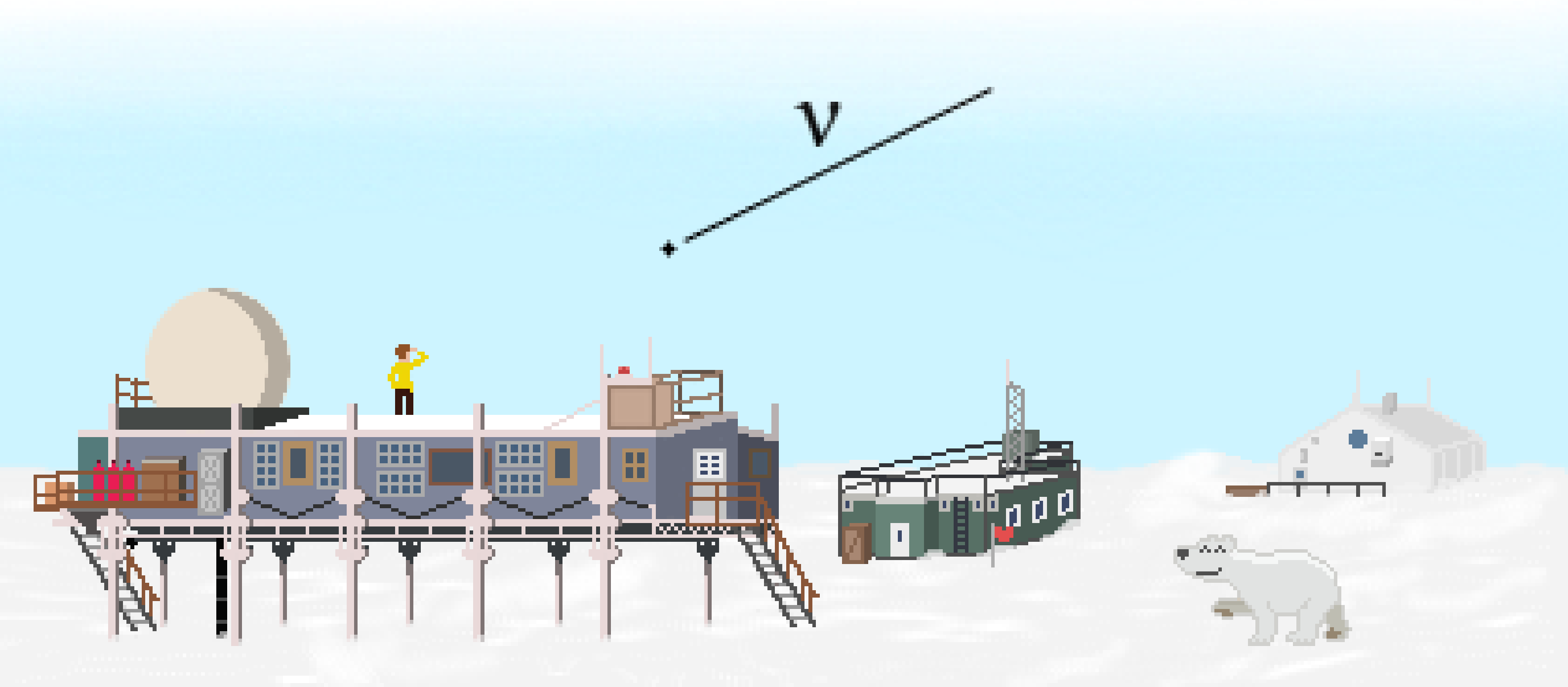
Vrije Universiteit Brussel

Faculteit Wetenschappen
Departement Natuurkunde

Calibration and comissioning of the Radio Neutrino Observatory in Greenland (RNO-G)

Jethro Stoffels
Promotor: Prof. Dr. Nick Van Eijndhoven
Co-promotor: Prof. Dr. Krijn de Vries
Supervisor: Prof. Dr. Simona Toscano

3 June 2022
Academic year: 2021 – 2022



Contents

Abstract

An important exploration in modern day astronomy revolves around the investigation of the diffuse cosmic neutrino flux, which has recently been observed by the IceCube neutrino observatory. A continuation of the observed flux is expected, but yet to be found at the highest of energies (PeV up to EeV energy range). Contemporary neutrino detectors fail to efficiently measure this neutrino flux at these energies since the neutrino flux steeply declines toward higher energies. In order to counter this, increasingly larger detection volumes are required. The required volume increase is so large that optical detection of Cherenkov light, as done by IceCube, from secondary particles after a neutrino interacts in a dense medium such as ice, is out of the question because of cost considerations. With its kilometer scale attenuation length, radio measurements naturally lend themselves to observe a large instrumented volume at once with only a rather sparse distribution of detector elements. For this reason, the Radio Neutrino Observatory in Greenland (RNO-G) is being developed and deployed in Greenland. This new radio detector array makes use of Askaryan emission, which is in the radio regime, produced by incident neutrinos interacting with the Greenlandic ice sheet. Last summer three station of the planned 35 were installed and started taking data. In this work the first samples of recorded data, taken during the first campaign, are explored to quantify its behaviour. In addition to that, a background investigation is performed to obtain its characteristic thermal noise temperature. Furthermore, the SαS distribution is employed to develop a model of the noise including transient impulsive outliers which add a non Gaussian component.

1 Introduction

1.1 Neutrino astronomy

In order to study the cosmos together with all the processes and mechanisms that occur within it, a wide variety of observation methods can be utilised. The most common way of detecting astronomical phenomena in all of mankinds history is through electromagnetic radiation. This way of detection uses the photon as its information carrier. Originally only light in the visible range was available but current day observations use different photon energies which require different detection methods (e.g. radio astronomy, X-ray astronomy, etc.). Although astronomy that utilises the photon as information carrier is the oldest and therefore most developed observational method, additional information is found in other detection channels. In this work we will focus on the neutrino as information carrier. Neutrino astronomy is a new field in physics which was opened by the IceCube collaboration in 2013 [1], when the discovery of a diffuse cosmic neutrino flux was announced. Contrary to photons, which at the highest energies (E $> 10^{15}$ eV) are absorbed and scattered to such a strong degree that they are unable to reach Earth from extragalactic distances, the neutrino remains largely unaffected at these energies and hence allows the study of the most extreme environments. As such, the neutrino is the ideal cosmic messenger to probe the most energetic processes in our universe.

Where charged particles or photons would fail (the latter due to scattering or absorption from dust clouds and the former due to their deflection in magnetic fields), the neutrino succeeds in remaining detectable and pointing back to its source even after traveling cosmological distances as is visualised in Figure 1. Another reason why neutrinos are such excellent information carriers to probe the high-energy extragalactic universe lies in their nature to barely interact with any material they pass through. Therefore, when neutrinos are produced in high-energy astrophysical events they can propagate outwards and reach Earth at near light speeds while remaining practically unhindered. Since the neutrino has no charge, its trajectory is not influenced by intergalactic magnetic fields or charge distributions.

However, the fact that the neutrino barely interacts with ambient matter is a double edged sword, it allows the neutrino to reach Earth unscathed but it makes detecting it quite challenging. In order to measure a significant amount of neutrinos coming from an astrophysical source one needs a large enough detector volume. On top of that we are met with a steeply declining neutrino flux at high energies, thus making detection of neutrinos at these energies even more difficult [2, 3]. A diffuse flux of neutrinos coming from astrophysical sources, as well as neutrinos created through interaction of ultra-high energy cosmic rays with the (extra-)galactic

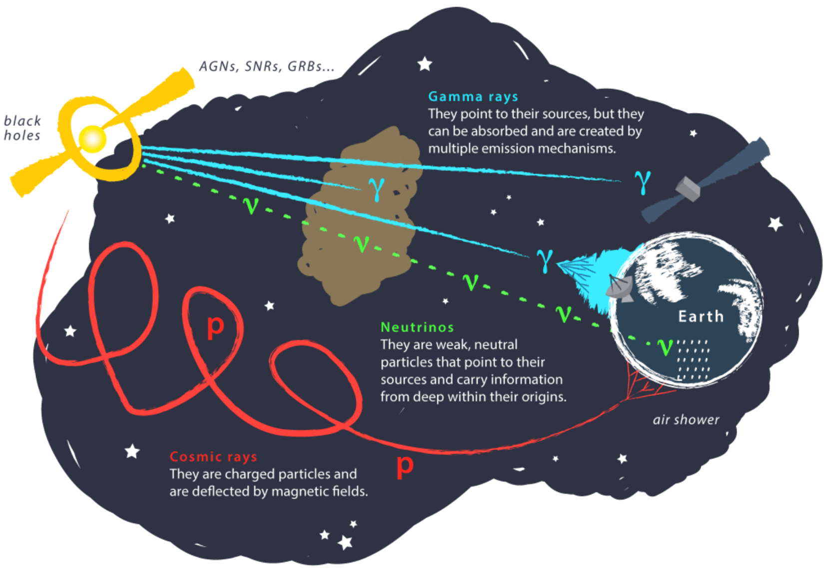


Figure 1: Artistic impression of the trajectories of different messenger particles as they travel through space. (Credits: Juan Antonio Aguilar and Jamie Yang. IceCube/WIPAC)

photon background is predicted over a large energy range [9]. Certain models for neutrino bursts in various scenarios (e.g. tidally disrupted stars [4], neutron star mergers [5], etc.) predict neutrino fluxes with energies in the PeV up to EeV range.

A current view of the high-energy multimessenger landscape consisting of neutrino flux predictions, neutrino flux observations made by IceCube, diffuse flux of γ rays by Fermi and the cosmic ray flux as measured by the Pierre-Auger observatory can be seen in Figure 2. The fact that the observed spectra show similar energy densities potentially points to a common origin. Not only is the study of high-energy neutrinos important for further investigation of the diffuse neutrino flux but the detection of these high-energy neutrinos originating from astrophysical sources provides clear evidence for cosmic ray acceleration in such sources. Note that, although cosmic-rays with energies in excess of 10^{19} eV have been detected at Earth, their origin is up to date a longstanding question.

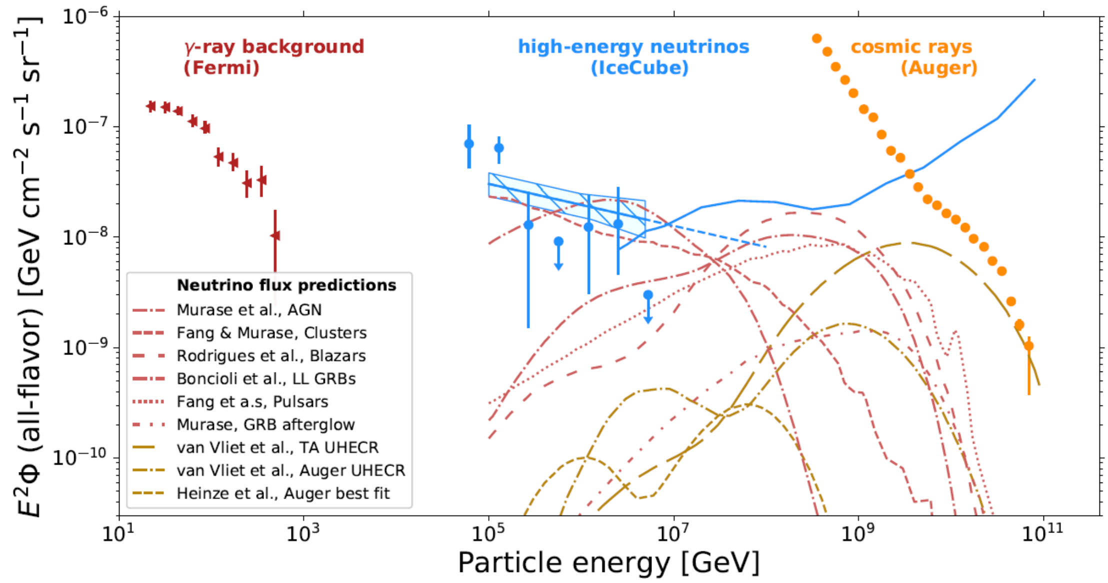


Figure 2: The multimessenger landscape of the high-energy universe. The lines colored in red represent the neutrino flux predictions originating from sources as well as interactions of ultra-high energy cosmic rays with the (extra-)galactic background light. Furthermore, γ-ray measurements from Fermi [6], neutrino measurements from IceCube (together with a fit to the muon neutrino flux) [7] and measurements of the ultra-high energy cosmic rays from the Pierre Auger Observatory [8] are shown in dark red, blue and orange respectively. (Credits: [15])

1.2 Neutrino astronomy: the IceCube experiment

In Figure 2, it is seen that a diffuse high-energy neutrino flux has been observed. This flux was detected by the IceCube Neutrino Observatory in 2013 [1]. IceCube [10] is an in-ice optical Cherenkov detector located at the South Pole. The detector is embedded in a cubic kilometer of ice and is composed of 5160 Digital Optical Modules (DOMs) attached to vertical strings arranged in a hexagonal grid. The DOMs are buried in the ice layer at a depth of 1450-2450 meter.

Whenever a cosmic neutrino traverses Earth, it has a probability of interacting with Earth's material through the weak interaction, hereby creating a shower of charged particles and/or its associated lepton. Neutrinos only interact through the weak force and at the TeV-PeV energy range, in which IceCube operates, these interactions are dominated by deep inelastic scattering. When a neutrino interacts with the ice, it can do so via a charged current interaction or a neutral current interaction. In the case of a charged current interaction, a $W^{\pm}$ boson is exchanged. A particle cascade is initiated and the lepton associated to the neutrino flavour (e, μ or τ in case of a ν_e, ν_μ or ν_τ respectively) is produced as well. The neutral current interaction occurs via the exchange of a Z^0 boson and initiates a particle cascade [11]. Should one of these processes happen close enough to Earth's surface, the cascade particles and/or the associated lepton end up traveling through the ice at relativistic speeds. When a particle travels through a material at speeds higher than the speed of light in that material, it will create Cherenkov light [12]. It is this light that is detected by the Digital Optical Modules (DOMs) in the IceCube observatory.

The detector has three main components, a surface array called "Icetop", a deeper "InIce" component and lastly a deep inner subdetector called "Deepcore". An artistic impression of the IceCube observatory can be seen in Figure 3.

The IceTop component consists of an array of 162 tanks filled with ice, capable of detecting cosmic ray induced airshowers. Each surface station has two ice tanks separated by 10 meters, located symmetrically with respect to the InIce string positions. Each ice tank contains two DOMs which allow for the detection of the airshower particle energy deposition. The deep InIce detector component consists of 86 vertical strings horizontally spaced by 125 meter. Such a vertical string contains 60 DOMs with a vertical spacing of 17 meter. The energy loss of particles in ice (about 300 MeV per meter) and the attenuation length of light in ice of 250 meter limits the energy threshold for reconstructable particle tracks to about 200 GeV. This limit arises due to the fact that at least 5 activated DOMs are needed for a good track reconstruction. The DeepCore strings are spaced differently, they have a horizontal spacing of 70 meter and a vertical spacing of 7 meter. Since the DeepCore has a much denser geometry, it is capable of reconstructing particle tracks for particles with much lower energies, this lowers the energy threshold for this component down to about 20 GeV.

The recording of Cherenkov light, originating from cascading particles traveling at speeds higher than the speed of light in the ice, allows reconstruction of the path and energy of the incoming particles. Cherenkov light is the electromagnetic equivalent of a sound shockwave and is also emitted in the shape of a cone [12]. The angle under which the Cherenkov light is emitted is the Cherenkov angle θ, which is given by $\cos(\theta) = \frac{1}{\beta n}$ where n is the index of refraction of the medium and $\beta = v/c$ of the relativistic particle. For highly relativistic velocities, we have $\beta \approx 1$, such that $\cos(\theta) \approx 1/n$, which is approximately constant and therefore allows for directional reconstruction of the particle. By measuring the timing properties, location and the deposited energy in the form of light, the path and energy of the original incident particle can be reconstructed [13]. More information on the

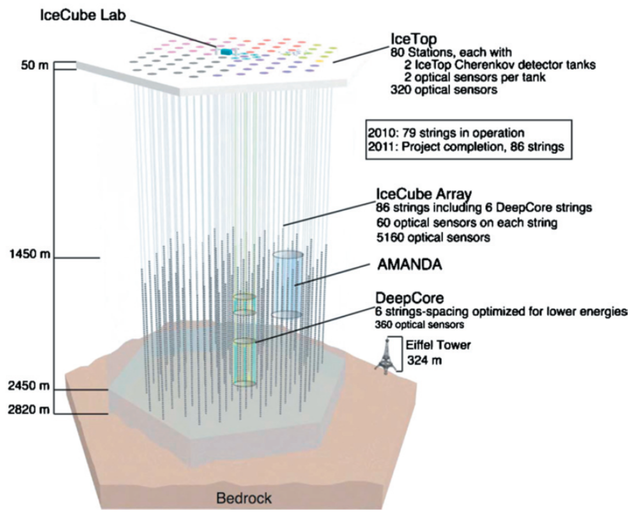


Figure 3: Artistic impression of the the IceCube Observatory and the geometry of its strings (Credits: [10]).

IceCube observatory can be found in source [10].

1.3 Probing the energy regime above 1 PeV

The IceCube Neutrino Observatory has served very well in pioneering cosmic neutrino astronomy and continues to be a great source of scientific information in this domain. However, just like any observatory it has its detection limitations. The amount of observable neutrinos is proportional to the detector volume and as mentioned above, we are facing a steeply declining neutrino flux for increasing energy. IceCube has measured the neutrino energy spectrum in the TeV-PeV energy range, with rather low statistics at the high-energy end. In order to study neutrinos with better statistics and at even higher energies, a large increase in detector volume is necessary. The required volume increase is so high that at this scale optical detection is no longer a possibility in view of cost considerations, and a switch to radio detectors is inescapable due to the long attenuation length (O(km)) of the radio signal compared to optical signals (O(100 m)). In order to expand the observable range of neutrino energies, which will allow for detection and to study the sources of the highest energy particles in the universe, a next generation instrument called IceCube-Gen2 is being developed [14]. The IceCube-Gen2 observatory will consist of the IceCube detector together with a new larger scale in-ice optical array, a low-energy core, a surface air shower array and an extended radio detector array.

With the predicted sensitivity of IceCube-Gen2, the diffuse neutrino flux can be studied in a much more extended energy range.

As mentioned above, The IceCube-Gen2 observatory will contain an optical component. This optical array aims to increase the observed amount of neutrinos in the PeV regime tenfold, allowing the detector to observe sources which are five times more faint [14]. The preliminary baseline design for the optical upgrade consists of adding 120 new strings (with 80 modules each) to the already existing IceCube strings, with the currently envisioned geometry and spacing this results in an additional 7.9 km^3 of instrumented detector volume equipped with upgraded modules which are capable of collecting up to three times more photons than the previous IceCube DOMs. More information on IceCube-Gen2 can be found in [14].

The addition of a radio detector set-up will open the detection window for neutrinos with energies above several tens of PeV, i.e. about two orders of magnitude above that of the IceCube Neutrino Observatory. Figure 4 shows the expected sensitivity of the radio component of the IceCube-Gen2 observatory compared to the current view of the multi-PeV neutrino landscape. The predictions shown by the grey bumps show the expected neutrino flux from ultra-high energy cosmic ray (UHECR) interactions with the extragalactic background light (EBL) and the cosmic microwave background radiation (CMBR). Based on kinematical arguments, the neutrino that is produced in a proton-gamma interaction gets on average about 4% of the initial proton energy. Since the photons of the EBL are more energetic than those of the CMBR, the proton threshold energy for neutrino production with EBL photons is lower than that for CMBR photons. Hence, the grey bump to the right corresponds to protons interacting with the CMBR, where the grey bump on the left is associated to interactions with the EBL.

However, additional R&D and field experience concerning the foreseen baseline design for IceCube-Gen2 still needs to be acquired. More specifically the development of targeted electronics and optimised antennas for an autonomous radio component is still in full swing. Much of the required experience and development for this will be acquired during the deployment and commissioning of the explorer radio array in Greenand (RNO-G).

1.4 First generation EeV (radio) neutrino detectors

The methodology and instrumentation used for radio arrays that aim to detect signals from incoming cosmic neutrinos is less mature and developed than for the optical array. In order to gain expertise and knowledge on this matter, proto-

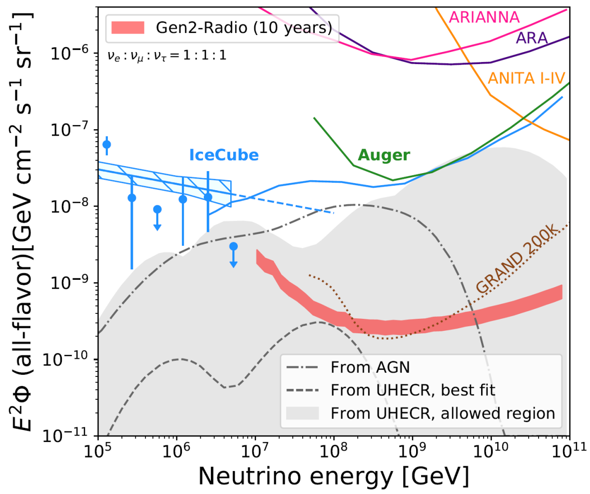


Figure 4: Current view of the multi-PeV neutrino landscape together with (limits on) the neutrino fluxes from different experiments and the expected sensitivity of the radio component of the IceCube-Gen2 observatory. (Credits: [14])

types and pilot-stage detectors were deployed. Among these are RICE (Radio Ice Cherenkov Experiment) [19], ARA (Askaryan Radio Array) [20, 21] and ARIANNA (Antarctic Ross Ice Shelf Antenna Neutrino Array) [22, 23]. Much valuable knowledge about the deployment and properties of such radio arrays was obtained from these experiments.

RICE, co-deployed with the AMANDA experiment [18] (IceCube's predecessor) at the South Pole, gave a valuable first experience with in-ice radio detectors, and provided the first neutrino limits from Askaryan radio detectors [24]. The ARA experiment is a direct successor to RICE. The ARA stations operate in dedicated dry holes with depths up to 200 m. Every ARA station consists of four receiver strings, each equipped with two vertically-polarized birdcage dipole antennas (Vpol) and two ferrite-loaded slot antennas (Hpol) to reconstruct the radio signals. The ARA experiment uses part of its antennas together as a phased array, a technique which it pioneered for radio detection of neutrinos [25]. The ARA stations also utilized calibration strings and surface antennas (on earlier stations). Apart from obtaining field experience, the ARA collaboration has also published constraints on the diffuse ultra-high energy neutrino flux, neutrinos from gamma-ray bursts and radio emissions from solar flares [21].

The ARIANNA experiment is centered around surface level antennas [22], making use of both direct signals as well as signals reflected from the sea water underneath the ice sheet at the Antarctic coast. Unrestricted by borehole geometry, high-gain log-periodic dipole antennas (LPDAs) are deployed in shallow slots in the snow-surface. These LPDAs possess broadband characteristics and dedicated polarization sensitivity, especially for horizontally polarized signals. ARIANNA pioneered autonomous low power stations, operated via wireless communication, which make use of renewable energy sources (e.g. solar panels and wind turbines), as well as the employment of surface detectors which can self-veto against air showers.

These efforts also gave much insight in technologically important aspects of operating under the harsh conditions such as those encountered at the South Pole. The balloon borne ANITA (Antarctic Impulsive Transient Antenna) [26] experiment has had four missions where it flew over Antarctica and scanned the surface from far away for upcoming neutrino radio signals coming from beneath the ice surface. Several hardware components from the ANITA experiment were incorporated into the ARA and ARIANNA experiments.

Another important experiment in the search for neutrinos in the EeV energy regime is the Pierre Auger Observatory [27], located in Argentina. This experiment is a hybrid detector, which consists of a large surface particle detector covering about 3000 km^2 together with a fluorescence detector. It can measure cosmic ray or neutrino induced air showers via a surface array of 1660 water tank Cherenkov detectors. The array consists of a main array, where the detectors are placed 1500 m away from each other and a more dense part, where the detectors are separated from each other by 750 m. Each Cherenkov surface detector is self-powered and can communicate wirelessly. The fluorescence detector consists of 24 telescopes which overlook the surface detector from four vantage points. The measurements from these two components are then combined for full airshower reconstruction. This way the Pierre Auger observatory is able to study the origin and characteristics of cosmic rays with energies above 100 PeV but no neutrinos were detected in this energy range [28]. The limits of detection for neutrino fluxes for these experiments can be seen in Figure 4.

2 RNO-G

The focus of this work will be the Radio Neutrino Observatory Greenland (RNO-G) [15]. RNO-G is a novel observatory under construction at Summit Station in Greenland that aims to measure astrophysical neutrinos at the EeV energy level. RNO-G will be a showcase for the scalability of the radio detection technology to arbitrarily large arrays, while having the world's best ultra high-energy neutrino sensitivity. Probing the universe at the highest neutrino energies also allows for fundamental research on topics such as cosmology, high-energy astrophysical processes, and increase our understanding of long lasting problems surrounding dark matter and fundamental symmetries of nature. When looking at the multi-PeV neutrino landscape in Figure 4, it is clear that IceCube is only sensitive up to a few PeV, after which there is a gap in the spectrum which is then followed by potential observations from the Pierre-Auger observatory. RNO-G aims to fill in this gap covering energies above tens of PeV. The predicted sensitivity of RNO-G within the multi PeV neutrino landscape can be seen in Figure 5.

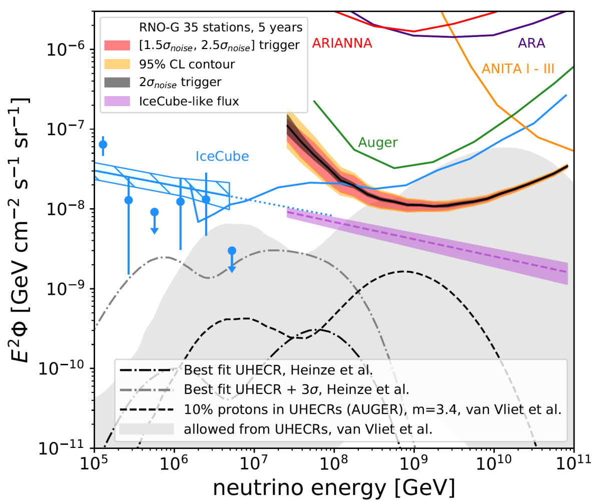


Figure 5: Current view of the multi-PeV neutrino landscape together with the neutrino fluxes from different experiments and the expected sensitivity of RNO-G. The red bands represent the differential sensitivity band spanning the interval from $1.5\sigma_{noise}$ to $2.5\sigma_{noise}$. In orange are the 95% contour levels. The sensitivity to be expected from a $2\sigma_{noise}$ trigger is shown in black. Lastly, the purple band shows the expected integrated sensitivity for an IceCube-like flux in the $1.5\sigma_{noise}$ to $2.5\sigma_{noise}$ range. (Credits: [15])

A visual representation of an RNO-G station and array design can be seen in Figure 6. The inter-station distance is chosen to be 1.25 km. Such a distance limits the effective volume but was opted for in order to restrict the logistical burden for installing the stations, while preserving the possibility of multi-station coincident events, which allows for unambiguous event identification and high precision reconstruction. RNO-G will consist of 35 stations that are installed over the span of four installation seasons. Three stations have already been installed during the first installation season in the summer of 2021 and are taking data. The data taking occurs in blocks of events called runs. Each of these three stations has already collected roughly 700 runs during the summer 2021 campaign [15]. The three stations that have been installed are station 11 , station 21 and station 22, numbered according to a matrix index like fashion; which have respectively also been named Nanoq, Amaroq and Avinngaq, reflecting animal names in Inuit language. Every station has 15 downhole antennas divided over three boreholes together with nine surface log-periodic dipole antennas (LPDAs) for a grand total of 24 antennas per station. One of the strings of each station contains a set of 4 compact Vpol antennas that serve as a phased array, just as was the case in one of the ARA stations.

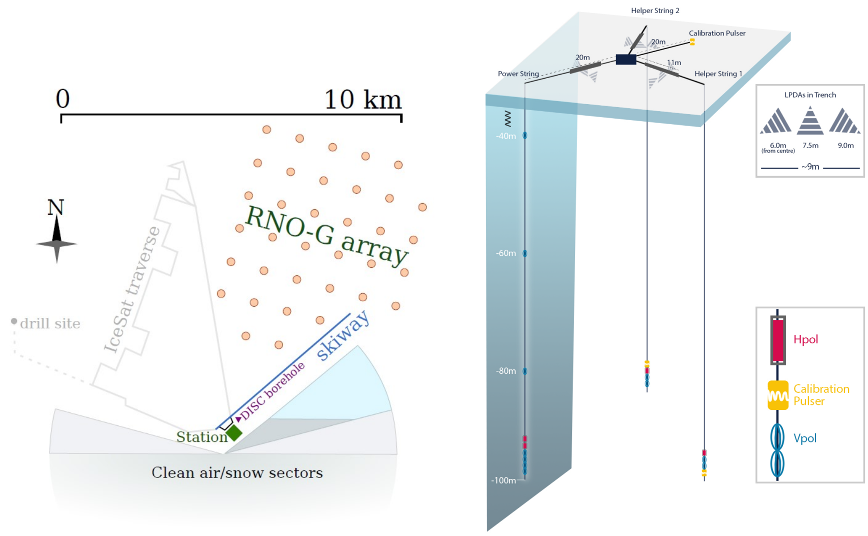


Figure 6: Top down view of the planned station array with 1.25 km spacing. Right: A single station containing three strings, each with Hpol and Vpol antennas plus LPDA surface antennas. It also contains two calibration pulsars. (Credits: [15])

Though RNO-G is a novel experiment that makes observations at the highest energies feasible, many of its aspects are preceded by other experiments. The technology and methodology used in the RNO-G experiment builds heavily upon the experience gained from RICE [19] , ARA [20] , ARIANNA [22] as well as the balloon-borne ANITA [26] experiments.

2.1 Method of detection

2.1.1 Askaryan radiation

Radio emission from a high-energy neutrino-induced particle cascade can be understood when considering the cascade in more detail. When a particle shower occurs in the ice many different interactions and processes take place as particles are scattered and interact with the ambient matter. The electromagnetic component of the particle shower changes over time. Compton scattering will upscatter electrons (hereby accelerating these electrons) while created positrons will annihilate with ambient electrons in the environment. This results in a net negative charge of relativistic electrons propagating in the shower front and a charge separation along the shower axis. The appearance of this net negative charge distribution at very high velocity leads to coherent Cherenkov radiation. The mechanism through which the above radiation process is generated is called the Askaryan effect [32]. Due to coherence this radiation is strongest at angles near the Cherenkov angle. The frequency of the emitted radiation is dependent on the shower geometry and is found to be coherent up to 1 GHz. The radiation has been observed to be a broad-band nanosecond-scale radio pulse [33].

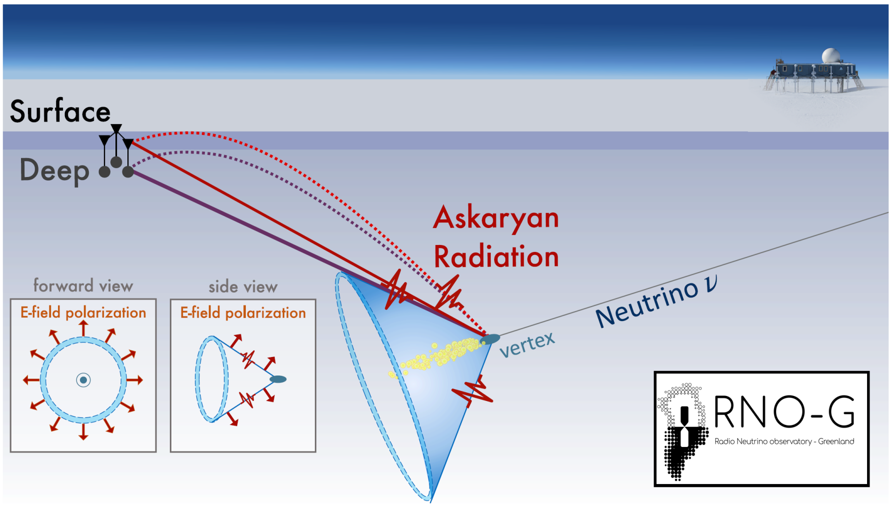


Figure 7: Graphical visualization of radio emission resulting from a neutrino interaction. Note the stronger emission near the Cherenkov angle and the bent trajectories, where the latter is due to a variable index of refraction in the top 150 meter of ice. Source: [15]

2.1.2 Background noise

Event purity is a very important factor for an experiment such as RNO-G and so background noise should be reduced to a minimum. The main sources of background in this case are incoherent thermal noise (resulting from the thermal agitation of charge carriers in the electronic equipment), impulsive anthropogenic noise

(noise coming from activities performed by living organisms e.g. the in situ science team using comms equipment etc.), radio impulses from cosmic ray air showers and galactic sky noise. In order to limit the background noise, the triggering is done from 100 m deep within the ice, at the bottom of the borehole and the LPDAs are used to veto non thermal background sources coming from the suface (both man-made sources and air shower remnants).

Besides the noise backgrounds described above, several other backgrounds are expected. One very interesting background is found in the radio emission from electron charge excess in cosmic-ray induced air showers, which is quite similar to the radio emission induced by neutrino interactions in the ice [15]. A main difference however is that due to the low density in air compared to ice, the cosmic ray particle can travel much longer and further. This in turn allows the geomagnetic field to exert force on the cosmic rays for a much longer time and hence bend them to a stronger degree than if the particle would travel through ice. Due to the stronger effect of the geomagnetic field on the particles in air than in ice, the air showers will exhibit more of a charge separation through the Lorentz force. This charge separation then causes so called geomagnetic emission, which in air showers is much larger compared to emission from the Askaryan effect. Nonetheless, due to its impulsive nature, signals from air showers are still quite similar to the denser in-ice showers and on top of that are much more frequent. This makes them suitable calibration signals if understood and tagged properly. Their similarity may be problematic since radiation of an air shower could refract into the ice and wrongly be interpreted as a neutrino signal. To prevent this, self-vetoing surface detectors, acting as dedicated air shower arrays, such as with ARIANNA are employed for both veto and calibration purposes.

2.1.3 Neutrino interactions

The incident neutrinos which interact with the ice will create particle showers induced by neutral-current $\nu_{e,\mu,\tau}$ ($\overline{\nu}_{e,\mu,\tau}$) or charged current ν_e ($\overline{\nu_e}$) interactions [10]. Since Askaryan emission is associated with electromagnetic showers, which also occur as within hadronic showers, the radio detector is sensitive to all neutrino flavors [29]. It should be noted that the radio detection method has a higher sensitivity for the purely electromagnetic showers caused by electron neutrinos (when looking below EeV energies) since the signature of both the charged and neutral current interactions of an electron neutrino are identical. At the highest energies these radio detection methods have about the same effective volumes for all neutrino flavours given a suitable detector geometry. The amplitude of the measured radio signal is directly proportional to the shower energy and therefore the energy of the initial incoming neutrino, this gives radio detection an energy threshold above the thermal noise floor at approximately 50 PeV per shower.

Another source of uncertainty comes from neutrino interactions in the ice. The normal model for bremsstrahlung and pair production of particles travelling through a medium is described by the Bethe Heitler formula [30]. This model however only describes interaction with a single isolated atom. For sufficiently high energies, these formulae are not valid anymore since adjacent interaction sites cannot be considered to be independent anymore due to quantum interference effects originating from the uncertainty principle. This effect is called the Landau–Pomeranchuk–Migdal (LPM) effect [16, 17] and in general results in an elongation of the shower profile.

2.1.4 Radio wave propagation and reconstruction

In environments such as the South Pole, propagation of radio waves through the ice is a non-trivial process [31]. The attenuation length depends on environmental parameters such as temperature and density c.q. depth. Another important issue is that the index of refraction is depth dependent. Therefore radio waves propagating through the ice do not move along a straight trajectory. While a part of the radiation propagates directly to the detector, another part is reflected on the discrete material transition between the ice surface and the air. This results in a reflected signal contribution at the detector, from which extra information can be inferred about the vertex location. The methodology of measuring both the direct and the reflected radio emission has already been studied and utilized by both the ARA and the ARIANNA experiments [53, 54]. Obtaining the shower energy can be done if one is able to reconstruct the vertex location. This can be done via timing of the wavefront, assuming the vertex is the origin of a spherical wave or by the amplitude ratio and timing difference of the directly detected emission and its reflection off the surface. Reconstruction of the arrival direction of the neutrino requires the arrival direction of the signal at the antenna and the position of the measured signal within the Cherenkov cone it was emitted in. The former can be obtained from timing measurements while the latter is obtained by measuring the polarization and frequency of the electric field since the polarization always points radially w.r.t. the shower axis (see Figure 7) and the frequency spectrum is dependent on the viewing angle (higher frequencies fall off strongest away from the Cherenkov cone) and therefore yields the viewing angle to the shower axis. Reconstruction, however, is not trivial. In environments such as the South Pole snow piles up on the surface and compresses itself under its own gravity and then recrystallizes, hereby creating the so called firn layer. In such a case the material density increases with depth resulting in an exponentially increasing index of refraction. More information on radio propagation in glacial ice can be found in [31].

2.2 Location

With the above aims for RNO-G in mind, its location should be one with thick homogeneous and cold ice as well as sufficient infrastructure (it has to be accessible either by plane or large vehicle) for installation and potential maintenance of the stations [15]. Since RNO-G also aims to develop technologies and build up experience for deployment of the radio component of IceCube-Gen2, it would be advantageous to operate in a site which bears close resemblance to the South Pole. Furthermore, the kilometer-scale attenuation length in ice naturally makes places with thick and cold ice favorable for deployment. This together with national infrastructure for the RNO-G host institutions practically only leaves Summit Station in Greenland and South Pole Station. The latter already houses the South Pole Telescope and the IceCube observatory, which are undergoing upgrades. Furthermore, due to the Covid pandemic, access to the South Pole became prohibited, meaning that currently there already is a large logistical burden present. Another aspect is that for source studies it is scientifically advantageous to have a radio array probing the highest neutrino energies in the same field of view as where IceCube is most sensitive at the lower energies. This would allow investigation of the same sources at an extended energy range, as outlined below. Because of these reasons, Summit Station in Greenland was chosen as the location for the RNO-G experiment.

Located on top of the Greenland ice sheet, Summit Station hosts a variety of scientific research programs. Its specific coordinates are 72°35′46” N and 38°25′19” W . Underneath it is some 3 km of glacial ice which has been measured to be very radio transparent [36]. The station itself consists of three main buildings: the ”Big House”, the Science & operations barn and the greenhouse/clinic. There is also the skiway, a berm storage area, distribution shack, emergency generator, some summer-use structures and ”Tent city”, where researcher visiting during the summer can stay. Furthermore, there is also an existing 740 m deep borehole called the ”DISC borehole” which can be used to send pulses to the array from different depths. This allows for checking the antenna positions and investigation of the radio properties of the ice.

The fact that RNO-G is located in Greenland has several implications. One of these implications is the sky coverage of the observatory, which is very important as it dictates what parts of the sky and therefore what sources it will be able to study. The continuous sky coverage of the detector, which is mainly determined by the geometry of its location in Greenland, and its large field of view allow RNO-G to study point sources of high-energy neutrinos [15]. RNO-G is mostly sensitive to an annulus of about 45 ° above the horizon at a certain point in time. As time passes this band of sensitivity moves due to Earth's rotation. This results in a larger total field of view of the detector. A visualisation of the field of view of RNO-G and multiple interesting point sources within it can be seen in Figure 8.

Given its unique field of view from the Northern Hemisphere, RNO-G will also yield strong scientific contributions. Its different field of view of the sky compared to that of other radio observatories such as ARA and the foreseen Gen2-radio array, makes it uniquely sensitive and complementary to other planned radio neutrino observatories in the Southern Hemisphere. The overlap of sky coverage with the IceCube Observatory allows the study of multiple interesting transient sources over a broad energy band. Even just a single event observed by a radio detector in the Northern hemisphere will yield a flux in a new energy regime and in the case of a complete absence of signals one can still extract an upper limit on the allowed flux through "multi-wavelength" neutrino observations.

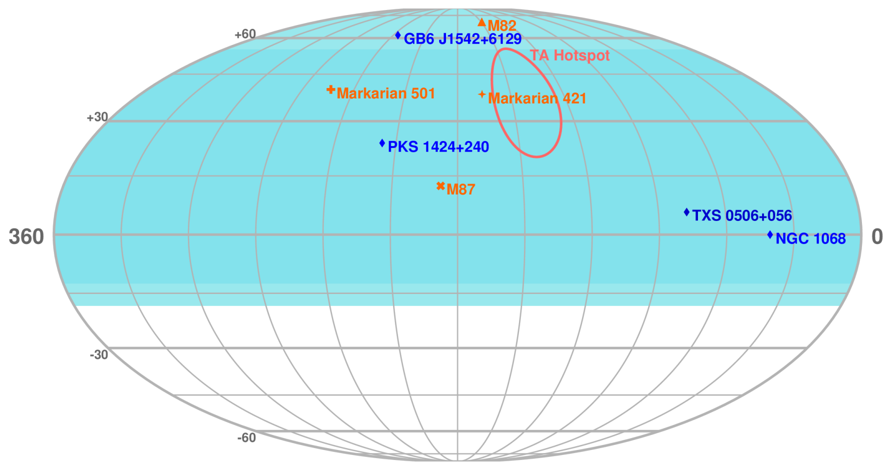


Figure 8: Visualisation of the field of view for RNO-G. In blue is its diurnally averaged total field of view, the sources in blue are possible sources detected by IceCube and in orange are other interesting candidates. The red circle denotes the so called "TA hotspot", a hotspot of ultra high energy cosmic rays observed by the Telescope Array [37]. (Credits: [15]).

2.3 Hardware and electronics

Given the circumstances in which RNO-G operates, its electronics are constrained in both their size (the boreholes have a radius of about 14.6 cm) as well as their power consumption. The design chosen for RNO-G consists of a deep component working in tandem with a surface component. A schematic of the system for a single RNO-G station can be seen in Figure 9. The power consumption depends on the operation mode of the system, where in a nominal mode, a single station uses 25 W.

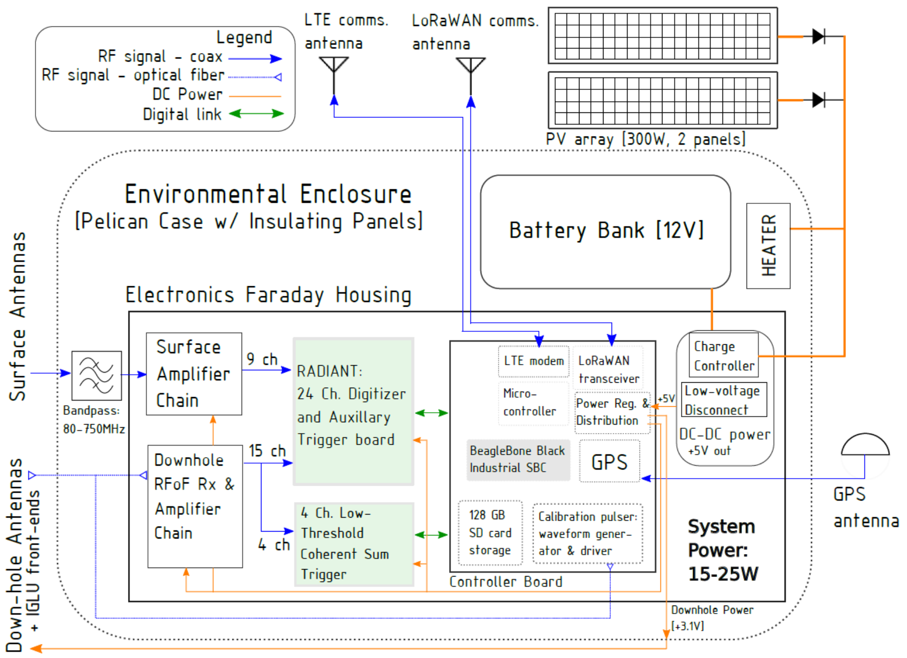


Figure 9: Schematic of the system for a single RNO-G station. Credits: [15].

2.3.1 Antennas

An RNO-G station utilises three different kinds of antennas [15]. The vertically-polarised (Vpol) antennas, the horizontally polarised (Hpol) antennas and the log-periodic dipole antennas (LPDAs) similar to the ARIANNA experiment. The latter are used by the surface components of an RNO-G station. The Vpol antennas are of the "fat dipole" design, which have an azimuthally symmetric beam pattern and have a usable bandwidth in the relevant frequency range for RNO-G (about 200-800 MHz). The "fat dipole" design has already proven efficient in previous neutrino experiments. For the Hpol antennas, so-called "cylindrical tri-slot"

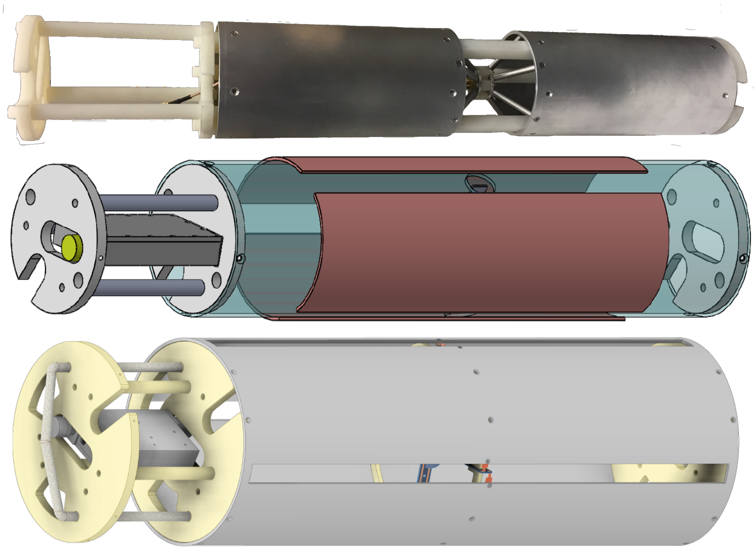


Figure 10: Top: picture of the "fat dipole" design of Vpol that is used in RNO-G stations. Mid and Bottom: different technical drawings for design options of the Hpol antennas. (Credits: [15])

antennas are utilised. The Hpol antennas suffer the most from the constraints due to the borehole dimensions. Due to this handicap for the Hpol, which is inherent to the vertical nature of the borehole geometry, other Hpol designs such as the Alford loop antenna [38] are being actively explored for future use. Both Hpol and Vpol antennas can be seen in Figure 10.

The signal from the Vpol and Hpol antennas can be combined for polarisation reconstruction for analysis purposes. Since the nine LPDAs are located on the surface, they are not constrained to the borehole dimensions and therefore have a larger gain than the other antennas. These allow the station to veto for signals coming from incident atmospheric neutrinos. The 24 channels all correspond to one of these three types of antennas mentioned and for identification purposes are labeled with a number from 0 to 23. Channels 0, 1, 2, 3, 5, 6, 7, 9, 10, 22 and 23 correspond to Vpols; channels 4, 8, 11 and 21 correspond to Hpols and channels 12, 13, 14, 15, 16, 17, 18, 19 and 20 correspond to LPDAs. A visualisation of the channel mapping can be found in Figure 11.

A short coaxial cable connects the feed of each antenna to a downhole front-end, after which the signal is amplified by a Low-Noise Amplifier (LNA) of the type IGLU. When the signal reaches the DAQ box at the surface via an RF over Fibre optical cable, it is again amplified by a Downhole Readout Amplifier Board

(DRAB) and converted back to an analog signal. Here the signals from the surface channels are also amplified, though these signals require less amplification than the signals from the downhole channels since they are stronger and only have to travel through the coaxial cable for roughly 20 m and therefore are much less attenuated. In order to protect the IGLU amplifiers from environmental factors and to reduce noise on all amplifiers, all amplifiers are encased in RF-tight housings. Every station is also equipped with a GPS, which allows it to synchronise event timing between stations as well as keep track of the interstation distance which is susceptible to variation due to the moving ice sheets. More information on the RNO-G antennas can be found in [15].

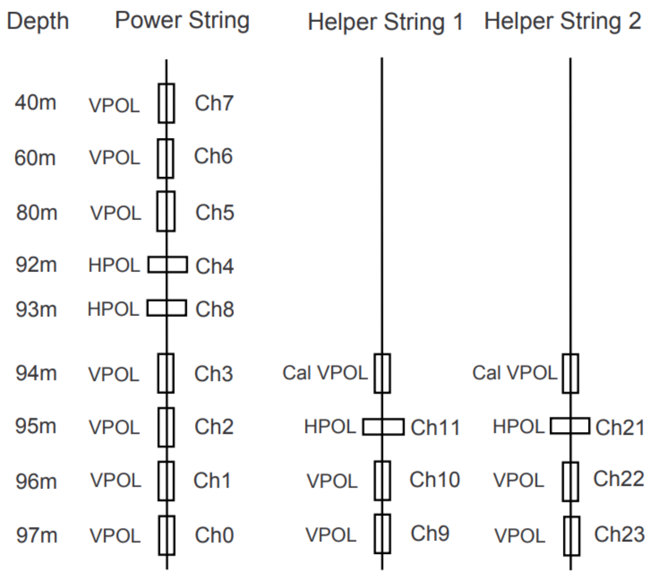


Figure 11: Channel Mapping of a single RNO-G station. Credits: Nick van Eijndhoven.

2.3.2 Triggering

The main trigger of RNO-G is inspired by the phased array from ARA but adapted to meet the criteria and constraints for operation in RNO-G (e.g. borehole dimensions, low power requirements, low cost per item, etc.) [15]. This trigger consists of a compact array of four Vpols and is located at the bottom of the main borehole string, at a depth of 100 meters. The phased array trigger sums single channel waveforms in a coherent manner while incorporating time delays corresponding to a range of angles of incident waves. This improves the trigger-level signal to noise ratio, which increases as the square root of the number of antennas present in the array. Using the four Vpol antennas from the phased array, together with the two Hpol antennas next to it, will allow for reconstruction of the frequency slope as well as the polarisation of the signal. Three deep boreholes with antennas are required for full directional reconstruction. Additionally, a surface calibration pulsar is foreseen as well. The LPDAs will be used for polarisation and timing measurements from surface events as well as a veto for airshower events.

2.3.3 Signal treatment

The radio signals that arrive at the antennas, are recorded and eventually stored as ADC counts. In order to be able to reconstruct the incident radiation from these ADC counts, different types of antennas focus on specific components of the radiation [15]. As mentioned before, the Hpols record the horizontally polarised component of the radiation while the Vpols do the same for the vertically polarised component of the incident radiation. By combining this information, the polarisation of the radio waves can be reconstructed. As can be seen in Figure 7, the polarisation of Askaryan radiation always points radially w.r.t. the shower axis. Measuring the polarisation allows to reduce the entire ring of potential neutrino origin to a possible patch in the sky. The radio signal is strongest at the Cherenkov cone itself. As the viewing angle increases, the signal strength weakens. However, this weakening is not the same for all frequencies, higher frequencies already fall off much stronger for lower viewing angles as a consequence of the shower profile elongation due to the LPM effect [39]. The measured frequency spectrum is therefore a function of the viewing angle as well. Consequently, a measure of the viewing angle can be inferred from the frequency slope of the signal. Combining the frequency slope and the polarisation of the signal leads to the off-Cherenkov signal angle and the neutrino arrival direction.

The Vpol antennas together with some Hpol antennas at different depths work together to obtain the azimuthal and zenith information for signals with any kind of polarisation. Two of the boreholes also host calibration pulsars, which are used for regular monitoring of the station performance and for calibration purposes.

2.3.4 Digitisation

Once the waveforms from the antennas reach the main system and have been amplified, they reach the RAdio DIgitizer and Auxiliary Neutrino Trigger (RADIANT) board, where the waveforms are digitized. Digitisation is done with a commercially available 12-bit ADC. The RADIANT board makes use of a LAB4D chip for every channel [34]; this chip is capable of recording at a rate of up to 3.2GSa/s and of performing multi-event buffering. For RNO-G it is chosen to operate in 2x 2048-sample buffers in order to obtain a near deadtime-less performance.

The RNO-G private LTE (long term evolution) network is responsible for data communication and operates at 900 MHz, which is the lowest frequency safely above our passband.

Each RADIANT board is controlled by a separate controller board. The RADIANT board reads the signals off of the antennas, converts them into a pedestal-subtracted signal in ADC counts, which is then transmitted to us via the LTE network. New pedestal values are computed from a dedicated pedestal data taking mode for every new run at the beginning of the run. Next to LTE the station and the detectors also make use of a LoRaWAN network in a 900 MHz NA ISM band (which almost conflicts with the LTE); on the detector side the LoRaWAN network provides timing for the microcontroller, commanding and uplink for housekeeping. In order to facilitate the making of plots for the housekeeping data from the pgsql database for the whole collaboration in the shortest possible time, Grafana [35] was chosen as it is relatively flexible and easy to use. Currently Grafana provides real time plots for some basic housekeeping quantities; possibilities are being explored to upgrade the interface as well as the amount of quantities represented in there. In case of a network outage, each station is also equipped with a local SD card for temporary data storage purposes. The data is sent from the RNO-G stations to Summit Station, where it is stored on a redundant disc array. A subset of the waveform data together with the instrument status data and the event metadata is sent to the University of Wisconsin via the Summit Station satellite link. RNO-G will use the JADE software [40], which was developed for and used by IceCube, for its data management.

2.3.5 Power

Since RNO-G is located in a desolate location, its stations have to be able to work autonomous to a large degree [15]. Hence power management is of large importance in the design and operation of the electronics. Every station is equipped with two solar panels, which have a total power output of 300 W in ideal conditions. Apart from that, the stations are also equipped with a 5kWh sealed lead-acid battery bank that can provide 24 W of power for three days when fully charged. Another option that is currently being explored by the RNO-G collaboration is to possibly add wind turbines in order to increase the reliability of consistent renewable energy at the stations; especially during polar nights, when the solar power is unavailable for long periods of time. The current plan consists of two side-by-side mounted turbines on a 6 meter tower together with a turbine mast box with rectification, charge controller, divert load and I2C monitoring [41]. In order to better manage power usage, the station has four different operation modes in which it can operate depending on the available power.

1. *Full-station mode*: features the power, trigger and data acquisition of all 24 channels together with the low-threshold trigger and the LTE data telemetry. This mode has a power usage of 24 Watts and is the most power consuming mode of the four.

2. *High-threshold mode*: same as the Full-station mode except it excludes the low-threshold trigger and only has minimal LTE data telemetry. This mode has a power consumption of about 17 Watts.

3. *Surface-only mode*: only foresees power, trigger and data acquisition in the nine surface LPDAs together with minimal LTE data telemetry. The power usage of this mode is about 6 Watts.

4. *Winter-over mode*: operating mode specifically reserved for during the polar night. Only the charge-controller, LoRaWAN network and station-control microcontroller have power in this mode. To further limit the power usage, data telemetry is kept to a minimum by only sending the essential housekeeping data. This results in a power consumption of about 70 mW.

The RNO-G station controller can autonomously decide to switch between these modes as well as receive commands remotely to switch to another mode.

3 Calibration efforts

This work focuses around the first RNO-G calibration efforts. The data is available for analysis in the form of root files [42] organised on a station and run basis. Each run has four associated root files, a file where the combined data is stored, one for the header, one for the daqstatus and one for the pedestal values. The combined data file consists of only a part of the full run data. Since this data is from the first RNO-G summer campaign, which occurred during the summer of 2021, the focus is on understanding and commissioning the detector, so no data analysis framework is yet set up. The following calibration efforts were thus developed within this thesis from scratch with the goals of performing a first exploration of this data.

The data analysis in this work is performed in Python [44], more specifically in Jupyter notebooks [45]. In order to retrieve the data stored in the root files, the Pyroot module [46] was utilised. Each run consists of a certain amount of events and each event consists of 2048 samples per antenna channel. For this work, only data runs containing randomly triggered noise data were used. There also exists pulser data (measurements when a calibration pulser was sending out signals) but these are not considered here.

On the deepest level, the main part of the raw data consist of consecutive ADC counts as measured during a single event (consisting of 2048 samples for each antenna channel). Since the data has not been investigated before, a general sweep of the data is required for all stations to identify any problems and potential anomalous behaviours. Once some problematic patterns in the data have been identified, their cause can be investigated and ideally these behaviours can be solved or calibrated out in the future.

3.1 Exploration of the data

A first general sweep of the data was performed in order to pinpoint any grotesque errors. This was done by visualising the average ADC values for all channels and all events in different runs. The results for a specific run and station can be seen in Figure 12.

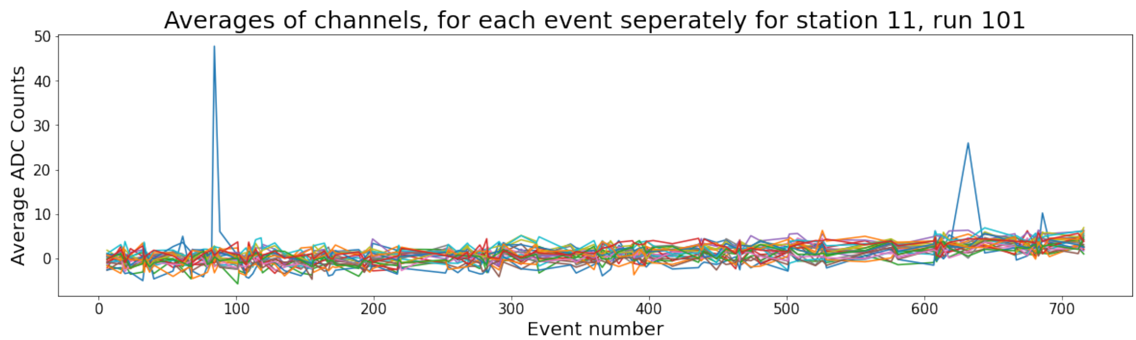


Figure 12: Average ADC count values for run 101, station 11 for each event. Each color represents a different channel of station 11.

Here it is instantly clear that for some events the average ADC count is way too high. On top of that, the average ADC counts per event seem to systematically increase over time. The three biggest spikes were investigated and turn out to all belong to channel 10 of station 11. From this it is clear that channel 10 can be quite problematic at times. In order to further investigate the strange behaviour of this channel, the timetrace of the measured signal from these events are investigated. The timetrace for event 632 can be seen in Figure 13. It can be seen from this timetrace that there is some kind of an offset that is present in certain blocks of the data, where these blocks are consistently of the same length. This phenomenon has been called "Waveform baselining", and is thought to be connected to the physical windows of the buffer in the LAB4D chip. Therefore a logical first place to look for potential errors is the pedestal values computed by the LAB4D chip, which will be discussed below.

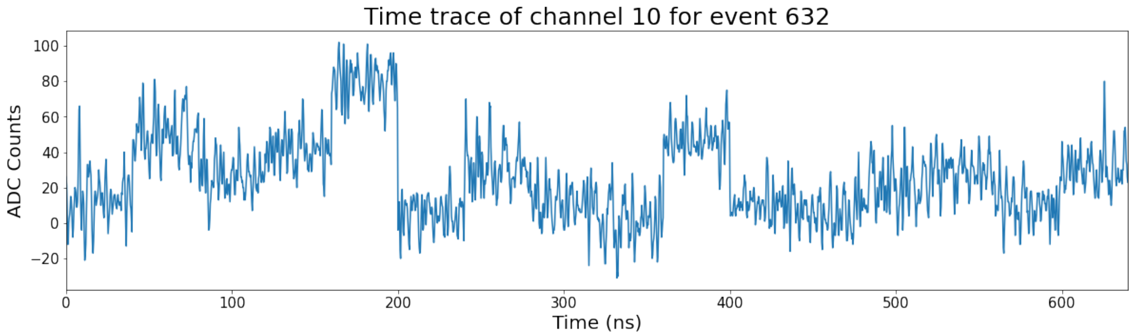


Figure 13: Timetrace of event 632 from run 101 of station 11, channel 10 .

3.2 Pedestal

When considering a waveform such as those in the data, one can look at many different aspects which characterise the waveform. Since a waveform can be seen as variations with respect to a central value, an important characteristic value is the mean value of the waveform. The mean value of a waveform is often called the "DC offset" or "DC bias". Since the LABRADOR ASIC ADCs, a class to which the LAB4D chip belongs, feature a usable signal voltage range of 0–2.5V, no negative signals may be propagated [34]. In order to work around this, a DC offset called a "pedestal" in this context is introduced to make sure the voltage remains positive. Each storage cell in the main storage array of the LAB4D chip develops a slightly different DC offset. In order to retrieve the original signal, the pedestals must be subtracted from the output. Therefore it is crucial that the pedestal values are monitored and stored as well. The pedestal values are taken at the beginning of the run for all channels and stay the same for all events within a run. Both the pedestal values for the problematic example run 101 from station 11 and the resulting timetrace when these pedestals are applied to the run data

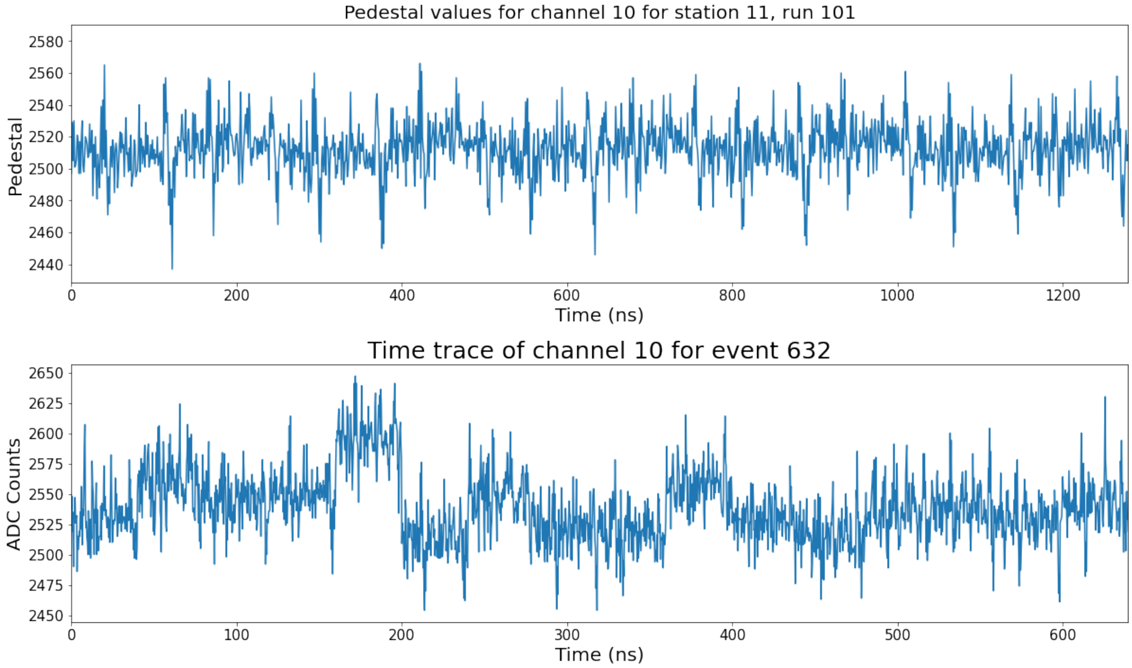


Figure 14: Top: pedestal values associated to station 11, run 101 by the LAB4D chip. Bottom: The timetrace from Figure 13 before the pedestal was subtracted.

can be seen in in Figure 14. Re-adding the pedestal clearly does not solve the unphysical behaviour of this kind of runs, therefore the cause of this does not lie in an incorrect subtraction or computation of the pedestal values but rather somewhere else. We further note the unexpected sharp spikes in the pedestal values in Figure 14, the cause of this problematic behaviour in the pedestal values is still under investigation.

Since the cause of the waveform baselining likely has nothing to do with the pedestals, we are forced to look at other possibilities. This behaviour might also be the result of component failure within the electronics. If this is the case, the problematic hardware component should be identifiable when inspecting the data on a basis of this component type. In this work a piece of code was developed which allows the investigation of the performance of the storage windows within the LAD4D chip. Though the performance of some physical windows in e.g. channel 10 of station 11 is definitely flawed, the waveform baselining does not appear to be attributable to a subgroup of malfunctioning storage windows as outlined hereafter.

3.3 Inspecting window problems

As mentioned above, the waveform baselining might be the result of component failure of some piece of hardware in the RNO-G station. Due to the blocky nature of the offsets in Figure 13, a prime suspect for component failure are the storage windows from the circular buffers in the LAB4D chip [34]. A block diagram of the LAB4D chip can be seen in Figure 15. The components of interest for this inspection are the storage arrays. The LAB4D chip has 32 sampling windows, each of these sampling windows can contain 128 samples for a grand total of 4096 samples. The LAB4D chip has undergone a firmware upgrade , which changed the data storage procedure. The relevant differences that this upgrade ensues are discussed further below. Depending on the firmware version, there are two buffers which can be read out at once, each buffer has 8 windows and consequently can hold 1024 samples. One event is stored onto two buffers and therefore 16 windows, which corresponds to 2048 samples. Since the chip works at a sampling rate of 3.2 Gsa/s, an event is always 640 ns long. The storage system consists of a primary storage array, a secondary storage array and a main storage array. The primary storage array in turn consists of two 64 sample/hold cells which sample the incoming RF signals. These then transport the signal to the secondary storage array, where each 64 sample/hold cell from the primary storage array communicates the samples to a separate A and a B part of the secondary storage array. When intermediary storage array 1 (consisting of A1 and B1) is full, i.e. 128 samples have been written to it, the focus switches over to intermediary storage array 2 and vice versa. Then both A components (B components) are combined to a 128 sample signal and communicated to the main storage array, where they are stored on a circular buffer with even (uneven) window numbers. After the windows are full, the event is written away. To summarise this, writing SW to indicate sampling windows, for the old firmware the data storage is structured as

$$
\begin{aligned}
&\underline{\text{Full capacity, used for pedestal recording}}:\\
&1\text{ Station} \rightarrow 24\text{ Channels}\\
&1\text{ Channel} \rightarrow 4\text{ Buffers} \xrightarrow{\text{8 SW / Buffer}} 32\text{ SW} \xrightarrow{\text{128 Sampling cells / SW}} 4096\text{ Samples}
\end{aligned}
$$

Such that:

$$
\begin{aligned}
&\underline{\text{Data taking mode}}:\\
&1\text{ Run} \rightarrow \text{x Events}\\
&1\text{ Event} \rightarrow 2\text{ Buffers} \xrightarrow{\text{8 SW / buffer}} 16\text{ SW} \xrightarrow{\text{128 Sampling cells / SW}} 2048\text{ Samples.}
\end{aligned}
$$

Where x denotes the amount of events in the considered run, which can vary per run. And for the new firmware:

$\underline{\text{Full capacity, used for pedestal recording}}$:

1 Station $\rightarrow$ 24 Channels

1 Channel $\rightarrow$ 2 Buffers $\xrightarrow{\text{16 SW / Buffer}}$ 32 SW $\xrightarrow{\text{128 Sampling cells / SW}}$ 4096 Samples

Such that:

$\underline{\text{Data taking mode}}$:

1 Run $\rightarrow$ x Events

1 Event $\rightarrow$ 1 Buffer $\xrightarrow{\text{16 SW / Buffer}}$ 16 SW $\xrightarrow{\text{128 Sampling cells / SW}}$ 2048 Samples.

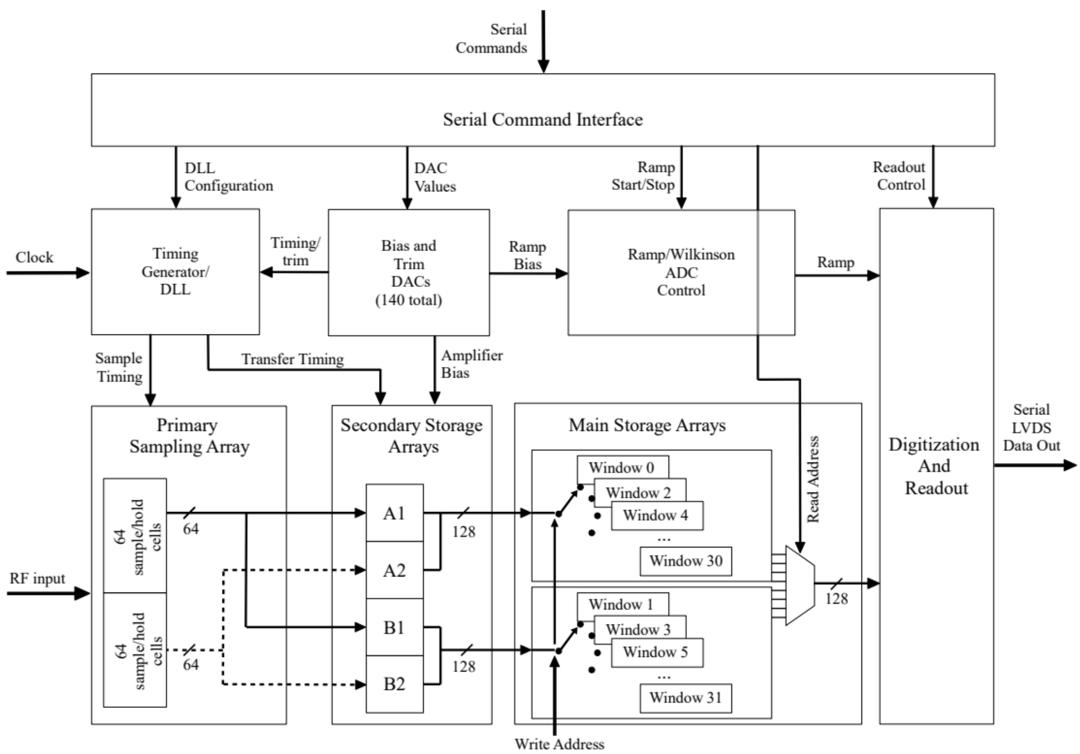


Figure 15: Block diagram for the LAB4D chip. (Credits: [34])

The fact that the whole connection through the storage arrays has two identical parts allows the system to work in a sort of "ping-pong" manner between the two connections, leading to a near deadtimeless performance. Circular buffers have the advantage that it does not matter where in the buffer it starts storing data, it will just fill the circular buffer up to the same length every time. This makes the system extra flexible concerning data storage. This does mean that the first window in a run sample is not always the same physical window. For analysis purposes, the starting window number of the physical window is also stored. However, the electronics have undergone a firmware upgrading and the storage buffers are different among these firmware versions. With the old firmware, there were two circular buffers, each consisting of eight windows while the new firmware works with only one buffer which has 16 windows.

The difference between the chronological sampling windows (order of windows as they appear in the waveform) and the physical sampling windows (with the order of windows as shown in Figure 15) is important to take into account when looking for malfunctioning windows and when re-adding the pedestal values to the waveform since the pedestal values are always stored in order of the physical windows. A list of two starting window numbers is stored for each event and channel, their interpretation depends on the firmware and can be seen in Figure 16.

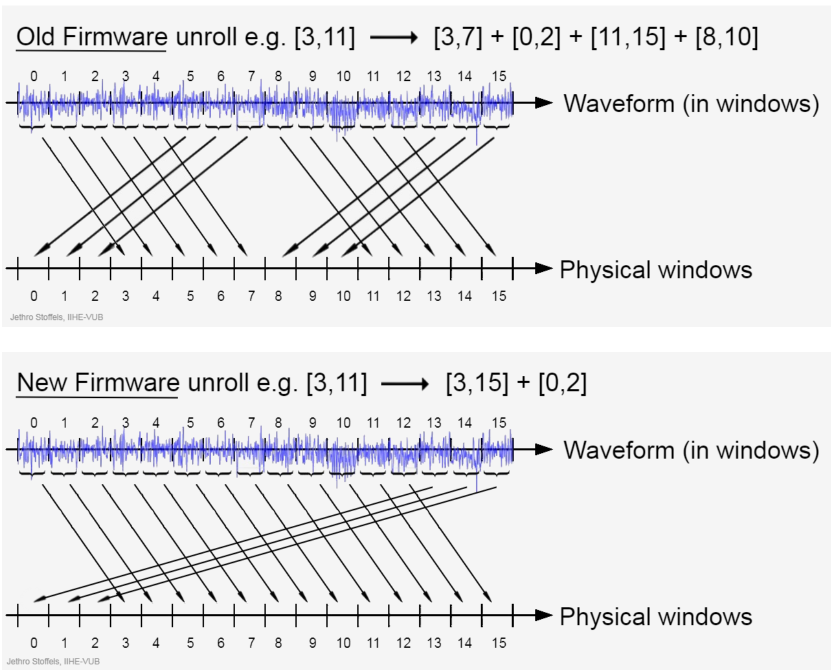


Figure 16: Top: The unrolling procedure for the old firmware of the LAB4D chip. Bottom: The unrolling procedure for the new firmware of the LAB4D chip.

Taking the correct unrolling into account, the average values of station 11, run 101 are computed on a per channel and per physical window basis. Inspecting a plot of these values allows us to see whether there is any anomalous behaviour in one specific physical window. If so, this could point towards a faulty physical window in the electronics of this station itself. Within this work, a procedure has been implemented that plots the average physical window values for all channels given some run data, and the results for channel 10 and channel 11 from station 11 during run 101 are shown in Figure 17. This plot was restricted to these channels since channel 10 showed to be the only deviant from the norm while all other channels have very similar behaviour as channel 11. Though not clearly visible for this specific run, it is frequently seen that the averages jump up and down on a per physical window basis such that the even (odd) physical window numbers lie higher than the odd (even) physical windows. This reflects the fact that the LAB4D chip works with a dual system, as mentioned before, where the odd and even physical windows are separated and experience a different bias voltage. The results from this procedure do not indicate a component failure for a specific physical window in channel 10 from station 11. Hence, in order to further study potential causes for the anomalous behaviour, another procedure has to be developed.

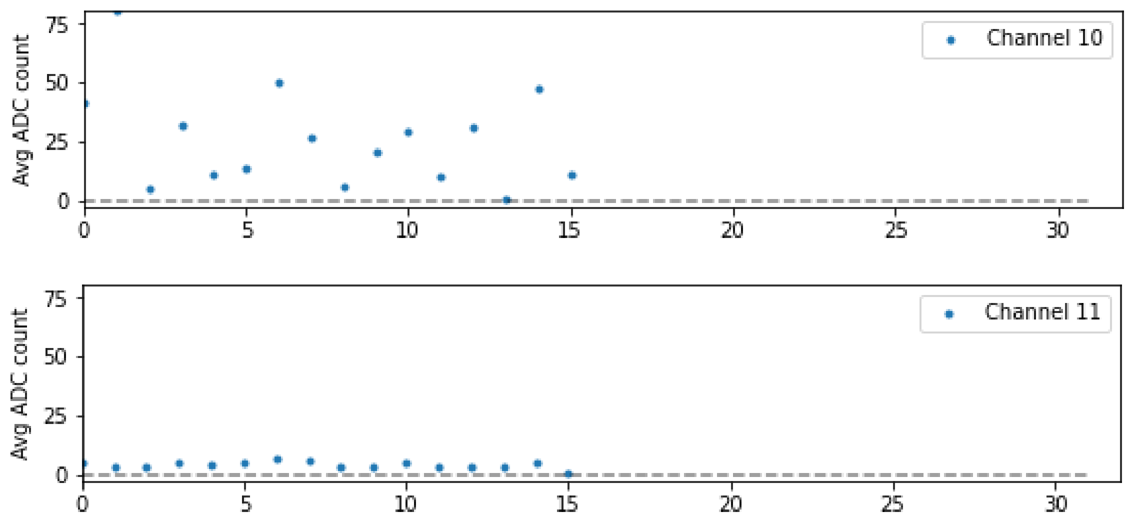


Figure 17: Average values of run 632 for channel 10 (top) and channel 11 (bottom) for each physical window number on the horizontal axis.

A different way to inspect the channels and their windows for strange behaviour is by looking at the distribution of the ADC counts of the noise data. Assuming a pure noise input, the expected distribution is Gaussian and has to be symmetrical. Any waveform baselining will be clearly visible in the form of a bump towards higher ADC counts added on top of the Gaussian-like distribution. The ADC count distributions for station 11, channels 9,10 and 11 during run 101 can be seen in Figure 18. All channels seem to have a similar Gaussian-like distribution such as those of channel 9 and 11, except for channel 10, as expected from our previous findings.

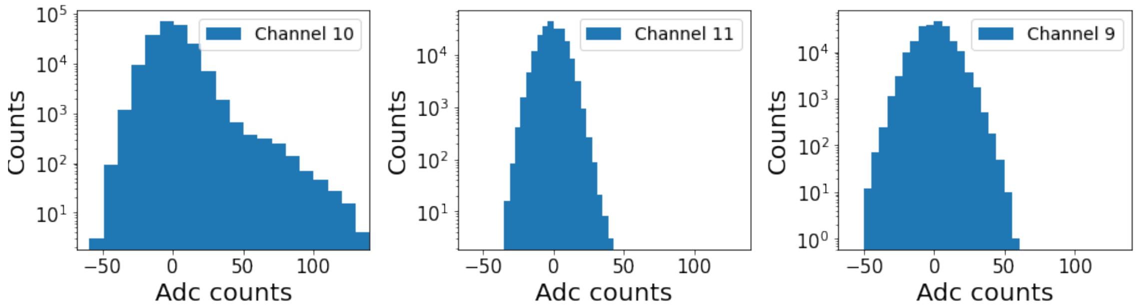


Figure 18: ADC count distribution of station 11, run 101 for channels 9, 10 and 11. The ADC count distributions for the other channels can be seen in Figure 28.

The ADC count distribution for each station in the case of multiple runs can be found in Appendix A. Such an ADC count distribution overview has also been made for each individual physical window. An example of this can be seen in Figure 19, where the ADC count distributions for physical windows 13, 14 and 15 of channel 10, station 11 during run 101 can be seen.

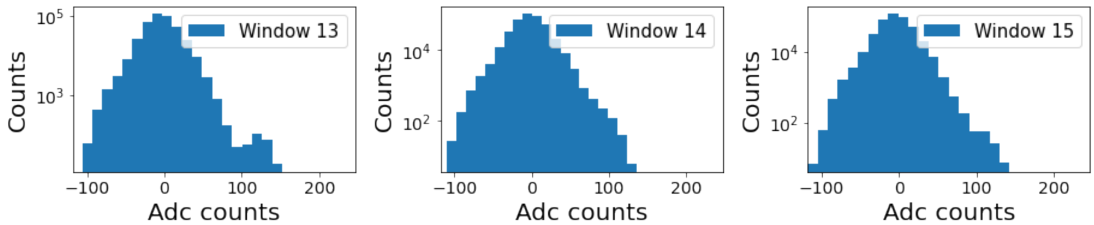


Figure 19: ADC count distributions for physical window 13, 14 and 15 of channel 10, station 11 during run 101. The ADC count distributions for the other windows can be seen in Figure 29.

As expected, the ADC count distribution of the physical windows of channel 10 show some Gaussian-like shapes together with a bump located at higher ADC counts. From these plots it is clearly visible that the problem is not due to one particular malfunctioning physical window, as these bumps are present in most of the physical windows. To enable a systematic investigation, an analysis framework has been developed within this thesis to make these figures for any combination of station number, channel number, physical window number and run number of interest. This allows for a sweep of new (or old) data obtaining the distributions such as shown in Figure 18 and if these distributions show any anomalies, one can make use of the ADC count distributions on a per physical window basis such as shown in Figure 19 in order to investigate the behaviour of the physical windows. Performing such a sweep on data from the other stations shows that station 21 has several bad runs with unphysical spikes as well, though these spikes are not instances of waveform baselining as is seen for station 11. Station 22 seems to behave as it should, without any abnormalities, over a long span of runs, as can be seen in Figure 32.

To conclude, using this procedure, the anomalous behaviour of channel 10 from station 11 has been visualised in a more meaningful and descriptive manner than the procedure based on the window average values as shown in Figure 17. Investigating the ADC count distributions has shown that station 21 also exhibits irratic behaviour, though upon further study of the associated problematic waveforms, it was concluded to result from a vastly different cause. No problems have been identified in the ADC count distributions of station 22. Though the cause for the anomalous behaviour in channel 10 from station 11 has not been deduced by usage of this procedure, it provides the stepping stone for quantification of noise power in well behaved channels.

3.4 Quantification of noise power

The fact that the noise data from station 22 behaves normal for long time periods extending over several runs makes its noise data a good candidate for quantification of the noise power. The type of noise that will be quantified in this work is thermal noise. As mentioned before, thermal noise is the result of thermal excitation of charge carriers in the electronic equipment, which results in artificial signals in the setup. In the case of the random voltage fluctuations in a resistor, it is also referred to as Johnson noise [43].

3.4.1 Thermal noise power

The analytical expression that describes thermal noise is dependent on the dimensions of the object in which the charge carriers are thermally excited. A simplified case can be made for thermal noise in a one-dimensional transmission line with single mode propagation (typically the Transverse Electromagnetic Mode TEM) [43]. The setup for which the power spectral density has to be computed is shown in Figure 20.

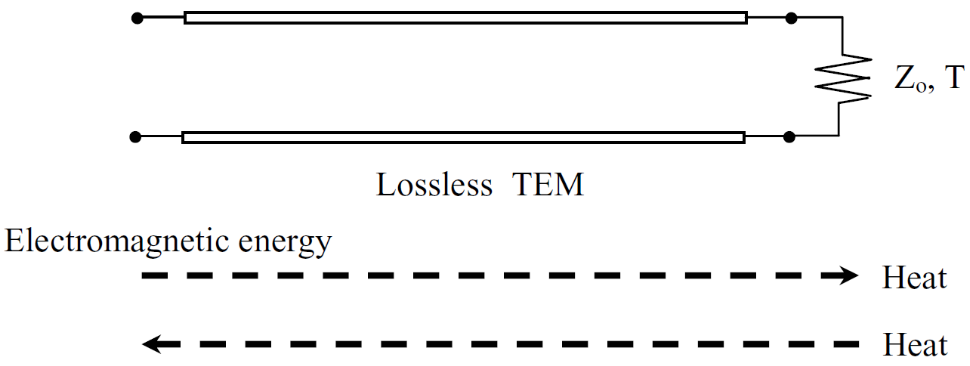


Figure 20: Visualisation of the setup for deriving the formula for the power spectral density of a TEM wave emerging from the left hand of a transmission line and the relation between the electromagnetic power and the heat dissipation. (Credits: [43])

Consider a closed container with length D which is in thermal equilibrium at a certain temperature T. This temperature is slightly coupled to the electromagnetic waves inside, i.e. the container is slightly lossy. The container may harbor an infinite amount of resonant modes and frequencies. A visualisation of this container can be found in Figure 21.

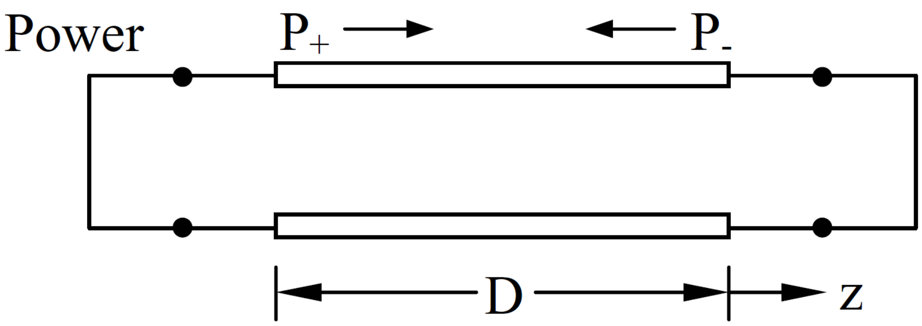


Figure 21: Visualisation of the closed and slightly lossy TEM transmission line resonator. (Credits: [43])

In order to calculate the thermal power spectral density P propagating to the right and left, the average energy density $W(\nu)$ has to be computed first, where ν is the frequency of the resonant mode. $W(\nu)$ can be calculated as

$$W(\nu) = \left(\frac{\text{Modes}}{\text{Frequency}}\right)\left(\frac{\text{Photons}}{\text{Mode}}\right)\left(\frac{\text{Energy}}{\text{Photon}}\right)\frac{1}{\text{D}}. \tag{1}$$

The "energy per photon" term is given by hν, where h is Planck's constant (h= $6.626\cdot10^{-34}$ J·s) and ν the frequency of the considered photon. In order to calculate the modes per frequency term, one should first note the fact that the voltage at each end of the transmission line is zero as a consequence of the short circuits there. This forces the resonant modes to be an integer multiple of half wavelengths along the transmission line. This constraint is visualised in Figure 22

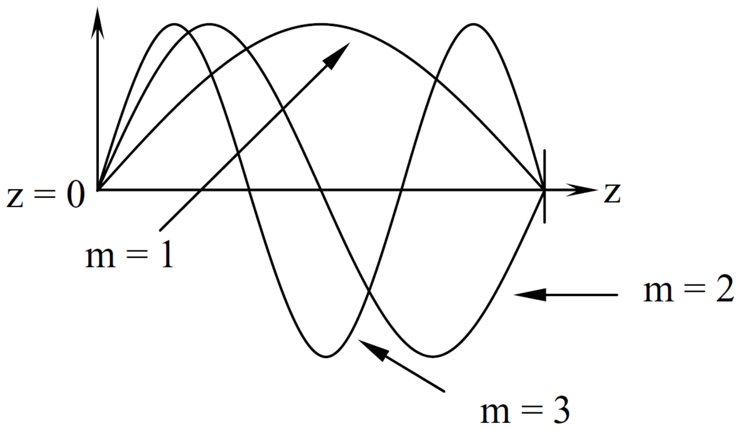


Figure 22: Constrained resonant modes possible in the closed slightly lossy TEM transmission line resonator. (Credits: [43])

Investigating a single frequency ν_{m}, the amount of half wavelengths that fit in length D is simply

$$\mathrm{m} = \frac{2\mathrm{D}}{\lambda_{\mathrm{m}}} = \frac{2\mathrm{D}\nu_{\mathrm{m}}}{\mathrm{v_p}}. \tag{2}$$

Where $\mathrm{v_p}$ is the phase velocity. The number of modes per frequency can be extracted from this by differentiating Equation 2 with respect to frequency

$$\frac{\mathrm{dm}}{\mathrm{d}\nu} = \frac{2\mathrm{D}}{\mathrm{v_p}}. \tag{3}$$

The average number of photons $\overline{\mathrm{n_j}}$ in the j*th* mode can now be obtained if an expression for the probability $\mathrm{p_j(n)}$, the probability of having n photons in the j*th* mode, is given. In the case of photons, this expression is obtained from Bose-Einstein statistics. From a statistical physics standpoint, the most likely configuration is dictated by combinatorics, with the probability distribution of photons among states proportional to the Boltzmann distribution exp(-nW_j/k_bT). Here, k_b is the Boltzmann constant (k_b =1.3806 $\cdot$ 10^{-23} J/K) and $W_j = h\nu_j$ such that the total energy in state j is given by nW_j, or:

$$p_j(n) = Q \cdot \exp(-nW_j/k_bT) \tag{4}$$

and now using the following equations

$$\sum_{n=0}^{\infty} p_j(n) = 1 \text{ and } \sum_{n=0}^{\infty} x^n = 1/(1-x) \text{ if } x < 1. \tag{5}$$

one can compute the proportionality factor Q via

$$\sum_{n=0}^{\infty} p_j(n) = Q \sum_{n=0}^{\infty} (\exp(-W_j/k_bT))^n = \frac{Q}{1-\exp(-W_j/k_b)} \tag{6}$$

In order for this to be unity, Q needs to be equal to $1 - \exp(-W_j/k_b)$. Therefore $p_j(n)$ is given by

$$p_j(n) = (1-\exp(-W_j/k_b))\exp(-nW_j/k_bT). \tag{7}$$

Now that an expression of the probability density of photons among states has been obtained, the average number of photons $\overline{n_j}$ in the j*th* mode can be computed:

$$\overline{n_j} = \sum_{n=0}^{\infty} n \cdot p_j(n) = (1-\exp(-W_j/k_bT)) \sum_{n=0}^{\infty} n\,(\exp(-W_j/k_bT))^n\,. \tag{8}$$

Which can be simplified by using that

$$\sum_{n=0}^{\infty} n \cdot x^n = x\frac{d}{dx}\sum_{n=0}^{\infty} x^n = x\frac{d}{dx}(1-x)^{-1} = \frac{x}{(1-x)^2}. \tag{9}$$

Using this in Equation 8 yields

$$\overline{n_j} = (1-\exp(-W_j/k_bT))\left[\exp(-W_j/k_bT)/\left(1-\exp(-W_j/k_bT)\right)^2\right] \tag{10}$$
$$= 1/\left(\exp(-W_j/k_bT)-1\right). \tag{11}$$

Now writing the energy W_j associated to state j as $h\nu_j$, the average number of photons $\overline{n_j}$ in the j*th* mode (from now on called the *photon state density*) is given by

$$\overline{n_j} = 1/\left(\exp(-h\nu_j/k_bT)-1\right). \tag{12}$$

Substitution of Equations 3, 12, and the photon energy $\mathrm{h}\nu$ into the first three terms of Equation 1 respectively yields

$$W(\nu) = \left(\frac{2\mathrm{D}}{\mathrm{v_p}}\right)\left(\frac{1}{\exp(-\mathrm{h}\nu/\mathrm{k_b T}) - 1}\right)(\mathrm{h}\nu)\,\frac{1}{\mathrm{D}} = \frac{2\mathrm{h}\nu}{\mathrm{v_p}\,(\exp(-\mathrm{h}\nu/\mathrm{k_b T}) - 1)} \tag{13}$$

for the average energy density. The energy density $\mathrm{W}(\nu)$ can be related to the power flow inside the closed resonator. Let W_+ and W_- represent the uncoupled energy flows to the right and left respectively. Due to symmetry, these are equal in magnitude, hence $\mathrm{W}(\nu) = \mathrm{W}_+ + \mathrm{W}_- = 2\mathrm{W}_+$. The power flowing to the right P_+ is simply equal to W_+ multiplied by the group velocity $\mathrm{v_g}$, which for a nondispersive transmission line is equal to the phase velocity $\mathrm{v_p}$. Combining this information one gets that

$$\mathrm{P}_+(\nu) = \frac{\mathrm{h}\nu}{\exp(-\mathrm{h}\nu/\mathrm{k_b T}) - 1} \tag{14}$$

This expression reflects the thermal noise per mode, its behaviour is visualised in Figure 23. From this figure it is clear that for low frequencies the thermal noise spectral density is constant as a function of frequency, i.e. for low frequencies there is a white-noise region. When the frequency increases, the thermal power spectral density eventually decreases and reaches zero. Between these two regions there is a transition point ν_0 ($\mathrm{f_0}$ in Figure 23). The value of the transition point ν_0 is depending on the temperature. The flat region is where $\mathrm{h}\nu << \mathrm{k_b T_0}$ (with $\mathrm{T_0} = \mathrm{h}\nu_0/\mathrm{k_b}$), which is called the "Rayleigh-Jeans approximation". Since it lies at the low frequency end of the spectrum, it is mainly applicable in domains dealing with the radio spectrum as opposed to the optical spectrum.

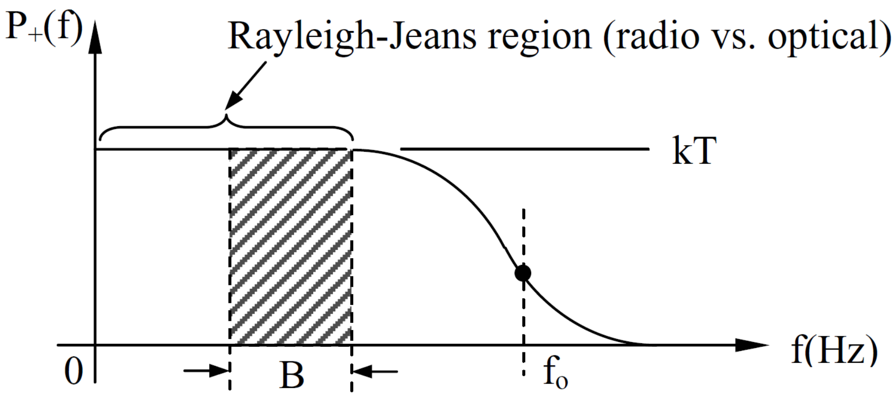


Figure 23: Thermal power spectral density P_+ in a TEM transmission line. In this figure, f denotes the frequency and k the Boltzmann constant. (Credit: [43])

In the Rayleigh-Jeans limit, Equation 14 can be simplified by using a standard Taylor expansion: $\exp(x) \approx 1+x+x^2/2+... \approx 1+x$ for $x << 1$. This simplification yields

$$P_+(\nu) \approx k_b T \tag{15}$$

If one now considers a frequency band $\Delta\nu$ which fully lies within the Rayleigh-Jeans limit (such as region B in Figure 23), the total thermal power propagating down the single mode transmission line from a matched load at temperature T is given by

$$P = k_b T \Delta\nu. \tag{16}$$

Where as mentioned before, P is the total thermal power within the considered bandwidth $\Delta\nu$ and T is the temperature of the matched load. The full derivation of this formula can be found in [43], Chapter 2. Since the formula was derived within the Rayleigh-Jeans limit, it is mostly used for radio systems.

3.4.2 ADC count to antenna voltage conversion

When looking at waveform data one can consider it in either the time domain of the frequency domain. Switching from one domain to another is done via the Fourier transform. There are many different implementations of the Fourier transformation, the most popular of which is the Fast Fourier Transform (FFT) as was developed by Cooley & Tukey in 1965 [47]. Fourier transforming a waveform from the noise data yields its frequency components, which allows for easier manipulation of magnitudes when the manipulation is frequency dependent. When an electric field arrives at the antennas, it is converted into a voltage in the system. The conversion factor for this is the antenna effective height. After that, as mentioned in Section 2.3, the voltage is amplified and then converted into ADC counts. In order to compute the thermal noise temperature, the measured noise data as it was recorded has to be transformed backwards through this whole chain of manipulations all the way up until it represents the voltage at the antenna. The stepwise plan of action for this is first converting the ADC counts from the recorded noise data back to units of voltage. Then, these voltages have to be divided by the amplification factor, which yields the voltages at the antenna. After that, the total noise power can be computed from these voltages, after which the thermal noise temperature can be retrieved via Equation 16.

Conversion of the recorded waveform data from ADC counts to voltage is done via voltage bias scans performed on the RADIANT boards in the lab. The voltage calibration factors turn out to be temperature dependent, this dependence has not been taken into account yet as it is still under research. As such, the voltage calibration factors used in this work are given by $V = 0.000618N - 0.008133$ [48], where N represents the ADC counts and V the associated voltage. After the conversion to ADC values, the amplification factor originating from the amplifier has to be removed, which is frequency dependent. In order to remove this, the magnitude of the amplifier response has to be known. This is retrieved via the NuRadioMC [1] package [55], which allows the retrieval of these amplifier responses from the hardware parameter database for different types of amplifiers (which is important to keep track of, since the type of amplifier used is channel dependent). The amplifier response (aka amplification factor) for the IGLU amplifier used for the downhole antennas can be seen in Figure 24.

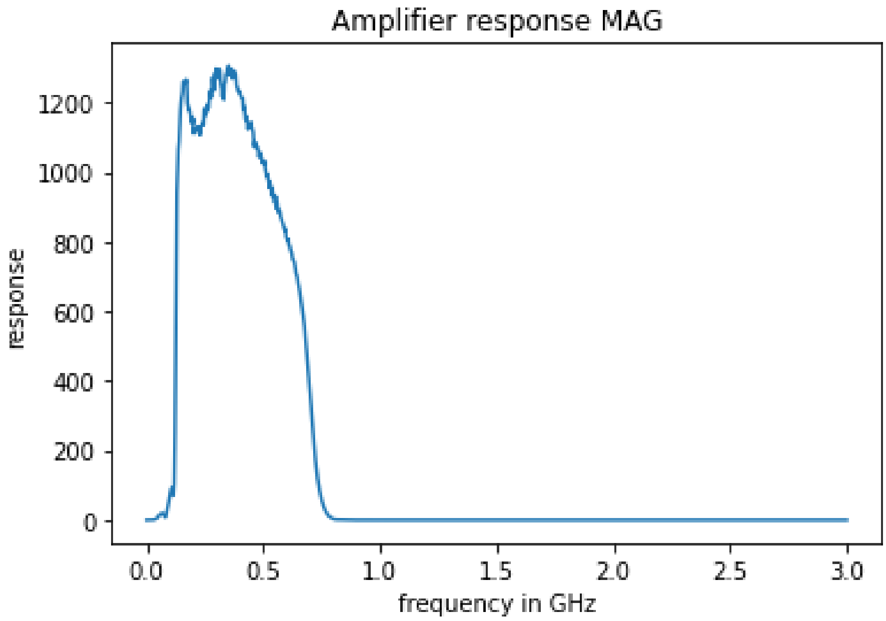


Figure 24: The amplifier response as a function of frequency for the IGLU type amplifier. Note that the response and therefore the amplification factor for frequencies outside the domain of study is zero in order to filter these frequencies out.

Dividing the amplitudes of the frequency spectrum of the noise data by this amplifier response (in the regime where the amplifier response is nonzero) then yields

[1]The NuRadioMC software provides an integrated package for both simulation and data analysis.

the signal frequency spectrum before it was amplified. In the regions where the amplifier response is zero, the frequency spectrum of the pre-amplified signal was set to zero. These components were surely nonzero before amplification, but in the context of noise estimation for signals detected by RNO-G, these components are not of interest anyway as this frequency regime is excluded from the RNO-G frequency range of study.

As can be seen in Figure 24, the amplification in the region of interest is of the order of 1300. The amplifier response however goes from such a large amplification factor to zero in a continuous manner, which as a result introduces unphysical jumps in the resulting pre-amplified frequency spectrum in the regimes where the amplification factor approaches zero but is not exactly zero. Such an enormous increase brought forth by these jumps would introduce a large amount of (unphysical) noise power. Therefore these jumps have to be cut off at a certain threshold value. It is to be expected that the amplifier response is of the order of 1300, therefore a frequency range can be defined for which the resulting pre-amplified frequency spectrum is surely physical. In this region, the median amplitude and the standard deviation can be determined. A threshold value is then obtained from the median value and a set amount of standard deviations added to this. The procedure developed for this work starts off in the middle of this frequency range that is surely physical and extends this range until two consecutive amplitudes are both higher than the aforementioned threshold (the constraint for two consecutive higher amplitudes was put in place in order to eliminate singular outliers within the physical frequency range). This then defines the frequency range within which the frequency spectrum of the pre-amplified waveform will be accepted, the frequency spectrum for domains outside of this range are set to zero. The resulting voltage waveform is the one present at the antenna. Now, using the frequency spectrum of this waveform, the power spectrum can be computed via $P = V^2/R$, where V is the amplitude of the waveform for each frequency and R is taken to be 50 Ω. Once the power spectrum is computed, the power is integrated over frequency in order to obtain the total power of the signal at the antenna. The bandwidth $\Delta\nu$ is then taken to be the one from the frequency range for which the pre-amplified voltage waveform frequency spectrum was kept after the cutoff. Equation 16 can then be solved for T.

Some care has to be taken with which data the thermal noise temperature is computed since attempting to compute this from runs which show anomalous behaviour will result in an incorrect temperature. To compute the thermal noise temperature from a run, the waveforms from all events are stitched together and Fourier transformed. After this the procedure outline above is followed to extract the thermal noise temperature from this. In order to increase statistics this was done for multiple runs. With these temperatures the mean temperature and a standard deviation can be computed.

In this work the chosen runs which do not exhibit extraordinary behaviour are runs 91 to 111 from station 22, channel 7. These yield an average thermal noise temperature of 290.14 ± 14.72 °K. The same can be done for the other stations and channels if no anomalous behaviour is present in the considered runs or if the problematic behaviour is calibrated out in the future.

With the procedure outlined above, the voltage at the antenna can be computed from the recorded data. However, the physical quantity that arrives at the antenna, which is effectively being measured, is the electric field. Therefore, in order to compare data and simulation, it is important that a conversion between the two is possible. The relation between the incident electric field and the measured voltage is given by

$$E_{in} = h_{eff} V. \tag{17}$$

Where E_{in} is the incident electric field, V the measured voltage and h_{eff} the so called "effective antenna height" [56]. The antenna effective height can depend on many different parameters (e.g. in loop antennas it depends on the amount of loops among others), but in this case it mainly depends on the frequency and the polar and azimuthal angles of incidence. The antenna effective heights for the three different kinds of antennas in RNO-G can be seen in Figure 25

Using the effective height, the incident electric field coming from a certain direction can be computed from the measured voltages and therefore allows for reconstruction of the incident radiation.

3.4.3 SαS distribution modeling

So far, the quest to quantify the noise in the electronics has led to a characteristic thermal noise temperature, which already is of great importance. Ideally, the noise as seen in the electronic equipment is purely of the Gaussian kind. However, the ADC count distributions are clearly seen to contain a more impulsive component, which deviates from a pure Gaussian noise. A good noise model is important for reconstruction studies and hence the next step in the quantification of noise is to quantitatively characterise this noise in a non Gaussian manner since the noise is observed to contain transient events. In order to do so, another kind of distribution is employed in this work, namely the so called "Symmetric alpha stable" (SαS)

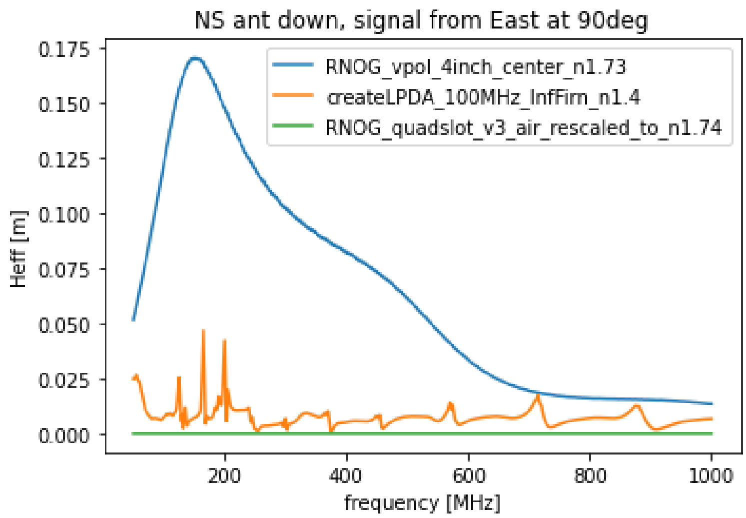


Figure 25: Antenna effective height for the three kinds of amplifiers present in an RNO-G station in case of radiation incident from the East under an angle of 90 degrees.

distribution [49]. The SαS distribution is a distribution that has a Gaussian-like shape but is more flexible regarding outliers. It can harbour much stronger and more complex tails than a Gaussian distribution. The SαS distribution is given by [49]

$$P(A, \mu, \beta, \alpha, \sigma) = \lim_{d \to \infty} \int_0^d f(A, \mu, \beta, \alpha, \sigma, t)dt. \tag{18}$$

When $\alpha \neq 1$, f is given by

$$f(A, \mu, \beta, \alpha, \sigma, t) = \frac{1}{\pi\sigma} exp(-t^\alpha) cos \left[\left(\frac{(A-\mu)t}{\sigma} \right) + \beta(t - t^\alpha) tan \left(\frac{\pi\alpha}{\sigma} \right) \right]. \tag{19}$$

In this parametrisation, $\mu \in \Re$ is the location parameter. It is equal to the median value of the distribution in case α is smaller than one, otherwise it is equal to the mean value of the distribution. Secondly, $\beta \in [-1, 1]$ is the skewness parameter, which is a measure of the asymmetry of the distribution around the mean. Thirdly, $\alpha \in [0, 2)$ is the tail index of the distribution, which is a measure of the heaviness of the tails of the distribution. The lower α is, the stronger the tails are and hence the more outliers it encompasses compared to a Gaussian distribution. Lastly, $\sigma > 0$ is the scale parameter. The scale parameter reflects the width of the distribution.

In order to obtain better performance for the numerical computation of the integral of Equation 18, the Gauss-Laguerre quadrature is utilised [50]. This method for the approximation of the value of integrals uses that

$$\int_0^\infty exp(-x)f(x)dx \approx \sum_{i=1}^{n} w_i f(x_i), \tag{20}$$

where x_i is the i-th root of the Laguerre polynomial $L_n(x)$ and the weights w_i are given by

$$w_i = \frac{x_i}{(n+1)^2(L_{n+1}(x_i))^2}. \tag{21}$$

A plot of the SαS distribution for a specific set of parameters as computed with the quad function from the Scipy package (as a brute force reference) and the aforementioned Gauss-Laguerre method can be seen in Figure 26.

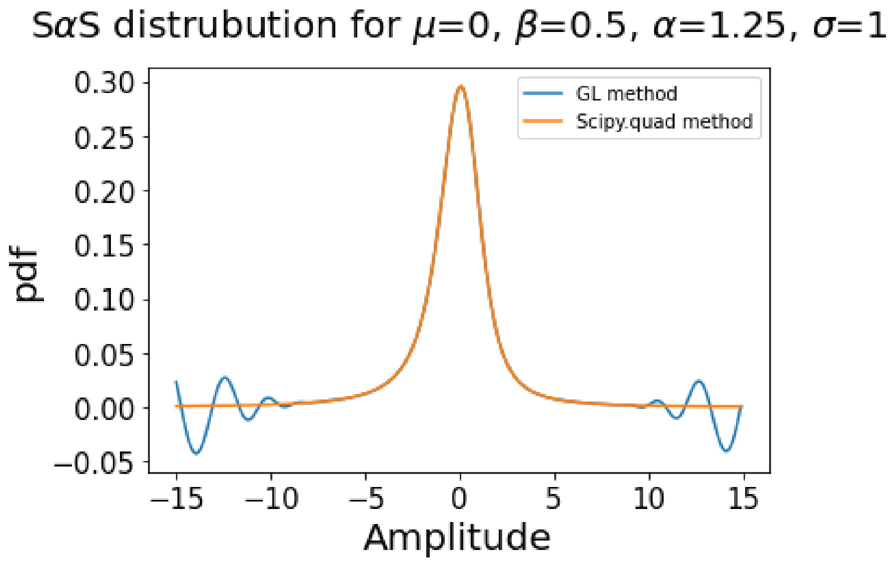


Figure 26: The SαS distribution for $\mu = 0$, $\beta = 0.5$, $\alpha = 1.25$ and $\sigma = 1$. In orange is this distribution where the integral is computed via the quad method from the Scipy package [51]. In blue is this distribution where the integral is computed via the aforementioned Gausss-Laguerre method (Equation 20).

As shown in Figure 26, the Gauss-Laguerre method shows some deviations further away from the center in the form of oscillatory behaviour. It is an inherent property of the Laguerre quadrature to undergo fast oscillation for large values away from the center. This oscillatory behaviour can be suppressed by adding a similar extra term which is shifted in the variable of the integrand [52]. This correction has not been taken into account in this work since the usage of the SαS distribution in order to quantify the noise has proven successful without implementing this term [49]. In the case of the computation of the distributions from Figure 26, the Gauss-Laguerre method was roughly 30% faster than the quad method from the Scipy package.

Fitting this distribution to the noise data from station 22, channel 7 during runs 99 to 103 yields a Full Width at Half Max (FWHM) of 32.3 ADC counts. The resulting fit as well as its parameter values can be seen in Figure 27. The same fit has been performed for all channels of station 22 in order to characterise its performance, and this fit and its parameters can be seen in Figure 33 and Table 1 respectively. The fitted parameters for this specific case show that the values for μ, β and α are quite stable across all channels. The value for σ however, seems to vary even by up to a factor of two in some cases, which could already have been inferred from a qualitative inspection of the plots in Figure 28.

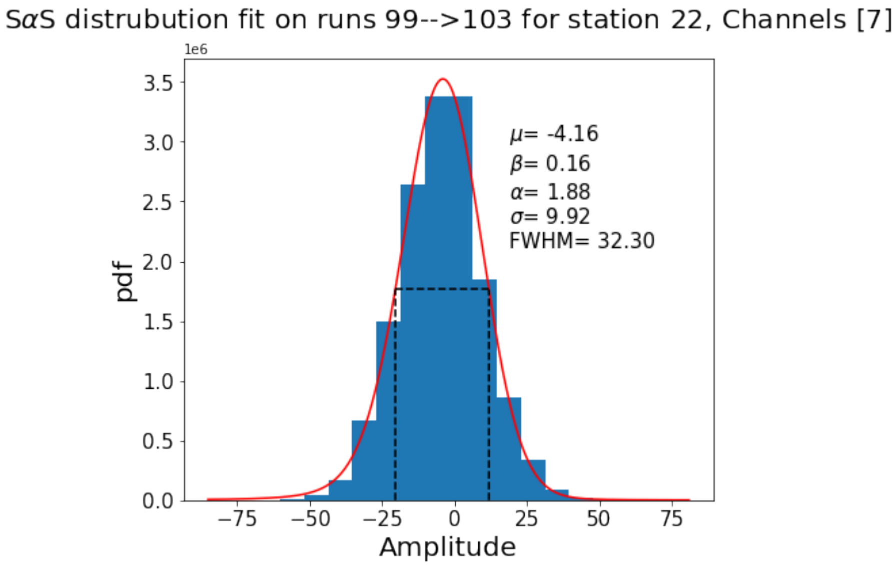


Figure 27: Fitting the SαS distribution to the noise data from station 22, channel 7 during runs 99 to 103.

3.5 Code developments

In order to perform the analyses above, various computational algorithms were developed within this thesis work. The language chosen for this task was Python, more specifically, the code is written in ten Jupyter Notebooks. Due to the repeated use of certain parts of the code, one of these notebooks is a collection of all the common functions and is written such that this notebook can just be imported at the beginning of a new notebook for easy and instantaneous access to all previously developed utilities. In Appendix B, a short description of the notebooks as they were developed in chronological order is given together with how they fit in this work.

4 Conclusion and outlook

In this work several important aspects of RNO-G and the importance of the RNO-G experiment in the context of modern day astronomy were highlighted. Contemporary experiments fail to efficiently measure the diffuse neutrino flux at the highest energies due to the steeply declining flux they are facing at these energies. In order to observe this low flux, a very large detector volume has to be instrumented, which forces us to utilise the radio detection technique instead of the current optical devices due to the much longer attenuation lengths of the radio signal. The RNO-G experiment aims to show the scalability of radio arrays to encompass the required volumes to detect these fluxes while maintaining an as of yet unparallelled sensitivity. During the last summer campaign, three of the in total 35 planned RNO-G stations have been installed and are taking data. This data first underwent a general sweep in order to identify any faulty runs and patterns in problematic runs with the aim of understanding their origin and to potentially solve the problem or calibrate these out in the future. One of these frequently reoccurring patterns which might be calibrated out in the future is the waveform baselining. This was mostly prevalent in station 11, channel 10. The origin of the waveform baselining was investigated and both a specific physical window malfunction as well as a pedestal subtraction have been ruled out as potential cause for this anomalous behaviour. Apart from a general sweep on a few fronts of the data, this work also presents a quantification and modeling of the noise. Firstly, a characterisation of the thermal noise was made via the thermal noise temperature, which was obtained by working backwards from the measured ADC counts to the voltage as measured at the antenna. The thermal noise temperature was found to be 290.14 ± 14.72 °K. Due to its clear deviations from purely Gaussian noise (caused by impulsive transient outliers), the more complex model of the SαS distribution was fitted to the noise data, from which a FWHM was obtained for a sequence of specific runs. The obtained SαS parameters show that the location parameter, skewness parameter and tail index of the distribution which characterises the thermal noise are quite consistent but the scale parameter seems to frequently vary by up to a factor of two.

With the new installation season at the horizon, the RNO-G experiment will soon see an increase of the amount of operational stations (for this summer a maximum of seven new stations will be installed). With these extra stations, the RNO-G array becomes the world's largest radio neutrino observatory, and a lot of extra data will be obtained and hence further calibration efforts will be required. General systematic calibration software needs to be developed such that an analysis of future data is easily possible. Anomalous behaviour such as the waveform baselining will have to be identified and corrected for. The complete RNO-G radio array will consist of 35 stations, so many more stations will be installed in the upcoming years, each possibly with their own problematic behaviour in data taking and much more calibration effort will need to be done. Once the stations have been properly calibrated, a first analysis can start in order to either determine the diffuse neutrino flux in the highest energy regime or put an upper limit to it.

References

[1] IceCube Collaboration. *Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector*. Science 342, 1242856 (2013). DOI: `https://doi.org/10.1126/science.1242856`. arXiv:1311.5238v2

[2] Zatsepin, G. T., Kuz'min, V. A.. *Upper limit of the spectrum of cosmic rays.* JETP Lett., Vol. 4, p.78. DOI: N/A. arXiv: N/A.

[3] Kenneth Greisen. *End to the Cosmic-Ray Spectrum?* Phys. Rev. Lett. 16, 748. DOI: `https://doi.org/10.1103/PhysRevLett.16.748`. arXiv: N/A.

[4] Claire Guépin, Kumiko Kotera, et. al.. *Ultra-High Energy Cosmic Rays and Neutrinos from Tidal Disruptions by Massive Black Holes*. A&A 616 (2018) A179. DOI: `https://doi.org/10.1051/0004-6361/201732392`. arXiv:1711.11274v3.

[5] Ke Fang, Brian D. Metzger. *High-Energy Neutrinos from Millisecond Magnetars formed from the Merger of Binary Neutron Stars*. 2017 ApJ 849 153. DOI: `https://doi.org/10.3847/1538-4357/aa8b6a`. arXiv:1707.04263v2

[6] The Fermi LAT collaboration. *The spectrum of isotropic diffuse gamma-ray emission between 100 MeV and 820 GeV*. 2015 ApJ 799 86. DOI: `https://doi.org/10.1088/0004-637X/799/1/86`. arXiv:1410.3696v1.

[7] IceCube Collaboration. *Differential limit on the extremely-high-energy cosmic neutrino flux in the presence of astrophysical background from nine years of IceCube data*. Phys. Rev. D 98, 062003 (2018). DOI: `https://doi.org/10.1103/PhysRevD.98.062003`. arXiv:1807.01820v2

[8] The Pierre Auger Collaboration. *The Pierre Auger Observatory: Contributions to the 34th International Cosmic Ray Conference (ICRC 2015)*. Phys. Rev. D 98, 062003. DOI: `https://doi.org/10.1103/PhysRevD.98.062003`. arXiv:1807.01820v2.

[9] V. S. Berezinsky and G. T. Zatsepin. *Cosmic rays at ultrahigh-energies (neutrino?)* Phys. Lett. B 28 (1969) 423–424. DOI: https://doi.org/10.1016/0370-2693(69)90341-4. arXiv: N/A.

[10] M. G. Aartsen, et al. (The IceCube Collaboration) . *The IceCube Neutrino Observatory: Instrumentation and Online Systems*. JINST 12 P03012 (2017), DOI: 10.1088/1748-0221/12/03/p03012, arXiv:1612.05093v2.

[11] M. G. Aartsen, et al. (The IceCube Collaboration) . *Measurements using the inelasticity distribution of multi-TeV neutrino interactions in IceCube.* Phys. Rev. D 99, 032004 (2019), DOI: https://doi.org/10.1103/PhysRevD.99.032004, arXiv:1808.07629v2.

[12] J. D. Jackson, *Classical Electrodynamics 3rd Edition*, Chapter 13.4.

[13] N. van Eijndhoven et al. . *Implementation of a Gauss Convoluted Pandel PDF for Track Reconstruction in Neutrino Telescopes.* Astroparticle Physics 28 (2007) 456. DOI: https://doi.org/10.1016/j.astropartphys.2007.09.001. arXiv:0704.1706v2

[14] M. G. Aartsen, et al. (The IceCube-Gen2 Collaboration) . *IceCube-Gen2: The Window to the Extreme Universe.* J.Phys.G 48 (2021) 6, 060501, DOI: 10.1088/1361-6471/abbd48, arXiv:1612.05093v2.

[15] J. A. Aguilar, et. al. . *Design and Sensitivity of the Radio Neutrino Observatory in Greenland (RNO-G).* JINST 16 P03025 (2021), DOI: 10.1088/1748-0221/16/03/P03025, arXiv:2010.12279v3

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

[17] P. Aurenche , F. Gelis , H. Zaraket . *Landau-Pomeranchuk-Migdal effect in thermal field theory.* Phys.Rev. D62 (2000) 096012. DOI: https://doi.org/10.1103/PhysRevD.62.096012, arXiv:hep-ph/0003326v2

[18] E. Andres, et. al. . *The AMANDA Neutrino Telescope: Principle of Operation and First Results.* Astropart.Phys.13:1-20,2000. DOI: `https://doi.org/10.1016/S0927-6505(99)00092-4`, arXiv:astro-ph/9906203v1

[19] I. Kravchenko, et. al. (The RICE Collaboration). *Performance and Simulation of the RICE detector.* Astropart.Phys.19:15-36,(2003). DOI: https://doi.org/10.1016/S0927-6505(02)00194-9, arXiv:astro-ph/0112372v1

[20] P. Allison, et. al. . *Design and Initial Performance of the Askaryan Radio Array Prototype EeV Neutrino Detector at the South Pole.* Astropart. Phys. 35 (2012) 457–477. DOI: `https://doi.org/10.1016/j.icarus.2012.05.028`. arXiv:1105.2854v2

[21] P. Allison, et. al. (ARA Collaboration). *Performance of two Askaryan Radio Array stations and first results in the search for ultra-high energy neutrinos.* Phys. Rev. D 93, 082003 (2016). DOI: `https://doi.org/10.1103/PhysRevD.93.082003` . arXiv:1507.08991v3

[22] S.W. Barwick, et. al. (ARIANNA Collaboration). *A First Search for Cosmogenic Neutrinos with the ARIANNA Hexagonal Radio Array.* Astropart.Phys. 70 (2015) 12-26. DOI: https://doi.org/10.1016/j.astropartphys.2015.04.002 . arXiv:1410.7352v2

[23] A. Anker, et. al. . *A search for cosmogenic neutrinos with the ARIANNA test bed using 4.5 years of data.* JCAP03(2020)053. DOI: https://doi.org/10.1088/1475-7516/2020/03/053 . arXiv:1909.00840v3

[24] I. Kravchenko, et. al. . *RICE Limits on the Diffuse Ultra-High Energy Neutrino Flux.* Phys.Rev.D73:082002,(2006). DOI: https://doi.org/10.1103/PhysRevD.73.082002. arXiv:astro-ph/0601148v1

[25] P. Allison, et. al.. *Design and Performance of an Interferometric Trigger Array for Radio Detection of High-Energy Neutrinos.* Nuclear Instruments and Methods in Physics Research A. DOI: `https://doi.org/10.1016/j.nima.2019.01.067`. arXiv:1809.04573v2

[26] P. Gorham, et. al. (ANITA collaboration). *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). DOI: `https://doi.org/10.1016/j.astropartphys.2009.05.003`. arXiv:0812.1920v1

[27] The Pierre Auger Collaboration . *The Pierre Auger Cosmic Ray Observatory.* Nuclear Instruments and Methods in Physics Research A, 798 (2015) 172-213. DOI: https://doi.org/10.1016/j.nima.2015.06.058. arXiv:astro-ph/0601148v1.

[28] Jonas Heinze, et. al.. *A new view on Auger data and cosmogenic neutrinos in light of different nuclear disintegration and air-shower models.* ApJ 873 88. DOI: `https://doi.org/10.3847/1538-4357/ab05ce`. arXiv:1901.03338v3

[29] Amy L. Connolly, Abigail G. Vieregg. *Radio Detection of High Energy Neutrinos.* Neutrino Astronomy, pp. 217-240 (2017). DOI: `https://doi.org/10.1142/9789814759410_0015`. arXiv:1502.01323

[30] H. Bethe, W. Heitler. *On the stopping of fast particles and on the creation of positive electrons.* Proc. R. Soc. Lond. A146:83–112. DOI: `https://doi.org/10.1098/rspa.1934.0140`. arXiv: N/A

[31] C. Deaconu et al. . *Measurements and Modeling of Near-Surface Radio Propagation in Glacial Ice and Implications for Neutrino Experiments.* Phys. Rev. D 98, 043010 (2018). arXiv:1805.12576v2.

[32] G. A. Askaryan, *Excess negative charge of an electron-photon shower and its coherent radio emission.* Soviet Physics JETP 14 (1962) 441.

[33] David Saltzberg, et. al.. *Observation of the Askaryan Effect: Coherent Microwave Cherenkov Emission from Charge Asymmetry in High Energy Particle Cascades.* Phys.Rev.Lett. 86 (2001) 2802-2805 . DOI: https://doi.org/10.1103/PhysRevLett.86.2802. arXiv:hep-ex/0011001v1

[34] Jarred M. Roberts, et. al.. *LAB4D: A Low Power, Multi-GSa/s, Transient Digitizer with Sampling Timebase Trimming Capabilities.* Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 925, p92-100 (2019) . DOI: https://doi.org/10.1016/j.nima.2019.01.091. arXiv:1803.04600v2

[35] Grafana Labs, *Grafana*, Available from `https://grafana.com/grafana/`

[36] J. Avva et al. . *An in situ measurement of the radio-frequency attenuation in ice at Summit Station.* J.Glaciol. 61 : 1005-1011, 2015. DOI: https://doi.org/10.3189/2015JoG15J057. arXiv:1409.5413v3.

[37] The Telescope Array Collaboration. *Indications of Intermediate-Scale Anisotropy of Cosmic Rays with Energy Greater Than 57 EeV in the Northern Sky Measured with the Surface Detector of the Telescope Array Experiment.* ApJ 790 L21 (2014) . DOI: https://doi.org/10.1088/2041-8205/790/2/L21. arXiv:1404.5890v3

[38] Thomas A. Milligan, *Modern Antenna Design*, Chapter 5-7.

[39] Christian Glaser et. al. . *NuRadioMC: Simulating the radio emission of neutrinos from interaction to detector.* Eur. Phys. J. C 80, 77 (2020) . DOI: https://doi.org/10.1140/epjc/s10052-020-7612-8. arXiv:1906.01670

[40] P. Meade, *jade: An End-To-End Data Transfer and Catalog Tool.* J. Phys.: Conf. Ser. 898 062050. 10.1088/1742-6596/898/6/062050. arXiv: N/A.

[41] Eric Oberla. *RNO-G 2022 Deployment Plans.* RNO-G Collaboration Meeting - Chicago 2022 . DOI: N/A. arXiv: N/A

[42] ROOT Team, *ROOT: Data Analysis Framework*, Available from: `https://root.cern/`

[43] David Staelin . *Receivers, antennas, and signals.* Massachusetts Institute of Technology, course number: 6.661. DOI: N/A . arXiv:N/A.

[44] Python Software Foundation, *Python*, Available from: `https://www.python.org/about/`

[45] Jupyter, *Project Jupyter*, Available from: `https://jupyter.org/`

[46] ROOT Team, *Python interface: PyROOT*, Available from: `https://root.cern/manual/python/`

[47] James W. Cooley, John W. Tukey. *An algorithm for the machine calculation of complex Fourier series.* Math. Comp. 19 (1965), 297-301 . DOI: https://doi.org/10.1090/S0025-5718-1965-0178586-1. arXiv: N/A

[48] Cosmin Deaconu, *RNO-G Waveform Calibration.* RNO-G meeting November 15, 2021. DOI: N/A. arXiv: N/A

[49] Wendy Docters . *Unraveling the mysteries of high-energy cosmic rays using radio detection.* [S.n.] (2015). electronic ISBN: 978-90-367-7511-3 . arXiv:N/A

[50] Herbert E. Salzer and Ruth Zucker . *Table of the zeros and weight factors of the first fifteen Laguerre polynomials.* Bull. Amer. Math. Soc. 55 (1949), 1004-1012. DOI: https://doi.org/10.1090/S0002-9904-1949-09327-8. arXiv: N/A

[51] Pauli Virtanen, et. al. . *SciPy 1.0–Fundamental Algorithms for Scientific Computing in Python.* Nature Methods 17, 261 (2020). DOI: `https://doi.org/10.1038/s41592-019-0686-2`. arXiv:1907.10121v1

[52] I.A. BELOV . *On the computation of the probability density function of α-stable distributions.* Mathematical Modelling and Analysis 2005, Proceedings of the 10th International Conference MMA2005&CMAM2, Pages 333–341. ISBN: 9986-05-924-0 . arXiv: N/A.

[53] P. Allison, et. al. . *Long-baseline horizontal radio-frequency transmission through polar ice.* JCAP12(2020)009. DOI: https://doi.org/10.1088/1475-7516/2020/12/009 . arXiv:1908.10689v3

[54] A. Anker, et. al. . *Neutrino vertex reconstruction with in-ice radio detectors using surface reflections and implications for the neutrino energy resolution.* JCAP11(2019)030. DOI: https://doi.org/10.1088/1475-7516/2019/11/030 . arXiv:1909.02677v2 .

[55] Christian Glaser, Anna Nelles, et. al. . *NuRadioReco: A reconstruction framework for radio neutrino detectors.* Eur. Phys. J. C (2019) 79: 464. DOI: https://doi.org/10.1140/epjc/s10052-019-6971-5 . arXiv:1903.07023v2 .

[56] Rudge, Alan W. . *The Handbook of Antenna Design.* USA: IET (1982) Vol. 1, p. 24. ISBN: 0-906048-82-6 . arXiv: N/A . .

Acknowledgements

I would like to start off by expressing my deepest gratitude to both my promotor and co-promotor, Nick Van Eijndhoven and Krijn de Vries respectively, for their excellent and irreplaceable guidance and inspiration during the making of this thesis. The countless suggestions that they have given and the friendly & healthy work environment they have provided me with are very much appreciated. This work would not have been possible without them and all the time and effort they have put into it. Furthermore, I want to thoroughly thank Simona Toscano for supervising this work as well. Her advice, assistance and contributions to this work (especially during the earlier stages of the thesis, but also later on) are numerous and played a crucial role. Next, I want to thank my parents for making this five year journey of studying physics possible and for supporting me along the way in every way, shape and form possible. Without their continuous love and support I could never have made it this far and so I will be forever grateful to them for all that they have done for me. Moreover, I would like to thank my high school physics teacher Olbren for cultivating my fascination for physics, providing me with the necessary tools to begin my physics journey, and, most importantly, for convincing me to take my shot at studying at a university. Without him, I would most likely not even have pursued a higher education. Next up, I also want to thank Joe. And lastly, I would like to thank the 9th-century Ethiopian goat-herder Kaldi for the invention of coffee, SWAMP for all the fun and memorable times, second order approximations for being relatively accurate (most of the time), uncle Iroh for his wisdom and Darco & Jesse for keeping me sane.

Appendix A

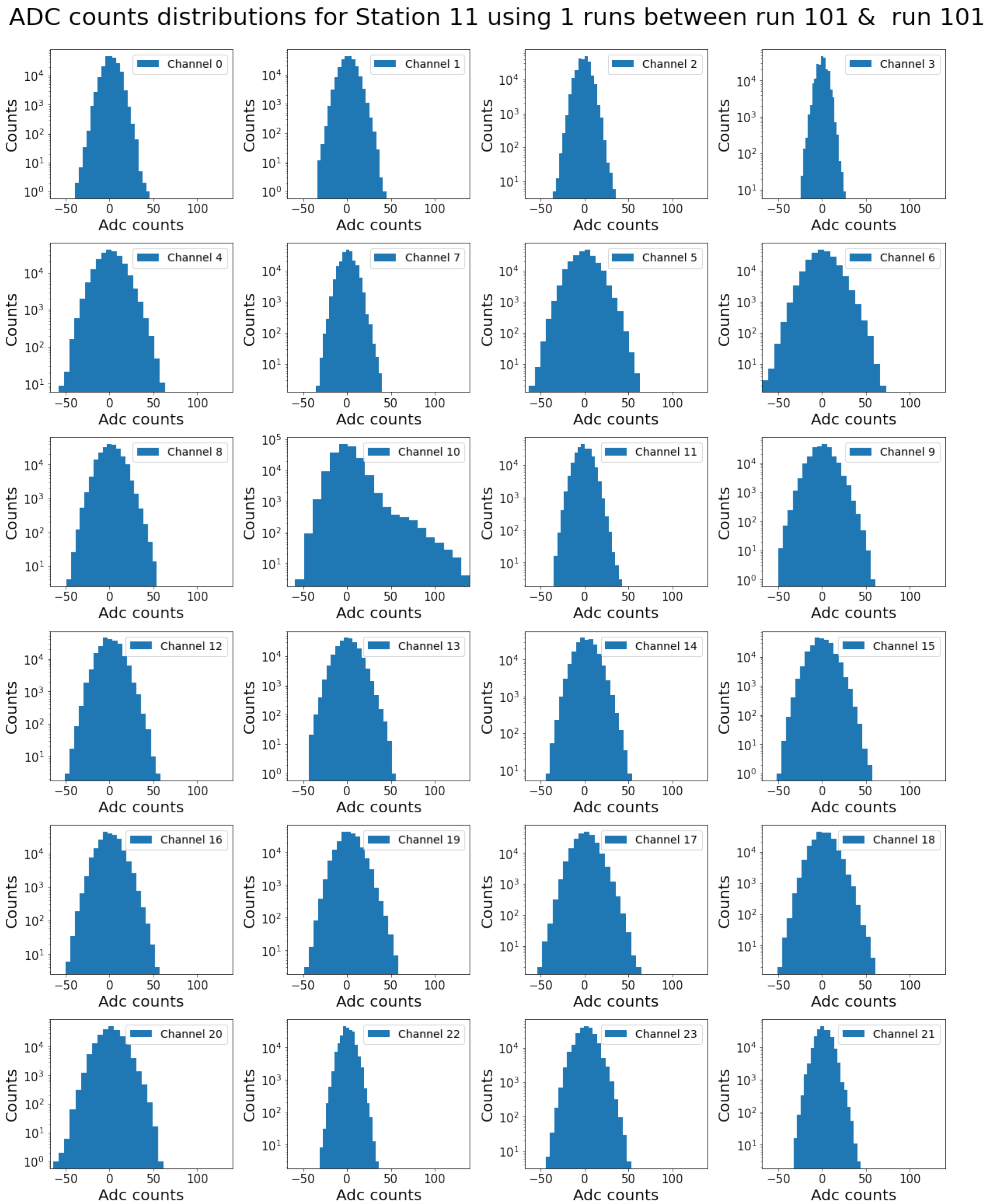


Figure 28: ADC count distribution of station 11, run 101 for all channels.

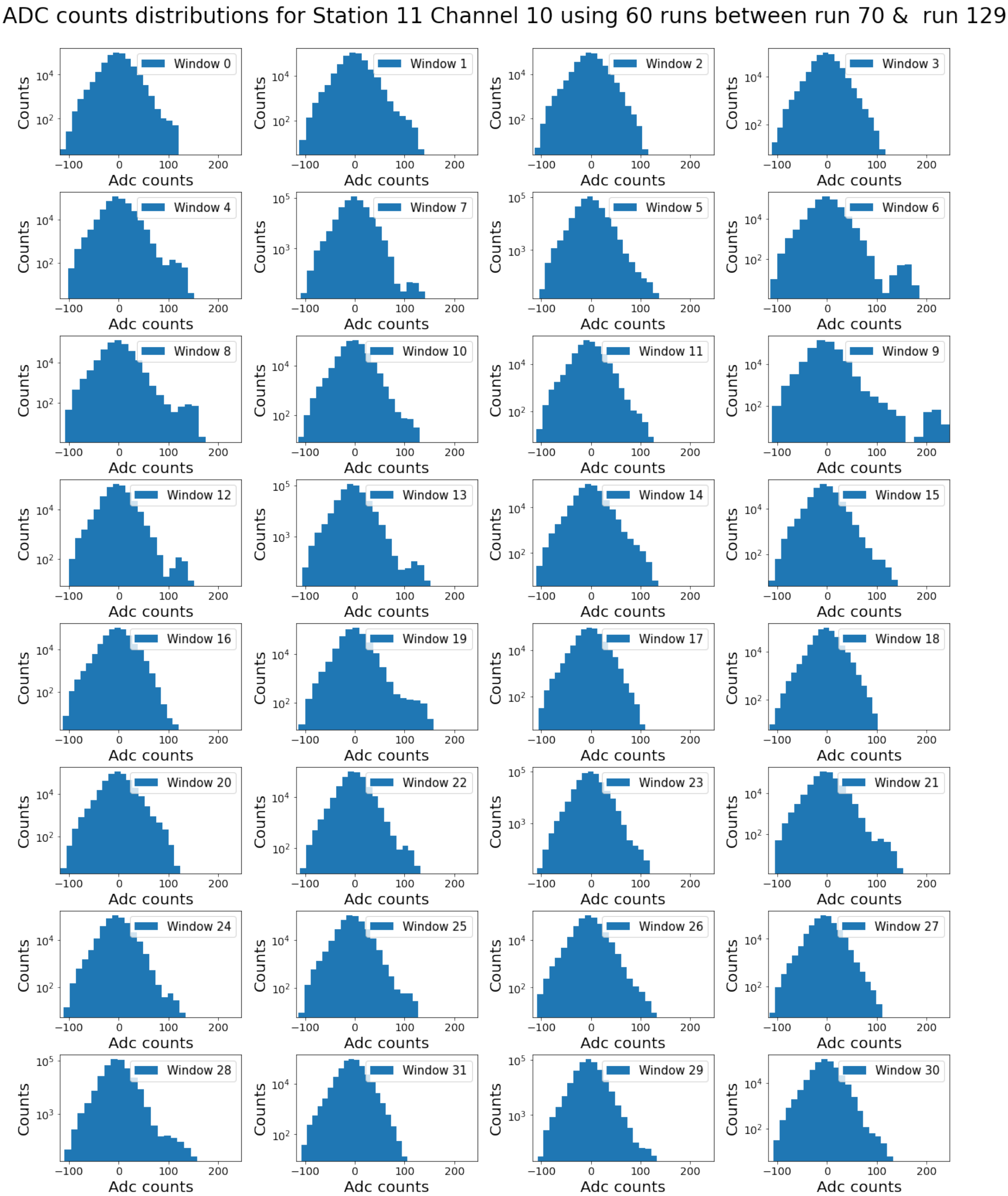


Figure 29: ADC count distributions for all physical windows of channel 10, station 11 during run 101

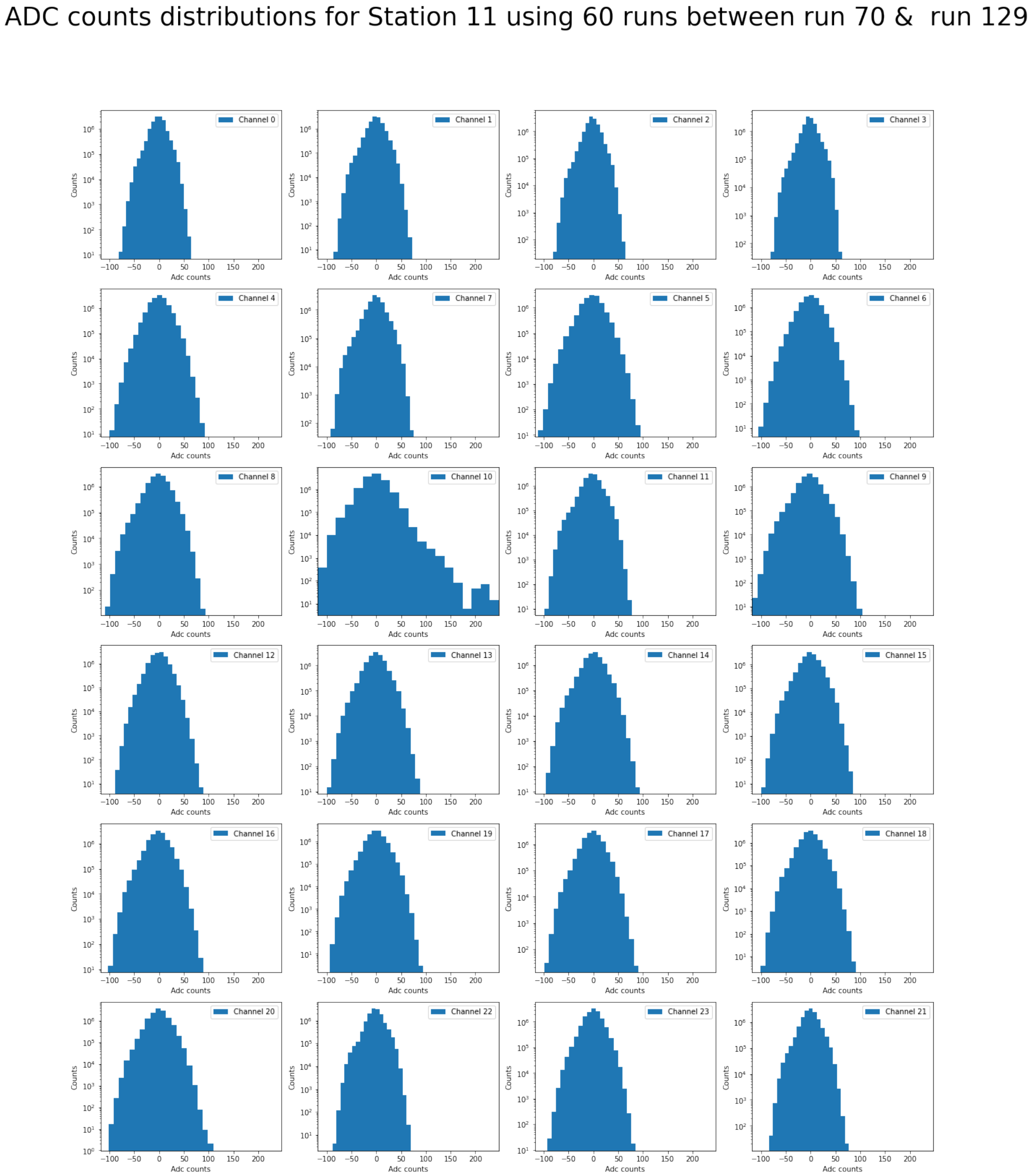


Figure 30: ADC count distribution of station 11, runs 70 to 129 for all channels.

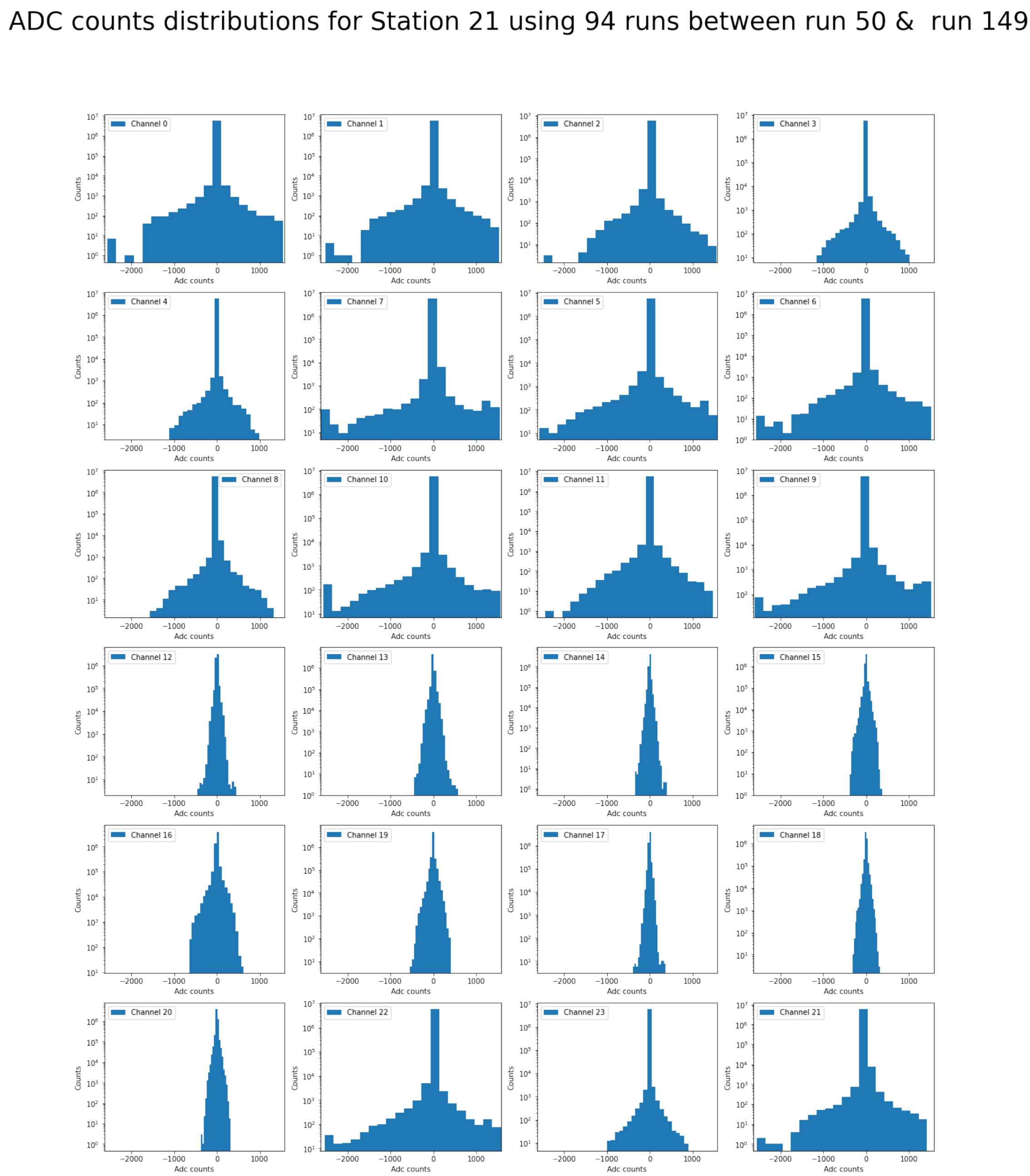


Figure 31: ADC count distribution of station 21, runs 50 to 149 for all channels.

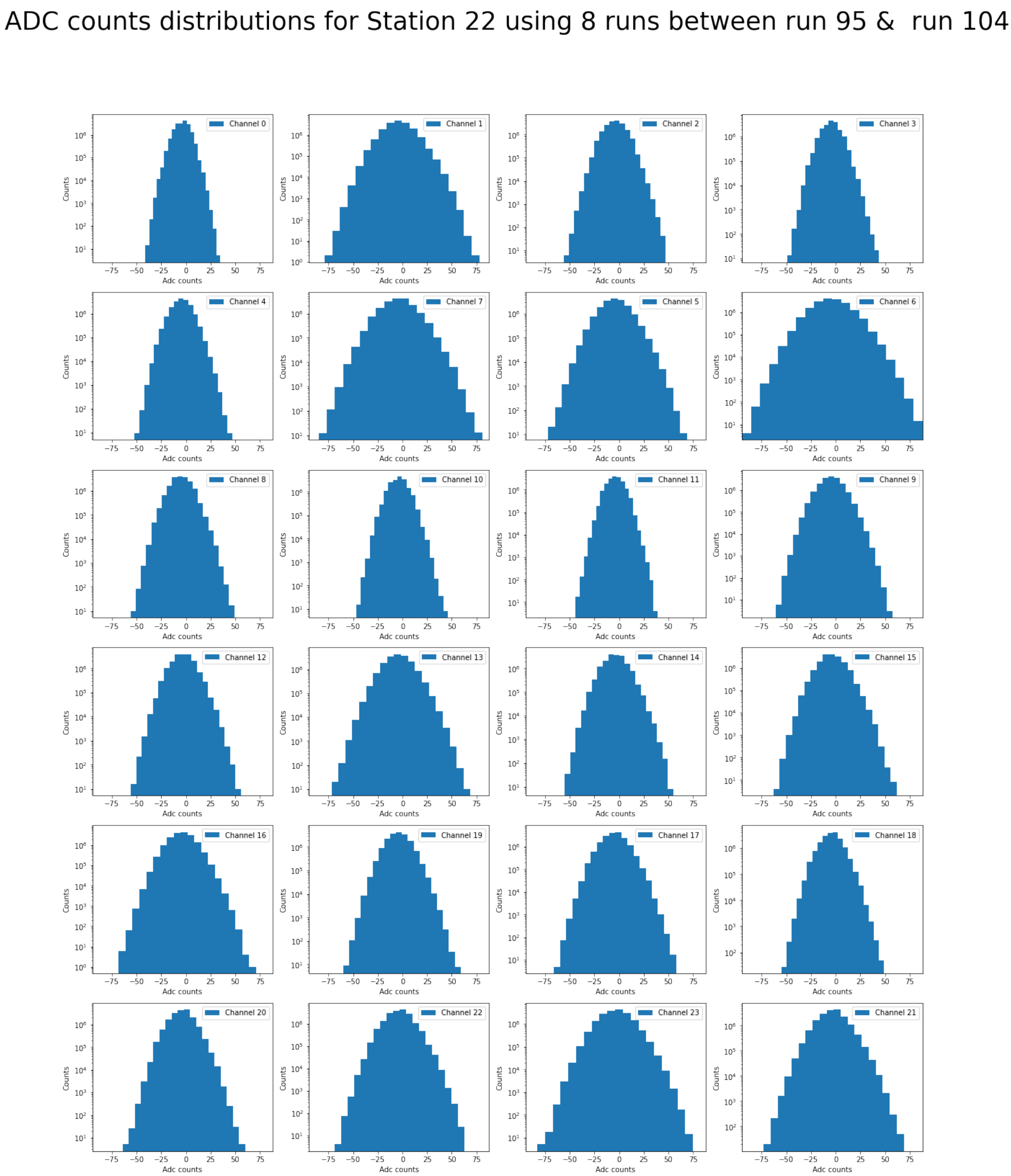


Figure 32: ADC count distribution of station 22, run 95-104 for all channels.

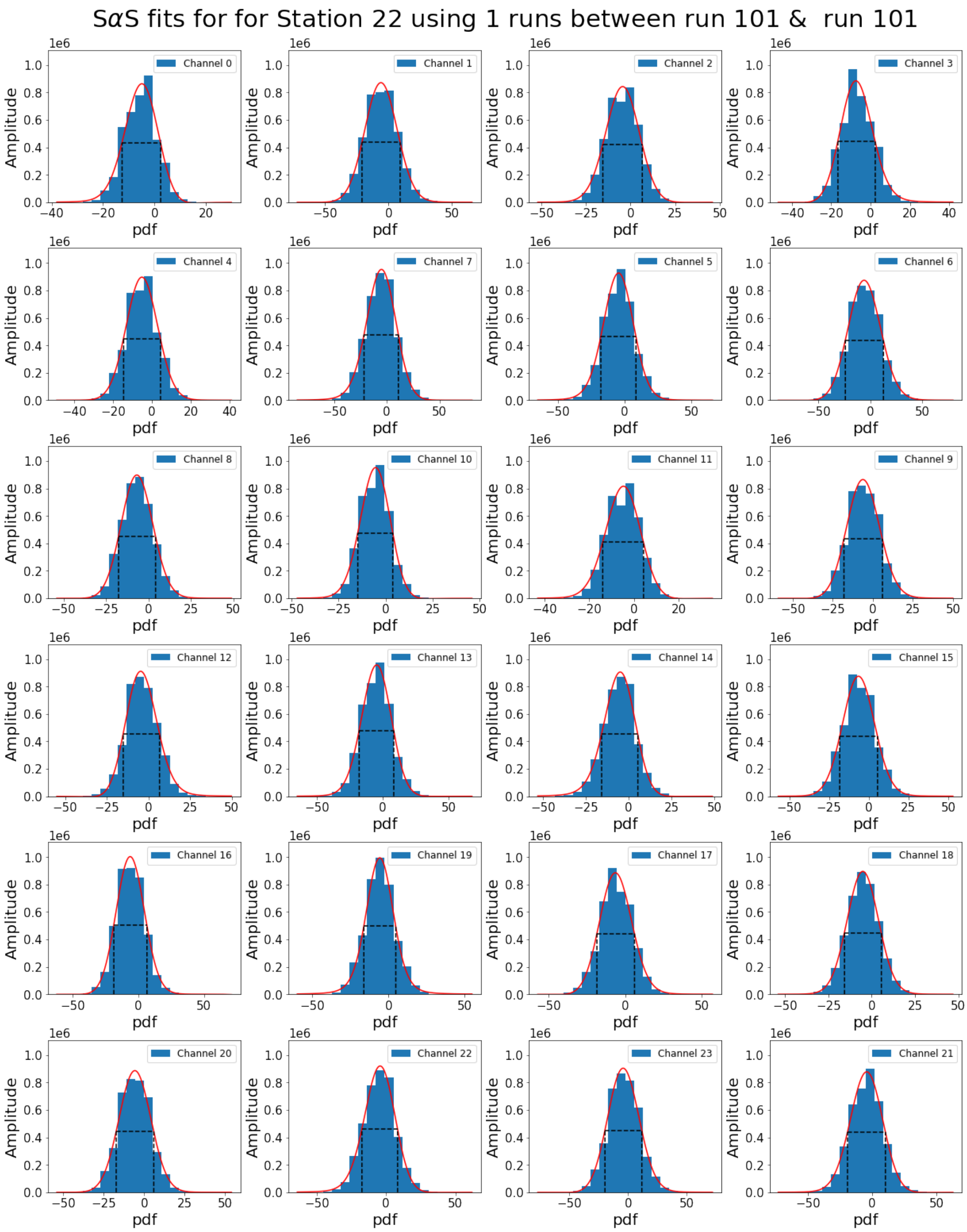


Figure 33: ADC count distributions and associated SαS fit for channels of station 22 during run 101

Channel Number	μ	β	α	σ	FWHM
0	-5.02691492	0.03306829	2.	4.49152467	15.0
1	-5.92526994	-0.04194721	1.99999988	9.04933032	30.1
2	-4.45255788	0.03961067	2.	6.64195243	22.1
3	-7.32201528	-0.08866701	1.95638132	5.74284141	19.0
4	-5.19064127	-0.00849574	1.98092833	5.73531013	19.0
5	-4.52785932	0.07488891	1.89396021	8.08213829	26.3
6	-5.88598915	-0.10242043	1.99999196	11.04354461	36.8
7	-5.22443812	0.16240755	1.91303555	9.95375934	32.6
8	-6.80195838	-0.02982147	1.99999962	6.55350004	21.8
9	-6.14290607	-0.05902266	2.	7.19810656	24.0
10	-5.44485236	0.05547112	2.	5.57648682	18.6
11	-4.84681144	0.05237231	2.	5.45577297	18.2
12	-4.4364126	-0.15070562	1.95957956	6.57691232	21.8
13	-4.71894078	0.08216526	1.97535849	7.93987437	26.3
14	-5.06437219	0.14658083	1.90501708	6.43400363	21.1
15	-6.46944598	-0.02149644	1.95487771	7.19106165	23.8
16	-6.1328057	-0.04545073	2.	7.57493145	25.2
17	-6.28501798	-0.08375895	1.98699503	7.35503116	24.5
18	-5.13907458	-5.79471638e-03	1.96465498	6.44700399	21.3
19	-5.74363063	3.27244940e-03	1.76489454	6.63576667	21.0
20	-5.76115131	-0.01298246	2.	6.96143581	23.2
21	-4.26526032	0.09329226	2.	8.84979806	29.5
22	-4.44188322	0.14711469	1.93488021	7.81972621	25.7
23	-4.15633938	-0.09150758	2.	9.41576349	31.4

Table 1: Table of the parameters which characterise the ADC count distribution of the channels from station 22 during run 101 as an SαS distribution.

Appendix B

Summary of the Jupyter notebooks which were developed in order to perform the analyses done in this work:

- <u>Functions.ipynb:</u>
  This notebook contains all python functions developed for this work in order to allow for easy importing. The most notable functions are summarised below:

  - *ADCDist(StNr,Runs,NBins=20,WPed=False)*: Produces a plot such as 28, given the station number StNr and a list of runs (Runs) for which this has to be computed. It also has the option to change the amount of bins NBins used in the histogram and the option WPed to include the pedestal values.
  - *ADCDistCh(StNr,ChNr,Runs,NBins=20,WPed=False)*: Produces the same kind of plot as the ADCDist function, e.g. Figure 29, but for a specific channel number ChNr of the station.
  - *TimeTraces(StNr,Runs,ChNr,WPed=False)*: Plots the timetraces of the waveforms measured by channel number ChNr in station StNr during the runs in list Runs with the possibility WPed of taking the Pedestals into account.
  - *SASFitRuns(StNr,Runs,ChNrs=range(0,24),Plot=True)*: Fits the SαS distribution to the ADC distribution of the channels in ChNrs from station StNr during the runs in list Runs with the option of returning a plot like 27
  - *SASFitStation(StNr,Runs,NBins=20, Plot=True)*: Does the same as the SASFitRuns function but characterises an entire station given the station StNr and the list of runs (Runs) as shown in Figure 33.
  - *ThermalNoiseTemp(StNr,Runs,ChNr,DateStr,Debug=False)*: Computes the thermal noise temperature associated to the data from the runs in the list Runs from the channel ChNr from station StNr and the date when the data was taken DateStr in the form "YYYYMMDD". There is also an option Debug to shown the intermediate plots during the entire conversion methodology as outlined in section 3.4.2.

- <u>ChannelAverages.ipynb:</u>
  A first exploration of the average values of station 11, from which Figure 12 came, was performed in this script. The first identification of the waveform baselining was done here as well.

- <u>CompareAvgEnvironment.ipynb:</u>
In this notebook a potential correlation between the waveform baselining and the ambient temperature was investigated using the environmental data from Grafana. However, when zooming in on the timescale of the anomalous run, the temperature appeared to have a constant value.

- <u>WindowCheck.ipynb:</u>
This script contains an investigation of the average values for each window on a per channel basis for an entire station. This investigation was done twice, once while keeping the window unrolling as described in Figure 16 in account (so as to investigate the physical windows) and once without keeping this into account (in order to investigate the chronological windows). This piece of code produces figures such as Figure 17. The same investigation was also performed but with the pedestal values re-added to the waveform data in order to investigate whether the waveform baselining originated from a pedestal miscalculation.

- <u>ADCCountDistrInspect.ipynb:</u> In this notebook, a first inspection of the ADC count distributions was performed to identify which stations act as expected under which runs and where anomalous behaviour is to be expected. It is here that Figures such as 28 and 29 were developed. Then, the timetraces associated to the abnormally behaving ADC count distributions were explored in more detail in order to investigate what kind of a problem these timetraces exhibit.

- <u>HistogramFit.ipynb:</u> This script focusses on fitting the SαS distribution to an ADC count distribution. The main part consists of an exploration of the SαS distribution followed by a manual fit of this distribution to a specific run of data and a characterisation of the noise for a full station. In the last part, these fits are performed via the use of an automated function in order to yield plots such as 27 and 33.

- <u>ADCToAntennaConversion.ipynb:</u> The entire conversion from ADC counts to voltage at the antenna as outlined in section 3.4.2. On top of that, the thermal power noise temperature was computed from some specific runs in this script.