrino candidates. The cut development is further divided in two stages, where the first stage defines initial cut values by modeling the RF data. However, the background is observed to be dominated by distribution tails that are often hard to described with the models. Consequently in the second stage we specifically model the tail events after applying initial cuts defined in the first stage. This allows us to properly estimate the background and tune final cut values accordingly.

Once the 10% unblinded sample is explored and final cut values set, the blinded data are opened with the final cut values, and the signal efficiency is estimated by applying the same set of cuts to signal simulation.

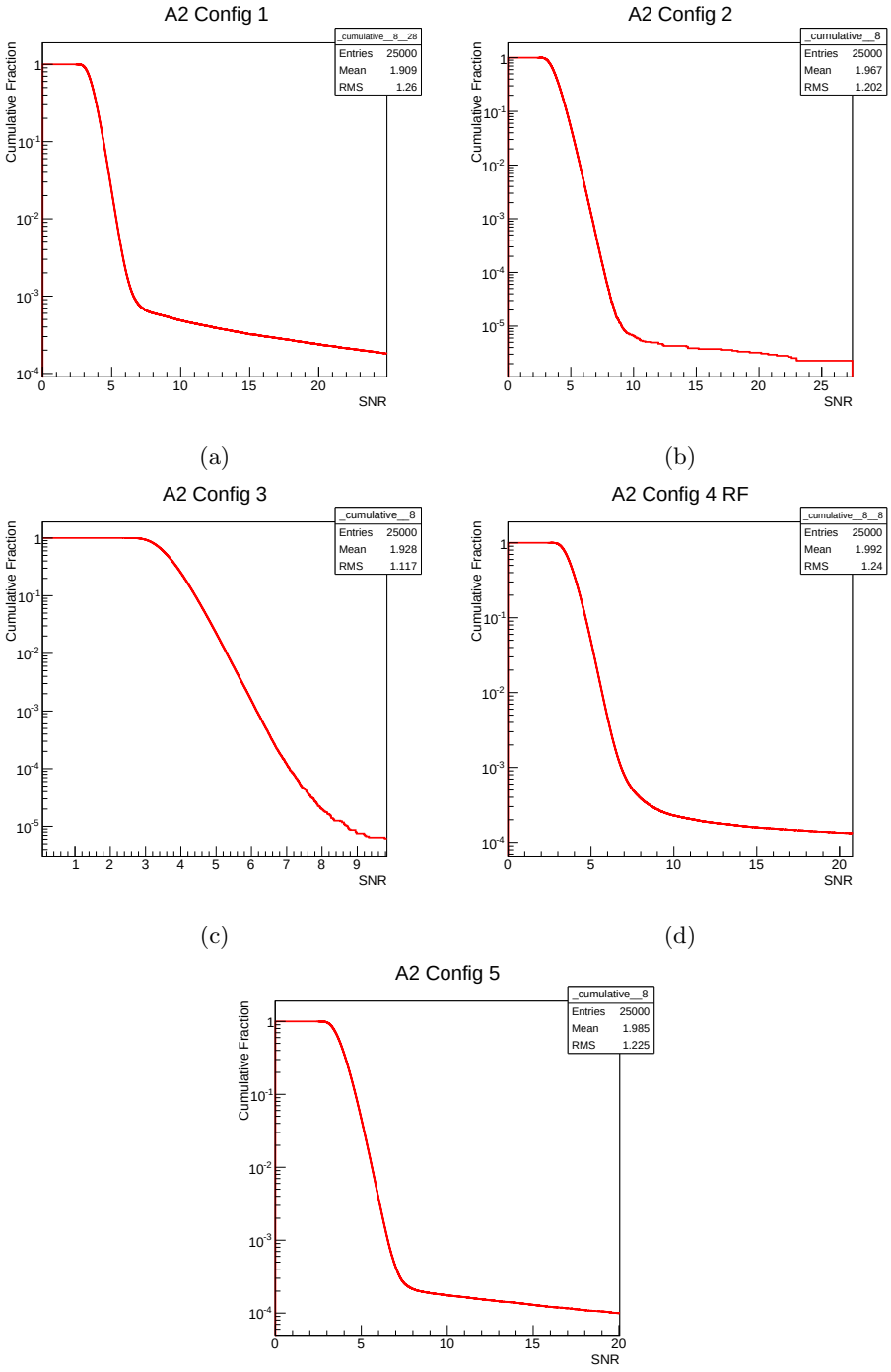


Figure 5.5: Fractional cumulative distribution of the event SNR of RF event in A2 10% sample. In each configuration, the 1% pass rate point is taken as the threshold in the n-channel event filter. The high-SNR tail typically consists of surface noise.

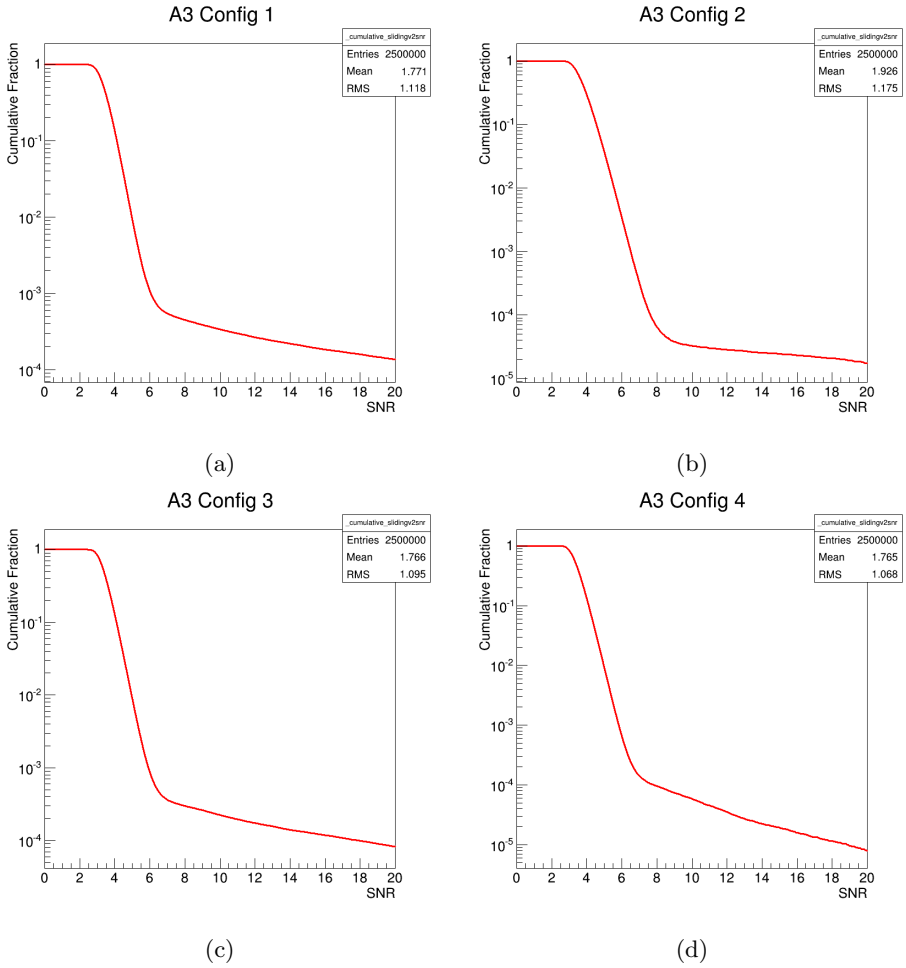


Figure 5.6: Fractional cumulative distribution of the event SNR of RF event in A3 10% sample. In each configuration, the 1% pass rate point is taken as the threshold in the n-channel event filter. The high-SNR tail typically consists of surface noise. Configuration 5 takes the threshold from configuration 1 since at trigger level they are the same.

After all cuts, in the case of non-zero surviving events in the 100% data, we study the surviving events closely to determine if they can be reported as neutrino candidates. In the case of null observation, we place an upper limit to the diffuse neutrino flux in the UHE regime.

5.6.1 Initial Analysis — Stage 1

Exploration of the A2 and A3 unblinded sample suggests five common analysis cuts. The coherence value from the reconstruction is used to rejects thermal events. The low frequency cut and the SNR cut are mainly devised to cut CW contamination. The calibration pulser geometric cut masks out a portion of the station-centric sky to prevent mistagged calibraion pulser events. The surface cut vetoes events originating for above the ice surface. There are, additionally, A2- or A3- specific cuts that deal with backgrounds unique to each stations. All cuts in the first stage of analysis are introduced in this section.

5.6.1.1 Low Frequency Cut

We observe in data two main categories of CW contamination. The first is the NOAA weather balloon signal near 400 MHz, occurring twice everyday. It will be dealt with in section 5.6.1.6. The second is CW interference near or below the lower end of the ARA frequency response, namely 150 MHz. The sources of such interference are unclear. Visually the CW can be identified by prominent spectral peaks in the frequency domain. In the time domain, the second category of CW events often manifests as powerful interference with amplitude up to several hundred millivolts. To veto such events, in the frequency domain we require that the maximum number of channels in either polarization with a maximum spectral peak below 170 MHz must be below three for an event to pass. This cut is empirically developed, and represents negligible signal efficiency loss below 10^{19} eV, and a $\sim 10\%$ loss at 10^{21} eV from off-cone signals or bright events where spectral leakage arises due to the artificial flattening out of the saturated waveforms.

5.6.1.2 Thermal Cut

At a station trigger rate of 5 Hz, each station expect $\geq 1.5 \times 10^8$ RF trigger every year, among which the majority are thermal events. Hence, to reach a target background level of ~ 0.01 event per 228 or 220 days, our cuts must bring the thermal background down by a factor of 10^{10}. Given that the n-channel event filter already filter event with a 1% pass rate, here we set cut values corresponding to a pass rate of 10^{-8}. The variable used to set the cut is the coherence value from the interferometric reconstruction. Signals in an event should contribute constructively to the coherence value when the pulses are lined up correctly. On the other hand, cross-correlating uncorrelated thermal noise tends to average down the coherence value . Since the coherence value of an event is defined as the maximum of all summed cross correlation values, it is expected to follow the generalized extreme value (GEV) distribution, which describes the probability

distribution of the maxima of ensembles of randomly distributed, independent values. The GEV probability distribution function (PDF) is chracterized by three parameters μ: location, σ: scale, ξ: shape as

$$\frac{1}{\sigma}\ t(x)^{\xi+1} \cdot e^{-t(x)}$$
$$t(x) = \begin{cases} (1+\xi(\frac{x-\mu}{\sigma}))^{-1/\xi} & if\ \xi \neq 0 \\ e^{-(x-\mu)/\sigma} & if\ \xi = 0 \end{cases} \tag{5.3}$$

We model the coherence distribution of the RF events after the event selection with the GEV distribution, and use the fitted parameters to extrapolate to obtain a cut value that allows only 10^{-8} pass rate. This cut value is the initial thermal cut. In practice, this fit is performed in two part of the sky map independently. This is because the reconstruction of RF events reveal systematic features in the reconstructed sky map throughout the data taking period, as illustrated in Fig. 5.7, where the accumulation of hot spots in the horizontal "band" around -43.5° in elevation angle points to a constant bias from the correlation of vertical baselines (TV-BV pairs). If the cross correlation function in a vertical baseline is plotted for a thermal event, for example in Fig. 5.8, one can observe high correlation around a lag of 75 ns. This lag corresponds to the cable delay difference between a top Vpol and a bottom Vpol, and indicates a low-level crosstalk issue in the signal chain for channels in the same physical string. The source of the crosstalk is currently being investigated. To treat this systematic bias, events that reconstruct to inside or outside of the $-43.5 \pm 2.5°$ zenith band are fitted separately, producing two fits and two cut values for each configuration. They are referred to as "in-band" and "out-of-band" respectively. Another subtly is that each event is reconstructed twice with the arrival time of the first and the second rays as provided by `Radiospline`, given that up to $\sim 30\%$ of of events are triggered only by the second ray in simulation. Out of the two reconstruction results, we take the direction and coherence from the one with a larger coherence value.

In Fig. 5.9, Fig. 5.10, Fig. 5.11 and Fig. 5.12, the data with the fitted PDFs are shown for out-of-band and in-band respectively. The data tails that are inconsistent with the fitted models represent non-thermal backgrounds that will need to be rejected with cuts to follow. The initial cut values for A2 and A3 are summarized in Tab. 5.6 and Tab. 5.7.

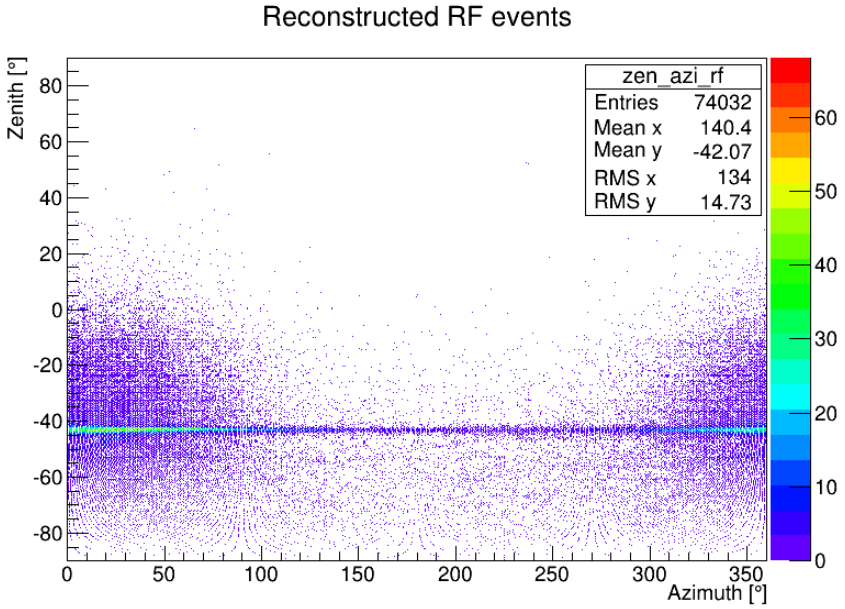


Figure 5.7: The reconstructed zenith and azimuth for the 15% RF sample from A3 configuration 2. The band near -43.5° is clearly visible. The non-uniform distribution of events in the azimuth is explained by the extra 100 ns of delays in string 2. Due to this extra delay, the recorded waveforms of all channels on string 2 systematically lead other channels, producing a bias of the interferometry towards the direction of the leading string in the absence of signal which would overpower the bias. The same phenomenon can be seen in A2, with the extra delay in string 3, as well as in simulation where ad hoc delays are added to reproduce the same bias.

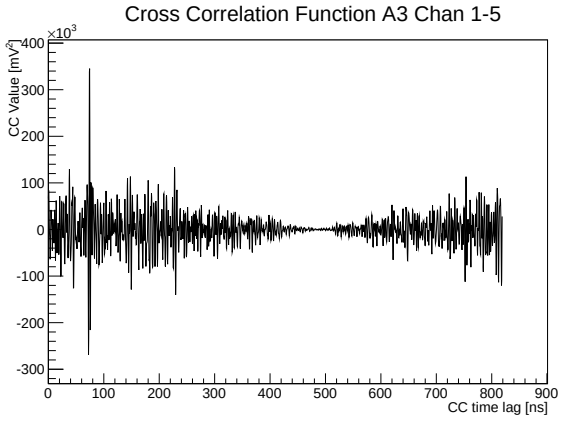


Figure 5.8: Cross correlation function of D2TV and D2BV in an A3 run3606 thermal event showing crosstalk features around a lag of 75 ns.

Table 5.6: Initial thermal cut values for A2 as extrapolated from the GEV fits of the coherence distribution to a thermal pass rate of 10^{-8}. The in-band cut values are higher than the out-of-band values as expected.

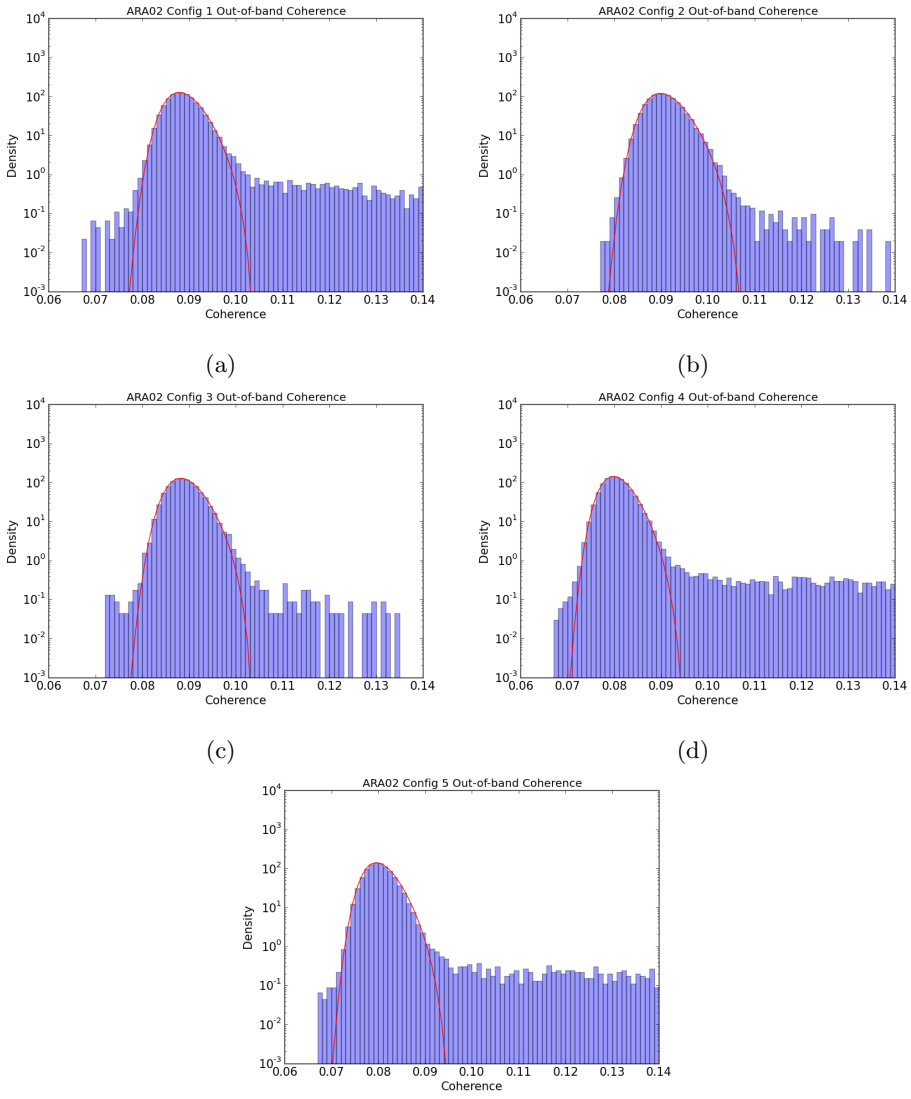
Config	In-band	Out-of-band
1	0.1080	0.1037
2	0.1093	0.1073
3	0.1084	0.1036
4	0.0990	0.0947
5	0.0971	0.0951

Table 5.7: Initial thermal cut values for A3 as extrapolated from the GEV fits of the coherence distribution to a thermal pass rate of 10^{-8}. The in-band cut values are higher than the out-of-band values as expected.

Config	In-band	Out-of-band
1	0.1128	0.1058
2	0.1103	0.1054
3	0.1295	0.1294
4	0.1351	0.1341
5	0.1381	0.1180

5.6.1.3 Calibration pulser Geometric Cut

Calibration pulser events can be mistagged by `AraRoot` as regular RF triggers. Given that we are able to reconstruct the direction of the calibration pulsers with roughly 1 degree precision, we mask out events reconstructed from the known direction of the calibration pulsers using the 40m direct-ray sky map. Using the first 100 tagged calibration pulser event in each run, we build histograms of the reconstructed zenith and azimuth of each calibration pulser as shown in Fig. 5.13, and apply 1° Gaussian smoothing before fitting the distribution tail with the normal distribution. The fit allows us to set the cut values giving only 10^{-10} leaked calibration pulser event in the whole dataset, assuming that all calibration pulser events are mistagged. This is a very aggressive assumption, and thus the cut is defined conservatively. This procedure is illustrated in Fig. 5.14, 5.15, 5.16 for the three cut regions in A2 and Fig. 5.17, 5.18 for the two cut regions in A3. Eventually, for each calibration pulser we defined a "box" region in the sky map and veto all events from within the box. Generally the boxes are $\sim 10^\circ \times 10^\circ$ in A2 and $8^\circ \times 8^\circ$ in A3. The cuts values are shown in Tab. 5.8.



(a) (b) (c) (d) (e)

Figure 5.9: Normalized out-of-band coherence value distribution of RF event in A2 10% sample. The red curve is the best-fit GEV PDF that is used to obtain the initial thermal cut value.

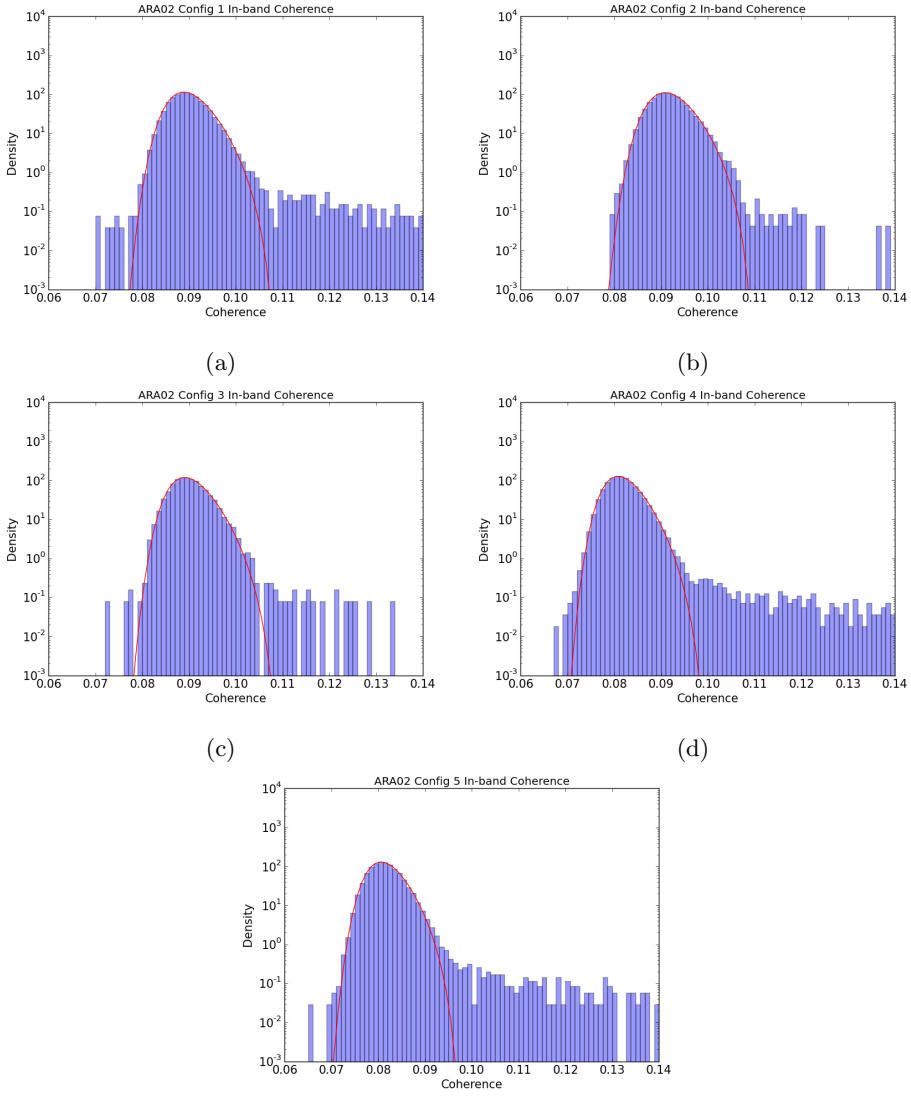
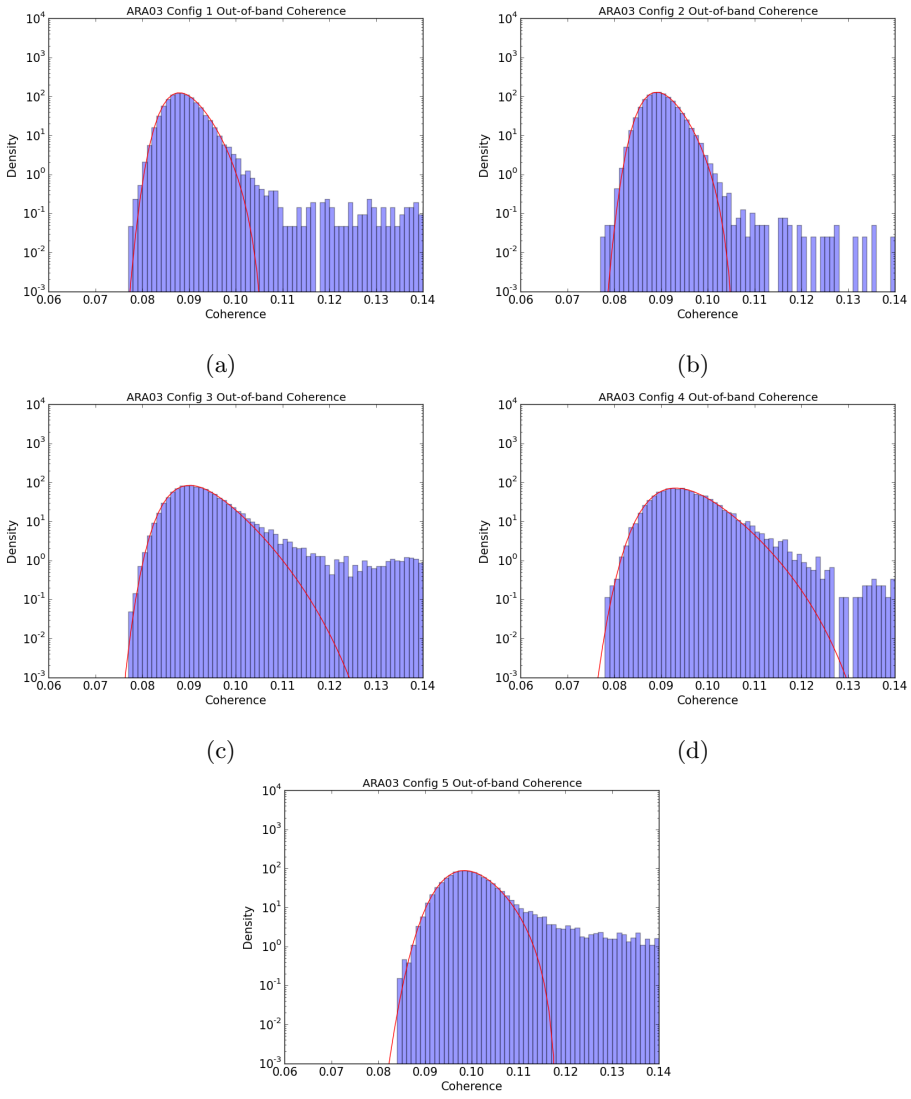


Figure 5.10: Normalized in-band coherence value distribution of RF event in A2 10% sample. The red curve is the best-fit GEV PDF that is used to obtain the initial thermal cut value.



(a) (b) (c) (d) (e)

Figure 5.11: Normalized out-of-band coherence value distribution of RF event in A3 10% sample. The red curve is the best-fit GEV PDF that is used to obtain the initial thermal cut value.

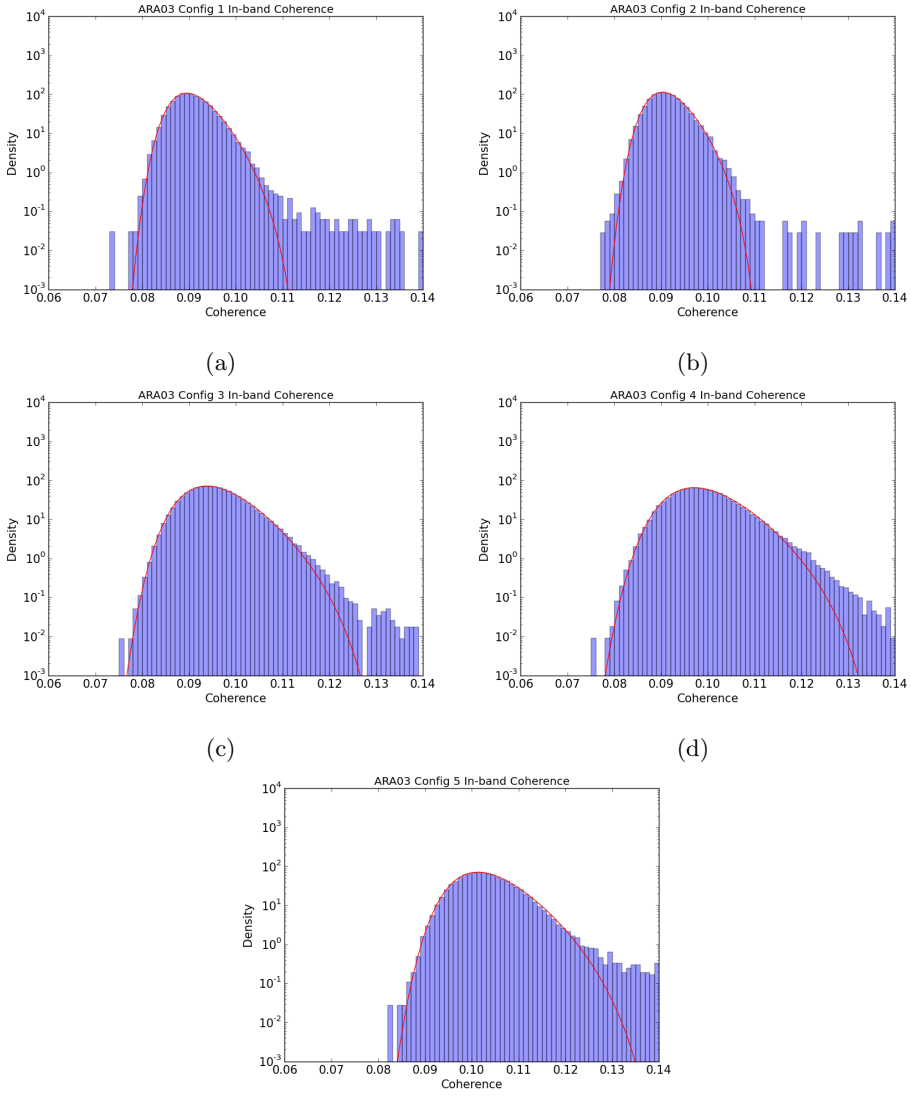
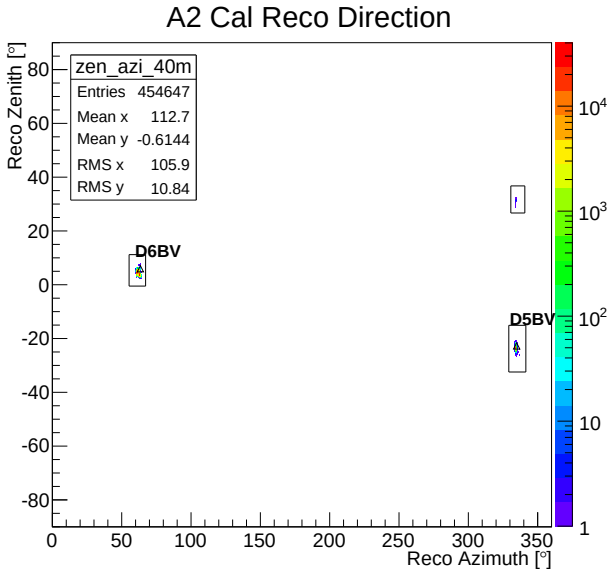
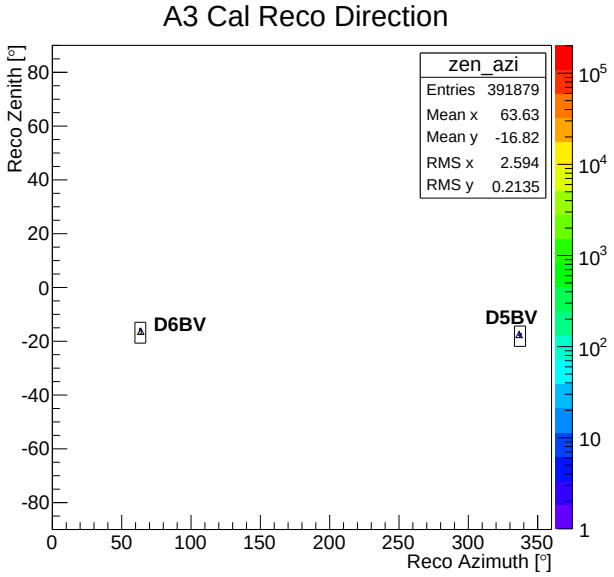


Figure 5.12: Normalized in-band coherence value distribution of RF event in A3 10% sample. The red curve is the best-fit GEV PDF that is used to obtain the initial thermal cut value.



(a)



(b)

Figure 5.13: The reconstructed direction of the first 100 tagged calibration pulser events in each run on the 40-m sky map. The clustering of events allows us to place geometric cuts to mask out the potential pulser leakage events. The procedure of defining the cut regions is described in text. (a): A2. Besides the two "boxes" corresponding to the two Vpol pulsers, due to weak signal and missing pulses in configuration 5, an additional box needs to be placed for the same configuration that corresponds to the mirror of D5BV. (b): A3.

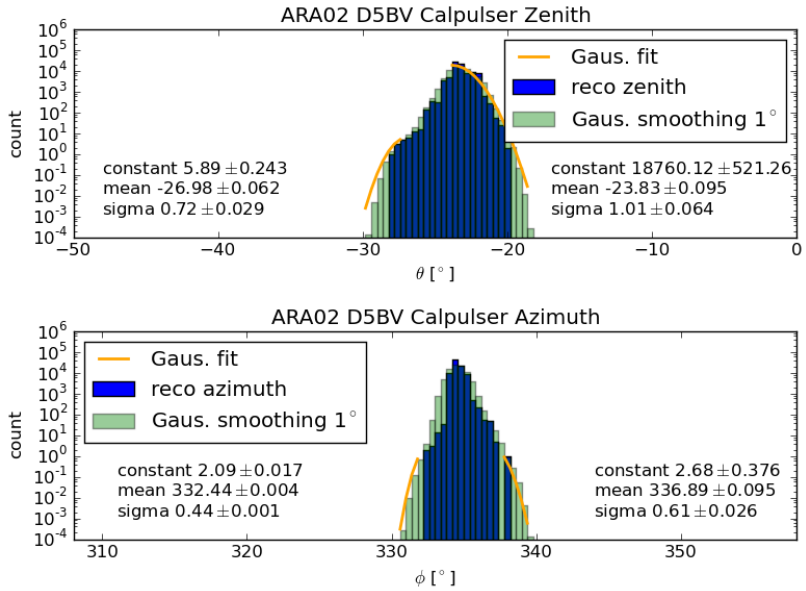


Figure 5.14: Gaussian fits to the reconstructed zenith (top) and azimuth (bottom) of A2 D5BV. Blue: reconstructed distribution. Green: 1° Gaussian smoothing applied. Yellow: Gaussian fit to the tails of smoothed histograms.

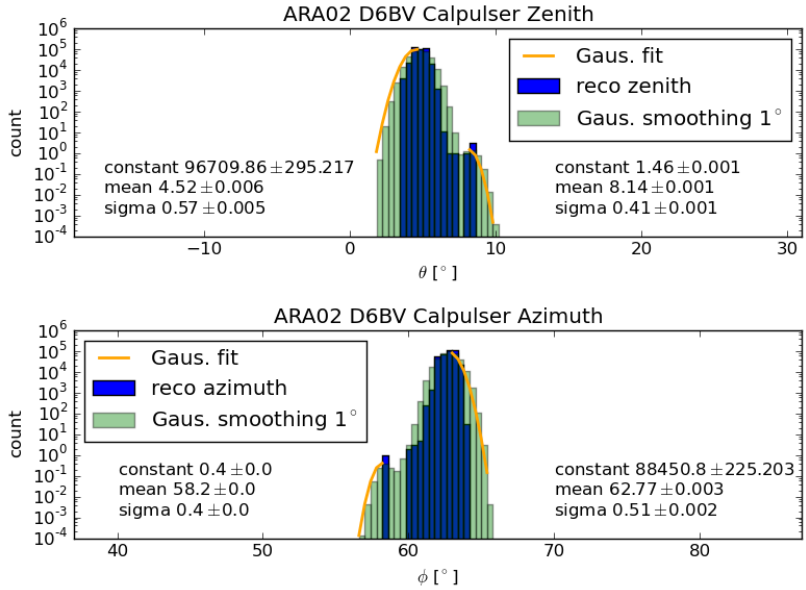


Figure 5.15: Gaussian fits to the reconstructed zenith (top) and azimuth (bottom) of A2 D6BV. Blue: reconstructed distribution. Green: 1° Gaussian smoothing applied. Yellow: Gaussian fit to the tails of smoothed histograms.

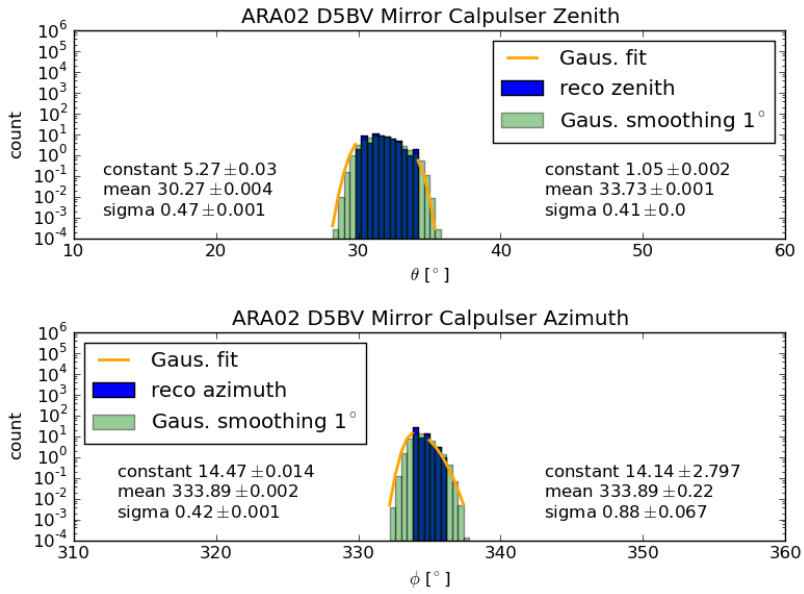


Figure 5.16: Gaussian fits to the reconstructed zenith (top) and azimuth (bottom) of the mirror hot spot from A2 D5BV. Blue: reconstructed distribution. Green: 1° Gaussian smoothing applied. Yellow: Gaussian fit to the tails of smoothed histograms. The geometric cut obtained here will only be used to check configuration 5 events.

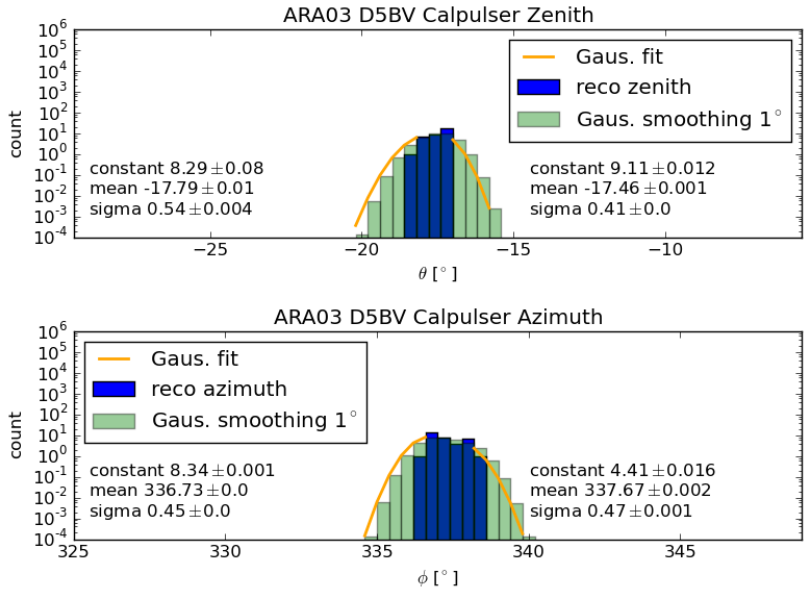


Figure 5.17: Gaussian fits to the reconstructed zenith (top) and azimuth (bottom) of A3 D5BV. Blue: reconstructed distribution. Green: 1° Gaussian smoothing applied. Yellow: Gaussian fit to the tails of smoothed histograms.

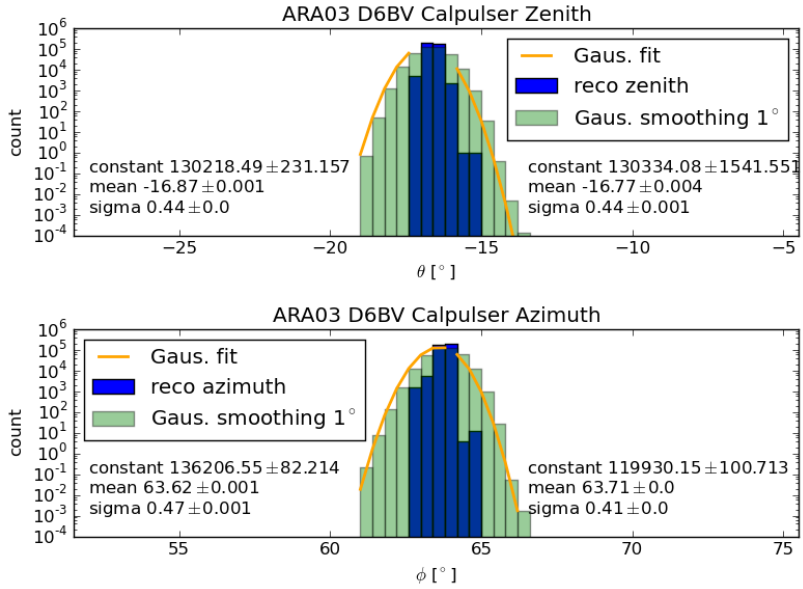


Figure 5.18: Gaussian fits to the reconstructed zenith (top) and azimuth (bottom) of A3 D6BV. Blue: reconstructed distribution. Green: 1° Gaussian smoothing applied. Yellow: Gaussian fit to the tails of smoothed histograms.

Table 5.8: Calibration pulser geometric cut values. Each geometric cut is an angular box and the four sides of each box are listed here.

	$\theta_{\min}$ [°]	$\theta_{\max}$ [°]	$\phi_{\min}$ [°]	$\phi_{\max}$ [°]	Note
A2 D5BV	-32.45	-15.09	329.18	341.48	
A2 D6BV	-0.50	11.20	55.32	67.23	
A2 D5BV Mirror	26.68	36.75	330.64	340.74	Only config 5
A3 D5BV	-21.91	-14.34	333.3	341.3	
A3 D6BV	-20.69	-12.95	59.53	67.34	

Note that in A2, in configurations where cable delays were missing, the later pulses are sometime truncated in the event. Additionally, pulser amplitudes from D5BV are observed to be generally much weaker than the other calibration pulser. Combining these two observations, events from D5BV in configuration 5 are sometime not reconstructed properly in the nominal direction. Fortunately, they still have a common timing pattern and would be concentrated in a "mirror" location that can be masked out with the same procedure. This is shown in Fig. 5.13. Another measure taken to address such events is the implementation of the "iterative reconstruction". While the default reconstruction use all functioning Vpol channels, in the iterative reconstruction we start from the three channels with the largest SNR values, and progress each round of reconstruction with the next channel in the SNR rank. Hence, each event will effectively be reconstructed six times. It was observed that for calibration pulser events that fail to properly point to the boxes defined above using the default 8-channel reconstruction, at least one iteration would successfully place the event back in the boxes. Thus, the same criterion is applied to RF triggers to cut leaking calibration pulser events — that if at least one iteration reconstructs to any of the calibration pulser boxes, the event is rejected.

The calibration pulser cuts are final and require no second-stage tuning, since the cuts are conservatively defined and the ensuing backgrounds are much lower than required ($O(0.01)$/year).

5.6.1.4 Surface Cut

The geometric cut at the surface is used primarily to reject anthropogenic noise, as well as potential down-going physics signals such as cosmic rays. We make the cut on events from above the surface because we expect neutrino events to predominantly yield up-coming signals. The cut on events from above the surface proceeds similarly to the calibration pulser geometric cut. In stage 1, we fit the distribution of events that pass the thermal cut near the transition between the air-ice boundary from an elevation angle of 38-46 degrees with a Gaussian distribution, and place an angular cut at a 3σ level. The quasi-planewave reconstruction proved to be effective when resolving the ICL rooftop pulser singal. Consequently we use the

quasi-planewave reconstructed zenith angle as the parameter in the Gaussian fit. Fig. 5.19 shows the fits and cut values for A2 and A3. For A2, the initial surface cut value is 35.96° in elevation, while for A3 it is 33.86°.

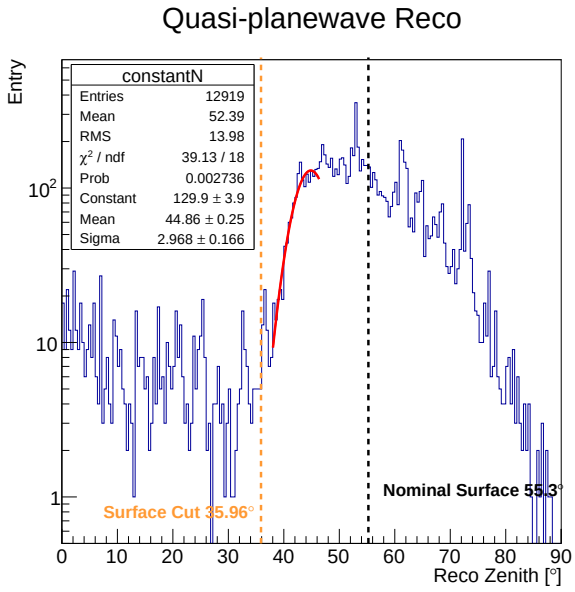
Certain events in A2 fail to point to above the surface cut as defined above, even though the timing of the apparent pulses indicate a down-going signal. This could be due to the source of signal being near the station, as opposed to "far" sources such as the rooftop pulser. This entails that the wavefront deviates from a planewave, and thus thus the inability for the quasi-planewave reconstruction to successfully model it. Here we once again use the iterative reconstruction, taking the iterations with at least five channels up to all channels, and perform a majority vote in 3-degree bins. This is shown in Fig. 5.20, and the same fitting procedure is performed in 34-43.2 degrees to give a cut value of 31.63°. An A2 event has to pass both the surface cut defined using the quasi-planewave reconstructed angle and the "majority voted" surface cut using the iterative reconstruction to avoid being rejected.

5.6.1.5 Surface Noisy Runs

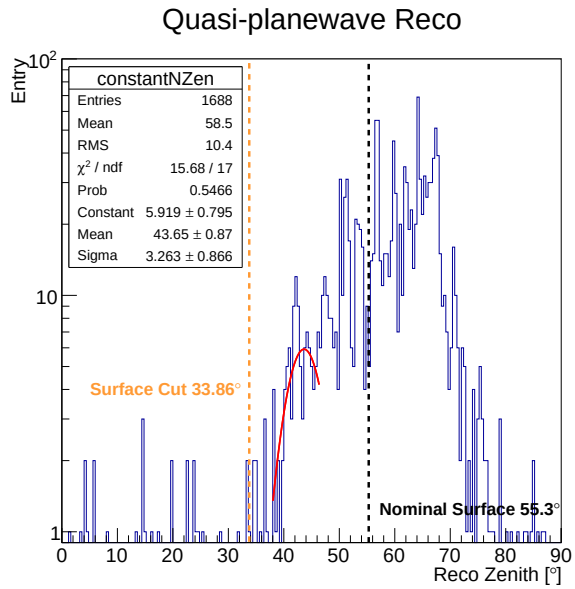
The fact that the surface cuts need to extend below the critical angle attest to the difficulty in modeling the transmission in the air-ice transition region, which directly translates to a difficulty in precisely rejecting backgrounds from this region. However, anthropogenic noise from the surface often cluster in time, and we can exclude runs containing such clusters to further suppress surface noise. At the stage 1 level, this is done by examining the number of non-thermal events rejected by the surface cut described in the previous section for each run, and cut runs where the number of surface events is larger than one in the 10% sample. This result of this counting exercise is shown in Fig. 5.21. Runs that are cut are referred to as the "surface-noisy" runs. This amounts to 46 runs for A2 and 19 runs for A3. In stage 2 the surface events in the 100% sample will be used to redefine the criterion for surface-noisy runs, with the new list of runs fixed before the final unblinding.

5.6.1.6 SNR Cut

To eliminate the contamination of CW emission, we already placed the low-frequency cut. To reject CW contamination in higher frequencies, we observe that such events, while producing high coherence values due to their CW nature, do not produce high SNR values. Therefore, we reject these events with an SNR cut. The cut is defined using a Weibull PDF fit to the event SNR distribution of remaining events after all the above cuts. The initial cut value is set to produce 0.1 backgrounds per year. Note that while this background level appears high, the cut is not final and will be tuned for lower background values in stage 2. The data-fitting is shown in Fig. 5.22 and Fig. 5.23. The initial cut values are shown in Tab. 5.9 and



(a)



(b)

Figure 5.19: (a): A2. (b): A3. The quasi-planewave reconstructed zenith angles of non-thermal events are plotted. The Gaussian fit is models the tail of the distribution in the air-ice transition region. The fitted parameters are used to define the surface cut as the -3σ point in the Gaussian function.

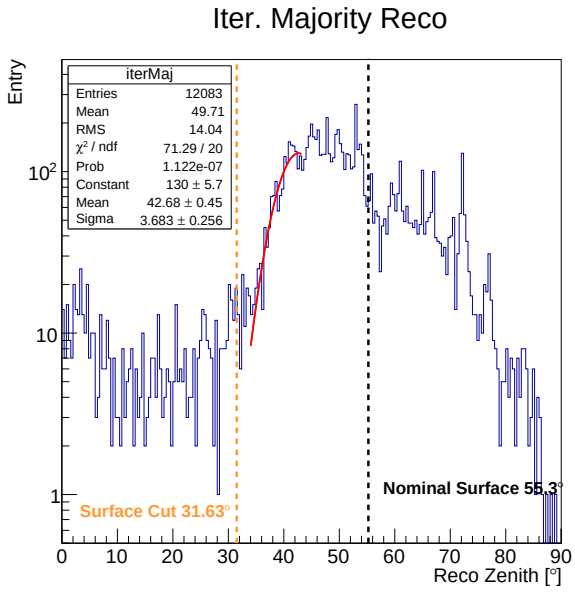


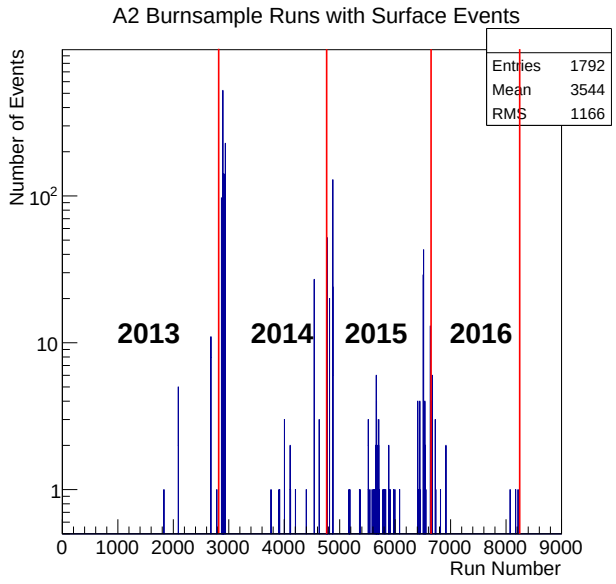
Figure 5.20: Majority-voted zenith angle in 3-degree bins from iterative reconstruction of A2 10% RF events. The procedure to define the cut value is the same as that of the quasi-planewave surface cut.

Table 5.9: Initial SNR cut values for A2 and A3 as extrapolated from the Weibull fits of the SNR distribution to a background level of 0.1 event per year.

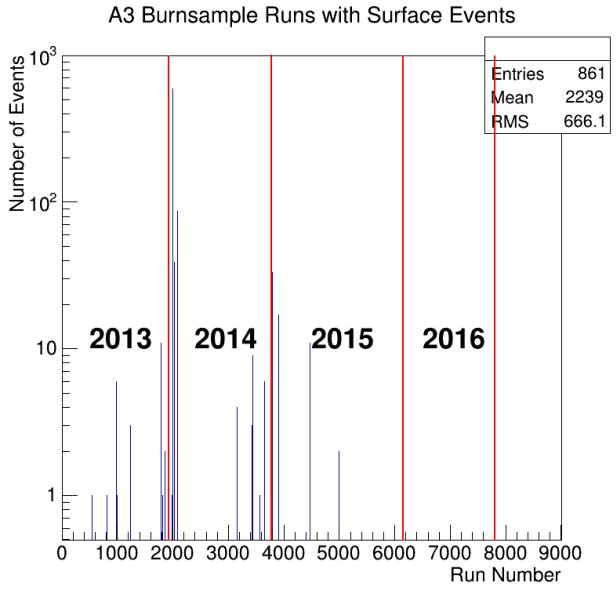
Config	A2	A3
1	8.414	8.716
2	8.938	7.898
3	8.377	same as 1
4	8.574	same as 1
5	9.723	same as 1

5.6.1.7 Station-specific Cuts

In additional to the cuts described so far, one extra cut exists for A3. This is the "spikey ratio" cut. Five events similar to that shown in Fig. 5.24 are found in configuration 2 in the 10% sample of 2013. The default reconstruction fails to identify a well-confined hot spot in the sky maps. However, the pulse times are consistent with the hypothesis that the signal is emitted from the vicinity of the BH-BV, suggesting a self-noise nature of such events. This source of this self-noise is still under investigation. To reject such events, we tailor the criteria and note that such events can be characterized by a large amplitude gradient from D1

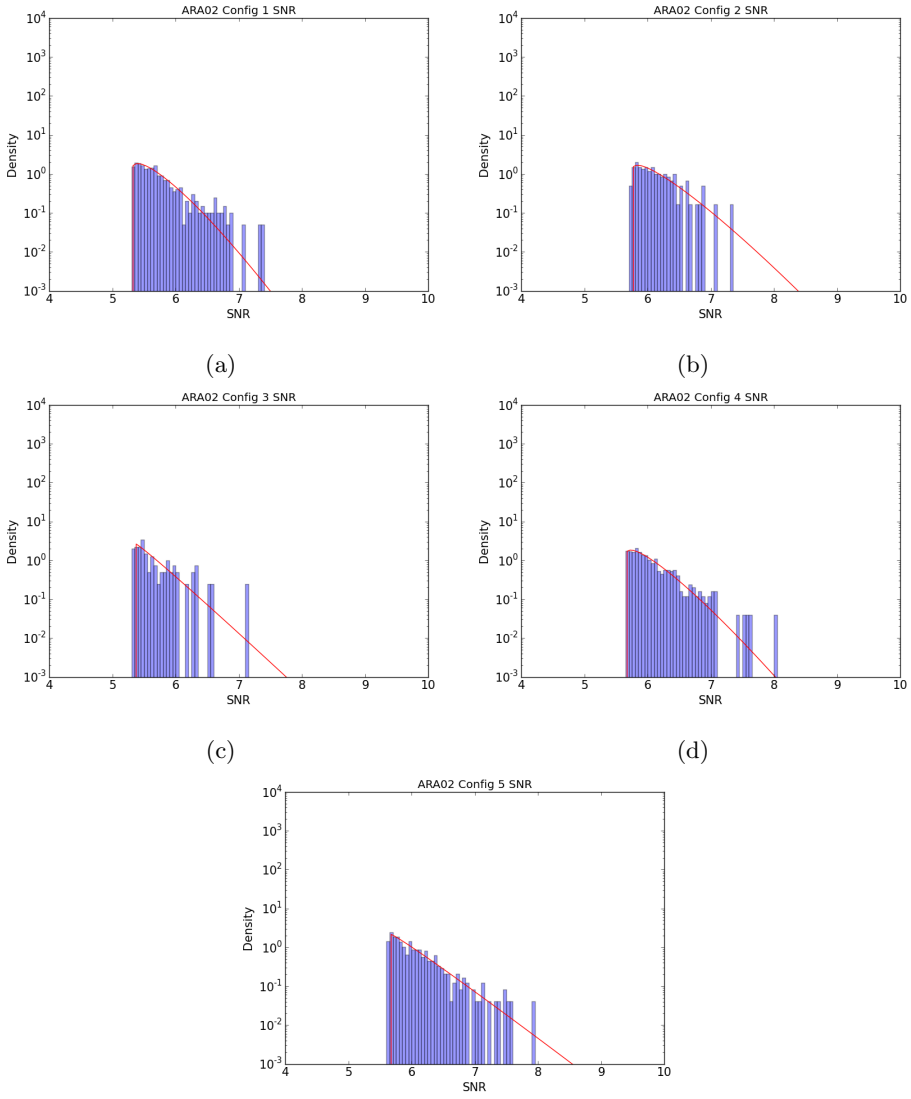


(a)



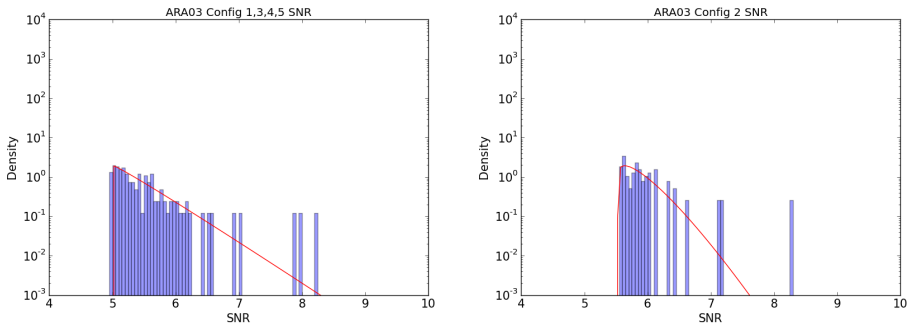
(b)

Figure 5.21: Number of events above the initial surface cut in the 10% sample for A2 (a) and A3 (b) in each run. The initial surface noisy run is defined as runs where at least two surface events are present in the 10% sample. The number of surface events can be seen to correlate strongly with the austral summers, indicating an origin of human-activity nature. It is also apparent that A2 received more noise from the surface than A3. This is reasonable since the location of A2 is closer to the South Pole skiway and other sites of surface activities.



(a) (b) (c) (d) (e)

Figure 5.22: Event SNR distributions of A2 10% sample after applying other initial cuts. The red curve is the best-fit Weibull model. (a)-(e): configuration 1 - 5.



(a) (b)

Figure 5.23: Event SNR distributions of A3 10% sample after applying other initial cuts. The red curve is the best-fit Weibull model. In (a), configurations 1, 3, 4, 5 are combined for the fit due to low individual statistics in each configuration. Configuration 2 is fitted independent of other configurations due to its $\sim 0.5\sigma$ higher n-channel filter threshold.

to other strings. Therefore, we follow the treatment from [94] and define the "spikey ratio" to used as a cut parameter. First for each string, an average peak amplitude $|V_{peak,avg}|$ is computed across its TV, BV, and BH. The spikey ratio is then defined as $|V_{peak,avg}|_{D1}/(|V_{peak,avg}|_{D2} + |V_{peak,avg}|_{D3} + |V_{peak,avg}|_{D4})/3)$. In the case where D4 is dropped in A3, the denominator becomes the average of the D2 and D3. If the spikey ratio is above a configuration-specific cut value, the event is rejected. The cut values are derived from exponential fits to the spikey ratio distributions of the 10% data, extrapolated to a background level of 0.01 event per year. This is shown in Fig. 5.25, and the cut values are listed in Tab. 5.10. Note that this background estimate does not translate to the final background since after the other cuts are applied it becomes negligible.

Table 5.10: Spikey ratio cut for each configuration in A3 as defined by the procedure described in text.

Config	Spikey Ratio Cut
1	2.4055
2	2.5267
3	3.5007
4	3.9877
5	3.6174

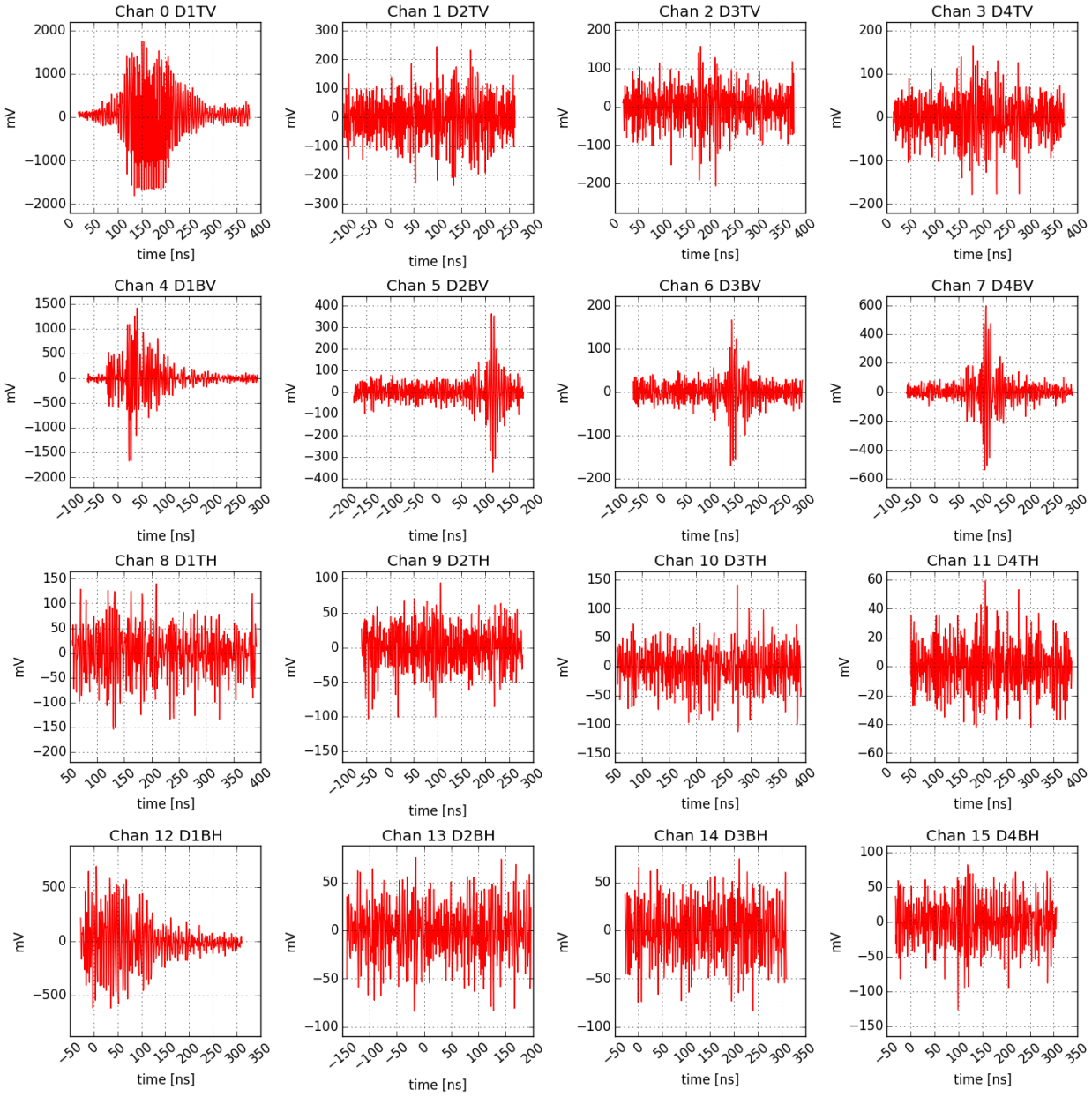


Figure 5.24: A3 run 1115 event 76709, a spikey-ratio event. The pulses seen in the BVs and D1TV are consistent with the signal being emitted from the location of BH-BV pair.

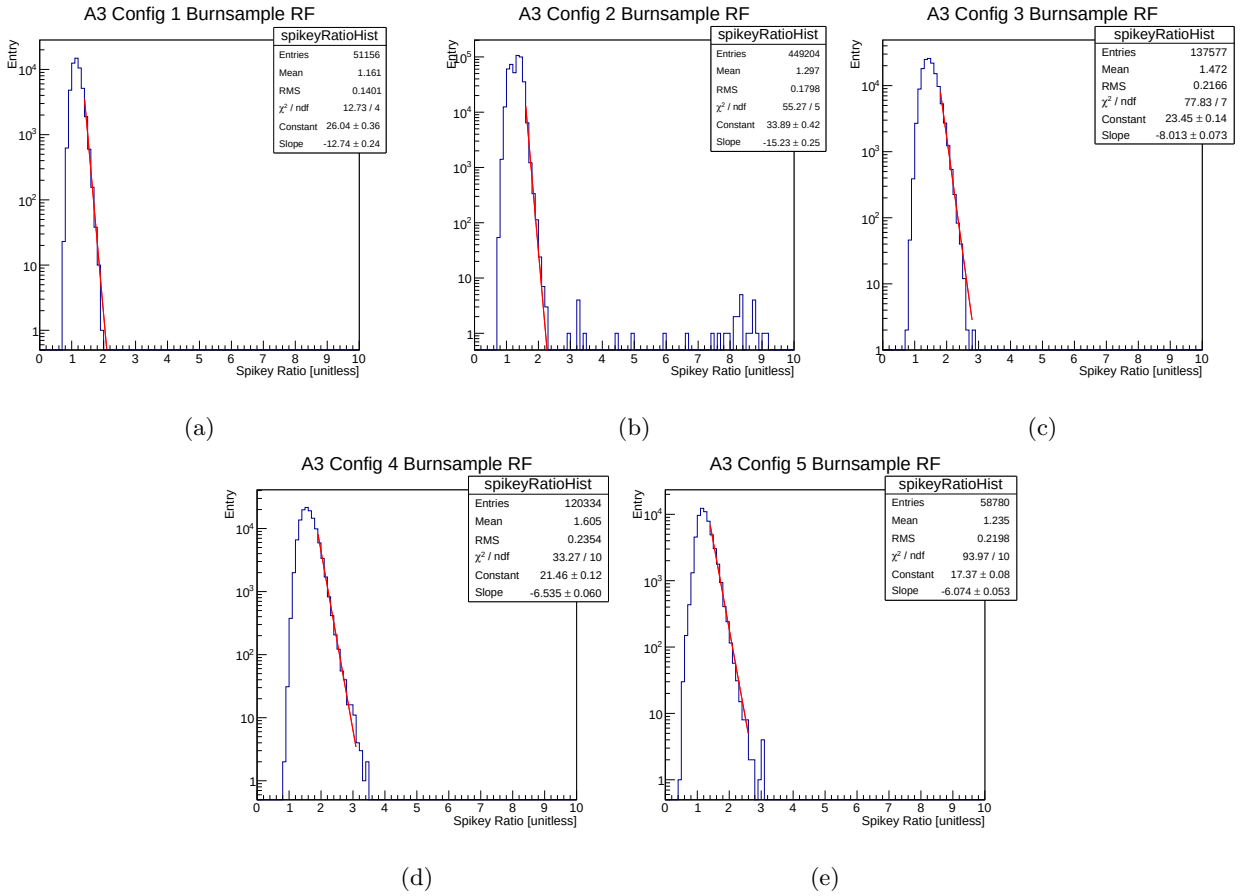


Figure 5.25: D1 spikey ratio distributions and exponential fits for A3 10% sample. (a) - (e): configuration 1 - 5. The red line is the exponential fit to the right tail of the distribution used to defined the cut value. In the 10% sample only configuration 2 contains events that would be removed by the cut. Nevertheless, the cut is applied to all configurations when unblinding.

5.6.2 Final Analysis — Stage 2

Using the above initial cuts, all events in the unblinded dataset are removed and no neutrino candidates are found. In addition, several kinds of problematic events or waveform anomalies are identified and treated. Assuming that all categories of backgrounds in the full dataset are represented in the 10% sample, the next step before "opening" the full dataset is to make sure the signal to background event ratio is at the ~ 10 level to maximal sensitivity. In practice, this requires that the background contribution from each cut be estimated, and the cut values be adjusted on the background-driven basis.

To estimate the background contribution from each cut, all cuts except the one to be estimated are applied to the unblinded sample. This is referred to as the "all-but-one" method. By modeling the distribution of the cut parameter from the events left, the survival probability over the cut value can be computed, which gives the number of backgrounds when multiplied by the number of events left. This number is weighted by the live time of the configuration it belongs to, so the live time also needs to be taken into consideration when compared with the target background rate. One difficulty that was encountered in this prescription is the lack of statistics when all-but-one cuts are applied, specifically for the thermal cut and the SNR cut, for the events removed by the former is usually also removable by the latter. This can be resolved by relaxing a cut that is already applied to allow more surviving events to populate the parameter space of the cut being evaluated, resulting in a more statistically robust estimate of the background distribution. This procedure and its detailed use will be described in the sections to follow.

Once the background contribution of each cut is evaluated, they are summed to give the total background. This effectively treats the backgrounds as uncorrelated and is conservative. The cuts for which we evaluate the backgrounds are:

- Thermal cut
- SNR cut
- Surface cut
- Calibration pulser geometric cut

The combined background rate should be ≈ 0.01 backgrounds per 228 (A2) or 220 (A3) days. Once the final cut values are set by adjusting for the desired level of backgrounds, the 100% dataset is analyzed for the number of surface noise in each run. This is used to refine the surface noisy run definition and produces a final list of surface noisy runs. When the 100% sample is used to search for in-ice signals, runs in this list will be skipped.

5.6.2.1 Thermal Cut

To estimate the backgrounds from the thermal cut, all other cuts were applied. In particular, the SNR cut is relaxed from its initial value in -2% steps. The step is enumerated from zero. The coherence distribution of the surviving events in each relaxation step is again fitted with a GEV PDF using a binned maximum likelihood method. To choose a step for background estimation, the following criteria are used χ:

- More than 100 events are left after all-but-one cuts
- Reduced-χ^2 is closest to 1

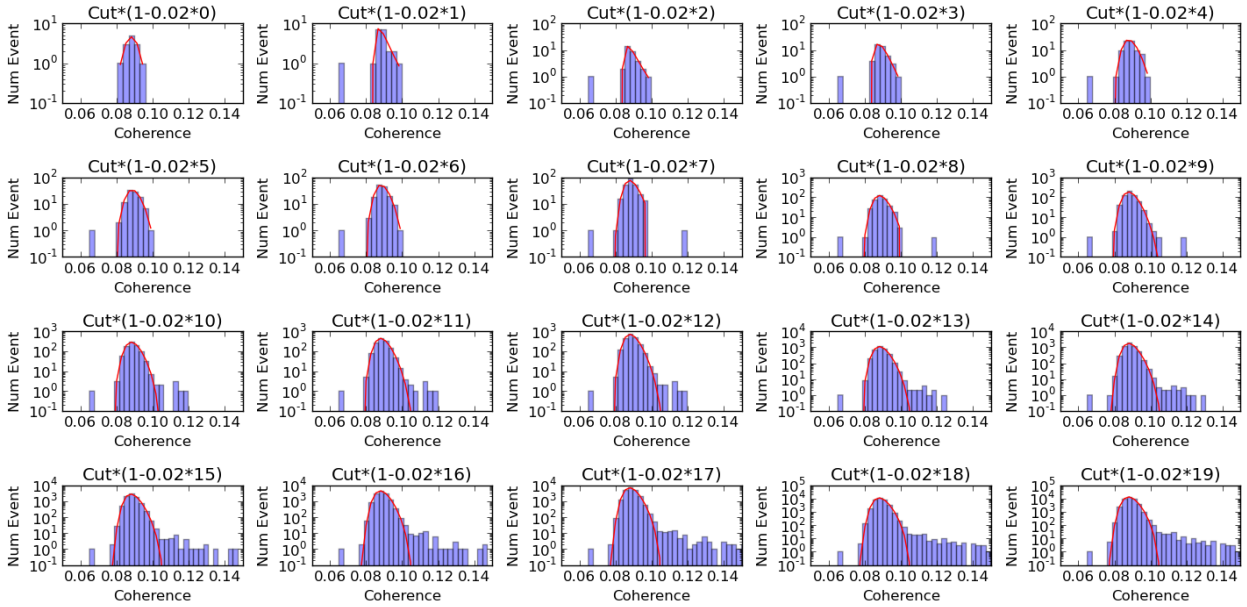


Figure 5.26: Coherence distribution of A2 configuration 1 when the SNR cut is relaxed in -2% steps. Top left to bottom right: step 0 to step 19. In each step, the red curve represents the best-fit GEV PDF.

For instance, the coherence distribution of the surviving events at each step of the SNR relaxation for A2 configuration 1 is shown in Fig. 5.26. In this example, the steps from 0 to 19 are examined, and the number of surviving events and the reduced χ^2 are listed in Tab. 5.11. Following the above selection criteria, step 7 is chosen to be representative of the background distribution, and its GEV fit will be used to estimate the background from the thermal cut. Starting from the initial value, the coherence thermal cut is tuned to reach a survival probability that would give a background count of 0.001 in 228 or 220 days of live time. In some cases the initial cut value already gives a sufficiently low background level. The cut values are left as-is

in these cases. In general the cut values are shifted only by $\leq 10\%$ after the tuning, since the GEV function drops very rapidly in the relevant range. The signal efficiency decrease over the shift is only a few percent.

Table 5.11: A2 configuration 1 SNR cut relaxation steps and the out-of-band coherence distribution of the remaining events with binned maximum likelihood GEV fit paramaters. The initial thermal cut is coherence ≥ 0.1037. After adjusting for the background using the fit from step 7 (cyan), it is redefined to be coherence ≥ 0.1089. d.o.f.: degree of freedom

Step	SNR Cut	N_{event}	N_{bin}	d.o.f.	χ^2	Reduced χ^2
0	8.4144	13	5	2	0.0587	0.02936
1	8.2461	21	7	4	1.908075	0.477019
2	8.0778	33	7	4	1.026768	0.256692
3	7.9096	45	7	4	1.114352	0.278588
4	7.7413	75	8	5	2.452805	0.490561
5	7.573	104	8	5	2.229827	0.445965
6	7.4047	149	8	5	2.799697	0.559939
7	7.2364	250	10	7	6.907178	0.98674
8	7.0681	392	12	9	8.678852	0.964317
9	6.8998	599	14	11	11.623514	1.056683
10	6.7315	917	18	15	18.355581	1.223705
11	6.5632	1415	26	23	27.868026	1.211653
12	6.395	2216	28	25	34.567003	1.38268
13	6.2267	3463	36	33	48.63725	1.473856
14	6.0584	5406	44	41	67.39646	1.643816
15	5.8901	8596	50	47	100.368691	2.135504
16	5.7218	13697	62	59	171.153863	2.900913
17	5.5535	21576	84	81	259.670682	3.205811
18	5.3852	34284	98	95	402.29833	4.234719
19	5.2169	40585	104	101	481.335326	4.765696

The step taken for each configuration is summarized in Tab. 5.12. The thermal cut is finalized as in Tab. 5.13 and Tab. 5.14.

Table 5.12: SNR cut relaxation steps chosen to evaluate the thermal cut background contribution for A2 and A3 following the criteria described in text. The steps and fits are evaluated for in-band and out-of-band data independently.

	A2		A3	
Config.	In-band	Out-of-band	In-band	Out-of-band
1	10	7	15	14
2	8	10	7	7
3	11	8	11	13
4	10	7	10	20
5	12	14	17	17

Table 5.13: Final thermal cut values for A2

Config	In-band	Out-of-band
1	0.1080	0.1089
2	0.1093	0.1089
3	0.1088	0.1042
4	0.0990	0.0947
5	0.0977	0.0951

Table 5.14: Final thermal cut values for A3

Config	In-band	Out-of-band
1	0.1128	0.1058
2	0.1113	0.1054
3	0.1295	0.1588
4	0.1351	0.1341
5	0.1381	0.1198

5.6.2.2 SNR Cut

To estimate the background contribution from the SNR cut, the coherence thermal cut is relaxed in -2% steps. The remaining events at each relaxation step is fitted with a Weibull PDF, as shown in Fig. 5.27. Due

to the presence of empty bins in the distribution, the reduced χ^2 is no longer straightforward to compute. Hence the step selection criteria are altered to be:

- More than 100 events are left after all-but-one cuts
- The increase rate in the number of surviving events encounters the first abrupt change

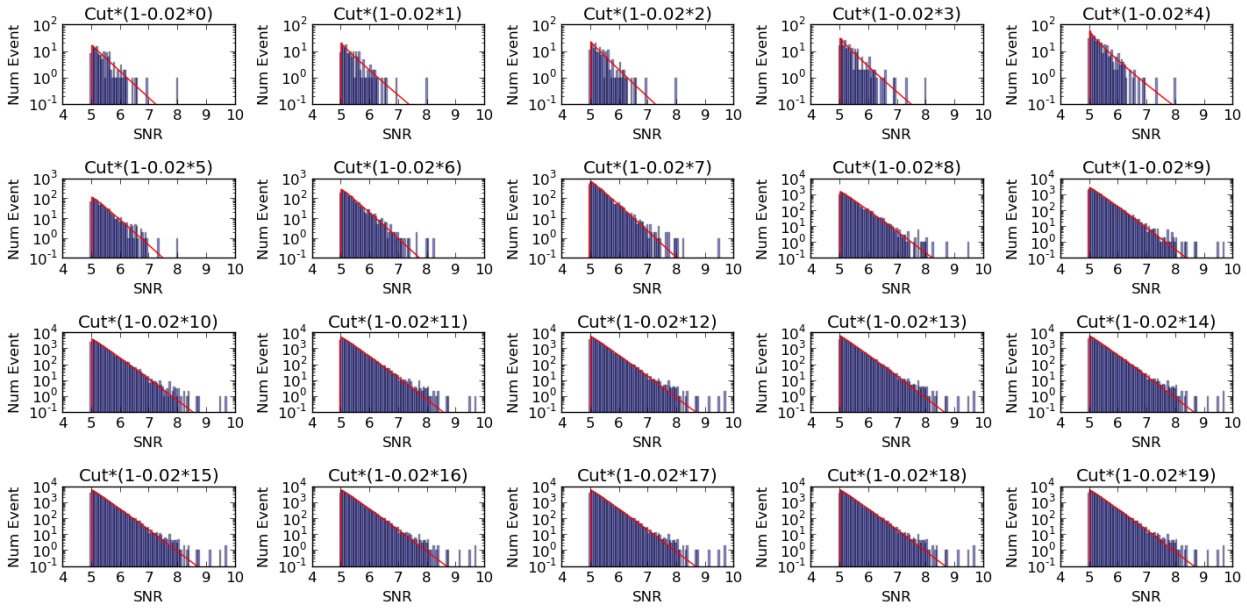


Figure 5.27: SNR distribution of A3 configuration 1,3,4,5 when the thermal cut is relaxed in -2% steps. Top left to bottom right: step 0 to step 19. In each step, the red curve represents the best-fit Weibull PDF.

The reason behind the choice of the second criterion is that the background is dominated by the events at the tail of the distribution. These events are not properly described by the fitted model in Stage 1, which tends to describe the thermal population near the distribution center. The point where the increase rate of the surviving events changes abruptly indicates the transition from the non-thermal tail events to the thermal events. Consequently the step where this change is encountered is chosen for the SNR distribution to be fitted. For example, in Fig. 5.28, the number of surviving events and its increase rate in each relaxation step for A3 configuration 1+3+4+5 are shown. The slope of the increase rate can be observed to change noticeably at the second data point, corresponding the step 2. Thus, the fitted Weibull PDF from step 2 will be used to calculate the number of backgrounds. In Tab. 5.15, the steps taken for each configuration are shown. The signal efficiency tends to scale strongly with the SNR, as shown in Fig. 5.29, with a $\sim 45\%$ drop in efficiency over an SNR range of 10 to 20 as estimated using 10^{18} eV simulation. Therefore, an

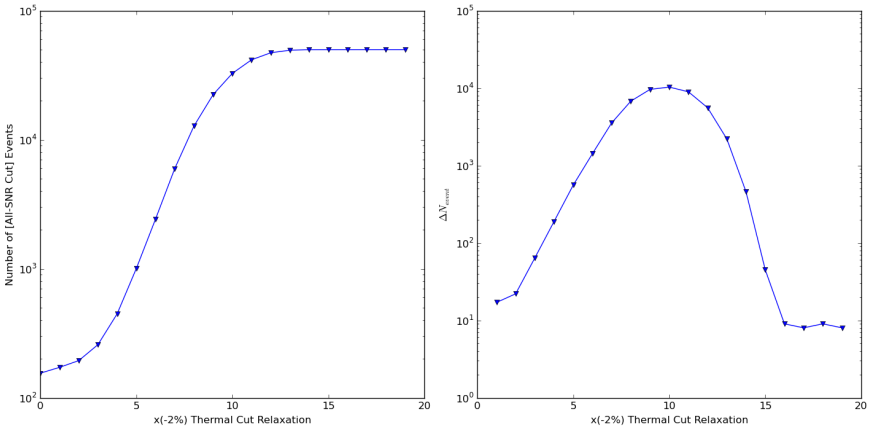


Figure 5.28: Left: A3 configuration 1,3,4,5 number of surviving events at each thermal cut relaxation step. All cuts except the SNR cut applied. Right: the increase rate of the number of surviving events at each step. The sudden change at step 2 is interpreted as the transition from non-thermal events, which dominates background estimation, to thermal events.

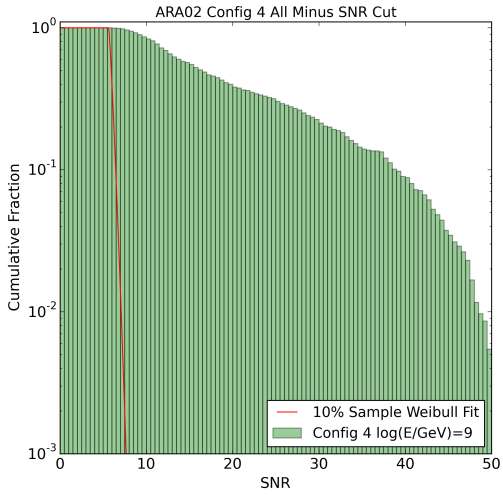


Figure 5.29: 10^{18} eV neutrino signal pass rate versus event SNR in A2 configuration 4. The red curve represents the Weibull fit to the event SNR distribution after all cuts are applied except the SNR cut.

optimization procedure is placed to search for the optimal cut value while maximizing the signal pass rate. This procedure takes into account both the SNR cut and the surface cut, and will described in the next section.

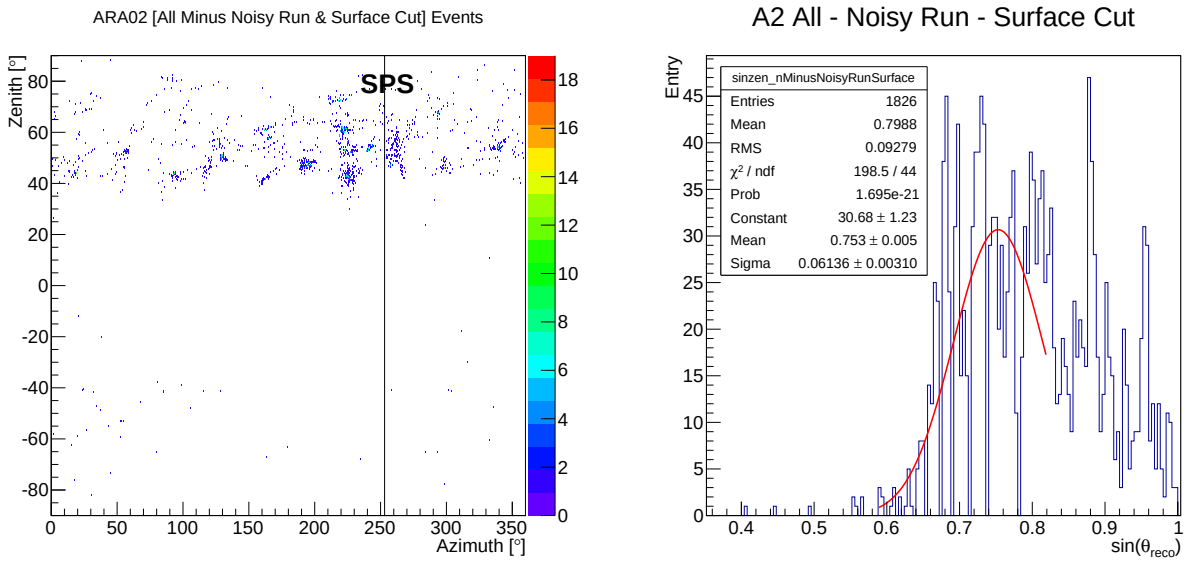
Table 5.15: Thermal cut relaxation steps chosen to evaluate the SNR background contribution for A2 and A3 following the criteria described in text.

Config	A2	A3
1	3	2
2	1	2
3	2	2
4	2	2
5	2	2

5.6.2.3 Surface Cut

The quasi-planewave reconstructed zenith angle is used as the parameter in which the background estimate will be carried out. The low-statistics issue is also present for this estimate. In A2, after all cuts except the surface cut are applied, only 53 events remain when all configurations are combined. In A3, < 20 events remain. To increase the number of surface events, rejected events in the initial surface noisy runs are put back. This makes the number of surface events close to 2000 for A2 and close to 1000 for A3. Following the

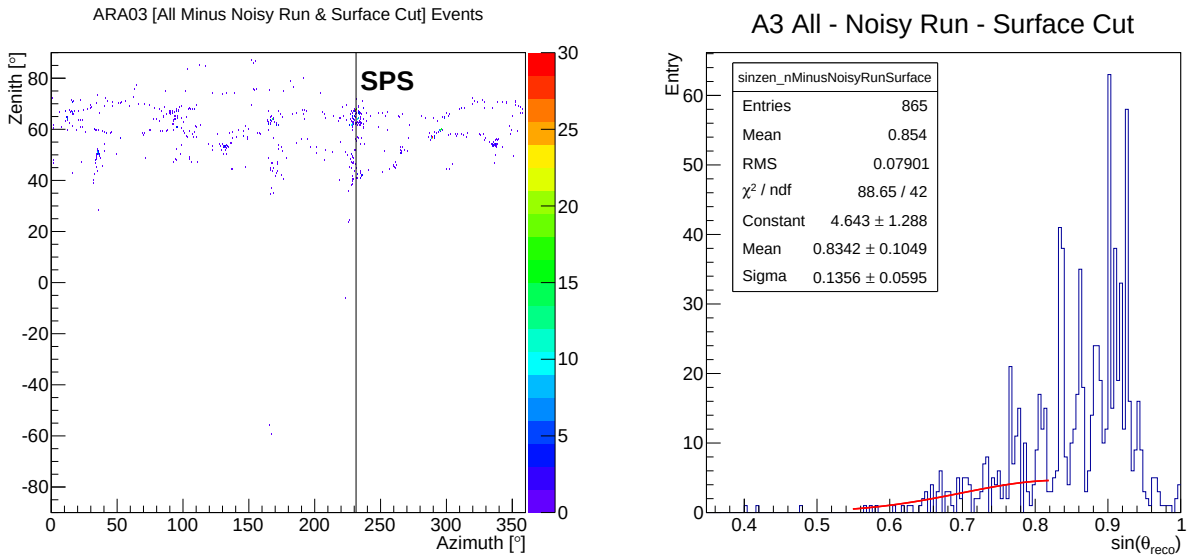
procedure in defining the initial surface cut, a Gaussian PDF is fitted to the quasi-planewave zenith in the air-ice transition region. The fits are shown in Fig. 5.30 and Fig. 5.31. The flatness of the tail in A3 leads to a much steeper surface cut, as will be shown in Tab. 5.17.



(a) (b)

Figure 5.30: Quasi-planewave reconstructed directions of surface events in A2 10% sample from all configurations. All initial cuts applied except the surface cut and the noisy surface runs. (a): local sky map of reconstructed directions. SPS indicates the azimuth of the South Pole Station, around which clusters of surface events are found. (b): the $sin(\theta_{reco})$ distribution of the events. The air-ice transition is fitted with a Gaussian function (red) with which the background is calculated. Note that although the Gaussian fit does not agree with data in the -1σ region, it describes the tail ($sin(\theta_{reco}) \approx 0.6$) of the data well. The optimization is carried out over the tail region.

With the Gaussian fit obtained, it is combined with the Weibull fit from the SNR cut acquired in the previous section for each configuration in an optimization process that maximizes the pass rate of 10^{18} eV simulated neutrino events while keeping the combined background rates to 0.011 in 228 or 220 days. This two-dimensional parameter scan can be visualized in Fig. 5.32 for A3 configuration 1. The final SNR and surface cut values are summarized in Tab. 5.16 and Tab. 5.17.



(a) (b)

Figure 5.31: Quasi-planewave reconstructed directions of surface events in A3 10% sample from all configurations. All initial cuts applied except the surface cut and the noisy surface runs. (a): local sky map of reconstructed directions. SPS indicates the azimuth of the South Pole Station, around which clusters of surface events are found. (b): the $sin(\theta_{reco})$ distribution of the events. The air-ice transition is fitted with a Gaussian function (red) with which the background is calculated. Note that although the Gaussian fit does not agree with data within the -1σ region, it describes the tail ($sin(\theta_{reco}) \approx 0.6 - 0.7$) of the data well. The optimization is carried out over the tail region.

Table 5.16: Final SNR cut values for A2 and A3 after a joint optimization on simulated 10^8 GeV neutrinos with the surface cut is performed.

Config	A2	A3
1	10.192	10.426
2	9.5037	9.0050
3	9.4592	8.9250
4	10.033	8.5672
5	10.112	8.8378

Table 5.17: Final quasi-planewave surface cut values for A2 and A3 after a joint optimization on simulated 10^8 GeV neutrinos with the SNR cut is performed. The cut values for A2 configurations 2, 3 and A3 configuration 4 remain unchanged from the initial values due to the number of surviving events being zero when all other cuts are applied.

Config	A2 [°]	A3 [°]
1	30.37	11.39
2	35.96	12.80
3	35.96	13.35
4	31.40	33.86
5	30.76	13.99

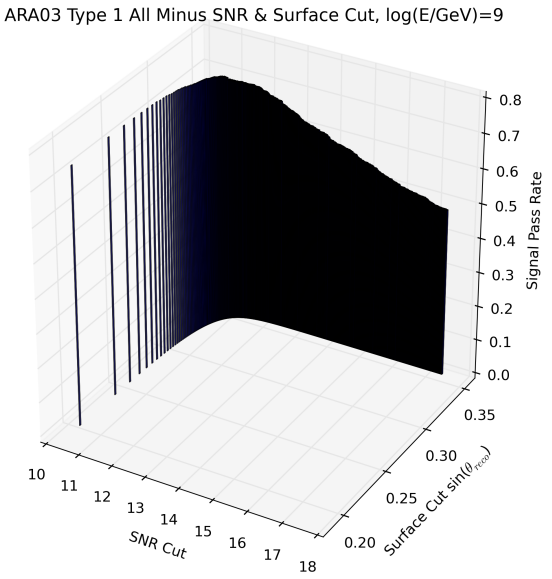
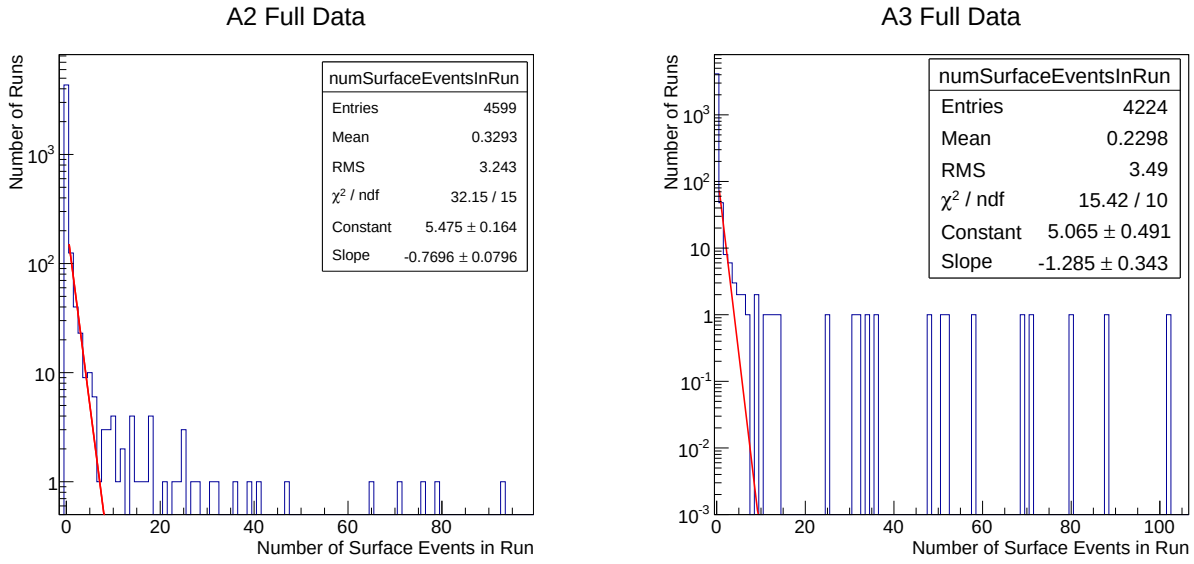


Figure 5.32: SNR and surface cut optimization with 10^{18} eV simulated signal for A3 configuration 1. The surface cut and the SNR cut are confined to combine for a total background of 0.01 events per 220 days. The combination of cut values that gives the highest signal pass rate as represented by the signal simulation is chosen as the final cuts.

5.6.2.4 Surface Noisy Runs

With the surface cut angle modified in the last section, for consistency the surface noise runs are redefined. In order to properly set the event threshold over which a run is classified as surface noisy, the 100% sample is reconstructed, and the number of non-thermal events that are rejected by the surface cut are counted for each run. This approach keeps the "blindness" of the in-ice phase space relevant to neutrinos, and allows only the inspection of the "surface box" where the events would be rejected eventually in any case. Fig. 5.33 shows the number of runs as histogrammed by the count of surface events. An exponential fit is performed on each histogram to extrapolate to a count corresponding to only 0.01 runs. For A2, this number is 13.10; for A3, 7.53. These numbers are rounded up to give a threshold of 14 and 8, over which the run is classified as surface noisy and excluded from being unblinded in the "neutrino box". Totally there are 35 surface noisy runs for A2 and 32 runs for A3. Again, listed runs will be excluded from unblinding the 100% sample (neutrino box opening).

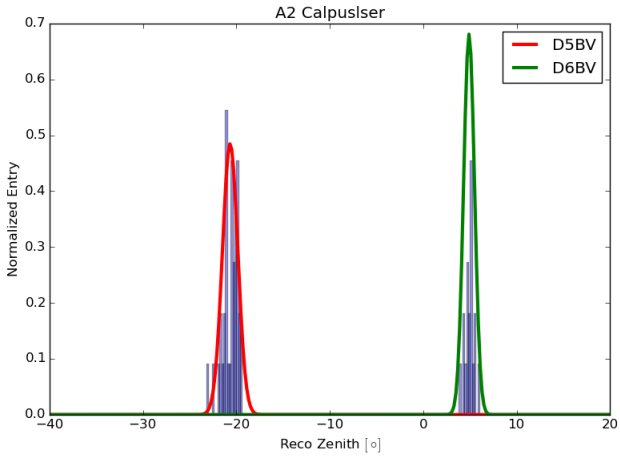


(a) (b)

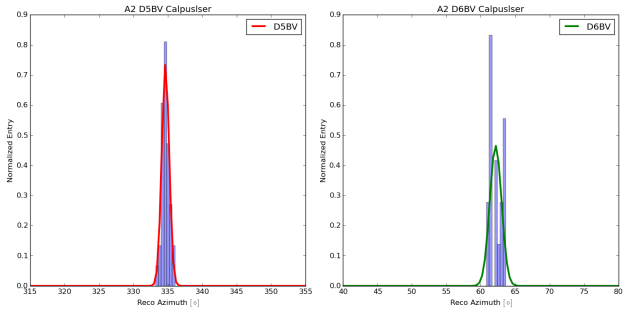
Figure 5.33: Number of runs histogrammed by the number of surface events in run with 100% data. (a): A3. (b): A3. The red line marks the exponential fit to the distribution and is used to extrapolate to a threshold value of 0.01 runs. The number of surface events per run is defined as the number of non-thermal events vetoed by the surface cut in each run using the 100% sample.

5.6.3 Final Background Estimate

So far the cut values of the thermal cut, the SNR cut, and the surface cut are tuned to give the desired background rate. The last contribution to the background that we consider is the calibration pulser geometric cut. Applying the all-but-one cut method, The zenith and azimuth of surviving events can be modeled by Gaussian PDFs, as shown in Fig. 5.34 and Fig. 5.35. Using the initial calibration pulser geometric cuts, the expected total number of backgrounds can be calculated. The contributed backgrounds in the full live time are roughly 3×10^{-7} events in A2 and 4×10^{-9} events in A3, negligible compared to other background sources.

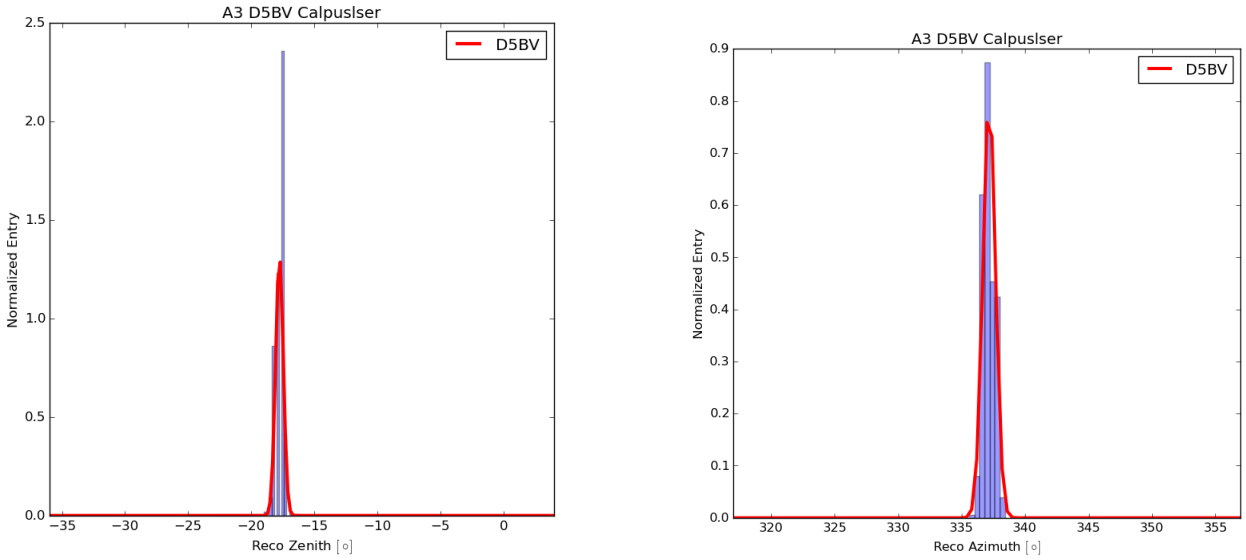


(a)



(b)

Figure 5.34: Zenith (a) and azimuth (b) distribution of A2 10% data after all cuts except the calibration pulser geometric cut is applied. The curves are the Gaussian PDF fits to the histogram, which in turn are used to compute the background level beyond the cut boundaries. Note that the events here are tagged as regular RF triggers, but in fact are from the calibration pulser. The calibration pulser geometric cut successfully mask out such events when applied.



(a) (b)

Figure 5.35: Zenith (a) and azimuth (b) distribution of A3 10% data after all cuts except the calibration pulser geometric cut is applied. The curves are the Gaussian PDF fits to the histogram, which in turn are used to compute the background level beyond the cut boundaries. Note that the events here are tagged as regular RF triggers, but in fact are from the calibration pulser. The calibration pulser geometric cut successfully mask out such events when applied. No mistagged events from D6BV are found in the data.

The background contribution from each cuts in the 100% sample can be summarized in the two following tables:

Table 5.18: Estimated total number of background events in A2 over the full live time in this analysis. The errors are statistical. The final error is calculated with standard Gaussian error propagation.

A2	N_{bkg}	+	-
Thermal	4.7237×10^{-3}	6.5580×10^{-3}	1.6777×10^{-3}
SNR	4.1160×10^{-2}	1.1748×10^{-2}	6.9902×10^{-3}
Surface	7.9384×10^{-3}	1.1001×10^{-2}	1.7839×10^{-2}
Cal. Geom.	3.1789×10^{-7}	6.5503×10^{-10}	6.5378×10^{-10}
Total	5.3822×10^{-2}	1.7379×10^{-2}	1.9233×10^{-2}

Table 5.19: Estimated total number of background events in A3 over the full live time in this analysis. The errors are statistical. The final error is calculated with standard Gaussian error propagation.

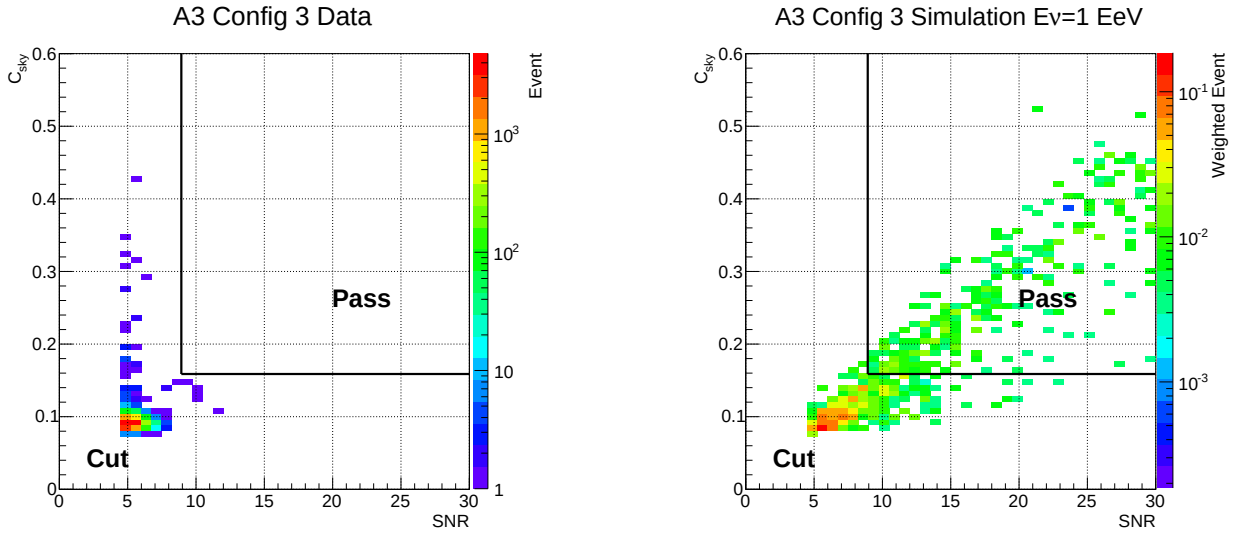
A3	N_{bkg}	+	-
Thermal	3.0405×10^{-3}	8.3666×10^{-3}	1.3054×10^{-3}
SNR	4.9792×10^{-2}	2.1636×10^{-2}	1.2093×10^{-2}
Surface	4.9183×10^{-4}	7.9531×10^{-3}	2.8754×10^{-4}
Cal. Geom.	3.6393×10^{-9}	8.9281×10^{-11}	8.7228×10^{-11}
Total	5.3324×10^{-2}	2.4523×10^{-2}	1.2166×10^{-2}

5.7 Effective Area And Sensitivities

The analysis cuts presented in the previous section have varied levels of impact on the neutrino signal efficiency. While the surface cut and the calibration pulser geometric cut veto event based on geometric properties, the thermal cut and the SNR cut combine to rejects the majority of the thermal events and CW contamination. This latter point is illustrated in Fig. 5.36. In Tab. 5.20 and Tab. 5.21, the signal efficiency at 10^{18} eV for individual cuts are tabulated. In most configurations, the SNR cut removes $30 \sim 40\%$ of signals, representing the least efficienct cut. It is followed by the surface cut, which removes up to $\sim 30\%$ of simulated neutrinos in A3, a direct consequence of the slowly-falling tail of events in the air-ice transition region, as seen in Fig. 5.31.

Table 5.20: 10^{18} eV signal efficiency for individual analysis cuts in A2. An efficiency of unity means the cut removed no signal. An efficiency of zero means the cut removes all signal.

A2	Thermal	SNR	Surface	Cal. Geom.
1	0.80	0.60	0.86	0.96
2	0.76	0.70	0.91	0.97
3	0.84	0.69	0.90	0.97
4	0.78	0.65	0.89	0.97
5	0.79	0.57	0.85	0.96



(a) (b)

Figure 5.36: An example of the two-dimensional parameter space formed by the coherence value (denoted as C_{sky} here) and the SNR, over which the final cust is made for A3 configuration 3, effectively forming a box cut. (a): the plane as observed in 10% data, showing events clustering at low-correlation and low-SNR. (b): the plane populated with simulated neutrinos at 10^{18} eV, showing events distributed throughout. Events at low-correlation and low-SNR are cut; events at higher values define the signal region, and pass the analysis.

Table 5.21: 10^{18} eV signal efficiency for individual analysis cuts in A3. An efficiency of unity means the cut removed no signal. An efficiency of zero means the cut removes all signal.

A3	Thermal	SNR	Surface	Cal. Geom.
1	0.79	0.60	0.72	0.99
2	0.75	0.68	0.76	0.99
3	0.53	0.65	0.73	1.00
4	0.59	0.67	0.88	0.99
5	0.78	0.67	0.73	0.99

In Fig. 5.37 and Fig. 5.38, the signal efficiency curves versus event SNR are shown. Confined by the SNR cut, the 50% point of the efficiency curve is typically $\geq$ an SNR of 9.5 for A2, and $\geq$ an SNR of 9 for A3. Above that, the efficiency of A2 generally plateaus around 80%. For A3, the variation among configurations is more apparent. For instance, in configuration 3, due the steep surface cut and a population of non-thermal events that forces the GEV fit to the coherence values to become wider, the final signal efficiency is only

around 50% above the SNR threshold. In other A3 configurations, the final efficiency is around $\geq$ 70%. Novel methods are being developed to replace the SNR cut and improve the coherence measure with a goal of recovering more signal. Additionally, a better understanding of the air-ice transition will help with the precise reconstruction of surface backgrounds, which can translate to more near-surface solid angles being recovered. Notice that the offset-block and cliff event detection algorithms have negligible impact on the signal efficiency. This justifies their design.

In Fig. 5.39, the live-time-weighted signal efficiency is plotted for each half-decade energy simulation. Again, the n-channel filter and the SNR cut play important roles in the deficiency at lower energies. The downturn of the efficiency at energies above $10^{20.5}$ eV is attributable to the low-frequency cut, as described in section 5.6.1.1. In general the efficiency rises to $\sim$35% by 10^{18} eV, and peaks near 10^{20} eV at between 50-60%, depending on the specific station.

The efficiency at each energy bin $\epsilon_{\text{ana}}(E)$ will be used to weight the energy-dependent trigger-level effective area, as obtained in section 3.5, to obtain the analysis-level effective area via

$$A_{\text{eff, ana}}(E) = \epsilon_{\text{ana}}(E) \cdot A_{\text{eff,trig}}(E). \tag{5.4}$$

The total analysis-level exposure Λ over all configurations in both stations can then be calculated, in each energy bin, by

$$\Lambda(E) = 4\pi \sum_{\text{station}=2}^{3} \sum_{\text{config}=1}^{5} (A_{\text{eff,ana,config}}(E) \cdot T_{\text{config}})_{\text{station}} \tag{5.5}$$

, where T_{config} denotes the live time of the specific configuration.

In the next section, it will be used to place a differential limit to the diffuse neutrino flux.

5.8 Analysis Results — Unblinding

The 100% datasets of A2 and A3 from 2013 to 2016 are unblinded with the event selection and cuts described in this chapter. The number of glitch events as listed in section 5.5.2 is found to be generally $\sim\leq 8 \times 10^{-5}$ of all RF triggers. In principle, any uncut events left require further investigation to determine if they are neutrino candidates. The number of total events and RF-tagged events in each configuration is shown in Tab. 5.22 and Tab. 5.23. For A2, there are two surviving events. For A3, there are four surviving events.

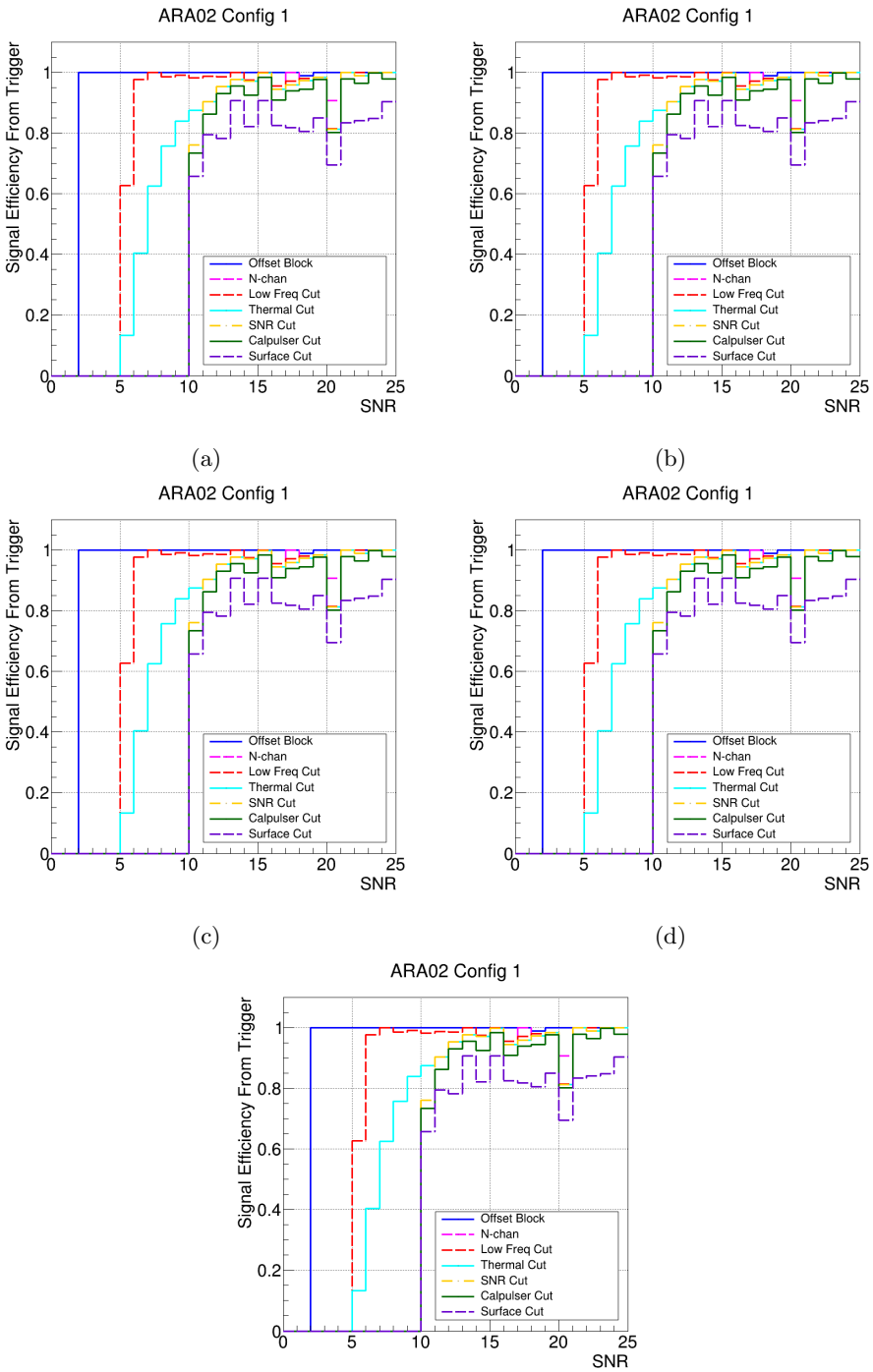


Figure 5.37: Monte Carlo estimated A2 signal efficiency at the analysis level versus the SNR of the event. (a)-(e): configuration 1 - 5. The last curve (surface cut) is the final signal analysis efficiency.

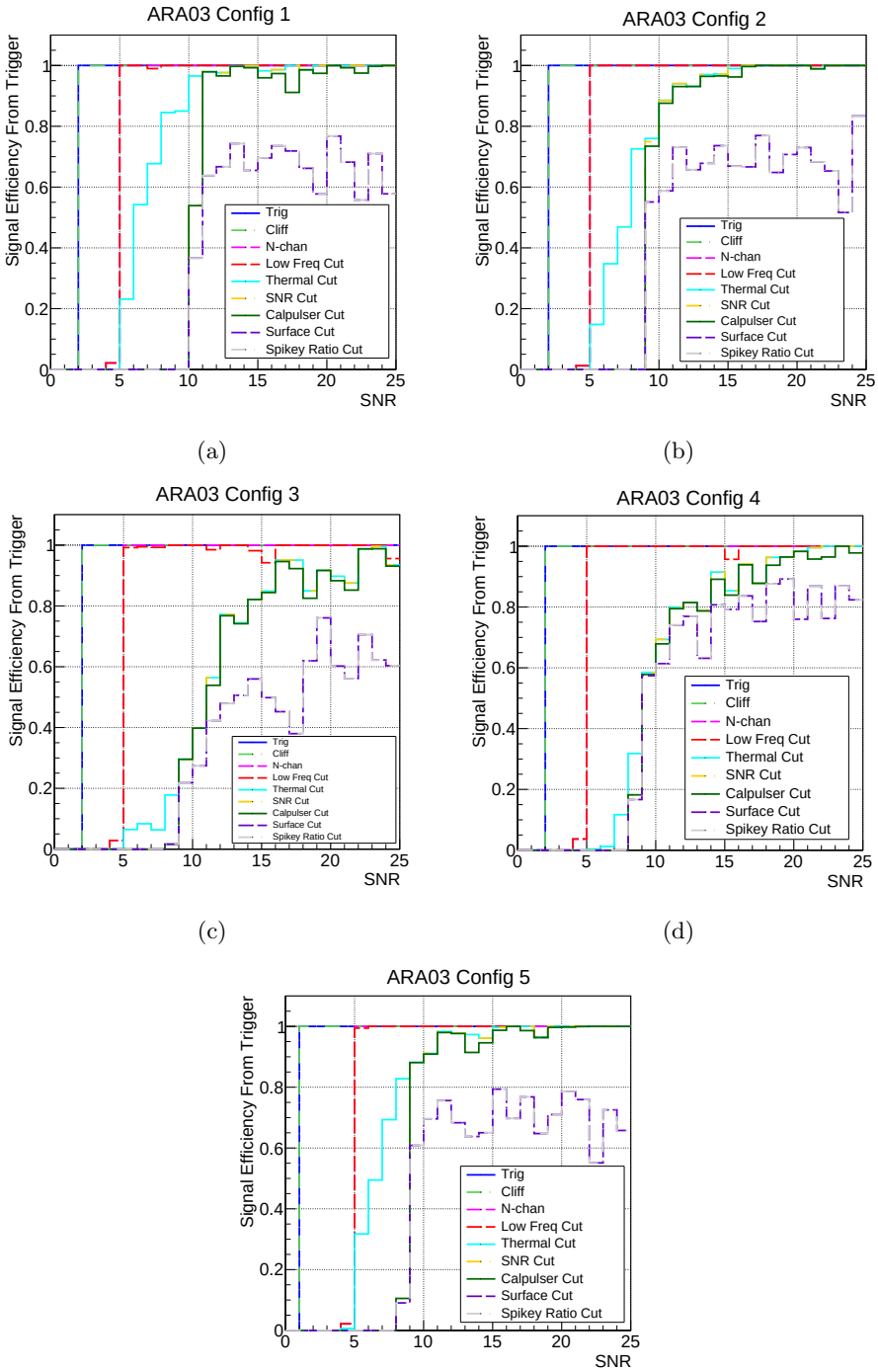
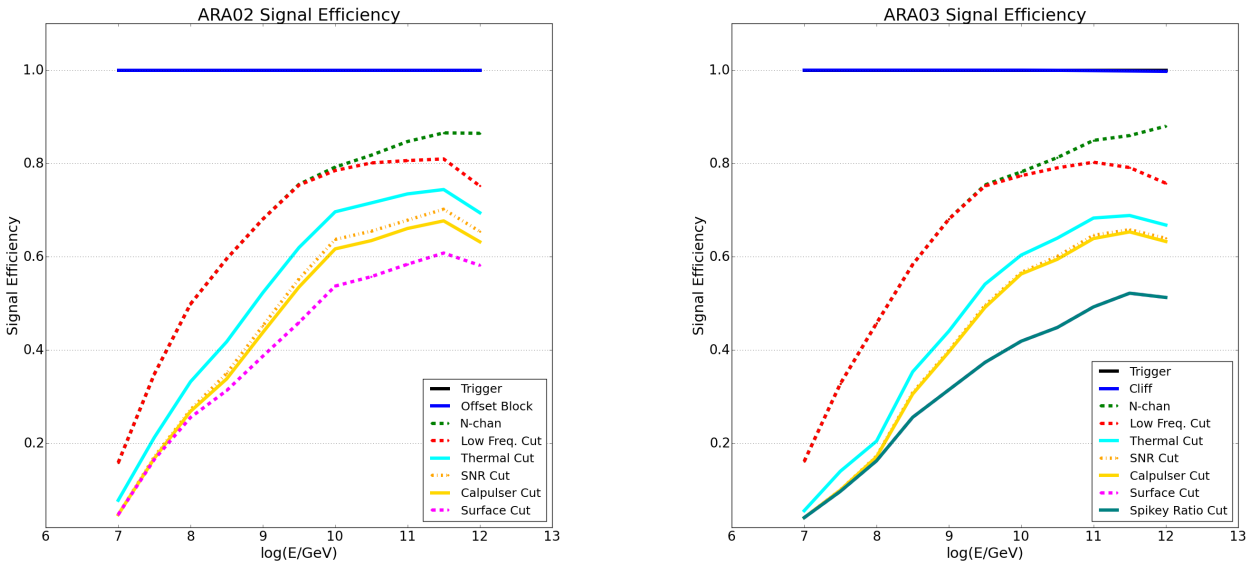


Figure 5.38: Monte Carlo estimated A3 signal efficiency at the analysis level versus the SNR of the event. (a)-(e): configuration 1 - 5. The last curve (spikey ratio cut) is the final signal analysis efficiency.



(a) (b)

Figure 5.39: Monte Carlo estimated A2 (a) and A3 (b) signal efficiency at the analysis level versus primary neutrino energy. The efficiency of each configuration are weighted by live time and combined.

Table 5.22: A2 number of total events and RF-tagged events in the 100% sample

Config	Total [million]	RF [million]
1	104	75
2	97	73
3	52	36
4	213	139
5	128	84

Table 5.23: A3 number of total events and RF-tagged events in the 100% sample

Config	Total [million]	RF [million]
1	43	31
2	89	71
3	200	146
4	170	126
5	117	83

The two events in A2 are

1. Run 3412 event 2959. Configuration 1, April 12 2014. Fig. 5.40

2. Run 8156 event 14165. Configuration 4, October 29 2016. Fig. 5.41

From the waveforms of the two event, clear bottom Vpol pulses can be observed. Indeed, if the quasi-planewave reconstruction is restricted to using the bottom Vpols, both events can be pointed back to an above-surface origin and rejected. Furthermore, for the first event, it can be verified that if the first block of samples that was discarded due to systematic offsets in each waveform is put back, a leading pulse in channel 0 can be found at a timing consistent with the down-going hypothesis as reconstructed by the bottom Vpols, as shown in Fig. 5.42. Given the evidence that the events are down-going with source directions above the surface cut, they are not considered as neutrino candidates. The failure of the default reconstruction to classify these events as from the surface can be attributed to the lack of signal in the top Vpols. This prohibits the formation of a TV-BV vertical baseline with accurate timing from signal which would ideally determine the zenith of the event. While the BVs have a depth difference of roughly 5 m among themselves, their power in determining the zenith is relatively weak compared to the longer (20 m) BV-TV baseline.

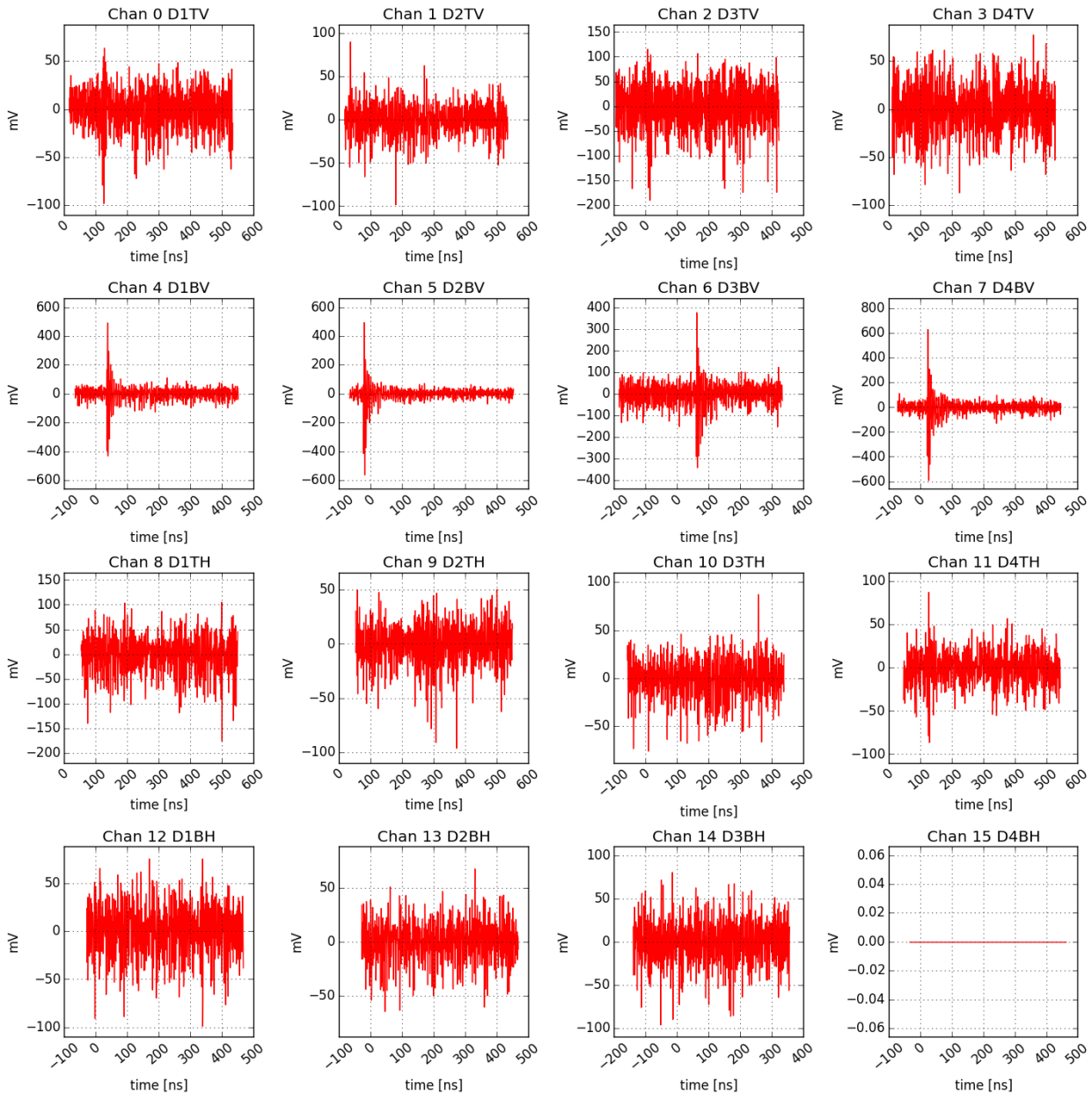


Figure 5.40: A2 run 3412 event 2959, a surviving event after unblinding. The BVs reconstruct to a source angle above the surface cut.

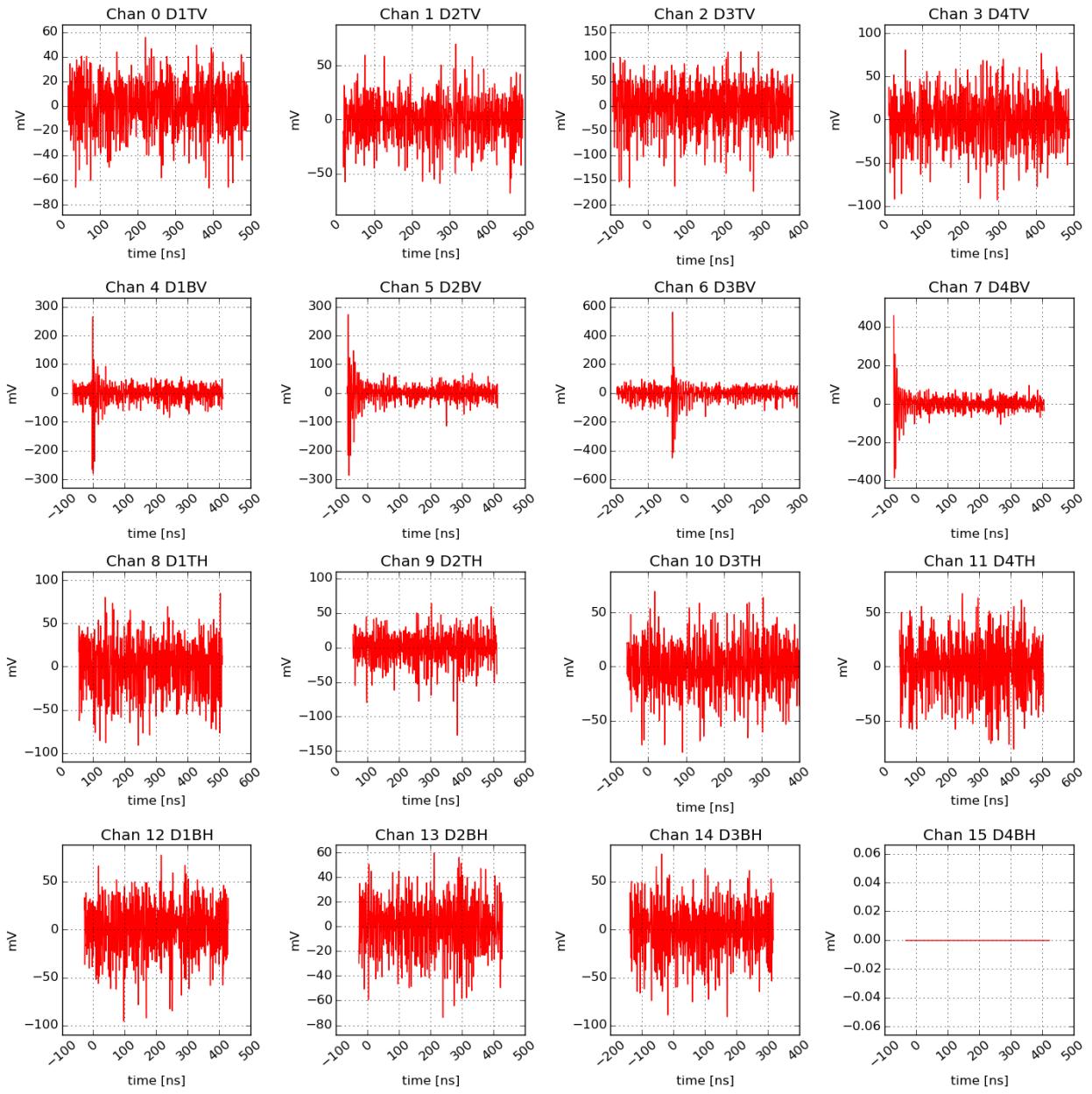


Figure 5.41: A2 run 8156 event 14165, a surviving event after unblinding. The BVs reconstruct to a source angle above the surface cut.

Therefore, in these events the zenith reconstruction is scrambled by noise. Once the signal-less top Vpol are excluded in the reconstruction to reduce noise, the zenith can be unambiguously resolved. A dedicated trigger study was carried out in response to these events. No systematic trigger misconfiguration or anomaly was found during the relevant data period that would otherwise require us to further account for them in the analysis.

The four events in A3 are

1. Run 6051 event 133654, 133842, 133779. Configuration 4, December 9 2015.

2. Run 6697 event 131864. Configuration 4, May 2 2016. Fig. 5.43

The first three events cluster in run 6051, and the waveforms show similar timing patterns. In addition, the run has six surface events in the 100% surface box, close to the threshold of eight for a surface-noisy run. Therefore, this run is considered an under-fluctuation of a surface noisy run, and the three surviving events are not considered sufficiently isolated to be neutrino candidates. The last event, as one can see in Fig. 5.43, has prominent pulses in the TVs from D1 to D3. Note that D4 is by default zeroed out in the reconstruction for the run. Again due to the lack of signal in the BVs, the default reconstruction is unable to unambiguously reject this event with the surface cut. However, hit-time based reconstruction algorithms developed by D.Z. Besson are able to point to event source to the surface. In addition, while there are visible glitches in the waveforms on D4, a pulse can be observed near the end of D4BV (channel 7), which is outside of the apparent glitch region. To incorporate this additional information in the interferometric reconstruction, a 100-ns snippet is taken around the end of D4TV to encapsulate the pulse, and it is inserted to the interferometry. The resulting quasi-planewave sky map is shown in Fig. 5.44. The result points to a down-going signal from above the surface cut with no ambiguity. Thus, the event is not considered a neutrino candidate.

In summary, no neutrino candidate is found in the four years of data from A2 and A3. The number of expected neutrino events as a result of the analysis from various flux models is shown in Tab. 5.24.

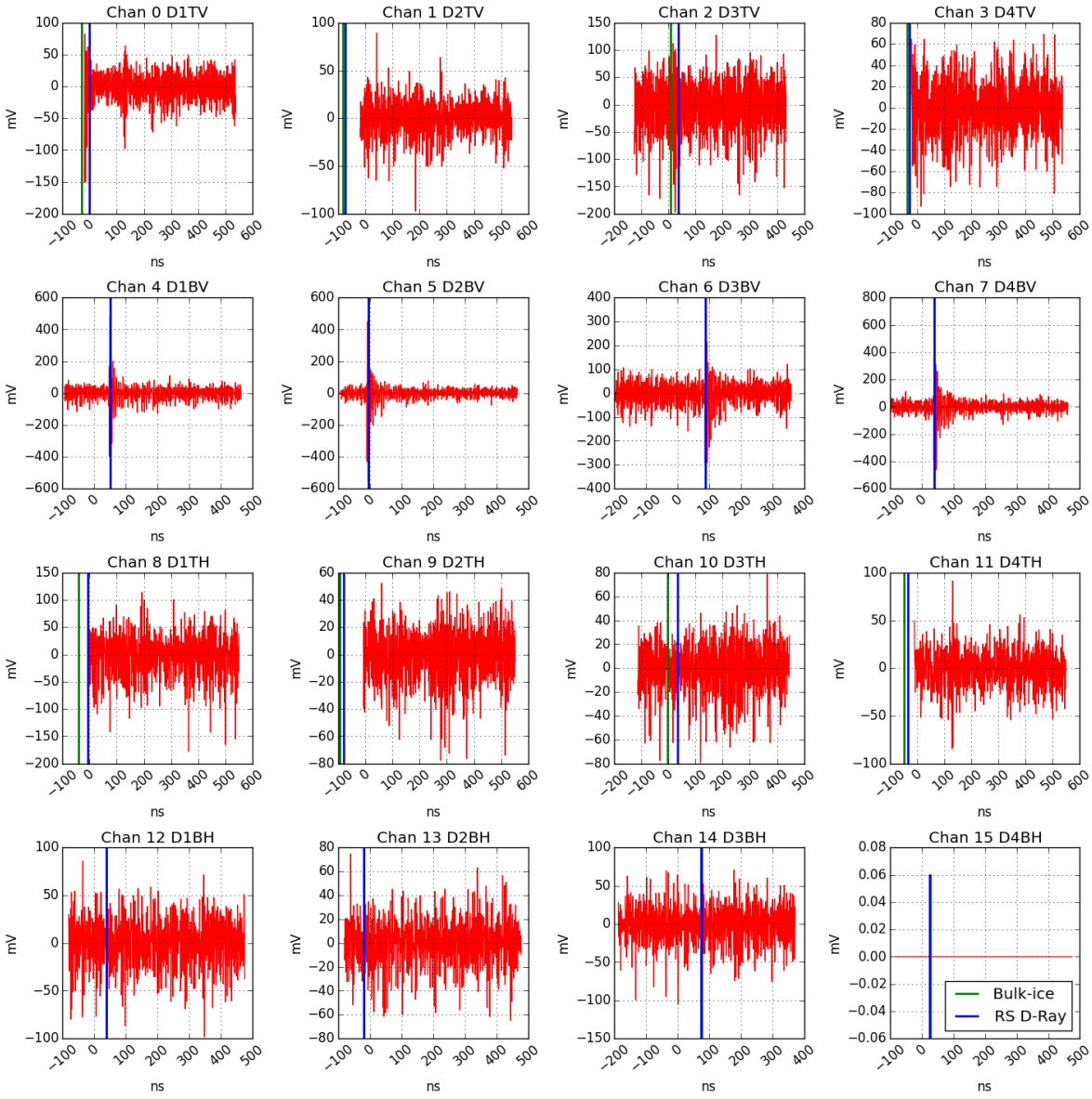


Figure 5.42: A2 run 3412 event 2959 with the first block of samples in each channel restored. Given the max-coherence reconstructed source direction using only the BVs, the signal arrival times with the quasi-planewave propagation hypothesis (green, "Bulk-ice") and with the in-ice `Radiospline` propagation hypothesis (blue, "RS D-Ray") are drawn as vertical lines in each channel. The pulse in the first block of channel 0 arrives at a time consistent with that predicted by either hypothesis, corroborating the conclusion that the signal is down-going.

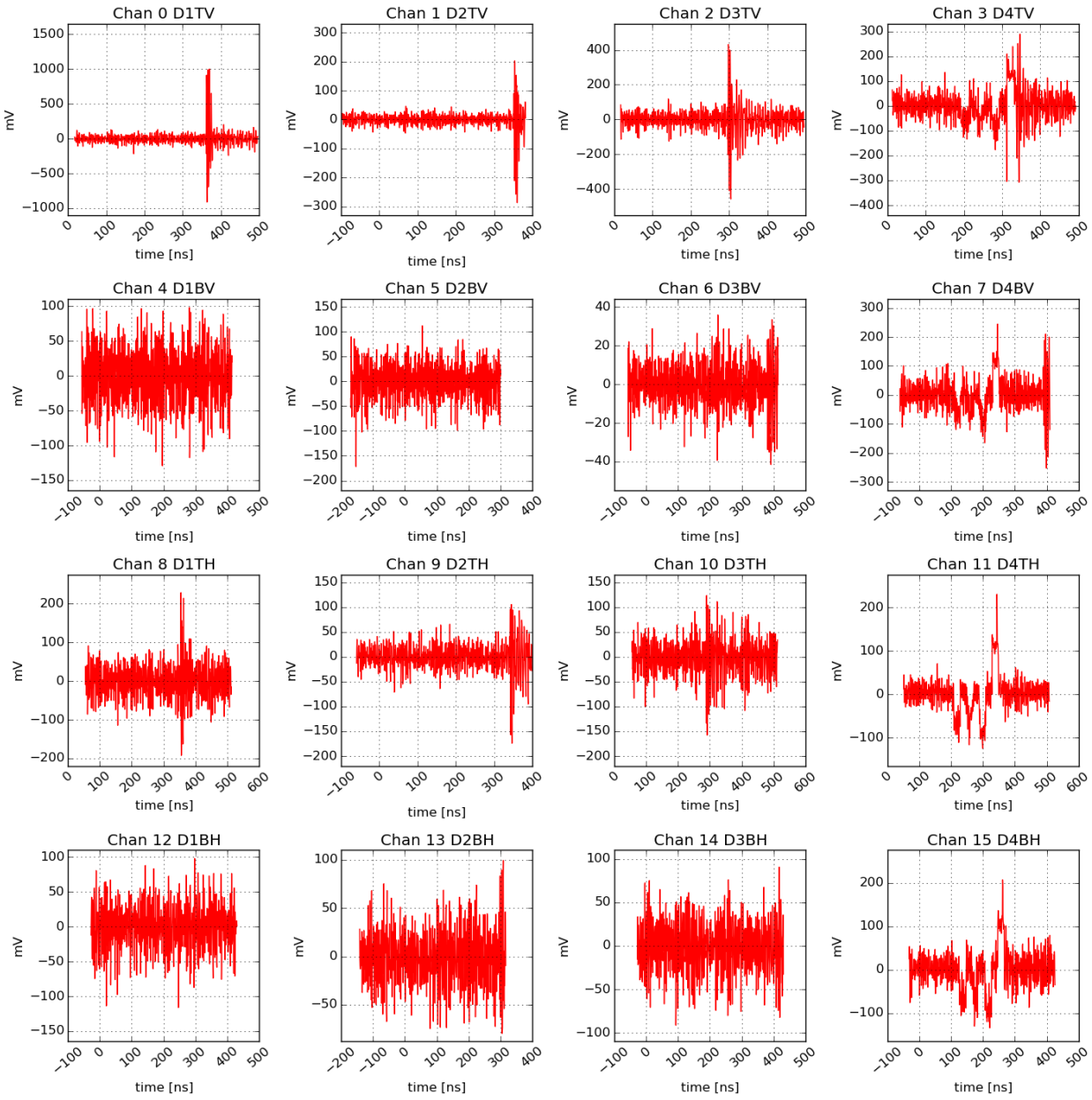


Figure 5.43: A3 run 6697 event 131864, a surviving event after unblinding. The TVs and the late pulse in D4BV reconstruct to a source angle above the surface cut. The extraction of the late pulse avoids the apparent problematic blocks of samples in D4BV

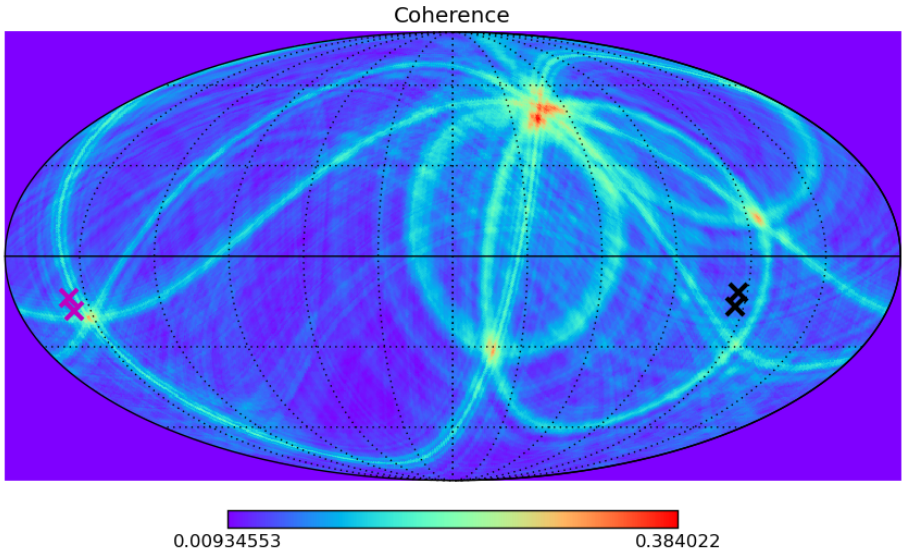


Figure 5.44: A3 run 6687 event 131864 quasi-planewave reconstruction skymap with channels 0, 1, 2, 7. The waveform in channel 7 (D4BV) is truncated to only use the last 100 ns. The crosses mark the directions of the local calibration pulsers. This sky map indicates a source direction above the surface cut.

Table 5.24: Number of neutrinos expected from this analysis with various neutrino flux models. Ahlers 10%/100%: [56]. Kotera min.: [111], grey-shaded region lower bound. IceCube 2016: [43], latest published. $\Phi_{\rm astro} = 0.90$, $\gamma = 2.13$. IceCube 2019: [44], latest available. $\Phi_{\rm astro} = 1.44$, $\gamma = 2.28$.

Model	N_ν
Ahlers 10% p	0.03741
Ahlers 100% p	0.1857
Kotera min.	0.1856
IceCube ν_μ 2016	0.3510
IceCube ν_μ 2019	0.1241

In the absence of detection, in Fig. 5.45 we compute the differential 90% confidence level (CL) upper limit on the diffuse flux of neutrinos. The 90% confidence level upper limit on the flux $EF(E)$ in the energy bin of width $\mathrm{dlog}_{10}E$ and central energy E can be computed by Eq. 5.6:

$$EF(E) = \frac{UL(E)}{\Lambda(E)\ \ell n(10)\ \mathrm{dlog}_{10}E} \tag{5.6}$$

where UL(E) is the Feldman-Cousins upper limit [112] for zero measured events on a background of $(5 \pm 2) \times 10^{-2}$. The value is taken to be UL(E) = 2.44, as commonly observed in the literature. In the figure, the limit is evaluated with $\mathrm{dlog}_{10}E = 1$, corresponding to decade wide bins in energy. $\Lambda(E)$ is the total exposure of the instrument, which sums over stations and configurations, taking into account analysis efficiencies. It is calculated according to Eq. 5.5.

The upper limit on the diffuse neutrino flux at 10^{18} eV is $EF(E) = 5.6 \times 10^{-16}$ cm^{-2}s^{-1}sr^{-1}. In the energy range of 10^{17} to 10^{19} eV, ARA places the most stringent upper limit among radio neutrino detectors. By 2022, the accrued exposure will grant ARA enough sensitivity to probe different mainstream GZK models.

5.9 Systematic Uncertainties

In this section, the systematic uncertainties considered in the analysis are discussed. The impact of these systematics on $A_{\rm eff}$ are shown in Fig. 5.47. The systematic uncertainties are categorized broadly in two classes. The first class is associated with theoretical uncertainties surrounding the neutrino-nucleon cross section and Askaryan emission, and are shown in Fig. 5.47 as solid bands. The second class is associated with uncertainties in our understanding of the detection medium and our instrument. They are shown as dashed/dotted lines in Fig. 5.47. A caveat to the discussion here is that the systematics are currently being studied by the collaboration and the results presented here are preliminary. The uncertainties will need to

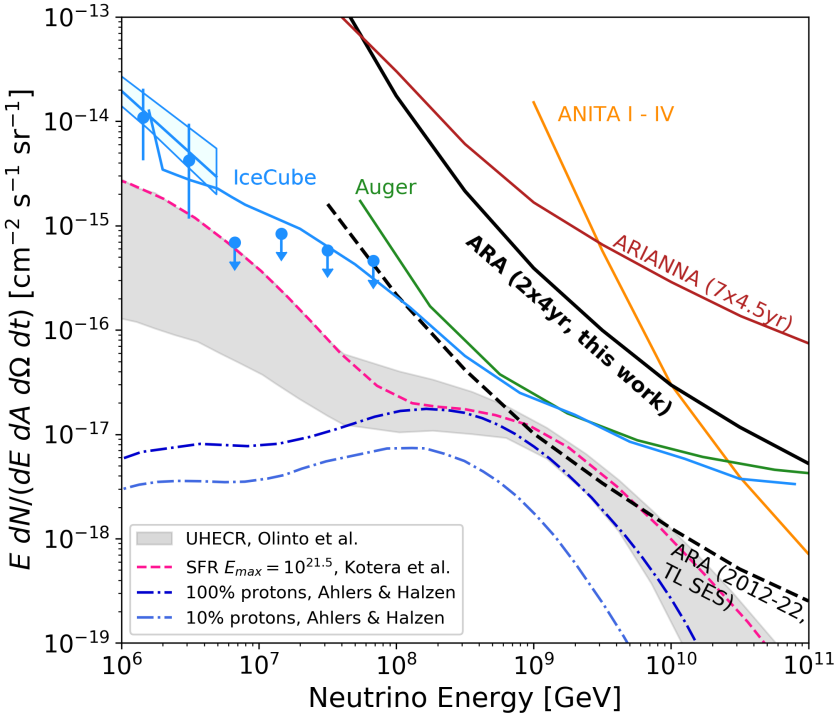


Figure 5.45: The 90% confidence-level upper limit on the diffuse flux of neutrinos set by this analysis (thick black line) [95]. The projected trigger-level single-event sensitivity for the five-station ARA5 array by 2022 is also presented as a black-dashed curve. Also shown are the latest limits and flux measurements from IceCube [43, 58], Auger [113], ANITA [70], and ARIANNA [114]. Shown for comparison are several benchmark cosmogenic neutrino flux models [56, 115, 116].

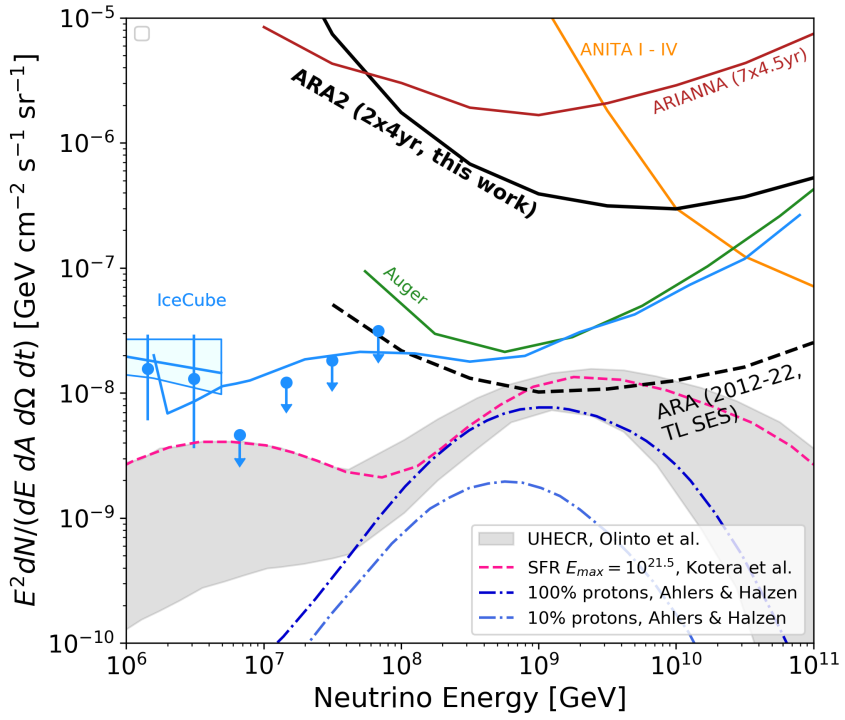


Figure 5.46: Same figure as Fig. 5.45, but scaled by an additional factor of energy.

be propagated to the analysis level, and treated as a correction to the upper limit with adequate statistical tools. Additionally, the study is a combined effort by Brian Clark, Jorge Torres, and me. It is presented in this thesis for completeness.

For the neutrino-nucleon cross section ($\sigma_{\nu-N}$), `AraSim` uses the model derived by Connolly *et al.* [99]. The upper and lower bounds for $\sigma_{\nu-N}$ are substituted for the central value in simulation to estimate the effect of the uncertainty on the simulated A_{eff} at the trigger level. Due to a lack of accelerator measurements, $\sigma_{\nu-N}$ is poorly constrained at very high energies, and so the associated uncertainties are large, and grow as a function of energy, exceeding 100% above 10^{21} eV. At 10^{18} eV the uncertainties due to the cross-section are estimated at -15%/+18%.

For the Askaryan emission, `AraSim` implements a modified version of the model derived by Alvarez-Muñiz *et. al.*, as described in section 3.2. The primary differences arise in `AraSim`'s non-inclusion of the LPM effect and in use of functional parameterizations for the shower profile instead of direct simulated shower profiles. The relative difference between waveform amplitudes produced by `AraSim`, and those derived from a full shower Monte Carlo are at most $\sim$ 12% [64]. The effect of this systematic uncertainty is conservatively estimated by reducing or increasing the simulated field amplitude by $\pm$12% and assessing the change in A_{eff} at the trigger level. The relative difference between the default parameterization and the scaled parameterization has a maximum value of about 25% near 10^{16} eV, and starts falling as energy increases. This is because at high energies the instrument acceptance becomes dominated by geometric effects (ray-tracing, etc.) and not signal amplitude. At 10^{18} eV the estimated uncertainties due to the Askaryan emission model are -11%/+13%.

In the second category of uncertainties, those arising from our detector response and from measurements of quantities such as the index of refraction in ice and the attenuation length of radio waves in ice are considered. These systematics will be included in the calculation of the final limit when the analysis-level effects become available in the future. Here, we consider uncertainties associated with (1) the attenuation length (L_α) of South Pole Ice and (2) the depth-dependent index of refraction ($n(z)$) of South Pole ice, (3) the calibration of the ARA signal chain, and (4) the triggering efficiency of the detector.

The model for the attenuation length (L_α) of South Pole ice was derived from data taken with the ARA Testbed prototype as described in section 1.4.2.2. Confidence bands providing an upper and lower limit on L_α are given in the model. To set upper/lower limits on our sensitivity, in `AraSim`, the upper and lower bounds for L_α are substituted for the central value. At 10^{18} eV the uncertainty on the effective area due to uncertainties in attenuation length are -13%/+44%.

The model for the depth-dependent index of refraction $n(z)$ was obtained by fitting data obtained by the RICE experiment, as described in section 1.4.2.2. We recalculate the sensitivity, setting all parameters to their upper and lower limits simultaneously. The lower (upper) limit generally corresponds to a slower

(faster) transition from surface to deep ice, and correspondingly have a smaller (larger) geometric acceptance for neutrinos. Additionally, since the ice-model assumption used to reconstruct the incoming direction of the RF emission is not changed, this systematic uncertainty also captures errors which may be present if the true ice model for radio wave propagation does not match that used for reconstruction. At 10^{18} eV the uncertainties on the effective area due to the index of refraction model are -4%/+4%.

We consider four sources of uncertainties that exist in the signal chain. They are the transmission coefficient t, the ambient noise power received N_{ant}, the signal chain noise power N_{sc}, and the antenna directivity D. The treatment used in the previous ARA result [94] is followed where we consider the system signal-to-noise ratio representing the ratio of input signal power to total system noise power in a given channel:

$$\text{SNR}_{\text{sys}} = \frac{tDP_{\text{sig}}}{tN_{\text{ant}} + N_{\text{sc}}} \tag{5.7}$$

with P_{sig} being the received signal power. The four sources of uncertainty translate to an uncertainty in SNR_{sys} by standard error propagation, which is then implemented as an uncertainty in the antenna gain G in code ($\Delta G = \Delta\text{SNR}_{\text{sys}}/P_{\text{sig}}$). In line with previous ARA work, only the case where the effective gain of the instrument is reduced is considered, providing a conservative estimate of our sensitivity. This is done due to the lack of sufficient calibration data at this time to constrain the upper bound on the gain. The Vpol antenna gain has an overall estimated uncertainty of -10%, while the Hpol antenna gain is estimated at -32%. The modified gain values are substituted in the simulation to assess the impact of this uncertainty, and the uncertainty of the effective area at 10^{18} eV is found to be -14%.

For the systematic uncertainty associated with the trigger efficiency of the detector as a function of SNR, ϵ(SNR), the simulated trigger efficiency ϵ_{sim}(SNR) is compared to the measured trigger efficiency in calibration pulser data ϵ_{dat}(SNR): $\Delta\epsilon = \epsilon_{\text{dat}}(\text{SNR}) - \epsilon_{\text{sim}}(\text{SNR})$. ϵ_{dat}(SNR) is measured by varying a tunable attenuator on the local calibration pulsers and counting the number of calibration pulser events recorded. Using `AraSim`, it can be found that the uncertainties on the trigger efficiency decreases the simulated effective area from between 2-5% depending on energy, and at 10^{18} eV the size of the effect is 3%.

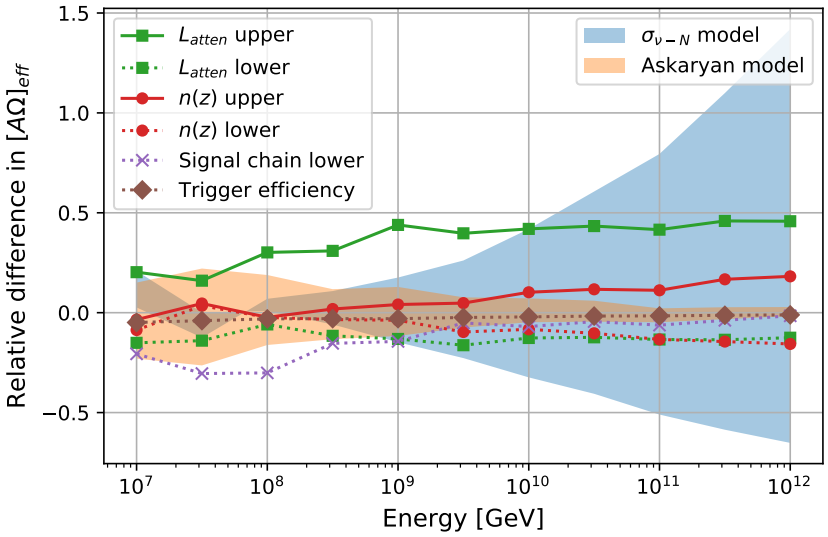


Figure 5.47: Trigger-level uncertainties between the central values used in the simulation and upper/lower bounds for each model parameters. Theoretical systematics: shaded region. Uncertainties associated with the detector and medium: dashed and solid lines. $[A\Omega]_{eff}$ denotes $4\pi A_{\text{eff}}$

Chapter 6

Conclusion And Outlook

The thesis is summarized in this chapter. Specifically, the tools and approaches that are developed for the neutrino search will be discussed, and potential improvements are pointed out. The scientific implication of the search result will also be explored. Finally, an outlook on future radio neutrino detectors is presented.

6.1 Conclusion

In this thesis, the scientific case for the quest for cosmogenic neutrinos is made. The unique role played by UHE neutrinos is stressed as a key component to the multimessenger approach towards unraveling the mystery of UHECR. Among different detection channels, radio techniques taking advantage of the Askaryan signal generated from neutrino-induced showers in radio-transparent ice proves to be a relatively low-cost but highly promising approach. ARA is built based upon such a detection principle, and is under a phased construction at the South Pole. The ARA station is a standalone, fully operative neutrino detector that is capable of measuring the Askaryan signal with multiple radio antennas that read out the signal electric fields with up to several GHz of sampling speed. The design of the antennas and the DAQ system is optimized for sensitivity in the relevant frequency range of 150 to 850 MHz, and allows for a degree-level directional reconstruction and the ability to differentiate between backgrounds and signals. An end-to-end Monte Carlo simulation software package `AraSim` is introduced as the main tool for generating neutrino simulation and estimating the detector sensitivity. An interferometric technique is developed that takes into account the ray-bending nature of the propagation of electromagnetic waves in the Antarctic ice due to a depth-dependent index of refraction. This technique is shown to successfully reconstruct simulated neutrinos and multiple in-ice or surface calibration sources to a desired level of precision. Using four years of data from two ARA stations, a blind analysis is performed to search for diffuse UHE neutrinos. The interferometric reconstruction is applied to events in the data, and its result and other waveform features are used to form a set of event selection and cuts that remove backgrounds and attain reasonable neutrino signal efficiency. Eventually, the full dataset is unblinded, and for each station we find zero signal events with an expected background level of 5×10^{-2} events. Since no neutrino candidate is detected, a 90% CL

upper limit is placed. This upper limit is the result of more than quadruple the live time of the previously published ARA analysis, and has the most constraining power in the energy range of 10^{17} to 10^{19} eV among all radio neutrino detectors. An expanded 5-station array will soon be able to probe the cosmogenic flux level predicted by realistic theoretical models. By 2022, it will have accumulated approximately four times the exposure of what this analysis has.

6.2 Outlook

The search motivates several potential improvements from which future experimenters may benefit:

1. More streamlined and centralized management of the time-dependent variation in station operation. This pertains to the management of the various configuration changes and special runs during the station live time. Historically these variations are poorly documented and hard to trace, hindering the setup and adjustments to analysis algorithms so that they can be applicable through the variations. A more streamlined and centralized control over the relevant station parameters will help significantly reduce the amount of overhead needed to analyze data. This can be extended to the firmware and software space as well.

2. Vetoing anthropogenic activity from the South Pole inl Austral summers. The analysis requires aggressive cuts on down-going signals, which eliminated 10-30% of neutrino events. Improvements to reconstruction algorithms to more confidently reject down-going events without requiring such substantial cuts on solid angle will improve the analysis efficiencies in the future. In addition, a better understanding of the radio transmission properties in the air-ice transition region will be very valuable. As the analysis showed, the events arrive below the critical angle, and the reconstruction of events in this region is sometime difficult. Being able to properly model the ray path in this region will eventually allow more sophisticated surface rejection criteria.

3. Taking full advantage of the physics-rich double-pulse events. Double pulse events arising from the capture of both the 1st and the 2nd ray are unique events that provide extra handles to properties such as the vertex position and the neutrino direction. In the current analysis, no algorithm is deployed to specifically search for these valuable events, and the reconstruction does not take into account the co-existence of the two rays.

4. More sophisticated event filter. In this analysis the event filter is a simple threshold-passing test for the event SNR, which severely impact the signal efficiency especially at lower energies. More sophisticated filters taking into account the timing pattern or the frequency content of an event may be able to reach the same level of noise rejection while still be sensitive to lower energy events.

The field is rapidly moving towards the next generation of radio neutrino detectors. Beyond the current ARA design, some of the latest developments are:

1. Phased array trigger [117]. The phased array technique relies on coherently summing the impulsive Askaryan signal over multiple receiving antennas. This promotes the SNR of events, and effectively lowers the energy threshold of the detector. This technique could prove to be especially valuable in the regard that it may fill the gap of sensitivity around 10 PeV and above between the IceCube astrophysical flux and radio detectors. In the latest A5 station, the phased trigger has already been implemented using 10 downhole Vpol antennas in a dedicated central string that provides an additional trigger stream external to the default multiplicity trigger.

2. Envelope detector and reconstruction. Motivated by the fact that interferometric reconstruction so far in ARA requires the use of Hilbert envelopes, new efforts have been devoted to the development of envelope detectors and reconstruction. An envelope waveforms does not require as high a digitization speed, which helps lower both cost and power consumption. This makes the technique more scalable. Whether the ability to reconstruct events with envelope waveforms is comparable with current techniques or not is being studied, as the loss of timing resolution and full spectral information could result in difficulties at the analysis level.

LIST OF REFERENCES

[1] V.F. Hess. Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten. *Phys. Z.*, 13:1084—-1091, 1912.

[2] Carl D. Anderson. The Positive Electron. *Physical Review*, 43(6):491–494, Mar 1933. doi:10.1103/PhysRev.43.491.

[3] Seth H. Neddermeyer and Carl D. Anderson. Note on the nature of cosmic-ray particles. *Phys. Rev.*, 51:884–886, May 1937. URL: https://link.aps.org/doi/10.1103/PhysRev.51.884, doi:10.1103/PhysRev.51.884.

[4] D H PERKINS. Nuclear Disintegration by Meson Capture. *Nature*, 159(4030):126–127, 1947. URL: https://doi.org/10.1038/159126a0, doi:10.1038/159126a0.

[5] G P S OCCHIALINI and C F POWELL. NUCLEAR DISINTEGRATIONS PRODUCED BY SLOW CHARGED PARTICLES OF SMALL MASS. *Nature*, 159(4032):186–190, 1947. URL: https://doi.org/10.1038/159186a0, doi:10.1038/159186a0.

[6] A. M. Hillas. Cosmic Rays: Recent Progress and some Current Questions. *arXiv e-prints*, pages astro–ph/0607109, Jul 2006. arXiv:astro-ph/0607109.

[7] ENRICO Fermi. On the origin of the cosmic radiation. *Phys. Rev.*, 75:1169–1174, Apr 1949. URL: https://link.aps.org/doi/10.1103/PhysRev.75.1169, doi:10.1103/PhysRev.75.1169.

[8] James J. Beatty and Stefan Westerhoff. The highest-energy cosmic rays. *Annual Review of Nuclear and Particle Science*, 59(1):319–345, 2009. URL: https://doi.org/10.1146/annurev.nucl.58.110707.171154, arXiv:https://doi.org/10.1146/annurev.nucl.58.110707.171154, doi:10.1146/annurev.nucl.58.110707.171154.

[9] Kenneth Greisen. End to the cosmic-ray spectrum? *Phys. Rev. Lett.*, 16:748–750, Apr 1966. URL: https://link.aps.org/doi/10.1103/PhysRevLett.16.748, doi:10.1103/PhysRevLett.16.748.

[10] G. T. Zatsepin and V. A. Kuzmin. Upper limit of the spectrum of cosmic rays. *JETP Lett.*, 4:78–80, 1966. [Pisma Zh. Eksp. Teor. Fiz.4,114(1966)].

[11] P Sokolsky and G B Thomson. Highest energy cosmic-rays and results from the HiRes experiment. *Journal of Physics G: Nuclear and Particle Physics*, 34(11):R401–R429, sep 2007. URL: https://doi.org/10.1088%2F0954-3899%2F34%2F11%2Fr01, doi:10.1088/0954-3899/34/11/r01.

[12] D. Ivanov et al. (Telescope Array Collab.). Energy spectrum measured by the telescope array. *Proceedings of Science (ICRC2019)*, page 298, 2019.

[13] V. Verzi et al. (Auger Collab.). Measurement of the energy spectrum of ultra-high energy cosmic rays using the pierre auger observatory. *Proceedings of Science (ICRC2019)*, page 450, 2019.

[14] A. A. Abdo, B. T. Allen, T. Aune, D. Berley, S. Casanova, C. Chen, B. L. Dingus, R. W. Ellsworth, L. Fleysher, R. Fleysher, M. M. Gonzalez, J. A. Goodman, C. M. Hoffman, B. Hopper, P. H. Hntemeyer, B. E. Kolterman, C. P. Lansdell, J. T. Linnemann, J. E. McEnery, A. I. Mincer, P. Nemethy, D. Noyes, J. Pretz, J. M. Ryan, P. M. Saz Parkinson, A. Shoup, G. Sinnis, A. J. Smith, G. W. Sullivan, V. Vasileiou, G. P. Walker, D. A. Williams, and G. B. Yodh. THE LARGE-SCALE COSMIC-RAY ANISOTROPY AS OBSERVED WITH MILAGRO. *The Astrophysical Journal*, 698(2):2121–2130, jun 2009. URL: https://doi.org/10.1088%2F0004-637x%2F698%2F2%2F2121, doi:10.1088/0004-637x/698/2/2121.

[15] X. X. Zhou et al. ON TEMPORAL VARIATIONS OF THE MULTI-TeV COSMIC RAY ANISOTROPY USING THE TIBET III AIR SHOWER ARRAY. *The Astrophysical Journal*, 711(1):119–124, feb 2010. URL: https://doi.org/10.1088%2F0004-637x%2F711%2F1%2F119, doi:10.1088/0004-637x/711/1/119.

[16] A. U. Abeysekara et al. All-sky measurement of the anisotropy of cosmic rays at 10 TeV and mapping of the local interstellar magnetic field. *The Astrophysical Journal*, 871(1):96, jan 2019. URL: https://doi.org/10.3847%2F1538-4357%2Faaf5cc, doi:10.3847/1538-4357/aaf5cc.

[17] A. Yushkov et al. (Auger Collab.). Mass composition of cosmic rays with energies above $10^{17.2}$ ev from the hybrid data of the pierre auger observatory. *Proceedings of Science (ICRC2019)*, page 482, 2019.

[18] W. Hanlon et al. (Telescope Array Collab.). Telescope array 10 year composition. *Proceedings of Science (ICRC2019)*, page 280, 2019.

[19] William Hanlon. Telescope Array Hybrid Composition and Auger-TA Composition Comparison. *EPJ Web Conf.*, 208:02001, 2019. arXiv:1810.00431, doi:10.1051/epjconf/201920802001.

[20] W. Pauli. Dear radioactive ladies and gentlemen. *Phys. Today*, 31N9:27, 1978.

[21] F. Reines. The neutrino: From poltergeist to particle. *Rev. Mod. Phys.*, 68:317–327, 1996. doi:10.1103/RevModPhys.68.317.

[22] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire. Detection of the free neutrino: A confirmation. *Science*, 124(3212):103–104, 1956. URL: http://www.jstor.org/stable/1751492.

[23] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562, doi:10.1103/PhysRevLett.81.1562.

[24] Shouvik Roy Choudhury and Sandhya Choubey. Updated Bounds on Sum of Neutrino Masses in Various Cosmological Scenarios. *JCAP*, 1809(09):017, 2018. arXiv:1806.10832, doi:10.1088/1475-7516/2018/09/017.

[25] M. Aker et al. An improved upper limit on the neutrino mass from a direct kinematic method by KATRIN. *Phys. Rev. Lett.*, 123(22):221802, 2019. arXiv:1909.06048, doi:10.1103/PhysRevLett.123.221802.

[26] Ali Ashtari Esfahani et al. Determining the neutrino mass with cyclotron radiation emission spectroscopyProject 8. *J. Phys.*, G44(5):054004, 2017. arXiv:1703.02037, doi:10.1088/1361-6471/aa5b4f.

[27] Z. Maki, M. Nakagawa, and S. Sakata. Remarks on the Unified Model of Elementary Particles. *Progress of Theoretical Physics*, 28(5):870–880, Nov 1962. doi:10.1143/PTP.28.870.

[28] Raymond Davis, Don S. Harmer, and Kenneth C. Hoffman. Search for neutrinos from the sun. *Phys. Rev. Lett.*, 20:1205–1209, May 1968. URL: https://link.aps.org/doi/10.1103/PhysRevLett.20.1205, doi:10.1103/PhysRevLett.20.1205.

[29] K. Hirata, T. Kajita, M. Koshiba, M. Nakahata, Y. Oyama, N. Sato, A. Suzuki, M. Takita, Y. Totsuka, T. Kifune, T. Suda, K. Takahashi, T. Tanimori, K. Miyano, M. Yamada, E. W. Beier, L. R. Feldscher, S. B. Kim, A. K. Mann, F. M. Newcomer, R. Van, W. Zhang, and B. G. Cortez. Observation of a neutrino burst from the supernova sn1987a. *Phys. Rev. Lett.*, 58:1490–1493, Apr 1987. URL: https://link.aps.org/doi/10.1103/PhysRevLett.58.1490, doi:10.1103/PhysRevLett.58.1490.

[30] R. M. Bionta, G. Blewitt, C. B. Bratton, D. Casper, A. Ciocio, R. Claus, B. Cortez, M. Crouch, S. T. Dye, S. Errede, G. W. Foster, W. Gajewski, K. S. Ganezer, M. Goldhaber, T. J. Haines, T. W. Jones, D. Kielczewska, W. R. Kropp, J. G. Learned, J. M. LoSecco, J. Matthews, R. Miller, M. S. Mudan, H. S. Park, L. R. Price, F. Reines, J. Schultz, S. Seidel, E. Shumard, D. Sinclair, H. W. Sobel, J. L. Stone, L. R. Sulak, R. Svoboda, G. Thornton, J. C. van der Velde, and C. Wuest. Observation of a neutrino burst in coincidence with supernova 1987a in the large magellanic cloud. *Phys. Rev. Lett.*, 58:1494–1496, Apr 1987. URL: https://link.aps.org/doi/10.1103/PhysRevLett.58.1494, doi:10.1103/PhysRevLett.58.1494.

[31] E. N. Alexeyev, L. N. Alexeyeva, I. V. Krivosheina, and V. I. Volchenko. Detection of the neutrino signal from SN 1987A in the LMC using the INR Baksan underground scintillation telescope. *Physics Letters B*, 205(2-3):209–214, Apr 1988. doi:10.1016/0370-2693(88)91651-6.

[32] K. Nomoto and T. Shigeyama. Supernova 1987A: constraints on the theoretical model. In Minas Kafatos and Andrew G. Michalitsianos, editors, *Supernova 1987A in the Large Magellanic Cloud*, pages 273–288, Jan 1988.

[33] Donald K. Perkins. *Introduction to High Energy Physics*. Cambridge University Press, 7 2000.

[34] John G Learned and Sandip Pakvasa. Detecting ν_τ oscillations at pev energies. *Astroparticle Physics*, 3(3):267 – 274, 1995. URL: http://www.sciencedirect.com/science/article/pii/0927650594000433, doi:https://doi.org/10.1016/0927-6505(94)00043-3.

[35] H. ATHAR, C. S. KIM, and JAKE LEE. Intrinsic and oscillated astrophysical neutrino flavor ratios revisited. *Modern Physics Letters A*, 21(13):1049–1065, 2006. URL: https://doi.org/10.1142/S021773230602038X, arXiv:https://doi.org/10.1142/S021773230602038X, doi:10.1142/S021773230602038X.

[36] S. Hümmer, M. Rüger, F. Spanier, and W. Winter. SIMPLIFIED MODELS FOR PHOTOHADRONIC INTERACTIONS IN COSMIC ACCELERATORS. *The Astrophysical Journal*, 721(1):630–652, aug 2010. URL: https://doi.org/10.1088%2F0004-637x%2F721%2F1%2F630, doi:10.1088/0004-637x/721/1/630.

[37] "S. Hümmer, M. Maltoni, W. Winter, and C. Yaguna". Energy dependent neutrino flavor ratios from cosmic accelerators on the hillas plot. *Astroparticle Physics*, 34(4):205 – 224, 2010. URL: http://www.sciencedirect.com/science/article/pii/S092765051000126X, doi:https://doi.org/10.1016/j.astropartphys.2010.07.003.

[38] P. Meszaros. Astrophysical sources of high-energy neutrinos in the icecube era. *Annual Review of Nuclear and Particle Science*, 67(1):45–67, 2017. URL: https://doi.org/10.1146/annurev-nucl-101916-123304, arXiv:https://doi.org/10.1146/annurev-nucl-101916-123304, doi:10.1146/annurev-nucl-101916-123304.

[39] Kohta Murase and Imre Bartos. High-energy multimessenger transient astrophysics. *Annual Review of Nuclear and Particle Science*, 69(1):477–506, 2019. URL: https://doi.org/10.1146/annurev-nucl-101918-023510, arXiv:https://doi.org/10.1146/annurev-nucl-101918-023510, doi:10.1146/annurev-nucl-101918-023510.

[40] Francis Halzen. High-energy neutrino astrophysics. *Nature Physics*, 13(3):232–238, 2017. URL: https://doi.org/10.1038/nphys3816, doi:10.1038/nphys3816.

[41] M.G. Aartsen et al. The IceCube neutrino observatory: instrumentation and online systems. *Journal of Instrumentation*, 12(03):P03012–P03012, mar 2017. URL: https://doi.org/10.1088%2F1748-0221%2F12%2F03%2Fp03012, doi:10.1088/1748-0221/12/03/p03012.

[42] M.G. Aartsen et al. Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data. *Phys. Rev. Lett.*, 113:101101, 2014. arXiv:1405.5303, doi:10.1103/PhysRevLett.113.101101.

[43] M. G. Aartsen et al. OBSERVATION AND CHARACTERIZATION OF a COSMIC MUON NEUTRINO FLUX FROM THE NORTHERN HEMISPHERE USING SIX YEARS OF ICECUBE DATA. *The Astrophysical Journal*, 833(1):3, dec 2016. URL: https://doi.org/10.3847%2F0004-637x%2F833%2F1%2F3, doi:10.3847/0004-637x/833/1/3.

[44] D.R. Williams et al. (IceCube Collab.). Results from icecube. *Proceedings of Science (ICRC2019)*, page 016, 2019.

[45] T. Carver et al. (IceCube Collab.). Ten years of all-sky neutrino point-source searches. *Proceedings of Science (ICRC2019)*, page 851, 2019.

[46] M. G. Aartsen et al. The icecube realtime alert system. *Astropart. Phys.*, 92:30–41, 2017. arXiv:1612.06028, doi:10.1016/j.astropartphys.2017.05.002.

[47] M. W. E. Smith et al. The Astrophysical Multimessenger Observatory Network (AMON). *Astropart. Phys.*, 45:56–70, 2013. arXiv:1211.5602, doi:10.1016/j.astropartphys.2013.03.003.

[48] S. D. Barthelmy, P. Butterworth, T. L. Cline, N. Gehrels, F. Marshall, T. Takeshima, V. Connaughton, R. M. Kippen, C. Kouveliotou, and C. R. Robinson. The grb coordinates network (gcn): A status report. *AIP Conference Proceedings*, 428(1):99–103, 1998. URL: https://aip.scitation.org/doi/abs/10.1063/1.55426, arXiv:https://aip.scitation.org/doi/pdf/10.1063/1.55426, doi:10.1063/1.55426.

[49] M.G. Aartsen et al. Multimessenger observations of a flaring blazar coincident with high-energy neutrino icecube-170922a. *Science*, 361(6398), 2018. URL: https://science.sciencemag.org/content/361/6398/eaat1378, arXiv:https://science.sciencemag.org/content/361/6398/eaat1378.full.pdf, doi:10.1126/science.aat1378.

[50] M. Ajello et al. 3fhl: The third catalog of HardFermi-LAT sources. *The Astrophysical Journal Supplement Series*, 232(2):18, sep 2017. URL: https://doi.org/10.3847%2F1538-4365%2Faa8221, doi:10.3847/1538-4365/aa8221.

[51] F. Acero et al. (Fermi-LAT Collaboration). Fermi Large Area Telescope Third Source Catalog. *Astrophysical Journal Supplements*, 218(2):23, Jun 2015. arXiv:1501.02003, doi:10.1088/0067-0049/218/2/23.

[52] V. S. Beresinsky and G. T. Zatsepin. Cosmic rays at ultra high energies (neutrino?). *Physics Letters B*, 28(6):423–424, Jan 1969. doi:10.1016/0370-2693(69)90341-4.

[53] Ralph Engel, David Seckel, and Todor Stanev. Neutrinos from propagation of ultrahigh energy protons. *Phys. Rev. D*, 64:093010, Oct 2001. URL: https://link.aps.org/doi/10.1103/PhysRevD.64.093010, doi:10.1103/PhysRevD.64.093010.

[54] J. P. Rachen. *Interaction Processes and Statistical Properties of the Propagation of Cosmic Rays in Photon Backgrounds*. PhD thesis, MPIfR Bonn, Germany, 1996.

[55] F. W. Stecker and M. H. Salamon. Photodisintegration of ultra–high-energy cosmic rays: A new determination. *The Astrophysical Journal*, 512(2):521–526, feb 1999. URL: https://doi.org/10.1086%2F306816, doi:10.1086/306816.

[56] Markus Ahlers and Francis Halzen. Minimal cosmogenic neutrinos. *Phys. Rev. D*, 86:083010, Oct 2012. URL: https://link.aps.org/doi/10.1103/PhysRevD.86.083010, doi:10.1103/PhysRevD.86.083010.

[57] F. Pedreira et al. (Auger Collab.). Bounds on diffuse and point source fluxes of ultra-high energy neutrinos with the pierre auger observatory. *Proceedings of Science (ICRC2019)*, page 979, 2019.

[58] M. G. Aartsen et al. Differential limit on the extremely-high-energy cosmic neutrino flux in the presence of astrophysical background from nine years of IceCube data. *Phys. Rev.*, D98(6):062003, 2018. arXiv:1807.01820, doi:10.1103/PhysRevD.98.062003.

[59] G. A. Askar'yan. Excess negative charge of an electron-photon shower and its coherent radio emission. *Sov. Phys. JETP*, 14(2):441–443, 1962. [Zh. Eksp. Teor. Fiz.41,616(1961)].

[60] G. A. Askar'yan. Coherent radio emission from cosmic showers in air and in dense media. *Sov. Phys. JETP*, 48:988–990, 1965.

[61] E. Zas, F. Halzen, and T. Stanev. Electromagnetic pulses from high-energy showers: Implications for neutrino detection. *Phys. Rev. D*, 45:362–376, Jan 1992. URL: https://link.aps.org/doi/10.1103/PhysRevD.45.362, doi:10.1103/PhysRevD.45.362.

[62] Jaime Alvarez-Muñiz, Andrés Romero-Wolf, and Enrique Zas. Čerenkov radio pulses from electromagnetic showers in the time domain. *Phys. Rev. D*, 81:123009, Jun 2010. URL: https://link.aps.org/doi/10.1103/PhysRevD.81.123009, doi:10.1103/PhysRevD.81.123009.

[63] Jaime Alvarez-Muñiz, Andrés Romero-Wolf, and Enrique Zas. Practical and accurate calculations of askaryan radiation. *Phys. Rev. D*, 84:103003, Nov 2011. URL: https://link.aps.org/doi/10.1103/PhysRevD.84.103003, doi:10.1103/PhysRevD.84.103003.

[64] E. Hong. *Searching for Ultra-high Energy Neutrinos with Data from a Prototype Station of the Askaryan Radio Array*. PhD thesis, The Ohio State University, USA, 2014.

[65] David Saltzberg, Peter Gorham, Dieter Walz, Clive Field, Richard Iverson, Allen Odian, George Resch, Paul Schoessow, and Dawn Williams. Observation of the askaryan effect: Coherent microwave cherenkov emission from charge asymmetry in high-energy particle cascades. *Phys. Rev. Lett.*, 86:2802–2805, Mar 2001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.86.2802, doi:10.1103/PhysRevLett.86.2802.

[66] P. W. Gorham, D. Saltzberg, R. C. Field, E. Guillian, R. Milinčić, P. Miočinović, D. Walz, and D. Williams. Accelerator measurements of the askaryan effect in rock salt: A roadmap toward teraton underground neutrino detectors. *Phys. Rev. D*, 72:023002, Jul 2005. URL: https://link.aps.org/doi/10.1103/PhysRevD.72.023002, doi:10.1103/PhysRevD.72.023002.

[67] P. W. Gorham, S. W. Barwick, J. J. Beatty, D. Z. Besson, W. R. Binns, C. Chen, P. Chen, J. M. Clem, A. Connolly, P. F. Dowkontt, M. A. DuVernois, R. C. Field, D. Goldstein, A. Goodhue, C. Hast, C. L. Hebert, S. Hoover, M. H. Israel, J. Kowalski, J. G. Learned, K. M. Liewer, J. T. Link, E. Lusczek, S. Matsuno, B. Mercurio, C. Miki, P. Miočinović, J. Nam, C. J. Naudet, J. Ng, R. Nichol, K. Palladino, K. Reil, A. Romero-Wolf, M. Rosen, L. Ruckman, D. Saltzberg, D. Seckel, G. S. Varner, D. Walz, and F. Wu. Observations of the askaryan effect in ice. *Phys. Rev. Lett.*, 99:171101, Oct 2007. URL: https://link.aps.org/doi/10.1103/PhysRevLett.99.171101, doi:10.1103/PhysRevLett.99.171101.

[68] Alexander Aab et al. Probing the radio emission from air showers with polarization measurements. *Phys. Rev.*, D89(5):052002, 2014. arXiv:1402.3677, doi:10.1103/PhysRevD.89.052002.

[69] P. W. Gorham et al. The Antarctic Impulsive Transient Antenna Ultra-high Energy Neutrino Detector Design, Performance, and Sensitivity for 2006-2007 Balloon Flight. *Astropart. Phys.*, 32:10–41, 2009. arXiv:0812.1920, doi:10.1016/j.astropartphys.2009.05.003.

[70] P. W. Gorham, P. Allison, O. Banerjee, L. Batten, J. J. Beatty, K. Belov, D. Z. Besson, W. R. Binns, V. Bugaev, P. Cao, C. H. Chen, P. Chen, Y. Chen, J. M. Clem, A. Connolly, L. Cremonesi, B. Dailey, C. Deaconu, P. F. Dowkontt, B. D. Fox, J. W. H. Gordon, C. Hast, B. Hill, S. Y. Hsu, J. J. Huang, K. Hughes, R. Hupe, M. H. Israel, K. M. Liewer, T. C. Liu, A. B. Ludwig, L. Macchiarulo, S. Matsuno, K. McBride, C. Miki, K. Mulrey, J. Nam, C. Naudet, R. J. Nichol, A. Novikov, E. Oberla, S. Prohira, B. F. Rauch, J. Ripa, J. M. Roberts, A. Romero-Wolf, B. Rotter, J. W. Russell, D. Saltzberg, D. Seckel, H. Schoorlemmer, J. Shiao, S. Stafford, J. Stockham, M. Stockham, B. Strutt, M. S. Sutherland, G. S. Varner, A. G. Vieregg, N. Wang, S. H. Wang, and S. A. Wissel. Constraints on the ultrahigh-energy cosmic neutrino flux from the fourth flight of anita. *Phys. Rev. D*, 99:122001, Jun 2019. URL: https://link.aps.org/doi/10.1103/PhysRevD.99.122001, doi:10.1103/PhysRevD.99.122001.

[71] S. W. Barwick et al. Design and Performance of the ARIANNA HRA-3 Neutrino Detector Systems. *IEEE Trans. Nucl. Sci.*, 62(5):2202–2215, 2015. arXiv:1410.7369, doi:10.1109/TNS.2015.2468182.

[72] E. Andres et al. Observation of high-energy neutrinos using Cherenkov detectors embedded deep in Antarctic ice. *Nature*, 410:441–443, 2001. doi:10.1038/35068509.

[73] V. V. Bogorodsky, C. R. Bentley, and P. E. Gudmandsen. *Radioglaciology*. D. Reidel Publisihing Company, 1985.

[74] Takeshi Matsuoka, Shuji Fujita, and Shinji Mae. Effect of temperature on dielectric properties of ice in the range 5-39 ghz. *Journal of Applied Physics*, 80(10):5884–5890, 1996. URL: https://doi.org/10.1063/1.363582, arXiv:https://doi.org/10.1063/1.363582, doi:10.1063/1.363582.

[75] Martin Rongen. *Calibration of the IceCube Neutrino Observatory*. PhD thesis, RWTH Aachen U., 2019. arXiv:1911.02016, doi:10.18154/RWTH-2019-09941.

[76] S. Fujita, T. Matsuoka, Y. Ishida, K. Matsuoka, and S. Mae. A summary of the complex dielectric permittivity of ice in the megahertz range and its applications for radar sounding of polar ice sheets. *Physics of Ice Core Records*, pages 185–212, 2000.

[77] S. Barwick, D. Besson, P. Gorham, and D. Saltzberg. South polar in situ radio-frequency ice attenuation. *Journal of Glaciology*, 51(173):231238, 2005. doi:10.3189/172756505781829467.

[78] P. Buford Price, Oleg V. Nagornov, Ryan Bay, Dmitry Chirkin, Yudong He, Predrag Miocinovic, Austin Richards, Kurt Woschnagg, Bruce Koci, and Victor Zagorodnov. Temperature profile for glacial ice at the south pole: Implications for life in a nearby subglacial lake. *Proceedings of the National Academy of Sciences*, 99(12):7844–7847, 2002. URL: https://www.pnas.org/content/99/12/7844, arXiv:https://www.pnas.org/content/99/12/7844.full.pdf, doi:10.1073/pnas.082238999.

[79] Ilya Kravchenko, David Besson, and Josh Meyers. In situ index-of-refraction measurements of the south polar firn with the rice detector. *Journal of Glaciology*, 50(171):522532, 2004. doi:10.3189/172756504781829800.

[80] K. C. Kuivinen, B. R. Koci, G. W. Holdsworth, and A. J. Gow. South pole ice core drilling, 1981-1982. *Antarctic Journal of the United States*, 17(5):89–91, 1982.

[81] M. Newcomb. https://icecube.wisc.edu/~araproject/radio/. Accessed: 2019-09-30.

[82] P. Allison et al. First Constraints on the Ultra-High Energy Neutrino Flux from a Prototype Station of the Askaryan Radio Array. *Astropart. Phys.*, 70:62–80, 2015. arXiv:1404.5285, doi:10.1016/j.astropartphys.2015.04.006.

[83] T. H. Hankins, R. D. Ekers, and J. D. O'Sullivan. A search for lunar radio erenkov emission from high-energy neutrinos. *Monthly Notices of the Royal Astronomical Society*, 283(3):1027–1030, 12 1996. URL: https://doi.org/10.1093/mnras/283.3.1027, arXiv:https://academic.oup.com/mnras/article-pdf/283/3/1027/3303825/283-3-1027.pdf, doi:10.1093/mnras/283.3.1027.

[84] P. W. Gorham, C. L. Hebert, K. M. Liewer, C. J. Naudet, D. Saltzberg, and D. Williams. Experimental limit on the cosmic diffuse ultrahigh energy neutrino flux. *Phys. Rev. Lett.*, 93:041101, Jul 2004. URL: https://link.aps.org/doi/10.1103/PhysRevLett.93.041101, doi:10.1103/PhysRevLett.93.041101.

[85] O. Scholten, S. Buitink, J. Bacelar, R. Braun, A. G. de Bruyn, H. Falcke, K. Singh, B. Stappers, R. G. Strom, and R. al Yahyaoui. Improved flux limits for neutrinos with energies above 10^{22} eV from observations with the westerbork synthesis radio telescope. *Phys. Rev. Lett.*, 103:191301, Nov 2009. URL: https://link.aps.org/doi/10.1103/PhysRevLett.103.191301, doi:10.1103/PhysRevLett.103.191301.

[86] C. W. James et al. The lunar askaryan technique with the square kilometre array. *Proceedings of Science (ICRC2015)*, page 291, 2015.

[87] J. D. Bray, R. D. Ekers, P. Roberts, J. E. Reynolds, C. W. James, C. J. Phillips, R. J. Protheroe, R. A. McFadden, and M. G. Aartsen. A lunar radio experiment with the Parkes radio telescope for the LUNASKA project. *Astroparticle Physics*, 65:22–39, May 2015. arXiv:1412.4418, doi:10.1016/j.astropartphys.2014.11.008.

[88] G.A. Askarian, B.A. Dolgoshein, A.N. Kalinovsky, and N.V. Mokhov. ACOUSTIC DETECTION OF HIGH-ENERGY PARTICLE SHOWERS IN WATER. *Nucl. Instrum. Meth.*, 164:267–278, 1979. doi:10.1016/0029-554X(79)90244-1.

[89] Robert Lahmann. Acoustic detection of neutrinos: Review and future potential. *Nuclear and Particle Physics Proceedings*, 273-275:406–413, 04 2016. doi:10.1016/j.nuclphysbps.2015.09.059.

[90] P. Allison, J. Auffenberg, R. Bard, J. J. Beatty, D. Z. Besson, S. Boeser, C. Chen, P. Chen, a. Connolly, J. Davies, M. DuVernois, B. Fox, P. W. Gorham, E. W. Grashorn, K. Hanson, J. Haugen, K. Helbing, B. Hill, K. D. Hoffman, M. Huang, M. H. a. Huang, a. Ishihara, a. Karle, D. Kennedy, H. Landsman, a. Laundrie, T. C. Liu, L. Macchiarulo, K. Mase, T. Meures, R. Meyhandan, C. Miki, R. Morse, M. Newcomb, R. J. Nichol, K. Ratzlaff, M. Richman, L. Ritter, B. Rotter, P. Sandstrom, D. Seckel, J. Touart, G. S. Varner, M. Z. Wang, C. Weaver, a. Wendorff, S. Yoshida, and R. Young. Design and Initial Performance of the Askaryan Radio Array Prototype EeV Neutrino Detector at the South Pole. page 25, may 2011. URL: http://arxiv.org/abs/1105.2854, arXiv:1105.2854, doi:10.1016/j.icarus.2012.05.028.

[91] J. A. Aguilar et al. The next-generation radio neutrino observatory — multi-messenger neutrino astrophysics at extreme energies. 2019. arXiv:1907.12526.

[92] K. Kurusu. http://ara.physics.wisc.edu/cgi-bin/DocDB/ShowDocument?docid=1694. Accessed: 2018-08-06.

[93] T. Meures. *Development of a sub-glacial array of radio antennas for the detection of the flux of GZK neutrinos*. PhD thesis, ULB - VUB, Belgium, 2014.

[94] P. Allison, R. Bard, J. J. Beatty, D. Z. Besson, C. Bora, C.-C. Chen, C.-H. Chen, P. Chen, A. Christenson, A. Connolly, J. Davies, M. Duvernois, B. Fox, R. Gaior, P. W. Gorham, K. Hanson, J. Haugen, B. Hill, K. D. Hoffman, E. Hong, S.-Y. Hsu, L. Hu, J.-J. Huang, M.-H. A. Huang, A. Ishihara, A. Karle, J. L. Kelley, D. Kennedy, I. Kravchenko, T. Kuwabara, H. Landsman, A. Laundrie, C.-J. Li, T. C. Liu, M.-Y. Lu, L. Macchiarulo, K. Mase, T. Meures, R. Meyhandan, C. Miki, R. Morse, J. Nam, R. J. Nichol, G. Nir, A. Novikov, A. O'Murchadha, C. Pfendner, K. Ratzlaff, M. Relich, M. Richman, L. Ritter, B. Rotter, P. Sandstrom, P. Schellin, A. Shultz, D. Seckel, Y.-S. Shiao, J. Stockham, M. Stockham, J. Touart, G. S. Varner, M.-Z. Wang, S.-H. Wang, Y. Yang, S. Yoshida, and R. Young. Performance of two askaryan radio array stations and first results in the search for ultrahigh energy neutrinos. *Phys. Rev. D*, 93:082003, Apr 2016. URL: https://link.aps.org/doi/10.1103/PhysRevD.93.082003, doi:10.1103/PhysRevD.93.082003.

[95] P. Allison et al. Constraints on the Diffuse Flux of Ultra-High Energy Neutrinos from Four Years of Askaryan Radio Array Data in Two Stations. 2019. arXiv:1912.00987.

[96] D. Seckel. http://ara.physics.wisc.edu/cgi-bin/DocDB/ShowDocument?docid=1508. Accessed: 2017-06-14.

[97] Atri Bhattacharya, Raj Gandhi, Werner Rodejohann, and Atsushi Watanabe. The Glashow resonance at IceCube: signatures, event rates and pp vs. $p\gamma$ interactions. *JCAP*, 1110:017, 2011. arXiv:1108.3163, doi:10.1088/1475-7516/2011/10/017.

[98] Raj Gandhi, Chris Quigg, Mary Hall Reno, and Ina Sarcevic. Ultrahigh-energy neutrino interactions. *Astropart. Phys.*, 5:81–110, 1996. arXiv:hep-ph/9512364, doi:10.1016/0927-6505(96)00008-4.

[99] Amy Connolly, Robert S. Thorne, and David Waters. Calculation of High Energy Neutrino-Nucleon Cross Sections and Uncertainties Using the MSTW Parton Distribution Functions and Implications for Future Experiments. *Phys. Rev.*, D83:113009, 2011. arXiv:1102.0691, doi:10.1103/PhysRevD.83.113009.

[100] T. K. Gaisser and A. M. Hillas. Reliability of the Method of Constant Intensity Cuts for Reconstructing the Average Development of Vertical Showers. In *International Cosmic Ray Conference*, volume 8 of *International Cosmic Ray Conference*, page 353, Jan 1977.

[101] G. Puppi, H.S. Bridge, and K. Greisen. *Progress in Cosmic Ray Physics. Vol. 3. Edited by J.G. Wilson, ... Contributors: K. Greisen, H.S. Bridge, R.W. Thompson, G. Puppi.* North-Holland Publishing C, 1956. URL: https://books.google.com/books?id=V_E8ygAACAAJ.

[102] Dmitry Chirkin and Wolfgang Rhode. Muon Monte Carlo: A High-precision tool for muon propagation through matter. 2004. arXiv:hep-ph/0407075.

[103] http://www.nec2.org/.

[104] http://dd6um.darc.de/QucsStudio/qucsstudio.html.

[105] B. A. Clark. *Optimization of a Search for Ultra-High Energy Neutrinos in Four Years of Data of ARA Station 2.* PhD thesis, The Ohio State University, USA, 2019.

[106] K. M. Gorski, Eric Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelman. HEALPix - A Framework for high resolution discretization, and fast analysis of data distributed on the sphere. *Astrophys. J.*, 622:759–771, 2005. arXiv:astro-ph/0409513, doi:10.1086/427976.

[107] J. Kelley et al. Interferometric neutrino event reconstruction in inhomogeneous media with the askaryan radio array. *Proceedings of Science (ICRC2017)*, page 1054, 2017.

[108] Nathan Whitehorn, Jakob van Santen, and Sven Lafebre. Penalized Splines for Smooth Representation of High-dimensional Monte Carlo Datasets. *Comput. Phys. Commun.*, 184:2214–2220, 2013. arXiv:1301.2184, doi:10.1016/j.cpc.2013.04.008.

[109] J. Kelley et al. Observation of two deep, distant (1.4, 4)km impulsive rf transmitters by the askaryan radio array (ara). *Proceedings of Science (ICRC2017)*, page 1030, 2017.

[110] P. Allison et al. Measurement of the real dielectric permittivity ϵ_r of glacial ice. *Astropart. Phys.*, 108:63–73, 2019. arXiv:1712.03301, doi:10.1016/j.astropartphys.2019.01.004.

[111] Kumiko Kotera and Angela V. Olinto. The Astrophysics of Ultrahigh Energy Cosmic Rays. page 44, 2011. URL: http://arxiv.org/abs/1101.4256, arXiv:1101.4256, doi:10.1146/annurev-astro-081710-102620.

[112] Gary J. Feldman and Robert D. Cousins. Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D*, 57:3873–3889, Apr 1998. URL: https://link.aps.org/doi/10.1103/PhysRevD.57.3873, doi:10.1103/PhysRevD.57.3873.

[113] Alexander Aab et al. Probing the origin of ultra-high-energy cosmic rays with neutrinos in the EeV energy range using the Pierre Auger Observatory. *JCAP*, 1910(10):022, 2019. arXiv:1906.07422, doi:10.1088/1475-7516/2019/10/022.

[114] A. Anker et al. A search for cosmogenic neutrinos with the ARIANNA test bed using 4.5 years of data. 2019. arXiv:1909.00840.

[115] A. V. Olinto, K. Kotera, and D. Allard. Ultrahigh Energy Cosmic Rays and Neutrinos. *Nucl. Phys. Proc. Suppl.*, 217:231–236, 2011. arXiv:1102.5133, doi:10.1016/j.nuclphysbps.2011.04.109.

[116] K. Kotera, D. Allard, and A. V. Olinto. Cosmogenic Neutrinos: parameter space and detectabilty from PeV to ZeV. *JCAP*, 1010:013, 2010. arXiv:1009.1382, doi:10.1088/1475-7516/2010/10/013.

[117] A.G. Vieregg, K. Bechtol, and A. Romero-Wolf. A technique for detection of PeV neutrinos using a phased radio array. *Journal of Cosmology and Astroparticle Physics*, 2016(02):005–005, feb 2016. URL: https://doi.org/10.1088%2F1475-7516%2F2016%2F02%2F005, doi:10.1088/1475-7516/2016/02/005.

APPENDIX
AraSim Setup Scripts

A.1 Introduction

The purpose of this appendix is to make reproducible the AraSim simulated samples used in the diffuse analysis of A2 and A3 as laid out in the thesis by presenting the setup scripts used to configure AraSim runs. The SVN revision number of the AraSim used to generate signal data is 746. At trigger level, A2 live time can be segmented into five different configurations, and A3 four configurations (see chapter [**?**]). Each configuration is simulated with 11 energy bins ranging from 10^{16} to 10^{21} eV in half-decade energy steps. For each configuration I present only the script for the 10^{16} eV step. Other energies can trivially be obtained by changing the `EXPONENT` parameter and adjusting the `POSNU_RADIUS` parameter according to Tab. 3.1. Notice that for configurations where the proper trigger delay correction is absent, the setup scripts do not reflect this absence. Rather, the missing delays are simulated in the code body. Such implementation is incorporated in the latest AraSim, which can be found on GitHub [1].

A.2 Setup Scripts

In the following scripts, lines that start with "//" are ignored by code. The configurations are referred to as `TYPE`s. For each station, configuration-specific parameters are located at the top of each script. Parameters that are specified explicitly in the scripts follow the default values as defined in AraSim.

A.2.1 A2 Configuration 1

```
//---ARA02 TYPE 1 SPECIFIC PARAMETERS---

DETECTOR=4                //actual deployed station
DETECTOR_STATION=2        //ARA02
WAVEFORM_LENGTH=1280      //number of samples in a waveform
TRIG_WINDOW=1.1E-7        //110ns
POWERTHRESHOLD=-6.22      //5Hz RF trigger rate with 110ns trigger window 3/8 Hpol/Vpol trigger
```

[1] https://github.com/ara-software/AraSim

```
//---ENVIRONMENT AND GENERAL SETTING---

RANDOM_MODE=1         //new seed every run
NFOUR=1024
TRIG_ANALYSIS_MODE=0  //0 = signal + noise, 1 = signal only, 2 = noise only


//---DETECTOR---

//DETECTOR=4                  //actual deployed station
//DETECTOR_STATION=2          //ARA02
number_of_stations=1
TIMESTEP=3.125E-10
core_x=10000
core_y=10000
READGEOM=1
BH_ANT_SEP_DIST_ON=0
BORE_HOLE_ANTENNA_LAYOUT=0  //HVHV
ANTENNA_MODE=1              //apply TM antenna measurements
ALL_ANT_V_ON=0              //apply Hpol and Vpol gains respectively


//---PRIMARY AND SHOWER---

EXPONENT=16
NNU=10000                   //number of neutrino events
ONLY_PASSED_EVENTS=0        //0 (default): AraSim throws NNU events whether or not they pass; 1:
    AraSim throws events until the number of events that pass the trigger is equal to NNU_PASSED
NU_NUBAR_SELECT_MODE=3      //random
SELECT_FLAVOR=0             //random 1:1:1 ratio
SELECT_CURRENT=2            //rndom nc or cc
SECONDARIES=1               //enable secondaries
TAUDECAY=1                  //account for tau decay in secondaries
SHOWER_MODE=2               //both EM and hadronic shower in t-domain signal
POSNU_RADIUS=3000
RAYSOL_RANGE=13000
INTERACTION_MODE=1          //Veff cylinder mode


//---SIGNAL AND PROPAGATION---
```

```
SIMULATION_MODE=1    //time-domain mode
NOFZ=1
USE_ARA_ICEATTENU=1  //use ARA-measured attenuation data
WAVE_TYPE=0          //plane wave in-ice


//---NOISE---

NOISE_WAVEFORM_GENERATE_MODE=0  //generate new noise waveforms for each event
NOISE_EVENTS=1                  //number of pure noise waveforms
NOISE=0                         //flat thermal noise
NOISE_CHANNEL_MODE=1            //using different noise temperature for each channel
APPLY_NOISE_FIGURE=1            //apply TM signal chain calibrated results


//---TRIGGER---

//TRIG_WINDOW=1.1E-7       //110ns
//POWERTHRESHOLD=-6.22     //5Hz RF trigger rate with 110ns trigger window 3/8 Hpol/Vpol trigger
TRIG_MODE=1              //use vpol, hpol separated trigger mode
N_TRIG_V=3
N_TRIG_H=3


//---READOUT---

FILL_TREE_MODE=1        //fill tree for useful posnu events
V_SATURATION=1.5        //readout voltage saturation 1.5V
V_MIMIC_MODE=0          //0 : global trig is located center of readout windows
DATA_SAVE_MODE=1
DATA_LIKE_OUTPUT=1      //save data to eventTree
//WAVEFORM_LENGTH=1280    //number of samples in a waveform
```

A.2.2 A2 Configuration 2

```
//---ARA02 TYPE 2 SPECIFIC PARAMETERS---

DETECTOR=4               //actual deployed station
DETECTOR_STATION=2       //ARA02
WAVEFORM_LENGTH=1280     //number of samples in a waveform
TRIG_WINDOW=1.1E-7       //110ns
POWERTHRESHOLD=-6.22     //5Hz RF trigger rate with 110ns trigger window 3/8 Hpol/Vpol trigger
```

```
//---ENVIRONMENT AND GENERAL SETTING---

RANDOM_MODE=1         //new seed every run
NFOUR=1024
TRIG_ANALYSIS_MODE=0  //0 = signal + noise, 1 = signal only, 2 = noise only


//---DETECTOR---

//DETECTOR=4                  //actual deployed station
//DETECTOR_STATION=2          //ARA02
number_of_stations=1
TIMESTEP=3.125E-10
core_x=10000
core_y=10000
READGEOM=1
BH_ANT_SEP_DIST_ON=0
BORE_HOLE_ANTENNA_LAYOUT=0  //HVHV
ANTENNA_MODE=1              //apply TM antenna measurements
ALL_ANT_V_ON=0              //apply Hpol and Vpol gains respectively


//---PRIMARY AND SHOWER---

EXPONENT=16
NNU=10000                   //number of neutrino events
ONLY_PASSED_EVENTS=0        //0 (default): AraSim throws NNU events whether or not they pass; 1:
    AraSim throws events until the number of events that pass the trigger is equal to NNU_PASSED
NU_NUBAR_SELECT_MODE=3      //random
SELECT_FLAVOR=0             //random 1:1:1 ratio
SELECT_CURRENT=2            //rndom nc or cc
SECONDARIES=1               //enable secondaries
TAUDECAY=1                  //account for tau decay in secondaries
SHOWER_MODE=2               //both EM and hadronic shower in t-domain signal
POSNU_RADIUS=3000
RAYSOL_RANGE=13000
INTERACTION_MODE=1          //Veff cylinder mode


//---SIGNAL AND PROPAGATION---
```

```
SIMULATION_MODE=1    //time-domain mode
NOFZ=1
USE_ARA_ICEATTENU=1  //use ARA-measured attenuation data
WAVE_TYPE=0          //plane wave in-ice


//---NOISE---

NOISE_WAVEFORM_GENERATE_MODE=0  //generate new noise waveforms for each event
NOISE_EVENTS=1                  //number of pure noise waveforms
NOISE=0                         //flat thermal noise
NOISE_CHANNEL_MODE=1            //using different noise temperature for each channel
APPLY_NOISE_FIGURE=1            //apply TM signal chain calibrated results


//---TRIGGER---

//TRIG_WINDOW=1.1E-7       //110ns
//POWERTHRESHOLD=-6.22     //5Hz RF trigger rate with 110ns trigger window 3/8 Hpol/Vpol trigger
TRIG_MODE=1              //use vpol, hpol separated trigger mode
N_TRIG_V=3
N_TRIG_H=3


//---READOUT---

FILL_TREE_MODE=1        //fill tree for useful posnu events
V_SATURATION=1.5        //readout voltage saturation 1.5V
V_MIMIC_MODE=0          //0 : global trig is located center of readout windows
DATA_SAVE_MODE=1
DATA_LIKE_OUTPUT=1      //save data to eventTree
//WAVEFORM_LENGTH=1280    //number of samples in a waveform
```

A.2.3 A2 Configuration 3

```
//---ARA02 TYPE 3 SPECIFIC PARAMETERS---

DETECTOR=4               //actual deployed station
DETECTOR_STATION=2       //ARA02
WAVEFORM_LENGTH=1280     //number of samples in a waveform
TRIG_WINDOW=1.1E-7       //110ns
POWERTHRESHOLD=-6.22     //5Hz RF trigger rate with 110ns trigger window 3/8 Hpol/Vpol trigger
```

```
//---ENVIRONMENT AND GENERAL SETTING---

RANDOM_MODE=1         //new seed every run
NFOUR=1024
TRIG_ANALYSIS_MODE=0  //0 = signal + noise, 1 = signal only, 2 = noise only


//---DETECTOR---

//DETECTOR=4                  //actual deployed station
//DETECTOR_STATION=2          //ARA02
number_of_stations=1
TIMESTEP=3.125E-10
core_x=10000
core_y=10000
READGEOM=1
BH_ANT_SEP_DIST_ON=0
BORE_HOLE_ANTENNA_LAYOUT=0  //HVHV
ANTENNA_MODE=1              //apply TM antenna measurements
ALL_ANT_V_ON=0              //apply Hpol and Vpol gains respectively


//---PRIMARY AND SHOWER---

EXPONENT=16
NNU=10000                   //number of neutrino events
ONLY_PASSED_EVENTS=0        //0 (default): AraSim throws NNU events whether or not they pass; 1:
    AraSim throws events until the number of events that pass the trigger is equal to NNU_PASSED
NU_NUBAR_SELECT_MODE=3      //random
SELECT_FLAVOR=0             //random 1:1:1 ratio
SELECT_CURRENT=2            //rndom nc or cc
SECONDARIES=1               //enable secondaries
TAUDECAY=1                  //account for tau decay in secondaries
SHOWER_MODE=2               //both EM and hadronic shower in t-domain signal
POSNU_RADIUS=3000
RAYSOL_RANGE=13000
INTERACTION_MODE=1          //Veff cylinder mode


//---SIGNAL AND PROPAGATION---
```

```
SIMULATION_MODE=1    //time-domain mode
NOFZ=1
USE_ARA_ICEATTENU=1  //use ARA-measured attenuation data
WAVE_TYPE=0          //plane wave in-ice


//---NOISE---

NOISE_WAVEFORM_GENERATE_MODE=0  //generate new noise waveforms for each event
NOISE_EVENTS=1                  //number of pure noise waveforms
NOISE=0                         //flat thermal noise
NOISE_CHANNEL_MODE=1            //using different noise temperature for each channel
APPLY_NOISE_FIGURE=1            //apply TM signal chain calibrated results


//---TRIGGER---

//TRIG_WINDOW=1.1E-7       //110ns
//POWERTHRESHOLD=-6.22     //5Hz RF trigger rate with 110ns trigger window 3/8 Hpol/Vpol trigger
TRIG_MODE=1              //use vpol, hpol separated trigger mode
N_TRIG_V=3
N_TRIG_H=3


//---READOUT---

FILL_TREE_MODE=1        //fill tree for useful posnu events
V_SATURATION=1.5        //readout voltage saturation 1.5V
V_MIMIC_MODE=0          //0 : global trig is located center of readout windows
DATA_SAVE_MODE=1
DATA_LIKE_OUTPUT=1      //save data to eventTree
//WAVEFORM_LENGTH=1280    //number of samples in a waveform
```

A.2.4 A2 Configuration 4

```
//---ARA02 TYPE 4 SPECIFIC PARAMETERS---

DETECTOR=4               //actual deployed station
DETECTOR_STATION=2       //ARA02
WAVEFORM_LENGTH=1664     //number of samples in a waveform
TRIG_WINDOW=1.7E-7       //170ns
POWERTHRESHOLD=-6.4     //5Hz RF trigger rate with 170ns trigger window 3/8 Hpol/Vpol trigger
```

```
//---ENVIRONMENT AND GENERAL SETTING---

RANDOM_MODE=1         //new seed every run
NFOUR=1024
TRIG_ANALYSIS_MODE=0  //0 = signal + noise, 1 = signal only, 2 = noise only


//---DETECTOR---

//DETECTOR=4                  //actual deployed station
//DETECTOR_STATION=2          //ARA02
number_of_stations=1
TIMESTEP=3.125E-10
core_x=10000
core_y=10000
READGEOM=1
BH_ANT_SEP_DIST_ON=0
BORE_HOLE_ANTENNA_LAYOUT=0  //HVHV
ANTENNA_MODE=1              //apply TM antenna measurements
ALL_ANT_V_ON=0              //apply Hpol and Vpol gains respectively


//---PRIMARY AND SHOWER---

EXPONENT=16
NNU=10000                   //number of neutrino events
ONLY_PASSED_EVENTS=0        //0 (default): AraSim throws NNU events whether or not they pass; 1:
    AraSim throws events until the number of events that pass the trigger is equal to NNU_PASSED
NU_NUBAR_SELECT_MODE=3      //random
SELECT_FLAVOR=0             //random 1:1:1 ratio
SELECT_CURRENT=2            //rndom nc or cc
SECONDARIES=1               //enable secondaries
TAUDECAY=1                  //account for tau decay in secondaries
SHOWER_MODE=2               //both EM and hadronic shower in t-domain signal
POSNU_RADIUS=3000
RAYSOL_RANGE=13000
INTERACTION_MODE=1          //Veff cylinder mode


//---SIGNAL AND PROPAGATION---
```

```
SIMULATION_MODE=1    //time-domain mode
NOFZ=1
USE_ARA_ICEATTENU=1  //use ARA-measured attenuation data
WAVE_TYPE=0          //plane wave in-ice


//---NOISE---

NOISE_WAVEFORM_GENERATE_MODE=0  //generate new noise waveforms for each event
NOISE_EVENTS=1                  //number of pure noise waveforms
NOISE=0                         //flat thermal noise
NOISE_CHANNEL_MODE=1            //using different noise temperature for each channel
APPLY_NOISE_FIGURE=1            //apply TM signal chain calibrated results


//---TRIGGER---

//TRIG_WINDOW=1.7E-7       //170ns
//POWERTHRESHOLD=-6.4     //5Hz RF trigger rate with 170ns trigger window 3/8 Hpol/Vpol trigger
TRIG_MODE=1              //use vpol, hpol separated trigger mode
N_TRIG_V=3
N_TRIG_H=3


//---READOUT---

FILL_TREE_MODE=1        //fill tree for useful posnu events
V_SATURATION=1.5        //readout voltage saturation 1.5V
V_MIMIC_MODE=0          //0 : global trig is located center of readout windows
DATA_SAVE_MODE=1
DATA_LIKE_OUTPUT=1      //save data to eventTree
//WAVEFORM_LENGTH=1664    //number of samples in a waveform
```

A.2.5 A2 Configuration 5

```
//---ARA02 TYPE 5 SPECIFIC PARAMETERS---

DETECTOR=4               //actual deployed station
DETECTOR_STATION=2       //ARA02
WAVEFORM_LENGTH=1664     //number of samples in a waveform
TRIG_WINDOW=1.7E-7       //170ns
POWERTHRESHOLD=-6.4     //5Hz RF trigger rate with 170ns trigger window 3/8 Hpol/Vpol trigger
```

```
//---ENVIRONMENT AND GENERAL SETTING---

RANDOM_MODE=1         //new seed every run
NFOUR=1024
TRIG_ANALYSIS_MODE=0  //0 = signal + noise, 1 = signal only, 2 = noise only


//---DETECTOR---

//DETECTOR=4                  //actual deployed station
//DETECTOR_STATION=2          //ARA02
number_of_stations=1
TIMESTEP=3.125E-10
core_x=10000
core_y=10000
READGEOM=1
BH_ANT_SEP_DIST_ON=0
BORE_HOLE_ANTENNA_LAYOUT=0  //HVHV
ANTENNA_MODE=1              //apply TM antenna measurements
ALL_ANT_V_ON=0              //apply Hpol and Vpol gains respectively


//---PRIMARY AND SHOWER---

EXPONENT=16
NNU=10000                   //number of neutrino events
ONLY_PASSED_EVENTS=0        //0 (default): AraSim throws NNU events whether or not they pass; 1:
    AraSim throws events until the number of events that pass the trigger is equal to NNU_PASSED
NU_NUBAR_SELECT_MODE=3      //random
SELECT_FLAVOR=0             //random 1:1:1 ratio
SELECT_CURRENT=2            //rndom nc or cc
SECONDARIES=1               //enable secondaries
TAUDECAY=1                  //account for tau decay in secondaries
SHOWER_MODE=2               //both EM and hadronic shower in t-domain signal
POSNU_RADIUS=3000
RAYSOL_RANGE=13000
INTERACTION_MODE=1          //Veff cylinder mode


//---SIGNAL AND PROPAGATION---
```

```
SIMULATION_MODE=1    //time-domain mode
NOFZ=1
USE_ARA_ICEATTENU=1  //use ARA-measured attenuation data
WAVE_TYPE=0          //plane wave in-ice


//---NOISE---

NOISE_WAVEFORM_GENERATE_MODE=0  //generate new noise waveforms for each event
NOISE_EVENTS=1                  //number of pure noise waveforms
NOISE=0                         //flat thermal noise
NOISE_CHANNEL_MODE=1            //using different noise temperature for each channel
APPLY_NOISE_FIGURE=1            //apply TM signal chain calibrated results


//---TRIGGER---

//TRIG_WINDOW=1.7E-7       //170ns
//POWERTHRESHOLD=-6.4     //5Hz RF trigger rate with 170ns trigger window 3/8 Hpol/Vpol trigger
TRIG_MODE=1              //use vpol, hpol separated trigger mode
N_TRIG_V=3
N_TRIG_H=3


//---READOUT---

FILL_TREE_MODE=1        //fill tree for useful posnu events
V_SATURATION=1.5        //readout voltage saturation 1.5V
V_MIMIC_MODE=0          //0 : global trig is located center of readout windows
DATA_SAVE_MODE=1
DATA_LIKE_OUTPUT=1      //save data to eventTree
//WAVEFORM_LENGTH=1664    //number of samples in a waveform
```

A.2.6 A3 Configuration 1

```
//---ARA03 TYPE 1 SPECIFIC PARAMETERS---

DETECTOR=4               //actual deployed station
DETECTOR_STATION=3       //ARA03
WAVEFORM_LENGTH=1280     //number of samples in a waveform
TRIG_WINDOW=1.1E-7       //110ns
POWERTHRESHOLD=-6.22     //5Hz RF trigger rate with 110ns trigger window 3/8 Hpol/Vpol trigger
```

```
//---ENVIRONMENT AND GENERAL SETTING---

RANDOM_MODE=1         //new seed every run
NFOUR=1024
TRIG_ANALYSIS_MODE=0  //0 = signal + noise, 1 = signal only, 2 = noise only


//---DETECTOR---

//DETECTOR=4                  //actual deployed station
//DETECTOR_STATION=2          //ARA02
number_of_stations=1
TIMESTEP=3.125E-10
core_x=10000
core_y=10000
READGEOM=1
BH_ANT_SEP_DIST_ON=0
BORE_HOLE_ANTENNA_LAYOUT=0  //HVHV
ANTENNA_MODE=1              //apply TM antenna measurements
ALL_ANT_V_ON=0              //apply Hpol and Vpol gains respectively


//---PRIMARY AND SHOWER---

EXPONENT=16
NNU=10000                   //number of neutrino events
ONLY_PASSED_EVENTS=0        //0 (default): AraSim throws NNU events whether or not they pass; 1:
    AraSim throws events until the number of events that pass the trigger is equal to NNU_PASSED
NU_NUBAR_SELECT_MODE=3      //random
SELECT_FLAVOR=0             //random 1:1:1 ratio
SELECT_CURRENT=2            //rndom nc or cc
SECONDARIES=1               //enable secondaries
TAUDECAY=1                  //account for tau decay in secondaries
SHOWER_MODE=2               //both EM and hadronic shower in t-domain signal
POSNU_RADIUS=3000
RAYSOL_RANGE=13000
INTERACTION_MODE=1          //Veff cylinder mode


//---SIGNAL AND PROPAGATION---
```

```
SIMULATION_MODE=1    //time-domain mode
NOFZ=1
USE_ARA_ICEATTENU=1  //use ARA-measured attenuation data
WAVE_TYPE=0          //plane wave in-ice


//---NOISE---

NOISE_WAVEFORM_GENERATE_MODE=0  //generate new noise waveforms for each event
NOISE_EVENTS=1                  //number of pure noise waveforms
NOISE=0                         //flat thermal noise
NOISE_CHANNEL_MODE=1            //using different noise temperature for each channel
APPLY_NOISE_FIGURE=1            //apply TM signal chain calibrated results


//---TRIGGER---

//TRIG_WINDOW=1.1E-7       //110ns
//POWERTHRESHOLD=-6.22     //5Hz RF trigger rate with 110ns trigger window 3/8 Hpol/Vpol trigger
TRIG_MODE=1              //use vpol, hpol separated trigger mode
N_TRIG_V=3
N_TRIG_H=3


//---READOUT---

FILL_TREE_MODE=1        //fill tree for useful posnu events
V_SATURATION=1.5        //readout voltage saturation 1.5V
V_MIMIC_MODE=0          //0 : global trig is located center of readout windows
DATA_SAVE_MODE=1
DATA_LIKE_OUTPUT=1      //save data to eventTree
//WAVEFORM_LENGTH=1280    //number of samples in a waveform
```

A.2.7 A3 Configuration 2

```
//---ARA03 TYPE 2 SPECIFIC PARAMETERS---

DETECTOR=4               //actual deployed station
DETECTOR_STATION=3       //ARA03
WAVEFORM_LENGTH=1280     //number of samples in a waveform
TRIG_WINDOW=1.1E-7       //110ns
POWERTHRESHOLD=-6.22     //5Hz RF trigger rate with 110ns trigger window 3/8 Hpol/Vpol trigger
```

```
//---ENVIRONMENT AND GENERAL SETTING---

RANDOM_MODE=1         //new seed every run
NFOUR=1024
TRIG_ANALYSIS_MODE=0  //0 = signal + noise, 1 = signal only, 2 = noise only


//---DETECTOR---

//DETECTOR=4                  //actual deployed station
//DETECTOR_STATION=2          //ARA02
number_of_stations=1
TIMESTEP=3.125E-10
core_x=10000
core_y=10000
READGEOM=1
BH_ANT_SEP_DIST_ON=0
BORE_HOLE_ANTENNA_LAYOUT=0  //HVHV
ANTENNA_MODE=1              //apply TM antenna measurements
ALL_ANT_V_ON=0              //apply Hpol and Vpol gains respectively


//---PRIMARY AND SHOWER---

EXPONENT=16
NNU=100                   //number of neutrino events
ONLY_PASSED_EVENTS=0        //0 (default): AraSim throws NNU events whether or not they pass; 1:
    AraSim throws events until the number of events that pass the trigger is equal to NNU_PASSED
NU_NUBAR_SELECT_MODE=3      //random
SELECT_FLAVOR=0             //random 1:1:1 ratio
SELECT_CURRENT=2            //rndom nc or cc
SECONDARIES=1               //enable secondaries
TAUDECAY=1                  //account for tau decay in secondaries
SHOWER_MODE=2               //both EM and hadronic shower in t-domain signal
POSNU_RADIUS=3000
RAYSOL_RANGE=13000
INTERACTION_MODE=1          //Veff cylinder mode


//---SIGNAL AND PROPAGATION---
```

```
SIMULATION_MODE=1    //time-domain mode
NOFZ=1
USE_ARA_ICEATTENU=1  //use ARA-measured attenuation data
WAVE_TYPE=0          //plane wave in-ice


//---NOISE---

NOISE_WAVEFORM_GENERATE_MODE=0  //generate new noise waveforms for each event
NOISE_EVENTS=1                  //number of pure noise waveforms
NOISE=0                         //flat thermal noise
NOISE_CHANNEL_MODE=1            //using different noise temperature for each channel
APPLY_NOISE_FIGURE=1            //apply TM signal chain calibrated results


//---TRIGGER---

//TRIG_WINDOW=1.1E-7       //110ns
//POWERTHRESHOLD=-6.22     //5Hz RF trigger rate with 110ns trigger window 3/8 Hpol/Vpol trigger
TRIG_MODE=1              //use vpol, hpol separated trigger mode
N_TRIG_V=3
N_TRIG_H=3


//---READOUT---

FILL_TREE_MODE=1        //fill tree for useful posnu events
V_SATURATION=1.5        //readout voltage saturation 1.5V
V_MIMIC_MODE=0          //0 : global trig is located center of readout windows
DATA_SAVE_MODE=1
DATA_LIKE_OUTPUT=1      //save data to eventTree
//WAVEFORM_LENGTH=1280    //number of samples in a waveform
```

A.2.8 A3 Configuration 3

```
//---ARA03 TYPE 3 SPECIFIC PARAMETERS---

DETECTOR=4               //actual deployed station
DETECTOR_STATION=3       //ARA03
WAVEFORM_LENGTH=1664     //number of samples in a waveform
TRIG_WINDOW=1.7E-7       //170ns
POWERTHRESHOLD=-6.4     //5Hz RF trigger rate with 170ns trigger window 3/8 Hpol/Vpol trigger
```

```
//---ENVIRONMENT AND GENERAL SETTING---

RANDOM_MODE=1         //new seed every run
NFOUR=1024
TRIG_ANALYSIS_MODE=0  //0 = signal + noise, 1 = signal only, 2 = noise only


//---DETECTOR---

//DETECTOR=4                  //actual deployed station
//DETECTOR_STATION=2          //ARA02
number_of_stations=1
TIMESTEP=3.125E-10
core_x=10000
core_y=10000
READGEOM=1
BH_ANT_SEP_DIST_ON=0
BORE_HOLE_ANTENNA_LAYOUT=0  //HVHV
ANTENNA_MODE=1              //apply TM antenna measurements
ALL_ANT_V_ON=0              //apply Hpol and Vpol gains respectively


//---PRIMARY AND SHOWER---

EXPONENT=16
NNU=10000                   //number of neutrino events
ONLY_PASSED_EVENTS=0        //0 (default): AraSim throws NNU events whether or not they pass; 1:
    AraSim throws events until the number of events that pass the trigger is equal to NNU_PASSED
NU_NUBAR_SELECT_MODE=3      //random
SELECT_FLAVOR=0             //random 1:1:1 ratio
SELECT_CURRENT=2            //rndom nc or cc
SECONDARIES=1               //enable secondaries
TAUDECAY=1                  //account for tau decay in secondaries
SHOWER_MODE=2               //both EM and hadronic shower in t-domain signal
POSNU_RADIUS=3000
RAYSOL_RANGE=13000
INTERACTION_MODE=1          //Veff cylinder mode


//---SIGNAL AND PROPAGATION---
```

```
SIMULATION_MODE=1    //time-domain mode
NOFZ=1
USE_ARA_ICEATTENU=1  //use ARA-measured attenuation data
WAVE_TYPE=0          //plane wave in-ice


//---NOISE---

NOISE_WAVEFORM_GENERATE_MODE=0  //generate new noise waveforms for each event
NOISE_EVENTS=1                  //number of pure noise waveforms
NOISE=0                         //flat thermal noise
NOISE_CHANNEL_MODE=1            //using different noise temperature for each channel
APPLY_NOISE_FIGURE=1            //apply TM signal chain calibrated results


//---TRIGGER---

//TRIG_WINDOW=1.7E-7       //170ns
//POWERTHRESHOLD=-6.4     //5Hz RF trigger rate with 170ns trigger window 3/8 Hpol/Vpol trigger
TRIG_MODE=1              //use vpol, hpol separated trigger mode
N_TRIG_V=3
N_TRIG_H=3


//---READOUT---

FILL_TREE_MODE=1        //fill tree for useful posnu events
V_SATURATION=1.5        //readout voltage saturation 1.5V
V_MIMIC_MODE=0          //0 : global trig is located center of readout windows
DATA_SAVE_MODE=1
DATA_LIKE_OUTPUT=1      //save data to eventTree
//WAVEFORM_LENGTH=1664    //number of samples in a waveform
```

A.2.9 A3 Configuration 4

```
//---ARA03 TYPE 4 SPECIFIC PARAMETERS---

DETECTOR=4               //actual deployed station
DETECTOR_STATION=3       //ARA03
WAVEFORM_LENGTH=1664     //number of samples in a waveform
TRIG_WINDOW=1.7E-7       //170ns
POWERTHRESHOLD=-6.4     //5Hz RF trigger rate with 170ns trigger window 3/8 Hpol/Vpol trigger
```

```
//---ENVIRONMENT AND GENERAL SETTING---

RANDOM_MODE=1         //new seed every run
NFOUR=1024
TRIG_ANALYSIS_MODE=0  //0 = signal + noise, 1 = signal only, 2 = noise only


//---DETECTOR---

//DETECTOR=4                  //actual deployed station
//DETECTOR_STATION=2          //ARA02
number_of_stations=1
TIMESTEP=3.125E-10
core_x=10000
core_y=10000
READGEOM=1
BH_ANT_SEP_DIST_ON=0
BORE_HOLE_ANTENNA_LAYOUT=0  //HVHV
ANTENNA_MODE=1              //apply TM antenna measurements
ALL_ANT_V_ON=0              //apply Hpol and Vpol gains respectively


//---PRIMARY AND SHOWER---

EXPONENT=16
NNU=10000                   //number of neutrino events
ONLY_PASSED_EVENTS=0        //0 (default): AraSim throws NNU events whether or not they pass; 1:
    AraSim throws events until the number of events that pass the trigger is equal to NNU_PASSED
NU_NUBAR_SELECT_MODE=3      //random
SELECT_FLAVOR=0             //random 1:1:1 ratio
SELECT_CURRENT=2            //rndom nc or cc
SECONDARIES=1               //enable secondaries
TAUDECAY=1                  //account for tau decay in secondaries
SHOWER_MODE=2               //both EM and hadronic shower in t-domain signal
POSNU_RADIUS=3000
RAYSOL_RANGE=13000
INTERACTION_MODE=1          //Veff cylinder mode


//---SIGNAL AND PROPAGATION---
```

```
SIMULATION_MODE=1    //time-domain mode
NOFZ=1
USE_ARA_ICEATTENU=1  //use ARA-measured attenuation data
WAVE_TYPE=0          //plane wave in-ice


//---NOISE---

NOISE_WAVEFORM_GENERATE_MODE=0  //generate new noise waveforms for each event
NOISE_EVENTS=1                  //number of pure noise waveforms
NOISE=0                         //flat thermal noise
NOISE_CHANNEL_MODE=1            //using different noise temperature for each channel
APPLY_NOISE_FIGURE=1            //apply TM signal chain calibrated results


//---TRIGGER---

//TRIG_WINDOW=1.7E-7       //170ns
//POWERTHRESHOLD=-6.4     //5Hz RF trigger rate with 170ns trigger window 3/8 Hpol/Vpol trigger
TRIG_MODE=1              //use vpol, hpol separated trigger mode
N_TRIG_V=3
N_TRIG_H=3


//---READOUT---

FILL_TREE_MODE=1        //fill tree for useful posnu events
V_SATURATION=1.5        //readout voltage saturation 1.5V
V_MIMIC_MODE=0          //0 : global trig is located center of readout windows
DATA_SAVE_MODE=1
DATA_LIKE_OUTPUT=1      //save data to eventTree
//WAVEFORM_LENGTH=1664    //number of samples in a waveform
```

APPENDIX
ARA Station 2 - Geometry Calibration of D3BV and D6BV

B.1 Motivation

This study stems from an effort to search for a diffuse neutrino flux with ARA station 2 (A2). In such an analysis, one source of background is the four local calibration pulsers deployed with the station. In an ARA station, in addition to the four central measurement strings (D1-4), two outer strings are deployed (D5-6), each equipped with one horizontally polarized antenna (TH) and one vertically polarized antenna (BV). Each of the outer strings has a controllable pulser unit to enable pulse generation from the designated antenna. Only one out of the four calibration pulsers is firing at a time, at a nominal rate of 1 PPS. Therefore, in a run the pulser events in principle should always trigger at the same time, with a deviation of several tens of nanosecond only. This synchronicity is how these events are tagged by the software. However, due to clock drifts that we currently have no clear explanation for, sometimes calibration pulser events can be mis-tagged as an regular RF trigger. Since once deployed, the topology and and the signal source is fixed, one way to reject such backgrounds is to mask out geometric regions on the local sky of the station that correspond to the directions of firing calibration pulsers. In this scenario, it is ideal if the angular reconstruction of these pulser events is robust and well-understood, so that the masked regions could be as small as possible. We will only be using Vpol channels for reconstruction in this study. All Vpols are used by default, unless otherwise specified.

The reconstruction skymap of a sample of calibration pulser events of A2 from 2013 to 2016 is shown in Fig. B.1. Events that reconstruct to the vicinity of D5BV show inconsistency in the sense that there are two groups of events with an azimuth difference of $\sim 5^\circ$. We will refer to events with smaller azimuth as being in the "left lobe", versus the events in the "right lobe", which reconstructs to larger azimuthal angles. Notice that the right lobe contains more events than the left, and is closer to the nominal true direction of D5BV. The occurrence of this inconsistency does not correspond to any specific time frame in the data considered, except of course it only appears when D5BV was firing. This disproves the hypothesis that the mis-reconstructed left lobe is due to missing pulses from periods where the trigger delays are not correctly implemented in the trigger logic.

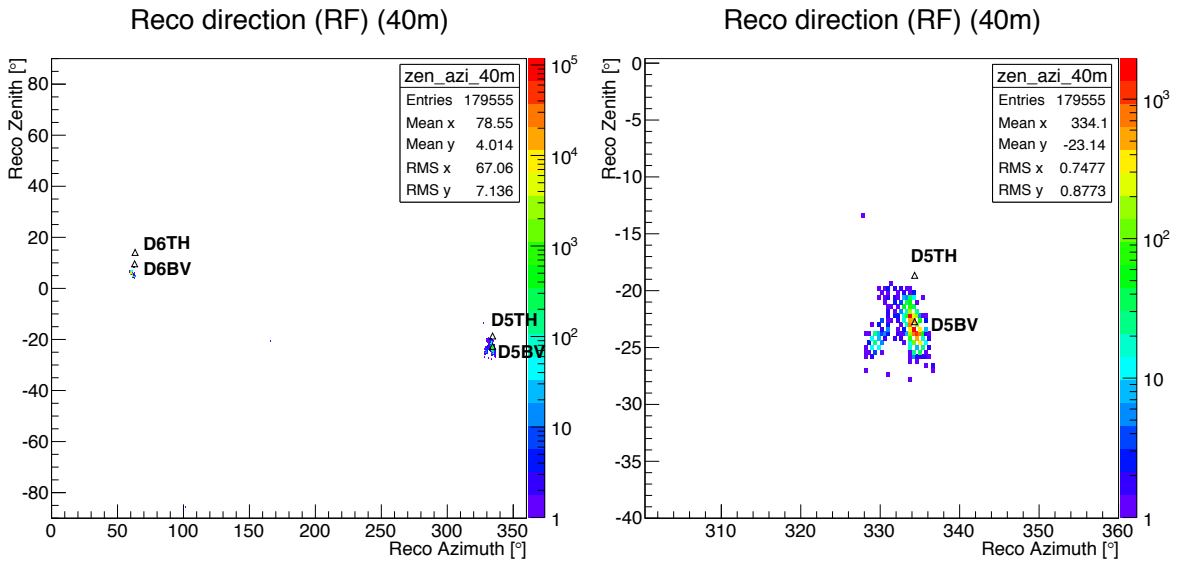


Figure B.1: A sample of A2 calibration pulser reconstruction. The convention for zenith is such that +90° is pointing vertically to the sky and -90° is vertically to the center of the Earth. The reconstruction assumes a source distance of 40m from the center of the local system, which sits 180m below the ice surface. (Right) The same figure with a focus on D5 calibration pulsers.

In Fig. B.2 and Fig. B.3, example events from each of the two lobes are shown. One may notice that from D5BV generally the top Vpols have weak or nearly no signal, while the bottom Vpols have stronger and more visible pulse signals. This could be explained by the larger VSWR for top antennas. We will examine systematically the arrival time of the pulses in these events. This will reveal the source of the reconstruction discrepancy.

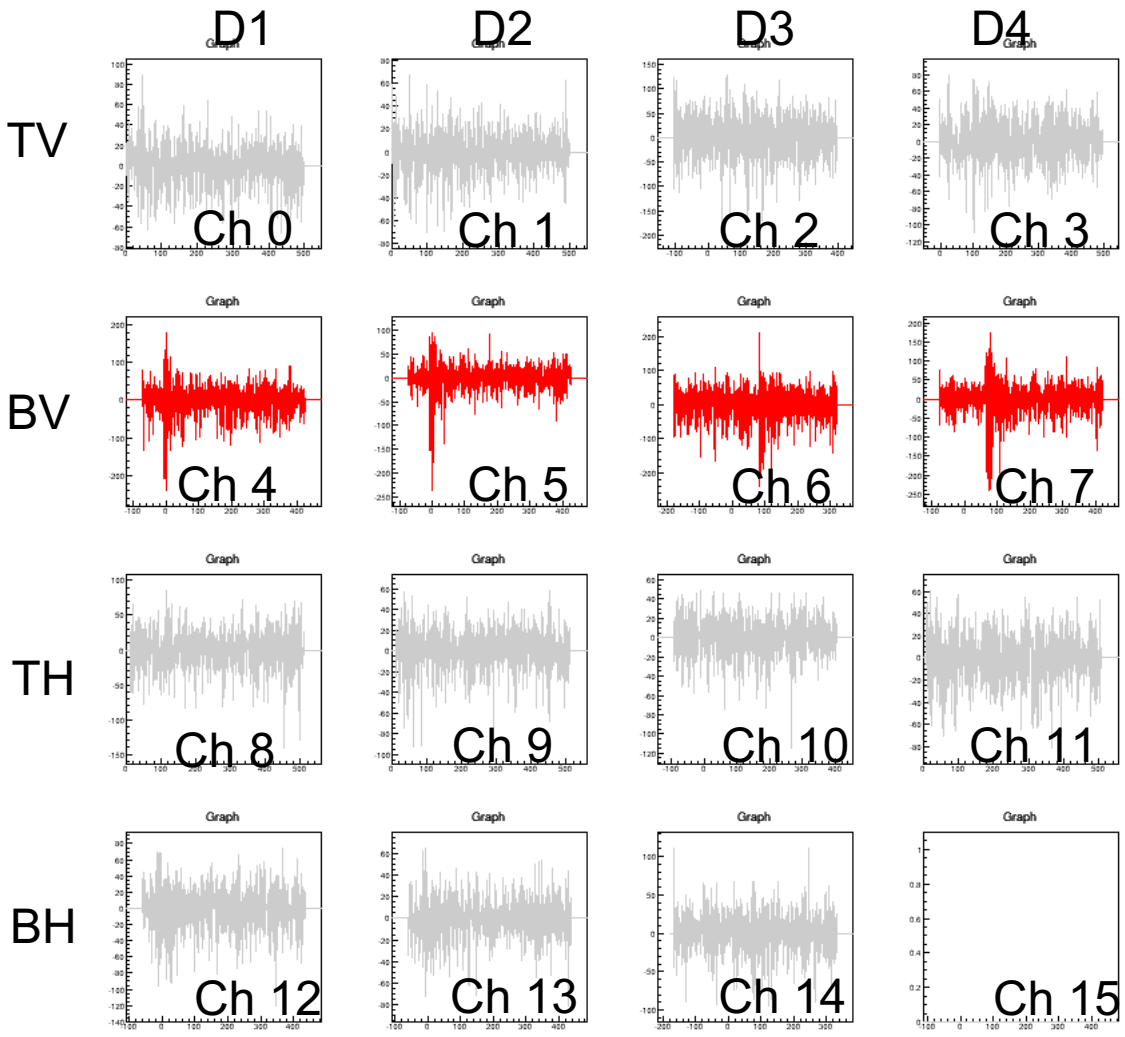


Figure B.2: A2 run 8111 event 3187 (left lobe). Colors are not meaningful here.

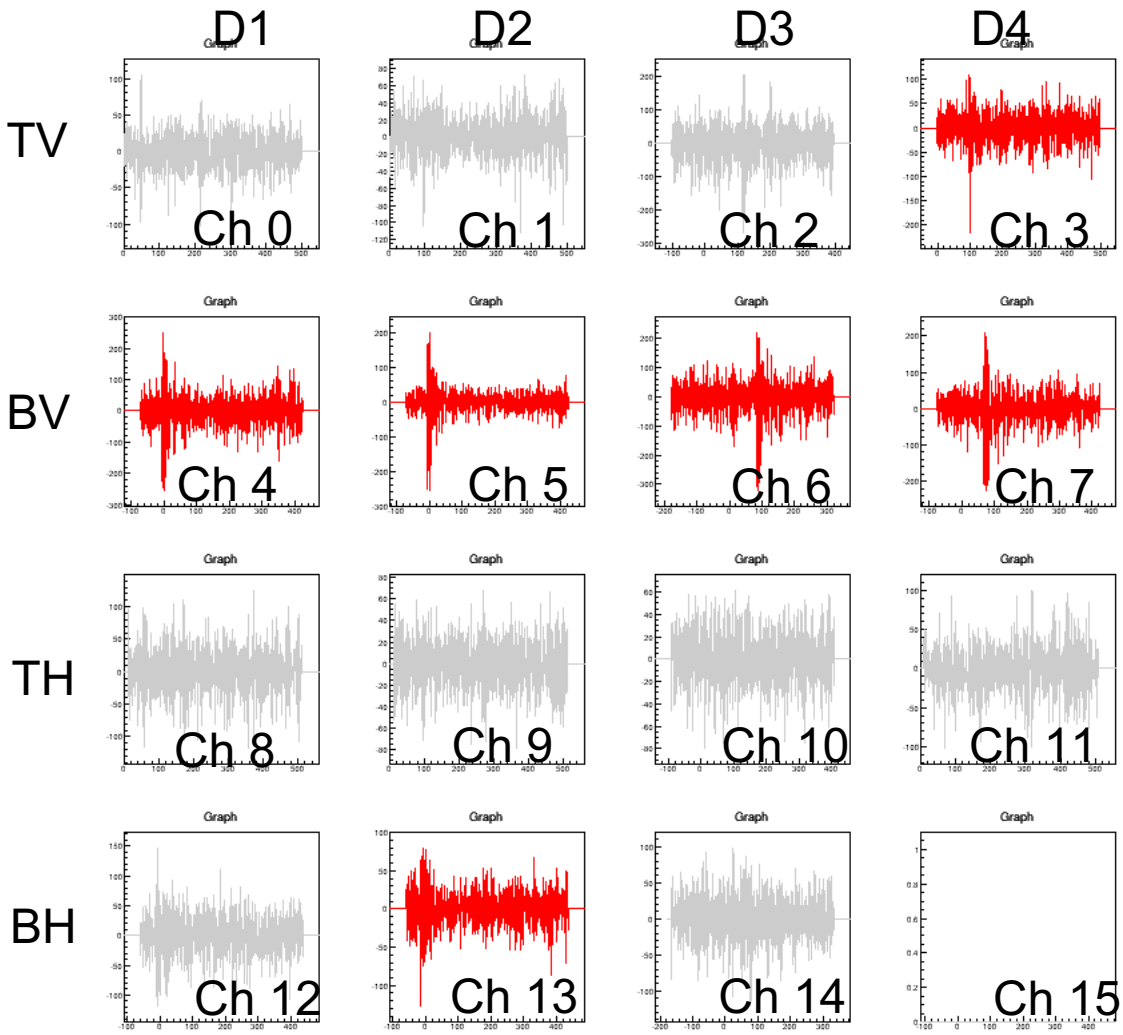


Figure B.3: A2 run 8111 event 800 (right lobe). Colors are not meaningful here.

B.2 Measured and Expected Pulse Arrival Time - Discrepancy

In Fig. B.4, the arrival time differences (dT) between every two Vpol channels are shown for all calibraton pulser events in A2 run8111. Each sub-figure is represents the dT for channel pair (row number, column number). These dT's are extracted as the peak time from the Hilbert envelope of the cross-correlation between channels. Using the known geometry, we can calculate the expected dT's, given our best knowledge of the ice model and its raytracing properties. Here, the expected dT's are computed using the Radiospline package, and are marked by the blue dashed lines. Two observations can be made - first of all, channel 1 (D2TV) sometimes register virtually no pulse. This could be a manifestation of the VSWR effect mentioned

above. Secondly, The dT's involving channel 6 (D3BV) show consistent discrepancies of $\sim$ 7ns from the expected values. The measured dT's of the rest of channel pairs are consistent with the expected to within $\sim$ 2ns. After inspecting closely the interferometry process of several events, we posit that the formation of the "left lobe" is due to a combination of weak top Vpol signals and the incorrect arrival time from channel 6. Since the top Vpols are weak, the reconstruction loses almost half of the timing information in the first place. Thus, occasionally when channel 6 dominates the interferometric traces, its inconsistent timing gives rise to a shift in the final reconstruction hot spot.

We can confirm that the mis-reconstruction of D5BV is indeed due to the anomalous timing of channel 6 by reconstructing the same events with channel 6 explicitly excluded. By doing so, the left lobe is eliminated, and all the events now reconstruct to the right lobe. At this point, to avoid mis-reconstruction, one can choose to proceed with the analysis with channel 6 discarded, or alternatively try to fix the anomaly. We will take the latter option here.

There are two possible explanations for this anomalous timing from channel 6. We can attribute it to an extra delay, or to a shift in its antenna location that is unaccounted for. In fact with data only from D5BV these two hypotheses are degenerate. Thus, we will include data from another Vpol calibration pulser - D6BV, and attempt to fit the delay and position of channel 6 combining these two data sets. Let us first review how the current geometry is obtained and calibrated in the next section.

B.3 Existing Geometry and Calibration

The deployment locations of ARA strings are marked by surveyors prior to drilling. Therefore, the surveyors' record represents the basis for the X-Y geometry of each ARA station. The Z coordinates of downhole antennas, on the other hand, are based on length measurement of cables. Combining data from all four calibration pulsers, Thomas Meures performed full station geometry calibration given the above basis. In his station fit, one measurement string and one calibration pulser string are fixed as references, and other three measurement strings can freely vary their XYZ position and an overall delay. For the calibration strings, the delay is unimportant and so is ignored. For details, refer to the Ph.D. thesis from Meures. The results from this work represents the current best knowledge of ARA station geometry, and are applied in the standard data analysis for A2 and A3. However, it was noted that the calibration results for A2 channel 6 and 14, which are the bottom antennas of D3, were unsatisfying and inconsistent with other channels. In the same thesis, analyses were carried out with these two channels completely removed.

In this study, we focus on obtaining a reasonable calibration on channel 6, as an addition to the aforementioned calibration.

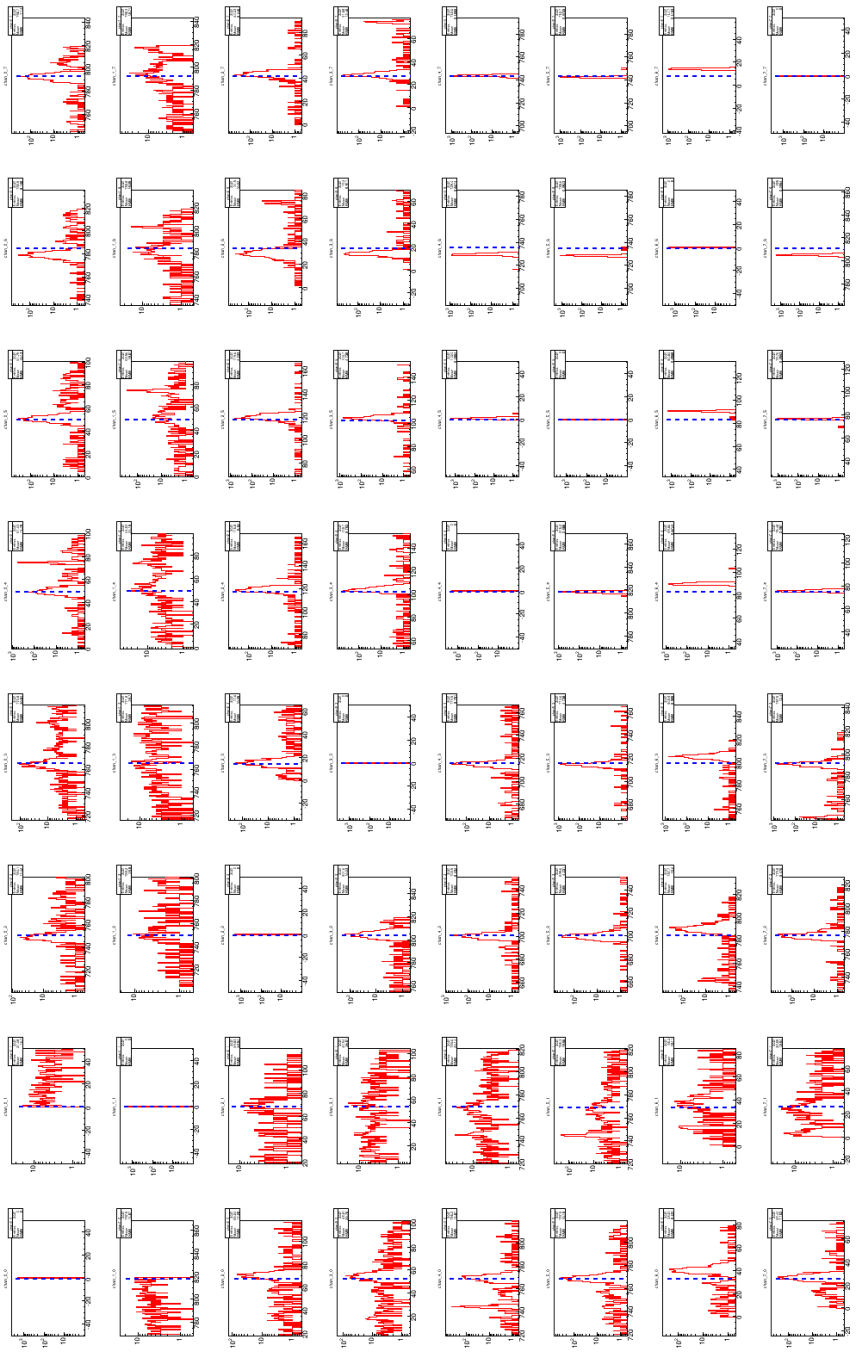


Figure B.4: D5BV dT [ns] distribution for all calibration pulser events in A2 run 8111. In each sub-figure, the dT is defined as $T_{row} - T_{column}$, where the row and column numbers start from zero. The blue-dashed line marks the expected dT from Radiospline.

B.4 Re-calibration

Considering that aside from channel 6 pairs, dT's among all other channel pairs are in agreement with the expected timing from D5BV, we will not attempt to carry out a full re-calibration of the station geometry. Instead, we will assume that the previous calibration for all channels except channel 6 is valid, and only seek to re-position channel 6 or re-estimate its delay.

Examining D6BV data also shed light on potential issues. The dT distribution for D6BV is shown in Fig. B.5. Note that channel 2 (D3TV), due to its proximity to D6BV, is saturated. This causes non-linearity in the digitizer, and therefore dT's involving channel 2 deviate from expectation. We can learn from the Fig. B.5 that

1. In general the dT's from any top-bottom Vpol pair is inconsistent with expectation, except point 2.
2. Channel 6, while being a bottom Vpol, has timing more consistent with those of the top Vpols.

Point 1 suggests a geometry modification such that the vertical separation between top and bottom Vpols should decrease, or that the pulser should be at a deeper position. Channel 6 still stands out as an anomaly. Since there is no indication from D5BV reconstruction that the top and bottom Vpols need further geometry correction, we consider the inconsistency solely attributable to missing calibration factors in the position of D6BV. We will try to find this missing correction. Given that the location and delay of channel 6 are also in doubt, we propose three different fit methods:

1. Free parameters: D6BV X,Y,Z and channel 6 Z, delay. Use D5BV + D6BV data sets to perform a single fit that determines the 5 parameters simultaneously.
2. Free parameters: D6BV X,Y,Z and channel 6 Z, delay. Use D6BV data set to fit for the 3 parameters of D6BV first, with channel 6 excluded at this stage. After the XYZ position of D6BV has been corrected, D5BV + D6BV data is used to re-fit the Z coordinate and delay of channel 6.
3. Free parameters: D6BV Z and channel 6 Z, delay. Use D6BV data set to fit for only the Z coordinate of D6BV first, with channel 6 excluded at this stage. After the Z position of D6BV has been corrected, D5BV + D6BV data is used to re-fit the Z and delay of channel 6.

In all three fits, channel 2 in D6BV data and channel 1 in D5BV data are not used due to aforementioned reasons. We use A2 run 6395 to represent D6BV, and run 8111 to represent D5BV. The measured dT's and RMS values are extracted from the peak positions and widths in each channel pair in Fig. B.4 and Fig. B.5. The expected dT's are from Radiospline. In each fit, we seek to minimize to reduced χ^2 as defined in Eq.B.1:

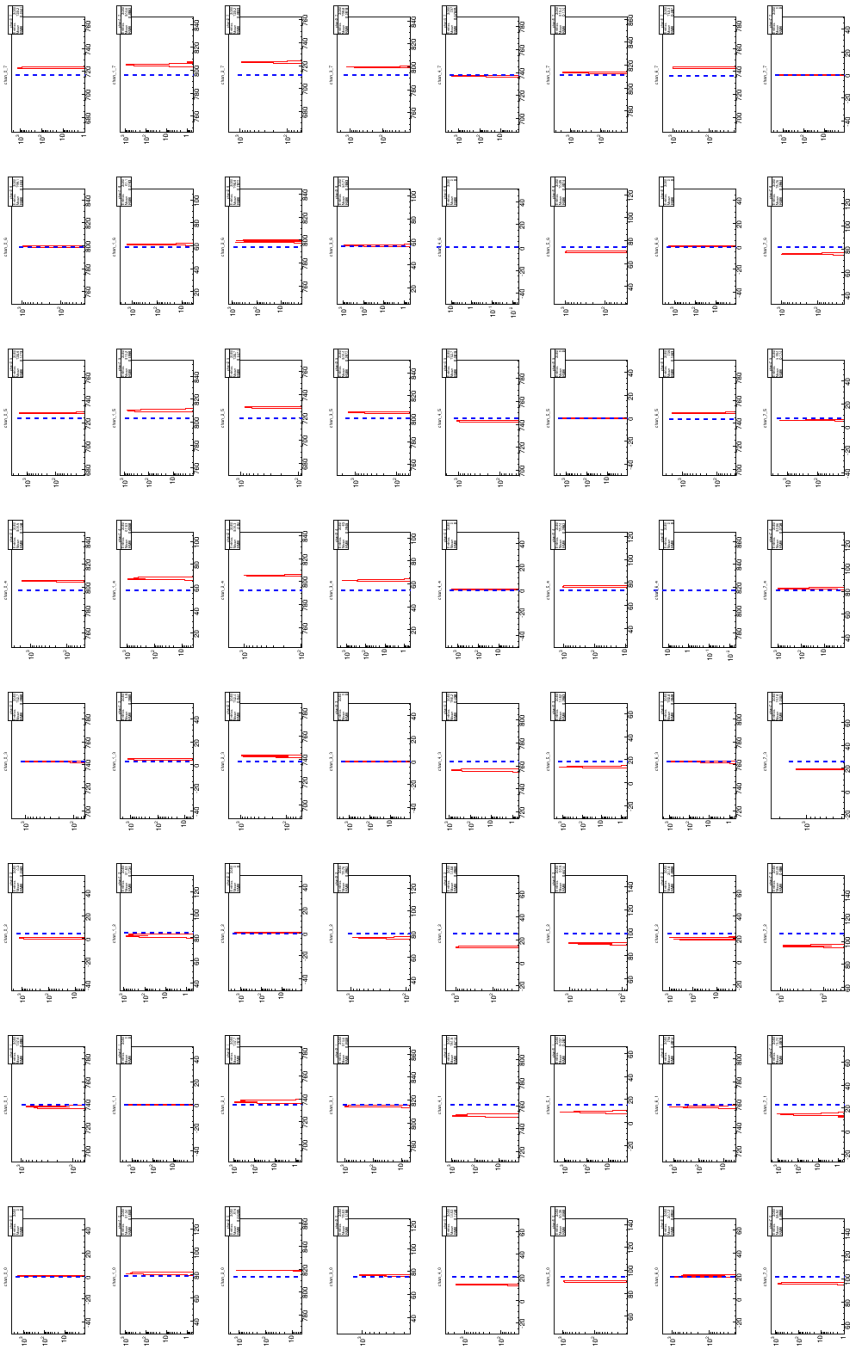


Figure B.5: D6BV dT [ns] distribution for all calibration pulser events in A2 run 6395.

$$\chi^2_\nu = (\sum_{i<j}(dT_{meas,ij} - dT_{exp,ij})^2/\sigma^2_{ij})/\nu \tag{B.1}$$

, where ν is the degree of freedom, which is equal to the number of selected distinct pairs minus the number of parameters being fitted.

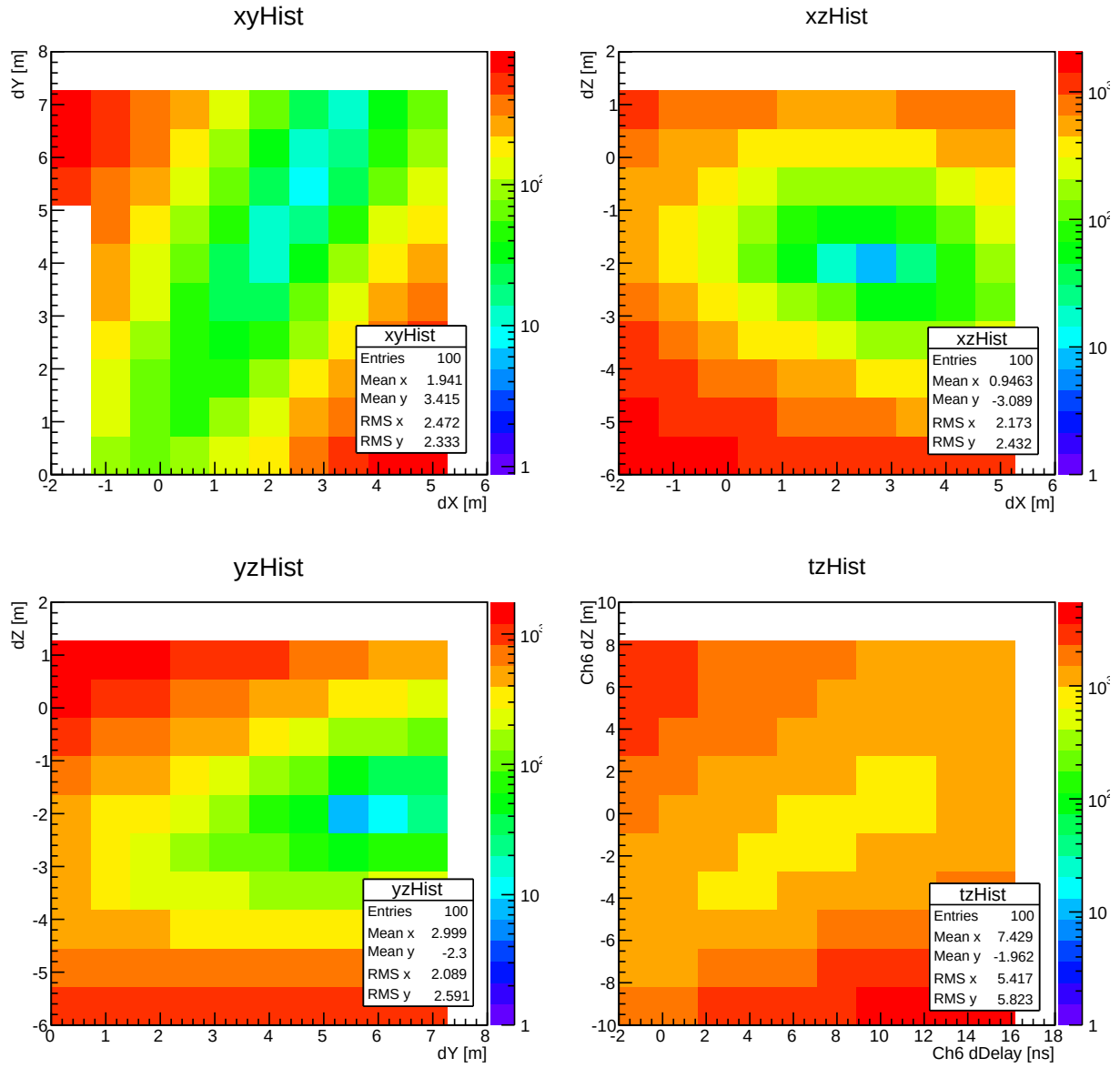


Figure B.6: Slices of the χ^2_ν space at the minimum from Fit 1. xyHist, xzHist, yzHist represent the X,Y,Z corrections to D6BV, and tzHist represent the Z and delay corrections to channel 6.

B.5 Fit Results

Fit 1 is visualized in Fig. B.6. Each sub-figure represents a 2-D slice of the χ^2_ν space intersecting with the best fit point. From the gradient of the contours we note that while D6BV X-Z and Y-Z are better constrained, minima in D6BV X-Y and channel 6 Z-Delay slices show much wider error regions. For D6BV X-Y, this indicates that with current data it is difficult to determine the X-Y location of the pulser, other than its relative azimuth ($\sim 63.35°$). For channel 6 Z-Delay, the contour suggests an increasing Z with an increasing delay, which only makes sense for incoming signal from D6BV (shallower than channel 6). This can be understood from the fact that the RMS values of the dT distributions from D6BV are much smaller than from D5BV, and so in this combined fit D6BV dominates.

Fit 2 is visualized in Fig. B.7 and Fig. B.8. Fig. B.7 shows the χ^2_ν contours of the 3-paramater fit of D6BV. A similar observation as Fit 1 can be made. Fig. B.8 shows the Z-Delay scan result after the above D6BV correction. The χ^2_ν space here in points to a clear minimum.

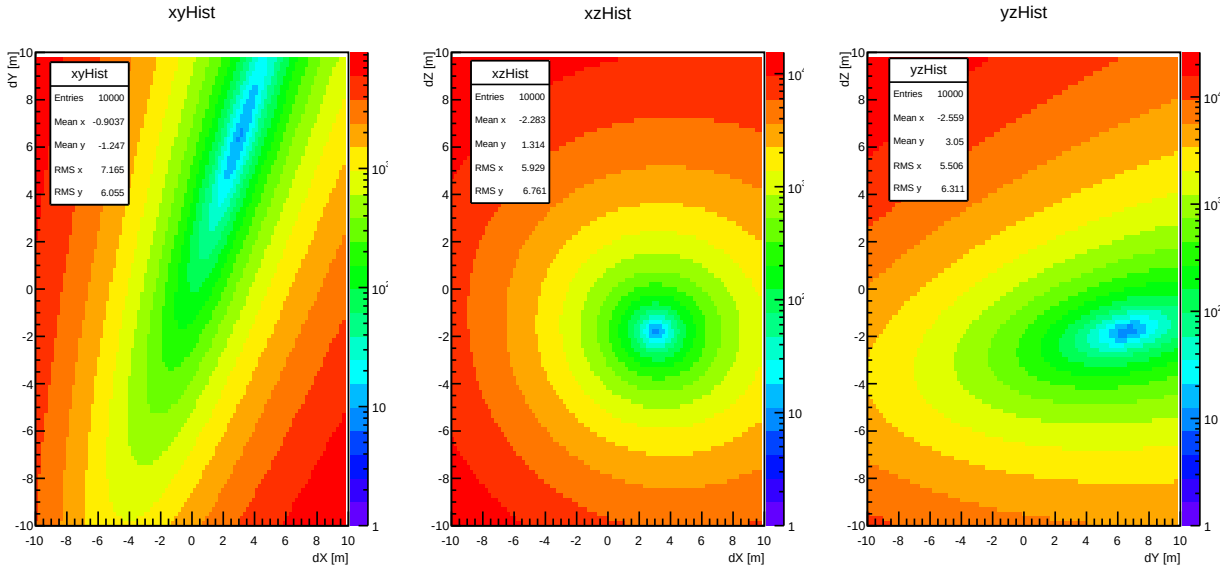


Figure B.7: Slices of the χ^2_ν space at the minimum for D6BV from Fit 2.

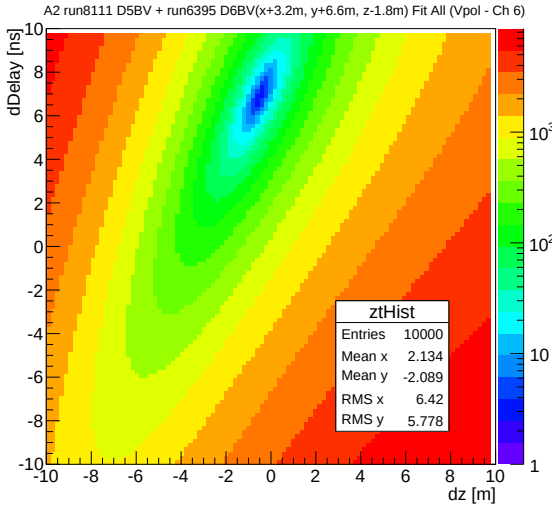


Figure B.8: χ^2_ν space for channel 6 from Fit 2, after corrections obtained in Fig. B.7 have been applied.

Fit 3 is visualized in Fig. B.9. As described earlier, fit to D6BV is reduced to a single parameter scan, after which the 2-D scan of channel 6 Z-Delay is performed.

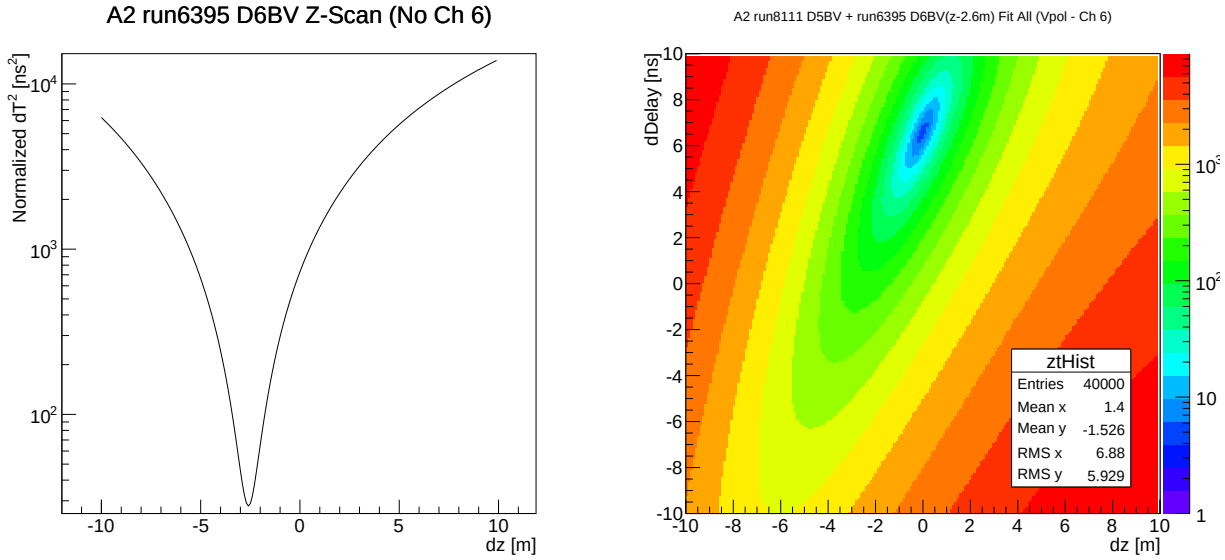


Figure B.9: (Left) χ^2_ν of D6BV Z from Fit 3. (Right) χ^2_ν space for channel 6 from Fit 3, after D6BV Z correction obtained in the left has been applied.

Fit results are summarized in Table.B.1. χ^2_ν values of each fit are listed in Table.B.2.

Table B.1: Fit results using D5BV and D5BV measured dT's. dX, dY, dZ and dDelay are deviations from the existing geometry. R, θ, ϕ are the spherical coordinates of D6BV in the A2 local system after each correction has been applied. The Z-coordinate of the origin of A2 local system is 180m below the local ice surface.

	D6BV						Ch 6	
Fit No.	dX [m]	dY [m]	dZ [m]	R [m]	θ[°]	ϕ[°]	dZ [m]	dDelay [ns]
1	2.8	5.6	-2	46.36	5.832	63.35	0	8
2	3.2	6.6	-1.8	47.44	5.941	63.46	-0.6	6.8
3	-	-	-2.6	40.07	5.889	63.33	0	6.6

Table B.2: Reduced χ^2 for each fit.

Fit No.	χ^2_ν
1	8.289
2	9.703, 2.997
3	27.91, 5.234

To validate these corrections, we will apply each of them and examine the reconstruction results. The reconstruction skymaps of D6BV and D5BV are shown in Fig. B.10 and Fig. B.11, respectively.

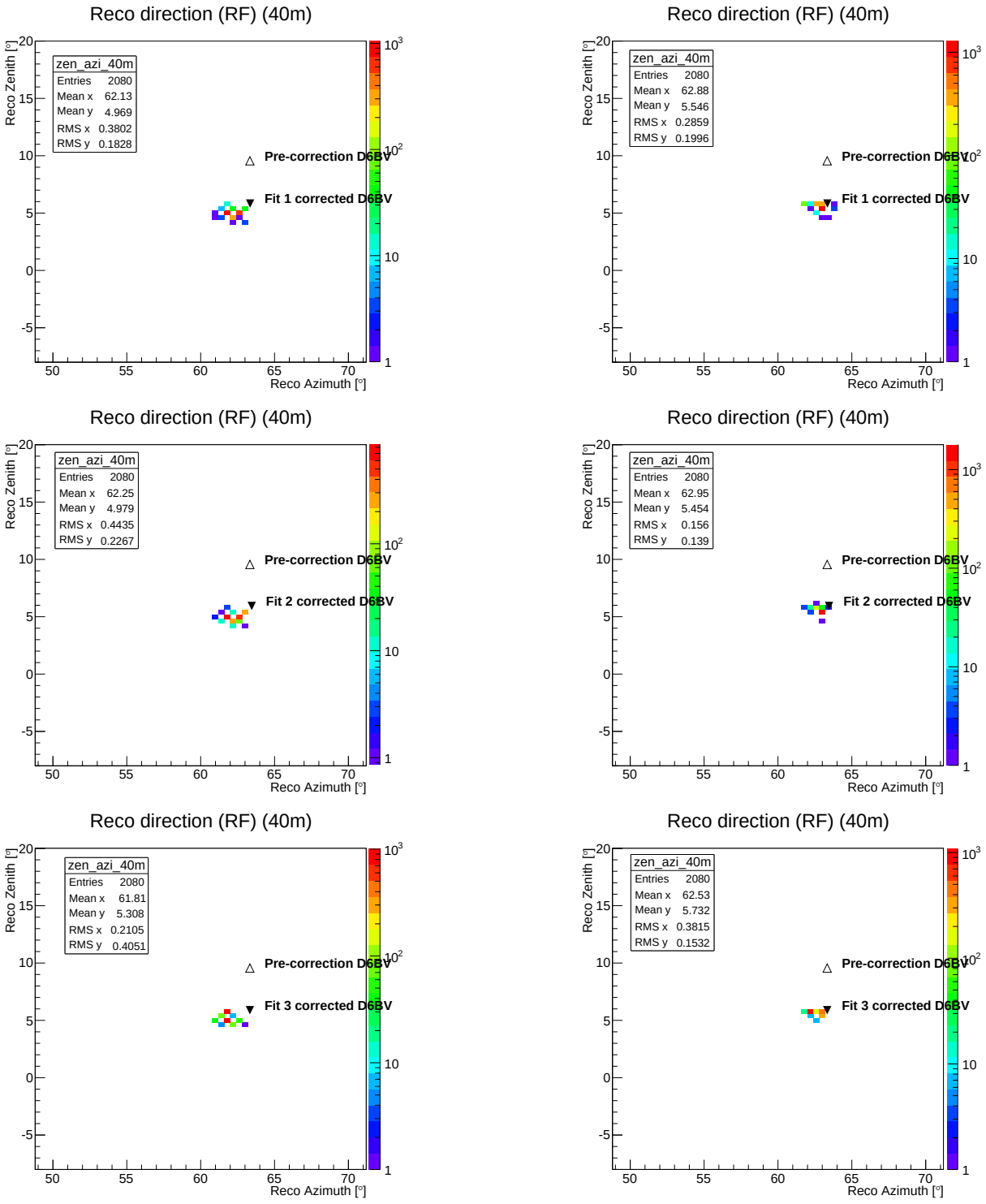


Figure B.10: (Left) D6BV pulser events reconstruction with all Vpol channels after applying corrections from Fit 1 to 3 (top → bottom). (Right) Same as left, but channel 2 is excluded in reconstruction.

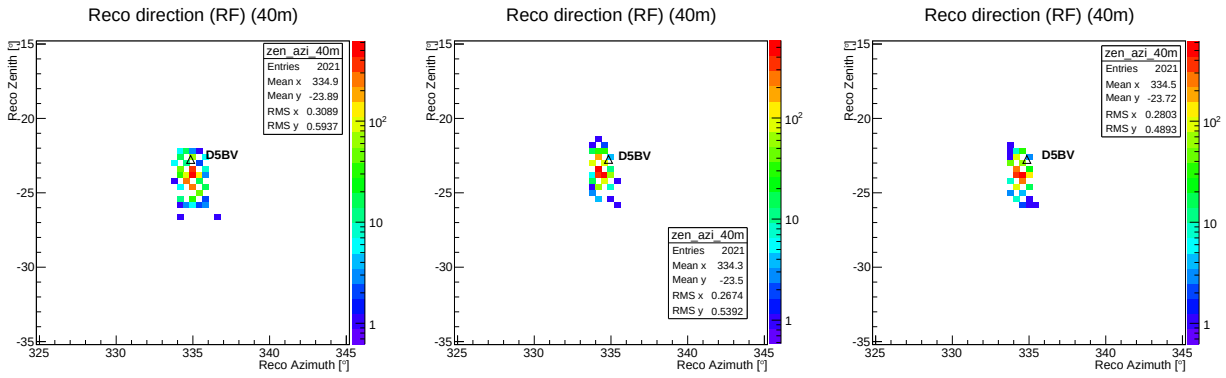


Figure B.11: (Left to right) D5BV pulser events reconstruction will all Vpol channels after applying corrections from Fit 1 to 3.

Table B.3 summarizes the angular deviation and RMS values of each fit. We also calculate the average zenith RMS values from the two calibration pulsers in table B.4. Note that the deviation for D6BV reconstruction is still $\sim>1$ degree after we used data to re-position the calibration pulser. This deviation can be attributed to the saturation in channel 2. In the context of the current diffuse analysis, we use all waveforms, including saturated one, so as to exploit as much timing information contained in the waveforms as possible. The trade-off is the $O(1^\circ)$ offset in the pointing ability of the reconstruction. The comparison with and without channel 2 is also shown in Fig. B.10. As a validity check, when channel 2 is excluded, the reconstruction can be observed to be correctly pointing to the nominal pulser direction to within 1° error.

Table B.3: Deviation of reconstruction from corrected calibration pulser directions and RMS values.

	D6BV			D5BV		
Fit No.	1	2	3	1	2	3
$d\theta[^\circ]$	-0.8633	-0.9618	-0.5805	-1.13	-0.74	-0.96
$d\phi[^\circ]$	-1.2156	-1.2063	-1.5216	0.045	-0.555	-0.355
$\sigma_\theta[^\circ]$	0.1828	0.2267	0.4051	0.5937	0.5392	0.4893
$\sigma_\phi[^\circ]$	0.3802	0.4435	0.2015	0.3089	0.2674	0.2803

Table B.4: Average zenith RMS values of D5BV and D6BV reconstructions from the three fits.

Fit No.	Avg. $\sigma_\theta[^\circ]$
1	0.38825
2	0.38295
3	0.4472

B.6 Discussion and Conclusion

For the purpose of setting mask regions on the local sky of A2, we are mostly concerned about the zenith and azimuth of D6BV. In this regard, the three fits are almost identical to within $\sim 1^\circ$. Channel 6 corrections among the three fits are also very close to each other. To decide on the corrections to apply, we will look at two key values: the mean χ^2_ν and the RMS of the zenith reconstruction. The mean χ^2_ν is the χ^2 per degree of freedom, and is indicative of the goodness of fit. The RMS of the zenith reconstruction determines how well we can localize anthropogenic backgrounds from the ice surface. A smaller RMS value in zenith reconstruction entails higher signal efficiency. Judging from these criteria, we choose Fit 2.

Looking forward, we plan to implement these corrections to all future A2 analyses. A SQLite file containing coordinates of all relevant station components is currently being maintained from which all analyzer should in principle obtain the station geometry. Corrections from Fit 2 will be added to this SQLite file. For analyses that mask out regions in the local sky of A2 to reject leaking calibration pulser events, such updated geometry should be used.

APPENDIX
IceCube Deep Pulser Locations

C.1 Introduction

The purpose of this document is to present the collected location information of the IceCube radio pulsers, known as deep pulsers, and of ARA station 2 and 3. These information will then be used to determine the line-of-sight directions of the pulsers relative to the two stations. The sources of information include South Pole surveys, ARA and IceCube internal slides, and ARA and IceCube internal documents. They will be specified wherever possible.

Probably more than needed number of digits are kept in this document.

C.2 Coordinate Systems

A more comprehensive document on ARA coordinate systems and their relations to IceCube and surveyor systems can be found at `https://www.phys.hawaii.edu/elog/ARA/91`. The document is the foundation of all the following contents.

We will consider five Cartesian coordinate systems: the IceCube global coordinate system, the ARA global and ARA2/3 station-centric coordinate systems, and the South Pole surveyor coordinate system. The IceCube system measures distances in meters, while the rest measures distances in feet.

The x axis of the IceCube system, the ARA global system, and the surveyor system aligns with Grid East. The y axis of these three systems aligns with Grid North. The x axes of the two ARA station-centric systems are parallel to the iceflow direction, which is 36.773°W with respect to Grid North. The y axes lie on the surface tangent to the geoid.

"Elevation" of the surveyor system is reported relative to a benchmark in the ice, the South Pole Benchmark (SPBM). However, the elevations of the origins of the IceCube and ARA global systems are defined with respect to "Mean Sea Level"(MSL). The conversion is:

$$Elevation_{MSL} = Elevation_{SPBM} + offset \tag{C.1}$$

The offset varies by year as the benchmark sinks due to packing snow pulling the benchmarking pipe. In this document, the 2011/2012 season value of 9263.98 ft will be used.

The surveyor system's origin for Easting and Northing sits at a point in the ice near the former Dome location. The Easting and Northing readings of the point are offset to be N 50,000' E 50,000'.

The origin of the IceCube system is defined as follows:

46500'E 52200'N and 2900' MSL

The origin of the ARA global coordinate system is defined as:

22400'E 53900'N and 9300' MSL

The origin of the ARA2 station-centric coordinate system expressed in ARA global system coordinates:

13126.70'E -8519.62'N and -18.72' elevation

The origin of the ARA3 station-centric coordinate system expressed in ARA global system coordinates:

9848.35'E -2835.19'N and -12.70' elevation

To calculate the relative directions of the deep pulsers, the approach is to first convert all ARA and IceCube coordinates to surveyor's coordinates, and then compute the relevant distances and angles.

C.3 ARA stations

I define the depth of the station center to be 180 meters below surface. This is equivalent to 590.55 ft.

C.3.1 ARA2

ARA2 center in surveyor's coordinates is:

$$\begin{pmatrix} 13126.70'E \\ -8519.62'N \\ -18.72'Ele - 590.55'Ele \end{pmatrix} + \begin{pmatrix} 22400'E \\ 53900'N \\ 9300'Ele - 9263.98'Ele \end{pmatrix} = \begin{pmatrix} 35527.70'E \\ 45380.38'N \\ -573.25'Ele \end{pmatrix} \tag{C.2}$$

C.3.2 ARA3

ARA3 center in surveyor's coordinates is:

$$\begin{pmatrix} 9848.35'E \\ -2835.19'N \\ -12.70'Ele - 590.55'Ele \end{pmatrix} + \begin{pmatrix} 22400'E \\ 53900'N \\ 9300'Ele - 9263.98'Ele \end{pmatrix} = \begin{pmatrix} 32248.35'E \\ 51064.81'N \\ -567.23'Ele \end{pmatrix} \tag{C.3}$$

C.4 IceCube deep pulsers

Three deep pulsers are installed, two on IceCube string 1, one on string 22. On string 1, one pulser is installed at a depth near 1450m, the other 2450m. I will refer to them as IceCube 1 Shallow (IC1S) and IceCube 1 Deep (IC1D), respectively. On string 22, the pulser is also near the 1450m mark. I refer to it as IceCube 22 Shallow (IC22S). To determine the positions of the deep pulsers, I take into account the following information:

1. IceCube hole surface coordinates - Coordinates of the IceCube hole surfaces measured at deployment, in which Eastings, Northings, and elevations in MSL are reported. These can be obtained from IceCube software and data, as described in `https://wiki.icecube.wisc.edu/index.php/How_to_calculate_surface_coordinates`. I will subtract the depth of a deep pulser from its surface hole elevation to get the corresponding MSL elevation.

2. Deep pulser design and deployment documentation - This includes the vertical dimension of the pulser antenna, and the deployment drawings. The vertical dimension halved can be found in slide 2 of `http://ara.physics.wisc.edu/cgi-bin/DocDB/ShowDocument?docid=92` to be 10 inches, or ∼0.254m. The deployment drawings can be found at `https://docushare.icecube.wisc.edu/dsweb/Get/Document-55620/Pulser_SpecialDeviceReview.ppt`.

The deployment drawing of the 1450m pulser is:

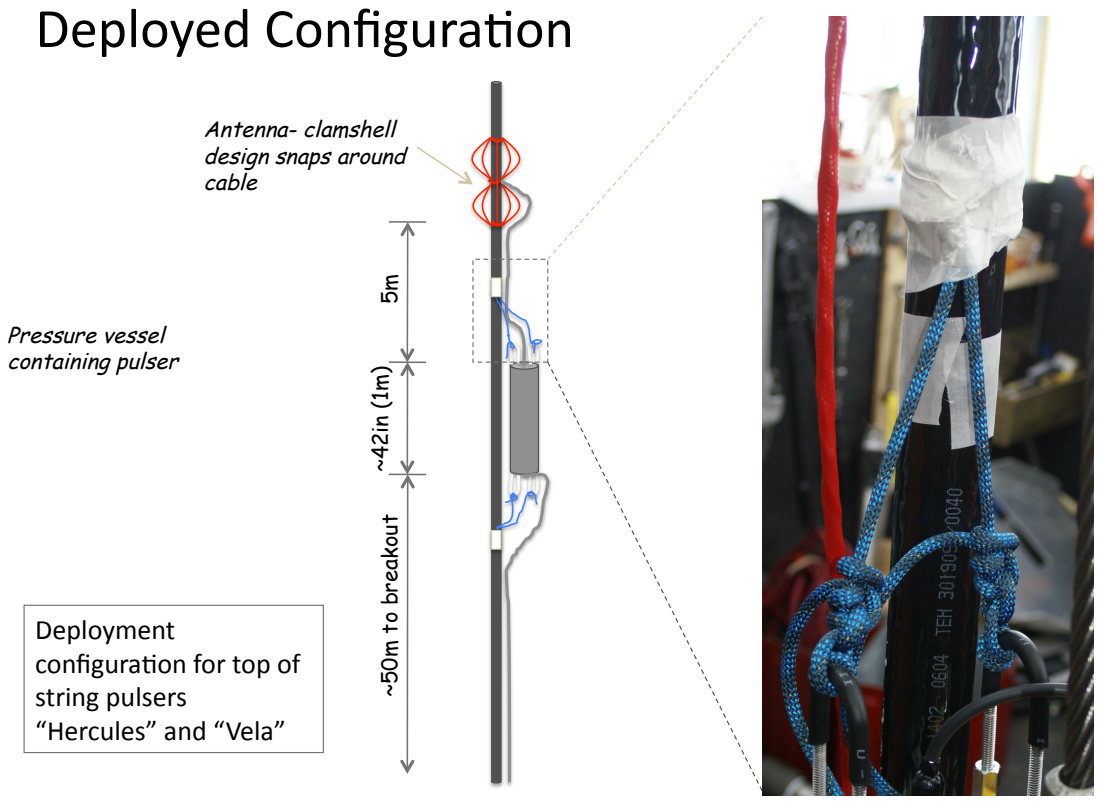


Figure C.1: Deployment drawing of the 1450m pulser.

From the drawing, the antenna center is then $50 + 1 + 5 + 0.254 = 56.254m$ above the top DOM (DOM 1).

The deployment drawing of the 2450m pulser is:

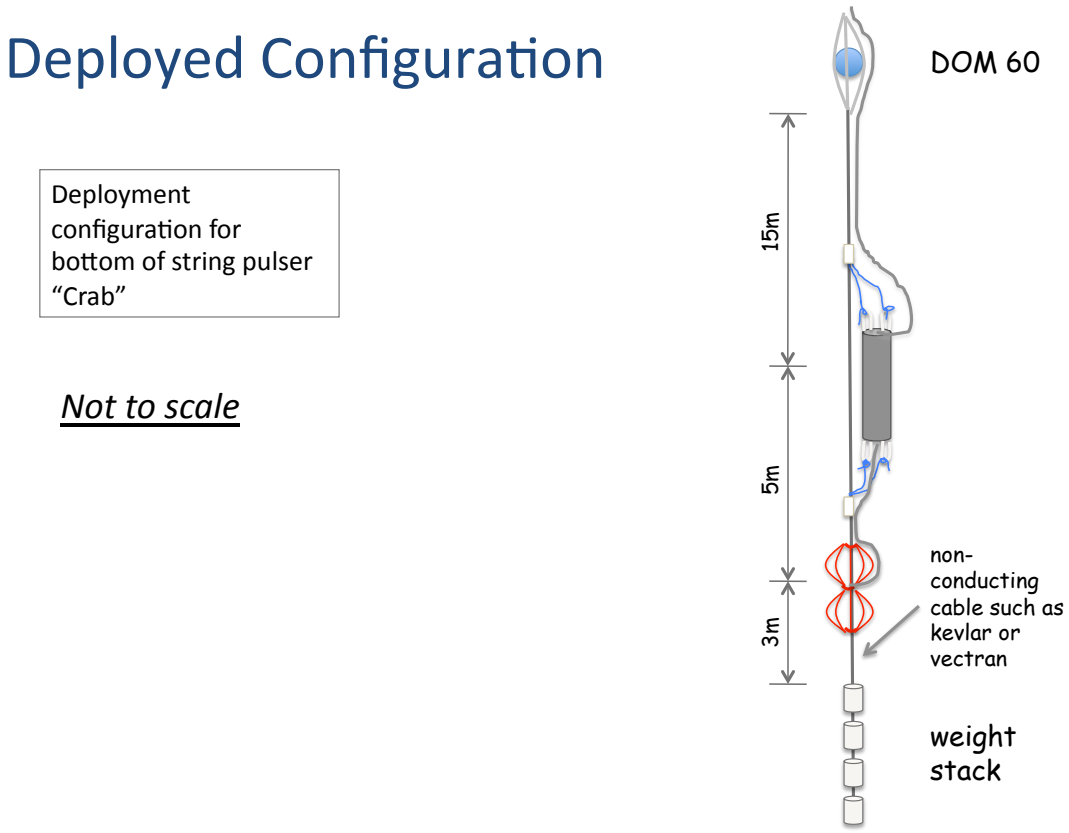


Figure C.2: Deployment drawing of the 2450m pulser.

From the drawing, the antenna center is $15 + 5 = 20m$ below the bottom DOM (DOM 60).

3. IceCube DOM positions - These positions are measured in the IceCube local coordinate system. In particular, I will use the positions of DOM 1 and 60. The data can be obtained from `http://icecube.wisc.edu/~dima/work/LBNL/reader/fat-other/doms/geo-f2k`, which is a result of IceCube stage 2 geometry calibration.

C.4.1 String 1

Hole 1 surface coordinates are (45659.6457'E, 50490.4199'N, 9282.865'MSL). IC1S and IC1D have identical Easting and Northing.

C.4.1.1 IC1S

DOM 1 on string 1 has a depth of 1452.04m below the surface. Therefore, the antenna MSL elevation is $9282.865' - (1452.04m - 56.254m) * 3.2808ft/m = 4703.5146'$. Using Eq. (1), this is -4560.4654' SPBM.

C.4.1.2 IC1D

DOM 60 on string 1 has a depth of 2452.47m below the surface. The antenna MSL elevation is $9282.865' - (2452.47m + 20m) * 3.2808ft/m = 1171.08677'$. Using Eq. (1), this is -8092.8932' SPBM.

C.4.2 String 22

Hole 22 surface coordinates are (44884.416'E, 51444.8819'N, 9282.854'MSL). IC22S has identical Easting and Northing.

C.4.2.1 IC22S

DOM 1 on string 22 has a depth of 1450.47m below the surface. Therefore, the antenna MSL elevation is $9282.854' - (1450.47m - 56.254m) * 3.2808ft/m = 4708.65452'$. Using Eq. (1), this is -4555.3255' SPBM.

Deep pulsers coordinates in surveyor's system are summarized in Table 1:

Table C.1: Surveyor's Coordinate System

	E [ft]	N [ft]	SPBM Elevation [ft]
IC1S	45659.6457	50490.4199	-4560.4654
IC1D	45659.6457	50490.4199	-8092.8932
IC22S	44884.416	51444.8819	-4555.3255

C.5 Relative Directions

C.5.1 Zenith and azimuth conventions

The zenith, denoted by θ, ranges from 0° to 180°. 0° is "up" towards the sky, and 180° is "down" towards the center of the Earth. The azimuth is denoted by ϕ, and goes from 0° to 360°, where 0° is along the +x axis. The angles to be calculated in the next section are "line-of-sight" angles.

C.5.2 Results

To compute the azimuth angle of a deep pulser to an ARA station, one must take into account the fact that the x axis of the ARA station-centric coordinate system aligns with the iceflow direction, which is essentially $90 + 36.773 = 126.773°$ in azimuth from the x axis of the surveyor system (Grid East). As a

result, we have:

$$if(\phi_{surveyor} - 126.773^\circ >= 0)\phi_{station-centric} = \phi_{surveyor} - 126.773^\circ \quad \text{(C.4)}$$

$$if(\phi_{surveyor} - 126.773^\circ < 0)\phi_{station-centric} = 360^\circ + \phi_{surveyor} - 126.773^\circ \quad \text{(C.5)}$$

Using the coordinates of the ARA stations and the deep pulsers, incorporating Eq. (4) and (5), I summarize the relative distances and directions of the deep pulsers to ARA2 and ARA3 in Table 2 and Table 3. R is the straight-line distance. R_{xy} is the horizontal distance. θ is the zenith. ϕ is the azimuth in the surveyor system. ϕ_{local} is the azimuth in the station-centric system.

Table C.2: Deep pulser relative position to ARA2

ARA2	R [ft]	R_{xy} [ft]	$\theta[^\circ]$	$\phi[^\circ]$	$\phi_{local}[^\circ]$
IC1S	12028.5920	11348.5284	-19.358521	26.7617889	259.988789
IC1D	13613.7478	11348.5284	-33.528835	26.7617889	259.988789
IC22S	11840.6907	11151.0104	-19.651873	32.9462641	266.173264

Table C.3: Deep pulser relative position to ARA3

ARA3	R [ft]	R_{xy} [ft]	$\theta[^\circ]$	$\phi[^\circ]$	$\phi_{local}[^\circ]$
IC1S	14004.9529	13423.5903	-16.566633	-2.4524122	230.774588
IC1D	15389.2294	13423.5903	-29.276284	-2.4524122	230.774588
IC22S	13255.9241	12641.7807	-17.508893	1.72284251	234.949843

Table 4 and 5 converts the distance unit to meters.

Table C.4: Deep pulser distance to ARA2

ARA2	R [m]	R_{xy} [m]
IC1S	3666.3143	3459.03145
IC1D	4149.47034	3459.03145
IC22S	3609.04254	3398.82796

Table C.5: Deep pulser distance to ARA3

ARA3	R [m]	R_{xy} [m]
IC1S	4268.70963	4091.51031
IC1D	4690.63713	4091.51031
IC22S	4040.40567	3853.21475

C.6 Acknowledgement

Thanks to Jim Haugen and Dmitry Chirkin for providing ARA and IceCube geometry information. Thanks also to Albrecht Karle and David Seckel for the discussion.